TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Proteomik und Bioanalytik

When chemical proteomics meets medicinal chemistry:
Guided drug discovery towards EPHA2 inhibitors

Stephanie Heinzlmeir

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

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Martin Klingenspor

Prüfer der Dissertation:
1. Prof. Dr. Bernhard Küster
2. Prof. Dr. Norbert Sewald
3. Prof. Dr. Stephan Sieber (schriftliche Beurteilung)
   Prof. Dr. Stephan Hacker (mündliche Prüfung)

Die Dissertation wurde am 19.09.2017 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 15.11.2017 angenommen.

“The world is your oyster. It’s up to you to find the pearls.”

- The pursuit of happiness -

Abstract

The receptor tyrosine kinase EPHA2 (Ephrin type-A receptor 2) is involved in a multitude of pathological settings and has gained rising interest as a therapeutic drug target, notably in cancer and infectious diseases. Both the elucidation of its biological function and its implication in disease is complicated by intertwined and context-dependent signaling outcomes. So far, investigation of EPHA2 biology was hampered by the lack of dedicated small molecules for pharmacological inhibition of EPHA2. Such chemical probes require high selectivity towards their target; a task which is often not achieved by small molecule kinase inhibitors targeting the highly conserved ATP binding pocket. The development of selective inhibitors, regardless of their use as chemical probes or therapeutic agents, benefits from early selectivity profiling which can support different stages of a medicinal chemistry project.

This work is based on a chemical proteomic approach (Kinobeads) that was used to survey 235 clinical kinase inhibitors for their target space. Here, a new selectivity metric is described which helps to apprehend the selectivity of this unparalleled dataset. CATDS (Concentration and Target Dependent Selectivity) is a simple but versatile selectivity scoring system and outperforms existing scores allowing for concentration- and target-dependent determination of compound selectivity – thus qualifying it for an application in medicinal chemistry and in particular for the discovery of EPHA2 inhibitors.

Within the Kinobeads drug screen dataset, 24 clinical kinase inhibitors were found to bind EPHA2 as an off-target with sub-micromolar affinities – showing the value of such a screen for lead generation. Subsequent structure-affinity-relationship analysis of nine co-crystal structures delineated molecular drug:EPHA2 interactions in full detail. An original combination of selectivity profiling, structure determination and kinome-wide sequence alignment allowed the development of a residue classification system for amino acids involved in drug:protein interaction – amino acids in the drug binding site of EPHA2 were categorized into key, scaffold, potency and selectivity residues. This classification scheme was also exerted for ABL1 and MELK using also published crystal structures, evidencing the general applicability of the method. This demonstrates that this scheme formalizes the medicinal chemistry process and enables rationalized inhibitor design towards high selectivity compounds.

After selection of the clinical BCR-ABL/SRC inhibitor Dasatinib as the lead structure for a medicinal chemistry project targeting EPHA2, hybrid structures of Dasatinib and the previously identified EPHA2 binders CHEMBL249097, PD-173955 and a known EPHB4 inhibitor were designed and synthesized. These inhibitors exploit both the ATP pocket entrance and the ribose pocket as binding epitopes in the kinase EPHA2. Inhibitor optimization was guided by early Kinobeads selectivity profiling of the newly synthesized inhibitor candidates. Concomitant protein crystallography of 17 inhibitor co-crystals delivered detailed insight into the atomic interactions that underlie the structure-affinity-relationship. The *in vivo* efficacy and an anti-proliferative effect of the inhibitor candidates was confirmed in the glioblastoma cell line SF-268 that was selected according to its favorable protein expression profile. These efforts led to the discovery of inhibitor candidate **4a** – featuring an improved selectivity profile (CATDS: 0.176) while maintaining potency against EPHA2 (K_D^{app}: 0.8 nM) and anti-cancer activity in SF-268 cells (EC_{50}: 92 nM).

The work presented in this thesis showcases the powerful combination of chemical proteomics and medicinal chemistry and introduces novel concepts of chemoproteomics-aided drug discovery.

Zusammenfassung

Die Rezeptortyrosinkinase EPHA2 (Ephrin type-A receptor 2) ist an der Entstehung einer Vielzahl von Erkrankungen beteiligt und hat zunehmendes Interesse als therapeutischer Ansatzpunkt, vor allem für Krebs- und Infektionskrankheiten, erlangt. Die Aufklärung sowohl ihrer biologischen Funktion als auch ihrer Wirkung im Krankheitsgeschehen wird durch die äusserst verzweigten und kontextabhaengigen Signalwege dieser Kinase erschwert. Bisher stehen keine dezidierten niedermolekularen Kinaseinhibitoren für die pharmakologische Hemmung von EPHA2 zur Verfügung, was die weitere Erforschung der biologischen Funktion dieser Kinase erschwert. Derartige chemische Sonden zeichnen sich durch eine hohe Selektivität gegenüber ihrem Zielprotein aus; dies wird jedoch von niedermolekularen Inhibitoren, welche an die hochkonservierte ATP Bindetasche binden, meist nicht erreicht. Die Entwicklung selektiver Inhibitoren, unabhängig von ihrer Verwendung als chemische Sonde oder therapeutische Wirkstoffe, profitiert von einer frühzeitigen Aufklärung ihres Selektivitätsprofils, da auf diese Weise verschiedene Stadien eines medizinalchemischen Projektes unterstützen werden können.

Diese Arbeit basiert auf einem chemisch-proteomischen Ansatz (Kinobeads), welcher verwendet wurde, um die Zielproteine von 235 klinischen Kinaseinhibitoren aufzuklären. Dies ermöglichte die Entwicklung einer neuen Selektivitätsmetrik, die dazu beiträgt, die Selektivität dieses einmaligen Datensatzes zu erfassen. CATDS (Concentration and Target Dependent Selectivity) ist ein einfaches, aber vielseitiges Selektivitäts-Scoring-System und übertrifft vorhandene Metriken, da es eine konzentrations- und zielprotein-abhängige Bestimmung der Selektivität ermöglicht und sich damit für eine Anwendung in der Medizinchemie und insbesondere für die Entdeckung von EPHA2 Inhibitoren eignet.

Basierend auf diesem Kinobeads Datensatz wurden 24 klinische Kinaseinhibitoren identifiziert, welche EPHA2 unbeabsichtigt mit submikromolaren Affinitäten binden - was den Wert eines solchen Screen für die Generierung neuer Leitstrukturen demonstriert. Mittels einer anschließenden Struktur-Affinitäts-Beziehungsanalyse von neun Co-Kristallstrukturen wurden die molekularen Wechselwirkungen zwischen diesen Inhibitoren und EPHA2 im Detail beschrieben. Eine neuartige Kombination aus Selektivitätsprofilierung, Strukturbestimmung und kinom-weiter Sequenzanalyse ermöglichte die Entwicklung eines Klassifizierungssystems für Aminosäuren, die an der Wirkstoff-Protein-Interaktion beteiligt sind: Aminosäuren in der betreffenden Bindungsstelle von EPHA2 wurden als Schlüssel-, Gerüst-, Potenz- und Selektivitätsreste kategorisiert. Dieses Klassifizierungsschema wurde anhand öffentlich zugänglicher Kristallstrukturen auch für ABL1 und MELK verwendet, was die allgemeine Anwendbarkeit der Methode belegt. Dies zeigt, dass dieses Schema die medizinalchemische Herangehensweise formalisiert und ein rationalisiertes Inhibitordesign für Verbindungen mit hoher Selektivität ermöglicht.

Der klinische BCR-ABL/SRC-Inhibitor Dasatinib wurde als Leitstruktur für ein EPHA2-spezifisches medizinalchemisches Projekt ausgewählt, woraufhin Hybridstrukturen von Dasatinib und den bekannten EPHA2-Bindern CHEMBL249097, PD-173955 und einem bekannten EPHB4-Inhibitor entworfen und synthetisiert wurden. Diese Inhibitoren nutzen sowohl den Eingang der ATP Bindetasche als auch die Ribose-Tasche als Bindungs-Epitope in der Kinase EPHA2 aus. Die Inhibitor-Optimierung wurde durch die frühzeitigen Aufklärung des Selektivitätsprofils der neu synthetisierten Inhibitor-Kandidaten mittels Kinobeads angeleitet. Gleichzeitig lieferte die kristallografische Analyse von 17 Inhibitor-Co-Kristallen detaillierte Einblicke in atomare Wechselwirkungen, welche der Struktur-Affinitäts-Beziehung zugrunde liegen. Die intrazelluläre Wirksamkeit und eine anti-proliferative Wirkung der Inhibitor-Kandidaten wurde in der Glioblastom-Zelllinie SF-268 bestätigt. Diese Zelllinie wurde anhand ihres vorteilhaften Protein-Expressionsprofil ausgewählt. Dies führte zur Entdeckung des Inhibitor-Kandidaten **4a** – welcher ein verbessertes Selektivitätsprofil (CATDS: 0.176) unter Beibehaltung der Potenz gegen EPHA2 (K_D^{app}: 0.8 nM) und anti-proliferative Aktivität in SF-268-Zellen (EC_{50}: 92 nM) aufweist.

Die in dieser Dissertation dargelegte Arbeit demonstriert die wirksame Kombination von chemischer Proteomik und medizinischer Chemie und präsentiert neue Konzepte der chemoproteomik-unterstützten Wirkstoffforschung.

Table of contents

Abstract i
Zusammenfassung iii
Table of contents v
Introduction 1
Experimental Procedures 39
Results and Discussion 63
General Discussion and Outlook 113
References 123
Abbreviations 137
Acknowledgement I
Publication record III
Curriculum vitae V
Appendix VII

Introduction

Table of contents

1 EPHA2 is a rising target within the EPH kinase family 3
  1.1 EPH family of receptor tyrosine kinases – an overview 3
  1.2 EPHA2 signaling - defining biological and disease state 6
  1.3 Role of EPHA2 in biology and disease 9
  1.4 Different molecular strategies are used to target EPHA2 12
2 Small molecule kinase inhibitors targeting the kinase domain 16
  2.1 Catalytic activity and structure of the protein kinase domain 16
  2.2 Kinase inhibitors comprise different binding modes 19
  2.3 Different metrics to calculate small molecule selectivity 22
3 Mass-spectrometry based target deconvolution 25
  3.1 Mass spectrometry based quantitative proteomics 26
  3.2 (Chemo-)proteomic methods used for target deconvolution 29
  3.3 Thermodynamic considerations about ligand:protein interactions 32
4 Objectives and outline 37

“Todo se puede hacer. Es cuestion de atraverse a ello o no.”
- Carlos Ruiz Zafon -

1 EPHA2 is a rising target within the EPH kinase family

Protein kinases are among the most important cellular players since they are responsible for receiving and processing external and internal stimuli to trigger a desired cellular response. In 2002, Manning *et al*[1] catalogued the genetic sequences of 518 human protein kinases revealing these enzymes to be one of the largest protein class of the human genome. These proteins build up a fine-tuned and intertwined signaling network that enables the cell to react in a flexible manner upon manifold stimuli, to change its phenotype and behavior according to the environmental requirements and to maintain cellular homeostasis by the regulation of protein activity and localization.[2] The important function of receiving and transmitting external signals is accomplished by 58 transmembrane receptor tyrosine kinases (RTK).[1] This protein kinase family is responsible for the detection of extracellular signals such as growth factors, cytokines and hormones, and also plays a pivotal role in cell-cell communication and tissue homeostasis. The largest subfamily of receptor tyrosine kinases are the erythropoietin-producing hepatocellular carcinoma (EPH) receptor kinases. The genes for EPH receptors and their respective ligands are highly conserved among the animal kingdom[3] which points towards their important contribution to cell biology and integrity. EPH receptors, and the ephrin type-A receptor 2 (EPHA2) in particular, have attracted growing interest over the past two decades (Figure 1a).[4] Their intriguing biology, implication in various pathological conditions and therapeutic targeting strategies are slowly being uncovered.[5]

1.1 EPH family of receptor tyrosine kinases – an overview

Since 1987 – when the first EPH receptors were discovered in a hepatoma cell line[6] – these receptor tyrosine kinases have developed from the status of orphan receptors with unknown biological function to being regarded as highly complex signaling mediators involved in a vast number of biological processes. The EPH receptor family includes 14 members and forms the largest subfamily of the 58 known transmembrane receptor tyrosine kinases. According to sequence similarity and ligand binding affinities, EPH receptors and their respective ligands, EPH receptor interacting proteins (ephrins), are divided into two major subclasses; nine EPHA receptors preferentially bind five ephrin-A ligands, whereas five EPHB receptors mainly interact with three ephrin-B ligands.[4,7,8] However, in the light of accumulating evidence concerning EPH/ephrin signaling, this division appears less and less accurate: EPH receptors likely interact in large heteromeric clusters and some EPH receptors have also been found to bind both types of ligands.[9,10]

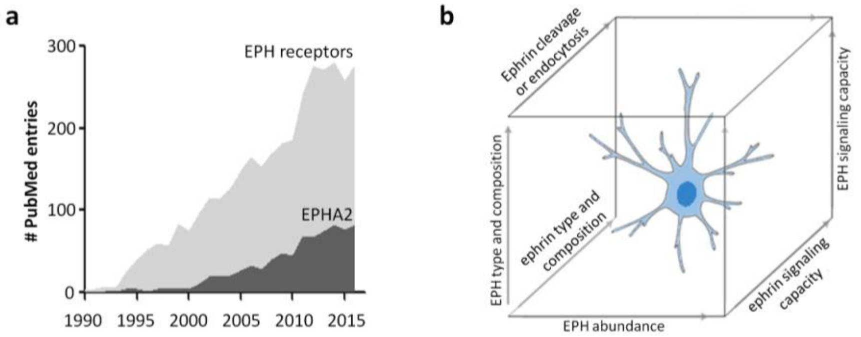


Figure 1 | EPH family of receptor tyrosine kinases. a) Number of PubMed entries for EPH receptors and EPHA2 in particular. **b)** The cellular effect and biological outcome of EPH signaling is the result of a complex interplay between abundance, composition, activation state and protein processing of both, EPH receptors and ephrin ligands. Modified from Lackmann *et al*.[4]

EPH receptor biological function. EPH receptor signaling was first found to regulate axon path-finding and neuronal cell migration in developing tissues.[11,12] In the following years, EPH signaling gained immense interest and researchers have uncovered that a multitude of biological processes, such as migration or cell-cell communication, depend on EPH activity.[13] It becomes more and more evident that EPH:ephrin signaling orchestrates a Yin and Yang regulation of chemoattractive and chemorepulsive actions. It provides a guidance system for motile cells along multidimensional spatial gradients of receptors and ligands[14] – leading them towards their predestined localization and into a dynamic range of fine-tuned cellular responses such as repulsion/attraction and motility/adhesion (Figure 1b).[4,15-17] Hence, the cellular effect is established rather by a precisely modulated combination of several EPH:ephrin pairs than by a single EPH:ephrin interaction.[4] EPH:ephrin signaling is particularly suited for cell-cell communication leading to cell positioning and sorting and, therefore, plays an important role during organ development and patterning of the vascular, skeletal and nervous system.[4] EPH receptors are implicated in the developing nervous system (axon guidance, spatial orientation, synapse formation)[18] but also in adult nervous tissue in which they regulate synapse function, neural stem cell properties[19] and memory formation[20]. Similarly, EPH signaling is involved in the developing and the adult cardiovascular system and angiogenesis.[21] It is not surprising that misregulation of EPH signaling leads to very diverse pathological outcomes and was found to be associated with several neurodegenerative diseases,[22] induction and persistence of pain, cardiovascular diseases, pathogen infection and cancer.[21]

Structural features of EPH receptor and EPH:ephrin binding. In contrast to many other RTK ligands, ephrins are bound to the plasma membrane enabling receptor:ligand interaction only upon close contact of two neighboring cells. Ephrin ligands are anchored to the membrane via a glycosylphosphatidylinositol (GPI) anchor (ephrin-A subtype) or via a transmembrane region linked to a cytoplasmic PSD95/Dlg/ZO1 (PDZ) domain (ephrin-B subtype).[8] EPH receptors are type 1 transmembrane proteins and consist of eight domains which exert diverse molecular and structural functions (Figure 2).[23] The highly conserved N-terminal extracellular ligand binding domain (LBD) is responsible for ligand recognition and binding which occurs in several steps: initially, an induced fit mechanism leads to the formation of high affinity interactions between ligands and receptors; these dimers further tetramerize via an additional low affinity interface;[8] eventually, higher-order oligomers are formed which can be composed of different EPH receptors and ligands and thus modulate different signaling outcomes.[8,15,24] The ligand binding domain (LBD) is connected to a cysteine-rich domain (including a Sushi and EGF-like domain) and two fibronectin-type II repeats which are responsible for receptor dimerization and extracellular receptor orientation.[25] At the cytoplasmic part of the receptor, the transmembrane domain and the juxtamembrane region are followed by the highly conserved catalytic tyrosine kinase domain. The juxtamembrane domain serves two main functions – providing a scaffold for src homology 2 domain containing downstream signaling proteins (e.g. VAV2[26]) and regulating kinase activity through conformational changes.[27] In the non-activated state, the juxtamembrane domain tightly interacts mainly with the N-lobe of the kinase domain which provokes the distortion of the central αC-helix motif and an inactivation of the kinase.[8] Upon ligand-stimulation, phosphorylation of both the juxtamembrane domain and the kinase domain leads to structural reorganization and releases the active kinase domain to perform its catalytic action.[28,29]

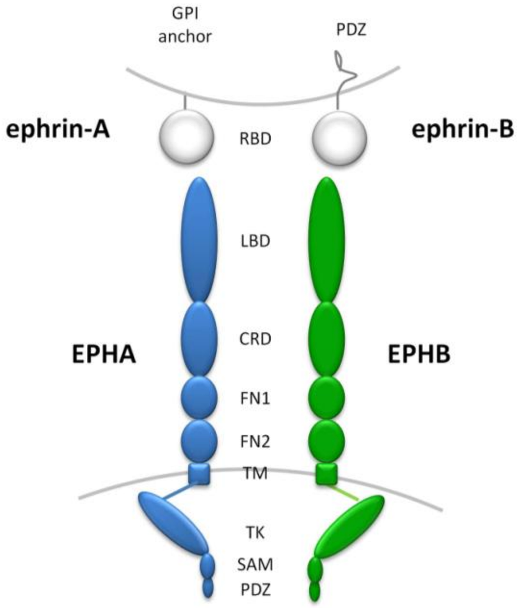


Figure 2 | Structural features of EPH receptor and EPH:ephrin binding. Structural composition of EPHA and EPHB receptors and their respective ephrin-A and ephrin-B ligands. EPH receptors are structured into an extracellular part containing a ligand-binding domain (LBD), a cysteine-rich domain (CRD), two fibronectin-type II repeats (FN1, FN2), the transmembrane domain (TM), and a cytoplasmic part composing of the tyrosine kinase domain (TK), a sterile alpha motif domain (SAM) and a PSD95/Dlg/ZO1 domain (PDZ). Modified from Guicciardo *et al.*[30]

EPH receptors possess a sterile alpha motif (SAM) domain – a unique feature among the 20 subfamilies of receptor tyrosine kinases.[31] SAM domains mediate protein:protein interactions and were found to regulate dimerization and kinase activity of different EPH receptors in diverse ways.[8] It is believed that the multi-faceted biological functions of EPH receptor signaling may be due to differential lateral receptor interactions[32] – rendering the SAM domain an important module for the regulation of diverse and even opposite cellular outcomes. How SAM domains function in all EPH receptors remains to be further investigated. However, it has already been found that in the absence of a ligand, the SAM motif stabilizes receptor dimerization in case of EPHA3,[33] whereas it performs the opposite function in EPHA2 in which the dimer is destabilized and kinase activity is suppressed.[23] Deletion of the SAM domain stabilizes EPHA2 dimers independently from ligand binding and leads to constitutive activation of EPHA2 similar to the ligand-bound form.[23,31] This also implies that interactions of an EPH receptor SAM domain with other binding partners might directly influence the dimerization and activity of the neighboring kinase domain.[23] Akin to ephrin-B ligands, EPH receptors also contain a C-terminal PDZ domain which provides a platform for anchoring the cell surface receptor to cytoskeleton components and for the formation and organization of signaling complexes.[34]

Receptor processing. After cell-cell contact and ligand stimulation, EPH:ephrin interactions are released, EPH receptors are internalized and partly degraded or recycled. Several scenarios of internalization have already been identified: i) RAC-mediated ruffling of the opposing cell membrane induces transendocytosis (as shown for EPHB4 and ephrin-B2)[35] where the complete EPH:ephrin signaling cluster is internalized by one of the cells, thus transferring associated components from one cell into another; ii) clathrin-mediated endocytosis (shown for both receptors[36] and ligands[37]) leads to persisting endocytotic signaling, similar to other RTKs; iii) proteolytic shedding via membrane bound proteases such as ADAM10 releases the EPH:ephrin association and allows cell repulsion.[38,39]

Given the complexity of EPH receptor structure and receptor processing, it appears prudent to assume versatile and fine-tuned signaling events driven by EPH:ephrin interactions. In-depth investigation of the roles of each receptor and ligand module is inevitable and will pave the way for elucidating the multi-layered and sometimes even opposite biological functions of EPH:ephrin signaling.

1.2 EPHA2 signaling - defining biological and disease state

One prominent and well-studied member of the EPH family is EPHA2. EPHA2 signaling is involved in a multitude of cellular processes modulating the biological function of several tissues and organs. This receptor tyrosine kinase comprises a very diverse set of signaling options to differentially modulate signaling according to the cellular context. It is known to perform bidirectional ligand-dependent signaling and ligand-independent signaling.[13] Further modulation of EPHA2 signaling is provided by crosstalk with other signaling pathways[40] or differential receptor processing by proteases or phosphatases. [29]

Ligand-dependent bidirectional signaling. EPHA2 preferentially interacts with ephrin-A1 and upon ligand binding, signaling can occur either in one direction into the receptor-bearing cell (forward signaling), in the ligand-associated cell (reverse signaling) or bidirectional into both cells (Figure 3).[13,41] If the same cell carries both receptor and ligand, signaling can be activated in parallel (receptor- and ligand-mediated signaling in same cell) or anti-parallel fashion (receptor signaling in both neighboring cells).[13] Upon ligand stimulation, EPHA2 receptors form dimers and eventually larger heteromeric clusters also involving other EPH receptors. This leads to autophosphorylation of the residues Y588 and Y594 in the juxtamembrane domain and the activation loop tyrosine Y722.[42] The associated conformational changes induce kinase catalytic activity leading to phosphorylation of the interacting receptors and downstream signaling molecules such as the protein kinase PTK2. Activation of the SRC/PTK2 complex results in RHO-dependent cytoskeletal rearrangement which reduces adhesive properties, elicits cell rounding and inhibits cell spreading.[43] Additionally, the intracellular SAM domain associates with different proteins such as GRB7[44] regulating cell migration or the lipid phosphatase SHIP2[45] leading to decreased EPHA2 kinase activity and endocytosis. EPHA2 was found to retain signaling also in endosomes where it associates with the RHO family guanine nucleotide exchange factor (GEF) TIAM1.[46] This in turn induces RAC1 activity which probably regulates endosomal trafficking, receptor recycling and degradation.

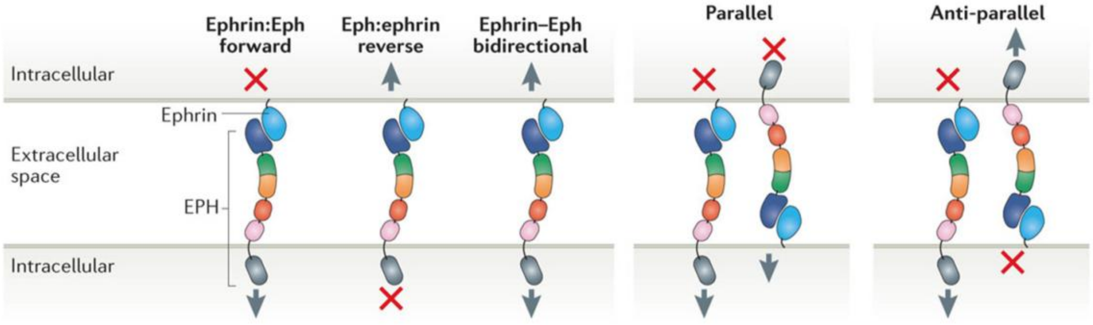


Figure 3 | Ligand-dependent bidirectional signaling. Upon ligand binding, EPH:ephrin signaling occurs in the receptor-associated cell (forward signaling), the ligand-associated cell (reverse signaling) or in both directions (bidirectional signaling). In cells carrying both receptor and ligand, signaling can be transduced in parallel (receptor- and ligand-mediated signaling in one of the interacting cells) or anti-parallel manner (receptor-mediated signaling in both cells). Figure adapted from Kania *et al*.[13]

Ligand-independent signaling. In the absence of ephrin-A1 (Figure 4a), receptor uptake and processing cannot take place leading to an overexpression and overrepresentation of EPHA2 at the membrane.[47] This state is distinctive of aberrant oncogenic EPHA2 signaling and characterized by abolished activation loop Y772 phosphorylation and elevated levels of S897 phosphorylation. Ligand-independent signaling involves a reciprocal regulatory loop with the oncogenic protein kinase AKT leading to phosphorylation of S897 and concomitant phosphorylation of the AKT activatory sites S473 and T308.[48] EPHA2 ligand-independent signaling via the PI3K/AKT signaling cascade mediates the assembly of the actin cytoskeleton at the leading edge of migrating cells which is the major mechanism responsible for the ligand-independent stimulation of cell migration and invasion.[48] Depending on the presence or absence of ephrin-A1 ligand, EPHA2 can be both: AKT substrate (ligand-independent signaling) and negative regulator of AKT (ligand-dependent signaling) (Figure 4b).[48] Ligand stimulation of EPHA2 with ephrin-A1 completely reverses this picture and leads to phosphorylation of the characteristic ligand-dependent phosphorylation site Y772, dephosphorylation of S897 and inactivation of AKT signaling by recruitment of phosphatases that inactivate AKT through dephosphorylation.[48,49] This binary function of S897 is uniquely associated to EPHA2 – no other RTK shows similar association of this phosphorylation site and cancer.[23] Contrarily, Zhou *et al*[50] found that pS897 is not mediated by AKT but by RSK signaling. It remains to be fully investigated whether both pathways stimulate pS897 in a concerted action or if these are mutually exclusive processes.

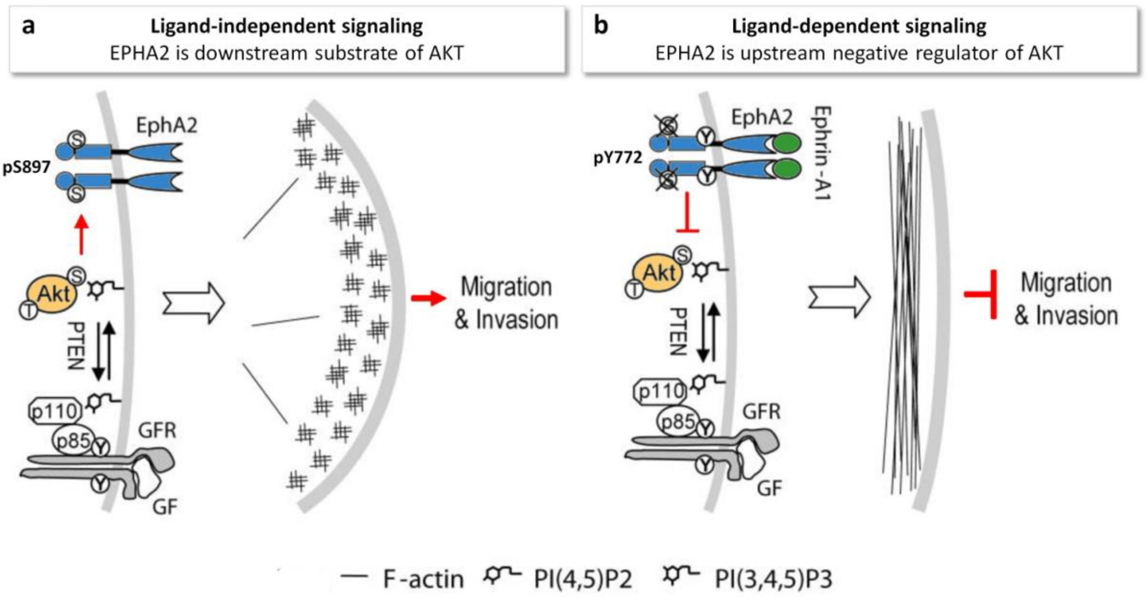


Figure 4 | EPHA2 ligand-dependent and -independent signaling. a) Overexpression of EPHA2 in absence of ligand leads to S897 phosphorylation by AKT and a migratory and invasive phenotype. Additionally, EPHA2 performs PTEN-mediated cross-talk with growth factor receptor (GFR) signaling and thus, results in increasing cell proliferation. Modified from Miao *et al.*[48] **b)** EPHA2:ephrin-A1 interaction inhibits oncogenic AKT/PI3K and GFR-driven activation of MAPK/ERK signaling. This leads to tumor-suppressive inhibition of migration and invasion.

Crosstalk with other signaling pathways. It was found that EPH receptors are able to perform pronounced cross-talk not only with PI3K/AKT signaling but also with other signaling pathways, such as RAS/ERK signaling, MAPK signaling, or WNT signaling, and to interact with many other proteins such as pY-binding adaptor proteins, PDZ domain containing proteins or modulators of the RAS and RHO family of GTPases.[4,40] EPHA2 can function in an oncogenic or tumor-suppressive manner which is partly also mediated by crosstalk with signaling pathways such as growth factor receptor (GFR) signaling, and downstream PI3K/AKT or MAPK signaling (Figure 4a,b).[40] For instance, it was shown that EPHA2 ligand-independent signaling augments tumor-promoting ERBB activity[51] involving the RAS/ERK signaling pathway, which in turn leads to an up-regulation of EPHA2 gene expression by a positive feedback loop.[52] Ligand-dependent signaling suppresses the effect of ERBB signaling by recruitment of RASA1 (p120 RAS-GAP).[53] This subsequently counteracts and reduces the ERBB-mediated activation of RAS-GTP. Thus, the activation of RAS/ERK and PI3K/AKT oncogenic pathways by growth factor receptors can be inhibited by stimulation of ligand-dependent EPHA2 signaling.[54] EPHA2 was also identified as a key downstream target of the MEK/ERK/RSK signaling pathway stimulated by EGF. Ligand-independent signaling is augmented upon EGF treatment and S897 phosphorylation is probably induced by MEK and RSK leading to increased cell proliferation.[55] This is surprising since a potential proliferative effect of EPH receptors is contradictorily debated in the field and EPHA2 was the only EPH receptor shown so far to influence cell proliferation. Recently, EPHA2 was also identified to be the corresponding receptor for the growth factor progranulin, which causes receptor activation and downstream signaling via MAPK pathway and AKT.[56,57]

Receptor shedding. Depending on the cellular background, the same EPH:ephrin interaction partners might drive cell-cell repulsion in one context and adhesion to cell or substrate in another context – also depending on extracellular proteolytic activity of proteases such as ADAM10.[29] As already described earlier, receptor shedding of the extracellular components of the EPH:ephrin signaling clusters is required to release the cell-cell contact and to perform cell-cell repulsion. EPHA2:ephrin-A1 signaling clusters are cleaved by ADAM10 leading to subsequent receptor internalization by clathrin-mediated endocytosis.[58] Contrarily, shedding of the extracellular ligand binding domain by another protease, the matrix metalloproteinase MMP14, converts EPHA2 to a ligand-insensitive form of the receptor leading to a phenotypic switch from ligand-dependent (tumor-suppressive) to ligand-independent (oncogenic) signaling.[53]

Phosphatases. Further modulation of EPHA2 signaling is provided by the activity of phosphatases such as the low molecular-weight tyrosine phosphatase ACP1. This phosphatase downregulates tyrosine phosphorylation of EPHA2[59] and, thus, negatively influences and terminates ephrin-A1 mediated signaling. This phosphatase is frequently upregulated in tumors and leads to tumor onset, tumor growth and a migratory phenotype.[60-62] Interestingly, another well-characterized oncoprotein, the phosphatase PTEN, is also involved in EPH signaling and comprises dual function in being an EPHA2 substrate and in dephosphorylating EPHA2 in a unique reciprocal feedback loop mechanism.[63]

1.3 Role of EPHA2 in biology and disease

As already discussed, the complex EPHA2 signaling events are governed not only by the spatial organization of receptor and concentrations of different ligands but also by other modulating factors such as phosphatase or protease activity. According to this, EPHA2 plays pivotal and diverse roles in a wide range of cellular processes and tissues. EPHA2 is mainly expressed in epithelial cells, e.g. in skin, lens, kidneys, lungs, liver, small intestine or colon,[42] where it regulates cell proliferation, differentiation, migration and tissue morphogenesis.[64] EPHA2 is involved in maintaining lens transparency and shape by controlling the organization of lens fiber cells.[65,66] Recently, it was proposed that EPHA2 influences lens fiber cell differentiation by cross-talk and modulation of the FGFR signaling pathway.[67] Interaction of EPHA2 and ephrin-A2 was found to act on osteoblast and osteoclast activity and negatively regulate bone formation.[68] In the skin, EPHA2 is implicated in epidermal differentiation and homeostasis of keratinocytes.[69] EPHA2 was also found to be expressed in cardiac progenitor cells and ephrin-A1 is found in cardiomyocytes functioning as niche cells. Here, EPHA2 signaling mediates cardiomyocyte protection and heart repair by migration of progenitor cells to the damaged myocardium.[70] Similarly, EPHA2 ligand-dependent signaling plays a role in the adult cardiovascular system[21] regulating angiogenesis and adult angiogenic remodeling.[71]

Owing to this complex and context-dependent role of the EPHA2 receptor and its contribution to tissue formation and homeostasis, it is obvious that dysregulation of EPHA2 can be associated with the formation and progression of many severe pathologies, such as lens cataract, pathogen infection or cancer.

Lens cataract. Inherited congenital cataract is a major reason of childhood blindness and can be caused by mutations in the EPHA2 gene.[72,73] Several point and frameshift mutations affecting the juxtamembrane domain (P584L), the SAM domain (T940I, D942fsXC71, A959T) or the PDZ domain (V972GfsX39) are implicated in the formation of lens cataract. SAM domain mutations result in decreased receptor stability and mislocalization within the cell, which infers alterations of folding or glycosylation state of the receptor. Other mutated forms of EPHA2 exert their pathological effect via aberrant signaling or disrupted interactions with other proteins.[72,73]

Pathogen infection. Successful infection of a host organism relies on several factors such as cell adherence and endocytosis, productive infection and replication, but also on the formation of a conducive environment by subverting the host response and achieving immune evasion. Interestingly, EPHA2 can play a role in all of these steps and thus, influences the outcome of viral, bacterial and parasitic infectious diseases at different levels. In 2011, Lupberger *et al*[74] were the first ones to show that EPHA2 can promote cell entry of *hepatitis C virus*.[75] Similarly, members of the *herpesviridae* family such as the *Kaposi's sarcoma-associated herpesvirus* were shown to invade the host cell by interaction and activation of EPHA2 leading to cellular uptake by EPHA2-regulated macropinocytosis.[76,77] Another example of how EPHA2 signaling may be exploited is given by enteropathogenic *Escherichia coli* strains (EPEC) which secrete effector proteins into host cells that lead to altered EPHA2 phosphorylation and MAPK signaling – implicating a pivotal role for these signaling pathways during host cell response and productive EPEC infection.[78] Intracellular bacterial pathogens such as *Chlamydia trachomatis* exploit EPHA2 as adherence/invasion mechanism and intracellular signaling receptor.[79] After cellular uptake, *C. trachomatis* bacteria are enclosed into a membrane-bound vacuole called "inclusion", where EPHA2 accumulates and triggers the PI3K/AKT signaling pathway which is essential for chlamydial development and replication. Other bacterial pathogens use EPHA2 signaling to modulate the host cell immune response in order to accomplish persistent infection. One striking example is the infection with *Mycobacterium tuberculosis* where EPHA2 is implicated in driving immune evasion – being one key virulence mechanism in persisting *M. tuberculosis* infection.[80] Similarly to viral or bacterial pathogens, EPHA2 is also involved in the infectious cycle of the malaria parasite *Plasmodium falciparum* by supporting immune evasion and productive infection in an early stage of malaria infection.[81] EPHA2 was found to be engaged by sporozoite proteins and thus likely plays a key role in establishing a permissive replication compartment and disease progress. Several of the mentioned studies showed that EPHA2 directed antibodies as well as small molecule inhibitors (e.g. Dasatinib) are suitable to abrogate EPHA2 function during pathogen infection[74,76,82] – opening a novel perspective on EPHA2 as potential target for therapeutic intervention.

Cancer. When EPHA2 was first reported and characterized in 1990,[83] Lindberg and Hunter already proposed a potential role of EPHA2 as oncoprotein. The following three decades of EPHA2 research have proven this initial hypothesis to be true. EPHA2 expression in adult healthy tissue is rather low while its overexpression and dysregulation have been strongly correlated to oncogenesis, metastasis, angiogenesis, resistance and poor clinical prognosis (Figure 5).[42,47] Despite the sustained efforts to interrogate EPHA2 function, the role of EPHA2 in cancer remains confusing: expression levels can be very heterogeneous, downstream signaling can be dominated by ligand-dependent and -independent signaling and depending on the cellular context EPHA2 can mediate tumor-suppressive or pro-oncogenic functions.[40,84] Moreover, it becomes more and more evident that the impact of EPH signaling has to be considered as a concerted action of the whole EPH family and not as stemming from individual enzymes.[85] Overexpression of EPHA2 has frequently been associated with tumor malignancy and poor prognosis in a large variety of cancer entities such as squamous head and neck cancer, glioblastoma, breast cancer, lung cancer, ovarian cancer, prostate cancer and many more.[47,86] High EPHA2 expression levels often correlate with decreased expression of the ligand ephrin-A1 and low Y772 / high S897 phosphorylation[21,48] – suggesting the non-canonical ligand-independent signaling pathway to be active (Figure 5). As discussed earlier, ligand-independent signaling involves crosstalk with the PI3K/AKT signaling pathway. This interplay seems to elicit the progression to more malignant and aggressive tumor phenotypes since the characteristic pS897 phosphosite is highly elevated in most aggressive tumors and particularly in the putative tumor stem cell population.[40,48,87] Contrarily, it was found that EPHA2 is capable of forming dimers even in the absence of ligands;[88] thereby downregulating the oncogenic S897 phosphosite and tumorigenic signaling. In this case, EPHA2 overexpression potentially leads to ligand-independent stimulation of the downstream effector protein PTK2[43] which is characteristic for ligand-dependent and tumor-suppressive signaling.

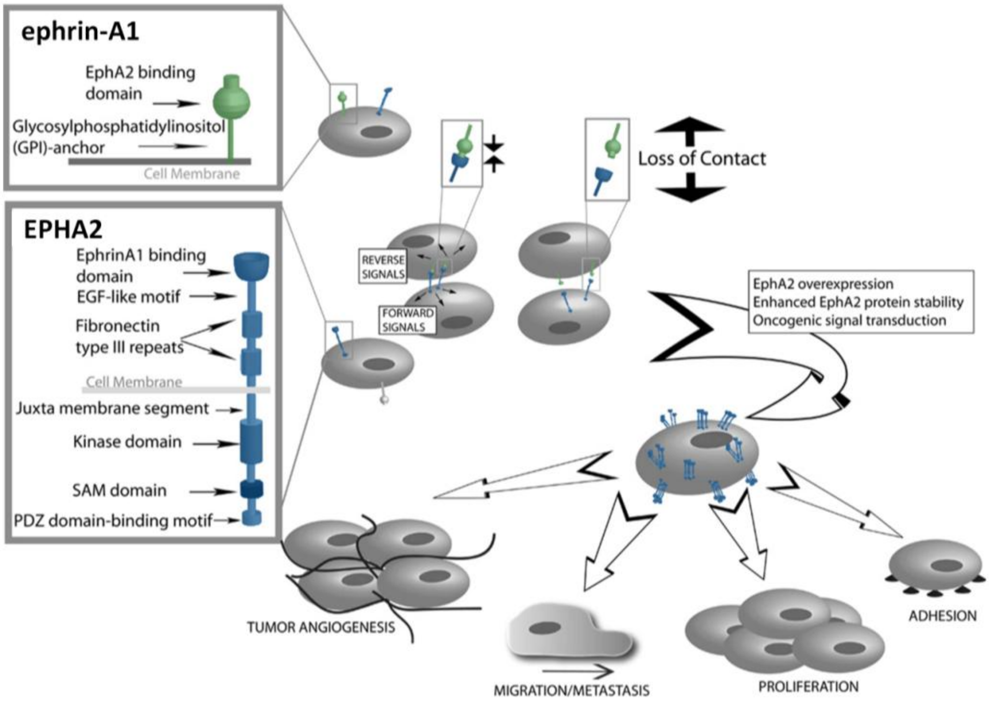

ephrin-A1
EphA2 binding domain
Glycosylphosphatidylinositol (GPI)-anchor
Cell Membrane
EPHA2
EphrinA1 binding domain
EGF-like motif
Fibronectin type III repeats
Cell Membrane
Juxta membrane segment
Kinase domain
SAM domain
PDZ domain-binding motif
REVERSE SIGNALS
FORWARD SIGNALS
Loss of Contact
EphA2 overexpression
Enhanced EphA2 protein stability
Oncogenic signal transduction
TUMOR ANGIOGENESIS
MIGRATION/METASTASIS
PROLIFERATION
ADHESION

Figure 5 | Implications of EPHA2 in oncogenesis and cancer progression. EPHA2/ephrin-A1 interaction results in non-pathologic reverse and forward signaling in the neighboring cells. Upon loss of contact to the ligand, EPHA2 is overrepresented on the membrane which leads to a switch to oncogenic signaling and results in increased tumor angiogenesis, a migratory and metastatic phenotype, upregulated proliferation and a change in the adhesive properties of cancer cells. Modified from Tandon *et al*.[47]

In addition to the expression level, also the mutational status of EPHA2 plays a pivotal role in conveying a more aggressive cancer phenotype.[89] As an example, the G391R mutation prevents proteolytic cleavage of the EPH:ephrin complex by the matrix metalloprotease MMP14.[90] This leads to an increase of EPHA2 at the cell membrane and promotes an invasive and anchorage-independent tumor type.[91] Moreover, EPHA2 was also linked to angiogenesis and vasculogenic mimicry which is the *de novo* formation of highly patterned microvascular channels by tumor cells.[92-94] Both contribute to adequate nourishment of the tumor tissue; thus, implying an important role of EPHA2 not only in tumorigenesis but also in sustenance of the tumor. Blood perfusion by angiogenic vessels and vasculogenic mimicry are also key requirements for the ability of tumors to spread tumor cells to other organs. Both, the formation of vascular mimicry as well as invasion of the neighboring tissue rely on a phenotype switch called epithelial-mesenchymal-transition.[95] During this process, single cells release their connectivity to the surrounding tumor tissue and matrix and change their properties towards a migrative phenotype. Upon expression of additional extracellular proteases, the cells become invasive and capable of entering the blood stream and forming metastatic sites in other organs remote from the primary tumor. Migration and invasion depend on S897 phosphorylation (ligand-independent EPHA2 signaling) and its effect on cell repulsion and the actin cytoskeleton.[40,48] Recently, it was shown that the RAB-coupling protein RCP is phosphorylated at S435 by LMTK3 and in conjunction with the RAS-related protein RAB14 controls EPHA2 trafficking and endocytosis.[96] Thus, EPHA2 and RCP cooperate in RAB14-dependent cell-cell repulsion and tumor dissemination driving the metastatic potential of tumor cells. Another very important angle of EPHA2 oncogenic function is its capability of promoting therapy resistance by the reprogramming of cell signaling. So far, this was mainly shown for antibody-based[97] and small molecule-based[98,99] HER2/EGFR-targeted therapies. EGFR inhibition, e.g. by the small molecule kinase inhibitor Gefitinib, was overcome by overexpression of EPHA2 leading to Gefitinib-resistant cells. Translational or pharmacological inhibition of EPHA2 restored Gefitinib sensitivity. Similarly, EPHA2 signaling was altered in response to BRAF inhibitor resistance.[100] In melanoma patients, ligand-independent EPHA2 signaling was shown to be responsible for the adaption of a targeted therapy-mediated metastatic phenotype. Along these lines, EPHA2 also promotes resistance against the anti-estrogen receptor drug Tamoxifen, probably through decreased activation of a PTK2 dependent pathway.[101] Another highly debated mechanism leading to therapy resistance is the existence of chemoresistant cancer stem cells.[102] In glioblastoma cell lines, EPHA2 seems to be implicated in the formation of highly malignant glioblastoma by promoting self-renewal of stem-cell like cells and by regulating stem-cell marker expression.[87,103]

1.4 Different molecular strategies are used to target EPHA2

As discussed above, EPHA2 signaling is implicated in plenty of biological processes as well as disease onset and progression. Hence, molecularly targeted therapies against EPHA2 seem promising and were proven beneficial in several pathological conditions, as already described in the previous chapter. However, EPHA2-targeted treatment may also result in on-target toxicities that should be considered carefully, especially during clinical evaluation of targeted therapies. Potential toxicities could include changes in bone homeostasis, immune function, pain sensing, memory and cognition.[21] Due to the high complexity of EPH signaling and the manifold interactions of different EPH:ephrin systems, it is conceivable that compensatory mechanisms engaging other EPH family members might prevent the severe abolition of important physiological processes driven by EPHA2. In certain cases, however, on-target side effects might also be advantageous for the clinical outcome; especially if EPHA2 inhibition leads to decreasing inflammation, atherosclerosis or pain during therapy. [21]

According to their complex structure comprising different domains and interaction sites, diverse strategies have evolved to target EPH receptors.[5,21,47,104,105] Some approaches exploit EPH receptor derived peptides for the generation of cancer vaccines and immunotherapy.[5] Other strategies include siRNA / antisense nucleotides that lead to transcriptional downregulation of EPH receptor or ligand. However, most EPH targeting agents act on the extracellular ligand binding domain eliciting either agonistic or antagonistic effects on EPH signaling:[105] Single or multimeric recombinant extracellular domains, antibodies, peptide agonists and antagonists, and small molecule agonists and antagonists.[106-108] In contrast, small molecule kinase inhibitors target EPH receptors intracellularly. The development of specific EPH receptor targeting small molecules is still in its infancy – only few molecules have been designed for EPH receptor inhibition such as NVP-BHG712 (EPHB4) or ALW-II-49-7 (EPHB2).[109] However, several kinase inhibitors are known to inhibit EPH receptor kinases as off-targets (e.g. Dasatinib).

Efforts to target EPHA2 encompass all these promising strategies as detailed hereafter (Figure 6).

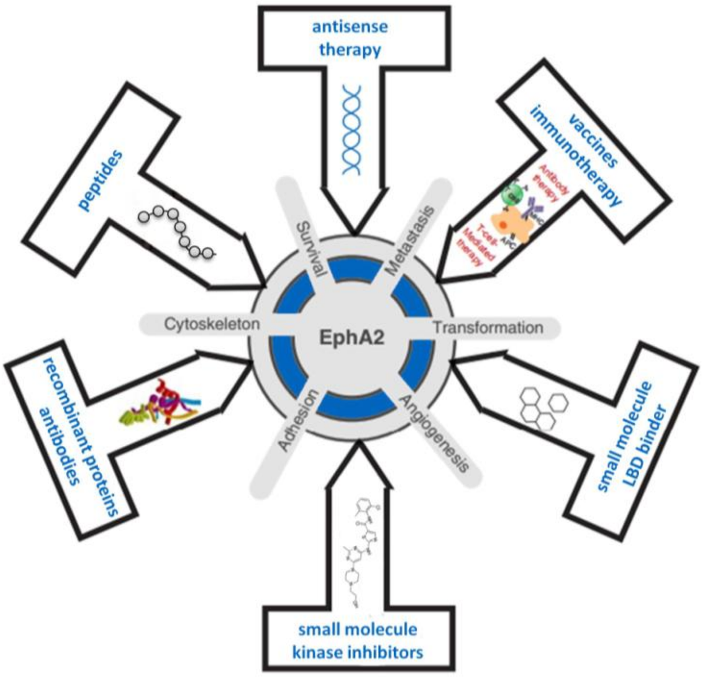


Figure 6 | Different strategies for targeting EPHA2. EPHA2 is involved in different hallmarks of cancer and can be targeted by different molecular strategies. Modified from Tandon *et al*[47].

Vaccines and immunotherapy. EPHA2 was identified as tumor-associated antigen in cancer, e.g. glioma[110] – providing the possibility of treating cancer with EPHA2-targeted immunotherapy. Hence, EPHA2-derived peptides have been utilized for vaccination in high-grade glioma patients in a phase I clinical trial.[111] The investigated peptide cocktail showed acceptable safety and modest clinical activity, thus, leaving room for further improvement. A similar study will soon enter a phase 2 clinical trial (NCT02754362). Another approach was the generation of dendritic cell vaccines, generated by pulses of EPHA2-derived peptides which are subsequently administered to mice.[112] This immunization led to the development of cytotoxic CD8$^+$ T cells specific for EPHA2 derived peptides and decreased tumor growth. Exploiting this strategy, a phase 2 clinical study was set up using EPHA2-specific dendritic cells (NCT01876212).[110,112] The concomitant administration of Dasatinib in this study should lead to degradation of EPHA2 and boost the presentation of EPHA2-derived antigens on the tumor cell surface, thus increasing the immune response and efficacy of the dendritic cell vaccine. Another novel strategy to specifically modulate the immune system towards tumor clearance is the generation of T cells expressing chimeric antigen receptors (CARs).[113] Immunization is performed using T cells that are genetically modified to produce a chimeric receptor recognizing EPHA2 as an epitope. In a preclinical setting, EPHA2 CARs have already been successfully used to kill EPHA2-positive glioma cells and glioma xenografts in mice and are currently investigated in a phase 1/2 clinical trial in EPHA2-positive malignant glioma (NCT02575261).[113] The clinical outcome of EPHA2-directed immunotherapy remains to be fully evaluated, but given the fact that some cancer entities – and especially glioma – are very hard to treat by conventional methods, even the modest clinical results obtained so far are of high value and promising novel therapy strategies.

Antisense therapy. Transcriptional downregulation of protein expression using siRNA or antisense oligonucleotides is regularly used in research to interrogate the function of a protein of interest. Given its high selectivity and efficiency this approach also entered the clinical research. However, efficient protein knockdown in patients mainly suffers from insufficient delivery *in vivo*. This can be overcome by novel drug delivery approaches such as nanoliposomes, nanocarriers, micelles or nanobubbles. One such approach (EPHARNA) targeting EPHA2 using

1,2-dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC) containing nanoliposomes recently entered clinical investigation in a phase 1 trial (NCT01591356).[114]

Antibodies. Antibodies comprise the advantage of binding the epitope of interest with high selectivity and affinity. Additionally, half-life in the living organism is known to be relatively high. However, production of antibody therapeutics is very tedious and cost-intensive and they have to be administered intravenously. Due to the outstanding role of EPHA2 in cancer formation and progression, it was the first EPH receptor to be considered for antibody development.[115] Several EPHA2 agonistic antibodies have already been discovered to activate anti-oncogenic signaling and degradation, but show contradictory results in preclinical and clinical evaluation.[116-119] However, combination treatment of EPHA2 targeting antibodies with rather unspecific chemotherapeutics such as tamoxifen, paclitaxel or docetaxel seems to be promising.[101,120] Another example, the agonistic antibody 1C1, did not affect cancer growth despite maintaining the expected inducing effects on EPHA2 ligand-dependent signaling.[105,121] This antibody was the first one to be used as drug delivery agent in clinical trials as 1C1-auristatin drug conjugate (MEDI-547), but failed in early phase 1 due to bleeding and coagulation (NCT00796055).[122] So far, three clinical trials have been started to examine the clinical outcome of EPHA2 targeted antibody treatment in patients using the monoclonal antibody DS-8895a (NCT02252211), the antibody-drug conjugate MEDI-547 (NCT00796055) and antibody-targeted nanoliposomes delivering a Docetaxel prodrug (NCT03076372).

Recombinant soluble receptor or ligand extracellular domains. Soluble EPHA2-Fc and ephrin-A1-Fc have been used extensively to investigate EPHA2 biological function.[123,124] It was shown that these soluble surrogates are capable of inhibiting proliferation, neovascularization and invasion in several tumor models. [5] Ephrin-A1 ligands were also coupled to cytotoxic drugs – to exploit them as molecularly targeted therapy and as drug delivery approach – and were shown to affect glioblastoma, breast and prostate cancer cell lines.[125]

Peptides. Peptidic agents occupy the interface of biologicals and small molecules and show characteristics of both areas. They are usually very selective, affine and easy to produce, whereas bioavailability and half-life *in vivo* can be drastically reduced due to proteolytic activity and absorption by plasma proteins. However, peptides provide the possibility to further enhance binding affinity, proteolytic stability and half-life[21] by affinity maturation, cyclization or inclusion of unnatural amino acids. Phage display identified EPHA2 binding peptides that promote receptor activation and internalization in an agonistic mechanism; this is surprising since other EPH receptor targeting peptides mostly act as antagonists.[126,127]

Small molecule ligand-binding domain interactors. Disruption of the protein-protein interaction between EPH receptors and their ligands is another interesting concept of targeting EPH signaling.[128-130] Due to the large size and flexibility of the EPH ligand binding domain, these antagonists often suffer from low binding affinities and low selectivity. This strategy has led to the discovery of several molecules and further optimization successfully increased binding affinities and cellular potencies. The most researched scaffold is lithocholic acid, but also 2,5-dimethylpyrrolyl benzoic acid derivatives,[107] disalicylic acid-furanyl derivatives[106] and urolithin D[131] have shown some promising effects. Derivatives of lithocholic acids were explored and conjugation to amino acids finally resulted in the discovery of UniPR129, the L-homo-Trp conjugate of lithocholic acid.[132-134] This compound efficiently disrupted protein:ligand interaction with relatively high potency (370 nM) and an inhibitory effect was confirmed with micromolar potency in functional cellular assays.

Small molecule kinase inhibitors. Small molecule kinase inhibitors have a successful history in the clinics and provide the possibility of comparably easy and cheap production as well as the potential for oral bioavailability. However, their main disadvantage is poor selectivity since most of the inhibitors target the highly conserved ATP binding pocket. For EPHA2, no dedicated inhibitors had been described in the beginning of this work.

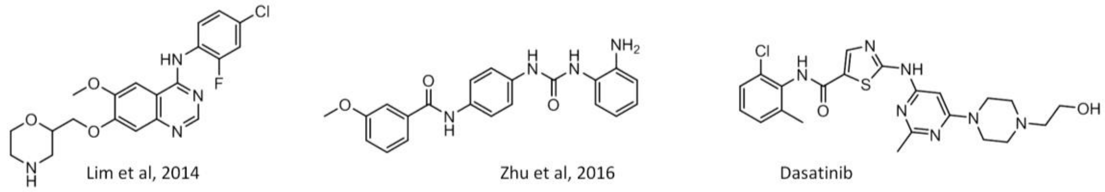

Lim et al, 2014
Zhu et al, 2016
Dasatinib

Figure 7 | Targeting EPHA2 using small molecule kinase inhibitors. So far, only few small molecule kinase inhibitors were developed to specifically target EPHA2. The approved dual SRC/ABL inhibitor Dasatinib was found to potently inhibit EPHA2 as an off-target.

Since then, Lim *et al*[135] built inhibitors based on a quinazoline scaffold and reached EPHA2 inhibitory activity of 160 nM (Figure 7). The chemical scaffold utilized in this study is very similar to known EGFR inhibitors such as Gefitinib or Erlotinib. Consequently, EGFR is one of the most potently hit off-targets of these inhibitors (100% inhibition at 1 µM) followed by KDR (76%), HER2 (22%), PDGFRA (17%), MAPK14 (16%), GSK3B (11%), IGF1R (7%) and MAPK11 (2%); a comprehensive target profile is not available yet. In another study by Zhu *et al,*[136] 1-(2-aminophenyl)-3-arylurea derivatives have been evaluated for their dual inhibition of EPHA2 and HDAC (Figure 7). These compounds were found to inhibit EPHA2 with modest affinity (0.5 – 9.3 µM) and showed antiproliferative effects in the micromolar range in HCT116, K562 and MCF7 cells. In most studies, pharmacological EPHA2 inhibition is obtained by the designated BCR-ABL/SRC inhibitor Dasatinib which is known to very potently inhibit EPHA2.[137] Dasatinib is a very promiscuous kinase inhibitor targeting a whole range of different protein kinases – rendering the usefulness of this kinase inhibitor for EPHA2 targeted inhibition questionable. In addition to Dasatinib, more kinase inhibitors were found to inhibit EPHA2 as an off-target but were not as broadly used as Dasatinib in EPHA2 research yet. These include but are not limited to other BCR-ABL or SRC inhibitors (e.g. Ponatinib,[138] Bosutinib,[139] Nilotinib,[139] Saracatinib[140]), VEGFR family inhibitors (e.g. Foretinib,[139] Tivozanib,[141] Regorafenib[142]), AURK inhibitors (e.g. Alisertib,[143] MLN-8054[139]), or ALK inhibitors (e.g. Crizotinib[139]). Applicability of small molecule kinase inhibitors against EPHA2 pro-oncogenic signaling is highly debated. Miao *et al*[48] reported that pro-oncogenic signaling of EPHA2 was independent from catalytic kinase activity. Other research groups, however, showed that kinase inhibitor treatment diminishes phosphorylation levels of EPHA2 S897 and AKT S473.[144] The response to small molecule kinase inhibitors might involve an indirect mechanism including cross-talk signaling events between EPHA2, BRAF/CRAF and CAV-1.[144] Although not fully validated, pharmacological EPHA2 inhibition has been the object of several clinical trials. These include administration of Dasatinib either as single (NCT01440998, NCT00563290, NCT02693535) or combination treatment with radiation therapy (NCT02661113, NCT00895960).

2 Small molecule kinase inhibitors targeting the kinase domain

EPHA2 is a rather novel target and no designated EPHA2 inhibitor had been developed nor entered clinical trials when this project was started – despite the fact that small molecule kinases inhibitors have experienced huge interest during the last two decades.[145] Since the approval of the BCR-ABL inhibitor Imatinib by the Food and Drug Administration (FDA) in 2001, more and more kinase inhibitors for different kinase targets were discovered and succeeded in clinical application. Currently, 37 kinase inhibitors are FDA-approved and routinely used in cancer patients; more than 200 kinase inhibitors are evaluated in different phases of clinical trials. The major pitfall of kinase inhibitors is their selectivity.[146] The majority of these molecules directly target the ATP binding pocket which is highly conserved among protein kinases and also other ATP binding proteins – rendering the elucidation of small molecule kinase inhibitor selectivity an important task to understand mode of action and clinical effects of these drugs.

2.1 Catalytic activity and structure of the protein kinase domain

Catalytic activity. The enzymatic reaction performed by protein kinases is the phosphorylation of substrate proteins at specific serine, threonine or tyrosine residues. Protein phosphorylation is broadly used by the cell in forwarding external and internal stimuli via highly complex and intertwined signaling cascades – ultimately leading to the downstream activation of effector proteins or protein expression to perform a certain cellular answer.[147] Tight interplay between kinases and phosphatases catalyzing dephosphorylation is required to fine-tune this multilayered signaling network and to control the biological outcome (Figure 8). Protein phosphorylation evolved to one of the key posttranslational modifications exploited by nature due to its versatile effects and reversibility. It introduces a negative charge to a substrate protein which is responsible for conformational changes or provides new interaction sites for protein-protein interactions.[147] Phosphorylation is the formation of a phosphate ester linkage by a nucleophilic substitution reaction (Figure 8b). The phospho-acceptor hydroxy amino acid side chain is activated by deprotonation creating a strong nucleophile that attacks the phospho-donor adenosine triphosphate (ATP) at the terminal γ-phosphate. Adenosine diphosphate (ADP) is released as stable leaving group.[147] ATP serves as cellular energy storage molecule (2 to 4 mM cellular concentration) and comprises three phosphates linked to each other in form of a phosphate anhydride. During phosphorylation, the terminal anhydride bond between the β-phosphate and the γ-phosphate is cleaved releasing approximately 8-12 kcal/mol of free energy which is the driving force of this reaction.[147] Nonetheless, protein phosphorylation requires the action of kinases catalyzing this reaction in a cellular context.

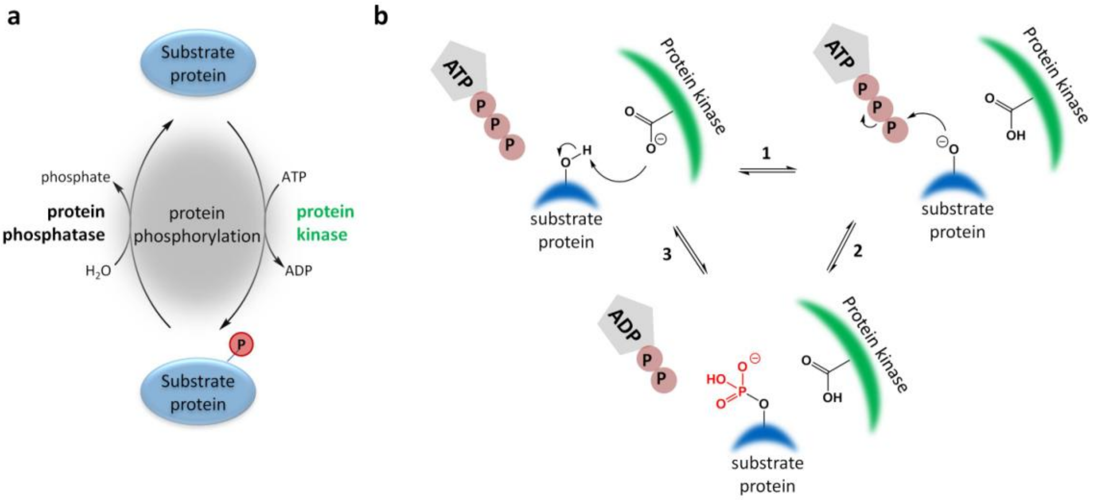


Figure 8 | Catalytic activity of protein kinases. a) Protein phosphorylation is conducted by protein kinases under consumption of the phosphate donor ATP. This process can be reversed by the activity of protein phosphatases under generation of free phosphate. **b)** Base-catalyzed nucleophilic attack of a hydroxy amino acid substrate residue on the γ-phosphate group of ATP: (1) substrate deprotonation by aspartate (HRD motif), (2) nucleophilic substitution reaction, (3) recovery of active center by deprotonation of catalytic Asp and release of products.

Catalytic action of protein kinases requires the presence of certain structural key elements. Typically, protein kinase domains are structured into two major lobes, the N-terminal N-lobe (V613-I691 in EPHA2) and the C-terminal C-lobe (L700-I875 in EPHA2) which are connected via the flexible hinge region (T692-A699 in EPHA2).[148] The interface between both lobes features the ATP binding pocket and the substrate binding site and thus, the catalytic center of this enzyme. This structural composition is highly dynamic and allows the formation of multiple conformations – each of them performing a certain function during the catalytic process.[149] The highest degree of structural flexibility is usually obtained in the inactivated state. Activation of the kinase domain by phosphorylation of the activation loop provides a first mean of stabilization and is required for the binding of ATP and the substrate protein in many kinases.[150] During the catalytic cycle, the kinase domain executes a typical movement known as kinase domain “breathing”, finally releasing ADP and the phosphorylated substrate protein.[148]

Structural insights. Key requirement of the active kinase conformation is the formation of a hydrophobic spine connecting N- and C-lobe and thus stabilizing the kinase structure towards an active conformation (Figure 9a). [148,150,151] This regulatory spine (R-spine) is comprised of four hydrophobic residues located in the catalytic loop (RS1: Tyr735 in EPHA2), the DFG motif (RS2: Phe758 in EPHA2), the αC-helix (RS3: Met667 in EPHA2) and at the beginning of the β4-sheet (RS4: Leu678 in EPHA2). The kinase is only capable of performing its enzymatic action when this structural feature is fully established. In spite of the common opinion, activation loop phosphorylation is not as essential for kinase activation as previously anticipated.[152] It has surprisingly different importance for the catalytic activity even of structurally similar kinases. Upon phosphorylation of the activation loop (Tyr772 in EPHA2), the negatively charged phosphate interacts with an arginine located in the HRD motif in the catalytic loop and thus, performs a conformational change establishing the hydrophobic R-spine. Thus, activation loop phosphorylation is only required in kinases comprising this arginine (RD kinases); others are not regulated by the activation loop phosphorylation.[153] The most important R-spine residue RS2 is the highly-conserved phenylalanine of the DFG-motif located at the beginning of the activation loop (D757- G759 in EPHA2). This phenylalanine is capable of executing a 180° conformational switch – shuttling between the DFG-in and DFG-out state.[2] In the inactivated conformation, this DFG motif is rotating quite flexibly, whereas the DFG-in conformation completes the R-spine and is thus characteristic for an active kinase conformation. Upon the formation of this first activated structure, ATP is capable of binding the pocket. The adenine part of ATP is accommodated in the hydrophobic nucleotide binding pocket (Figure 9b) formed mainly by the N-lobe and completes another structural element called the catalytic C-spine.[151] This leads to the recruitment of the Gly-rich loop located at the tip of an N-lobal β-sheet. The Gly-rich loop forms a "lid" to correctly position ATP within the pocket and interacts with both, the β- and γ-phosphate groups of ATP. Another highly-conserved structural feature is the AxK motif which is also found in the same β-sheet next to the Gly-rich loop (A644-K646 in EPHA2).[148] The AxK lysine engages the α- and β-phosphate groups of ATP and participates in the catalytic cycle. Additionally, it forms a characteristic salt bridge with a glutamic acid located within the αC-helix (E663 in EPHA2) and thus, also plays a critical role in stabilizing the active kinase domain.[148] The main interaction hub within the kinase domain is the aspartic acid located within the DFG motif (D757 in EPHA2). Correct formation of the R-spine and DFG-in conformation positions this aspartic acid towards the catalytic site and enables interaction with all three ATP phosphate groups – either directly or indirectly via coordination of two magnesium ions.[151] In order to fulfill the catalytic action, the substrate peptide needs to be positioned in close proximity to the γ-phosphate of ATP. Interaction with the acceptor hydroxy group is obtained by the aspartate provided from the HRD motif (D739 in EPHA2). This amino acid abstracts a proton from the protein substrate hydroxy group and thus, facilitates nucleophilic attack at the ATP γ-phosphate (Figure 8b).[154] The HRD motif fulfills several functions in linking the phospho-acceptor site (Asp, D739 in EPHA2), to the activation loop (Arg, R738 in EPHA2) and the catalytic DFG motif (His, H737 in EPHA2).[155] Additionally, a conserved lysine situated in the catalytic loop can help to engage both the γ-phosphate and the substrate hydroxy group. Only upon concerted action of all these structural elements, can the transfer of the γ-phosphate group to the substrate protein be achieved.

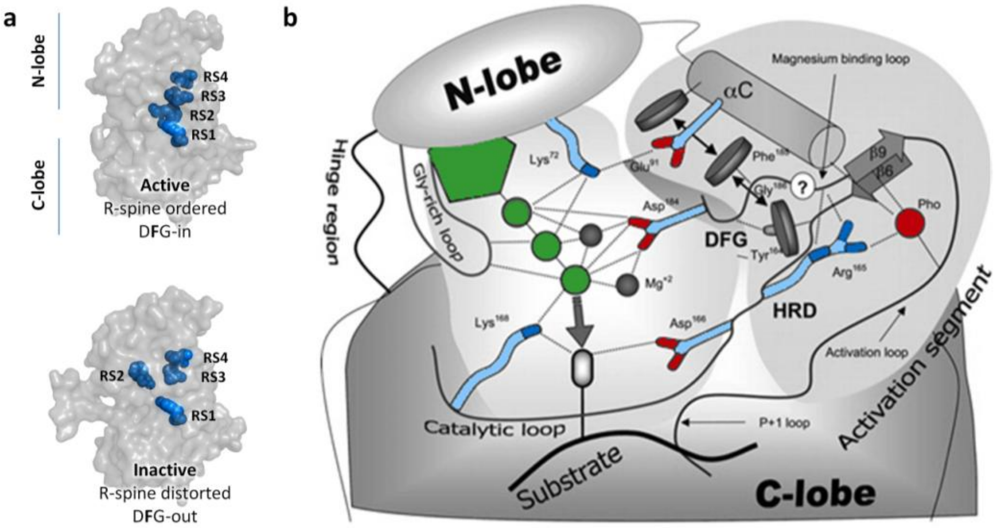
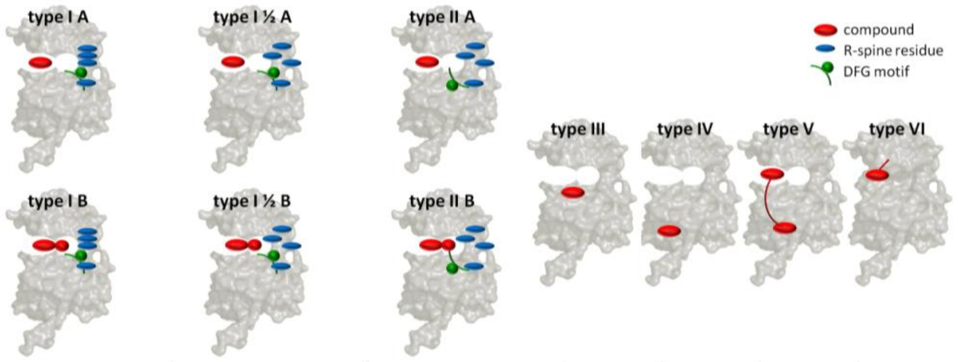

a
N-lobe
C-lobe
RS4
RS3
RS2
RS1
Active
R-spine ordered
DFG-in
RS2
RS4
RS3
RS1
Inactive
R-spine distorted
DFG-out
b
N-lobe
Hinge region
Gly-rich loop
Lys^{72}
Glu^{91}
αC
Magnesium binding loop
Phe^{185}
Gly^{186}
?
β9
β6
Pho
Asp^{184}
DFG
Tyr^{164}
Arg^{165}
Mg^{+2}
HRD
Lys^{168}
Asp^{166}
Activation loop
Activation segment
Catalytic loop
P+1 loop
Substrate
C-lobe

Figure 9 | Structural prerequisites define enzymatic activity of protein kinases. a) The protein kinase domain is structured into two major lobes, the N-terminal N-lobe and the C-terminal C-lobe. Active kinase conformation is characterized by an inwards orientation of the DFG phenylalanine (RS2) complementing and stabilizing the regulatory R-spine, whereas inactive kinase conformation comprises a distorted R-spine. **b)** Several catalytic residues and motifs within the ATP pocket are involved in binding and positioning ATP for the subsequent transfer of γ-phosphate to the substrate protein. Modified from Kornev *et al*[148].

Another important aspect of inhibitor binding is the role of regulatory domains or interaction partners that may influence binding affinity as well as binding kinetics.[156] For example, receptor tyrosine kinases comprise a juxtamembrane segment which can have autoinhibitory effects on the catalytic function (e.g. in EPH,[157] PDGF,[158] or VEGF receptors[156,159]). Interestingly, the presence of juxtamembrane domain was shown to increase the affinity of several kinase inhibitors most probably by favorable entropic contribution.[159,160] Other examples for regulation through other domains include SH2 domains (e.g. in SRC, ABL, HCK)[161,162] or pseudokinase domains (e.g. in JAK family)[163,164]. In case of the CDK family of kinases, the formation of a CDK-cyclin complex is necessary to establish a catalytically active ATP binding pocket.[165] This aspect is of special importance when it comes to the biochemical evaluation of kinase inhibitors and comparison of different assay formats. Accordingly, it was shown that results obtained with full length protein kinases can vary greatly from results obtained with recombinant kinase domains.[166]

2.2 Kinase inhibitors comprise different binding modes

Inhibitor binding modes. Given the high structural flexibility of kinase domains, it seems natural that inhibitors can adopt different binding modes. The increasing amount of inhibitor-bound kinase crystal structures allows drawing a more and more detailed picture of how inhibitors engage their targets. Hence, previous categorizations of inhibitor binding modes have been extended and refined during the past years.[167] The most recent attempt to classify compound binding resulted in the categorization into type I, type I ½ (A and B), type II (A and B), type III, type IV, V and type VI inhibitors (Figure 10).[154] Inhibitors are divided into those that bind to the ATP pocket (types I, I ½, II), to the neighboring phospho-acceptor site (type III[168]), or to an allosteric site remote from the catalytic center (type IV[169]). Bivalent inhibitors targeting both the ATP pocket and an allosteric binding site are categorized into type V[170,171] and covalent (reversible and irreversible) inhibitors into binding type VI[172]. The large majority of small molecule kinase inhibitors are ATP-competitive requiring a more fine-tuned classification for binding to this pocket. To do so, several structural features that define the activation state of a kinase domain were taken into account: the DFG motif (in: active, out: inactive), the αC-helix orientation (in: active, out: inactive) and the formation of the R-spine (ordered: active, distorted: inactive). Additionally, the binding cleft can be roughly divided into a front cleft, the gate area and the back cleft;[173,174] the most recent update on binding types distinguishes inhibitors engaging the back cleft (A) or not (B). This is of interest, since engagement of this additional cleft may have influences on the residence time of inhibitors at their target protein.[154] Compounds binding to a completely activated kinase domain classified by DFG-in, αC-helix in and linear R-spine form the group of type I inhibitors. Due to historical reasons, this is the largest group of inhibitors and includes many clinical compounds such as Gefitinib, Vantedanib, Ruxolitinib, or Palbociclib to name a few. Zuccotto *et al*[175] introduced type I ½ binders which showed binding to "hybrid" structures comprising the activity determining DFG-in conformation but also αC-helix out and distorted R-spine formation which are characteristic for an inactive conformation. Depending on the occupancy of the back pocket, this kind of inhibitors are termed type I ½ A (back pocket engaged) and I ½ B (back pocket not engaged). Examples for type I ½ A inhibitors include Lapatinib, Vemurafenib or Lenvatinib; examples for type I ½ B inhibitors are Erlotinib, Sunitinib (binding CDK) and Crizotinib (binding MET). Another quite prominent group of inhibitors follow type II binding mode. These structures are defined by inactive DFG-out, αC-helix out and distorted R-spine leading to complete abrogation of catalytic activity. Again, this class of inhibitors is divided into type II A (e.g. Imatinib, Sorafenib, Nilotinib, Ponatinib) or type II B (e.g. Nintedanib, Sunitinib) depending on their ability to extend to the back pocket or not.



	I		I ½		II		III	IV	V	VI
	A	B	A	B	A	B				
location	ATP pocket	ATP pocket	ATP pocket	ATP pocket	ATP pocket	ATP pocket	substrate site	allosteric	bivalent	ATP pocket
binding	reversible	reversible	reversible	reversible	reversible	reversible	reversible	reversible	reversible	covalent
DFG	in	in	in	in	out	out	-	-	-	-
αC helix	in	in	out	out	out	out	-	-	-	-
R-spine	ordered	ordered	distorted	distorted	distorted	distorted	-	-	-	-
back cleft engaged	no	yes	no	yes	no	yes	-	-	-	-

Figure 10 | Small molecule kinase inhibitors bind kinases with different binding modes.[154] Kinase inhibitors can be categorized according to their ability of binding different kinase conformations and different positions into type I, type I½, type II, type III (phosphoacceptor site), type IV (allosteric), type V (bivalent) and type VI (covalent).

Conformational selectivity. The picture is further complicated by the fact that inhibitors are not conformation selective. This means that the same drug may bind in different conformations to another or even the same kinase. This is often observed for type I inhibitors, such as Erlotinib which binds its main target EGFR in both, type I and type I ½ B binding mode (as shown by protein crystallography).[154] Other examples are provided by Dasatinib being a type I inhibitor for ABL but a type I ½ A inhibitor for LYN, or Bosutinib targeting SRC in type I mode and ABL in type II B mode.[154] This can be due to several technical and biochemical reasons: i) crystal structures only prevent a snapshot of one conformation favorable for crystallization and do not provide information about the flexibility of the kinase domain; ii) from nuclear magnetic resonance (NMR) studies it is known that the DFG motif can very flexibly flip between in and out state especially in the inactive state and also in presence of in inhibitors;[176,177] iii) different kinases may provide different binding sites for the inhibitors resulting in different binding modes; iv) different kinases may intrinsically prefer different conformations over others when expressed recombinantly for crystallization experiments. This raises the question if such categorization mainly based on crystal structures is meaningful, or if other ways to describe compound binding would be more suitable. NMR might be an alternative and it is indeed striking that NMR resolves the flexibility of changing kinase domain conformations – yet, it cannot provide detailed insight into the molecular interactions and fully recapitulate the influence of other regulating protein domains.

Binding type selectivity. Since type II inhibitors exploit an additional hydrophobic pocket which is only present in inactive conformation and not very conserved, these inhibitors have long been considered to be more selective than type I inhibitors.[178] This is controversially debated in the field and several studies already invalidated this hypothesis.[179,180] The additional chemical moiety extending to the back pocket might lead to prolonged residence times, but does not contribute necessarily to drug selectivity in terms of a compound's target space.[181] Another attempt to increase compound selectivity are irreversible inhibitors of class VI which bind their intended target covalently by a Michael addition.[172] In this case, compound binding can be divided into two steps: Non-covalent interactions position the molecule in the active center and enable the latter covalent modification between the nucleophilic Cys-SH and an unsaturated carbonyl moiety.[182] This strategy is only suitable for kinase inhibitors comprising an accessible cysteine in the active center (such as EGFR or BTK)[183] and prolongs residence time of the drug. However, other kinases can still be bound by non-covalent interactions which makes this approach less suitable as a strategy to increase drug selectivity *per se*. Type III inhibitors are a relatively small class of very selective (mainly MEK-)inhibitors occupying the phospho-acceptor site next to the ATP binding pocket (e.g. TAK-733, Trametinib).[154] This is a very successful approach to increase selectivity but is not be amenable for all kinases.

Structural implications for ATP-competitive inhibitors. ATP-competitive small molecule kinase inhibitors (type I, I½ and II) comprise several chemical features to provide interactions with particular structural features of the catalytic kinase domain (Figure 11).[184,185] One such very common motif is the hinge binder moiety which is responsible for two direct binding interactions to the backbone of the hinge residue (in EPHA2: Met695). Often, this moiety is composed of a hydrogen donor (NH) and a hydrogen acceptor (N) bridged by a single carbon unit. Most of the times, the inhibitors interact with several hinge residues and about 50% of kinase inhibitors engage the neighboring gatekeeper residue (in EPHA2: Thr692) either by direct or hydrophobic interactions.[154] The gatekeeper is located in the center of the ATP binding pocket and controls access to a hydrophobic pocket which is deeply buried behind the DFG motif.

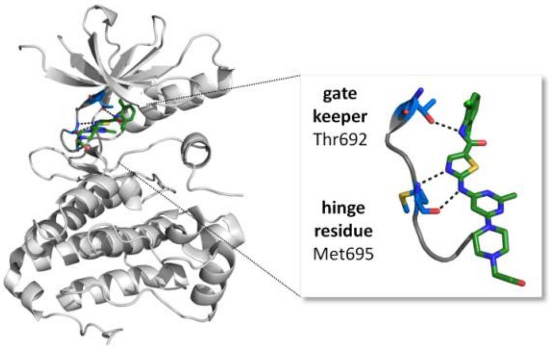


Figure 11 | Structural implications for ATP-competitive inhibitors. The hinge residue (Met695 in EPHA2) and the gatekeeper (Thr692 in EPHA2) are of special importance for small molecule binding to the ATP pocket.

The gatekeeper residues have gained immense interest in small molecule inhibitor research since mutation of these residues often lead to drug resistant phenotypes in the clinic (e.g. T315I in BCR-ABL).[186] ATP competitive inhibitors engage the same binding cleft as ATP and thus, often interact with catalytically active residues and residues forming the catalytic C-spine and/or the regulatory R-spine.[154] This may influence binding affinity and also residence time of these compounds.[181] As described above, protein kinases undergo a multitude of different conformations during their activation and the catalytic activity. The inactive kinase conformation is generally less conserved and provides the opportunity of targeting additional hydrophobic pockets and sub-pockets which are not accessible in the activated state. This implies the possibility to target different conformations by different inhibitor designs.

2.3 Different metrics to calculate small molecule selectivity

Compound selectivity in general. Compound selectivity is a major issue starting from drug discovery and target validation up to the clinical application – especially for ATP competitive inhibitors. Due to the high conservation of the ATP pocket and the large amount of proteins comprising ATP pockets, the design of selective ATP competitive inhibitors remains challenging.[146] There are different opinions towards the necessity of selectivity for the success of a drug in the clinics. On the one hand, many approved and efficient kinase inhibitors, such as Dasatinib, display a very promiscuous target profile, which might be valuable for the clinical outcome and successful therapy (polypharmacology).[187] In a clinical background, compound selectivity might explain the mode of action of a certain drug and explain adverse and advantageous effects of drug treatment. On the other hand, the large amount of target proteins might cause toxic side effects and renders this inhibitor a poor chemical probe to study its inhibitory effect on particular proteins.[188] High selectivity was defined to be one of the key characteristics of chemical probes (chemicalprobes.org) and the development of chemical probes for the majority of kinases is still in its infancy. Independently from the application of an inhibitor as a therapeutic agent or a chemical probe, selectivity profiling remains crucial to understand a drug's mode of action and to anticipate a phenotypic effect.[180]

Calculation of compound selectivity is a challenging task due to different reasons. First of all, the desired answer from selectivity calculation very much depends on the prior research question which varies for different research areas.[180] In this respect, compound selectivity could be calculated for instance to i) receive a target-independent global view on compound selectivity (e.g. identification of chemical probes in a large selectivity screen), ii) explain a certain mode of action or phenotype (e.g. adverse and advantageous effects of clinical drug treatment), iii) characterize the ability of compounds to selectively hit a particular target (e.g. important for medicinal chemistry).[180] Another challenge for selectivity calculation is the diversity of available data sets and assay formats used to profile compounds. Data can be obtained in either single dose treatments or full dose responses which results in different levels of usability of the data. Also, different assay formats such as activity or binding assays may not necessarily have the same outcome. The comparability of different data sets is also very much dependent on the utilized input material. For instance, purified recombinant constructs of kinase domains may not have the same properties than full length kinases in complex cell lysates.[166] Given the increasing number of kinase drugs and the technological advances in screening them, more and more large compound selectivity screens were published within the last decade.[139,141,143,189-192] This created the demand but also facilitated the development of several strategies to calculate compound selectivity.

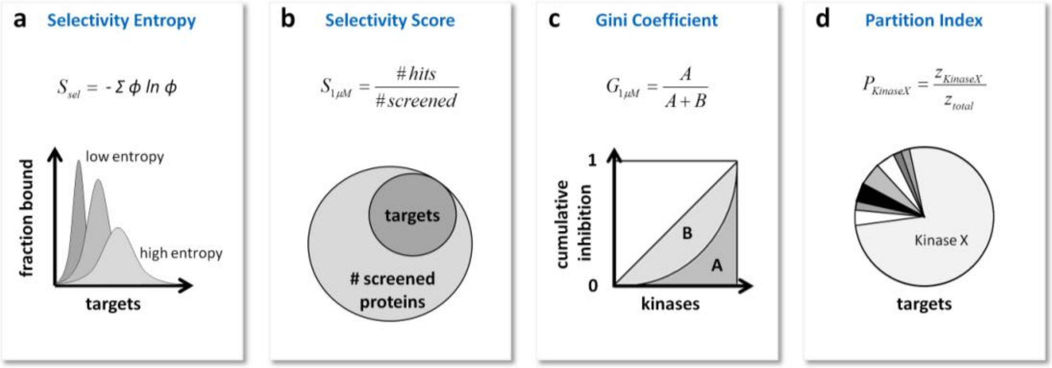


Figure 12 | Measures of small molecule selectivity. a) Selectivity entropy is a measure for the distribution of an inhibitor amongst all target proteins. **b)** The selectivity score divides the number of targets at a certain threshold concentration by the number of total screened proteins. **c)** The Gini coefficient calculates the cumulative inhibition of all target proteins at a certain threshold concentration in relation to all tested proteins. **d)** The partition index describes which fraction of an inhibitor is bound by a particular target protein in a pool of target proteins.

Selectivity entropy. Uitdehaag *et al.* introduced the selectivity entropy[193,194] to enable global selectivity determination for compounds irrespective of particular target proteins (Figure 12a). The selectivity entropy assumes that a small amount of drug would distribute in an excess of target proteins according to their K_D values in a Boltzmann distribution. The width of this distribution reflects the amount of different energy states (equivalent to different target proteins) of the compound, and thus, describes its selectivity. The selectivity entropy can be calculated by theoretical entropy calculation according to Boltzmann's entropy formula. Very unselective compounds have a broad distribution leading to a high selectivity entropy value whereas selective compounds display a tight distribution and low selectivity entropy values. In this approach, only target proteins have an impact on the distribution, non-targeted proteins are not considered. Hence, this scoring system enables the calculation of compound selectivity independent of the tested panel size – providing a way to compare compound data derived from different screens. The major drawback of this method, however, is the fact that it is not possible to determine compound selectivity with regards to a particular target protein. This would be necessary for many applications as for example in medicinal chemistry which renders this selectivity metric unsuitable for this kind of research question.

Selectivity Score. Perhaps the simplest way to measure selectivity is given by the selectivity score (Figure 12b).[139,191] It is calculated by dividing the number of target proteins of a drug at an arbitrary concentration (e.g. 100 nM) by the number of tested proteins. It ranges from 0 (very selective) to 1 (very unselective). One drawback of this scoring system is clearly the strong dependence on the panel size used for screening. Depending on how many proteins are tested, the denominator may change drastically and lead to very different selectivity scores even for the same inhibitor. This also means that selectivity scores can only be compared within one data set and are not suitable for cross-comparison of different data sets. Another disadvantage evolves from the fact that target proteins are only counted and the experimentally determined values are not taken into account. Thus, all targets are treated equally even though a compound might affect them with very different potencies. Also, selectivity scores cannot be calculated with respect to a particular target which is a major disadvantage of this score. To sum up, these problems impair comparability between different screens and reduce the selectivity score to a rather imprecise and untargeted measure of selectivity.

Gini coefficient. The Gini coefficient[195] uses the percent inhibition data at a single inhibitor concentration to calculate the relative inhibition fraction of each target protein (Figure 12c). It orders and normalizes the single data points to derive a cumulative fraction inhibition plot which describes the inhibitor's selectivity profile depending on the background set of tested kinases as a Lorenz curve. The gini coefficient is calculated as the ratio of the area outside this distribution and the complete area, resulting in a value between 0 (unselective; all tested proteins are inhibited equally) and 1 (selective; only 1 target protein). The gini coefficient was developed to determine selectivity of compounds tested at single compound concentrations which makes it more vulnerable to technical variation. Moreover, every tested protein is included in the calculation which leads to a strong dependence of the gini coefficient on the panel size of the respective assay. Similarly to the selectivity score, this impairs comparison of selectivity data between different datasets and also leads to poor performance with very small[196] but also very large screening panels (all values close to 1). A major drawback of the method is the fact that it is possible to calculate the gini coefficient in a concentration-dependent but not target-dependent manner.

Partition index. Similarly to the selectivity entropy, the partition index[196] is based on the theoretical distribution a compound would have in a complex mixture of protein targets, where proteins are in excess (Figure 12d). The hypothesis is, that the compound will distribute between the protein targets according to their affinities (K_D values). Hence, the partition index describes the relative affinity for a reference kinase compared to all measured affinities. The higher the affinity for a certain target protein over other targets, the more inhibitor molecules will bind to this target preferentially. Very selective compounds will have a high partition index (close to 1) whereas very unselective compounds will result in a low partition index (close to 0). The partition index only considers targeted proteins for its calculation which makes is less vulnerable to the tested panel size. Also, it can be used to determine selectivity of a compound towards single or groups of proteins which makes it more broadly applicable than the selectivity entropy. One drawback of this approach is the requirement of full dose response data as it relies on K_D values. Also, the theoretical background of this calculation is quite artificial. Usually, activity or binding assays are set up in a way that protein concentration is much lower than the expected K_D and not every protein will be tested at the same protein concentration. This is also true for a biological system where kinase concentrations are very low and may differ a lot. Accordingly, the theoretical basis of this scoring systems does not reflect the assay situation nor the application in a biological system which could be problematic for its applicability.

Each of these scores is capable of reflecting a certain angle of compound selectivity – being it a global selectivity independent of target or concentration (selectivity entropy), concentration-dependent selectivity (selectivity score, gini coefficient), or target-dependent selectivity (partition index). However, none of the scores would be satisfactory to calculate selectivity in a concentration- and target-dependent manner and in the case of screening data where the number of proteins in not constant. This is the case if selectivity profiling is performed using chemical proteomic approaches as described in the next chapter.

3 Mass-spectrometry based target deconvolution

Very often, the target profiles of compounds are not completely resolved – even for compounds that are already used or evaluated in the clinic. Comprehensive drug target deconvolution is crucial to understand a drug's mode of action and its usability as therapeutic agent or chemical probe.[180] One aspect of improved target profiling is the possibility of drug repurposing in other biological contexts: for instance, kinase inhibitors designed for treating oncogenic diseases could also be useful in inflammatory disorders, (auto)immune disorders, or neurodegenerative diseases.[154,197,198] So far, inhibitor development mainly focuses on about 10% of all protein kinases which are designated as *bona fide* drug targets;[199] FDA-approved inhibitors even only target about two dozen protein kinases.[185] Yet, more and more protein kinases are found to be implicated in disease which will lead to an increased demand of targeted inhibitors in the near future.[188,199] Improved characterization of drug target space might fill this gap by enabling the usage of a compound for the targeted inhibition of formerly unknown target proteins (e.g. using Dasatinib for EPH receptor inhibition).[180] Target deconvolution is also required to elucidate drug targets after phenotypic screening.[200] In this case, compounds are screened against a certain phenotypic readout, but the molecular target and the molecular mode of action are most often unknown.

Hence, there is a high demand of comprehensive drug target profiling which is covered by several technologies for target space elucidation.[200] One possibility includes cloning and expression of protein libraries to examine compound:protein interaction for instance in phage display or yeast-three-hybrid experiments. These are cost-and labor-intensive *in vitro* technologies and require previous knowledge about the target space.[200] In case of kinase inhibitors, drug target deconvolution is often performed using *in vitro* activity screens.[201] Enzymatic kinase activity can be measured by different means: consumption of ATP, generation of ADP, formation of phosphopeptides or decrease of non-phosphorylated peptides. Typical readouts include radioactivity, luminescence, fluorescence, (time resolved) fluorescence resonance energy transfer, or the behavior of peptides in a thermophoretic gradient or microfluidic platform. All these approaches offer a comprehensive view on the molecular target space of compounds, but completely decouple the compound:target interaction from its relevance in a biological system.

Most compounds exert their function on proteins which may be subject to post-translational modifications, protein processing, expression as different isoforms and different functional complexes. Thus, proteomic experiments performed in a relevant biological background are very well-suited for the evaluation of the target space of compounds.[200,202,203] Mass spectrometry-based quantitative proteomics enables the simultaneous identification and quantification of a few thousand proteins in a single measurement – providing a powerful technology platform for interrogation of drug:protein interaction.[204] Accordingly, several chemical proteomic approaches, such as affinity matrices, have been developed for mass spectrometry-based target deconvolution, and will be specified in more detail hereafter.[202,205-207]

3.1 Mass spectrometry based quantitative proteomics

Bottom-up proteomics workflow. Bottom-up proteomics is the most frequently used proteomics approach to identify and quantify proteins derived from complex biological samples and is based on the measurement of peptides instead of full length proteins (top-down proteomics).[208,209] Following this workflow, proteins are first extracted from their cellular context, either by native or denaturing lysis conditions. Next, the proteins are digested to peptides by treatment with a proteolytic enzyme that specifically cleaves at certain amino acid motifs.[210-212] Most often, this is accomplished by the enzyme trypsin which introduces a proteolytic cleavage C-terminal of lysines and arginines.[209] An on-line ion-pairing reversed-phase liquid chromatography (LC) system is used for separation of the tryptic peptides prior to mass-spectrometric measurement.[213-215] Every peptide is retained to different extent according to its hydrophobicity and size and elutes at a specific point of the LC gradient. Upon elution, the peptides are ionized and subjected to the mass spectrometer which measures the mass-to-charge ratio and intensity of each peptide.

According to sample type and experimental setup, different enrichment or fractionation steps may be performed at different stages of this standard workflow. If required, a potential enrichment at protein level can be performed after protein extraction, e.g. by affinity matrices, chemical probes or antibodies. In case of very complex samples, additional orthogonal off-line fractionation (e.g. by strong anion exchange,[216] or mixed mode[217]) of peptides may be required to enhance the protein coverage. Another possibility are enrichment strategies on peptide level which are often utilized for post-translationally modified peptides such as phosphorylated (e.g. via immobilized metal affinity chromatography[218]) or acetylated (e.g. via immunoprecipitation) peptides.

For chemoproteomic target deconvolution in this work, a standard bottom-up workflow was performed (Figure 13), including native protein extraction from cell culture, protein enrichment using an affinity matrix, in-gel tryptic digestion,[219,220] on-line peptide separation by reverse phase chromatography, peptide ionization and mass spectrometric acquisition.

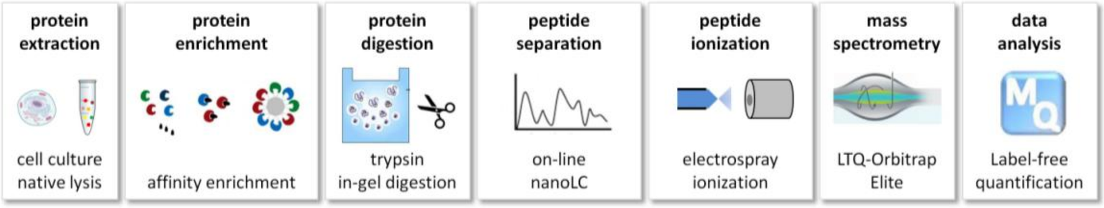


Figure 13 | Bottom-up proteomics workflow for small molecule selectivity profiling. Proteins are extracted under native lysis conditions and affinity enrichment is performed using Kinobeads. Enriched proteins are desalted via a short SDS-PAGE and in-gel digested. The resulting peptide mixture is separated by an on-line liquid chromatography setup (nanoLC) and peptides are ionized via electrospray ionization. Mass-to-charge ratios (m/z) of peptide precursors and fragments are measured by tandem mass spectrometry in a LTQ-Orbitrap Elite. Peptide and protein identification and label-free quantification is performed using MaxQuant. Modified from Steen *et al*[209].

Mass spectrometry. In a simplified view, mass spectrometers can be described as balances which measure the molecular weight of the respective analyte as a mass-to-charge ratio.[221] It roughly consists of three modules: an ion source, a mass analyzer and a detection system.

Technological development has generated a variety of different ionization systems (e.g. electrospray, matrix-assisted laser desorption) and mass analyzers (e.g. ion traps, time-of-flight analyzers, orbitraps). In this study, a hybrid ion trap mass spectrometer, namely the Orbitrap Elite (Figure 14),[222,223] was used and the following section will give an overview on the underlying technology.

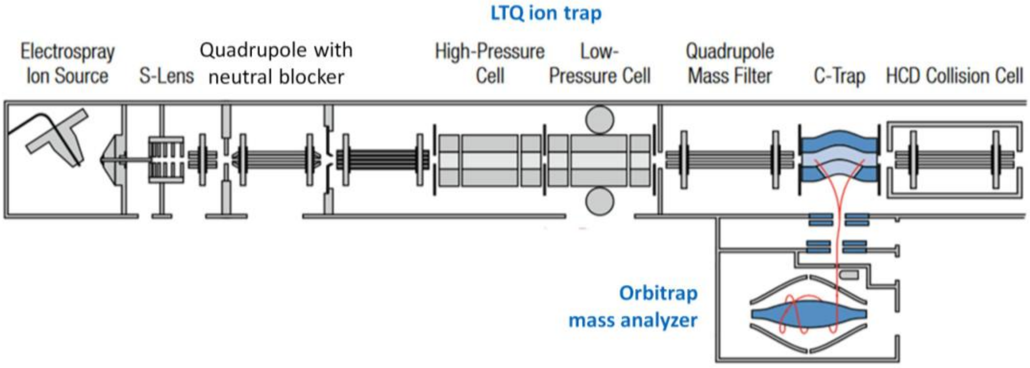


Figure 14 | Schematic overview of the LTQ-Orbitrap Elite.[222,223] Peptides are ionized via electrospray and uncharged species are removed in the first quadrupole. Peptides are collected in the linear 2D ion trap (LTQ) and further transferred to the C-trap. Here, the peptides are injected into the orbitrap mass analyzer where m/z ratios are determined (MS1). Selected peptides are fragmented in the HCD collision cell and fragments are subjected to the orbitrap mass analyzer again (MS2). Figure modified from Thermo Fisher Scientific.

Ionization of the peptides is accomplished by electrospray ionization (ESI).[224-226] Addition of a (usually positive) charge is necessary to make them manipulable, guide them through a complex setup of electromagnetic fields, and for latter detection of the analytes.[227] As the peptides elute from the nanoLC[228], they pass a thin capillary and electric potential applied to this capillary leads to a charged mobile phase. At a certain voltage, a Taylor cone forms at the end of the capillary and the peptides are emitted as a precise jet which subsequently disperses into single droplets.[229-231] The mobile phase evaporates which leads to droplet fission and finally to the release of charged species. Addition of dimetlhylsulfoxide to the solvent was found to enhance ionization efficiency, probably by decreasing the surface tension of the formed nanodroplets.[232] Ionized peptides enter the mass spectrometer through a transfer capillary and are guided through the S-lens towards a rotated multipole comprising a neutral blocker where neutral uncharged species are removed. Peptides are subjected to a high sensitivity linear (2D) ion trap (LTQ) where a defined amount of charges is collected. [221] Peptides are further transferred into the C-trap which focuses the ions for subsequent injection into the high resolution orbitrap mass analyzer.[233-235] Peptides are injected perpendicular into an electromagnetic field established by an oval outer electrode and a central spindle electrode. Charged species start to oscillate around the central spindle electrode and adopt stable trajectories with certain oscillation frequencies according to their mass-to-charge ratio. The movement of charged species within the applied electromagnetic field leads to transients which are acquired by the mass spectrometer and deconvoluted by fast fourier transformation to generate a mass spectrum.[233,236]

In tandem mass spectrometry,[238] two different types of mass spectra are acquired – MS1 spectra of all peptides eluting at the same time from the LC column, and MS2 spectra of peptide fragments of selected peptide precursors (Figure 15a).[209]

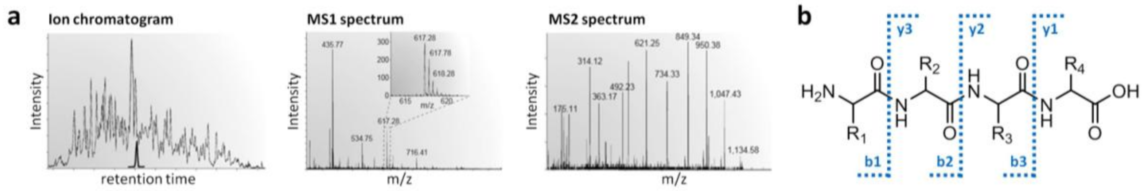


Figure 15 | Mass spectrometry based proteomics. a) Tandem mass spectrometry measures peptides precursors in MS1 spectra and peptide fragments in MS2 spectra. **b)** HCD fragmentation of peptides produces characteristic b- and y-ion series.[237]

After a MS1 spectrum has been generated for the peptide precursors, single precursor ions are collected and transferred to the higher-energy collisional dissociation (HCD) cell. HCD fragmentation occurs by beam-type activation where peptides are accelerated by current offsets und fragmentation happens upon collision with nitrogen atoms.[239] The generated fragments are then subjected to the C-trap and fragment masses are again read out in the orbitrap mass analyzer. Different fragmentation techniques result in the formation of different ions generated by different breakpoints within the amino acid chain. HCD fragmentation leads to breakage between the carbonyl C=O and the nitrogen N-H of the peptide bonds connecting the single amino acids (Figure 15b). This generates characteristic b- (charge retained at N-terminus) and y-ion series (charge retained at C-terminus).[237] Ideally, the distance between two neighboring peaks corresponds to the mass of a single amino acid and is used to derive the amino acid sequence of this particular precursor peptide. In data-dependent acquisition, the mass spectrometer alternates between MS1 and MS2 spectra and picks a predefined number of peptide precursors for subsequent fragmentation based on intensity.[240,241] This is advantageous for an unbiased and comprehensive view of the sample proteome but may lead to inconsistent identification of low abundant peptide species between different MS runs.[242] Inclusion lists containing the m/z ratio and elution times of known peptides of interest can partly help to reduce this problem.

Peptide identification and quantification. The amino acid sequence can be derived from MS2 fragment spectra but is often not resolved completely. *De novo* sequencing of peptides requires full sequence coverage which is mostly not achieved in the mass-spectrometric measurement. This problem was overcome in 1990 when researchers started to use databases for matching their experimental spectra. There is only an infinitesimal amount of theoretically possible sequences present in nature which greatly reduces the search space and also allows matching of spectra with incomplete sequence information.[209] A predefined database is digested *in silico* to yield theoretical tryptic peptides. This list of potentially present peptides is then filtered for the list of experimentally measured and charge state-deconvoluted peptide precursor masses. This reduces complexity for the subsequent step where spectra of *in silico* tryptic peptides are predicted and compared with the acquired MS2 fragment spectra. Each peptide-spectrum-match (PSM) is scored according to how well the predicted and the experimental spectra fit to each other. In practice, the experimental spectra are not only searched against the specified database but also against a decoy database which contains reversed versions of the *in silico* generated tryptic peptides.[243-246] This enables the determination of a false discovery rate (FDR) and a score cutoff to filter the outcome for high confidence identifications.[247-249] One major drawback of this peptide identification strategy is the requirement of a comprehensive and well-curated sequence database for proteomic experiments; limiting the approach to sequenced organisms. Several search engines, such as Mascot or Andromeda (MaxQuant),[247,250-252] were developed to accomplish peptide identification and subsequent matching to proteins. Unambiguous assignment of peptides to one particular protein is often not possible if the identified peptide is present in several related proteins or protein isoforms (referred to as razor peptides in MaxQuant). Hence, proteomic experiments often report protein groups of related proteins if no unique peptide was found that allows the unambiguous assignment to a single protein. This also shows, that proteomic experiments using this database search strategy are usually not capable of distinguishing distinct protein isoform derived from alternative splicing or by post-translational protein processing.

Typically, not only the identification of peptides and proteins is of interest but also an estimation of the abundance within the sample.[253,254] Several label-free and label-based techniques have been developed so far to enable quantitative proteomics; in this project MS1-based label-free quantification was utilized.[255-257] Mass-spectrometry derived intensities *per se* do not qualify for quantification as different peptides yield different signals – mainly depending on their physicochemical properties such as hydrophobicity or presence of basic residues. Nevertheless, the behavior of the same peptide should be identical between different mass-spectrometric runs allowing for relative quantification of peptide abundance between two samples. MS1-based label-free quantification (MaxLFQ in MaxQuant[250-252]) uses the precursor intensities and elution times of the MS1 spectra to infer an elution peak of each peptide precursor. LFQ intensities are derived from integration of the peak area in MS1 space. Identification of the peptide species is still obtained from the MS2 spectra. Data-dependent acquisition may lead to missing quantification values and missing identifications, especially for low abundant peptides. This can be overcome to some extent by retention time alignment of several runs which allows utilizing sequence information obtained from one sample also for the same precursor mass found in another sample (Match-between-runs option in MaxQuant).[247] This quantification strategy facilitates relative quantification of the same protein over several samples but does not enable comparison of different proteins within the same sample. MaxLFQ intensities are normalized by a “delayed normalization” approach[250] which is a sophisticated total sum normalization calculating the median of all measured peptide ratios. It is based on the prerequisite of a dominant population of proteins that does not change notably between experimental conditions.

3.2 (Chemo-)proteomic methods used for target deconvolution

Proteomic approaches for target deconvolution. An optimal (chemo)proteomic experiment for target deconvolution of a ligand[258] would measure the interaction between the ligand and the protein i) without any prior modification of one of the binding partner, ii) without an additional chemical probe that might interfere with the binding equilibrium, and iii) in the biological environment where the ligand is used. Yet, no (chemo)proteomic experiment is capable of providing a comprehensive view on the target profile accomplishing all these requirements in the same time. Recently, a novel proteomic approach called cellular thermal shift assay (CETSA) was developed and raised high expectations towards achieving this goal (Figure 16a).[259-261] Here, cells or cell lysates are treated with a ligand and subsequently subjected to a temperature gradient. Interaction with the ligand is assumed to lead to stabilization (or destabilization) of the protein resulting in a shift in the protein melting temperature. Unfortunately, the results obtained by CETSA so far were not as comprehensive as formerly anticipated. It appears that not all proteins are amenable to thermal shifts which might be due to protein size, binding kinetics with high off-rates or other reasons. Thus, CETSA is favored by an unbiased experimental setup but the picture it provides is not comprehensive enough to evaluate the target space of a certain ligand. In the past, large efforts were undertaken in the area of activity-based protein profiling (ABPP, Figure 16b).[262-265] Chemical probes for different enzymes such as hydrolases,[266] dehydrogenases,[267] proteases[268] or phosphatases[269] were developed that specifically target and covalently bind important amino acids involved in the catalytic activity of the enzyme. A similar approach is photoaffinity labeling (PAL) where a chemical probe reversibly binds to its target proteins and is then covalently linked by photoactivation of a crosslinker such as diazirine (Figure 16c).[270,271] Both ABPP and PAL probes usually comprise an additional handle that enables stable enrichment of target proteins e.g. by click chemistry or streptavidin-biotin interaction.[204] Usually, these chemical probes are also relatively small which facilitates their application within cells. However, they are not optimal to screen for binding affinities of reversible ligands in competitive experiments as the covalent binding of the probe strongly influences the binding equilibrium of drug and protein. In this case, the application of affinity matrices has been proven advantageous (Figure 16d).

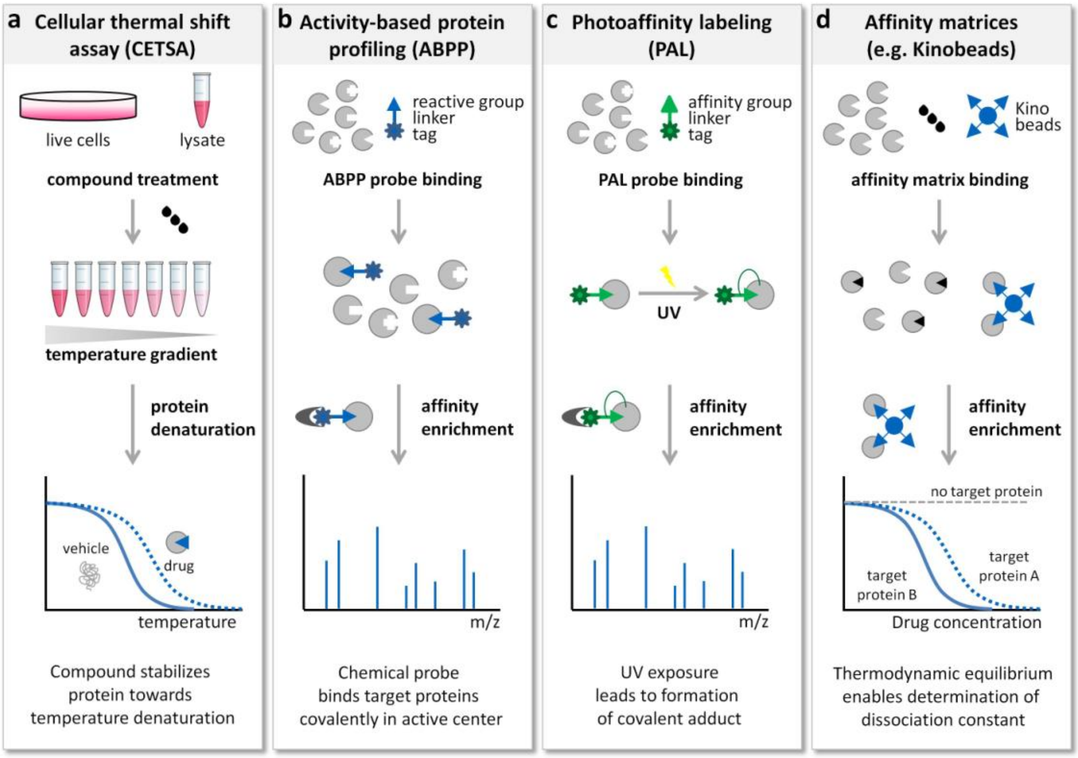

a Cellular thermal shift assay (CETSA)
live cells
lysate
compound treatment
temperature gradient
protein denaturation
vehicle
drug
temperature
Compound stabilizes protein towards temperature denaturation
b Activity-based protein profiling (ABPP)
reactive group
linker
tag
ABPP probe binding
affinity enrichment
m/z
Chemical probe binds target proteins covalently in active center
c Photoaffinity labeling (PAL)
affinity group
linker
tag
PAL probe binding
UV
affinity enrichment
m/z
UV exposure leads to formation of covalent adduct
d Affinity matrices (e.g. Kinobeads)
Kino beads
affinity matrix binding
affinity enrichment
no target protein
target protein A
target protein B
Drug concentration
Thermodynamic equilibrium enables determination of dissociation constant

Figure 16 | Proteomic methods used for small molecule target deconvolution. a) The cellular thermal shift assay (CETSA) measures changes of protein melting temperature that are provoked by the (de)stabilizing effects of ligands binding to their target proteins. **b)** Activity-based protein profiling (ABPP) employs chemical probes that covalently react with their target proteins. **c)** Photoaffinity labeling (PAL) utilizes chemical probes that reversibly bind to their targets and are subsequently crosslinked via a UV-reactive chemical moiety. **d)** Affinity matrices, such as Kinobeads, bind and enrich target proteins reversibly and under thermodynamic equilibrium which enables the determination of the binding affinity of a small molecule inhibitor.

Here, ligands are immobilized on a solid support such as sepharose beads and enrich their target proteins from complex cell lysate. Traditionally, a linkable version of the ligand of interest is designed which should optimally not influence the binding properties of the ligand to its target proteins.[205-207] This is (at least partially) accomplished by prior analysis of the underlying structure-affinity relationships and optimization of linker lengths. However, this strategy optimizes the probe for binding to its main target; the influence of linker length on binding to unknown off-targets is disregarded. Also, immobilization of ligands reduces their degrees of freedom and leads to an entropic contribution to the binding affinity which cannot be overcome by the probe design. Additionally, the clear identification of the enriched sub-proteome is often hampered by high background due to unspecific binding. This can be optimized by evaluation of the optimal coupling density which increases the signal-to-noise ratio of such experiments, and by dose-dependent competitive experiments with the non-modified ligand. In this scenario, specific binders can be identified by their dose-dependent behavior which greatly reduces false-positive target selection. Optimally, an affinity matrix is created in such way, that it allows the enrichment of complete sub-proteomes and not only the specific targets of a specific ligand of interest.[204] An unbiased and comprehensive matrix facilitates competition experiments with many different compounds and renders the labor-intensive design of separate probes unnecessary. Examples for this strategy include matrices based on the histone deacetylase inhibitor Vorinostat that are used to investigate HDAC complexes[272] or the Kinobeads affinity matrix[137,273] which enriches protein kinases and other nucleotide-binding proteins.

Target deconvolution of kinase inhibitors. Kinobeads affinity matrix was first described in 2007 and originally consisted of seven immobilized broad spectrum inhibitors capable of enriching 165 to 183 human kinases in different human cell lines.[137] Both the set of probes and the used cell lysate mix were further optimized towards increased coverage of the human kinome: the Kinobeads γ setup utilized in this study consists of a mix of five immobilized ATP-competitive small molecule kinase inhibitors.[273] These compounds are immobilized on sepharose beads via NHS chemistry (probe linked via NH2 group) or via a PyBroP-mediated amide coupling reaction (probe linked via COOH group). The utilized compounds include Purvalanol B, linkable PD-173955 and Cpd19 which feature a broad selectivity profile facilitating the enrichment of a multitude of protein kinases. This broad selectivity matrix is complemented by a few more specific probes enriching EGFR (linkable Vandetanib) and AKT kinase family (Cpd15) that are not covered by the unselective probes. Furthermore, kinome coverage is enhanced by mixing four different cancer cell lysates: K-562, MV-4-11, SK-N-BE(2), COLO 205. This increases the target space amenable to drug selectivity profiling and allows identification of potential toxic off-targets like the ferrochelatase FECH[274] or other non-kinase targets – thus, expanding the mode of action analysis for small molecule kinase inhibitors.[275] Target deconvolution of kinase inhibitors is accomplished by a competitive setup where the lysate is treated with increasing concentrations of an inhibitor of interest. Both the compound and the Kinobeads compete for the same binding site leading to reduced binding to Kinobeads and a dose-dependent loss of signal intensity in the subsequent mass spectrometric readout.

A similar approach, the KiNativ technology, uses desthiobiotinylated ATP to covalently link ATP-binding proteins and subsequently enrich them via streptavidin.[276,277] This technology also enables broad assessment of target proteins beyond the protein kinase family.[278] In contrary, the traditionally used screening panels that mainly consist of recombinantly expressed kinase domains do not offer this possibility. Screening of isolated protein kinase domains facilitates high throughput and reproducible results, but does not accomplish close-to-nature conditions as lysate-based technologies do. Complex native cell lysates offer full-length proteins that have been functional in a cellular context, thus, carrying all their posttranslational modifications and binding partners; this may also lead to differing results in recombinant screens and lysate-based assays.[166]

The Kinobeads γ approach also comprises some drawbacks as several protein kinases are not accessible to enrichment by the presented experimental setup. This can be due to insufficient affinity to the matrix, which is the case for the PI3K family, and can be overcome by addition of other probes such as linkable Omipalisib. Another reason for incomplete coverage is the used cell line mix that does not cover all protein kinases such as VEGF and FGF receptor families. Utilization of lysates from other cell lines or tissues (e.g. placenta) can fill these gaps.

3.3 Thermodynamic considerations about ligand:protein interactions

Binding events between ligands, compounds and affinity matrices alike, and proteins and their affinities are governed by thermodynamic principles such as the binding equilibrium, association and dissociation kinetics and enthalpic/entropic contributions to the Gibbs free energy. These principles build the basis of how binding affinity and residence time of ligand:protein interactions are effected in a complex system of protein, ligand and their environment. Understanding these principles is not only relevant for the interaction of a protein with a compound of interest but also for the interaction of a protein to the utilized chemical probe.

Equilibrium and affinity. Determination of the affinity of ligand-protein interactions is key to many research questions and especially important for compound priorization in drug discovery projects. The interaction of a target protein P and a ligand L to a complex P.L can be described in a simple formula (1).

$$P + L \rightleftharpoons P.L \tag{1}$$

If binding equilibrium is established, the dissociation constant K_D describes the propensity of a complex to dissociate into its individual complex components, as indicated in formula (2) with [P], [L], [P.L] referring to the concentration of the target protein P, the ligand L and their complex P.L.

$$dissociation\ constant\ K_D = \frac{[P]*[L]}{[P.L]} \tag{2}$$

In the simple case of two-membered equilibrium consisting of a single ligand L binding to a single target protein P, the K_D can be described as the concentration of ligand L that is necessary to occupy half of the binding sites of target protein P (3). In this case, the K_D has molar units and is reciprocal to the affinity of a protein-ligand interaction.

$$\frac{[P.L]}{[P]total} = \frac{[P.L]}{[P]+[P.L]} = 0.5 \quad \rightarrow \quad [P] = [P.L] \quad \rightarrow \quad K_D = [L] \tag{3}$$

Experimental determination of the K_D can only be accomplished if the protein concentration [P] is lower than the K_D. If this prerequisite is not met and the total protein concentration $[P]_{total}$ is higher than $K_D = [L]$, then $[P]_{total}$ cannot be divided into equal parts of [P] + [P.L] as more ligand L would be required to fill half of the binding sites allocated by protein target P. In this scenario, K_D determination would be independent from ligand concentration [L] and result in $K_D = ½ [P]_{total}$; irrespective of the true binding affinity of the ligand. K_D determination by Kinobeads is based on the assumption that the target protein concentration is below K_D concentration. This is a black hole of the assay, since the absolute quantification of each single protein concentration in the lysate is not feasible. However, kinases are in general very low abundant proteins such that this should not cause major problems in the current setting.

Optimally, chemical probes should not influence the binding equilibrium of a compound and its target protein, but this is only barely the case. Thus, what is measured is not the K_D of a two-state equilibrium but an assay-dependent EC_{50} value (half maximal effective concentration) of a three-member equilibrium (Figure 17a). Protein P is depleted from the binding equilibrium by the Kinobeads which leads to a re-adjustment of the binding equilibrium of ligand L and protein P. Supposing that the affinity and the maximum concentration of the complex [P.L] remains the same within the sample, an increase of ligand concentration [L] is required to yield the same K_D value (according to formula (2)). Thus, protein depletion provokes a shift of the measured EC_{50} values towards higher concentrations compared to the actual K_D value. In case of Kinobeads, the degree of protein depletion from the lysate depends on i) protein abundance in the cell lysate, ii) affinity to the Kinobeads probes, and iii) the apparent Kinobeads probe concentration. Given the complexity when working with native lysates, none of these parameters is known and their determination would be tedious and time-consuming. Still, being able to calculate the degree of protein depletion from the lysate would enable the estimation of the true dissociation constant from the experimentally obtained values.

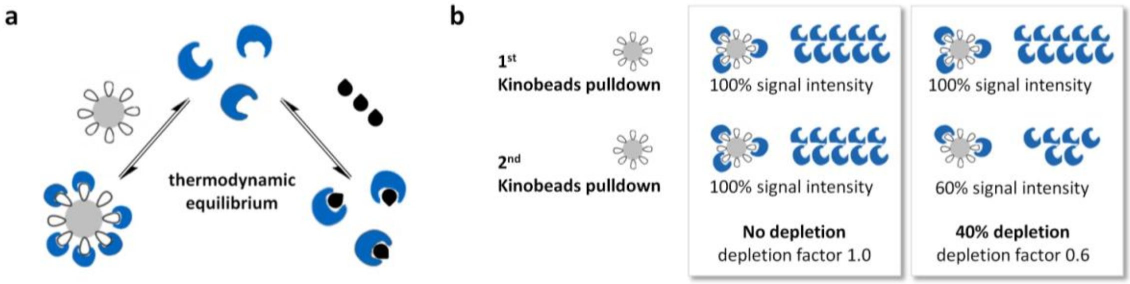


Figure 17 | Determination of dissociation constants using Kinobeads. a) In a competitive Kinobeads pulldown, a three-membered thermodynamic equilibrium is established between the affinity matrix, the target proteins and the kinase inhibitor of interest. **b)** The utilized chemical probe depletes protein from the lysate. The relative degree of protein depletion can be evaluated via a subsequent pulldown of pulldown. The determined depletion factor is used for EC_{50} to K_D conversion.

A way of estimating the binding affinity in such a complex experimental setup was developed by Sharma *et al.*[279] This approach is a valuable option for EC_{50} to K_D conversion, as described in more detail hereafter.

Kinobeads bind a fraction f(P) of the total protein amount $[P]_{total}$; the remaining fraction of free protein [P] can be described by 1-fraction f(P)), as depicted in formula (4).

$$[P]_{total} = [P] + [probe.P] = (1 - fraction\, f(P)) + fraction\, f(P) \tag{4}$$

The signal intensity measured in the mass spectrometer derives from fraction f(P) which reflect the protein amount bound by Kinobeads [probe.P]. Moreover, we assume that the fraction f(P) of a target P captured by the beads remains constant for subsequent pulldowns. Thus, a simple experiment can help to determine the protein depletion by a subsequent pulldown of pulldown experiment of the DMSO treated control lysate (Figure 17b). As described in formula (5), the first pulldown experiments depletes the total protein concentration by fraction f(P); the protein concentration available for the second pulldowns is the fraction 1-f(P)). That means that if the protein is not depleted substantially by the Kinobeads, then fraction f(P) is negligible and the fraction 1-f(P) equals $[P]_{total}$, and a subsequent pulldown experiment using the same lysate yields the same signal intensity as the first pulldown. However, if a larger amount of protein is depleted by the first pulldown experiment, the protein concentration available for the second pulldowns is the reduced fraction 1-f(P). The signal derived from the second pulldown decreases by the same fraction f(P) as protein has been depleted in the first experiment: fraction f(1-f(P)). This enables the calculation of a depletion factor r by dividing the intensity of the second pulldown experiment by the intensity of the first experiment as described in formula (5). A value close to 1 indicates low or no protein depletion, whereas a value close to 0 suggests high degree of protein depletion by the utilized chemical probe.

$$depletion\ r = \frac{pulldown\ 2}{pulldown\ 1} = \frac{signal\ intensity\ 2}{signal\ intensity\ 1} = \frac{fraction\ f\ (1-f(P))}{fraction\ f\ (P)} = 1 - fraction\ f(P) \tag{5}$$

Cheng and Prusoff[280] developed formula (6) which enables the conversion of an experimentally obtained EC_{50} value to the actual K_D constant for reversible competitive inhibition experiments.

$$K_{D\ (ligand)} = \frac{K_{D\ (probe)}}{K_{D\ (probe)} + [probe]} * EC_{50\ (ligand)} \tag{6}$$

Here, prior knowledge about the binding affinity of a target towards the Kinobeads $K_{D(probe)}$ and the concentration of the immobilized probe [probe] is required. The binding affinities of proteins towards Kinobeads is not known and not determined during the experiment. But, when combining the knowledge about protein depletion (4-5), the Cheng-Prusoff equation (6), and the binding equilibrium between Kinobeads [probe] and targets [P] (7), the conversion of EC_{50} to K_D is feasible for each individual protein in a simple manner as described below (8-11):

$$K_{D\ (probe)} = \frac{[probe]*[P]}{[probe.P]} \tag{7}$$

$$[probe] = K_{D\ (probe)} * \frac{[probe.P]}{[P]} = K_{D\ (probe)} * \frac{fraction\ f(P)}{(1-fraction\ f(P))} = K_{D\ (probe)} * \frac{1-r}{r} \tag{8}$$

$$K_{D\ (ligand)} = \frac{K_{D\ (probe)}}{K_{D\ (probe)} + [probe]} * EC_{50\ (ligand)} = \frac{K_{D\ (probe)}}{K_{D\ (probe)} + \left(K_{D\ (probe)} * \frac{1-r}{r}\right)} * EC_{50\ (ligand)} \tag{9}$$

$$K_{D\ (ligand)} = \frac{1}{1+\left(\frac{1-depletion\ r}{depletion\ r}\right)} * EC_{50\ (ligand)} = \frac{1}{\frac{r}{r}+\left(\frac{1-r}{r}\right)} * EC_{50\ (ligand)} = \frac{1}{\left(\frac{r+1-r}{r}\right)} * EC_{50\ (ligand)} \tag{10}$$

$$\boldsymbol{K_{D\ (ligand)} = depletion\ r * EC_{50\ (ligand)}} \tag{11}$$

Due to the native lysates used for Kinobeads selectivity profiling, the determined dissocation constant for a particular protein might cover binding events towards different protein isoforms, complexes or different activation states of the very same protein. In order to account for this ambiguity, the dissociation constant is referred to as apparent dissociation constant K_D^{app} in the following sections.

Kinetics and residence time. A major drawback of the Kinobeads technology is the missing link between binding equilibrium and binding kinetics.[281-285] The Kinobeads technology allows for the determination of binding affinities but not of association and dissociation constants. The relation between thermodynamics and kinetics is described by (12).

$$dissociation\ constant\ K_D = \frac{dissociation\ rate\ k_{off}}{association\ rate\ k_{on}} \tag{12}$$

In practice, affinity often corresponds to association/dissociation rates. Weak binders with K_D values in the micromolar-millimolar range often display rapid association and dissociation kinetics (half-lives in microsecond time scale), whereas nanomolar binders often feature half-lives from seconds to minutes or even hours.[154] However, strong affinity does not necessarily correlate with slow k_{off} rates. The time-scale of the dissociation constant correlates in a reciprocal manner with the ligand's residence time τ (13).[281]

$$residence\ time\ \tau = \frac{1}{k_{off}} \tag{13}$$

Residence times can be significantly different – even for ligands with similar binding affinities. One such example is given by the EGFR inhibitors Gefitinib (K_D: 0.4 nM), Erlotinib (K_D: 0.7 nM), and Lapatinib (K_D: 3 nM). Despite very similar binding affinities in the low nanomolar range, the residence times of these inhibitors varied from less than 10 min for gefitinib and erlotinib up to 300 min for Lapatinib.[286] It is hypothesized that this might be due to the different binding modes of these inhibitors, but this hypothesis still needs to be proven. For the clinical application of these inhibitors where constant drug dosing is difficult, prolonged drug residence times are beneficial for the therapy outcome and might even be more important for drug efficacy than the affinity of a drug for its target. Slow k_{off} rates are also beneficial for the Kinobeads experimental setup. Upon removal of the binding equilibrium during the washing step, fast k_{off} rates lead to loss of bound proteins and eventually to a decreased signal. The current setup allows profiling of binders in the nanomolar and up to low micromolar range. Weaker affinities are mostly too transient to survive the experimental procedure.

Enthalpy and entropy. Understanding the interaction between a ligand and a protein requires the view on the protein-ligand-solvent system as a multi-faceted and interdependent thermodynamic setting.[287] The stability – and thus, binding affinity – of a given ligand-protein complex is determined by the Gibbs free energy that is released upon binding as described in (14).

$$Gibbs\ free\ energy\quad \Delta G^0 = -RTlnK_A = RTlnK_D \tag{14}$$

The change of Gibbs free energy can be described by enthalpic ΔH and entropic ΔS contributions (15). Ligand-protein binding is influenced by both effects to different extents depending on the properties of both the ligand and the protein.

$$\Delta G = \Delta H - T\Delta S \tag{15}$$

Enthalpy is the measurement of energy or heat content in a thermodynamic system. In a simplified view, changes in binding enthalpy are driven by the formation of novel non-covalent interactions upon binding (e.g. hydrophobic van-der-Waals forces, hydrogen bonds or ion pairs). Establishment of these bonds leads to a release of enthalpy in an exothermic manner and positively contributes to the binding affinity. In contrary, entropy is a measure of the randomness and the disorder of molecules in a system. Entropy is affected by changes in degrees of freedom in a system: changes in the water shell surrounding protein and ligand, conformational entropy changes upon binding, or the change in translational or rotational degrees of freedom. Entropy contributes positively to Gibbs free energy if a higher disorder of the system is obtained, for instance if water molecules are released by ligand binding. Different ligands may bind with different enthalpic or entropic contributions and both can be favored by targeted ligand design.[288] Introduction of novel interaction sites will enhance enthalpic contribution (decrease enthalpy ΔH), whereas the generation of less flexible or symmetric ligands will favor entropic contribution to binding affinity (decrease entropy ΔS reduces penalty of entropy loss).

These considerations also influence the experimental settings of the Kinobeads technology and explain different thermodynamic and kinetic properties of immobilized and free compounds. Compared to the immobilized probe, the non-modified ligand features more degrees of freedom; binding of the non-modified ligand to its target protein signifies a large reduction in degrees of freedom. The immobilized compound on the other side has already lost translational and rotational degrees of freedom due to the immobilization. Binding of this more rigid compound will lead to a comparably lower entropic penalty on the Gibb's free energy ΔG (formula (15)), whereas the enthalpic contribution mainly established by molecular protein:ligand interactions should be the same. According to formula (14), lower Gibb's free energy corresponds to lower K_D. In turn, this affinity change should also affect the kinetic rates (formula (12)) and either correspond to a lower k_{off} or higher k_{on} rate. It is not clear yet which factors contribute to association and dissociation kinetics, but it was shown that restricting the conformational flexibility of compounds can lead to increasing k_{on} rates and that increasing molecule size can reduce k_{off} rates. Hence, it appears very likely that immobilization of a ligand affects kinetics in a way that binding happens faster or residence time is longer in comparison to the free ligand. Noteworthy, slow association/dissocation kinetics can distort measurement of binding affinity as they prevent establishment of a binding equilibrium. Thus, it is recommended to increase incubation times to at least 3-fold of the longest residence time,[289,290] which is most probably not achieved in typical pulldown experiments. This means that pre-incubation of the non-modified compound is inevitable for the experimental success, especially for classical pulldown experiments where a linkable version and a non-modified version of the same molecule are tested.

4 Objectives and outline

The receptor tyrosine kinase EPHA2 is implicated in a multitude of biological and pathological conditions and receives rising interest as therapeutic target. Selective small molecule kinase inhibitors could be a valuable approach to further validate EPHA2 as a drug target and provide means of therapeutic inhibition of EPHA2. The discovery of selective inhibitors benefits from early selectivity profiling that can be accomplished by mass spectrometry based chemical proteomics. The objective of this thesis was to harness the power of chemical proteomics in order to guide medicinal chemistry efforts aiming at the discovery of EPHA2 inhibitors. It illustrates how the Kinobeads technology supports several steps of drug discovery including lead selection and optimization, and further evaluation of the affinity and selectivity of newly synthesized molecules.

In this work, Kinobeads were utilized to elucidate the target space of 235 clinical kinase inhibitors. This dataset enabled the development of a novel selectivity scoring system named CATDS (*Concentration And Target Dependent Selectivity*) which facilitates the ranking of compounds according to their selectivity towards a particular target (chapter 1). The same dataset was further used to identify clinical EPHA2 inhibitors and enabled the selection of an appropriate lead structure for a medicinal chemistry program aiming at discovering better EPHA2 inhibitors (chapter 2). Target space and affinity information derived from the Kinobeads drug screen were merged with structure-affinity-relationship analysis and kinase sequence alignment to aid the development of a residue classification system (chapter 3). This analysis helped to categorize the potential of different amino acid residues within the ATP pocket for rational design of a selective kinase inhibitor. Both the residue classification and the possibility to calculate compound selectivity enabled the synthesis of dedicated EPHA2 inhibitors (chapter 4). Here, the molecular scaffold was decorated with different chemical moieties targeting particular residues within the ATP pocket of EPHA2. Finally, the biochemical and biological evaluation of the EPHA2 inhibitor candidates revealed promising candidates for targeted EPHA2 inhibition (chapter 5).

Experimental Procedures

Table of contents

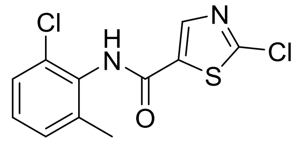
1 Inhibitor synthesis 41
2 Biochemistry, mass spectrometry and data processing 57
3 Cell culture 62

"I have not failed. I have just found 1000 ways that didn't work."
- Nikola Tesla -

1 Inhibitor synthesis

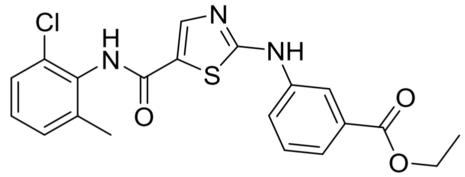
All chemicals and solvents were purchased from commercial suppliers (Sigma-Aldrich Co, VWR International, Carl Roth GmbH & Co.KG, Alfa Aeasar, Fluorochem Ltd) and were used without further purification. Flash chromatography was performed on an Interchim puriFlash evo 430 system. Reverse phase LC-MS was performed utilizing a TriArt C18 column and an Agilent 1100 series HPLC system, applying a 5-95% gradient of solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in water) for 20 min, followed by a 20 min plateau of 95% solvent B. Chromatography signals were detected at λ: 254nm and 365 nm. Mass analysis was accomplished by an amazon speed ETD ion trap mass spectrometer in positive mode. NMR spectra were recorded on a 400 MHz or 500 MHz Spectrometer (Bruker) at the Chair of Food Chemistry and Molecular Sensory Science (Prof. Hofmann) and the Chair of Biochemistry (Prof. Eisenreich, Prof. Groll) at the Technical University of Munich. Chemical shifts were recorded in parts per million (ppm) and NMR signals were described as following: s (single), d (doublet), t (triplet), q (quadruplet), and m (multiplet).

Preparation of 2-chloro-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (II).



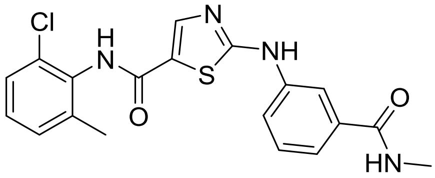
A solution of 2-chlorothiazole (compound **I**, 717 mg, 6.0 mmol) in tetrahydrofuran (10 ml), under argon, was cooled to -78 °C. A 2.5 M solution of n-butyllithium in hexanes (2.64 ml, 6.6 mmol) was added dropwise. The reaction was stirred at -78 °C for 15 min. A solution of 2-chloro-6-methylphenyl isocyanate (899 µl, 6.6 mmol) in THF (5 ml) was then added. The reaction was kept at -78 °C for 2 h and subsequently quenched with a solution of saturated aq. NH_4Cl (10 ml). The reaction was extracted with ethyl acetate (30 ml). The resulting organic layer was washed with saturated aq. NaCl (25 ml), dried over $MgSO_4$ and concentrated in vacuo. The obtained crude was purified by crystallization: petroleum ether was added to a solution of the crude in ethyl acetate to yield compound **II** (1.33 g, 77%) as a pale yellow solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 10.42 (s, 1H), 8.46 (s, 1H), 7.42 (dd, J = 7.0, 2.4 Hz, 1H), 7.30 (m, 2H), 2.23 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d6) δ 157.48, 154.46, 142.71, 138.53, 137.29, 132.54, 132.07, 129.21, 128.69, 127.14, 18.17; m/z [M+H]+ 286.9

Preparation of ethyl 3-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)benzoate (III).



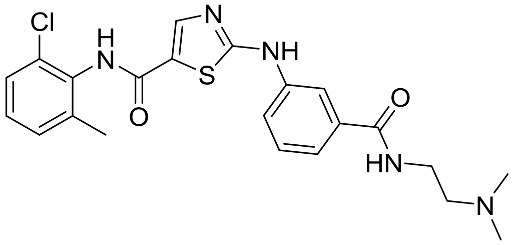
Compound **II** (230 mg, 0.8 mmol), ethyl 3-aminobenzoate (143 µl, 0.96 mmol) and (+)-camphor-10-sulfonic acid (186 mg, 0.8 mmol) were dissolved in 2-propanol (1.5 ml) and irradiated with microwave at 120 °C for 3 hours under argon and constant stirring. The reaction mixture was diluted with ethyl acetate (25 ml), washed with saturated aq. $NaHCO_3$ (25 ml) and saturated aq. NaCl (25 ml). The combined aqueous layers were further extracted with ethyl acetate (2 x 25 ml). The combined organic fractions were dried over $MgSO_4$ and concentrated in vacuo. Dichloromethane (2 ml) was added to the crude material and the obtained solid was collected by filtration and washed with dichloromethane (10 ml) to yield compound **III** (220 mg, 66%) as a bright white solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 10.85 (s, 1H), 9.95 (s, 1H), 8.29 (s, 1H), 8.16 (s, 1H), 7.94 (m, 2H), 7.60 (d, J = 7.8 Hz, 1H), 7.50 (t, J = 7.9 Hz, 1H), 7.40 (dd, J = 7.5, 1.7 Hz, 2H), 7.28 (m, 3H), 4.33 (q, J = 7.1 Hz, 2H), 2.24 (s, 3H), 1.33 (t, J = 7.1 Hz, 3H); ^{13}C NMR (101 MHz, DMSO-d6) δ 166.33, 165.58, 159.11, 142.24, 140.65, 138.76, 133.34, 132.36, 130.68, 129.50, 129.04, 128.23, 127.01, 123.12, 122.57, 121.92, 117.90, 60.82, 18.26, 14.15; m/z [M+H]+ : 416.0

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(methylcarbamoyl)phenyl)amino)thiazole-5-carboxamide (1a).



Compound **II** (10 mg, 35 µmol), 3-amino-N-methylbenzamide (6 mg, 38 µmol) and (+)-camphor-10-sulfonic acid (186 mg, 52 µmol) were dissolved in 2-propanol (0.2 ml) and irradiated with microwave at 140 °C for 3 hours under argon and constant stirring. The reaction mixture was diluted with ethyl acetate (10 ml), washed with saturated aq. $NaHCO_3$ (10 ml) and saturated aq. NaCl (10 ml). The combined aqueous layers were further extracted with ethyl acetate (2 x 10 ml). The combined organic fractions were dried over $MgSO_4$ and concentrated in vacuo. Dichloromethane (0.5 ml) was added to the crude material and the obtained solid was collected by filtration and washed with dichloromethane (10 ml) to yield inhibitor **1a** (4 mg, 28%) as a white solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 10.75 (s, 1H), 9.94 (s, 1H), 8.41 (q, J = 4.3 Hz, 1H), 8.15 (s, 1H), 8.08 (s, 1H), 7.79 (d, J = 7.1 Hz, 1H), 7.41 (m, 3H), 7.28 (m, 2H), 2.78 (d, J sq= 4.5 Hz, 3H), 2.23 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d6) δ 170.20, 166.58, 159.17, 142.33, 140.42, 138.77, 135.65, 133.37, 132.37, 129.04, 129.00, 128.23, 127.01, 120.39, 120.35, 120.08, 116.76, 26.29, 18.28; m/z [M+H]+ : 401.0

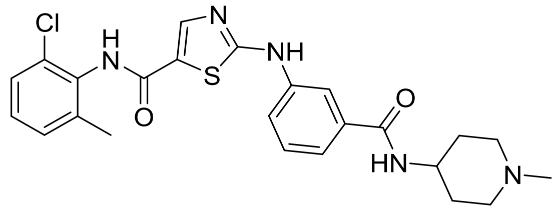
Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((2-(dimethylamino)ethyl)carbamoyl)phenyl)amino) thiazole-5-carboxamide (1b).



DABAL-Me_3 (61 mg, 0.24 mmol) was suspended in tetrahydrofuran (0.5 ml) and N,N-dimethylethylenediamine (392 µl, 3.6 mmol) was added dropwise at room temperature. The reaction was heated to 40 °C and stirred for 1 hour under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (1 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 120 °C for 15 hours, then quenched with saturated aq. NaCl (25 ml) and extracted with a solution of 10% (2 N NH_3 in methanol) in dichloromethane (2 x 25 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. Purification on silica (12 g cartridge, dichloromethane neat to 10% (2 N NH_3 in methanol) in DCM) yielded inhibitor **1b** (47 mg, 85%) as a white solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.12 (s, 1H), 8.07 (s, 1H), 7.75 (d, J = 7.8 1H), 7.50 (d, J = 7.8 1H), 7.44 (t, J = 7.8 Hz, 1H), 7.35 (dd, J = 7.2, 2.0 Hz, 1H), 7.24 (m, 2H), 3.56 (t, J = 6.8 Hz, 2H), 2.65 (t, J = 6.8 Hz, 2H), 2.37 (s, 6H), 2.32 (s, 3H), 2.10 (s, 1H); ^{13}C NMR (101 MHz, Methanol-d4) δ 170.11, 169.50, 162.44, 143.83, 141.96, 140.35, 136.82, 134.28, 134.24, 130.38, 130.15, 130.09, 129.60, 128.34, 122.48, 122.45, 118.45, 59.17, 45.44, 38.49, 18.68; m/z [M+H]+ : 458.1

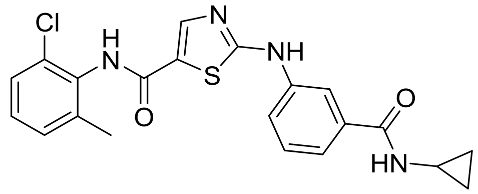
Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((1-methylpiperidin-4-yl)carbamoyl)phenyl)amino) thiazole-5-carboxamide (1c).

DABAL- Me_3 (61 mg, 0.24 mmol) was suspended in tetrahydrofuran (0.5 ml) and 4-amino-1-methylpiperidine (100 µl, 0.8 mmol) was added dropwise at room temperature. The reaction



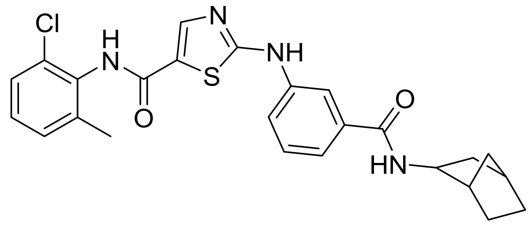
was heated to 40 °C and stirred for 1 hour under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (1 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 120 °C for 15 hours, then quenched with saturated aq. NaCl (25 ml) and extracted with a solution of 10% (2 N NH_3 in methanol) in dichloromethane (2 x 25 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. Purification on silica (dichloromethane neat to 10% (2 N NH_3 in methanol) in dichloromethane) yielded inhibitor **1c** (9 mg, 16%) as a white solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.07 (m, 2H), 7.73 (dd, J = 7.8, 2.2 Hz, 1H), 7.46 (m, 2H), 7.35 (dd, J = 7.4, 2.1 Hz, 1H), 7.25 (m, 2H), 3.92 (m, 1H), 2.96 (m, 2H), 2.34 (m, 8H), 1.98 (m, 2H), 1.74 (m, 2H); ^{13}C NMR (101 MHz, Methanol-d4) δ 169.83, 169.56, 162.42, 143.75, 141.83, 140.35, 137.15, 134.27, 134.24, 130.34, 130.15, 129.60, 128.34, 123.77, 122.63, 122.44, 118.54, 55.56, 45.97, 45.90, 32.06, 18.68; m/z [M+H]+ : 484.1

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(cyclopropylcarbamoyl)phenyl)amino)thiazole-5-carboxamide (1d).



DABAL-Me_3 (154 mg, 0.6 mmol) was suspended in tetrahydrofuran (0.5 ml) and cyclopropylamine (251 µl, 3.6 mmol) was added dropwise at room temperature. The reaction was heated to 40°C and stirred for 30 min under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 100 °C for 1 hour, then quenched with 2 N HCl (25 ml) and extracted with ethyl acetate (3 x 25 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. Purification on silica (12 g cartridge, dichloromethane neat to 10% methanol in dichloromethane) yielded inhibitor **1d** (32 mg, 63%) as a white solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 10.75 (s, 1H), 9.95 (s, 1H), 8.44 (d, J = 4.2 Hz, 1H), 8.14 (s, 1H), 8.04 (s, 1H), 7.80 (d, J = 6.9 Hz, 1H), 7.40 (m, 3H), 7.27 (m, 2H), 2.85 (m, 1H), 2.23 (s, 3H), 0.69 (m, 2H), 0.57 (m, 2H); ^{13}C NMR (101 MHz, DMSO-d6) δ 167.52, 166.61, 159.22, 142.35, 140.40, 138.81, 135.60, 133.40, 132.40, 129.08, 128.94, 128.27, 127.05, 122.84, 120.60, 120.18, 116.87, 23.10, 18.30, 5.75; m/z [M+H]+ : 426.1

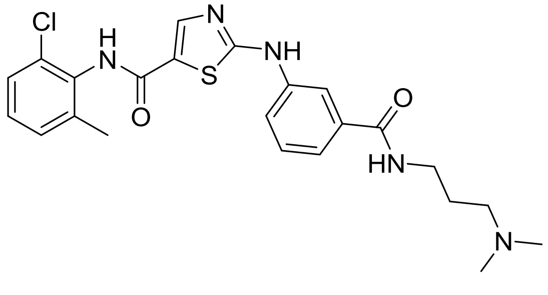
Preparation of 2-((3-((1*S*,4*R*)-bicyclo[2.2.1]heptan-2-ylcarbamoyl)phenyl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (1e).



DABAL- Me_3 (92 mg, 0.36 mmol) was suspended in tetrahydrofuran (0.5 ml) and exo-2-norbornanamine (428 µl, 3.6 mmol) was added dropwise at room temperature. The reaction was heated to 40 °C and stirred for 30 min under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 100 °C for 2 hours, then quenched with 2 N HCl (25 ml) and extracted with ethyl acetate (3 x 25 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. Purification on silica (dichloromethane neat to 10% methanol in dichloromethane) yielded inhibitor **1e** (52 mg, 91%) as a white solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.06 (s, 1H), 8.03 (s, 1H), 7.70 (d, J = 7.4, 2.1 Hz, 1H), 7.43 (m, 2H), 7.35 (dd, J = 7.2, 2.2 Hz, 1H), 7.23 (m, 2H), 3.81 (m, 1H), 2.31 (s, 3H), 1.79

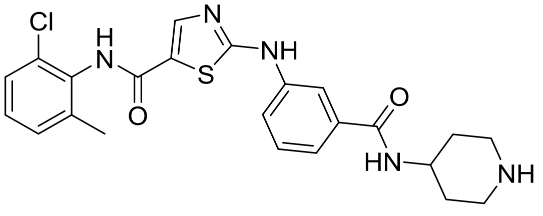
(m, 1H), 1.56 (m, 5H), 1.25 (m, 4H); ^{13}C NMR (101 MHz, Methanol-d4) δ 169.95, 169.64, 162.43, 143.72, 141.73, 140.35, 137.48, 134.28, 134.26, 130.31, 130.14, 129.58, 128.33, 123.76, 122.74, 122.30, 118.56, 55.19, 43.63, 39.86, 37.06, 36.23, 29.36, 27.61, 18.67; m/z [M+H]+ : 481.1

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((3-(dimethylamino)propyl)carbamoyl)phenyl)amino) thiazole-5-carboxamide (1f).



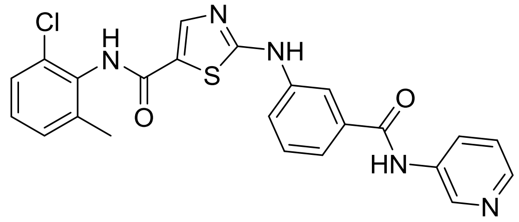
DABAL- Me$_3$ (92 mg, 0.36 mmol) was suspended in tetrahydrofuran (0.5 ml) and 4-amino-1-methylpiperidine (146 µl, 1.2 mmol) was added dropwise at room temperature. The reaction was heated to 40 °C and stirred for 30 min under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (1 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 100 °C for 4 hours, then quenched with saturated aq. NaCl (25 ml) and extracted with a solution of 10% (2 N NH$_3$ in methanol) in dichloromethane (2 x 25 ml). The combined organic phase was dried over MgSO$_4$ and concentrated in vacuo. Purification on silica (dichloromethane neat to 10% (2 N NH$_3$ in methanol) in dichloromethane) yielded inhibitor **1f** (44 mg, 78%) as a white solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.10 (s, 1H), 8.07 (s, 1H), 7.74 (d, J = 7.2 Hz, 1H), 7.45 (m, 2H), 7.35 (dd, J = 7.0, 2.1 Hz, 1H), 7.24 (m, 2H), 3.43 (t, J = 7.0 Hz, 2H), 2.46 (t, J = 7.0 2H), 2.32 (s, 3H), 2.30 (s, 6H), 1.83 (q, J = 7.0 2H); ^{13}C NMR (101 MHz, Methanol-d4) δ 170.12, 169.56, 162.45, 143.81, 141.98, 140.35, 137.07, 134.29, 134.26, 130.40, 130.18, 130.15, 129.59, 128.35, 122.42, 122.41, 118.39, 58.33, 45.40, 39.38, 28.04, 18.66; m/z [M+H]+ : 472.1

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(piperidin-4-ylcarbamoyl)phenyl)amino)thiazole-5-carboxamide (1g).



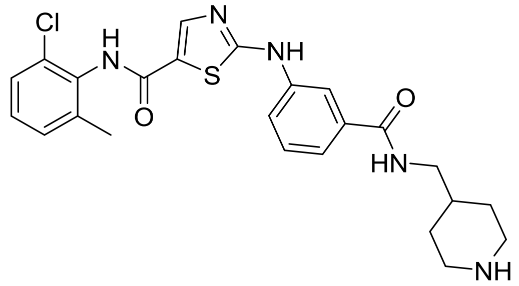
DABAL-Me$_3$ (123 mg, 0.48 mmol) was suspended in tetrahydrofuran (0.5 ml) and 4-aminopiperidine (252 µl, 2.4 mmol) was added dropwise at room temperature. The reaction was heated to 40 °C and stirred for 30 min under argon. Compound **III** (100 mg, 0.24 mmol) was dissolved in tetrahydrofuran (1 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 120 °C for 2 hours, then quenched with saturated aq. NaHCO$_3$ (25 ml) and extracted with a solution of 10% (2 N NH$_3$ in methanol) in dichloromethane (2 x 25 ml). The combined organic phase was dried over MgSO$_4$ and concentrated in vacuo. Dichloromethane (1 ml) was added to the crude material and the obtained solid was collected by filtration and washed with dichloromethane (10 ml) to yield inhibitor **1g** (50 mg, 44%) as a yellow solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 9.95 (s, 1H), 8.26 (d, J = 7.8 Hz, 1H), 8.15 (s, 1H), 8.03 (s, 1H), 7.84 (d, J = 7.8 1H), 7.43 (m, 3H), 7.28 (m, 2H), 3.83 (m, 1H), 2.98 (m, 2H), 2.55 (m, 2H), 2.23 (s, 3H), 1.74 (m, 2H), 1.44 (m, 2H); ^{13}C NMR (101 MHz, DMSO) δ 166.58, 165.41, 159.17, 142.32, 140.33, 138.76, 135.94, 133.39, 132.37, 129.03, 128.83, 128.20, 127.00, 122.77, 120.67, 120.02, 117.04, 47.26, 45.10, 32.58, 18.27; m/z [M+H]+ : 470.1

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(pyridin-3-ylcarbamoyl)phenyl)amino)thiazole-5-carboxamide (1h).



DABAL-Me$_3$ (153 mg, 0.6 mmol) and 3-aminopyridine (338 mg, 3.6 mmol) were suspended in tetrahydrofuran (0.5 ml) at room temperature. The reaction was heated to 40 °C and stirred for 30 min under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 140 °C for 6 hours, then quenched with saturated aq. NaCl (25 ml) and extracted with ethyl acetate (3 x 25 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. Dichloromethane (1 ml) was added to the crude material and the obtained solid was collected by filtration and washed with dichloromethane (10 ml) to yield inhibitor **1h** (12.7 mg, 23%) as a white yellow solid: ^{1}H NMR (500 MHz, DMSO-d6) δ 10.84 (s, 1H), 10.48 (s, 1H), 9.95 (s, 1H), 8.93 (d, J = 2.5 Hz, 1H), 8.32 (dd, J = 4.7, 1.4 Hz, 1H), 8.19 (m, 3H), 7.89 (dd, J = 8.2, 2.4 Hz, 1H), 7.62 (d, J = 7.7 Hz, 1H), 7.53 (t, J = 7.9 Hz, 1H), 7.40 (dd, J = 8.4, 4.7 Hz, 2H), 7.27 (m, 2H), 2.24 (s, 3H); ^{13}C NMR (126 MHz, DMSO-d6) δ 166.45, 165.92, 159.12, 144.58, 142.27, 141.95, 140.57, 138.75, 135.76, 135.43, 133.34, 132.35, 129.17, 129.02, 128.21, 127.28, 126.99, 123.50, 123.02, 121.02, 120.77, 117.01, 18.25; m/z [M+H]+ : 464.1

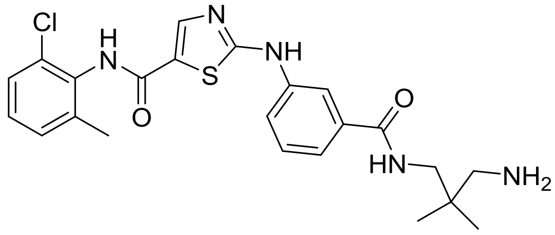
Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((piperidin-4-ylmethyl)carbamoyl)phenyl)amino) thiazole-5-carboxamide (1i).



DABAL-Me$_3$ (123 mg, 0.48 mmol) was suspended in tetrahydrofuran (0.5 ml) and 4-(aminomethyl)piperidine (306 µl, 2.4 mmol) was added dropwise at room temperature. The reaction was heated to 40 °C and stirred for 30 min under argon. Compound **III** (100 mg, 0.24 mmol) was dissolved in tetrahydrofuran (1 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 120 °C for 3 hours, then quenched with saturated aq. $NaHCO_3$ (25 ml) and extracted with a solution of 10% (2 N NH_3 in methanol) in dichloromethane (2 x 25 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. Dichloromethane (1 ml) was added to the crude material and the obtained solid was collected by filtration and washed with dichloromethane (10 ml) to yield inhibitor **1e** (60 mg, 52%) as a yellow solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 9.95 (s, 1H), 8.44 (t, *J* = 5.8 Hz, 1H), 8.15 (s, 1H), 8.06 (s, 1H), 7.82 (d, *J* = 7.8 Hz, 1H), 7.52 – 7.34 (m, 3H), 7.34 – 7.17 (m, 2H), 3.35 (s, 2H), 3.12 (t, *J* = 6.2 Hz, 2H), 2.94 (d, *J* = 12.0 Hz, 2H), 2.60 (s, 1H), 2.43 (t, *J* = 12.0 Hz, 2H), 2.24 (s, 3H), 1.61 (d, *J* = 12.0 Hz, 2H), 1.05 (dd, *J* = 12.0, 3.8 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d6) δ 167.05, 166.75, 159.66, 142.80, 140.89, 139.26, 136.37, 133.88, 132.87, 129.53, 129.40, 128.71, 127.50, 123.28, 120.98, 120.51, 117.38, 47.42, 46.17, 36.71, 31.22, 18.76; m/z [M+H]+ : 484.1

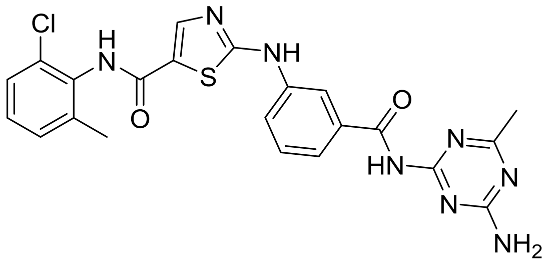
Preparation of 2-((3-((3-amino-2,2-dimethylpropyl)carbamoyl)phenyl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (1j).

DABAL-Me$_3$ (96 mg, 0.36 mmol) and 2,2-dimethylpropane-1,3-diamine (288 µl, 2.4 mmol) were suspended in tetrahydrofuran (0.5 ml). The reaction was heated to 40 °C and stirred for 30 min under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and



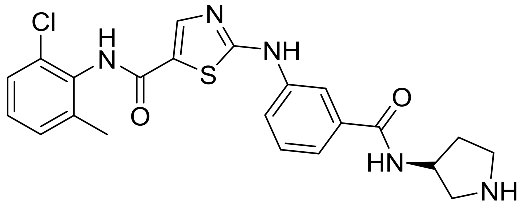
added to the reaction tube. The mixture was further irradiated with microwave at 100 °C for 8 hours, then quenched with saturated aqueous NaHCO$_3$ (25 ml) and extracted with a solution of 10% (2 N NH$_3$ in methanol) in dichloromethane (3 x 25 ml). The combined organic phase was dried over MgSO$_4$ and concentrated in vacuo. Purification on silica (12 g cartridge, dichloromethane neat to 10% (2N NH$_3$ in methanol) in dichloromethane) yielded inhibitor **1j** (11.7 mg, 21%) together with the cyclized side product *N*-(2-chloro-6-methylphenyl)-2-((3-(5,5-dimethyl-1,4,5,6-tetrahydropyrimidin-2-yl)phenyl)amino)thiazole-5-carboxamide (3.4 mg, 5.2%): ^{1}H NMR (500 MHz, Methanol-d4) δ 8.14 (s, 1H), 8.07 (s, 1H), 7.72 (d, J = 7.7 Hz, 1H), 7.46 (dt, J = 15.5, 7.7 Hz, 2H), 7.34 (dd, J = 9.1, 7.7 Hz, 1H), 7.27 – 7.19 (m, 2H), 3.29 (s, 2H), 2.52 (s, 2H), 2.31 (s, 3H), 1.00 (s, 6H); ^{13}C NMR (126 MHz, Methanol-d4) δ 171.12, 169.64, 162.57, 143.90, 142.10, 140.48, 137.08, 134.39, 130.58, 130.29, 129.75, 128.49, 123.96, 122.62, 122.61, 118.56, 50.17, 47.91, 37.53, 30.93, 24.11, 18.83; m/z [M+H]+ : 472.2

Preparation of 2-((3-((4-amino-6-methyl-1,3,5-triazin-2-yl)carbamoyl)phenyl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (1k).



DABAL-Me$_3$ (102 mg, 0.40 mmol) and 6-methyl-1,3,5-triazine-2,4-diamine (316 mg, 2.5 mmol) were suspended in tetrahydrofuran (0.5 ml). The reaction was heated to 40 °C and stirred for 1 h under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 115 °C for 2.8 h and at 120°C for 12.8 h, then quenched with saturated aqueous NaHCO$_3$ (25 ml) and extracted with a solution of 10% (2 N NH$_3$ in methanol) in dichloromethane (3 x 25 ml). The combined organic phase was washed with ddH$_2$O (5 x 50 ml), dried over MgSO$_4$ and concentrated in vacuo. Purification on silica (4 g cartridge, dichloromethane neat to 10% (2N NH$_3$ in methanol) in dichloromethane) yielded inhibitor **1k** (5.2 mg, 2.6%). ^{1}H NMR (400 MHz, DMSO-d6) δ 9.97 (s, 1H), 8.17 – 8.10 (m, 2H), 7.55 (d, *J* = 8.1 Hz, 1H), 7.26 (d, *J* = 7.5 Hz, 2H), 2.25 (s, 3H), 2.23 (s, 3H), 7.37 – 7.31 (m, 1H), 10.82 (s, 1H), 10.62 (s, 1H), 7.86 (dd, *J* = 7.9, 2.6 Hz, 1H), 7.45 (t, *J* = 7.9 Hz, 1H), 10.03 – 9.90 (m, 1H), 7.39 (dd, *J* = 7.5, 2.0 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-d6) δ 176.26, 167.38, 166.61, 165.77, 164.40, 159.33, 142.40, 140.48, 138.89, 135.32, 133.45, 132.48, 129.23, 129.18, 128.40, 127.14, 123.07, 121.88, 121.24, 117.56, 24.93, 18.38; m/z [M+H]+ : 495.1

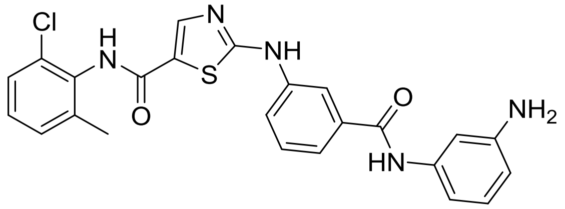
Preparation of (*R*)-*N*-(2-chloro-6-methylphenyl)-2-((3-(pyrrolidin-3-ylcarbamoyl)phenyl)amino)thiazole-5-carboxamide (1l).



DABAL- Me$_3$ (92 mg, 0.36 mmol) and (*S*)-pyrrolidin-3-amine (250 mg, 1.57 mmol) were dissolved in tetrahydrofuran (0.5 ml). The reaction was heated to 40 °C and stirred for 1 h under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 100 °C for 6 h, at 110 °C for 2h and at 120 °C for 2 h. After addition of an extra eq.

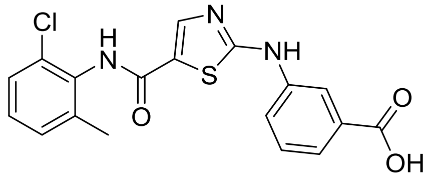
DABAL- Me_3 and tetrahydrofuran (0.5 ml) the reaction was further irradiated at 120 °C for 20 h. Reaction was quenched with saturated aqueous $NaHCO_3$ (25 ml) and extracted with a solution of 10% (2 N NH_3 in methanol) in dichloromethane (3 x 25 ml and 3 x 50 ml). The combined organic phase was washed with ddH_2O (5 x 50 ml), dried over $MgSO_4$ and concentrated in vacuo. Purification on silica (4 g cartridge, dichloromethane neat to 10% (2N NH_3 in methanol) in dichloromethane) yielded inhibitor **1l** (9.4 mg, 17%): ^{1}H NMR (500 MHz, Methanol-*d₄*) δ 8.06 (s, 1H), 7.93 (d, *J* = 8.6 Hz, 1H), 7.68 – 7.59 (m, 1H), 7.44 (t, *J* = 7.9 Hz, 1H), 7.35 (dd, *J* = 7.6, 2.0 Hz, 1H), 7.23 (ddt, *J* = 12.6, 9.1, 6.6 Hz, 3H), 3.86 – 3.75 (m, 1H), 3.75 – 3.61 (m, 2H), 3.61 – 3.39 (m, 2H), 2.31 (s, 3H), 2.28 – 2.10 (m, 1H), 1.84 (ddt, *J* = 18.4, 12.3, 6.3 Hz, 1H).; ^{13}C NMR (101 MHz, Methanol-d_4) δ 162.47, 143.93, 141.91, 140.34, 138.70, 134.27, 134.24, 130.42, 130.14, 129.60, 128.34, 123.74, 122.27, 121.02, 120.94, 117.90, 117.83, 45.81, 34.67, 32.99, 18.68, 17.10; m/z [M+H]+ : 456.2

Preparation of 2-((3-((3-aminophenyl)carbamoyl)phenyl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (1m).



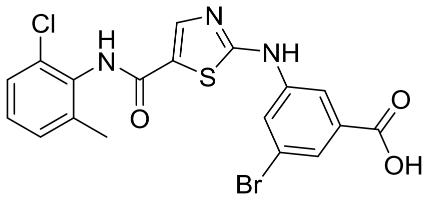
DABAL- Me_3 (106 mg, 0.41 mmol) and benzene-1,3-diamine (260 mg, 2.40 mmol) were dissolved in tetrahydrofuran (1 ml). The reaction was heated to 40 °C and stirred for 1 h under argon. Compound **III** (50 mg, 0.12 mmol) was dissolved in tetrahydrofuran (0.5 ml) and added to the reaction tube. The mixture was further irradiated with microwave at 100 °C for 2 h. The reaction was quenched with saturated aqueous $NaHCO_3$ (25 ml) and extracted with a solution of 10% (2 N NH_3 in methanol) in dichloromethane (4 x 25 ml). The combined organic phase was washed with ddH_2O (2 x 100 ml), dried over $MgSO_4$ and concentrated in vacuo. Inhibitor **1m** was precipitated as grey powder by dichloromethane (27.5 mg, 48%): ^{1}H NMR (500 MHz, Methanol-d4) δ 8.17 (s, 1H), 8.08 (s, 1H), 7.76 (d, J = 7.3 Hz, 1H), 7.55 (dd, J = 15.0, 5.4 Hz, 1H), 7.48 (t, J = 7.9 Hz, 1H), 7.34 (dt, J = 11.2, 5.5 Hz, 1H), 7.29 – 7.19 (m, 2H), 7.16 (d, J = 1.7 Hz, 1H), 7.09 (t, J = 8.0 Hz, 1H), 7.00 – 6.94 (m, 1H), 6.57 – 6.51 (m, 1H), 2.32 (s, 3H); ^{13}C NMR (101 MHz, Methanol-d_4) δ 169.52, 168.74, 162.44, 149.40, 143.76, 141.89, 140.52, 140.35, 137.85, 134.26, 134.25, 130.45, 130.34, 130.15, 129.61, 128.35, 123.84, 122.84, 122.51, 118.58, 113.09, 112.15, 109.38, 18.68; m/z [M+H]+ : 478.1

Preparation of 3-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)benzoic acid (Va).



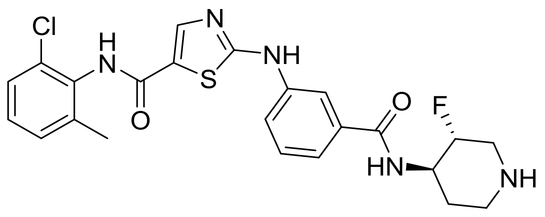
Compound **II** (368 mg, 1.28 mmol) and (+)-camphor-10-sulfonic acid (328 mg, 1.41 mmol) were dissolved in tert-butanol (2.5 ml). 3-Aminobenzoic acid (193 mg, 1.41 mmol) was added and the mixture was irradiated with microwave for 3.5 h at 120 °C under argon. The mixture was diluted with ethyl acetate (10 ml), washed with sat. aq. NH_4Cl (20 ml) and sat. aq. NaCl (20 ml). The combined aqueous phases were extracted with ethyl acetate (3 x 20 ml). Compound **Va** was obtained as white solid by filtration of the combined organic phases (383.6 mg, 77%): ^{1}H NMR (360 MHz, DMSO-d6) δ 13.00 (s, 1H), 10.88 (s, 1H), 9.98 (s, 1H), 8.30 (d, *J* = 2.2 Hz, 1H), 8.17 (s, 1H), 7.88 (dd, *J* = 8.1, 2.4 Hz, 1H), 7.58 (d, *J* = 7.0 Hz, 1H), 7.46 (td, *J* = 7.9, 1.7 Hz, 1H), 7.40 (dd, *J* = 7.4, 2.2 Hz, 1H), 7.33 – 7.20 (m, 2H), 2.24 (s, 3H); ^{13}C NMR (91 MHz, DMSO-d6) δ 167.13, 166.37, 159.14, 142.27, 140.56, 138.76, 133.38, 132.38, 131.60, 129.25, 129.00, 128.19, 126.98, 122.96, 122.83, 121.66, 118.10, 18.26; m/z [M+H]+ : 388.0

Preparation of 3-bromo-5-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)benzoic acid (Vb).



Compound **II** (1.3 g, 4.63 mmol) and (+)-camphor-10-sulfonic acid (1.0 g, 4.63 mmol) were dissolved in tert-butanol (10 ml) and put under argon atmosphere. 3-Amino-5-bromobenzoic acid (1.0 g, 4.63 mmol) was added slowly, and the mixture was irradiated with microwave at 120 °C for 3.5 h under constant stirring. It was diluted with ethyl acetate (10 ml), washed with saturated aqueous NH_4Cl (20 ml) and with saturated aq. NaCl (20 ml). the combined aqueous phases were extracted with ethyl acetate (3 x 20 ml). The organic layer was dried over $MgSO_4$ and the solvent was evaporated. The crude was dissolved in ethyl acetate (5 ml), cooled down and crystallized by adding chilled petroleum ether. After filtration, compound **Vb** (2.0 g, 92%) was obtained as a brown solid: ^{1}H NMR (500 MHz, Methanol-d4) δ 8.28 (t, J = 1.8 Hz, 1H), 8.16 (t, J = 1.6 Hz, 1H), 8.10 (s,1H), 7.77 (t, J = 1.4 Hz, 1H), 7.35 (dd, J = 7.5, 1.4 Hz, 1H), 7.24 (m, 2H), 2.32 (s, 3H); ^{13}C NMR (126 MHz, Methanol-d_4) δ 167.13, 166.86, 163.46, 160.98, 142.38, 141.80, 138.93, 133.59, 132.84, 128.74, 128.21, 126.94, 125.61, 123.93, 123.20, 122.16, 117.40, 17.26; m/z [M+H]+ 465.8

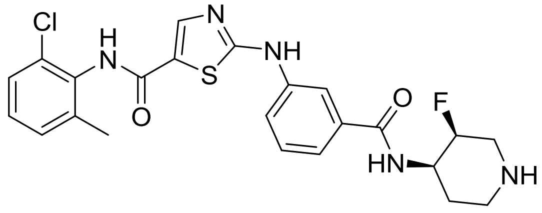
Preparation of N-(2-chloro-6-methylphenyl)-2-((3-((3,4-trans-3-fluoropiperidin-4-yl)carbamoyl)phenyl)amino)thiazole-5-carboxamide racemate (2a).



Compound **Va** (59 mg, 0.15 mmol) and *tert*-butyl-3,4-trans-4-amino-3-fluoro-piperidine-1-carboxylate racemate (31 mg, 0.14 mmol) were dissolved in dimethylformamide (1 ml) and cooled to 0 °C. *N,N*-diisopropylethylamine (36.1 µl, 0.21 mmol) and triethylamine (28.6 µl, 0.21 mmol) were added to the reaction mixture. PyBrop (96.8 mg, 0.21 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0°C and also added to the mixture. The reaction was quenched after 30 min by addition of 15 ml ddH_2O. Dichloromethane (15 ml) was added and the organic phase was washed with ddH_2O (5 x 15 ml). The combined aqueous phase was extracted with dichloromethane (3 x 10 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. The boc-protected intermediate was purified via flash chromatography using a gradient of 0-100% ethyl acetate in petroleum ether (65.2 mg): ^{1}H NMR (360 MHz, Methanol-d_4) δ 8.09 (s, *J* = 2.2 Hz, 1H), 8.07 (s, 1H), 7.80 – 7.69 (d, 1H), 7.52 – 7.41 (m, 2H), 7.39 – 7.32 (ddz, 1H), 7.28 – 7.18 (m, 2H), 4.90 (m, 1H), 4.55 (m, *J* = 193.8 Hz, 1H), 4.42 (m, *J* = 9.0, 4.7, 0.0 Hz, 1H), 4.34 – 4.15 (m, 2H), 3.98 (d, *J* = 14.0 Hz, 1H), 2.32 (s, 3H), 2.04 (dd, *J* = 11.9, 7.3 Hz, 1H), 1.69 – 1.56 (m, 1H), 1.48 (d, *J* = 1.3 Hz, 9H).

The boc-protected intermediate (46 mg, 0.08 mmol) was suspended in dichloromethane (3 ml) and cooled down to 0 °C. Trifluoroacetic acid was added and the reaction mixture was stirred at RT until completion. Solvent and trifluoroacetic acid were removed by evaporation. The crude was repeatedly dissolved in acetonitrile and evaporated to remove remaining trifluoroacetic acid. Trituration with diethylether yielded inhibitor **2a** as a white powder (32.7 mg, 85%): ^{1}H NMR (360 MHz, Methanol-d_4) δ 8.13 (t, *J* = 1.9 Hz, 1H), 8.08 (s, 1H), 7.72 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.53 (td, 1H), 7.46 (t, *J* = 7.9 Hz, 1H), 7.35 (dd, *J* = 7.2, 2.3 Hz, 1H), 7.28 – 7.18 (m, 2H), 5.12 – 5.01 (m, 1H), 4.49 – 4.35 (m, 1H), 3.65 (ddd, *J* = 20.8, 13.3, 3.4 Hz, 1H), 3.52 – 3.35 (m, 2H), 3.30 – 3.20 (m, 1H), 2.42 – 2.33 (m, 1H), 2.31 (s, 3H), 2.09 – 1.91 (m, 1H); ^{13}C NMR (91 MHz, Methanol-d_4) δ 170.56, 169.58, 160.65, 143.34, 141.78, 140.32, 136.54, 134.20, 130.48, 130.15, 129.62, 128.33, 123.03, 122.94, 118.80, 87.40, 85.43, 45.49, 45.20, 42.22, 25.47, 25.43, 18.6; m/z [M+H]+ : 488.1

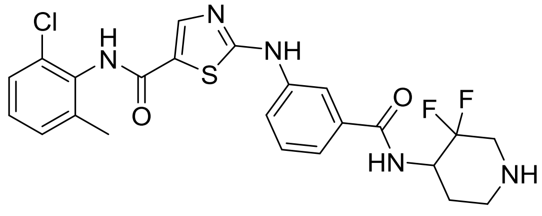
Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((3,4-cis-3-fluoropiperidin-4-yl)carbamoyl)phenyl)amino)thiazole-5-carboxamide racemate (2b).



Compound **Va** (58 mg, 0.15 mmol) and *tert*-butyl-3,4-cis-4-amino-3-fluoro-piperidine-1-carboxylate racemate (32 mg, 0.15 mmol) were dissolved in dimethylformamide (1ml) and cooled to 0 °C. *N,N*-diisopropylethylamine (36.1 µl, 0.21 mmol) and triethylamine (28.6 µl, 0.21 mmol) were added to the reaction mixture. PyBrop (103 mg, 0.22 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0 °C and also added to the mixture. The reaction was quenched after 40 min by addition of 15 ml ddH_2O. Dichloromethane (15 ml) was added and the organic phase was washed with ddH_2O (5 x 15 ml). The combined aqueous phase was extracted with dichloromethane (3 x 10 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. The crude was purified via flash chromatography (0-100% ethyl acetate in petroleum ether) and the boc-protected intermediate was obtained as a white powder (28 mg, 33%).

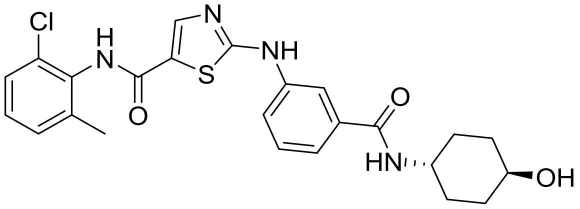
Boc-protected intermediate (15 mg, 0.03 mmol) was suspended in dichloromethane (1.5 ml) and cooled down to 0 °C. A solution of 2N trifluoroacetic acid in dichloromethane (1 ml) was added dropwise and stirred at room temperature for 30 min and at 35 °C for 30 min. Organic solvent and trifluoroacetic acid were removed by evaporation. Trituration with diethyl ether yielded inhibitor **2b** as a white solid (quantitative yield): ^{1}H NMR (360 MHz, Methanol-d4) δ 8.17 – 8.09 (m, 1H), 8.06 (s, 1H), 7.78 – 7.66 (m, 1H), 7.55 – 7.43 (m, 2H), 7.35 (dd, J = 7.1, 2.5 Hz, 1H), 7.29 – 7.20 (m, 2H), 5.15 (d, J = 47.6 Hz, 1H), 4.41 (dddd, J = 31.1, 12.5, 4.8, 2.0 Hz, 1H), 3.74 (ddt, J = 11.8, 9.6, 2.6 Hz, 1H), 3.60 – 3.36 (m, 2H), 3.24 (dd, J = 13.2, 3.3 Hz, 1H), 2.32 (s, 3H), 2.25 (dd, J = 13.2, 4.4 Hz, 1H), 2.09 (d, J = 13.8 Hz, 1H); m/z [M+H]+ : 488.1

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((3,3-difluoropiperidin-4-yl)carbamoyl)phenyl)amino)thiazole-5-carboxamide (2c).



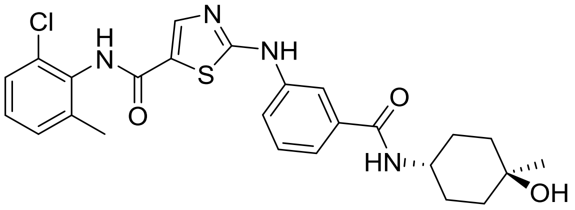
Compound **Va** (58 mg, 0.15 mmol) and 3,3-difluoropiperidine-4-amine dihydrochloride (32 mg, 0.15 mmol) were dissolved in dimethylformamide (1ml) and cooled to 0 °C. *N,N*-diisopropylethylamine (34.0 µl, 0.19 mmol) and triethylamine (26.9 µl, 0.19 mmol) were added to the reaction mixture. PyBrop (103 mg, 0.22 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0 °C and also added to the mixture. The reaction was quenched after 140 min by addition of ddH_2O (15 ml). The reaction mixture was diluted with 15% (1.33N NH_3 in methanol) in dichloromethane (20 ml) and washed with ddH_2O (6 x 20 ml). The aqueous phase was extracted with 15% (1.33N NH_3 in methanol) in dichloromethane (2 x 15 ml). The combined organic phase was dried over $MgSO_4$ and concentrated in vacuo. The crude was purified by flash chromatography (dichloromethane to 15% (1.33 M NH_3 in methanol) in dichloromethane) yielding inhibitor **2c** as white solid (2.88 mg, 4%): ^{1}H NMR (500 MHz, DMSO-d6) δ 10.83 (s, 1H), 9.98 (s, 1H), 8.15 (s, 1H), 7.79 (s, 1H), 7.66 (d, *J* = 8.1 Hz, 1H), 7.51 – 7.34 (m, 2H), 7.33 – 7.21 (m, 2H), 2.01 – 1.68 (m, 2H), 7.10 – 6.94 (m, 1H), 3.54 (d, *J* = 36.0 Hz, 1H), 3.31 – 3.05 (m, 2H), 2.23 (s, 3H), 2.08 – 1.67 (m, 2H), 1.51 (d, *J* = 11.5 Hz, 1H), 1.23 (s, 1H); m/z [M+H]+ : 506.1

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(((1*R*,4*R*)-4-hydroxycyclohexyl)carbamoyl)phenyl)amino)thiazole-5-carboxamide (2d).



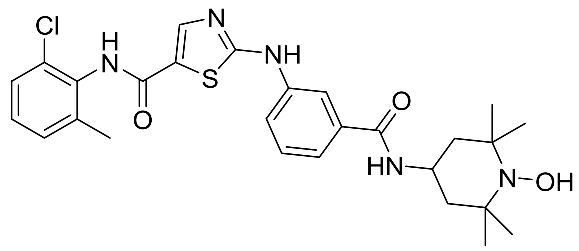
Compound **Va** (96 mg, 0.25 mmol) and (1*R*,4*R*)-4-aminocyclohexan-1-ol (44 mg, 0.38 mmol) were dissolved in dimethylformamide (1.5ml) and cooled to 0 °C. *N,N*-diisopropylethylamine (67.8 µl, 0.39 mmol) and triethylamine (53.7 µl, 0.39 mmol) were added to the reaction mixture. PyBroP (180 mg, 0.39 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0 °C and also added to the mixture. The reaction was quenched after 60 min by addition of saturated aq. NaCl (15 ml). The reaction mixture was diluted with 10% methanol in dichloromethane (15 ml) and the white precipitate containing the product was filtered. Additionally, crude product was extracted with 10% methanol in dichloromethane (3 x 25 ml) from the aqueous phase, dried over $MgSO_4$ and concentrated in vacuo. 25% of combined crude and precipitate were purified using flash chromatography (0-15% methanol in dichloromethane). Inhibitor **2d** was obtained as a white powder (22.3 mg, 74%): ^{1}H NMR (400 MHz, DMSO-d6) δ 10.76 (s, 1H), 9.95 (s, 1H), 8.23 – 8.12 (m, 2H), 8.05 – 8.00 (m, 1H), 7.87 – 7.81 (m, 1H), 7.49 – 7.37 (m, 3H), 7.33 – 7.22 (m, 2H), 4.56 (d, *J* = 4.4 Hz, 1H), 3.79 – 3.65 (m, 1H), 3.44 – 3.36 (m, 1H), 2.24 (s, 3H), 1.84 (t, *J* = 14.1 Hz, 4H), 1.44 – 1.31 (m, 2H), 1.27 (d, *J* = 10.3 Hz, 2H); ^{13}C NMR (126 MHz, DMSO-d6) δ 169.64, 162.41, 143.76, 140.33, 137.26, 134.24, 130.32, 130.15, 129.61, 128.34, 123.73, 122.53, 122.31, 118.46, 70.52, 61.54, 49.93, 47.87, 35.03, 31.45, 20.87, 18.70, 14.47, 9.24; m/z [M+H]+ : 485.2

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(((1*R*,4*R*)-4-hydroxy-4-methylcyclohexyl)carbamoyl)phenyl)amino)thiazole-5-carboxamide (2e).



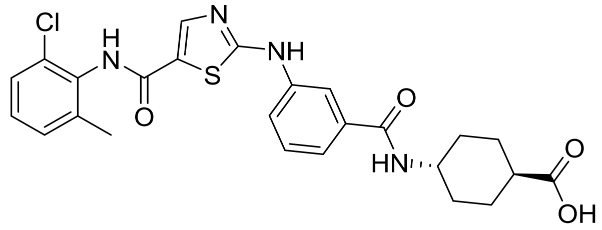
Compound **Va** (100 mg, 0.26 mmol) and (1*R*,4*R*)-4-amino-1-methylcyclohexan-1-ol (50 mg, 0.39 mmol) were dissolved in dimethylformamide (1.5ml) and cooled to 0 °C. *N,N*-diisopropylethylamine (67.8 µl, 0.39 mmol) and triethylamine (53.7 µl, 0.39 mmol) were added to the reaction mixture. PyBroP (180 mg, 0.39 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0 °C and also added to the mixture. The reaction was quenched after 80 min by addition of saturated aq. NaCl (10 ml). Reaction mixture was diluted with 10% methanol in dichloromethane (10 ml). Product was extracted with 10% methanol in dichloromethane (3 x 20 ml) from the aqueous phase. Combined organic phases were washed with ddH_2O (5 x 20ml), dried over $MgSO_4$ and concentrated in vacuo. Inhibitor **2e** was crystallized by addition of ice cold diethyl ether (70 mg, 54%): ^{1}H NMR (400 MHz, DMSO-d6) δ 10.75 (s, 1H), 9.94 (s, 1H), 8.15 (d, *J* = 6.5 Hz, 2H), 8.01 (t, *J* = 1.8 Hz, 1H), 7.89 – 7.77 (m, 1H), 7.52 – 7.34 (m, 3H), 7.34 – 7.16 (m, 2H), 3.83 – 3.77 (m, 2H), 2.23 (s, 3H), 1.75 – 1.71 (m, 4H), 1.46 (td, *J* = 11.3, 10.2, 3.3 Hz, 4H), 1.15 (s, 3H); ^{13}C NMR (126 MHz, DMSO-d6) δ 169.64, 162.41, 143.76, 140.33, 137.26, 134.24, 130.32, 130.15, 129.61, 128.34, 123.73, 122.53, 122.31, 118.46, 70.52, 61.54, 49.93, 47.87, 35.03, 31.45, 20.87, 18.70, 14.47, 9.24; m/z [M+H]+ : 499.2

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-((1-hydroxy-2,2,6,6-tetramethylpiperidin-4-yl)carbamoyl)phenyl)amino)thiazole-5-carboxamide (2f).



Compound **Va** (100 mg, 0.26 mmol) 4-amino-2,2,6,6-tetramethylpiperidin-1-ol (67 mg, 0.39 mmol) were dissolved in dimethylformamide (1.5 ml) and cooled to 0 °C. *N,N*-diisopropylethylamine (67.8 μl, 0.39 mmol) and triethylamine (53.7 μl, 0.39 mmol) were added to the reaction mixture. PyBroP (180 mg, 0.39 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0 °C and also added to the mixture. The reaction was quenched after 80 min by addition of saturated aq. NaCl (15 ml). The reaction mixture was diluted with 10% methanol in dichloromethane (15 ml) and crude product was extracted with 10% methanol in dichloromethane (3 x 25 ml) from the aqueous layer. Organic layer was washed with ddH$_2$O (5 x 20 ml), dried over MgSO$_4$, concentrated in vacuo and purified via flash chromatography (0-100% ethyl acetate in petroleum ether). The obtained product was recrystallized from ethyl acetate with ice cold ether and yielded inhibitor **2f** (41.2 mg, 29%) as a pale orange powder: m/z [M+H]+ : 541.2

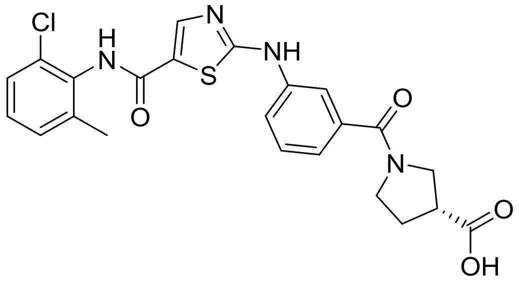
Preparation of (1*R*,4*R*)-4-(3-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)benzamido)cyclohexane-1-carboxylic acid (2g).



To a solution of compound **Va** (75 mg, 0.19 mmol) in dimethylformamide (2 ml) was added HATU (110 mg, 0.29 mmol), *N,N*-diisopropylethylamine (50 μl), and methyl (1*R*,4*R*)-4-aminocyclohexane-1-carboxylate (40 mg, 0.25 mmol). The reaction was stirred for 2 hours at room temperature, after which another 80 mg (0.51 mmol) of methyl (1*R*,4*R*)-4-aminocyclohexane-1-carboxylate and 100 μl of *N,N*-diisopropylethylamine was added. The reaction was stirred for another 2 hours at room temperature, after which 110 mg of HATU and 50 μl of *N,N*-diisopropylethylamine was added. The reaction was allowed to stir overnight at room temperature, after which it was quenched by the addition of water. The aqueous mixture was extracted twice with dichloromethane. The organic layers were dried and adsorbed onto silica gel and purified (0-100% ethyl acetate in hexane) to obtain the methyl ester (60 mg) as a dry film.

To a solution of the intermediate ester (60 mg, 0.11 mmol) in tetrahydrofuran (1 ml) and methanol (0.5 ml) was added lithium hydroxide (0.5 ml, 0.50 mmol). The reaction was stirred at room temperature for 1 hour, after which it was taken up with water and acidified with 1N HCl. The aqueous layer was extracted twice with ethyl acetate. The organic layers were combined, dried, and concentrated to give inhibitor **2g** (29.0 mg, 50%) as a white powder: ^{1}H NMR (400 MHz, DMSO-d6) δ 10.76 (s, 1H), 9.94 (s, 1H), 8.25 (d, J = 7.9 Hz, 1H), 8.15 (s, 1H), 8.08 – 7.98 (m, 1H), 7.90 – 7.77 (m, 1H), 7.53 – 7.35 (m, 3H), 7.33 – 7.21 (m, 2H), 3.75 – 3.72 (m, 2H), 2.24 (s, 3H), 2.21 – 2.11 (m, 1H), 2.01 – 1.85 (m, 4H), 1.48 – 1.32 (m, 4H); ^{13}C NMR (101 MHz, DMSO-d6) δ 166.58, 165.60, 159.15, 156.15, 142.31, 140.32, 138.76, 136.03, 129.02, 128.82, 127.00, 120.65, 119.98, 117.01, 67.97, 64.89, 47.53, 45.85, 37.93, 28.74, 26.38, 25.93, 25.85, 18.27; m/z [M+H]+ : 513.1

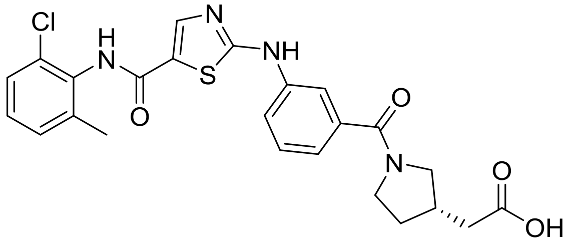
Preparation of (R)-1-(3-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)benzoyl)pyrrolidine-3-carboxylic acid (2h).



A solution of compound **Va** (75 mg, 0.19 mmol) in dimethylformamide (2 ml) was cooled to 0 °C on an ice bath. To this solution was added (R)-ethyl pyrrolidine-3-carboxylate (31 mg, 0.21 mmol), *N,N*-diisopropylethylamine (0.051 ml, 0.29 mmol), triethylamine (0.040 ml, 0.29 mmol), and a solution of PyBrop (135 mg, 0.29 mmol) previously cooled to 0 °C. This solution was allowed to stir at 4 °C for 2.5 hours, after which it was quenched by addition of water. The aqueous layer was extracted twice with dichloromethane. The organic layers were combined, dried with Na_2SO_4, and adsorbed to silica gel. The crude mixture was purified via column chromatography (0-100% ethyl acetate in hexane) to yield the ethyl ester (45.0 mg) as a dry film.

To a solution of the intermediate ester (44 mg, 0.09 mmol) in tetrahydrofuran (1 ml) and ethanol (0.5 ml) was added lithium hydroxide (0.5 ml, 0.50 mmol). This reaction was stirred at room temperature for 1 hour, after which it was taken up with water, acidified using 1 M HCl, and extracted twice into ethyl acetate. The organic layers were combined and concentrated to yield inhibitor **2h** (23 mg, 55%) as a pale yellow powder: ^{1}H NMR (400 MHz, DMSO-d6) δ 10.78 (s, 1H), 9.95 (s, 1H), 8.15 (s, 1H), 7.88 (s, 1H), 7.68 (d, *J* = 8.1 Hz, 1H), 7.42 (td, *J* = 7.9, 7.2, 3.1 Hz, 2H), 7.33 – 7.23 (m, 2H), 7.15 (t, *J* = 6.4 Hz, 1H), 3.80 – 3.62 (m, 1H), 3.59 (d, *J* = 7.0 Hz, 1H), 3.50 (s, 3H), 3.12 (dt, *J* = 21.9, 7.1 Hz, 1H), 2.24 (s, 3H), 2.10 (ddq, *J* = 28.0, 21.5, 6.8, 6.3 Hz, 2H)

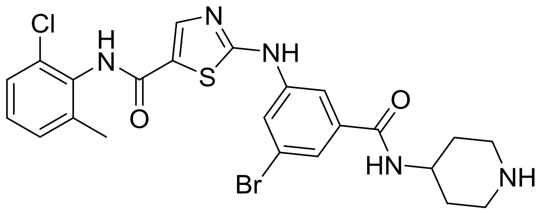
Preparation of (*S*)-2-(1-(3-((5-((2-chloro-6-methylphenyl)carbamoyl)thiazol-2-yl)amino)benzoyl)pyrrolidin-3-yl)acetic acid (2i).



To a solution of (*S*)-methyl 2-(pyrrolidin-3-yl)acetate (117 mg, 0.82 mmol) in dimethylformamide (1.5 ml) was added triethylamine (0.043 ml, 0.31 mmol) *N,N*-diisopropylethylamine (0.054 ml, 0.31 mmol) and compound **Va** (80 mg, 0.21 mmol). This solution was cooled to 0 °C, after which a solution of PyBroP (144 mg, 0.31 mmol) in dimethylformamide (1 ml), also cooled to 0 °C, was added slowly. The reaction was stirred at 0 °C for 30 minutes, quenched with water and extracted with dichloromethane. The organic layers were combined, dried, and concentrated to give the ester intermediate as a dry film.

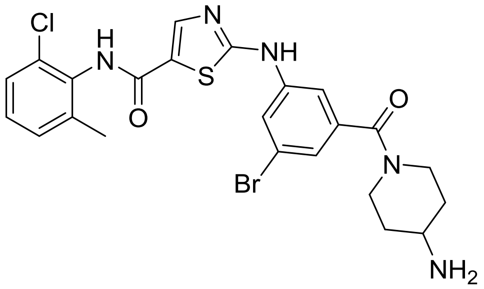
To a solution of the ester (112 mg, 0.22 mmol) in tetrahydrofuran (2 ml) and methanol (1 ml) was added lithium hydroxide (1.1 ml, 1.1 mmol). The reaction was stirred at room temperature for 1 hour, after which it was taken up with water, acidified with 1N HCl, and extracted twice into ethyl acetate. The combined organic layers were dried and concentrated to yield inhibitor **2i** (78 mg, 72%) as a white powder: ^{1}H NMR (400 MHz, DMSO-d6) δ 12.18 (s, 1H), 10.77 (s, 1H), 9.95 (s, 1H), 8.15 (s, 1H), 7.90 – 7.79 (m, 1H), 7.73 – 7.59 (m, 1H), 7.41 (dtd, *J* = 6.6, 3.9, 1.6 Hz, 2H), 7.32 – 7.24 (m, 2H), 7.15 (dd, *J* = 11.5, 7.6 Hz, 1H), 3.72 (dd, *J* = 12.1, 7.3 Hz, 1H), 3.63 – 3.53 (m, 1H), 3.52 – 3.42 (m, 2H), 3.13 (dd, *J* = 11.3, 8.1 Hz, 1H), 2.42 (t, *J* = 6.8 Hz, 1H), 2.33 (dd, *J* = 7.0, 3.0 Hz, 1H), 2.24 (s, 3H), 2.15 – 1.98 (m, 1H), 1.58 (ddd, *J* = 14.9, 12.1, 8.5 Hz, 1H)

Preparation of 2-((3-bromo-5-(piperidin-4-ylcarbamoyl)phenyl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (3a).



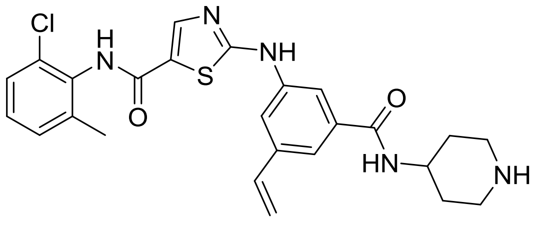
Compound **Vb** (200 mg, 0.43 mmol) and 4-aminopiperidine (90.2 µl, 0.86 mmol) were dissolved in dimethylformamide (1.5 ml) and cooled to 0 °C. After adding *N,N*-diisopropylethylamine (149.7 µl, 0.86 mmol) and triethylamine (119.8 µl, 0.86 mmol), PyBroP (300.3 mg, 0.64 mmol) was dissolved in dimethylformamide (0.5 ml), cooled to 0 °C and also added to the mixture. The reaction was quenched after 30 min by adding saturated aq. NaCl (8 ml). The extraction was done with ethyl acetate (3 x 10 ml) and 15% (1.33 M NH$_3$ in methanol) in dichloromethane (2 x 10 ml). The organic layer was dried over MgSO$_4$ and concentrated in vacuo. The crude was purified by flash chromatography (dichloromethane to 15% (0.2 N NH$_3$ in methanol) in dichloromethane) yielding inhibitor **3a** (141 mg, 60%) as a clear yellow solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.09 (s, 1H), 8.00 (t, J = 1.7 Hz, 1H), 7.74 (t, J = 1.4 Hz, 1H), 7.36 (dd, J = 7.2, 2.0 Hz, 1H), 7.25 (m, 2H), 7.21 (t, J = 1.5 Hz, 1H), 4.58 (d, J = 7.0 Hz, 1H), 3.77 (d, J = 10.1 Hz, 1H), 3.19 (t, J = 10.0 Hz, 1H), 3.01 (m, 1H), 2.96 (t, J = 10.0 Hz, 1H), 2.32 (s, 3H), 2.00 (d, J = 8.0 Hz, 1H), 1.88 (d, J = 8.0 Hz, 1H), 1.42 (t, J = 9.8 Hz, 2H); m/z [M+H]+ : 548.0

Preparation of 2-((3-(4-aminopiperidine-1-carbonyl)-5-bromophenyl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (3c).



Compound **Vb** (19 mg, 0.04 mmol), *N,N*-diisopropylethylamine (14.1 µl, 0.08 mmol) and 4-aminopiperidine (10 µl, 0.1 mmol) were dissolved in dimethylformamide (0.5 ml) and cooled down to 0 °C. The reaction was started by adding HATU (23 mg, 0.06 mmol) and after 30 min quenched with NH$_4$Cl (10 ml). Ethyl acetate (3 x 10 ml) was used for extraction, the resulting organic layer was dried over MgSO$_4$ and concentrated in vacuo. The obtained crude was purified by flash chromatography (dichloromethane to 15% (0.2 N NH$_3$ in methanol) in dichloromethane) yielding inhibitor **3c** (4.8 mg, 22%) and inhibitor **3a** as a side product: ^{1}H NMR (400 MHz, Methanol-d_4) δ 8.15 (t, *J* = 1.9 Hz, 1H), 8.10 (s, 1H), 8.00 (t, *J* = 1.8 Hz, 1H), 7.63 (t, *J* = 1.6 Hz, 1H), 7.36 (dd, *J* = 7.2, 2.3 Hz, 1H), 7.28 – 7.20 (m, 2H), 4.15 (ddd, *J* = 11.0, 6.9, 4.0 Hz, 1H), 3.48 (dt, *J* = 13.1, 3.7 Hz, 2H), 3.16 (td, *J* = 12.8, 3.1 Hz, 2H), 2.32 (s, 3H), 2.25 – 2.17 (m, 2H), 1.93 – 1.81 (m, 2H); m/z [M+H]+ : 458.0

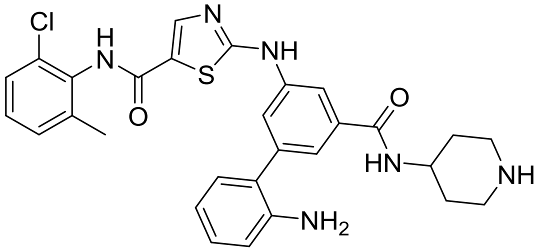
Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(piperidin-4-ylcarbamoyl)-5-vinylphenyl)amino)thiazole-5-carboxamide (3d).



Compound **Vb** (100 mg, 0.22 mmol) and tributyl(vinyl)tin (94 µl, 0.32 mmol) were dissolved in dioxane (2.5 ml). Toluene (1.25 ml) was added and the mixture was degassed for 30 min by introducing argon. After adding Pd(PPh$_3$)$_4$ (25 mg, 0.02 mmol), the reaction was conducted in an oilbath for 4 h at 110 °C under argon atmosphere and constant stirring. Distilled water (5 ml) and saturated aq. NaCl (5 ml) were used for quenching, ethyl acetate (3 x 10 ml) for extraction. The organic layer was dried over $MgSO_4$ before evaporating the solvent. For purification, the crude was dissolved in ethyl acetate (4 ml), cooled down and crystallized by adding chilled petroleum ether. Filtration yielded the crude carboxylic acid (53 mg) as a pale yellow solid.

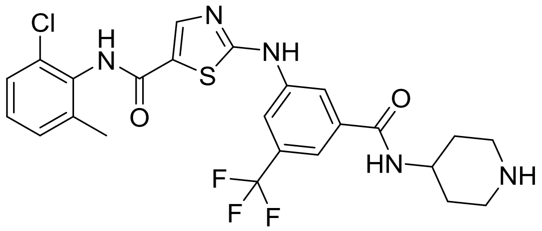
The crude carboxylic acid (16 mg) and 4-aminopiperidine (8.2 µl, 0.08 mmol) were dissolved in dimethylformamide (0.2 ml) and cooled to 0 °C before adding *N,N*-diisopropylethylamine (13.6 µl, 0.08 mmol) and triethylamine (10.9 µl, 0.08 mmol). A solution of PyBroP (27 mg, 0.06 mmol) in dimethylformamide (0.2 ml) was also cooled down to 0 °C and added. The reaction was quenched after 30 min by the addition of distilled water (2 ml) and saturated aq. NaCl (8 ml) and extracted with 15% (1.33 M NH_3 in methanol) in dichloromethane (3 x 10 ml). The organic layer was dried over $MgSO_4$ and concentrated in vacuo. Flash chromatography (dichloromethane to 15% (0.2 N NH_3 in methanol) in dichloromethane) yielded inhibitor **3d** (5.3 mg, 27%) as a clear, slightly yellow solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.07 (s, 1H), 7.73 (t, J = 1.7 Hz, 1H), 7.70 (t, J = 1.7 Hz, 1H), 7.36 (dd, J = 7.2, 2.0 Hz, 1H), 7.25 (m, 2H), 7.16 (t, J = 1.2 Hz, 1H), 6.78 (dd, J = 18.0, 10.9 Hz, 1H), 5.87 (d, J = 17.7 Hz, 1H), 5.36 (d, J = 11.1 Hz, 1H), 4.63 (d, J = 9.8 Hz, 1H), 3.85 (d, J = 9.0 Hz, 1H), 3.21 (t, J = 13.7 Hz, 1H), 3.12 (tt, J = 11.0, 4.1 Hz, 1H), 2.96 (t, J = 11.7 Hz, 1H), 2.32 (s, 3H), 2.03 (m, 1H), 1.91 (m, 1H), 1.47 (t, J = 8.8 Hz, 2H); m/z [M+H]+ : 496.2

Preparation of 2-((2'-amino-5-(piperidin-4-ylcarbamoyl)-[1,1'-biphenyl]-3-yl)amino)-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (3e).



Inhibitor **3b** (25 mg, 0.046 mmol), 2-aminobenzeneboronic acid (6 mg, 0.046 mmol) and $Pd(PPh_3)_4$ (5 mg, 0.046 mmol) were dissolved in dimethylformamide (0.5 ml) and degassed for 30 min by introducing argon. After adding K_2CO_3 (12 mg, 0.09 mmol), the mixture was irradiated with microwave for 5 h at 115 °C under argon. Afterwards it was diluted with methanol (10 ml), filtered through celite and concentrated in vacuo. Flash purification (dichloromethane to 15 % (1.33 M NH_3 in methanol) in dichloromethane) yielded inhibitor **3e** (8.2 mg, 32%) as a clear yellow solid: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.07 (s, 1H), 7.85 (s, 1H), 7.67 (s, 1H), 7.35 (d, J = 7.1 Hz, 1H), 7.25 (m, 2H), 7.16 (s, 1H), 7.12 (m, 2H), 6.86 (d, J = 8.0 Hz, 1H), 6.79 (t, J = 7.5 Hz, 1H), 4.74 (m, 1H), 4.04 (m, 1H), 3.45 (tt, J = 11.5, 3.7 Hz, 1H), 3.29 (m, 1H), 2.99 (m, 1H), 2.31 (s, 3H), 2.13 (m, 1H), 2.05 (m, 1H), 1.63 (m, 2H); ^{13}C NMR (101 MHz, Methanol-*d*$_4$) δ 172.20, 169.23, 164.87, 162.42, 145.39, 143.98, 143.09, 142.54, 140.33, 137.80, 134.22, 131.16, 130.17, 130.07, 129.66, 128.36, 127.63, 123.84, 122.50, 121.49, 119.62, 117.46, 115.94, 47.08, 41.51, 36.95, 31.65, 30.75, 18.68; m/z [M+H]+ : 561.2

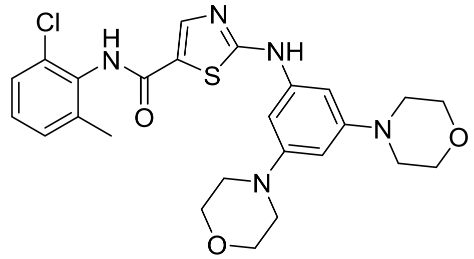
Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(piperidin-4-ylcarbamoyl)-5-(trifluoromethyl)phenyl)amino)thiazole-5-carboxamide (3b).



Compound **II** (100 mg, 0.35 mmol) was dried with toluene and dissolved in tert-butanol (2 ml). 3-Amino-5-(trifluoromethyl)benzoic acid (107 mg, 0.52 mmol) and (+)-camphor-10-sulfonic acid (81 mg, 0.35 mmol, dried with toluene) were added before putting the mixture under argon and irradiating it for 3.5 h at 120 °C and constant stirring in a microwave. The crude was diluted with ethyl acetate (8 ml), washed with sat. aq. NH_4Cl (10 ml), washed with saturated aq. NaCl (10 ml) and extracted with ethyl acetate (3 x 10 ml). Afterwards the organic layer was dried over $MgSO_4$ and the solvent was evaporated. For purification, it was dissolved in ethyl acetate (2 ml), cooled down and crystallized by adding chilled petroleum ether. Filtration yielded the crude carboxylic acid (45 mg) as a white solid.

A solution of the crude carboxylic acid (45 mg) and 4-aminopiperidine (20.7 µl, 0.20 mmol) in dimethylformamide (0.3 ml) was cooled to 0 °C before *N,N*-diisopropylethylamine (34.4 µl, 0.20 mmol) and triethylamine (27.5 µl, 0.20 mmol) were added. PyBroP (69.0 mg, 0.15 mmol) was also dissolved in dimethylformamide (0.3 ml), cooled to 0 °C and added to the mixture. Quenching of the reaction was achieved with distilled water (3 ml) and saturated aq. NaCl (7 ml). After 30 min, extraction was done with 15% (1.33 M NH_3 in methanol) in dichloromethane (3 x 10 ml). After washing the organic layer with saturated aq. NaCl (20 ml), it was dried over $MgSO_4$ and concentrated in vacuo. Flash purification (dichloromethane to 15% (0.2 N NH_3 in methanol) in dichloromethane) yielded inhibitor **3b** (35.5 mg, 67%) as a pale yellow solid: ^1H NMR (400 MHz, Methanol-d4) δ 8.11 (s, 2H), 8.06 (s, 1H), 7.33 (m, 2H), 7.25 (m, 2H), 4.63 (m, 1H), 3.77 (m, 1H), 3.23 (m, 1H), 3.11 (tt, J = 11.0, 3.9 Hz, 1H), 2.99 (m, 1H), 2.32 (s, 3H), 2.03 (m, 1H), 1.91 (m, 1H), 1.46 (m, 2H); ^{13}C NMR (101 MHz, Methanol-d_4) δ 170.54, 168.34, 164.87, 162.29, 143.76, 142.92, 140.32, 138.95, 134.21, 133.15, 130.18, 129.67, 128.37, 124.97, 120.08, 117.77, 116.47, 47.63, 42.12, 36.94, 34.65, 33.85, 31.64, 18.66; m/z [M+H]+ : 538.2

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3,5-dimorpholinophenyl)amino)thiazole-5-carboxamide (4a).

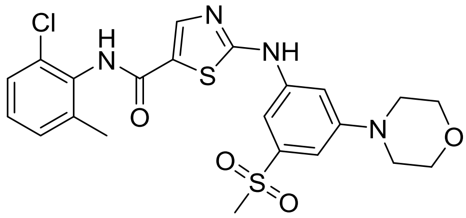


1,3-Difluoro-5-nitrobenzene (5.0 g, 31.43 mmol) and morpholine (10 ml) were irradiated with microwave at 160 °C for 2 hours. The reaction mixture was diluted with ethyl acetate (100 ml), and the organic phase was washed with saturated aq. NaCl (3 x 50 ml), dried over Na_2SO_4 and evaporated. The crude product was purified by flash silica chromatography (0-30% ethyl acetate in petroleum ether) to afford intermediate 4-(5-nitro-1,3-phenylene)dimorpholine (4.0 g) as a yellow solid.

Palladium on activated charcoal (73 mg, 0.68 mmol) and 4,4'-(5-nitro-1,3-phenylene)dimorpholine (2.0 g, 6.82 mmol) were stirred in ethanol (50 ml) under an atmosphere of hydrogen for 12 hours. The reaction mixture was filtered through celite and solvent was removed under reduced pressure. The crude 3,5-dimorpholinoaniline (1.65 g) was used in the next step without further purification.

Compound **II** (545 mg, 1.90 mmol) and hydrogen chloride (11 µl) were added to a solution of 3,5-dimorpholinoaniline (500 mg, 1.90 mmol) in 2-propanol (10 ml). The mixture was stirred overnight at 95 °C, then diluted with ethyl acetate (50 ml), and washed with saturated aq. NaCl (3 x 20 ml). The organic layer was dried over Na_2SO_4, filtered and evaporated. The crude product was purified by C18-flash chromatography (0-30% acetonitrile in water) to afford inhibitor **4a** (256 mg, 26%) as a white solid: ^1H NMR (400 MHz, DMSO-d6) δ 9.83 (s, 1H), 8.09 (s, 1H), 7.39 (dd, *J* = 7.5, 2.0 Hz, 1H), 7.32 – 7.19 (m, 2H), 6.65 (s, 2H), 6.21 (s, 1H), 3.84 – 3.62 (m, 8H), 3.15 – 2.99 (m, 8H), 2.23 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d6) δ 167.35, 161.77, 152.56, 141.55, 138.78, 133.47, 129.01, 128.16, 126.99, 122.24, 118.66, 97.79, 97.15, 66.12, 48.71, 18.27; m/z [M+H]+ : 514.2

Preparation of *N*-(2-chloro-6-methylphenyl)-2-((3-(methylsulfonyl)-5-morpholinophenyl)amino)thiazole-5-carboxamide (4b).



Copper(I) iodide (7.13 g, 37.45 mmol) was added to 1-fluoro-3-iodo-5-nitrobenzene (5.0 g, 18.73 mmol) and methanesulfonic acid sodium salt (3.32 g, 28.09 mmol) in dimethylformamide (50 ml) under nitrogen. The resulting suspension was stirred at 110 °C for 12 hours. The reaction mixture was diluted with ethyl acetate (200 ml), and washed with saturated aq. NaCl (3 x 50 ml). The organic layer was dried over Na_2SO_4, filtered and evaporated. The crude product was purified by flash silica chromatography (0-20% ethyl acetate in petroleum ether) to afford 1-fluoro-3-(methylsulfonyl)-5-nitrobenzene (3.77 g) as a yellow solid.

1-Fluoro-3-(methylsulfonyl)-5-nitrobenzene (3.67 g, 16.74 mmol) was added to morpholine (30 ml) in dimethylsulfoxide (5 ml) under nitrogen. The resulting suspension was stirred at 110 °C for 12 hours. The reaction mixture was diluted with ethyl acetate (200 ml), and washed with saturated aq. NaCl (3 x 50 ml). The organic layer was dried over Na_2SO_4, filtered and evaporated. The crude product was purified by flash silica chromatography (0-30% ethyl acetate in petroleum ether) to afford 4-(3-(methylsulfonyl)-5-nitrophenyl)morpholine (4.17 g) as a yellow solid.

Palladium on activated charcoal (74 mg, 0.70 mmol) and 4-(3-(methylsulfonyl)-5-nitrophenyl)morpholine (2.0 g, 6.99 mmol) in ethanol (50 ml) were stirred under an atmosphere of hydrogen for 12 hours. The reaction mixture was filtered through celite. The solvent was removed under reduced pressure. The crude product was purified by flash silica chromatography (0-10% methanol in dichloromethane) to afford 3-(methylsulfonyl)-5-morpholinoaniline (860 mg) as a white solid.

Hydrogen chloride (11 μl) was added to compound **II** (1.68 g, 5.85 mmol) and 3-(methylsulfonyl)-5-morpholinoaniline (1.0 g, 3.90 mmol) in 2-propanol (30 ml) warmed to 95 °C. The resulting solution was stirred at 95 °C for 12 hours, then diluted with ethyl acetate (200 ml), and washed with saturated aq. NaCl (3 x 100 ml). The organic layer was dried over Na_2SO_4, evaporated and purified by flash C18-flash chromatography (0-30% acetonitrile in water) to afford inhibitor **4b** (67 mg, 34%) as a white solid: ^{1}H NMR (400 MHz, DMSO-d6) δ 9.95 (s, 1H), 8.16 (s, 1H), 7.69 – 7.65 (m, 1H), 7.51 – 7.46 (m, 1H), 7.41 (dd, *J* = 7.5, 2.0 Hz, 1H), 7.33 – 7.23 (m, 2H), 7.09 (t, *J* = 1.8 Hz, 1H), 3.78 (dd, *J* = 6.2, 3.6 Hz, 4H), 3.21 (d, *J* = 3.7 Hz, 7H), 2.24 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d6) δ 166.29, 161.60, 159.05, 152.14, 142.41, 142.22, 141.82, 138.77, 133.31, 132.38, 129.04, 128.27, 127.02, 107.36, 106.59, 105.85, 65.88, 47.81, 43.47, 18.25; m/z [M+H]+ : 507.1

2 Biochemistry, mass spectrometry and data processing

Compound immobilization. Probes 1, 5, 13, 19 were immobilized on NHS-activated sepharose (GE Healthcare, Freiburg, Germany) beads through covalent linkage via a primary amine as described previously.[273] Coupling was performed in DMSO at final coupling densities of 2 µmol/ml beads for compounds 5, 13 and 19 or 1 µmol/ml beads for compound 1. Briefly, equilibrated sepharose beads (15 ml settled beads) were mixed with the respective compound, triethylamine (225 µl) was added and the mixture was incubated on an end-over-end shaker for 20 h in the dark. Free NHS-groups on the beads were blocked by adding 750 µl amino ethanol and incubation on an end-over-end shaker for 20 h in the dark. Cbz protection group of immobilized compound 13 was removed after equilibrating the beads in dimethylformamide by 6 alternating washings using *N*-methyl-2-pyrrolidon (10 ml), 1,8-Diazabicyclo(5.4.0)undec-7-ene (7.5 ml), and β-mercaptoethanol (7.5 ml). Probe 7 was immobilized on reversed sepharose beads through covalent linkage via a carboxylic functional group as described previously.[273] Equilibrated beads (15 ml settled beads) were reversed by addition of a mixture containing triethylamine (225 µl), ethylendiamine (40.2 µl) and amino ethanol (144.9 µl) in dimethylformamide and incubation on an end-over-end shaker for 20 h in the dark. Reversed beads were washed and equilibrated with DMSO prior to addition of compound (coupling density 2 µmol/ml beads), *N,N*-diisopropylethylamine (1.5 ml of a 200 mM dilution in dimethylformamide) and coupling reagent Bromotripyrrolidinophosphonium hexafluorophosphate (PyBroP, 1.5 ml of a 100 mM dilution in dimethylformamide). Coupling was performed on an end-over-end shaker for 20 h in the dark and blocking of remaining binding sites was accomplished by addition of NHS acetate blocking solution and incubation on an end-over-end shaker for 20 h in the dark. NHS acetate blocking solution was prepared by mixing equal amounts of 200 mM *N,N'*-Dicyclohexylcarbodiimide and 200 mM *N*-hydroxysuccinimide in acetonitrile, addition of 11.4 ul acetic acid per 1 ml of mix and incubation at RT overnight. Coupled beads were washed and stored in ethanol at 4°C in the dark. The coupling reaction was monitored by LC-MS.

Selectivity profiling using kinase affinity matrices. 96-well plate Kinobeads γ competition assays have been performed as described previously.[273,274] 5 mg mixed protein lysate (K-562, MV-4-11, SK-N-BE(2) and COLO 205) was pre-incubated with compound dilutions in DMSO (0, 3, 10, 30, 100, 300, 1000, 3000, 30000 nM final concentrations) for 45 min at 4 °C. Subsequently, each lysate was incubated with 35 µl settled and equilibrated Kinobeads for 30 min at 4 °C. DMSO control lysate was recovered for the pulldown of pulldown experiment. Beads were washed and bound proteins were eluted with 40 µl 2x LDS sample buffer (NuPAGE, Invitrogen) containing 50 mM DTT. Reduced disulfide bridges were alkylated using 4 µl chloroacetamide (final concentration 55 mM). Proteins were concentrated and desalted by a short SDS-PAGE using a 4-12% gel (NuPAGE, Invitrogen) and in-gel digested according to standard procedures.

LC-MS/MS analysis. Peptides generated by in-gel trypsin digestion were analyzed via LC-MS/MS on a nanoLC-Ultra 1D+ (Eksigent) coupled to a LTQ-Orbitrap Elite mass spectrometer (Thermo Scientific) as described previously.[273,274] Peptides were delivered to a trap column (100 µm x 2 cm, packed in house with Reprosil-Pur C18 AQ 5 µm resin, Dr. Maisch) for 10 min at a flow rate of 5 µl/min in 100% solvent A_0 (0.1% v/v FA in HPLC grade water). Peptides were then separated on an analytical column (75 µm x 40 cm, packed in-house with Reprosil-Gold C18, 3 µm resin, Dr. Maisch) using a 100 min gradient ranging from 4–32% solvent B (0.1% v/v FA and 5% v/v DMSO in acetonitrile) in solvent A (0.1% v/v FA and 5% v/v DMSO in HPLC grade water) at a flow rate of 300 nl/min. The mass spectrometer was operated in data dependent mode, automatically switching between MS and MS2 spectra. Up to 15 peptide precursors were subjected to fragmentation by high energy collision-induced dissociation (HCD) and analyzed in the Orbitrap. Dynamic exclusion was set to 20 s. A kinase peptide inclusion list (3 most intense peptides per kinase) was enabled.

Peptide and protein identification and quantification. Label free quantification was performed using MaxQuant (version 1.5.3.30)[251] by searching MS2 data against all canonical protein sequences as annotated in the Swissprot reference database (human proteins only, 20193 entries, downloaded 22.03.2016, internally annotated with PFAM domains) using the embedded search engine Andromeda[252] as described previously.[273,274] Carbamidomethylated cysteine was set as fixed modification; variable modifications included phosphorylation of serine, threonine or tyrosine, oxidation of methionine, and N-terminal protein acetylation. Trypsin/P was specified as proteolytic enzyme with up to two allowed missed cleavage sites. Precursor and fragment ion tolerances were 10 ppm and 20 ppm, respectively. Label-free quantification[250] and match-between-runs options were enabled and results were filtered for a minimal length of seven amino acids, 1% peptide and protein FDR as well as common contaminants and reverse identifications. For consistent peptide identification and protein grouping, the MS data for each compound was supplemented with 15 standard DMSO controls. Each compound was analyzed separately.

Label free quantification of Kinobead Drug Screen data of clinical kinase inhibitors was performed using MaxQuant (version 1.4.0.5)[251] by searching MS data against a human UniProt reference database (version 22.07.13, 88354 entries) using the search engine Andromeda [252] as described previously [273,274]. Carbamidomethylated cysteine was used as fixed modification; variable modifications included phosphorylation of serine, threonine or tyrosine, oxidation of methionine, and N-terminal protein acetylation. Trypsin/P was specified as proteolytic enzyme with up to two allowed miscleavage sites. Precursor tolerance was set to 10 ppm and fragment ion tolerance was set to 0.5 Da. Label-free quantification[250] and match-between-runs options were enabled and results were filtered for a minimal length of seven amino acids, 1% peptide and protein FDR as well as common contaminants and reverse identifications. Each compound was processed and analyzed separately.

Data processing. LFQ intensities were normalized to DMSO controls and EC_{50} values were deduced by a four-parameter log-logistic regression using an in-house pipeline based on the drc add-on package in R.[291] The depletion factor was calculated and applied to transform EC_{50} to K_D^{app} values as described previously.[273,274] Binding affinities are reported as pK_D^{app} values which is the negative logarithm (base 10) of the K_D^{app} value in mol/l.

Target selection criteria. Targets were manually annotated. A protein was considered a high-confidence target if the binding curve showed a sigmoidal shape with a dose-dependent decrease in binding to the Kinobeads. Proteins that only showed an effect at the highest inhibitor dose where not annotated as targets. The number of unique peptides and MSMS spectra were also included as target selection criteria. Peptide intensity in DMSO controls and MS/MS data quality was also taken into account. Proteins with low peptide counts, MS/MS spectral counts or MS1 intensity that nonetheless showed a reasonable dose response curve fit were considered as potential targets. In addition, if an inhibitor also interacted with similar kinases (*e.g.*, CDK family) it was also considered as a potential target. Low-confidence targets were excluded from further analysis. Note that for some targets, curve fitting with our data processing pipeline was not possible resulting in no or very high K_D^{app} values. Targets were considered as direct Kinobeads binders if annotated in Uniprot.org as a protein or lipid kinase. Furthermore, nucleotide binders, helicases, ATPases and GTPases, FAD (*e.g.*, NQO2) and heme (*e.g.*, FECH) containing proteins were also considered as potential direct binders. Most other target proteins are interaction partners/adaptor proteins of the kinases and are termed indirect Kinobeads binders.

Data analysis and visualization. For the calculation of the selectivity score S a threshold was set at 10 times K_D^{app} (EPHA2), the number of kinase targets was counted and divided by all identified protein kinases (n: 320) as a reference.[191] Gini coefficient was calculated using the K_D^{app} (EPHA2) of the respective compound as threshold concentration.[195] Figures and tables were produced in GraphPad Prism 5 (version 5.01) and Excel. Gene ontology (GO) term analysis was performed on the basis of GO_Biological Process terms reported by UniProt and complemented by literature research.[292] Kinometrees were generated by an in-house Java application. Kinometree illustration reproduced courtesy of Cell Signaling Technology, Inc (www.cellsignal.com). Sequence motifs were created using Weblogo (version 2.8.2).[293]

Immunoblot analysis. Antibodies against EPHA2 (C20, sc-924), EPHA2 pS897 (#6347), AKT1 (#4691), AKT1 pS473 (#4060), α-tubulin (E-19, sc-12462), β-Actin (C4, sc-47778) were purchased from Santa Cruz or Cell Signaling technology. Half of the obtained Kinobead eluate or 120 ug cell lysate was separated by 4-12% NuPAGE gel electrophoresis and transferred onto PVDF membranes (Novex, Life Technologies). Membranes were blocked for 1h in 2% BSA in 1x Tris buffered saline at room temperature and probed over night at 4 °C with the respective primary antibody. Antibody binding was detected using fluorophore-conjugated secondary antibodies (LI-COR) using an Odyssey scanner (LI-COR Biosciences).

NCI60 kinome profiling. NCI60 kinome data were obtained from Gholami *et al*[294] and processed in MaxQuant (version 1.4.0.5) applying the same settings as above and including iBAQ quantification. MaxQuant results were analyzed using the MaxQuant associated software suite Perseus (version1.5.6.0).[295] iBAQ intensities were log2 transformed and normalized by protein-wise Z-score transformation.

Kinase activity assay (performed by Reaction Biology, Malvern, USA). Dose dependent inhibition of EPHA2 kinase activity was validated for 15 selected inhibitors using a radioactive filter binding assay performed at the K_M(ATP) of EPHA2 (50 μM) as published previously[189] (outsourced to Reaction Biology).

Docking (performed by Dr. Xiaofeng Liu, East China University of Science and Technology, Shanghai, China). The binding poses of Dasatinib, CHEBI-513815 and PD-173955 in the ATP-binding site of EPHA2 were predicted by molecular docking using Glide (Schrödinger, Inc.). The crystal structure of EPHA2 in complex with ANP was retrieved from the Protein Data Bank (PDB: 1MQB) and prepared using the Protein Preparation Wizard in Maestro (Schrödinger, Inc.) to remove non amino acid molecules, add hydrogen atoms, and assign the protonation states for the polar residues. The scoring grid was generated by enclosing the residues 14 Å around Dasatinib in the binding site. The docking was performed in Glide SP mode and the top ranked pose by GScore of each inhibitor was retained for visual analysis of interactions. Docking studies were kindly performed by Dr. Xiaofeng Liu, East China University of Science and Technology, Shanghai, China.

Recombinant expression and purification of EPHA2 kinase domain (performed by Dr. Santosh Lakshmi Gande, Johann Wolfgang Goethe-Universität, Frankfurt, Germany). The detailed protocols for isotopic labeling procedures, expression and purification conditions of EPHA2 were published elsewhere.[296] In brief, the gene encoding the catalytic domain of human EPHA2 (residues D596-G900) was synthesized at GenScript USA Inc., USA and was optimized for its expression in insect cells. Synthesized EPHA2 gene was sub-cloned in pTriEx 1.1 with cleavable N-terminal Flag-His tag. Recombinant baculovirus incorporating the kinase cDNA construct was generated by homologous recombination with BacMagic DNA (Novagen), in Spodoptera frugiperda (Sf9) cells. High titer viral stocks were generated by infection of Sf9 cells, which were then used for the expression of EPHA2 by infection of Sf9 cells in large-scale. Recombinantly expressed EPHA2 was isolated from the cellular extracts by passage over NiNTA resin, followed by removal of Flag- polyhistidine tags with TEV protease (in-house) and inverse NiNTA. The protein was further purified by size-exclusion chromatography with 20 mM Tris, 0.2 M NaCl, 5 mM $MgCl_2$, 3 mM TCEP, the peak fractions were concentrated to 9.4 mg/ml and stored at –80 °C. Purified EPHA2 was not phosphorylated during the expression in Sf9 cells and isolation procedures, as determined by mass spectrometry. Expression and purification were kindly performed by Dr. Santosh Lakshmi Gande, Johann Wolfgang Goethe-University, Frankfurt, Germany.

Protein crystallization and structure determination of EPHA2 (performed by Verena Linhard and Dr. Denis Kudlinzki, Johann Wolfgang Goethe-Universität, Frankfurt, Germany). Purified EPHA2 protein was concentrated up to 6–10 mg/ml in a buffer containing 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM $MgCl_2$ and 3 mM TCEP. After the addition of different ligands up to a final concentration of 1 mM (5 mM for MLN-8054 and Danusertib; 10 mM for ATPγS), the protein sample was diluted 1:1 with crystallization buffer (respectively 3:1 for MLN-8054) in crystallization drops of 500–800 nl volume on a 96-well microplate. Crystals were grown as sitting drops at 291 K against 50 µl reservoir solution. Rod-shaped crystals (0.1–0.2 mm) or plate-like cuboids (0.05–0.1 mm) appeared after 1–2 weeks and grew to their final size within 4 weeks. Crystallization conditions were composed from three different building blocks and varied for each ligand between the following ranges: 30–40% precipitant solution (stock solution consisting of 25% PEG 1000, 25% PEG 3350, 25% MPD), 0.05–0.3 M amino acids solution (stock solution consisting of 0.2 M glutamate, 0.2 M glycine, 0.2 M serine, 0.2 M alanine, 0.2 M lysine) or 0.05–0.3 M carboxylic acids solution respectively (stock solution consisting of 0.2 M sodium formate, 0.2 M ammonium acetate, 0.2 M sodium citrate, 0.2 M sodium/potassium tartrate, 0.2 M sodium oxamate) and 0.1 M buffer pH 5.5-8.5 (Bis-Tris, Tris, Hepes, MES, Bicine or Tris/Bicine). Prior to flash-cooling some crystals had to be soaked with 20% ethylene glycol in mother liquor for cryoprotection. Diffraction data at resolutions between 1.10–2.10 Å had been collected at beamline BL14.1 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron storage ring (Berlin, Germany)[297] at beamline P13 operated by the EMBL Hamburg at the PETRA III synchrotron source (DESY, Hamburg Germany)[298], at X06DA (PXIII) beamline operated by the Swiss Light Source (PSI, Villigen, Switzerland) and at PROXIMA-1 beamline operated by the Synchrotron SOLEIL (Saint-Aubin, France). The data was processed using *XDSAPP*.[299] The structure was determined by molecular replacement with *Phaser*[300] using the crystal structure of the human EPHA2 kinase domain (PDB: 1MQB)[301] as a search model. Model building was performed using *Coot*[302] and the structure was refined and validated using *PHENIX*.[303] Drug-protein interaction analysis was performed using LigPlot+.[304] Figures containing molecular graphics were prepared using *PyMOL* (Schroedinger). Structure-solution and refinement statistics can be found in Tables S2e and S7b. Protein crystallography studies were kindly performed by Dr. Denis Kudlinzki and Verena Linhard, Johann Wolfgang Goethe-University, Frankfurt, Germany.

Crystallization and structure determination of MELK (performed by Giulia Canevari, Elena Casale, Stefanie Re Depaolini, and Eduard Rudolf Felder, Oncology, Nerviano Medical Sciences Srcl, Milan, Italy). Protein production was performed as previously described.[305] Crystallization experiments were established using the dephosphorylated MELK protein (residues 2-340) plus 1 mM DTT and 0.5 mM inhibitor dissolved in DMSO. Crystals were grown in hanging drops prepared by mixing equal volumes (1-2 µL) of protein solution at a concentration of 7 mg/mL and reservoir solution containing 10-20% PEG 3350 or PEG 4000, 600 mM NaCl and 100 mM Bis-Tris pH 6.5. After setting the drops, 1 µL 14.3 M β-mercaptoethanol was added to the 500 µL reservoir. Streak seeding was essential in order to obtain diffraction quality crystals. Two different crystal morphologies were observed. MELK complexed with Nintedanib, PF-3758309, K-252a and Defactinib yielded rectangular, prism-shaped crystals as described[305] whilst the MELK-BI-847325 complex produced cubic crystals. Crystals were transferred in a cryo-protectant solution (15% PEG 3350, 600 mM NaCl and 100 mM Bis-Tris pH 6.5 and 30% glycerol) prior to flash-cooling in liquid nitrogen. X-ray diffraction data were collected at the European

Synchrotron Radiation Facility (ESRF) in Grenoble (beam line ID23-1). Crystals of MELK complexed with Nintedanib, PF-3758309, K-252a and Defactinib were of space group P212121; whilst crystals of MELK complexed with BI-847325 belonged to space group I4132. Using an internal MELK structure as a search model, data were processed with iMOSFLM[306,307] or XDS[308] merged with SCALA[306,309] and phased via molecular replacement with Phaser[300,306]. Refinement of the structures was achieved using REFMAC[306,310] and COOT[306,311]. Full data collection and refinement statistics are reported in Table S6c. Protein crystallography studies were kindly performed by Giulia Canevari, Elena Casale, Stefanie Re Depaolini, and Eduard Rudolf Felder, Oncology, Nerviano Medical Sciences Srcl, Milan.

Data deposition. EPHA2 structural data is accessible via PDB protein data bank (https://www.rcsb.org/pdb): apo (PDB: 5I9U), ATPγS (PDB: 5I9V), ADPNP (PDB: 5I9W), Dasatinib (PDB: 5I9Y), Danusertib (PDB: 5I9Z), Cpd66 (PDB: 5IA2), Bosutinib (PDB: 5I9X), PD-173955 (PDB: 5IA3), Alisertib (PDB: 5IA0), MLN-8054 (PDB: 5IA1), Golvatinib (PDB: 5IA5), Foretinib (PDB: 5IA4), inhibitor 1g (pdb: 5NJZ), inhibitor 1j (pdb: 5NK0), inhibitor 1k (pdb: 5NK1), inhibitor 1l (pdb: 5NK3), inhibitor 1m (pdb: 5NK5), inhibitor 2a (pdb: 5NK7), inhibitor 2b (pdb: 5NK2), inhibitor 2c (pdb: 5NK4), inhibitor 2d (pdb: 5NK6), inhibitor 2e (pdb: 5NK9), inhibitor 2f (pdb: 5NK8), inhibitor 2g (pdb: 5NKA), inhibitor 2h (pdb: 5NKC), inhibitor 2i (pdb: 5NKD), inhibitor 3a (pdb: 5NKE), inhibitor 3b (pdb: 5NKF), inhibitor 3d (pdb: 5NKG), inhibitor 3e (pdb: 5NKH), inhibitor 4a (pdb: 5NKB), inhibitor 4b (pdb: 5NKI). MELK structural data is accessible via PDB protein data bank (https://www.rcsb.org/pdb): K-252a (PDB: 5M5A), Nintedanib (PDB: 5MAF), PF-3758309 (PDB: 5MAG), Defactinib (PDB: 5MAH), BI-847325 (PDB: 5MAI). The mass spectrometry data and all obtained dose response curves for each inhibitor selectivity profiling have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [312] with the dataset identifier PXD006193, PXD004112 and PXD005336.

Sequence alignment. Aligned kinase domain sequences were obtained from supplemental information published by Creixell et al [313]. Sequence alignment for kinases constituting the target space of the compounds studied in this report was restricted to target proteins with affinities higher than 1 μM. Protein groups in selectivity profiling were split and counted separately for sequence alignment of target proteins.

NMR spectroscopy (performed by Dr. Sridhar Sreeramulu, Johann Wolfgang Goethe-Universität, Frankfurt, Germany). NMR experiments were performed at measuring temperature of 298 K on a Bruker 950-MHz spectrometer equipped with room temperature TXI-HCN probes. NMR samples were prepared with 10% v/v D_2O to lock the spectrometers, and 3-(trimethylsilyl)-2,2,3,3-tetradeuteropropionic acid (TSP-d4; 1 mM) was used as an internal standard for spectral referencing. The processing and analysis of NMR spectra were done in Topspin version 3.1 (Bruker Biospin). Sample conditions: 0.1–0.15 mM protein in 20 mM Tris pH8, 200 mM NaCl, 5 mM $MgCl_2$, 3 mM TCEP. Ligands were added as 100 mM stock solutions in d6-DMSO to yield a final concentration of 0.5 mM. Protein crystallography studies were kindly performed by Dr. Sridhar Sreeramulu, Johann Wolfgang Goethe-University, Frankfurt, Germany.

3 Cell culture

Cell lines and reagents. K-562, MV-4-11 and COLO 205 cells were cultivated in RPMI medium 1640 (RPMI1640, Biochrom GmbH) and SK-N-BE(2) was grown in Dulbecco's modified Eagle medium/Ham's F12 (DMEM/F12, Biochrom GmbH), all supplemented with 10% (v/v) fetal bovine serum (Biochrom GmbH) and 1% antibiotic/antimycotic solution (Sigma Aldrich). SF-268 cells were grown in Iscove's Modified Dulbecco's Medium (IMDM, Biochrom GmbH) supplemented with 10% (v/v) fetal bovine serum. Cell lines were kindly provided by the NCI-Frederick Cancer DCTD Tumor/Cell line repository, were not found in the database of commonly misidentified cell lines (ICLAC) and were tested negatively for mycoplasm contamination. Cell lysis and Bradford assay were performed as described previously.[273] Kinase affinity matrices were prepared in house as published elsewhere.[273]

Cpd66 was synthesized in four steps according to the original synthetic route [314] using the advanced 8-bromo-3-methyl-3,7-dihydro-purine-2,6-dione building block (KaïronKem). PD-173955 [315] was obtained as a camphorsulfonate salt in one step, by reacting the advanced 6-(2,6-dichloro-phenyl)-2-methanesulfonyl-8-methyl-8H-pyrido[2,3-d]pyrimidin-7-one intermediate (DIV00341, Diverchim) with 3-(methylthio)aniline (3 eq) in the presence of camphorsulfonic acid (3 eq) in dry isopropanol for six hours at 120 °C under microwave irradiation, followed by filtration with methanol and recrystallization with dichloromethane/ethyl acetate. The other small molecule kinase inhibitors were commercially sourced from Selleckchem, MedChemExpress, Active Biochem, Abmole, Merck or LC Labs.

Protein knockdown by small interfering RNA (siRNA). SF268 cells were seeded in 96-well plates (1000 cells per well), incubated for 24 h and medium was exchanged before addition of siRNA (Qiagen: SI00063553, SI0030081, SI00300188, SI02223508). siRNA was diluted in Opti-MEM medium (Gibco, Life Technologies) to a concentration of 70 nM. Upon addition of INTERFERin^{TM} (PolyPlus, peqlab), the mixture was vortexed for 10 s and incubated at room temperature for 10 min before addition to the cells at a final concentration of 20 nM. The different siRNA preparations were tested individually and as mix of all four siRNA. Controls were performed using 1 nM AllStars scrambled (Qiagen, SI03650318) and CellDeath siRNA (Qiagen, SI04381048), INTERFERin^{TM} only and Opti-MEM medium only. Cell viability was determined according to the manufacturer after 6 days by fluorescence readout of alamarBlue (Pierce) in a FluoStar Omega plate reader (excitation: 544 nm, emission: 590 nm). All experiments were repeated in technical triplicates and biological triplicates. Relative cell viability was normalized to the INTERFERin control. Statistical analysis was performed using an unpaired two-tailed t-test (significance level $p<0.001$).

Kinase inhibitor treatment. SF268 cells were seeded in 96-well plates (1000 cells per well) and incubated for 24 h before treatment. Drug dilutions were prepared in DMSO (0, 3, 10, 30, 100, 300, 1000, 3000, 30000 nM final concentrations) and added to the cells to a final DMSO concentration of 0.3%. Cells were incubated with drug for 72 hours and cell viability was determined according to the manufacturer by fluorescence readout of alamarBlue (Pierce) in a FluoStar Omega plate reader (excitation: 544 nm, emission: 590 nm). Drug treatments were performed in biological replicates each containing three technical replicates.

Results and Discussion

Table of contents

1 Kinobeads screen enables determination of inhibitor selectivity 65
  1.1 Kinobeads screen of clinical kinase inhibitors 65
  1.2 Measuring drug selectivity 68
  1.3 CATDS is a thermodynamics-based metric suitable for Kinobeads data 68
  1.4 Calculation of CATDS for different applications 69
  1.5 Comparison of CATDS with other scoring systems 76
2 EPHA2 as target of clinical kinase inhibitors 81
  2.1 Chemical proteomics reveals EPHA2 as common off-target 81
  2.2 Target profiling of EPHA2 inhibitors enables rational drug repurposing 82
3 Structure-based residue classification for inhibitor design 85
  3.1 Global analysis of ligand bound EPHA2 structures 85
  3.2 Identification of drug-protein interactions in EPHA2 87
  3.3 Classification scheme for amino acids involved in drug binding 89
  3.4 Expanding the classification scheme 93
4 Design and synthesis of dedicated EPHA2 inhibitors 97
  4.1 New EPHA2 inhibitor design by hybridization of known EPH binders 97
  4.2 Synthesis of dedicated EPHA2 inhibitors 99
5 Biochemical and cellular evaluation of EPHA2 inhibitors 101
  5.1 EPHA2 inhibitors bind EPHA2 with high affinity 101
  5.2 Novel EPHA2 inhibitors display improved selectivity 104
  5.3 Protein crystallography reveals epitopes for inhibitor binding 106
  5.4 Cellular activity of EPHA2 inhibitors in SF-268 cells 108
  5.5 EPHA2 inhibitor 4a potently reduces SF-268 cell viability 110

“When you are in jail, a good friend will be trying to bail you out.
A best friend will be in the cell next to you saying: Damn, that was fun.”
- Groucho Marx -

„Sometimes the most ordinary things could be made extraordinary,
simply by doing it with the right people.”
- Nicholas Sparks -

1 Kinobeads screen enables determination of inhibitor selectivity

1.1 Kinobeads screen of clinical kinase inhibitors

Since the approval of Imatinib in 2001, many efforts have led to the discovery of a large amount of potent kinase inhibitors targeting a surprisingly small number of validated oncogenic protein kinases. Especially in industry, drug discovery is often performed by the screening of large compound libraries. Most of the tested inhibitor candidates fail during preclinical development before they are evaluated in clinical trials. Compounds that enter clinical trials are, thus, rather mature compounds that already succeeded preclinical evaluation and comprise favorable characteristics for the application in human. Several hundred kinase inhibitors are currently being tested in clinical trials – however, a large number of these drugs have not been profiled systematically yet or the information is not available in public resources.

The Kinobeads technology was used in a competitive pulldown setup for dose-resolved selectivity profiling of 235 clinical kinase inhibitors. A full and detailed description of this screen was presented in the PhD thesis of Dr. Susan Kläger (March 2017). Briefly, 235 clinical kinase inhibitors were profiled against a total of 246 protein and lipid kinases derived from a mix of four cancer cell lines (Figure 18). The resulting matrix depicts the “druggable kinome” and can be searched in both directions, in a target-centric or compound-centric angle. Target-centric evaluation aids in the identification of inhibitors binding a particular protein; compound-centric evaluation provides insights into the target spectrum of a particular inhibitor.

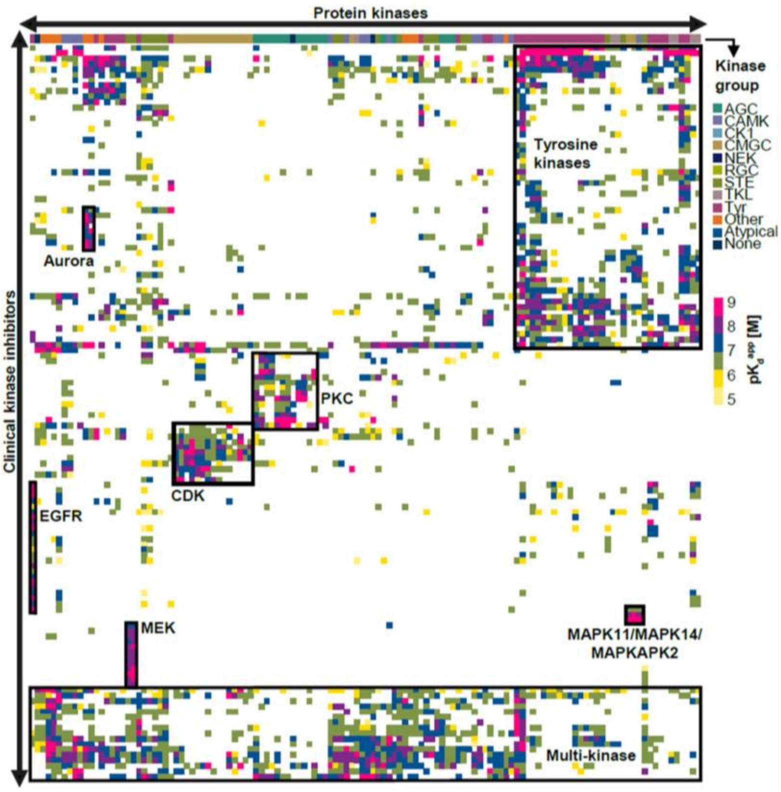


Figure 18 | View on the “druggable” kinome. Hierarchical clustering of kinase targets against clinical kinase drugs (Color code depicts the pK_D^{app} of drug-target interactions). The Kinobeads drug screen identified highly selective (e.g. EGFR or MEK inhibitors) and relatively unselective inhibitors (e.g., tyrosine kinase and multi-kinase inhibitors).

The experimental setup of the Kinobeads technology allows for the identification of additional non-protein kinase targets, including seven metabolic kinases, 23 other nucleotide binding proteins, five FAD containing proteins and one heme containing protein (Figure 19a). This is of special interest since these proteins are not part of conventional screening panels, but could potentially impact a drug's mode of action or explain adverse side effects. Examples for such non-kinase off-targets include glycogen phosphorylase (PYGL, PYGM, PYGB)[316], ribosyldihydronicotinamide dehydrogenase (NQO2)[137], or ferrochelatase (FECH)[274]. The evaluated inhibitors were originally designed for a total of only 114 kinase targets – leaving a large space of off-target inhibition and thus, considerably increasing our understanding of the "druggable kinome" (Figure 19b). Comparison to publically available data from other kinase inhibitor screens[143,189] or commonly used databases such as CHEMBL or LINCS showed that many identified drug:protein interactions have not been reported yet – rendering this screen highly valuable for the scientific community.

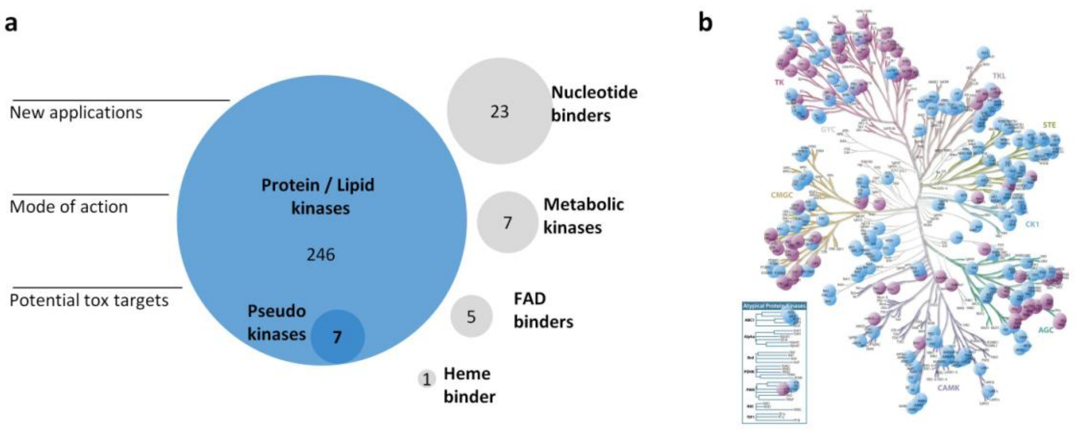


Figure 19 | Target space of clinical kinase inhibitors revealed by Kinobeads. a) Target space includes protein and lipid kinases, but also other nucleotide binding proteins, metabolic kinases, FAD containing proteins and one heme containing protein. **b)** Targets revealed by the Kinobeads screen include many more protein kinases (blue dots) than the originally intended kinase targets (purple dots).

Selectivity of kinase inhibitors was analyzed according to their binding type, their designated targets and the clinical phase they reached. In the past, type 2 inhibitors were thought to be more selective than type 1 inhibitors owing to the exploitation of an additional less conserved hydrophobic pocket. However, the Kinobeads data showed that type 2 span a similar selectivity range than type 1 inhibitors: both inhibitor designs can yield very selective but also very unselective molecules. Type 3 inhibitors (mainly targeting MAP2K1, MAP2K2) were the most selective. Allosteric (type 4) inhibitors (such as Rapamycin) did not score in this assay since they did not alter the ATP pocket in a way that prevented binding to the Kinobeads. Selectivity was also not driven by the designated target; analysis of EGFR and MET inhibitors showed that their designated inhibitors can be both, very selective and very promiscuous. Selectivity did also not correlate to the clinical progress of drugs. This can be discussed from two angles: selectivity seems not to be required for clinical success of a drug candidate, and scientific progress did not succeed in generating more selective drugs for clinical evaluation.

The results of the Kinobeads drug screen illustrate that elucidation of the target space of a drug is crucial to understand its mode of action, to explain positive or adverse side effects, or to discover possibilities to use this drug for other applications (drug repurposing) as presented in detail in the PhD thesis of Dr. Susan Kläger.

This work uses the Kinobeads drug screen data to identify new starting points for medicinal chemistry efforts aiming at novel kinase targets. One such novel target is the receptor tyrosine kinase EPHA2. Large parts of this work are based on the Kinobeads drug screen data and present how this data set was used to inform a medicinal chemistry program in different ways – by generating a novel metric to calculate compound selectivity, identifying clinical EPHA2 inhibitors, developing a way to connect structure-affinity-relationships to selectivity, and choosing an appropriate lead structure for the discovery of dedicated EPHA2 inhibitors.

1.2 Measuring drug selectivity

As already described in the introduction, selectivity is a quite vague term as it depends a lot on the particular question a researcher is asking. Large inhibitor screening studies might not be interested in the selectivity against a particular protein target but would rather be interested in a more global determination of inhibitor selectivity independent from certain reference kinases. Yet, in a clinical research project, drug selectivity might be understood as selectivity of an anticipated effect of a drug against other (undesired) side effects. In such circumstances, selective targeting of a group of proteins determining a certain mode of action might be of even higher interest than a single reference kinase or the global selectivity of a drug. In a medicinal chemistry approach as presented in this thesis, it is desirable to determine the selectivity of novel molecules for a certain designated target of interest.

Besides the problem of an unclear definition of selectivity *per se*, the determination of selectivity is also very much dependent on the availability and quality of assay data. The apparent target spectra of a drug may vary greatly depending on the used drug concentration, depending on the size of the tested protein set and the assay format. Additional targets may arise at higher drug concentrations leading to a more unselective target profile. Also, testing the drug against a more unbiased larger screening panel will probably result in the identification of more (maybe unexpected) target proteins than a screen against a selected subpopulation of targets as often done in medicinal chemistry programs. Frequently, compounds are only tested in single doses and lack full dose response data, which renders these data prone to the individual errors of each single measurement.

Due to the multifaceted limitations of selectivity determination, a selectivity metric is required that is flexible enough to be utilized for different research questions and can cope with different assay and value formats. Thus, the novel scoring system *concentration and target dependent selectivity* (CATDS) was elaborated which enables the calculation of drug selectivity in a target- and concentration-dependent manner. CATDS is based on a thermodynamic background and can be flexibly applied to a range of topics relevant for basic biology, drug discovery and clinical research. It is highly compatible with different conditions imposed by differing data input and the desired output, as detailed hereafter.

1.3 CATDS is a thermodynamics-based metric suitable for Kinobeads data

Thermodynamic background of CATDS. If an inhibitor is added to a pool of target proteins (such as in cells or cell lysates) and the thermodynamic equilibrium is established, the inhibitor will bind to its target proteins according to the applied drug concentration and the respective binding affinities. It is assumed that the concentration of each target protein is much lower than the specific dissociation constant of the drug-protein interaction which assures that the actual protein concentration is not affecting the determination of K_D^{app} values. It is also assumed that the binding of a drug to its target proteins does not reduce the effective concentration of a drug, ensuring that each individual drug-protein binding equilibrium can be established at the applied drug concentration. Both assumptions are prerequisites for determining binding affinities in a dose response experiment. If these basic assumptions are met, we can derive the relative target engagement of each individual protein at each particular drug concentration directly from its dose response curve (e.g. 50% target engagement if the drug concentration is equal to the K_D^{app}).

With increasing drug concentration, more and more protein binding sites will be occupied by the inhibitor and target engagement will eventually near 100%. This is true for any target protein present e.g. in a cell albeit the concentration at which this happens will be different between proteins as their affinities for the drug are not the same. This has important consequences for determining drug selectivity. This becomes clear, if we consider two hypothetical inhibitors A and B: Inhibitor A has one target with a K_D^{app} of 1 nM, ten targets with a K_D^{app} of 100 nM and no further targets beyond 100 nM; Inhibitor B that has two targets with a K_D^{app} of 1 nM each and no further targets beyond 1 nM. Both inhibitors have different selectivities depending on the applied drug concentration. Inhibitor A is more selective at 1 nM, whereas inhibitor B is more selective at 100 nM. As a consequence, drug selectivity should generally be determined as a function of the applied inhibitor concentration.

Requirements for Kinobeads data. As described in the introduction, Kinobeads selectivity profiling is based on a thermodynamic equilibrium that is established between the compound, the protein and the Kinobeads affinity matrix. Profiling is performed in a complex cell lysate comprising a pool of target proteins – thus, reflecting the thermodynamic background of CATDS also in the experimental setting. This renders CATDS a suitable selectivity metric especially for (but not limited to) Kinobeads data. However, a few details need to be considered when using CATDS for this type of data. The identified target space of a drug obviously depends on the set of assayed proteins. As discussed earlier, in this screen, >300 probable direct binders to Kinobeads including kinases and other ATP- or cofactor-binding proteins were identified. All these were taken into account when calculating compound selectivity; in contrast, proteins not classified as direct binders (e.g. complex partners) were not considered for selectivity calculation. Another issue derives from the fact that Kinobeads experiments use native cell lysate for measuring drug-protein interactions. Often, a bottom plateau can be observed for some targets (i.e. dose response curve not reaching zero despite high compound dose). Also, the top plateau of the fit does not necessarily start at 1 which is usually due to technical reasons, for instance if the DMSO control used for normalization showed low MS intensity. Therefore, dose response curves should be bottom- and top-corrected for CATDS calculation to enable the use of the full effect size range between 0 and 1.

1.4 Calculation of CATDS for different applications

General concept. The large body of data accumulated in the Kinobeads screen, enabled the conception and testing of the *Concentration and Target Dependent Selectivity* (CATDS). CATDS measures the engagement of a specific protein target at a particular drug concentration relative to all target protein engagements of that drug at the same concentration. It is calculated by dividing the sum of the target engagements of the target protein(s) of interest by the sum of all target engagements (including target protein(s) of interest) at a particular concentration. Target engagement at any concentration can be derived from dose response curve fits, such as those generated by the clinical inhibitor screen (Figure 20).

$$CATDS = \frac{\sum(target\ engagement)_{target(s) of\ interest}}{\sum(target\ engagement)_{all\ targets}}$$

To illustrate the concept, the relatively selective EGFR inhibitor Gefitinib was chosen as an example (Figure 20). The Kinobeads selectivity profile revealed seven target proteins (including

EGFR as the most potently hit target; K_D^{app} of 413 nM). In order to determine the selectivity of Gefitinib for its designated target, EGFR is chosen as the target of interest and CATDS is calculated at the respective K_D^{app} concentration.

$$CATDS_{\,EGFR}\ (Gefitinib) = \frac{\sum(target\ engagement)_{EGFR}}{\sum(target\ engagement)_{EGFR,GAK,RIPK2,RIPK3,MET,STK10,FECH}}$$

The degree of engagement for each target protein is derived from the curve fit at the chosen concentration (here K_D^{app}, Figure 20). By definition, 50% of all EGFR molecules are occupied with Gefitinib at 413 nM thus the numerator is 0.5. The denominator is the sum of the target engagements of all seven targets of Gefitinib.

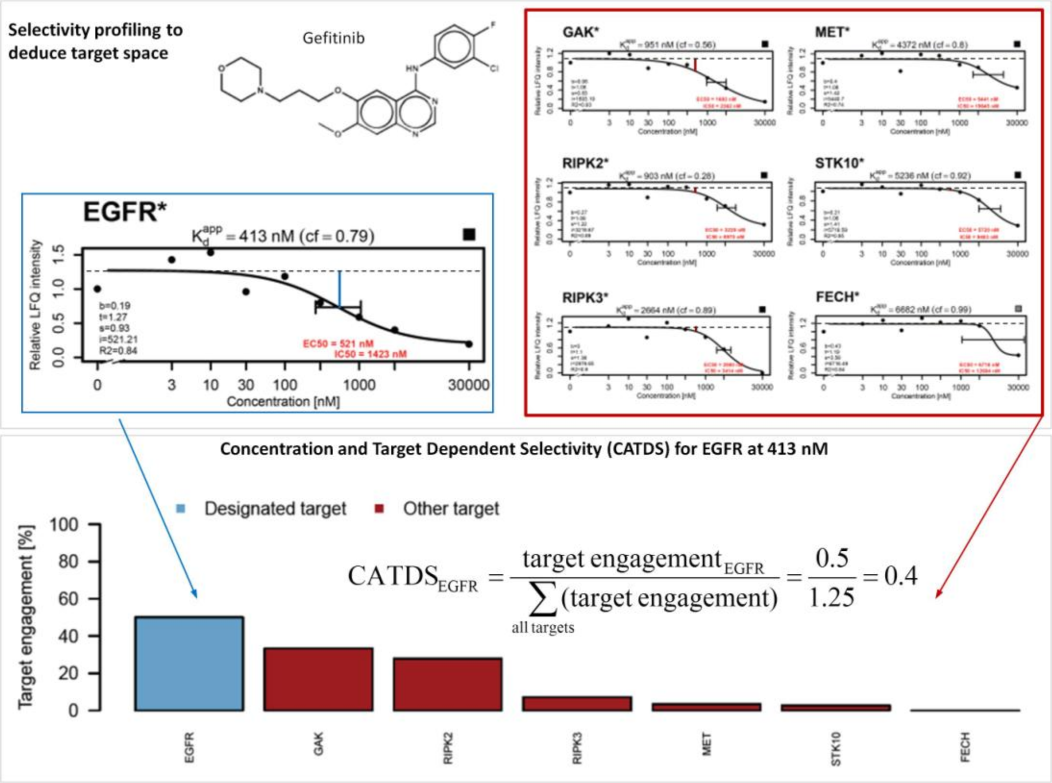


Figure 20 | Determination of compound selectivity by the Concentration and Target Dependent Selectivity (CATDS) score. An example of a CATDS calculation is provided for Gefitinib (upper panel). Kinobeads profiling determined seven targets for Gefitinib (blue box for EGFR and red box for the six other targets). The target engagement (blue vertical bar in the dose-response plot for EGFR; red vertical bar in dose response plot for other targets) is shown for all targets. The histogram in the middle panel uses the same color code. The selectivity of Gefitinib for EGFR ($CATDS_{EGFR}$) can be calculated by dividing the target engagement of EGFR (here at its K_D^{app} of 413 nM, *i.e.* 0.5, blue bar) by the sum of all target engagements (blue and red bars) at the same concentration. More generally, using the fitted curves from the Kinobead assay, $CATDS_{EGFR}$ can be calculated at any concentration (lower panel, blue dashed line) to monitor the selectivity across the entire concentration range.

Different flavors of CATDS. CATDS can be calculated for a single target protein of interest or groups of target proteins of interest ($\text{CATDS}_{\text{target}}$, $\text{CATDS}_{\text{multi-target}}$). Obviously, the protein of interest can be the protein(s) the inhibitor was originally developed for ($\text{CATDS}_{\text{designated}}$, $\text{CATDS}_{\text{multi-designated}}$). CATDS is also applicable to a more global view on drug selectivity regardless of a particular target ($\text{CATDS}_{\text{most-potent}}$) and provides the possibility to calculate the selectivity for several modes of action ($\text{CATDS}_{\text{MoA}}$).

$\text{CATDS}_{\text{target}}$ describes the selectivity of a compound towards a particular target. $\text{CATDS}_{\text{target}}$ can be useful in medicinal chemistry programs or to choose the most selective inhibitor for a biological or biochemical experiment. It is calculated by dividing the target engagement of a certain fixed protein of interest by the sum of all target engagements. $\text{CATDS}_{\text{target}}$ is determined for each particular compound at the respective K_D^{app} concentration of the selected target protein.

As an example of interest for the evaluation of EPHA2 inhibitors, Dasatinib comprises 66 target proteins in total and binds EPHA2 as off-target with an affinity of 6 nM, as determined by Kinobeads. $\text{CATDS}_{\text{EPHA2}}$ enables the calculation of the selectivity of Dasatinib for EPHA2 at K_D^{app}: 6 nM. Dasatinib hits many more targets at this concentration which is reflected by a low $\text{CATDS}_{\text{EPHA2}}$ of 0.05.

$$CATDS_{EPHA2}(Dasatinib) = \frac{\sum(target\ engagement)_{EPHA2}}{\sum(target\ engagement)_{66\ target\ proteins\ in\ total}} = 0.05$$

To exemplify a potential application of $\text{CATDS}_{\text{target}}$, $\text{CATDS}_{\text{CHEK1}}$ was calculated to characterize and select valuable CHEK1 inhibitors among the 19 compounds identified to target CHEK1 in the Kinobeads drug scree. Initially, only 4 of them (AZD-7762, PF-477736, Rabusertib, SCH-900776) were designated CHEK1 inhibitors (Figure 21). Rabusertib was found to be the most selective molecule for CHEK1 in this screen, whereas the RTK inhibitor Sunitinib was the most unselective molecule targeting CHEK1 among 68 other kinases. Moreover, this analysis depicts that the selectivity of each compound is a function of drug concentration. Rabusertib is not the most affine (K_D^{app}: 43 nM) but the most selective CHEK1 inhibitor ($\text{CATDS}_{\text{CHEK1}}$: 1) as CHEK1 is the only target identified in the drug screen. Both, SCH-900776 and PF-477736 comprise relatively high selectivity ($\text{CATDS}_{\text{CHEK1}}$: 0.74 and 0.73) which is reached when using the compounds at low concentrations (K_D^{app} (SCH-900776): 11 nM and K_D^{app} (PF-477736): 1 nM). At higher concentrations, off-target inhibitions start to play a role and the compounds become less selective for CHEK1. AZD-7762 is a potent designated CHEK1 inhibitor (K_D^{app}: 5 nM) but with a total of 70 targets not very selective. Still, CHEK1 is one of the most potently hit targets leading to an acceptable selectivity of $\text{CATDS}_{\text{CHEK1}}$: 0.32 at K_D^{app} concentration. Hence, CATDS can be used to find a suitable compound and the concentration window that would provide highest target selectivity: this analysis showcases that the most potent compound is not necessarily the optimal choice to achieve optimal selectivity.

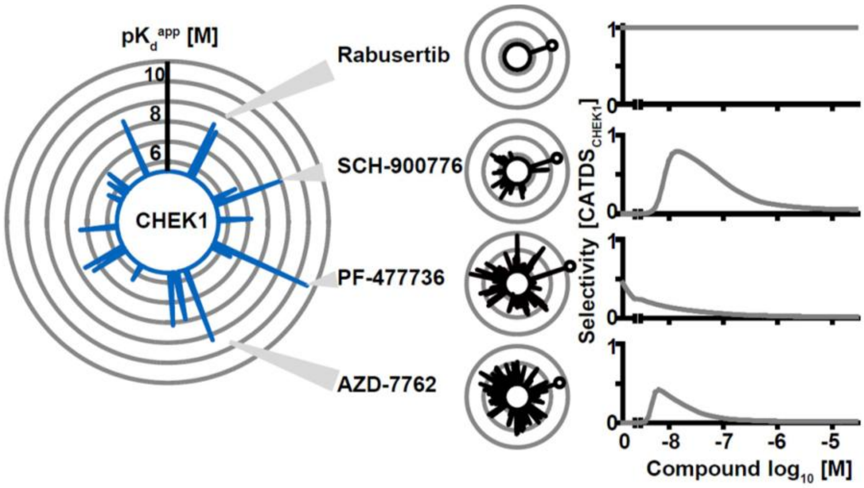


Figure 21 | Using $CATDS_{target}$ to identify selective CHEK1 inhibitors. The large radar plot shows all CHEK1 inhibitors identified in the Kinobeads drug screen (each spike is a drug and the length of the spike is indicative of binding affinity). The smaller plots depict the number and affinity of targets for Rabusertib, SCH-900776, PF-477736 and AZD-7762. The plot to the right shows that the selectivity of each compound ($CATDS_{CHEK1}$) is a function of drug concentration. AZD-7762 is a potent CHEK1 inhibitor, but is not selective at any concentration. PF-477736 and SCH-900776 are selective at lower doses and Rabusertib is selective at all doses as no other targets beside CHEK1 were observed in this screen.

$CATDS_{designated}$ is a particular case of $CATDS_{target}$, where the protein of interest is the designated target the inhibitor was designed for. $CATDS_{designated}$ helps to evaluate the selectivity for the protein which is intended to be inhibited by a certain inhibitor. As described for $CATDS_{target}$, it is calculated by dividing the target engagement of the designated target protein by the sum of all target engagements. The calculation is performed at the K_D^{app} concentration of this particular target.

For example, Dasatinib is approved for the treatment of BCR-ABL positive chronic myeloid leukemia. The molecular cause of this pathological condition is a genetic fusion of BCR and ABL1 causing constitutive activation of ABL1. The designated target of Dasatinib is ABL1 which is bound with a K_D^{app} of 5 nM. As already described for EPHA2, there are many targets hit at this concentration leading to a low $CATDS_{designated}$ of 0.06 (Figure 22).

$$CATDS_{designated}(Dasatinib) = \frac{\sum(target\ engagement)_{ABL1}}{\sum(target\ engagement)_{66\ target\ proteins\ in\ total}} = 0.06$$

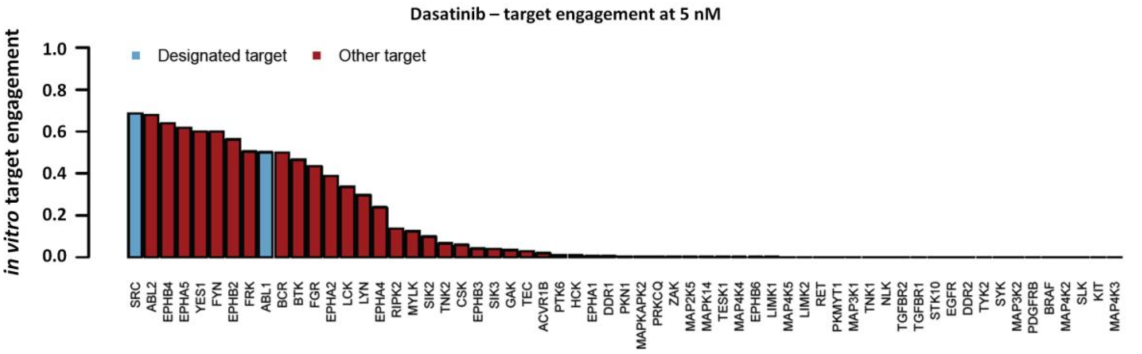


Figure 22 | *In vitro* target engagement of protein targets. The histogram shows the *in vitro* target engagement of Dasatinib targets at 5 nM compound concentration (as determined by Kinobeads selectivity profiling). CATDS calculation is based on the *in vitro* target engagement of all targets at a particular concentration. CATDS$_{\text{designated}}$ is calculated for ABL1 at its K_D^{app} concentration (5 nM). CATDS$_{\text{multi-designated}}$ is determined for SRC and ABL1 at the concentration of the least potently hit target (here: ABL1 at 5 nM).

CATDS$_{\text{most-potent}}$ is a particular case of CATDS$_{\text{target}}$ that should enable a more global view on compound selectivity, regardless of a certain target of protein. CATDS$_{\text{most-potent}}$ can be used to assess the general applicability of a compound as a chemical probe or to analyze large inhibitor screens where no focus is set on a particular target or concentration. It is calculated by dividing the target engagement of the most potent target protein of a compound by the sum of all target engagements. The calculation is performed at the K_D^{app} concentration of this particular target. The most potent target of an inhibitor is not necessarily the protein an inhibitor was originally designed for.

The most potently hit target of Dasatinib is SRC (K_D^{app} of 3 nM) which is one of its designated targets but not the main target it is used for in therapy. Calculating CATDS$_{\text{most-potent}}$ for Dasatinib yields low selectivity of 0.10 which mirrors the high amount of off-targets at this concentration range. Thus, Dasatinib is not a very selective inhibitor for none of its (off-)targets and not suitable as chemical probe.

$$CATDS_{most-potent}(Dasatinib) = \frac{\sum(target\ engagement)_{SRC}}{\sum(target\ engagement)_{66\ target\ proteins\ in\ total}} = 0.10$$

CATDS$_{\text{multi-target}}$ can be calculated if the selectivity of a compound against a group of proteins is of interest. To do so, the sum of binding reduction of all targets of interest is used to calculate drug selectivity. In this scenario, special emphasis needs to be put on the concentration that is used for selectivity determination. In order to include all target proteins to CATDS calculation, the K_D^{app} concentration of the least potently hit target seems preferable, since this concentration ensures that at least 50 % of each target is occupied by the drug.

Dasatinib is a rather promiscuous compound that binds many EPH receptors as off-targets, which could qualify Dasatinib as a valuable pan-EPH receptor inhibitor. ***CATDS$_{\text{multi-target}}$*** can be calculated in order to evaluate the selectivity of Dasatinib for the EPH family of kinases.

$$CATDS_{multi-target}\ (Dasatinib) = \frac{\sum(target\ engagement)_{\text{EPH receptors}}}{\sum(target\ engagement)_{\ 66\ target\ proteins\ in\ total}} = 0.30$$

CATDS$_{multi\text{-}designated}$ is a particular case of multi-target selectivity calculation and is calculated by dividing the sum of target engagements of all designated target proteins by the sum of all target engagements. Many inhibitors are intended to address several target proteins at once and thus, selectivity for the intended use of that compound should consider all designated targets relative to all targets of this compound. The calculation is performed at the K_D^{app} concentration of the least potent designated target protein to ensure that all designated target proteins are included in the calculation.

Dasatinib is described as a dual BCR-ABL/SRC inhibitor. It targets ABL1 and SRC with an apparent affinity of 5 nM and 3 nM, respectively. CATDS$_{\text{multi-designated}}$ considers the summed binding reduction of ABL1 and SRC and is calculated at K_D^{app} of ABL1 (5 nM; Figure 22). Considering both designated targets increases the selectivity of Dasatinib (CATDS$_{\text{multi-designated}}$: 0.14) as in comparison to the selectivity against ABL1 alone (CATDS$_{\text{designated}}$: 0.06).

$$CATDS_{multi-designated}(Dasatinib) = \frac{\sum(target\ engagement)_{ABL1,SRC}}{\sum(target\ engagement)_{66\ target\ proteins\ in\ total}} = 0.14$$

CATDS$_{MoA}$ is designed to distinguish the selectivity of a compound for certain mode of actions. Key requirement for this analysis are different protein subsets that distinguish certain modes of actions. The considered target space is restricted to these proteins, other proteins that are not relevant for any of the analyzed modes of actions are excluded for this analysis. CATDS$_{\text{MoA}}$ is calculated by dividing the sum of target engagements of selected proteins by the sum of target engagements in this (restricted) set of target proteins. The calculation is performed at the K_D^{app} concentration of the most potent target protein of the selected target group. This is based on the assumption that single inhibition of one member is sufficient to influence the respective mode of action and not all members need to be inhibited. However, the concentration used for CATDS calculation may be adjusted to the K_D^{app} concentration of the least potently hit target if required.

With regards to Dasatinib, it is interesting to determine how selective this drug is in inhibiting EPHA receptor signaling amongst all EPH tyrosine kinases. To do so, the considered target space of Dasatinib is restricted to all EPH receptor kinases and the EPHA receptors constitute the group of target proteins of interest for CATDS calculation. CATDS is calculated at K_D^{app} of the most potently hit EPHA receptor (EPHA5 at 4 nM).

$$CATDS_{MoA-EPH}\ (Dasatinib) = \frac{\sum(target\ engagement)_{\text{EPHA receptors}}}{\sum(target\ engagement)_{EPH\ \text{receptors}}} = 0.48$$

CATDS$_{\text{MoA}}$ was found of particular relevance to analyze inhibitors targeting cyclin-dependent kinases (CDK). In the Kinobeads drug screen, 12 CDKs were identified to be targeted by 36 small molecules. These CDKs were categorized into several modes of action: cell cycle (CDK1/2/4/5/6), dual role in cell cycle and transcription (CDK7), transcription (CDK9/11/12) or other processes (atypical; CDK16/17/18). CATDS$_{\text{MoA}}$ was calculated for each inhibitor and each mode of action; the background target space was restricted to all CDKs. Clustering of the calculated CATDS$_{\text{MoA}}$ (Figure 23) revealed that most designated CDK inhibitors were not very selective for certain modes of actions. Interestingly, some compounds binding CDKs as off-targets showed rather high selectivity for a certain biological role. This implies that more selective CDK inhibitor design towards a particular mode of action might be feasible. However, it needs to be considered that this analysis focuses on CDK selectivity only and that many of these non-designated inhibitors (e.g. K-252a) have very promiscuous target profiles apart from CDK inhibition.

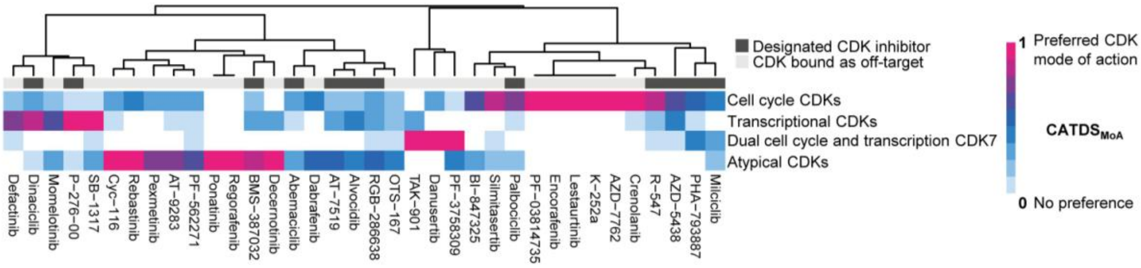
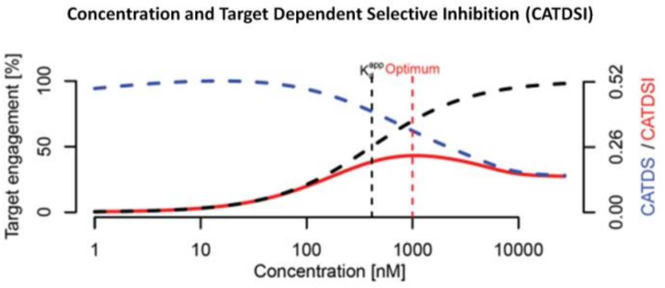


Figure 23 | Using CATDS$_{MoA}$ to analyse CDK inhibitor mode of action. Clustering of inhibitors targeting CDKs according to CATDS$_{MoA}$. Compounds either had no preference (white-light blue) or targeted the same biological process (pink; e.g., the cell cycle). Off-target CDK inhibitors (light grey bar) often inhibited one potential CDK mode-of-action.

Calculation of CATDSI. As discussed earlier, drug selectivity can vary with increasing drug concentration – a fact which underscores the necessity of calculating the CATDS across the tested concentration range to identify the optimal concentration for maximal selectivity (Figure 24). However, the effect of a drug is determined by the engagement of its target protein[317], as estimated by the dose response curve in this screen. That means, that concentrations below K_D^{app} might result in high selectivity for a certain target, but drug efficacy may be very low. In order to determine the most selective yet effective concentration of a drug, the CATDS can be multiplied with the target engagement at each drug concentration to generate the *Concentration And Target Dependent Selective Inhibition* (CATDSI). The maximum of this curve highlights the drug concentration at which a balance between selectivity and potency of a drug for a certain target is reached (Figure 24, red line).

$$CATDSI = CATDS \times \sum(target\ engagement)_{target(s)\ of\ interest}$$



$$\text{CATDSI} = \text{CATDS} \times \text{target engagement}$$

Figure 24 | Expansion of compound selectivity calculation to the Concentration and Target Dependent Selective Inhibition (CATDSI). A further useful metric is the Concentration and Target Dependent Selective Inhibition (CATDSI) which provides an estimate of the optimal concentration at which the highest selectivity and highest target engagement can be obtained. Thus CATDSI is the CATDS multiplied by the target engagement of the selected target (lower panel, dashed black line).

1.5 Comparison of CATDS with other scoring systems

As described in the introduction, selectivity score, gini coefficient, selectivity entropy and partition index are metrics which cover some aspects of selectivity and are valuable to answer particular questions. The CATDS metric however combines several important aspects of selectivity determination which have not been addressed in their entirety by the other selectivity metrics published so far (Figure 25). In this chapter, CATDS is compared to these main metrics and is shown to outperform all the others in terms of versatility, precision and stability.

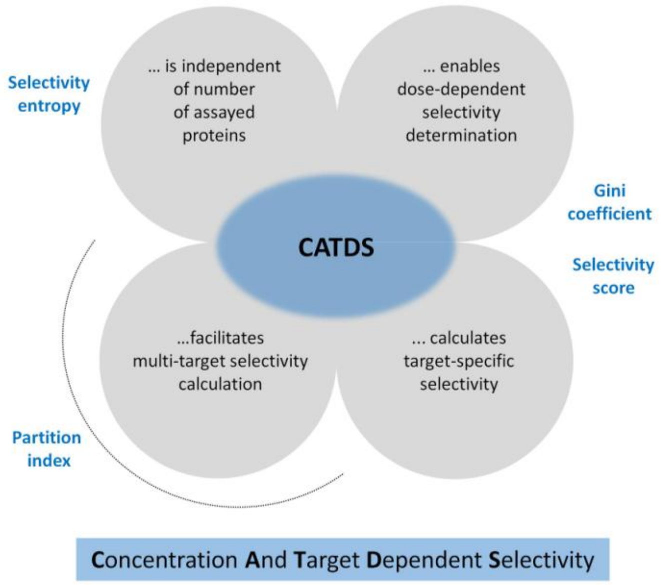

Selectivity entropy
... is independent of number of assayed proteins
... enables dose-dependent selectivity determination
Gini coefficient
CATDS
Selectivity score
...facilitates multi-target selectivity calculation
... calculates target-specific selectivity
Partition index
Concentration And Target Dependent Selectivity

Figure 25 | Comparison of CATDS to other selectivity metrics. Comparison of CATDS features to selectivity entropy, selectivity score, gini coefficient and partition index. CATDS encompasses features of all the other metrics but only CATDS provides an assay-independent selectivity measure for single and multiple proteins of interest at any concentration.

CATDS versus selectivity entropy. Selectivity entropy calculates compound selectivity irrespective of a particular target or concentration. In order to assess global selectivity as obtained by the calculation of the selectivity entropy, $\text{CATDS}_{\text{most-potent}}$ is the most suitable CATDS score. $\text{CATDS}_{\text{most-potent}}$ is calculated at the K_D^{app} concentration of the most potent identified target protein in order to provide a view of compound selectivity independent of its designated target protein. It (inversely) follows the same trend as observed in the selectivity entropy (Figure 26a) – showing that CATDS is equally capable of covering this aspect of global compound selectivity. In contrast to CATDS however, selectivity entropy does not allow to determine the selectivity of a compound for a particular target protein. Comparison of $\text{CATDS}_{\text{multi-designated}}$ and selectivity entropy highlights this advantage of CATDS. It reveals that there are compounds with very low selectivity entropy values (allegedly quite selective molecules) but it ignores the possibility that another target is more potently inhibited than the intended target (i.e. the selectivity entropy values are artificially low; Figure 26b). In contrast, $\text{CATDS}_{\text{multi-designated}}$ is able to distinguish whether a compound is selective for its intended use (blue) or if another protein is targeted more potently (black).

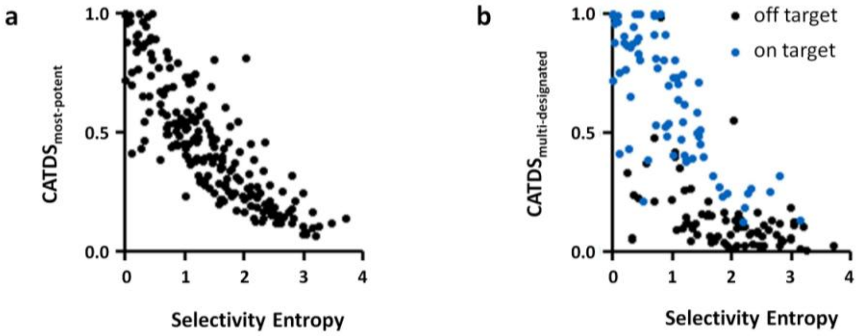


Figure 26 | Comparison of CATDS to selectivity entropy. a) Selectivity entropy is comparable to $\text{CATDS}_{\text{most-potent}}$. **b)** The graph indicates that the selectivity entropy is unable to differentiate compounds that are on-target (blue circles) or off-target (black circles) proteins, whereas $\text{CATDS}_{\text{multi-designated}}$ allows this distinction.

CATDS versus selectivity score. The selectivity score is a bit more informative than the selectivity entropy as it allows for a concentration-dependent view on compound selectivity. As exemplified in Figure 27a, *in vitro* target engagement is calculated at a certain concentration; targets reaching the cutoff are counted whereas other targets are not. This leads to the problem that the selectivity score is highly dependent on the screened panel size, as already discussed in the introduction. The selectivity score S changes massively if the same inhibitor (here: Dasatinib at 5 nM) is screened against 300 proteins as done in Kinobeads (S: 0.033) or if it is screened for instance against all 58 receptor tyrosine kinases (S: 0.172), even if the same number of targets was identified. Kinobeads allow testing against a rather large panel size of more than 300 direct binders leading to a compression of the selectivity score to low values mainly between 0.1 and 0.003 (Figure 27b, left panel). A comparison of $\text{CATDS}_{\text{designated}}$ (threshold concentration: K_D^{app} of the most potent designated target) and $-\log_{10}$ transformed selectivity score resolves the selectivity score distribution and shows that both scores follow the same trend to some extent (Figure 27b, right panel). However, the tight distribution of the selectivity scores shows that this score is not capable of distinguishing well between selective and non-selective compounds, especially when large panels are screened. All inhibitors appear to be rather selective (low selectivity score) whereas the CATDS distribution covers the entire score range (0-1) and adequately reflects different inhibitor selectivity towards their designated targets. Another major drawback of the selectivity score is the fact that it simply counts the number of targets at a certain concentration and does not include dose-dependent inhibition or the slope of the measured binding curve.

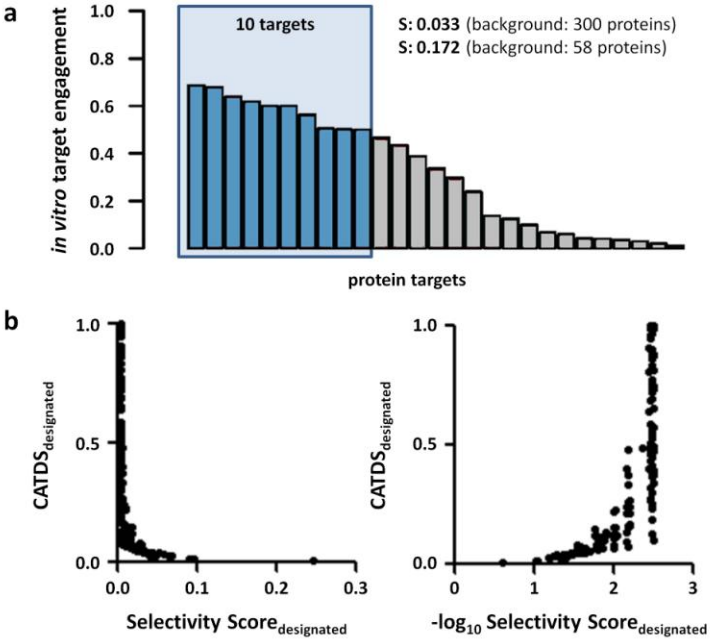


Figure 27 | Comparison of CATDS to selectivity score. a) Excerpt of Dasatinib target profile at 5 nM drug concentration. Selectivity score uses a concentration cutoff to distinguish targets from non-targets. Targets are counted and divided by the total number of screend proteins. All assayed proteins are included in calculation, which leads to high dependence on screening background. **b)** Scatter plots showing the comparison of CATDS to selectivity score and $-\log_{10}$ selectivity score. Each dot refers to one tested drug. $\text{CATDS}_{\text{designated}}$ follows the same trend as the $-\log_{10}$ selectivity score.

CATDS versus gini coefficient. The Gini coefficient also allows for concentration-dependent selectivity calculation and is more precise than the selectivity score as it takes dose-dependent inhibition including slope information into account (Figure 28a). No artificial threshold is drawn as done in selectivity score calculation and the *in vitro* target engagements of all assayed proteins at a certain concentration are considered. Since also non-inhibited proteins are included in gini coefficient calculation, the distribution of gini coefficients also suffers from the large background set of protein, similarly to the selectivity score. This leads to a compressed scale with values mainly between 0.8 and 1 which hampers clear differentiation of selective and unselective compounds (Figure 28b, left panel). The CATDS, however, distributes between 0 and 1 enabling a stable and comparable determination of compound selectivity independent from the tested panel size. Another disadvantage of the gini coefficient is its inability to calculate selectivity towards a certain target protein. To estimate target-specific selectivity by the Gini coefficient, one could use the K_D^{app} of the protein of interest as a threshold concentration. Here, $CATDS_{designated}$ is compared to the Gini coefficient calculated at the same threshold concentration (K_D^{app} of the most potent designated target). CATDS and Gini coefficient show a clear correlation (Figure 28b, right panel). However, a "target-dependent" calculation of gini (or the selectivity score) only facilitates the selection of an appropriate threshold concentration; it does not calculate the selectivity of a compound towards a particular target at this concentration.

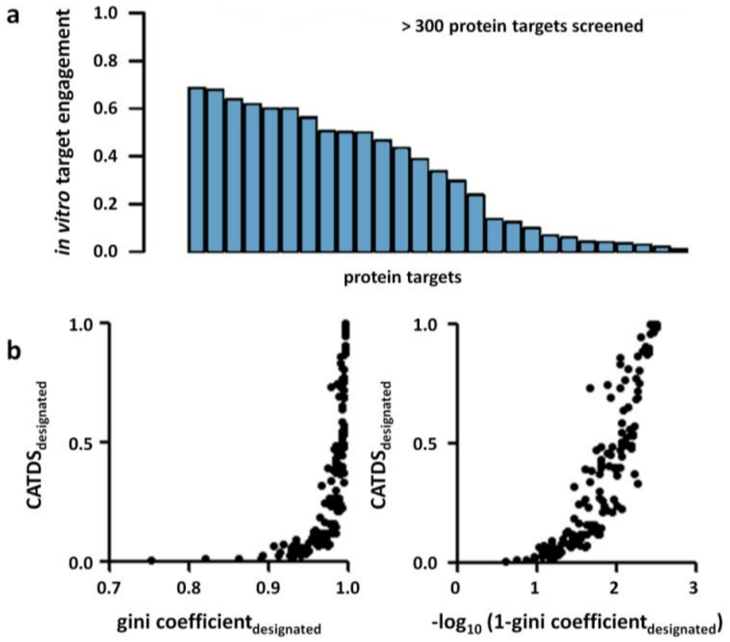


Figure 28 | Comparison of CATDS to other gini coefficient. a) Gini coefficient uses in vitro target engagement of all assayed proteins at a acertain threshold concentration. All assayed proteins are included in calculation, which leads to high dependence on screening background. **b)** $CATDS_{designated}$ correlates to the gini coefficient.

CATDS versus partition index. In contrast to the selectivity score and the gini coefficient, the partition index can be calculated in a target- but not concentration-dependent manner. It is the only score that allows to calculate target-dependent compound selectivity and can also be calculated for multiple targets. Indeed, both scoring systems follow the same trend for single (Figure 29a, reference: most-potent target protein) and multiple targets (Figure 29b, reference: multi-designated at K_D^{app} concentration of most potent target). Both, the partition index and the CATDS are based on full dose response data. However, the main advantage of CATDS over the partition index is that the CATDS calculation accounts for the curve shape of the dose response which is more precise than the K_D^{app} value that is utilized by the partition index. Thus, CATDS considers binding effects that might be reflected by different Hill slopes and that are not resolved by the partition index. Most importantly, the partition index does not allow for determining drug selectivity at individual drug concentrations which is a clear advantage of CATDS and can be observed when calculating both scores at 10 nM and 1 µM compound concentration (Figure 29c).

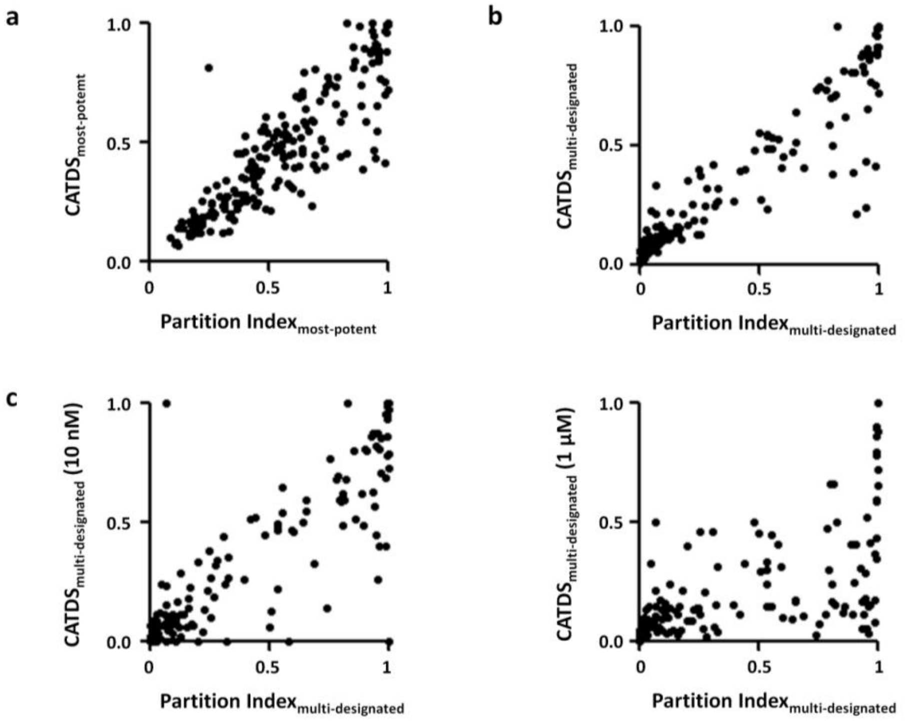


Figure 29 | Comparison of CATDS to partition index. Partition index allows target-dependent selectivity determination. There is high correlation between **a)** $\text{CATDS}_{\text{most-potent}}$ and the partition $\text{index}_{\text{most-potent}}$; and **b)** $\text{CATDS}_{\text{multi-designated}}$ and partition $\text{index}_{\text{multi-designated}}$. **c)** The comparison of $\text{CATDS}_{\text{multi-designated}}$ calculated at 10 nM (left) and 1,000 nM (right) with the partition $\text{index}_{\text{multi-designated}}$ demonstrates that the partition index cannot account for the fact that selectivity is a function of drug dose.

CATDSI. As described previously, the CATDSI facilitates an estimation of a drug concentration that provides a balance between selectivity and effect against a particular target. It requires a target- and concentration-dependent view on drug selectivity which is only available from CATDS selectivity determination. None of the other selectivity metrics is capable of providing such information – further increasing the value and novelty of CATDS and CATDSI calculation for compound characterization.

To summarize, the Kinobeads drug screen data of clinical kinase inhibitors enabled the development of a novel selectivity metric (CATDS) that was found to be more versatile and stable than previously described selectivity metrics. CATDS provides manifold ways to elucidate compound selectivity according different perspectives; in this thesis, $\text{CATDS}_{\text{EPHA2}}$ will be used to evaluate the selectivity of clinical kinase inhibitors and newly synthesized inhibitors towards the receptor tyrosine kinase EPHA2.

2 EPHA2 as target of clinical kinase inhibitors

2.1 Chemical proteomics reveals EPHA2 as common off-target

In order to identify EPHA2 inhibitors, the Kinobeads technology was used to determine the target space and binding affinities of 235 clinical kinase inhibitors (Table S1a) and three tool compounds known to target EPH family proteins (Cpd66, LDN-211904, PD-173955)[137,273,279]. The clinical kinase inhibitors that were tested are approved drugs or are administered to humans as part of clinical trials (phase 1-3) and therefore provide a highly optimized pharmacological profile. Although none of the clinical drugs tested was originally intended to target EPHA2, a total of 32 molecules showed inhibition of EPHA2 with apparent affinities ranging in K_D^{app} values from 2.8 nM to 6.1 μM (Figure 30, affinities given as pK_D^{app} values: $-\log10(K_D^{app})$; Table 1). Twenty four clinical inhibitors and the three tool compounds showed sub-micromolar K_D^{app} values and, surprisingly, twelve of these compounds had not been previously known to target EPHA2. Dasatinib (2.8 nM), Ponatinib (6.9 nM) and Danusertib (9.0 nM) showed the highest affinity in this screen. Fifteen inhibitors were validated in a HotSpot kinase activity assay, confirming that molecules binding to Kinobeads also showed inhibition of EPHA2 enzymatic activity (Table 1). As far as described in the literature, the identified set of EPHA2 inhibitors comprised ATP competitive type 1 and type 2 inhibitors providing a rich set of molecular features for the analysis of the chemical landscape of EPHA2 inhibition (see Table S1a-b for more detailed information about the inhibitor set; and Table S1c-d for binding data and respective binding curves).

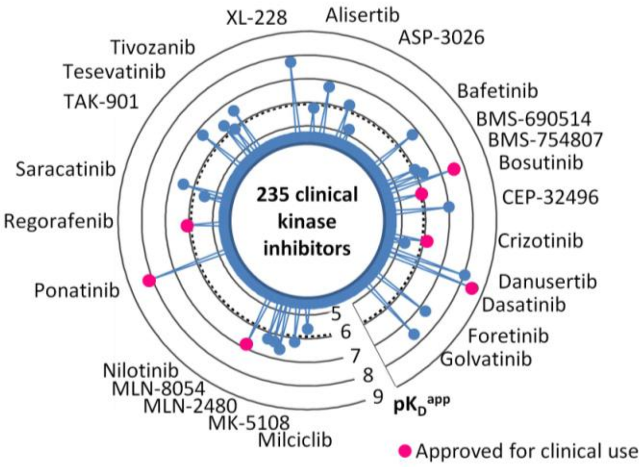


Figure 30 | Identification of EPHA2 inhibitors by chemical proteomics. Radar plot summarizing the results of chemical proteomic selectivity profiling of 235 clinical small molecule kinase inhibitors leading to the identification of 24 drugs with binding constants below 1 μM for EPHA2 (expressed as $-\log10\ K_D^{app} = pK_D^{app}$). Compounds approved for clinical use are depicted in pink.

Table 1 |Data summary of clinical kinase inhibitors identified to target EPHA2. K_D^{app} values were determined by Kinobeads binding assay, IC_{50} values were determined by HotSpot kinase activity assay. Selectivity was calculated by $CATDS_{EPHA2}$.

Compound	K_D^{app} [nM]	EPHA2 known target	Clinical status	Inhibitor type	IC_{50} [nM]	X-Ray resolution [Å]	PDB code	$CATDS_{EPHA2}$
Dasatinib	2.8	yes	approved	1	0.8	1.3	5I9Y	0.05
Ponatinib	6.9	yes	approved	2	1.3	-	-	0.06
Danusertib	9	no	phase 2	1	4.4	1.7	5I9Z	0.04
XL-228	19	no	phase 1	1	4.0	-	-	0.02
Golvatinib	25	no	phase 1/2	2	13	1.8	5IA5	0.11
Bosutinib	30	yes	approved	1	2.0	1.4	5I9X	0.03
Foretinib	52	yes	phase 2	2	1.9	1.8	5IA4	0.04
CEP-32496	99	yes	phase 1	2	28	-	-	0.07
Alisertib	187	yes	phase 3	up	48	1.9	5IA0	0.09
Bafetinib	192	no	phase 2	2	17	-	-	0.06
TAK-901	198	no	phase 1	-	4.8	-	-	0.02
Tivozanib	259	yes	phase 3	2	-	-	-	0.10
MK-5108	265	no	phase 1	-	-	-	-	0.05
Saracatinib	338	yes	phase 3	1	-	-	-	0.02
Tesevatinib	458	no	phase 2	-	-	-	-	0.03
Nilotinib	415	yes	approved	2	33	-	-	0.07
MLN-2480	484	no	phase 1	up	-	-	-	0.05
BMS-754807	525	no	phase 2	1	-	-	-	0.04
MLN-8054	530	yes	phase 1	up	-	2.1	5IA1	0.10
ASP-3026	640	yes	phase 1	-	-	-	-	0.02
Milciclib	672	no	phase 2	-	-	-	-	0.01
Crizotinib	772	yes	approved	1	-	-	-	0.04
Regorafenib	869	yes	approved	2	-	-	-	0.05
BMS-690514	961	no	phase 2	1	-	-	-	0.02
Cpd66	19	yes	tool	1	1.0	1.6	5IA2	-
PD-173955	46	yes	tool	1	7.9	1.8	5IA3	-
LDN-211904	165	no	tool	1	4.1	-	-	-

2.2 Target profiling of EPHA2 inhibitors enables rational drug repurposing

Competitive Kinobeads pulldowns determine the range of kinases that can be bound by a test compound and thus provides information about its target space and selectivity. Collectively, the 27 EPHA2 inhibitors targeted 133 protein kinases (K_D^{app} <1 μM; Table S1c-d) but the number of targets, binding affinities and selectivities (expressed by the $CATDS_{EPHA2}$) varied greatly. For example, while Tivozanib showed modest affinity for EPHA2 (K_D^{app}: 259 nM), it was one of the most selective inhibitors ($CATDS_{EPHA2}$: 0.10). In contrast, Danusertib has much higher affinity for EPHA2 (K_D^{app}: 9 nM) but was much less selective ($CATDS_{EPHA2}$: 0.04). Because of the complex relationship between compound structure, affinity and selectivity, it remains important to evaluate and compare the target space of each small molecule. Therefore, the question arised, which kinases are co-targeted by a particular compound in addition to EPHA2 and how often this occurs (Figure 31, more detailed information in Table S1d). Not unexpectedly, compounds that target EPHA2 also bind other tyrosine kinases (TK) including members of the EPH receptor family. Almost all compounds also targeted ABL1 and ABL2 because many of these are designated BCR-ABL inhibitors. SRC family kinases were also commonly hit as well as DDR1 or RET. From the tyrosine kinase like branch (TKL), RIPK2 and ZAK were the most frequently observed targets. More surprisingly, phylogenetically quite different kinases such as SIK2, GAK, AURKA and AURKB were also often targeted together with EPHA2 confirming prior observations that overall ATP binding site sequence conservation is not necessarily a dominating factor affecting drug selectivity[318,319].

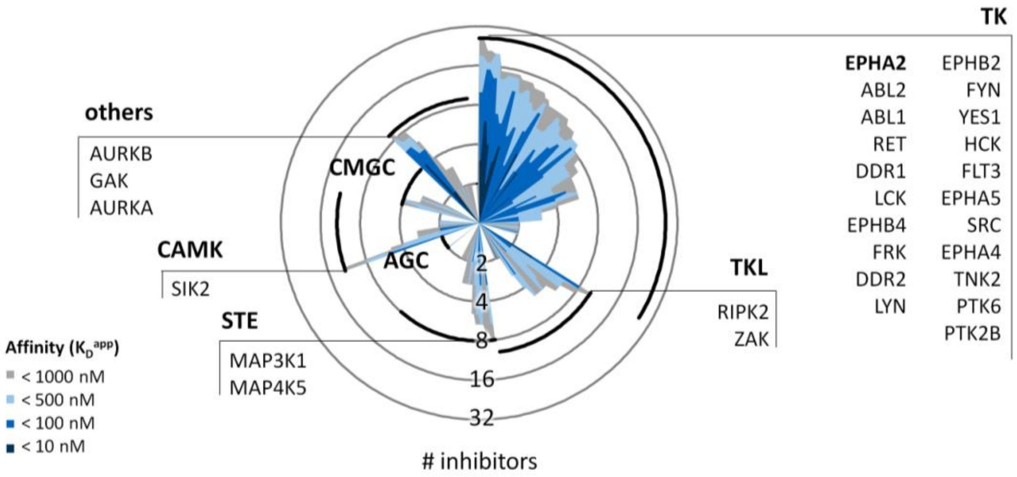


Figure 31 | Distribution of kinases co-targeted by 24 clinical EPHA2 inhibitors and 3 tool compounds. Kinases are rank ordered by phylogenetic sequence similarity and by the number of inhibitors targeting each individual kinase (affinity bins from dark blue to grey according to increasing K_D^{app}).

Given the range of proteins co-targeted by EPHA2 inhibitors, it was to be expected that these participate in numerous biological processes and co-targeting may, therefore, lead to desirable or undesirable functional outcomes. In order to investigate this further, the most frequently observed co-targets (number of inhibitors $n \geq 8$, $K_D^{app} < 1\ \mu M$) were extracted and a functional gene ontology (GO)[292] and literature analysis for the resulting 28 proteins (Figure 32) was performed. As expected, many proteins were involved in cancer-related processes (proliferation, cellular differentiation, cell cycle regulation, apoptosis, angiogenesis, migration, adhesion) and a surprisingly large number of commonly co-targeted proteins were implicated in inflammatory processes and the immune response as well as in the neuronal system. Furthermore, literature research highlighted tumor suppressive effects for some co-targeted kinases (n=4; e.g. FRK) and one potential toxicity target, ZAK, which has been implicated in the formation of cutaneous squamous cell carcinoma or keratoacanthomas[320]. Given the range of kinases identified – including many that may mediate desirable but also undesirable biological effects (e.g. targeting tumor suppressors[321], immune system regulators[322] or cell cycle mediators[323]) – it appears prudent to evaluate kinase selectivity as comprehensively as possible even in the early phases of drug discovery.

Selectivity profiling allows the rational selection of appropriate drugs with beneficial target profile for drug repurposing in EPHA2 dependent pathologies. As an example, EPHA2 can cause drug resistance to EGFR inhibitors in non-small cell lung cancer (NSCLC) cells[98,99] and targeting EPHA2 by a single agent or in combination with other drugs may constitute a viable therapeutic option. Unexpectedly, this study uncovered that the dual MET and VEGFR inhibitor Golvatinib is also a very potent EPHA2 inhibitor (K_D^{app}: 25 nM; IC_{50}: 13 nM). This is noteworthy as both MET and EPHA2 have been implicated in resistance formation against Gefitinib in NSCLC. It is therefore tempting to hypothesize that Golvatinib may overcome both these resistance mechanisms simultaneously. In fact, Golvatinib has already been shown to be active in combination with Gefitinib in a Gefitinib-resistant NSCLC setting; but the study did not investigate if this was owing to EPHA2 inhibition[324]. Another potential opportunity for drug repurposing of EPHA2 inhibitors could be the prevention of pathogen infections. Viruses[75], bacteria[79,80] and parasites[81] have all been shown to abuse EPHA2 internalization as an entry mechanism into the cell and a positive inhibitory effect on pathogen entry by abrogating ligand dependent EPHA2 signaling and endocytosis has already been found for Dasatinib[79]. Other inhibitors targeting fewer immunologically important kinases (such as Alisertib or Tivozanib) could prove to be even more useful for this application.

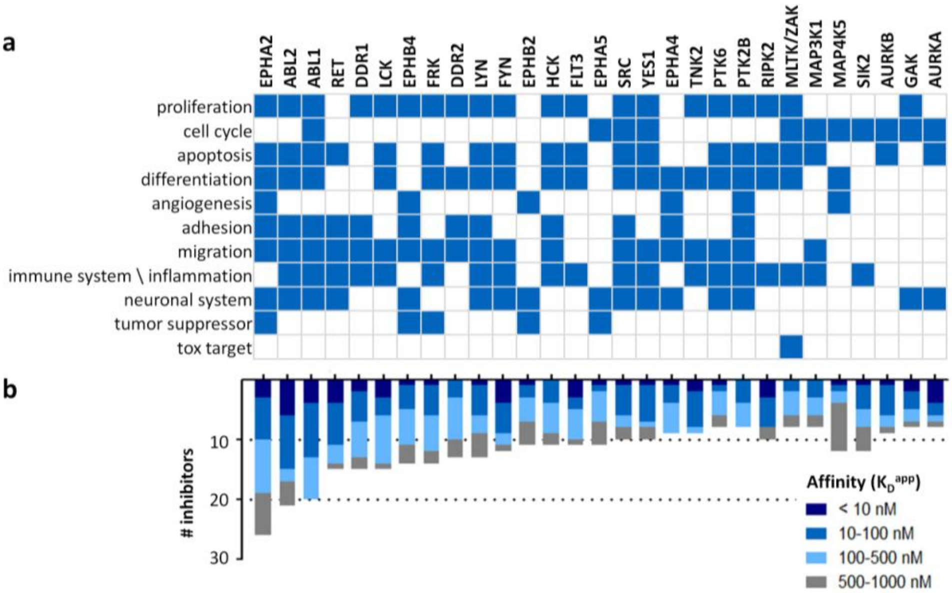


Figure 32 | Gene Ontology (GO) term analysis of commonly co-targeted kinases. a) 28 protein kinases were found to be commonly inhibited along with EPHA2 (n≥8 compounds, K_D^{app}<1 μM). GO term analysis combined with literature search showed that these proteins are implicated in many biological processes related to cancer (e.g. proliferation, apoptosis, differentiation). However, the identified co-targets were also involved in other biological processes such as the immune or the neuronal system. **b)** Bar plot depicting the number of inhibitors and affinity bins for each target.

There are many facets to and opinions about the frequently observed polypharmacology of kinase inhibitors, and these data also illustrates that selectivity is not a requirement for the success of a drug in the clinic. However, broad selectivity profiling is still crucial to understand a drug's target space and mode of action including potential toxic side effects resulting from inhibition of undesired off-targets. Knowledge about the target profile of each drug can provide guidance for its appropriate use and the anticipated phenotypic outcome. The relevance of this outcome for chemical biology and translational research is underscored by the notion that about a dozen clinical trials have been launched that aim to exploit the inhibitory effect of Dasatinib or other agents on EPHA2 (www.clinicaltrials.gov, NCT00563290, NCT01440998, NCT00895960, NCT01876212, NCT02693535, NCT02661113, NCT01591356, NCT02575261, NCT02252211, NCT00796055, NCT02754362). Should it turn out that EPHA2 inhibition is indeed a good therapeutic concept, these data can, in the short term, help clinicians to consider other EPHA2 inhibitors with more desirable target or safety profiles.

3 Structure-based residue classification for inhibitor design

3.1 Global analysis of ligand bound EPHA2 structures

In order to understand the structural basis of ligand binding and its effects on kinases, NMR spectroscopy and protein crystallography was performed using recombinant non-phosphorylated EPHA2 kinase domain. These experiments were performed by members of the research group for Biomolecular Magnetic Resonance (headed by Prof. Harald Schwalbe) at the Institute for Organic Chemistry and Chemical Biology at Johann Wolfgang Goethe-Universitaet in Frankfurt. The highly conserved Asp-Phe-Gly-motif (DFG motif) in the kinase activation loop is indicative of the active (DFG-in) and inactive (DFG-out) state of the kinase, which can be unambiguously classified by NMR spectroscopy[177]. Using 15N-Phe labeling and 1H, 15N TROSY NMR of wildtype EPHA2 and mutant Phe758Tyr EPHA2, it become apparent that ligand bound EPHA2 can adopt both conformations (Figure 33a).

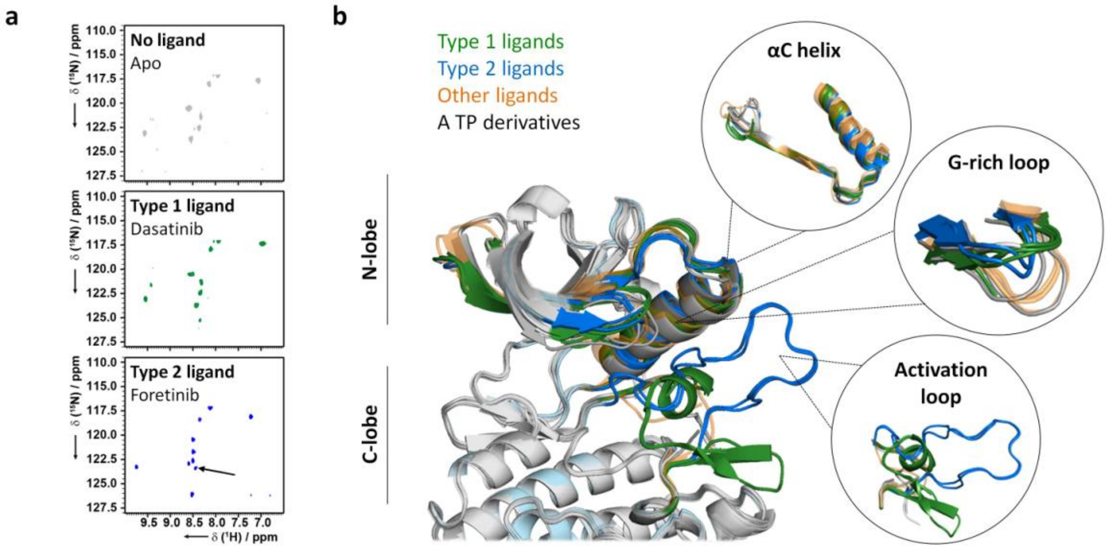


Figure 33 | Global structural characterization of EPHA2-drug interactions by NMR and crystallography. a) ^{1}H, ^{15}N TROSY solution NMR spectra of ^{15}N-Phe labeled EPHA2 without ligand (upper panel), in the presence of the DFG-in ligand Dasatinib (middle panel), and the DFG-out ligand Foretinib (lower panel). The arrow points to an additional signal corresponding to Phe758 of the DFG motif which can only be observed for DFG-out ligands thereby confirming that EPHA2 can adopt DFG-in as well as DFG-out conformations when bound by a ligand. **b)** Overlays of co-crystal structures of EPHA2 and nine kinase inhibitors. The data shows an overall high degree of structural alignment with differences in detail induced by the respective ligands (type 1: green; type 2: blue; others: orange, ATP derivatives: grey).

The absence of the backbone amide signal for Phe758 in the DFG motif indicates a conformational equilibrium at intermediate timescale causing severe line broadening in apo and DFG-in ligand (e.g. Dasatinib) bound protein (Figure 33a). A DFG-out ligand should suppress the DFG-in/out interconversion leading to the detection of the amide signal for Phe758, which was indeed found in the presence of DFG-out ligands like Foretinib (Figure 33a, indicated by arrow). In order to gain further structural insights, EPHA2 was co-crystallized with seven designated type 1 inhibitors (Bosutinib, Cpd66, Danusertib, Dasatinib, PD-173955; DFG-up: Alisertib, MLN-8054), two ATP derivatives (ATPγS [AGS], ADPNP [ANP]) and two designated type 2 inhibitors (Foretinib, Golvatinib) and at high resolution of between 1.2 and 2.0 Å (Table 1, Table S2e). The binding modes of the different ligands induced N- and C-lobe mobility and rearranged functional key motifs such as the Gly-rich loop, the activation loop and the αC helix (Figure 33b).The Gly-rich loop conformation was also altered by ligand binding (Figure 33b); while hydrogen bonds between Gly622/Glu623 and ATP derivatives (Figure 33b, grey; PDB: 5I9V, 5I9W) induced a strong loop bending compared to the ligand-free state, Alisertib and MLN-8054 rearranged the loop mainly by additive hydrophobic effects (Figure 33b, orange; PDB: 5IA0, 5IA1). In contrast, the type 2 ligand structures (Figure 33b, blue) revealed that Phe624 of the Gly-rich loop formed a tight T-shaped pi-stack with Phe758 thereby stabilizing the DFG motif in a slightly distorted conformation (PDB: 5IA4, 5IA5). Unsurprisingly, all type 1 ligands bound EPHA2 in DFG-in mode and this part of the structure aligned very well although the following activation loop orientation varied (Figure 33b, green). As previously reported for c-MET[325], the designated type 2 ligands Foretinib and Golvatinib were found to bind a hydrophobic pocket between the αC helix and the activation loop and induced an inactive DFG-out conformation (Figure 33b, blue). The αC helix orientation in EPHA2 determines the activation state of the kinase and is dependent on the formation of a Lys646/Glu663 salt bridge (Figure 34) [326], which was found in all type 1 ligand bound (PDB: 5I9Y, 5I9Z, 5IA2, 5I9X, 5IA3) or ligand-free EPHA2 structures (PDB: 5I9U). Type 2 inhibitor binding prevents the salt bridge formation (PDB: 5IA4, 5IA5) and instead shifted the αC helix by 1.3 Å, resulting in the abrogation of kinase activity.

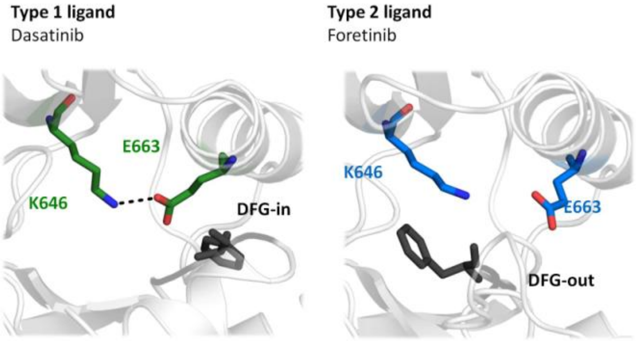


Figure 34 | Salt bridge formation. Type 1 ligands (DFG-in) support formation of a salt bridge between Lys646 and Glu663 (exemplified by Dasatinib, left panel), whereas this salt bridge is disrupted for type 2 ligands binding the DFG-out conformation of the kinase (exemplified by Foretinib, right panel).

MLN-8054 binding increased the flexibility of the activation loop resulting in a flipped DFG-in mode with Phe758 in an orthogonal orientation. Interestingly, Alisertib, which is structurally very similar to MLN-8054, adopted the typical DFG-in mode but also induced a shift of the αC helix similar to that found in type 2 ligand bound structures. In this case, the analysis of the R-spine (Figure 35a) ordering enabled more insight into the observed compound binding type and revealed Alisertib to bind as type 1 inhibitor (ordered R-spine) and MLN-8054 as type 1 ½ inhibitor (distorted R-spine). This is also reflected when comparing the novel EPHA2 co-crystal structures to the analysis by Vijayan et al. of D1 (Phe758-Asn744) and D2 (Phe758-Glu663) αC distances of all kinase structures in PDB[178]. Here, it was found that all drugs except MLN-8054 fell within the expected type 1 and type 2 binding modes (Figure 35b). Consistent with the observed DFG-flip, αC distances in the MLN-8054 structure did not cluster with the canonical DFG-in or DFG-out groups, reinforcing the notion that kinase flexibility and activation loop conformation are highly ligand-dependent.

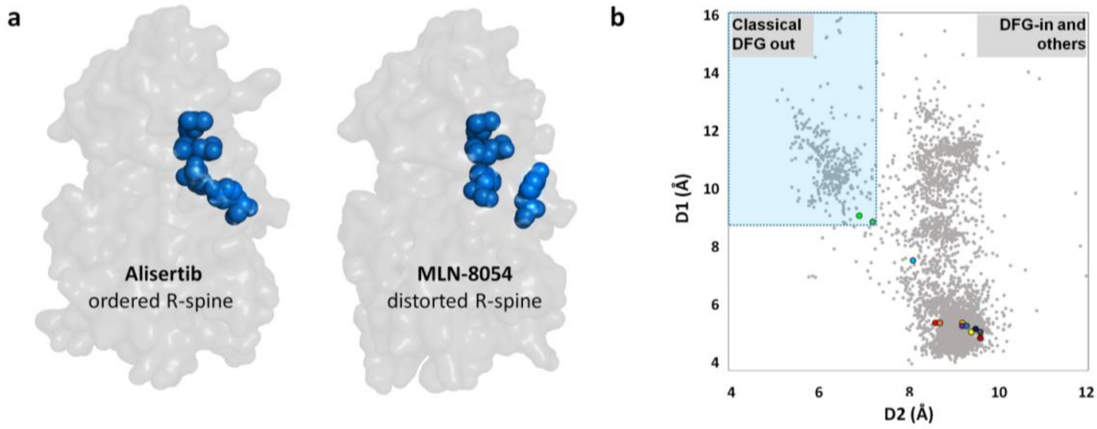


Figure 35 | Binding type confirmation. a) Binding type assignment of the AURK inhibitors MLN-8054 and Alisertib is not possible based on the classical structural features such as the DFG motif or the αC helix orientation. Analysis of the R-spine residues is necessary for unambiguous binding type assignment. Alisertib binds in type 1 binding mode, whereas MLN-8054 binds EPHA2 in type 1 ½ binding mode. **b)** Scatterplot of the D1 and D2 distances. D1 is the DFG Phe758 to Asn744 αC distance and D2 is the DFG Phe758 to salt-bridge Glu663 distance. Blue box points are assigned as classical DFG-out binders, while others are assigned to be DFG-in and other binders. Background 4221 kinase domain structures, light red: without ligand, brown: AGS, orange: ANP, deep teal: Alisertib, dark red: Bosutinib, dark grey: Cpd66, purple: Danusertib, yellow: Dasatinib, light green: Foretinib, dark green: Golvatinib, dark blue: PD-173955, cyan: MLN-8054.

3.2 Identification of drug-protein interactions in EPHA2

The nine co-crystal structures obtained for EPHA2 and its inhibitors enabled a detailed characterization of the amino acid residues involved in ligand-protein interactions (Table S2a-b). Collectively, the data revealed 36 amino acid residues that can be engaged in interactions with common or unique molecular features of the ligands and these amino acids were distributed across distinct areas of the kinase domain including the nucleotide pocket, the ribose pocket, the Gly-rich loop, the activation loop and the solvent-exposed ATP pocket entrance (Figure 36, Table Table S2a-b). As expected, ligand binding was strongly facilitated by hydrogen bond interactions within the hinge region (Figure 36, dark blue; Thr692–Gly698, hinge residue Met695) and the N-lobal beta sheet comprising the Gly-rich loop (Figure 36, red; Lys619–Lys646). Type 2 inhibitor binding was mediated by additional non-polar interactions with the αC helix (Figure 36, orange; Glu663, Ile666, Phe670, Ile675, Ile676, Ile690) and the catalytic loop (Figure 36, violet; Leu730, Tyr735, His737, Arg743, Asn744) that form the DFG-out specific hydrophobic pocket. As observed for the ATP analogs AGS and ANP, Bosutinib and Dasatinib established tighter H-bond interactions with the gatekeeper residue Thr692 (Figure 36, purple) compared to the other ligands which only exhibited non-bonded interactions. All ligands interacted directly with the catalytic residues Lys646 and Glu663 either by direct hydrogen bonding, water-mediated contacts or non-bonded interactions. These interactions were often enabled by oxygen atoms provided by carbonyl (e.g. in Dasatinib), methoxy (e.g. in Danusertib), hydroxyl (e.g. in Cpd66) or by halogen functionalities (e.g. in Bosutinib). The activation loop (Figure 36, light blue; Leu746, Val755–Phe758) harboring the functionally important DFG motif was also targeted by all inhibitors. The ribose pocket was commonly found to be engaged by compounds mainly via water-mediated and non-bonded interactions with Ala699. Residues Lys702 and Glu706 (Figure

36, dark green) are located at the ATP pocket entrance and form H-bond contacts with hydroxyl moieties (Dasatinib), carboxylic acids (Alisertib) or water-mediated interactions with tertiary amines (Bosutinib, Danusertib).

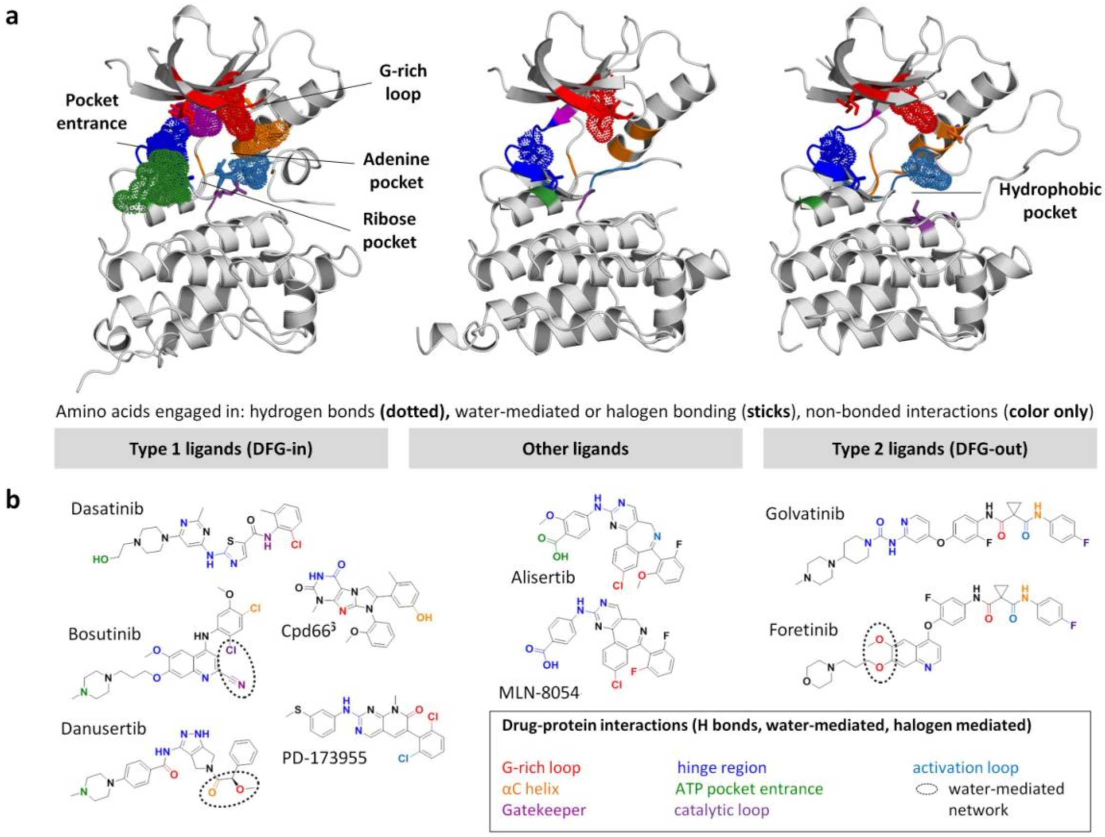

a
Pocket entrance
G-rich loop
Adenine pocket
Ribose pocket
Hydrophobic pocket
Amino acids engaged in: hydrogen bonds **(dotted)**, water-mediated or halogen bonding (**sticks**), non-bonded interactions (**color only**)
Type 1 ligands (DFG-in)
Other ligands
Type 2 ligands (DFG-out)
b
Dasatinib
Bosutinib
Danusertib
Cpd66[3]
PD-173955
Alisertib
MLN-8054
Golvatinib
Foretinib
Drug-protein interactions (H bonds, water-mediated, halogen mediated)
G-rich loop
αC helix
Gatekeeper
hinge region
ATP pocket entrance
catalytic loop
activation loop
water-mediated network

Figure 36 | Detailed analysis of drug-protein interactions elucidate the structural basis for EPHA2 inhibition. a) EPHA2 inhibitors establish interactions in different regions within the ATP pocket including the Gly-rich loop (red), the αC helix (orange), the gatekeeper Thr692 (purple), the hinge region (dark blue), the ATP pocket entrance (green), the catalytic loop comprising the HRD motif (violet) and the activation loop containing the DFG motif (light blue). Amino acids engaged in hydrogen bonding are shown as dotted patterns, water-mediated or halogen bonding is shown as sticks and hydrophobic non-bonded interactions are shown in plain color. **b)** Structures of the nine EPHA2 inhibitors for which co-crystal structures were obtained and grouped according to binding mode. Chemical moieties interacting with amino acids in a directed manner via hydrogen bonds, water-mediated bonds or halogen bonds are colored according to the color coding used for the crystal structures in panel a. Moieties engaged in water-mediated networks are highlighted by dashed circles.

Interestingly, we were able to recapitulate the conserved water-mediated H-bond network observed for Bosutinib in SRC [327] and found similar networks for Danusertib as well as for Foretinib in EPHA2 (dashed ellipse). A comparison to protein structures containing the investigated set of ligands and their primary targets revealed a highly conserved drug-protein interaction pattern (Table S2c-d). Still, minor differences exist which might determine the selectivity of a particular compound. For example, binding of Alisertib or MLN-8054 to their main target AURKA (PDB: 2WTW) is characterized by an interaction with Arg137 located at the N-lobal ATP pocket entrance of AURKA; but EPHA2 lacks this residue and the compounds engaged an interaction with Lys702 and Ala699 instead (Figure 37, this work, PDB: 5IA0, 5IA1).

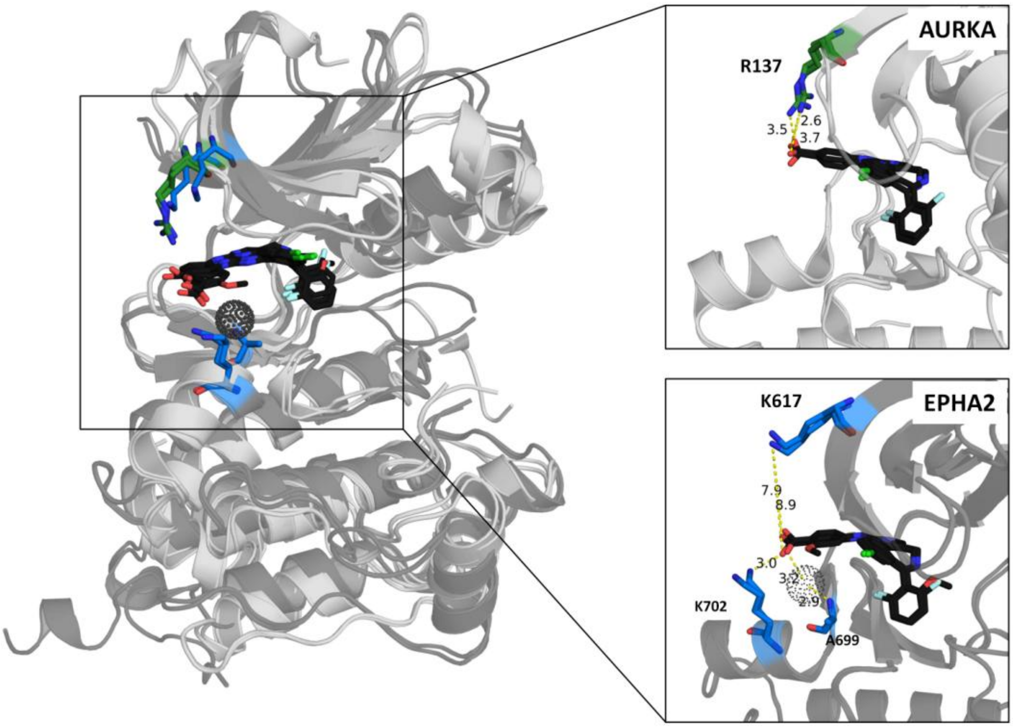


Figure 37 | Comparison of drug protein interaction in EPHA2 and AURKA. Alisertib and the structurally related MLN-8054 adopt a different pose in the entrance of the ATP pocket in EPHA2 and AURKA. Both molecules comprise a carboxylic moiety pointing towards the solvent exposed ATP pocket entrance. In the main target AURKA, this moiety establishes a salt bridge to Arg137 located at the N-lobe (upper zoom in, showing MLN-8054, PDB: 2X81 and 2WTW). EPHA2 lacks this Arg, but features a neighboring Lys which is not capable of establishing a similar interaction. Instead we found that the carboxylic moiety interacts tightly with Lys702 located at the C-lobe and establishes a water mediated bond to the backbone of Ala699 (lower zoom in).

3.3 Classification scheme for amino acids involved in drug binding

With drug-protein interaction information involving 36 amino acids of EPHA2 determined from nine co-crystal structures and chemical proteomic selectivity profiles for 27 EPHA2 inhibitors and their 133 target proteins in hand, the next step was to identify amino acids in the EPHA2 kinase domain that would inform the rational development of selective molecules. By integrating the biochemical and structural information with sequence alignments of human kinase domains[313] (Table S3a–b), a classification scheme was developed that categorizes the amino acids involved in drug binding into key residues, scaffold residues, potency residues and selectivity residues (Figure 38; Table S4a). In the following, this scheme is explained using Dasatinib as an example.

Amino acid frequency analysis: from the sequence alignment of all 530 human kinase domains, the amino acid frequency was calculated at each particular sequence position. For each inhibitor, the amino acid frequencies was also calculated at each position but only including those kinases that were targeted by that inhibitor (affinity of better than 1 µM; i.e. the target space of this inhibitor).

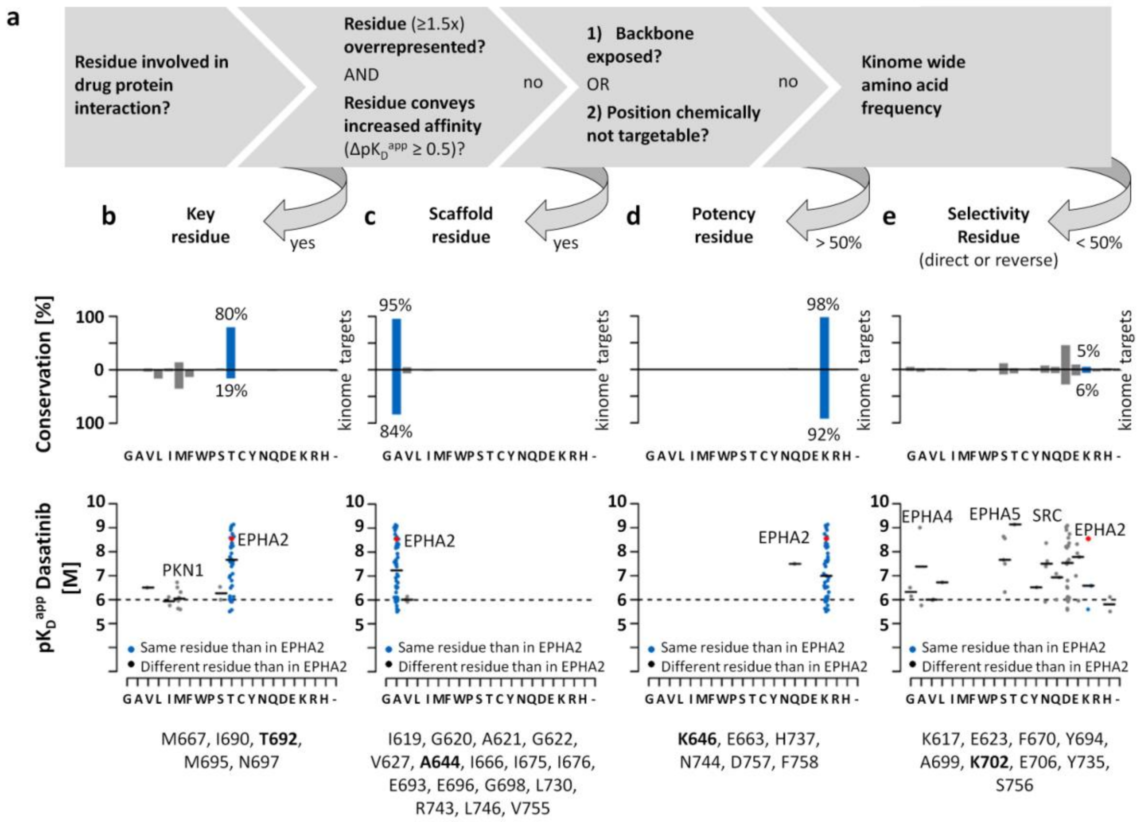


Figure 38| Classification scheme for amino acids in EPHA2 to guide inhibitor design. a) All amino acid residues found to be involved in drug-protein interactions by crystallography were classified into the categories of key, scaffold, potency and selectivity residues according to the flow chart shown depending on whether or not they showed preferential binding, have targetable side chains or are conserved across the kinome. **b)** Key residues are overrepresented in targets of a particular drug and binding to these residues can result in higher (median) affinity compared to other amino acids in the same position. Using Dasatinib as an example, it is evident that most Dasatinib targets feature a Thr residue in position 692 while only 19% of all human kinases have a Thr in this position (upper panel). It is also clear from this analysis, that targets featuring Thr at this position tend to have higher affinity for Dasatinib (lower panel, EPHA2 marked as a red dot) than targets with other amino acids in this position (e.g. Met in PKN1). A similar characteristic is found by the amino acids shown in the bottom list. **c)** Scaffold residues (e.g. Ala644) are characterized by very high conservation across the kinome (upper panel), unfavorable side chain orientation (mostly backbone exposed) or limited chemical targetability (small, nonpolar amino acids). **d)** Potency residues (e.g. Lys646) are physically accessible (side chain exposed) and chemically targetable but also tend to be highly conserved across the kinome (upper panel) as most are important for catalysis. Such residues cannot contribute to selectivity. **e)** Selectivity residues (e.g. Lys702) show low conservation both within the target space of an inhibitor as well as across the kinome (upper panel). Their side chains are exposed towards the ATP pocket and are generally chemically tractable leaving room for the development of selective compounds.

Key residues: First, it was examined for each sequence position whether the amino acid found in EPHA2 was overrepresented (at least by a factor 1.5) in the target space of the drug compared to the frequency of this residues in all kinases. For example, the residue Thr692 was found in 80% of all Dasatinib targets but only in 19% of all human kinases (Figure 38b, upper panel). Then, it was analyzed whether Dasatinib targets harboring the same amino acid as found in EPHA2 are bound with higher affinities than Dasatinib targets carrying different amino acids in the same position. For example, Dasatinib targets containing a Thr had a higher (median) affinity to the compound than proteins that harbor a different amino acid in this position (e.g. PKN1). The criterion for this was a minimal difference of 0.5 in pK_D^{app} to the most conserved amino acid in all kinases. A similar pattern was also observed for Met667, Ile690, Met695, and Asn697 of EPHA2 and other compounds targeting this protein (Figure 38b bottom list; for all structure-selectivity-relationship plots see Dataset S1). The properties of key residues may offer the possibility to optimize the selectivity as well as the potency of compounds in a targeted way.

Scaffold residues: Next, the pocket architecture was analyzed with respect to the physical accessibility (orientation of side chains) and the chemical targetability of the amino acids involved in drug binding (Figure 38c). Sequence positions which either show i) backbone exposition to the drug or ii) are in general chemically inert were categorized as “scaffold residues” because they may contribute to ligand binding primarily by formation of a hydrophobic pocket to fit the molecule. It was found that scaffold residues are often highly conserved (exemplified by Ala644 which is found in 95% of all Dasatinib targets and in 84% of all kinases; Figure 38c, upper panel). Such residues are generally difficult to engage in a directed drug-protein interaction which is why they are of lesser value for medicinal chemistry.

Potency residues: these residues contain a chemically targetable and physically accessible side chain, are highly conserved in human kinases (criterion: presence in at least 50% of all kinases) and are often directly involved in the catalytic process. For example, Lys646 is found in 98% of all Dasatinib targets and in 92% of all kinases (Figure 38d upper panel). While compound interactions with these residues can drive potency, they cannot contribute to selectivity.

Selectivity residues: In contrast to potency residues, selectivity residues are characterized by low amino acid conservation across the kinome. For example, Lys702 is found in 5% of all Dasatinib targets and in 6% of all kinases (Figure 38e upper panel). Selectivity residues feature an exposed side chain which often allows for additional drug interactions for example via hydrogen bonding or salt bridges to drive selectivity against other kinases (e.g. an Asp residue in SRC, a Thr residue in EPHA5, an Ala residue in EPHA4 all of which have very different chemical properties than Lys702 of EPHA2, Figure 38e bottom panel). Within this class, Ala699 is a special case (Figure 39): it is not directly targetable but most human kinases contain a hydrophilic or charged residue at this position, which could potentially become accessible to inhibitors (e.g. a Ser residue in EPHA5 or an Asn residue in ABL1). These residues were termed “reverse selectivity residues” and such interactions should be avoided in order to gain selectivity for EPHA2.

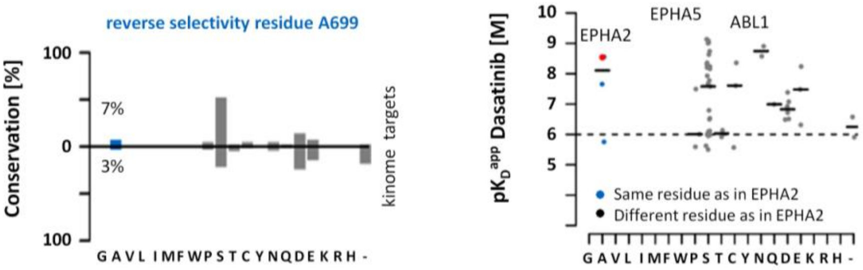


Figure 39 | Reverse selectivity residue A699. The “reverse selectivity” residue Ala699 is (unlike the straight selectivity residues) not directly targetable, but most human kinases contain a hydrophilic or charged residue (e.g. Ser, Asp) in this position, which could potentially become accessible to inhibitors and should therefore be avoided when aiming at making selective EPHA2 inhibitors.

An interesting example is provided by the selectivity residues Tyr694 and Tyr735 which are described autophosphorylation sites of EPHA2.[59] Tyr694 was found to interact with all co-crystallized inhibitors and it could be hypothesized that permanent phosphorylation of this site could lead to drug resistance. Tyr735 phosphorylation was previously shown to mediate interaction with PIK3R1 and to drive the oncogenic potential of EPHA2.[328] Since this residue would only be accessible by type 2 inhibitors, sustained Tyr735 phosphorylation might prevent binding of such molecules and might motivate the development of compounds that would target this phosphorylated form of the protein.

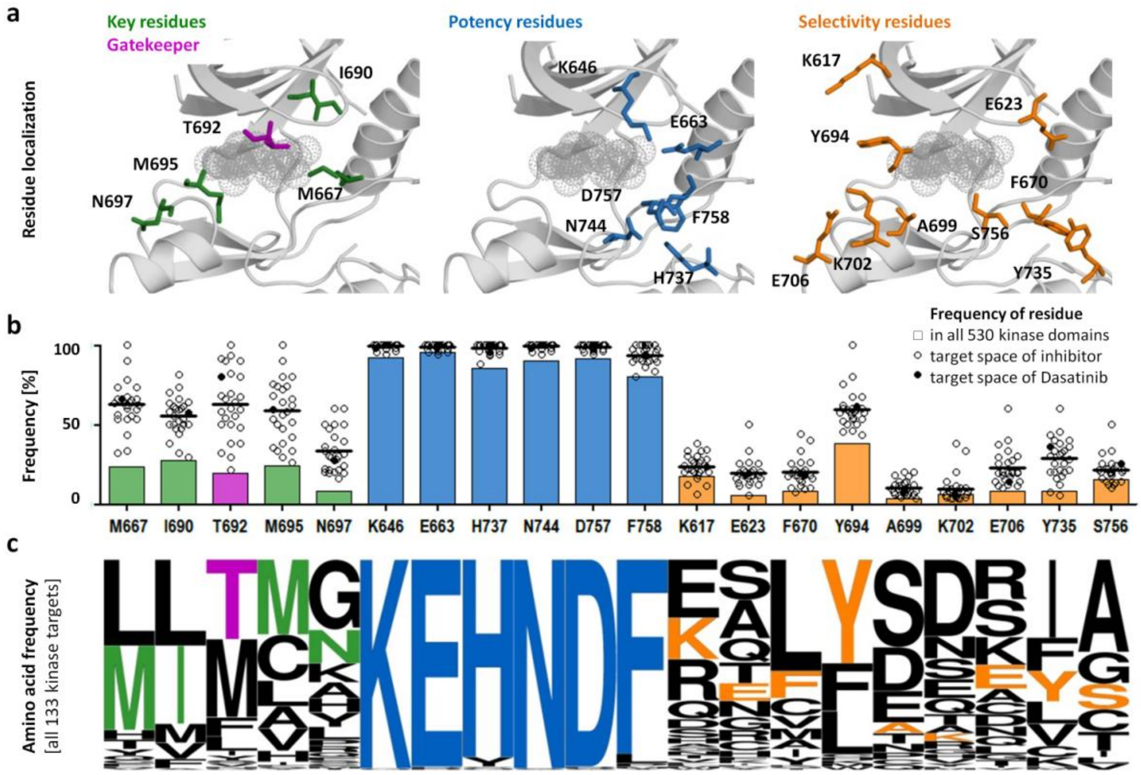


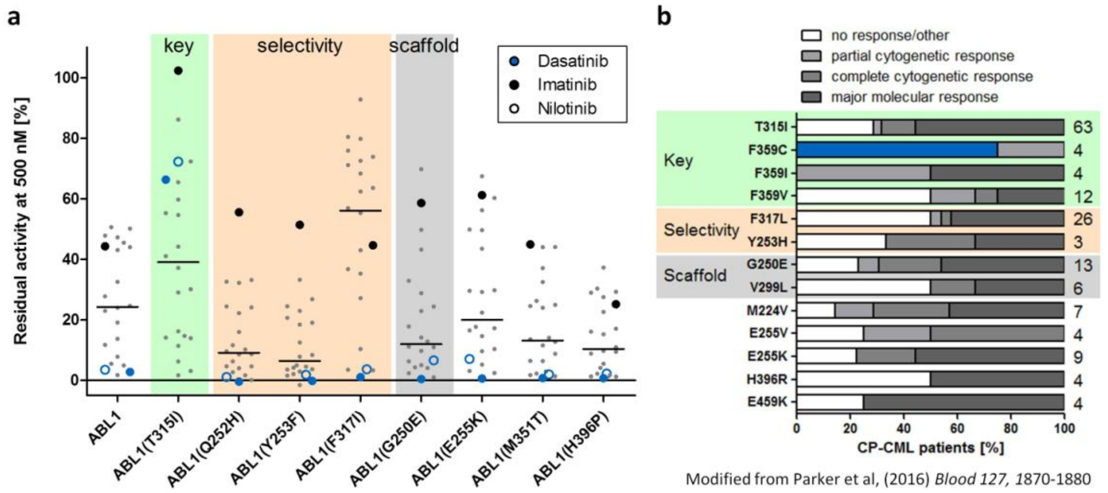
Figure 40 |Structure and sequence space for the development of selective EPHA2 inhibitors. a) Spatial positioning of key, potency and selectivity residues within the EPHA2 drug binding site (AGS-EPHA2 co-crystal). **b)** Frequency analysis of each amino acid in a particular position across the kinome (filled bars) or within the target space of each of the 27 identified EPHA2 inhibitors (each inhibitor is illustrated as an open circle; Dasatinib is highlighted as a black dot). It is evident that the amino acids found in EPHA2 are also preferentially found in targets of EPHA2 inhibitors (circles well-spaced from the bar) and that this effect is larger for key residues than for selectivity residues. **c)** Frequency analysis of amino acids across the 133 human kinases targeted by the 27 EPHA2 inhibitors. Key residues (in green and purple) are frequently found in the target proteins of an inhibitor but are generally not highly conserved across the kinome (comparison to bars in panel b). Potency residues (blue) are highly conserved and thus add little selectivity to a drug-protein interaction. In contrast, selectivity residues (orange) generally exhibit low conservation in the target space of EPHA2 inhibitors. Scaffold residues are omitted for clarity.

Figure 40a illustrates the position of all 20 amino acids classified as key, potency or selectivity residues in the drug binding site of the EPHA2 crystal structure (see Figure S1 and Table S4b for the full analysis including scaffold residues). The above data showed that a classification scheme can be developed on a limited number of compound co-crystal structures and chemoproteomic selectivity profiles. In order to test if the derived classifications hold, the analysis was extended to all 27 EPHA2 inhibitors and their targets identified in the chemical proteomics experiments. It is evident from Figure 40b, that key residues are overrepresented in the target space of most inhibitors (open circles) compared to all kinases (bars) and that potency residues are highly conserved across the kinome and target space. Selectivity residues (orange) generally exhibit lower frequency in the target space of EPHA2 inhibitors likely because these compounds were never intended to target this protein thus leaving further room for the development of more selective compounds for EPHA2. More specifically, when aggregating all 133 target proteins of the 27 EPHA2 inhibitors into an amino acid frequency plot (Figure 40c), it becomes possible to prioritize residues for medicinal chemistry starting e. g. from existing compounds as lead structures. In particular, Lys617 would be an attractive residue as this Lys occurs rarely in other kinases and would thus likely reduce the target space of an inhibitor addressing this residue. Similar arguments apply for Glu623, Lys702 or Glu706 as these residues are also chemically very different to other amino acids in this position.

3.4 Expanding the classification scheme

In the light of the steadily growing number of ligand bound protein crystal structures and compound selectivity data, this classification scheme may be broadly applicable also for other protein kinases (or even other protein classes). To substantiate this hypothesis, the same scheme was applied to the protein kinases ABL1 and MELK.

ABL1. ABL1 is one of the most extensively studied kinases as it forms part of the BCR-ABL translocation and causes the development of chronic myelogenic leukemia. The structural data bank PDB comprises 78 ABL1 structures with 53 different ligands. Kinobeads selectivity profiling data was obtained for 10 of those inhibitors and 96 protein kinases were identified to be co-targeted together with ABL1 in a submicromolar affinity range (Table S5a). A detailed PISA and LigPlot+ analysis was performed (Table S5b) on 15 publically available co-crystal structures (3UE4, 2V7A, 2GQG, 1M52, 3OXZ, 3IK3, 3QRI, 3QRJ, 1IEP, 2HYY, 3PYY, 1OPJ, 3CS9, 2E2B, 2F4J) of ABL1 wildtype and mutant protein (T315I: 2V7A, 3IK3, 3QRJ; H396P: 2F4J). The analysis encompassed three DFG-in, two DFG-outlike and ten DFG-out ligands and yielded 42 residues involved in drug protein interaction. Residue classification was performed as described for EPHA2 in the previous chapter (for results see Table S5c-e and Dataset S2) and identified 6 “key residues” (M290, I313, T315, M318, Y320 and F359), 21 scaffold residues (L248, G249, G250, G251, V256, A269, V270, K285, V289, I293, L298, V299, E316, T319, G321, L354, I360, R367, L370, V379 and G383), 8 selectivity residues (Q252, Y253, E282, F317, N322, D325, E329, reverse selectivity residue A380) and 7 potency residues (K271, E286, H361, R362, N368, D381 and F382). Both, ABL1 and EPHA2, classifications lead to very similar results which substantiates the general applicability of the classification workflow. Key and potency residues were annotated at the same sequence positions. The only difference was found at the ABL1 key residue Phe359 which was classified as selectivity residue in EPHA2 (Tyr735); a difference which is probably due to the better coverage of drug protein interactions involved in type 2 ligand binding in the ABL1 structures. ABL1 kinase inhibitors were among the first ones to be approved for clinical use (Imatinib, 2001) and since that time many ABL1 mutations have been found in CML patients. Residues which are frequently involved in mutations are T315, F317, G250, M351, F359, E255, V299, F359, Y235, M244, H396, E255, F359[329,330]. The gatekeeper mutation T315I is probably the most common and most studied ABL1 mutation and mediates resistance against several kinase inhibitors (e.g. Imatinib, Dasatinib, Nilotinib, Figure 41a)[331].



Modified from Parker et al, (2016) *Blood 127,* 1870-1880

Figure 41 | Activity assay data and effect of ABL1 mutations on treatment response. a) Duong-Ly [18] et al performed a HotSpot kinase inhibitor screen profiling 183 kinase inhibitors against 76 recombinant mutant kinases. For ABL1 it was found that 22 inhibitors reduced ABL1 activity by at least 50% at an applied dosage of 500 nM. The mutations that have been tested include key, selectivity, scaffold and non-categorized residues. The key residue mutation T315I showed the highest impact on drug efficacy of designated ABL1 inhibitors (Dasatinib, Imatinib, Nilotinib), whereas other mutations affecting scaffold (G250E) or selectivity residues (Q252H) did not lead to decreased drug potencies. Mutation of the selectivity residue F317 led to diminished drug efficacy of several inhibitors, but did not have an impact on designated ABL1 inhibitors (Dasatinib, Imatinib, Nilotinib). **b)** Parker et al[17] observed different response rates of patients carrying different ABL1 mutations during Ponatinib treatment. Ponatinib is a designated ABL1 inhibitor designed to overcome the most common ABL1 mutation affecting the gatekeeper (T3135I). In this study, the most pronounced effect was observed for the F359C mutation (highlighted in blue) which was categorized as a key residue and seems to have a major impact on the clinical outcome of Ponatinib treatment.

Mutation of F359 (to Val, Cys or Ile) was found to mediate poor clinical response rates during Ponatinib treatment[331] (Figure 41, F359C highlighted in blue). Both residues are categorized as key residues in our classification scheme which reinforces the hypothesis that these residues might have a major impact on drug binding and removal of these residues might lead to drug resistance. Conversely, other mutations affecting scaffold residues in our classification (e.g. G250E) or selectivity residues (e.g. Q252H, Y253F) were shown to influence drug efficacy[331] and treatment response[330] to a lower extent or not at all.

MELK. The classification scheme was also used to categorize amino acid residues in the structurally different serine-threonine kinase MELK. Little is known about the structure-activity-relationship (SAR) of MELK and its inhibitors as information on MELK as a target and compounds from medicinal chemistry programs focusing on MELK have just begun to emerge in the literature[332-334]. The Kinobeads drug screen can be used as a rich source of molecules for SAR elucidation and identify potential lead structures for future medicinal chemistry campaigns. Here, 16 compounds were found to target MELK as an off-target and seven of those drugs bound and inhibited the kinase with sub-micromolar affinity (Table S6a; Nintedanib, PF-3758309, K-252a, Lestaurtinib, CC-401, Defactinib, BI-847325). Co-crystallization experiments yielded high resolution crystal structures of the MELK kinase domain in complex with the inhibitors Nintedanib, PF-3758309, K-252a, Defactinib and BI-847325 (Figure 42a, for data processing and structure refinement statistics see Table S6c; PDB codes: 5MAF, 5MAG, 5M5A, 5MAH, 5MAI respectively). These compounds comprise a diverse set of different pharmacophores, which enabled the determination of a range of different drug-protein interactions. In addition, Nintedanib and BI-847325 are of particular interest as they share large parts of the molecular scaffold but differ significantly in their binding affinities (53 nM and 918 nM, respectively) and selectivity profiles in the Kinobeads assay. Ligand interaction analysis identified a total of 26 drug-protein interactions which were located at different positions within the ATP pocket (e.g. G-rich loop, hinge region, activation loop, catalytic loop, etc.). By applying the classification scheme to the obtained MELK data set, 14 scaffold residues (I17, G18, T19, G20, F22, A23, V25, A38, L61, E87, P90, G92, E136, L139), 5 potency residues (K40, E57, N137, D150, F151), 5 selectivity residues (E15, C70, Y88, C89, E93) and 2 reverse selectivity residues (L86, I149); Figure 42b-c; Table S6b, Dataset S3) were identified.

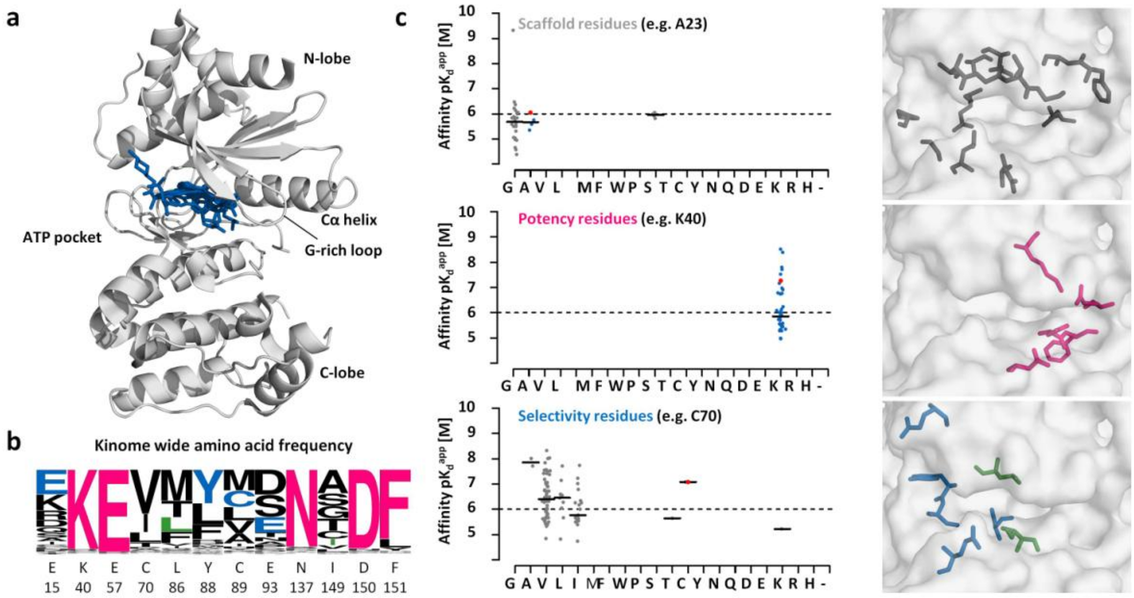


Figure 42 | Residue classification for serine/threonine kinase MELK. a) Co-crystal structures of MELK with Nintedanib, K-252a, PF-3758309, Defactinib and BI-847325 (superimposed compound structures in the ATP pocket in blue). **b)** The sequence logo shows kinome wide frequency of drug-interacting residues. **c)** Drug-interacting residues are classified as scaffold (grey), potency (pink) and selectivity (blue) residues (middle panel) and are localized within the ATP-pocket (right panel). A complete list of all residues plus classification is provided in Table S6b.

No key residues could be assigned which is probably due to different factors: i) the number of analyzed crystals is still too low, ii) pharmacophore diversity between the compounds is quite high, iii) affinity of the inhibitors is not extremely high and iv) the intended targets of the inhibitors (VEGFR/FGFR/PDGFR, PAK4, PRKC, MEK/AURK) are structurally very different from MELK. The potency residue E57 forms direct interactions to Nintedanib and the potency residue N137 to PF-3758309 and K-252a, respectively, which might be responsible for the higher affinity of these compounds compared to Defactinib and BI-847325. Particularly, the potency residue E57 appears to have a strong influence when comparing Nintedanib and the structurally similar inhibitor BI-847325. Co-crystal structures of both compounds revealed very different binding modes of Nintedanib (DFG in, αC helix in, type 1) and BI-847325 (DFG in, αC helix out, type 1½). The shift of the αC helix is induced by the more bulky ethylamide substituent of BI-847325, leads to a disruption of the characteristic salt bridge between K40 and E57 (salt bridge indicates active kinase conformation) and prevents BI-847325 from establishing an interaction with E57 as observed for Nintedanib. This might explain the large affinity difference towards MELK. The five selectivity residues (E15, C70, Y88, C89, E93) could potentially facilitate MELK inhibitor design as exemplified by the direct interactions engaged between Nintedanib and E15 (kinome-wide conservation level: 6%) or K-252a and E93 (kinome-wide conservation level: 15%). The selectivity residues C70 and C89 are readily accessible in the center of the ATP pocket and open up the possibility to design irreversible inhibitors for MELK. Given the low conservation level of both cysteines (C70: 1%, C89: 19%) and the fact that this combination of cysteines is present in only one more protein kinase (TBCK), one can anticipate that such an irreversible inhibitor would not only gain in potency and increase the drug's residence time but also provide selectivity towards MELK.

To summarize, a novel approach combining Kinobeads selectivity profiling data, structure-affinity-relationship analysis derived from protein crystallography, and kinome-wide sequence alignment enabled the development of a residue classification scheme that categorizes amino acids according to their utility for targeted inhibitor design. This residue classification was developed for EPHA2 but is also applicable to other proteins. It enables a novel and interesting perspective on structure-affinity/selectivity-relationships and offers a formalized concept to the discovery of more affine and more selective inhibitors. In this thesis, this residue classification was utilized to guide inhibitor design towards the discovery of dedicated EPHA2 inhibitors.

4 Design and synthesis of dedicated EPHA2 inhibitors

4.1 New EPHA2 inhibitor design by hybridization of known EPH binders

Dasatinib features a very broad target profile with low nanomolar affinities mainly towards tyrosine kinases (TK) such as the ABL family, SRC family or EPH family (Figure 43). In total, Dasatinib targets 44 protein kinases with sub-micromolar affinity, including many proteins of the TKL, AGC, CAMK subclasses of kinases.

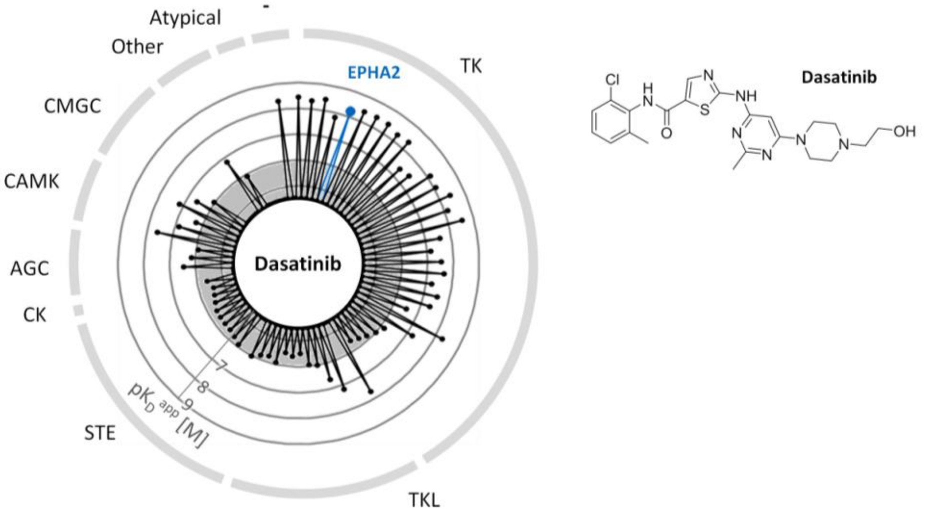


Figure 43 | Dasatinib target profile. The dual BCR-ABL/SRC inhibitor Dasatinib comprises a very broad selectivity profile targeting 44 proteins with sub-micromolar affinities across several kinase families. Dasatinib is the most potent known EPHA2 inhibitor and was chosen as molecular scaffold for the development of dedicated EPHA2 inhibitors.

It is hypothesized that a judicious choice of substituents could improve selectivity while maintaining potency against EPHA2. A minimal pharmacophore was identified as a scaffold on which to append relevant chemical moieties found in other EPH receptor binders. The previously reported crystal structure that determined the binding mode of Dasatinib to EPHA2 (PDB: 5I9Y) established that the pyrimidine moiety of Dasatinib is not engaged in EPHA2 binding. In docking experiments, it was found that two other known EPHA2 inhibitors, CHEMBL249097[335] and PD-173955[139,301] (Figure 44a), position a phenyl moiety in lieu of the pyrimidine (Figure S2). Docking experiments were kindly performed by Dr. Xiaofeng Liu at Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai, China. Hetero substitutions were therefore excluded from this aromatic ring in the inhibitor design. The X-ray structure of EPHA2 further highlighted two sites amenable for chemical modification – the ribose pocket and the ATP pocket entrance which can both accommodate spatially demanding groups (schematic representation in Figure 44b).

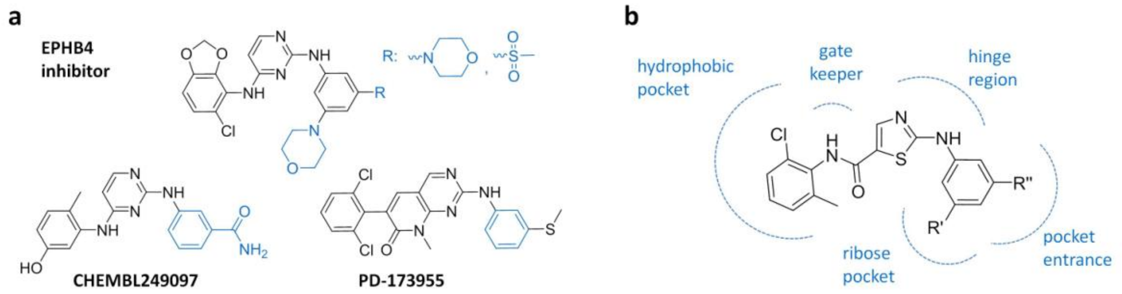


Figure 44 | EPHA2 inhibitor design by hybridization of known EPH receptor binders. a) Hybridization of known EPH binders (Dasatinib, EPHB4 inhibitors, CHEMBL249097, PD-173955) motivated the introduction of morpholino and methylsulfonyl moieties, the substitution of Dasatinib's pyrimidine by an aryl moiety, and the introduction of an amide bond in meta-position. **b)** Inhibitor design comprised a N-(2-chloro-6-methylphenyl)-2-(arylamino)thiazole-5-carboxamide scaffold binding the nucleotide binding pocket and chemical modifications which were introduced at positions R' and R'' to engage interactions within the ribose pocket and/or the ATP pocket entrance, respectively.

On the basis of this analysis, a hybrid structure of Dasatinib and previously described EPHB4 inhibitors[336] was designed (Figure 44a) to investigate the size of both pockets by introduction of bulky morpholino and methylsulfonyl substituents. In contrast to the ribose pocket, the pocket entrance is rich in selectivity residues, which are defined as side chain exposed and chemically targetable residues with low conservation across the kinome (notably Lys702, Glu706 and Lys617).[301] Direct interactions of a compound with these charged residues could potentially drive selectivity towards EPHA2 and away from many other kinases. Accordingly, introduced chemical moieties included primary and secondary amines, hydroxyl groups or carboxylic acid groups. Docking of CHEMBL249097 also revealed a putative interaction between its amide and the backbone of Glu696 located in the hinge region. Since this residue was not engaged by Dasatinib, this interaction was probed by including an amide bond at the meta-position of the aryl. This led to the consideration of four sub-series of the *N*-(2-chloro-6-methylphenyl)-2-(arylamino)thiazole-5-carboxamide series (Figure 45): series 1, to interact with Glu696 and to probe for selectivity residues interactions by introducing an amine; series 2, to engage the same residues at the pocket entrance by other chemical moieties such as fluorinated piperidines, hydroxyl groups or carboxylic acids; series 3, to examine the spatial properties of the ribose pocket by introduction of sterically demanding substituents, and series 4 to probe for dual exploitation of the pocket entrance and the ribose pocket by bulky substituents.

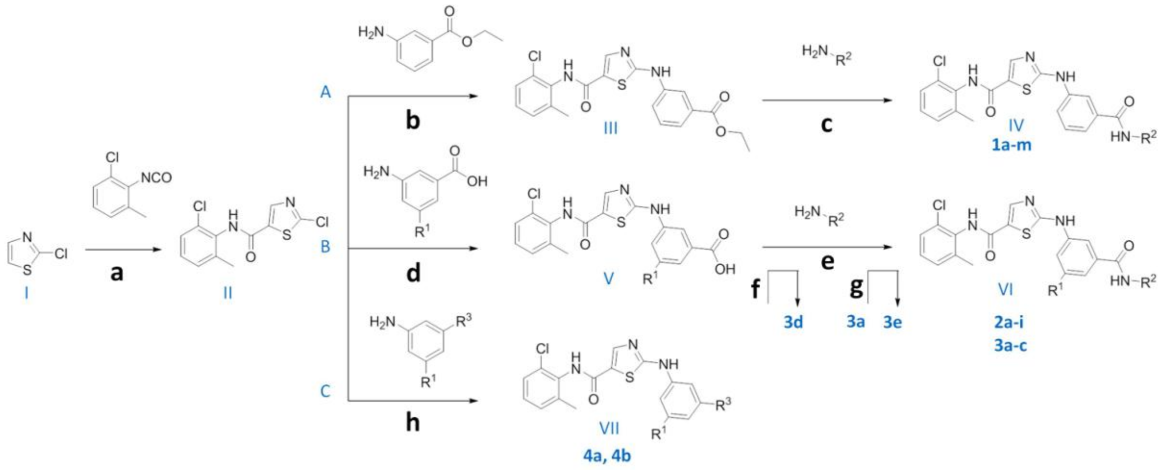


Figure 45 | Synthesis scheme for EPHA2 inhibitors. a) Compound I, n-butyllithium, THF, -78 °C, 15 min, argon; 2-chloro-6-methylphenyl isocyanate, THF, -78 °C, 2h, argon (70-80% yield). **b)** Compound II, ethyl 3-aminobenzoate, (+)-camphor-10-sulfonic acid, 2-propanol, MW, 120 °C, 3h, argon (66-76% yield). **c)** DABAL-Me$_3$, primary amine, THF, 40 °C, 30 min; Compound III, THF, MW, 100-140 °C, 2-6 h (3-91 % yields). **d)** Compound II, (+)-camphor-10-sulfonic acid, 5-substituted 3-aminobenzoic acids, tert-butanol, MW, 120 °C, 3.5 h (28-92% yield). **e)** Compound V, primary amine, DMF, 0 °C; DIEA, TEA, PyBroP, DMF, 30-140 min (4-74% yield). **f)** Compound V, tributyl(vinyl)tin, dioxane, toluene, Pd(PPh$_3$)$_4$, 110 °C, 4h, argon (59% yield). **g)** Inhibitor 3a, 2-aminobenzenboronic acid, Pd(PPh$_3$)$_4$, DMF; K$_2$CO$_3$, MW, 115 °C, 5h, argon (32% yield). **h)** 3,5-substituted aniline, 2-propanol, compound II, HCl, ON, 95 °C (26-43%).

4.2 Synthesis of dedicated EPHA2 inhibitors

To obtain these four series, the common intermediate 2-chloro-*N*-(2-chloro-6-methylphenyl)thiazole-5-carboxamide (**II**) was synthesized following the procedure reported for the synthesis of Dasatinib[337]: 2-chlorothiazole (**I**) was regioselectively deprotonated and metalated at position C$_5$ by n-butyllithium. The intermediate was trapped with 2-chloro-6-methylphenyl isocyanate with 70-80% yield by nucleophilic addition of the metalated thiazole at the isocyanate group. **II** was further derivatized by acid-catalyzed nucleophilic heteroaromatic substitution with either (+)-camphor-10-sulfonic-acid (general methods A and B) or hydrogen chloride (general method C); reaction with ethyl 3-aminobenzoate yielded **III** (Method A; 66-76% yield); reaction with 5-substituted 3-aminobenzoic acids yielded **V** (Method B; 28-92% yield); reaction with 3,5-disubstituted anilines yielded **VII** (Method C; 26-34% yield). **IV** was obtained from **III** by reaction with a primary amine in the presence of DABAL-Me$_3$ [338] in a microwave reactor at 100-140°C [339] (inhibitors **1a-m**; 3-91 % yields). The amides **VI** were obtained from the carboxylic acids **V** by reaction with a primary amine and the phosphonium salt PyBroP. Briefly, a base (DIEA) abstracts the proton of the carboxyl group followed by nucleophilic substitution at the positively charged phosphorous atom of PyBroP. The primary amine adds to the activated carbon which leads to the elimination of the stable phosphoric acid triamide in an addition-elimination reaction to produce **VI** (inhibitors **2a-i**, **3a**, **3b**; 4-74% yields). Compound **3c** was obtained by reacting **V** with 4-aminopiperidine and the guanidinium salt HATU as a coupling reagent. Similar to PyBroP, HATU first forms an activated ester with the carboxylic acid before the nucleophilic amine is added to afford product **3c.** HATU led to the formation of both regioisomers and reduced the yield of the desired product (22% yield). Compound **V** was reacted with tributyl(vinyl)tin in a palladium-catalyzed coupling reaction according to Stille to afford the carboxylic acid precursor of inhibitor **3d**. Briefly, the aryl halide performs an oxidative addition to the catalyst $Pd(PPh_3)_4$ and subsequently, tributyl(vinylt)tin exchanges the bromine in a transmetalation step. Reductive elimination releases the product and the unmodified catalyst. The final inhibitor **3d** was produced by coupling this precursor to 4-aminopiperidine (27% yield). Another inhibitor, **3e**, was synthesized by reacting the bromine containing inhibitor **3a** with 2-aminobenzene boronic acid in a palladium-catalyzed Suzuki coupling reaction. Similar to the Stille reaction mechanism, the aryl halide adds to the palladium catalyst $Pd(PPh_3)_4$ before an intermediate is formed due to reaction with a base (K_2CO_3). The intermediate reacts in a transmetalation step with base-activated 4-aminobenzene boronic acid. Subsequent reductive elimination affords inhibitor **3e** (32% yield). Compound **VII** was afforded in an HCl-catalyzed coupling reaction by linking intermediate **II** to a 3,5-substituted aniline derivative. This aniline derivative was generated from 3,5-halogenated nitrobenzene by nucleophilic aromatic substitution at C_3 and C_5, followed by reduction of the nitro group to the primary amine. Coupling to intermediate **II** afforded the inhibitors **4a** (26% yield) and **4b** (34% yield). According to this synthesis routes, a total of 29 EPHA2 inhibitor candidates was obtained.

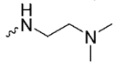
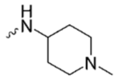
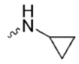
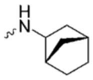
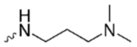
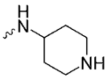
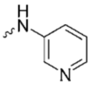
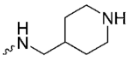
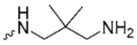
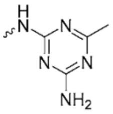
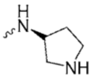
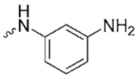
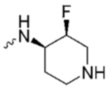
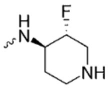
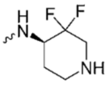
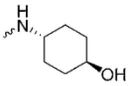
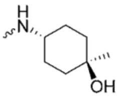
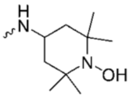
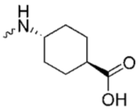
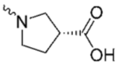
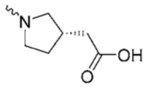
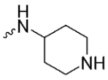
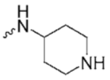
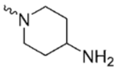
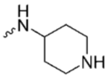
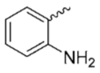
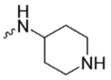
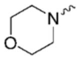
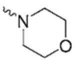
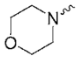
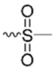
5 Biochemical and cellular evaluation of EPHA2 inhibitors

5.1 EPHA2 inhibitors bind EPHA2 with high affinity

The affinities of all 29 inhibitor candidates were determined by single dose competitive Kinobeads [137,273] pulldowns coupled to either quantitative Western Blotting (10 μM compound concentration) or mass spectrometry read out (3 μM compound concentration) to calculate the relative binding inhibition of EPHA2 compared to a DMSO control (Table 2).

These results reveal that inhibitors targeting charged residues at the ATP pocket entrance by an amine (**1a-m**; 40-95% binding inhibition), a hydroxyl group (**2d-f**; 43-67% binding inhibition) or a carboxyl group (**2g-l**; 40-93% binding inhibition) showed varying binding inhibition of EPHA2. The most promising candidates presenting binding inhibitions of more than 90% contained either a secondary amine (**1g**: 95%, **1l**: 94%) or a carboxylic acid functional group (**2h**: 93%, **2i**: 93%). For further inhibitor development, the piperidine moiety of inhibitor **1g** was favored over the other modifications as it was the only one without chiral center. The modifications introduced at R' (Figure 43c) allowed to investigate the spatial properties of the ribose pocket using groups of increasing steric hindrance (inhibitors **3a-e**, **4a-b**). Inhibitor **3e** comprising an aniline moiety showed high binding inhibition (96%) which indicated that larger hydrophobic groups potentially establishing hydrophobic interactions are tolerated at this position. Inhibitors **4a** and **4b** contain a morpholine or methylsulfonyl group as substituents pointing towards the ribose pocket and showed similar behavior with very high binding inhibition of 98% and 95%, respectively. The inhibitory activity of selected compounds (**2c**, **2f**, **2i**, **4b**) was confirmed by recombinant activity assays, establishing that the Kinobeads binding assay can indeed aid in discovering functional inhibitors of EPHA2 (Table 3).[166]

Table 2 | Synthesis and affinity of EPHA2 inhibitors. Chemical modifications were introduced at R’ and R’’ according to synthesis methods A, B or C. Binding inhibition was determined in single dose Kinobeads pulldowns.

Inhibitor	R′	R″	Method	Yield [%]	Binding Inhib. [%][a]
dasatinib	–	–	–	–	100[b]
1a	H	methyl	A	28	75[c]
1b	H		A	85	68[c]
1c	H		A	16	74[c]
1d	H		A	63	57[c]
1e	H		A	91	< 50[c]
1f	H		A	78	< 50[c]
1g	H		A	44	95[c]
1h	H		A	23	< 50[c]
1i	H		A	52	< 50[c]
1j	H		A	21	83[c]
1k	H		A	3	64[c]
1l	H		A	17	94[c]
1m	H		A	48	59[c]
2a	H		B	63	50[b]
2b	H		B	quant	64[b]
2c	H		B	4	95[b]
2d	H		B	74	56[b]
2e	H		B	54	43[b]
2f	H		B	29	67[b]
2g	H		B	50	40[b]
2h	H		B	55	93[b]
2i	H		B	72	93[b]
3a	Br		B	60	89[c]
3b	CF_3		B	67	88[c]
3c	Br		B	22	53[c]
3d			B	27	97[c]
3e			B	32	96[c]
4a			C	26	98[b]
4b			C	34	95[b]

[a] as determined by quantitative mass spectrometry at 3 µM
[b] as determined by quantitative Western Blot analysis at 10 µM.

Table 3 | Properties of EPHA2 inhibitor candidates. Promising inhibitors were further characterized according to their binding affinity, inhibition of enzymatic activity, inhibition of cell viability, and their selectivity as described by the number of sub-micromolar targets and $CATDS_{EPHA2}$. Crystallography data was obtained for 20 inhibitors in high resolution. Cellular EC50 were determined in three biological replicates each containing three technical replicates.

Inhibitor	**Affinity pK_D^{app} [M]**	**Activity pIC50 [M]**	**Cellular $EC_{50} \pm SD$ [μM]**	**# targets < 1μM**	**Selectivity $CATDS_{EPHA2}$**	**PDB code**	**X-ray resolution [Å]**
Dasatinib	8.23	-	0.097 ± 0.020	44	0.047	5I9Y	1.23
1c	-	-	2.67 ± 0.53	-	-	-	-
1g	7.78	-	4.41 ± 2.69	22	0.074	5NJZ	1.77
1j	-	-	-	-	-	5NK0	1.60
1k	-	-	-	-	-	5NK1	1.55
1l	7.42	-	-	26	0.057	5NK3	1.82
1m	-	-	-	-	-	5NK5	1.35
2a	6.33	-	1.99 ± 0.24	21	0.032	5NK7	1.89
2b	6.74	-	1.54 ± 0.79	23	0.033	-	-
2c	8.70	9.00	1.40 ± 0.13	26	0.141	5NK4	1.45
2d	6.80	-	2.38 ± 0.75	20	0.051	5NK6	1.27
2e	6.14	-	2.32 ± 0.30	16	0.037	5NK9	1.59
2f	6.64	8.52	3.10 ± 0.94	16	0.057	5NK8	1.76
2g	5.46	-	> 5	13	0.026	5NKA	1.38
2h	7.73	-	> 5	17	0.063	-	-
2i	7.58	7.74	> 5	23	0.066	-	-
3a	-	-	-	-	-	5NKE	1.39
3c	-	-	-	-	-	5NKF	1.10
3d	7.68	-	-	28	0.066	5NKG	1.10
3e	7.59	-	-	22	0.056	5NKH	1.29
4a	9.12	-	0.092 ± 0.021	31	0.176	5NKB	1.50
4b	7.73	8.47	0.506 ± 0.034	26	0.050	5NKI	1.68

5.2 Novel EPHA2 inhibitors display improved selectivity

To investigate whether the new inhibitors would be more selective for EPHA2 than the parent lead compound Dasatinib, 16 compounds were subjected to full dose response measurements using Kinobeads with quantitative mass spectrometry readout (Table S7a). As mentioned earlier, Dasatinib bound 44 targets with sub-micromolar affinity (Figure 46a). In contrast, all novel EPHA2 inhibitors had fewer sub-micromolar targets ranging from 13 (inhibitor **2g**) to 31 (inhibitor **4a**). Closer inspection of the target profiles revealed decreased affinities and fewer targets in all kinase groups, but particularly in the TK, TKL, STE and AGC branches. Interestingly, many of the 16 new inhibitors showed substantially reduced binding affinities towards members of the SRC family of kinases (SRC, YES1, FYN, LCK, LYN, HCK, FRK) which are among the main targets of the designated dual ABL/SRC inhibitor Dasatinib.

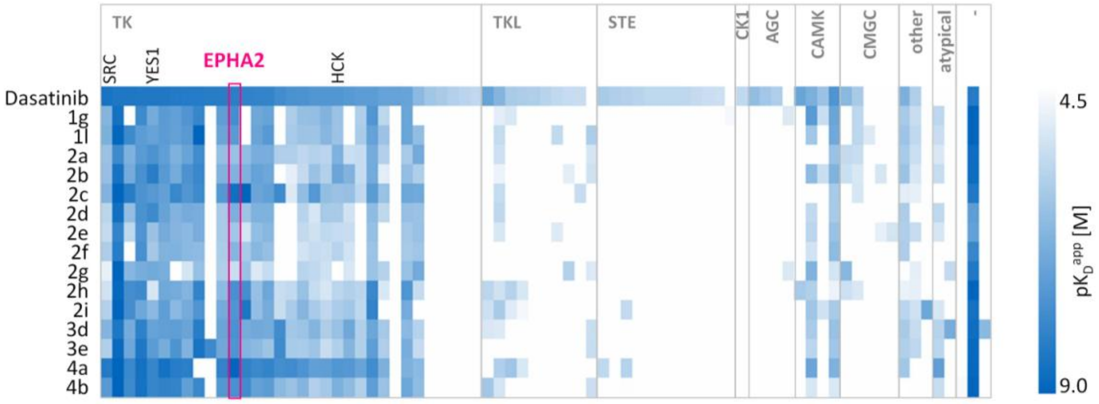


Figure 46 | Target profiles of EPHA2 inhibitors. Heatmap showing selectivity profiles and apparent binding affinities (pK$_D^{app}$ [M]) of Dasatinib and EPHA2 inhibitor candidates. Target kinases are sorted according to kinase phylogeny. EPHA2 is highlighted by a pink box.

Next, the selectivity of the inhibitors was to be quantified. As described above, different metrics have been described previously for this purpose including the selectivity score[191], the gini coefficient[195] and the selectivity entropy[193]. As already discussed earlier, they all suffer from different drawbacks such as dependence on the number of kinases tested in the experiment, negligence of binding affinities or non-protein centric selectivity calculation. Therefore, CATDS$_{EPHA2}$ (Figure 47) was utilized which helped to overcome these limitations and to calculate drug selectivity for a certain target of interest (here: EPHA2) in a concentration-dependent manner (here: at the K$_D^{app}$ of EPHA2).

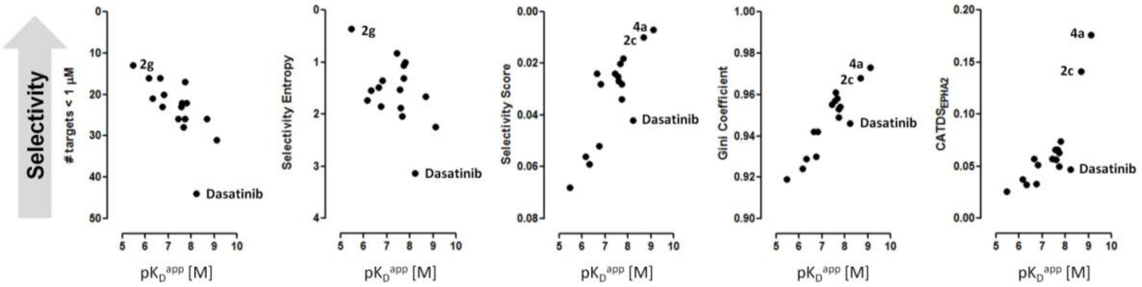


Figure 47 | Comparison of different selectivity metrics used in the literature for calculating compound selectivity. Scores that cannot be calculated with respect to a fixed target protein (target counting, selectivity entropy) indicate inhibitor 2g as the most selective inhibitor. Yet, the potency for EPHA2 is very low which disqualifies this inhibitor as a promising candidate for EPHA2 inhibition. Selectivity score and Gini coefficient require the determination of a threshold concentration and were calculated using the 10x K_D^{app} or K_D^{app} of EPHA2, respectively. However, these values describe the selectivity of a compound at this concentration but not directly in relation to the desired target of interest. CATDS overcomes these limitations and calculates a compound selectivity with respect to a concentration and a certain target. It is therefore especially valuable for medicinal chemistry programs where selectivity of compounds towards their intended target(s) is crucial for inhibitor selection.

This work depicts how CATDS can assist a medicinal chemistry program to evaluate the difference in selectivity of several inhibitors in the light of a particular target protein. As discussed above, *CATDS* adopts values between 0 (unselective) and 1 (selective) and can be interpreted as the percentage of an applied compound that binds to a particular target at a particular concentration. Here, simple target counting without consideration of the affinities towards EPHA2 or the other targets would identify inhibitor **2g** (13 targets below 1 μM) as most selective inhibitor (Figure 48a). This interpretation would, however, be misleading because the compound is also the weakest of all the inhibitors (K_D^{app}(EPHA2): 3.4 μM, Figure 48b). The same effect is visible with selectivity determination using the selectivity entropy (Figure 47). In contrast, the CATDS score highlighted inhibitor **4a** ($CATDS_{EPHA2}$: 0.176, K_D^{app}(EPHA2): 0.8 nM) and **2c** ($CATDS_{EPHA2}$: 0.141, K_D^{app}(EPHA2): 2 nM) to be the most selective EPHA2 inhibitors in the panel (Figure 48c, Table 3). Selectivity score and Gini were also capable of identifying **4a** as most selective compound but the selectivity distribution across all inhibitors was not as unambiguous as by using $CATDS_{EPHA2}$ (Figure 47). To summarize, even though inhibitor **2g** only bound 13 proteins and inhibitor **4a** targeted 31 proteins at a concentration set to 1 μM, the latter is still the more selective compound as it engages fewer drug-protein interactions at the K_D^{app} of EPHA2 (Figure 48d, Figure S3 for radar plots of all inhibitors). This analysis illustrates that it is important to consider the dose dependence of compound selectivity because it can guide the choice of dose in a biological system.

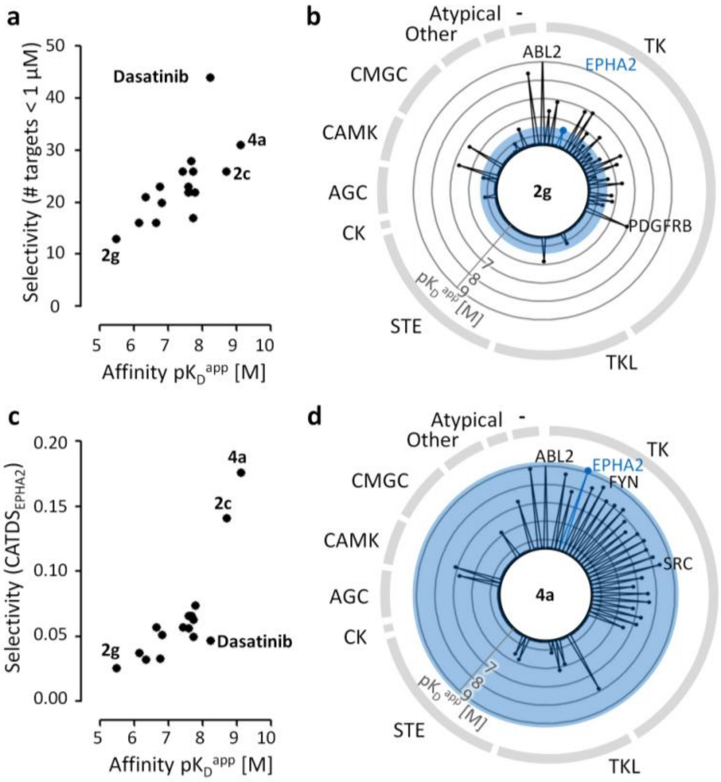


Figure 48 | Selectivity of EPHA2 inhibitors. a) Selectivity of EPHA2 inhibitors as obtained by simple target counting qualifies inhibitor 2g as most selective inhibitor with lowest amount of targets. **b)** Radar plot depicting the target profile and binding affinities inhibitor 2g which showed the lowest number of targets but also lowest affinity for EPHA2. Targets that are inhibited together with EPHA2 are outside the blue circle. **c)** Selectivity as determined by $\mathrm{CATDS}_{\mathrm{EPHA2}}$. Inhibitor 4a was found to be the most potent and most selective EPHA2 inhibitor in our panel. **d)** Radar plot depicting the target space and binding affinities of inhibitor 4a that has the highest affinity and selectivity for EPHA2. Targets that are inhibited along with EPHA2 are outside the blue circle.

5.3 Protein crystallography reveals epitopes for inhibitor binding

Co-crystallization experiments were performed for 17 dedicated EPHA2 inhibitors and using recombinant EPHA2 kinase domain (Asp596-Gly900) expressed in Sf9 cells (Table 3, Table S7b for crystal structure data collection and refinement statistics).[296] These experiments were performed by members of the research group for Biomolecular Magnetic Resonance (headed by Prof. Harald Schwalbe) at the Institute for Organic Chemistry and Chemical Biology at Johann Wolfgang Goethe-Universitaet in Frankfurt. As anticipated, all inhibitors bound to the ATP binding site in the typical type 1 (DFG-in) conformation, and the *N*-(2-chloro-6-methylphenyl)-2-(methylamine)thiazole-5-carboxamide part of the molecule - common to all inhibitors, including Dasatinib - was equally positioned in all inhibitor-EPHA2 co-crystals (Figure 49a). The direct interactions to residues Met667, the gatekeeper Thr692 and Met695 (hinge region), as previously reported for Dasatinib, were also maintained. More interestingly, an overlay of all obtained structures showed that the inhibitors engaged the pocket entrance as well as the ribose pocket, as originally intended (Figure 49b). Residues R' and R'' do not adopt similar conformations in the different crystal structures but distribute across the pockets. As observed e.g. for inhibitor **2c**, the aryl moiety could adopt different rotational poses such that the R' moieties targeted either of the two pockets.

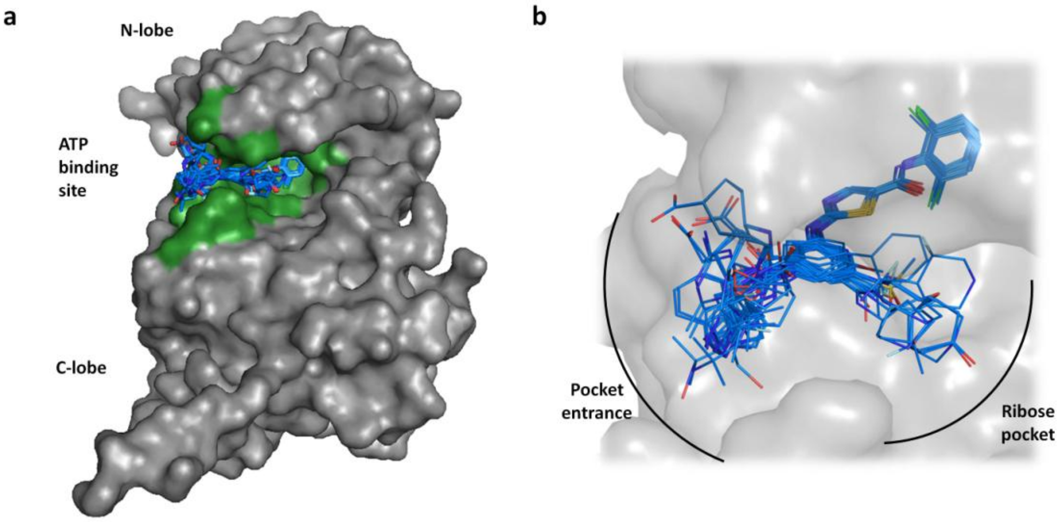


Figure 49 | Inhibitor-EPHA2 co-crystal structures validate medicinal chemistry concept. a) All 17 EPHA2 inhibitor candidates bound in the type 1 conformation (DFG-in) in the ATP binding site located between the N lobe and the C lobe. **b)** The inhibitors engaged both the ATP pocket entrance and the ribose pocket.

Inhibitors **1a-m**, **2a-i** and **3a-e** contain an amide linkage and this additional moiety, as compared to Dasatinib, was engaged in binding interactions with EPHA2 in the majority of inhibitors (Figure 50a). This interaction was mostly established with Glu696 located in the hinge region, as intended by the hybrid inhibitor design (11 molecules out of 15). In addition to Glu696, the amide bond also engaged Tyr694 via direct side chain interaction (e.g. inhibitors **2f-g**). Unexpectedly, it was found that the amide bond orientation was very flexible and could also form backbone interactions with Ile619 and Ala699 within the ribose pocket (e.g. inhibitors **2e**, **1l**). Targeting what was expected to be "selectivity residues" also proved successful, since some amine, hydroxyl or carboxyl functional groups, which were positioned in different spatial orientations, were able to extend the molecular interactions to the charged residues Lys617, Glu706 and Lys702 located at the entrance of the ATP pocket (Figure 50b). Inhibitors carrying an amine often engaged Glu706 and Lys702 in a water-mediated interaction (e.g. inhibitors **1g**, **2a**, **3e**, **3b**, **3d**). Lys617 located in the N-lobe was targeted by a primary amine (inhibitor **1j**). The carboxyl containing inhibitor **2g** looped back towards the ATP pocket to interact with Tyr694. Molecules **2d** and **2e** comprising a hydroxyl group unexpectedly engaged the ribose pocket which positioned the hydroxyl group in an orientation distant to the target residues, whereas the hydroxyl group of molecule **2f** directly interacted with Glu706 with a distance of 2.5 Å.

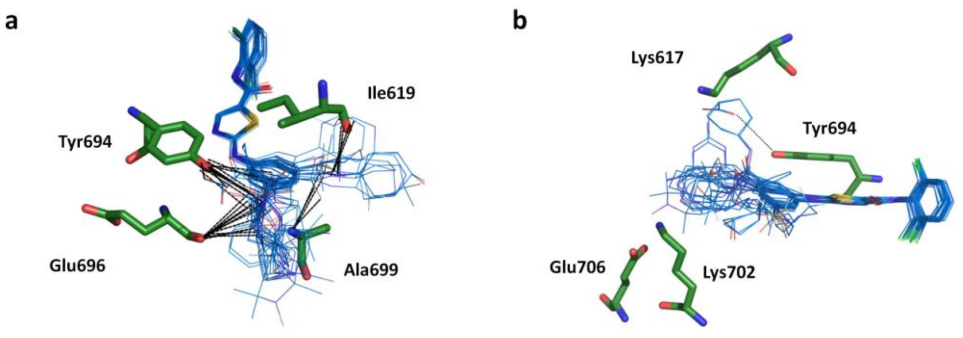


Figure 50 | Direct interactions to ATP pocket entrance. a) In most inhibitors, the introduced amide bond formed direct side chain interactions with Tyr694 and Glu696 located towards the ATP pocket entrance. Some inhibitors engaged a distinct conformation involving backbone interactions to Ile619 and Ala699 in the ribose pocket was observed. **b)** As intended, selectivity residues Lys617, Lys702, Glu706 were often engaged in direct or water-mediated interactions by inhibitors carrying an amine, hydroxyl or carboxyl group. Surprisingly, inhibitors 2g and 2h carrying a carboxylic moiety adapted an unexpected bent conformation and interacted with Tyr694 (e.g. inhibitor 2h, indicated by sticks).

As expected, direct interactions within the ribose pocket were scarce and inhibitor-EPHA2 interaction in this binding site was mainly driven by hydrophobic contacts and advantageous space-filling with sterically demanding groups (e.g. inhibitors **3e**, **4a**). Interestingly, inhibitor **3b** features a precise symmetric orientation of its trifluoro group towards the carbonyl oxygen of the Ile619 backbone. Due to the strong electronegativity of oxygen, CF--O interactions are typically known to be of rather repulsive than attractive nature.[340] However, the carbonyl oxygen directly points towards the positively polarized C atom of the CF_3 unit. It is therefore tempting to speculate that this might positively affect inhibitor binding. Inhibitor **4a** was found to be the most affine (K_D^{app}: 0.8 nM) and most selective inhibitor for EPHA2 ($CATDS_{EPHA2}$: 0.176, i.e. 17.6% of **4a** bound to EPHA2, which is 4 times more than Dasatinib). Somewhat surprisingly, it did not engage in additional direct interactions but occupied both the ATP entrance as well as the ribose pocket. Occupying both pockets with the same residue might reduce degrees of freedom and therefore favor inhibitor binding due to entropic contribution.

5.4 Cellular activity of EPHA2 inhibitors in SF-268 cells

The anti-proliferative potential of the EPHA2 inhibitor candidates were examined in a human cancer cell line that is growth-dependent on EPHA2. Kinome-wide protein expression profiling data of the NCI60 panel of cancer cell lines previously obtained using the Kinobeads technology[294] was utilized to identify a cellular system that would be specifically suitable for testing our inhibitor panel (Figure 51). A suitable cell line was defined as featuring high EPHA2 expression and low abundance of other kinases targeted by the EPHA2 inhibitors. As mass-spectrometry intensities are biased by the affinity of the Kinobeads for the different kinases, the intensity change of the proteins across the NCI60 panel were analyzed in relation to the mean protein intensity level (z-score transformation). The resulting value depicts the number of standard deviations by which the protein intensity in a particular cell line differs from the mean intensity across the panel and indicates relative high (positive values) or low (negative values) abundance of a protein. In the literature, the prostate cancer cell line PC-3 is often used as

EPHA2 dependent cellular system. However, for testing the EPHA2 inhibitors this cell line was found to be non-advantageous as it revealed high expression levels of several other prominent targets of these inhibitors, such as ABL1, SRC, YES1 and other EPH inhibitors (e.g. EPHB4, EPHB2, EPHA5). The glioblastoma cell line SF-268 showed the highest EPHA2 expression level among the NCI60 panel and also featured an advantageous overall kinome profile.

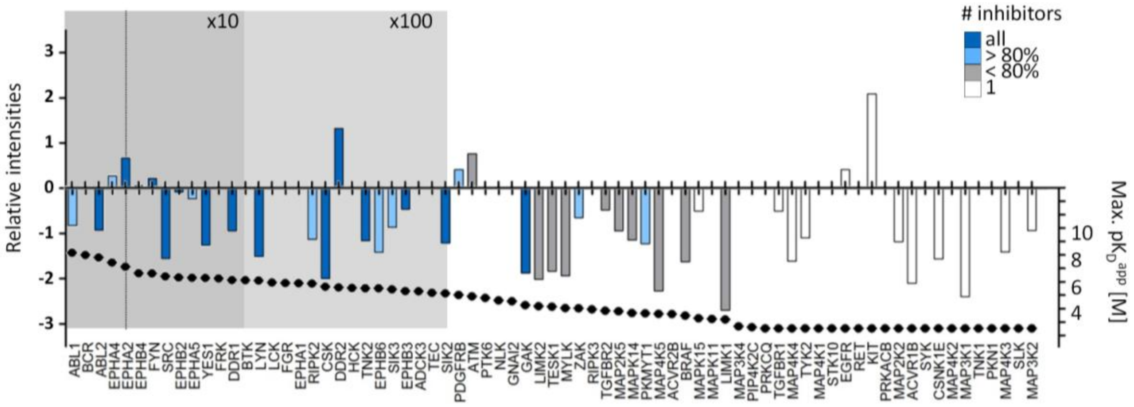


Figure 51 | SF-268 glioblastoma cells are suitable for testing the *in vivo* effect of EPHA2 inhibitors. Kinome profiling of the NCI60 panel of cancer cell lines revealed that SF-268 cells feature high EPHA2 expression and low abundance of other targets (barchart) frequently hit by our EPHA2 inhibitors (bar color). Inhibitor targets were sorted according to the maximum observed affinity (pK$_d^{app}$ [M]) within our dataset and theoretical therapeutic windows were set at 10x and 100x K$_d^{app}$ of EPHA2.

All 77 proteins that were identified as targets of all EPHA2 inhibitors generated in this study as well as Dasatinib were combined and sorted according to the maximum observed affinity, regardless of which inhibitor showed this affinity (pK$_D^{app}$, Figure 51, black dots). Target proteins that bound with high affinity were targeted by a large proportion (> 80 %) of the dedicated EPHA2 inhibitors (Figure 51, bar color). The distribution of relative intensities (Figure 51, bar chart) revealed that almost all potently inhibited targets other than EPHA2 – as for example the main targets of Dasatinib ABL and SRC - are relatively weakly expressed in this cell line compared to all other cell lines of the NCI60 panel. Within a therapeutic window of 10 x K$_D^{app}$ of EPHA2 (indicated by the dark grey box), only EPHA4 and FYN show slightly higher expression and may thus be candidates for relevant off-targets in this cell line. Within a therapeutic window of 100 x K$_D^{app}$ of EPHA2 (indicated by the light grey box) only DDR2 was another such candidate. Western blot analysis of SF-268 cells showed that this cell line expresses phosphorylated EPHA2 (at S897) and AKT1 (at S473) which is indicative for aberrant and tumor-promoting ligand-independent EPHA2 signaling (Figure 52a).[48] Next, protein knockdown experiments by siRNA were performed to demonstrate a functional dependence of SF-268 cell viability on EPHA2 expression (Figure 52b). Four independent siRNAs were examined separately and as a mixture and revealed a reduction of cell viability of about 60% after 6 days of treatment. These results showed that SF-268 cells are suitable for testing EPHA2-directed inhibitors and that cell viability reduction is an appropriate readout for assessing the response of the cell to EPHA2 inhibition.

5.5 EPHA2 inhibitor 4a potently reduces SF-268 cell viability

Having established the glioblastoma cell line SF-268 as a suitable biological model, we measured the anti-proliferative effect of the inhibitor candidates synthesized in this study. Cellular activity was examined utilizing a resazurin-based cell viability assay where reduced metabolic activity of cells leads to decreased reduction of resazurin to resorufin. Inhibitor treatment was performed in full dose response experiments with eight inhibitor concentrations ranging from 3 nM to 30 μM. Most of the compounds inhibited cell viability with low micromolar potency (**2c**: 1.4 μM – **1g**: 4.4 μM,Table 3), whereas inhibitors **2g-i** did not show any anti-proliferative effect. We attribute the latter to the presence of a charged carboxyl moiety that likely impairs cellular uptake. Future work on these compounds could therefore include derivatization into ester prodrugs with enhanced permeability. Inhibitor **1g** bearing a piperidine at position R' affected SF-268 cell viability with a very low potency of 4.4 μM. This result can be rationalized by comparison to meta-fluorinated versions of this compound (inhibitors **2a-2c**). Cellular potency increased 2- to 3-fold for the singly fluorinated (EC$_{50}$(**2a**): 1.99 μM, EC$_{50}$(**2b**): 1.54 μM) and doubly fluorinated inhibitor **2c** (EC$_{50}$: 1.4 μM). This indicates that the positive charge of the piperidine at physiological pH (pK$_a$ of around 11) could be responsible for the diminished cellular effect of inhibitor **1g**. Indeed, fluorination in meta-position to the secondary amine should reduce the pK$_a$ of piperidine from 11 to 9 (singly fluorinated) and 7 (doubly fluorinated) and the obtained molecules showed more activity in the viability assay. Interestingly, inhibitors **4a** and **4b** were much more potent (**4a**: 92 nM, **4b**: 506 nM) and similar to the potency of the lead compound Dasatinib (97 nM) (Figure 52c). It cannot be entirely ruled out that this potency is not only due to the inhibition of EPHA2, but this excellent activity designated **4a** as the most promising EPHA2 inhibitor in the current compound set with a strong demonstrated anti-proliferative effect in a relevant EPHA2 overexpressing cell-line.

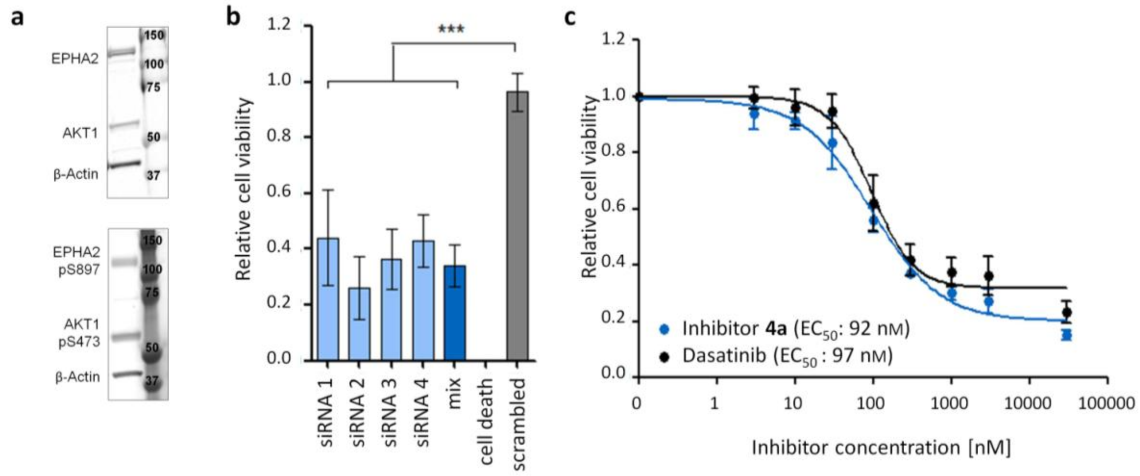


Figure 52 | Inhibitor 4a shows anti-proliferative effects in the SF-268 glioblastoma cell line. a) Western Blot analysis revealed the presence of tumor-promoting ligand-independent signaling which is characterized by phosphorylation of EPHA2 S897 and AKT1 S473. This signaling acts in a negative feedback loop, where EPHA2 phosphorylates AKT1, which in turns phosphorylates EPHA2 S897. This signaling promotes migrative cell behavior and may eventually also lead to metastasis. **b)** Protein knockdown of EPHA2 by four independent siRNAs confirmed that SF-268 cell viability is dependent on EPHA2 expression. Experiments were performed in three biological replicates each containing three technical replicates. Significance was calculated by a two-sided unpaired student's t-test. **c)** SF-268 cells were treated with increasing concentrations of inhibitor 4a resulting in potent reduction of cell viability similar to Dasatinib.

To summarize, novel dedicated EPHA2 inhibitors were successfully designed and synthesized. The residue classification scheme assisted in constructing inhibitors specifically targeting amino acids that could convey selectivity towards EPHA2. Protein crystallography elucidated the molecular details of compound:protein binding and allowed structure-affinity/selectivity-relationship analysis. Lead optimization was governed by simultaneous Kinobeads selectivity profiling and selectivity determination using the novel selectivity metric CATDS. This combinatorial approach illustrates how chemical proteomics can guide a medicinal chemistry project in order to generate more selective, yet potent small molecule kinase inhibitors.

General Discussion and Outlook

Table of contents

1 What is a good drug? 115
2 What is a good way to discover new drugs? 118
3 What makes a good drug target? 121
4 Yin and Yang 122

“**Me explicaste de la quietud nace la inspiración y del movimiento surge la creatividad. ”**

- Isabel Allende -

“Las cosas tienen vida propia. Todo es cuestión de despertarle el ánima.”

- Gabriel García Márquez -

1 What is a good drug?

(Mis)belief in kinase inhibitors. A change of perspective is required when using small molecule kinase inhibitors. Scientists and clinicians need to accept and be fully aware of the fact that most kinase inhibitors are inhibiting more than only one target or target family.[341,342] The progress from a single-protein view to a system-wide evaluation could pave the way for better understanding of drug perturbation and an informed targeted application also of promiscuous inhibitors. To reach this aim, selectivity profiling must be performed as comprehensively as possible and profiling data need to be easy accessible. It is also crucial, that scientists and clinicians understand their biological system and a drug's mode of action as a complex interplay of many inhibitory effects rather than simplifying it to the inhibition of a single kinase. This perspective is of special interest for EPHA2 and the EPH family in general. Since EPH signaling is believed to function via the concerted action of many different EPH receptors and several downstream signaling pathways,[9,10] the abrogation of a particular branch of signaling events by a multi-targeted inhibitor could be even more beneficial than targeting a particular EPH receptor alone.

Selectivity vs. Polypharmacology. Facing the reality of multi-component phenotypes and the hurdles to create selective compounds, the question arises whether we really need single-targeted compounds?[341] The merit of highly selective compounds is under debate in the scientific community and needs to be discussed from several angles. Looking at clinical inhibitors that are already used in human for targeted therapy, polypharmacology does not seem to be a major issue.[342] Many approved kinase drugs (e.g. Dasatinib, Midostaurin, Brigatinib) have a very broad target profile inhibiting numerous kinases across the whole kinome. In a clinical background, off-target inhibition could even be considered beneficial if proteins are affected that could potentially overcome the inhibition of a particular kinase (for instance other family members) or that could lead to therapy resistance (as known for MET or EPHA2). Multi-targeted drugs could also provide the opportunity of several modes of action that function in a concerted manner in defeating the disease.[341-343] As an example, the inhibition of pro-inflammatory immune response could be beneficial during cancer therapy. However, if off-target inhibition mediates toxic or adverse side effects, it could be detrimental for therapy outcome and will most probably lead to failure during clinical evaluation.[187]

Yet, modern clinical research is more and more evolving the idea of precision medicine.[344,345] The basis for this approach is the understanding that a tumor is rather characterized by its molecular composition than by the entity it emerged from. In precision medicine, the disease-driving molecular background of each patient is characterized to enable molecular tumor boards to suggest appropriate treatment options. This approach could profit from highly selective compounds that specifically target the pathogenic condition each patient suffers from.[146] If the goal of selective drugs for each kinase or kinase family is ever reached, combination treatments of several such drugs could be optimized for each patient and allow for fine-tuned individualized therapy.[342] However, it is still a long way until precision medicine may evolve to be the standard of care; but if we succeed in providing adequate diagnostic tools and therapies this might be the future of clinical healthcare.

From a scientific point of view, highly selective compounds provide the great opportunity to be used as chemical probes in target validation studies.[146,188] In this regard, they offer a complementary approach to genetic methods such as RNAi or CRISPR/Cas.[342] Pharmaceutical inhibition has the advantage of being reversible, temporally controllable and interfering less with the overall cellular protein expression than the typically used genetic approaches. The choice of an appropriate chemical probe is difficult as it requires a thorough understanding of the molecule, its selectivity and it's mode of action.[188] Unfortunately, many compounds are used as chemical probes in scientific literature which are not suitable *per se* or not applied in an appropriate concentration which could lead to false hypotheses. The chemical probe portal initiative (www.chemicalprobes.org) provides the research community with a selection of chemical probes for the interrogation of particular proteins and protein classes and defined criteria a compound must fulfill to be suitable for usage as chemical probe: potency, selectivity, evidence for target engagement and an understanding of the mechanism by which the probe modulates the target. Still, many of these probes have multiple targets and appropriate application needs to be carefully evaluated for each research question.[346] Highly selective molecules are valuable but the development of truly selective inhibitors for each and every kinase is tedious. Considering the efforts that have been undertaken by experienced scientists in the past decades,[347,348] the question remains whether it is even feasible to reach this goal.

Measuring drug selectivity. Calculating drug selectivity is not an obvious task and requires a precisely asked research question. The requirements for selectivity calculation may vary greatly depending on whether a researcher is interested in a specific target, in a specific concentration or global selectivity regardless of target or concentration.[180] The conventional selectivity metrics (such as the selectivity entropy,[193,194] the selectivity score,[139,191] the gini coefficient,[195] and the partition index[196]) are applicable to rather defined research questions but not flexible enough to elucidate compound selectivity from different angles. The selectivity entropy provides a global view on drug selectivity which is completely unrelated to a particular target or concentration. The selectivity score and the gini coefficient account for certain concentrations but cannot be calculated with regards to a particular target of interest. On the contrary, the partition index is capable of determining the selectivity towards a certain target but cannot resolve the concentration-dependent angle of selectivity. But, a target- and concentration-dependent view on drug selectivity is crucial to understand how selectivity influences a drug's mode of action at different concentrations. CATDS was developed to overcome these limitations and proved to be a versatile and practical, yet simple to calculate, scoring system for the determination of compound selectivity. It can use various input data formats (activity or binding, single dose or full dose response). CATDS is also flexible in terms of choice of which set of proteins or drug dose is used in the calculation. Therefore, it can be applied for asking very different questions regarding compound selectivity. Kinobeads data appears to be particularly amenable to CATDS. The experimental setup uses native cell lysate which comprises a very complex mixture of endogenous proteins at vastly varying concentrations. Thus, Kinobeads data approximates the dose-dependent binding behavior of a drug within a native cellular environment and directly reflects the underlying thermodynamic equilibrium. The combination of selectivity information derived from CATDS with the measured percentage of binding reduction at each concentration yielded the *Concentration and Target Dependent Selective Inhibition* (CATDSI), which can be understood as an approximation of the balance between selectivity and effect. In the future, CATDSI could inform the rational design of single drug and especially combination treatments in many ways. The selectivity profiling of clinical kinase inhibitors builds the basis of rational selection of combination treatments.[342] Given a set of proteins of interest to be targeted by combination treatment, CATDSI can help to identify appropriate inhibitors for this task. Additionally, it facilitates the selection of appropriate drug concentrations where the desired inhibitory effect and selectivity on the selected targets are most balanced. Combination of several drugs should lead to an (at least) additive inhibitory effect on the respective drug target. Thus, combination treatment could also be useful when a particular protein should be targeted but no chemical probe is available. If several inhibitors with opposing target profiles were used at low dosage, this additive effect on the target of interest could provide highly selective inhibition despite using non-selective inhibitors. In such a scenario, CATDSI could help to select appropriate concentrations of each inhibitor to gain most selectivity for the target of interest.

Affinity vs. residence time. CATDS is based on thermodynamic principles which is advantageous for using it for Kinobeads data, but also holds few limitations. Affinity, as determined as K_d^{app} in a thermodynamic equilibrium, is the result of enthalpic and entropic changes upon drug:protein interaction. However, affinity is not capable of describing the underlying kinetics driving this interaction.[282,283,349] Kinetic constants, mainly the k_{off} rates, define the residence time of a drug within its target protein.[281] This greatly influences selectivity, as prolonged residence times will have a more pronounced inhibitory effect than observed for transient drug:protein interactions.[349] Importantly, the observed affinity can be the same for both, prolonged and transient interactions – a fact which blurs and hampers the selectivity calculation. In order to cope with this problem, selectivity calculation using CATDS could be expanded by residence time information as third dimension, next to target and concentration dependence. How to integrate this information in a meaningful manner requires careful evaluation, but it might be possible to i) use residence time cutoffs for the selection of drug:target interactions to be included in CATDS calculation (residence time as third input parameter), or to ii) multiply the influence of different drug:protein interactions on the selectivity calculation by weighting factors based on residence time. The integration of residence time would also be beneficial for the calculation of CATDSI. Adding this third dimension would enable washout experiments where the selectivity of a drug should change over time according to its residence time. This could greatly impact and fine-tune our way of performing drug treatments on cells and might facilitate data analysis. Unfortunately, there is no study available yet where residence time of small molecule kinase inhibitors was determined in large-scale and system-wide manner.

Deciphering binding kinetics. Chemical proteomics could be used to overcome this limitation in the future by integrating kinetic measurements into the standardized Kinobeads workflow. A conceivable experimental setup could use the drug-treated lysates that remain after Kinobeads pulldowns and contain the target proteins and the respective compound. After establishing a thermodynamic equilibrium between target proteins and drugs, a covalent lysine-reactive probe could be added to the lysate. This chemical probe could be based for instance on a reactive sulfonyl fluoride moiety (as shown by Zhao *et al*[350]) or an ATP-derivative (as used by the KiNativ technology[276,277]). Covalent addition to the highly conserved lysine in the ATP pocket can only take place if the target protein is not occupied by the drug. The apparent reaction rate of this covalent addition should correlate to the dissociation rate of the drug:protein interaction. However, this is only true if the experimental setting is designed in such a way that the rate-determining factor is not the reaction rate of the covalent addition but the residence time of the reversible drug:protein interaction. Samples should be collected in a time-dependent fashion and covalently bound target proteins could be enriched via an enrichment handle such as alkynes (click chemistry) or biotin (streptavidin-biotin interaction) and further processed for LC-MSMS analysis. A targeted proteomics readout measuring few proteotypic peptides per captured kinases could be useful in this setting to gain sensitivity on the mass spectrometric readout. The kinetic information that could be potentially gained by such an experiment would not only fine-tune CATDS selectivity calculation but could also greatly enhance our understanding of drug selectivity, relationship of kinetics and molecular structure, and further inform medicinal chemistry efforts towards more selective and more effective compounds.

2 What is a good way to discover new drugs?

Traditional Drug Discovery. Pharma industry most often use high-throughput screening (HTS) of large compound libraries, which is cost-intensive and often not accessible for academic research.[351,352] The identified lead structures are then further optimized towards higher target affinity and improved pharmacokinetic properties (ADMET: absorption, distribution, metabolism, excretion, toxicology). This hit-to-lead optimization is often performed by traditional structure-based design which relies on co-crystal structures of the target protein and the designed inhibitor candidates.[353] Medicinal chemists develop their molecular scaffold towards particular pockets or amino acids identified in the crystal structure, or try to retrospectively explain the activity of their best candidates based on structure-activity-relationship (SAR) analysis. Mostly, this strategy is employed to gain molecular interactions and thus, affinity towards the desired target. Yet, the impact of the desired drug:protein interaction on the selectivity of the novel ligands is often not considered. Usually, the selectivity of inhibitor candidates is only tested for very mature compounds and is not part of the lead optimization process. Many drugs fail during clinical evaluation, mainly due to insufficient efficacy or toxic side effects – both of which could be prevented by more thorough preclinical evaluation also including early selectivity profiling of inhibitor candidates.

Chemoproteomics-aided drug discovery. Selectivity profiling by chemical proteomics used as guidance for the discovery of more selective inhibitors is a novel concept bearing several advantages as compared to the classical drug discovery approach. This work presents how chemical proteomics informed multiple steps of drug discovery including lead discovery, inhibitor design and hit-to-lead optimization, and phenotypic model selection. Lead discovery was facilitated by a Kinobeads screen of 235 clinical kinase inhibitors. The examined compound set included mature preclinically developed kinase inhibitors, thus, excluding typical drawbacks observed for leads derived from HTS, as for instance the detection of PAINS (pan-assay interference compounds[354]) or structures comprising unfavorable ADMET properties. Here, Dasatinib was selected as lead scaffold mainly due to its high affinity, but the wealth of selectivity data generated in the Kinobeads drug screen would also allow for choosing another appropriate lead structure comprising a favorable selectivity profile. Inhibitor design and lead optimization were guided by selectivity profiling, SAR analysis and cell viability experiments. Promising candidates were developed further if they showed enhanced selectivity towards EPHA2; on a later stage, inhibitor design focused on enhanced cellular permeability. Early selectivity profiling at the phase of lead optimization facilitates rationalized decisions towards more selective compounds and also allows for the early exclusion of inhibitor candidates that target unfavorable off-targets (e.g. FECH,[274] ZAK[320]). Additionally, information gained from selectivity profiling helped to select SF-268 cells as a phenotypic model according to the requirements of the presented inhibitor set. These glioblastoma cells comprise a kinase profile which is favorable for the evaluation of the inhibitor candidates – EPHA2 is highly overexpressed whereas other prominent targets of these inhibitors are relatively low abundant, which increases the possibility that the observed cellular effect might be more dependent on inhibition of EPHA2 rather than on other off-targets. This concept can and should be applied also for other research projects. When studying a particular inhibitor, its target profile should be examined comprehensively to choose an appropriate phenotypic model according to its expression profile. When studying a certain biological model, thorough characterization of the cellular model should be the basis for the selection of an appropriate inhibitor according to its target profile.

Residue classification. A combined approach of selectivity profiling, protein crystallography and kinase sequence alignment was used to categorize “druggable” amino acids according to their potential impact on affinity and selectivity of novel ligands. Interactions with key residues were proven to modulate selectivity of a drug towards a certain target class, interactions with potency residues were expected to increase affinity only, whereas engagement of selectivity residues could enhance both affinity and selectivity. Scaffold residues may build up hydrophobic pockets which could harbor hydrophobic moieties but do not allow for direct interactions. This residue classification should help medicinal chemists to evaluate and prioritize potential drug:protein interactions in their ligand design. It was developed on the example of the receptor tyrosine kinase EPHA2, but is also applicable for other kinases as shown for the non-receptor tyrosine kinase ABL1 and the serine/threonine kinase MELK. Residue classification requires co-crystal structures as starting point and thus, profits from a combination of diverse crystal structures comprising distinct conformations and different molecular entities. Higher diversity enhances the chance of finding more molecular interactions and broadens the chemical space that could inform future ligand design. In fact, more than 5000 human kinase structures with more than 3000 ligands are publically available in the structural database PDB covering a broad range of chemical space. Residue classification could greatly benefit from expanding the drug:protein interaction analysis to all available structures as it might render the generation of more and diverse co-crystal structures less important. It might be possible to determine selectivity and key residues on the basis of the selectivity profiles without co-crystallizing these compounds with the target of interest. Also, the residue classification approach is not necessarily limited to kinase drug discovery but could be suitable for small molecule drug discovery in general.

Residue classification was found to be a useful approach for rational inhibitor design, but cannot be considered as a stand-alone approach. Co-crystal structures deliver useful insights into drug:protein binding on the molecular level – an information which is not available in such extent by any other technology. However, protein crystals do not fully represent the native system;[355] especially for highly dynamic proteins such as kinases. Under native conditions, the kinase domain switches dynamically between several functional conformations and many more conformational sub-states.[356] Hence, freezing one conformation in a protein crystal necessarily discards many other conformations that might be informative for ligand design.[355] Additionally, it needs to be considered that residues, especially those at the protein surface and ATP pocket entrance, might not be sufficiently accessible for the compound due to solvent interactions that might cover the targetable side chain moieties. With regards to EPHA2, this problem mainly affects selectivity residues located at the pocket entry which might explain why targeting those residues did not work as efficiently as anticipated. Another limiting aspect of this approach is its restriction to molecular interactions. The formation of hydrogen bonds, salt bridges, water-mediated interactions or van-der-Waals forces increases the enthalpic contribution to ligand binding. However, the stability (and thus affinity) of a ligand:protein complex is described by the Gibb’s free energy that is released upon ligand binding and which is the result of enthalpic and entropic contributions.[287] Rational inhibitor design can also favor entropic contributions on ligand binding by generating more rigid or symmetric ligands – an information that cannot be delivered by the presented residue classification approach. Regarding the EPHA2 inhibitors presented in this thesis, the potential low-entropy ligand comprising a symmetrical dimorpholino moiety (inhibitor 4a) was most successful. Isothermal titration calorimetry could prove whether the positive results of this inhibitor are indeed due to favorable entropic properties.

Novel EPHA2 inhibitor. Inhibitor 4a was the most promising inhibitor candidate in this study: it was potent against EPHA2, showed improved selectivity compared to the lead structure Dasatinib, and was found to have an anti-proliferative effect on SF-268 glioblastoma cells.

Nonetheless, it still potently hits many off-targets such as members of the SRC or EPH family – rendering this compound less favorable as a probe with regards to its overall target profile. In spite of that, CATDS still identified this compound as being the most selective EPHA2 inhibitor which is mainly driven by its very high affinity towards its designated target. Utilizing this inhibitor as chemical probe for selective inhibition of EPHA2 can only be fulfilled at low inhibitor concentrations and requires careful evaluation of how other targets might also be influenced in the experimental background. This also further exemplifies how drug concentration influences inhibitor selectivity. Inhibitor 4a showed a pronounced anti-proliferative effect on SF-268 cells at a potency of around 100 nM; selectivity of 4a at 100 nM is rather low but given the fact that the molecule needs to cross the membrane to perform its action, the “apparent” intracellular concentration is probably not as high as the applied compound concentration.

The development of inhibitor 4a is at an early stage and much more information needs to be gathered to fully characterize the properties of this compound. The next step would include the evaluation of target engagement within a cell and the attribution of a phenotypic effect caused by the inhibition of this target.[317] This evaluation can be tedious and complicated through the facts that most inhibitors have multiple targets and that phosphorylation sites (which are often used for proving kinase engagement) are often promiscuously phosphorylated by several kinases and not as specific as anticipated.[313] However, this step is crucial during drug development and may prevent promising compounds to fail in the clinic due to insufficient efficacy. Optimally, selectivity profiling should also be complemented by other preclinical approaches such as testing toxicity on non-cancer and primary cells, optimizing ADMET properties, and phenotypic screening in several relevant cellular models and disease backgrounds. In case of EPHA2, it would be of interest to evaluate inhibitors not only in an oncogenic background, but also on its influence in, for instance, pathogen infection or neurological disorders.

Forward pharmacology. The workflow presented in this thesis follows a *reverse pharmacology* approach, where a protein of interest is selected, lead structures for targeting this protein are identified in a screen and further optimized during drug development. Another even more historical approach, *forward pharmacology*, relies on phenotypic screening where compounds are tested for a particular biological response without prior knowledge about the molecular target. Nowadays, this approach experiences gaining interest again:[357] mainly because many drugs generated in the reverse approach fail in clinical trials due to insufficient efficacy, and because better target deconvolution strategies have been developed over the years.[200] Combining phenotypic screening, chemical proteomics and bioinformatics could revolutionize our way of designing new ligands. A possible scenario would be to start from a cellular phenotype of interest and screen various inhibitors for their phenotypic effect.[358] Optimally, these inhibitors are well profiled by chemical proteomics and comprise target profiles that cover different target families with slight overlaps. A computational approach can be used to attribute the observed effect size of the inhibitor treatment to their target profiles and identify which protein targets are responsible for the observed effect. This analysis allows for characterization of the molecular contributors to the observed phenotype but also enables the rational selection of appropriate molecules to interrogate a certain phenotype instead of single proteins. It could even be useful to predict the phenotypic effect a compound might have according to its target profile. One such combinatory approach was developed by Gujral *et al*[358] and was already tested successfully in-house in order to evaluate the anti-inflammatory effect of SIK2 inhibitors (identified by Kinobeads drug screen) in an inflammatory background.

3 What makes a good drug target?

Target validation. Drug attrition during clinical trials is a major and cost-intensive problem of the pharmaceutic industry. As already discussed previously, drugs usually fail due to toxic side effects or low efficacy.[359] To overcome this problem, intensive characterization is required on both sides, the compound and the target.[204,360] Target validation describes the verification of a predicted molecular target of a drug – i.e. that inhibition of this target leads to the desired phenotypic effect and that the target is "druggable". This may include the elucidation of structure-activity-relationships and evaluation of target engagement *in vitro*, in cells and in *in vivo* animal models (e.g. mouse, zebrafish).[361] Classical molecular biology methods such as the generation of drug-resistant mutants and genetic knockdown experiments are also frequently used. Additionally, chemical probes provide the possibility of pharmacological knockdown experiments.[188] New chemical biology approaches to interrogate protein function include proteolysis targeting chimeric molecules (PROTAC).[362] These probes consist of two moieties, one targeting the protein of interest and one recruiting E3 ubiquitin ligases. It leads to ubiquitination and subsequent proteasomal degradation of the target protein and provides the opportunity to mediate protein knockdown in a time controllable manner. Most drugs act on proteins which may be expressed as several isoforms, comprise different post-translational modifications or form diverse complexes. Thus, target validation on proteome level is more meaningful than genetic disruption of a protein. Proteomics can for instance be useful for monitoring the downstream signaling of the presumed molecular target in order to prove target engagement within the cell.

EPHA2 target validation. Despite the fact that several preclinical studies and first results of clinical trials showed positive effects of EPHA2 inhibition in an oncological background[144] as well as during pathogen infection[74], EPHA2 is still not fully accepted as *bona fide* drug target and may require more thorough target validation. As described earlier, EPHA2 signaling is very heterogeneous and its biological and pathological function is not understood to full extent yet. If the underlying biology and the complex interplay between EPH receptors and their signaling pathways is too complicated and promiscuous, this kinase family might be de-validated as valuable drug targets. On the contrary, these intertwined signaling patterns might define very specific biological outcome which could also be seen as possibility to develop fine-tuned therapeutics against certain branches of EPH receptor signaling.

Innovative approaches need to be developed to further investigate EPHA2 biology and its utility as a drug target. Small molecule kinase inhibitors are only one potential tool to enable pharmacological knockdown of kinase function. Other interesting options would be constrained peptides or small molecules targeting different structural modules such as the extracellular ligand binding domain, the intracellular SAM domain or the PDZ domain. Peptide motifs targeting the ligand binding domain have been discovered by phage display in the past, and are extensively used in the literature for the interrogation of EPHA2 signaling and for targeted drug delivery to cancer cells. One option to further develop this strategy could be constrained peptides[363] that offer increased plasma endurance, enhanced proteolytic stability, and potentially higher affinity due to more favorable entropic properties. Using the CLIPS technology,[364] such peptides have already been developed and showed preliminary positive results on *in vitro* target engagement of EPHA2, stability and cell migration (data not shown). This technology could also be used to investigate the function of the intracellular SAM domain. This domain can be targeted by helical peptides which could be functional in inhibiting the protein:protein interaction of EPHA2 and downstream interaction partners such as INPPL1 or ANKS1A.[365] The PDZ domain is another structural module of EPHA2 which has not been investigated a lot in the past and where no clear biological function has been attributed so far. PDZ domains can be targeted by peptides[366] as well as small molecules[367] and the development of such tools could help to shed light on their role in EPHA2 signaling. The presented work has expanded the toolbox by populating the small molecule angle of EPHA2 targeting strategies and to help evaluating the potential of EPHA2 as a drug target. All these technologies may however be useful and required to understand how EPHA2 functions in different cellular contexts, and hence validate it as a target.

4 Yin and Yang

The combination of multi-disciplinary chemical biology and traditional medicinal chemistry is emerging and provides the opportunity of thriving and innovative approaches to trigger advances on both sides, biology and chemistry.[368] Alleviating the boundaries between these traditional disciplines will generate new tools to interrogate biological function and to investigate the influence of drug treatment on the cellular level. Optimally, the interplay between chemical biology and medicinal chemistry functions as a Yin and Yang mechanism, where chemical biology facilitates the identification of novel molecular targets and lead structures, and medicinal chemistry helps to create better compounds for further investigation or treatment. In this context, this thesis is representative on how chemical proteomics can be used to support different stages of a medicinal chemistry project and provides novel concepts of chemoproteomics-aided drug discovery.

References

1 Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934, doi:10.1126/science.1075762 (2002).

2 Roskoski, R., Jr. A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacol Res* **100**, 1-23, doi:10.1016/j.phrs.2015.07.010 (2015).

3 Drescher, U. Eph family functions from an evolutionary perspective. *Curr Opin Genet Dev* **12**, 397-402 (2002).

4 Lackmann, M. & Boyd, A. W. Eph, a protein family coming of age: more confusion, insight, or complexity? *Sci Signal* **1**, re2, doi:10.1126/stke.115re2 (2008).

5 Boyd, A. W., Bartlett, P. F. & Lackmann, M. Therapeutic targeting of EPH receptors and their ligands. *Nat Rev Drug Discov* **13**, 39-62, doi:10.1038/nrd4175 (2014).

6 Hirai, H., Maru, Y., Hagiwara, K., Nishida, J. & Takaku, F. A novel putative tyrosine kinase receptor encoded by the eph gene. *Science* **238**, 1717-1720 (1987).

7 Gale, N. W. *et al.* Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. *Neuron* **17**, 9-19 (1996).

8 Himanen, J. P. & Nikolov, D. B. Eph signaling: a structural view. *Trends Neurosci* **26**, 46-51 (2003).

9 Guo, F. Y. & Lesk, A. M. Sizes of interface residues account for cross-class binding affinity patterns in Eph receptor-ephrin families. *Proteins* **82**, 349-353, doi:10.1002/prot.24414 (2014).

10 Pasquale, E. B. Eph-ephrin promiscuity is now crystal clear. *Nat Neurosci* **7**, 417-418, doi:10.1038/nn0504-417 (2004).

11 Drescher, U. *et al.* In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. *Cell* **82**, 359-370 (1995).

12 Henkemeyer, M. *et al.* Nuk controls pathfinding of commissural axons in the mammalian central nervous system. *Cell* **86**, 35-46 (1996).

13 Kania, A. & Klein, R. Mechanisms of ephrin-Eph signalling in development, physiology and disease. *Nat Rev Mol Cell Biol* **17**, 240-256, doi:10.1038/nrm.2015.16 (2016).

14 Reber, M., Burrola, P. & Lemke, G. A relative signalling model for the formation of a topographic neural map. *Nature* **431**, 847-853, doi:10.1038/nature02957 (2004).

15 Janes, P. W., Nievergall, E. & Lackmann, M. Concepts and consequences of Eph receptor clustering. *Semin Cell Dev Biol* **23**, 43-50, doi:10.1016/j.semcdb.2012.01.001 (2012).

16 Lin, B., Yin, T., Wu, Y. I., Inoue, T. & Levchenko, A. Interplay between chemotaxis and contact inhibition of locomotion determines exploratory cell migration. *Nat Commun* **6**, 6619, doi:10.1038/ncomms7619 (2015).

17 Batson, J., Astin, J. W. & Nobes, C. D. Regulation of contact inhibition of locomotion by Eph-ephrin signalling. *J Microsc* **251**, 232-241, doi:10.1111/jmi.12024 (2013).

18 Hruska, M. & Dalva, M. B. Ephrin regulation of synapse formation, function and plasticity. *Mol Cell Neurosci* **50**, 35-44, doi:10.1016/j.mcn.2012.03.004 (2012).

19 Conway, A. *et al.* Multivalent ligands control stem cell behaviour in vitro and in vivo. *Nat Nanotechnol* **8**, 831-838, doi:10.1038/nnano.2013.205 (2013).

20 Dines, M. & Lamprecht, R. The Role of Ephs and Ephrins in Memory Formation. *Int J Neuropsychopharmacol* **19**, doi:10.1093/ijnp/pyv106 (2016).

21 Barquilla, A. & Pasquale, E. B. Eph receptors and ephrins: therapeutic opportunities. *Annu Rev Pharmacol Toxicol* **55**, 465-487, doi:10.1146/annurev-pharmtox-011112-140226 (2015).

22 Chen, Y., Fu, A. K. & Ip, N. Y. Eph receptors at synapses: implications in neurodegenerative diseases. *Cell Signal* **24**, 606-611, doi:10.1016/j.cellsig.2011.11.016 (2012).

23 Singh, D. R. *et al.* The SAM domain inhibits EphA2 interactions in the plasma membrane. *Biochim Biophys Acta* **1864**, 31-38, doi:10.1016/j.bbamcr.2016.10.011 (2017).

24 Nikolov, D. B., Xu, K. & Himanen, J. P. Eph/ephrin recognition and the role of Eph/ephrin clusters in signaling initiation. *Biochim Biophys Acta* **1834**, 2160-2165, doi:10.1016/j.bbapap.2013.04.020 (2013).

25 Chavent, M., Seiradake, E., Jones, E. Y. & Sansom, M. S. Structures of the EphA2 Receptor at the Membrane: Role of Lipid Interactions. *Structure* **24**, 337-347, doi:10.1016/j.str.2015.11.008 (2016).

26 Hunter, S. G. *et al.* Essential role of Vav family guanine nucleotide exchange factors in EphA receptor-mediated angiogenesis. *Mol Cell Biol* **26**, 4830-4842, doi:10.1128/MCB.02215-05 (2006).

27 Hubbard, S. R. Theme and variations: juxtamembrane regulation of receptor protein kinases. *Mol Cell* **8**, 481-482 (2001).

28 Binns, K. L., Taylor, P. P., Sicheri, F., Pawson, T. & Holland, S. J. Phosphorylation of tyrosine residues in the kinase domain and juxtamembrane region regulates the biological and catalytic activities of Eph receptors. *Mol Cell Biol* **20**, 4791-4805 (2000).

29 Holmberg, J. & Frisen, J. Ephrins are not only unattractive. *Trends Neurosci* **25**, 239-243 (2002).

30 Gucciardo, E., Sugiyama, N. & Lehti, K. Eph- and ephrin-dependent mechanisms in tumor and stem cell dynamics. *Cell Mol Life Sci* **71**, 3685-3710, doi:10.1007/s00018-014-1633-0 (2014).

31 Shi, X. *et al.* A role of the SAM domain in EphA2 receptor activation. *Sci Rep* **7**, 45084, doi:10.1038/srep45084 (2017).

32 Seiradake, E. *et al.* Structurally encoded intraclass differences in EphA clusters drive distinct cell responses. *Nat Struct Mol Biol* **20**, 958-964, doi:10.1038/nsmb.2617 (2013).

33 Singh, D. R. *et al.* Unliganded EphA3 dimerization promoted by the SAM domain. *Biochem J* **471**, 101-109, doi:10.1042/BJ20150433 (2015).

34 Lee, H. J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal* **8**, 8, doi:10.1186/1478-811X-8-8 (2010).

35 Marston, D. J., Dickinson, S. & Nobes, C. D. Rac-dependent trans-endocytosis of ephrinBs regulates Eph-ephrin contact repulsion. *Nat Cell Biol* **5**, 879-888, doi:10.1038/ncb1044 (2003).

36 Yoo, S., Shin, J. & Park, S. EphA8-ephrinA5 signaling and clathrin-mediated endocytosis is regulated by Tiam-1, a Rac-specific guanine nucleotide exchange factor. *Mol Cells* **29**, 603-609, doi:10.1007/s10059-010-0075-2 (2010).

37 Parker, M. *et al.* Reverse endocytosis of transmembrane ephrin-B ligands via a clathrin-mediated pathway. *Biochem Biophys Res Commun* **323**, 17-23, doi:10.1016/j.bbrc.2004.07.209 (2004).

38 Janes, P. W. *et al.* Adam meets Eph: an ADAM substrate recognition module acts as a molecular switch for ephrin cleavage in trans. *Cell* **123**, 291-304, doi:10.1016/j.cell.2005.08.014 (2005).

39 Atapattu, L. *et al.* Antibodies binding the ADAM10 substrate recognition domain inhibit Eph function. *J Cell Sci* **125**, 6084-6093, doi:10.1242/jcs.112631 (2012).

40 Miao, H. & Wang, B. EphA receptor signaling--complexity and emerging themes. *Semin Cell Dev Biol* **23**, 16-25, doi:10.1016/j.semcdb.2011.10.013 (2012).

41 Davy, A. & Soriano, P. Ephrin signaling in vivo: look both ways. *Dev Dyn* **232**, 1-10, doi:10.1002/dvdy.20200 (2005).

42 Park, J. E., Son, A. I. & Zhou, R. Roles of EphA2 in Development and Disease. *Genes (Basel)* **4**, 334-357, doi:10.3390/genes4030334 (2013).

43 Taddei, M. L. *et al.* Kinase-dependent and -independent roles of EphA2 in the regulation of prostate cancer invasion and metastasis. *Am J Pathol* **174**, 1492-1503, doi:10.2353/ajpath.2009.080473 (2009).

44 Borthakur, S., Lee, H., Kim, S., Wang, B. C. & Buck, M. Binding and function of phosphotyrosines of the Ephrin A2 (EphA2) receptor using synthetic sterile alpha motif (SAM) domains. *J Biol Chem* **289**, 19694-19703, doi:10.1074/jbc.M114.567602 (2014).

45 Leone, M., Cellitti, J. & Pellecchia, M. NMR studies of a heterotypic Sam-Sam domain association: the interaction between the lipid phosphatase Ship2 and the EphA2 receptor. *Biochemistry* **47**, 12721-12728, doi:10.1021/bi801713f (2008).

46 Boissier, P., Chen, J. & Huynh-Do, U. EphA2 signaling following endocytosis: role of Tiam1. *Traffic* **14**, 1255-1271, doi:10.1111/tra.12123 (2013).

47 Tandon, M., Vemula, S. V. & Mittal, S. K. Emerging strategies for EphA2 receptor targeting for cancer therapeutics. *Expert Opin Ther Targets* **15**, 31-51, doi:10.1517/14728222.2011.538682 (2011).

48 Miao, H. *et al.* EphA2 mediates ligand-dependent inhibition and ligand-independent promotion of cell migration and invasion via a reciprocal regulatory loop with Akt. *Cancer Cell* **16**, 9-20, doi:10.1016/j.ccr.2009.04.009 (2009).

49 Yang, N. Y. *et al.* Crosstalk of the EphA2 receptor with a serine/threonine phosphatase suppresses the Akt-mTORC1 pathway in cancer cells. *Cell Signal* **23**, 201-212, doi:10.1016/j.cellsig.2010.09.004 (2011).

50 Zhou, Y. *et al.* Crucial roles of RSK in cell motility by catalysing serine phosphorylation of EphA2. *Nat Commun* **6**, 7679, doi:10.1038/ncomms8679 (2015).

51 Brantley-Sieders, D. M. *et al.* The receptor tyrosine kinase EphA2 promotes mammary adenocarcinoma tumorigenesis and metastatic progression in mice by amplifying ErbB2 signaling. *J Clin Invest* **118**, 64-78, doi:10.1172/JCI33154 (2008).

52 Macrae, M. *et al.* A conditional feedback loop regulates Ras activity through EphA2. *Cancer Cell* **8**, 111-118, doi:10.1016/j.ccr.2005.07.005 (2005).

53 Koshikawa, N. *et al.* Proteolysis of EphA2 Converts It from a Tumor Suppressor to an Oncoprotein. *Cancer Res* **75**, 3327-3339, doi:10.1158/0008-5472.CAN-14-2798 (2015).

54 Miao, H. *et al.* Activation of EphA receptor tyrosine kinase inhibits the Ras/MAPK pathway. *Nat Cell Biol* **3**, 527-530, doi:10.1038/35074604 (2001).

55 Hamaoka, Y., Negishi, M. & Katoh, H. EphA2 is a key effector of the MEK/ERK/RSK pathway regulating glioblastoma cell proliferation. *Cell Signal* **28**, 937-945, doi:10.1016/j.cellsig.2016.04.009 (2016).

56 Chitramuthu, B. & Bateman, A. Progranulin and the receptor tyrosine kinase EphA2, partners in crime? *J Cell Biol* **215**, 603-605, doi:10.1083/jcb.201610097 (2016).

57 Neill, T. *et al.* EphA2 is a functional receptor for the growth factor progranulin. *J Cell Biol* **215**, 687-703, doi:10.1083/jcb.201603079 (2016).

58 Greene, A. C. *et al.* Spatial organization of EphA2 at the cell-cell interface modulates trans-endocytosis of ephrinA1. *Biophys J* **106**, 2196-2205, doi:10.1016/j.bpj.2014.03.043 (2014).

59 Balasubramaniam, D., Paul, L. N., Homan, K. T., Hall, M. C. & Stauffacher, C. V. Specificity of HCPTP variants toward EphA2 tyrosines by quantitative selected reaction monitoring. *Protein Sci* **20**, 1172-1181, doi:10.1002/pro.646 (2011).

60 Kikawa, K. D., Vidale, D. R., Van Etten, R. L. & Kinch, M. S. Regulation of the EphA2 kinase by the low molecular weight tyrosine phosphatase induces transformation. *J Biol Chem* **277**, 39274-39279, doi:10.1074/jbc.M207127200 (2002).

61 Chiarugi, P. *et al.* LMW-PTP is a positive regulator of tumor onset and growth. *Oncogene* **23**, 3905-3914, doi:10.1038/sj.onc.1207508 (2004).

62 Locard-Paulet, M. *et al.* Phosphoproteomic analysis of interacting tumor and endothelial cells identifies regulatory mechanisms of transendothelial migration. *Sci Signal* **9**, ra15, doi:10.1126/scisignal.aac5820 (2016).

63 Wei, Q. *et al.* Structures of an Eph receptor tyrosine kinase and its potential activation mechanism. *Acta Crystallogr D Biol Crystallogr* **70**, 3135-3143, doi:10.1107/S1399004714021944 (2014).

64 Elena, B. P. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. *Nature reviews. Cancer* **10**, 165-180, doi:10.1038/nrc2806 (2010).

65 Shi, Y., De Maria, A., Bennett, T., Shiels, A. & Bassnett, S. A role for epha2 in cell migration and refractive organization of the ocular lens. *Invest Ophthalmol Vis Sci* **53**, 551-559, doi:10.1167/iovs.11-8568 (2012).

66 Cheng, C. & Gong, X. Diverse roles of Eph/ephrin signaling in the mouse lens. *PLoS One* **6**, e28147, doi:10.1371/journal.pone.0028147 (2011).

67 Lee, S., Shatadal, S. & Griep, A. E. Dlg-1 Interacts With and Regulates the Activities of Fibroblast Growth Factor Receptors and EphA2 in the Mouse Lens. *Invest Ophthalmol Vis Sci* **57**, 707-718, doi:10.1167/iovs.15-17727 (2016).

68 Matsuo, K. & Otaki, N. Bone cell interactions through Eph/ephrin: bone modeling, remodeling and associated diseases. *Cell Adh Migr* **6**, 148-156, doi:10.4161/cam.20888 (2012).

69 Perez White, B. E. *et al.* EphA2 proteomics in human keratinocytes reveals a novel association with afadin and epidermal tight junctions. *J Cell Sci* **130**, 111-118, doi:10.1242/jcs.188169 (2017).

70 Goichberg, P. *et al.* The ephrin A1-EphA2 system promotes cardiac stem cell migration after infarction. *Circ Res* **108**, 1071-1083, doi:10.1161/CIRCRESAHA.110.239459 (2011).

71 Ogawa, K. *et al.* The ephrin-A1 ligand and its receptor, EphA2, are expressed during tumor neovascularization. *Oncogene* **19**, 6043-6052, doi:10.1038/sj.onc.1204004 (2000).

72 Park, J. E. *et al.* Human cataract mutations in EPHA2 SAM domain alter receptor stability and function. *PLoS One* **7**, e36564, doi:10.1371/journal.pone.0036564 (2012).

73 Dave, A. *et al.* Epha2 Mutations Contribute to Congenital Cataract through Diverse Mechanisms. *Mol Vis* **22**, 18-30 (2016).

74 Lupberger, J. *et al.* EGFR and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. *Nat Med* **17**, 589-595, doi:10.1038/nm.2341 (2011).

75 Colpitts, C. C., Lupberger, J., Doerig, C. & Baumert, T. F. Host cell kinases and the hepatitis C virus life cycle. *Biochim Biophys Acta* **1854**, 1657-1662, doi:10.1016/j.bbapap.2015.04.011 (2015).

76 Hahn, A. S. *et al.* The ephrin receptor tyrosine kinase A2 is a cellular receptor for Kaposi's sarcoma-associated herpesvirus. *Nat Med* **18**, 961-966, doi:10.1038/nm.2805 (2012).

77 Boshoff, C. Ephrin receptor: a door to KSHV infection. *Nat Med* **18**, 861-863, doi:10.1038/nm.2803 (2012).

78 Scholz, R. *et al.* Novel Host Proteins and Signaling Pathways in Enteropathogenic E. coli Pathogenesis Identified by Global Phosphoproteome Analysis. *Mol Cell Proteomics* **14**, 1927-1945, doi:10.1074/mcp.M114.046847 (2015).

79 Subbarayal, P. *et al.* EphrinA2 receptor (EphA2) is an invasion and intracellular signaling receptor for Chlamydia trachomatis. *PLoS Pathog* **11**, e1004846, doi:10.1371/journal.ppat.1004846 (2015).

80 Khounlotham, M., Subbian, S., Smith, R., 3rd, Cirillo, S. L. & Cirillo, J. D. Mycobacterium tuberculosis interferes with the response to infection by inducing the host EphA2 receptor. *J Infect Dis* **199**, 1797-1806, doi:10.1086/599096 (2009).

81 Kaushansky, A. *et al.* Malaria parasites target the hepatocyte receptor EphA2 for successful host infection. *Science* **350**, 1089-1092, doi:10.1126/science.aad3318 (2015).

82 Chakraborty, S., Veettil, M. V., Bottero, V. & Chandran, B. Kaposi's sarcoma-associated herpesvirus interacts with EphrinA2 receptor to amplify signaling essential for productive infection. *Proc Natl Acad Sci U S A* **109**, E1163-1172, doi:10.1073/pnas.1119592109 (2012).

83 Lindberg, R. A. & Hunter, T. cDNA cloning and characterization of eck, an epithelial cell receptor protein-tyrosine kinase in the eph/elk family of protein kinases. *Mol Cell Biol* **10**, 6316-6324 (1990).

84 Pasquale, E. B. Eph-ephrin bidirectional signaling in physiology and disease. *Cell* **133**, 38-52, doi:10.1016/j.cell.2008.03.011 (2008).

85 Al-Ejeh, F. *et al.* Eph family co-expression patterns define unique clusters predictive of cancer phenotype. *Growth Factors* **32**, 254-264, doi:10.3109/08977194.2014.984807 (2014).

86 Miyazaki, T., Kato, H., Fukuchi, M., Nakajima, M. & Kuwano, H. EphA2 overexpression correlates with poor prognosis in esophageal squamous cell carcinoma. *Int J Cancer* **103**, 657-663, doi:10.1002/ijc.10860 (2003).

87 Miao, H. *et al.* EphA2 promotes infiltrative invasion of glioma stem cells in vivo through cross-talk with Akt and regulates stem cell properties. *Oncogene* **34**, 558-567, doi:10.1038/onc.2013.590 (2015).

88 Singh, D. R. *et al.* EphA2 Receptor Unliganded Dimers Suppress EphA2 Pro-tumorigenic Signaling. *J Biol Chem* **290**, 27271-27279, doi:10.1074/jbc.M115.676866 (2015).

89 Sjoblom, T. *et al.* The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268-274, doi:10.1126/science.1133427 (2006).

90 Sugiyama, N. *et al.* EphA2 cleavage by MT1-MMP triggers single cancer cell invasion via homotypic cell repulsion. *J Cell Biol* **201**, 467-484, doi:10.1083/jcb.201205176 (2013).

91 Li, J. J. & Xie, D. The roles and therapeutic potentials of Ephs and ephrins in lung cancer. *Exp Cell Res* **319**, 152-159, doi:10.1016/j.yexcr.2012.08.005 (2013).

92 Wang, H. *et al.* Vasculogenic Mimicry in Prostate Cancer: The Roles of EphA2 and PI3K. *J Cancer* **7**, 1114-1124, doi:10.7150/jca.14120 (2016).

93 Wang, W. *et al.* Epithelial-mesenchymal transition regulated by EphA2 contributes to vasculogenic mimicry formation of head and neck squamous cell carcinoma. *Biomed Res Int* **2014**, 803914, doi:10.1155/2014/803914 (2014).

94 Folberg, R., Hendrix, M. J. & Maniotis, A. J. Vasculogenic mimicry and tumor angiogenesis. *Am J Pathol* **156**, 361-381, doi:10.1016/S0002-9440(10)64739-6 (2000).

95 Zhang, J. *et al.* Vasculogenic mimicry and tumor metastasis. *J BUON* **21**, 533-541 (2016).

96 Gundry, C. *et al.* Phosphorylation of Rab-coupling protein by LMTK3 controls Rab14-dependent EphA2 trafficking to promote cell:cell repulsion. *Nat Commun* **8**, 14646, doi:10.1038/ncomms14646 (2017).

97 Zhuang, G. *et al.* Elevation of receptor tyrosine kinase EphA2 mediates resistance to trastuzumab therapy. *Cancer Res* **70**, 299-308, doi:10.1158/0008-5472.CAN-09-1845 (2010).

98 Koch, H., Busto, M. E., Kramer, K., Medard, G. & Kuster, B. Chemical Proteomics Uncovers EPHA2 as a Mechanism of Acquired Resistance to Small Molecule EGFR Kinase Inhibition. *J Proteome Res* **14**, 2617-2625, doi:10.1021/acs.jproteome.5b00161 (2015).

99 Amato, K. R. *et al.* EPHA2 Blockade Overcomes Acquired Resistance to EGFR Kinase Inhibitors in Lung Cancer. *Cancer Res* **76**, 305-318, doi:10.1158/0008-5472.CAN-15-0717 (2016).

100 Paraiso, K. H. *et al.* Ligand-independent EPHA2 signaling drives the adoption of a targeted therapy-mediated metastatic melanoma phenotype. *Cancer Discov* **5**, 264-273, doi:10.1158/2159-8290.CD-14-0293 (2015).

101 Gokmen-Polar, Y. *et al.* Dual targeting of EphA2 and ER restores tamoxifen sensitivity in ER/EphA2-positive breast cancer. *Breast Cancer Res Treat* **127**, 375-384, doi:10.1007/s10549-010-1004-y (2011).

102 Abdullah, L. N. & Chow, E. K. Mechanisms of chemoresistance in cancer stem cells. *Clin Transl Med* **2**, 3, doi:10.1186/2001-1326-2-3 (2013).

103 Binda, E. *et al.* The EphA2 receptor drives self-renewal and tumorigenicity in stem-like tumor-propagating cells from human glioblastomas. *Cancer Cell* **22**, 765-780, doi:10.1016/j.ccr.2012.11.005 (2012).

104 Noberini, R., Lamberto, I. & Pasquale, E. B. Targeting Eph receptors with peptides and small molecules: progress and challenges. *Semin Cell Dev Biol* **23**, 51-57, doi:10.1016/j.semcdb.2011.10.023 (2012).

105 Lodola, A., Giorgio, C., Incerti, M., Zanotti, I. & Tognolini, M. Targeting Eph/ephrin system in cancer therapy. *Eur J Med Chem*, doi:10.1016/j.ejmech.2017.07.029 (2017).

106 Noberini, R. *et al.* A disalicylic acid-furanyl derivative inhibits ephrin binding to a subset of Eph receptors. *Chem Biol Drug Des* **78**, 667-678, doi:10.1111/j.1747-0285.2011.01199.x (2011).

107 Noberini, R. *et al.* Small molecules can selectively inhibit ephrin binding to the EphA4 and EphA2 receptors. *J Biol Chem* **283**, 29461-29472, doi:10.1074/jbc.M804103200 (2008).

108 Giorgio, C. *et al.* Lithocholic acid is an Eph-ephrin ligand interfering with Eph-kinase activation. *PLoS One* **6**, e18128, doi:10.1371/journal.pone.0018128 (2011).

109 Choi, Y. *et al.* Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg Med Chem Lett* **19**, 4467-4470, doi:10.1016/j.bmcl.2009.05.029 (2009).

110 Hatano, M. *et al.* EphA2 as a glioma-associated antigen: a novel target for glioma vaccines. *Neoplasia* **7**, 717-722 (2005).

111 Pollack, I. F. *et al.* Antigen-specific immunoreactivity and clinical outcome following vaccination with glioma-associated antigen peptides in children with recurrent high-grade gliomas: results of a pilot study. *J Neurooncol* **130**, 517-527, doi:10.1007/s11060-016-2245-3 (2016).

112 Yamaguchi, S. *et al.* Immunotherapy of murine colon cancer using receptor tyrosine kinase EphA2-derived peptide-pulsed dendritic cell vaccines. *Cancer* **110**, 1469-1477, doi:10.1002/cncr.22958 (2007).

113 Chow, K. K. *et al.* T cells redirected to EphA2 for the immunotherapy of glioblastoma. *Mol Ther* **21**, 629-637, doi:10.1038/mt.2012.210 (2013).

114 Wagner, M. J. *et al.* Preclinical Mammalian Safety Studies of EPHARNA (DOPC Nanoliposomal EphA2-Targeted siRNA). *Mol Cancer Ther* **16**, 1114-1123, doi:10.1158/1535-7163.MCT-16-0541 (2017).

115 Tanaka, T. *et al.* A Modified Adenovirus Vector-Mediated Antibody Screening Method Identifies EphA2 as a Cancer Target. *Transl Oncol* **10**, 476-484, doi:10.1016/j.tranon.2017.04.001 (2017).

116 Kiewlich, D. *et al.* Anti-EphA2 antibodies decrease EphA2 protein levels in murine CT26 colorectal and human MDA-231 breast tumors but do not inhibit tumor growth. *Neoplasia* **8**, 18-30, doi:10.1593/neo.05544 (2006).

117 Coffman, K. T. *et al.* Differential EphA2 epitope display on normal versus malignant cells. *Cancer Res* **63**, 7907-7912 (2003).

118 Bruckheimer, E. M. *et al.* Antibody-dependent cell-mediated cytotoxicity effector-enhanced EphA2 agonist monoclonal antibody demonstrates potent activity against human tumors. *Neoplasia* **11**, 509-517, 502 p following 517 (2009).

119 Lee, J. W. *et al.* EphA2 targeted chemotherapy using an antibody drug conjugate in endometrial carcinoma. *Clin Cancer Res* **16**, 2562-2570, doi:10.1158/1078-0432.CCR-10-0017 (2010).

120 Merritt, W. M. *et al.* Clinical and biological impact of EphA2 overexpression and angiogenesis in endometrial cancer. *Cancer Biol Ther* **10**, 1306-1314 (2010).

121 Jackson, D. *et al.* A human antibody-drug conjugate targeting EphA2 inhibits tumor growth in vivo. *Cancer Res* **68**, 9367-9374, doi:10.1158/0008-5472.CAN-08-1933 (2008).

122 Annunziata, C. M. *et al.* Phase 1, open-label study of MEDI-547 in patients with relapsed or refractory solid tumors. *Invest New Drugs* **31**, 77-84, doi:10.1007/s10637-012-9801-2 (2013).

123 Dobrzanski, P. *et al.* Antiangiogenic and antitumor efficacy of EphA2 receptor antagonist. *Cancer Res* **64**, 910-919 (2004).

124 Wykosky, J. *et al.* Soluble monomeric EphrinA1 is released from tumor cells and is a functional ligand for the EphA2 receptor. *Oncogene* **27**, 7260-7273, doi:10.1038/onc.2008.328 (2008).

125 Wykosky, J., Gibo, D. M. & Debinski, W. A novel, potent, and specific ephrinA1-based cytotoxin against EphA2 receptor expressing tumor cells. *Mol Cancer Ther* **6**, 3208-3218, doi:10.1158/1535-7163.MCT-07-0200 (2007).

126 Wang, S. *et al.* Novel targeted system to deliver chemotherapeutic drugs to EphA2-expressing cancer cells. *J Med Chem* **55**, 2427-2436, doi:10.1021/jm201743s (2012).

127 Mitra, S. *et al.* Structure-activity relationship analysis of peptides targeting the EphA2 receptor. *Biochemistry* **49**, 6687-6695, doi:10.1021/bi1006223 (2010).

128 Giorgio, C. *et al.* Biochemical characterization of EphA2 antagonists with improved physico-chemical properties by cell-based assays and surface plasmon resonance analysis. *Biochem Pharmacol* **99**, 18-30, doi:10.1016/j.bcp.2015.10.006 (2016).

129 Incerti, M. *et al.* Amino acid conjugates of lithocholic acid as antagonists of the EphA2 receptor. *J Med Chem* **56**, 2936-2947, doi:10.1021/jm301890k (2013).

130 Wu, B. *et al.* HTS by NMR of combinatorial libraries: a fragment-based approach to ligand discovery. *Chem Biol* **20**, 19-33, doi:10.1016/j.chembiol.2012.10.015 (2013).

131 Giorgio, C. *et al.* The ellagitannin colonic metabolite urolithin D selectively inhibits EphA2 phosphorylation in prostate cancer cells. *Mol Nutr Food Res* **59**, 2155-2167, doi:10.1002/mnfr.201500470 (2015).

132 Russo, S. *et al.* Synthesis and structure-activity relationships of amino acid conjugates of cholanic acid as antagonists of the EphA2 receptor. *Molecules* **18**, 13043-13060, doi:10.3390/molecules181013043 (2013).

133 Tognolini, M. *et al.* Structure-activity relationships and mechanism of action of Eph-ephrin antagonists: interaction of cholanic acid with the EphA2 receptor. *ChemMedChem* **7**, 1071-1083, doi:10.1002/cmdc.201200102 (2012).

134 Hassan-Mohamed, I. *et al.* UniPR129 is a competitive small molecule Eph-ephrin antagonist blocking in vitro angiogenesis at low micromolar concentrations. *Br J Pharmacol* **171**, 5195-5208, doi:10.1111/bph.12669 (2014).

135 Lim, C. J. *et al.* 4-Substituted quinazoline derivatives as novel EphA2 receptor tyrosine kinase inhibitors. *Bioorg Med Chem Lett* **24**, 4080-4083, doi:10.1016/j.bmcl.2014.07.081 (2014).

136 Zhu, Y. *et al.* Synthesis and Biological Evaluation of 1-(2-Aminophenyl)-3-arylurea Derivatives as Potential EphA2 and HDAC Dual Inhibitors. *Chem Pharm Bull (Tokyo)* **64**, 1136-1141, doi:10.1248/cpb.c16-00154 (2016).

137 Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* **25**, 1035-1044, doi:10.1038/nbt1328 (2007).

138 O'Hare, T. *et al.* AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell* **16**, 401-412, doi:10.1016/j.ccr.2009.09.028 (2009).

139 Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat Biotechnol* **29**, 1046-1051, doi:10.1038/nbt.1990 (2011).

140 Green, T. P. *et al.* Preclinical anticancer activity of the potent, oral Src inhibitor AZD0530. *Mol Oncol* **3**, 248-261, doi:10.1016/j.molonc.2009.01.002 (2009).

141 Gao, Y. *et al.* A broad activity screen in support of a chemogenomic map for kinase signalling research and drug discovery. *Biochem J* **451**, 313-328, doi:10.1042/BJ20121418 (2013).

142 Wilhelm, S. M. *et al.* Regorafenib (BAY 73-4506): a new oral multikinase inhibitor of angiogenic, stromal and oncogenic receptor tyrosine kinases with potent preclinical antitumor activity. *Int J Cancer* **129**, 245-255, doi:10.1002/ijc.25864 (2011).

143 Metz, J. T. *et al.* Navigating the kinome. *Nat Chem Biol* **7**, 200-202, doi:10.1038/nchembio.530 (2011).

144 Huang, J. *et al.* Cross-talk between EphA2 and BRaf/CRaf is a key determinant of response to Dasatinib. *Clin Cancer Res* **20**, 1846-1855, doi:10.1158/1078-0432.CCR-13-2141 (2014).

145 Wu, P., Nielsen, T. E. & Clausen, M. H. FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol Sci*, doi:10.1016/j.tips.2015.04.005 (2015).

146 Muller, S., Chaikuad, A., Gray, N. S. & Knapp, S. The ins and outs of selective kinase inhibitor development. *Nat Chem Biol* **11**, 818-821, doi:10.1038/nchembio.1938 (2015).

147 Hunter, T. Why nature chose phosphate to modify proteins. *Philos Trans R Soc Lond B Biol Sci* **367**, 2513-2516, doi:10.1098/rstb.2012.0013 (2012).

148 Kornev, A. P., Haste, N. M., Taylor, S. S. & Eyck, L. F. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A* **103**, 17783-17788, doi:10.1073/pnas.0607656103 (2006).

149 Pucheta-Martinez, E. *et al.* An Allosteric Cross-Talk Between the Activation Loop and the ATP Binding Site Regulates the Activation of Src Kinase. *Sci Rep* **6**, 24235, doi:10.1038/srep24235 (2016).

150 Taylor, S. S., Shaw, A. S., Kannan, N. & Kornev, A. P. Integration of signaling in the kinome: Architecture and regulation of the alphaC Helix. *Biochim Biophys Acta* **1854**, 1567-1574, doi:10.1016/j.bbapap.2015.04.007 (2015).

151 Kornev, A. P., Taylor, S. S. & Ten Eyck, L. F. A helix scaffold for the assembly of active protein kinases. *Proc Natl Acad Sci U S A* **105**, 14377-14382, doi:10.1073/pnas.0807988105 (2008).

152 Steichen, J. M. *et al.* Global consequences of activation loop phosphorylation on protein kinase A. *J Biol Chem* **285**, 3825-3832, doi:10.1074/jbc.M109.061820 (2010).

153 Johnson, L. N. & Lewis, R. J. Structural basis for control by phosphorylation. *Chem Rev* **101**, 2209-2242 (2001).

154 Roskoski, R., Jr. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol Res* **103**, 26-48, doi:10.1016/j.phrs.2015.10.021 (2016).

155 Adams, J. A. Activation loop phosphorylation and catalysis in protein kinases: is there functional evidence for the autoinhibitor model? *Biochemistry* **42**, 601-607, doi:10.1021/bi020617o (2003).

156 McTigue, M. *et al.* Molecular conformations, interactions, and properties associated with drug efficiency and clinical performance among VEGFR TK inhibitors. *Proc Natl Acad Sci U S A* **109**, 18281-18289, doi:10.1073/pnas.1207759109 (2012).

157 Wybenga-Groot, L. E. *et al.* Structural basis for autoinhibition of the Ephb2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. *Cell* **106**, 745-757 (2001).

158 Wodicka, L. M. *et al.* Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry. *Chem Biol* **17**, 1241-1249, doi:10.1016/j.chembiol.2010.09.010 (2010).

159 Solowiej, J. *et al.* Characterizing the effects of the juxtamembrane domain on vascular endothelial growth factor receptor-2 enzymatic activity, autophosphorylation, and inhibition by axitinib. *Biochemistry* **48**, 7019-7031, doi:10.1021/bi900522y (2009).

160 Klein, T. *et al.* Structural and dynamic insights into the energetics of activation loop rearrangement in FGFR1 kinase. *Nat Commun* **6**, 7877, doi:10.1038/ncomms8877 (2015).

161 Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **109**, 275-282 (2002).

162 Lamontanara, A. J., Georgeon, S., Tria, G., Svergun, D. I. & Hantschel, O. The SH2 domain of Abl kinases regulates kinase autophosphorylation by controlling activation loop accessibility. *Nat Commun* **5**, 5470, doi:10.1038/ncomms6470 (2014).

163 Hammaren, H. M., Virtanen, A. T. & Silvennoinen, O. Nucleotide-binding mechanisms in pseudokinases. *Biosci Rep* **36**, e00282, doi:10.1042/BSR20150226 (2016).

164 Min, X. *et al.* Structural and Functional Characterization of the JH2 Pseudokinase Domain of JAK Family Tyrosine Kinase 2 (TYK2). *J Biol Chem* **290**, 27261-27270, doi:10.1074/jbc.M115.672048 (2015).

165 Jeffrey, P. D. *et al.* Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**, 313-320, doi:10.1038/376313a0 (1995).

166 Rudolf, A. F., Skovgaard, T., Knapp, S., Jensen, L. J. & Berthelsen, J. A comparison of protein kinases inhibitor screening methods using both enzymatic activity and binding affinity determination. *PLoS One* **9**, e98800, doi:10.1371/journal.pone.0098800 (2014).

167 Dar, A. C. & Shokat, K. M. The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annu Rev Biochem* **80**, 769-795, doi:10.1146/annurev-biochem-090308-173656 (2011).

168 Gavrin, L. K. & Saiah, E. Approaches to discover non-ATP site kinase inhibitors. *Med Chem Commun* **4**, 41-51, doi:10.1039/c2md20180a (2013).

169 Adrian, F. J. *et al.* Allosteric inhibitors of Bcr-abl-dependent cell proliferation. *Nat Chem Biol* **2**, 95-102, doi:10.1038/nchembio760 (2006).

170 Lamba, V. & Ghosh, I. New directions in targeting protein kinases: focusing upon true allosteric and bivalent inhibitors. *Curr Pharm Des* **18**, 2936-2945 (2012).

171 Parang, K. *et al.* Mechanism-based design of a protein kinase inhibitor. *Nat Struct Biol* **8**, 37-41, doi:10.1038/83028 (2001).

172 Liu, Q. *et al.* Developing irreversible inhibitors of the protein kinase cysteinome. *Chem Biol* **20**, 146-159, doi:10.1016/j.chembiol.2012.12.006 (2013).

173 Liao, J. J. Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *J Med Chem* **50**, 409-424, doi:10.1021/jm0608107 (2007).

174 van Linden, O. P., Kooistra, A. J., Leurs, R., de Esch, I. J. & de Graaf, C. KLIFS: a knowledge-based structural database to navigate kinase-ligand interaction space. *J Med Chem* **57**, 249-277, doi:10.1021/jm400378w (2014).

175 Zuccotto, F., Ardini, E., Casale, E. & Angiolini, M. Through the "gatekeeper door": exploiting the active kinase conformation. *J Med Chem* **53**, 2681-2694, doi:10.1021/jm901443h (2010).

176 Levinson, N. M. *et al.* A Src-like inactive conformation in the abl tyrosine kinase domain. *PLoS Biol* **4**, e144, doi:10.1371/journal.pbio.0040144 (2006).

177 Vogtherr, M. *et al.* NMR characterization of kinase p38 dynamics in free and ligand-bound forms. *Angew Chem Int Ed Engl* **45**, 993-997, doi:10.1002/anie.200502770 (2006).

178 Vijayan, R. S. *et al.* Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type II inhibitors. *J Med Chem* **58**, 466-479, doi:10.1021/jm501603h (2015).

179 Zhao, Z. *et al.* Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery? *ACS Chem Biol* **9**, 1230-1241, doi:10.1021/cb500129t (2014).

180 Sutherland, J. J., Gao, C., Cahya, S. & Vieth, M. What general conclusions can we draw from kinase profiling data sets? *Biochim Biophys Acta* **1834**, 1425-1433, doi:10.1016/j.bbapap.2012.12.023 (2013).

181 Wentsch, H. K. *et al.* Optimized Target Residence Time: Type I1/2 Inhibitors for p38alpha MAP Kinase with Improved Binding Kinetics through Direct Interaction with the R-Spine. *Angew Chem Int Ed Engl* **56**, 5363-5367, doi:10.1002/anie.201701185 (2017).

182 Schwartz, P. A. & Murray, B. W. Protein kinase biochemistry and drug discovery. *Bioorg Chem* **39**, 192-210, doi:10.1016/j.bioorg.2011.07.004 (2011).

183 Milik, S. N., Lasheen, D. S., Serya, R. A. T. & Abouzid, K. A. M. How to train your inhibitor: Design strategies to overcome resistance to Epidermal Growth Factor Receptor inhibitors. *Eur J Med Chem*, doi:10.1016/j.ejmech.2017.07.023 (2017).

184 Tanramluk, D., Schreyer, A., Pitt, W. R. & Blundell, T. L. On the origins of enzyme inhibitor selectivity and promiscuity: a case study of protein kinase binding to staurosporine. *Chem Biol Drug Des* **74**, 16-24, doi:10.1111/j.1747-0285.2009.00832.x (2009).

185 Noble, M. E., Endicott, J. A. & Johnson, L. N. Protein kinase inhibitors: insights into drug design from structure. *Science* **303**, 1800-1805, doi:10.1126/science.1095920 (2004).

186 Gibbons, D. L., Pricl, S., Kantarjian, H., Cortes, J. & Quintas-Cardama, A. The rise and fall of gatekeeper mutations? The BCR-ABL1 T315I paradigm. *Cancer* **118**, 293-299, doi:10.1002/cncr.26225 (2012).

187 Reddy, A. S. & Zhang, S. Polypharmacology: drug discovery for the future. *Expert Rev Clin Pharmacol* **6**, 41-47, doi:10.1586/ecp.12.74 (2013).

188 Arrowsmith, C. H. *et al.* The promise and peril of chemical probes. *Nat Chem Biol* **11**, 536-541, doi:10.1038/nchembio.1867 (2015).

189 Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H. & Peterson, J. R. Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat Biotechnol* **29**, 1039-1045, doi:10.1038/nbt.2017 (2011).

190 Elkins, J. M. *et al.* Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat Biotechnol* **34**, 95-103, doi:10.1038/nbt.3374 (2016).

191 Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat Biotechnol* **26**, 127-132, doi:10.1038/nbt1358 (2008).

192 Fabian, M. A. *et al.* A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotechnol* **23**, 329-336, doi:10.1038/nbt1068 (2005).

193 Uitdehaag, J. C. & Zaman, G. J. A theoretical entropy score as a single value to express inhibitor selectivity. *BMC Bioinformatics* **12**, 94, doi:10.1186/1471-2105-12-94 (2011).

194 Uitdehaag, J. C. *et al.* A guide to picking the most selective kinase inhibitor tool compounds for pharmacological validation of drug targets. *Br J Pharmacol* **166**, 858-876, doi:10.1111/j.1476-5381.2012.01859.x (2012).

195 Graczyk, P. P. Gini coefficient: a new way to express selectivity of kinase inhibitors against a family of kinases. *J Med Chem* **50**, 5773-5779, doi:10.1021/jm070562u (2007).

196 Cheng, A. C., Eksterowicz, J., Geuns-Meyer, S. & Sun, Y. Analysis of kinase inhibitor selectivity using a thermodynamics-based partition index. *J Med Chem* **53**, 4502-4510, doi:10.1021/jm100301x (2010).

197 Cohen, P. & Alessi, D. R. Kinase drug discovery--what's next in the field? *ACS Chem Biol* **8**, 96-104, doi:10.1021/cb300610s (2013).

198 Levitzki, A. Tyrosine kinase inhibitors: views of selectivity, sensitivity, and clinical performance. *Annu Rev Pharmacol Toxicol* **53**, 161-185, doi:10.1146/annurev-pharmtox-011112-140341 (2013).

199 Edwards, A. M. *et al.* Too many roads not taken. *Nature* **470**, 163-165, doi:10.1038/470163a (2011).

200 Lee, J. & Bogyo, M. Target deconvolution techniques in modern phenotypic profiling. *Curr Opin Chem Biol* **17**, 118-126, doi:10.1016/j.cbpa.2012.12.022 (2013).

201 Wang, Y. & Ma, H. Protein kinase profiling assays: a technology review. *Drug Discov Today Technol* **18**, 1-8, doi:10.1016/j.ddtec.2015.10.007 (2015).

202 Raida, M. Drug target deconvolution by chemical proteomics. *Curr Opin Chem Biol* **15**, 570-575, doi:10.1016/j.cbpa.2011.06.016 (2011).

203 Cravatt, B. F., Simon, G. M. & Yates, J. R., 3rd. The biological impact of mass-spectrometry-based proteomics. *Nature* **450**, 991-1000, doi:10.1038/nature06525 (2007).

204 Schirle, M., Bantscheff, M. & Kuster, B. Mass spectrometry-based proteomics in preclinical drug discovery. *Chem Biol* **19**, 72-84, doi:10.1016/j.chembiol.2012.01.002 (2012).

205 Daub, H. Quantitative proteomics of kinase inhibitor targets and mechanisms. *ACS Chem Biol* **10**, 201-212, doi:10.1021/cb5008794 (2015).

206 Rix, U. & Superti-Furga, G. Target profiling of small molecules by chemical proteomics. *Nat Chem Biol* **5**, 616-624, doi:10.1038/nchembio.216 (2009).

207 Bantscheff, M., Scholten, A. & Heck, A. J. Revealing promiscuous drug-target interactions by chemical proteomics. *Drug Discov Today* **14**, 1021-1029, doi:10.1016/j.drudis.2009.07.001 (2009).

208 Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347-355, doi:10.1038/nature19949 (2016).

209 Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5**, 699-711, doi:10.1038/nrm1468 (2004).

210 Giansanti, P., Tsiatsiani, L., Low, T. Y. & Heck, A. J. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat Protoc* **11**, 993-1006, doi:10.1038/nprot.2016.057 (2016).

211 Tsiatsiani, L. & Heck, A. J. Proteomics beyond trypsin. *FEBS J* **282**, 2612-2626, doi:10.1111/febs.13287 (2015).

212 Swaney, D. L., Wenger, C. D. & Coon, J. J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J Proteome Res* **9**, 1323-1329, doi:10.1021/pr900863u (2010).

213 Zhou, F., Sikorski, T. W., Ficarro, S. B., Webber, J. T. & Marto, J. A. Online nanoflow reversed phase-strong anion exchange-reversed phase liquid chromatography-tandem mass spectrometry platform for efficient and in-depth proteome sequence analysis of complex organisms. *Anal Chem* **83**, 6996-7005, doi:10.1021/ac200639v (2011).

214 Manadas, B., Mendes, V. M., English, J. & Dunn, M. J. Peptide fractionation in proteomics approaches. *Expert Rev Proteomics* **7**, 655-663, doi:10.1586/epr.10.46 (2010).

215 Mallick, P. & Kuster, B. Proteomics: a pragmatic perspective. *Nat Biotechnol* **28**, 695-709, doi:10.1038/nbt.1658 (2010).

216 Ritorto, M. S., Cook, K., Tyagi, K., Pedrioli, P. G. & Trost, M. Hydrophilic strong anion exchange (hSAX) chromatography for highly orthogonal peptide separation of complex proteomes. *J Proteome Res* **12**, 2449-2457, doi:10.1021/pr301011r (2013).

217 Yu, P. *et al.* Trimodal Mixed Mode Chromatography That Enables Efficient Offline Two-Dimensional Peptide Fractionation for Proteome Analysis. *Anal Chem*, doi:10.1021/acs.analchem.7b01356 (2017).

218 Ruprecht, B. *et al.* Comprehensive and reproducible phosphopeptide enrichment using iron immobilized metal ion affinity chromatography (Fe-IMAC) columns. *Mol Cell Proteomics* **14**, 205-215, doi:10.1074/mcp.M114.043109 (2015).

219 Winkler, C., Denker, K., Wortelkamp, S. & Sickmann, A. Silver- and Coomassie-staining protocols: detection limits and compatibility with ESI MS. *Electrophoresis* **28**, 2095-2099, doi:10.1002/elps.200600670 (2007).

220 Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**, 2856-2860, doi:10.1038/nprot.2006.468 (2006).

221 Savaryn, J. P., Toby, T. K. & Kelleher, N. L. A researcher's guide to mass spectrometry-based proteomics. *Proteomics* **16**, 2435-2443, doi:10.1002/pmic.201600113 (2016).

222 Michalski, A. *et al.* Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Mol Cell Proteomics* **11**, O111 013698, doi:10.1074/mcp.O111.013698 (2012).

223 Pachl, F., Ruprecht, B., Lemeer, S. & Kuster, B. Characterization of a high field Orbitrap mass spectrometer for proteome analysis. *Proteomics* **13**, 2552-2562, doi:10.1002/pmic.201300076 (2013).

224 Nilsson, T. *et al.* Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods* **7**, 681-685, doi:10.1038/nmeth0910-681 (2010).

225 Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71 (1989).

226 Wilm, M. & Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal Chem* **68**, 1-8 (1996).

227 Wilm, M. S. & Mann, M. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int J Mass Spectrom Ion Processes* **136**, 167-180, doi:doi.org/10.1016/0168-1176(94)04024-9 (1994).

228 Karas, M., Bahr, U. & Dulcks, T. Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. *Fresenius J Anal Chem* **366**, 669-676 (2000).

229 Wu, X., Oleschuk, R. D. & Cann, N. M. Characterization of microstructured fibre emitters: in pursuit of improved nano electrospray ionization performance. *Analyst* **137**, 4150-4161, doi:10.1039/c2an35249d (2012).

230 Kebarle, P. & Verkerk, U. H. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev* **28**, 898-917, doi:10.1002/mas.20247 (2009).

231 Cech, N. B. & Enke, C. G. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom Rev* **20**, 362-387, doi:10.1002/mas.10008 (2001).

232 Hahne, H. *et al.* DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat Methods* **10**, 989-991, doi:10.1038/nmeth.2610 (2013).

233 Makarov, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* **72**, 1156-1162 (2000).

234 Perry, R. H., Cooks, R. G. & Noll, R. J. Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom Rev* **27**, 661-699, doi:10.1002/mas.20186 (2008).

235 Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *J Mass Spectrom* **40**, 430-443, doi:10.1002/jms.856 (2005).

236 Scigelova, M., Hornshaw, M., Giannakopulos, A. & Makarov, A. Fourier transform mass spectrometry. *Mol Cell Proteomics* **10**, M111 009431, doi:10.1074/mcp.M111.009431 (2011).

237 Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* **11**, 601, doi:10.1002/bms.1200111109 (1984).

238 McLafferty, F. W. Tandem mass spectrometry. *Science* **214**, 280-287 (1981).

239 Olsen, J. V. *et al.* Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* **4**, 709-712, doi:10.1038/nmeth1060 (2007).

240 Michalski, A., Cox, J. & Mann, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J Proteome Res* **10**, 1785-1793, doi:10.1021/pr101060v (2011).

241 Nesvizhskii, A. I., Vitek, O. & Aebersold, R. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Methods* **4**, 787-797, doi:10.1038/nmeth1088 (2007).

242 Tabb, D. L. *et al.* Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* **9**, 761-776, doi:10.1021/pr9006365 (2010).

243 Nesvizhskii, A. I. & Aebersold, R. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug Discov Today* **9**, 173-181, doi:10.1016/S1359-6446(03)02978-7 (2004).

244 Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **75**, 4646-4658 (2003).

245 Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383-5392 (2002).

246 Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4**, 207-214, doi:10.1038/nmeth1019 (2007).

247 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319, doi:10.1038/nprot.2016.136 (2016).

248 Jeong, K., Kim, S. & Bandeira, N. False discovery rates in spectral identification. *BMC Bioinformatics* **13 Suppl 16**, S2, doi:10.1186/1471-2105-13-S16-S2 (2012).

249 Nesvizhskii, A. I. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *J Proteomics* **73**, 2092-2123, doi:10.1016/j.jprot.2010.08.009 (2010).

250 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526, doi:10.1074/mcp.M113.031591 (2014).

251 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).

252 Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**, 1794-1805, doi:10.1021/pr101065j (2011).

253 Bantscheff, M., Lemeer, S., Savitski, M. M. & Kuster, B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem* **404**, 939-965, doi:10.1007/s00216-012-6203-4 (2012).

254 Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* **389**, 1017-1031, doi:10.1007/s00216-007-1486-6 (2007).

255 Grossmann, J. *et al.* Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *J Proteomics* **73**, 1740-1746, doi:10.1016/j.jprot.2010.05.011 (2010).

256 Chelius, D. & Bondarenko, P. V. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* **1**, 317-323 (2002).

257 Bondarenko, P. V., Chelius, D. & Shaler, T. A. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal Chem* **74**, 4741-4749 (2002).

258 Wang, J. *et al.* Target identification of natural and traditional medicines with quantitative chemical proteomics approaches. *Pharmacol Ther* **162**, 10-22, doi:10.1016/j.pharmthera.2016.01.010 (2016).

259 Axelsson, H., Almqvist, H., Seashore-Ludlow, B. & Lundback, T. in *Assay Guidance Manual* (eds G. S. Sittampalam *et al.*) (2004).

260 Martinez Molina, D. & Nordlund, P. The Cellular Thermal Shift Assay: A Novel Biophysical Assay for In Situ Drug Target Engagement and Mechanistic Biomarker Studies. *Annu Rev Pharmacol Toxicol* **56**, 141-161, doi:10.1146/annurev-pharmtox-010715-103715 (2016).

261 Savitski, M. M. *et al.* Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784, doi:10.1126/science.1255784 (2014).

262 Heal, W. P., Dang, T. H. & Tate, E. W. Activity-based probes: discovering new biology and new drug targets. *Chem Soc Rev* **40**, 246-257, doi:10.1039/c0cs00004c (2011).

263 Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383-414, doi:10.1146/annurev.biochem.75.101304.124125 (2008).

264 Lanning, B. R. *et al.* A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat Chem Biol* **10**, 760-767, doi:10.1038/nchembio.1582 (2014).

265 Leung, D., Hardouin, C., Boger, D. L. & Cravatt, B. F. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. *Nat Biotechnol* **21**, 687-691, doi:10.1038/nbt826 (2003).

266 Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* **96**, 14694-14699 (1999).

267 Ge, X. & Sem, D. S. A chemical proteomic probe for detecting dehydrogenases: catechol rhodanine. *Methods Mol Biol* **803**, 55-64, doi:10.1007/978-1-61779-364-6_5 (2012).

268 Staub, I. & Sieber, S. A. Beta-lactam probes as selective chemical-proteomic tools for the identification and functional characterization of resistance associated enzymes in MRSA. *J Am Chem Soc* **131**, 6271-6276, doi:10.1021/ja901304n (2009).

269 Kumar, S. *et al.* Activity-based probes for protein tyrosine phosphatases. *Proc Natl Acad Sci U S A* **101**, 7943-7948, doi:10.1073/pnas.0402323101 (2004).

270 Hatanaka, Y. Development and leading-edge application of innovative photoaffinity labeling. *Chem Pharm Bull (Tokyo)* **63**, 1-12, doi:10.1248/cpb.c14-00645 (2015).

271 Smith, E. & Collins, I. Photoaffinity labeling in target- and binding-site identification. *Future Med Chem* **7**, 159-183, doi:10.4155/fmc.14.152 (2015).

272 Bantscheff, M. *et al.* Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat Biotechnol* **29**, 255-265, doi:10.1038/nbt.1759 (2011).

273 Médard, G. *et al.* Optimized chemical proteomics assay for kinase inhibitor profiling. *J Proteome Res* **14**, 1574-1586, doi:10.1021/pr5012608 (2015).

274 Klaeger, S. *et al.* Chemical Proteomics Reveals Ferrochelatase as a Common Off-target of Kinase Inhibitors. *ACS Chem Biol* **11**, 1245-1254, doi:10.1021/acschembio.5b01063 (2016).

275 Munoz, L. Non-kinase targets of protein kinase inhibitors. *Nat Rev Drug Discov* **16**, 424-440, doi:10.1038/nrd.2016.266 (2017).

276 Patricelli, M. P. *et al.* In situ kinase profiling reveals functionally relevant properties of native kinases. *Chem Biol* **18**, 699-710, doi:10.1016/j.chembiol.2011.04.011 (2011).

277 Patricelli, M. P. *et al.* Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **46**, 350-358, doi:10.1021/bi062142x (2007).

278 Lemeer, S., Zorgiebel, C., Ruprecht, B., Kohl, K. & Kuster, B. Comparing immobilized kinase inhibitors and covalent ATP probes for proteomic profiling of kinase expression and drug selectivity. *J Proteome Res* **12**, 1723-1731, doi:10.1021/pr301073j (2013).

279 Sharma, K. *et al.* Proteomics strategy for quantitative protein interaction profiling in cell extracts. *Nat Methods* **6**, 741-744, doi:10.1038/nmeth.1373 (2009).

280 Cheng, Y. & Prusoff, W. H. Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **22**, 3099-3108 (1973).

281 Copeland, R. A., Pompliano, D. L. & Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* **5**, 730-739, doi:10.1038/nrd2082 (2006).

282 Corzo, J. Time, the forgotten dimension of ligand binding teaching. *Biochem Mol Biol Educ* **34**, 413-416, doi:10.1002/bmb.2006.494034062678 (2006).

283 Ferruz, N. & De Fabritiis, G. Binding Kinetics in Drug Discovery. *Mol Inform* **35**, 216-226, doi:10.1002/minf.201501018 (2016).

284 Guo, D., Heitman, L. H. & AP, I. J. The Role of Target Binding Kinetics in Drug Discovery. *ChemMedChem* **10**, 1793-1796, doi:10.1002/cmdc.201500310 (2015).

285 Pan, A. C., Borhani, D. W., Dror, R. O. & Shaw, D. E. Molecular determinants of drug-receptor binding kinetics. *Drug Discov Today* **18**, 667-673, doi:10.1016/j.drudis.2013.02.007 (2013).

286 Wood, E. R. *et al.* A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res* **64**, 6652-6659, doi:10.1158/0008-5472.CAN-04-1168 (2004).

287 Du, X. *et al.* Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci* **17**, doi:10.3390/ijms17020144 (2016).

288 Schauperl, M. *et al.* Binding Pose Flip Explained via Enthalpic and Entropic Contributions. *J Chem Inf Model* **57**, 345-354, doi:10.1021/acs.jcim.6b00483 (2017).

289 Motulsky, H. J. & Mahan, L. C. The kinetics of competitive radioligand binding predicted by the law of mass action. *Mol Pharmacol* **25**, 1-9 (1984).

290 Fleck, B. A., Hoare, S. R., Pick, R. R., Bradbury, M. J. & Grigoriadis, D. E. Binding kinetics redefine the antagonist pharmacology of the corticotropin-releasing factor type 1 receptor. *J Pharmacol Exp Ther* **341**, 518-531, doi:10.1124/jpet.111.188714 (2012).

291 Ritz, C. & Streibig, J. C. Bioassay Analysis using R. *J. Statist. Software* **12** (2005).

292 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29, doi:10.1038/75556 (2000).

293 Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190, doi:10.1101/gr.849004 (2004).

294 Moghaddas Gholami, A. *et al.* Global proteome analysis of the NCI-60 cell line panel. *Cell Rep* **4**, 609-620, doi:10.1016/j.celrep.2013.07.018 (2013).

295 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

296 Gande, S. L. *et al.* Expression and purification of EPHA2 tyrosine kinase domain for crystallographic and NMR studies. *Chembiochem*, 1439-7633, doi:10.1002/cbic.201600483 (2016).

297 Mueller, U. *et al.* The macromolecular crystallography beamlines at BESSY II of the Helmholtz Zentrum Berlin: Current status and perspectives. *Eur. Phys. J. Plus* **130**, 141-150 (2015).

298 Cianci, M. *et al.* P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J Synchrotron Radiat* **24**, 323-332, doi:10.1107/S1600577516016465 (2017).

299 Krug, M., Weiss, M., Heinemann, U. & Mueller, U. XDSAPP: a graphical user interace for the convenient processing of diffraction data using XDS. *J. Appl. Cryst.*, 568-572 (2012).

300 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).

301 Heinzlmeir, S. *et al.* Chemical Proteomics and Structural Biology Define EPHA2 Inhibition by Clinical Kinase Drugs. *ACS Chem Biol* **11**, 3400-3411, doi:10.1021/acschembio.6b00709 (2016).

302 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).

303 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-367, doi:10.1107/S0907444912001308 (2012).

304 Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786, doi:10.1021/ci200227u (2011).

305 Canevari, G. *et al.* Structural insight into maternal embryonic leucine zipper kinase (MELK) conformation and inhibition toward structure-based drug design. *Biochemistry* **52**, 6380-6387, doi:10.1021/bi4005864 (2013).

306 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).

307 Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271-281, doi:10.1107/S0907444910048675 (2011).

308 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).

309 Evans, P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82, doi:10.1107/S0907444905036693 (2006).

310 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255, doi:10.1107/S0907444996012255 (1997).

311 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

312 Vizcaino, J. A. *et al.* The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* **41**, D1063-1069, doi:10.1093/nar/gks1262 (2013).

313 Creixell, P. *et al.* Unmasking Determinants of Specificity in the Human Kinome. *Cell* **163**, 187-201, doi:10.1016/j.cell.2015.08.057 (2015).

314 Lafleur, K., Huang, D., Zhou, T., Caflisch, A. & Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J Med Chem* **52**, 6433-6446, doi:10.1021/jm9009444 (2009).

315 Moasser, M. M., Srethapakdi, M., Sachar, K. S., Kraker, A. J. & Rosen, N. Inhibition of Src kinases by a selective tyrosine kinase inhibitor causes mitotic arrest. *Cancer Res* **59**, 6145-6152 (1999).

316 Kaiser, A. *et al.* The cyclin-dependent kinase (CDK) inhibitor flavopiridol inhibits glycogen phosphorylase. *Arch Biochem Biophys* **386**, 179-187, doi:10.1006/abbi.2000.2220 (2001).

317 Simon, G. M., Niphakis, M. J. & Cravatt, B. F. Determining target engagement in living systems. *Nat Chem Biol* **9**, 200-205, doi:10.1038/nchembio.1211 (2013).

318 Posy, S. L. *et al.* Trends in kinase selectivity: insights for target class-focused library screening. *J Med Chem* **54**, 54-66, doi:10.1021/jm101195a (2011).

319 Huang, D., Zhou, T., Lafleur, K., Nevado, C. & Caflisch, A. Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* **26**, 198-204, doi:10.1093/bioinformatics/btp650 (2010).

320 Vin, H. *et al.* Sorafenib suppresses JNK-dependent apoptosis through inhibition of ZAK. *Mol Cancer Ther* **13**, 221-229, doi:10.1158/1535-7163.MCT-13-0561 (2014).

321 Kandouz, M., Haidara, K., Zhao, J., Brisson, M. L. & Batist, G. The EphB2 tumor suppressor induces autophagic cell death via concomitant activation of the ERK1/2 and PI3K pathways. *Cell Cycle* **9**, 398-407 (2010).

322 Hernandez-Florez, D. & Valor, L. Protein-kinase inhibitors: A new treatment pathway for autoimmune and inflammatory diseases? *Reumatol Clin* **12**, 91-99, doi:10.1016/j.reuma.2015.06.004 (2015).

323 Bavetsias, V. & Linardopoulos, S. Aurora Kinase Inhibitors: Current Status and Outlook. *Front Oncol* **5**, 278, doi:10.3389/fonc.2015.00278 (2015).

324 Takeuchi, S. *et al.* Dual inhibition of Met kinase and angiogenesis to overcome HGF-induced EGFR-TKI resistance in EGFR mutant lung cancer. *Am J Pathol* **181**, 1034-1043, doi:10.1016/j.ajpath.2012.05.023 (2012).

325 Qian, F. *et al.* Inhibition of tumor cell growth, invasion, and metastasis by EXEL-2880 (XL880, GSK1363089), a novel inhibitor of HGF and VEGF receptor tyrosine kinases. *Cancer Res* **69**, 8009-8016, doi:10.1158/0008-5472.CAN-08-4889 (2009).

326 Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol Cell* **3**, 629-638 (1999).

327 Levinson, N. M. & Boxer, S. G. A conserved water-mediated hydrogen bond network defines bosutinib's kinase selectivity. *Nat Chem Biol* **10**, 127-132, doi:10.1038/nchembio.1404 (2014).

328 Fang, W. B., Brantley-Sieders, D. M., Hwang, Y., Ham, A. J. & Chen, J. Identification and functional analysis of phosphorylated tyrosine residues within EphA2 receptor tyrosine kinase. *J Biol Chem* **283**, 16017-16026, doi:10.1074/jbc.M709934200 (2008).

329 Khorashad, J. S. *et al.* BCR-ABL1 compound mutations in tyrosine kinase inhibitor-resistant CML: frequency and clonal relationships. *Blood* **121**, 489-498, doi:10.1182/blood-2012-05-431379 (2013).

330 Parker, W. T. *et al.* The impact of multiple low-level BCR-ABL1 mutations on response to ponatinib. *Blood* **127**, 1870-1880, doi:10.1182/blood-2015-09-666214 (2016).

331 Duong-Ly, K. C. *et al.* Kinase Inhibitor Profiling Reveals Unexpected Opportunities to Inhibit Disease-Associated Mutant Kinases. *Cell Rep* **14**, 772-781, doi:10.1016/j.celrep.2015.12.080 (2016).

332 Ji, W. *et al.* OTSSP167 Abrogates Mitotic Checkpoint through Inhibiting Multiple Mitotic Kinases. *PLoS One* **11**, e0153518, doi:10.1371/journal.pone.0153518 (2016).

333 Toure, B. B. *et al.* Toward the Validation of Maternal Embryonic Leucine Zipper Kinase: Discovery, Optimization of Highly Potent and Selective Inhibitors, and Preliminary Biology Insight. *J Med Chem* **59**, 4711-4723, doi:10.1021/acs.jmedchem.6b00052 (2016).

334 Chung, S. *et al.* Preclinical evaluation of biomarkers associated with antitumor activity of MELK inhibitor. *Oncotarget*, doi:10.18632/oncotarget.7685 (2016).

335 Bamborough, P., Drewry, D., Harper, G., Smith, G. K. & Schneider, K. Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *J Med Chem* **51**, 7898-7914, doi:10.1021/jm8011036 (2008).

336 Bardelle, C. *et al.* Inhibitors of the tyrosine kinase EphB4. Part 3: identification of non-benzodioxole-based kinase inhibitors. *Bioorg Med Chem Lett* **20**, 6242-6245, doi:10.1016/j.bmcl.2010.08.100 (2010).

337 Das, J. *et al.* 2-aminothiazole as a novel kinase inhibitor template. Structure-activity relationship studies toward the discovery of N-(2-chloro-6-methylphenyl)-2-[[6-[4-(2-hydroxyethyl)-1- piperazinyl)]-2-methyl-4-pyrimidinyl]amino)]-1,3-thiazole-5-carboxamide (dasatinib, BMS-354825) as a potent pan-Src kinase inhibitor. *J Med Chem* **49**, 6819-6832, doi:10.1021/jm060727j (2006).

338 Novak, A., Humphreys, L. D., Walker, M. D. & Woodward, S. Amide bond formation using an air-stable source of $AlMe_3$. *Tetrahedron Lett* **47**, 5767-5769 (2006).

339 Glynn, D., Bernier, D. & Woodward, S. Microwave acceleration in DABAL-Me_3-mediated amide formation. *Tetrahedron Lett* **49**, 5687-5688 (2008).

340 Zhou, P., Zou, J., Tian, F. & Shang, Z. Fluorine bonding--how does it work in protein-ligand interactions? *J Chem Inf Model* **49**, 2344-2355, doi:10.1021/ci9002393 (2009).

341 Metz, J. T. & Hajduk, P. J. Rational approaches to targeted polypharmacology: creating and navigating protein-ligand interaction networks. *Curr Opin Chem Biol* **14**, 498-504, doi:10.1016/j.cbpa.2010.06.166 (2010).

342 Knight, Z. A., Lin, H. & Shokat, K. M. Targeting the cancer kinome through polypharmacology. *Nat Rev Cancer* **10**, 130-137, doi:10.1038/nrc2787 (2010).

343 Anighoro, A., Bajorath, J. & Rastelli, G. Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* **57**, 7874-7887, doi:10.1021/jm5006463 (2014).

344 Larry Jameson, J. & Longo, D. L. Precision Medicine—Personalized, Problematic, and Promising. *Obstetrical & Gynecological Survey* **70**, 612-614, doi:10.1097/01.ogx.0000472121.21647.38 (2015).

345 Hollingsworth, S. J. Precision medicine in oncology drug development: a pharma perspective. *Drug Discov Today* **20**, 1455-1463, doi:10.1016/j.drudis.2015.10.005 (2015).

346 Blagg, J. & Workman, P. Choose and Use Your Chemical Probe Wisely to Explore Cancer Biology. *Cancer Cell* **32**, 9-25, doi:10.1016/j.ccell.2017.06.005 (2017).

347 Li, B., Liu, Y., Uno, T. & Gray, N. Creating chemical diversity to target protein kinases. *Comb Chem High Throughput Screen* **7**, 453-472 (2004).

348 Fedorov, O., Sundstrom, M., Marsden, B. & Knapp, S. Insights for the development of specific kinase inhibitors by targeted structural genomics. *Drug Discov Today* **12**, 365-372, doi:10.1016/j.drudis.2007.03.006 (2007).

349 Tonge, P. J. Drug-Target Kinetics in Drug Discovery. *ACS Chem Neurosci*, doi:10.1021/acschemneuro.7b00185 (2017).

350 Zhao, Q. *et al.* Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J Am Chem Soc* **139**, 680-685, doi:10.1021/jacs.6b08536 (2017).

351 Pereira, D. A. & Williams, J. A. Origin and evolution of high throughput screening. *Br J Pharmacol* **152**, 53-61, doi:10.1038/sj.bjp.0707373 (2007).

352 Saeidnia, S., Gohari, A. R. & Manayi, A. Reverse Pharmacognosy and Reverse Pharmacology; Two Closely Related Approaches for Drug Discovery Development. *Curr Pharm Biotechnol* **17**, 1016-1022 (2016).

353 Kuhn, P., Wilson, K., Patch, M. G. & Stevens, R. C. The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr Opin Chem Biol* **6**, 704-710 (2002).

354 Pouliot, M. & Jeanmart, S. Pan Assay Interference Compounds (PAINS) and Other Promiscuous Compounds in Antifungal Research. *J Med Chem* **59**, 497-503, doi:10.1021/acs.jmedchem.5b00361 (2016).

355 Dunlop, K. V., Irvin, R. T. & Hazes, B. Pros and cons of cryocrystallography: should we also collect a room-temperature data set? *Acta Crystallogr D Biol Crystallogr* **61**, 80-87, doi:10.1107/S0907444904027179 (2005).

356 Kornev, A. P. & Taylor, S. S. Dynamics-Driven Allostery in Protein Kinases. *Trends Biochem Sci* **40**, 628-647, doi:10.1016/j.tibs.2015.09.002 (2015).

357 Zheng, W., Thorne, N. & McKew, J. C. Phenotypic screens as a renewed approach for drug discovery. *Drug Discov Today* **18**, 1067-1073, doi:10.1016/j.drudis.2013.07.001 (2013).

358 Gujral, T. S., Peshkin, L. & Kirschner, M. W. Exploiting polypharmacology for drug target deconvolution. *Proc Natl Acad Sci U S A* **111**, 5048-5053, doi:10.1073/pnas.1403080111 (2014).

359 Smith, C. Drug target validation: Hitting the target. *Nature* **422**, 341, 343, 345 passim, doi:10.1038/422341a (2003).

360 Bunnage, M. E., Gilbert, A. M., Jones, L. H. & Hett, E. C. Know your target, know your molecule. *Nat Chem Biol* **11**, 368-372, doi:10.1038/nchembio.1813 (2015).

361 Blake, R. A. Target validation in drug discovery. *Methods Mol Biol* **356**, 367-377 (2007).

362 Toure, M. & Crews, C. M. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angew Chem Int Ed Engl* **55**, 1966-1973, doi:10.1002/anie.201507978 (2016).

363 Han, X. *et al.* Discovery and characterization of a novel cyclic peptide that effectively inhibits ephrin binding to the EphA4 receptor and displays anti-angiogenesis activity. *PLoS One* **8**, e80183, doi:10.1371/journal.pone.0080183 (2013).

364 Timmerman, P., Puijk, W. C. & Meloen, R. H. Functional reconstruction and synthetic mimicry of a conformational epitope using CLIPS technology. *J Mol Recognit* **20**, 283-299, doi:10.1002/jmr.846 (2007).

365 Mercurio, F. A. *et al.* Targeting EphA2-Sam and Its Interactome: Design and Evaluation of Helical Peptides Enriched in Charged Residues. *Chembiochem* **17**, 2179-2188, doi:10.1002/cbic.201600413 (2016).

366 Seisel, Q., Radisch, M., Gill, N. P., Madden, D. R. & Boisguerin, P. Optimization of the process of inverted peptides (PIPEPLUS) to screen PDZ domain ligands. *Bioorg Med Chem Lett* **27**, 3111-3116, doi:10.1016/j.bmcl.2017.05.045 (2017).

367 Choi, J. *et al.* Identification of small-molecule compounds targeting the dishevelled PDZ domain by virtual screening and binding studies. *Bioorg Med Chem* **24**, 3259-3266, doi:10.1016/j.bmc.2016.03.026 (2016).

368 Plowright, A. T. *et al.* Joining Forces: The Chemical Biology-Medicinal Chemistry Continuum. *Cell Chem Biol*, doi:10.1016/j.chembiol.2017.05.019 (2017).

Abbreviations

RTK	receptor tyrosine kinases
EPH	erythropoietin-producing hepatocellular carcinoma receptor kinases
EPHA2	ephrin type-A receptor 2
Ephrin	EPH receptor interacting protein
GPI	glycosylphosphatidylinositol
PDZ	PSD95/Dlg/ZO1
LBD	ligand binding domain
SAM	sterile alpha motif
GEF	guanine nucleotide exchange factor
EPEC	enteropathogenic *Escherichia coli*
CARs	chimeric antigen receptors
FDA	Food and Drug Administration
ATP	adenosine triphosphate
ADP	adenosine diphosphate
N-lobe	N-terminal lobe
C-lobe	C-terminal lobe
R-spine	regulatory spine
C-spine	catalytic spine
NMR	nuclear magnetic resonance
K_D	dissociation constant
LC	liquid chromatography
ESI	electrospray ionization
LTQ	linear (2D) ion trap
HCD	higher-energy collisional dissociation
PSM	peptide-spectrum-match
FDR	false discovery rate
LFQ	label-free quantification
CETSA	cellular thermal shift assay
ABPP	activity-based protein profiling
PAL	photoaffinity labeling
EC_{50}	half maximal effective concentration
K_D^{app}	apparent dissociation constant
k_{on}	association rate
k_{off}	dissociation rate
CATDS	Concentration And Target Dependent Selectivity

Proteins and gene names are based on UniProt and HUGO accessions

“The most memorable people in life will be the friends who loved you when you weren’t very lovable”

-Aidan Chambers-

Acknowledgement

Da liegt es nun: dieses Stückchen Text – und kommt so unscheinbar daher, obwohl es so viel Zeit, Energie, und Schweiss gekostet hat. Leider darf ich nur meinen eigenen Namen auf das Titelblatt schreiben, aber in Wirklichkeit haben sehr viel mehr Menschen auf vielfältigste Weise dazu beigetragen und mir dabei geholfen dieses Kapitel meines Lebens zu meistern. Ich bin sehr sehr dankbar und glücklich, dass ich in dieser Zeit so viele helfende Hände, lächelnde Gesichter, aufmunternde Worte und Zuneigung erfahren durfte und mich so viele und so besondere Menschen auf meinem Weg begleiten.

Ganz besonders möchte ich meiner Familie und Humberto danken! Ihr seid die, die in den letzten Jahren am meisten aushalten mussten: viel zu wenig Zeit, viel zu viel Stress, sämtliche Stimmungsschwankungen; und das unglaubliche daran ist, dass ihr das alles einfach mitgemacht habt und ihr mich trotzdem gern habt. Danke für eure unerschütterliche Zuversicht, dass ihr mich bis in die Nacht hinein mit Essen versorgt, dass ihr mir immer den Rücken frei haltet und dass ihr mich immer wieder erdet wenn es grad wieder zuviel wird. Zu sehen wie stolz ihr seid, macht auch mich stolz, und zwar vor allem darauf dass ich euch an meiner Seite habe! Das gilt auch für meine langjährigen Freunde, allen voran Maria, Biggi und Anne, die in den letzten Jahren mehr zurückstecken mussten als mir lieb ist.

Ein ganz besonders grosses Danke geht natürlich an dich, Bernhard! Ich schätze mich wirklich glücklich, dass ich diese Arbeit bei dir am Lehrstuhl schreiben durfte. All die Dinge, die ich von dir gelernt habe, und die Möglichkeiten, die du mir eröffnet hast, sind unbezahlbar. Ich bin zudem ernsthaft beeindruckt davon, dass du mir immer ein bisschen mehr zutraust als ich mir selber jemals zutrauen würde. Ohne dein Vertrauen und deine Unterstützung und das „mal-freundlich-aus-der-comfort-zone-rausschubsen“ wäre ich längst nicht so weit gekommen. Ich kann nur sagen, du bist in vielerlei Hinsicht ein Vorbild, nicht nur wissenschaftlich sondern auch menschlich. Danke!

Thank you, Guillaume! You have not just been an excellent supervisor; you have also been a wonderful person, the “clown in the lab” that cheered me up when things didn’t work, and a critical counterpart that helped me looking at things from different angles. You were the one that convinced me that I could do this weird chemical stuff and I really hope you didn’t regret this – even though you had a hard time teaching me chemistry over and over again. Thanks for being so patient, for explaining without making me feel stupid, and for being a very good friend at all times!

Danke auch euch, Susan und Mathias! Für laaaaange laaaaange Abende, von denen ich keinen einzigen missen möchte. Susan, was hatten wir für eine grossartige Zeit... Ich glaube ich werde für den Rest meines Lebens an deinen tanzenden Haarschopf hinter meinem Bildschirmrand denken, wenn ich Enrique Iglesias Lieder höre. Danke für die vielen Stunden, in denen wir miteinander gelacht, gefeiert aber auch geweint haben! Mathias, wo soll ich anfangen... Vielleicht bei den unzähligen Nächten, die du mit mir hier gesessen hast damit ich nicht nachts alleine hier arbeiten muss; oder dass du mir immer wieder beim „zu-Ende-denken“ hilfst und dass du figures, für die ich zwei Wochen brauche, in zwei Minuten generierst. Danke für all die grossen und kleinen Dinge, Gesten und Worte! Und Leute... ¡Viva la revolución!

Of course, I want to send a very big and warm „**THANK YOU**“ to all my current and former colleagues at the chair of proteomics and at BayBioMS and my students! I’m the luckiest person to not just go to work every day, but to be able to spend every day among my friends! Thank you, for answering all my weird “Steffi questions” and for patiently explaining everything, for your helping hands and your support when needed, for inspiring and also funny conversations, and especially for keeping this whole thing running and working. Thank you for generating such a fruitful, energetic and joyful environment. Thanks for having a great time and unforgettable moments also besides the lab: awesome barbecues, spontaneous 90ies parties, certain movie and cinema nights, sporting activities in the institute, amazing days in Brixen, Zugspitze, Heidelberg, Thessaloniki, Kopenhagen, Tegernsee and Boston. You are awesome!

Ich möchte mich zudem bei den Mitgliedern meines Prüfungskommittes Prof. Dr. Sewald, Prof. Dr. Hacker, Prof. Dr. Sieber und Prof. Dr. Klingenspor für die Bereitschaft bedanken meine Arbeit zu begutachten.

Danke.

Publication record

Main Publications of this thesis:

- Klaeger S*, **Heinzlmeir S***, Wilhelm M*, Polzer H, Vick B, König PA, Reinecke M, Ruprecht B, Petzoldt S, Meng C, Zecha J, Reiter K, Qiao H, Helm D, Koch H, Schoof M, Canevari G, Casale E, Re Depaolini S, Feuchtinger A, Wu Z, Schmidt T, Rueckert L, Becker W, Huenges J, Garz A, Gohlke B, Zolg DP, Kayser G, Vooder T, Preissner R, Hahne H, Tonisson N, Kramer K, Götze K, Bassermann F, Schlegl J, Ehrlich H, Aiche S, Walch A, Greif P, Schneider S, Felder ER, Ruland J, Médard G, Jeremias I, Spiekermann K, Kuster B, (2017) The target landscape of clinical kinase drugs. Science. 358(6367). pii: eaan4368
- **Heinzlmeir S**, Lohse J, Treiber T, Kudlinzki D, Linhard V, Gande SL, Sreeramulu S, Saxena K, Liu X, Wilhelm M, Schwalbe H, Kuster B, Médard G, (2017) Chemoproteomics-Aided Medicinal Chemistry for the Discovery of EPHA2 Inhibitors. ChemMedChem. 12(12):999-1011. doi: 10.1002/cmdc.201700217.
- **Heinzlmeir S***, Kudlinzki D*, Sreeramulu S, Klaeger S, Gande SL, Linhard V, Wilhelm M, Qiao H, Helm D, Ruprecht B, Saxena K, Médard G, Schwalbe H, Kuster B, (2016) Chemical Proteomics and Structural Biology Define EPHA2 Inhibition by Clinical Kinase Drugs. ACS Chem Biol. 11(12):3400-3411. doi: 10.1021/acschembio.6b00709

Additional publications during PhD:

- Frejno M, Zenzini Chiozzi R, Wilhelm M, Koch H, Zheng R, Klaeger; S, Ruprecht B, Meng C, Kramer K, Jarzab A, **Heinzlmeir S**, Johnstone E, Domingo E, Kerr D, Jesinghaus M, Slotta-Huspenina J, Weichert W, Knapp S, Feller SM, Kuster B, (2017) Pharmacoproteomic characterisation of human colon and rectal cancer. Mol Syst Biol. 13(11):951
- Gande SL, Saxena K, Sreeramulu S, Linhard V, Kudlinzki D, **Heinzlmeir S**, Reichert AJ, Skerra A, Kuster B, Schwalbe H, (2016) Expression and Purification of EPHA2 Tyrosine Kinase Domain for Crystallographic and NMR Studies. Chembiochem. 17(23):2257-2263. doi: 10.1002/cbic.201600483.
- Slynko I, Schmidtkunz K, Rumpf T, Klaeger S, **Heinzlmeir S**, Najar A, Metzger E, Kuster B, Schüle R, Jung M, Sippl W, (2016) Identification of Highly Potent Protein Kinase C-Related Kinase 1 Inhibitors by Virtual Screening, Binding Free Energy Rescoring, and in vitro Testing. ChemMedChem. 11(18):2084-94. doi: 10.1002/cmdc.201600284.
- Ostrovskyi D, Rumpf T, Eib J, Lumbroso A, Slynko I, Klaeger S, **Heinzlmeir S**, Forster M, Gehringer M, Pfaffenrot E, Bauer SM, Schmidtkunz K, Wenzler S, Metzger E, Kuster B, Laufer S, Schüle R, Sippl W, Breit B, Jung M, (2016) Tofacitinib and analogs as inhibitors of the histone kinase PRK1 (PKN1). Future Med Chem. 8(13):1537-51. doi: 10.4155/fmc-2016-0132.
- Klaeger S, Gohlke B, Perrin J, Gupta V, **Heinzlmeir S**, Helm D, Qiao H, Bergamini G, Handa H, Savitski MM, Bantscheff M, Médard G, Preissner R, Kuster B, (2016) Chemical Proteomics Reveals Ferrochelatase as a Common Off-target of Kinase Inhibitors. ACS Chem Biol. 11(5):1245-54. doi: 10.1021/acschembio.5b01063.
- Ruprecht B, Zecha J, **Heinzlmeir S**, Médard G, Lemeer S, Kuster B, (2015) Evaluation of Kinase Activity Profiling Using Chemical Proteomics. ACS Chem Biol. 10(12):2743-52. doi: 10.1021/acschembio.5b00616.

- Médard G, Pachl F, Ruprecht B, Klaeger S, **Heinzlmeir S**, Helm D, Qiao H, Ku X, Wilhelm M, Kuehne T, Wu Z, Dittmann A, Hopf C, Kramer K, Kuster B, (2015) Optimized chemical proteomics assay for kinase inhibitor profiling. J Proteome Res. 14(3):1574-86. doi: 10.1021/pr5012608.
- Mergner J, **Heinzlmeir S**, Kuster B, Schwechheimer C, (2015) DENEDDYLASE1 deconjugates NEDD8 from non-cullin protein substrates in Arabidopsis thaliana. Plant Cell. 27(3):741-53. doi: 10.1105/tpc.114.135996.
- Höfener M*, **Heinzlmeir S***, Kuster B, Sewald N, (2014) Probing SH2-domains using Inhibitor Affinity Purification (IAP). Proteome Sci. 12:41. doi: 10.1186/1477-5956-12-41.
- Ku X, **Heinzlmeir S**, Helm D, Médard G, Kuster B, (2014) New affinity probe targeting VEGF receptors for kinase inhibitor selectivity profiling by chemical proteomics. J Proteome Res. 13(5):2445-52. doi: 10.1021/pr401247t.
- Ku X, **Heinzlmeir S**, Liu X, Médard G, Kuster B, (2014) A new chemical probe for quantitative proteomic profiling of fibroblast growth factor receptor and its inhibitors. J Proteomics. 96:44-55. doi: 10.1016/j.jprot.2013.10.031.
- Hakenjos JP, Bejai S, Ranftl Q, Behringer C, Vlot AC, Absmanner B, Hammes U, **Heinzlmeir S**, Kuster B, Schwechheimer C, (2013) ML3 is a NEDD8- and ubiquitin-modified protein. Plant Physiol. 163(1):135-49. doi: 10.1104/pp.113.221341.

* Authors contributed equally to this work (underlined in case Stephanie Heinzlmeir was co-first author). Authors appear in the order of the original publication.

Patent application:

Heinzlmeir S, Médard G, Kuster B, Schwalbe H, Kudlinzki D. European Patent Application number EP17165807.3 (10 April, 2017), unpublished

Curriculum vitae

Personal details

Name	Stephanie Heinzlmeir
Date of birth	18.08.1986 in Dachau
Adress	Emil-Erlenmeyer-Forum 5, 85354 Freising
Email	stephanie.heinzlmeir@tum.de

Education

Since 12/2012	Doctoral candidate, Chair of Proteomics and Bioanalytics Technical University of Munich and German Cancer Consortium (DKTK) Thesis: "When Chemical Proteomics meets Medicinal Chemistry: Guided Drug Discovery towards EPHA2 inhibitors"
10/2010 – 11/2012	Master of Science (M.Sc.), Biochemistry Technical University of Munich Thesis: "Synthesis and Evaluation of a FGFR specific Kinobeads Probe for Small Molecule Inhibitor Drug Profiling" (Chair of Proteomics and Bioanalytics, Prof. Küster)
10/2006 – 09/2009	Bachelor of Science (B.Sc.), Biochemistry Technical University of Munich Thesis: "Construction and comparative *in vitro* analysis of membrane-targeted GFP-fusion proteins delivered by MVA vectors" (Insitute of Virology, Prof. Drexler)

Work and Research Experience

Since 01/2017	part-time support for head of division "Proteomics", Bavarian Biomolecular Mass Spectrometry Center (BayBioMS) Technical University of Munich
04/2016 – 12/2016	Parental leave substitution, head of division "Proteomics", Bavarian Biomolecular Mass Spectrometry Center (BayBioMS) Technical University of Munich
2009/2010/2011	short-term employments at Agrobiogen GmbH

Conference and workshop contributions

06/2013	1st DKTK Retreat, Heidelberg (poster)
05/2014	2nd DKTK Retreat, Heidelberg (poster)
08/2014	EMBO Chemical Biology 2014, Heidelberg (poster)
11/2014	ICBS 2014, San Francisco (poster)
01/2015	Status Seminar Chemical Biology 2015, Frankfurt a. Main (poster)
04/2015	COST conference Chemical proteomics, Oxford (poster)
07/2015	2nd DKTK Munich Cancer Retreat, Herrsching (poster)
10/2015	ICBS 2015, Berlin (speed talk, poster)
01/2016	Status Seminar Chemical Biology 2016, Frankfurt a. Main (speed talk, poster)
06/2016	3rd DKTK Munich Cancer Retreat, Herrsching (talk)
08/2016	EMBO Chemical Biology 2016, Heidelberg (talk)
10/2016	iNEXT user meeting, Madrid (invited talk)
12/2016	Helmholtz Drug Innovation in Academia, Heidelberg (poster)
02/2017	BayBioMS meets SFB proteomics course for plant scientists, Freising (organizer)
07/2017	4th DKTK Munich Cancer Retreat, Herrsching (talk)
09/2017	ETN MagicBullet Summer School, Como (invited workshop)

Appendix

Table of contents

1 Supplementary Figures IX
2 Supplementary Tables XIV
3 Supplementary Datasets XV

“Vui huift vui.”
- unknown-

1 Supplementary Figures

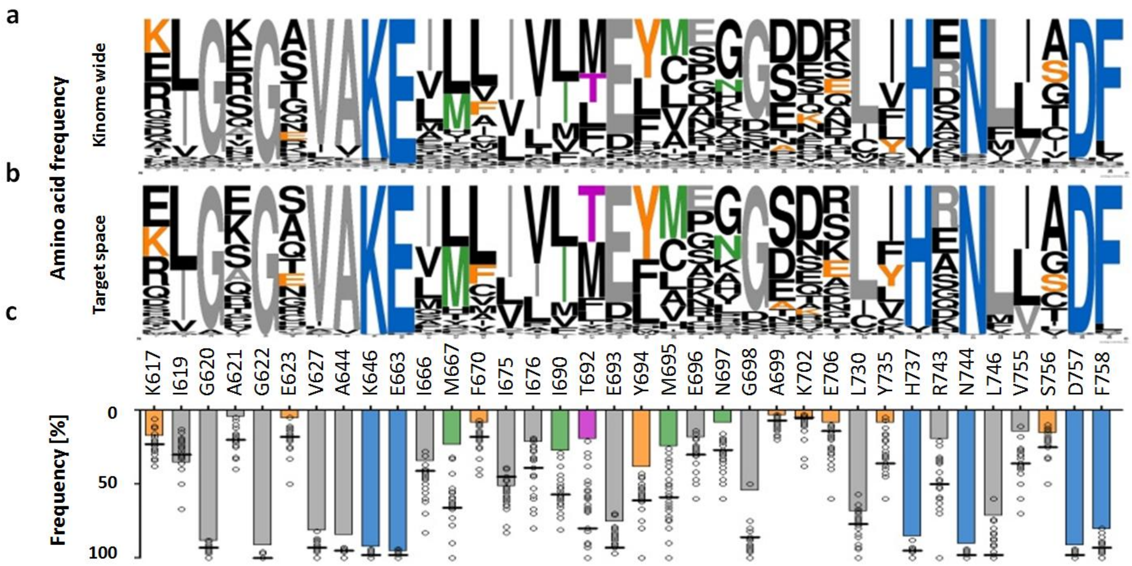


Figure S1 | Overrepresentation of EPHA2 amino acids within the target space of EPHA2 inhibitors. All 36 residues engaged in drug-protein interactions were categorized into five key residues (green and purple (gatekeeper), six potency residues (blue), nine selectivity residues (orange) and sixteen scaffold residues (grey). a) Sequence alignment of all human protein kinases illustrates the amino acid frequency at each identified position within the whole kinome (amino acids identified in EPHA2 highlighted in the respective colors). Key residues and selectivity residues were found to be rather underrepresented within the kinome wide distribution, whereas scaffold and potency residues were often highly conserved amino acids. b) Sequence alignment of the combined target spaces of the identified EPHA2 inhibitors showed that mainly key residues and to a lower extent also selectivity residues are enriched within the target space in comparison to the kinome-wide distribution. c) The kinome wide frequency of the amino acid found in EPHA2 (bar plot) was compared to the frequency of the same amino acid within the target space of each EPHA2 inhibitor (dot plot, each dot represents the target space of one inhibitor). The amino acids contained in EPHA2 were found to be overrepresented to different extents in comparison to the kinome.

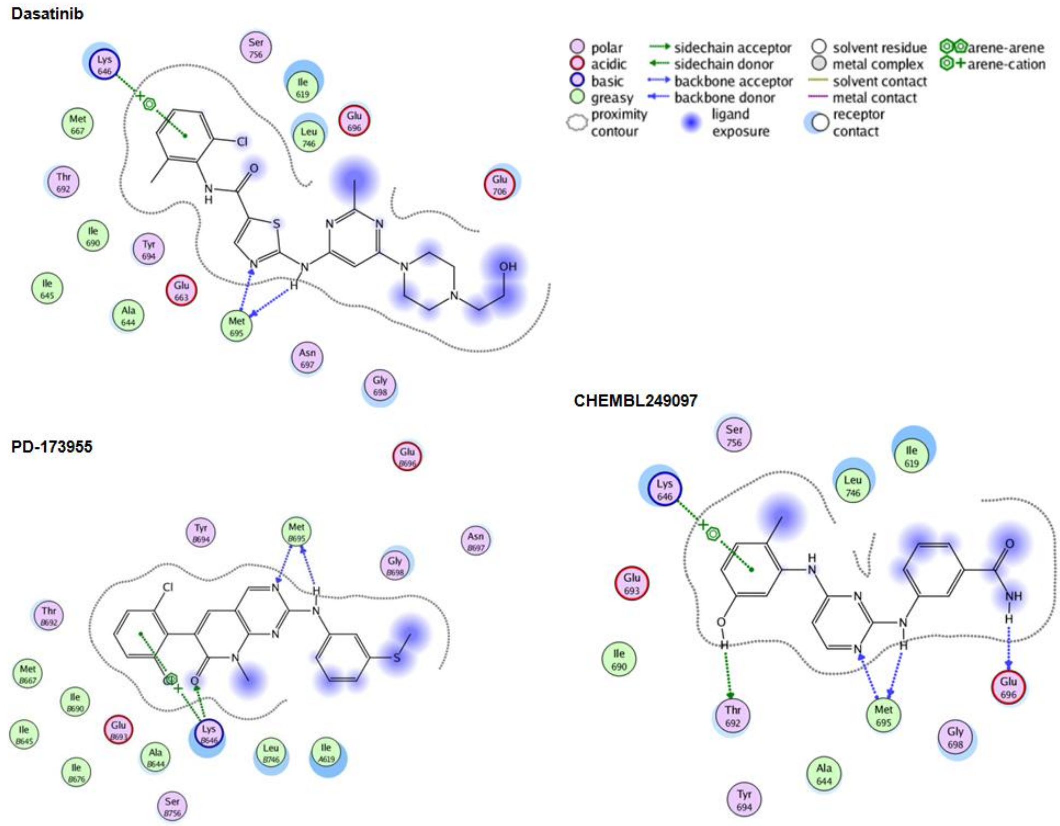

Dasatinib
PD-173955
CHEMBL249097
polar
acidic
basic
greasy
proximity contour
sidechain acceptor
sidechain donor
backbone acceptor
backbone donor
ligand exposure
solvent residue
metal complex
solvent contact
metal contact
receptor contact
arene-arene
arene-cation
Lys 646
Ser 756
Ile 619
Glu 696
Leu 746
Met 667
Thr 692
Glu 706
Ile 690
Tyr 694
Ile 645
Glu 663
Ala 644
Met 695
Asn 697
Gly 698
Glu 693
Cl
OH

Figure S2 | Docking studies of known EPH inhibitors. Dasatinib, CHEMBL249097 and PD-173955 were docked into EPHA2 kinase domain (PDB: 1MQB). Interaction analysis revealed that all inhibitors engaged direct interactions with Lys646 and Met695. Dasatinib's pyrimidine moiety did not form any additional interaction which infers that it could be replaced by an aryl as found in CHEMBL249097 and PD-173955. Interestingly, CHEMBL249097 targeted Glu696 via a direct side chain interaction between Glu696 and its terminal amide. This amide group was introduced in inhibitor design to establish this interaction also in our novel EPHA2 inhibitors.

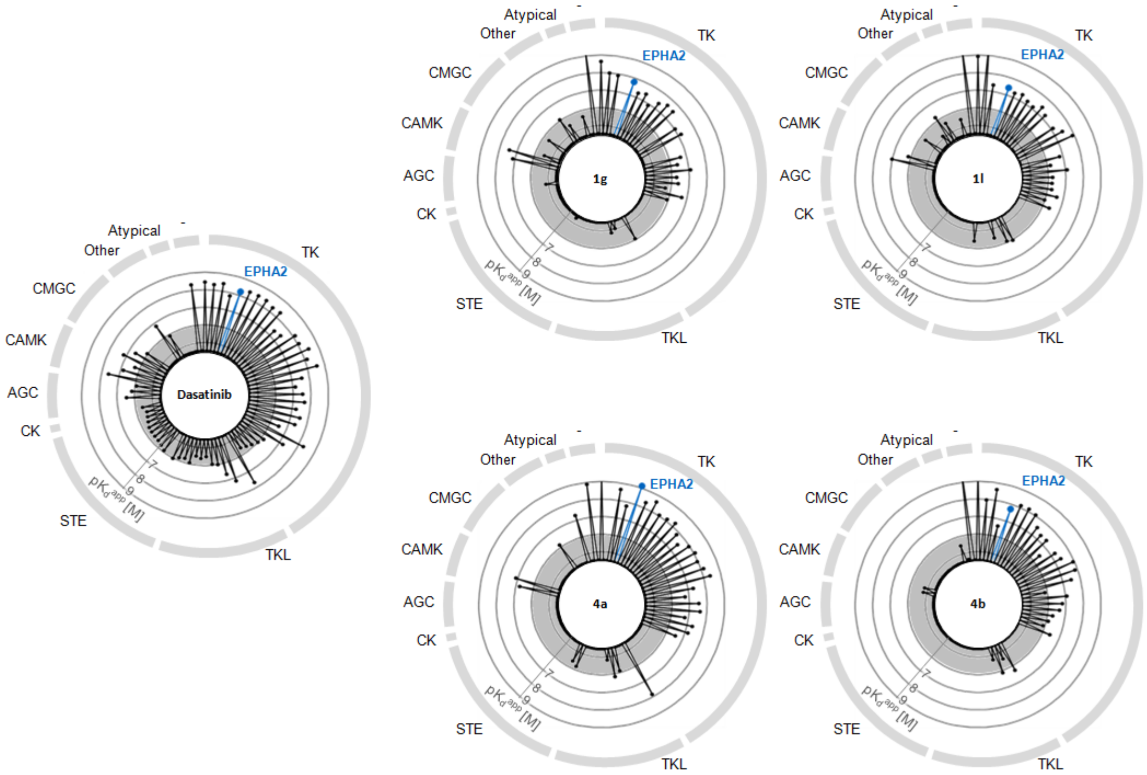


Figure S3 | Target profiles of Dasatinib and inhibitors 1g, 1l, 4a and 4b. Radar plot depicting the target space and binding affinities of inhibitors. Grey circle depicts affinity cutoff of 1µM.

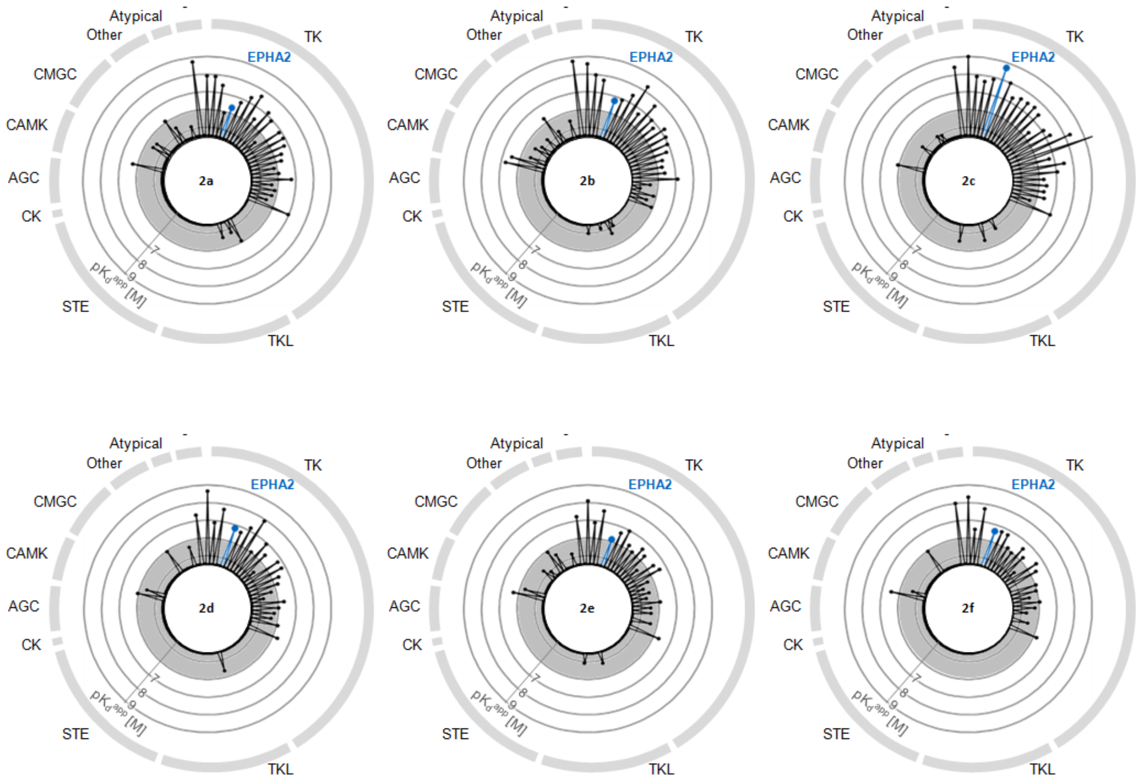


Figure S4 | Target profiles of Dasatinib and inhibitors 2a, 2b, 2c, 2d, 2e, 2f. Radar plot depicting the target space and binding affinities of inhibitors. Grey circle depicts affinity cutoff of 1µM.

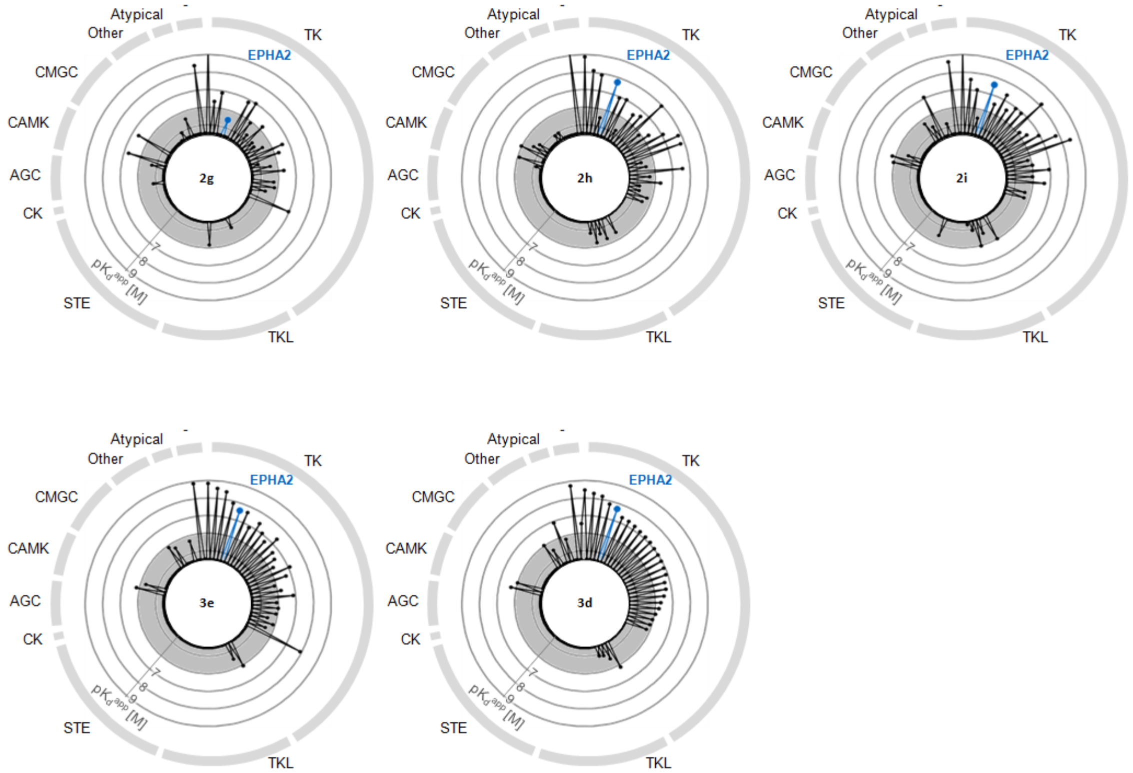


Figure S5 | Target profiles of Dasatinib and inhibitors 2g, 2h, 2i, 3e, 3d. Radar plot depicting the target space and binding affinities of inhibitors. Grey circle depicts affinity cutoff of 1µM.

2 Supplementary Tables

Table S1 | Selectivity profiling of 235 clinical inhibitors

a) List of all analyzed kinase drugs.

b) Overview over all 32 identified EPHA2 drugs and three tool compounds.

c) Single data points and detailed curve fit information for each drug-protein interaction.

d) Selectivity profiling of submicromolar clinical EPHA2 inhibitors and tool compounds.

Table S2 | Protein crystallography of clinical inhibitors

a) Overview drug-EPHA2 interactions.

b) Drug-protein interaction analysis of drug-EPHA2 crystals.

c) Overview reference structures.

d) Drug-protein interaction analysis of drug-EPHA2 crystals and reference structures.

e) Crystal structure data collection and refinement statistics.

Table S3 | Kinome-wide sequence alignment

a) Sequence Alignment, modified from Creixell *et al* (PubMed ID: 26388442).

b) Extraction of EPHA2 residues of interest.

Table S4 | Residue classification EPHA2

a) Residue classification EPHA2.

b) Frequency of EPHA2 amino acid.

Table S5 | Residue classification ABL1

a) Selectivity profiling of clinical ABL1 inhibitors.

b) Drug-protein interaction analysis of drug-ABL1 crystals.

c) Extraction of ABL1 residues of interest.

d) Frequency of ABL1 amino acid.

e) Overview drug-ABL1 interactions and residue classification.

Table S6 | Residue classification MELK

a) Selectivity profiling of clinical MELK inhibitors.

b) Overview drug-MELK interactions and residue classification.

c) Crystal structure data collection and refinement statistics.

Table S7 | Dedicated EPHA2 inhibitors

a) Selectivity profiles of dedicated EPHA2 inhibitors

b) Crystal structure data collection and refinement statistics.

3 Supplementary Datasets

Dataset S1 | Structure selectivity relationship analysis EPHA2

Dataset S2 | Structure selectivity relationships analysis ABL1

Dataset S3 | Structure selectivity relationships analysis MELK

“Discovery follows discovery,
Each both answering and raising questions,
Each ending a long search,
And each providing new instruments for a new search.”
- J. Robert Oppenheimer -