Technical University of Munich
TUM School of Life Sciences

TUM

Pushing the boundaries of affinity-based target deconvolution: Promiscuous immobilization of nature-made molecules and proteome-wide medicinal chemistry

Polina Prokofeva

Complete reprint of the dissertation approved by the TUM School of Life Sciences of the

Technical University of Munich for the award of the

doctoral degree (Dr. rer. nat.)

Chair: Prof. Dr. Corinna Dawid

Examiners:

1.   Prof. Dr. Bernhard Kuster
2.   Prof. Dr. Norbert Sewald

The dissertation was submitted to the Technical University of Munich on 22.03.2023 and

accepted by the TUM School of Life Sciences on 14.06.2023.

"Serendipity was my tour guide assisted by caprice."
- Pico Iyer -

Abstract

Mass spectrometry-based chemical proteomics enables studying a drug's action in complex biological systems by identifying its protein targets with the help of diverse chemical probes.

Bioactive Natural Products (NPs) provide endless inspiration for drug research. However, identifying their protein targets with the help of tailored probes in a classic target deconvolution realm is often hampered by the complexity of their chemical structure or lack of structure-activity relationships. Consequently, a general approach that enables to probe NPs for protein binding while omitting extensive probe synthesis is highly desirable. The first part of this work describes the efforts toward establishing a chemical toolbox for preparing NP-based tailored affinity matrices via their native functional groups and various promiscuous bead immobilization chemistries. Benchmarking of gold(I)-catalyzed alkoxylation and [2+2]-cycloaddition, ruthenium-catalyzed metathesis, Mitsunobu reaction, and UV-induced photo-crosslinking were assayed here for on-bead immobilization of unmodified tacrolimus. A novel photo-immobilization protocol was further exploited for the preparation of matrices of 30 more bioactive NPs and protein cofactors, where 80% of tested unmodified molecules were successfully immobilized on beads and obtained tailored affinity matrices were subjected to chemoproteomic target deconvolution. The latter enabled the identification of protein targets of NPs in human cancer and bacteria cell lysates, as well as probing protein interactors of coenzyme A and flavin adenine dinucleotide.

In the second part of this work, the two kinase inhibitor chemotypes, pyrimidopyridones and lophines, were employed to demonstrate that chemical proteomics enables deciphering structure-affinity relationships for a given pharmacophore not only against one target of interest but on a proteome-wide level. Here several linkable analogs of each chemotype were chemically synthesized and utilized in a competition-based chemoproteomic selectivity profiling against their bead-immobilized avatars. This set of molecules and matrices was used to interrogate which proteins were bound by which compound in a fully target-agnostic manner. The small set of seven pyrimidopyridones demonstrated that it is possible to unravel drastic changes in the target profiles of chemotype analogs upon introducing modifications in their chemical structures. The approach was extended to a newly synthesized library of 35 linkable lophines, whose target profiles revealed very tight structure-affinity relationships. Eleven molecules showed submicromolar affinity, mostly for kinases. Unexpectedly, four compounds were found to possess a unique selectivity for either RIPK2, ALK5, or SPR. An activity assay confirmed that the novel probe inhibited SPR, an unexpected non-kinase target. In addition, the cellular submicromolar potency of the ALK5 binder was evidenced by monitoring the inhibition of its downstream pathway.

This work demonstrates complementary applications of chemoproteomics in the study of small molecule protein targets and discusses how such methods can be implemented into the chemoproteomics-aided drug discovery pipeline. The first part offers a tool to probe NPs for protein binders as a potential follow-up of a phenotypic screen. Results of the second part prompt the generalization of proteome-wide SAR to many chemotypes with the view of proposing chemical probes for proteins that do not have any and expanding the addressable proteome.

Zusammenfassung

Die chemische Proteomik auf der Grundlage der Massenspektrometrie ermöglicht die Untersuchung der Wirkung von Arzneimitteln in komplexen biologischen Systemen durch die Identifizierung ihrer Protein-Zielstrukturen mit Hilfe verschiedener chemischer Sonden.

Bioaktive Naturprodukte (NPs) bieten ein großes Potential für die Arzneimittelforschung. Häufig wird jedoch die Identifizierung ihrer Protein-Zielstrukturen mittels maßgeschneiderter chemischer Sonden durch die Komplexität ihrer Strukturen oder das Fehlen von Struktur-Wirkungs-Beziehungen behindert. Daher wird dringend ein Ansatz zur Untersuchung von Proteinbindung durch NPs benötigt, der keine umfangreiche Synthesechemischer Sonden erfordert.

Der erste Teil dieser Arbeit beschreibt, wie eine Palette an chemischen Werkzeugen aufgebaut wurde, mit deren Hilfe auf Basis verschiedener selektiver und unselektiver Kopplungsreaktionen aus NPs maßgeschneiderte Affinitätsmatrizen hergestellt werden können. In diesem Kontext wurden mehrere Protokolle zur Immoblisierung von unmodifiziertem Tacrolimus auf Beads etabliert: Gold(I)-katalysierte Alkoxylierung, [2+2]-Cycloaddition, Ruthenium-katalysierte Metathese, Mitsunobu-Reaktion sowie UV-induzierte Photo-Kopplungsreaktion. Das neuartige Photo-Immobilisierungsprotokoll wurde für die Herstellung von Matrizen aus 30 weiteren bioaktiven NPs und Protein-Kofaktoren genutzt, wobei 80% der unmodifizierten Moleküle erfolgreich auf der Matrix immobilisiert wurden. Die so erhaltenen maßgeschneiderten Affinitätsmatrizen wurden einem chemoproteomischen Selektivitäts-Assay unterzogen. Dies ermöglichte die Identifizierung von Protein-Zielstrukturen der NPs in menschlichen Krebs- und Bakterienzell-Lysaten sowie die Untersuchung von Protein-Bindungspartnern von Coenzym A und Flavin-Adenin-Dinukleotid.

Im zweiten Teil dieser Arbeit wurde anhand zweier Chemotypen von Kinaseinhibitoren, Pyrimidopyridonen und Lophinen, gezeigt, dass mittels chemischer Proteomik die Struktur-Affinitäts-Beziehungen für ein bestimmtes Pharmakophor auf proteomweiter Ebene untersucht werden kann. Hierzu wurden mehrere kopplungsfähige Analoga jedes Chemotyps synthetisiert und in einem wettbewerbs-basierten Selektivitäts-Assay gleichzeitig mit ihrer eigenen Affinitätsmatrix eingesetzt. Dieses Set an Molekülen und Matrizen wurde verwendet, um zu untersuchen, welche Proteine von welcher Verbindung in einer vollkommen ziel-unabhängigen Weise gebunden wurden. Anhand der sieben getesteten Pyrimidopyridone konnte gezeigt werden, dass der Einbau chemischer Modifikationen in der Tat zu einer drastischen Veränderung in den Selektivitäts-Profilen von Chemotyp-Analoga führen kann. Der Ansatz wurde auf eine neu synthetisierte Bibliothek von 35 kopplungsfähigen Lophinen ausgeweitet, deren Selektivitäts-Profile sehr enge Struktur-Affinitäts-Beziehungen erkennen ließen. Elf Moleküle wiesen eine submikromolare Affinität gegenüber ihren Zielstrukturen auf, hauptsächlich für Kinasen. Überraschenderweise zeigten vier Verbindungen eine einzigartige Selektivität gegenüber RIPK2, ALK5 oder SPR auf, wobei SPR keine Kinase ist und daher ein besonders unerwartetes Ergebnis darstellte. Ein Aktivitäts-Assay bestätigte, dass SPR tatsächlich durch die neuartige chemische Sonde gehemmt wird. Die zelluläre submikromolare Wirksamkeit des ALK5-Binders konnte durch die Hemmung des dem Protein nachgeschalteten Signalwegs nachgewiesen werden.

Diese Arbeit demonstriert komplementäre Anwendungen der Chemoproteomik zur Untersuchung von Protein-Zielstrukturen niedermolekularer Substanzen, und erörtert, wie solche Methoden in die Chemoproteomik-gestützte Arzneimittelforschung implementiert werden können. Der erste Teil beschreibt ein zeiteffizientes Werkzeug zur Untersuchung von NPs auf ihr Vermögen, an Proteine zu binden, welches im Anschluss an phänotypische Screens eingesetzt werden kann. Die Ergebnisse des zweiten Teils regen dazu an, proteomweite SAR-Studien auf zusätzliche Chemotypen auszuweiten, um neue chemische Sonden für Proteine zu finden und dadurch das durch Wirkstoffe erreichbare Proteom zu erweitern.

Table of contents

Contents

# Introduction

1 Medicinal chemistry is target-centric

1.1 Drug discovery efforts are phenotype- or target-oriented

Medicinal chemistry plays an indispensable role in drug discovery and is devoted to understanding drugs' mechanisms of action and pharmacology, as well as establishing structure-activity relationships (SAR) of drug-target interactions with an intent to make molecules that cure disease.[1]

In very simplified terms, one could describe the primary objective of drug discovery as finding a pair of a biological target that plays a role in the disease and a therapeutic agent that interacts with it. In reality, prior to any drug development campaign, a lot of effort is dedicated to understanding of physiological processes that cause or arise from a disease, so-called pathophysiology. Such laborious studies often aim not only to identify the characteristic disease phenotypes but also to uncover the key players, typically proteins, responsible for them. Most traditional targets are enzymes, intra- and extra-cellular receptors, structural proteins, or transcription factors. In cases when such proteins can be addressed with therapeutics, they evolve as targets in following drug discovery campaigns. The process of finding a "druggable" biological target, termed target identification, lies at the heart of modern target-based drug discovery (Fig. 1). All potential targets must be validated i.e. prove their relevance to the disease and that modulation of their behavior or function can provide a desired therapeutic outcome. This, however, heavily depends on having relevant, physiologically realistic disease models and phenotypic assays. Currently, target validation is often performed through knockout experiments, where a system that lacks the expression and activity of one or more candidate genes is generated, and the response of a system to such alterations is monitored.[2] The challenge of discovering and validating targets is reflected in the failure of clinical trials, where drug candidates fail to show efficacy in treatments. The reason for these failures often lies in the fundamental hypothesis of a particular target modulating the disease in a patient being wrong.[3] Furthermore, about a half of marketed drugs target in total only five main protein families: kinases, proteases, G protein-coupled receptors, nuclear receptors, and ion channels, whilst other parts of the proteome remain therapeutically unaddressed.[4]

The phenotype-based drug discovery approach involves exposing cells, isolated tissues, or animal models to small molecules and determining whether some of the tested molecules exert the desired effect – specific changes in the observed phenotype. Here the effect of the molecules is determined through the phenotypic response before the biological target that underlies it is identified (Fig. 1a). The drug's action is monitored in its relevant biological context (cellular environment) rather than with biochemical assays on a purified target, representing an advantage of such an approach over the target-based. Nevertheless, to assess a compound's utility for potential therapeutic use, the molecular mechanisms behind observed phenotypic effects have to be delineated. More precisely, the question of which protein targets are responsible for the observed phenotype has to be addressed. To this day, the actual deconvolution of biological targets that cause observed phenotypic changes is the biggest challenge of phenotype-based drug discovery, especially when the effect stems from synergistic polypharmacology. The complexity of phenotypic assays, higher costs, and smaller throughput are other limitations associated with

such cell-based screens. In comparison, the target-based approach holds an advantage as the drug's intended molecular action is predicated by the biological target, simplifying the conduction of evaluation assays and interpretation of the results. Consequently, the majority of drug discovery campaigns currently initiate as target-specific. In the last few years, phenotypic screens are experiencing a renaissance, however, much effort is dedicated to the parallel identification of compound targets.[5] Recently, a pharma collaborative consortium, JUMP-CP, was formed to implement the cell painting technology into a large-scale compound phenotypic screen.[6] Such an assay employs multiple dyes to stain and microscopically visualize cellular organelles to facilitate tracking various structural changes upon drug perturbation within live cells. Here the mass spectrometry-based proteomics readout was added to complement the insight into the compound mode of action by monitoring the changes in protein expression upon drug exposure. The individual changes in protein expression caused by the treatment may help uncover the effects of the drug on the cell cycle and affected by the treatment signaling cascades, which in turn should allow restricting of the number of potential drug targets for the following evaluation.
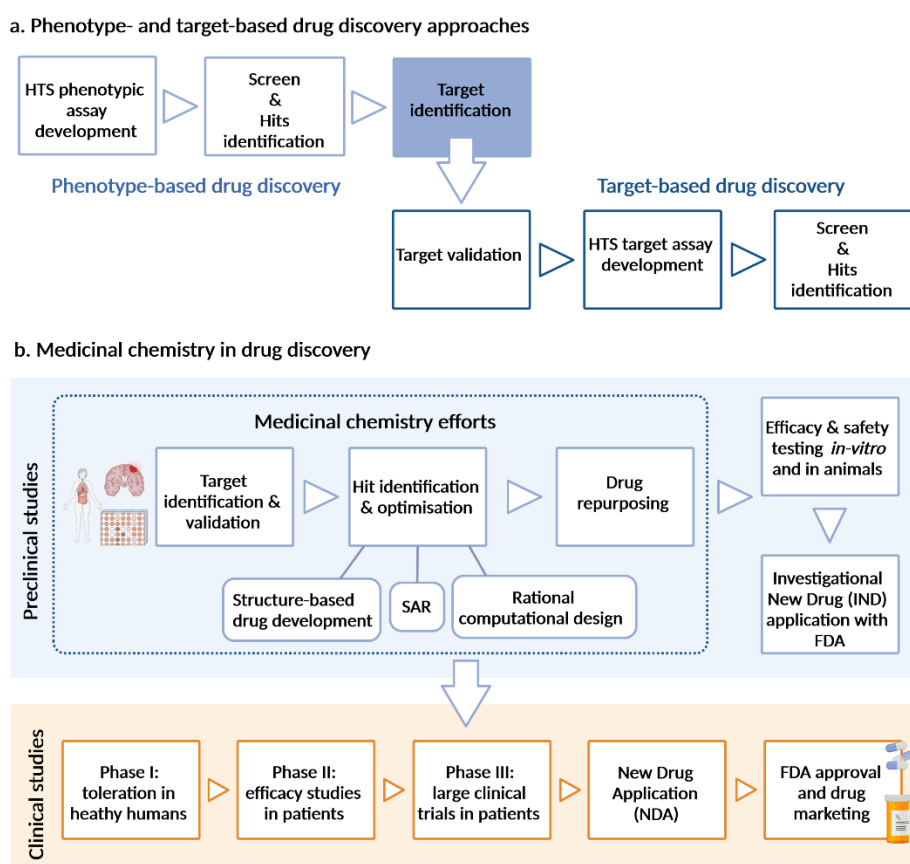
Figure 1. Schematic representation of drug discovery pipeline. a) Phenotype- vs. target-based drug discovery. b) Medicinal chemistry within drug discovery. Adapted from Terstappen et al. (panel a), Lombardino and Lowe III, and Tucker et al. (panel b).[1,5,7] Figure created with Biorender.com

Primary medicinal chemistry efforts within drug discovery campaigns revolve around finding a suitable bioactive molecule to effectively address a particular target. This often includes the development of assays that would allow screening for suitable candidates as well as finding and evolving a promising chemical scaffold by means of experimental and/or computational SAR.

Primarily, establishing chemical scaffold-target SAR aims to optimize on-target potency and is often accomplished with the help of so-called rational structure and computational design (Fig. 1b). Medicinal chemists also explore the drug repurposing avenues that allow re-using existing drugs to address other therapeutic targets or treat another disease (further discussed in the following chapter). In a nutshell, medicinal chemistry is an interdisciplinary field that encompasses the discovery of novel bioactive compounds (hits) and their optimization into leads that can be further employed in drug development campaigns (Fig. 1b).

1.2 Compound leads are obtained through screening

A 1988 publication by DesJarlais et al. in the Journal of Medicinal Chemistry states: "Finding novel leads from which to design drug molecules has traditionally been a matter of screening and serendipity".[8] Nowadays, within the realm of target-based drug design, the picture is not much different: once a "druggable" target is identified, the compound hits are obtained with the help of combinatorial chemistry, high throughput screens (HTS) and structure-based drug design (Fig. 1). Combinatorial chemistry here allows for the parallel generation of a large array of structurally diverse compounds and a great expansion of the screening space. When a new target needs to be addressed, the same libraries of bioactive compounds are often re-screened for hits, as each individual screen is usually focused on one specific target or a small number of related ones. For instance, for decades, large libraries of bioactive natural molecules, synthetic compounds, or fragments are screened in high throughput fashion for the desired on-target effect using *in-vitro* assays with recombinant proteins or in large-scale phenotypic assays.[2] For example, the immunosuppressant drug cyclosporin A was identified this way in 1988.[9] There is a number of assays currently available for specific target classes, pathways, networks, and cellular phenotypes. Large initiatives such as EU-Openscreen and DE-Openscreen enable the performing of such HTS campaigns in academia by bringing together facilities and know-how. By definition, HTS is tailored to quickly evaluate large compound collections in parallel (hundreds of compounds at a time); therefore, such large screening assays are designed and developed to be simple yet efficient and allow maximal automation of the experiment. Commonly, such screens monitor either 1) substance concentration (enzyme substrate or product), 2) receptor-ligand complex concentration, or 3) distribution of cellular markers. Most difficulties associated with HTS assay optimization usually arise from maintaining the balance between assay sensitivity and statistical performance. Automated data processing is equally crucial as typical screens are set up to process more than 100 000 samples per day. Traditional HTS tests each compound within a library in a single concentration dilution, commonly as high as 10 µM.[4] Quantitative screens (qHTS) are a natural evolution of classical single-dose assays and are performed to test compounds with multiple concentrations, which consequently limits the throughput of molecules that can be analyzed in parallel.[10] Almost two-thirds of therapeutic targets were reported to be comprised of enzymes and receptors.[4] Protein kinases, proteases, nuclear receptors, G protein-coupled receptors (GPCRs), and ion channels are commonly assayed in HTS with a number of available automated assays. For example, protein kinases catalyze the transfer of a phosphoryl group from ATP to hydroxyl groups of amino acid side chains, which can be analyzed by monitoring the ATP depletion via luciferase, an ATP-dependent photoenzyme.[11] An antibody-based assay allows for

measuring kinase activity as a function of ADP formation, where the readout is performed using an ADP-specific antibody.[12] Proteases are commonly probed as purified enzymes or in lysates with profluorescent or FRET-based substrates. Utilized probes are typically dye-coupled short peptides that comprise a protease cleavage site and become fluorescent after the protease cleavage. Cellular protease activity can be assayed with FRET-based methods using fluorescent proteins or dyes linked to a protease consensus sequence. Overall, HTS can be versatile and tuned to assay a specific target, protein interaction, or phenotype using biochemical and imaging microscopy techniques. Pharma companies have also adopted HTS format for the further evaluation of the drug's ADMET/DMPK (absorption, distribution, metabolism, excretion, toxicity/drug metabolism, and pharmacokinetics) activities.[4]

Structure-based virtual screenings (SBVS) have gained a lot of attention in recent years for the potential to substitute HTS and facilitate hit generation. Here large compound collections are computationally analyzed for the ability to bind the target. SBVS computes the best interaction mode between two molecules to form a stable complex, where various scoring functions estimate the force of non-covalent interactions between a ligand and its molecular target. This approach is naturally limited to the protein targets with solved 3D structures. Virtual screens often precede the experimental ones aiming to narrow the compound library.[13]

Assessment of the drug candidates' polypharmacology and mode of action is essential and can only be achieved via thorough evaluation of phenotypic effects and unbiased identification of direct and complex binders along with other interactors in representative disease models. Even though large screens can unravel early hits that perform well against a particular target, yet again, as already mentioned, the absence of a native biological context may result in unanticipated or undetected off-target effects or even overestimate the hypothesized relevance to the underlying disease process. Such screens cannot always adequately represent biological levels of target proteins, which, in turn, can potentially generate wrong hypotheses for on-target drug effects. Initial screens are also often not considering such drug parameters as cell permeability, metabolism, or potential competition with other cellular molecules for the target engagement. HTS is frequently performed with the same large bioactive compound libraries for novel targets under investigation, where the outcome is purely relying on serendipitous hits. After completion of the screen, only a considerably small number of compounds retain in a selected pool of candidates; a handful of those evolve into leads through the initial stages of development, and eventually, a couple of best-evolved candidates undergo thorough efficacy and safety testing, and only single compounds proceed into extensive multi-stage clinical evaluation (Fig. 1). Not even considering that overall the screen of thousands of compounds would lead to one single molecule evolving as a drug while all other molecules from the panel remain neglected, the initially selected candidates often drop out of campaigns when previously obtained activity data from computational or experimental HTS appear to be no longer valid *in-vivo*.[3,14]

Repurposing existing drugs offers an alternative avenue in drug development and bypasses extensive screening efforts. The main objective of repurposing is to reassign available and thoroughly investigated drugs to engage another target and/or treat another disease. This strategy offers lower failure risks (as the analyzed candidates have already been evaluated for safety) and holds the potential to reveal novel targets that can be further exploited.[15] Repurposing

efforts enabled the extension of the BCR-ABL inhibitor imatinib use from the treatment of chronic myeloid leukemia (CML) to gastrointestinal stroma tumors after discovering other targets of the drug, namely c-KIT, and PDGFR. Imatinib was the first FDA-approved kinase inhibitor (2001), and interestingly, it was also the first drug developed by targeting a specific protein kinase to be approved. Surprisingly, other kinase inhibitors that got FDA approval earlier than 2001, such as fasudil and rapamycin, were developed and approved without knowing of their respective targets.[16] Imatinib illustrates the prime example of drug repurposing and demonstrates how the drug development efforts have largely switched gears in the last 30 years from a phenotype to target. Another excellent example of a repurposed drug is thalidomide. After the post-approval discovery of teratogenesis, the drug developed to treat the morning sickness of pregnant women was withdrawn. However, it was later shown that thalidomide decreases the levels of tumor necrosis factor-α; therefore, it found another use in the treatment of erythema nodosum leprosum (ENL).[17,18] It has been equally repurposed for treating multiple myeloma, and derivatives such as lenalidomide and pomalidomide have been developed for treating blood cancers. Since it has been reported to primarily target the E3 ubiquitin ligase cereblon, thalidomide found yet another application in PROTAC (proteolysis targeting chimeras) approach.[19] PROTACs consist of three building blocks: the affinity warhead that directs the chimera into the pocket of protein of interest (POI), a linker, and the E3 ligase recruiter. Such bi-functional molecules enable *in-cellulo* chemical knockdown of a protein by recruiting the degradation machinery to the POI. Thalidomide is an excellent illustration of how thorough investigation can reveal unexpected actions of the molecule that can be beneficially exploited for another application.

1.3 Medicinal chemists evolve hits into leads by SAR

Obtained initial HTS hits are further evolved and optimized into leads with improved properties by medicinal chemists. This is typically done through establishing structure-activity relationships (SAR) between the selected from the HTS chemical scaffold and the target, where closely structurally related library analogs are either synthesized or computationally generated (so-called QSAR) and evaluated for on-target activity.[20,21] An understanding of the SAR for a series of molecules by rationally exploring the chemical space involves deciphering the functional groups or molecular building blocks that influence (and to which extent) the compound activity and/or other important properties. In other words, SAR identifies relationships between compound structures and properties. SAR for a compound series can help simultaneously optimize on-target activity or potency and other drug candidates' physicochemical and biological properties, for example, reducing toxicity or ensuring sufficient bioavailability.

Over the past ten years, the idea of representing and analyzing SAR as a landscape has gained focus. Here the SAR relationships are visualized in a three-dimensional space, typically with structure similarity represented in the X-Y plane and the activity plotted along the Z-axis (Fig. 2). Such visualization allows exploring the SAR as a landscape of varying topography: structurally similar compounds with similar activities will produce smooth or flat regions (continuous), whereas differences in the activity of structurally similar compounds will create so-called activity cliffs (discontinuity). Such cliffs are often viewed as the most exciting parts within the SAR dataset, as they decipher small structural changes that greatly influence compound activity. SAR

discontinuity (activity cliffs) is still not entirely amenable to the analysis via computational methods (computational SAR or QSAR), where the machine learning or statistical approaches still greatly underperform compared to the experimental methods.[22,23]
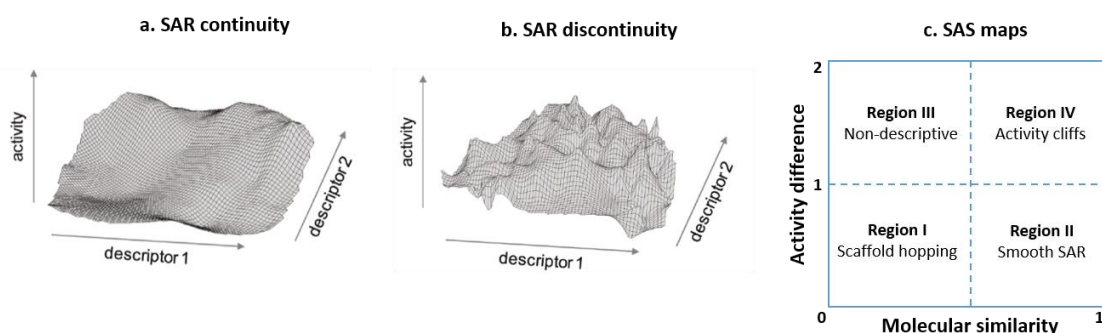
Figure 2. SAR representations. SAR landscapes: a set of compounds is analyzed for activity in a 2D projection of chemical reference space with a third potency dimension; the distances in 2D space reflect compound similarity. Target potency values of individual compounds on the vertical axis allow the generation of a coherent surface (landscape) through interpolation. The two idealized hypersurfaces representing SAR characters are shown: a) SAR continuity, b) discontinuity. SAS maps: c) general SAS maps scheme, where regions I and II are associated with scaffold hopping and smooth/continuous SAR, region III does not provide relevant information (different compounds with different activity), and region IV depicts activity cliffs (similar compounds with different activity). Adapted from Wassermann et al.[22] and Saldivar-Gonzalez et al.[24]

One common approach to quantifying the SAR landscape is structure-activity similarity maps (SAS maps). SAS maps are pairwise plots of compound similarity against the activity difference (or similarity) (Fig. 2c). Such plots can be divided into four regions:

1. Region I with structurally different compounds that have similar activity (or target potency)
2. Region II of smooth SAR with similar compounds and activities
3. Non-descriptive region III, as it comprises different compounds with different activities (SAR can not be elucidated for those)
4. Region IV depicts compounds of similar structure but different activity (activity cliffs).

The compound activity similarity is defined as

$$S_{Act}(A,B) = 1 - \frac{|\text{Act}(A) - \text{Act}(B)|}{Act_{max} - Act_{min}} \quad (1)$$

where Act(A) and Act(B) are the activities or potencies of compounds A and B, and Act_{max} and Act_{min} are the maximal and minimal activities within the compound set, respectively.[23] SAS maps utilize compound similarity as another main descriptor. There are many different approaches to calculating structure similarity; however, the Tanimoto coefficient is by far the most widely used one. Here the molecular structure is separated into fingerprint bits, where each moiety (or a particular functional group) is represented by one bit. In order to compute the similarity between two chemical structures, the two respective fingerprints (i.e., the ensemble of bits or bit-vectors) are compared.[25]

At the same time, the presence or absence of activity cliffs can be represented with the structure-activity landscape index (SALI):

$$SALI_{A,B} = \frac{|\text{Act(A)} - \text{Act(B)}|}{1 - sim(A,B)} \qquad (2)$$

where Act(A) and Act(B) are measured activities of compounds A and B and sim(A,B) is the Tanimoto similarity between the two molecules.

Peltason and Bajorath suggested an alternative metric of the SAR landscape for a pair of molecules, SAR index or SARI.[22] It is composed of two separately calculated scores and is defined as

$$SARI_{A,B} = \frac{1}{2}(score_{cont} + (1 - score_{disc})) \qquad (3)$$

where $score_{cont}$ is a continuity score calculated as the potency weighted arithmetic mean of pairwise compound dissimilarity within a set. It strongly weights structurally different compounds having similarly high potency. Thus, it accounts for continuous regions within the activity landscape. The discontinuity score $score_{disc}$ emphasizes structurally similar compounds with large potency differences and hence accounts for activity cliffs. Because the discontinuity score is designed to monitor the presence of activity cliffs, only pairs of compounds with at least one order of magnitude difference in potency and a similarity exceeding a predefined threshold are considered. SARI helps to classify the SAR into three categories: continuous (high SARI), discontinuous (low SARI), and heterogeneous. SARI is, by default, target-specific and assesses the activity of multiple compounds to identify activity cliffs for a defined target, whereas SALI focuses on the comparison of individual pairs of molecules. SALI essentially corresponds to the SARI discontinuity score: SALI is a pairwise score with infinite value range that emphasizes large potency differences between similar compounds, whereas the SARI discontinuity score takes average potency differences of all pairs of similar compounds into account and is normalized.

One considerable limitation in the current SAR elucidation paradigm is the fundamental foundation of SAR being target-tailored. Typically, broad compound selectivity and off-targets are assessed only across a focused panel of representative proteins and only for a handful of molecules after the initial SAR evaluation for on-target activity. Consequently, the global potential of generated SAR compound libraries to bind proteins (and alter protein function) is mostly unaddressed since the molecules that did not score in the initial on-target SAR screens remain unexploited. One notable example is the study from Lo et al.[26] on the development of selective PI3KC2α inhibitors. Initial efforts of the work were dedicated to identifying of a chemical scaffold that can interrogate the PI3KC2α function by means of HTS. More than 35,000 compounds were subjected to an enzymatic activity assay with purified recombinant protein, narrowing down the compound pool to 352 initial hits. A small subset of these hits (48 total) were selected for a selectivity profiling against a panel of lipid kinases in a similar recombinant enzyme activity assay setup, which led to an identification of a selective inhibitor (IC50 PI3KC2α =2.6 μM) containing a rare for kinase inhibitors pteridinone scaffold.

Further medchem chemotype evolution of multiple key-scaffold substituents led to the discovery of three compounds with low nanomolar potency of PI3KC2α inhibition. In the rounds of medchem molecular evolution, the synthesized library analogs were solely tested for PI3KC2α

inhibition but were not subjected to broad selectivity profiling. Only the three final compounds (termed PITCOINs by the authors) were profiled against a large set of human kinases. Although these compounds showed low nanomolar PI3KC2α inhibition potencies in the recombinant enzyme activity assay, the chemoproteomic Kinobeads target profiling in lysate followed by mass spectrometry readout of these compounds demonstrated much poorer target engagement with affinities ranging between 3 to nearly 30 μM for the three PITCOINs. This example also highlights how the measured on-target effects can differ simply by using another readout (here, activity assay against purified PI3KC2α vs Kinobeads in lysate). Meanwhile, all generated SAR library analogs were never assayed against any other target or target class other than PI3KC2α. It is the classical route within the current target-oriented drug research; nevertheless, it appears especially prodigal within the context of this work, as all the chemical effort invested in the synthesis and purification of the compound library simply produced a range of "inactive" analogs. There is, however, a fair chance that these compounds might be able to engage another target, which will not be explored, and all these molecules will forever remain uncharacterized.

1.4 Tuning drugs to act in a complex environment is challenging

Drugs can interact with a plethora of partners in a biological system: proteins, lipids, metabolites, sugars, DNA, RNA, or other small molecules.[27] Some of these interactions may cause unexpected or undesired side effects. On the other hand, the ability to engage multiple targets or resistant mutants, so-called drug promiscuity, can be advantageous. For example, a poly-mutant drug that can address rapidly mutating targets in infectious diseases (like HIV or cancer) will undeniably hold a better potential for prolonged positive treatment effects.[28] In cases when several pathways can be accessible for a given signaling cascade, targeting members in each parallel pathway may be advantageous to successfully block the downstream signal. Therefore, one of the other challenges in drug research is maintaining the fine balance between covering multiple disease-relevant targets and narrowing down the selectivity to avoid the engagement of undesired interactors.

Kinase inhibitors are one of the prominent examples of such a balance. Targeting kinases with small molecule inhibitors is non-trivial, and one can rather speak about family or subtype binding selectivity instead of absolute selectivity for a single kinase. These targets are typically homologs binding the same molecule ATP and catalyzing the transfer of a phosphate group to a substrate. Meaning it is challenging to target one specific kinase, but tuning the binding profile to a smaller subset of kinases is possible and can sometimes even be favorable.[29] Selective small molecules are undeniably valuable for general research and help discover new insights into kinase signaling pathways. However, most clinically successful kinase inhibitors demonstrate broad selectivity profiles.[30]

Conceptually, designing a drug with a particular selectivity profile is much more complex than for a single target. It involves identifying and exploiting binding similarities across a selection of targets rather than finding key differences between targets and decoys that could tune the selectivity of the molecule for one particular target. Selectivity alone does not define whether the drug will become a successful treatment option. Nevertheless, drug promiscuity can be genuinely beneficial when controlled, meaning all drug targets (potency and selectivity) and potential toxicity are adequately monitored and evaluated. It is becoming increasingly common to probe

molecules for off-target interactions with ion channels, cytochrome P450s (CYPs), and other proteins that can potentially cause adverse side effects.[31,32]

There are several experimental and computational approaches to probe for off- and on-target interactions to assess the drug's action. Although it is not possible to perform wide selectivity screens against all plausible interaction partners, it is possible to create screening panels that can be used to gain selectivity insights. One common aspect of many target profiling methods is the assumption that physical affine interaction is a prerequisite to the observed functional effects. Therefore, they intend to determine the affinities between the bioactive molecule and its protein targets. Common methods include affinity chromatography, where protein targets are identified by immunodetection or mass spectrometry, expression-cloning methods, such as yeast three-hybrid, phage (or mRNA display), and protein (protein chips) as well as "reverse transfected" cell microarrays.[5] Some experimental and computational methods for potency and selectivity profiling, mainly target deconvolution with the help of unbiased quantitative mass spectrometry, will be discussed in the following chapters.

1.5 Reversible interactions that define ligand binding affinity

The affinity is commonly defined as the ability to form a stable protein-ligand interaction, and its strength is determined from the ensemble of hydrogen bonds, salt and water bridges, hydrophobic interactions, and pi-stacking formed between two interacting molecules (Fig. 3a). Target-oriented drug design is tailored to finding high-affinity ligands that bind the target and alter its disease-associated function (catalytic activity or protein-protein interactions); therefore, the correct binding affinity estimation is crucial in drug development. The analysis of all stable interactions between the bioactive small molecule ligand and the protein pocket can most of the time shed light on the observed binding strength and potentially explain selectivity.

For example, staurosporine is a broad-selective kinase inhibitor that, among other targets, binds CSK with micromolar affinity, whereas its binding affinity to CDK2 is much higher (low nanomolar range). The position of Staurosporine in the hydrophobic N-terminal ATP-binding pocket of CSK is very similar to CDK2, however, some key differences can explain the orders of magnitude binding affinity differences between the two enzymes.[33] For instance, Staurosporine forms only three hydrogen bonds with CSK, whereas it can form four of such interactions with CDK2. Among those are the nitrogen of the lactam ring donating an amide bond to the Glu95, the oxygen on the same lactam donating to the backbone amide of Met97, and the secondary amine nitrogen is donating to the carbonyl of the backbone Arg146 (Fig. 3b). Analogously for CDK2: there are the same three hydrogen bonds with Leu83, Glu81, and Gln131, respectively (Fig. 3c). The missing hydrogen bond should be forming between the secondary amine nitrogen and the Ser273 (the bond of the methylamino nitrogen and the Asp86 in CDK2), however, the latter is already engaged with the amide backbone of Asp276. Therefore, the interaction with Staurosporine cannot be established. Additionally, the methylamino group of Staurosporine is expected to be protonated at physiological pH, hinting towards a possible stabilizing charge-charge interaction with the side chain. Since the Ser273 in CSK is already engaged in the bond with the backbone, such charge interaction could not be formed, whereas it can take place for CDK2. Other kinases such as MAPK, CK1 and CK2 also possess a Ser in the analogous positions and are not potently inhibited by Staurosporine, supporting the evidence that differences in binding affinities arise from the bond between the kinase side chain and methylamino-nitrogen. Interestingly, PLIP analysis of both

Staurosporine-bound complexes suggests that staurosporine interaction with CDK2 is additionally stabilized with 8 hydrophobic interactions, while in the case of CSK, there are only 5 of them (Fig. 3d,e). Staurosporine example highlights the importance of the assessment of drug-target binding affine interactions as it can explain the potency and potential selectivity for a particular protein.
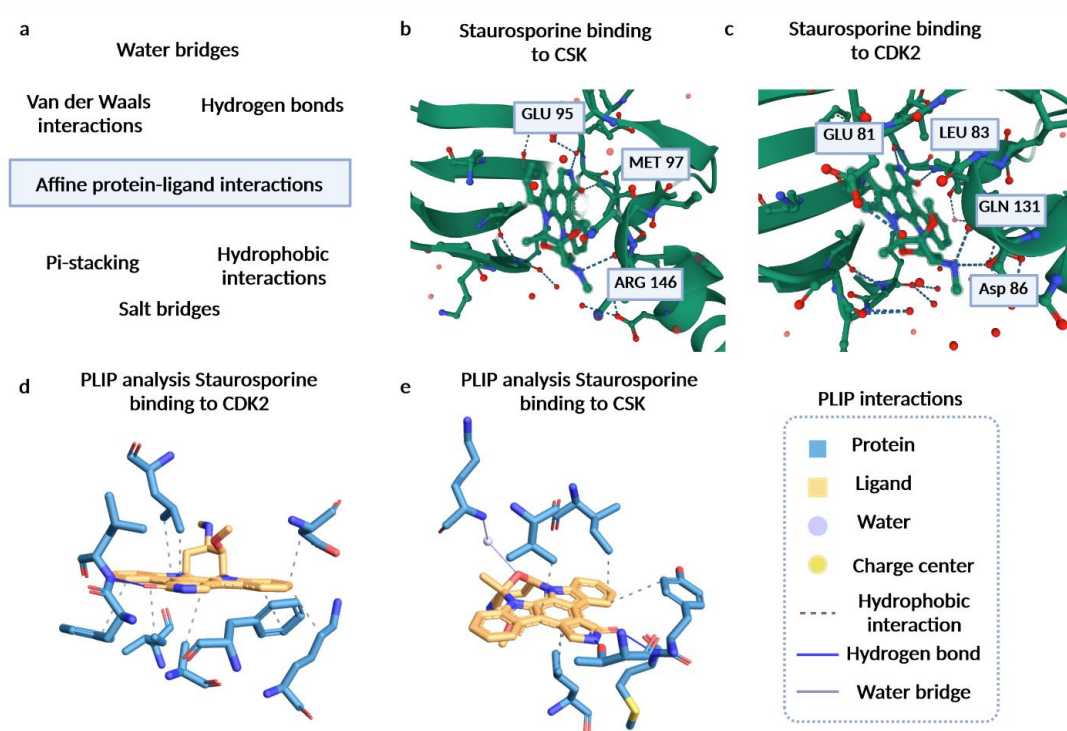
Figure 3. Affine interactions and design of affinity probes. a) Non-covalent interactions that determine drug-target binding affinity. b) Affine interactions of staurosporine with C-terminal SRC kinase (CSK). c) Staurosporine binding to CDK2. Binding interactions (hydrogen bonds) were reported in PDBs 1AQA and 1BYG, respectively d) Analysis of binding interactions of staurosporine with CDK2 and CSK performed with PLIP using PDB 1AQA and 1BYG.[34] Figure created with Biorender.com

1.6 Structure-based rational design helps to sculpt better drugs for a given protein pocket

Rational or structure-based drug design is a commonly used approach that utilizes a combination of computational and experimental methods to sculpt better drugs for specific molecular targets, determine structure-activity relationships, and facilitate lead optimization based on structural data.

Three standard models explain protein-ligand binding mechanisms: the lock-and-key, induced fit, and conformational selection. The lock-and-key model is based on the assumption that both the protein and the small molecule ligand are rigid and their shapes are perfectly complementary. Thus, only the ligand (the key) can correctly fit into the binding site of a protein of interest (the lock). This model fails to explain the experimentally acquired data when the ligand can bind a protein even though their shapes do not match. The second model, the induced fit, assumes a

protein undergoes a conformation change upon ligand binding, inducing the stable binding interaction. This model considers protein flexibility only within and proximal to the ligand binding site. Therefore, it is not suitable to account for major protein conformational changes. Both of the described models treat proteins as one stable conformation. The conformational selection model is based on the premise that proteins are dynamic and exist in an equilibrium of different conformational states of variable population distributions and that the ligand binds the most complementary one, ultimately shifting the equilibrium towards this state.[35]

Structure-based drug design is built on the fundamental assumption that a drug exerts its bioactivity through the specific binding to a particular protein target. This binding leads to the alternation of protein function that, in an ideal case, cures the disease. A prerequisite for such potent and selective binding is the complementarity of protein targets' and drug ligands' shapes. Structure-based drug design, therefore, is a variety of methods that allow optimizing a ligand for a perfect fit in its targets' active site.[36]

Among other techniques, molecular docking is a widely used medchem computational approach that helps understand ligand-protein interactions by placing the ligand in the binding site of the target. The possibility of mapping any small molecule into the pocket of any protein of interest offers a unique opportunity for preliminary assessment of potential binding interaction and facilitation of rational probe design. The primary purpose of docking is to understand the binding interaction's premise and to help design a better probe to complement the protein target and form a stable affine association. The concept of shape complementarity here can be exploited for the design of either narrow or broad drug selectivity by requiring, for example, a high degree of complementarity for the indented target(s), simultaneously not matching the shapes of undesirable interactors. Furthermore, systematic exploration of binding pockets with the help of computational methods, such as docking, can delineate key selectivity and affinity residues responsible for forming a stable affine interaction with the drug/bioactive small molecule or the critical differences between the binding pocket residues of the target compared to the pockets of other interactors. For example, the analysis of the ROCK-1 ATP-binding pocket has identified five key residues responsible for inhibitor binding, namely Met123, Ala142, Asp158, Ile186, and Phe327. Sequence alignment with 491 other kinases and ligand docking indicated that these residues were comparably common; however, the specific combination of residues in ROCK-1 appeared unique, with Phe327 as a part of the C-terminal strand being found only in a very small subset of kinases, which altogether allows for the design of selective ROCK-1 inhibitors.[27,37]

Proteins are not static; their flexibility and dynamics largely determine their recognition and functional properties. Enzymes are known to rearrange their catalytic sites for optimal reactivity; similarly, the protein-protein complex formation can involve structural changes beyond the lock-and-key model or fundamental principles of surface accessibility. Specific conformational changes induced upon the formation of a complex with the binding partner can be analyzed with computational methods. For example, computational modeling has demonstrated that the binding selectivity of Imatinib to SRC and ABL human tyrosine kinases with high sequence homology in the ATP-binding site is greatly influenced by the conformational change occurring after the binding event and not solely depends on affine interactions with the protein pocket. The high binding affinity to ABL is predominantly the result of the induced fit, which does not form

after binding to SRC. Here, the NMR titrations, energy state calculations, and real-time binding kinetics experiments showed that the induced fit conformation of the ABL-imatinib complex is largely stabilized, explaining why inhibitor binding to ABL has three orders of magnitude higher affinity compared to SRC.[38] It is apparent that understanding receptor and ligand flexibility is an essential part of selectivity design. Moreover, proteins rarely function in isolation but rather continuously interact with multiple partners to carry out their diverse tasks. This implies that proteins must possess a certain degree of flexibility to adjust for all necessary interactions. Structure simulations nowadays (including flexible docking[39]) can account for protein flexibility and help decipher other modes of targeting proteins, for example, by allosteric regulation. Essentially, allosteric regulation involves ligand binding in the sites other than active, which in turn causes variations in protein motion (can be a conformational change into the inactive state), often associated with an inhibitory effect. Allosteric targeting can be very advantageous for resolving target specificity: for protein families with conserved active sites (e.g., protein kinases or G-protein-coupled receptors), allosteric sites are usually less constrained, opening avenues for selective targeting of specific proteins.[40] Additionally, allosteric targeting can provide an alternative solution for proteins for which targeting active sites proved laborious.

To sum up, computational tools, including modeling and docking, can nowadays account for protein motions and interactions; they currently find wide applications in medicinal chemistry drug research, providing additional unique insights into protein function, molecular recognition, and small molecule ligand selectivity to facilitate drug design.

2 Proteomic profiling and target deconvolution techniques establish the targets and off-targets of drugs

The chemical proteomics field combines chemical tools and biorthogonal techniques in an effort to discover new drugs and agnostically identify their targets. As mentioned, most drugs exert their action not only by engaging a single target but rather through complex perturbations of biological systems and polypharmacology. Mass spectrometry-based chemical proteomics, in particular, offers a range of unbiased quantitative assays for target deconvolution, some of which will be discussed hereafter.

2.1 MS-based proteomics techniques for target deconvolution

In recent years, two major groups of mass spectrometry-coupled approaches have emerged as powerful tools for target deconvolution. "Label-free" techniques (DARTS[41], Lip-MS[42], SPROX[43], CETSA/TPP[44,45]) measure changes in protein biophysical properties upon ligand binding and require the quantitative analysis of complex proteomes at sufficient depth. On the other hand, AfBPP[46], PAL[47–49], and ABPP[49–51] rely on chemical probes that enrich protein targets that are then identified and quantified (Fig. 4). In order to place the AfBPP, the main approach used in this work, within the panel of other widely applied MS-coupled target deconvolution methods, each of them will be discussed shortly below.
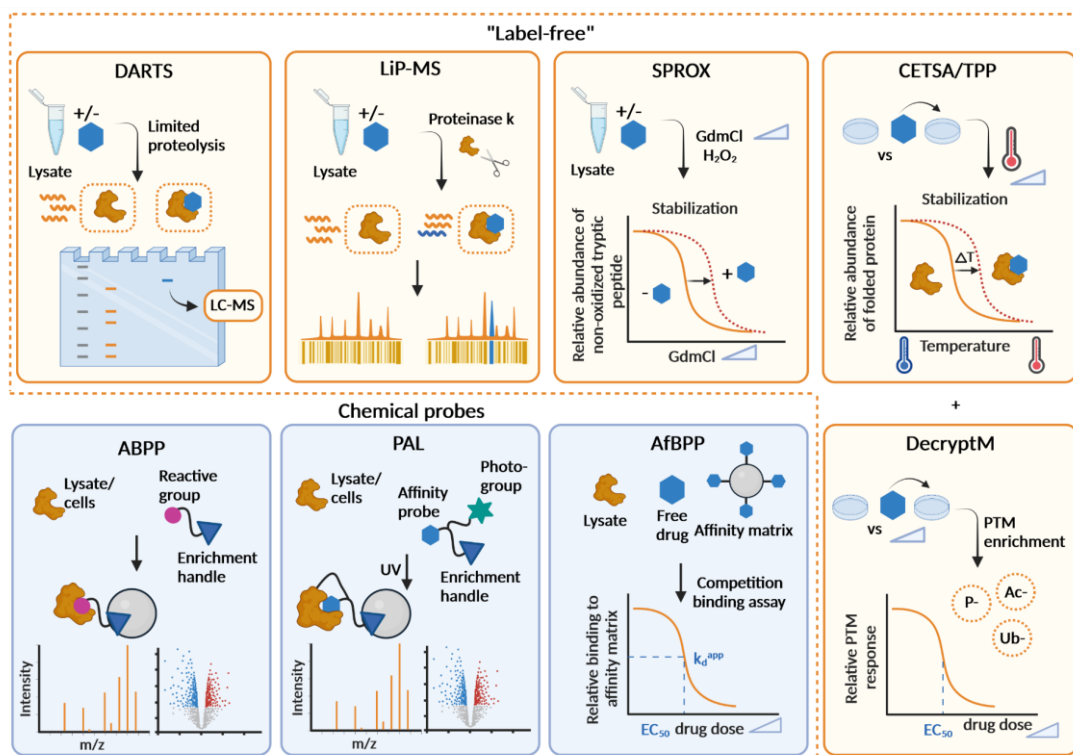
Figure 4. MS-based proteomics target deconvolution methods. Figure created with Biorender.com

Drug Affinity Responsive Target Stability (DARTS) assay is based on protein thermodynamic stabilization upon ligand binding and consecutive resistance to proteolysis. DARTS is performed

by treating cell lysates either with the compound of interest or vehicle control, followed by limited protein digestion with proteases such as thermolysin or Pronase. Subsequently, the samples are separated by SDS-PAGE and stained to identify protein bands protected from proteolysis by the small molecule. Individual bands are then analyzed by MS to identify corresponding proteins.[41,52]

Limited proteolysis coupled to mass spectrometry (LiP-MS) is another structural proteomics method that relies on the limited proteolysis by the sequence-unspecific protease proteinase K (PK). In a typical Lip-MS experiment, lysates are treated either with the compound of interest or vehicle control and subjected to partial digestion with PK. The binding of the small molecule can alter the protease accessibility by occupying the binding site, changing protein rigidity, or altering the tertiary and quaternary structure, thus producing condition-specific cleavage products. The fragments are further digested by trypsin, and peptide abundances are analyzed by MS, yielding structural fingerprints for all detected proteins. These fingerprints are used to identify the regions where structural changes occur.[42,53]

The Stability of Proteins from Rates of Oxidation (SPROX) is one of the protein stability methods focused on identifying and quantifying methionine-containing peptides. It assesses the thermodynamic properties of proteins and protein-ligand complexes via the hydrogen peroxide–mediated oxidation of methionine residues as a function of the chemical denaturant concentration. Chemical denaturant (e.g., guanidine hydrochloride or urea) introduced to lysates shifts the equilibrium between the folded protein, where the methionine residues are buried in solvent inaccessible-regions, and its unfolded form. In the unfolded protein, methionine residues are solvent-exposed and can undergo oxidation with hydrogen peroxide. Methionine oxidation, in turn, can be further mediated by the presence of a ligand. SPROX experiments allow determining the folding free energy of proteins and the K_D values of protein-ligand complexes.[43,54]

Another protein stability method used for unbiased identification of direct and indirect drug interactors is a cellular thermal shift assay (CETSA) coupled to MS-readout termed thermal proteome profiling (TPP). This method capitalizes on resistance to heat-induced protein unfolding upon ligand binding. In a TPP experiment, samples are treated with the compound of interest or vehicle control, followed by heat exposure in a range of temperatures. Soluble proteins are then collected and processed in a typical bottom-up proteomic workflow (see following chapters) to afford the so-called "melting curves". Ligand binding typically stabilizes the protein, consecutively shifting the melting temperatures to a higher range. The major advantage of this technique compared to any aforementioned method is the ability to assess the drug binding not only in lysates but also in living cells as well as *in vivo*.[44,45,55] However, it is important to note that the meltome has so far been fairly poorly characterized, and it is unclear which targets are truly stabilized by the ligand binding and what stabilization changes for each particular target are "significant enough" to manifest the interaction.

A newly developed by the group "label-free" technique to functionally annotate proteomes and characterize a drug's action, DecryptM, analyzes the cellular response upon drug perturbation through the dose-dependent effects on post-translational protein modifications (PTMs). Here, cells are treated with increasing concentrations of a drug or vehicle control, and the respective changes in the levels of PTMS (i.e., phospho-, acetyl-, or ubi-) are analyzed after the MS readout.[56]

Activity-based protein profiling (ABPP) is one of the chemical proteomic approaches that uses small-molecule chemical probes to understand the interaction mechanisms between compounds and targets. A typical ABPP probe contains a reactive group that directly (covalently) interacts with target proteins, a linker, and a reporter group, introduced to either visualize or enrich labeled proteins. Commonly used reporters are either fluorescent groups for rapid gel screening and monitoring of small molecules localization in cells or animals, or chemical moieties, which can be modified by click chemistry (alkyne or azide) for protein enrichment and MS detection to identify target proteins.[49–51]

Photo-affinity labeling (PAL) uses affinity probes functionalized with an affinity warhead that directs the probe into the pockets of protein targets, an enrichment handle, and a photocrosslinking group. Upon light activation, photo groups (e.g., diazirines, benzophenones, aryl azides) generate highly reactive species that presumably unselectively insert into proximal molecules, resulting in a covalent modification. Probe-bound proteins are then enriched out of whole proteomes and subjected to MS analysis.[47,48]

Affinity-based proteome profiling (AfBPP) uses affinity matrices to enrich protein binders out of complex proteome mixtures that are further analyzed by MS. Such affinity matrices can either be 1) solid support (beads) functionalized with a linkable analog of a drug (classic target deconvolution)[49] or 2) a mixture of promiscuous affinity probes covering an entire or at least a large part of a sub-proteome of choice (e.g., kinome and kinobeads, HDACs and corresponding beads).[57,58] Since AfBPP was the main method utilized in this work, it will be further discussed in more detail in the following chapter.

The main advantage of the aforementioned "label-free" MS-coupled target deconvolution methods is the ability to use native unmodified small molecules. However, they require significantly more MS measurement time. At the same time, the assumed biophysical property changes that evidence the ligand-protein interaction are by far not always measurable in such complex systems and/or represent a minor fraction of the analyzed proteomes. Meanwhile, probe-based techniques circumvent this issue by virtue of the target enrichment step. Ironically, the main bottleneck of the latter approach lies in the design and synthesis of probes. These probes are expected to bind and enrich all protein targets of the molecule/drug whose target space is being investigated. Commonly affinity probes are prepared by selectively installing enrichment handles on the molecule by means of chemical synthesis.

The obtained pure and characterized linkable analog can then either be directly immobilized on beads or functionalized with another enrichment-enabling moiety. Here, the lack of binding information or SAR hinders the rational installation of an enrichment handle, linker, or a photo-group. As a result, the ability of the obtained probe to engage the unknown targets may be compromised. Available data on the small molecule (SM) binding mode and/or computational analysis of possible interactions with protein targets can vastly facilitate the preparation of SM-affinity probes. Crystal structures of ligand-protein complexes reveal the SM solvent-exposed regions that are ideal for the installation of an immobilization/enrichment handle or additional conjugation groups (e.g., photoreactive groups for PAL) with full preservation of the protein binding capability. In the absence of either, one can assume that the functionalization of SM can

partially or entirely impair its ability to engage proteins. In theory, a number of SM analogs have to be synthesized and evaluated for their binding capability, which can only be achieved for molecules with the previously delineated binding profiles or a target hypothesis. Additionally, the high structural complexity of the studied bioactive molecule (e.g., bioactive natural products) or drug candidate may render the chemical synthesis of probes especially laborious. On the other hand, not all complex interactions can be monitored with chemical probes. Some target interactions may be too transient to capture, or the drug needs to be metabolized to exert its bioactivity. While chemical probes can facilitate the delineation of selectivity profiles, whether or not they can be implemented is, unfortunately, quite compound- and target-specific.

In summary, it is important to note that none of the described experimental proteomic methods is universal, none of them can fully account for and monitor conformational changes upon small molecule ligand binding that can contribute to the differences in measured binding affinities. However, they definitely are capable of manifesting the binding event and are great tools to shed light on compound selectivity. Generally speaking, all described methods are complementary and can enrich the insights on drug target profiles when combined.

2.2 Competition binding assay for chemoproteomic target deconvolution

Considering issues associated with affinity probe/matrix design, affinity pulldown coupled with MS readout is still a very attractive target deconvolution approach. Potent bioactive molecules are expected to show strong binding affinities for their protein targets; therefore, an affinity matrix based on the chemical structure of the compound in question can directly enrich molecular targets out of biological samples (lysates), which can then be further identified and quantified agnostically with unbiased mass-spectrometry readout. The use of probes prepared based on the structure of the analyzed compound falls into the context of so-called classic target deconvolution with tailored affinity matrices. The main advantages of such an assay are 1) reduced sample complexity and, consequently, reduced MS measurement time, and, 2) more importantly, an affinity matrix can be set to compete for target binding with the unmodified molecule. Meaning that the competing molecule will occupy the binding pocket of a protein target and, therefore, will not be available for bead enrichment. This assay performed with a range of competitor concentrations yields dose-response competition curves that characterize binding strength (Fig. 5a). These specific binders can then be annotated as potential targets and off-targets of the analyzed compound, providing selectivity profiles across all bead-bound proteins.
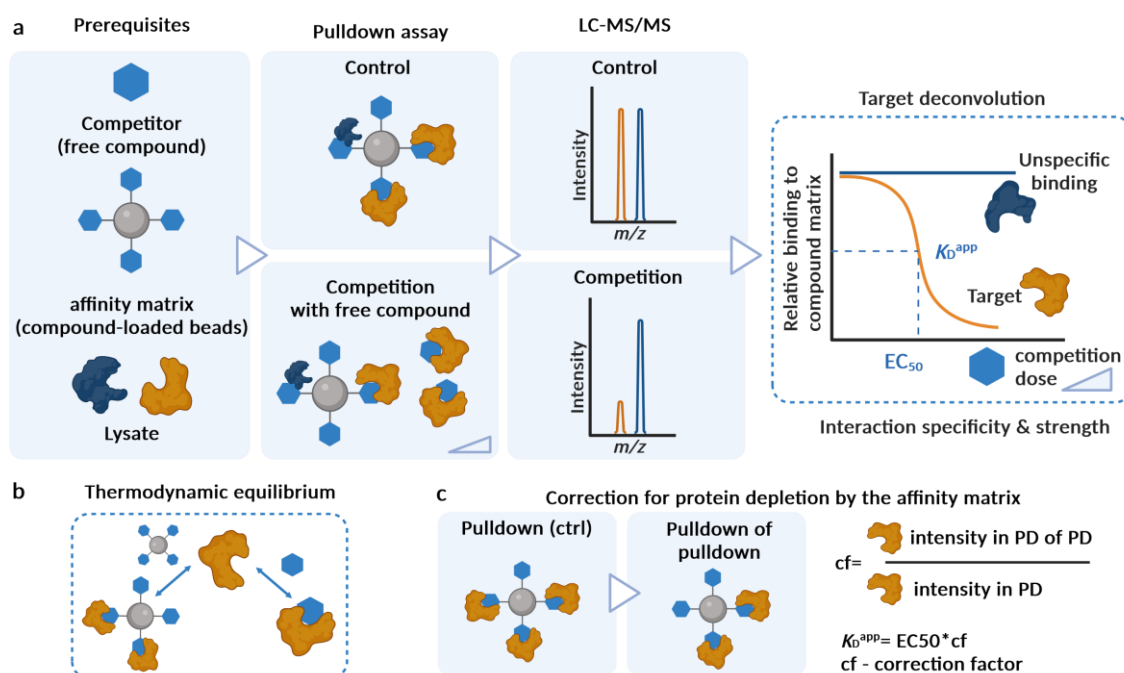
Figure 5. Dose-dependent competition pulldown experiment. a) Schematic representation of the binding assay with increasing concentrations of a profiled molecule; EC$_{50}$: effective concentration of natural product necessary to reduce protein binding to beads by 50%, K_D^{app}: an apparent binding constant that measures the affinity of observed interaction. b) Thermodynamic equilibrium between the protein, protein-affinity matrix, and protein-ligand complexes. c) Calculation of correction for protein depletion by the affinity matrix. Figure created with Biorender.com

In order to quantify the strength of observed binding interactions in such an assay, certain kinetics and thermodynamics relationships must be considered. The kinetic and thermodynamic considerations described below have been previously in part discussed in the Ph.D. thesis of Dr. Stephanie Wilhelm.[59]

A classic Michaelis-Menten binding kinetics of a protein (P) and ligand (L) association is described with the following equation:

$$P + L \rightleftharpoons PL \qquad (4)$$
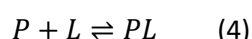
where PL represents the protein-ligand complex, whose formation and dissociation is determined by the kinetic rate constants k$_{on}$ and k$_{off}$:

$$P + L \xrightarrow{k_{on}} PL \text{ and } P + L \xleftarrow{k_{off}} PL \qquad (5)$$
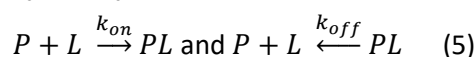
At equilibrium, the binding reaction is balanced by the dissociation:

$$k_{on}[P][L] = k_{off}[PL] \qquad (6)$$

where square brackets represent the concentrations at equilibrium. The equilibrium binding constant is then defined as:

$$K = \frac{k_{on}}{k_{off}} = \frac{[PL]}{[P][L]} = \frac{1}{K_D} \quad (7)$$

or

$$K_D = \frac{k_{off}}{k_{on}} = \frac{[P][L]}{[PL]} \quad (8)$$

where K_D is a dissociation constant. Based on these assumptions, the fast binding accompanied by a slow dissociation will result in a low dissociation constant and, respectively, high binding affinity.

Affinity probe (AP) disrupts the two-state equilibrium between the protein P and the profiled ligand L by depleting the protein from lysate (Fig. 5b). Therefore, a direct dissociation constant of the P-L interaction cannot be correctly estimated. Moreover, this assay does not allow measuring individual protein [P] and the protein-ligand complex [PL] concentrations. However, the assay-dependent half-maximal effective concentrations or EC_{50} values of a three-member equilibrium (affinity probe, protein, ligand) can be directly derived from dose-response curves. Here, measured by MS protein intensities are assumed to be a proxy of the actual protein concentrations in the analyzed sample. Bead binding removes an unknown fraction of the protein from the binding equilibrium with the ligand, which causes the re-equilibration of the remaining protein P and ligand L. Assuming that the binding affinity (K_D) between P and L should remain unchanged, the re-equilibration of [P] would cause a shift in the measured EC_{50} values (according to eq.8), as the higher [L] will be required to achieve the same K_D value. The degree of depletion is a function of protein concentration in lysate, protein affinity to the probe functionalized on beads, and the apparent loading (concentration) of the probe on beads. Even though the individual protein and complex concentrations cannot be easily estimated in this assay, the degree of protein depletion can be assessed by performing a sequential pulldown (pulldown of pulldown, or PD of PD) in the lysate that has already been subjected to the incubation with the affinity matrix (Fig. 5c). The correction for protein depletion should allow for accurate estimation of dissociation constants and is calculated as

$$cf = \frac{P \text{ intensity in sequential pulldown (PD of PD)}}{P \text{ intensity in pulldown (PD)}} \quad (9)$$

where cf stands for correction factor.

The Cheng-Prusoff[60] relationship between the experimentally derived EC_{50} and K_D values for reversible competitive inhibitors (L) is defined as:

$$K_D(L) = \frac{K_D(AP)}{K_D(AP)+[AP]} * EC_{50}(L) \quad (10)$$

where [AP] stands for affinity probe loading on beads. The binding affinity of a protein to the probe immobilized on beads ($K_D(AP)$) is unknown, however, the equation 10 can be transformed to account for protein depletion by the beads instead (that is experimentally determined according to the eq. 9). The binding affinity of a protein P to the probe AP on beads is defined as:

$$K_D(AP) = \frac{[P][AP]}{[PAP]} \quad (11)$$

where [PAP] reflects the amount of protein bound by the affinity matrix. The competition binding assay is performed under the premise that the affinity matrix (AP) binds only a fraction of a total available protein amount [P]$_{total}$:

$$[P]_{total} = [P] + [PAP] = ([P]_{total} - [PAP]) + [PAP] \quad (12)$$

$$or$$

$$[P]_{total} = (1 - f(P)) + f(P) = 1 \quad (13)$$

It is important to note that this assumption is only valid if the protein amount is greatly exceeding the amount of the affinity probe; therefore, the protein is not greatly depleted out of the binding equilibrium with the ligand. Based on that, one also assumes that the affinity matrix captures the same amount of protein P in the first pulldown as in the sequential pulldown of pulldown, however, the amount of protein available for the PD of PD is depleted after the first pulldown by [PAP]:

$$cf = \frac{[PAP([P]_{total} - [PAP])]}{[PAP]} = [P]_{total} - [PAP] \quad (14)$$

$$or$$

$$cf = 1 - f(P) \quad (15)$$

The concentration of the probe [AP] can thus be defined as:

$$[AP] = K_D(AP) * \frac{[PAP]}{[P]_{total}} \quad (16)$$

$$or$$

$$[AP] = K_D(AP) * \frac{f(P)}{1 - f(P)} = K_D(AP) * \frac{1 - cf}{cf} \quad (17)$$

Then the ligand dissociation constant (eq. 10) is determined as the following:

$$K_D(L) = \frac{K_D(AP)}{K_D(AP) + (K_D(AP) * \frac{1-cf}{cf})} * EC_{50}(L) = \frac{1}{1 + \frac{1-cf}{cf}} * EC_{50}(L) \quad (18)$$

$$or$$

$$K_D(L) = cf * EC_{50}(L) \quad (19)$$

where both parameters are determined in the competition assay: the EC$_{50}$ values are directly derived from competition curves, and correction factors are calculated as ratios of protein intensities from two consecutive pulldowns in the same lysate.

The binding event always occurs in some sort of solvent. From the thermodynamics perspective, the driving forces that enable the protein-ligand binding result from various affine interactions and energy exchanges between the protein, ligand, water, and buffer ions. A thermodynamic potential that quantifies the capacity of a system to do maximum or reversible work at isothermal and isobaric conditions and characterizes the driving forces of the interaction is Gibbs free energy. The binding can only occur if a system's change in Gibbs free energy (ΔG) at equilibrium state

(isobaric, isothermal) is negative. Therefore, ΔG has a great influence on the stability of any interaction and binding affinity.[35] The standard binding free energy ΔG^0 measured at 1 atm pressure, 298 K, and effective protein and ligand concentrations of 1 M is related to the dissociation constant as the following:

$$\Delta G^0 = -RTlnK = RTlnK_D \quad (20)$$

where R is the universal gas constant. From eq. 20, it becomes clear that the lower the dissociation constant, in other words, the higher the binding affinity, the more negative is the ΔG^0. Thus, the association and dissociation rates, along with binding and dissociation constants, determine the stability and the affinity of the protein-ligand binding interaction.

The binding free energy (ΔG) at any moment of time can be defined with enthalpic and entropic contributions as the following:

$$\Delta G = \Delta H - T\Delta S \quad (21)$$

where ΔH and ΔS are respective changes in the enthalpy and entropy upon ligand binding. Binding enthalpy ΔH can be treated as the changes in energy of a system upon formation and breakage of noncovalent interactions during the binding event between the protein, the ligand and the solvent. Entropy, in turn, is a measure of heat distribution within the system. The second law of thermodynamics states that heat flows spontaneously from higher to lower temperature regions within one system. Since this process, in general, lowers the degree of order of a system, entropy can be treated as a measure of disorder. ΔS, thus, describes the overall changes in degrees of freedom in a system and can be defined with three separate terms as:

$$\Delta S = \Delta S_{solvent} + \Delta S_{conformation} + \Delta S_{r-t\ degree\ of\ freedom} \quad (22)$$

where $\Delta S_{solvent}$ describes entropy changes after solvent release due to ligand binding, $\Delta S_{conformation}$ represents changes in conformational freedom for ligand and protein upon binding, and $\Delta S_{r-t\ degree\ of\ freedom}$ defines losses in translational and rotational degrees of freedom of a compound upon complex formation.

The binding assay utilized in this work is largely affected by the differences in thermodynamic and kinetic properties of profiled free and functionalized on beads compounds. The binding of the free compound to a protein causes losses of degrees of freedom. Meanwhile, the compound immobilized on beads by default possesses less rotational and translational freedom due to the immobilization step itself, prior to any protein binding. As a result, the binding of these two compounds (free vs. more rigid immobilized) will have different entropic contributions to the Gibbs free energy, with the immobilized compound having a lower entropic penalty. Different ΔS will lead to two unequal ΔG (eq. 21) for the free and immobilized compound, implying different affinities for binding of the same protein (eq. 20). In the context of ligand binding kinetics, it has been previously shown that more rigid compounds have increased association rates and slower

dissociation.[61] It is likely that the immobilization of a compound on beads will affect its binding kinetics in a similar fashion. Moreover, the association and dissociation rates define the residence time of a ligand in a binding site that can greatly differ for different protein-ligand pairs. Residence time τ negatively correlates with the dissociation rate of a protein-ligand complex:

$$\tau = \frac{1}{k_{off}} \quad (23)$$

For clinical drugs, the long residence time may be of greater importance than the binding affinity itself as it can prolong the therapeutic effect and be beneficial for treatment outcomes or allow for lowering the treatment doses. In the context of compound immobilization, one may speculate that a slower dissociation rate of the affinity probe-protein complex will result in prolonged residence time and/or affect the establishment of binding equilibrium, thus affecting the experimental measurement of dissociation constants. In order to compensate for possible kinetic consequences of compound immobilization, a pre-incubation step of the free molecule with lysate is performed prior to bead enrichment.

To conclude, although such competition binding assay offers comprehensive selectivity profiles among all bead-bound proteins, it can unfortunately not account for conformational changes and kinetics of observed binding interactions. As discussed, the kinetics of affine interactions is not only a function of concentrations but is also greatly influenced by protein dynamics and association/dissociation rates. Proteins are dynamic; they exist in equilibrium between different states and often undergo conformational changes to perform a biological function or upon ligand binding. This binding assay is not capable of distinguishing the cases when 1) the ligand binding blocked the active site and caused the inhibition of the enzymatic activity and/or 2) induced a conformational change that prevents the protein from functioning and explains measured high affinity of the interaction or 3) the binding does not affect protein function/activity at all. Thus, all observed binding events must be validated with orthogonal techniques. It is important to remember that the ligand binding alone does not imply the consequent changes in protein function. Similarly, additional structural analyses, kinetics, or protein dynamics experiments would have to be carried out to explain the differences in observed binding affinities.

3 Mass spectrometry-based quantitative shotgun proteomics

Proteomics, in general, deals with the determination of gene and cellular function at the protein level. Microscopic and cellular imaging, chip arrays, and genetic interaction experiments are widely applied approaches within the field. Mass spectrometry (MS)-based proteomics has emerged as a versatile technique that allows for unbiased quantitative analysis of complex proteomes, protein interactions, and modifications and proves to be an indispensable tool for molecular and cellular biology.[62,63] Mass spectrometers measure the molecular weight of the analyte as a mass-to-charge ratio and consist of three main building blocks: an ion source that converts analyte molecules into gas phase ions, a mass analyzer that separates analytes based on m/z values, and a detector that records the number of ions at each m/z value.[64]

3.1 Typical bottom-up proteomics workflow applied in this work

Two alternative approaches are used in MS-based proteomics: either intact proteins are analyzed in so-called top-down proteomics, or in bottom-up (or shotgun) proteomics proteins are identified and quantified from their peptides.[65] Since solely the bottom-up proteomic approach was used in this work, it will be detailed hereafter. In a typical bottom-up workflow that includes a protein enrichment step (Fig. 6.a), proteins to be analyzed are isolated from tissue samples or cell lysates, or parts of the proteomes are enriched with the help of chemical probes, and enzymatically digested, commonly by trypsin. Trypsin is a sequence-specific protease that cuts peptides at the carboxyl side of lysine and arginine, generating peptides with C-terminally protonated lysine or arginine residues providing an advantage for subsequent peptide sequencing. In order to gain analytical depth, obtained peptide mixtures are typically decomplexified prior to the MS analysis by means of liquid chromatography (LC). Coupling of a reversed-phase high-pressure liquid chromatography system (RP HPLC) to a mass spectrometer affords a so-called online separation of peptides, where peptide eluates are directly subjected to the MS analysis. Such setup is typically used to separate peptides based on their hydrophobicity: the LC is equipped with a column loaded with a non-polar stationary phase (most commonly hydrophobic C18 chains), while a polar mobile phase contains an amphiphilic ion-pairing reagent like formic acid. Hydrophobic amino acids of peptide chains interact with the stationary phase and, therefore, retain on the column, while the polar peptide chains engage in interactions mediated by the ion-pairing reagent, which further increases the resolution depth. Elution from the column is achieved by the mobile phase gradient with increasing organic solvent content (typically acetonitrile). More hydrophilic or polar peptides elute first, while more hydrophobic or non-polar peptides get retained on the column and elute later with higher concentrations of acetonitrile. MS measurements are carried out in the gas phase of ionized analytes; therefore, eluted peptides are delivered to the mass spectrometer interface and undergo electrospray ionization (ESI) at the ion source. Peptides elute from an electrospray capillary (emitter) under high voltage and ionize, forming a Taylor cone (jet). The jet further disperses into single droplets (plume), enabling ions to transfer into the gas phase. Upon mobile phase evaporation and droplets fission, charged peptides eventually enter the MS.
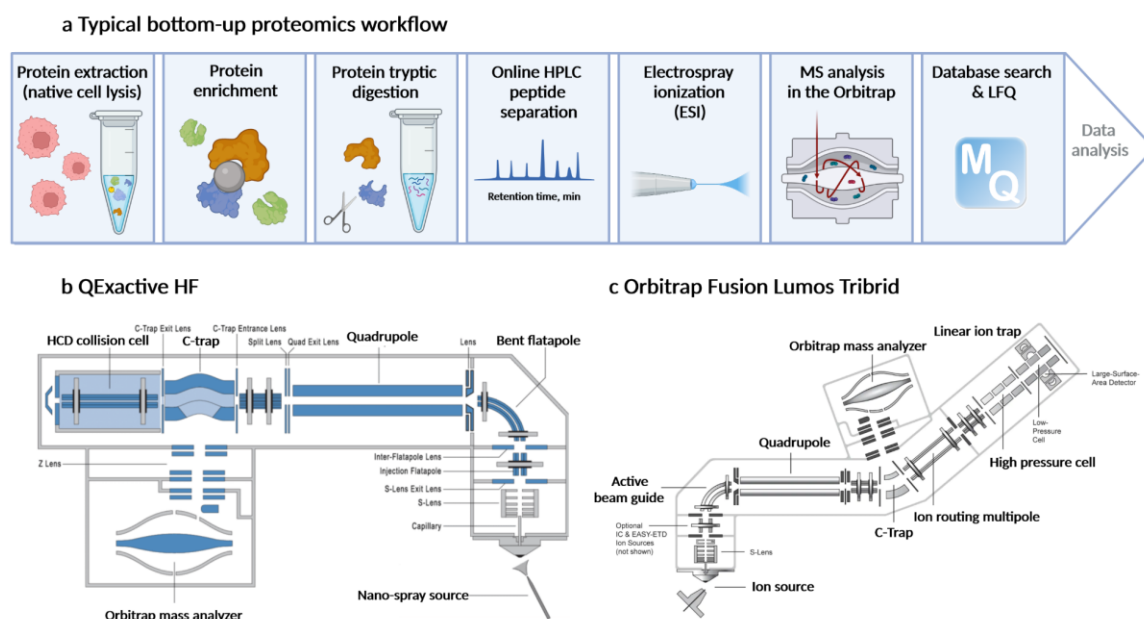
Figure 6. Bottom-up proteomics workflow and MS instruments used in this study. a) Cells or tissues are lysed under non-denaturing conditions, and proteins are enriched using affinity matrices, followed by a tryptic in-solution digestion. Obtained peptide mixtures are separated by online liquid chromatography and ionized via electrospray (ESI). b) Schematic representation of Thermo Fisher QExactive HF[TM]. c) Schematic representation of Thermo Fisher Orbitrap Fusion[TM] Lumos[TM] Tribrid[TM] utilized for data acquisition in this work. Schemes adapted from Thermo Fisher Scientific, figure created with BioRender.com

3.2 Data-dependent acquisition of the Orbitrap mass spectrometer

In tandem or shotgun proteomics, the full mass spectrum of all peptides eluting at a time point is acquired (MS1 spectrum). Then, a defined number of precursor peptides are selected for consecutive fragmentation, usually based on their abundance, and a series of tandem or MS/MS scans are recorded for each of the isolated precursors (Fig. 7). In LC-MS/MS data-dependent acquisition (DDA), a selected number of most abundant precursors within a certain mass range is picked for fragmentation, and the MS instrument records cycles of MS-MS/MS spectra along the LC gradient.

Precursor ions can be fragmented into different types of fragment ions, depending on the fragmentation techniques applied.[66] However, for Orbitrap-type MS instruments in this study, collision-induced dissociation (CID) and higher-energy collisional dissociation (HCD) are the most commonly used, where fragmentation is afforded either by energetic collision with inert gas (HCD) in the collision cell or by internal destabilization of the analyte molecules and dissociation upon collision with other analyte molecules (CID) in the ion trap.[67] HCD fragment spectra are always recorded in the Orbitrap, while CID spectra are normally analyzed in the ion trap. Both of these fragmentation techniques yield predominantly b- and y-ion series upon the peptide bond breakage between the carbonyl- and the amino- groups of neighboring amino acids (Fig. 7c). Y-ions carry the charge on the peptide C-terminus, whereas b-ions possess charged N-terminus. Consequently, the distance between neighboring peaks in the MS2 spectrum corresponds to the mass of a single amino acid and is used to reconstitute the amino acid sequence and identify the peptide.

In a full scan (MS1) mode, ionized peptides are captured and focused for effective ion transmission using a series of lenses (S-lens, S-lens exit lens, injection flatapole, inter-flatapole lens). Uncharged neutral species are filtered out in the bent-flatapole, while all charged peptides are passed through the quadrupole (Fig. 6b). Ions are collected into packets and stabilized in the C-trap, and sent to the Orbitrap for detection. In an Orbitrap analyzer, ions are trapped in orbital motion, where the frequency of rotation is related to the mass-to-charge ratio, i.e. ions with different oscillation frequencies have different m/z ratios. Frequency measurement undergoes an internal Fourier transformation to afford a full scan or MS1 spectrum of all peptides eluting at this time point.[68]

In the MS/MS or tandem analysis, upon ionization, stabilization, and filtration of neutral species, only the ions of specific m/z, typically the most abundant ones in the preceding full scan, are being selected and isolated by the quadrupole. Isolated ions are then sent to a collision cell for collision-induced fragmentation into fragment ions. Fragment ions are collected into packets in the C-trap, stabilized and sent to the Orbitrap for detection. This process is repeated for all ions to be fragmented. Samples prepared in the course of this work were analyzed on QExactive HF and Orbitrap Fusion Lumos. Tandem spectra for QExactive HF are acquired in the Orbitrap analyzer, while Fusion Lumos Tribrid is additionally equipped with an ion trap and an ion routing multipole, which enables the recording of tandem spectra in both Orbi- and ion traps (Fig. 6.b,c).[68–70]
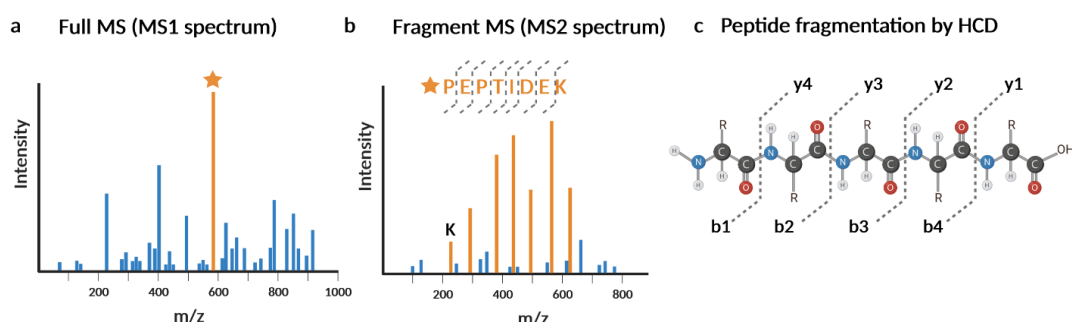
Figure 7. Tandem bottom-up (shotgun) proteomics. a) When ionized peptides enter the MS instrument, the full MS or MS1 spectrum is acquired b) Precursor peptide from MS1 spectrum is selected for consecutive fragmentation, and a series of tandem or MS2 spectra are recorded c) higher collisional dissociation (HCD) fragmentation of peptides generates b- and y-ion series. Figure created with BioRender.com

3.3 Protein identification with database search

In shotgun DDA proteomics, proteins are typically identified and quantified from their peptides via database search.[71] This process can be described as a matching experimentally acquired fragment spectra to theoretical spectra computed from *in-silico* digestion of protein sequences derived from genomic or transcriptomic data.[72] The major limitation of this identification approach arises from the fact that such databases can only be generated for well-sequenced organisms. Multiple search engines enable database search, in the course of this study, the MaxQuant software with embedded probability-based Andromeda search engine was utilized.[73–75]

Typically, experimental spectra are pre-processed by the search engine, including the improvement of signal-to-noise ratios, de-isotoping, and peptide charge stage deconvolution. Later, the pre-processed experimental MS2 spectra are matched to the *in-silico* generated

fragment spectra within the pre-filtered range of respective peptide precursor masses. Such exercise yields peptide-spectrum-matches (PSMs), which are characterized by probability scores (computed probability of the identified match being a random event). Such scores are commonly used as a measure of confidence for PSM identifications, where high scores mean a low probability of false identification.

In order to control for the wrong (random) matches, experimental fragment spectra are also searched against a collection of *in-silico* truncated and/or reversed peptide sequences, a so-called decoy database. The estimated false-discovery rate (FDR, usually a ratio of decoy hits to target hits) improves the overall identification quality.[76] It is also important to note that setting correct search parameters is essential to ensure the accuracy of obtained identifications. Digestion enzyme, fixed and variable peptide modifications, precursor ions fragmentation type, and mass error tolerance have to be specified to efficiently narrow the search space and minimize the risk of false/random matches.

Obtained PSMs are further assigned to corresponding proteins. Only unique peptides can unequivocally identify a particular protein; for peptides present in several related proteins and/or protein isoforms, the so-called shared or razor peptides, distinct protein identification is difficult. Usually, such cases are reported as protein groups, which include all proteins that share the identified peptide sequence.

3.4 Label-free protein quantification

Protein quantification is a process that infers the abundance of proteins in the analyzed sample and can be performed with one of the well-established label-based or label-free approaches.[77] In the course of this study, proteins were quantified by means of MS1-based label-free quantification (LFQ).

Label-free methods allow relative quantification of abundance changes in two or more biological samples. Generally, two major groups of label-free approaches are divided by the method used for data extraction. In so-called spectral counting methods, the quantification can be inferred by counting the number of peptides or spectra assigned to a given protein. Relative quantification is achieved by comparing the number of identified MS/MS spectra based on the assumption that an increase in protein abundance typically results in an increased number of its proteolytic peptides and, consecutively, an increased number of total MS/MS spectra for each protein.[63]

Another group of methods, particularly widely applied for data generated via LC-MS/MS, measures quantitative values by extracting the area of the precursor ions' chromatographic peaks - the area under the curve (AUC). It has been observed that signal intensity from ESI correlates with ion concentration. The peak area of a particular m/z from a mass spectrum reflects the number of ions for that m/z detected by the mass spectrometer at any given time. In other words, intensity-based LFQ infers protein quantification from signal responses of intact peptides. Previous studies have demonstrated that the correlation between signal response and concentration of measured peptide remains linear over four orders of magnitude.[63,78,79] Although acquired peptide MS intensities can not be used for absolute quantification due to the dependence of an individual signal on peptides' biophysical properties (hydrophobicity, acidity, or ability to ionize), the behavior of an individual peptide can be assumed to remain identical

between different MS measurements, and, therefore relatively quantified (termed MaxLFQ in MaxQuant).[80]

In MaxQuant software, peptide intensities are normalized using a sophisticated algorithm of delayed normalization, where the function calculates the normalization factors from summed logarithms of peptide pairwise ratios between different samples.[80] In a nutshell, after normalization, the total sum LFQ intensities of different samples are very comparable. This algorithm is based on the assumption that the global protein population between all analyzed samples is similar; therefore, such normalization is not possible for very different samples. For computationally expensive normalizations (a large amount of samples, >10) MaxQuant utilizes a Fast LFQ algorithm, which basically only uses a subset of pairwise comparisons.[80]

Although straightforward at first glance, such a quantification method heavily depends on several technical aspects. Elution time offsets (differences in peptide retention times between samples) and irreproducibility of DDA acquisition (the same peptide is not selected for fragmentation in all of the samples) may drastically harden the quantification. To overcome these issues, a range of solutions can be applied: *in-silico* retention time alignment can be performed to facilitate the comparison of peptide elution profiles. A common method to tackle another LFQ constraint, i.e., the stochastic MS/MS sequencing, is the match between runs algorithm.[81] In brief, this algorithm analyses MS1 scans of aligned MS runs for matches of precursor masses, isotopic patterns, and charge states and assigns missing identifications for the runs where the peptide precursor was not picked for fragmentation.

Despite the limitations above, LFQ methods remain a very attractive alternative to label-based quantification as they have no limitations regarding the number and/or kind of samples. Furthermore, LFQ is time- and cost-efficient and applies to samples that can not be directly isotopically labeled. Additionally, LFQ does not suffer from the potential variabilities introduced by chemical labeling or tags.

4 Man-made and natural molecules to probe the druggable proteome

A "druggable" proteome is defined as a fraction of proteins that can be functionally modulated by a drug. To this day, the number of druggable proteins remains small: even though there are more than 20,000 canonical human proteins in the human proteome, the ChEMBL database features around 5,000 proteins with druggable pockets, and only less than a thousand of those are therapeutic targets of FDA approved drugs.[82] The challenge of identifying new druggable targets arises partially from the fact that medicinal chemistry is indeed based on the premise of a stable interaction between the pocket of a target and a ligand (lock-and-key). A number of computational modeling methods described above vastly facilitate drug design for "classical" target pockets. In cases when there is no "lock" to fill with the "key", i.e., the conventional catalytic binding site is not available, drug design quickly becomes challenging. Thus, a large number of proteins that were proven to play a pivotal role in diseases remained "undruggable" for a very long time, prime examples being transcription factors or various oncoproteins like MYC, Notch, or KRas.[83–85] nother bottleneck for expanding the druggable proteome lies in the difficulty of identifying and validating novel therapeutic targets. It appears much easier to exploit previously investigated and validated targets and develop new therapeutic agents to interrogate those than to establish and develop novel functional assays that would allow identifying novel unexpected disease regulators. As a result, HTS screens are typically launched around a relatively limited selection of targets, where the potential of screened molecular libraries to inhibit unexpected proteins is usually never addressed. Thus, since there is no apparent need to develop new assays and probes to validate novel enzymes as pharmacologically actionable, no screens are initiated to find agents for these potentially interesting targets. Finally, not all the results of individual phenotypic screens (especially in pharma) are publically available; therefore it is challenging to narrow down the initial screening libraries for each individual target. Correspondingly, it is not common knowledge which proteins can, in principle, be interrogated by the chemical scaffolds typically included in the bioactive compound screening collections or, at the very least, this information is not easily accessible and requires a lot of literature and public resources (such as PubChem) mining.[86]

Equally important, commercial drug sources oftentimes provide redundant or simply wrong information regarding the target space of particular molecules. For example, two commercial molecules, namely MC1568 and LMK235, are marketed as class IIa HDAC inhibitors. However, the recent study by Lechner et al. has demonstrated that these inhibitors show no class IIa target engagement at all, which highly questions their use as chemical probes.[58] Furthermore, this study is an excellent example of how chemical proteomics combined with unbiased mass spectrometry can help obtain the comprehensive selectivity profiles of drugs and expand the druggable proteome. Here, the metalloenzyme-enriching affinity matrices were used to establish the target landscapes of commercially available HDAC inhibitors, providing unprecedented insights into the selectivity of these drugs. One notable finding of this screen was the identification of a previously unknown common off-target of more than half of the analyzed HDACis – MBLAC2. This study vastly contributed to the functional characterization of the previously poorly studied enzyme MBLAC2 as well as shed light on the observed phenotypes upon treatment with HDACis that could not be explained by HDAC inhibition, and finally provided a useful biochemical tool that would allow following-up on the role of this enzyme in the accumulation of extracellular vesicles.

Another example when sub-proteome-enriching affinity matrices contributed to the expansion of the druggable proteome is the study by Klaeger et al., where the Kinobeads technology allowed screening the target space of more than 200 clinical kinase inhibitors.[30] Apart from unexpected kinase off-targets, this study also uncovered a previously unreported binder of 29 kinase-targeting drugs, ferrochelatase (FECH), that most likely explains the photosensitivity developed in patients treated with these inhibitors.[87] The original screen of clinical kinase inhibitors with Kinobeads was followed up in the study carried out by Reinecke et al., where more than a thousand tool compounds (published kinase inhibitor set (PKIS), PKIS2, kinase chemogenomic set, and Roche library) were screened for targets in a similar fashion.[88] This screen revealed hundreds of potential novel chemical probes for 73 different kinases. Several compounds were found to bind an understudied kinase PKN3 and were employed in the following functional phosphoproteomic assay to decipher the potential downstream substrates of PKN3. A medchem study carried out by Heinzlmeir et al. used the results of the Kinobeads screen to select compound leads for their further chemical optimization into selective EPHA2 inhibitors, where chemoproteomic readout served as a tool of the affinity and selectivity evaluation of newly synthesized molecules.[89] Obtained EPHA2 binding affinities of a novel compound set together with protein crystallography allowed to characterize the binding pocket of EPHA2 and decipher affinity- and selectivity-defining residues.[90]

In general, broad-enriching affinity probes, like Kinobeads or HDAC-beads, are a great tool to support different stages of drug discovery, i.e., identification of drug targets as well as lead selection and optimization. However, they can only be used under the premise that the matrix enriches all possible targets and off-targets of the profiled molecules. In cases when the matrix does not enrich the protein targets of the profiled molecule, the assay will not be able to identify such binding interactions. Similarly, when the targets are entirely unknown, using sub-proteome enriching probes might not be able to explain the observed cellular phenotype. The target landscape of the entire chemical space is unknown; therefore, if the target deconvolution has to be completely hypothesis-free, the use of tailored probes is inevitable. This is true for both man- and nature-made bioactive molecules.

Natural products, an evolutionary enriched pool of biologically active compounds, are a vital resource for drug discovery. The simple fact that NPs are produced in nature bioactive by default to serve a certain biological function provides endless inspiration for drug design and the development of novel therapeutic agents. Of all small molecule drugs approved between 1981 and 2019, nearly 70% were either native natural products, derivatives thereof, or NP-inspired molecules.[91] Numerous NPs have proven to be useful up to this day as antibacterial agents, anticancer drugs, and immunosuppressive therapeutics. High three-dimensional structure complexity and vast scaffold diversity, which make NPs so attractive for drug research, simultaneously are the major bottlenecks that often restrict their exploitation. NPs are frequently included in phenotypic drug screens, and many NPs with anti-cancer activities have been identified this way. Nevertheless, target deconvolution quickly becomes unattainable in the context of bioactive NPs, since often isolation of NP in the required quantity, chemical synthesis, derivatization and/or late-stage functionalization are arduous or simply impossible. Therefore, for a vast number of NPs, that possess a proven phenotypic effect, the mode of action is still poorly understood. A simple assay that would allow for agnostic "hypothesis-free" identification of NP targets without extensive chemical synthesis of probes is highly desirable.

Apart from agnostic target deconvolution, tailored probes have the potential to answer the question of what is "druggable" in principle. Since they should be able to engage all targets of the analyzed molecules, combined with MS-based chemical proteomics, they can decipher the proteome-wide chemical space of individual pharmacophores. This would imply a rather molecule-centric focus compared to traditional target- or phenotype-based approaches, where the main objective would be to uncover the selectivity potential of one chemical scaffold across the entire proteome rather than finding the selective agents capable of interrogating the targets of interest. In other words, one would ask which proteins one molecule can address instead of asking this molecule hundreds of times whether it can engage one particular protein. On a pharmacophore level, this can be done through by creating a focused library of closely structurally related analogs and respective tailored affinity probes, which together can be used in a chemoproteomic proteome-wide selectivity profiling assay. This profiling exercise should not only enable target identification of all the molecules and establishment of the structure-affinity relationships for all expected and unexpected protein targets, shedding light on selectivity but simultaneously allowing for pharmacophore repurposing and potentially unlocking the untargeted proteome.

5 Objectives and outline

Classic target deconvolution using tailored affinity matrices in combination with unbiased quantitative mass spectrometry is a powerful tool to identify protein-binding selectivity profiles of bioactive small molecules. Bioactive natural products are a vital source of inspiration for drug design. Nevertheless, the agnostic identification of protein targets of these compounds is hampered by the lack of general expeditious methods to probe for protein binding. On the other hand, target deconvolution with tailored probes should theoretically allow probing the integrity of chemical space for each individual evaluated molecule. Such a molecule-centric view on medicinal chemistry and deciphering small molecule bioactivity would allow for hypothesis-free target identification with the potential to shift drug discovery efforts from target- or phenotype-based focus toward a thorough investigation of general targeting abilities of different chemotypes. The main objective of this work is to prove that chemical proteomics can achieve this paradigm shift. The results chapter of this thesis is divided into two parts and discusses the novel molecule-centric approach.

The first part of this work was dedicated to setting up and evaluating the profiling assay that would allow for unbiased agnostic target deconvolution of natural products using tailored affinity probes. Multiple chemistries were initially evaluated for immobilizing unmodified natural products on beads, including gold-catalyzed alkoxylation and cycloaddition, ruthenium-catalyzed metathesis, Mitsunobu reaction as well as light-induced photo-immobilization. The latter, surprisingly, proved to be the most robust. Upon its establishment, the general experimental assay was systematically evaluated for its scope and applicability to target deconvolution of NPs of different reported bioactivity and biosynthetic origin (Part 1).

The second part of the work (Part 2), a proof-of-concept study, challenged mass spectrometry-based chemical proteomics to unravel the ability to address protein targets of the two privileged pharmacophores, namely 2,4,5-triarylimidazole and pyrimidopyridone. Two small focused libraries of analogs and respective tailored affinity matrices were chemically synthesized, and the selectivity profiles of each molecule against all bead-bound proteins were elucidated. Systematic exploration of the binding profiles allowed establishing proteome-wide structure-affinity relationships for obtained protein targets as well as identifying affinity and selectivity cliffs.

Experimental Procedures

Contents

6 Chemical synthesis of 2,4,5-triarylimidazole compound library

All chemicals and solvents were purchased from commercial suppliers (Sigma-Aldrich Co, Merck, VWR International, Carl Roth GmbH & Co.KG, Alfa Aeasar, Fluorochem Ltd) and were used without further purification. The progress of chemical reactions was monitored by thin-layer chromatography (TLC). Crude reaction mixtures were spotted on normal phase TLC plates (TLC Silica gel 60 F254, Merck) for separation. Spots were visualized with UV at λ = 254 nm and 365 nm. Reaction monitoring was additionally performed on an amazon speed ETD ion trap mass spectrometer coupled to an Ultimate 1100 HPLC system (Agilent Technologies). For sample preparation, crude reaction mixtures were diluted with 0.1% formic acid (FA) in acetonitrile (ACN) and filtered through a syringe filter. For analysis, a 10-90% gradient of solvent B (0.1 % FA in ACN) in solvent A (0.1 % FA in deionized water) was applied to a C-18 reversed-phase column for 30 min. Mass analysis was accomplished in the positive ion mode (ESI+). Flash chromatography was performed on an Interchim puriFlash evo 430 system. Chromatography signals were detected at λ = 254 and 365 nm. NMR spectra were recorded on a 400 MHz Spectrometer (Bruker) at the Chemistry Department of the Technical University of Munich. Chemical shifts were recorded in parts per million (ppm), and NMR signals were described as follows: s (singlet), d (doublet), t (triplet), q (quadruplet), and m (multiplet).

Briefly, triarylimidazole scaffold was obtained through the Debus-Radziszewski reaction of the diketone intermediate with cyanobenzaldehyde in the presence of ammonium acetate (Scheme 1.III); the nitrile was then further reduced into amine with lithium aluminium hydrate (Scheme 1.IV). Synthesis of diketone intermediates (Scheme 1.I+II) showed to be the most versatile step in the overall synthetic route and was performed following one of the procedures detailed below.

6.1 General chemical procedures used in this work
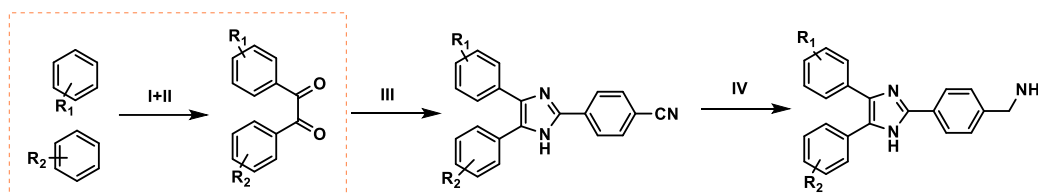
Scheme 1. General synthetic route for 2,4,5-triarylimidazoles

6.1.1 Synthesis of the 1,2-diketone intermediates: one-pot via Sonogashira coupling followed by internal alkyne oxidation (I+II one-pot)
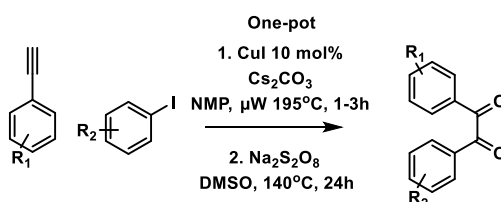
Scheme 2. One-pot procedure for 1,2-diketone intermediate synthesis

Equimolar amounts of alkyne and iodide starting material were dissolved in 10 mL NMP in a microwave vial. Two equivalents of caesium carbonate and 10 mol% of copper iodide were added under moderate stirring. The vial was sealed and heated to 195°C for 60-180 minutes under

microwave irradiation. After reaction completion, the alkyne intermediate obtained via Sonogashira coupling was further used without work up or purification. 10 mL of DMSO was added to a microwave vial followed by the addition of 2.5 equivalents of sodium persulfate. Reaction mixture was further stirred at 140°C in an oil bath for 24 hours. After reaction completion, the crude mixture was diluted with EtOAc and filtered through a patch of celite to remove insoluble salts and the catalyst. The mixture was further extracted with EtOAc and water. The combined organic layers were washed with water and brine, and dried over MgSO₄. After evaporation of solvents, the crude product was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

6.1.2 Synthesis of 1,2-diketone intermediates from ester enolates (I,II)
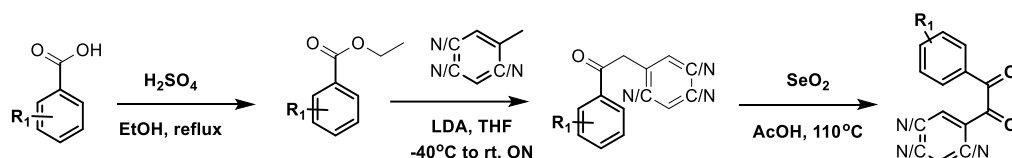
Scheme 3. Synthetic procedure via ester enolates

Benzoic acid starting material was dissolved in ethanol, concentrated sulfuric acid was added dropwise up to 10% final volume in the reaction mixture. Reaction was stirred under reflux until completion (6h to overnight). Reaction mixture was then cooled down to room temperature, extracted with diethyl ether and water, dried over MgSO₄. After evaporation of solvents, the crude product was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

1 equivalent of picoline was dissolved in anhydrous THF under argon at -40°C. 1 equivalent of 1M LDA in THF was slowly added dropwise under vigorous stirring followed by an addition of 1 equivalent of ester intermediate dissolved in THF. Reaction mixture was allowed to stir under argon and warm up to room temperature overnight. In case of incomplete conversion of the starting material 1 equivalent of LDA was added. Upon completion, the reaction mixture was extracted with EtOAc and aqueous saturated solution of ammonium chloride. Aqueous phase was further extracted with EtOAc. Combined organic fraction was washed with brine, dried over MgSO₄ and concentrated. After evaporation of solvents, the crude product was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

1 equivalent of a ketone intermediate was dissolved in glacial acetic acid followed by an addition of 1 equivelent of selenium dioxide. Reaction mixture was stirred at 110°C in an oil bath until completion. Upon completion, reaction mixture was allowed to cool down to room temperature, then excess of acetic acid was evaporated. The residual crude mixture was extracted with EtOAc and NaHCO₃ aqueous saturated solution. Combined organic fraction was washed with brine, dried over MgSO₄ and concentrated. After evaporation of solvents, the crude product was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

6.1.3 Synthesis of methoxymethylbemzamide intermediates via carboxylic acid activation by HATU and subsequent N-acylation (I,II)
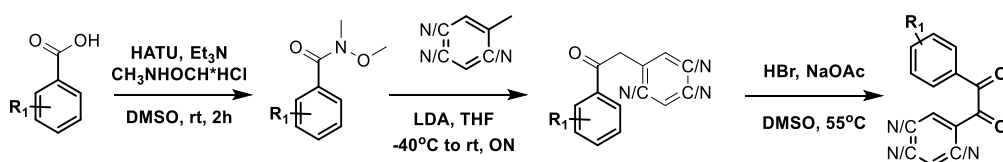
Scheme 4. Diketones synthesis via methoxymethylbenzamide intermediates

The mixture of carboxylic acid starting material, 3 mol eq of triethylamine, 3 mol eq of HATU, 1.2 mol eq of dimethylhydroxylamine in DMSO was stirred for 2-4 hours under argon. Upon reaction completion crude mixture was extracted with EtOAc and water, organic fraction was washed with brine and dried over MgSO4. After evaporation of solvents, the crude product was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

Picoline was dissolved in THF under argon and cooled to -40°C, 3 mol eq of LDA was added dropwise. Methoxymethylbenzamide intermediate (1 mol eq) solved in THF was then added dropwise, and reaction mixture was stirred under argon until completion (from 1.5h to overnight). Upon completion reaction mixture was warmed up to room temperature, extracted with EtOAc and water, combined organic fraction was washed with brine and dried over MgSO4. After evaporation of solvents, the crude product was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

The ketone intermediate was dissolved in DMSO and heated up to 55°C in an oil bath, HBr 48% aq sol was added dropwise (final concentration of 6% HBr). Reaction mixture was stirred at 55°C until completion, then cooled down to room temperature, then poured into aqueous solution of NaOAc (2.1g per 100 mL water), stirred for 30 min, extracted with EtOAc and water, washed with brine and dried over MgSO4. After evaporation of solvents, the diketone was purified with flash chromatography with a gradient of 100 % PE to 100 % EtOAc.

6.1.4 1,2-diketones synthesis through Sonogashira coupling followed by internal alkyne oxidation with palladium (II) iodide and DMSO (I,II)
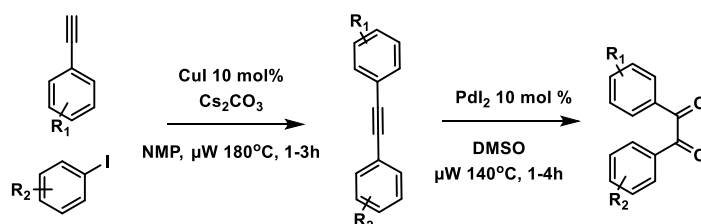
Scheme 5. 1,2-Diketone synthesis through Sonogashira coupling and subsequent alkyne oxidation

One equivalent of Iodine and Alkyne were dissolved in 10 mL NMP in a microwave vial. Two equivalents caesium carbonate and 10 mol% copper iodide were added under light stirring. The vial was sealed and heated to 180 °C for 60-180 min under microwave irradiation. After completion, the crude mixture was diluted with EtOAc and filtered through celite to remove the solid catalyst. The mixture was further extracted with EtOAc and water. The combined organic layers were washed with water and brine, and dried over MgSO4. Solvents were evaporated and the crude product was purified with flash chromatography using a gradient of 100 % PE to 100 % EtOAc.

Internal alkynes were dissolved in 5 mL DMSO in a microwave vial, 10 mol% palladium (II) iodide was added. The reaction mixture was heated to 140 °C under microwave irradiation for 1-4 hours. Upon completion, the crude mixture was diluted with EtOAc and filtered through celite to remove the catalyst. The mixture was further extracted with EtOAc and water. The combined organic layers were washed with water and brine, and dried over MgSO4. After evaporation of solvents, crude product was purified with flash chromatography using a gradient of 100 % PE to 100 % EtOAc.

6.1.5 Debus-Radziszewski nitrile imidazole intermediates synthesis (III)
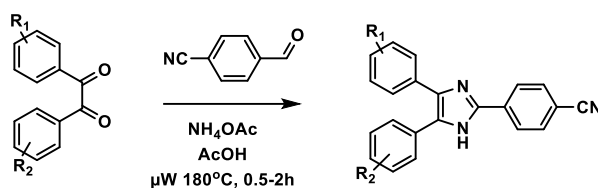
Scheme 6. Multicomponent Debus-Radziszewski reaction

1,2-Diketone was dissolved in 3 mL glacial acetic acid in a microwave vial. One equivalent 4-cyanobenzaldehyde and 10 equivalents ammonium acetate were added under stirring. The mixture was heated to 180 °C for 30 minutes under microwave irradiation. After completion, the reaction mixture was transferred dropwise to 20 mL chilled ammonium hydroxide 50 % v/v aqueous solution. The mixture was further extracted with EtOAc and water, the combined organic layers were washed with water and brine and dried over MgSO$_4$. Solvents were evaporated and the resulting residue purified by flash chromatography with a gradient of 100 % DCM to 10 % MeOH in DCM.

6.1.6 Demethylation with boron tribromide (IIIa)
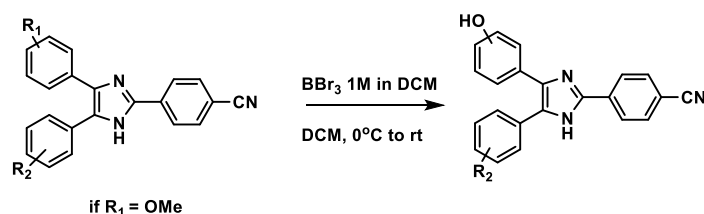
Scheme 7. Methyl ether cleavage with boron tribromide

One equivalent of nitrile imidazole intermediate was dissolved in 5 mL of DCM under argon at 0 °C. Two equivalents (per each methyl ether) of a 1M boron tribromide in DCM solution was added dropwise and stirred for several hours. After completion the reaction was quenched with EtOAc and water. The reaction mixture was extracted with EtOAc and water, washed with water and brine, and dried over MgSO$_4$. The crude product was purified by flash chromatography with a gradient of 100 % DCM to 10 % MeOH in DCM.
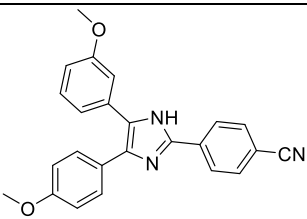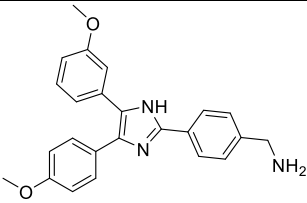
6.1.7 Synthesis of primary amines via LAH reduction (IV)
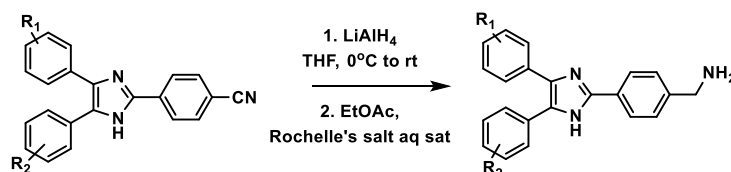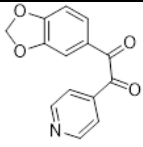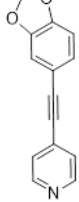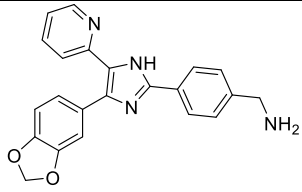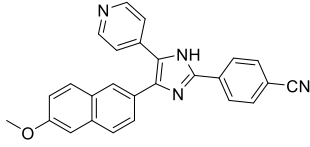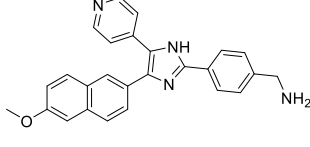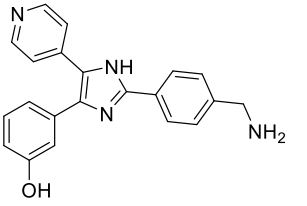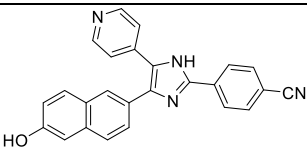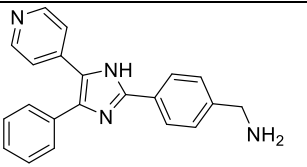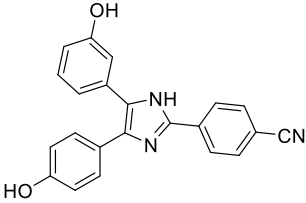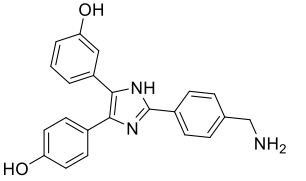

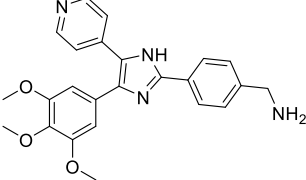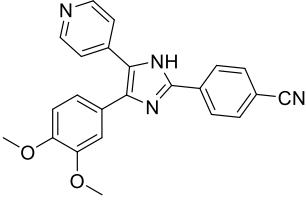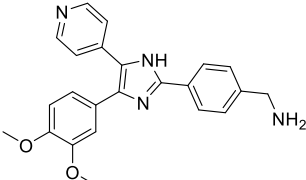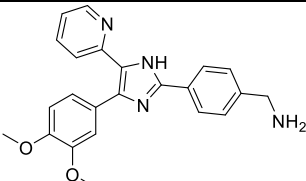**Scheme 8.** Nitrile reduction with lithium aluminium hydride

Nitrile imidazole intermediate was dissolved in 5 mL anhydrous THF at 0°C. 2.5 equivalents of lithium aluminiumhydride (LAH) were added and the mixture was stirred on ice for 5 hours with a bubbler. After completion, LAH was quenched by slow addition of EtOAc, water and aqueous saturated solution of Rochelle´s salt (potassium sodium tartrate). The mixture was further extracted with EtOAc and dried over MgSO$_4$. After evaporation of solvents, the crude product was purified by strong cation exchange chromatography. After loading on the SCX column, the crude product was washed with water, MeOH, toluene and EtOAc. Amine was then eluted using 1 M ammonia solution. Solvent was evaporated, and the product was further purified with flash chromatography using a gradient of 100 % DCM to 0.1N ammonia in 10 % MeOH 90% DCM.
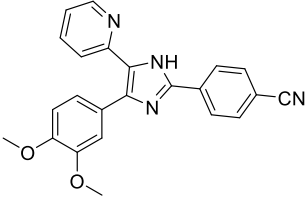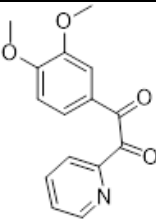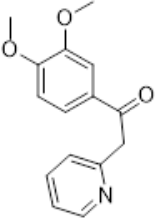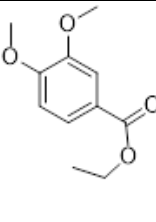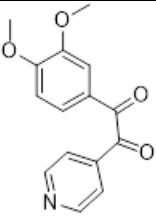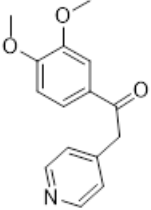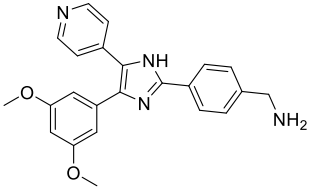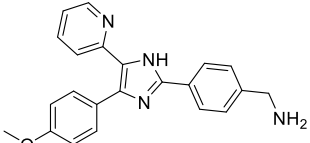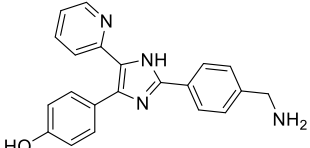
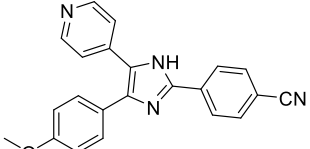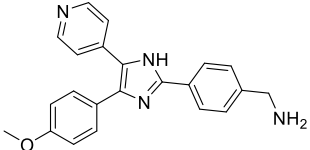## 6.2 Characterization of synthesized molecules

Name	Structure	Analysis
(4-(4-(4-fluorophenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Pale yellow solid (45 mg, 44%): ^{1}H NMR (400 MHz, Methanol-d_4) δ 8.45 (d, J = 6.3 Hz, 2H), 8.00 (d, J = 8.2 Hz, 2H), 7.59 – 7.53 (m, 4H), 7.51 (d, J = 8.3 Hz, 2H), 7.23 (t, J = 8.7 Hz, 2H), 3.89 (s, 2H). ^{13}C NMR (101 MHz, MeOD) δ 166.44, 163.99, 151.01, 150.06, 145.63, 133.04, 132.96, 130.43, 129.94, 128.11, 124.06, 117.88, 117.67, 47.18; m/z [M+H]+: 345.097
4-(4-(4-fluorophenyl)-5-(pyridin-2-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.53 (ddd, J = 4.9, 1.8, 1.0 Hz, 1H), 8.09 – 8.02 (m, 2H), 7.74 – 7.68 (m, 2H), 7.66 (dd, J = 8.4, 5.6 Hz, 2H), 7.57 (td, J = 7.8, 1.8 Hz, 1H), 7.47 (d, J = 8.1 Hz, 1H), 7.22 – 7.10 (m, 3H); m/z [M+H]+: 341.411
(4-(4-(4-fluorophenyl)-5-(pyridin-3-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (33 mg, 67%): ^{1}H NMR (400 MHz, Chloroform-d) δ 11.88 (d, J = 134.2 Hz, 1H), 8.70 (s, 1H), 8.37 (s, 1H), 7.85 (d, J = 8.0 Hz, 1H), 7.36 (s, 2H), 7.24 – 7.08 (m, 2H), 7.01 – 6.83 (m, 2H), 6.64 (s, 1H), 4.68 (s, 1H), 4.32 (d, J = 6.1 Hz, 1H), 1.85 (s, 2H). m/z [M+H]+: 345.185
(4-(4-(benzo[d][1,3]dioxol-5-yl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (9 mg, 27%): ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.00 (s, 1H), 8.46 (d, J = 5.6 Hz, 2H), 8.31 – 8.18 (m, 2H), 7.98 (d, J = 8.4 Hz, 2H), 7.59 (s, 2H), 7.50 (d, J = 1.7 Hz, 1H), 7.48 (d, J = 1.7 Hz, 2H), 7.04 (d, J = 5.5 Hz, 2H), 6.07 (s, 2H), 5.70 (s, 2H); ^{13}C NMR (101 MHz, MeOD) δ 193.45, 150.07, 131.28, 128.43, 126.83, 123.17, 110.24, 104.29, 102.90, 53.29; m/z [M+H]+: 371.213
4-(4-(benzo[d][1,3]dioxol-5-yl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow solid (74 mg, 68%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.16 (s, 1H), 8.50 (s, 2H), 8.33 – 8.19 (m, 2H), 7.52 (d, J = 5.3 Hz, 2H), 7.16 – 6.99 (m, 2H), 6.11 (s, 2H). ^{13}C NMR (101 MHz, DMSO) δ 149.74, 147.57, 133.97, 132.83, 125.68, 118.81, 110.45, 108.75, 101.43; m/z [M+H]+: 367.153

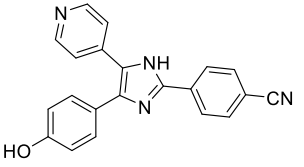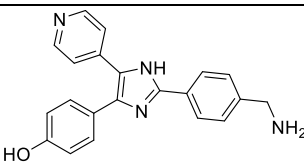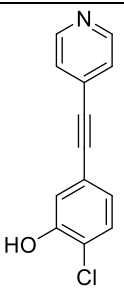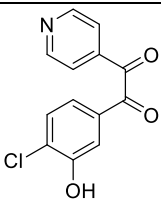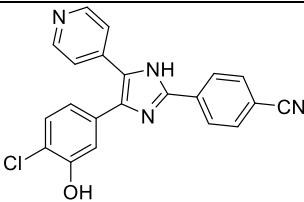|		
1-(benzo[d][1,3]dioxol-5-yl)-2-(pyridin-4-yl)ethane-1,2-dione		Prepared according to procedure 2. Yellow solid (121 mg, 70%): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.85 (d, *J* = 4.0 Hz, 2H), 7.75 (d, *J* = 4.0 Hz, 2H), 7.54 − 7.41 (m, 2H), 6.94 − 6.81 (m, 1H), 6.10 ((d, *J* = 3.3 Hz, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 193.16, 190.97, 154.13, 151.30, 149.02, 139.05, 128.36, 127.35, 122.35, 108.69, 108.50, 102.55.
4-(benzo[d][1,3]dioxol-5-ylethynyl)pyridine		Prepared according to procedure 2. White solid (167 mg, 75%): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.57 (d, *J* = 3.7 Hz, 2H), 7.34 (d, *J* = 3.7 Hz, 2H), 7.09 (dd, *J* = 8.3, 3.3 Hz, 1H), 6.99 (d, *J* = 3.3 Hz, 1H), 6.86 − 6.74 (m, 1H), 6.01 (d, *J* = 3.2 Hz, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 149.87, 148.86, 147.75, 131.72, 127.06, 125.53, 115.43, 111.81, 108.78, 101.66, 94.23, 85.42.
(4-(4-(benzo[d][1,3]dioxol-5-yl)-5-(pyridin-2-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (65 mg, 88%): 1H NMR (400 MHz, Methanol-d4) δ 8.57 (ddd, J = 4.9, 1.8, 1.0 Hz, 1H), 7.98 (d, J = 8.4 Hz, 1H), 7.72 (td, J = 7.8, 1.8 Hz, 1H), 7.55 − 7.35 (m, 3H), 7.28 (ddd, J = 7.5, 4.9, 1.1 Hz, 1H), 7.11 − 6.89 (m, 2H), 6.85 (d, J = 8.0 Hz, 1H), 5.98 (s, 2H), 3.85 (s, 2H). 13C NMR (101 MHz, MeOD) δ 150.24, 149.25, 149.06, 148.20, 144.52, 138.10, 129.74, 129.08, 128.98, 127.26, 127.22, 123.97, 123.68, 123.38, 110.15, 109.38, 102.64; m/z [M+H]+: 371.194
4-(4-(6-methoxynaphthalen-2-yl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. White solid (71 mg, 80%): ^{1}H NMR (400 MHz, DMSO-*d*$_6$) δ 13.34 (s, 1H), 8.46 (d, *J* = 5.2 Hz, 2H), 8.29 (d, *J* = 8.1 Hz, 2H), 8.09 (s, 1H), 7.98 (d, *J* = 8.2 Hz, 2H), 7.92 (dd, *J* = 17.1, 8.7 Hz, 2H), 7.65 − 7.46 (m, 3H), 7.45 − 7.37 (m, 1H), 7.25 (dd, *J* = 9.0, 2.5 Hz, 1H), 3.92 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 149.53, 133.93, 132.87, 129.62, 127.84, 127.36, 127.00, 125.70, 120.75, 119.44, 118.80, 110.53, 106.07, 55.31. m/z [M+H]+: 403.198
(4-(4-(6-methoxynaphthalen-2-yl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. White solid (13 mg, 23%): 1H NMR (400 MHz, DMSO-d6) δ 12.93 (s, 1H), 8.44 (s, 2H), 8.09 − 8.02 (m, 3H), 7.89 (d, J = 8.7 Hz, 2H), 7.55 (dd, J = 8.5, 1.8 Hz, 1H), 7.51 (d, J = 5.2 Hz, 2H), 7.46 (d, J = 8.0 Hz, 2H), 7.39 (d, J = 4.6 Hz, 1H), 7.22 (d, J = 9.2 Hz, 1H), 3.91 (s, 3H), 3.78 (s, 2H). m/z [M+H]+: 407.230

3-(2-(4-(aminomethyl)phenyl)-5-(pyridin-4-yl)-1H-imidazol-4-yl)phenol		Prepared according to procedure 7. White solid (8 mg, 26%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.58 (s, 1H), 8.48 (d, J = 6.8 Hz, 2H), 8.28 − 7.63 (m, 3H), 7.61 − 7.43 (m, 1H), 7.42 (dq, J = 14.7, 7.7 Hz, 1H), 7.34 − 7.19 (m, 3H), 7.02 − 6.83 (m, 2H), 6.84 (d, J = 8.1 Hz, 1H), 4.47 (d, J = 5.8 Hz, 2H). m/z [M+H]+: 343.166
4-(4-(6-hydroxynaphthalen-2-yl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 6. White solid (23 mg, 96%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.29 (s, 1H), 9.97 (s, 1H), 8.59 − 8.35 (m, 2H), 8.28 (d, J = 8.2 Hz, 3H), 8.02 (s, 1H), 7.97 (d, J = 8.2 Hz, 2H), 7.82 (t, J = 8.6 Hz, 2H), 7.54 (d, J = 5.2 Hz, 2H), 7.48 (dd, J = 8.5, 1.8 Hz, 1H), 7.20 (d, J = 2.3 Hz, 1H), 7.16 (dd, J = 8.8, 2.4 Hz, 1H). ^{13}C NMR (101 MHz, DMSO) δ 156.27, 149.68, 144.24, 133.97, 132.86, 132.66, 129.76, 127.89, 126.71, 125.68, 120.71, 119.49, 118.82, 110.46, 108.79. m/z [M+H]+: 389.143
(4-(4-phenyl-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (11 mg, 23%): m/z [M+H]+: 327.086
4-(5-(3-methoxyphenyl)-4-(4-methoxyphenyl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow solid (129 mg, 68%): 1H NMR (400 MHz, Methanol-d4) δ 10.03 (s, 1H), 8.20 (d, J = 8.3 Hz, 1H), 8.06 − 7.94 (m, 2H), 7.56 (dd, J = 17.3, 8.2 Hz, 1H), 7.43 (dd, J = 8.8, 6.6 Hz, 2H), 7.25 (q, J = 7.9 Hz, 1H), 7.08 (td, J = 5.7, 5.1, 3.2 Hz, 2H), 6.95 (dd, J = 8.6, 6.0 Hz, 2H), 6.91 − 6.80 (m, 1H); m/z [M+H]+: 382.153
(4-(5-(3-methoxyphenyl)-4-(4-methoxyphenyl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (19 mg, 47%): 1H NMR (400 MHz, DMSO-d6) δ 12.53 (d, J = 14.7 Hz, 1H), 8.22 − 7.84 (m, 3H), 7.58 − 7.37 (m, 4H), 7.37 − 7.16 (m, 2H), 7.16 − 7.11 (m, 2H), 7.04 (t, J = 9.0 Hz, 2H), 6.90 (d, J = 8.2 Hz, 2H), 6.81 − 6.74 (m, 1H), 3.82 (d, J = 7.6 Hz, 3H), 3.75 (d, J = 5.0 Hz, 3H), 3.68 (s, 2H); m/z [M+H]+: 386.170
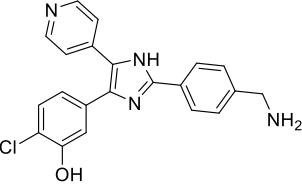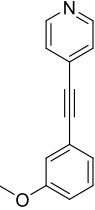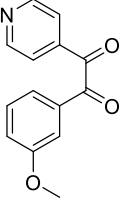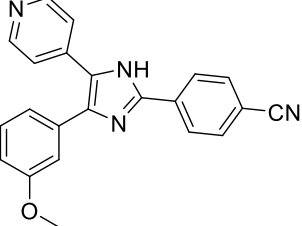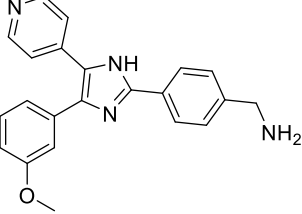
Compound	Structure	Description
4-(4-(3-hydroxyphenyl)-5-(4-hydroxyphenyl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 6. Grey solid (86 mg, 89%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.82 (d, J = 3.9 Hz, 1H), 9.64 (d, J = 60.7 Hz, 1H), 9.35 (d, J = 55.1 Hz, 1H), 8.22 (dd, J = 8.4, 6.1 Hz, 2H), 7.93 – 7.90 (m, 2H), 7.39 – 7.15 (m, 2H), 7.11 – 7.01 (m, 1H), 6.99 – 6.88 (m, 2H), 6.84 (d, J = 8.6 Hz, 1H), 6.79 – 6.58 (m, 2H). ^{13}C NMR (101 MHz, DMSO) δ 157.47, 157.16, 142.79, 137.23, 136.16, 134.45, 132.75, 130.08, 129.97, 129.03, 128.53, 125.39, 125.29, 118.95, 117.72, 115.48, 115.03, 113.83, 113.57, 109.80, 109.76. m/z [M+H]+: 354.095
3-(2-(4-(aminomethyl)phenyl)-4-(4-hydroxyphenyl)-1H-imidazol-5-yl)phenol		Prepared according to procedure 7. White solid (6 mg, 11%): m/z [M+H]+: 358.112
(4-(5-(pyridin-4-yl)-4-(3,4,5-trimethoxyphenyl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (5 mg, 12%): 1H NMR (400 MHz, Methanol-d4) δ 8.51 – 8.43 (m, 2H), 8.10 (d, J = 8.4 Hz, 2H), 7.67 – 7.53 (m, 4H), 6.82 (s, 2H), 4.15 (s, 2H), 3.81 (d, J = 16.3 Hz, 9H). 13C NMR (101 MHz, MeOD) δ 154.96, 150.14, 148.21, 136.83, 131.43, 130.38, 127.60, 123.47, 107.43, 61.25, 56.71, 44.38; m/z [M+H]+: 417.198
4-(4-(3,4-dimethoxyphenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Pale yellow solid (425 mg, 63%): 1H NMR (400 MHz, DMSO-d6) δ 13.14 (s, 1H), 8.48 (s, 1H), 8.31 – 8.20 (m, 2H), 8.00 – 7.92 (m, 2H), 7.55 (d, J = 5.2 Hz, 2H), 7.11 (d, J = 14.4 Hz, 3H), 3.83 (s, 3H), 3.73 (s, 3H). 13C NMR (101 MHz, DMSO) δ 149.70, 148.73, 133.97, 132.84, 125.68, 121.56, 120.68, 118.83, 111.95, 110.43, 55.57; m/z [M+H]+: 383.136
(4-(4-(3,4-dimethoxyphenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (72 mg, 55%): m/z [M+H]+: 387.165
(4-(4-(3,4-dimethoxyphenyl)-5-(pyridin-2-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (100 mg, 57%): m/z [M+H]+: 387.183
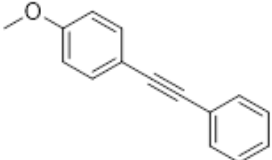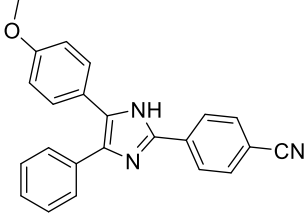
4-(4-(3,4-dimethoxyphenyl)-5-(pyridin-2-yl)-1H-imidazol-2-yl)benzonitrile		Pale yellow solid (1.33g, 56%): 1:1 mixture of 2 isomers ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.22 (s, 1H), 12.93 (s, 1H), 8.70 (d, *J* = 4.5 Hz, 1H), 8.50 – 8.44 (m, 1H), 8.35 (d, *J* = 8.5 Hz, 2H), 8.28 (d, *J* = 8.4 Hz, 2H), 8.00 – 7.90 (m, 5H), 7.87 – 7.74 (m, 2H), 7.55 – 7.45 (m, 2H), 7.33 (dd, *J* = 7.6, 5.0 Hz, 1H), 7.28 – 7.20 (m, 3H), 7.17 (dd, *J* = 8.2, 2.0 Hz, 1H), 7.03 (d, *J* = 8.3 Hz, 1H), 6.98 (d, *J* = 8.3 Hz, 1H), 3.81 (s, 3H), 3.79 (s, 2H), 3.75 (s, 3H), 3.70 (s, 2H).
1-(3,4-dimethoxyphenyl)-2-(pyridin-2-yl)ethane-1,2-dione		Prepared according to procedure 3. Yellow solid (1.68 g, 66%): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.68 (dt, *J* = 4.7, 1.4 Hz, 1H), 8.20 (dt, *J* = 7.8, 1.2 Hz, 1H), 7.94 (td, *J* = 7.7, 1.6 Hz, 1H), 7.63 (d, *J* = 1.8 Hz, 1H), 7.56 – 7.49 (m, 1H), 7.41 (dd, *J* = 8.3, 1.8 Hz, 1H), 6.88 (d, *J* = 8.3 Hz, 1H), 3.95 (dd, *J* = 3.8, 1.0 Hz, 6H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 195.13, 194.72, 154.90, 152.04, 150.01, 149.73, 137.38, 128.14, 126.59, 126.19, 123.39, 110.44, 110.12, 56.35, 56.21.
1-(3,4-dimethoxyphenyl)-2-(pyridin-2-yl)ethan-1-one		Prepared according to procedure 3. Thick yellow oil (2.74 g, 37%): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.55 (ddd, *J* = 4.9, 1.9, 0.9 Hz, 1H), 7.75 (dd, *J* = 8.4, 2.0 Hz, 1H), 7.67 – 7.57 (m, 2H), 7.31 (dd, *J* = 7.8, 1.0 Hz, 1H), 7.16 (ddd, *J* = 7.6, 4.9, 1.1 Hz, 1H), 6.88 (d, *J* = 8.4 Hz, 1H), 4.45 (s, 2H), 3.92 (d, *J* = 6.5 Hz, 6H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 195.65, 155.78, 153.57, 149.68, 149.14, 136.70, 129.86, 124.19, 124.03, 121.99, 110.89, 110.18, 56.20, 56.10, 48.48.
ethyl 3,4-dimethoxybenzoate		Prepared according to procedure 2. Thick transparent oil (19.4 g, 84%): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.68 (dd, *J* = 8.4, 1.9 Hz, 1H), 7.54 (d, *J* = 2.0 Hz, 1H), 6.88 (d, *J* = 8.5 Hz, 1H), 4.40 – 4.30 (m, 2H), 3.93 (s, 6H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 166.55, 153.00, 148.71, 123.62, 123.18, 112.08, 110.33, 60.94, 56.13, 14.53
1-(3,4-dimethoxyphenyl)-2-(pyridin-4-yl)ethane-1,2-dione		Prepared according to procedure 3. Yellow solid (470 mg, 19%): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.90 – 8.83 (m, 2H), 7.60 (d, *J* = 2.0 Hz, 1H), 7.48 (dd, *J* = 8.4, 2.0 Hz, 1H), 6.92 (d, *J* = 8.4 Hz, 1H), 3.97 (d, *J* = 1.3 Hz, 6H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 193.33, 191.46, 155.63, 151.27, 149.97, 139.21, 126.78, 125.69, 122.41, 110.58, 110.39, 56.48, 56.29.

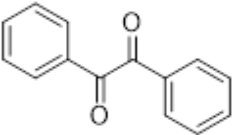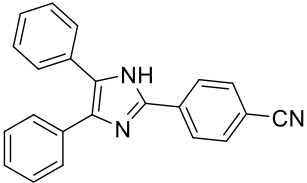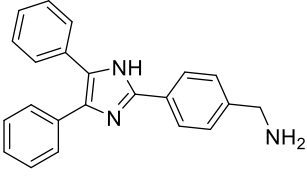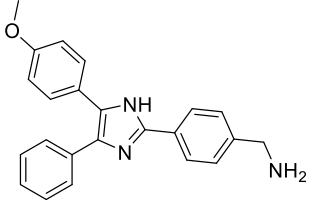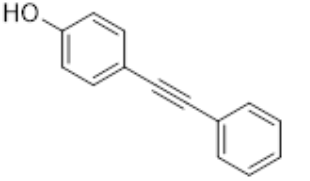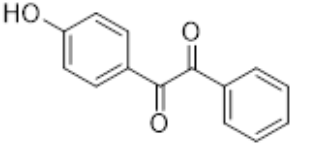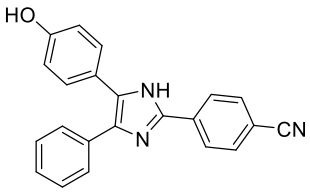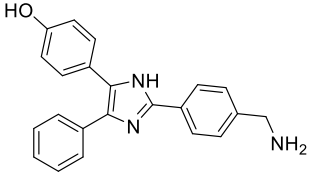|		
1-(3,4-dimethoxyphenyl)-2-(pyridin-4-yl)ethan-1-one		Prepared according to procedure 3. Yellow solid (2.36 g, 32%):^{1}H NMR (400 MHz, Chloroform-d) δ 8.60 – 8.50 (m, 2H), 7.62 (dd, J = 8.4, 2.0 Hz, 1H), 7.54 (d, J = 2.0 Hz, 1H), 7.23 – 7.16 (m, 2H), 6.90 (d, J = 8.4 Hz, 1H), 4.24 (s, 2H), 3.95 (s, 3H), 3.92 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 194.64, 153.87, 150.06, 149.38, 144.01, 129.51, 124.87, 123.48, 110.55, 110.16, 56.24, 56.10, 44.32.
(4-(4-(3,5-dimethoxyphenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (4 mg, 21%): ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 12.81 (s, 1H), 8.50 (s, 2H), 8.32 – 7.66 (m, 4H), 7.55 (d, J = 4.7 Hz, 2H), 7.43 (d, J = 8.1 Hz, 1H), 6.70 (dd, J = 4.7, 2.2 Hz, 2H), 6.59 (s, 1H), 4.55 (d, J = 3.8 Hz, 2H), 3.74 (s, 6H); m/z [M+H]+: 388.165
(4-(4-(4-methoxyphenyl)-5-(pyridin-2-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (5 mg, 14%): ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 8.55 (d, J = 22.2 Hz, 1H), 8.29 – 7.99 (m, 2H), 7.98 – 7.69 (m, 2H), 7.59 (d, J = 8.1 Hz, 2H), 7.47 – 7.31 (m, 2H), 7.33 – 7.14 (m, 1H), 7.07 – 6.86 (m, 2H), 3.80 (s, 3H), 3.78 (s, 2H). ^{13}C NMR (101 MHz, DMSO) δ 158.70, 143.87, 136.39, 129.96, 128.42, 127.36, 125.27, 121.60, 113.46, 60.75, 55.14, 45.22; m/z [M+H]+: 357.164
4-(2-(4-(aminomethyl)phenyl)-5-(pyridin-2-yl)-1H-imidazol-4-yl)phenol		Prepared according to procedure 7. Pale yellow solid (10 mg, 42%): ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 12.51 (s, 1H), 9.53 (s, 1H), 8.74 – 8.31 (m, 1H), 8.30 – 7.96 (m, 2H), 7.90 (d, J = 8.4 Hz, 1H), 7.76 (s, 1H), 7.44 (dd, J = 20.8, 7.5 Hz, 2H), 7.27 (dd, J = 48.0, 13.8 Hz, 2H), 6.79 (dd, J = 8.7, 3.4 Hz, 2H); m/z [M+H]+: 343.152
4-(4-(4-methoxyphenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 6. Yellow solid (235 mg, 38%): ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 13.15 (s, 1H), 8.48 (d, J = 5.2 Hz, 2H), 8.31 – 8.19 (m, 2H), 7.96 (d, J = 8.4 Hz, 2H), 7.57 – 7.42 (m, 4H), 7.07 (d, J = 8.2 Hz, 2H), 3.83 (s, 3H); m/z [M+H]+: 353.110
(4-(4-(4-methoxyphenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (11 mg, 26%): ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 13.21 (s, 1H), 8.61 – 8.49 (m, 4H), 8.15 (d, J = 8.0 Hz, 2H), 7.73 – 7.59 (m, 4H), 7.51 (d, J = 8.1 Hz, 2H), 7.09 (s, 2H), 4.06 (d, J = 5.6 Hz, 2H), 3.84 (s, 3H); m/z [M+H]+: 357.160
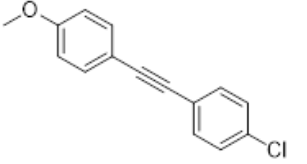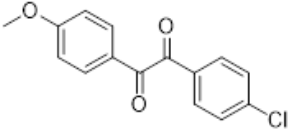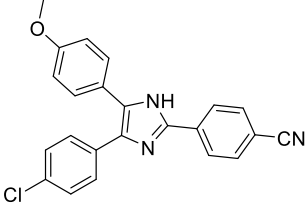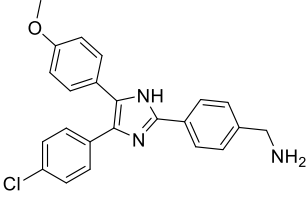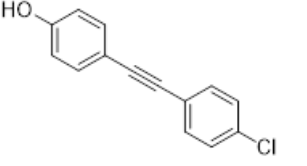
4-(4-(4-hydroxyphenyl)-5-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 6. Yellow solid (50 mg, 51%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.11 (s, 1H), 9.85 (s, 1H), 8.50 – 8.41 (m, 2H), 8.24 (d, J = 8.4 Hz, 2H), 7.95 (d, J = 8.2 Hz, 2H), 7.58 – 7.44 (m, 2H), 7.35 (d, J = 8.5 Hz, 2H), 6.91 (d, J = 8.5 Hz, 2H). ^{13}C NMR (101 MHz, DMSO) δ 158.09, 149.62, 143.74, 141.96, 134.19, 134.03, 132.82, 130.44, 125.55, 120.47, 118.82, 115.74, 110.30; m/z [M+H]+: 339.072
4-(2-(4-(aminomethyl)phenyl)-5-(pyridin-4-yl)-1H-imidazol-4-yl)phenol		Prepared according to procedure 7. Pale yellow solid (3mg, 9%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.57 (d, J = 6.0 Hz, 1H), 8.22 – 7.52 (m, 1H), 7.54 – 7.13 (m, 13H), 2.54 (s, 2H); m/z [M+H]+: 343.150
2-chloro-5-(pyridin-4-ylethynyl)phenol		Prepared according to procedure 4. Yellow solid (121 mg, 45 %): 1H-NMR (400 MHz, Chloroform-d) δ [ppm] = 8.58 (d, J = 6.0 Hz, 2H), 8.24 (s, 1H), 7.35 (d, J = 6.1 Hz, 2H), 7.29 (d, J = 8.2 Hz, 1H), 7.20 (d, J = 1.8 Hz, 1H), 7.00 (dd, J = 8.2, 1.9 Hz, 1H). 13C-NMR (101 MHz, Chloroform-d) δ [ppm] = 152.44, 149.72 131.32, 129.68, 125.6, 125.54, 122.05, 121.71, 119.60, 93.13, 86.83. LC-MS [M+H] = 229.981.
1-(4-chloro-3-hydroxyphenyl)-2-(pyridin-4-yl)ethane-1,2-dione		Prepared according to procedure 4. Yellow solid, (23 mg, 17 %): ^{1}H-NMR (300 MHz, Chloroform-d) δ [ppm] = 9.21(s, OH), 8.92 (d, J = 6.08 Hz, 2H), 7.81 (d, J = 6.09 Hz, 2H), 7.66 (d, 1H), 7.52 (d, 1H), 7.22 (s, 1H). LC-MS [M+H] = 262.032
(5-(4-chloro-3-hydroxyphenyl)-4-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 4. Yellow solid (15 mg, 5 %): ^{1}H NMR (400 MHz, Methanol-d_4) δ [ppm] = 10.07 (s, 1H), 8.47 (d, J = 5.58 Hz 2H), 8.18 (d, J = 8.54 Hz, 2H), 7.85 (d, J = 8.54 Hz, 2H), 7.60 (d, J = 5.36 Hz, 2H), 7.43 (dd, J = 8.24, 20.12 Hz, 1H), 7.06 (s, 1H), 6.98 (dd, J = 1.99, 8.18 Hz,1H). ^{13}C NMR (101 MHz, MeOD) δ [ppm] = 192.11, 153.45, 148.78, 138.66, 133.66, 132.52, 131.46, 130.28, 129.21, 128.37, 127.96, 127.11, 126.01, 125.91, 121.99, 121.87, 120.39, 118.09, 116.51, 111.89. LC-MS [M+H] = 373.217
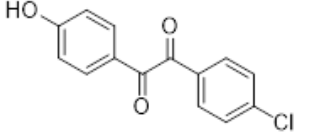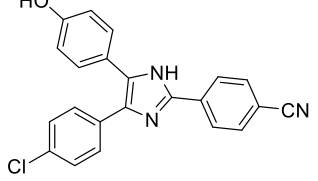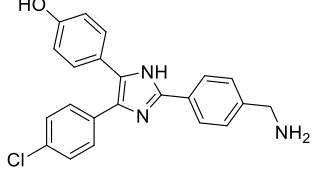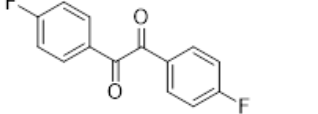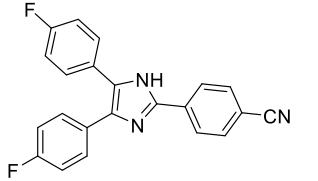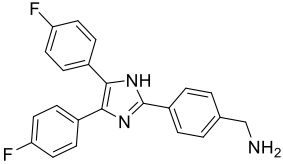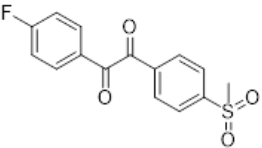
5-(2-(4-(aminomethyl)phenyl)-4-(pyridin-4-yl)-1H-imidazol-5-yl)-2-chlorophenol		Prepared according to procedure 5. Yellow solid (12 mg, 79 %): ^{1}H NMR (400 MHz, Methanol-d_4) δ [ppm] = 8.43 (ddt, J = 6.6, 4.8, 1.6 Hz, 2H), 7.99 (d, J = 8.1 Hz, 2H), 7.59 (d, J = 5.5 Hz, 2H), 7.49 (d, J = 8.1 Hz, 2H), 7.35 (d, J = 7.6 Hz, 1H), 7.28 (t, J = 7.9 Hz, 1H), 7.00 (d, J = 2.0 Hz, 1H), 3.90 (m, 2H). ^{13}C NMR (101 MHz, MeOD) δ [ppm] = 46.08, 116.81, 118.55, 120.60, 121.05, 123.26, 123.33, 127.30, 128.33, 129.20, 129.23, 129.85, 131.17, 131.42, 150.00, 150.10, 159.15. LC-MS [M+H] = 377.412
4-((3-methoxyphenyl)ethynyl)pyridine		Prepared according to procedure 4. Yellow solid (102 mg, 36 %): ^{1}H NMR (300 MHz, Chloroform-d) δ [ppm] = 8.53 (q, J = 7.2 Hz, 2H), 7.26 (s, 2H), 7.05 (d, 1H), 6.44 (s, 3H), 5.68 (t, 2H). ^{13}C NMR (75 MHz, CDCl$_3$) δ [ppm] = 13.56, 20.05, 40.33, 59.57, 169.95. LC-MS [M+H] = 210.207
1-(3-methoxyphenyl)-2-(pyridin-4-yl)ethane-1,2-dione		Prepared according to procedure 4. Yellow solid (90 mg, 18 %): 1H NMR (400 MHz, Chloroform-d) δ [ppm] = 9.19 (d, J = 6.7 Hz, 2H), 8.88 (d, J = 6.0 Hz, 2H), 7.81 (d, J = 6.7 Hz, 1H), 7.77 (d, J = 6.1 Hz, 1H), 7.44 (d, J = 7.9 Hz, 1H), 7.00 (s, 1H), 4.22 (t, J = 6.1 Hz, 3H). LC-MS [M+H] = 242.040
4-(5-(3-methoxyphenyl)-4-(pyridin-4-yl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow solid (147 mg): 1H NMR (500 MHz, Chloroform-d) δ [ppm] = 8.17 (dd, J = 5.7, 3.3 Hz, 2H), 7.99 (dd, J = 5.7, 3.3 Hz, 2H), 7.93 (s, 2H), 7.73 (s, 2H), 7.51 (s, 1H), 2.01 (s, 3H). LC-MS [M+H] = 353.143
(4-(5-(3-methoxyphenyl)-4-(pyridin-4-yl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (46 mg, 31 %): ^{1}H NMR (400 MHz, Methanol-d_4) δ [ppm] = 8.36 (d, J = 6.3 Hz, 2H), 7.93 (d, J = 8.2 Hz, 2H), 7.51 (d, J = 6.3 Hz, 2H), 7.42 (d, J = 8.3 Hz, 2H), 7.30 (t, J = 8.1 Hz, 1H), 7.00 (s, 1H), 7.02 (d, J = 1.5 Hz, 2H), 6.94 (dd, J = 8.1, 2.9 Hz, 1H), 3.80 (s, 2H), 3.74 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ [ppm] = 46.38, 54.80, 55.78, 115.31, 115.38, 122.21, 123.31, 127.25, 128.98, 129.58, 131.09, 133.81, 143.73, 144.90, 149.07, 150.01, 161.47. LC-MS [M+H] = 357.197
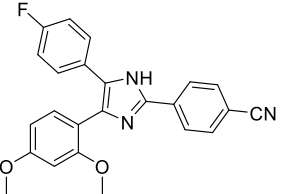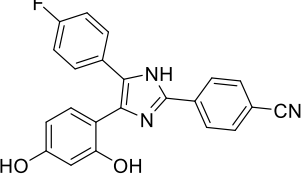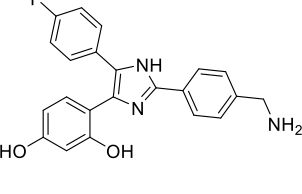
Benzil		Prepared according to procedure 1. White solid (80 mg, 38 %): 1H NMR (400 MHz, Chloroform-d) δ 7.98 (d, J = 7.2 Hz, 4H), 7.66 (t, J = 7.4 Hz, 2H), 7.52 (t, J = 7.8 Hz, 4H); 13C NMR (101 MHz, CDCl3) δ 194.55, 134.88, 133.02, 129.91, 129.02, 77.33, 77.21, 77.01, 76.70; m/z [M+DMSO+H]+ : 279.101
4-(4,5-diphenyl-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. White solid (33 mg, 27 %): 1H NMR (400 MHz, Acetone-d6) δ 8.16 (d, J = 8.5 Hz, 2H), 7.74 (d, J = 8.5 Hz, 2H), 7.46 (s, 4H), 7.24 (s, 6H); 13C NMR (101 MHz, Acetone) δ 206.02, 144.02, 134.48, 132.58, 131.13, 128.35, 125.67, 118.48, 111.13, 13.38, 10.38; m/z [M+H]+ : 321.127
(4-(4,5-diphenyl-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (4 mg, 25 %): 1H NMR (400 MHz, Methanol-d4) δ 8.46 (s, 2H), 7.99 (d, J = 8.4 Hz, 1H), 7.89 (dt, J = 8.2, 4.2 Hz, 1H), 7.81 (d, J = 8.4 Hz, 1H), 7.44 – 7.32 (m, 5H), 7.22 (ddd, J = 16.2, 7.2, 4.2 Hz, 6H), 3.80 (s, 2H); 13C NMR (101 MHz, MeOD) δ 148.25, 144.87, 137.71, 133.32, 130.41, 129.05, 128.98, 128.92, 128.66, 127.97, 127.12, 126.95, 124.22, 45.5; m/z [M+H]+ : 325.158
1-methoxy-4-(phenylethynyl)benzene		Prepared according to procedure 4. Brown solid (161 mg, 65 %): 1H NMR (400 MHz, Chloroform-d) δ 7.98 (t, J = 8.5 Hz, 4H), 7.67 (t, J = 7.4 Hz, 1H), 7.53 (t, J = 7.8 Hz, 2H), 7.00 (d, J = 8.9 Hz, 2H), 3.91 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 194.83, 193.14, 164.99, 134.70, 133.23, 132.39, 129.91, 128.95, 126.13, 114.37, 55.65; m/z [M+H]+ : 241.021
4-(5-(4-methoxyphenyl)-4-phenyl-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Orange solid (210 mg, 89 %): 1H NMR (400 MHz, DMSO-d6) δ 12.93 (s, 1H), 8.24 (d, J = 8.4 Hz, 2H), 7.94 (d, J = 8.6 Hz, 2H), 7.56 (d, J = 7.1 Hz, 1H), 7.51 (d, J = 6.9 Hz, 1H), 7.45 (dd, J = 12.0, 8.7 Hz, 3H), 7.32 (t, J = 7.4 Hz, 1H), 7.24 (d, J = 7.3 Hz, 1H), 7.04 (d, J = 8.7 Hz, 1H), 6.90 (d, J = 8.8 Hz, 1H), 3.79 (d, J = 22.1 Hz, 3H); 13C NMR (101 MHz, DMSO) δ 159.65, 158.75, 143.92, 143.72, 138.57, 137.92, 135.33, 134.83, 134.80, 133.23, 131.23, 130.40, 130.05, 129.16, 128.87, 128.83, 128.70, 128.40, 127.64, 127.40, 127.09, 125.92, 125.87, 123.29, 119.39, 114.66, 114.20, 110.43, 110.39, 55.70, 55.51; m/z [M+H]+ : 352.140
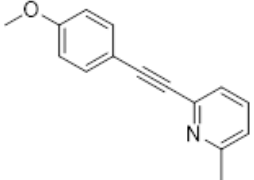
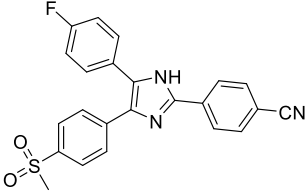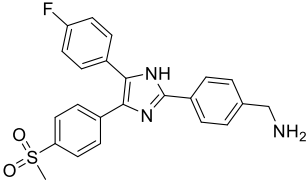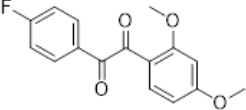|		
(4-(5-(4-methoxyphenyl)-4-phenyl-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (80 mg, 75 %): 1H NMR (400 MHz, Methanol-d4) δ 10.04 (s, 1H), 8.60 (s, 2H), 8.11 (d, J = 8.4 Hz, 2H), 7.98 (td, J = 17.8, 7.5 Hz, 2H), 7.50 (t, J = 8.2 Hz, 3H), 7.42 (d, J = 8.7 Hz, 2H), 7.34 (q, J = 9.6, 8.3 Hz, 2H), 6.94 (d, J = 8.8 Hz, 2H), 4.68 (s, 2H), 3.83 (s, 3H); 13C NMR (101 MHz, MeOD) δ 173.31, 159.35, 142.13, 139.62, 135.83, 129.35, 128.91, 128.61, 128.08, 128.04, 127.83, 126.96, 125.66, 125.55, 125.46, 113.57, 63.44, 54.32, 48.23, 48.02, 47.80, 47.59, 47.38, 47.16; m/z [M+H]+ : 339.162
4-(phenylethynyl)phenol		Prepared according to procedure 4. Yellow solid (473 mg, 62 %): 1H NMR (400 MHz, Chloroform-d) δ 7.70 (dt, J = 8.1, 1.6 Hz, 2H), 7.37 (dd, J = 13.4, 7.1 Hz, 5H), 7.14 (d, J = 8.7 Hz, 2H); 13C NMR (101 MHz, CDCl3) δ 157.30, 141.70, 129.74, 129.62, 129.46, 128.69, 128.36, 128.33, 123.36, 116.94, 116.56, 92.43, 77.34; m/z [M+H]+ : 197.034
Preparation of 1-(4-hydroxyphenyl)-2-phenylethane-1,2-dione		Prepared according to procedure 4. Yellow oil (150 mg, 39 %): 1H NMR (400 MHz, Chloroform-d) δ 10.05 (s, 1H), 8.14 (d, J = 7.6 Hz, 2H), 7.82 (d, J = 7.0 Hz, 2H), 7.52 (dd, J = 7.4, 3.8 Hz, 3H), 7.15 − 7.07 (m, 2H); 13C NMR (101 MHz, CDCl3) δ 194.88, 194.88, 155.52, 155.10, 133.73, 130.20, 129.65, 129.34, 128.90, 128.55, 128.50, 128.07, 120.75, 115.29; m/z [M+H]+ : 235.019
4-(5-(4-hydroxyphenyl)-4-phenyl-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow solid (50 mg, 23 %): 1H NMR (400 MHz, Methanol-d4) δ 8.23 (d, J = 8.5 Hz, 1H), 8.07 (d, J = 8.5 Hz, 1H), 7.59 (dd, J = 6.8, 3.0 Hz, 1H), 7.53 (d, J = 2.7 Hz, 1H), 7.40 (d, J = 8.6 Hz, 1H), 6.91 (d, J = 8.6 Hz, 1H); 13C NMR (101 MHz, MeOD) δ 159.29, 141.84, 133.12, 133.12, 131.18, 130.03, 130.03, 129.78, 129.62, 128.89, 128.89, 128.36, 128.36, 127.66, 127.66, 127.01, 117.29, 115.69, 115.69, 115.22; m/z [M+H]+ : 338.148
4-(2-(4-(aminomethyl)phenyl)-4-phenyl-1H-imidazol-5-yl)phenol		Prepared according to procedure 7. Yellow solid (4 mg, 13 %): m/z [M+H]+ : 342.175

1-chloro-4-((4-methoxyphenyl)ethynyl)benzene		Prepared according to procedure 4. Yellow solid (367 mg, 70 %): 1H NMR (400 MHz, Chloroform-d) δ 7.47 (dd, J = 10.7, 8.7 Hz, 4H), 7.33 (d, J = 8.5 Hz, 2H), 6.90 (d, J = 8.8 Hz, 2H), 3.85 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 159.79, 133.87, 133.06, 132.63, 128.63, 122.14, 115.02, 114.06, 90.36, 86.98, 55.32; m/z [M+H]+ : 243.063
1-(4-chlorophenyl)-2-(4-methoxyphenyl)ethane-1,2-dione		Prepared according to procedure 4. Brown solid (152 mg, 38 %):1H NMR: δ 3.79 (3H, s), 7.05 (2H, ddd, J = 8.3, 1.2, 0.4 Hz), 7.55 (2H, ddd, J = 8.7, 1.8, 0.4 Hz), 7.86-7.96 (4H, 7.89 (ddd, J = 8.3, 1.8, 0.4 Hz), 7.93 (ddd, J = 8.7, 1.8, 0.4 Hz); 13C NMR (101 MHz, CDCl3) δ 192.99, 160.42, 135.68, 132, 131.67, 131.2, 128.96, 126.1, 114.01, 114.01, 55.46; m/z [M+H]+ : 273.042
4-(4-(4-chlorophenyl)-5-(4-methoxyphenyl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Brown solid (95 mg, 44 %): 1H NMR (400 MHz, DMSO-d6) δ 12.99 (s, 1H), 8.23 (d, J = 8.5 Hz, 2H), 7.94 (d, J = 8.4 Hz, 2H), 7.60 − 7.49 (m, 2H), 7.41 (dd, J = 25.0, 8.6 Hz, 4H), 7.06 (d, J = 8.7 Hz, 2H), 3.82 (s, 3H); 13C NMR (101 MHz, DMSO) δ 193.04, 193.04, 159.79, 134.67, 134.18, 133.64, 133.25, 131.54, 130.48, 130.38, 129.22, 129.06, 128.91, 128.77, 125.93, 122.98, 119.35, 114.78, 114.32, 110.55, 55.72, 49.06; m/z [M+H]+ : 386.122
(4-(4-(4-chlorophenyl)-5-(4-methoxyphenyl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Brown solid (12 mg, 21 %): 1H NMR (400 MHz, Methanol-d4) δ 9.97 (s, 1H), 8.18 (d, J = 8.4 Hz, 2H), 8.00 (d, J = 8.4 Hz, 2H), 7.48 (t, J = 8.8 Hz, 2H), 7.39 (t, J = 7.8 Hz, 2H), 7.32 (t, J = 8.2 Hz, 2H), 6.95 (t, J = 7.0 Hz, 2H), 4.70 (s, 2H), 3.83 (s, 3H); 13C NMR (101 MHz, MeOD) δ 192.05, 136.12, 135.21, 129.88, 129.52, 129.19, 129.19, 128.20, 128.15, 126.68, 125.63, 113.82, 113.76, 78.18, 77.85, 63.13, 54.40; m/z [M+H]+ : 369.00
4-((4-chlorophenyl)ethynyl)phenol		Prepared according to procedure 4. Brown oil (187 mg, 36 %): 1H NMR (400 MHz, Chloroform-d) δ 7.93 (d, J = 8.5 Hz, 1H), 7.69 (d, J = 8.6 Hz, 1H), 7.56 (d, J = 8.6 Hz, 2H), 7.48 (d, J = 8.6 Hz, 2H), 7.26 (s, 2H), 6.86 (d, J = 7.6 Hz, 2H); 13C NMR (101 MHz, CDCl3) δ 155.96, 133.28, 132.64, 129.90, 129.66, 128.62, 127.73, 122.13, 120.78, 116.27, 115.55, 115.27, 99.99; m/z [M+H]+ : 227.006
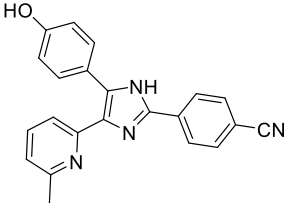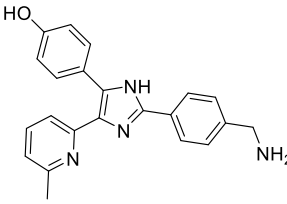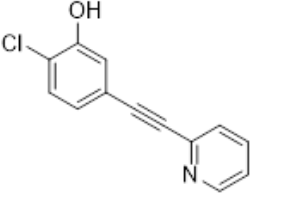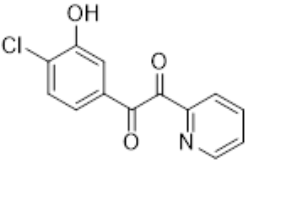
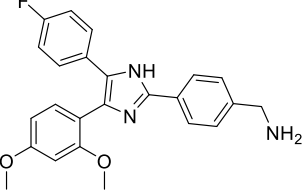|		
1-(4-chlorophenyl)-2-(4-hydroxyphenyl)ethane-1,2-dione		Prepared according to procedure 4. Brown oil (64 mg, 31 %): 1H NMR (400 MHz, Chloroform-d) δ 10.02 (s, 1H), 7.93 (d, J = 27.7 Hz, 2H), 7.61 (d, J = 8.2 Hz, 2H), 7.55 (d, J = 8.3 Hz, 2H), 7.31 (d, J = 8.0 Hz, 2H); 13C NMR (101 MHz, CDCl3) δ 193.48, 170.17, 142.72, 132.36, 131.88, 131.47, 128.86, 127.78, 127.17, 103.03; m/z [M+H]+ : 258.994
4-(4-(4-chlorophenyl)-5-(4-hydroxyphenyl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow oil (30 mg, 33 %): 1H NMR (400 MHz, Methanol-d4) δ 8.15 (d, J = 8.5 Hz, 2H), 7.83 (d, J = 8.5 Hz, 2H), 7.51 (d, J = 8.6 Hz, 2H), 7.33 (dd, J = 13.4, 8.6 Hz, 4H), 6.84 (d, J = 8.6 Hz, 2H); 13C NMR (101 MHz, MeOD) δ 157.54, 144.16, 133.99, 132.42, 132.42, 129.60, 129.60, 129.07, 129.07, 128.17, 128.17, 125.69, 125.69, 118.19, 115.20, 115.20, 111.27; m/z [M+H]+ : 372.112
4-(2-(4-(aminomethyl)phenyl)-4-(4-chlorophenyl)-1H-imidazol-5-yl) phenol		Prepared according to procedure 7. Yellow solid (5 mg, 33 %): 1H NMR (400 MHz, Methanol-d4) δ 10.04 (s, 1H), 8.57 (d, J = 18.4 Hz, 2H), 8.20 (d, J = 8.3 Hz, 1H), 8.09 (d, J = 8.3 Hz, 1H), 7.98 (dd, J = 36.4, 8.3 Hz, 2H), 7.52 (d, J = 7.0 Hz, 2H), 7.34 (d, J = 6.4 Hz, 4H), 6.83 (d, J = 8.6 Hz, 2H), 4.43 (s, 2H); 13C NMR (101 MHz, MeOD) δ 192.05, 157.27, 142.95, 135.71, 133.70, 129.86, 129.61, 129.47, 129.05, 128.61, 128.27, 128.00, 127.77, 126.96, 126.78, 125.66, 125.58, 115.19, 115.12, 114.95, 60.38; m/z [M+H]+ : 376.148
1,2-bis(4-fluorophenyl)ethane-1,2-dione		Prepared according to procedure 1. Yellow solid (125 mg, 15 %): 1H NMR (400 MHz, Chloroform-d) δ 7.72 (dd, J = 5.7, 3.3 Hz, 4H), 7.55 (dd, J = 5.7, 3.3 Hz, 4H); 13C NMR (101 MHz, CDCl3) δ 169.53, 167.85, 152.62, 117.40; m/z [M+H]+ : 248.969
4-(4,5-bis(4-fluorophenyl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow solid (20 mg, 12 %): 1H NMR (400 MHz, Methanol-d4) δ 10.10 (s, 1H), 8.09 (d, J = 8.5 Hz, 2H), 7.97 (d, J = 8.3 Hz, 2H), 7.74 (d, J = 8.5 Hz, 4H), 7.67 (d, J = 8.2 Hz, 4H); 13C NMR (101 MHz, MeOD) δ 191.34, 169.84, 152.65, 146.59, 139.20, 132.70, 131.68, 129.55, 126.86, 118.26, 117.40, 111.50; m/z [M+H]+ : 357.136

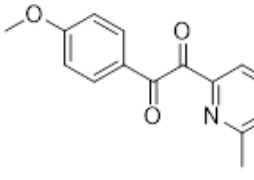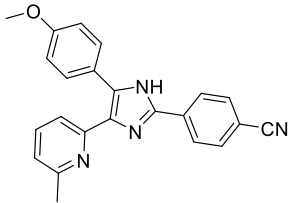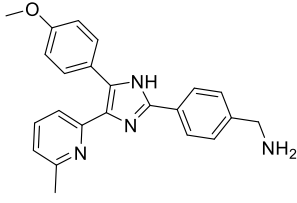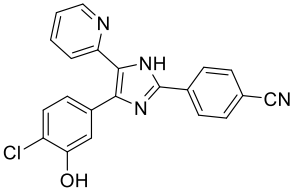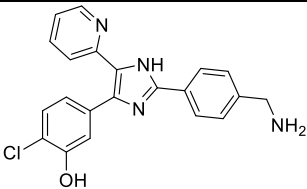|		
(4-(4,5-bis(4-fluorophenyl)-1H-imidazol-2-yl)phenyl)methanamine		Prepared according to procedure 7. Yellow solid (2 mg, 20 %): 1H NMR (400 MHz, Methanol-d4) δ 8.50 (s, 2H), 7.78 (d, J = 8.2 Hz, 2H), 7.46 (d, J = 8.1 Hz, 2H), 7.43 (s, 4H), 7.34 (q, J = 8.1 Hz, 4H), 4.05 (s, 2H); 13C NMR (101 MHz, MeOD) δ 180.29, 161.41, 141.67, 137.78, 128.90, 128.13, 127.80, 126.85, 126.70, 120.30, 63.57, 48.30; m/z [M+H]+ : 362.162
1-(4-fluorophenyl)-2-(4-(methylsulfonyl)phenyl)ethane-1,2-dione		Prepared according to procedure 1. Yellow solid (200 mg, 65 %): 1H NMR (400 MHz, Chloroform-d) δ 8.20 (d, J = 8.5 Hz, 2H), 8.13 (d, J = 8.5 Hz, 2H), 8.07 (dd, J = 8.8, 5.3 Hz, 2H), 7.25 (t, J = 8.6 Hz, 2H), 3.12 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 191.17, 165.81, 145.68, 136.72, 133.01, 132.91, 130.78, 128.11, 116.54, 44.28; m/z [M+H]+ : 307.021
2-(4-ethynylphenyl)-5-(4-fluorophenyl)-4-(4-(methylsulfonyl)phenyl)-1H-imidazole		Prepared according to procedure 5. Yellow solid (25 mg, 9 %): 1H NMR (400 MHz, Methanol-d4) δ 10.10 (s, 1H), 8.21 (d, J = 8.6 Hz, 2H), 7.94 (d, J = 7.9 Hz, 2H), 7.88 (d, J = 8.6 Hz, 2H), 7.80 (d, J = 8.1 Hz, 2H), 7.56 (dd, J = 8.8, 5.4 Hz, 2H), 7.21 (t, J = 8.5 Hz, 2H), 3.16 (s, 3H); 13C NMR (101 MHz, MeOD) δ 195.85, 181.43, 133.76, 132.50, 132.07, 131.67, 130.98, 130.62, 128.07, 127.30, 125.94, 119.00, 118.09, 111.78, 42.92; m/z [M+H]+ : 418.095
(4-(5-(4-fluorophenyl)-4-(4-(methylsulfonyl)phenyl)-1H-imidazol-2-yl)phenyl)methan amine		Prepared according to procedure 7. Yellow solid (6 mg, 40 %): 1H NMR (400 MHz, Methanol-d4) δ 10.06 (s, 1H), 8.23 (t, 2H), 8.05 (t, 2H), 7.93 (d, J = 6.6 Hz, 2H), 7.79 (t, J = 9.4 Hz, 2H), 7.59 – 7.53 (m, 2H), 7.20 (t, J = 8.7 Hz, 2H), 3.38 (s, 2H), 3.16 (s, 3H); 13C NMR (101 MHz, MeOD) δ 182.05, 159.63, 142.45, 138.17, 130.56, 130.47, 129.89, 128.45, 128.00, 127.26, 127.03, 125.86, 117.60, 45.11, 42.94; m/z [M+H]+ : 422.148
1-(2,4-dimethoxyphenyl)-2-(4-fluorophenyl)ethane-1,2-dione		Prepared according to procedure 1. Yellow solid (150 mg, 21 %): 1H NMR (400 MHz, Chloroform-d) δ 8.03 (d, J = 8.8 Hz, 1H), 7.93 (dd, J = 8.8, 5.4 Hz, 2H), 7.16 (t, J = 8.6 Hz, 2H), 6.66 (dd, J = 8.8, 2.3 Hz, 1H), 6.39 (d, J = 2.2 Hz, 1H), 3.89 (s, 3H), 3.56 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 192.93, 192.42, 166.95, 164.77, 162.44, 132.65, 131.95, 131.85, 117.09, 116.09, 115.87, 107.11, 98.34, 55.79; m/z [M+H]+ : 289.010

4-(4-(2,4-dimethoxyphenyl)-5-(4-fluorophenyl)-1H-imidazol-2-yl)benzonitrile		Prepared according to procedure 5. Yellow solid (67 mg, 32 %): 1H NMR (400 MHz, Chloroform-d) δ 10.14 (s, 1H), 7.97 (d, J = 8.5 Hz, 2H), 7.69 (d, J = 8.5 Hz, 2H), 7.59 (d, J = 8.3 Hz, 2H), 7.22 (d, J = 8.5 Hz, 1H), 7.01 (t, J = 8.8 Hz, 2H), 6.57 (s, 1H), 6.44 (dd, J = 8.6, 2.4 Hz, 1H), 3.88 (s, 3H), 3.84 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 163.34, 161.11, 160.90, 157.31, 134.00, 132.67, 131.24, 129.60, 125.21, 118.76, 115.38, 115.17, 111.41, 105.23, 99.26, 55.82, 55.49; m/z [M+H]+ : 400.157
(4-(4-(2,4-dimethoxyphenyl)-5-(4-fluorophenyl)-1H-imidazol-2-yl)phenyl)methanamine		repared according to procedure 7. Yellow solid (6 mg, 40 %): 1H NMR (400 MHz, Methanol-d4) δ 9.92 (s, 1H), 8.06 (d, J = 8.4 Hz, 1H), 7.88 (dd, J = 16.1, 8.4 Hz, 4H), 7.43 (d, J = 8.2 Hz, 1H), 7.37 (td, J = 8.8, 5.5 Hz, 2H), 7.10 (dd, J = 8.4, 6.1 Hz, 1H), 6.91 (td, J = 8.8, 6.4 Hz, 2H), 6.54 (dd, J = 4.6, 2.3 Hz, 1H), 6.48 (ddd, J = 8.3, 5.8, 2.4 Hz, 1H), 3.75 (d, J = 2.5 Hz, 2H), 3.57 (s, 3H), 3.25 (s, 3H); 13C NMR (101 MHz, MeOD) δ 192.07, 161.86, 158.63, 158.00, 141.76, 131.91, 129.87, 128.46, 126.93, 125.47, 125.11, 114.57, 114.29, 104.70, 102.97, 99.98, 98.40, 54.43, 51.86; m/z [M+H]+ : 403.122
4-(4-(2,4-dihydroxyphenyl)-5-(4-fluorophenyl)-1H-imidazol-2-yl)benzo nitrile		Prepared according to procedure 6. Yellow solid (25 mg, 56 %): 1H NMR (400 MHz, Methanol-d4) δ 8.12 (d, J = 8.4 Hz, 4H), 7.82 (d, J = 8.4 Hz, 4H), 7.57 (dd, J = 8.7, 5.5 Hz, 4H), 7.10 (t, J = 8.7 Hz, 4H), 7.00 (d, J = 8.4 Hz, 2H), 6.43 (d, J = 2.4 Hz, 2H), 6.27 (dd, J = 8.5, 2.4 Hz, 2H); 13C NMR (101 MHz, MeOD) δ 163.44, 161.00, 158.73, 156.94, 142.99, 133.97, 132.43, 130.18, 129.57, 129.49, 125.46, 118.27, 114.89, 114.67, 110.99, 109.47, 106.59, 102.77; m/z [M+H]+ : 371.107
4-(2-(4-(aminomethyl)phenyl)-5-(4-fluorophenyl)-1H-imidazol-4-yl)benzene-1,3-diol		Prepared according to procedure 7. Yellow solid (5 mg, 33 %): 1H NMR (400 MHz, Methanol-d4) δ 10.03 (s, 1H), 8.57 (s, 1H), 8.14 (d, J = 1.9 Hz, 2H), 8.07 (d, J = 8.4 Hz, 2H), 7.96 (dd, J = 16.5, 7.6 Hz, 2H), 7.58 (dd, J = 5.8, 2.9 Hz, 2H), 7.48 (d, J = 8.3 Hz, 1H), 7.20 (t, J = 8.7 Hz, 1H), 7.14 – 7.07 (m, 2H), 7.01 (d, J = 8.6 Hz, 1H), 6.45 – 6.39 (m, 1H), 6.25 (ddt, J = 8.1, 4.4, 2.8 Hz, 1H), 3.37 (s, 2H); 13C NMR (101 MHz, MeOD) δ 185.49, 173.67, 166.81, 156.92, 130.73, 129.93, 129.84, 129.66, 129.57, 128.58, 127.68, 125.33, 115.01, 114.83, 114.61, 108.38, 106.49, 102.78, 99.98, 51.86; m/z [M+H]+ : 375.138

Compound	Structure	Data
2-((4-methoxyphenyl)ethynyl)-6-methylpyridine		Prepared according to procedure 4. Yellow solid (400 mg, 80 %): 1H NMR (400 MHz, Chloroform-d) δ 7.55 (dd, J = 8.4, 6.7 Hz, 3H), 7.32 (d, J = 7.7 Hz, 1H), 7.08 (d, J = 7.7 Hz, 1H), 6.92 − 6.84 (m, 2H), 3.83 (s, 3H), 2.58 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 160.07, 158.86, 143.04, 136.30, 133.59, 124.14, 122.23, 114.48, 114.00, 89.04, 87.78, 55.30, 24.62; m/z [M+H]+ : 224.010
1-(4-methoxyphenyl)-2-(6-methylpyridin-2-yl)ethane-1,2-dione		Prepared according to procedure 4. Yellow solid (210 mg, 46 %): 1H NMR (400 MHz, Chloroform-d) δ 8.00 (d, J = 8.0 Hz, 1H), 7.92 (d, J = 8.9 Hz, 2H), 7.80 (t, J = 7.7 Hz, 1H), 7.38 (d, J = 7.8 Hz, 1H), 6.99 (d, J = 8.9 Hz, 2H), 3.91 (s, 3H), 2.52 (s, 3H); 13C NMR (101 MHz, CDCl3) δ 195.58, 195.58, 194.68, 194.68, 164.58, 159.18, 151.32, 137.05, 132.00, 127.79, 126.65, 120.57, 114.17, 55.58, 24.31; m/z [M+H]+ : 256.091
4-(5-(4-methoxyphenyl)-4-(6-methylpyridin-2-yl)-1H-imidazol-2-yl)benzo nitrile		Prepared according to procedure 5. Yellow solid (107 mg, 38 %): 1H NMR (400 MHz, Acetone-d6) δ 8.34 (d, J = 8.5 Hz, 2H), 7.87 (d, J = 8.5 Hz, 2H), 7.73 (d, J = 8.6 Hz, 2H), 7.63 (t, J = 7.7 Hz, 1H), 7.12 (d, J = 7.6 Hz, 1H), 7.01 (d, J = 8.8 Hz, 2H), 3.87 (s, 3H), 2.48 (s, 3H); 13C NMR (101 MHz, Acetone) δ 205.41, 205.21, 205.01, 159.57, 159.57, 157.58, 143.76, 143.76, 136.61, 134.48, 134.48, 132.50, 130.28, 125.86, 121.08, 118.47, 118.47, 118.24, 113.38, 111.19, 111.19, 54.71, 23.36; m/z [M+H]+ : 367.148
(4-(5-(4-methoxyphenyl)-4-(6-methylpyridin-2-yl)-1H-imidazol-2-yl)phenyl)methan amine		Prepared according to procedure 7. Yellow solid (12 mg, 79 %): 1H NMR (400 MHz, Methanol-d4) δ 8.57 (s, 1H), 8.20 (d, J = 8.5 Hz, 2H), 8.08 (d, J = 8.3 Hz, 2H), 8.01 (d, J = 6.6 Hz, 1H), 7.85 (d, J = 8.5 Hz, 1H), 7.66 − 7.58 (m, 1H), 7.55 (d, J = 8.1 Hz, 1H), 7.39 − 7.33 (m, 2H), 7.27 (d, J = 7.6 Hz, 2H), 7.17 (dd, J = 7.5, 4.6 Hz, 1H), 6.84 (dd, J = 8.6, 2.2 Hz, 2H), 4.08 (s, 1H), 2.58 (s, 3H); 13C NMR (101 MHz, MeOD) δ 173.40, 160.96, 158.02, 157.42, 136.93, 136.89, 132.41, 129.74, 128.59, 126.95, 126.02, 126.00, 125.73, 121.71, 121.51, 121.40, 114.97, 43.43, 22.60; m/z [M+H]+ : 371.183

4-(5-(4-hydroxyphenyl)-4-(6-methylpyridin-2-yl)-1H-imidazol-2-yl)benzo nitrile		Prepared according to procedure 6. Yellow solid (70.2 mg, 94 %): 1H NMR (400 MHz, DMSO-d6): δ 2.38 (3H, s), 7.01 (1H, dd, J = 7.9, 2.0 Hz), 7.08 (2H, ddd, J = 8.6, 1.5, 0.4 Hz), 7.60 (2H, ddd, J = 8.6, 1.8, 0.4 Hz), 7.70 (1H, dd, J = 7.9, 6.0 Hz), 7.85 (1H, dd, J = 6.0, 2.0 Hz), 7.92 (2H, ddd, J = 8.8, 1.6, 0.4 Hz), 7.99 (2H, ddd, J = 8.8, 1.6, 0.4 Hz); 13C NMR (101 MHz, DMSO-d6) δ 157.82, 157.7, 151.87, 148.25, 137.48, 133.32, 132.1, 129.05, 129.04, 128.98, 127.12364, 123.40, 119.6, 118.5, 115.01; 112.2, 24.33; m/z [M+H]+ : 353.115
4-(2-(4-(aminomethyl)phenyl)-4-(6-methylpyridin-2-yl)-1H-imidazol-5-yl) phenol		Prepared according to procedure 7. Yellow solid (7 mg, 13.8 %): 1H NMR (400 MHz, Methanol-d4) δ 9.00 (s, 1H), 8.85 (s, 1H), 8.15 (d, J = 8.3 Hz, 2H), 7.73 (t, J = 7.8 Hz, 1H), 7.63 (d, J = 8.3 Hz, 2H), 7.57 (s, 1H), 7.49 (d, J = 8.8 Hz, 2H), 7.33 (d, J = 7.9 Hz, 1H), 7.28 (d, J = 7.7 Hz, 1H), 7.02 (d, J = 8.8 Hz, 2H), 4.22 (s, 2H), 3.86 (s, 3H), 2.64 (s, 3H); 13C NMR (101 MHz, MeOD) δ 173.08, 165.46, 160.20, 157.44, 150.76, 149.59, 145.93, 138.50, 134.08, 129.82, 129.24, 126.36, 123.51, 122.26, 119.96, 119.26, 113.87, 54.44, 42.66, 21.89; m/z [M+H]+ : 357.205
2-chloro-5-(pyridin-2-ylethynyl)phenol		Prepared according to procedure 4. Yellow solid (90 mg, 13.6 %): 1H NMR (400 MHz, DMSO-d6) δ 10.61 (s, 1H), 8.61 (d, J = 4.8 Hz, 1H), 7.86 (td, J = 7.8, 1.8 Hz, 1H), 7.65 (dt, J = 7.8, 1.1 Hz, 1H), 7.45 – 7.40 (m, 2H), 7.14 (d, J = 1.9 Hz, 1H), 7.05 (dd, J = 8.2, 1.9 Hz, 1H); 13C NMR (101 MHz, DMSO) δ 153.66, 150.64, 142.51, 137.29, 130.94, 127.84, 124.12, 123.89, 121.98, 121.35, 119.52, 89.55, 87.88; m/z [M+H]+ : 229.945
1-(4-chloro-3-hydroxyphenyl)-2-(pyridin-2-yl)ethane-1,2-dione		Prepared according to procedure 4. Brown solid (70.6 mg, 77.1 %): 1H NMR (400 MHz, Methanol-d4) δ 8.66 (d, J = 4.7 Hz, 1H), 8.23 (d, J = 7.8 Hz, 1H), 8.11 (td, J = 7.7, 1.6 Hz, 1H), 7.95 – 7.87 (m, 1H), 7.69 (ddd, J = 7.7, 4.7, 1.3 Hz, 1H), 7.50 (d, J = 8.2 Hz, 1H), 7.46 (d, J = 2.0 Hz, 1H), 7.30 (dd, J = 8.2, 2.0 Hz, 1H); 13C NMR (101 MHz, MeOD) δ 209.11, 209.11, 195.55, 151.16, 149.56, 148.14, 137.78, 132.82, 130.46, 128.59, 123.74, 122.87, 122.87, 121.98, 121.20, 115.58, 48.27, 48.06, 47.84, 47.63, 47.42, 47.21, 46.99, 29.34; m/z [M+H]+ : 261.972

4-(5-(4-chloro-3-hydroxyphenyl)-4-(pyridin-2-yl)-1H-imidazol-2-yl)benzo nitrile		Prepared according to procedure 5. Yellow solid (49 mg, 48.7 %): 1H NMR (400 MHz, DMSO-d6) δ 10.59 (s, 1H), 8.70 (d, J = 5.4 Hz, 2H), 8.37 (d, J = 8.2 Hz, 3H), 8.13 (t, J = 7.6 Hz, 2H), 8.02 (d, J = 8.2 Hz, 3H), 7.74 (d, J = 8.2 Hz, 2H), 7.60 (s, 1H), 7.45 (d, J = 8.2 Hz, 2H), 7.31 (s, 2H), 7.09 (dd, J = 8.2, 2.0 Hz, 2H); 13C NMR (101 MHz, DMSO) δ 179.99, 164.66, 163.27, 157.41, 153.71, 153.71, 136.46, 133.36, 131.25, 130.62, 127.24, 126.87, 124.04, 123.33, 120.95, 119.13, 117.36, 110.51, 40.62, 40.41, 40.20, 39.99, 39.78, 39.57, 39.36; m/z [M+H]+ : 373.101
5-(2-(4-(aminomethyl)phenyl)-4-(pyridin-2-yl)-1H-imidazol-5-yl)-2-chloro phenol		Prepared according to procedure 7. Yellow solid (7 mg, 23.2 %): 1H NMR (400 MHz, Methanol-d4) δ 10.06 (s, 1H), 8.61 (s, 2H), 8.21 (d, J = 8.3 Hz, 2H), 8.14 (d, J = 7.9 Hz, 1H), 8.12 − 7.98 (m, 2H), 7.79 (s, 1H), 7.58 (d, J = 8.5 Hz, 1H), 7.51 (d, J = 8.0 Hz, 1H), 7.39 − 7.30 (m, 2H), 7.12 (s, 1H), 7.02 (s, 1H), 4.92 (s, 2H); 13C NMR (101 MHz, MeOD) δ 175.51, 153.75, 153.14, 136.86, 132.70, 132.45, 129.69, 128.61, 128.30, 126.98, 125.89, 122.34, 120.30, 120.13, 116.49, 61.42; m/z [M+H]+ : 377.137

7 Biochemistry and cell culture

7.1 Cell culture and lysis

Culture and lysis of human cancer and bacteria cells
K-562 cells were grown in Iscove's Modified Dulbecco's medium (PAN Biotech) supplemented with 10% fetal bovine serum (PAN Biotech) in 1000 mL suspension with vigorous agitation (90 rpm, 37 °C, 5% CO2). SW620 and HeLa cells were grown adherently in high glucose Dulbecco's Modified Eagles's Medium (PAN Biotech) supplemented with 10% fetal bovine serum (37 °C, 5% CO2). Cells were tested for mycoplasma contamination in-house.

E. coli K12 strain and DSM 30083 strain were inoculated from overnight culture in 1:100 dilution and grown in autoclaved Tryptic soy broth (Oxoid) medium in 200 mL suspension with vigorous agitation (200 rpm, 37 °C, 6 hours).

Human and bacteria cells were washed with PBS and lysed in CP lysis buffer containing 0.8% IGEPAL CA-630 (Sigma), 50mM Tris-HCl pH 7.5, 5% glycerol, 150 mM NaCl, 1.5 mM $MgCl_2$, 25 mM NaF, 1 mM Na_3VO_4, 1 mM DTT, phosphatase inhibitors cocktail (prepared in-house according to Phosphatase inhibitor cocktail 1, 2 and 3 from Sigma-Aldrich) and 1 eq. of protease inhibitor (SigmaFast, Sigma). Bacteria lysate was additionally sonicated. Lysates were cleared by ultracentrifugation. Protein concentration was determined with Bradford assay and adjusted to 5 mg/mL.

Lysis of placenta tissue

Placenta tissue aliquots were partially thawed. The tissue was then transferred onto a precooled petri dish and cut into fine pieces using a scalpel, excess of blood was removed after short centrifugation at high speed (5 min, 4 °C, 3900 rpm). The tissue was washed 5-10 times with cold PBS supplemented with protease inhibitor (1 x SigmaFast protease inhibitor tablet S8820 per 1L of PBS). The pellet was re-suspended in one volume CP lysis buffer and subjected to bead beating (2x 5500x 30s, 1x 6500x 10s). The lysate was incubated on ice for 30 min and frozen at –80 °C. Before use, lysate was cleared by centrifugation (1h at 3900 rpm, 30 min at 52000 rpm). Protein concentration was determined with Bradford assay and adjusted to 5 mg/mL.

7.2 Preparation of affinity matrices

7.2.1 Functionalization of NHS-activated Sepharose beads with alkyne, alcohol and amine for promiscuous immobilzation of natural products

Note: all bead immobilization procedures were performed with anhydrous solvents unless stated otherwise

Alkyne beads were prepared via the reaction between NHS-activated sepharose beads with propargylamine. In brief, for the beads bearing 2 μmol/mL alkyne 1 mL of settled Sepharose beads was washed in 3x 10 mL DMSO, then equilibrated in 1 mL of DMSO, 20 μL of 100 mM propargylamine stock was added to the beads followed by the addition of 15 μL triethylamine. Beads were incubated overnight on an end-over-end shaker at room temperature in the dark. Remaining free NHS groups were blocked with 50 μL ethanolamine overnight. Beads were washed with DMSO and ethanol and stored as 1:1 slurry in ethanol at 4˚C until further use.

Analogously, **amino-** and **alcohol-beads** were prepared via the reaction of NHS-activated sepharose beads with ethylenediamine or ethanolamine, respectively, in the presence of trietylamine. Here, however, the entire 20 μL/mL loading capacity of NHS-sepharose beads was converted to NH_2- or OH-.

7.2.2 Gold-catalyzed alkoxylation and cycloaddition for immobilization of tacrolimus on beads

Immobilization reactions were conducted in dry solvents under Argon unless stated otherwise

In the common for both immobilizations procedure, 1 mL of settled 2 µmol/mL alkyne beads were washed 4x10 mL of the assayed solvent (either ACN or DCM), 1 mol eq to the amount of alkyne on beads was added as a stock solution in the same solvent (e.g. 200 µL of 10 mM stock). 1-20% of respective gold catalyst (IPrAuNTf$_2$ or tBuXPhos Au(MeCN)SbF$_6$) was added as 10 mM stock solution in corresponding solvent. Reaction mixture was incubated on a thermoshaker at the assayed temperature (rt, 30 ^{0}C, 50 ^{0}C) and reaction progression was monitored via LC-MS. Upon reaction completeion, excess of reagents was removed via 3x10 mL washing with respective solvent, followed by 2x10 mL wash in ethanol. Beads were stored as 1:1 slurry in ethanol at 4˚C until further use.

7.2.3 Enyne metathesis for immobilization of tacrolimus on beads

A 1:1 slurry of washed alkyne beads (2 µmol/mL) was prepared in either ACN, DCM or ddH2O. 1 mol eq of tacrolimus was added as a 10 mM stock solution in the respective solvent, followed by the addition of 1-20 mol % of ruthenium catalyst (AquaMet or NtiroGrela) as 100 mM stock solutions in corresponding solvent. Reaction mixture was incubated on a thermoshaker in the range of temperatures (rt, 30 ^{0}C, 40 ^{0}C or 50 ^{0}C) and reaction progression was monitored via LC-MS. Upon reaction completeion, excess of reagents was removed via 3x10 mL washing with respective solvent, followed by 2x10 mL wash in ethanol. Beads were stored as 1:1 slurry in ethanol at 4˚C until further use.

7.2.4 Mitsunobu reaction for immobilization of tacrolimus on beads

A 1:1 slurry of washed NH$_2$- or OH-beads was prepared in either ACN or DCM. 10 µL of 100 mM tacrolimus stock in respective solvent was added, followed by the addition of 1 mol eq of triphenylphosphine, and 0.8 mol eq of 1,1'-(azodicarbonyl)dipiperidine (both as 100 mM stocks in respective solvent). Reaction mixture was incubated on a thermoshaker at the assayed temperature (30 ^{0}C or 50 ^{0}C) and reaction progression was monitored via LC-MS. Upon reaction completeion, excess of reagents was removed via 3x10 mL washing with respective solvent, followed by 2x10 mL wash in ethanol. Beads were stored as 1:1 slurry in ethanol at 4˚C until further use.

7.2.5 Functionalization of diazirine-loaded beads with natural products and cofactors

NHS-activated sepharose beads were functionalized with alkyl CF3 linker 2-(3-(trifluoromethyl)-3H-diazirin-3-yl)ethan-1-amine (CAS 2095409-03-1) or phenyl CF3 linker (4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)methanamine (CAS 1258874-29-1) with a coupling density of 1.5 µmol per mL beads. In brief, 1 mL of settled beads was washed in 3x 10 mL DMSO, then equilibrated in 1 mL of DMSO, 15 µL of 100 mM diazirine linker was added to the beads followed by the addition of 15 µL triethylamine. Beads were incubated for 20 h on an end-over-end shaker at room temperature in the dark. Remaining free NHS groups were blocked with 50 µL ethanolamine overnight. Analogously, Very Low density aminoethyl 4 rapid run beads (4RR-AVL4-25, Agarose Bead Technologies, ABT) were functionalized with SDA or LC-SDA linkers with a coupling density of 1.5 µmol per mL beads. Remaining free amino groups were blocked with 100 µL of 100mM NHS acetate in the presence of 20 µL triethylamine for 20 h in the dark. Beads functionalized with different diazirine linkers were combined in 1:1:1:1 ratio and equilibrated in mTHF. 500 µL of 1:1 beads-mTHF slurry was transferred into glass vials. 1 molar equivalent of compound in mTHF to linker was added, 10 µL of supernatant was taken for coupling control. Solvent was evaporated and dried bead-compound mixture was irradiated for 30 min with 365 nm 150 J/cm2 in an UV oven. Beads were then re-suspended in m-THF to the initial volume of 1:1 slurry, and 10 µL of

supernatant was taken for coupling control. MeOH was used as a substitute to mTHF for mTHF-insoluble compounds.

7.2.6 Preparation of RNA-loaded beads

NHS-activated sepharose beads were functionalized with custom-synthesized primary amine-containing RNA sequence (Sigma Aldrich) with 10 nM per mL beads coupling density. In brief, 0.5 mL of settled beads was washed three times with 5 mL deionized water, re-suspended in 0.5 mL (1:1) of 100 mM HEPPS (EPPS, Sigma Aldrich) buffer pH=8.5 NaOH, 10 μL of 500 nM RNA stock in 50 mM NaCl 1 mM EDTA was added and beads were incubated for 3 hours on end-over-end shaker at room temperature. Remaining free NHS groups were blocked with 100 μL of 1.5 % hydroxylamine stock in deionized water (incubation 1 h). Obtained RNA-loaded beads were washed 4 times with 5 mL of deionized water, then re-suspended 1:1 in deionized water and stored until further use.

7.2.7 Immobilization of amino-compounds for SAR study

NHS-activated sepharose beads were functionalized with in-house synthesized amino-compounds library analogues with a coupling density of 2 μmol per mL beads. In brief, 1 mL of settled beads was washed in 3x 10 mL DMSO, then equilibrated in 1 mL of DMSO, 20 μL of 100 mM compound stock was added to the beads followed by the addition of 15 μL triethylamine. Beads were incubated for 20 h on an end-over-end shaker at room temperature in the dark. Remaining free NHS groups were blocked with 50 μL ethanolamine overnight. Beads were washed with DMSO and ethanol and stored as 1:1 slurry in ethanol at 4°C until further use. To evaluate the influence of the compound density on beads, initially each library analog was functionalized on beads with 0.5, 1 and 2 μmol per mL beads densities.

7.2.8 Immobilization of BET bromodomain inhibitor I-BET726

NHS-activated sepharose beads were first reversed to bear amino-groups. In brief, 1 mL of settled beads was washed in 3x 10 mL DMSO, then equilibrated in 1 mL of DMSO, a 4:1 v:v mixture of aminoethanol and ethylenediamine was added to the beads reaching the final concentration of 200 μM in bead slurry, followed by the addition of 15 μL triethylamine. Beads were incubated overnight on an end-over-end shaker at room temperature in the dark. Beads were washed with 3x10 mL DMSO, and 2x10 mL DMF, and resuspended 1:1 in DMF. 10 μL of 100 mM I-BET726 stock was added, followed by the addition of 100 μL 200 mM diisopropylethylamine, 20 μL triethylamine and 100 μL 100 mM HATU. Beads were incubated overnight on an end-over-end shaker at room temperature in the dark. Remaining free amino groups were blocked with with 100 μL of 100mM NHS-acetate in the presence of 20 μL triethylamine overnight in the dark. Beads were washed with DMSO and ethanol and stored as 1:1 slurry in ethanol at 4°C until further use.

7.2.9 LC-MS coupling controls and calculation of compound conversion

The LC-MS measurement of compound loading controls was performed on an Ultimate 1100 HPLC system (Agilent Technologies) coupled via an ESI-interface to an Amazon Speed ETD ion-trap mass spectrometer (Bruker Daltonics). For sample preparation, 10 μl of coupling reaction supernatants were diluted in 90 μl of 0.1% formic acid in acetonitrile. 5-15 μl of the analyte solution was injected. A standard gradient was applied to a C18 analytical column ranging from 10% buffer B (0.1% formic acid in acetonitrile) in buffer A (0.1% formic acid in HPLC grade water) to 90% buffer B over a 30 min gradient. The mass spectrometer operated in the positive ion mode (ESI+). The detection range was set to 200 – 1600 m/z for natural products and to 100 – 800 m/z for linkers, respectively. Compound conversion (% immobilization) was calculated by manual integration of corresponding chromatographic peaks before and after UV irradiation using Compass Data Analysis software (Bruker Daltonics).

7.3 **Competition pulldown**

For competition pulldowns in 96-well plates, 500 µL of cell lysates (protein concertation in lysate adjusted to 5 mg/mL, 2.5 mg of protein per pulldown) were pre-incubated with 9 compound concentrations (ranging from vehicle DMSO to 100 µM) for 45 min at 4 oC in an end-over-end shaker. Affinity matrices were added to a 96-well filter plate (18 µL settled beads per well) and pre-equilibrated with 1 mL CP buffer (CP lysis buffer without phosphatase inhibitors, protease inhibitor and IGEPAL) and 2 mL of 0.4% IGEPAL CP buffer (CP lysis buffer without phosphatase and protease inhibitors, containing 0.4% IGEPAL detergent). Subsequently, compound-lysate mixtures were incubated with beads functionalized with the corresponding compound for 30 min at 4oC in an end-over-end shaker. To assess the degree of protein depletion from the lysates, a second pulldown was performed with fresh beads and the unbound protein flow through of the vehicle DMSO control lysates.

The beads were then consecutively washed with 1 mL of 0.4% IGEPAL CP buffer, 2 mL of 0.2% IGEPAL CP buffer and 3 mL of CP buffer. Proteins were denatured with 40 µL 8M urea, 10 mM DTT in 40 mM Tris HCl pH 7.4 (30 min, 40 oC) and alkylated with 4 µL 550 mM chloroacetamide (RT, 30 min). Urea concentration was diluted to 1 M by the addition of 250 µL 40 mM Tris HCl pH 7.4, proteins were digested overnight with 30 µL of 10 ng/µL trypsin. Digested peptides were eluted, acidified with 7 µL 10 % formic acid (final pH≈3) and subjected to C18 StageTip desalting.

7.4 **C18 StageTip desalting**

StageTips were prepared by packing 5xC18 material disks into a 200 µL pipette tip. C18 material was activated with 200 µL of acetonitrile, followed by washing with 200 µL of 0.1% formic acid in 50% acetonitrile and 200 µL of 0.1% formic acid in deionized water. Every washing step was followed by centrifugation at 1000 g. Acidified peptide digests were loaded twice onto StageTip columns and centrifuged at 500g. Columns then were washed twice with 200 µL of 0.1% in deionized water. Peptides were collected by double elution with 40 µL of 0.1% in 50% acetonitrile, frozen and dried in vacuum prior to LC-MS/MS analysis.

7.5 **Matrix loading density evaluation and silver gel staining**

In order to visually assess the level of unspecific protein binding by affinity matrices loaded with increasing densities of LC-SDA linker or the amino compounds from SAR study corresponding affinity matrices were subjected to incubation with 500 µL of K562 5 mg/mL lysate for 30 min at 4 oC in an end-over-end shaker (18 µL settled beads per well). The beads were then washed with 2 mL of 0.4% IGEPAL CP buffer and 2 mL of 0.2% IGEPAL CP buffer. Proteins were denatured and reduced within 30 min with 40 µL of 2xNuPAGE LDS Sample Buffer (Invitrogen, NP0007) containing 50 mM DTT at 50 oC. Harvested eluates were alkylated with 4 µL of 500 mM chloroacetamide. Half of the eluate was loaded onto NuPAGE 4-12% Bis-Tris gel (Invitrogen, NP0335) and separated at 200 Volt for 45 min in an electrophoresis chamber filled with 1x NuPAGE® MOPS SDS Running Buffer (Invitrogen, NP0001-02). Proteins were fixed on a gel slab by incubation with 50:5:45 v/v/v methanol : acetic acid : water for 1h, and subsequently washed with deionized water for 1h. Gels were sensitised by incubation with 0.02 % sodium thiosulfate for 2 min, rinsed with water and stained with cold 0.1 % AgNO3 for 30 minutes at 4 ˚C. Silver nitrate solution was then discarded, gels were rinsed with water and developed with 0.04 % formaldehyde in 2 % sodium carbonate. When a sufficient degree of staining was obtained, the staining was quenched by discarding the developing solution and replacement with 1 % acetic acid.

7.6 **LC-MS/MS analysis**

7.6.1 Molecule-centric proteome-wide pharmacophore SAR proof of concept study

Pulldown samples were analysed on an Orbitrap Q Exactive HF (Thermo Fisher Scientific) mass spectrometer coupled an online Dionex Ultimate3000 equipped with nano HPLC pump (Thermo Fischer Scientific). Peptides were delivered to a trap column within 10 min in 0.1% formic acid in HPLC grade water at 5 µL/min. Peptides were then separated on an analytical column over a 50 min gradient ranging 4-32 % solvent B (0.1 % formic acid, 5 % DMSO in acetonitrile) in solvent A (0.1% formic acid, 5 % DMSO in HPLC grade water) at 300 nL/min. Mass spectrometer operated in data dependent mode, and MS1 spectra were acquired over a mass-to-charge ratio of 360-1300 m/z at a resolution of 60,000 in the Orbitrap. Maximum injection time was set to 10 ms and automatic gain control (AGC) target value to 3e6. Top 12 most abundant peptide precursors were isolated (isolation window of 1.7 m/z, maximum injection time of 75 ms, and AGC target value of 2e5), fragmented by HCD using 25% normalized collision energy (NCE) and analysed in the Orbitrap at a resolution of 15,000. The dynamic exclusion was set to 30s.

7.6.2 Promiscuous immobilization of natural compounds study

Samples were analysed on an Orbitrap Fusion Lumos Tribrid (Thermo Fisher Scientific) mass spectrometers coupled an online Dionex Ultimate3000 equipped with either nano HPLC or micro flow Vanquish UHPLC pumps (Thermo Fischer Scientific).

In total three measurement methods were used in the course of this study. 1) Peptides were delivered to a trap column within 10 min in 0.1% formic acid in HPLC grade water at 5 µL/min. Peptides were then separated on an analytical column over a 50 min gradient ranging 4-32 % solvent B (0.1 % formic acid, 5 % DMSO in acetonitrile) in solvent A (0.1% formic acid, 5 % DMSO in HPLC grade water) at 300 nL/min. Mass spectrometer operated in data dependent mode, and MS1 spectra were acquired over a mass-to-charge ratio of 360-1300 m/z at a resolution of 60,000 in the Orbitrap. Maximum injection time was set to 50 ms and automatic gain control (AGC) target value to 5e5. Top 12 most abundant peptide precursors were isolated (isolation window of 1.7, maximum injection time of 22ms, and AGC value of 1e5), fragmented by HCD using 28% normalized collision energy (NCE) and analysed in the Orbitrap at a resolution of 15,000. The dynamic exclusion was set to 20s. In course of the project, the MS measurement method was optimized following the Orbitrap Fusion Lumos upgrade (LC parameters unchanged). 2) Raw data was recorded in data dependent mode, and MS1 spectra were acquired over a mass-to-charge ratio of 360-1300 m/z at a resolution of 60,000 in the Orbitrap. Maximum injection time was set to 50ms and automatic gain control (AGC) target value to 4e5. A fixed cycle time of 2s was selected and peptide precursors were isolated (isolation width of 1.2Th, maximum injection time of 50ms, AGC value of 2e5), fragmented by HCD using 30% normalized collision energy (NCE) and analysed in the Orbitrap at a resolution of 30,000. The dynamic exclusion was set to 30s. 3) Peptides were directly loaded onto an analytical column and separated over a 15 min gradient ranging 3-28 % solvent B (0.1 % formic acid, 5 % DMSO in acetonitrile) in solvent A (0.1% formic acid, 5 % DMSO in HPLC grade water) at 50 µL/min. The mass spectrometer operated in data dependent mode, the MS1 spectra were acquired over a mass-to-charge ratio of 360-1300 m/z at a resolution of 120,000 in the Orbitrap. Maximum injection time was set to 50 ms and automatic gain control (AGC) target value to 4e5. A fixed cycle time of 0.6s was selected and peptide precursors were isolated (isolation window of 0.4, maximum injection time of 10ms, and AGC value of 1e5), fragmented by HCD using 32% normalized collision energy (NCE) and analysed in the IonTrap. The dynamic exclusion was set to 12s.

7.7 **Peptide and protein identification and quantification**

Peptide and protein identification and quantification was performed with MaxQuantv.1.5.3.30. Acquired raw MS data was searched against all canonical protein sequences annotated in Swissprot reference databases (Human and Escherichia Coli, respectively) using the integrated search engine Andromeda. Trypsin/P was specified as the proteolytic enzyme with up to two missed cleavage sites allowed. Cysteine carbamidomethylation was set as fixed modification, methionine oxidation and protein N-terminal acetylation were set as variable modifications. IBAQ, label free quantification (LFQ) and match between runs (MBR) were enabled. All searches were performed with 1% PSM and protein FDR. Other search parameters were set as default.

7.8 **Competition pulldown data analysis**

For plotting of dose response competition curves (four-parameter log-logistic regression) an in-house built R pipeline was utilized. Relative binding for every compound concentration and correction (correction factor) for protein depletion was calculated as a ratio of LFQ intensity for every compound concentration to the DMSO control. The correction factor is calculated as a ratio of LFQ intensities from the second consecutive pulldown with the flowthrough of the DMSO vehicle pulldown to the LFQ intensities in the DMSO vehicle control.

EC50 values were derived from dose response curves. Apparent binding constants $K_D^{app}s$ were calculated by multiplying EC50 values with a correction factor ($K_D^{app} = EC50 * correction\ factor$). Targets or interactors were annotated manually if the resulting binding curve showed typical sigmoidal shape. Unique peptides and acquired MSMS spectral counts for each condition that showed a dose-dependent decrease were used as additional level of evidence for binder annotation, together with competition curves based on protein intensity.

7.9 **Immunoblot analysis**

HeLa cells were seeded in 6-well plates ($0.3x10^6$ cells per well) and incubated for 24h before treatment. Compound dilutions were prepared in DMSO (0, 0.1, 0.5, 1, 10, and 100 µM final concentrations) and added to the cells to a final DMSO concentration of 0.3%. Cells were incubated with drug for 2 hours, washed with PBS and lysed with CP lysis buffer containing 0.8% IGEPAL CA-630 (Sigma), 50mM Tris-HCl pH 7.5, 5% glycerol, 150 mM NaCl, 1.5 mM MgCl2, 25 mM NaF, 1 mM Na3VO4, 1 mM DTT, phosphatase inhibitors cocktail (prepared in-house according to Phosphatase inhibitor cocktail 1, 2 and 3 from Sigma-Aldrich) and protease inhibitor (SigmaFast, Sigma). Antibodies against Phospho-SMAD2 (Ser465/467), SMAD2, GAPDH, and β-Actin were purchased from Cell Signaling technology. 60-100µg of protein (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).

7.10 **Sepiapterin reductase (SPR) enzyme activity assay**

Recombinant human Sepiapterin reductase (SPR) was purchased from R&D systems (E. coli-derived, 10209-SP-020), L-sepiapterin was purchased from Santa Cruz Biotechnology (sc-506156), NADPH was purchased from Sigma-Aldrich (10107824001). Inhibition assay was performed in a dose range triplicates in 200µL total volume of 0.1µg SPR, 50µM L-sepiapterin, 100µM NADPH and respective concentration of tested compound in 50mM potassium phosphate pH 6.5. In brief, 1 µL of 200-fold concentrated stock solutions of tested compound in DMSO (0, 1, 3, 10, 30, 100, 300, 1000, 3000 and 30000nM final concentrations) were added in triplicates to a 96-well plate and diluted with 160µL of 50mM potassium phosphate pH 6.5. Then 0.5µL of enzyme (0.2mg/mL)

was added to each well, followed by the addition of 20μL 1mM NADPH (100μM final concentration). Reaction was initiated by the addition of 20μL 0.5mM L-sepiapterin (50μM final concertation) and incubated for 1h at 37°C. Finally, the absorbance at 420 nm was measured (consumption of L-sepiapterin) and activity of sepiapterin reductase (SPR) was calculated as:

$$SPR\ activity = \frac{\Delta Abs}{time, min}$$

Where ΔAbs is defined as:

ΔAbs=median triplicate absorbance(cpd conc X)-median triplicate absorbance(DMSO ctrl)

SPR activity for different compound concentrations was plotted relative to DMSO ctrl (100%).

Results and Discussion

8 Part 1. Natural molecules in affinity-based target deconvolution

Part of this work was included in the publication: Prokofeva, P.; Höfer, S.; Hornisch, M.; Abele, M.; Kuster, B.; Médard, G. Merits of Diazirine Photo-Immobilization for Target Profiling of Natural Products and Cofactors. ACS Chem. Biol. 2022. [92] Some material published in the article (including several figure panels) are reprinted in this thesis with permission from ACS Chem.Biol. Copyright 2022 American Chemical Society.

Two master thesis projects (Stefanie Höfer and Maximilian Hornisch) were carried out in the context of this work (supervisor: Polina Prokofeva).

8.1 A chemical toolbox for diversity-oriented immobilization of natural products (NPs)

To accelerate research of natural molecule action on the proteome, an expeditious general protocol for target deconvolution is highly desirable. Potent bioactive natural molecules are expected to show strong binding affinities for their protein targets. Therefore, an affinity matrix based on the chemical structure of the compound in question should efficiently enrich molecular targets from biological samples (lysates). In turn, these can then be identified using mass spectrometry. AfBPP affinity matrices are prepared from chemically synthesized linkable analogs of bioactive small molecules. Such analogs often contain amino- or carboxyl- groups that allow for their direct immobilization on beads via nucleophilic addition-elimination reaction and amide bond formation (Fig. 8).
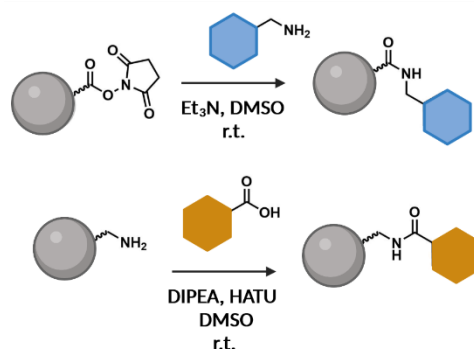
Figure 8. Schematic representation of common immobilization strategies via amino- or carboxyl- groups.

As previously mentioned in the introduction chapter, preparation of NP-based affinity probes is largely hampered by the complexity of their chemical structure and the frequent absence of SAR that would indicate a good anchor point for functionalization (Fig.9). Moreover, late-stage functionalization is typically not possible due to the difficulty of NP isolation in sufficient amounts. A way to circumvent these bottlenecks is to take advantage of various native NP functional groups and perform a single immobilization step of the unmodified NP using different chemistries. This approach holds an entirely opposite philosophy compared to selective functionalization and immobilization of chemically prepared linkable analogs, as, in this case, the chemo- and regio-selectivity is not sought. Here, promiscuity of the immobilization is a key asset: multiple immobilization chemistries via various moieties of the NP would produce a range of bead-coupled analogs. The promiscuous nature of immobilization statistically increases the probability of having a higher fraction of a compound on beads with complete preservation of its target engagement ability. Such an approach would allow for unbiased target deconvolution, it does not require any prior SAR knowledge, and simultaneously eliminates the cumbersome chemical synthesis of

linkable NP analogs. Besides, in this way, low quantities of unmodified molecules should become amenable for target deconvolution. Hence, the first part of this work was dedicated to evaluating different chemistries that could be employed for the on-bead immobilization of NPs, where tacrolimus (or FK506) was selected as a model compound (Fig. 10).
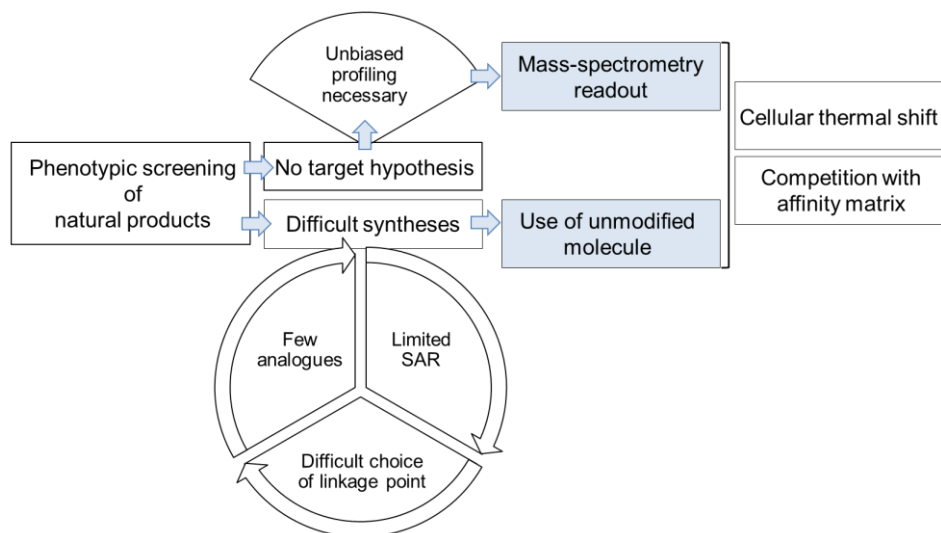
Figure 9. Challenges of natural product target deconvolution can be addressed using mass-spectrometry as a readout together with chemical proteomics approaches using unmodified molecules. The difficulty of the synthesis creates a vicious circle, where limited or no SAR is available to choose an anchorage point and to synthesize a linkable analog: the use of the unmodified molecule is then desirable. After phenotypic screening, an unbiased proteome-wide method is necessary: mass-spectrometry based proteomics readout hence appears as a natural choice. Adapted from the DFG grant proposal by Dr. Médard.

Photo-affinity labeling (PAL) combined with quantitative proteomics has received much attention in the last years and put a vast number of photo-activatable probes in the spotlight as tools to study protein-drug interactions in complex proteomes. The approach utilizes probes that are small molecules whose targets are being investigated, commonly functionalized with benzophenone, aryl azide, or diazirine photo-crosslinkers. Upon photo-exposure, highly reactive adducts are formed that directly covalently modify proximal molecules.[93] In the context of drug-protein interactions, the proteins covalently modified by the probe are the likely targets of the tested compounds. Implementation of photo-induced chemistry of PAL, which in theory allows covalent insertion into any C-H or X-H bond on the molecule, has been shown to be of potential interest for affinity matrices preparation in the context of NPs target profiling.

In an effort to circumvent laborious affinity probes synthesis, Kanoh et al. pioneered the implementation of UV-induced photo-immobilization of small molecules. Classically, the PAL approach utilizes chemical probes functionalized with photo-activatable moieties that covalently bind its target upon light activation. Kanoh et al. have repurposed the photo-activatable chemical moieties to photo-crosslink molecules onto beads rather than proteins to PAL probes. Kanoh et al. prepared an affinity matrix functionalized with a diazirine handle and utilized it as a tool to immobilize the immunosuppressive drug of natural origin tacrolimus (FK506).[94–96] Upon UV irradiation, the diazirines that decorate the beads convert to carbenes and covalently insert (presumably randomly and promiscuously) into proximal C-H or X-H bonds of bead-adsorbed molecules (Fig. 10). This approach adopts the "promiscuity is a key" philosophy mentioned above:

instead of seeking for selective insertion of an enrichment anchor, one takes advantage of promiscuous unselective carbene insertion chemistry, which allows the use of the unmodified natural product. A Novartis team has recently followed up on this technique and applied it to the study of a natural product with an unknown mode of action, chivosazole F. By competition between immobilized chivosazole F and either a vehicle control (such as DMSO) or 100 µM chivosazole F dosed into a HEK293T cell lysate, they identified actin-containing protein complexes as interactors of chivosazole F.[97] A recent publication by Melder et al. has shown that phenyl diazirine decorated cellulose membranes can also be used for target screening of bioactive natural molecules. They have compared the protein enrichment by membranes loaded with cyclosporine A, tacrolimus, sirolimus, or lenalidomide with unmodified control membranes and recapitulated known binders of the drugs.[98] Inspired by this elegant in its simplicity concept, this work sought to explore the scope of unselective photo-immobilization for the preparation of NP-affinity probes more systematically.

The study by Ertl and Schumann has computationally evaluated the distribution of common functional groups for the NPs within The Dictionary of Natural Products (DNP). They reported the most frequent NP functional group being the alcohol hydroxy group, which was present in 61% of all molecules, followed by alkene (40%), ether (35%), ester (29%), and the phenolic hydroxy group (28%).[99] Therefore, this work equally evaluated other chemistries that would allow for NP immobilization via some of these common moieties. These included gold-catalyzed alkoxylation for the functionalization via hydroxy groups, gold-catalyzed cycloaddition via the alkene groups, ruthenium-mediated metathesis, and finally, the versatile Mitsunobu reaction for the immobilization via hydroxy groups on beads functionalized with different nucleophiles (Fig. 10).
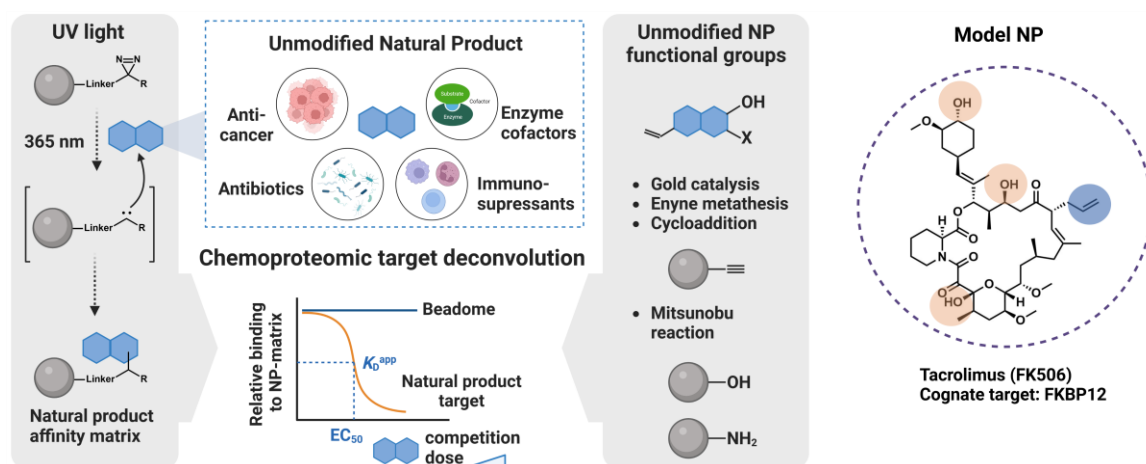
Figure 10. Schematic representation of the agnostic target deconvolution of natural products (NPs) with tailored affinity probes: immobilization of unmodified molecules via various functional groups and the application of obtained affinity matrices in the dose-dependent chemoproteomic profiling. In order to benchmark several immobilization chemistries tacrolimus (or FK506) was utilized as a model compound.

8.2 Benchmarking immobilization strategies with tacrolimus

The selection of chemical immobilization approaches mentioned above for the NP target deconvolution were evaluated on tacrolimus and relied on established and robust protocols for

affinity pulldown coupled to MS readout (Fig. 11).[46] The efficiency of compound immobilization was determined from LC-MS coupling controls, where the bead supernatant LC-MS/MS profiles before and after the immobilization reaction were compared, and compound conversion was determined as a ratio of integrated respective compound peaks. The performance of obtained tacrolimus-loaded affinity matrices was evaluated based on their individual abilities to enrich FKBP12, the cognate target of tacrolimus, out of K562 cell lysate.
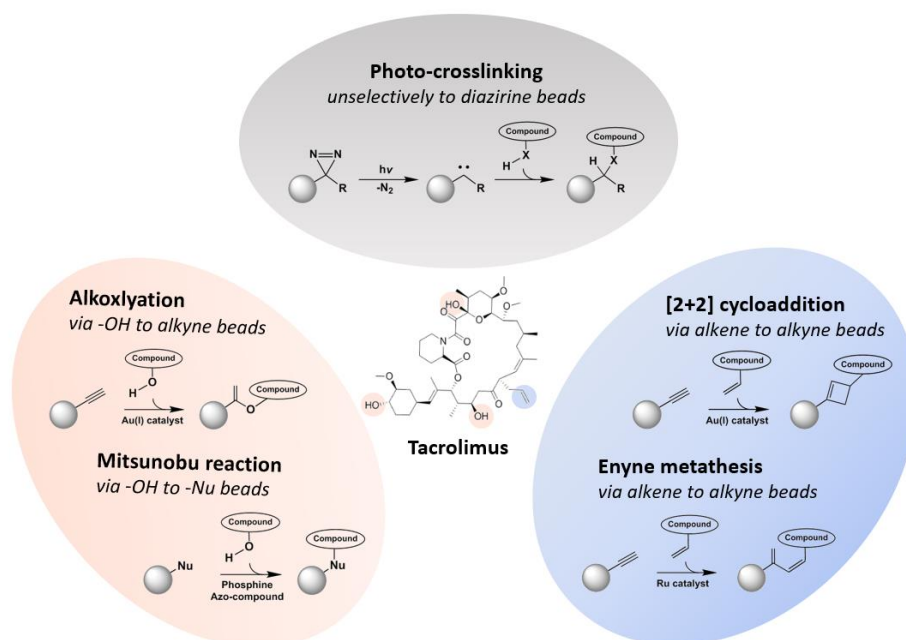
Figure 11. Schematic representation of the chemistries evaluated for the immobilization of tacrolimus with respective types of functionalities on affinity matrices used. Adapted from the Master Thesis of Stefanie Höfer (supervised by Polina Prokofeva).

8.2.1 UV-induced photo-immobilization on diazirines beads

First, to optimize UV-induced (365 nm) immobilization, the optimal irradiation time was determined for beads loaded with 2 μmol/mL succinimidyl 6-(4,4'-azipentanamido)hexanoate (LC-SDA) and using a two-fold molar excess of tacrolimus. Six irradiation times were evaluated from 10 min up to 1 h. Tacrolimus conversion reached a plateau after 30 min, which was then set for all subsequent experiments (Fig. 12a,c). To assess the optimal compound-to-linker ratio, tacrolimus was titrated to the 2 μmol/mL LC-SDA beads in 0.5, 1.0, 1.5 and 2.0 molar equivalents. There was no obvious difference in tacrolimus loading to the beads within the tested range (Fig. 12a,d). Pulldowns using the obtained matrices enriched FKBP12 in comparable quantities. One dose competition assays with 10 μM tacrolimus as competitor confirmed the specificity of the enrichment with slightly better performance for the affinity matrix prepared with equimolar amounts of linker and compound (Fig. 12f). Thus, the equimolar compound-to-linker ratio was set for all subsequent experiments.
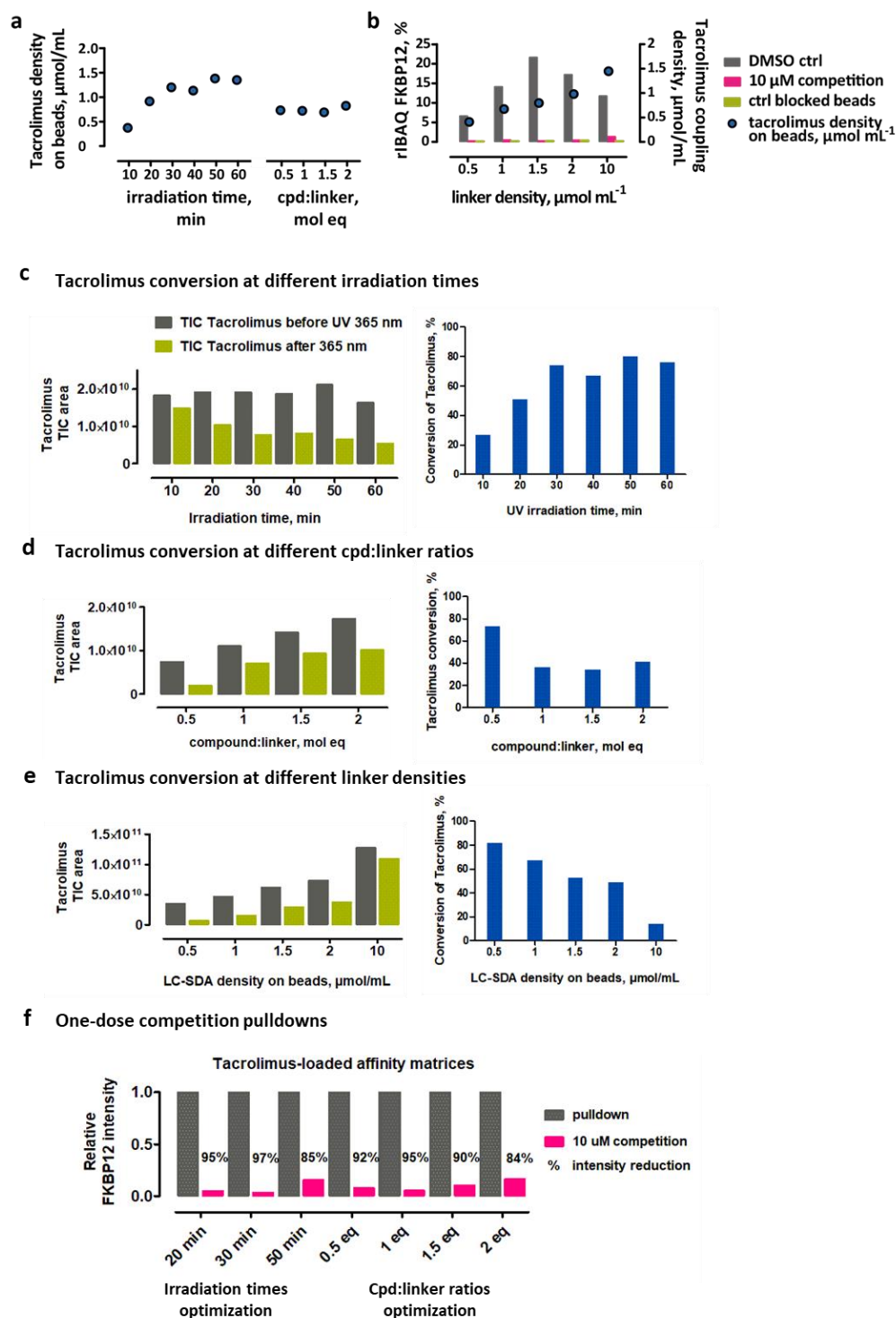
Figure 12. Optimization of photo-crosslinking conditions using tacrolimus as a model system a) tacrolimus apparent loading on beads (blue circles) as a function of irradiation time or compound:linker ratio. 2 molar excess of tacrolimus was added to 2 µmol mL^{-1} LC-SDA-loaded beads for irradiation time tests, for cpd:linker optimization respective mol amount of tacrolimus was titrated to 2 µmol mL^{-1} LC-SDA-loaded beads. Tacrolimus final loading on beads was determined from LC-MS coupling controls. b) Relative amount of FKBP12 bound to beads (expressed as the fraction of FKBP12 vs total protein intensity-based absolute quantification, rIBAQ, %) as a function of linker density in the presence of DMSO (grey), in competition with 10 µM free tacrolimus (pink) and bound to beads that do not display immobilized tacrolimus (green, blocked beads). Blue circles indicate the density of tacrolimus on beads. c-e) LC-MS coupling controls (grey-green) and respective tacrolimus conversion (blue) for different optimization steps. f) One-dose competition assay with obtained affinity matrices.

The influence of the diazirine linker density on beads was further evaluated. Not surprisingly, immobilization on beads decorated with higher LC-SDA densities led to higher coupling density (in mol amount) of tacrolimus (Fig. 12b,e). However, the reaction showed to be rather inefficient (lower conversion) with increasing diazirine densities. For instance, in the case of 10 μmol/mL LC-SDA loading, conversion of tacrolimus was found to be 3.5-times less than with 2 μmol/mL LC-SDA (49%) and nearly 6-times less compared to 0.5 μmol/mL LC-SDA (82%, Fig. 12e). Possibly, at higher linker densities light-induced activation of the increased amount of diazirine moieties is insufficient. Relative abundance of FKBP12 within the measured proteome amounted to at least 6% (0.1% for the control beads) for all affinity matrices and even reached an impressive 20% (0.2% for control beads) for the beads prepared with 1.5 μmol/mL LC-SDA and tacrolimus (Fig. 12b). At higher linker densities, overall increased unspecific background binding led to lower relative FKBP12 abundance (Fig. 13). Single-dose competition pulldown experiments with 10 μM Tacrolimus for all affinity matrices also defined 1.5 μmol/mL as the optimum linker coupling density. At this density, indeed, the ratio of measured FKBP12 was maximal between the vehicle and the drug-treated samples (Fig. 12b).
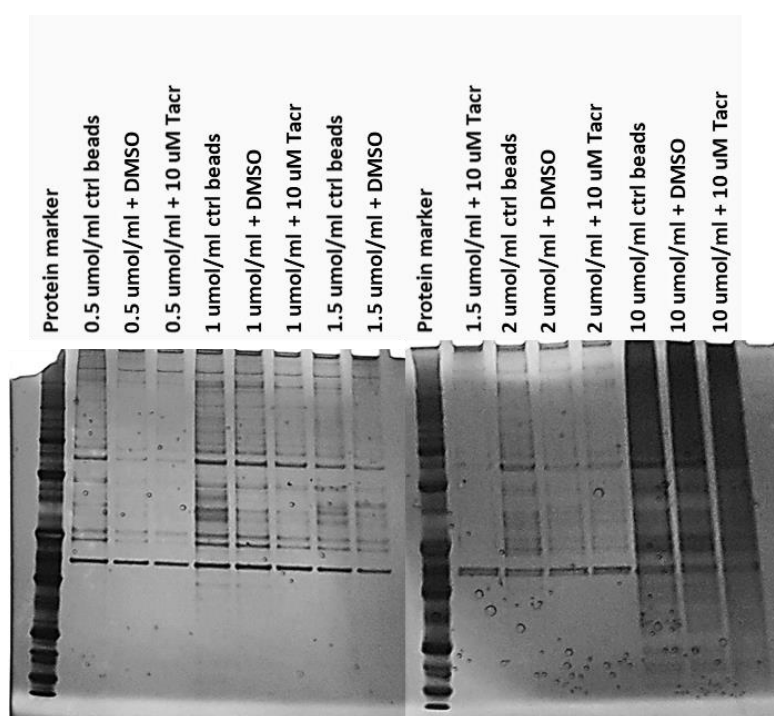
Figure 13. Influence of diazirine linker density on background protein binding: eluates from linker density optimization pulldowns were resolved on a gel by SDS-PAGE, and proteins were stained with silver nitrate (see methods for details). Control beads were prepared by subjecting Sepharose beads decorated with respective densities of LC-SDA linker to 30 min 365nm irradiation in the absence of tacrolimus. High linker density (10 μmol/mL) leads to an apparent increase of unspecific protein binding (smear on the gel) unrelated to tacrolimus loading on the beads. Adapted from the Master Thesis of Stefanie Höfer (supervised by Polina Prokofeva).

The chemical nature of the photo-activatable linker can influence the efficiency of the carbene insertion. It has been shown that, upon UV irradiation, diazirines can undergo diazo-isomerization, resulting in undesired linear diazo-intermediates.[100] Length, bulkiness or hydrophobicity of the spacer can additionally influence (un)specific enrichment of proteins from cell lysate. Therefore, new tacrolimus-loaded matrices were prepared using four commercially available photo-

activatable linkers (Fig. 14a). Beside LC-SDA, the shorter SDA and two trifluoromethyl diazirine molecules, namely the alkyl CF_3 linker 2-(3-(trifluoromethyl)-3H-diazirin-3-yl)ethan-1-amine (CAS 2095409-03-1) and the analogous phenyl CF_3 linker (4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl)methanamine (CAS 1258874-29-1) were evaluated (Fig. 14a). Under the optimized reaction conditions, in this experiment, tacrolimus conversion stretched between 25% (LC-SDA) and 60% (alkyl CF_3), with a more efficient conversion for the two CF_3 diazirine linkers (Fig. 14b). Enrichment of FKBP12, however, did not correlate with apparent Tacrolimus loading and ranged between 3.4% for the densest aliphatic trifluoromethyl-based beads and 27% for LC-SDA (Fig. 14b,c). It has to be deduced that either the carbene insertion follows different regioselectivity rules for the different linkers and/or that the chains of atoms that separate the Tacrolimus molecule from the bead surfaces have a strong impact on the binding to FKBP12. The approach builds on the unselective nature of carbene insertion. Hence, the immobilization may render an unknown fraction of the compound inactive, leading to a loss of protein binding. Another fraction might, however, be attached via diverse anchor points that maintain bioactivity. This experiment indicates that such balance is linker-dependent.

With an intention to generalize the experimental procedure, it was essential to maximize the chances of bioactivity-maintaining immobilization. Therefore, the immobilization of tacrolimus on a mixture of beads displaying the four diazirine linkers was tested. In this pulldown, FKBP12 constituted more than a third of the entire enriched sub-proteome (Fig. 14c). Crucially, the dose-response curves of native FKBP12 obtained using individual linker-derived or mixed linker-based tacrolimus matrices all exhibited typical inhibitory sigmoidal shapes that allowed to determine consistent affinities between 39 nM and 140 nM (Fig. 14d).
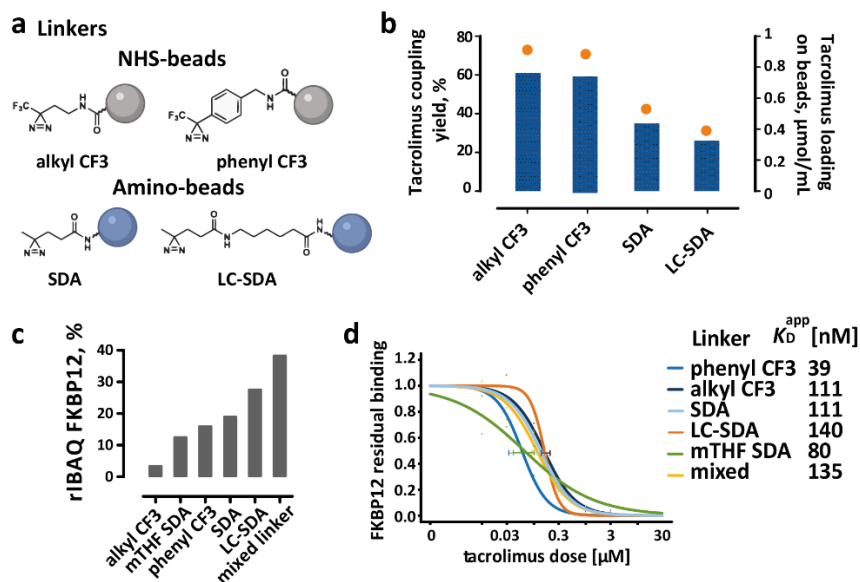
Figure 14. Optimization of photo-crosslinking conditions using tacrolimus as a model system part 2. a) Structures of the diazirine linkers evaluated for photo-crosslinking of tacrolimus. b) Tacrolimus conversion rates and respective density on beads. c) Relative amount of FKBP12 vs total protein bound to beads for the different linkers and the mixture of all linkers. d) Residual binding of FKBP12 to beads as a function of increasing doses of free tacrolimus used as a competitor.

All optimization experiments discussed so far employed diazirine beads evaporated to dryness, which, in fact, largely alters beads' morphology and causes them to shrink. In order to compare the affinity matrices produced on dried beads to immobilization in suspension, an additional

experiment was performed, where tacrolimus was added to the SDA-loaded beads in radical-friendly mTHF suspension and subjected to UV irradiation followed by a full dose competition pulldown. The enrichment of FKBP12 here was reduced by more than half (from 28% rFKBP12 to 12%) (Fig. 14c), and the dose-dependent competition resulted in a poorer quality curve (Fig. 14d). Based on these results, the assay was fixed with 1.5 µmol/mL mixed diazirine-loaded beads and 1 molar equivalent of a compound, that upon mixing are evaporated to dryness and exposed to 30 min of UV irradiation at 365 nm.

Overall, all evaluated tacrolimus affinity matrices prepared through UV-induced carbene insertion allowed for consistent FKBP12 identification as tacrolimus target. The addition of four diazirine linkers should further contribute to the promiscuity of such immobilizations, as it is likely for each linker to have different regioselectivity of insertion. It is definitely a very interesting question where the carbene insertion takes place and whether certain functional groups of NPs may be more susceptible. In the previous publication, Kanoh et al. performed an experiment to characterize the regioselectivity of carbene insertion using the photocleavable diazirine linker, chemically very similar to the phenyl CF3 linker in this study.[96] In their work, ten fairly small model compounds, each containing one of the functional groups commonly found in bioactive small molecules (i.e., alcohol, amine, thiol, carboxylic acid, and aromatics), were photo-immobilized. The use of the second UV wavelength allowed the cleaving of the compounds off the beads, and cleavage products were then analyzed. It has been observed that compounds containing an alcohol, a carboxylic acid, or an aromatic ring produced multiple different conjugates. While some functional groups (amine and thiol) reacted predominantly with the photogenerated carbene to produce an X–H insertion product (X = N or S). In addition, in the study by Woo et al. it was shown that alkyl and aryl-diazirines have different labeling preferences.[101] One may speculate that NPs with multiple functional groups previously reported being amenable for carbene insertion will likely produce immobilized analogs via various linkages. Given that four different diazirines were used in this work, photo-crosslinking could potentially produce multiple different carbene insertions allowing for truly promiscuous immobilization.

8.2.2 Gold(I)-catalyzed intermolecular alkoxylation on alkyne beads.

Under gold(I)-mediated catalysis, unsaturated carbon bonds, particularly alkynes, can undergo electrophilic activation and react with a plethora of functional groups, which generally allows for the construction of molecular complexity.[102] One notable example of a transformation that can take place upon gold(I)-catalysis is the intermolecular alkoxylation with the addition of nucleophiles, such as –OH and –NH_2. Trinchillo et al. have reported an extensive evaluation study of such alkoxylation, where the N-heterocyclic carbene-based gold(I) complex [(NHC)AuNTf$_2$] proved to be an effective catalyst (Fig. 15a).[103] This catalyst was employed here for benchmarking tacrolimus immobilization on alkyne-functionalized beads via its native hydroxyl groups. The alkyne beads were prepared through the nucleophilic addition-elimination reaction of NHS-ester-bearing sepharose beads with propargylamine following the general approach illustrated in Fig. 8.
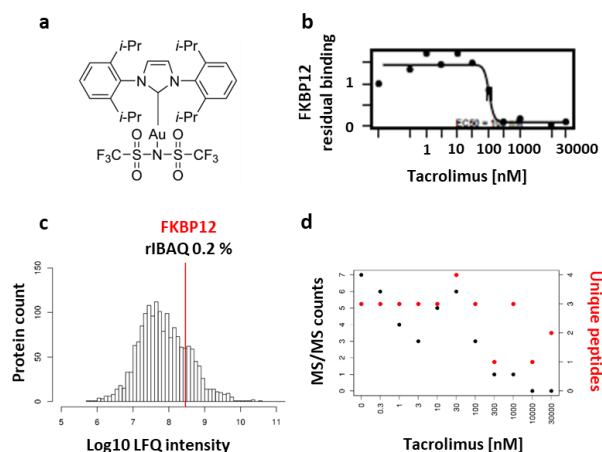
Figure 15. Gold(I)-catalyzed intermolecular alkoxylation on alkyne beads. a) Chemical structure of the gold(I) complex [(NHC)AuNTf2] used for internal alkoxylation. b) FKBP12 dose response curve from full-dose competition experiment. c) FKBP12 enrichment out of K562 lysate. d) Number of MS-identified unique peptides and MS2 spectra in each competition condition.

Initial trial experiments were performed under the exact conditions reported by Trinchillo et al., i.e., 2 µmol/mL alkyne-derivatized beads were reacted with a 4-fold molar excess of tacrolimus in the presence of 1 mol% catalyst for 1 h at 30 °C in anhydrous acetonitrile. These conditions led to 29% conversion of tacrolimus; however, in a subsequent pulldown experiment, FKBP12 was not sufficiently enriched from K562 cell lysate. Consequently, no effect was observed upon two-dose competition with 10 µM and 30 µM tacrolimus. In order to improve the immobilization efficiency, several alterations to the experimental procedure were assayed. For instance, the experiment was repeated with prolonged reaction time (overnight instead of 1 h), which, surprisingly, resulted in no conversion of tacrolimus. Temperature increase to 50 °C also did not lead to the increase in compound conversion. Since large excess of NP did not improve the coupling efficiency, for all subsequent experiments, the amount of tacrolimus was set equimolar to the alkyne on beads. The exchange of solvent to anhydrous dichloromethane resulted in 11% tacrolimus conversion. And finally, the increase of the catalyst amount from 1 mol% to 10 mol% led to an increase in tacrolimus conversion up to 37%. Beads prepared with 10 mol% catalyst were then subjected to a full-dose competition pulldown experiment in K562 lysate (Fig. 15b). With a relative IBAQ abundance of 0.2%, FKBP12 enrichment was more than 10 times lower compared to the photo-crosslinking immobilizations (from 3% up to 40%, Fig. 15c, 13c). Low target enrichment resulted in lower quality of a dose-response curve of the competition experiment. The number of peptides (3) and spectral counts (7) for FKBP12, consequently, were significantly lower than in the photo-crosslinking approach (unique peptides of at least 10 and MS/MS counts exceeding 200, Fig. 15d). One may speculate that gold-catalyzed alkoxylation produced a partially inactive tacrolimus analog on beads that was unable to engage FKBP12 in lysate fully.

8.2.3 Gold(I)-catalyzed [2+2] cycloaddition on alkyne beads.

Another gold(I)-catalyzed alkyne transformation is the [2+2] cycloaddition with alkenes. This reaction results in the formation of cyclobutenes. Gold(I)-catalysis was primarily used for intramolecular cyclization; however, it has been fairly recently employed in a range of intermolecular reactions mediated by bulky ligands on cationic gold (I) complexes.[104] Based on the results reported by Homs et al.,[104] the tBuXPhos Au(MeCN)SbF6 catalyst was assayed in this study for the immobilization of tacrolimus on alkyne beads via its alkene moieties (Fig. 16).
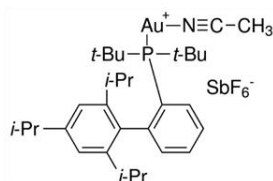
Figure 16. Chemical structure of tBuXPhos Au(MeCN)SbF$_6$ applied in the [2+2] cycloaddition reaction.

Reported by Homs et al. reaction conditions were utilized for the initial pilot immobilization experiment, where 2 µmol/mL alkyne-loaded beads were reacted with 2-fold molar excess of tacrolimus in the presence of 3 mol% gold(I)-catalyst in anhydrous acetonitrile at 30 °C overnight, resulting in 75% tacrolimus conversion. Obtained affinity matrix was subjected to a two-dose competition experiment with 10 and 30 µM of free tacrolimus. Despite the rather high conversion of tacrolimus, FKBP12 was found to be fairly poorly enriched (rIBAQ 0.2%), which was further reflected in the low number of MS/MS counts and peptide numbers (4 and 3, respectively). Nevertheless, target enrichment was clearly decreased upon competition with 10 µM and 30 µM tacrolimus (Table 1).

FKBP12	DMSO	10 µM competition	30 µM competition
-log2 FC LFQ intensity	-	3.3	5.1
Unique peptides	3	2	1
MS/MS counts	4	1	0

Table 1. Results from two-dose competition assay with tacrolimus-beads from gold(I)-catalyzed cycloaddition.

The reported above experimental conditions could not be reproduced in the subsequent experiment, where tacrolimus showed no more than 15% conversion. Furthermore, the exchange of acetonitrile to dichloromethane to improve reaction efficiency did not improve the compound conversion. Finally, the gradual increase of gold(I)-catalyst from 3 mol% to 10% and up to a maximum 20 mol% under the same conditions did not produce an affinity matrix that could demonstrate better FKBP12 enrichment (rIBAQ FKBP12 < 0.2 %, data not shown). Hence, this particular immobilization strategy proves to be rather irreproducible and requires further optimization. The use of another catalyst may improve the overall reproducibility and compound conversion. Therefore, the gold-catalyzed cycloaddition approach was not further followed up in this work, and ruthenium catalysis was assayed instead.

8.2.4 Enyne metathesis on alkyne beads.

Another immobilization strategy that utilizes alkyne-functionalized beads - ruthenium-catalyzed enyne metathesis - was evaluated in this work as a potential alternative to gold-catalyzed cycloaddition. Enyne metathesis yields 1,3-dienes, can be performed both intra- and intermolecularly and is mainly utilized for the synthesis of olefins.[105] Two different air-stable Hoveyda-type ruthenium-catalysts were evaluated in this study: Nitro-Grela and water-soluble AquaMet (Fig. 17a,b).[106,107]

In all experiments, 10 mol% of the catalyst and equimolar compound-to-linker amounts were applied. The initial experiments were performed in anhydrous acetonitrile at 30 °C overnight.

Here, no conversion of tacrolimus was observed with either catalyst, which did not improve upon temperature increase to 50 °C. The substitution of acetonitrile to water:methanol mixture (2:1) for AquaMet-catalyzed immobilization led to the increase of tacrolimus conversion to 39%. Pure water could potentially increase reaction efficiency even more; however, it could not be used here due to the insolubility of tacrolimus. For Nitro-Grela catalysis, the reaction solvent was switched from acetonitrile to anhydrous dichloromethane. Due to technical issues, there are no LC-MS coupling controls for this particular experiment. Here tacrolimus immobilization (not quantitative) was confirmed via pulldown experiments (Fig. 17c). The two matrices prepared with either catalyst could efficiently enrich FKBP12 from cell lysate. In the case of the AquaMet-experiment, FKBP12 comprised more than one-third of the sub-proteome enriched by the matrix (38%, Fig. 17c). An even higher enrichment was achieved with Nitro-Grela, resulting in a relative iBAQ value of 60%. Effective enrichment was also concluded from the relatively high MS/MS counts of 77 (AquaMet) and 94 (Nitro-Grela). Interestingly, both affinity matrices derived from enyne metathesis showed relatively low unspecific background binding compared to other immobilization strategies, as they overall enriched fewer proteins (Fig. 17c,d). Tacrolimus-matrix from the AquaMet catalyzed metathesis was subjected to a full dose competition assay, where a clear dose-dependent reduction of FKBP12 was observed (Fig. 17e). Tacrolimus-beads from Nitro-Grela catalyzed immobilization were also probed in a competition assay; the results are summarized in Table 2.
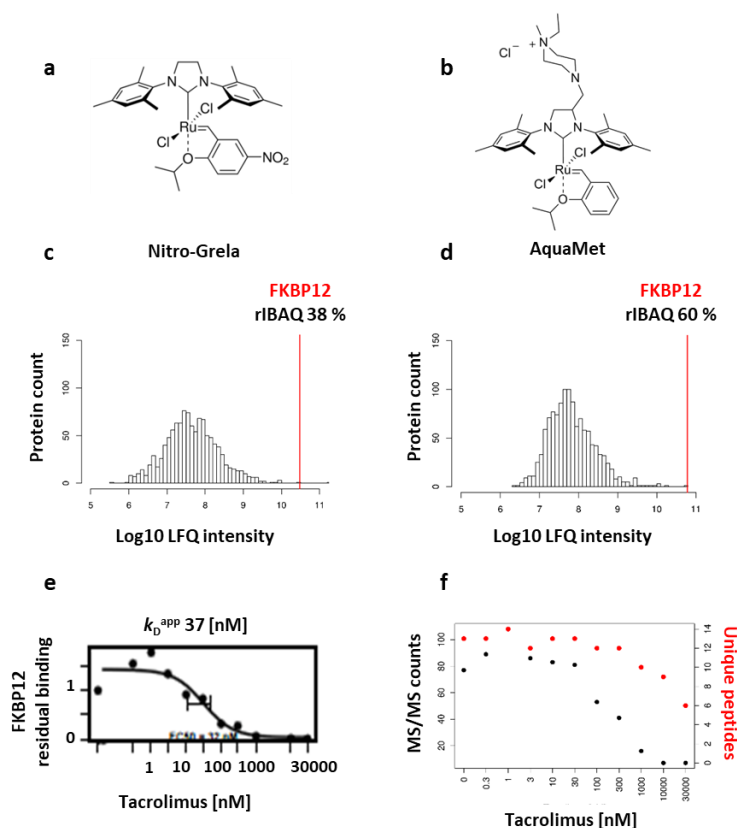
Figure 17. Enyne metathesis on alkyne beads with tacrolimus. a,b) Chemical structures of Nitro-Grela and AquaMet. Both ruthenium-catalysts were tested in the enyne metathesis reaction. c,d) FKBP12 enrichment out of K562 lysate. e) FKBP12 dose response curve from full-dose competition experiment with tacrolimus beads prepared though AquaMet-mediated catalysis. f) Number of MS-identified unique peptides and MS2 spectra in each competition condition (beads from AquaMet-mediated catalysis).

	AquaMet			Nitro-Grela		
FKBP12	**DMSO**	**10 µM competition**	**30 µM competition**	**DMSO**	**10 µM competition**	**30 µM competition**
-log2 FC LFQ intensity	-	6.3	7.9	-	6	8
Unique peptides	9	9	6	13	6	5
MS/MS counts	77	7	7	98	10	4

Table 2. Results from two-dose competition assays with tacrolimus-beads from both enyne metatheses.

Both affinity matrices showed very efficient competition for FKBP12 binding with free tacrolimus. The competition at 30 µM was very comparable for both experiments (-log2FC LFQi of 8); however, overall the affinity matrix obtained with Nitro-Grela outperformed the AquaMet one. These results indicate that both catalysts can be employed to prepare tailored affinity matrices for NP that feature terminal alkenes in their chemical structures.

8.2.5 Mitsunobu reaction on nucleophile beads.

The Mitsunobu reaction is a substitution of primary and secondary alcohols with nucleophiles, mediated by a redox combination of a triphosphine and a dialkyl azodicarboxylate. It is important to note that in most cases, chiral secondary alcohols undergo a complete inversion of stereochemistry in the course of the reaction, which generally enables the synthesis of molecules with inverted stereogenic centers in high enantiomeric purity.[108] Due to its scope, stereospecificity, and mild reaction conditions, it is frequently employed in the NP synthesis.[109] Herein, this reaction was applied to immobilize tacrolimus via its hydroxy groups to nucleophile-derivatized, both -OH and -NH$_2$, beads. The common Mitsunobu reaction reagents, triphenylphosphine (TPP) and azodicarboxylic acid dipiperidide (ADDP) were used. The nucleophile-bearing beads were prepared by reacting the NHS-ester sepharose beads with ethanolamine or ethylenediamine, respectively.

	-OH beads			-NH$_2$ beads		
FKBP12	**DMSO**	**10 µM competition**	**30 µM competition**	**DMSO**	**10 µM competition**	**30 µM competition**
-log2 FC LFQ intensity	-	3	3	-	4.5	7.5
Unique peptides	7	2	7	8	2	2
MS/MS counts	11	3	3	19	4	4

Table 3. Results from two-dose competition assay with tacrolimus-beads from Mitsunobu reaction.

Mitsunobu reaction was performed for both amine- and alcohol- 2 µmol/mL functionalized beads with the molar equivalent of tacrolimus at 30 °C overnight in dichloromethane. Obtained affinity matrices were utilized in two-dose competition pulldowns with tacrolimus. Both matrices enriched FKBP12 much better than previously discussed gold-catalyzed immobilizations, with rIBAQ values of 2.4% for hydroxy-functionalized beads and 7.2% for the amine-derivatized matrix.

Tacrolimus immobilized on NH_2-matrix, consequently, performed better in a competition experiment (Table 3). These results could be reproduced in the subsequent experiment indicating that 1) the immobilizations on $-NH_2$ and -OH potentially produced tacrolimus-matrices with different insertion regioselectivity, which would explain the differences in the efficiency of the competition; 2) the Mitsunobu reaction can, indeed, be used for the immobilization of NPs.

8.3 All tested immobilization chemistries allowed tacrolimus coupling on beads.

These initial pilot experiments evaluated five different immobilization strategies on tacrolimus (a short summary of the results can be found in Figure 18). All immobilization approaches allowed a certain degree of tacrolimus coupling to the beads while obtained matrices enriched very different amounts of FKBP12 out of K562 lysate (Fig. 18a). The latter can either be explained by 1) overall poor tacrolimus conversion or 2) the possibility that a particular Immobilization afforded tacrolimus coupling via an essential for target binding moiety. In order to fully define whether the tested chemistries are generally applicable for the immobilization of different NPs, these experiments have to be repeated on a range of compounds of different chemical structures. UV-induced photo-immobilization was selected for this study for its potential to be the most promiscuous, yet, surprisingly, it appeared to be the most reproducible among all other chemistries evaluated. Thus, the scope of this approach was evaluated on the broader range of NPs, which will be discussed in the following chapters.
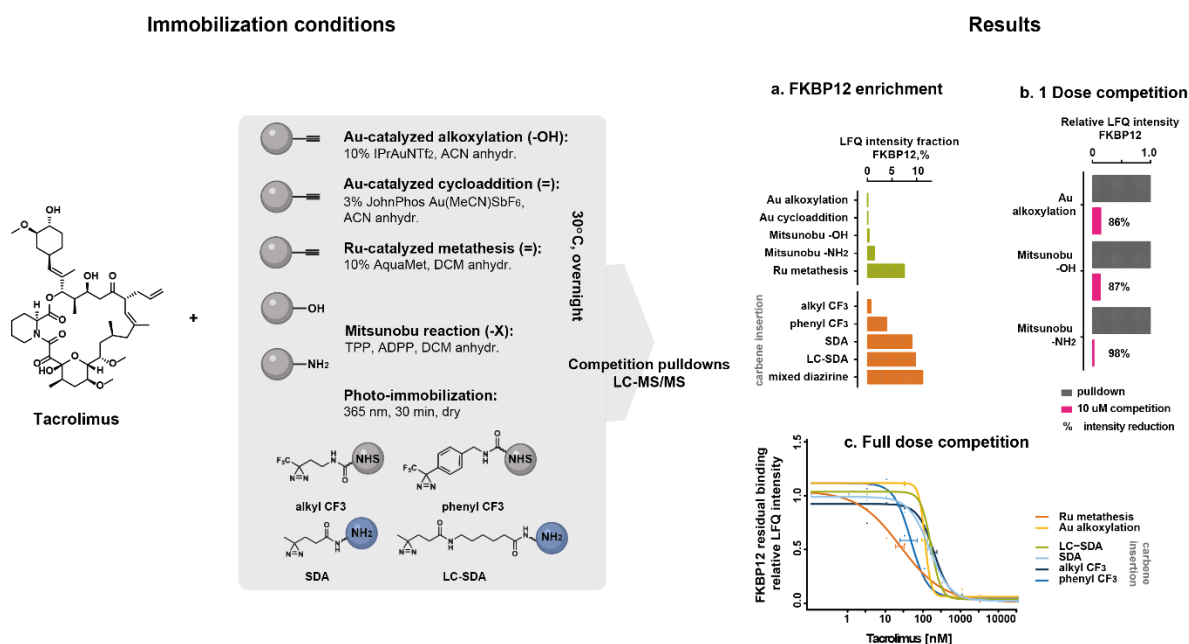
Figure 18. Pilot benchmarking of five chemistries for immobilization of tacrolimus. a) FKBP12 enrichment by tacrolimus-loaded affinity matrices from different immobilization chemistries pilot experiments as a fraction of total LFQ intensity in pulldown. b) Reduction of FKBP12 affinity in a single dose competition pulldown. c) Dose response curves of FKBP12 for a selection of affinity matrices.

An interesting take on using different chemistries for the immobilizations via multiple functional groups would be to use them for establishing SAR and determining functional groups especially vital for target engagement. For instance, different chemistries would produce affinity matrices, where a compound is linked to the beads via different moieties. Those individual matrices would then be subjected to a full dose competition experiment and obtained targets compared. In case

when individual competitions produce different sets of targets or affinities of observed binding interactions for the same targets would largely differ, it could be possible to decipher which functional groups on the molecule were responsible or essential for establishing affine interactions with a given target or multiple targets. Conversely, to maximize the chances of obtaining an affinity matrix with preserved bioactivity of a given immobilized NP, it would be rational to employ a mixture of all affinity matrices prepared using different immobilization chemistries for competition profiling.

8.4 Immobilization may affect the ability to engage protein targets

Comparison between promiscuous photo-crosslinking and conventional selective immobilization via a single defined functional group is more comprehensive if the analyzed compound is known to bind more than one target (unlike tacrolimus). Staurosporine is a well-characterized natural broad-selective kinase inhibitor. It possesses a methylamino group that allows for direct functionalization on beads (Fig. 19a). A matrix obtained by the selective immobilization was included in the first version of Kinobeads as one of the baits to enrich kinases out of cell lysates. The selectivity profiling of staurosporine with this version of Kinobeads identified 66 kinase targets.[110] The compound was later subjected to target profiling with CETSA-MS, where 51 kinases showed thermal shifts.[111]
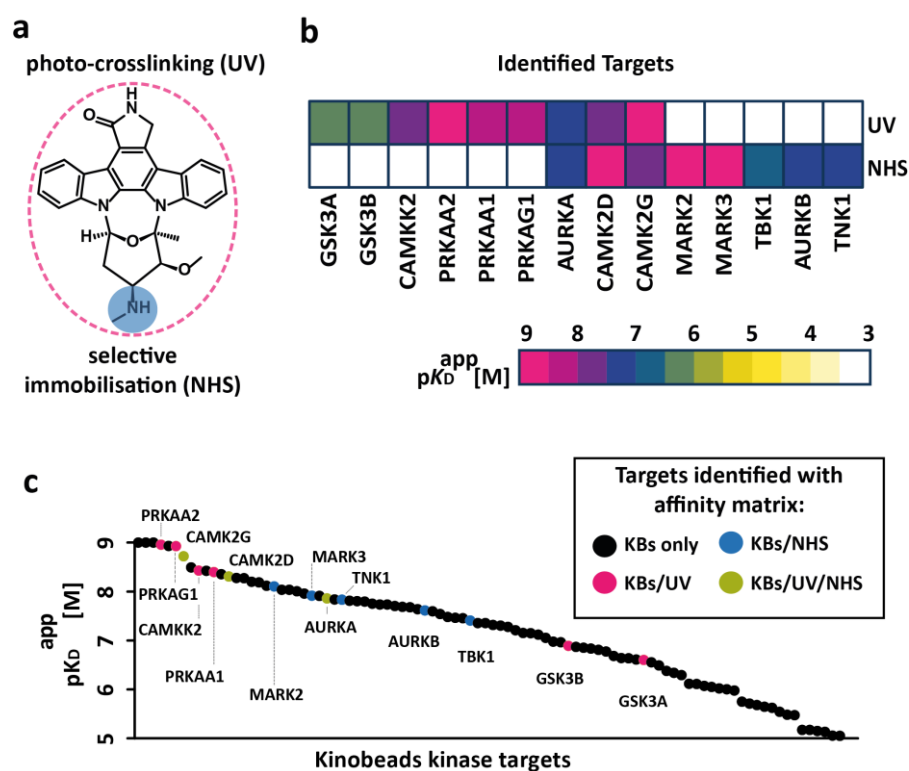
Figure 19. Comparison of target proteins identified by selective secondary amine vs photo- immobilization of staurosporine or Kinobeads a) Chemical structure of staurosporine with the secondary amine used for immobilization highlighted in blue (selective immobilization). The pink dashed circle depicts that promiscuous UV-induced immobilization may occur anywhere in the compound. b) Heatmap depicting apparent binding constants of protein targets obtained by photo-crosslinked or selectively immobilized staurosporine. c) Apparent binding constants of kinase targets of staurosporine obtained by Kinobeads (KBs) profiling only (black dots), selectively immobilized staurosporine and Kinobeads (NHS, blue dots), UV-immobilized staurosporine and Kinobeads (pink) or by each of the three approaches (green dots).

In this work, staurosporine was immobilized on beads following photo-induced carbene insertion and selective amine acylation approaches. The two obtained affinity matrices were subjected to a full dose competition pulldown in SW-620 lysate (Fig. 19a,b). Both matrices enriched a rather surprisingly small number of kinases, and only three targets were identified with similar affinities in both assays (AURKA, CAMK2D, and CAMK2G). Five kinases (AURKB, MARK2, MARK3, TBK1, and TNK1) were only found using NHS-immobilized staurosporine, and CAMKK2, GSK3A, GSK3B, PRKAA1, PRKAA2, and PRKAG1 were exclusively observed using photo-crosslinking to beads. Noticeably, these two tailored affinity matrices underperformed compared to the Kinobeads (this version of Kinobeads does not include immobilized staurosporine, ref. 57) profiling in this study that allowed the identification of 94 kinase targets with affinity below 10 μM, including all the above targets from tailored matrices pulldowns (Fig. 19c).[57] Interestingly, both tailored matrices enriched staurosporine targets from the entire range of affinities (according to Kinobeads profiling). None of the two matrices showed preference for binding of more or less affine targets. Based on crystal structures of multiple Staurosporine-kinase complexes, it is reasonable to assume that immobilization of the molecule through its methylamino group can result in a substantial change in protein binding: acylation annihilates the usually observed hydrogen bond between the charged amine and the protein pocket residues. The carbene insertion also largely affects its ability to engage targets but evidently produces linked analog(s) of staurosporine with different kinase preferences. Despite multiplying the chances of producing bioactive linked analogs, some molecules like staurosporine remain reluctant to modifications, which naturally limits target deconvolution assays with this method and affinity probes in general.
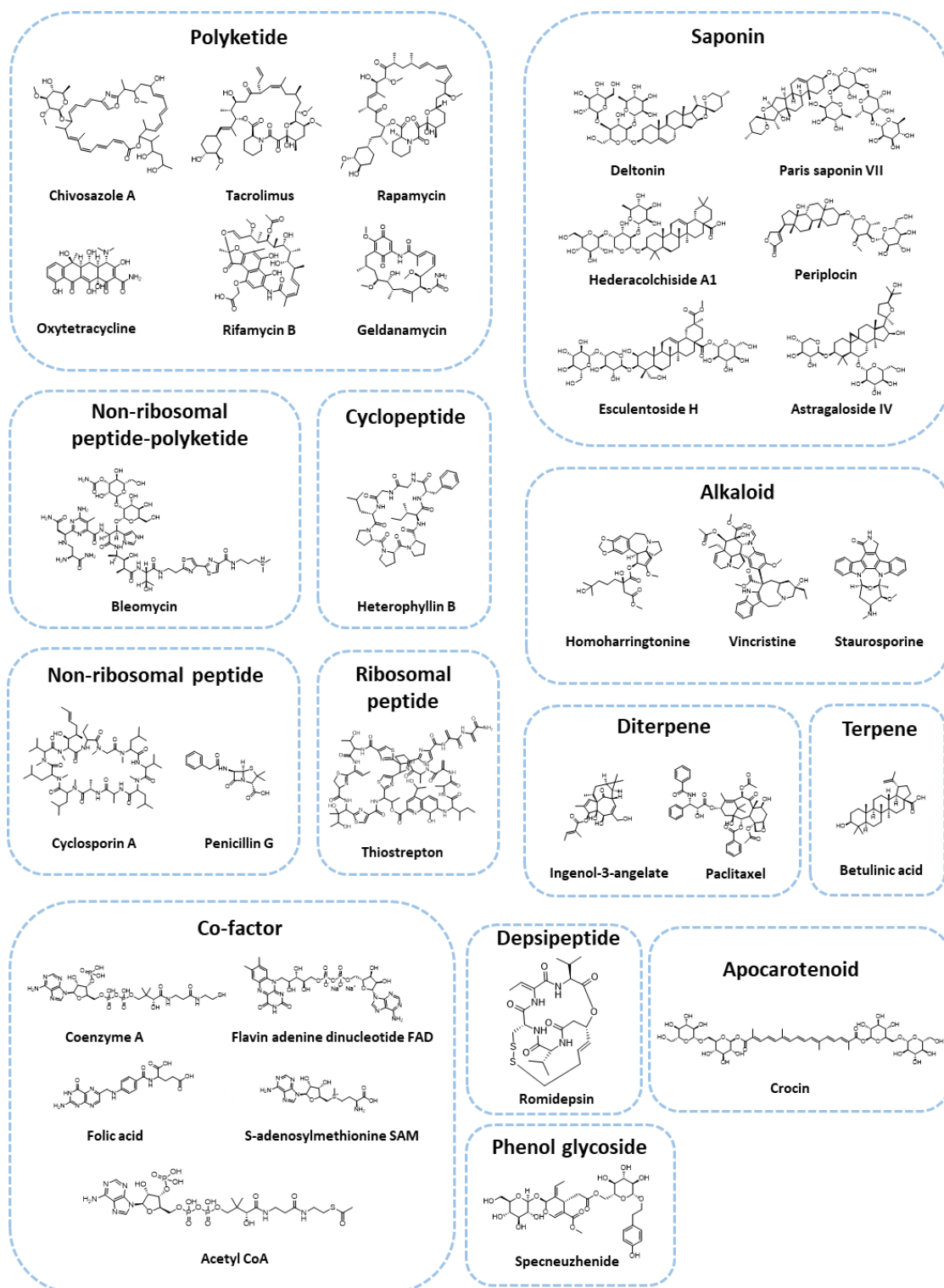
Figure 20. Chemical structures of NPs used in this study

8.5 Promiscuous photo-crosslinking immobilization can serve as quick tool to probe for protein binders of natural molecules

To probe the scope of the target deconvolution assay with photo-crosslinked matrices, 29 other commercially available natural molecules of different biosynthetic origins were tested (31 together with tacrolimus and staurosporine discussed earlier). Among those were three common antibiotics, five co-factors and other compounds that have been reported to possess anticancer activity in at least one assay (Fig. 20). Reassuringly, 23 compounds (80%) could indeed be immobilized on four-diazirine-linker beads, two compounds failed to react with diazirine beads, while further four compounds do not possess a UV chromophore or did not generate the UV or MS signals necessary for quantifying their conversion (Table 4). The coupling efficiency range spanned between 15% (coenzyme a) to quantitative immobilization (geldanamycin) (Fig. 21a). Obtained affinity matrices were utilized in a full-dose competition profiling followed by LC-MS/MS readout. Photo-crosslinked affinity matrices bound between several hundred to several thousand proteins in SW-620 lysate (Fig. 21b).
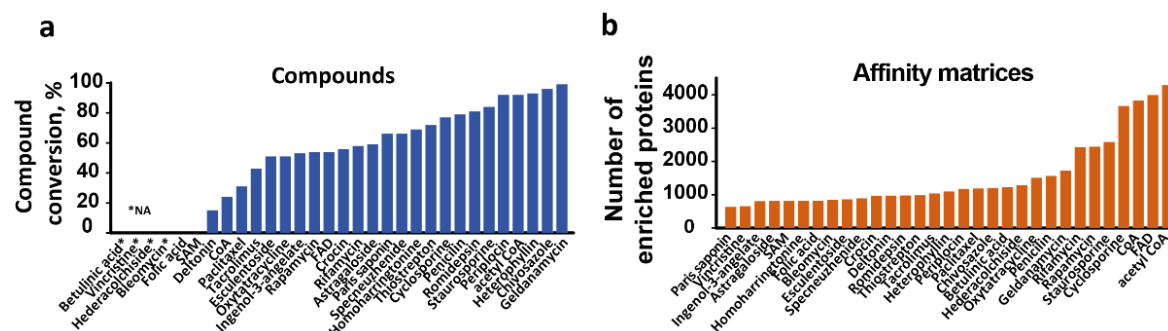
Figure 21. Broad assessment of photo-immobilization for target deconvolution of natural molecules. a) Immobilization efficiency upon UV irradiation for 31 natural compounds. b) Number of proteins bound by each affinity matrix. c) Chemical structures of all molecules employed in this study.

In this evaluation, protein targets were recapitulated for previously discussed tacrolimus, staurosporine, and seven other compounds (Table 4). For compounds without described protein binders, this assay also did not uncover any. There may be many reasons for not capturing targets in addition to the ones outlined above. For instance, the protein may not be expressed in the cells we used here; they may not retain their natural fold under lysis conditions, or the interaction may be too transient to capture. In addition, some NPs may be pro-drugs inactive in lysates, while others may not bind proteins as their normal mode of action. Bleomycin, for example, covalently binds to nucleic acids and catalyzes their oxidation which, in turn, leads to the scission of DNA and cell death. Six of the 31 compounds tested here are saponins. These form complexes with cholesterol and create pores in the cell membrane to kill cells. Such compounds would not be expected to score in our assay because of their particular mode of action. Among the successful cases, the known binders of the immunosuppressant cyclosporin A were recapitulated in human cells (Fig. 22a).[112] Similarly, the targets of the macrocyclic polyketide antibiotics rifamycin B and geldanamycin (Fig. 22 b,c) were identified in bacterial *E. coli* and human colorectal cancer SW-620 lysates, respectively.[113–115] Besides the cognate target, DNA-directed RNA polymerase subunit beta (rpoB), the competition experiment with increasing concentrations of free rifamycin B also showed a superimposing dose-response behavior for the subunit alpha (rpoA), likely co-enriched and co-competed with rpoB. Geldanamycin is a heat-shock protein 90 (HSP90) inhibitor. With the competition pulldown, main HSP90 isoforms were identified as Geldanamycin targets, namely the cytosolic HSP90AA1 and HSP90AB1, constituting 3% and 13%, respectively, of total IBAQ in the

DMSO control pulldown. Together, these results indicate that photo-crosslinking immobilization deserves consideration in future antibiotics discovery: it is easily deployable in human and bacterial lysates and could, therefore, help uncover targets of novel antibiotics.

No	Name	Class	Conversion, %	Identified targets/ interactors in this study	known binders/ MoA
1	Betulinic acid	Terpene	NA		triggers the mitochondrial pathway of apoptosis in cancer cells
2	Vincristine	Alkaloid	NA		inhibition of the microtubule structures within the cell
3	Bleomycin	Non-ribosomal peptide/polyketide	NA		oxidation of nucleic acids
4	Hederacolchiside A1	Saponin	NA		cholesterol
5	Geldanamycin	Polyketide	99	x	HSP90
6	Chivosazole F	Polyketide	96		Actin
7	Heterophyllin B	Cyclic peptide	93		
8	Staurosporine	Alkaloid	92	x	kinases
9	Periplocin	Saponin	84		cholesterol
10	Cyclosporin A	Non-ribosomal peptide	79	x	cyclophilins
11	Homoharringtonine	Alkaloid	72		ribosome, inhibition of protein synthesis
12	Thiostrepton	Ribosomal peptide	69		ribosome
13	Paris saponin VII	Saponin	66		cholesterol
14	Specneuzhenide	Phenol glycoside	66		
15	Astragaloside	Saponin	59		cholesterol
16	Crocin	Apocarotenoid	56		
17	Ingenol-3-angelate	Diterpene	54		induction of necrosis followed by PKC-driven immune response
18	Esculentoside H	Saponin	51		cholesterol
19	Rapamycin	Polyketide	51	x	FKBP12
20	Tacrolimus	Polyketide	43	x	FKBP12
21	Paclitaxel	Diterpene	31		tubulin
22	Deltonin	Saponin	24		cholesterol
23	Romidepsin	Depsipeptide	81		HDACs
24	Coenzyme A	co-factor	15	x	acyltransferases
25	Acetyl-CoA	co-factor	92	x	acyltransferases
26	FAD	co-factor	54	x	flavoenzymes
27	SAM	co-factor	0		SAM-dependent methyltransferases
28	Folic acid	co-factor	0		folate binding proteins
29	Oxytetracycline	Polyketide	53		30S ribosomal subunit
30	Rifamycin B	Polyketide	58	x	RNA polymerases
31	Penicilin G	Non-ribosomal peptide	77		transpeptidases

Table 4. Conversion on beads for all the compounds in this study, including known MoA information. Betulinic acid, vincristine, bleomycin and hederacolchiside A1 did not produce LC-MS signals from which the conversion could be quantified (NA entries).
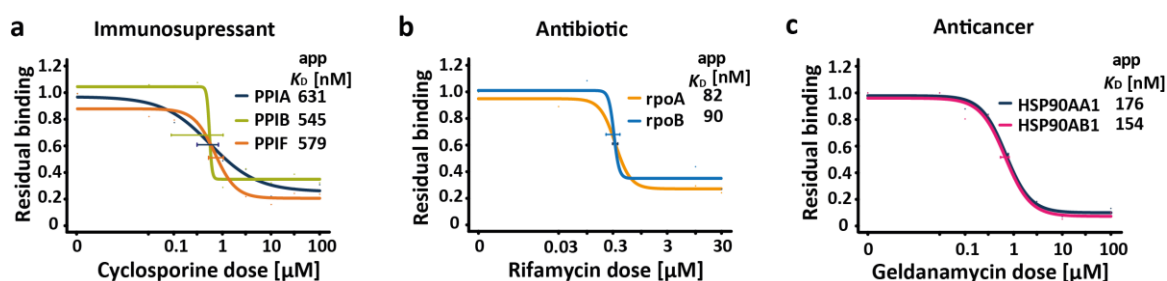
Figure 22. Successful target deconvolution examples c) Dose response curves for cyclosporine A binders PPIA, PPIF and PPIB. d) Same as panel c but for Rifamycin B. e) Same as panel c but for Geldanamycin.

8.6 Photo-immobilization allows for profiling of protein cofactor interactomes

8.6.1 UV immobilization of CoA and acetyl-CoA identifies interactions with acyl-transferases and hotdog domain proteins

Besides target deconvolution for drug discovery, another possible exploitation of the unselective immobilization strategy consists of delineating co-factor interactomes. Such an objective has been the focus of inspiring publications, where the affinity probes have been prepared from functionalized chemically synthesized co-factors analogs to explore, e.g., NAD and acyl-CoA interactomes.[116,117] For instance, Levy et al. created a lysine-CoA matrix to investigate the specificity of protein binders to different acyl-CoAs.[118] With this approach termed CATNIP (CoA/Acetyl-transferase interaction profiling), multiple CoA binders that compete with distinct acyl-CoA derivatives were identified. More specifically, a Lys-CoA probe was set to compete with multiple CoA metabolites at 30 μM dose, which led to the identification of 166 proteins that showed at least a two-fold reduction of binding to free coenzyme A. From those, 126 proteins (out of 3660 total enriched proteins) were also quantified in a pulldown with the coenzyme A affinity matrix prepared following the novel multi-linker immobilization procedure; however only seven proteins displayed the typical sigmoidal competition curves, unequivocally identifying them as CoA binders with affinities below 100 μM. Furthermore, the competition assay here identified 11 other interactors of CoA that have not been highlighted in the CATNIP study.

In order to compare CoA binding selectivity to the one of acetyl-CoA, the latter was also immobilized on diazirines beads via light-induced carbene insertion. In total, 21 interactors were identified in this assay with affinities below 100 μM (Fig. 23a). CoA and acetyl-CoA shared 9 protein binders, most of which are known acyl-CoA interactors such as the acyltransferases NAT14, NAA40, NAA15, NAA10, NAA16, NAA50, SAT1 and SAT2 (Fig. 23a,d-e). Interestingly, several acyltransferases showed selective binding to one of the cofactors. For example, NAA25 showed typical dose-dependent intensity reduction exclusively in competition with acetyl-CoA, whereas the CoA matrix failed to stably enrich it from SW-620 cell lysate (Fig. 23b-c). In order to rule out whether the immobilization of the two acyl-CoAs followed different regioselectivity, in turn affecting which proteins were competing for binding, an additional competition experiment would have to be performed, where each of the cofactors would be set to compete against a mixed cofactor matrix.

Another group of specific binders identified in the competition with both CoAs contains a hotdog domain, found in several human thioesterases, which links them to acetyl-CoA biochemistry. CoA and acetyl-CoA matrices enriched in total seven out of twelve human HotDog domain-containing proteins reviewed by UniProt (i.e. THEM4, ACOT13, ACOT7, ACOT8, ACOT9, HSD17B4 and THEM6). It is especially interesting to note the competition of hotdog domain-containing THEM4 and ACOT7 by coenzyme A. THEM4/CTMP has been reported as an oncogene in breast cancer, while ACOT7 appears to play an important role in inflammatory disease and breast and lung cancer development.[119–121] With such an easily obtained affinity matrix that specifically binds the HotDog domain, small molecules could be screened and potentially develop into HotDog binders and inhibitors. The chemoproteomic assay would then provide both the affinity and selectivity of the acyl-CoA mimics among a panel of acyl-CoA binding proteins.
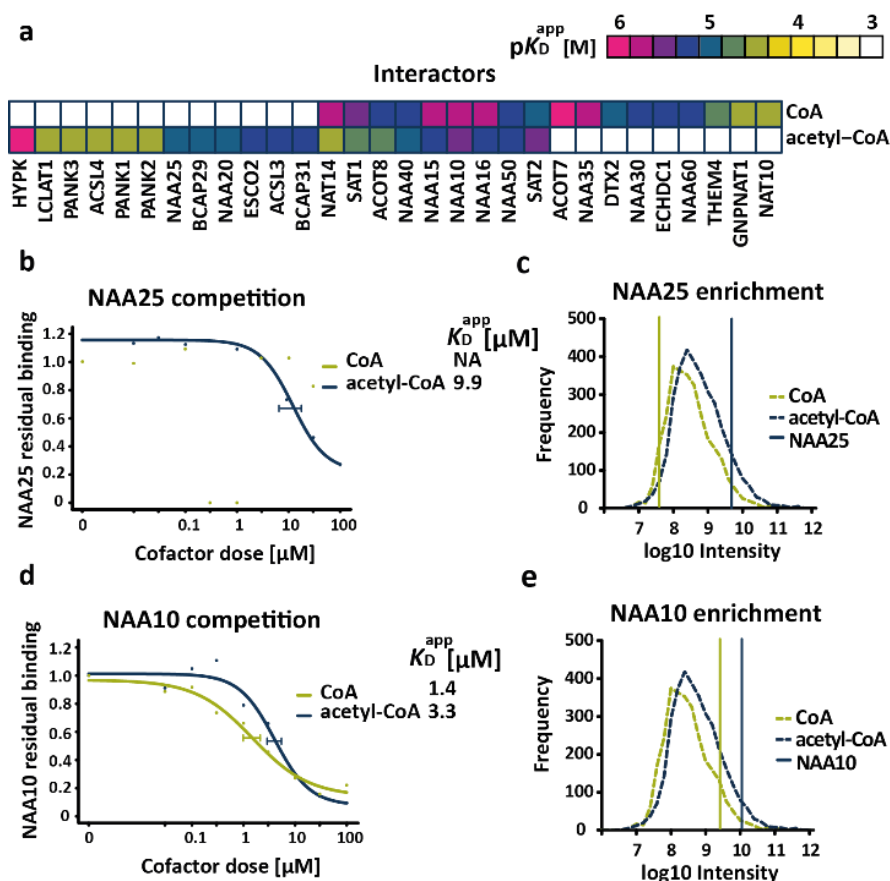
Figure 23. Binders of the protein cofactors CoA and acetyl-CoA. a) Heatmap of apparent interaction constants of proteins binders of CoA and acetyl-CoA. b) Residual binding of NAA25 to CoA and acetyl-CoA beads in response to increasing concentrations of free CoA and acetyl-CoA respectively. c) Intensity distribution of all proteins bound to CoA or acetyl-CoA beads (dotted lines). Solid lines mark the position of NAA25 in these distributions. d) Same as panel b but for NAA10. e) Same as panel c but for NAA10.

8.6.2 The cofactor flavin adenine dinucleotide (FAD) potently interacts with dozens of RNA-binding proteins

The interactome of flavin adenine dinucleotide (FAD) was equally explored with the help of photo-immobilization. Surprisingly, the obtained FAD matrix pulled down more than 600 RNA-binding proteins, from which more than 40 displayed typical dose-dependent intensity reduction in the competition assay with free FAD. Go enrichment analysis also confirmed that RNA binding term was greatly enriched in the entire pool of bead-bound proteins as well as in the competed proteins fraction (Fig. 24 a,b). Among all bead-bound proteins were 17 known FAD binding flavoproteins[122] (GCDH, ACADVL, KDM1A, ETFDH, ETFA, SDHA, PPOX, CYB5R1, SQLE, POR, CYB5R3, AIFM1, ACOX1, ILVBL, AGPS, ACOX3, DLD), but none of them were competed by free FAD. Flavoenzymes use flavins as prosthetic groups and are known to bind the flavin cofactor very tightly, some even covalently as part of the holo-enzyme[123]. Particularly the covalent binders would not be expected to bind a FAD matrix and may indicate that the proteins mentioned above are likely poor binders of FAD with affinities weaker than the highest competition dose.
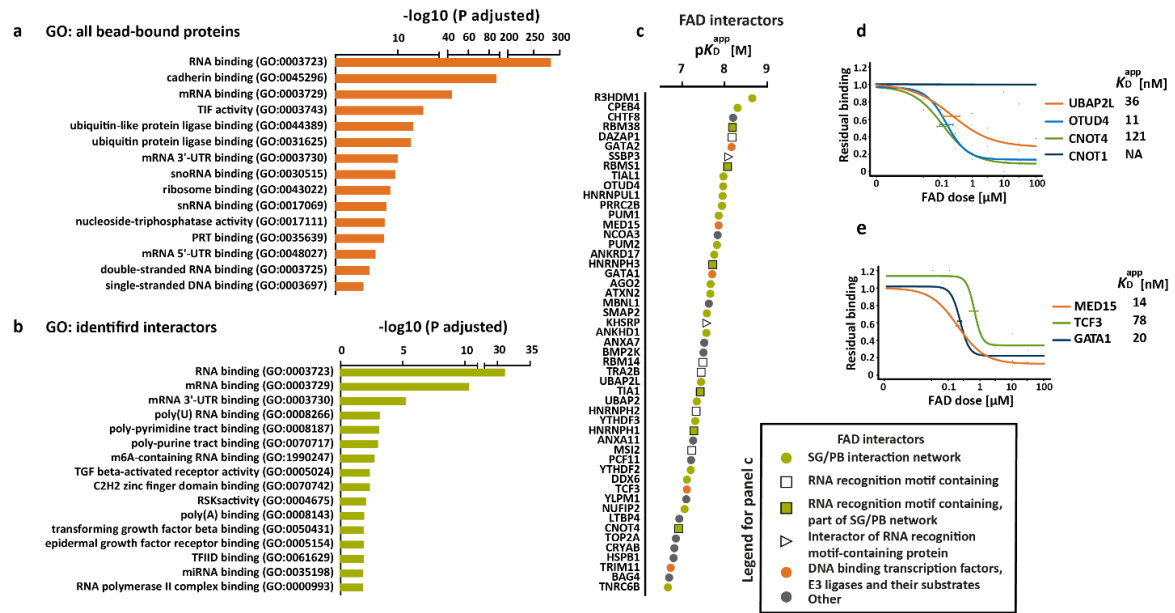
Figure 24. FAD interactors a) GO enrichment for all proteins bound by the FAD-loaded beads (first 15 terms, unique peptides ≥3, -log10(adjusted P value) ≥2.5). b) GO enrichment for proteins that were competed with free FAD in a full-dose competition pulldown with FAD matrix (same filter as for panel a). c) Affinity ranking of all identified direct or indirect FAD interactors in full dose competition pulldown assay. d) Dose response curves for identified FAD binders UBAP2L, OTUD4, CNOT4. CNOT1 was bound by the matrix but not competed up to 100 μM. e) Same as panel d but for MED15, TCF3 and GATA1.

More than half of potential FAD interactors found in the competition experiments were previously reported as important for stress granule (SG) formation. Youn et al.[124] published a BioID profiling study of mRNA-associated stress granules (SGs), where several proteins were reported as key regulators of stress granules formation (OTUD4, UBAP2L, CSDE1, and PRRC2C) or proteins recruited to SGs in response to stress (G3BP1, PABPC1, and eIF4A1). All of the above were identified in pulldown with photo-crosslinked FAD matrix in this work; however, only two proteins showed potent competition with free FAD, namely UBAP2L and OTUD4 (Fig. 24 c,d). Several of the other proteins involved in SGs that were identified as FAD interactors in this work are the members and the interactors of the CCR4-NOT deadenylases complex, namely E3 ubiquitin ligase CNOT4, GW182, TNRC6B, and the endonuclease Argonaute AGO2.[125] More than ten found interactors contain canonical RNA-binding domains, such as RRM, PUM-HD or KH, or are known to be involved in mRNA metabolism or regulation of transcription (HNRNPH1, MBNL1, CPEB4, PUM1, etc. Fig. 24c)[126–128]. While other RNA-binding domain containing interactors have seemingly unrelated annotated functions (ATXN2, ANKRD17). Several further proteins are rather poorly functionally annotated (R3HDM1, SMAP2, PRRC2B), and the results obtained here may potentially place them into the functional context of mRNA metabolism, stress granule formation, or regulation of transcription. The DNA-binding transcription factors GATA1, GATA2, TCF3, the E3 ligase TRIM11, and a member of the mediator complex MED15 also showed potent direct or indirect FAD binding. It has been shown that TRIM11 negatively regulates MED15 stability and reduces its transcriptional activity, which, again, links FAD to ubiquitin-regulated cellular processes.[129]

Among the few known consensus sequences for RNA-binding proteins, the PUF motif 5'-UGUANAUA-3' was selected to investigate whether FAD is binding the Pumilio homologs PUM1 and PUM2 in their RNA-binding pocket, as both these proteins showed dose-response competition

in the FAD versus FAD-beads experiment (Fig. 25a).[130,131] The custom-synthesized oligomer was immobilized on Sepharose beads with 10 nM/mL RNA loading density and subjected to competition profiling vs. free FAD and free oligomer (Fig. 25c). It is important to note that RNA-loaded beads were able to efficiently enriched PUM homologs out of K562:placenta lysate, in fact, much better compared to a range of other matrices used in this work (Fig. 25b,c). The results of competition profiling indicate the binding of FAD to PUM homologs is most likely allosteric as it does not influence the ability of PUMs to engage the RNA (Fig. 25d,e).
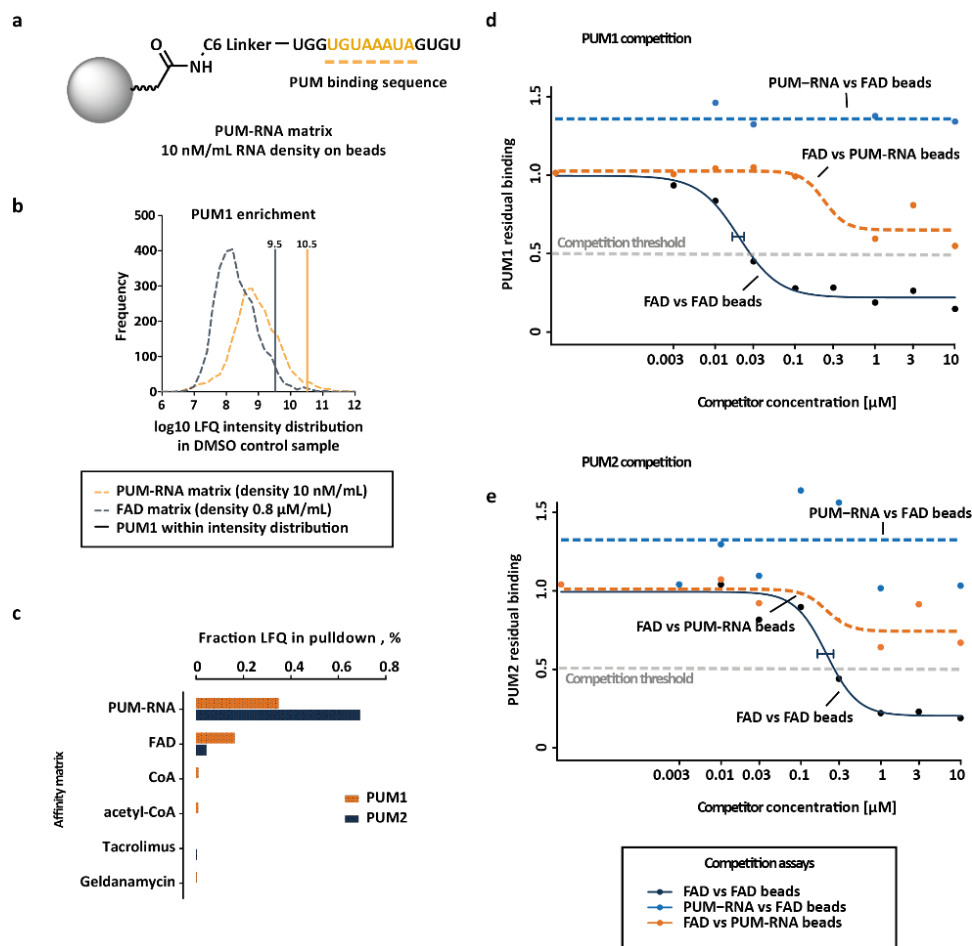
Figure 25. Pumilio homologs enrichment and competition with RNA-probe in K562:placenta (1:1) lysate. a) beads were loaded with 10 nM/mL of RNA probe based on PUM binding sequence UGUAAAUA[132] (see methods for details) b) PUM1 enrichment by FAD and PUM-RNA matrices. PUM-RNA matrix was loaded with 80 times less probe compared to FAD beads, but still efficiently enriched PUM1 out of K562:placenta (1:1) lysate; c) comparison of PUM1 and PUM2 intensity fraction in pulldowns by multiple affinity matrices; d) PUM1 curves in three competition assays. Dose-dependent competition (more than 50% intensity decrease) only observed in the assay with FAD beads vs free FAD; e) same as panel d but for PUM2.

It is important to follow up on these findings and explain the nature of the identified FAD interactions. For example, it would be essential to show which of the above proteins are direct binders of FAD, which are being pulled down as members of stable complexes, to which part of the FAD molecule the proteins bind, and how such binding may be involved in regulating processes such as mRNA deadenylation or DNA binding.

8.7 **Conclusions of Part 1.**

In conclusion, several chemistries for the diversity-oriented immobilization for target deconvolution of NPs were successfully evaluated in this work. The use of tailored affinity probes always comes hand in hand with the danger of losing compound bioactivity upon immobilization, as demonstrated for staurosporine. However, the approach proposed in this work should maximize the probability of retaining the bioactivity of the functionalized compound. The main advantage here is the use of unmodified NP and, therefore, the circumvention of extensive chemical synthesis of affinity probes. Photo-immobilization, in particular, holds a promise of being useful for agnostic target deconvolution as it proved to be very reproducible in pilot experiments with model compound tacrolimus and later enabled the preparation of affinity matrices for 30 other NPs for which the bioactivity has not yet been confirmed in the cells used. The simplicity and generality of the experimental procedure as well as the quantitative characteristic of the proteomic readout may make it the first "go-to method" to probe natural molecules for protein binding. Despite the seemingly low overall target deconvolution success rate across all 31 tested NPs, none of the other published alternatives, such as the aforementioned DARTS[41], SPROX[43], Lip-MS[133], CETSA[45]/TPP[44], AfBPP[134], PAL[48] or ABPP[50] is universal either. At the same time, the approach proposed here is technically very straightforward and can be deployed to a wide range of biological systems, so it is certainly worth considering for NP target deconvolution.

Target deconvolution exercises rationally tend to initiate after the discovery of a novel NP bioactivity in a phenotypic assay. Chemoproteomic-assisted target deconvolution in this context will be strengthened by the integration of such an assay prior to any selective immobilization effort. A cell viability screen, for example, would enable the selection of a relevant biological system for profiling as well as a reasonable concentration range, where the NP displays its bioactivity. Moreover, a combination of multiple successful immobilization chemistries for one NP can afford affinity matrices with multiple linkage points; those combined can be used for agnostic target deconvolution and establishment of target SAR. The experiment with staurosporine indicated that two tailored affinity matrices prepared using two immobilization chemistries produced complementary target profiles; combined, they should enable recapitulation of all protein targets from each of the two individual competition assays. This can be extended on any new NP with unknown targets succeeding the initial phenotypic assays, where multiple immobilization chemistries should be tested in parallel, and individual affinity matrices would be utilized separately as well as in a mixture. Comparing obtained binding profiles for individual matrices should allow for finding moieties most likely participating in target binding, which would not be possible using probe-free techniques like CETSA.

9 Part 2. Molecule-centric drug discovery

A master thesis (Verena Filz), a Bachelor thesis (Karlis Blums) and internship (Yuchun Li, Katrin Eisenmenger, Roman Graf, Stefanie Nadler, Stefanie Hoefer) projects were carried out in the context of this work under the supervision of Polina Prokofeva.

9.1 Proteome-wide structure-affinity relationships (SAR) with affinity-based proteome profiling (AfBPP)

Usually, target deconvolution follows the discovery of an interesting phenotype for a particular molecule. This is true for bioactive NPs and was discussed in the previous chapter of this thesis. There, only the molecules previously reported to possess activity in at least one phenotypic assay were employed for chemoproteomic-assisted target deconvolution. Generally, when the assayed compound does not display an effect on any given biological system, one may argue that there is no reason for target deconvolution. However, when a compound affects the cellular phenotype, there is merit in learning what molecular action leads to the observed effects. Classic target deconvolution methods were discussed in the introduction part of this thesis, in particular, the methods that utilize chemical probes based on the structure of the molecule. Target deconvolution with affinity matrices, for instance, typically requires chemical synthesis of linkable analog(s) that can be used for immobilization on beads and preparation of corresponding tailored affinity matrices. These linker-containing analogs are then tested to confirm that the activity is maintained, i.e., that the linkage is not impeding essential interactions between the phenotype-responsible target and the pharmacophore. The linkable analog can then be immobilized onto a solid matrix for the chemoproteomic experiment in a relevant lysate, which consists of a competition between the immobilized analog (i.e., the tailored affinity matrix) and the parental bioactive molecule, followed by mass spectrometry-based bottom-up proteomics.

The study presented in the following chapter reversed the usual sequence of events: first, a library of structurally close linkable molecules was prepared, then the target deconvolution of each molecule was performed, independently of any phenotypic or activity assay. The chemoproteomic assay hence acted as 1) a screening of novel molecules against the proteome, 2) a selectivity profiling of each of these molecules, and 3) a proteome-wide SAR of the chemotype (Fig. 26). Instead of making a linkable analog for one interesting bioactive molecule, here a small library of affinity probes of one particular chemotype was generated, which allowed for exploration of the SAR for a particular chemical scaffold.

Conceptually, the design of the analog library here is based on the premise that the bead-linked part of the chemotype is solvent-exposed, while the opposite side dives deep into the protein binding pockets. Thus, chemotype diversification at the deep-reaching moieties would affect the selectivity profiles of each individual analog and may create affinity and selectivity cliffs. Such a novel molecule-centric target-agnostic drug discovery approach should enable establishing of the simultaneous structure-affinity relationships for all expected and unexpected protein targets of a given chemotype, shedding light on selectivity and allowing for pharmacophore repurposing (Fig. 26).
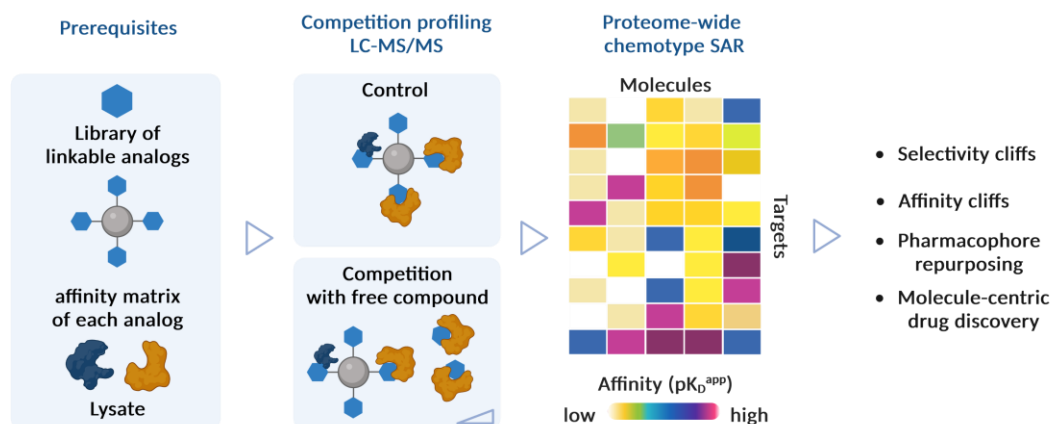
Figure 26. Schematic representation of chemoproteomic-assisted proteome-wide SAR for a given pharmacophore. A library of linkable analogs of a known bioactive molecule is synthesized and immobilized on Sepharose beads. Each tailored matrix is used in pulldown experiments to identify the targets of the molecule. In the obtained data matrix, where the color of each cell reflects the affinity of each protein/molecule interaction, one column constitutes the profile of one molecule. By comparing the binding profiles obtained for the different analogs, the proteome-wide SAR of the pharmacophore can be established.

9.2 Proof of concept: proteome-wide selectivity profiling of pyrimidopyridone chemotype

The applicability of the chemoproteomics-assisted molecule-centric approach for establishing proteome-wide chemotype selectivity and SAR was first evaluated for a kinase-targeting pyrimidopyridone scaffold (Fig. 27). Owing to their broad kinase enriching selectivity, several analogs of the chemotype were previously internally evaluated as baits for Kinobeads, notably Kinobeads compound 1 and both VI molecules (Fig. 27). Interesting to note, these molecules were only tested for their ability to enrich kinases out of native cell lysate, but their targets were never deconvoluted. One of the analogs within the small chemotype library is a cyclin-dependent kinase 4 and 6 (CDK4/6) inhibitor palbociclib, approved for the treatment of hormone receptor (HR) positive, human epidermal growth factor receptor 2 (HER2)-negative advanced breast cancer. Based on previously published literature on target deconvolution of palbociclib as well as available crystal structures, it is known that immobilization via its piperazine moiety does not affect the ability of the obtained probe to engage protein targets (Fig. 28).[135] It was hypothesized that the other six molecules in this study adopt a similar binding mode of their kinase targets and can, therefore, be analogously functionalized on beads via their aminoethoxyaniline groups. In total seven pyrimidopyridone analogs were evaluated here: their respective affinity matrices were prepared through direct immobilization via primary or secondary amines on NHS-activate Sepharose beads, and all molecules were subjected to a dose-dependent chemoproteomic selectivity profiling in HeLa cell lysate (Fig. 29). HeLa cells were previously shown to express one of the widest panels of kinases in a single entity and were, therefore, chosen as a biological system for profiling in this study (protein expression data was analyzed from ProteomicsDB.org[136,137]). Each profiling assay consisted of a competition with 9 doses (up to 30 µM) of the free molecule, followed by multiple washing steps and on-bead tryptic digestion. Obtained peptide eluates were subjected to a C18 StageTip desalting cleanup and finally to MS measurement. The raw data were searched against a Uniprot Human sequence database using MaxQuant software with an embedded Andromeda search engine. The intensity of each quantified protein group in each

competition dose was plotted relative to DMSO control, and individual proteins that showed typical sigmoidal intensity decrease with increasing concentration of the competing molecule were annotated as potential targets of the compound. Each dose-response curve allowed for the calculation of an EC_{50} and corresponding apparent binding constant K_D^{app} that defines the specificity and strength of each binding interaction among all bead-bound proteins.
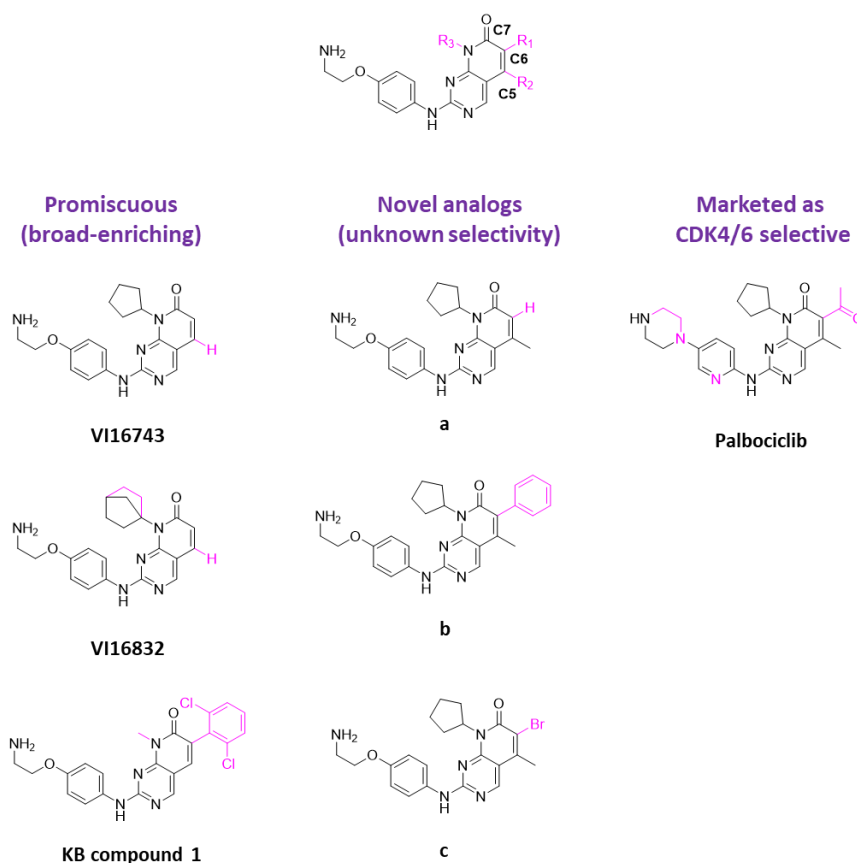
Figure 27. Evaluated pyrimidopyridone chemotype analogs. Molecules a, b and c were prepared and characterized by Dr. Guillaume Médard and generously provided for the selectivity analysis in this work.
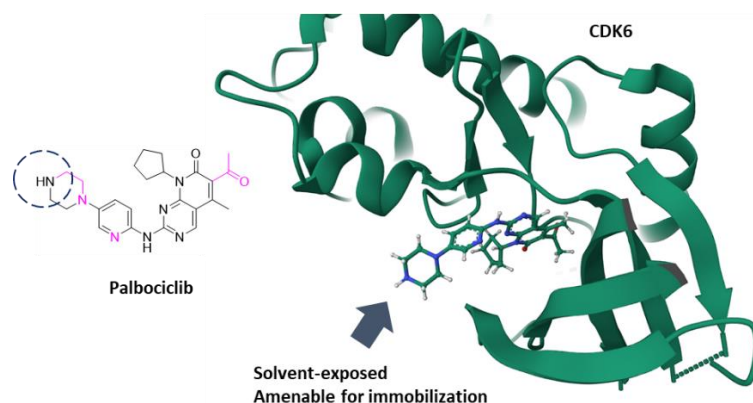
Figure 28. Palbociclib and its binding to the designated target CDK6 (PDB: 5L2I). The secondary amine of the piperazinyl substituent is solvent-exposed, therefore, should be amenable for bead functionalization.

The results of the chemoproteomic profiling showed that seven analogs bound in total 191 different targets (Fig. 29a), while CDK6 was the only common target of all molecules (Fig. 29d). Kinobeads compound 1 has the most dissimilar chemical features compared to other molecules analyzed here. While each analog demonstrated distinct binding profiles, compound 1, unsurprisingly, bound the most unique selection of protein targets (36 unique). It would be wrong to make direct conclusions regarding its targets SAR, as no analogs differ from compound 1 by a single-step substitution within the assayed set. However, it is possible to notice that the addition of a bulky R_1 substituent as well as removal thereof from R_2 and R_3 drastically affects selectivity of the molecule. Overall, pyrimidopyridone scaffold proved to be rather promiscuous, and the kinase targets were identified across six families of the kinome phylogenetic tree (Fig. 30).
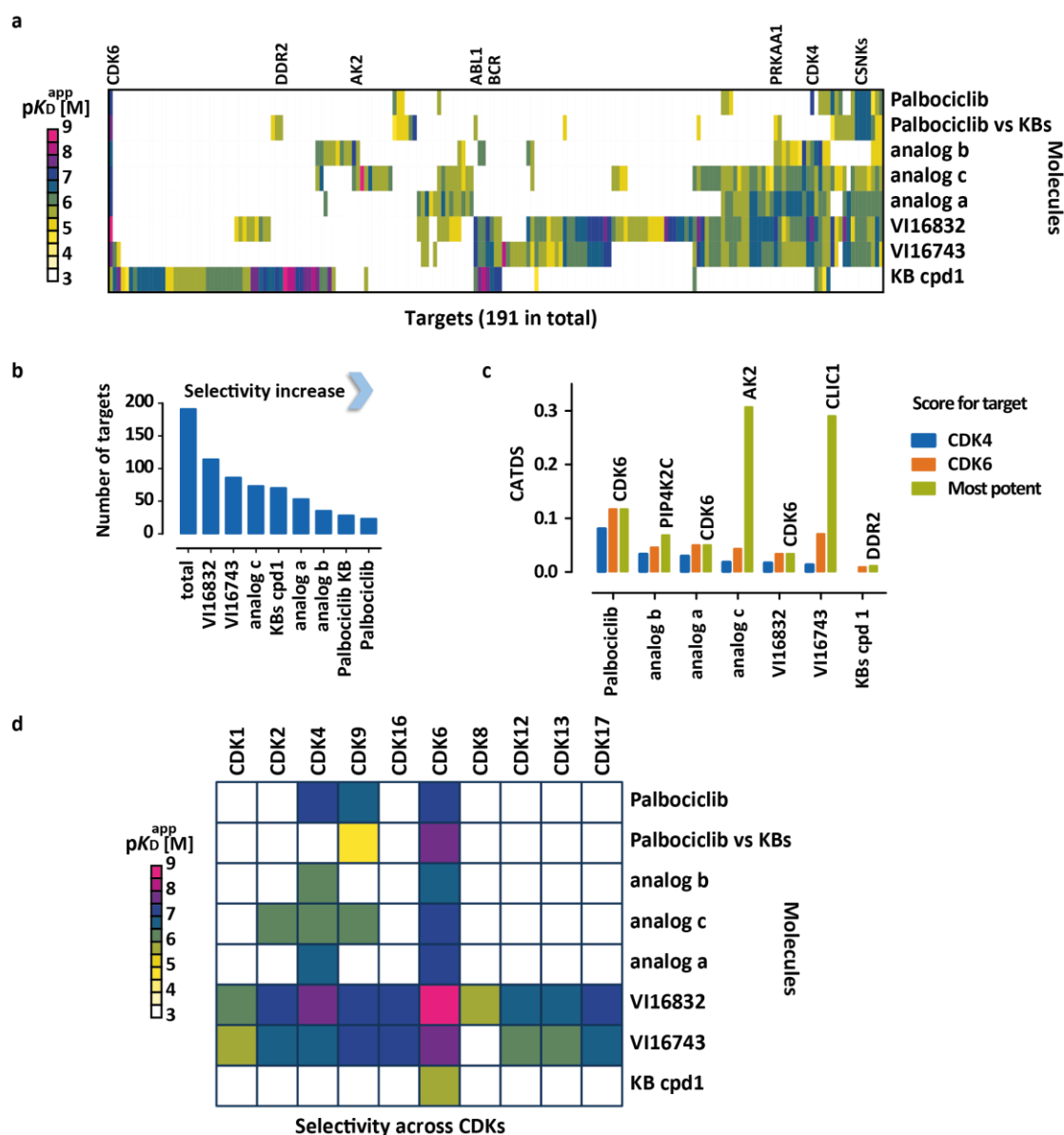
Figure 29. Chemoproteomic selectivity profiling of pyrimidopyridone scaffold analogs in HeLa lysate. a) Heatmap depicts targets and their corresponding binding affinities identified for each analog. b) Selectivity estimation based on the number of identified targets for each compound. c) CATDS scores for CDK4,6 and the most potent target of each compound. d) Heatmap of CDK targets, similar to panel a.
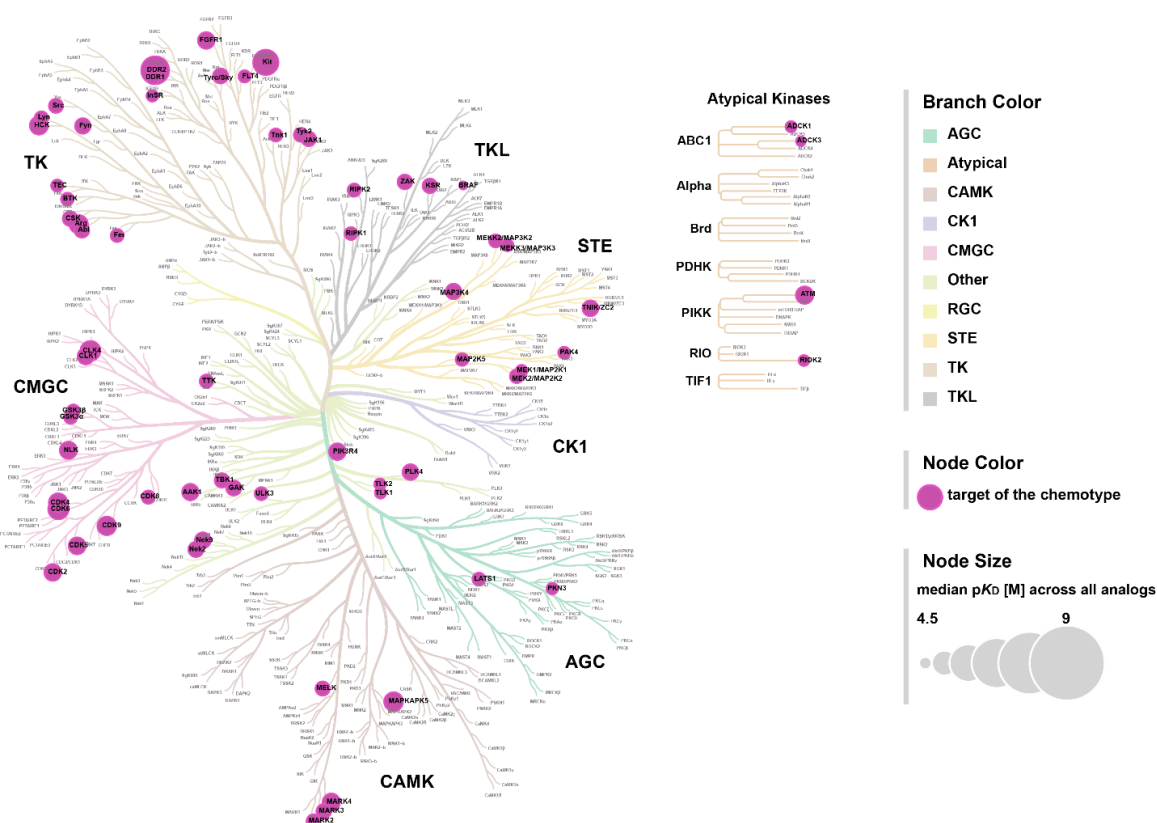
Figure 30. Kinase pyrimidopyridone chemotype targets across the kinome phylogenetic tree. Nodes indicate kinases identified as targets of the chemotype; the nodes' size depicts median binding affinity across all analyzed molecules. The graphic representation was created using the publically available Coral tool by Metz et al..[138]

Among all seven molecules, palbociclib, indeed, proved to be the most selective with 23 targets <30 µM (Fig. 29b). Nonetheless, assessing compound selectivity solely by the number of targets it can address is prone to misleading interpretation, as often, the binding affinity significantly differs for different protein targets. A score that accounts for selectivity for a particular target at a particular drug dose was developed by Heinzlmeier et al. CATDS, or concentration- and target-dependent score, was created based on chemoproteomic dose-dependent competition profiling data. It enables the selectivity calculation dependent on a particular target engagement at a selected drug concentration.[59] The score ranges between values of 0 and 1, where values close to 1 indicate very selective compounds. It is defined as:

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

where the score is calculated from target engagement (value derived from the curve fit) of a particular target(s) of interest at the respective K_D^{app} concentration relative to the total target engagement.

CATDS scores calculated for two designated palbociclib targets, CDK4 and 6, are considerably low (Fig. 29c), meaning that palbociclib binds a range of other targets with similar potency. Undeniably, among all analyzed analogs, palbociclib shows the best selectivity for CDK4 and 6, but it is unjust to proclaim this drug selective *per se*. Palbociclib was found to bind eight other kinase targets with affinities < 1 µM and 12 more with affinities < 10 µM.

This study is not the first report that assessed the selectivity of palbociclib using chemoproteomics methods. For instance, the publication by Sumi et al. reports the immobilization of palbociclib and the use of this affinity matrix in a single high 20 μM dose competition assay for target identification in H157 cell lysate.[135] Qualitatively, the targets presented in this study are in good accordance with those reported by Sumi et al., as the majority of targets were equally identified here. Those include CDK4/6, PRKAA1, MAPK8, GSK3B, several CSNKs, MAPK9, MAPK1, PIP4K2C, TTK, and FER. Only a few previously reported targets, such as TBK1, PTK1, CDK11B, and AAK1 were not identified in this work. Interestingly, TBK1, CDK11B, and AAK1 were bound by the affinity matrix but did not show characteristic competition to unequivocally identify as targets of palbociclib. PDXK, PTK2B, and YES1, although found competing for binding here, were not a part of the target spectrum of palbociclib in the report by Sumi et al. The difference in selectivity profiles might be explained by the use of a different cell lysate or the fact that the other study assessed the targets using one competition dose, which might have led to false positive identifications. It is also important to note that the experiment reported by Sumi et al. could not comprehensively assess the binding affinities to each individual target (i.e., EC50 and $K_D{}^{app}$), unlike this study, as only one high dose of 20 μM was used in their competition assay.

In the meantime, compared to specifically CDK4/6 inhibition-optimized palbociclib, broad-binding analogs c and VI16743 showed a much higher selectivity for their most potent targets, AK2 and CLIC1, respectively. Both VI analogs demonstrated the broadest selectivity across different CDKs (Fig. 29d) and broader binding across other kinase families, where each compound showed dose-dependent competition for more than 90 proteins. It appears that the lack of substituents on C6 in combination with 8-cyclopentyl-5-methylpyridopyrimidinone render the chemotype significantly unselective (referred to chemotype scheme in Fig. 27). Comparison of binding profiles obtained for VI16743 vs. analog a against their respective tailored beads revealed that the addition of a single methyl-group at R_2 position leads to a substantial loss of targets (Fig. 31a), as 45 less proteins showed typical competition for bead binding with the increasing concentrations of the free compound. The three binders that appear as targets identified uniquely for analog a, i.e., CSNK1E, RPS6KA1, and RPS6KA3, in fact, also displayed minor competition for bead binding in the profile of the VI compound. However, here the intensity loss at the highest assayed competition dose did not exceed 50% compared to the control; these proteins, therefore, were not considered potential binders of VI16743. Notably, adding methyl at R_2 led to a loss of 12 potent VI binders, including CDK9, CDK16, and CDK17, where the VI16743 beads barely or not at all enriched any of these 12 proteins. On the other hand, there were no drastic changes in the binding affinities calculated for targets shared between the two analogs. Further addition of bulkier substituents on R_1 resulted in more diverse changes on recovered selectivity profiles (Fig. 31b), and only 20 proteins were identified as common targets of the analogs a, b, and c. The addition of a phenyl group (analog b) produced a compound capable of engaging 9 unique targets, including MAP4K4, MAP2K1 and 2, DDR1, KSR1, SIK2, BCR, ABL, and FECH. Neither analog b nor c could bind PIK3C3, OSBPL3, MAO3K2, TYK2, DYNLL1 or EML4. Furthermore, analog c, in general, demonstrated a more unique panel of targets among the 3 (22 unique, including CDK2 and CDK9), with the most potent binding to AK2 ($K_D{}^{app}$ < 1 nM) and ATM ($K_D{}^{app}$ 217 nM).

Target profile comparison
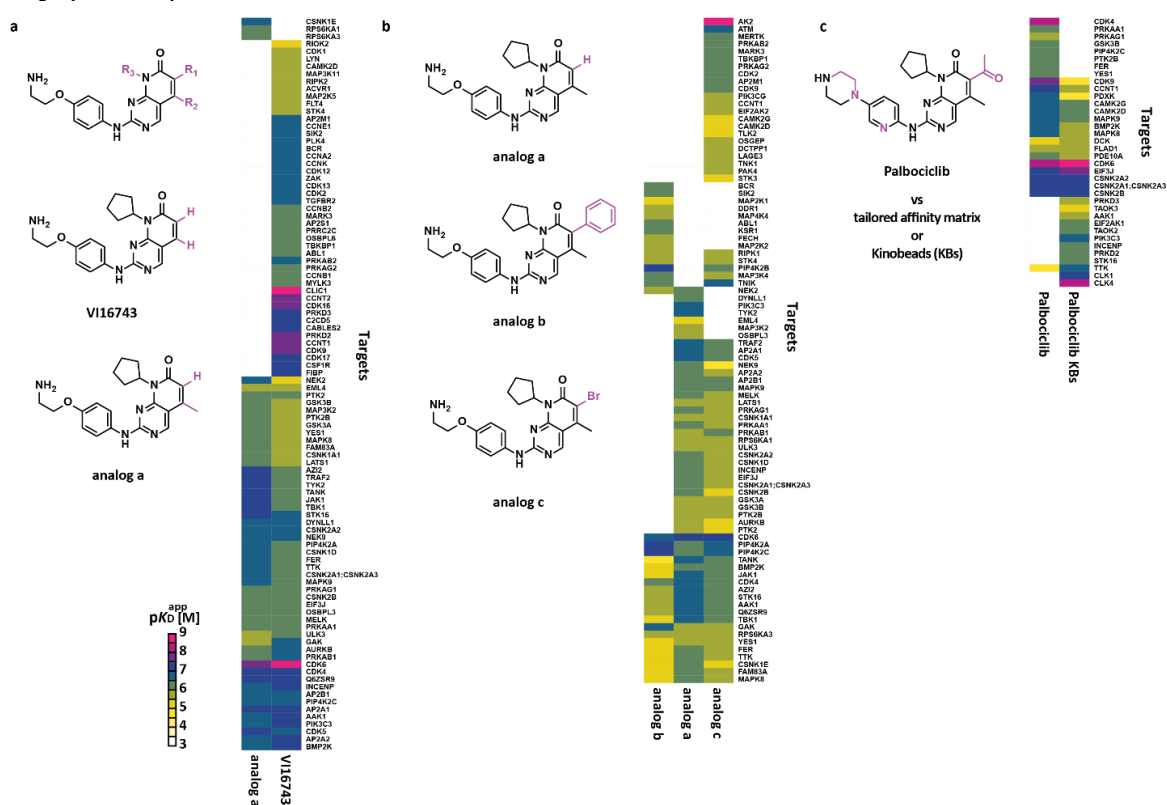
Figure 31. a and b) Comparison of target profiles between the close structural analogs (1 step substitutions). Heatmaps depict targets and their corresponding binding affinities identified for each analog profiled against its respective immobilized avatar. c) Comparison of targets found in competition pulldowns of palbociclib vs. tailored affinity matrix or Kinobeads.

Ironically, the current version of Kinobeads does not stably enrich CDK4.[57] This was reflected in the selectivity profiling of palbociclib against Kinobeads matrix in this work, where CDK4, the designated target of the drug, could not be identified (Fig. 31c). The use of a tailored affinity probe, however, allowed for stable enrichment and subsequent target identification. Moreover, several proteins that clearly displayed the competition in the profiling against the tailored probe were not annotated as targets in the assay against Kinobeads (Fig.32). These results are likely not arising from technical variation as both competition assays were performed on the same day using one batch of HeLa lysate. Surprisingly, although Kinobeads were able to enrich CDK9, the competition with palbociclib recapitulated much weaker binding affinity compared to the competition against palbociclib beads (Fig. 32). Altogether, these findings may indicate that the nature of target engagement between palbociclib and certain molecules on Kinobeads are different, i.e., the binding might potentially take place via different pockets or protein conformations. However, this statement must be taken cautiously and validated with further experiments. Moreover, this profiling suggests that previous Kinobeads screens could not identify the entire target spectrum of the analyzed drugs and simultaneously hints that adding a tailored probe to the Kinobeads mixture could benefit profiling across the proteins that the Kinobeads matrix cannot enrich.

On the other hand, the comparison of the two palbociclib profiles also demonstrates that several targets were lost upon on-bead immobilization (CLK1 and 4, STK16, PRKD2 and 3, TAOK2 and TAOK3, AAK1, EIF2AK1, PIK3C3), and the affinity matrix, indeed, could no longer enrich those out of lysate. It can be hypothesized that the piperazinyl immobilization perturbs the formation of

stable affine interactions crucial for binding these proteins. Overall, this comparison displays the shortcomings of both selectivity profiling approaches and prompts the incorporation of tailored affinity probes and broad-enriching ones to ensure stable enrichment of all plausible protein binders.
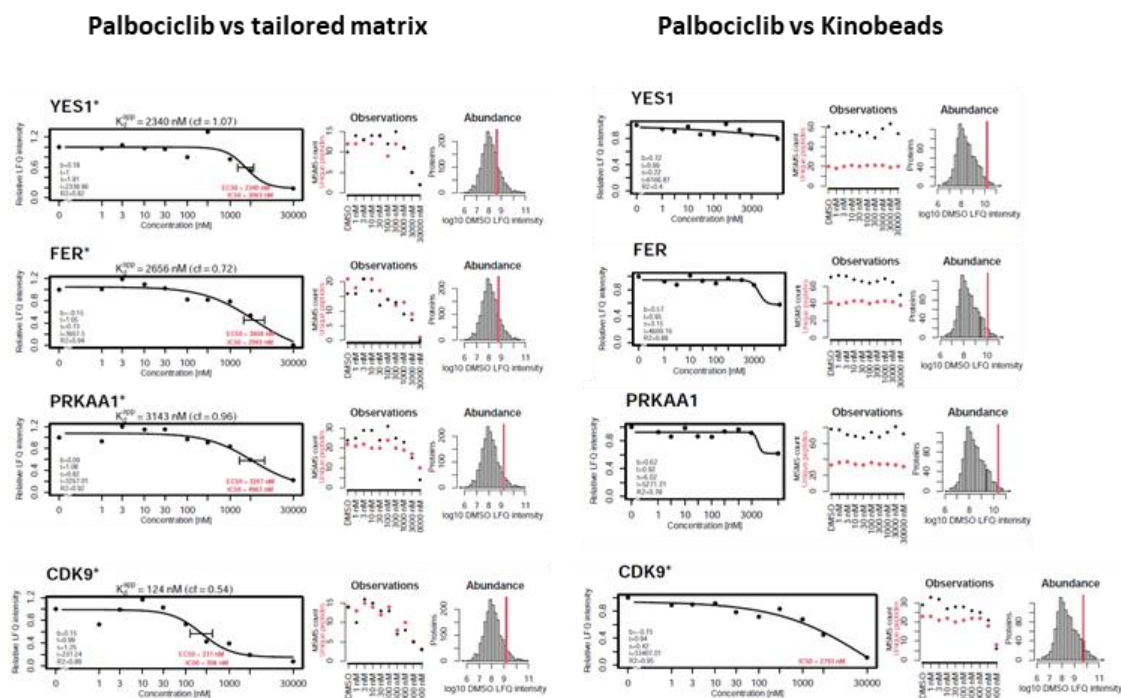
Figure 32. Dose responses for palbociclib targets. The profiling against a tailored affinity matrix shows dose-dependent competition. Profiling against Kinobeads does not display the same trend: for several well-enriched targets the competition is not taking place.

Analogs a, c, and both VIs bound most of their targets with affinities close to 1 µM (Fig. 33a), whereas the overall number of potent interactions remained small. Compound 1, conversely, demonstrated consistent protein binding across the entire span of the affinity range in this experiment. This data is in accordance with calculated CATDS selectivity scores, where the big difference between binding affinities of the most potent interactors of VI compounds to the rest actually renders the molecules more selective for these protein targets compared to the overall "more selective" palbociclib (in the context of CDK4 and 6 inhibition).

Analysis of protein enrichment by tailored affinity matrices here indicates that high protein intensity in a simple triplicate pulldown can potentially be indicative of specific target binding (Fig. 33b). The proteins that were identified with high intensities in a simple enrichment experiment were equally identified as potential targets of the respective molecules in the following competitive profiling. However, it is essential to note that there is no linear correlation between high protein enrichment and calculated binding affinities in a competition experiment, as demonstrated here for compound 1 (Fig. 33c). All proteins that were annotated as targets of compound 1 were indeed identified with LFQ intensities above the median intensity of the DMSO control; nevertheless, no linear correlation of protein enrichment to obtained target pK_D values was observed. The amount of a protein captured on beads primarily depends on its general abundance (expression) in the utilized biological system and the affinity of the protein to the immobilized molecule. For example, in the extreme case, a very low abundant target might not appear in the top tier of the most abundant bead-bound proteins even if it possesses high binding

affinity to the matrix due to its low quantity in lysate compared to other binders. Calculating enrichment factors relative to total proteomes rather than directly comparing protein intensities in a pulldown would unravel more evident dependencies between target enrichment and binding affinities.
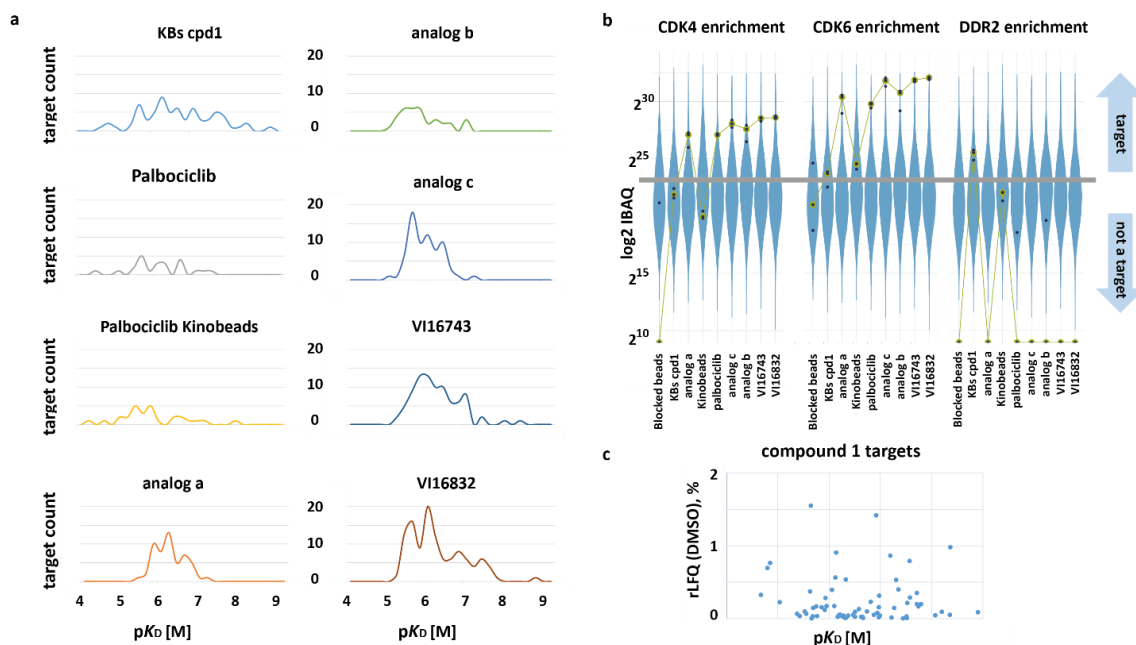
Figure 33. **Potency and target enrichment of chemotype analogs**. a) target count across binding affinities of all compounds. b) differential protein enrichment by the affinity matrices c) protein enrichment represented as a fraction of total protein intensity (rLFQ, %) vs. target pK_D.

Altogether, the analysis of seven pyrimidopyridone scaffold analogs demonstrated that it is possible to uncover the influence of the changes in the chemical structure on the selectivity profile of each individual molecule. However, with this small set of compounds, it was challenging to establish structure-affinity relationships for the chemotype as several analogs that would enable it to complete the chain of single-step structural modifications were not included. In order to evaluate the proteome-wide chemotype SAR approach more systematically, it was extended to another chemical scaffold, namely 2,4,5-triarylimidazoles (lophines). The results of this profiling will be discussed in the following chapter.

9.3 Proteome-wide SAR of lophine chemotype

The proteome-wide SAR concept preliminary explored on the pyrimidopyridone scaffold was extended to the 2,4,5-tri(hetero)arylimidazole (or lophine) chemotype. The choice was driven by the diversity of targets, primarily kinases, this chemotype can engage (Fig. 34). The scaffold was initially designed to inhibit p38; however, several commercially available inhibitors were reported to selectively inhibit other kinases, such as ALK5 (TGFBR1), B-Raf, and Tie-2 (TEK). Moreover, adezmapimod (SB203580), SB610251B, and FHPI have been previously reported to engage BET bromodomain-containing protein BRD4, indicating that the chemotype equally reaches beyond kinase binding.[139,140]

Adezmapimod (SB203580)
p38 inhibitor (300-500 nM)

FHPI (SB202190)
p38 inhibitor (50-100 nM)
BRD4-1 inhibitor (2.5 μM)

SB610251B
p38, RET, SRC inhibitor
BRD4-1 inhibitor (5.7 μM)

L-779450
B-RAF inhibitor (2.4 nM)

hGlur antagonist (53 nM)

Tie2 inhibitor (94 nM)

SB431542
ALK5 inhibitor (94 nM)

Vactosertib
ALK4/5 inhibitor (13/11 nM)
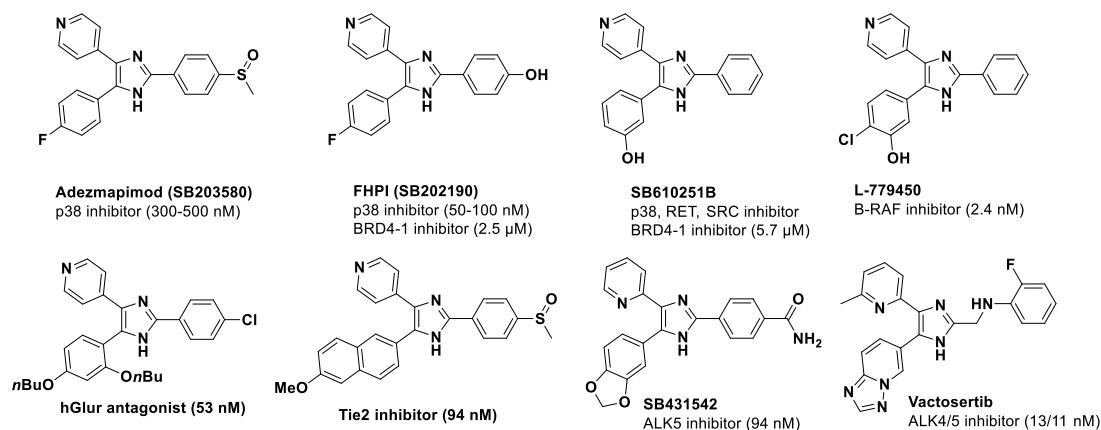
Figure 34. The 2,4,5-tri(hetero)arylimidazole chemotype: examples of inhibitors within the chemical scaffold and their reported activities.

MAPK14 (or p38α) is a serine/threonine kinase, an essential component of the MAP kinase signaling pathway, known to phosphorylate a broad range of proteins through ATP binding. Upon binding to MAPK14, the adenine of ATP forms two H-bonds with His107 and Met109 (Fig. 35a). In addition, three residues, namely, Lys53, Glu71, and Asp168, coordinate the triphosphate group, which enables the efficient catalysis of phosphorylation transfer to the substrate. Adezmapimod has been shown to bind MAPK14 in the ATP-binding cleft. Here, the most crucial for inhibitory activity interaction is the hydrogen bond formed between the nitrogen of the para-pyridine ring and the backbone amide NH of Met109 (Fig. 35b). The para-fluoro-substituted ring interacts with the hydrophobic pocket close to the gatekeeper of an adjacent selectivity pocket (Glu71 and Asp168), namely Thr106, however, the inhibitor occupies only a part of the entire available hydrophobic space (close to His107).[141,142] The imidazole nitrogen additionally forms a hydrogen bond with the amino-group of Lys53, whereas the 2-aryl group extends into the phosphate-binding region, where it can engage in π-π stacking with Tyr35.[143]

The adezmapimod activity was initially associated solely with MAPK14 inhibition; however, it has been proven to bind other kinases with similar affinities. For instance, a broad-selective kinase activity assay revealed GAK, CSNK1A1, MAPK8, MAPK9, and RIPK2 as targets of the compound.[144] Co-crystallization of RIPK2 with adezmapimod showed that the binding is defined analogously to MAPK14 by the hydrogen bond between the backbone amide of Met98 of the hinge and the para-pyridine nitrogen on the 4-aryl ring. The para-pyridine and imidazole occupy the adenine of the ATP binding site, while para-fluoro-substituted aryl fills a part of the hydrophobic back pocket. The para-sulfoxide group of the 2-aryl ring is extended through the sugar pocket of RIPK2 and is solvent-exposed.[145]

Since the methyl- of the sulfinyl-group on the 2-aryl ring is solvent-exposed for both kinase-adezmapimod complexes, it can be considered a position for installing an immobilization handle. Such functionalization should preserve the most vital affine interactions with the kinase pocket. It was, therefore, decided to functionalize all novel library analogs synthesized in this study with a primary-amine handle on the 2-aryl ring, which would allow their direct immobilization on beads (Fig. 35c, immobilization on Sepharose beads according to Fig. 8). The linkable analog of adezmapimod was previously utilized in the study that refined the characterization of the conformation-dependent binding of kinases to immobilized inhibitors.[146] The affinity matrix was employed to compare the enrichment of kinases from SK-N-BE(2) lysate prepared from cells with and without pervanadate treatment. The results of this study demonstrated that a) the linkable analog is capable of enriching kinases in lysate, and b) apart from a very small number of exceptions, the majority of kinases do not show differential conformation- or activity-dependent binding. From the collection of inhibitors in Fig. 34, one may hypothesize that the chemotype selectivity towards different kinase targets is driven primarily by the substituents on the 4- and 5-aryl rings. It should, therefore, be possible to agnostically probe the general ability of this pharmacophore to engage protein targets by collecting the selectivity profiles of analogs with varying substituents on 4,5-aryls.
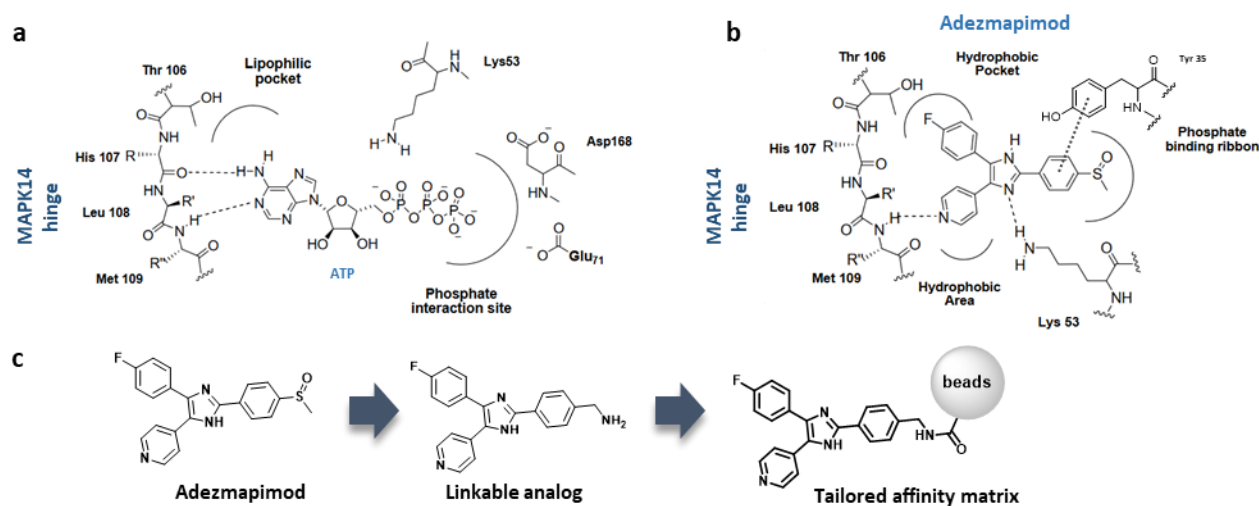
Figure 35. **Triarylimidazole scaffold MAPK14 binding.** a) ATP binding. b) Adezmapimod binding. c) the example of a linkable adezmapimod analog from this study's newly synthesized compound library. Panels a and b adapted from Bagley et al. and Poon et al.[142,143]

9.4 Synthesis of a 2,4,5-tri(hetero)arylimidazole library of analogs for target deconvolution

The small library of 35 linkable lophines (2,4,5-triarylimidazoles) was synthesized, where the obtained analogs can be divided into five separate compound series (Fig. 36). Owing to the number of targets bound by the chemotype analogs (Fig. 34) featuring a para-pyridine, a [pPy] series with a variety of aryl substitution was prepared. Other series comprise di-aryls ([Ph-] series), meta-pyridine ([mPy] series), and ortho-pyridines. The latter were prepared with ([oPy-3Me] series) or without ([oPy] series) the methyl-substituent encountered in vactosertib (compound structure can be found in Fig. 34).

[pPy] series   [oPy] series   [oPy-3Me] series   [mPy] series   [Ph-] series
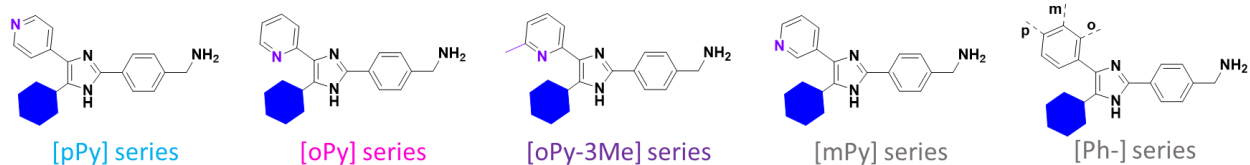
Figure 36. The library of 35 linkable analogs of 2-[Ph-pCH2NH2]-4/5-disubstituted imidazoles divided in 5 series.

Synthesis of the tri(hetero)aryl imidazole scaffold was achieved via the multicomponent Debus–Radziszewski reaction from an ad-hoc diketone, 4-cyanobenzaldehyde and ammonium acetate in acetic acid at 180°C in a microwave reactor (Fig. 37).[147]
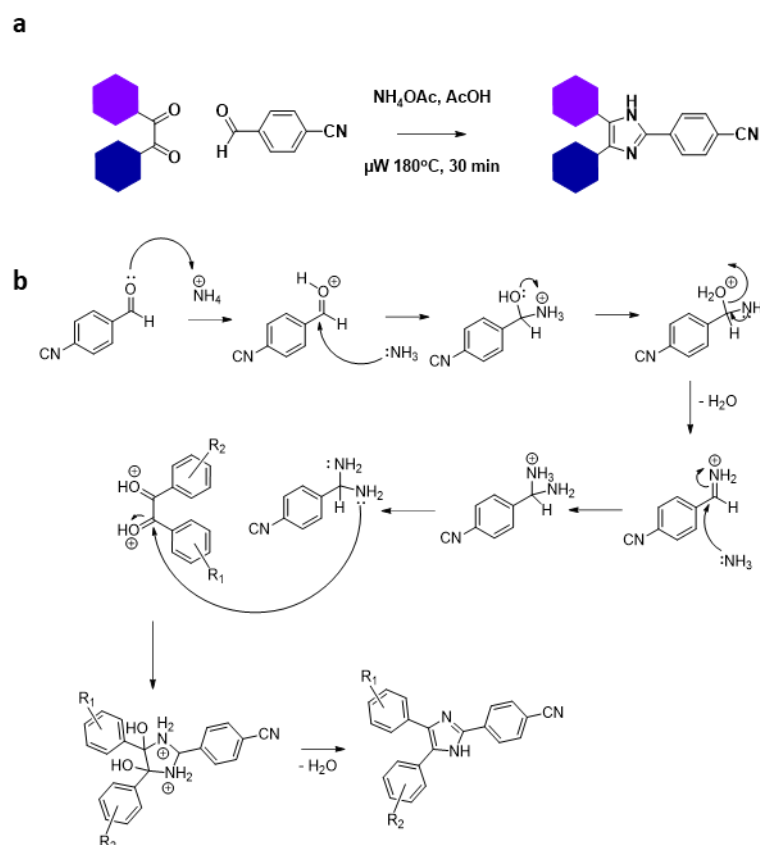
Figure 37. The tri-component Debus–Radziszewski reaction to obtain lophine analogs. a) general reaction scheme. b) the plausible reaction mechanism.[148]

The nitrile intermediates were then reduced with LiAlH$_4$ to obtain the primary amine ready for immobilization (fig. 38).[149] Several synthesized analogs featured methylated phenols, which were deprotected into phenols using BBr$_3$ before reducing the nitrile.[150] Hence, after the imidazole forming MCR, two different nitrile analogs were obtained (Fig. 38).
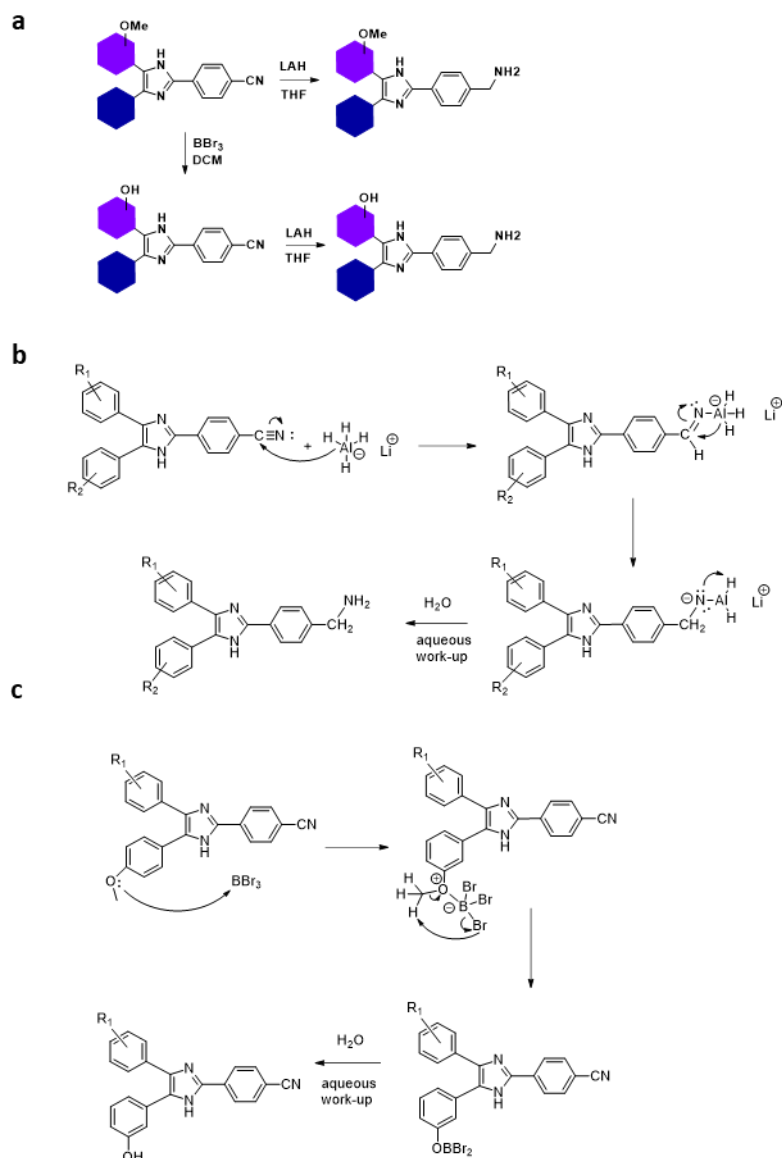
Figure 38. Nitrile reduction with LAH and BBr₃ demethylation.

The synthesis of the diketones proved to be the most variable part of the synthetic route; therefore, different types of protocols according to the available starting material and their reactivity were applied. For instance, a number of analogs were obtained via the copper(I)-catalyzed Sonogashira coupling[151] followed by palladium(II)-catalyzed and DMSO-mediated alkyne oxidation[152] sequence. A novel one-pot procedure that allows for direct oxidation of the Sonogashira coupling product into a diketone without isolation or purification steps was also assayed. Although it proved successful for the phenyl compound series, none of the pyridine-containing analogs could be successfully oxidized without isolation of the alkyne. Upon forming the alkyne intermediate, the one-pot reaction no longer progressed in various assayed conditions. Therefore, alternative synthetic routes specifically for the pyridine compound series were sought. In order to circumvent the alkyne oxidation step altogether, an LDA-lithiation of methyl pyridines for the addition of benzoic ester or Weinreb amide followed by oxidation of the enol with either selenium dioxide in acetic acid or hydrobromic acid in DMSO was applied (Fig. 39a).
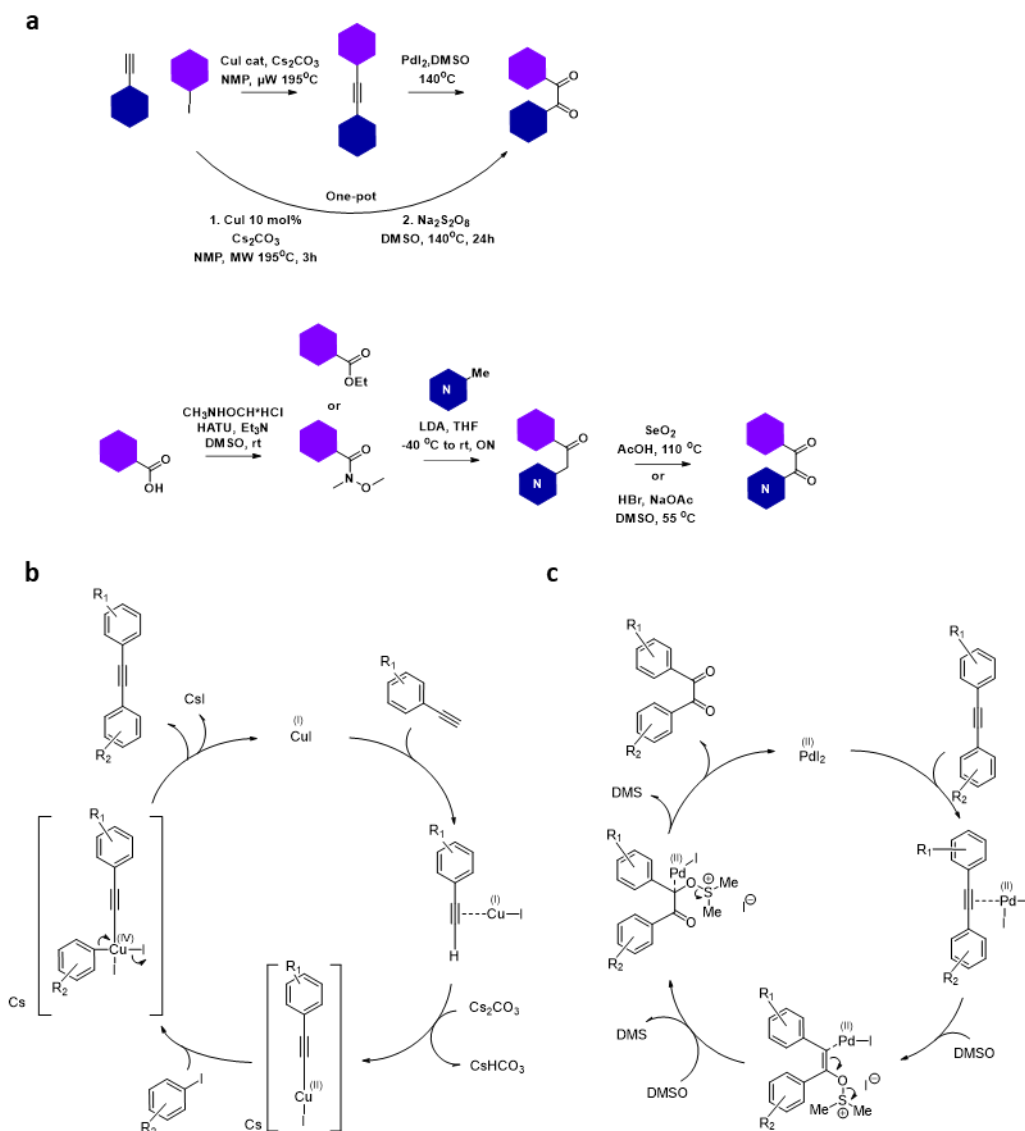
Figure 39. General preparation scheme of the diketones necessary for the Debus-Radziszewski reaction. a) three synthetic strategies applied in this work. b) reaction mechanism of copper(I)-mediated Sonogashira coupling. c) mechanism of palladium(II)- and DMSO- mediated alkyne oxidation.

In total, 35 linkable analogs were obtained, while another five, whose nitriles were successfully prepared, could not be isolated in a pure form, primarily due to their high hydrophilicity (Fig. 40).

a

Legend for cells: ■ = violet, ▨ = yellow, ░ = grey

	[Ph]	[Ph-pF]	[Ph-pOH]	[Ph-pOMe]	[Ph-mOH]	[Ph-mOMe]	[Ph-mOHpCl]	[Ph-mpOCH2O]	[Ph-mOHpOH]	[Ph-mOMepOMe]	[Ph-mOHmOH]	[Ph-mOMemOMe]	[Ph-mOHmOHpOH]	[Ph-mOMemOMepOMe]	[Ph-oOHpOH]	[Ph-oOMepOMe]	[2Napht]	[2Napht-6OH]	[2Napht-6OMe]	[Ph-pSO2Me]	[Ph-pCl]
[pPy]	■	■	■	■	■	■	■	■	▨	■	▨	■	▨			■	▨	■	■		
[mPy]		■	■																		
[oPy]			■	■		■	■	▨	■												
[oPy-3Me]				■	■																
[Ph]				■	■																
[Ph-pF]	░		■	■												■	■			■	
[Ph-pOH]	░	░			■															■	
[Ph-pOMe]	░	░	░			■															■

b

Reagents/conditions (left to right):
CuI 10%, Cs$_2$CO$_3$, NMP, 180°C μW, 3h — 14% →
PdI$_2$ 10%, DMSO, 140°C μW, 3h — 77% →
NH$_4$OAc, AcOH, 180°C μW, 1h — 49% →
LiAlH$_4$, THF, 0 °C, 3h — 23%

c

Reagents/conditions:
CH$_3$NHOCH$_3$·HCl, HATU, Et$_3$N, DMSO, rt, 2h — 82% →
LDA, THF, -40°C to rt, ON — 32% →
HBr, NaOAc, DMSO, 55°C — 20% →
63% ↓
NH$_4$OAc, AcOH, μW 180°C, 15 min — 55%
BBr$_3$ 1M in DCM, DCM, 0°C to rt — 54% ←
LiAlH$_4$, THF, 0°C to rt
could not be isolated
LiAlH$_4$, THF, 0°C to rt

Figure 40. a) Table summarizing the compounds that have been obtained and characterized. The violet squares show the obtained 2-[Ph-pCH$_2$NH$_2$]-imidazoles, while the yellow squares show the obtained 2-[Ph-pCN]-imidazoles that would not yield the reduced compounds in a pure form. The grey color represents the compounds that have not been made but would be necessary for a sequence of single-step structure variations. b) Synthetic route using Sonogashira coupling followed by PdI2/DMSO oxidation to produce the diketone that yielded e.g. 2-[Ph-pCH$_2$NH$_2$]-4/5[oPy][Ph-mOHpCl] imidazole. c) Synthetic route featuring a Weinreb amide intermediate of a dimethylated catechol allowing access to the dimethylated catechol nitrile and the unprotected catechol nitrile. After reducing such species, only the protected catechol could be isolated with sufficient purity.

9.5 Target deconvolution results reveal the proteome-wide SAR of lophine chemotype

Each of the 35 newly synthesized linkable molecules was directly immobilized on NHS-functionalized Sepharose beads using the procedure illustrated in Fig. 8. In order to control for optimal bead loading, each compound was immobilized in three coupling densities of 0.5, 1, and 2 μmol/mL, and subjected to a pulldown experiment in the mixed (1:1) placenta and K562 lysate. The protein eluates were then resolved on SDS-page gel and visualized with silver nitrate staining

(analogous to the previous study). For none of the 35 molecules, the highest tested density caused drastic unspecific binding; therefore, for consistency, only the 2 µmol/mL loaded matrices were used in the subsequent competition experiments. Analogously to the experimental setup applied for the pyrimidopyridone scaffold, each profiling assay consisted of the competition pulldown with 9 doses (up to 30 µM) of the free analog against its corresponding tailored affinity matrix. About a third (11 molecules out of 35 profiled) of the aminomethyl-analogs were found to bind in a total of 44 different protein targets with at least micromolar affinities (Fig. 41). Unexpectedly, four of those had a unique target either among the two kinases TGFBR1 (ALK5), RIPK2 or, more surprisingly, sepiapterin reductase (SPR).
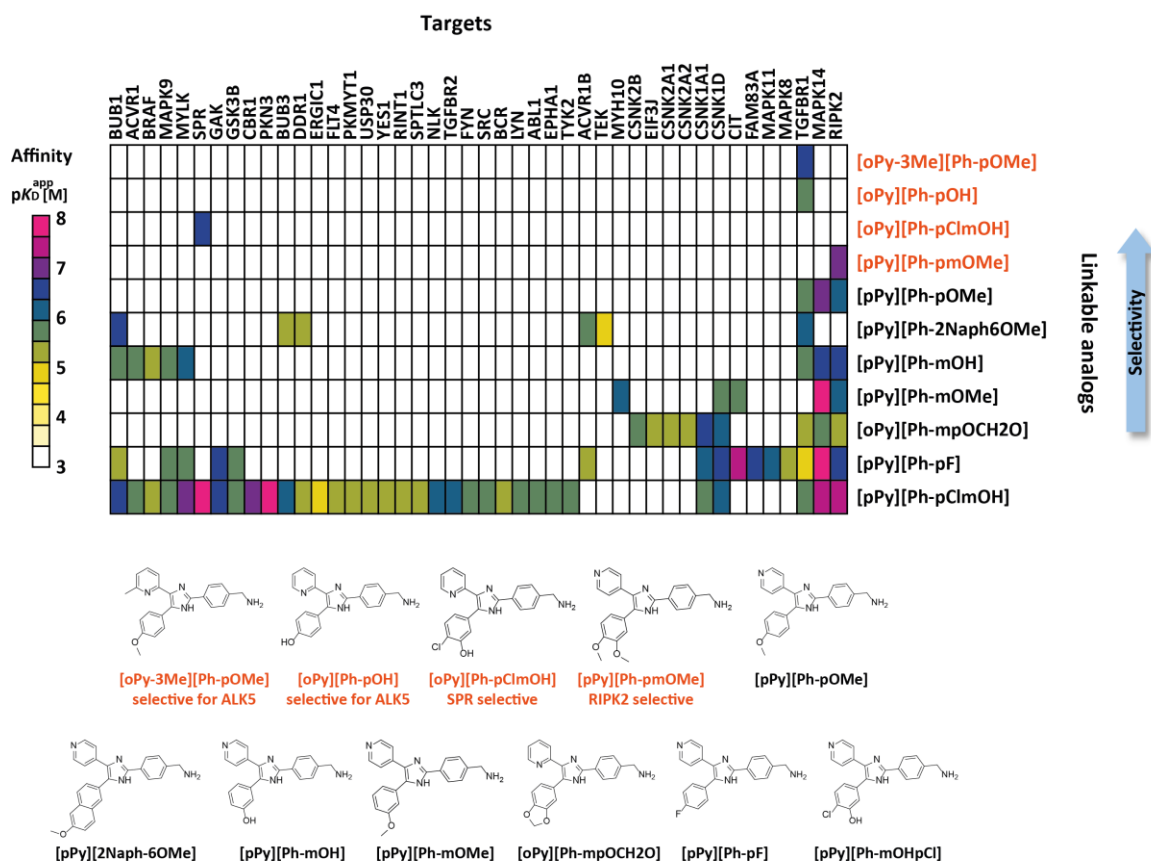
Figure 41. Heatmap showing the affinity (pK_D^{app}) of the 11 molecules with identified targets. Molecules marked in orange showed selective binding to one protein in this assay.

The profiling confirmed the importance of pyridine substituents on the aryl-rings since the analogs that did not contain any ([Ph] series) also failed to bind proteins in this assay specifically. From the range of commercially available inhibitors of this chemotype (Fig. 41), one may conclude that the pharmacophore is predefined to target MAPK14 (p38α) and ALK5 (TGFBR1). The newly obtained data confirms this statement, as 6 out of 11 active compounds were found to bind MAPK14, and 8 were binding ALK5. The following chapter discusses the structure-affinity relationships that could be established based on this profiling.

9.6 Structure-affinity and structure-selectivity relationships for ALK5

All but one profiled ortho-pyridine analogs demonstrated ALK5 binding (Fig. 41), which may define this substitution as vital for targeting ALK5. Ogunjimi et al. have investigated the structural basis of SB431542 (Fig. 34) selectivity towards ALK5 over other closely related kinases. They have superimposed the structure of the ALK2 kinase domain (PDB 3H9R) and the ALK5-SB431542 complex and identified the single residue Ser280, among all directly contacting the bound inhibitor, to be different between the two kinases (Thr in ALK2). Further mutation experiments (Ser to Thr) confirmed that Ser280 is a prerequisite for SB431542-mediated inhibition of ALK5.[153] The original publication did not directly propose that Ser residue is engaged in the hydrogen bond with the inhibitor; however, all the main findings indicate that this is likely the case. The two newly synthesized compounds that showed selective ALK5 binding here together with SB431542 (profiled against the linkable analog beads), namely [oPy-3Me][Ph-pOMe] and [oPy][Ph-pOH], are most likely adapting analogous interactions with the protein, where the ortho-pyridine is likely engaged in the hydrogen bond with Ser280.

Seven out of eight linkable analogs demonstrated reduced ALK5 binding affinities compared to SB431542 (K_D^{app} 0.6 µM, Fig. 42), which might potentially be explained by the loss of the stabilizing interactions via the amide group of the 2-aryl. Further substitution of an amine by a nitrile for the [oPy-3Me][Ph-pOMe] analog, which showed potent and selective ALK5 engagement (K_D^{app} 0.2 µM), led to a drastic drop in the binding affinity (>10 µM). The nitrile analog was profiled here against the [oPy-3Me][Ph-pOMe] tailored matrix and showed barely any competition for ALK5 binding at 30 µM, which further suggests that the interaction with the para substituents of the 2-aryl ring is indeed essential for potent ALK5 binding. The transition from the 2-aryls that are engaging into the stabilizing interaction to the ones that do not can thus be considered an affinity cliff for ALK5.

SB431542 was profiled here both against the corresponding tailored affinity matrix ([oPy][Ph-mpOCH$_2$O]NH$_2$) and Kinobeads, where it demonstrated selective ALK5 binding with affinity <1µM. Surprisingly, a linkable analog of SB431542 ([oPy][Ph-mpOCH$_2$O]NH$_2$) was showing a much weaker ALK5 binding affinity (>5 µM) along with MAPK14 and RIPK2 in the similar affinity range (CATDS$_{ALK5}$ 0.08, Fig. 42). This molecule was also able to engage five casein kinases, comprising two isoforms of casein kinase I and three isoforms of casein kinase II. Among all library analogs that were found to bind ALK5, this particular molecule demonstrated one of the weakest binding affinities. At the same time, its binding to CSNK1 was an order of magnitude more potent (Fig. 41). These findings suggest that the amide of the 2-aryl renders SB431542 not only potent but also more selective for ALK5. Additionally, all molecules with unique selective ALK5 binding (CATDS score of 1) possessed an ortho-pyridine on the 4-aryl, whereas the para-pyridine analogs engaged more protein targets at the ALK5 K_D^{app} concentration (lower CATDS scores). Finally, a transition from [oPy][Ph-mpOCH$_2$O] (linkable SB431542 analog) to [pPy][Ph-mpOCH$_2$O] produced a molecule deprived of any specific protein binding. Altogether, these findings suggest that the combination of the ortho-pyridine in the deep pocket and the amide of the 2-aryl at the pocket entrance predominantly drives potent and selective ALK5 binding. In that context, it would be particularly interesting to probe the amide versions of the two analogs that displayed selective ALK5 binding in this profiling.
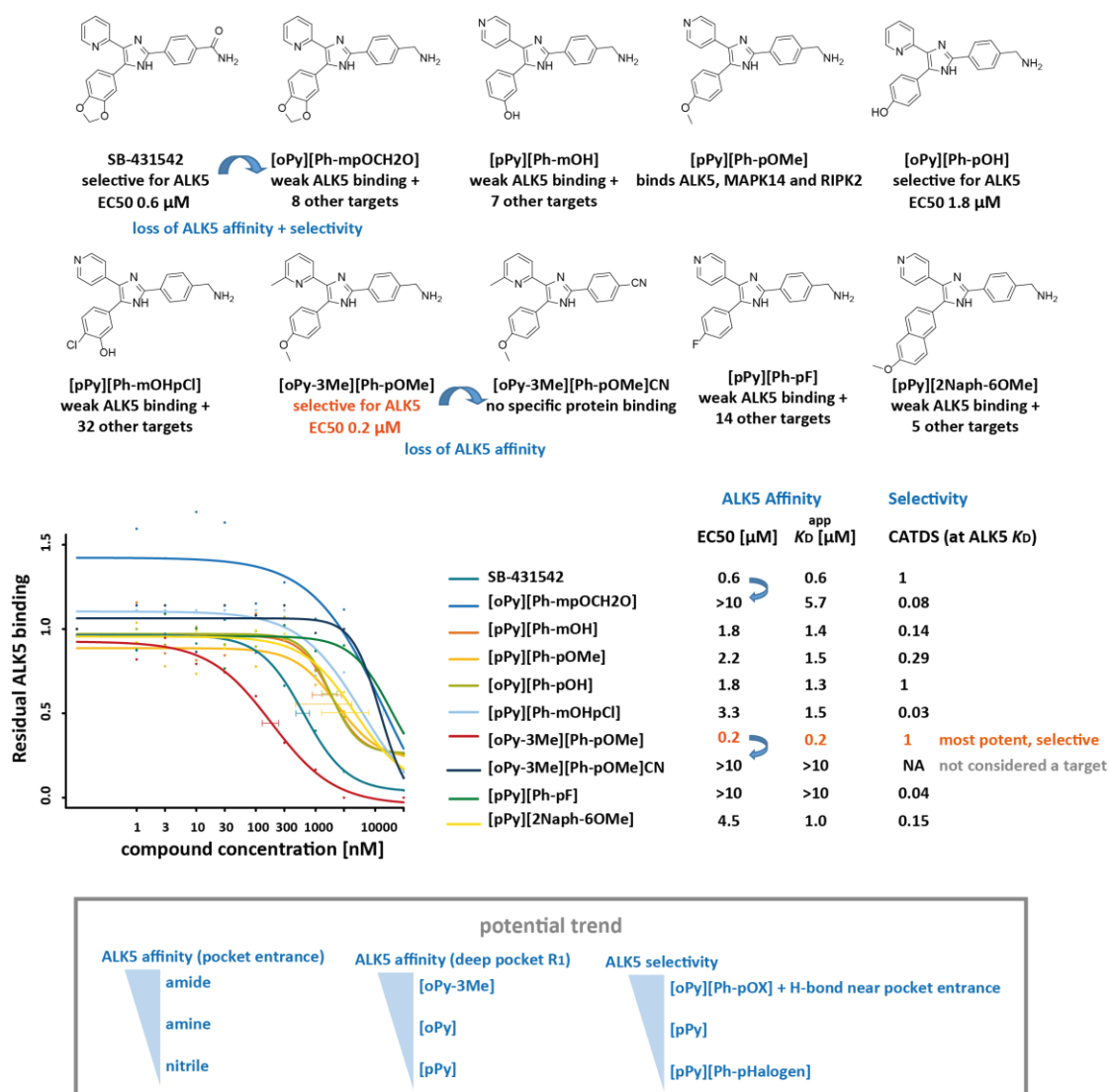
Figure 42. Binding of several analogs to ALK5 (TGFBR1). The [oPy-3Me][Ph-pOMe] analog demonstrates higher ALK5 binding affinity than SB-431542, interestingly 2-aryl substitutions of an amide with a primary amine and the latter with a nitrile lead to a drastic loss of ALK5 affinity (affinity cliff).

The analysis of the SB431542-ALK5 complex (PDB 3TZM) by three different structural platforms showed certain discrepancies (Fig. 43) in the proposed affine target-ligand interactions. Although all three tools identified a hydrogen bond between the oxygen of benzodioxolyl- and the His283, all other proposed interactions varied. For instance, the ligand binding tool from PDB suggested that the oxygen of the 2-aryl amide is engaged in the stabilizing interaction with Asp333 through a bridging water molecule, while the nitrogen of the same group is forming a hydrogen bond with Lys335 (Fig. 43a). In a meantime, ProteinPlus analysis reports no amide oxygen interaction and PLIP suggests that both oxygen and nitrogen of the amide are forming a hydrogen bond with the same Lys335 (Fig. 43b). The imidazole nitrogen, according to PDB tool, is stabilized with water, however the same nitrogen according to ProteinPlus is donating to the Lys232, while PLIP suggests both interactions: a hydrogen bond with Lys232 and a stabilizing water bridge to Asp351.[34,154] Most importantly, PLIP analysis solely indicates a formation of the hydrogen bond between the ortho-pyridine and the oxygen of Ser280 (Fig. 43c). This proposed interaction together with the report from Ogunjimi et. al. can potentially explain why the ortho-pyridine analogs in this study were able to potently and selectively engage into ALK5 binding.[153]
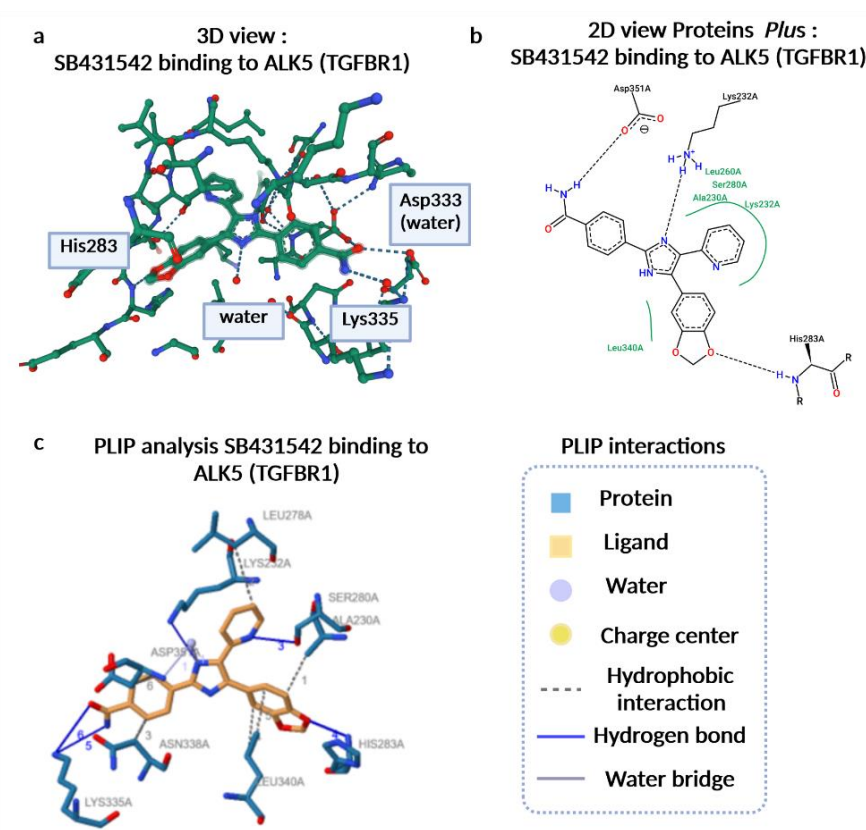
Figure 43. Analysis of the interactions between ALK5 and SB431542 based on the crystal structure PDB 3TZM. a) binding interactions according to PDB. b) analysis by proteinsPlus. c) same crystal structure analyzed in PLIP.

9.7 Selectivity of lophine pharmacophore

Calculation of CATDS scores offers an opportunity to investigate structure-selectivity relationships for the active lophine analogs and all of their identified protein binders (Fig. 44). Apart from the recapitulated apparent unique binding of the four molecules of either RIPK2, ALK5, or SPR, the other selectivity insights can be readily obtained from such an analysis. For instance, [pPy][Ph-pOMe] and [pPy][Ph-mOMe] engage MAPK14 with quite a good selectivity compared to other specifically bound proteins. Likewise, the [pPy][Ph-pClmOH] displays specific binding of 33 proteins; however, its engagement of PKN3 can be considered relatively selective at its respective K_D^{app} (Fig. 44). The exploration of the chemotype from both the selectivity and binding affinity perspective may potentially offer compound hits possessing potent engagement and desired selectivity trends towards particular targets of interest or a combination thereof. Further structure optimization efforts could then be applied to improve either of the desired molecule's protein binding properties.
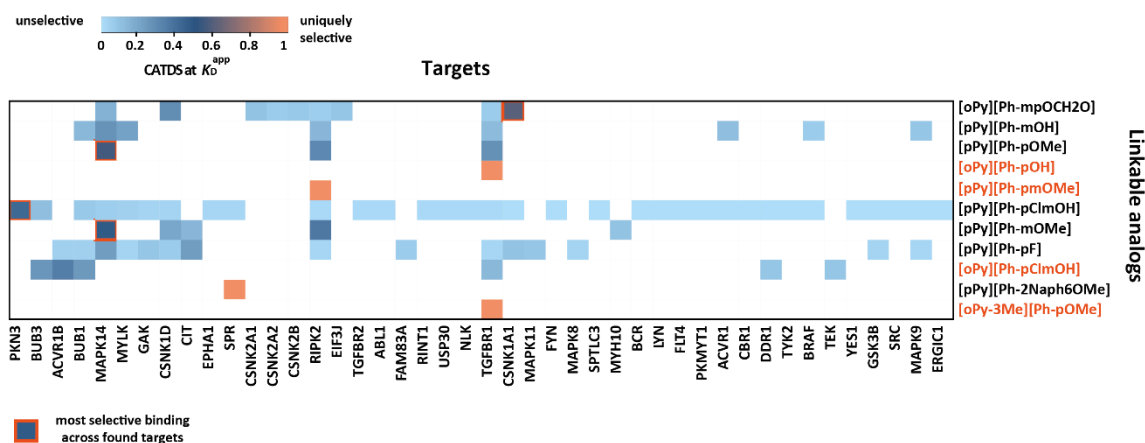
Figure 44. Selectivity map across 11 analogs that showed specific protein binding. Each cell represents a calculated CATDS score at the respective K_D^{app} of the target. Scores close to 1 indicate very selective binding.

For the linkable analog of adezmapimod ([pPy][Ph-pF]), the competition against the tailored matrix identified 15 targets with the most potent MAPK14 engagement (K_D^{app} 9 nM, CATDS$_{MAPK14}$ 0.25, Fig. 45). This profiling recapitulated all previously reported adezmapimod targets along with a few new interactors, namely BUB1, FAM83A, MYLK, and CIT. [144] The only other compound that showed similarly high MAPK14 binding affinity across five identified targets was the [pPy][Ph-mOMe] (K_D^{app} 15 nM, CATDS$_{MAPK14}$ 0.49), a methylated linkable analog of SB610251B (the structure can be found in Fig. 34). The binding, most likely, is analogously to adezmapimod is defined by the H-bond of para-pyridine and the amide of Met109, the other mOMe substituent is likely small enough to efficiently occupy the available hydrophobic space near the gatekeeper, while the imidazole nitrogen is engaged in the H-bond with Lys53. All other MAPK14-binding analogs showed decreased binding affinity, possibly due to the suboptimal orientation in the ATP-binding cleft. Moreover, only two ([pPy][2Naph-6OMe] and [pPy][Ph-mOMepOMe]) para-pyridine-containing analogs could not bind MAPK14, most likely due to the bulkiness of the second methylether- substituents.

According to the profiling results, para-pyridine appears to be not only crucial for MAPK14 but also for RIPK2 binding, as all of the [pPy] analogs engaged RIPK2 with a single exception of a molecule with a bulkier 5-aryl substituent, namely [2Naph-6OMe]. [pPy][Ph-mOHpCl] is essentially a close analog of SB203580, where para-fluoro is substituted by a para-chloro-group, and a hydroxy-group in meta position is added. These structural changes improve the affinity for RIPK2 from 336 nM (K_D^{app}) to 51 nM. However, it has to be noted that such substitutions of the 5-aryl apparently have a significant impact on the compound selectivity, as for this molecule, the profiling assay identified 33 targets (Fig. 45). Interestingly, this particular analog appeared to be the sole and potent PKN3 binder (K_D^{app} 16 nM) within the profiled library. A one-step pyridine substitution from [pPy][Ph-mOHpCl] to [oPy][Ph-mOHpCl] led to a drastic gain in selectivity (the pair of compounds can be considered a selectivity cliff), where the sole target identified was a non-kinase protein sepiapterin reductase (SPR, EC50 308 nM, K_D^{app} 207 nM) (Fig. 45). This target was also found for the former analog (EC50 224 nM, K_D^{app} 15 nM) among the 33 specifically bound proteins, indicating that for SPR binding the pyridine position likely plays a minor role. The substitution to ortho-pyridine rather allows the elimination of the targets, for which the interaction with the para-pyridine is vital. The removal of the para-chloro substituent ([pPy][Ph-

mOH] analog) rendered the molecule much more selective (6 targets) at the cost of slightly decreased binding affinities for MAP14 and RIPK2. The substitution of meta-hydroxy by meta-methyl further slightly decreased the RIPK2 affinity (from K_D^{app} 182 nM to 431 nM) while restoring the potent MAPK14 binding (K_D^{app} 15 nM), and the molecule specifically bound in total 5 protein targets. The next analog with the single para-methyl ether ([pPy][Ph-pOMe]) showed binding MAPK14 (120 nM) along with RIPK2 (870 nM) and ALK5 (1.5 µM). Here the switch from para-methyl ether ([pPy][Ph-pOMe]) to meta-methyl ether ([pPy][Ph-mOMe]) led to a complete loss of ALK5 binding (Fig. 45), whereas a further addition of the para-methyl ether rendered the molecule completely selective for RIPK2 (K_D^{app} 95 nM).
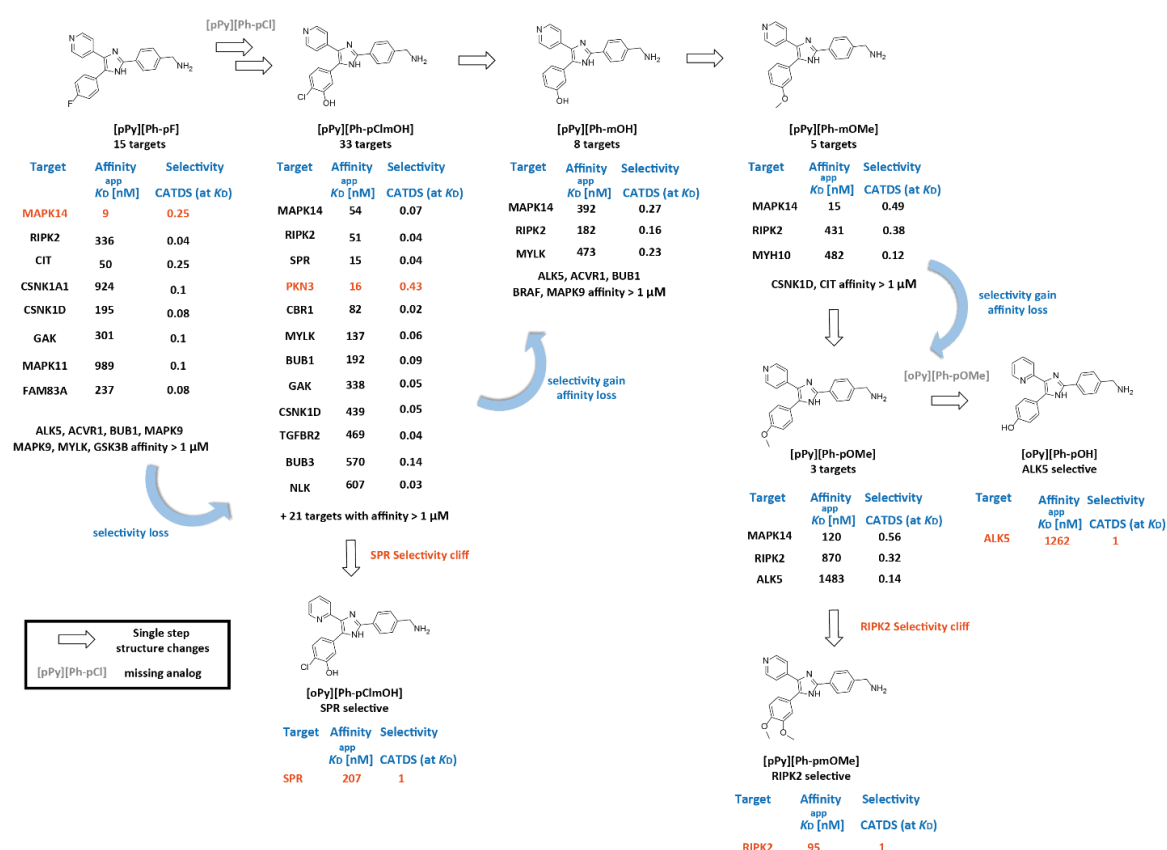
Figure 45. The selectivity and affinity changes along the one-step structure substitutions.

These results demonstrate that even within such a small set of structurally similar and quite selective compounds, the selectivity trends of each analog cannot be directly anticipated. However, here the single-step substitutions allow for the establishment of structure-affinity and structure-selectivity relationships (represented here as CATDS scores), unraveling potential novel chemical probes or molecules with improved selectivity towards a potentially desired binding profile.

9.8 Correlation of affinity and enrichment allows conceiving proteome-wide screening

In order to compare protein enrichment across all 35 SAR affinity matrices, the triplicate pulldowns were performed for each immobilized molecule. This comparison is possible for all bead-bound proteins (~2000 proteins quantified in each pulldown), including the targets found by the competition experiments (e.g., MAPK14, RIPK2, SPR, and ALK5/TGFBR1) or the ones that did not appear as bound by the free molecules such as BRD4 (Fig. 46a). For SPR, i19 and i28 stood out as the best matrices to enrich the protein and i28 did not enrich any of the other four selected proteins, while i19 also enriched MAPK14, RIPK2, and TGFR1/ALK5 at high levels. The high protein enrichment was in correlation with the potency and selectivity of the molecules in the competitive experiment that identified #28 ([oPy][Ph-mOHpCl]) with SPR as its only target, while #19 ([pPy][Ph-mOHpCl]) was rather unselective and engaged 33 proteins in total (Fig. 45).
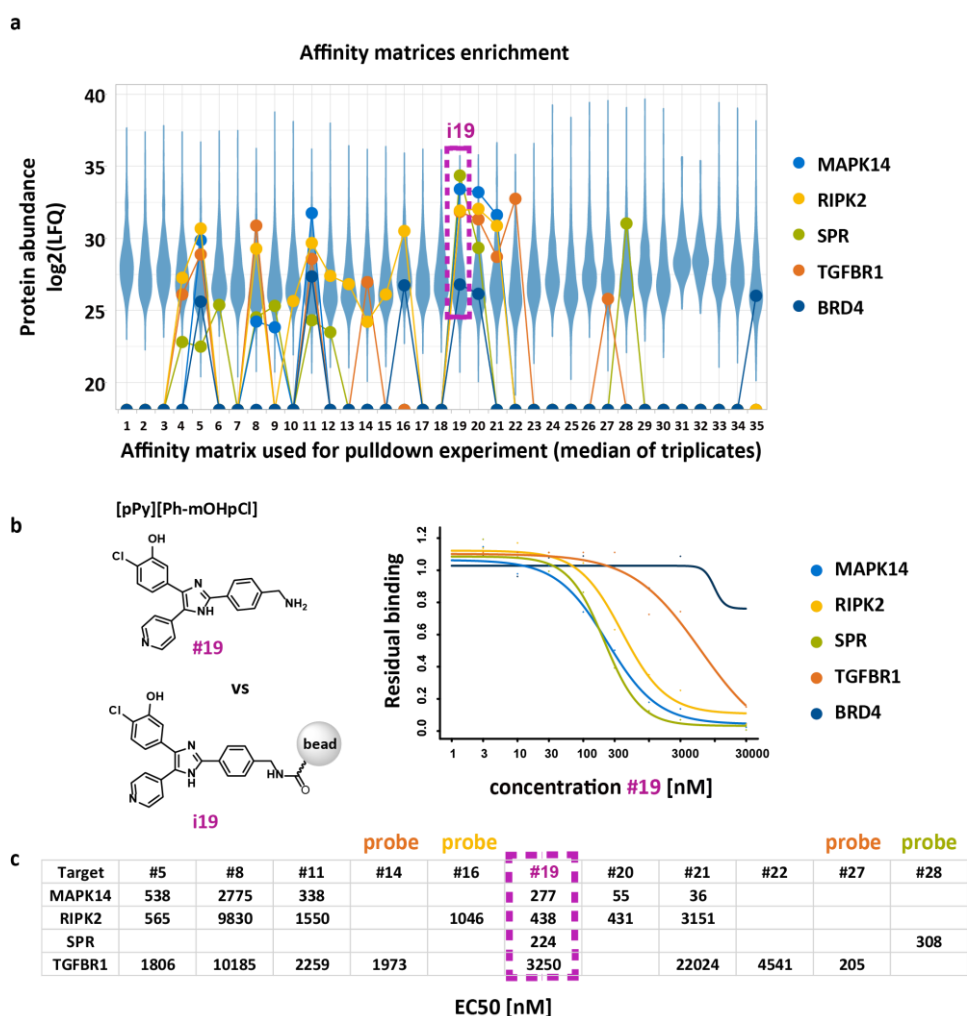
Figure 46. Differential enrichment of proteins determined by triplicate pulldowns by each of the 35 members of the affinity matrices library is indicative of the targets found by competitive target deconvolution experiments. a) Profile plot showing the median mass-spectrometry intensities of five selected proteins (MAPK14, RIPK2, SPR, TGFBR1/ALK5, and BRD4) after the triplicate pulldowns by each of the 35 affinity matrices marked in color on top of the intensities of all other proteins quantified in the pulldown, distributed in the respective blue violin. b) Exemplary dose-response curves obtained by competitive target deconvolution experiment of one (#19 = [pPy][Ph-mOHpCl] aka linkable SB476429-A) of the linkable SAR library analogs. c) Observed binding inhibitions for molecule #19 detailed in panel b compared to the other 10 molecules whose target deconvolution has at least revealed one target for the sub-selection of targets MAPK14, RIPK2, SPR, and TGFBR1/ALK5.

The other three compounds (i14, 16, and 27) that showed selective competition for one of the selected proteins demonstrated high corresponding protein enrichment in triplicate pulldowns. For less selective compounds, e.g., i11, the protein enrichment also corresponded to the obtained affinities in competition experiments. For a particular case of i11, the MAPK14 was one of the most abundant proteins in the triplicate pulldowns, which correlated with the highest binding affinity to this kinase (EC50 338 nM) in the full dose competition profiling, followed by RIPK2 (1.5 µM) and TGFBR1/ALK5 (2.2 µM) (Fig. 46c). The results of this profiling suggest that high protein enrichment by the tailored affinity matrix can be indicative of high binding affinity and potential target engagement by the analyzed molecule.

The possible cross-reactivity of kinase inhibitors with BET bromodomain-containing proteins has been previously explored for multiple inhibitors of the lophine chemotype. For instance, a study by Ember et al. describes an extensive co-crystallization screening campaign, where kinase inhibitor libraries were crystallized with the first bromodomain of BRD4.[155] The authors report that 14 KIs with 10 distinct chemical scaffolds can act as potential ligands of BRD4-1. Among those, the reported crystal structure of SB610251-B (non-linkable lophine analog with [pPy][Ph-mOH]) in BRD4-1 shows that the nitrogen of the para-pyridine binds a key water molecule closely interacting with Tyr 97 in the protein pocket (PDB 4O7E, Fig. 47d). This same water is bound by the nitrogen of the 3,5-dimethylisoxazole motif found in many BET inhibitors, notably IBET-151 (PDB 3ZYU, not shown here) or one of the nitrogens of the triazole motif found in JQ-1 (PDB 3MXF, not shown here). The PLIP analysis suggests that the SB10251-B interaction is further stabilized via the hydrogen bond of the pyridine nitrogen with Asn140 (Fig. 47d). Ember et al. equally calculated the inhibitory activity and binding potential for identified BRD4-1 interactors. The authors reported the IC50 of 5.7 µM for SB610251-B and 2.5 µM for SB202190 (contains pOH group on 2-Ph compared to linkable analog from the lophine library here), which was not supported by the competition profiling data obtained in this work (Fig. 47c). However, the triplicate enrichment by the SAR library matrices indicate that the [pPy][Ph-mOHpCl] substitution (compound #11 on Fig. 46a) is one of the most suitable for BRD4 binding in the para-pyridine series.

Despite the reported off-target binding and X-ray structures of the chemotype (original studies feature primarily adezmapimod and its closest analogs) with BRD4, none of the BET proteins have been found to bind any of the lophines analog molecules with affinities <30 µM in the chemoproteomic experiments here.[140] In order to confirm that the analog immobilization or pocket entrance substitutions do not affect bromodomain-containing protein binding and ensure stable bromodomain-proteins enrichment, a BET-bromodomain inhibitor I-BET726 was functionalized on beads and added to adezmapimod-analog affinity matrix (Fig. 47a). This mixed affinity matrix was used in a competition profiling of bromosporine along with adezmapimod-type (another name - SB203580) compounds synthesized in the course of this study. In competition against the mixed matrix, bromosporine showed the expected dose-response curves for various bromodomain-containing proteins, whereas none of the other SB203580-analog compounds displayed any BET competition within the tested concentration range (up to 1 mM here) (Fig. 47b,c). Overall, there were 117 bromosporine targets identified in this competition assay. Interestingly, this broad-selective bromodomain inhibitor showed competitive binding to several kinases that were identified as targets of the lophine chemotype, e.g. GAK, BUB1, GSK3B, MAPK8, and ALK5/TGFBR1 (Fig. 47c).
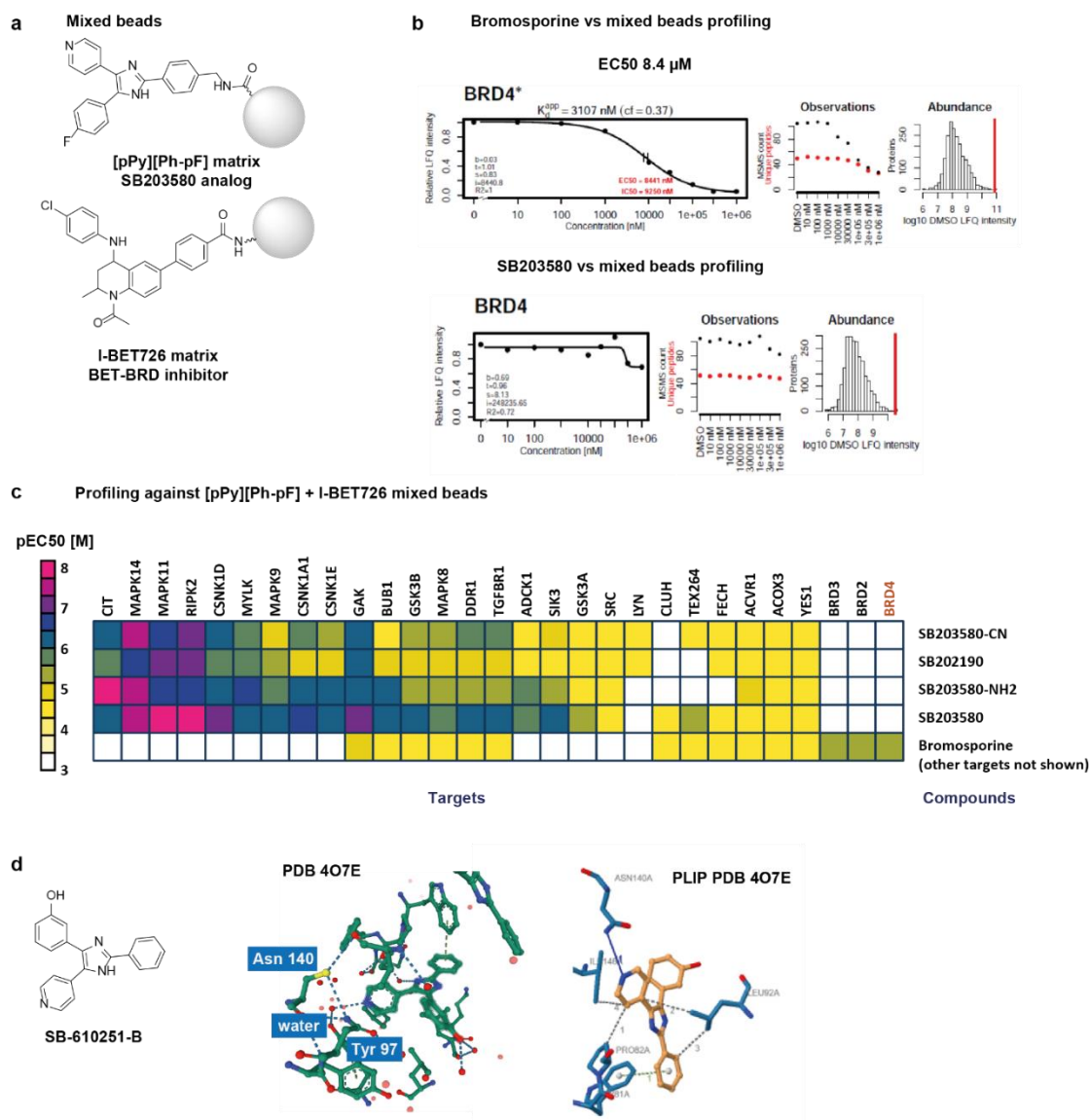
Figure 47. Cross-reactivity of lophine chemotype with BET-bromodomains. a) Mixed affinity matrix utilized for profiling. b) comparison of BRD4 dose-response competition curves in the profiling against mixed beads from panel a. c) Heatmap of targets identified in the competitive profiling of adezmapimod analogs (SB203580) and bromosporine. Only the latter demonstrated selective BET-BRD binding. d) Analysis of SB610251-b binding with BRD4-1.

It is possible that bromodomain-containing proteins, similarly to HDACs, possess slower binding kinetics to the inhibitors evaluated in this work than kinases.[58] This could potentially explain why the competition assay here could not identify BRDs as targets of the lophine analogs as well as why bromosporine showed such low binding affinities towards its designated targets. However, further experiments would have to be performed to evaluate whether this hypothesis holds.

9.9 **Pocket entrance study of the [pPy][Ph-pF] pharmacophore**

A range of commercially available p38 inhibitors contain the [pPy][Ph-pF] substitutions on the 4- and 5-aryls of the imidazole. In order to complement the two molecules made in the course of this study (the nitrile intermediate and the immobilizable methanamine), 7 of those molecules were purchased, and the 9 inhibitors were then profiled against their common affinity matrix (Fig. 48).

This experiment served to extend and generalize the previously discussed analysis for ALK5 and establish an entrance pocket SAR for [pPy][Ph-pF] pharmacophore, by deciphering the influence of the 2-substitutions of the imidazole on the selectivity and affinity for protein targets. Under the assumption that the 4- and 5-aryl substituents are more deterministic of target selectivity for this chemical scaffold, the immobilized molecule used in the deep-pocket probing should be capable of enriching most of the targets that these non-linkable molecules would engage. Thus, these analogs do not need to be immobilized, allowing evaluation of whether 2-substitutions may fine-tune the affinity and/or selectivity towards specific targets enriched by the common analogous affinity matrix. The [pPy][Ph-pF] affinity matrix bound 72 kinases across more than three thousand quantified proteins. Although only a fraction of those was confirmed to bind the adezmapimod analog in the competition experiment truly specifically, the 72 kinases represent the entire assayed kinase target repertoire.
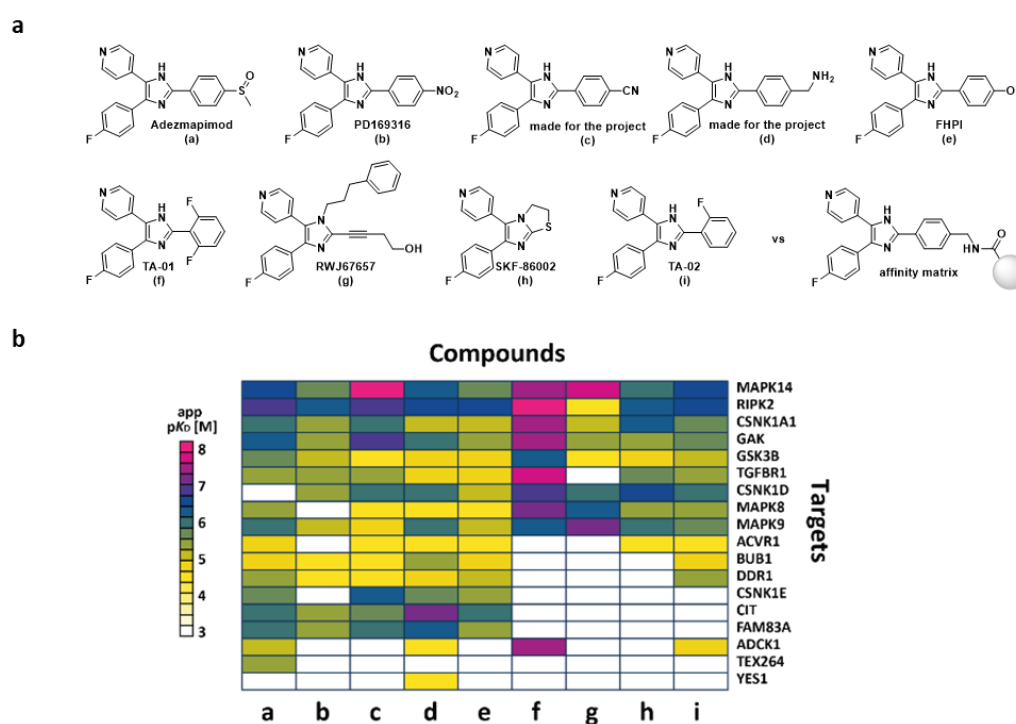
Figure 48. Pocket entrance SAR. a) Molecules used for the pocket entrance study. b) Nine molecules with the same 4/5 substitutions were profiled against the identical affinity matrix to generate the SAR of the entrance pockets of the targets.

In total, 18 different proteins were bound specifically by those 9 molecules (across the panel of 72 kinases and more than three thousand bead-bound proteins). Similarly to the case of ALK5 mentioned earlier, the 2-imidazole substitutions showed strong effects on the binding affinities of the molecules. Overall, the compound selectivity was affected to a lesser extent as no new targets were identified, and compounds f-i lost several weaker binders compared to a-e. The CATDS selectivity scores for all of the compounds profiled here against all identified targets help reveal the most selective binders for individual proteins (Fig. 49a). As expected from the initial deep pocket exploration, the ability to engage MAPK14 and RIPK2 was not compromised; however, the molecules showed clear differences in apparent binding constants. For instance, the analog that demonstrated the highest MAPK14 binding affinity was nitrile c, while the analog that showed the best selectivity for MAPK14 was compound g, with the highest difference in binding affinities between MAPK14 and other targets. RWJ67657 (g) differs from the other analyzed molecules by

the lack of the 2-aryl, which is replaced by a 2-butynol, and the additional 1-phenylpropyl substituent. The latter indicates that the loss of stabilizing hydrogen bond from the imidazole ring and the π-π stacking of the 2-aryl does not affect MAPK14 binding but helps to lower the affinity towards other binders like RIPK2, CSNK1A1, and GAK. For analogs b, e, and h MAPK14 was by far not the most potent binder, possibly due to their inability to form a stabilizing hydrogen bond interaction from the 2-imidazole substituent (-NO2 cannot form H-bonds, -OH possibly too far in the distance, same for h). These results suggest compound g as the best probe for MAPK14 engagement among all analyzed molecules within the chemotype.
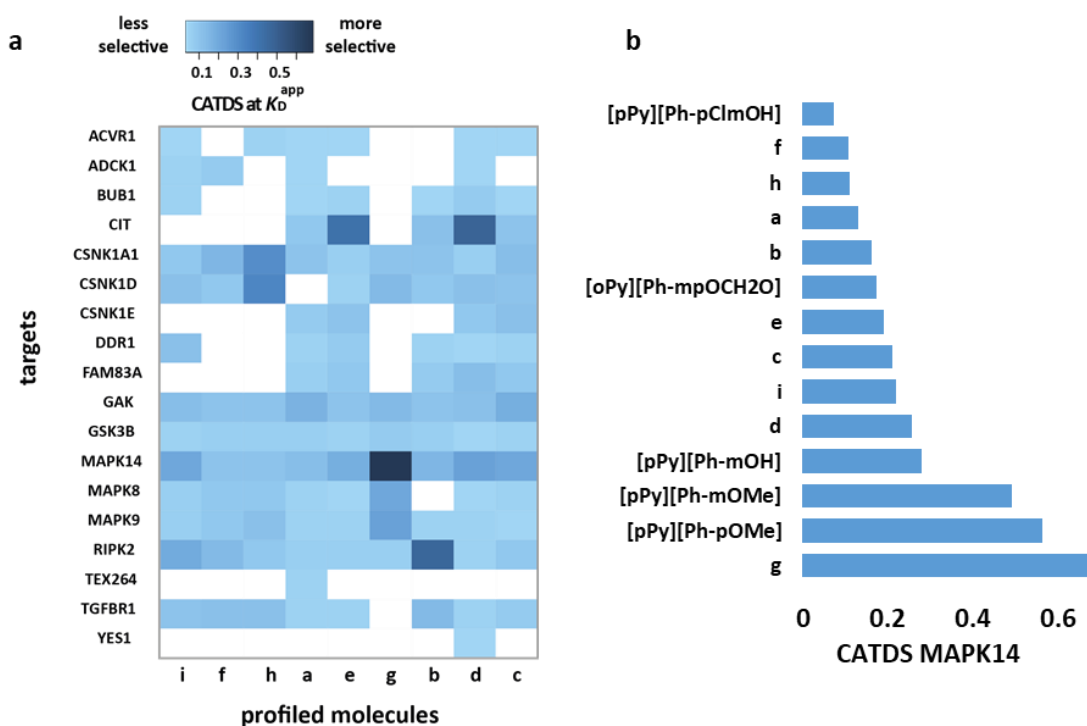
Figure 49. Pocket entrance structure-selectivity relationships. a) CATDS scores across the profiled molecules for all found targets b) comparison of the MAPK14 CATDS scores for the molecules in this profiling as well as previous deep-pocket SAR study.

Surprisingly, TA-01 (f) showed a very potent competition of all of its 10 binders, with overall at least twice higher K_D^{app} values compared to its closest analog TA-02 (i), which only differs by a single fluoro-substituent of the 2-aryl, and other molecules within the set. Although this compound is a commercially available inhibitor of casein kinase 1 and p38 (MAPK14), it demonstrated the most potent binding to RIPK2, as well its binding affinity to ALK5 was the highest among all analogs here, and all 35 linkable compounds from the deep pocket exploration profiling. The compound, however, appears quite unselective for individual targets, as it engages all of its binders within a similar affinity range (Fig. 49a). The difference in the obtained binding affinities for the compounds f and i are likely not the results of the experimental mistake, as all the discussed profiles in this chapter were conducted in parallel on the same day using the same batch of lysate. It can be hypothesized that the TA-01, unlike TA-02, is potentially unable to stabilize its binding interactions with the protein in case of the 2-aryl rotation. Finally, the target profiles of compounds a-e and I, that feature identical 4- and 5-aryls, were largely conserved, indicating that these binders can tolerate a range of 2-aryl substituents.

These results recapitulate that the target repertoires of the lophines chemotype analogs are much more influenced by the structural changes of 4- and 5-aryls. Then, the pocket entrance modifications may help shape the compound affinity towards a particular protein, whose binding is defined by the substituents reaching deep into the protein pocket. Such an experiment enabled obtaining a valuable target SAR (rows) along with the selectivity of each individual compound (columns of the heatmap) in the context of entrance pocket modifications. This approach may become particularly advantageous on the later stages of drug development, where it would allow for simultaneous profiling of drug candidates selectivity and fine-tuning the binding affinity towards a set of desired protein targets using a single appropriate tailored affinity matrix.

9.10 Mixed chemotype affinity matrix for profiling of non-linkable analogs

Sub-proteome enriching affinity matrices can be employed for selectivity profiling of drugs, for example, previously mentioned Kinobeads and HDACs beads.[57,58] In the context of proteome-wide SAR, a chemotype-targets enriching affinity matrix should allow for profiling of direct and indirect binders of chemotype analogs that can not be directly immobilized on beads or further help interrogate the influence of pocket entrance substitutions. In other words, by combining the tailored probes that should stably enrich all potential pharmacophore targets as a mixed affinity matrix, it should become possible to extend the chemotype deep pocket and pocket entrance SAR on any novel non-linkable analogs within the studied chemotype (any novel 2-, 4-, 5- substituents). In this study, the mixed chemotype affinity matrix was prepared from the three amino-analogs with the most complementary target profiles recapitulated in the original deep-pocket selectivity profiling (Fig. 41, 50), as they specifically enriched all of the 44 targets of the chemotype. This matrix was employed to probe several commercially available kinase inhibitors for off-target binding as well as to analyze the target space of the lophine-nitriles for which the amino-analogs could not be isolated in sufficient purity (Fig. 50a).

In total, ten different protein targets were identified for eight analyzed molecules. The three nitrile analogs showed micromolar binding affinities to MAPK14 and RIPK2, and each displayed additional binding of either MAPK8, GAK, MAPK9, or CIT, further expanding the pharmacophore SAR. Interestingly, all 8 compounds bound both MAPK14 and RIPK2, indicating that these two proteins must have very similar ATP-binding pockets, and highlights the difficulty of designing selective inhibitors for either of the two kinases. The commercially available imidazoles profiled here featured different substituents than the original lophines analogs. Here the 2-aryls were either replaced by a neopentane or the 3-N of the imidazole core featuring a methylcyclohexanol- or methylpiperidine- substituents. Most of these compounds are marketed as ALK5 selective inhibitors. Unfortunately, ALK5 was not stably enriched by the mixed affinity matrix, likely due to the lower ALK5 expression in the newly prepared batch of lysate and the simultaneous decrease in compound loading on beads. Here the affinity matrix was prepared with twice lower coupling density of the compound loading (1 µmol/mL vs. 2 µmol/mL) with expectations to lower protein depletion by the beads, as the original SAR selectivity profiling demonstrated a relatively high effect of protein depletion and correction factors close to zero. Nevertheless, other library targets were stably enriched in the new lysate, which allowed for off-target profiling of commercial molecules. The transition from 3- methylcyclohexanol (SB 239063) of the most unselective compound in this dataset to 3-methylpiperidine (SB 242235) increased the selectivity and led to a

loss of seven protein binders while preserving the affinity for MAPK14, RIPK2, and CSNK1D. SB 525334 and SB 505124 demonstrated the same binding profiles, where the RIPK2 was the most affine binder of both molecules. Provided that at least half of these molecules are marketed as selective ALK5 inhibitors, it is essential to note that chemoproteomic profiling here demonstrates that none of these molecules is purely selective, and all bind other disease-relevant kinases.
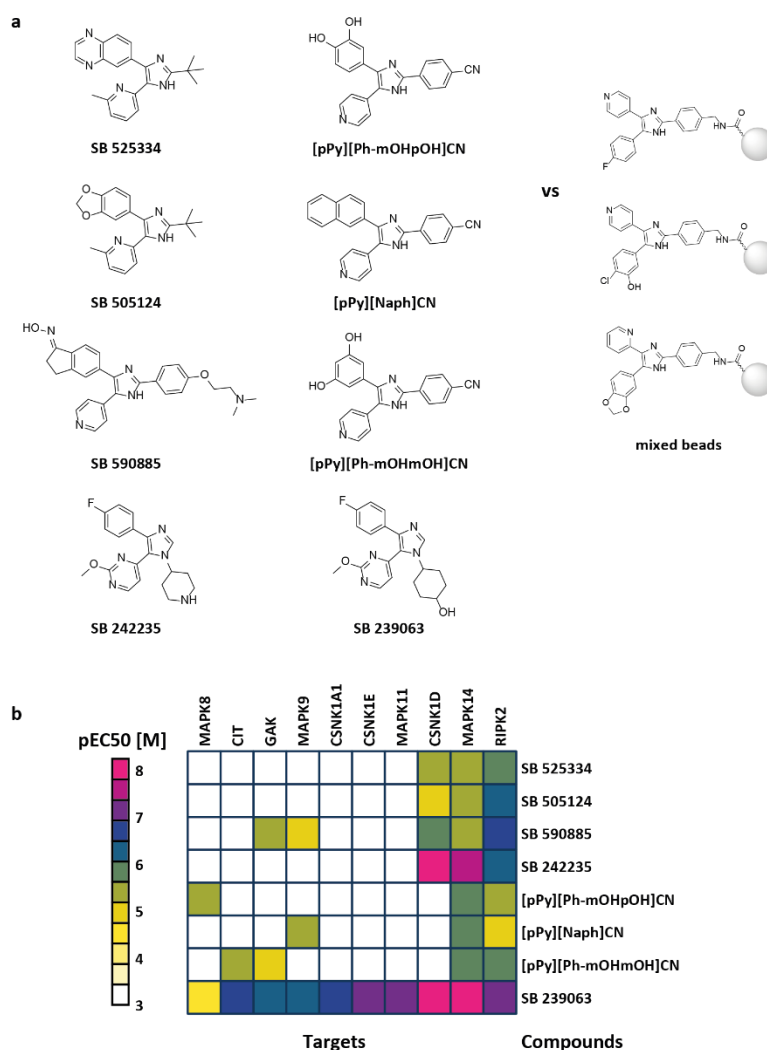
Figure 50. Mixed beads for chemotype analogs profiling. a) Several non-linkable compounds of lophine chemotype profiled against mixed chemotype beads. The mixed beads were prepared by combining three linkable SAR library analogs with complementary target profiles. b) Target heatmap, where the color of cells indicates the strength of observed binding (pEC50).

9.11 **From chemoproteomic binding assay to cell-active leads**

9.11.1 Novel lophine library analogs engage ALK5 in cellulo

Transforming growth factor β (TGFβ) plays a crucial role in cell proliferation, differentiation, and matrix formation. TGFβ binds to various activin-receptor-like kinase (ALK) receptors, which induces phosphorylation of downstream SMAD proteins. The latter form complexes with SMAD4 and translocate to the nucleus, mediating gene transcription. TGFβ was also shown to activate various SMAD-independent downstream signaling pathways, such as p38 or JNK (Fig. 51).[156–158]
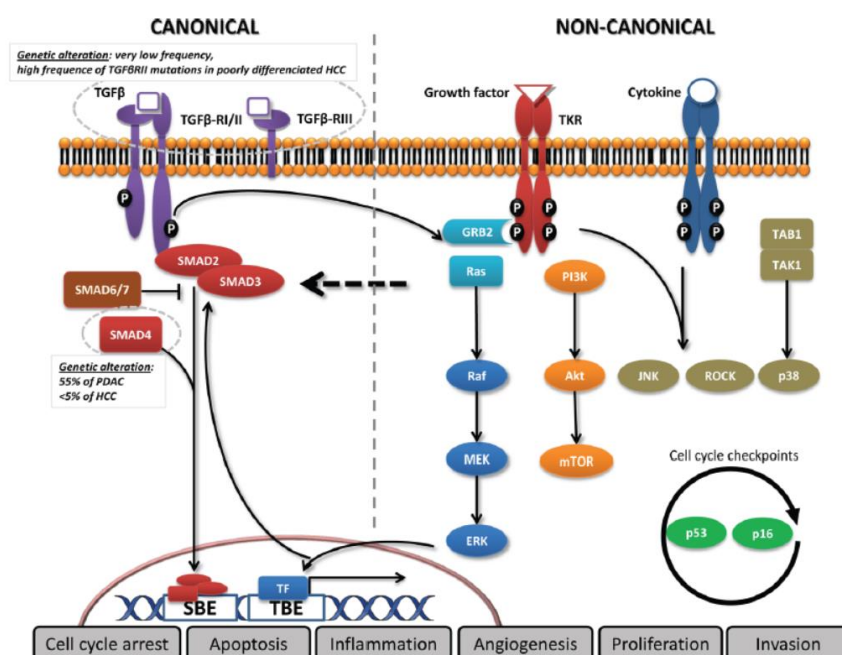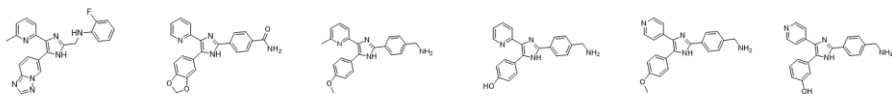
Figure 51. TGFβ signaling. Adapted from Neuzillet et. al.[158]

TGFβ paradoxically acts as both a tumor suppressor and tumor promoter.[159,160] It has been shown to inhibit the proliferation of epithelial and lymphoid cells and induce apoptosis. Generally, TGFβ signaling to the cell cycle in the context of tumor suppression is primarily mediated via activation of the SMAD pathway and subsequent induction of p21 and/or p27. The serine/threonine kinase TGFBR1 (ALK5) transduces TGFβ signals by phosphorylating SMAD2 at Ser465 and Ser467, mediating the growth inhibitory effect of TGFβ in various cancer cells. At the same time, TGFβ was proven to enhance epithelial-to-mesenchymal transition (EMT), often associated with enhanced tumor invasiveness and metastasis. Additionally, it was reported that late-stage human carcinomas often become resistant to TGFβ growth inhibition due to certain defects in the TGFβ signaling pathway. For instance, TGFβ was shown to activate the second major signaling cascade comprising MAP kinases (ERK, JNK, p38) to promote tumor angiogenesis and metastasis in the orthotopic mouse xenograft model. In prostate and ovarian cancer cells, overexpressed TGFβ stimulates DNA methyltransferases expression (DNMT), associated with an aggressive phenotype and poor prognosis. In clinical evaluation, in early-stage tumors or non-diseased conditions, low concentrations of TGFβ promote proliferation, migration, and expression of matrix metalloproteinases (MMPs) in endothelial cells through ALK1 activation of SMAD1/5, resulting in ID1 transcription. In high concentrations, the canonical SMAD2/3 pathway is activated by ALK5, inducing the expression of plasminogen activator inhibitor and fibronectin, thereby promoting angiogenesis and resulting in tumor progression. This suggests that any antiangiogenic effects of

TGFβ are overridden in advanced disease and highlights the importance of ALK5 as a therapeutic target.[161–164] Several small molecule ALK5 inhibitors were clinically evaluated to treat various cancers. Prime examples of such drugs are galunisertib (rectal adenocarcinoma, hepatocellular carcinoma, non-small cell lung cancer) and vactosertib (myeloproliferative neoplasm, osteosarcoma, multiple myeloma). The former molecule developed by Eli Lily was discontinued in 2020 for unlisted reasons. Interestingly, both drugs are commercially available for general research and are marketed as selective ALK5 inhibitors. However, chemoproteomic profiling of both drugs against Kinobeads revealed that each of them specifically bound more than 10 other kinases apart from ALK5, including GAK and RIPK2.[30] Uniquely selective ALK5 probes in this context might help analyze whether the sole inhibition of ALK5 is potentially beneficial for the treatment outcome, while comparison to other less selective inhibitors may explain the contribution of off-target binding to the observed drug effects.

In order to confirm that newly obtained selective ALK5 binders in this study were able to engage ALK5 in live cells, the levels of SMAD2 phosphorylation relative to the total amount of SMAD2 upon treatment were monitored by western blot. Here HeLa cells were treated with increasing concentrations of ALK5-binding compounds (4 in total) from the lophine library as well as SB431542, and the readout was performed with p-SMAD2 and its non-phosphorylated counterpart antibodies (table 5).

Compound name	Vactosertib	SB431542	[oPy-3Me][Ph-pOMe]	[oPy][Ph-pOH]	[pPy][Ph-pOMe]	[pPy][Ph-mOH]
ALK5 affinity (chemoproteomics assay), µM	0.014	0.63	0.18	2.0	2.2	1.8
Other targets	ACVR1B, RIPK2, GAK, TFFBR2, SIK3, FLT4…	none	none	none	MAPK14, RIPK2	MAPK14, RIPK2, MYLK, MAPK9, BRAF, ACVR1, BUB1
SMAD2 phosphorylation inhibition in HeLa cells (EC50, µM)	-	0.45	0.35	0.88	10	13

Table 5. Comparison of four novel ALK5 binders against known TGFBR1 inhibitors Vactosertib and SB431542. The affinity for ALK5 and the list of other targets were determined by chemoproteomics competition assay (against Kinobeads for Vactosertib, its immobilized methanamine analog for SB431542). Cellular inhibition of SMAD2 phosphorylation were measured using western blots (anti dual pSer465/467 antibody) after HeLa cell treatment.

Treatment of HeLa cells with the four most potent binders within the lophine SAR library or SB431542 indeed reduced SMAD2 phosphorylation, as evidenced by western blot (Fig. 52a-c). [oPy-3Me][Ph-pOMe] analog demonstrated the *in-cellulo* EC50 of 350 nM, thus proving itself as a suitable chemical probe for TGFBR1. While less potent than Vactosertib in the chemoproteomic assay in lysate, it is uniquely selective (Vactosertib was found to bind 14 other kinases in the Kinobeads competition profiling) and can engage its target *in-cellulo* with the same submicromolar potency as SB431542 (Table 5, Fig. 52d). Other three compounds demonstrated less potent inhibition of SMAD2 phosphorylation, and were able to reduce pSMAD2 levels by at least 65% at the highest dose. At 1 µM dose, [oPy-3Me][Ph-pOMe] reduced pSMAD2 level by more than 80%, and at 10 µM, only 1% of initial pSMAD2 could be detected. At the highest dose (100 µM), pSMAD2

was abolished entirely. Other tested molecules underperformed in comparison and SB431542 reduced 70% of pSMAD2 at 1 µM treatment (Fig. 52a).
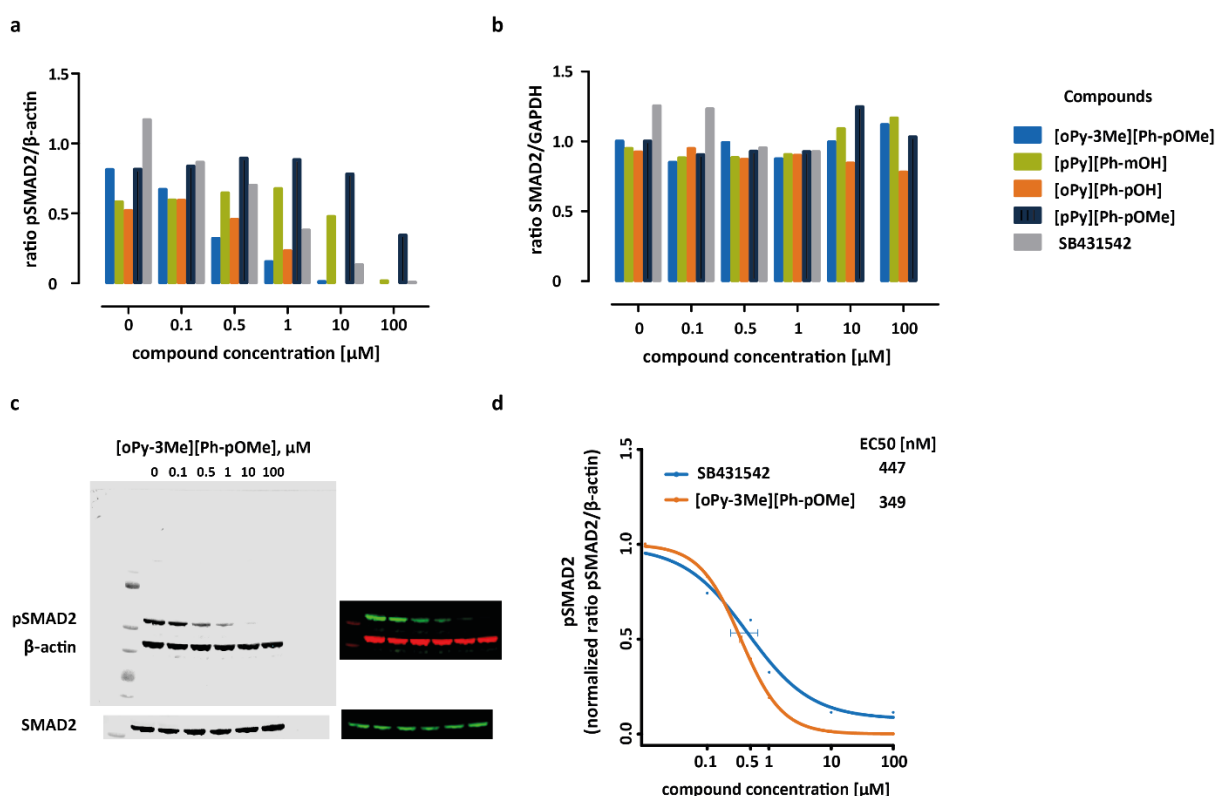
Figure 52. Western blots of HeLa cells treatment with ALK5 binders. a) Quantification of Western blot bands for pSMAD2 relative to b-actin loading control. Quantification was performed using Image J software. b) Same as panel a but for non-phosphorylated SMAD2. Overall, SMAD2 levels are not affected by the compound treatment. c) Full blot image of the [oPy-3Me][Ph-pOMe] analog treatment. d) Normalized quantified bands from treatments with [oPy-3Me][Ph-pOMe] analog and SB431542 plotted in dose response format for comparison. Selective SAR library analog does not underperform compared to commercially available inhibitor.

To sum up, this set of experiments prescribes [oPy-3Me][Ph-pOMe] analog as a suitable chemical probe, additional to SB431542, for ALK5. It is less potent than vactosertib but is uniquely selective and can engage its target in-cellulo with submicromolar potency.

9.11.2 Sepiapterin reductase (SPR) selective binder is indeed an inhibitor

Two analog molecules of the lophine library showed potent binding to sepiapterin reductase (SPR), a non-kinase terminal enzyme in the biosynthesis of tetrahydrobiopterin (BH4) (Fig. 53a). The latter is an essential cofactor for the synthesis of monoamine neurotransmitters, like dopamine and serotonin, and nitric oxide species.[165] SPR is associated with several diseases, including chronic pain, brain dysfunction, and cancer. BH4 deficiency in the context of brain dysfunction, for example, is associated with neurotransmitter-responsive disorders, such as movement and muscle tone impairment, motor dysfunction, epileptic seizures, and mental retardation. The role of BH4 and SPR in cancer is still debatable; however, several studies have demonstrated that SPR inhibition decreases the proliferation of neuroblastoma (NB) cells and growth inhibition of NB tumors in-vivo.[166,167] It has also been shown that SPR is essential for the proliferation of mature T-cells and that its inhibition links to immunosuppressive tumor

environment.[168] Nonetheless, by far the most focus the enzyme has received after the discovery of its role in neuropathic pain. For instance, several neuropathic pain models showed high upregulation of SPR and other enzymes in the BH4 biosynthetic pathway within injured sensory neurons. Here targeted inhibition of SPR significantly reduced chronic pain by inhibiting BH4 production. The main advantage of SPR inhibition over other enzymes within the BH4 synthetic pathway is actually the partial but not complete disruption of BH4 production. The total impairment of BH4 biosynthesis upon inhibition of GTPCH (another pathway enzyme) apparently results in severe side effects. In contrast, inhibition of SPR still allows for significantly reduced BH4 production via a parallel SPR-independent enzymatic route, resulting in a positive reduction of chronic pain without those side effects.[169,170] All of the above highlights the importance of SPR as a therapeutic target and explains the recent interest in developing SPR inhibitors.

SPR was identified as a potential target of the two analogs in this study, i.e., the unselective [pPy][Ph-mOHpCl] that bound 32 other proteins along with SPR and a selective [oPy][Ph-mOHpCl], where the single change from para- to ortho-pyridine made the molecule completely selective for SPR (a pair of molecules can thus be considered a selectivity cliff for SPR). The nitrile intermediate with the same [oPy][Ph-mOHpCl] substituents on 4- and 5-aryls equally demonstrated selective and potent SPR binding with similar affinity (Fig. 53b). All three compounds showed submicromolar SPR binding (Fig. 53b), however the binding alone could not indicate the inhibition of its enzymatic activity. An *in-vitro* activity assay for recombinant SPR was therefore established to complement the binding data of the chemoproteomics assay.

A potent SPR inhibitor QM385, thoroughly characterized in the study by Cronin et al., was kindly provided by Prof. Kai Johnsson (Max Planck Institute for Medical Research, Heidelberg) to serve as a benchmark molecule to confirm the enzyme inhibition in the newly established activity assay. QM385 represents an entirely different chemotype; however, it showed a very potent SPR binding in the chemoproteomics profiling assay against [oPy][Ph-mOHpCl] affinity matrix (Fig. 53c) in this study. These results indicated that although less potently, the novel [oPy][Ph-mOHpCl] analog engaged SPR in the same binding pocket as QM385.

The activity assay was established based on the SPR-mediated NADPH-dependent reduction of sepiapterin into dihydrobiopterin (BH2) (Fig. 53d), where the addition of an SPR inhibitor should ideally impede this transformation. Due to the overlapping absorbance maxima of NADP and BH2, the absorbance of sepiapterin served as a direct readout. The activity of SPR was calculated from sepiapterin absorbance as the following:

$$SPR\ activity = \frac{\Delta Abs}{\text{Time, min}}$$

where the absorbance difference ΔAbs is defined as the difference between the median triplicate sepiapterin absorbance at the compound concentration X and the median triplicate sepiapterin absorbance of the DMSO control sample.

Both QM385 and [oPy][Ph-mOHpCl] analog showed submicromolar SPR inhibition, proving that the novel selective molecule obtained in this study is indeed an SPR inhibitor (Fig. 53e). The EC50 values obtained in the recombinant enzyme activity assay corresponded to values obtained in the chemoproteomic profiling in lysate: 5 nM vs. 7 nM for QM385 and 369 nM vs. 308 nM for the [oPy][Ph-mOHpCl] analog. The affinity of SPR binding for the novel inhibitor could be improved; however, the current molecule already offers an opportunity for an easily obtained PROTAC SPR

degrader. Furthermore, a suitable linkage position for the E3 recruiting ligand is available, *i.e.*, the primary amine function allows for direct replacing the Sepharose beads by such ligands. This simple chemistry should readily produce functional PROTACs, whereas such linkage installation position is less clear for QM385. This is of particular relevance since it has been shown that SPR promotes hepatocellular carcinoma (HCC) progression independently of its enzymatic activity.[171] Comparison of the cellular effects induced by the PROTAC versus its related inhibitor could hence help comprehend the enzymatic and non-enzymatic function of SPR.
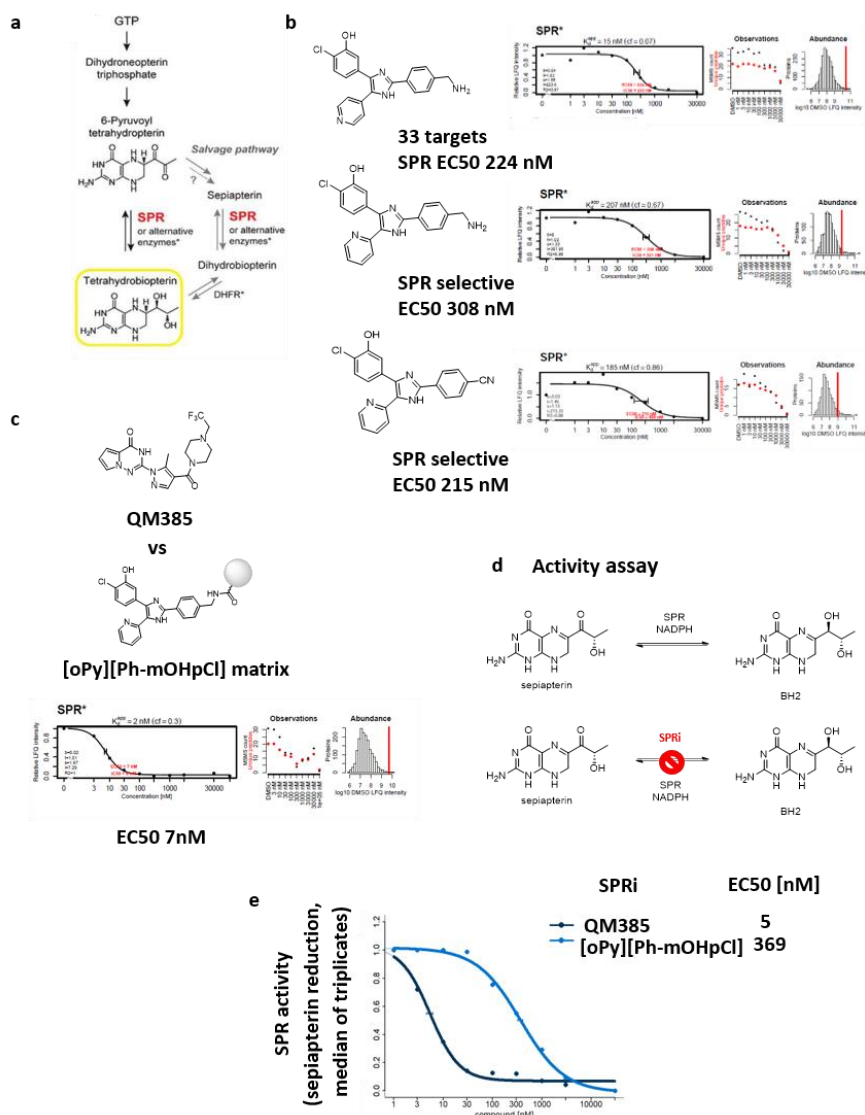
Figure 53. Sepiaterin reductase activity inhibition assay. a) BH4 biosynthetic pathway. b) SPR competition curves for two analogs in the lophine library from chemoproteomic profiling. Figure for panel a adapted from Haruki et. al.[165] b) SPR dose-response curve from the competitive profiling of [oPy][Ph-mOHpCl] analog against the tailored matrix c) same as panel b but for competition of QM385 against [oPy][Ph-mOHpCl] analog matrix. d) general scheme for enzymatic activity assay applied in this work: the addition of SPRi should impede the transformation of sepiapterin into BH2. e) Results of the activity assay, both QM385 and lophine analog inhibited the enzymatic activity of SPR.

9.12 **Conclusions of Part 2.**

A novel molecule-centric take on drug discovery was assayed in this work. It has been shown that it is indeed possible to employ chemoproteomic methods to assess the general ability of a particular chemotype to address protein targets on a proteome-wide level. The two chemotypes, pyrimidopyridones and lophines, were assayed here: several linkable analogs of each chemotype were immobilized on beads, and obtained affinity matrices were utilized in a competition-based chemoproteomic selectivity profiling. The assay was based on the premise that each tailored affinity probe could enrich the protein targets of each compound. In order to achieve that, the molecules were functionalized on beads within the part of the scaffold that is expected not to bear deep in the binding pockets of proteins. With the obtained set of molecules and matrices, it was possible to evaluate which proteins were bound by which compound in a fully agnostic manner. The matrix served as the enrichment of the target proteins out of cell lysate, while the free analog was set to compete with the matrix in increasing concentrations (up to 30 µM). Mass spectrometry-based bottom-up proteomics readout identified and quantified the hundreds of proteins retained specifically and unspecifically by the affinity matrix. When the free molecule tightly bound a protein, a typical dose-dependent sigmoidal curve upon concentration increase was observed, revealing it as a potential target of the molecule. For both scaffolds, it could be shown that the minor changes in the chemical structure greatly influenced the selectivity profile of each molecule, proving that the chemoproteomics-assisted proteome-wide SAR approach is generally applicable to different chemotypes. A complete interdisciplinary workflow applied in this work is represented in Fig. 54.
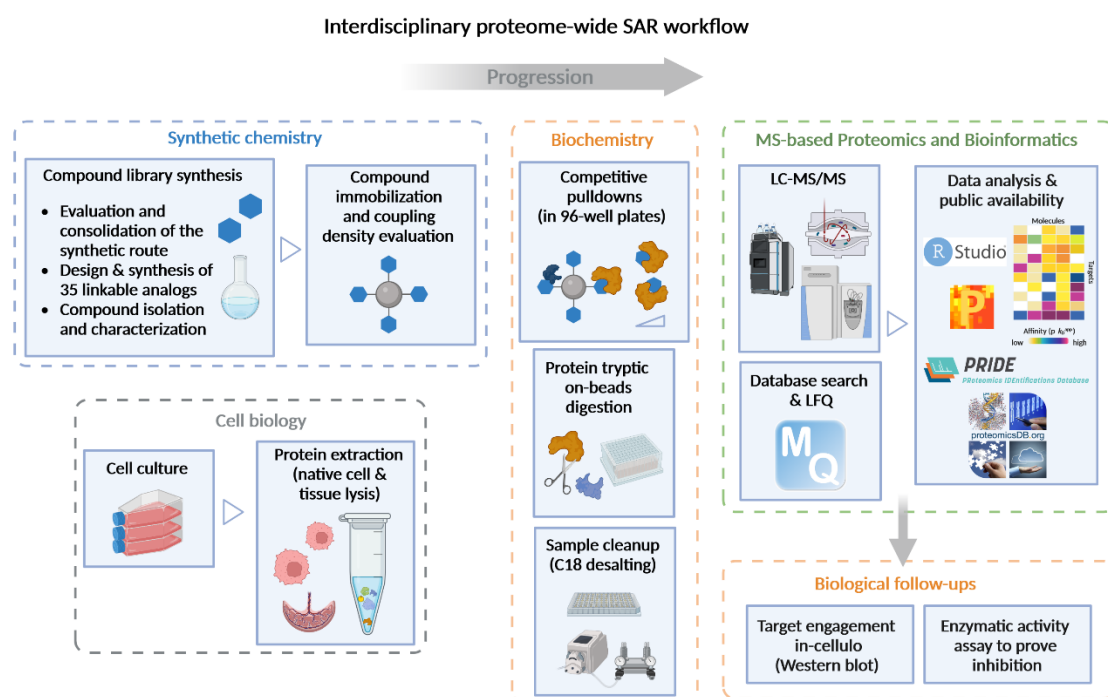
Figure 54. Schematic representation of proteome-wide SAR workflow comprising a chemical synthesis of linkable analogs and compound immobilization on beads, cell culture and preparation of lysates, competition pulldown assays, and sample cleanup, LC-MS/MS data acquisition and data processing using MaxQuant, data analysis and deposition of data on public resources.

The pyrimidopyridone chemotype proved to be more promiscuous than the lophine scaffold, as it could engage a much higher number of protein targets. Interestingly, despite being chemically similar, two analogs within the first scaffold, namely compound 1 and VI16832, demonstrated nearly complementary target profiles. The selectivity profiles of these two molecules combined encompass nearly 90% of all diverse targets observed for the pyrimidopyridone chemotype. Moreover, the scaffold historically evolved within the context of the development of selective CDK4 and 6 inhibitors. Unsurprisingly, all seven analogs tested in this work were able to engage both kinases, apart from a single exception, where compound 1 did not display CDK4 binding. However, it is interesting that two analogs within seven profiles showed much higher selectivity towards other protein targets, namely AK2 for newly synthesized analog c and CLIC1 for VI16743, in the assayed concentration range. Furthermore, apart from palbociclib, the drug specifically optimized for CDK4/6 inhibition, the other six molecules demonstrated fairly poor selectivity for CDK4/6. Both VI16832 and VI16743 were found to engage a broad panel of CDKs, and the former molecule was the most potent CDK4 and 6 binder among all profiled compounds.
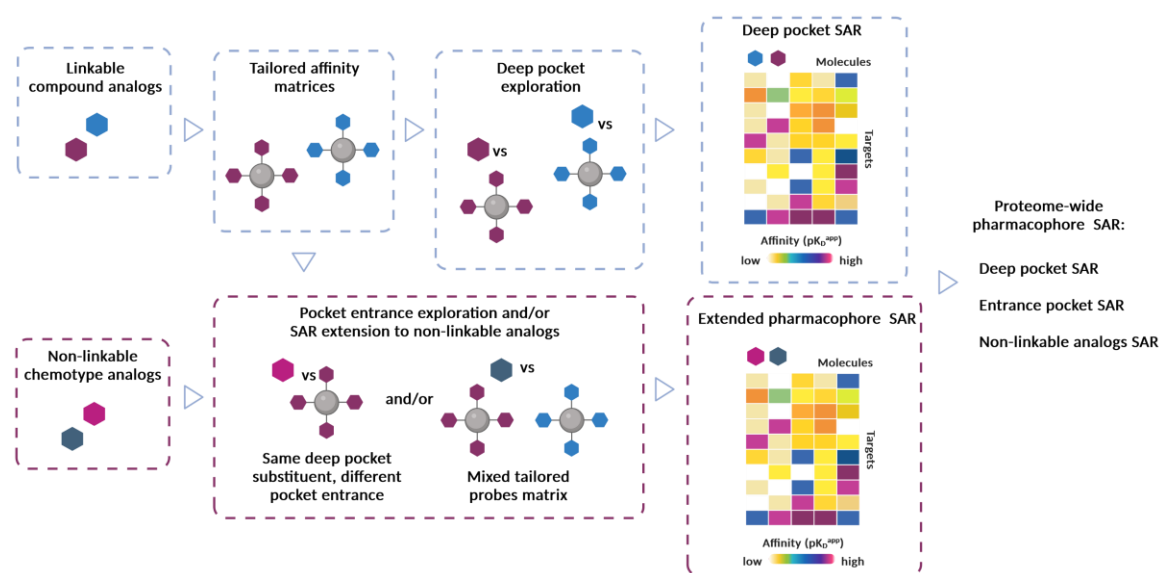
Figure 55. Schematic representation of the extended proteome-wide pharmacophore SAR

A focused library of 35 linkable lophine analogs has been prepared in a 4- to 5-step synthetic sequence involving a key multicomponent Debus–Radziszewski reaction. Each of these molecules featured a benzylamine as imidazole 2-substituent for subsequent preparation of respective tailored affinity matrices (Fig. 55). In a competition-based selectivity profiling, eleven molecules showed submicromolar affinity, primarily for kinases. This chemoproteomic profiling exercise simultaneously allowed for agnostic target identification of each analog and deep pocket SAR for the lophine chemotype. This work recapitulated all reported interactors and uncovered novel potential targets for directly linkable analogs of commercially available molecules with previously characterized targets. Unexpectedly, four of the 11 active molecules were found to have a single target: one exclusively bound RIPK2 (receptor-interacting serine/threonine-protein kinase 2), two molecules were binding only ALK5/TGFBR1 (TGF-beta receptor type-1), and surprisingly one compound was a selective SPR (sepiapterin reductase) binder. An activity assay confirmed that sepiapterin reductase was indeed inhibited by the molecule. In addition, ALK5 binders proved capable of engaging the kinase in-cellulo by monitoring the inhibition of its downstream pathway.

Moreover, the lophine pharmacophore SAR was further extended to explore the influence of scaffold pocket entrance substitutions on recovered selectivity profiles with the help of previously obtained tailored affinity matrices. This experiment demonstrated, in agreement with the available literature, that ALK5 binding is stabilized by an affine interaction from the 2-aryl substituents used to immobilize library analogs. Finally, it was demonstrated that the proteome-wide SAR could be further completed by profiling non-linkable chemotype analogs using the mixture of tailored pharmacophore affinity matrices (Fig. 55).

Applying the proteome-wide SAR on different pharmacophores may potentially greatly expand both the panel of available small molecule probes and the panel of novel druggable proteins. However, several limitations of this technique need to be addressed. Firstly, it has to be noted that the approach, by default, is challenging to implement, as it requires chemical synthesis of compound library analogs. The major limitation lies in knowing how analog molecules can be functionalized on beads without affecting their ability to bind proteins. For both scaffolds assayed in this work, the literature search and analysis of available crystal structures was sufficient to ensure that compound immobilization should not largely impede target engagement. However, for novel chemotypes with unknown binding modes, the direct implementation of this approach may generate affinity probes with compromised binding abilities. Therefore, in such cases, the preliminary SAR of multiple immobilization positions and/or acquisition of structural data (X-ray of protein compound-complex) is indispensable. Secondly, using the mixed pharmacophore matrix for extended non-linkable analog SAR has to be carried out with caution, as there is a risk of not enriching all of the targets of the profiled compounds. Here analyzed molecules are not identical to the probes functionalized on beads; therefore, the premise of enriching all protein targets by default can no longer be met. In this particular case, it is essential to ensure that at least the entire target panel of the chemotype from the initial library profiling is properly enriched by the mixed matrix to minimize the risk of not identifying targets of non-linkable analogs. This alone does not eliminate the risk of not having all protein targets bound by the mixed matrix, but it should help create a comprehensive enough profiling panel.

Overall, this work demonstrated the value of the novel concept: agnostic target deconvolution and evaluation of the effect of modifications on a chemotype selectivity in its binding pocket reveals drastic changes in the target profiles. Impressively for a kinase pharmacophore, the lophine chemotype also features sepiapterin reductase inhibitors, including a molecule deprived of any other activity. These results prompt the generalization of this approach to many chemotypes with the view of proposing chemical probes for proteins that do not have any and expanding the druggable proteome.

General Discussion and Outlook

Contents

10 Any pharmacophore holds the potential to afford a chemical probe

Proteome-wide SAR of lophine pharmacophore discussed in this thesis has demonstrated that four compounds out of profiled 35 displayed unexpected unique selectivity for three diverse protein targets. Notably, one of those, SPR, is a non-kinase and is not by default expected to be targeted by a kinase inhibitor scaffold. Suppose the assay was tailored to a particular protein target. In that case, those findings might have been completely missed as the tested molecules would simply produce negative results in an assay against any other protein. On the other side of the spectrum, several profiled pyrimidopyridone scaffold analogs showed very broad selectivity with more than a hundred protein targets each. These findings showcase the different and diverse target spectra between the two kinase inhibitor chemotypes.

Moreover, for both analyzed scaffolds, minor changes in the chemical structure showed great influence on the selectivity profiles of each individual molecule. For example, within the few profiled molecules of the pyrimidopyridone scaffold, palbociclib, in particular, although still not uniquely selective, demonstrated much higher selectivity towards its designated targets – CDK4/6. One may argue that palbociclib was optimized specifically for CDK4/6 through the rounds of SAR.[172] It is, however, conceivable that even without a particular target selection upon multiple structural alterations, it would be possible to obtain a selective compound within the same chemotype, similar to the lophine scaffold example. It is not a prerequisite, but rather a possibility that proteome-wide SAR of any given pharmacophore might lead to serendipitous discovery of novel molecules with completely unexpected selectivity trends, including novel probes possessing unique selectivity towards initially unanticipated protein targets. Here proteome-wide AfBPP profiling would help identify all protein binders of molecules in a target-agnostic fashion. Meaning that the approach can start as target-agnostic but simultaneously allow obtaining probes of unique selectivity.

Starting developing small molecule inhibitors for a validated protein target is undeniably reasonable. However, the results of this work suggest that adopting a molecule-centric point of view and profile molecules in a target-agnostic fashion might unravel unexpected compound selectivities. Employing proteome-wide pharmacophore SAR holds the potential to simultaneously 1) develop compounds with improved selectivity tailored to a specific target spectrum of interest, 2) profile all pharmacophore analogs for off-targets and simultaneously collect comprehensive selectivity profiles, 3) learn which protein targets a given pharmacophore can address in principle and which targets can be followed up for preparing the molecules with refined selectivity, 4) serendipitously unravel novel chemical probes, and 5) find new hits for previously untargeted proteins. The main premise of this approach is to use tailored affinity matrices that would enrich all potential protein targets of the profiled compounds. By design, this approach is applicable to study the influence of compound structural changes on affine interactions established deep within the binding pocket of a target. However, it will not directly be able to evaluate the contribution of chemical substituents at the immobilization site. It is, however, possible to utilize the original tailored probes to profile a range of non-linkable analogs with different substituents at the presumed solvent-exposed region of the scaffold. This should allow for deciphering how such substitutions affect the selectivity and binding affinity of the analyzed compounds, i.e., pocket entrance SAR, for any given pharmacophore of interest.

11 Compounds that bind without affecting protein function can be repurposed with the help of chemical proteomics

The major challenge that follows selectivity profiling using affinity probes still lies in confirming that observed binding events affect the activity or function of the identified potential protein binder. The protein can be validated as a target of the analyzed molecule using complementary techniques, where, for example, the enzyme activity is clearly inhibited upon the compound binding. For instance, the SAR study described in this thesis afforded a selective SPR binder, which inhibited the enzyme's activity in a complementary activity assay. Similarly, a novel ALK5 probe was able to inhibit the phosphorylation of ALK5 substrate SMAD2, which was verified via immunoblotting. These results are in line with the notion that proteins are typically considered targets only when the binding event results in direct consequences on a normal function of the protein. The latter is usually true when the molecule either binds directly in the active site of a protein or near it. In some cases, allosteric binding away from the active site induces conformational changes, affecting protein function.[173]

The other applications can be explored for the molecules that show potent binding in AfBPP competition assays but, at the same time, do not affect the function/activity of a protein. For instance, such compounds can be utilized as affinity warheads of protein-targeting chimeras (PROTACs). One may even speculate that in such cases the observed loss of protein activity or a phenotype can be easier attributed to the event of proteasomal degradation induced by the PROTAC and not enzymatic inhibition with the small molecule itself. Independent of potential therapeutic applications, such PROTACs can be very valuable chemical tools for general research. In particular, highly selective compounds can produce clean phenotypes upon removing a single protein from a biological system of interest. The use of a PROTAC, in this case, would serve as a potential alternative to difficult-to-control genetic experiments, where to study a role of a particular gene of interest, the gene product needs to be selectively removed from the system through gene editing or siRNA knockdowns. The comparison between the consequences of degradation versus enzymatic inhibition here can be assayed with complementary small molecule inhibitors that are binding orthosterically.

Compounds that potently bind to any protein pockets that do not (completely or at all) affect protein function can also be employed for protein visualization. For instance, small molecule-fluorophore conjugates can serve as quick tools to monitor the cellular localization or distribution of a protein of interest. This approach appears much simpler than other classic visualization methods, such as selective gene fusion with green fluorescent protein (GFP), for example.[174] Here, the small molecule warhead would guide the fluorophore-functionalized probe to the allosteric binding pocket of a protein and enable its visualization in diverse biological systems. Such an experiment would not require any additional genetic experiments and would allow visualizing native and active proteins.

Similarly, in cases when affine ligand binding does not affect protein function, such molecules can be repurposed for ligand-directed chemical labeling of proteins. Here, the fact that small molecule binding does not cause the loss of protein function is of major advantage. The general concept utilizes small molecule ligands that are fused with several types of reactive groups. When the ligand directs the probe into the protein binding pocket, this simultaneously enables a chemical

reaction between the reactive group of the probe and an amino acid located on the protein surface, affording selective and site-specific covalent chemical labeling of individual proteins.[175] Several types of reactive groups have already been incorporated into protein surfaces for various applications using this technique, for example, common photo-crosslinkers for detection of stabilized protein complexes or fluorophores for imaging of individual proteins or protein interactions.[176]

It has been shown that compounds, which display binding to surface receptors, can be fused to different types of bioactive cargo and facilitate drug delivery. For instance, receptor-mediated endocytosis has been previously evaluated for the ability to boost the transport of therapeutic agents like monoclonal antibodies, peptides, nucleic acids, or small molecule drugs across membranes, including the blood-brain barrier.[177] Here, small molecule warhead binding to surface receptors induces the cellular uptake of the chimeric molecule, eventually providing the opportunity for the bioactive cargo to exhibit its therapeutic effect in a disease-relevant environment. Conceptually, it is possible to envision that for allosteric surface receptor binders a similar approach could be assayed. Fusing a molecule that binds surface receptors within a pocket other than the active site with a bioactive cargo should still in theory enable its facilitated cellular uptake. If the allosteric binding alone would not be enough to trigger the membrane transport of the chimera, in the inevitable case of a specific ligand binding, the piggybacking chimera could be transported along with it. Evidently, this theoretical concept requires experimental validation; however, if proven right, it could serve as an opportunity for compound repurposing.

In summary, the compounds, which do not "score" in conventional functional assays despite evident protein binding, can be repurposed for another application taking advantage of their lack of effect on protein function. Such compounds can be functionalized into diverse chemical probes that may serve as tools for not only basic research but also drug discovery.

12 Protein enrichment may serve as a proxy indication of potential target binding

Proteome-wide SAR can only be comprehensive and target-agnostic when performed using tailored probes that do not affect the ability of a molecule to engage its targets, in the case of reversibly binding compounds – tailored affinity matrices. The premise is that each matrix is capable of enriching all potential protein targets of the molecule in a given biological system, where the native folding and activity of proteins are preserved. It is rational to assume that high affinity protein binders should also demonstrate high protein enrichment by the matrix, which is supported by the results of the pharmacophore SAR profiling discussed in this thesis. Therefore, it should be possible to roughly estimate that a given protein might be a binder of a molecule by analyzing the protein enrichment after a simple pulldown experiment with the affinity matrix. Of course, such assays would not be able to estimate the binding affinity of the interaction correctly. However, they can help narrow the potential high affine protein binders list. In the target-agnostic context, the simple enrichment assay would allow the identification of individual compound analogs that demonstrate high affinity towards unexpected targets or compounds with distinctive selectivity towards a particular protein, while other matrices would fail to enrich it out of lysate.

When the focus of profiling is tailored towards a specific target of interest, such a simple experiment would help to quickly eliminate compound analogs that do not show high affinity towards the designated protein target or show potential undesired broad selectivity. Parallel analysis of enrichment by multiple immobilized compound analogs would help to identify unspecific protein binders as they most likely would demonstrate constant level of enrichment across different matrices.

For newly isolated natural products (NPs) of unknown activity, a panel of immobilization techniques discussed in this thesis can be applied to quickly probe them for potential protein binding with a simple protein enrichment experiment. This would allow for a simultaneous screen for potential binders in very different biological systems. For instance, the same affinity probe can, in parallel, be used in lysates of human, bacterial, or plant origins, enabling quick assessment of potential interactors in all of these organisms. Consequently, such an assay would suggest plausible applications for the novel NP. The compounds that actually enrich proteins out of the utilized biological systems compared to respective total proteomes can be followed up in a proper target deconvolution exercise.

It is possible to envisage the implementation of an enrichment assay in a drug discovery pipeline as a screening setup for hit identification. When coupled to mass spectrometry, such an assay would allow the screening of a large panel of molecules immobilized on beads against complex proteomes, i.e., thousands of potential protein targets at a time. In order to valorize the screening, it would only be reasonable to expand the size of profiled affinity matrix libraries largely. This can be achieved by implementing on-bead synthesis, which can be performed in a high-throughput manner affording large panels of compound analogs on beads ready for profiling against complex proteomes. Only a selected number of interesting analogs would then need to be synthesized for further selectivity analysis and structure optimizations. Finally, since the assay can be performed in any lysate of any sequenced organism, the approach can be directly extended to various diseases and/or species. For instance, adding a stringent washing step after the original binding assay would allow reusing the same beads for subsequent pulldowns in another lysate. In this way, a large panel of on-bead prepared affinity matrices can be screened for protein binding against a large number of very diverse protein targets.

13 Mass spectrometry-based (chemo)proteomics enriches our understanding of how drugs work

Knowing the drug selectivity profile is not essential for it to be clinically successful, but it can definitely help reduce undesired side effects and better explain MoA. A recent review by Matthias Mann and Marcus Bantscheff discuss the statistic of drug attrition in phase II and III clinical trials and state that over 50% of drugs get withdrawn from evaluation due to low efficacy, insufficient target validation, and off-target effects.[178] Sadly, even nowadays, many studies report inadequate or/and incomplete interpretations of drug effects on biological systems. For instance, a recent Nature Cancer publication by Zeleke et al. assesses the inhibition of HDAC6 in the context of breast cancer treatment.[179] The study assessed the sensitivity of about 3,000 human breast cancers to HDAC6 inhibition with the HDAC6 "selective" inhibitor ricolinostat and even expanded the evaluation to a phase 1b dose-escalation clinical trial for patients with metastatic breast cancer.

The whole work is based on the assumption that ricolinostat pharmacology arises from selective HDAC6 inhibition, even though multiple earlier studies reported that this drug's anticancer activity is driven by the binding of other enzymes and is completely independent of HDAC6. Not only has it been previously reported that ricolinostat shows binding to more than 10 other enzymes, including several HDACs, apart from HDAC6, but also that the drug still kills cancer cells after a complete HDAC6 knockout.[58,180–182] Moreover, the work by Lin et al. reported that neither a knockout nor a knockdown of HDAC6 actually affects cell fitness across 12 different cell lines.[180] It appears rather clear that the drug's pharmacological effect must stem from the inhibition of targets other than HDAC6. Considering that unselective HDAC inhibitors demonstrated severe side effects in cancer patients and the data mentioned above, the clinical use of ricolinostat should also be assessed with great caution.[183] In this particular example, chemoproteomic selectivity profiling is indispensable as it greatly helps deduce the contribution of off-target binding to the overall observed drug effects. A similar case was reported in 2017, where the MS-coupled ABPP approach identified off-targets of a fatty acid amide hydrolase (FAAH) inhibitor BIA 10-2474. This drug caused considerable controversy earlier in 2016 after the death of one person and severe brain damage to four more volunteers in a dose escalation clinical trial.[184] Apparently, the severe side effects of the higher dose treatment could be linked to the off-target inhibition of several lipases involved in neuronal lipid metabolism, which caused substantial changes in the lipid networks of cortical neurons.[185] The latter findings were extended to the profiling of individual BIA 10-2474 metabolites targets by Huang et al. using ABPP coupled to MS readout in live cells.[186] This study identified aldehyde dehydrogenase ALDH2, the enzyme normally required for clearance of cellular formaldehyde for oxidative stress damage prevention, as a target of des-methylated BIA 10-2474. On the other hand, truly selective FAAH inhibitors, such as PF-04457845, were reported to exhibit good safety profiles in clinical trials.[187–190] Taken together, the off-target binding of BIA 10-2474 could be directly linked to the consequential nervous system dysfunction of the trial volunteers. Here the off-targets were elucidated with the help of MS-based chemoproteomics only after the tragic consequences of clinical trials in humans. It seems rather obvious that such selectivity profiling exercises must be included in the earlier stages of drug development before it leaves the *in-vitro* evaluation.

The panel of available probe-based and "label-free" techniques (AfBPP, ABPP, PAL, CETSA/TPP, SPROX, DARTS etc.) that enable the study of drug's action by the analysis of protein affinity, activity, stability, and folding were previously discussed in the Introduction of this thesis. Such methods support target deconvolution and selectivity profiling of hits or drug candidates and cover both reversible and irreversible protein binders. Selectivity profiling is not the sole potential application of MS-based proteomics in drug research. The possibility of in-depth measurement of complex proteomes enables us to assess the changes of function, abundance, and signaling networks of proteins upon various perturbations. Nowadays, MS-based proteomics can go beyond direct measurements of drug-target engagements and extend to the assessment of drug-modulated changes of protein expression, turnover, interactions and/or localization. Generally speaking, technological advances in MS instrumentation combined with chemical biology are greatly expanding the panel of available analytical tools that could support nearly every step of early drug discovery.[178]

14 Mass spectrometry-based (chemo)proteomics is the future of drug discovery

MS-based chemoproteomics can support design and evaluation of drug candidates with fine-tuned selectivity from the early stages of development. The targets and selectivity of the initial screening hits can be determined using one of the probe-based techniques mentioned above. The initial target identification and validation that follows up the phenotypic screens can be complemented by analyzing complex total proteomes upon classic genetic perturbation experiments. Here the analysis can be performed not only on the basic protein abundance level but also on the PTM level; therefore, the effects of genetic changes on signaling wiring networks within biological systems can be evaluated in parallel. Later on, the SAR study with the selected for medchem optimization molecules can further be supported with chemoproteomic proteome-wide selectivity profiling, where the advanced candidates' selectivity can be cross-evaluated in different biological contexts. Moreover, drug perturbation's direct and indirect influence on protein expression, PTMs, and turnover can also be analyzed. Such experiments could, for example, enrich our understanding of how the biological system adapts the signaling upon its perturbation with the molecule and/or eventually develops resistance to treatment.

The existing compound libraries can undergo proteome- or at least sub-proteome-wide selectivity profiling using appropriate methods (depending on the interaction nature of the molecule), and the acquired data can serve as a resource for hit selection when a new drug design campaign initiates. Such profiling effort would eliminate the need for repetitive screens with the same compounds each time a new target needs to be addressed. In other words, even for already existing molecules, MS-based proteomics would provide a comprehensive insight into what these molecules can actually do and hint at how they can be used. A selection of chemoproteomic screening tools already exists, allowing for sub-proteome profiling of reversibly binding small molecules against the kinome[30,191], metalloenzymes[58], bromodomain-containing proteins[192] and PARPs[193]. An excellent example of large-scale profiling was presented in the Ph.D. thesis of Dr. Maria Reinecke, where the Kinobeads technology was applied to simultaneously profile the selectivity of over 1.300 compounds from three bioactive compound libraries against a wide panel of kinases.[88] One can envisage parallel incorporation of a complementary sub-proteome selectivity profiling of the same compounds against a different panel of enzymes, for example, HDACs, which would dramatically enlarge the assayed panel of targets. This large dataset can then be used to select a molecule candidate to address a particular target or a selection thereof. Then, only a small number of such selected candidates would undergo further structure optimization steps sparing the initial extensive screening efforts. Individual promiscuous pharmacophores could be evaluated similarly to the SAR study presented in this thesis. The SAR against one single target using classic activity assays against recombinant target enzyme would not permit deciphering the actual protein binding spectrum of each SAR analog and any pharmacophore in general. In comparison, the extension to a proteome-wide evaluation enables the identification of unexpected targets and unique compound selectivities that would be ignored entirely following the classic single-target SAR.

In recent years, the chemical space of bioactives was greatly expanded with the help of fragment-based screening. The approach utilizes libraries of low-molecular-weight compounds (below 300 Da) and aims to identify structurally simple building blocks that can later be optimized into the ligands that fulfill specific selectivity and potency requirements. Such screening is based on the premise that a relatively small number of active fragments can represent a large fraction of chemical space. Initially, the technique did not extend beyond the low affine fragments assayed *in-vitro* against purified proteins, where the readout was performed through X-Ray crystallography or NMR spectroscopy.[194] In 2016, however, the group of Benjamin Cravatt described a chemoproteomic method of profiling fragments that target catalytic nucleophiles, i.e., reactive cysteine residues, using generic residue-specific reactive electrophile probes (such as iodo- or chloroacetamide, for instance).[195] This study further demonstrated that the particular covalent fragments could be advanced into selective and potent ligands. Since then, the fragment-based ABPP profiling incorporated mass spectrometry as the main analysis method and was extended to other covalent reactive residue binders, including serines, methionines, tyrosines, and lysines.[196–198] These efforts identified ligands for hundreds of proteins, including the ones that could never be targeted before and for a long time were believed to be "undruggable". In this context, the assay was equally applied to the ABPP-based screens of more structurally complex covalent ligands but using similar general reactive electrophilic probes. For instance, a study by Boike et al. performed a cysteine-reactive ABPP screen that identified a unique ligandable site within an intrinsically disordered region of MYC that could be addressed with a covalent binder.[199] This nuclear transcription factor is believed to be the most frequently amplified oncogene in human cancers; therefore, much effort historically was dedicated to developing MYC inhibitors. Unfortunately, the inhibition through direct binding remained unresolved for a long time due to the lack of conventional binding pockets within the intrinsically disordered MYC. The study by Boike et al. for the first time reported a covalent ligand that could functionally impair MYC transcription activity. It is quite apparent that here chemoproteomic ABPP approach not only expands the panel of available agents for known targets but can also serve as a resource for identifying novel disease-relevant targets. The fragment-based approach was equally extended to a high throughput screening of reversibly binding fragments in combination with PAL. In this setup, affine fragments are modified with light-activatable moieties allowing photo-crosslinking to the bound proteins. Upon covalent stabilization of the binding interaction, potential protein binders can be analyzed analogously to a classic covalent fragment screening.[200] Dr. Sebastian Essig presented at the EMBO Chemical Biology 2022 conference about a novel platform initiated by Bayer.[201] Essentially, the company decided to adopt a covalent fragment-based drug discovery approach, where privileged by medicinal chemists pharmacophores are expanded into screening panels of fragment analogs and are analyzed for potential ligandability using chemoproteomics. This campaign is a primary example of how chemical proteomics nowadays can be readily incorporated into drug research in industrial settings and facilitate drug design.

This thesis did not cover the topic of proteomics-assisted precision medicine. However, the treatment decision-making process can potentially benefit from the MS-based proteomic analysis of patient samples. Such analysis could help prescribe the best treatment for a specific patient in the function of their respective protein expression profile.[202,203] And even though this application is not directly related to drug research, here the combination of patient expression profiles

together with chemoproteomics drug selectivity data can potentially help tailor the treatment for each specific case and repurpose existing drugs to address another disease.
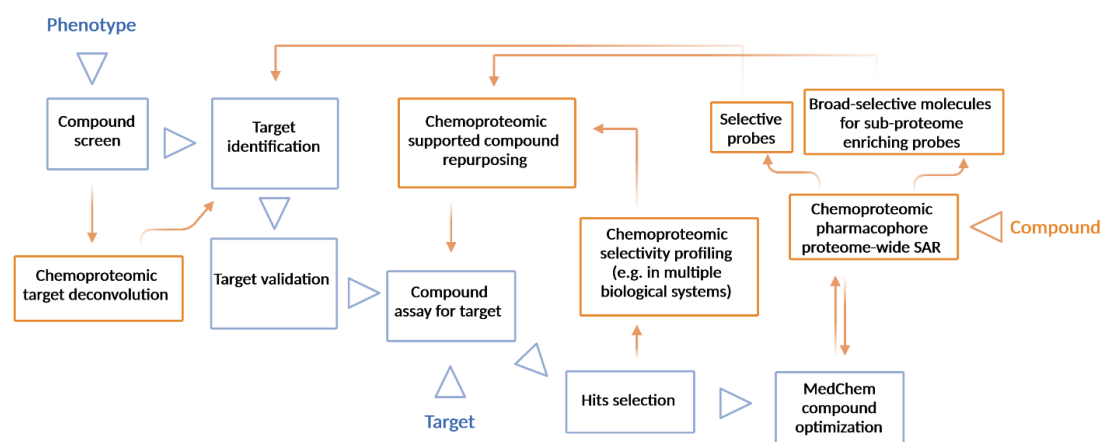
Figure 56. Chemoproteomics-assisted drug discovery.

The incorporation of proteomics at the early stages of drug discovery holds the potential to diverge from the classical target and phenotype and truly shift the paradigm towards the molecule-centric approach (Fig.56). Addressing the molecules with directly available selectivity profiles would vastly facilitate the assignment of their possible application. In the case of newly isolated natural products, the light-induced unselective immobilization could support the preparation of affinity probes to assess potential protein binders quickly. Other small molecules or fragments can be assayed using broad-selective affinity- or activity-based probes. All collected data could be deposited on a unified resource, enabling the search and selection of molecules or fragments with a desired target spectrum. Upon this evaluation, identified truly selective compounds can further serve as chemical probes, while the promiscuous ones can be employed for selectivity profiling of analogous pharmacophores and help expand the druggable proteome by identifying novel druggable proteins.

Even within the realm of classical phenotype-based discovery, adding a single step of chemoproteomic target deconvolution can assist in target identification and validation. In the target-oriented approach, the proteome-wide selectivity profiling of screening hits would drastically improve the selection of molecule candidates for further optimization stages toward improved potency, selectivity and/or pharmacokinetics. Here proteome-wide SAR would additionally enable affinity fine-tuning of more advanced candidates and simultaneously afford novel molecules with unprecedented selectivity.

Finally, chemoproteomics can be adapted as a standalone screening method. As previously mentioned, ABPP and AfBPP would employ small molecule or fragment probes for protein binding evaluation. Such experiments could quickly assess the potential protein targets in diverse biological systems and generally contribute to the exploratory evaluation of pharmacophore potential to address protein targets. The screening experiment can be set up both as target-oriented or completely target-agnostic, utilizing the same technique to answer various possible questions.

Altogether, MS-based (chemo)proteomic methods nowadays provide a powerful alternative to existing classic approaches within drug research and are currently being more and more frequently incorporated in the pipelines of big pharma companies. The unprecedented ability of MS-coupled approaches to comprehensively analyze complex biological systems makes them indispensable for the future of drug discovery.

Acknowledgements

Our lab in Freising has indeed become a home to me. And, of course, it could only be this way because of the fantastic people who worked with me side by side. So, unfortunately, I won't be able to call out each and one of you on this short page of text, but please know: it wouldn't be even a quarter as fun without you there.

First, I would like to thank Bernhard, who let me work in this great lab and learn more than I could ever imagine. Thank you for all the support and guidance you gave me.

My dear students: SG, Maxi, Verena, Simon, Yuchun, Roman, Katrin, Steffi N, and Karlis, thank you all for trusting me and deciding to complete the essential staples of your academic careers under my supervision. You all helped me immensely to become a teacher, and I also learned a ton from working with you. I hope you did enjoy your time in our lab and are proud of what we've accomplished together. Finally, I wish you all the very best for your future, wherever it may take you.

My dear friends: Julietti, NK, Leader Lee, Miri, Dirty Durian, Lapouschka, Fire-s, Mr. L, Nicolaj, MasterSteph, Tobi, Patroklos, Steffi, Mathias, Christoph, Giovanni, Sarah, the Mustascheers, you all made it into my heart. You made my day and shared many steps of this journey with me. You wouldn't believe how much I cherish all of you. The memories we created together I will not be able to forget.

SG, you deserve a separate paragraph here. My first ever student, and now one of my very closest friends, thank you for always understanding me so well and being so genuine and authentic. We have started the immobilization project together, which resulted in my first ever publication. I was hoping from day one, I could convince you to stick around. I'm so incredibly happy to call you my friend today.

My mom, who loves me unconditionally and, no matter what I do, is convinced I'm the best, thank you for this. Your love and encouragement pushed me through the most challenging times when I was doubting my decisions or feeling lonely. My mom was the person who taught me the importance of education and pushed me into pursuing the university degree. She was also the one who encouraged me to study abroad and experience the world outside of my hometown and my blissful comfort zone. This thesis is the pinnacle of her support and investment into my education. I thank my dad for being there for my mom when I couldn't. Finally, my closest friends beyond the lab, Anna P., Katya, and Nina, thank you for bearing with me even though I was many kilometers away. The distance never made me love you any less.

Last, and the most important (I decided to leave you for the very end following these people who always eat the best piece last, weirdos), Guillaume. Without you, none of this would even be possible. Thank you for entrusting me with your project, believing in and encouraging me, and, most of all, for getting angry every time I doubted myself. People in the lab were always a bit jealous that I had you on my side to teach, guide, and brainstorm with me. Well, it's true; I was truly lucky to be your Ph.D. student. Although you're not my official supervisor on paper, you are my only true mentor who taught me so incredibly much. Thank you. I would never stand where I do today without you. "We rock the place! Waka-waka!"

Each member of the Chair of Proteomics and Bioanalytics, past, present, and future: thank you for making this lab work. You are writing your own life stories, and I hope that for all of you, the chapter titled "Kusterlab" will be or was one of the happiest as it was for me.

Publication record

Main publications:

Prokofeva, P.; Höfer, S.; Hornisch, M.; Abele, M.; Kuster, B.; Médard, G. Merits of Diazirine Photo-Immobilization for Target Profiling of Natural Products and Cofactors. ACS Chem. Biol. 2022, 17, 11, 3100–3109

Additional publications during PhD:

Lo, W. T.; Belabed, H.; Kücükdisli, M.; Metag, J.; Roske, Y.; Prokofeva, P.; Ohashi, Y.; Horatscheck, A.; Cirillo, D.; Krauss, M.; Schmied, C.; Neuenschwander, M.; von Kries, J. P.; Médard, G.; Kuster, B.; Perisic, O.; Williams, R. L.; Daumke, O.; Payrastre, B.; Severin, S.; Nazaré, M.; Haucke, V. Development of Selective Inhibitors of Phosphatidylinositol 3-Kinase C2α. Nat. Chem. Biol. 2022, 19, 18-27.

Lechner, S.; Malgapo, M. I. P.; Grätz, C.; Steimbach, R. R.; Baron, A.; Rüther, P.; Nadal, S.; Stumpf, C.; Loos, C.; Ku, X.; Prokofeva, P.; Lautenbacher, L.; Heimburg, T.; Würf, V.; Meng, C.; Wilhelm, M.; Sippl, W.; Kleigrewe, K.; Pauling, J. K.; Kramer, K.; Miller, A. K.; Pfaffl, M. W.; Linder, M. E.; Kuster, B.; Médard, G. Target Deconvolution of HDAC Pharmacopoeia Reveals MBLAC2 as Common Off-Target. Nat. Chem. Biol. 2022, 18, 812-820. .

Zecha, J.; Bayer, F. P.; Wiechmann, S.; Woortman, J.; Berner, N.; Müller, J.; Schneider, A.; Kramer, K.; Abril-gil, M.; Hopf, T.; Reichart, L.; Chen, L.; Fynn, M.; Lechner, S.; Samaras, P.; Eckert, S.; Lautenbacher, L.; Reinecke, M.; Hamood, F.; Prokofeva, P.; Vornholz, L.; Falcomatà, C.; Dorsch, M.; Schröder, A.; Venhuizen, A.; Wilhelm, S.; Médard, G.; Stoehr, G.; Ruland, J.; Barbara, M.; Saur, D.; Buchner, M.; Ruprecht, B.; Hahne, H.; The, M.; Kuster, B. Decrypting Drug Actions and Protein Modifications by Dose- and Time-Resolved Proteomics. Science (80-.). 2023, 3925, 1–15

Curriculum vitae

Personal details

Name   Polina Prokofeva

Date of birth   19.04.1993

Address   Chair of Proteomics and Bioanalytics, Emil-Erlenmeyer-Forum 5,
                    85354 Freising, Germany

Email   Polina.prokofeva@tum.de

Education

2017-2023   **Doctoral candidate**, Chemical Proteomics

                    At Technical University of Munich, Munich, Germany

                    Thesis: "Pushing the boundaries of affinity-based target deconvolution: Promiscuous immobilization of nature-made molecules and proteome-wide medicinal chemistry"

2014-2016   **Master of Science**, Organic chemistry

                    At University Grenoble Alpes, France

                    Thesis: "Synthesis of Iridois"

2014-2016   **Bachelor of Science**, Chemistry

                    At Saint Petersburg State University, Institute of Chemistry, Russia

Teaching

Courses:

2018-2022   Intensive Course Proteomics, Technical University of Munich

2020-present   Chemical Biology lecture by Dr. Guillaume Medard, guest lecturer

Master projects:

2022   Drug target deconvolution by chemical proteomics: versatile live-cell competition assay with PAL probes, Technical University of Munich

2020   Target deconvolution of anticancer natural products by chemical proteomics, Technical University of Munich

2019   Chemical synthesis and target profiling of 2,4,5-tri(hetero)arylimidazole protein inhibitors among the native human proteome by chemical proteomics, Technical University of Munich

2019   Target identification of natural products by chemical proteomics: on-bead immobilization of Tacrolimus and Chivosazole, Technical University of Munich

Internship projects:

2022   Chemoproteomic target profiling of 3,4,5-triarylimidazole kinase inhibitors, Technical University of Munich

2022   Chemical Synthesis of linkable analogues of SB-431542 kinase inhibitor, Technical University of Munich

2021   Synthesis and chemical proteomic profiling of Palbociclib analogues, Technical University of Munich

2021   Chemical Synthesis of Palbociclib analogues, Technical University of Munich

2018   Synthesis and Chemical proteomic Target profiling of Raf-Kinase Inhibitor Analogues, Technical University of Munich

2018   Target profiling of ALK5 inhibitor SB-431542 and two analogues among the native proteome using chemical proteomics, Technical University of Munich

Bachelor projects:

2019   Target profiling of Bromodomain inhibitors JQ-1 and I-BET726 among the native human proteome using Chemical Proteomics, Technical University of Munich

Conference contributions:

2022   Expeditious Chemoproteomic Target Deconvolution of Nature-made Molecules (poster), 21st World Congress of the Human Proteome Organization (HUPO 2022), Cancun, Mexico

2022   Photo-immobilisation for target profiling of natural products and cofactors (talk), EMBO conference: Chemical Biology 2022, Heidelberg, Germany

2020   Proteome-wide SAR unravels a sepiapterin reductase probe (poster), EMBL Conference: Expanding the Druggable Proteome with Chemical Biology, Heidelberg, Germany

2019   Proteome-wide analysis of Structure-Affinity Relationships of drugs (poster), Keystone symposia: Proteomics and its Application to Translational and Precision Medicine, Stockholm, Sweden

2018   Proteome-wide Structure-Affinity Relationships (poster), International Chemical Biology Society ICBS2018, Vancouver, Canada

References

(1)   Lombardino, J. G.; Lowe, J. A. The Role of the Medicinal Chemist in Drug Discovery - Then and Now. *Nat. Rev. Drug Discov.* **2004**, *3* (10), 853–862. https://doi.org/10.1038/nrd1523.

(2)   Avery, V. M.; Camp, D.; Carroll, A. R.; Jenkins, I. D.; Quinn, R. J. The Identification of Bioactive Natural Products by High Throughput Screening (HTS). *Compr. Nat. Prod. III* **2020**, 410–429. https://doi.org/10.1016/B978-008045382-8.00062-9.

(3)   Heneghan, C.; Goldacre, B.; Mahtani, K. R. Why Clinical Trial Outcomes Fail to Translate into Benefits for Patients. *Trials* **2017**, *18* (1), 1–7. https://doi.org/10.1186/s13063-017-1870-2.

(4)   Inglese, J.; Johnson, R. L.; Simeonov, A.; Xia, M.; Zheng, W.; Austin, C. P.; Auld, D. S. High-Throughput Screening Assays for the Identification of Chemical Probes. *Nat. Chem. Biol.* **2007**, *3* (8), 466–479. https://doi.org/10.1038/nchembio.2007.17.

(5)   Terstappen, G. C.; Schlüpen, C.; Raggiaschi, R.; Gaviraghi, G. Target Deconvolution Strategies in Drug Discovery. *Nat. Rev. Drug Discov.* **2007**, *6* (11), 891–903. https://doi.org/10.1038/nrd2410.

(6)   JUMP-CP consortium https://jump-cellpainting.broadinstitute.org/.

(7)   Tucker, J. A.; Martin, M. P. Recent Advances in Kinase Drug Discovery Part i: The Editors' Take. *Int. J. Mol. Sci.* **2021**, *22* (14). https://doi.org/10.3390/ijms22147560.

(8)   DesJarlais, R. L.; Seibel, G. L.; Kuntz, I. D.; Sheridan, R. P.; Venkataraghavan, R.; Scott Dixon, J. Using Shape Complementarity as an Initial Screen in Designing Ligands for a Receptor Binding Site of Known Three-Dimensional Structure. *J. Med. Chem.* **1988**, *31* (4), 722–729. https://doi.org/10.1021/jm00399a006.

(9)   Ho, S.; Clipstone, N.; Timmermann, L.; Northrop, J.; Graef, I.; Fiorentino, D.; Nourse, J.; Crabtree, G. R. *The Mechanism of Action of Cyclosporin A and FK506*; 1996; Vol. 80.

(10)   Attene-Ramos, M. S.; Austin, C. P.; Xia, M. *High Throughput Screening*, Third Edit.; Elsevier, 2014; Vol. 2. https://doi.org/10.1016/B978-0-12-386454-3.00209-8.

(11)   Singh, P.; Harden, B. J.; Lillywhite, B. J.; Braod, P. M. Identification of Kinase Inhibitors by an ATP Depletion Method. *Assay Drug Dev. Technol.* **2004**, *2* (2), 161–169. https://doi.org/10.1089/154065804323056503.

(12)   Lowery, R. G.; Kleman-Leyer, K. Transcreener™: Screening Enzymes Involved in Covalent Regulation. *Expert Opin. Ther. Targets* **2006**, *10* (1), 179–190. https://doi.org/10.1517/14728222.10.1.179.

(13)   Wu, G.; Haw, R. *Functional Interaction Network Construction and Analysis for Disease Discovery*; 2017; Vol. 1558. https://doi.org/10.1007/978-1-4939-6783-4_11.

(14)   Hughes, J. P.; Rees, S. S.; Kalindjian, S. B.; Philpott, K. L. Principles of Early Drug Discovery. *Br. J. Pharmacol.* **2011**, *162* (6), 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x.

(15)   Pushpakom, S.; Iorio, F.; Eyers, P. A.; Escott, K. J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; Norris, A.; Sanseau, P.; Cavalla, D.; Pirmohamed, M. Drug Repurposing: Progress, Challenges and Recommendations. *Nat. Rev. Drug Discov.* **2018**, *18* (1), 41–58. https://doi.org/10.1038/nrd.2018.168.

(16)   Cohen, P.; Cross, D.; Jänne, P. A. Kinase Drug Discovery 20 Years after Imatinib: Progress and Future Directions. *Nat. Rev. Drug Discov.* **2021**, *20* (7), 551–569. https://doi.org/10.1038/s41573-021-00195-4.

(17)   Sarno, E. N.; Sampaio, E. P.; Kaplan, G.; Miranda, A.; Nery, J. A. C.; Miguel, C. P.; Viana, S. M. The Influence of Thalidomide on the Clinical and Immunologic Manifestation of Erythema Nodosum Leprosum. *J. Infect. Dis.* **1993**, *168* (2), 408–414. https://doi.org/10.1093/infdis/168.2.408.

(18)   Haslett, P. A. J.; Roche, P.; Butlin, C. R.; Macdonald, M.; Shrestha, N.; Manandhar, R.; LeMaster, J.; Hawksworth, R.; Shah, M.; Lubinsky, A. S.; Albert, M.; Worley, J.; Kaplan, G. Effective Treatment of Erythema Nodosum Leprosum with Thalidomide Is Associated with Immune Stimulation. *J. Infect. Dis.* **2005**, *192* (12), 2045–2053. https://doi.org/10.1086/498216.

(19)   Ito, T.; Handa, H. Molecular Mechanisms of Thalidomide and Its Derivatives. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **2020**, *96* (6), 189–203. https://doi.org/10.2183/PJAB.96.016.

(20)   Kwon, S.; Bae, H.; Jo, J.; Yoon, S. Comprehensive Ensemble in QSAR Prediction for Drug Discovery. *BMC Bioinformatics* **2019**, *20* (1), 1–12. https://doi.org/10.1186/s12859-019-3135-4.

(21)   Guha, R. On Exploring Structure-Activity Relationships. *Methods Mol. Biol.* **2013**, *993*, 81–94. https://doi.org/10.1007/978-1-62703-342-8_6.

(22)   Wassermann, A. M.; Wawer, M.; Bajorath, J. Activity Landscape Representations for Structure-Activity Relationship Analysis. *J. Med. Chem.* **2010**, *53* (23), 8209–8223. https://doi.org/10.1021/jm100933w.

References

(23)   Guha, R. Exploring Structure-Activity Data Using the Landscape Paradigm. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (6), 829–841. https://doi.org/10.1002/wcms.1087.

(24)   Saldívar-González, F. I.; Naveja, J. J.; Palomino-Hernández, O.; Medina-Franco, J. L. Getting SMARt in Drug Discovery: Chemoinformatics Approaches for Mining Structure-Multiple Activity Relationships. *RSC Adv.* **2017**, *7* (2), 632–641. https://doi.org/10.1039/c6ra26230a.

(25)   Bajusz, D.; Rácz, A.; Héberger, K. Why Is Tanimoto Index an Appropriate Choice for Fingerprint-Based Similarity Calculations? *J. Cheminform.* **2015**, *7* (1), 1–13. https://doi.org/10.1186/s13321-015-0069-3.

(26)   Lo, W. T.; Belabed, H.; Kücükdisli, M.; Metag, J.; Roske, Y.; Prokofeva, P.; Ohashi, Y.; Horatscheck, A.; Cirillo, D.; Krauss, M.; Schmied, C.; Neuenschwander, M.; von Kries, J. P.; Médard, G.; Kuster, B.; Perisic, O.; Williams, R. L.; Daumke, O.; Payrastre, B.; Severin, S.; Nazaré, M.; Haucke, V. Development of Selective Inhibitors of Phosphatidylinositol 3-Kinase C2α. *Nat. Chem. Biol.* **2022**. https://doi.org/10.1038/s41589-022-01118-z.

(27)   Huggins, D. J.; Sherman, W.; Tidor, B. Rational Approaches to Improving Selectivity in Drug Design. *J. Med. Chem.* **2012**, *55* (4), 1424–1444. https://doi.org/10.1021/jm2010332.

(28)   Mei, Y.; Yang, B. Rational Application of Drug Promiscuity in Medicinal Chemistry. *Future Med. Chem.* **2018**, *10* (15), 1835–1851. https://doi.org/10.4155/fmc-2018-0018.

(29)   Bhullar, K. S.; Lagarón, N. O.; McGowan, E. M.; Parmar, I.; Jha, A.; Hubbard, B. P.; Rupasinghe, H. P. V. Kinase-Targeted Cancer Therapies: Progress, Challenges and Future Directions. *Mol. Cancer* **2018**, *17* (1), 1–20. https://doi.org/10.1186/s12943-018-0804-2.

(30)   Klaeger, S.; Heinzlmeir, S.; Wilhelm, M.; Polzer, H.; Vick, B.; Koenig, P. A.; 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. K.; Gohlke, B. O.; Zolg, D. P.; Kayser, G.; Vooder, T.; Preissner, R.; Hahne, H.; Tõnisson, N.; Kramer, K.; Götze, K.; Bassermann, F.; Schlegl, J.; Ehrlich, H. C.; Aiche, S.; Walch, A.; Greif, P. A.; Schneider, S.; Felder, E. R.; Ruland, J.; Médard, G.; Jeremias, I.; Spiekermann, K.; Kuster, B. The Target Landscape of Clinical Kinase Drugs. *Science (80-.).* **2017**, *358* (6367). https://doi.org/10.1126/science.aan4368.

(31)   Farid, R.; Day, T.; Friesner, R. A.; Pearlstein, R. A. New Insights about HERG Blockade Obtained from Protein Modeling, Potential Energy Mapping, and Docking Studies. *Bioorganic Med. Chem.* **2006**, *14* (9), 3160–3173. https://doi.org/10.1016/j.bmc.2005.12.032.

(32)   Obach, R. S.; Walsky, R. L.; Venkatakrishnan, K.; Gaman, E. A.; Houston, J. B.; Tremaine, L. M. The Utility of in Vitro Cytochrome P450 Inhibition Data in the Prediction of Drug-Drug Interactions. *J. Pharmacol. Exp. Ther.* **2006**, *316* (1), 336–348. https://doi.org/10.1124/jpet.105.093229.

(33)   Lamers, M. B. A. C.; Antson, A. A.; Hubbard, R. E.; Scott, R. K.; Williams, D. H. Structure of the Protein Tyrosine Kinase Domain of C-Terminal Src Kinase (CSK) in Complex with Staurosporine. *J. Mol. Biol.* **1999**, *285* (2), 713–725. https://doi.org/10.1006/jmbi.1998.2369.

(34)   Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein-Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43* (W1), W443–W447. https://doi.org/10.1093/nar/gkv315.

(35)   Du, X.; Li, Y.; Xia, Y. L.; Ai, S. M.; Liang, J.; Sang, P.; Ji, X. L.; Liu, S. Q. Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **2016**, *17* (2), 1–34. https://doi.org/10.3390/ijms17020144.

(36)   Morra, G.; Genoni, A.; Neves, M.; Merz Jr., K.; Colombo, G. Molecular Recognition and Drug-Lead Identification: What Can Molecular Simulations Tell Us? *Curr. Med. Chem.* **2009**, *17* (1), 25–41. https://doi.org/10.2174/092986710789957797.

(37)   Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science (80-.).* **2002**, *298* (5600), 1912–1934. https://doi.org/10.1126/science.1075762.

(38)   Agafonov, R. V.; Wilson, C.; Otten, R.; Buosi, V.; Kern, D. Energetic Dissection of Gleevec's Selectivity toward Human Tyrosine Kinases. *Nat. Struct. Mol. Biol.* **2014**, *21* (10), 848–853. https://doi.org/10.1038/nsmb.2891.

(39)   Santos, K. B.; Guedes, I. A.; Karl, A. L. M.; Dardenne, L. E. Highly Flexible Ligand Docking: Benchmarking of the DockThor Program on the LEADS-PEP Protein-Peptide Data Set. *J. Chem. Inf. Model.* **2020**, *60* (2), 667–683. https://doi.org/10.1021/acs.jcim.9b00905.

(40)   Harini, K.; Jayashree, S.; Tiwari, V.; Vishwanath, S.; Sowdhamini, R. Ligand Docking Methods to Recognize Allosteric Inhibitors for G-Protein-Coupled Receptors. *Bioinform. Biol. Insights* **2021**, *15*. https://doi.org/10.1177/11779322211037769.

(41)   Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A.; Wohlschlegel, J. A.; Vondriska, T. M.; Pelletier, J.; Herschman, H. R.; Clardy, J.; Clarke, C. F.; Huang, J. Target Identification Using Drug Affinity Responsive Target Stability (DARTS). *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (51), 21984–21989. https://doi.org/10.1073/pnas.0910040106.

(42)   Schopper, S.; Kahraman, A.; Leuenberger, P.; Feng, Y.; Piazza, I.; Müller, O.; Boersema, P. J.; Picotti, P.

References

Measuring Protein Structural Changes on a Proteome-Wide Scale Using Limited Proteolysis-Coupled Mass Spectrometry. *Nat. Protoc.* **2017**, *12* (11), 2391–2410. https://doi.org/10.1038/nprot.2017.100.

(43) West, G. M.; Tang, L.; Fitzgerald, M. C. Thermodynamic Analysis of Protein Stability and Ligand Binding Using a Chemical Modification- and Mass Spectrometry-Based Strategy. *Anal. Chem.* **2008**, *80* (11), 4175–4185. https://doi.org/10.1021/ac702610a.

(44) Franken, H.; Mathieson, T.; Childs, D.; Sweetman, G. M. A.; Werner, T.; Tögel, I.; Doce, C.; Gade, S.; Bantscheff, M.; Drewes, G.; Reinhard, F. B. M.; Huber, W.; Savitski, M. M. Thermal Proteome Profiling for Unbiased Identification of Direct and Indirect Drug Targets Using Multiplexed Quantitative Mass Spectrometry. *Nat. Protoc.* **2015**, *10* (10), 1567–1593. https://doi.org/10.1038/nprot.2015.101.

(45) 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.* **2016**, *56*, 141–161. https://doi.org/10.1146/annurev-pharmtox-010715-103715.

(46) 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. Optimized Chemical Proteomics Assay for Kinase Inhibitor Profiling. *J. Proteome Res.* **2015**, *14* (3), 1574–1586. https://doi.org/10.1021/pr5012608.

(47) Smith, E.; Collins, I. Photoaffinity Labeling in Target-and Binding-Site Identification. *Future Med. Chem.* **2015**, *7* (2), 159–183. https://doi.org/10.4155/fmc.14.152.

(48) Murale, D. P.; Hong, S. C.; Haque, M. M.; Lee, J. S. Photo-Affinity Labeling (PAL) in Chemical Proteomics: A Handy Tool to Investigate Protein-Protein Interactions (PPIs). *Proteome Science*. BioMed Central Ltd. June 24, 2017. https://doi.org/10.1186/s12953-017-0123-3.

(49) Kubota, K.; Funabashi, M.; Ogura, Y. Target Deconvolution from Phenotype-Based Drug Discovery by Using Chemical Proteomics Approaches. *Biochim. Biophys. Acta - Proteins Proteomics* **2019**, *1867* (1), 22–27. https://doi.org/10.1016/j.bbapap.2018.08.002.

(50) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383–414. https://doi.org/10.1146/annurev.biochem.75.101304.124125.

(51) Wang, S.; Tian, Y.; Wang, M.; Wang, M.; Sun, G. B.; Sun, X. B. Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. *Front. Pharmacol.* **2018**, *9* (APR), 1–9. https://doi.org/10.3389/fphar.2018.00353.

(52) Dal Piaz, F.; Vera Saltos, M. B.; Franceschelli, S.; Forte, G.; Marzocco, S.; Tuccinardi, T.; Poli, G.; Nejad Ebrahimi, S.; Hamburger, M.; De Tommasi, N.; Braca, A. Drug Affinity Responsive Target Stability (DARTS) Identifies Laurifolioside as a New Clathrin Heavy Chain Modulator. *J. Nat. Prod.* **2016**, *79* (10), 2681–2692. https://doi.org/10.1021/acs.jnatprod.6b00627.

(53) Pepelnjak, M.; de Souza, N.; Picotti, P. Detecting Protein–Small Molecule Interactions Using Limited Proteolysis–Mass Spectrometry (LiP-MS). *Trends Biochem. Sci.* **2020**, *45* (10), 919–920. https://doi.org/10.1016/j.tibs.2020.05.006.

(54) Strickland, E. C.; Geer, M. A.; Tran, D. T.; Adhikari, J.; West, G. M.; Dearmond, P. D.; Xu, Y.; Fitzgerald, M. C. Thermodynamic Analysis of Protein-Ligand Binding Interactions in Complex Biological Mixtures Using the Stability of Proteins from Rates of Oxidation. *Nat. Protoc.* **2013**, *8* (1), 148–161. https://doi.org/10.1038/nprot.2012.146.

(55) Mateus, A.; Kurzawa, N.; Becher, I.; Sridharan, S.; Helm, D.; Stein, F.; Typas, A.; Savitski, M. M. Thermal Proteome Profiling for Interrogating Protein Interactions. *Mol. Syst. Biol.* **2020**, *16* (3), 1–11. https://doi.org/10.15252/msb.20199232.

(56) Zecha, J.; Bayer, F. P.; Wiechmann, S.; Woortman, J.; Berner, N.; Müller, J.; Schneider, A.; Kramer, K.; Abril-gil, M.; Hopf, T.; Reichart, L.; Chen, L.; Fynn, M.; Lechner, S.; Samaras, P.; Eckert, S.; Lautenbacher, L.; Reinecke, M.; Hamood, F.; Prokofeva, P.; Vornholz, L.; Falcomatà, C.; Dorsch, M.; Schröder, A.; Venhuizen, A.; Wilhelm, S.; Médard, G.; Stoehr, G.; Ruland, J.; Barbara, M.; Saur, D.; Buchner, M.; Ruprecht, B.; Hahne, H.; The, M.; Kuster, B. Decrypting Drug Actions and Protein Modifications by Dose- and Time-Resolved Proteomics. *Science (80-.).* **2023**, *3925*, 1–15.

(57) Reinecke, M.; Ruprecht, B.; Poser, S.; Wiechmann, S.; Wilhelm, M.; Heinzlmeir, S.; Kuster, B.; Médard, G. Chemoproteomic Selectivity Profiling of PIKK and PI3K Kinase Inhibitors. *ACS Chem. Biol.* **2019**, *14* (4), 655–664. https://doi.org/10.1021/acschembio.8b01020.

(58) Lechner, S.; Malgapo, M. I. P.; Grätz, C.; Steimbach, R. R.; Baron, A.; Rüther, P.; Nadal, S.; Stumpf, C.; Loos, C.; Ku, X.; Prokofeva, P.; Lautenbacher, L.; Heimburg, T.; Würf, V.; Meng, C.; Wilhelm, M.; Sippl, W.; Kleigrewe, K.; Pauling, J. K.; Kramer, K.; Miller, A. K.; Pfaffl, M. W.; Linder, M. E.; Kuster, B.; Médard, G. Target Deconvolution of HDAC Pharmacopoeia Reveals MBLAC2 as Common Off-Target. *Nat. Chem. Biol.* **2022**. https://doi.org/10.1038/s41589-022-01015-5.

References

(59) Heinzlmeir, S. Lehrstuhl Für Proteomik Und Bioanalytik When Chemical Proteomics Meets Medicinal Chemistry : Guided Drug Discovery towards EPHA2 Inhibitors, 2017.

(60) Yung-Chi, C.; Prusoff, W. H. Relationship between the Inhibition Constant (KI) and the Concentration of Inhibitor Which Causes 50 per Cent Inhibition (I50) of an Enzymatic Reaction. *Biochem. Pharmacol.* **1973**, *22* (23), 3099–3108. https://doi.org/10.1016/0006-2952(73)90196-2.

(61) Fleck, B. A.; Hoare, S. R. J.; Pick, R. R.; Bradbury, M. J.; Grigoriadis, D. E. Binding Kinetics Redefine the Antagonist Pharmacology of the Corticotropin-Releasing Factor Type 1 Receptor. *JPET* **2012**, *341* (2), 518–531.

(62) Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics: Abstract: Nature. *Nature* **2003**, *422* (6928), 198–207.

(63) 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.* **2012**, *404* (4), 939–965. https://doi.org/10.1007/s00216-012-6203-4.

(64) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 483–490. https://doi.org/10.1016/j.cbpa.2008.07.024.

(65) Gao, Y.; Yates, J. R. Protein Analysis by Shotgun Proteomics. *Mass Spectrom. Chem. Proteomics* **2019**, 1–38. https://doi.org/10.1002/9781118970195.ch1.

(66) Espadas, G.; Borràs, E.; Chiva, C.; Sabidó, E. Evaluation of Different Peptide Fragmentation Types and Mass Analyzers in Data-Dependent Methods Using an Orbitrap Fusion Lumos Tribrid Mass Spectrometer. *Proteomics* **2017**, *17* (9), 1–9. https://doi.org/10.1002/pmic.201600416.

(67) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J. J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. 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* **2012**, *11* (3), 1–11. https://doi.org/10.1074/mcp.O111.013698.

(68) Scheltema, R. A.; Hauschild, J. P.; Lange, O.; Hornburg, D.; Denisov, E.; Damoc, E.; Kuehn, A.; Makarov, A.; Mann, M. The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-Filter, High-Performance Quadrupole and an Ultra-High-Field Orbitrap Analyzer. *Mol. Cell. Proteomics* **2014**, *13* (12), 3698–3708. https://doi.org/10.1074/mcp.M114.043489.

(69) Levy, M. J.; Washburn, M. P.; Florens, L. Probing the Sensitivity of the Orbitrap Lumos Mass Spectrometer Using a Standard Reference Protein in a Complex Background. *J. Proteome Res.* **2018**, *17* (10), 3586–3592. https://doi.org/10.1021/acs.jproteome.8b00269.

(70) Tu, C.; Li, J.; Shen, S.; Sheng, Q.; Shyr, Y.; Qu, J. Performance Investigation of Proteomic Identification by HCD/CID Fragmentations in Combination with High/Low-Resolution Detectors on a Tribrid, High-Field Orbitrap Instrument. *PLoS One* **2016**, *11* (7), 1–14. https://doi.org/10.1371/journal.pone.0160160.

(71) Cottrell, J. S. Protein Identification Using MS/MS Data. *J. Proteomics* **2011**, *74* (10), 1842–1851. https://doi.org/10.1016/j.jprot.2011.05.014.

(72) Nesvizhskii, A. I. A Survey of Computational Methods and Error Rate Estimation Procedures for Peptide and Protein Identification in Shotgun Proteomics. *J. Proteomics* **2010**, *73* (11), 2092–2123. https://doi.org/10.1016/j.jprot.2010.08.009.

(73) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant Computational Platform for Mass Spectrometry-Based Shotgun Proteomics. *Nat. Protoc.* **2016**, *11* (12), 2301–2319. https://doi.org/10.1038/nprot.2016.136.

(74) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372. https://doi.org/10.1038/nbt.1511.

(75) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011**, *10* (4), 1794–1805. https://doi.org/10.1021/pr101065j.

(76) Jeong, K.; Kim, S.; Bandeira, N. False Discovery Rates in Spectral Identification. *BMC Bioinformatics* **2012**, *13 Suppl 1* (Suppl 16). https://doi.org/10.1186/1471-2105-13-S16-S2.

(77) Grossmann, J.; Roschitzki, B.; Panse, C.; Fortes, C.; Barkow-Oesterreicher, S.; Rutishauser, D.; Schlapbach, R. Implementation and Evaluation of Relative and Absolute Quantification in Shotgun Proteomics with Label-Free Methods. *J. Proteomics* **2010**, *73* (9), 1740–1746. https://doi.org/10.1016/j.jprot.2010.05.011.

(78) 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.* **2002**, *74* (18), 4741–4749. https://doi.org/10.1021/ac0256991.

(79) Chelius, D.; Bondarenko, P. V. Quantitative Profiling of Proteins in Complex Mixtures Using Liquid

References

Chromatography and Mass Spectrometry. *J. Proteome Res.* **2002**, *1* (4), 317–323. https://doi.org/10.1021/pr025517j.

(80) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **2014**, *13* (9), 2513–2526. https://doi.org/10.1074/mcp.M113.031591.

(81) Lim, M. Y.; Paulo, J. A.; Gygi, S. P. Evaluating False Transfer Rates from the Match-between-Runs Algorithm with a Two-Proteome Model. *J. Proteome Res.* **2019**, *18* (11), 4020–4026. https://doi.org/10.1021/acs.jproteome.9b00492.

(82) Kana, O.; Brylinski, M.; Rouge, B.; Rouge, B. Elucidating the Druggability of the Human Proteome with EFindSite. *J Comput Aided Mol Des.* **2020**, *33* (5), 509–519. https://doi.org/10.1007/s10822-019-00197-w.Elucidating.

(83) Lobry, C.; Oh, P.; Mansour, M. R.; Thomas Look, A.; Aifantis, I. Notch Signaling: Switching an Oncogene to a Tumor Suppressor. *Blood* **2014**, *123* (16), 2451–2459. https://doi.org/10.1182/blood-2013-08-355818.

(84) Llombart, V.; Mansour, M. R. Therapeutic Targeting of "Undruggable" MYC. *eBioMedicine* **2022**, *75*, 103756. https://doi.org/10.1016/j.ebiom.2021.103756.

(85) Huang, L.; Guo, Z.; Wang, F.; Fu, L. KRAS Mutation: From Undruggable to Druggable in Cancer. *Signal Transduct. Target. Ther.* **2021**, *6* (1), 1–20. https://doi.org/10.1038/s41392-021-00780-4.

(86) Butkiewicz, M.; Wang, Y.; Bryant, S. H.; W Lowe Jr, E.; Weaver C, D.; Meiler, J. High-Throughput Screening Assay Datasets from the PubChem Database. *Chem. Informatics* **2017**, *03* (01), 1–7. https://doi.org/10.21767/2470-6973.100022.

(87) Klaeger, S.; Gohlke, B.; Perrin, J.; Gupta, V.; Heinzlmeir, S.; Helm, D.; Qiao, H.; Bergamini, G.; Handa, H.; Savitski, M. M.; Bantscheff, M.; Médard, G.; Preissner, R.; Kuster, B. Chemical Proteomics Reveals Ferrochelatase as a Common Off-Target of Kinase Inhibitors. *ACS Chem. Biol.* **2016**, *11* (5), 1245–1254. https://doi.org/10.1021/acschembio.5b01063.

(88) Reinecke, M. Identifying Small Molecule Probes for Kinases by Chemical Proteomics, Technical University of Munich, 2020.

(89) Heinzlmeir, S.; Lohse, J.; Treiber, T.; Kudlinzki, D.; Linhard, V.; Gande, S. L.; Sreeramulu, S.; Saxena, K.; Liu, X.; Wilhelm, M.; Schwalbe, H.; Kuster, B.; Médard, G. Chemoproteomics-Aided Medicinal Chemistry for the Discovery of EPHA2 Inhibitors. *ChemMedChem* **2017**, *12* (12), 999–1011. https://doi.org/10.1002/cmdc.201700217.

(90) Heinzlmeir, S.; Kudlinzki, D.; Sreeramulu, S.; Klaeger, S.; Gande, S. L.; Linhard, V.; Wilhelm, M.; Qiao, H.; Helm, D.; Ruprecht, B.; Saxena, K.; Médard, G.; Schwalbe, H.; Kuster, B. Chemical Proteomics and Structural Biology Define EPHA2 Inhibition by Clinical Kinase Drugs. *ACS Chem. Biol.* **2016**, *11* (12), 3400–3411. https://doi.org/10.1021/acschembio.6b00709.

(91) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

(92) Prokofeva, P.; Hoefer, S.; Hornisch, M.; Abele, M.; Kuster, B.; Medard, G. Merits of Diazirine Photo-Immobilization for Target Profiling of Natural Products and Cofactors. *ACS Chem. Biol.* **2022**. https://doi.org/10.1021/acschembio.2c00500.

(93) Smith, E.; Collins, I. Photoaffinity Labeling in Target-and Binding-Site Identification. *Future Medicinal Chemistry*. Future Science Ltd February 1, 2015, pp 159–183. https://doi.org/10.4155/fmc.14.152.

(94) Kanoh, N. Photo-Cross-Linked Small-Molecule Affinity Matrix as a Tool for Target Identification of Bioactive Small Molecules. *Natural Product Reports*. Royal Society of Chemistry May 1, 2016, pp 709–718. https://doi.org/10.1039/c5np00117j.

(95) Kanoh, N.; Honda, K.; Shimizu, S.; Muroi, M.; Osada, H. *Photo-Crosslinked Small Molecule Affinity Matrix for Facilitating Forward and Reverse Chemical Genetics Supporting Information*.

(96) Kanoh, N.; Takayama, H.; Honda, K.; Moriya, T.; Teruya, T.; Simizu, S.; Osada, H.; Iwabuchi, Y. Cleavable Linker for Photo-Cross-Linked Small-Molecule Affinity Matrix. *Bioconjug. Chem.* **2010**, *21* (1), 182–186. https://doi.org/10.1021/bc900316q.

(97) Filipuzzi, I.; Thomas, J. R.; Pries, V.; Estoppey, D.; Salcius, M.; Studer, C.; Schirle, M.; Hoepfner, D. Direct Interaction of Chivosazole F with Actin Elicits Cell Responses Similar to Latrunculin A but Distinct from Chondramide. *ACS Chem. Biol.* **2017**, *12* (9), 2264–2269. https://doi.org/10.1021/acschembio.7b00385.

(98) Melder, F. T.; Lindemann, P.; Welle, A.; Trouillet, V.; Heißler, S.; Nazaré, M.; Selbach, M. Compound Interaction Screen on a Photoactivatable Cellulose Membrane (CISCM) Identifies Drug Targets. *ChemMedChem* **2022**. https://doi.org/10.1002/cmdc.202200346.

(99) Ertl, P.; Schuhmann, T. A Systematic Cheminformatics Analysis of Functional Groups Occurring in Natural Products. *J. Nat. Prod.* **2019**, *82* (5), 1258–1263. https://doi.org/10.1021/acs.jnatprod.8b01022.

References

(100) Das, J. Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chem. Rev.* **2011**, *111* (8), 4405–4417. https://doi.org/10.1021/cr1002722.

(101) West, A. V.; Muncipinto, G.; Wu, H. Y.; Huang, A. C.; Labenski, M. T.; Jones, L. H.; Woo, C. M. Labeling Preferences of Diazirines with Protein Biomolecules. *J. Am. Chem. Soc.* **2021**, *143* (17), 6691–6700. https://doi.org/10.1021/jacs.1c02509.

(102) Dorel, R.; Echavarren, A. M. Gold(I)-Catalyzed Activation of Alkynes for the Construction of Molecular Complexity. *Chem. Rev.* **2015**, *115* (17), 9028–9072. https://doi.org/10.1021/cr500691k.

(103) Trinchillo, M.; Belanzoni, P.; Belpassi, L.; Biasiolo, L.; Busico, V.; D'Amora, A.; D'Amore, L.; Del Zotto, A.; Tarantelli, F.; Tuzi, A.; Zuccaccia, D. Extensive Experimental and Computational Study of Counterion Effect in the Reaction Mechanism of NHC-Gold(I)-Catalyzed Alkoxylation of Alkynes. *Organometallics* **2016**, *35* (5), 641–654. https://doi.org/10.1021/acs.organomet.5b00925.

(104) Homs, A.; Obradors, C.; Leboeuf, D.; Echavarren, A. M. Dissecting Anion Effects in Gold(I)-Catalyzed Intermolecular Cycloadditions. *Adv. Synth. Catal.* **2014**, *356* (1), 221–228. https://doi.org/10.1002/adsc.201300704.

(105) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Metathesis Reactions in Total Synthesis. *Angew. Chemie - Int. Ed.* **2005**, *44* (29), 4490–4527. https://doi.org/10.1002/anie.200500369.

(106) Grela, K.; Harutyunyan, S.; Michrowska, A. A Highly Efficient Ruthenium Catalyst for Metathesis Reactions. *Angew. Chemie - Int. Ed.* **2002**, *41* (21), 4038–4040. https://doi.org/10.1002/1521-3773(20021104)41:21<4038::AID-ANIE4038>3.0.CO;2-0.

(107) Skowerski, K.; Białecki, J.; Tracz, A.; Olszewski, T. K. An Attempt to Provide an Environmentally Friendly Solvent Selection Guide for Olefin Metathesis. *Green Chem.* **2014**, *16* (3), 1125–1130. https://doi.org/10.1039/c3gc41943f.

(108) Fletcher, S. The Mitsunobu Reaction in the 21st Century. *Org. Chem. Front.* **2015**, *2* (6), 739–752. https://doi.org/10.1039/c5qo00016e.

(109) Swamy, K. C. K.; Kumar, N. N. B.; Balaraman, E.; Kumar, K. V. P. P. Mitsunobu and Related Reactions: Advances and Applications. *Chem. Rev.* **2009**, *109* (6), 2551–2651. https://doi.org/10.1021/cr800278z.

(110) Bantscheff, M.; Eberhard, D.; Abraham, Y.; Bastuck, S.; Boesche, M.; Hobson, S.; Mathieson, T.; Perrin, J.; Raida, M.; Rau, C.; Reader, V.; Sweetman, G.; Bauer, A.; Bouwmeester, T.; Hopf, C.; Kruse, U.; Neubauer, G.; Ramsden, N.; Rick, J.; Kuster, B.; Drewes, G. Quantitative Chemical Proteomics Reveals Mechanisms of Action of Clinical ABL Kinase Inhibitors. *Nat. Biotechnol.* **2007**, *25* (9), 1035–1044. https://doi.org/10.1038/nbt1328.

(111) Savitski, M. M.; Reinhard, F. B. M.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Molina, D. M.; Jafari, R.; Dovega, R. B.; Klaeger, S.; Kuster, B.; Nordlund, P.; Bantscheff, M.; Drewes, G. Tracking Cancer Drugs in Living Cells by Thermal Profiling of the Proteome. *Science (80-.).* **2014**, *346* (6205). https://doi.org/10.1126/science.1255784.

(112) Davis, T. L.; Walker, J. R.; Campagna-Slater, V.; Finerty, P. J.; Finerty, P. J.; Paramanathan, R.; Bernstein, G.; Mackenzie, F.; Tempel, W.; Ouyang, H.; Lee, W. H.; Eisenmesser, E. Z.; Dhe-Paganon, S. Structural and Biochemical Characterization of the Human Cyclophilin Family of Peptidyl-Prolyl Isomerases. *PLoS Biol.* **2010**, *8* (7). https://doi.org/10.1371/journal.pbio.1000439.

(113) Floss, H. G.; Yu, T. W. Rifamycin - Mode of Action, Resistance, and Biosynthesis. *Chem. Rev.* **2005**, *105* (2), 621–632. https://doi.org/10.1021/cr030112j.

(114) Campbell, E. A.; Korzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S. A. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **2001**, *104* (6), 901–912. https://doi.org/10.1016/S0092-8674(01)00286-0.

(115) Roe, S. M.; Prodromou, C.; O'Brien, R.; Ladbury, J. E.; Piper, P. W.; Pearl, L. H. Structural Basis for Inhibition of the Hsp90 Molecular Chaperone by the Antitumor Antibiotics Radicicol and Geldanamycin. *J. Med. Chem.* **1999**, *42* (2), 260–266. https://doi.org/10.1021/jm980403y.

(116) Šileikytė, J.; Sundalam, S.; David, L. L.; Cohen, M. S. Chemical Proteomics Approach for Profiling the NAD Interactome. *J. Am. Chem. Soc.* **2021**, *143* (18), 6787–6791. https://doi.org/10.1021/jacs.1c01302.

(117) Montgomery, D. C.; Garlick, J. M.; Kulkarni, R. A.; Kennedy, S.; Allali-Hassani, A.; Kuo, Y. M.; Andrews, A. J.; Wu, H.; Vedadi, M.; Meier, J. L. Global Profiling of Acetyltransferase Feedback Regulation. *J. Am. Chem. Soc.* **2016**, *138* (20), 6388–6391. https://doi.org/10.1021/jacs.6b03036.

(118) Levy, M. J.; Montgomery, D. C.; Sardiu, M. E.; Montano, J. L.; Bergholtz, S. E.; Nance, K. D.; Thorpe, A. L.; Fox, S. D.; Lin, Q.; Andresson, T.; Florens, L.; Washburn, M. P.; Meier, J. L. A Systems Chemoproteomic Analysis of Acyl-CoA/Protein Interaction Networks. *Cell Chem. Biol.* **2020**, *27* (3), 322-333.e5. https://doi.org/10.1016/j.chembiol.2019.11.011.

(119) Liu, Y. P.; Liao, W. C.; Ger, L. P.; Chen, J. C.; Hsu, T. I.; Lee, Y. C.; Chang, H. T.; Chen, Y. C.; Jan, Y. H.; Lee, K. H.;

References

Zeng, Y. H.; Hsiao, M.; Lu, P. J. Carboxyl-Terminal Modulator Protein Positively Regulates Akt Phosphorylation and Acts as an Oncogenic Driver in Breast Cancer. *Cancer Res.* **2013**, *73* (20), 6194–6205. https://doi.org/10.1158/0008-5472.CAN-13-0518.

(120) Forwood, J. K.; Thakur, A. S.; Guncar, G.; Marfori, M.; Mouradov, D.; Meng, W.; Robinson, J.; Huber, T.; Kellie, S.; Martin, J. L.; Hume, D. A.; Kobe, B. Structural Basis for Recruitment of Tandem Hotdog Domains in Acyl-CoA Thioesterase 7 and Its Role in Inflammation. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (25), 10382–10387. https://doi.org/10.1073/pnas.0700974104.

(121) Jung, S. H.; Lee, H. C.; Hwang, H. J.; Park, H. A.; Moon, Y. A.; Kim, B. C.; Lee, H. M.; Kim, K. P.; Kim, Y. N.; Lee, B. L.; Lee, J. C.; Ko, Y. G.; Park, H. J.; Lee, J. S. Acyl-CoA Thioesterase 7 Is Involved in Cell Cycle Progression via Regulation of PKCζ-P53-P21 Signaling Pathway. *Cell Death Dis.* **2017**, *8* (5), e2793. https://doi.org/10.1038/cddis.2017.202.

(122) Lienhart, W. D.; Gudipati, V.; MacHeroux, P. The Human Flavoproteome. *Arch. Biochem. Biophys.* **2013**, *535* (2), 150–162. https://doi.org/10.1016/j.abb.2013.02.015.

(123) Hefti, M. H.; Vervoort, J.; Van Berkel, W. J. H. Deflavination and Reconstitution of Flavoproteins: Tackling Fold and Function. *Eur. J. Biochem.* **2003**, *270* (21), 4227–4242. https://doi.org/10.1046/j.1432-1033.2003.03802.x.

(124) Youn, J. Y.; Dunham, W. H.; Hong, S. J.; Knight, J. D. R.; Bashkurov, M.; Chen, G. I.; Bagci, H.; Rathod, B.; MacLeod, G.; Eng, S. W. M.; Angers, S.; Morris, Q.; Fabian, M.; Côté, J. F.; Gingras, A. C. High-Density Proximity Mapping Reveals the Subcellular Organization of MRNA-Associated Granules and Bodies. *Mol. Cell* **2018**, *69* (3), 517-532.e11. https://doi.org/10.1016/j.molcel.2017.12.020.

(125) Shirai, Y. T.; Suzuki, T.; Morita, M.; Takahashi, A.; Yamamoto, T. Multifunctional Roles of the Mammalian CCR4-NOT Complex in Physiological Phenomena. *Front. Genet.* **2014**, *5* (AUG), 1–11. https://doi.org/10.3389/fgene.2014.00286.

(126) Hegele, A.; Kamburov, A.; Grossmann, A.; Sourlis, C.; Wowro, S.; Weimann, M.; Will, C. L.; Pena, V.; Lührmann, R.; Stelzl, U. Dynamic Protein-Protein Interaction Wiring of the Human Spliceosome. *Mol. Cell* **2012**, *45* (4), 567–580. https://doi.org/10.1016/j.molcel.2011.12.034.

(127) Wan, C.; Borgeson, B.; Phanse, S.; Tu, F.; Drew, K.; Clark, G.; Xiong, X.; Kagan, O.; Kwan, J.; Bezginov, A.; Chessman, K.; Pal, S.; Cromar, G.; Papoulas, O.; Ni, Z.; Boutz, D. R.; Stoilova, S.; Havugimana, P. C.; Guo, X.; Malty, R. H.; Sarov, M.; Greenblatt, J.; Babu, M.; Derry, W. B.; Tillier, E. R.; Wallingford, J. B.; Parkinson, J.; Marcotte, E. M.; Emili, A. Panorama of Ancient Metazoan Macromolecular Complexes. *Nature* **2015**, *525* (7569), 339–344. https://doi.org/10.1038/nature14877.

(128) Paul, S.; Dansithong, W.; Jog, S. P.; Holt, I.; Mittal, S.; Brook, J. D.; Morris, G. E.; Comai, L.; Reddy, S. Expanded CUG Repeats Dysregulate RNA Splicing by Altering the Stoichiometry of the Muscleblind 1 Complex. *J. Biol. Chem.* **2011**, *286* (44), 38427–38438. https://doi.org/10.1074/jbc.M111.255224.

(129) Ishikawa, H.; Tachikawa, H.; Miura, Y.; Takahashi, N. TRIM11 Binds to and Destabilizes a Key Component of the Activator-Mediated Cofactor Complex (ARC105) through the Ubiquitin-Proteasome System. *FEBS Lett.* **2006**, *580* (20), 4784–4792. https://doi.org/10.1016/j.febslet.2006.07.066.

(130) Afroz, T.; Skrisovska, L.; Belloc, E.; Guillén-Boixet, J.; Méndez, R.; Allain, F. H. T. A Fly Trap Mechanism Provides Sequence-Specific RNA Recognition by CPEB Proteins. *Genes Dev.* **2014**, *28* (13), 1498–1514. https://doi.org/10.1101/gad.241133.114.

(131) Van Nostrand, E. L.; Freese, P.; Pratt, G. A.; Wang, X.; Wei, X.; Xiao, R.; Blue, S. M.; Chen, J. Y.; Cody, N. A. L.; Dominguez, D.; Olson, S.; Sundararaman, B.; Zhan, L.; Bazile, C.; Bouvrette, L. P. B.; Bergalet, J.; Duff, M. O.; Garcia, K. E.; Gelboin-Burkhart, C.; Hochman, M.; Lambert, N. J.; Li, H.; McGurk, M. P.; Nguyen, T. B.; Palden, T.; Rabano, I.; Sathe, S.; Stanton, R.; Su, A.; Wang, R.; Yee, B. A.; Zhou, B.; Louie, A. L.; Aigner, S.; Fu, X. D.; Lécuyer, E.; Burge, C. B.; Graveley, B. R.; Yeo, G. W. A Large-Scale Binding and Functional Map of Human RNA-Binding Proteins. *Nature* **2020**, *583* (7818), 711–719. https://doi.org/10.1038/s41586-020-2077-3.

(132) Galgano, A.; Forrer, M.; Jaskiewicz, L.; Kanitz, A.; Zavolan, M.; Gerber, A. P. Comparative Analysis of MRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the MiRNA Regulatory System. *PLoS One* **2008**, *3* (9). https://doi.org/10.1371/journal.pone.0003164.

(133) Feng, Y.; De Franceschi, G.; Kahraman, A.; Soste, M.; Melnik, A.; Boersema, P. J.; De Laureto, P. P.; Nikolaev, Y.; Oliveira, A. P.; Picotti, P. Global Analysis of Protein Structural Changes in Complex Proteomes. *Nat. Biotechnol.* **2014**, *32* (10), 1036–1044. https://doi.org/10.1038/nbt.2999.

(134) Schirle, M.; Bantscheff, M.; Kuster, B. Mass Spectrometry-Based Proteomics in Preclinical Drug Discovery. *Chem. Biol.* **2012**, *19* (1), 72–84. https://doi.org/10.1016/j.chembiol.2012.01.002.

(135) Sumi, N. J.; Kuenzi, B. M.; Knezevic, C. E.; Remsing Rix, L. L.; Rix, U. Chemoproteomics Reveals Novel Protein and Lipid Kinase Targets of Clinical CDK4/6 Inhibitors in Lung Cancer. *ACS Chem. Biol.* **2015**, *10* (12), 2680–2686. https://doi.org/10.1021/acschembio.5b00368.

(136) Samaras, P.; Schmidt, T.; Frejno, M.; Gessulat, S.; Reinecke, M.; Jarzab, A.; Zecha, J.; Mergner, J.; Giansanti, P.;

References

Ehrlich, H. C.; Aiche, S.; Rank, J.; Kienegger, H.; Krcmar, H.; Kuster, B.; Wilhelm, M. ProteomicsDB: A Multi-Omics and Multi-Organism Resource for Life Science Research. *Nucleic Acids Res.* **2020**, *48* (D1), D1153–D1163. https://doi.org/10.1093/nar/gkz974.

(137) Frejno, M.; Meng, C.; Ruprecht, B.; Oellerich, T.; Scheich, S.; Kleigrewe, K.; Drecoll, E.; Samaras, P.; Hogrebe, A.; Helm, D.; Mergner, J.; Zecha, J.; Heinzlmeir, S.; Wilhelm, M.; Dorn, J.; Kvasnicka, H. M.; Serve, H.; Weichert, W.; Kuster, B. Proteome Activity Landscapes of Tumor Cell Lines Determine Drug Responses. *Nat. Commun.* **2020**, *11* (1), 1–12. https://doi.org/10.1038/s41467-020-17336-9.

(138) Metz, K. S.; Deoudes, E. M.; Berginski, M. E.; Jimenez-Ruiz, I.; Aksoy, B. A.; Hammerbacher, J.; Gomez, S. M.; Phanstiel, D. H. Coral: Clear and Customizable Visualization of Human Kinome Data. *Cell Syst.* **2018**, *7* (3), 347-350.e1. https://doi.org/10.1016/j.cels.2018.07.001.

(139) Ciceri, P.; Müller, S.; O'Mahony, A.; Fedorov, O.; Filippakopoulos, P.; Hunt, J. P.; Lasater, E. A.; Pallares, G.; Picaud, S.; Wells, C.; Martin, S.; Wodicka, L. M.; Shah, N. P.; Treiber, D. K.; Knapp, S. Dual Kinase-Bromodomain Inhibitors for Rationally Designed Polypharmacology. *Nat. Chem. Biol.* **2014**, *10* (4), 305–312. https://doi.org/10.1038/nchembio.1471.

(140) Feng, L.; Wang, G.; Chen, Y.; He, G.; Liu, B.; Liu, J.; Chiang, C. M.; Ouyang, L. Dual-Target Inhibitors of Bromodomain and Extra-Terminal Proteins in Cancer: A Review from Medicinal Chemistry Perspectives. *Med. Res. Rev.* **2022**, *42* (2), 710–743. https://doi.org/10.1002/med.21859.

(141) Sullivan, J. E.; Holdgate, G. A.; Campbell, D.; Timms, D.; Gerhardt, S.; Breed, J.; Breeze, A. L.; Bermingham, A.; Pauptit, R. A.; Norman, R. A.; Embrey, K. J.; Read, J.; VanScyoc, W. S.; Ward, W. H. J. Prevention of MKK6-Dependent Activation by Binding to P38α MAP Kinase. *Biochemistry* **2005**, *44* (50), 16475–16490. https://doi.org/10.1021/bi051714v.

(142) Bagley, M. C.; Davis, T.; Murziani, P. G. S.; Widdowson, C. S.; Kipling, D. Use of P38 MAPK Inhibitors for the Treatment of Werner Syndrome. *Pharmaceuticals* **2010**, *3* (6), 1842–1872. https://doi.org/10.3390/ph3061842.

(143) Poon, J. F.; Alao, J. P.; Sunnerhagen, P.; Dinér, P. Azastilbenes: A Cut-off to P38 MAPK Inhibitors. *Org. Biomol. Chem.* **2013**, *11* (27), 4526–4536. https://doi.org/10.1039/c3ob27449g.

(144) Godl, K.; Wissing, J.; Kurtenbach, A.; Habenberger, P.; Blencke, S.; Gutbrod, H.; Salassidis, K.; Stein-Gerlach, M.; Missio, A.; Cotten, M.; Daub, H. An Efficient Proteomics Method to Identify the Cellular Targets of Protein Kinase Inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (26), 15434–15439. https://doi.org/10.1073/pnas.2535024100.

(145) Charnley, A. K.; Convery, M. A.; Lakdawala Shah, A.; Jones, E.; Hardwicke, P.; Bridges, A.; Ouellette, M.; Totoritis, R.; Schwartz, B.; King, B. W.; Wisnoski, D. D.; Kang, J.; Eidam, P. M.; Votta, B. J.; Gough, P. J.; Marquis, R. W.; Bertin, J.; Casillas, L. Crystal Structures of Human RIP2 Kinase Catalytic Domain Complexed with ATP-Competitive Inhibitors: Foundations for Understanding Inhibitor Selectivity. *Bioorganic Med. Chem.* **2015**, *23* (21), 7000–7006. https://doi.org/10.1016/j.bmc.2015.09.038.

(146) Ruprecht, B.; Zecha, J.; Heinzlmeir, S.; Médard, G.; Lemeer, S.; Kuster, B. Evaluation of Kinase Activity Profiling Using Chemical Proteomics. *ACS Chem. Biol.* **2015**, *10* (12), 2743–2752. https://doi.org/10.1021/acschembio.5b00616.

(147) Wolkenberg, S. E.; Wisnoski, D. D.; Leister, W. H.; Wang, Y.; Zhao, Z.; Lindsley, C. W. Efficient Synthesis of Imidazoles from Aldehydes and 1,2-Diketones Using Microwave Irradiation. *Org. Lett.* **2004**, *6* (9), 1453–1456. https://doi.org/10.1021/ol049682b.

(148) Patel, G.; Dewangan, D. K.; Bhakat, N.; Banerjee, S. Green Approaches for the Synthesis of Poly-Functionalized Imidazole Derivatives: A Comprehensive Review. *Curr. Res. Green Sustain. Chem.* **2021**, *4* (September), 100175. https://doi.org/10.1016/j.crgsc.2021.100175.

(149) Amundsen, L. H.; Nelson, L. S. Reduction of Nitriles to Primary Amines with Lithium Aluminum Hydride. *J. Am. Chem. Soc.* **1951**, *73* (1), 242–244. https://doi.org/10.1021/ja01145a082.

(150) Benton, F. L.; Dillon, T. E. The Cleavage of Ethers with Boron Bromide. I. Some Common Ethers. *J. Am. Chem. Soc.* **1942**, *64* (5), 1128–1129. https://doi.org/10.1021/ja01257a035.

(151) He, H.; Wu, Y. J. Copper-Catalyzed Cross-Coupling of Aryl Iodides and Aryl Acetylenes Using Microwave Heating. *Tetrahedron Lett.* **2004**, *45* (16), 3237–3239. https://doi.org/10.1016/j.tetlet.2004.02.124.

(152) Gao, A.; Yang, F.; Li, J.; Wu, Y. Pd/Cu-Catalyzed Oxidation of Alkynes into 1,2-Diketones Using DMSO as the Oxidant. *Tetrahedron* **2012**, *68* (25), 4950–4954. https://doi.org/10.1016/j.tet.2012.04.069.

(153) Ogunjimi, A. A.; Zeqiraj, E.; Ceccarelli, D. F.; Sicheri, F.; Wrana, J. L.; David, L. Structural Basis for Specificity of TGFβ Family Receptor Small Molecule Inhibitors. *Cell. Signal.* **2012**, *24* (2), 476–483. https://doi.org/10.1016/j.cellsig.2011.09.027.

(154) Fährrolfes, R.; Bietz, S.; Flachsenberg, F.; Meyder, A.; Nittinger, E.; Otto, T.; Volkamer, A.; Rarey, M. Proteins

References

Plus: A Web Portal for Structure Analysis of Macromolecules. *Nucleic Acids Res.* **2017**, *45* (W1), W337–W343. https://doi.org/10.1093/nar/gkx333.

(155) Ember, S. W. J.; Zhu, J. Y.; Olesen, S. H.; Martin, M. P.; Becker, A.; Berndt, N.; Georg, G. I.; Schonbrunn, E. Acetyl-Lysine Binding Site of Bromodomain-Containing Protein 4 (BRD4) Interacts with Diverse Kinase Inhibitors. *ACS Chem. Biol.* **2014**, *9* (5), 1160–1171. https://doi.org/10.1021/cb500072z.

(156) van Caam, A.; Madej, W.; Garcia de Vinuesa, A.; Goumans, M. J.; ten Dijke, P.; Blaney Davidson, E.; van der Kraan, P. TGFβ1-Induced SMAD2/3 and SMAD1/5 Phosphorylation Are Both ALK5-Kinase-Dependent in Primary Chondrocytes and Mediated by TAK1 Kinase Activity. *Arthritis Res. Ther.* **2017**, *19* (1), 1–13. https://doi.org/10.1186/s13075-017-1302-4.

(157) Yue, J.; Mulder, K. M. Transforming Growth Factor-β Signal Transduction in Epithelial Cells. *Pharmacol. Ther.* **2001**, *91* (1), 1–34. https://doi.org/10.1016/S0163-7258(01)00143-7.

(158) Neuzillet, C.; de Gramont, A.; Tijeras-Raballand, A.; de Mestier, L.; Cros, J.; Faivre, S.; Raymond, E. Perspectives of TGF-ß Inhibition in Pancreatic and Hepatocellular Carcinomas. *Oncotarget* **2014**, *5* (1), 78–94. https://doi.org/10.18632/oncotarget.1569.

(159) Principe, D. R.; Doll, J. A.; Bauer, J.; Jung, B.; Munshi, H. G.; Bartholin, L.; Pasche, B.; Lee, C.; Grippo, P. J. TGF-β: Duality of Function between Tumor Prevention and Carcinogenesis. *Journal of the National Cancer Institute*. 2014. https://doi.org/10.1093/jnci/djt369.

(160) Zarzynska, J. M. Two Faces of TGF-Beta1 in Breast Cancer. *Mediators Inflamm.* **2014**, *2014*. https://doi.org/10.1155/2014/141747.

(161) Monsivais, D.; Peng, J.; Kang, Y.; Matzuk, M. M. Activin-like Kinase 5 (ALK5) Inactivation in the Mouse Uterus Results in Metastatic Endometrial Carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (9), 3883–3892. https://doi.org/10.1073/pnas.1806838116.

(162) Safina, A.; Vandette, E.; Bakin, A. V. ALK5 Promotes Tumor Angiogenesis by Upregulating Matrix Metalloproteinase-9 in Tumor Cells. *Oncogene* **2007**, *26* (17), 2407–2422. https://doi.org/10.1038/sj.onc.1210046.

(163) Baldwin, R. L.; Friess, H.; Yokoyama, M.; Lopez, M. E.; Kobrin, M. S.; Büchler, M. W.; Korc, M. Attenuated ALK5 Receptor Expression in Human Pancreatic Cancer: Correlation with Resistance to Growth Inhibition. *Int. J. Cancer* **1996**, *67* (2), 283–288. https://doi.org/10.1002/(SICI)1097-0215(19960717)67:2<283::AID-IJC21>3.0.CO;2-B.

(164) Whitman, M.; Raftery, L. TGFβ Signaling at the Summit. *Development* **2005**, *132* (19), 4205–4210. https://doi.org/10.1242/dev.02023.

(165) Haruki, H.; Hovius, R.; Pedersen, M. G.; Johnsson, K. Tetrahydrobiopterin Biosynthesis as a Potential Target of the Kynurenine Pathway Metabolite Xanthurenic Acid. *J. Biol. Chem.* **2016**, *291* (2), 652–657. https://doi.org/10.1074/jbc.C115.680488.

(166) Yco, L. P.; Geerts, D.; Mocz, G.; Koster, J.; Bachmann, A. S. Effect of Sulfasalazine on Human Neuroblastoma: Analysis of Sepiapterin Reductase (SPR) as a New Therapeutic Target. *BMC Cancer* **2015**, *15* (1), 1–11. https://doi.org/10.1186/s12885-015-1447-y.

(167) Mooney, M. R.; Geerts, D.; Kort, E. J.; Bachmann, A. S. Anti-Tumor Effect of Sulfasalazine in Neuroblastoma. *Biochem. Pharmacol.* **2019**, *162* (January), 237–249. https://doi.org/10.1016/j.bcp.2019.01.007.

(168) Cronin, S. J. F.; Seehus, C.; Weidinger, A.; Talbot, S.; Reissig, S.; Seifert, M.; Pierson, Y.; McNeill, E.; Longhi, M. S.; Turnes, B. L.; Kreslavsky, T.; Kogler, M.; Hoffmann, D.; Ticevic, M.; da Luz Scheffer, D.; Tortola, L.; Cikes, D.; Jais, A.; Rangachari, M.; Rao, S.; Paolino, M.; Novatchkova, M.; Aichinger, M.; Barrett, L.; Latremoliere, A.; Wirnsberger, G.; Lametschwandtner, G.; Busslinger, M.; Zicha, S.; Latini, A.; Robson, S. C.; Waisman, A.; Andrews, N.; Costigan, M.; Channon, K. M.; Weiss, G.; Kozlov, A. V.; Tebbe, M.; Johnsson, K.; Woolf, C. J.; Penninger, J. M. The Metabolite BH4 Controls T Cell Proliferation in Autoimmunity and Cancer. *Nature* **2018**, *563* (7732), 564–568. https://doi.org/10.1038/s41586-018-0701-2.

(169) Wu, Y.; Chen, P.; Sun, L.; Yuan, S.; Cheng, Z.; Lu, L.; Du, H.; Zhan, M. Sepiapterin Reductase: Characteristics and Role in Diseases. *J. Cell. Mol. Med.* **2020**, *24* (17), 9495–9506. https://doi.org/10.1111/jcmm.15608.

(170) Latremoliere, A.; Latini, A.; Andrews, N.; Cronin, S. J.; Fujita, M.; Gorska, K.; Hovius, R.; Romero, C.; Chuaiphichai, S.; Painter, M.; Miracca, G.; Babaniyi, O.; Remor, A. P.; Duong, K.; Riva, P.; Barrett, L. B.; Ferreirós, N.; Naylor, A.; Penninger, J. M.; Tegeder, I.; Zhong, J.; Blagg, J.; Channon, K. M.; Johnsson, K.; Costigan, M.; Woolf, C. J. Reduction of Neuropathic and Inflammatory Pain through Inhibition of the Tetrahydrobiopterin Pathway. *Neuron* **2015**, *86* (6), 1393–1406. https://doi.org/10.1016/j.neuron.2015.05.033.

(171) Wu, Y.; Du, H.; Zhan, M.; Wang, H.; Chen, P.; Du, D.; Liu, X.; Huang, X.; Ma, P.; Peng, D.; Sun, L.; Yuan, S.; Ding, J.; Lu, L.; Jiang, J. Sepiapterin Reductase Promotes Hepatocellular Carcinoma Progression via FoxO3a/Bim Signaling in a Nonenzymatic Manner. *Cell Death Dis.* **2020**, *11* (4). https://doi.org/10.1038/s41419-020-2471-7.

References

(172) Li, J. J.; Johnson, D. S. Innovative Drug Synthesis. *Innov. Drug Synth.* **2015**, 1–341. https://doi.org/10.1002/9781118819951.

(173) Goodey, N. M.; Benkovic, S. J. Allosteric Regulation and Catalysis Emerge via a Common Route. *Nat. Chem. Biol.* **2008**, *4* (8), 474–482. https://doi.org/10.1038/nchembio.98.

(174) Webb, C. D.; Resnekov, O. Use of Green Fluorescent Protein for Visualization of Cell-Specific Gene Expression and Subcellular Protein Localization in Bacillus Subtilis. *Methods Enzymol.* **1999**, *302* (20), 136–153. https://doi.org/10.1016/s0076-6879(99)02015-7.

(175) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. Ligand-Directed Tosyl Chemistry for Protein Labeling in Vivo. *Nat. Chem. Biol.* **2009**, *5* (5), 341–343. https://doi.org/10.1038/nchembio.157.

(176) Tsukiji, S.; Hamachi, I. Ligand-Directed Tosyl Chemistry for in Situ Native Protein Labeling and Engineering in Living Systems: From Basic Properties to Applications. *Curr. Opin. Chem. Biol.* **2014**, *21*, 136–143. https://doi.org/10.1016/j.cbpa.2014.07.012.

(177) Borsari, C.; Trader, D. J.; Tait, A.; Costi, M. P. Designing Chimeric Molecules for Drug Discovery by Leveraging Chemical Biology. *J. Med. Chem.* **2020**, *63* (5), 1908–1928. https://doi.org/10.1021/acs.jmedchem.9b01456.

(178) Meissner, F.; Geddes-McAlister, J.; Mann, M.; Bantscheff, M. The Emerging Role of Mass Spectrometry-Based Proteomics in Drug Discovery. *Nat. Rev. Drug Discov.* **2022**, *21* (9), 637–654. https://doi.org/10.1038/s41573-022-00409-3.

(179) Zeleke, T. Z.; Pan, Q.; Chiuzan, C.; Onishi, M.; Li, Y.; Tan, H.; Alvarez, M. J.; Honan, E.; Yang, M.; Chia, P. L.; Mukhopadhyay, P.; Kelly, S.; Wu, R.; Fenn, K.; Trivedi, M. S.; Accordino, M.; Crew, K. D.; Hershman, D. L.; Maurer, M.; Jones, S.; High, A.; Peng, J.; Califano, A.; Kalinsky, K.; Yu, J.; Silva, J. Network-Based Assessment of HDAC6 Activity Predicts Preclinical and Clinical Responses to the HDAC6 Inhibitor Ricolinostat in Breast Cancer. *Nat. Cancer* **2022**. https://doi.org/10.1038/s43018-022-00489-5.

(180) Lin, A.; Giuliano, C. J.; Palladino, A.; John, K. M.; Abramowicz, C.; Yuan, M. Lou; Sausville, E. L.; Lukow, D. A.; Liu, L.; Chait, A. R.; Galluzzo, Z. C.; Tucker, C.; Sheltzer, J. M. Off-Target Toxicity Is a Common Mechanism of Action of Cancer Drugs Undergoing Clinical Trials. *Sci. Transl. Med.* **2019**, *11* (509). https://doi.org/10.1126/scitranslmed.aaw8412.

(181) Depetter, Y.; Geurs, S.; De Vreese, R.; Goethals, S.; Vandoorn, E.; Laevens, A.; Steenbrugge, J.; Meyer, E.; de Tullio, P.; Bracke, M.; D'hooghe, M.; De Wever, O. Selective Pharmacological Inhibitors of HDAC6 Reveal Biochemical Activity but Functional Tolerance in Cancer Models. *Int. J. Cancer* **2019**, *145* (3), 735–747. https://doi.org/10.1002/ijc.32169.

(182) Winkler, R.; Mägdefrau, A. S.; Piskor, E. M.; Kleemann, M.; Beyer, M.; Linke, K.; Hansen, L.; Schaffer, A. M.; Hoffmann, M. E.; Poepsel, S.; Heyd, F.; Beli, P.; Möröy, T.; Mahboobi, S.; Krämer, O. H.; Kosan, C. Targeting the MYC Interaction Network in B-Cell Lymphoma via Histone Deacetylase 6 Inhibition. *Oncogene* **2022**, *41* (40), 4560–4572. https://doi.org/10.1038/s41388-022-02450-3.

(183) Cengiz Seval, G.; Beksac, M. A Comparative Safety Review of Histone Deacetylase Inhibitors for the Treatment of Myeloma. *Expert Opin. Drug Saf.* **2019**, *18* (7), 563–571. https://doi.org/10.1080/14740338.2019.1615051.

(184) Butler, D.; Callaway, E. Scientists in the Dark after French Clinical Trial Proves Fatal. *Nature* **2016**, *529* (7586), 263–264. https://doi.org/10.1038/nature.2016.19189.

(185) van Esbroeck, A. C. M.; Janssen, A. P. A.; Cognetta, A. B.; Ogasawara, D.; Shpak, G.; van der Kroeg, M.; Kantae, V.; Baggelaar, M. P.; de Vrij, F. M. S.; Deng, H.; Allarà, M.; Fezza, F.; Lin, Z.; van der Wel, T.; Soethoudt, M.; Mock, E. D.; den Dulk, H.; Baak, I. L.; Florea, B. I.; Hendriks, G.; De Petrocellis, L.; Overkleeft, H. S.; Hankemeier, T.; De Zeeuw, C. I.; Di Marzo, V.; Maccarrone, M.; Cravatt, B. F.; Kushner, S. A.; van der Stelt, M. Activity-Based Protein Profiling Reveals off-Target Proteins of the FAAH Inhibitor BIA 10-2474. *Science (80-.).* **2017**, *356* (6342), 1084–1087. https://doi.org/10.1126/science.aaf7497.

(186) Huang, Z.; Ogasawara, D.; Seneviratne, U. I.; Cognetta, A. B.; am Ende, C. W.; Nason, D. M.; Lapham, K.; Litchfield, J.; Johnson, D. S.; Cravatt, B. F. Global Portrait of Protein Targets of Metabolites of the Neurotoxic Compound BIA 10–2474. *ACS Chem. Biol.* **2019**, *14* (2), 192–197. https://doi.org/10.1021/acschembio.8b01097.

(187) Li, G. L.; Winter, H.; Arends, R.; Jay, G. W.; Le, V.; Young, T.; Huggins, J. P. Assessment of the Pharmacology and Tolerability of PF-04457845, an Irreversible Inhibitor of Fatty Acid Amide Hydrolase-1, in Healthy Subjects. *Br. J. Clin. Pharmacol.* **2012**, *73* (5), 706–716. https://doi.org/10.1111/j.1365-2125.2011.04137.x.

(188) Huggins, J. P.; Smart, T. S.; Langman, S.; Taylor, L.; Young, T. An Efficient Randomised, Placebo-Controlled Clinical Trial with the Irreversible Fatty Acid Amide Hydrolase-1 Inhibitor PF-04457845, Which Modulates Endocannabinoids but Fails to Induce Effective Analgesia in Patients with Pain Due to Osteoarthritis of Th. *Pain* **2012**, *153* (9), 1837–1846. https://doi.org/10.1016/j.pain.2012.04.020.

(189) Postnov, A.; Schmidt, M. E.; Pemberton, D. J.; de Hoon, J.; van Hecken, A.; van den Boer, M.; Zannikos, P.; van

References

der Ark, P.; Palmer, J. A.; Rassnick, S.; Celen, S.; Bormans, G.; van Laere, K. Fatty Acid Amide Hydrolase Inhibition by JNJ-42165279: A Multiple-Ascending Dose and a Positron Emission Tomography Study in Healthy Volunteers. *Clin. Transl. Sci.* **2018**, *11* (4), 397–404. https://doi.org/10.1111/cts.12548.

(190) Pawsey, S.; Wood, M.; Browne, H.; Donaldson, K.; Christie, M.; Warrington, S. Safety, Tolerability and Pharmacokinetics of FAAH Inhibitor V158866: A Double-Blind, Randomised, Placebo-Controlled Phase I Study in Healthy Volunteers. *Drugs R. D.* **2016**, *16* (2), 181–191. https://doi.org/10.1007/s40268-016-0127-y.

(191) Zhao, Q.; Ouyang, X.; Wan, X.; Gajiwala, K. S.; Kath, J. C.; Jones, L. H.; Burlingame, A. L.; Taunton, J. Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J. Am. Chem. Soc.* **2017**, *139* (2), 680–685. https://doi.org/10.1021/jacs.6b08536.

(192) Li, X.; Wu, Y.; Tian, G.; Jiang, Y.; Liu, Z.; Meng, X.; Bao, X.; Feng, L.; Sun, H.; Deng, H.; Li, X. D. Chemical Proteomic Profiling of Bromodomains Enables the Wide-Spectrum Evaluation of Bromodomain Inhibitors in Living Cells. *J. Am. Chem. Soc.* **2019**, *141* (29), 11497–11505. https://doi.org/10.1021/jacs.9b02738.

(193) Knezevic, C. E.; Wright, G.; Remsing Rix, L. L.; Kim, W.; Kuenzi, B. M.; Luo, Y.; Watters, J. M.; Koomen, J. M.; Haura, E. B.; Monteiro, A. N.; Radu, C.; Lawrence, H. R.; Rix, U. Proteome-Wide Profiling of Clinical PARP Inhibitors Reveals Compound-Specific Secondary Targets. *Cell Chem. Biol.* **2016**, *23* (12), 1490–1503. https://doi.org/10.1016/j.chembiol.2016.10.011.

(194) Hajduk, P. J.; Greer, J. A Decade of Fragment-Based Drug Design: Strategic Advances and Lessons Learned. *Nat. Rev. Drug Discov.* **2007**, *6* (3), 211–219. https://doi.org/10.1038/nrd2220.

(195) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-Páez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; Wolan, D. W.; Cravatt, B. F. Proteome-Wide Covalent Ligand Discovery in Native Biological Systems. *Nature* **2016**, *534* (7608), 570–574. https://doi.org/10.1038/nature18002.

(196) Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F. Global Profiling of Lysine Reactivity and Ligandability in the Human Proteome. *Nat. Chem.* **2017**, *9* (12), 1181–1190. https://doi.org/10.1038/nchem.2826.

(197) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; Chang, C. J. Redox-Based Reagents for Chemoselective Methionine Bioconjugation. *Science (80-.).* **2017**, *355* (6325), 597–602. https://doi.org/10.1126/science.aal3316.

(198) Hahm, H. S.; Toroitich, E. K.; Borne, A. L.; Brulet, J. W.; Libby, A. H.; Yuan, K.; Ware, T. B.; McCloud, R. L.; Ciancone, A. M.; Hsu, K. L. Global Targeting of Functional Tyrosines Using Sulfur-Triazole Exchange Chemistry. *Nat. Chem. Biol.* **2020**, *16* (2), 150–159. https://doi.org/10.1038/s41589-019-0404-5.

(199) Boike, L.; Cioffi, A. G.; Majewski, F. C.; Co, J.; Henning, N. J.; Jones, M. D.; Liu, G.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Nomura, D. K. Discovery of a Functional Covalent Ligand Targeting an Intrinsically Disordered Cysteine within MYC. *Cell Chem. Biol.* **2021**, *28* (1), 4-13.e17. https://doi.org/10.1016/j.chembiol.2020.09.001.

(200) Keeley, A.; Petri, L.; Ábrányi-Balogh, P.; Keserű, G. M. Covalent Fragment Libraries in Drug Discovery. *Drug Discov. Today* **2020**, *25* (6), 983–996. https://doi.org/10.1016/j.drudis.2020.03.016.

(201) Essig, S. *Towards a Chemical Biology Platform for the Systematic Discovery and Evaluation of Novel Covalent Chemistry*; 2022.

(202) Correa Rojo, A.; Heylen, D.; Aerts, J.; Thas, O.; Hooyberghs, J.; Ertaylan, G.; Valkenborg, D. Towards Building a Quantitative Proteomics Toolbox in Precision Medicine: A Mini-Review. *Front. Physiol.* **2021**, *12* (August). https://doi.org/10.3389/fphys.2021.723510.

(203) Duarte, T. T.; Spencer, C. T. Personalized Proteomics: The Future of Precision Medicine. *Proteomes* **2016**, *4* (4). https://doi.org/10.3390/proteomes4040029.