Development of a Compound Interaction Screen on a photoactivatable Cellulose Membrane (CISCM) to identify drug targets

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

FIONA TERESA INES MELDER

2023

This thesis was completed in the period from January 2018 to July 2023 at the Max-Delbrück-Centrum for Molecular Medicine in Berlin under the supervision of Prof. Dr. Matthias Selbach.

1st reviewer:	Prof. Dr. Matthias Selbach
2nd reviewer:	Prof. Dr. Gerhard Wolber
Date of defence:	24.01.2024

ACKNOWLEDGEMENTS

At this point, I want to sincerely thank all the people who made this thesis possible, who supported, inspired and guided me along this journey.

First I want to thank Matthias. Thank you for letting me see research through your eyes and for showing me how playful and fun science can be. Thank you for trusting me with this project, for letting me work with a lot of freedom, and for teaching me what it means to work in a self-dependent way, even though this might have required a lot of patience from your side and me getting lost until I found my way. Thank you for the always open doors and for creating a space where asking questions and open discussion were encouraged. No matter how big a problem appeared, you always seemed to find a solution and showed me how to approach challenges pragmatically. Some of the slogans you repeated significantly impacted my thinking and will for sure accompany me on my future career, such as ‘it doesn’t matter how much you invested so far, is it really relevant in the present?’ or ‘don’t overcomplicate, what is the easiest approach to solve this problem?’.

Next, I want to thank the present and former members of the Selbach lab that I had the pleasure to work with: Boris, Katrina, Michal, Amir, Florian, Anna, Christopher, Trëndelina, Robert, José, Martha, Kamila, Carlos, Fritz, Henrik, Joao, Mirjam, Christian, Marianna, Atakan and Maximilian. From each of you I learned a lot, not only scientifically but also how to approach challenges, how to deal with frustration, how to organize myself in research, and how to approach a project of such a long time-span and with so many uncertainties. Thank you for creating such a welcoming and supportive atmosphere, for all the lunch and coffee breaks, the in- and outside of science discussions, the good vibes, and for having me a hand whenever I asked for help.

I was blessed to work on an interdisciplinary project and in close collaboration with the lab of Marc Nazaré. Big thanks to Peter Lindemann for the synthetic work, for your insights and all the discussions on the project. Thank you Marc for attending all of my committee meetings and for always having an open ear for discussion and chemistry-related questions. Thanks also to Jerôme Paul for the inspiring discussions on the hallways and for lending me your brain for some minutes.

I also want to express my gratitude to Prof. Gerhard Wolber for being part of all my committee meetings, for the scientific input and for agreeing to exert the university supervision. Thank you for reading my thesis and for participating in my defence.

Special thanks also go to Martha Hergeselle, who helped me with cell culture and in the biology wetlab, Christian Sommer for keeping the lab together and for saving several of my mass spectrometry runs, Mohamed Haji for the automated sample preparation on the Agilent Bravo and to the people of the proteomics community (Philipp Mertins lab, Ilaria Piazza lab, Fabian Coscia lab) at the MDC for all the scientific input and the fruitful discussions. I also want to thank the people of the library on the campus Berlin Buch who let me use their facility to write major parts of this thesis. Special thanks also to Sabine and Petra who helped me with all administrative hurdles.

I also want to explicitly thank some people who supported me not only scientifically but also mentally along this journey. Thank you Trëndelina, Robert, Florian, Anna, Carlos and Christopher for always having me a shoulder and encouraging words in frustrating moments, for making my work more fun, for discussing scientific challenges with me when I got stuck, for inspiring me and for becoming my friends outside of the lab. Thank you Mohamed for lifting me up in difficult times of the PhD, for showing me a very pragmatic way on how to approach mass spectrometry, for sharing your secret tricks and helping me in despairing mass spectrometry machine moments. Thank you all and Anyta for reading parts of my thesis.

Despite all this support, my journey through this PhD was not always easy – especially during lockdown times - and sometimes it was a blessing being able to share my doubts and sorrows with people who already went through a similar journey: thank you Lisa, Jana and Christopher for always lending me an ear. Thanks also to my close friends outside of science for always having my back, for being patient and understanding in times I could not nurture our friendship, and for reminding me that there is a life outside of the lab: Thank you Kate, Vera, Juliette, Micah, Teresa and Susi for going through this emotional rollercoaster together with me and for creating a safe space. I also want to thank my dance coach Sarah Balzat (Yerba Buena Art Community) for teaching me how to build resilience, how to create my own strategy to push through difficulties, and how to transform feelings of failure into constructive

learning. You somehow always manage to bring out the best in me and were so much more than just a dance teacher and friend.

Last but not least, I want to thank my family, but especially my mom Célénie, my sister Louisa and my aunt Karin who supported me all along the way. Thank you for your compassion and unconditioned support, for listening to me, for enduring long frustrated phone calls, and for making me laugh again even if I called being down in tears. Thank you for believing in me and for being there.

DECLARATION OF INDEPENDENCE

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me. This dissertation has not yet been presented to any other examination authority in the same or a similar form and has not yet been published.

Berlin, _____

_____

Fiona Teresa Ines Melder

TABLE OF CONTENTS

SUMMARY

Proteins represent the main target class of bioactive compounds.[1] The identification of a drug's protein target is crucial to understand drug function on a molecular level. Major progress has been achieved in both the construction of large scale screening libraries of structures with drug-like properties as well as in high-throughput methods to screen the bioactivity of those substances against a phenotype of interest. Such phenotypic screens allow to identify hit compounds that alter the disease phenotype in a desired manner.[2,3] However, the following deconvolution of the mode of action (MoA) and identification of corresponding protein targets remains a key challenge in drug discovery.[2,4]

Here we developed a novel target identification (target-ID) approach, which combines the power of photoaffinity labeling (PAL) and quantitative affinity purification mass spectrometry (qAP-MS). This compound interaction screen on a photoactivatable cellulose membrane (CISCM) allows us to probe the interaction of many proteins with multiple compounds in parallel without compound derivatization or structure-activity relationship (SAR) studies being required. To achieve this, we equipped cellulose membranes with a diazirine PAL-probe and photoimmobilized a selection of well-studied bioactive compounds with ultraviolet (UV) light. The resulting photocrosslinked compound cellulose array was then incubated with protein extract and specific targets were identified via quantitative affinity purification mass spectrometry.

For the natural structure analogues cyclosporine A (CsA), tacrolimus (FK506) and sirolimus (rapamycin) this reliably identified known protein interactors, while the known targets of small fragment-like compounds, such as (S)-thalidomide, lenalidomide, methotrexate and metformin, could not be detected. As another limitation, the target tubulin beta (TUBB) could not be identified as a specific interactor of the pharmaceuticals vinblastine and paclitaxel. The observed results were nearly unaffected by the increase of the spacer-arm length, separating the diazirine from the cellulose surface, and by photocrosslinking the compounds isolated from the cellulose membranes with subsequent immobilization of resulting photocrosslinking products. Physisorbed CsA on diazirine-modified non-UV irradiated cellulose membranes could identify the known interactors Peptidyl-prolyl cis-trans isomerases A and F (PPIA,

PPIF), whereas physisorbed CsA, sirolimus and FK506 on non-functionalized cellulose membranes could not identify any of their known protein interactors.

In summary, we developed a novel compound interaction screen that can rapidly screen for drug targets in a parallel fashion without prior derivatization of the drugs. CISCM reliably identifies the targets of natural products and current limitations of our method are due to the known limitations of diazirine-based photocrosslinking strategies and affinity purification mass spectrometry.

ZUSAMMENFASSUNG

Proteine stellen die vorwiegende Target-Klasse bioaktiver Verbindungen dar.[1] Die Identifizierung des Targets einer bioaktiven Substanz ist entscheidend, um die Funktion dieser Substanz auf molekularer Ebene zu verstehen. Es wurden erhebliche Fortschritte erzielt sowohl in der Konstruktion großer Screening-Bibliotheken, die wirkstoffartige Substanzen enthalten, als auch in Hochdruchsatz-Methoden zur Untersuchung der Bioaktivität dieser Substanzen gegenüber eines relevanten Phänotyps. Solche phänotypischen Screenings ermöglichen die Identifizierung von Wirkstoffverbindungen, die den Krankheitsphänotyp in gewünschter Weise verändern.[2,3] Die anschließende Aufschlüsselung des Wirkmechanismus und die Identifizierung der entsprechenden Protein Targets bleiben jedoch eine zentrale Herausforderung in der Arzneimittelentwicklung.[2,4]

Hier wird ein von uns neuartig entwickelter Ansatz zur Target Identifizierung (target-ID) vorgestellt, der die Vorteile der Photoaffinitätsmarkierung (*photoaffinity labeling*, PAL) und der quantitativen Affinitätsreinigungs-Massenspektrometrie (*quantitative affinity purification mass spectrometry*, qAP-MS) kombiniert. Dieser Substanz-Interaktionsscreen auf einer photoaktivierbaren Zellulosemembran (*compound interaction screen on a photoactivatable cellulose membrane*, CISCM) ermöglicht es uns, die Interaktion vieler Proteine mit mehreren Verbindungen parallel zu untersuchen, ohne dass eine Derivatisierung der Substanz oder Untersuchungen der Strukturaktivitätsbeziehung (*structure activity relationship*, SAR) erforderlich sind. Um dies zu erreichen, haben wir Zellulosemembranen mit einer Diazirin-PAL-Sonde ausgestattet und eine Auswahl gut erforschter bioaktiver Substanzen mittel Ultraviolettem Licht photoimmobilisiert. Das resultierende Substanz-Zellulose-Array wurde dann mit Proteinextrakt inkubiert und spezifische proteinogene Interaktionspartner wurden mithilfe quantitativer Affinitätsreinigungs-Massenspektrometrie identifiziert.

Für die natürlichen Strukturanaloga Cyclosporin A (CsA), Tacrolimus (FK506) und Sirolimus (Rapamycin) wurden auf diese Weise bekannte proteinogene Interaktionspartner zuverlässig identifiziert, während die bekannten Targets kleiner fragmentartiger Verbindungen wie (S)-Thalidomid, Lenalidomid, Methotrexat und Metformin nicht nachgewiesen werden konnten. Als weitere Einschränkung konnte

das Target Tubulin-Beta (TUBB) nicht als spezifischer Interaktionspartner der Wirkstoffe Vinblastin und Paclitaxel identifiziert werden. Die beobachteten Ergebnisse wurden kaum beeinflusst durch das Verlängern des Platzhalters, welcher das Diazirin von der Zellulose-Oberfläche trennt, und durch die Photoreaktion der Substanzen in Isolation von den Zellulosemembranen mit anschließender Immobilisierung der resultierenden Photoreaktionsprodukte. Physisorbiertes CsA auf Diazirin-modifizierten, nicht UV-bestrahlten Zellulosemembranen konnte die bekannten Interaktionspartner PPIA und PPIF identifizieren, während für physisorbiertes CsA, Sirolimus und FK506 auf unfunktionalisierten Zellulosemembranen keine der bekannten Proteininteraktoren identifizieren konnten.

Zusammenfassend haben wir einen neuartigen Substanz-Interaktionsscreen entwickelt, der in der Lage ist die proteinogenen Interaktionspartner vieler Wirkstoffe gleichzeitig zu suchen, ohne dass eine vorherige Derivatisierung der Wirkstoffe erforderlich ist. CISCM identifiziert zuverlässig die Targets natürlicher Strukturanaloga und die derzeitigen Einschränkungen unserer Methode beruhen auf den bekannten Einschränkungen von Diazirin-basierten Photoreaktionen und der Affinitätsreinigungs-Massenspektrometrie.

STATEMENT OF CONTRIBUTION

This work would have not been possible without the help of several people, whom I will list here and mention again in the corresponding method section of chapter 5.

The synthesis and analysis of all photocrosslinkers was performed by Dr. Peter Lindemann (Nazaré Lab, FMP Berlin) as well as the functionalization of the cellulose membranes (until cellulose product 3 displayed in Figure 18) of experiments described in the following figures: Figure 19, Figure 20, Figure 24, Figure 25, Figure 28, Figure 29, Figure 30 and Figure 33. X-Ray photoelectron spectroscopy (XPS) measurements were performed and analysed by Vanessa Trouillet (IAM-ESS and KNMFi, KIT). Attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR) measurements were performed by Stefan Heißler (KNMFi, KIT) and analysed together with Peter Lindemann. Time-of-flight secondary ion mass spectrometry (ToF-SIMS) measurements were performed and analysed by Dr. Alexander Welle (KNMFi, KIT). Martha Hergeselle (Selbach Lab, MDC Berlin) prepared the HEK293T cell lysates and Mohamed Haji (Proteomics Core Facility, Mertins Lab, MDC Berlin) performed the automated sp3 sample processing. All the created data was discussed and interpreted with me.

All other experiments, data analysis and all figures were created, designed and performed by me under the supervision of Prof. Dr. Matthias Selbach (MDC Berlin), and with input from the members of our lab, collaboration partners and members of the committee meeting (Prof. Dr. Gerhard Wolber, FU Berlin; Dr. Marc Nazaré, FMP Berlin). The project was supervised by Prof. Dr. Matthias Selbach, who is also the originator of the project idea.

TABLE OF FIGURES

ABBREVIATIONS

abbreviation	meaning
AAC	azide-activated cellulose
ABC	ammonium bicarbonate
AP	affinity purification
ATR-FTIR	Attenuated total reflection Fourier-transform infrared spectroscopy
CCCP	compound-centric chemical proteomics
CISCM	compound interaction screen on a photoactivatable cellulose membrane
CM	cellulose membrane
CsA	cyclosporine A
CuAAC	Copper(I)-catalysed alkyne-azide cycloaddition
Cyp	Cyclophilin
DDA	data-dependent acquisition
DIA	data-independent acquisition
DMAP	4-Dimethylaminopyridine
DMF	dimethylformamide
DNA	deoxyribonucleic Acid
DTT	dithiothreitol
EA	ethanolamine
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
eq.	equivalent
ESI	electospray ionisation
FC	fold change
FDA	food and drug administration
HEK293T	human embryonic kidney cells 293, antigen T
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
HTS	high-throughput screening
iBAQ	intensity based absolute quantification
IMiD	immunomodulatory imide drugs
IR	infrared

IRE	internal reflection element
KIT	Karslruhe Institute of Technology
LC-MS/MS	liquid chromatography tandem mass spectrometry
LFQ	label-free quantification
LiP-Smap	limited proteolysis-small molecule mapping
LysC	lysyl endopeptidase
m/z	mass-to-charge ratio
MDC	Max Delbrück center for molecular medicine in the Helmholtz association
MG	molecular glue
MoA	mode of action
MQ	MaxQuant
MS	mass spectrometry
MTX	methotrexate
MW	molecular weight
NAC	NHS-activated cellulose
NHS	*N*-hydroxysuccinimide
PAL	photoaffinity labeling; photoaffinity label
PDB	protein data bank
PDE6D	phosphodiesterase 6D
PEG	polyethyleneglycol
PISA	proteome integral solubility alteration
PP	pulsed proteolysis
PPI	peptidyl-prolyl cis-trans isomerase
PRISMA	protein interaction screen on peptide matrix
PTM	post-translational modification
qAP-MS	quantitative affinity purification mass spectrometry
RIPA	radioimmunoprecipitation assay
SAR	structure-activity relationship
SILAC	stable isotope labelling by amino acids in cell culture
SPROX	stability of proteins from rates of oxidation
StageTip	Stop-and-Go extraction tip
target-ID	target identification

TEMPO	tetramethylpiperidinyloxyl
TFA	trifluoroacetic acid
TFC	TPD-functionalized cellulose
THF	tetrahydrofuran
TMT	tandem mass tag
ToF-SIMS	time-of-flight secondary ion mass spectrometry
TPD	trifluoromethylphenyldiazirine
TPP	thermal proteome profiling
TUBB	target tubulin beta
UV	ultraviolet
XlogP	enhanced atomic logarithmic partition coefficient
XPS	X-Ray photoelectron spectroscopy

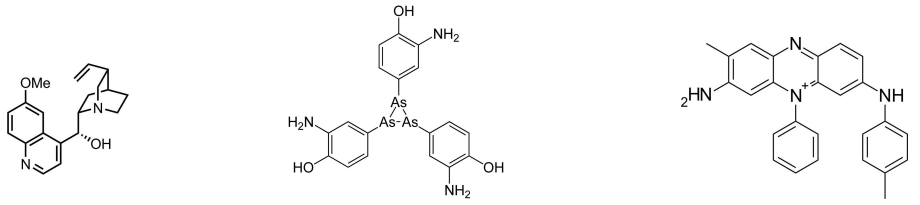
Science doesn't always go forwards. It's a bit like doing a Rubik's cube. You sometimes have to make more of a mess with a Rubik's cube before you can get it to go right.

Jocelyn Bell

1. Introduction

1.1. Drug discovery

The wish to treat and cure human disease dates back far in time to the first advanced civilisations, when drugs of plant, animal or mineral origin were used. Over time, these were further developed and collected systematically. One example of an early encyclopaedia of traditional medicine is the Bencau Gangmu written by Li Shizhen and published in 1590 in China.[5] With advancing technology, the understanding of such natural medicines significantly increased and allowed to identify bioactive natural structures, whereas the emergence of synthetic organic chemistry opened the door for their improvement and total-synthesis. On the basis of alkaloids and other substances isolated from plants, the synthesis of pharmaceuticals was intended. This often went hand in hand with the discovery of novel dyes or vice versa, as it can be demonstrated by the famous example of aims to synthesize quinine[6] - a potent substance to treat malaria – that led to the discovery of the first aniline dye mauveine (Figure 1).[7] But also the other way around, the following flourishing dye industry had a significant impact on the discovery of drugs. For instance, Robert Koch investigated the antibacterial and antiparasitic properties of several dyes and Paul Ehrlich developed with Salvarsan the first chemotherapeutic agent and a treatment against syphilis, which was produced by the dye factory Hoechst (Figure 1).[8] During this research Paul Ehrlich studied the relationship between the chemical composition of drugs and their mode of action (MoA) in an organism,[9] which formed the basis for the later postulated concept of the structure-activity relationship (SAR). Another milestone in understanding the molecular action of bioactive substances was the postulation of the lock-and-key model by Emil Fischer in 1894, which described the interaction between enzymes and their substrates.[10] All these efforts laid the groundwork for our modern understanding of drug function and modern drug design.



Quinine Salvarsan Mauveine

Figure 1. Structures of quinine, salvarsan, and mauveine.

1.1.1. Drug discovery approaches

In the 19th-century, many drug discoveries were related to coincidence, whereas modern drug discovery and development works in a more systematic and high-throughput manner. Generally, the drug development process can be separated into four different stages: drug discovery, preclinical testing, clinical testing and FDA-approval (Figure 2, top). In drug discovery, a library of compounds is screened using high-throughput methods to identify a ‚hit' compound that shows the desired bioactivity in the disease of interest. This ‚hit' compound is further derivatized to an optimised lead structure, which then enters the preclinical phase to test its toxicity and efficiency profile in the cellular and organismic level and to identify a safe dose in animals for the following clinical testing in humans. Only a very small amount of compounds identified as hits in high-throughput screening (HTS) make it through those first steps to be tested in patients. Clinical testing can be separated into three phases with increasing patient cohort sizes. Only when proven safe and efficient during these clinical trials are they considered for FDA-approval.

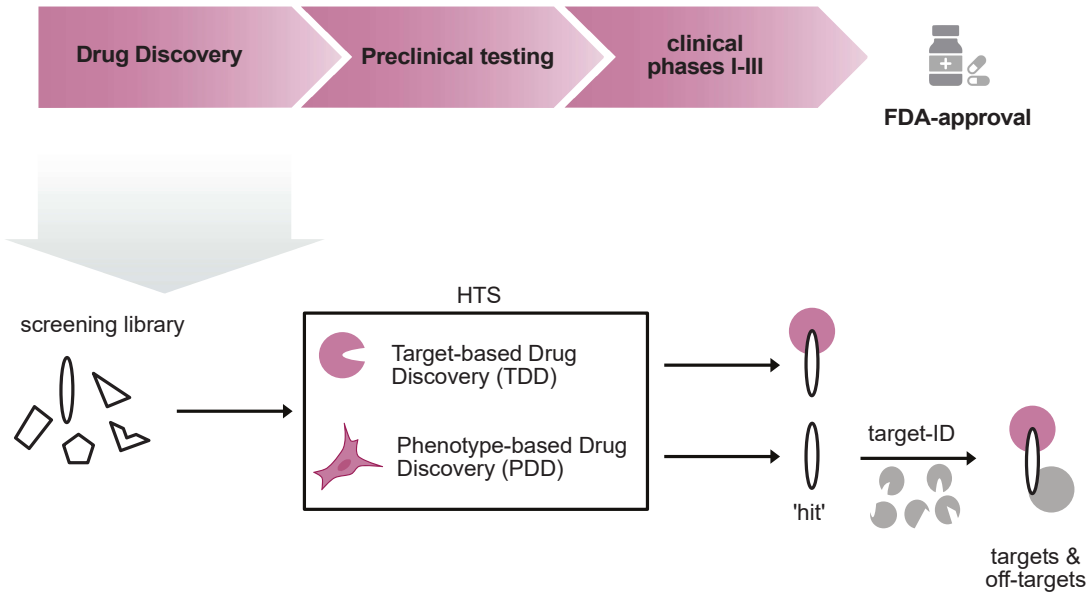


Figure 2. Drug discovery and development. Top: Different stages of drug development starting from drug discovery until its approval by the FDA. **Bottom:** The two main drug discovery approaches. A screening library containing compounds with suspected bioactivity is screened against one or few targets of interest (target-based drug discovery, TDD) or against a phenotype of interest (phenotype-based drug discovery, PDD) in a high-throughput format (high-throughput screening, HTS). The identified 'hit' compound from PDD requires the identification of molecular targets and off-targets.

In more detail, during drug discovery, the ‚hit' compound is validated by orthogonal methods and undergoes a first limited optimisation to identify a lead structure (hit to lead optimization, H2L) on which to follow up. This lead structure enters more intense structuralsoptimization (lead optimisation) for improved potency, better pharmacodynamics (PD, effect of drug on organism), pharmacokinetics (PK, effect of organism on drug) and absorption, distribution, metabolism and excretion (ADME) properties. For better toxicity and efficiency profiles, it is advantageous to identify off-target activities of the lead structure. The structure is sequentially side chain modified and corresponding positive and negative effects on target binding affinities and selectivities are studied.

Two main drug discovery approaches are differentiated to identify a ‚hit' compound (Figure 2, bottom): target-based drug discovery (TDD) and phenotypic drug discovery (PDD).[2,3,11] In TDD, also called reverse pharmacology, a protein that is a key player in a disease of interest and that might be a critical intervention point is chosen and screened against a library of substances with suspected bioactivity. The aim of this approach is to identify hit substances that can functionally modulate the malfunctioning protein of interest. The TDD approach has become very attractive due to the advancements in both, the understanding of disease mechanisms and in high-throughput screening technologies.[11] However, it focuses only on known druggable proteins in their isolated form. This includes proteins capable of binding orally bioavailable drug-like substances, that are functionally linked to a disease, and that can be modulated without disturbing other not disease related processes. Even though its number is expected to be much higher, the currently known druggable proteins represent only a minor fraction of the entire proteome.[2,12] Furthermore, proteins rarely carry out their function in isolation, instead they are typically part of complex protein networks. Their malfunctioning can result in a cascade of dysregulated protein signaling, leading to the pertubation of the cellular system and eventually leading to disease. Therefore, studying drug action with TDD approaches only displays part of a bigger cellular picture.

In contrast, in PDD, a library of compounds is screened against a model system of the disease, such as a cellular assay, and changes in its phenotype are monitored. Compounds that alter the phenotype in a desired manner are identified as hits. Consequently, the modulation of a disease phenotype rather than the modulation of

single protein function is taken as a measure. Hence, the impact of a compound on a complex living system is investigated in its entirety.[3] No a priori knowledge of targets or mode of action (MoA) is required for phenotypic screening, and proteins, annotated as undruggable, are screened in parallel to known druggable proteins. This has led to the discovery of new chemical entities and ‚first-in-class' drugs, such as the majority of clinically used antimalarial drugs.[13] Furthermore, hit compounds of PDD screens are more likely to be bioactive when tested in vivo. However, in contrast to TDD, the molecular target or targets of a hit compound that lead to the desired phenotypic effect is not known, nor are possible off-targets that might lead to toxicity of the hit compound when tested in vivo or that might alter its efficiency. This off-target activity of a hit compound can also be interesting for compounds that do not follow the principle of ‚one gene, one drug, one disease'.[3] For such compounds, promiscous binding plays a key role in reaching their desired pharmaceutical effect. For some compounds, off-target activity might also be relevant for their potential use in different disease indications. Identification of off-targets or entire target classes of a hit compound can shed more light into this phenomenon, also known as polypharmacology, or be used to repurpose established drugs.[14–17] Additionally, the knowledge of the target linked to the desired pharmacological effect allows to optimize the hit's binding affinity amongst other properties during lead-optimization. The lack of target information has also been linked to late stage attrition in the drug discovery process.[18] The identification of protein targets (target-ID) and the underlying MoA of the identified hit compound, however, is a tedious and time-consuming effort and remains a major bottleneck of PDD approaches and will be discussed in more detail in chapter 1.2.2.

Independent on the chosen drug discovery approach and despite major investments into ‚Omics' technologies, drug discovery and development remains a long and twisted road with low sucess rates, time spans over a decade from a hit compound to its approval as a drug by the food and drug administration (FDA) and an average cost of US$2.8 billion.[19] However, innovation in automation, the implementation of artificial intelligence and drug repurposing approaches may shorten this process and increase success rates in the future.[15,19–21]

1.1.2. Natural drugs

Natural products, as already mentioned, have been the major starting point for the discovery and development of numerous drugs.[22,23] Their structure is optimized by evolution to – amongst others - regulate endogenous defence mechanisms against other organisms. This made drug leads especially relevant in cancer and infectious diseases. Probably the most famous early natural drug using the defence mechanism of an organism is the antibiotic penicillin, discovered and developed by Alexander Fleming in the late 1920s, which started the golden age of antibiotic discovery. But also in recent drug discovery, natural products play a significant role. As an example, between 1981 and 2019 1394 small molecule drugs were approved, of which approximately two-thirds (66.7%) were natural products or related to such (Figure 3a).[22] The proportion was even higher for anti-infectives and anticancer drugs. Of the 185 approved anticancer drugs within that same time span only 15.7% were purely synthetic, while the rest could be attributed to natural products and derivatives or synthetic structures with a natural product inspired pharmacophore or natural substrate mimics.[22]

This major contribution of natural products and analogues to pharmacotherapy originates from their remarkable advantages over purely synthetic molecules, which are a result of millenia of optimization by evolutionary pressure. Compared to synthetic molecules they have a higher molecular mass, more H-bond acceptors and donors, a greater number of sp^3 hybridized carbon atoms and oxygen atoms, however less nitrogen and halogen atoms, higher ring content, diminished aromaticity, increased stereochemical content, smaller calculated octanol-water partition coefficients (cLogP values) and hence increased hydrophilicity, and higher molecular rigidity.[23–25] The latter is particularly advantagous for targeting protein-protein interactions whereas the increased saturation of carbon atoms might lead to improved solubility.[24] Natural products, when compared to the purely synthetic drug-like space, have a broader scaffold and functional diversity, a higher structural complexity oftentimes containing ring structures and several stereocenters and provide a wider biologically relevant chemical space.[24] The selection-driven optimization of metabolite structures led to high specificity of metabolite action on pathways of bacterial or animal cells, resulting in higher hit rates for natural product screening when compared to synthetic library screening.[26] Furthermore the oftentimes better compatibility of natural products with

cellular transporters results in more favourable ADME properties.[27] Natural products are the major source of molecules ‚beyond Lipinski's rule of five' - a drug class of increasing significance.[23]

On the other hand, to optimize the properties of natural product leads, their derivatization is often required – a mostly challenging enterprise. This is addressed by recent synthetic efforts, such as late-stage diversification, which aim to improve the synthetic tractability of natrual compounds.[28]

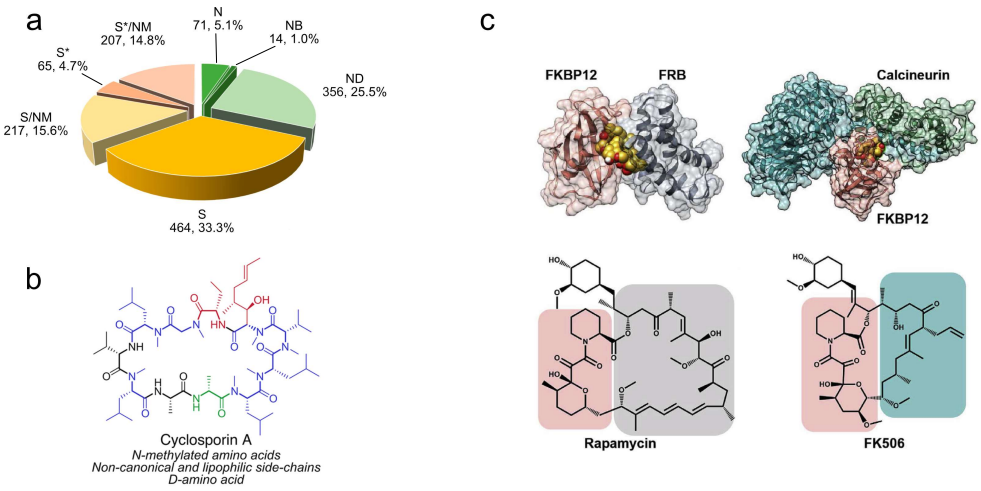


Figure 3. Natural drugs. (a) Prevalence of natural drugs amongst FDA-approvals between 1981 and 2020: Natural product (N), Natural product botanical (NB), Natural product with semisynthetical modification (ND), completely synthetic structure (S), synthetic structure with natural product derived pharmacophore (S*). **(b)** Structure of cyclosporine A with highlighted structural parts: amino acids containing non-proteogenic side-chains (red), D-amino acids (green) and N-methylated amino acids (blue). (c) Structures of rapamycin and FK506 (bottom) with interaction domains highlighted in salmon for FKBP1A, grey for FRB (FKBP-rapamycin binding domain of mTOR), and teal for calcineurin and the corresponding complexes (top). Modified from[22,29,30]

As emphasized, natural products cover a broad chemical space. Their very diverse structures can be classified into chemical scaffolds, such as terpenoides, polyketides, phenylpropanoides, alkaloides and peptides to name a few.[28] Probably the most famous peptide drug of natural origin is insulin, which was approved in the early 1920s for the treatment of diabetes mellitus and which to date remains the commercially most successful peptide drug.[31] However, it took another three decades until the next peptide drug (corticotropin) was released and almost another three decades until the shorter hormones desmopressin, leuprolide and oxytocin were approved. In 1983, the first cyclic peptide was approved: the immunosuppressant cyclosporine A (CsA). Its immunosuppressive action was discovered in 1972 during a screening program for antimicrobial compounds and in its first approval it was administered to prevent organ

rejection in kidney transplant patients.[31] Since then it has also been approved for several other therapeutic applications including liver, heart, skin and bone marrow transplantation and for the treatment of certain autoimmune diseases.[32] CsA is a macrocyclic peptide consisting of eleven amino acids, which build the cyclic backbone of the peptide (Figure 3b). Seven of the amino acids are N-methylated, which allows the cyclic peptide to passively diffuse through biological membranes, a remarkable property when compared to other macrocyclic peptide drugs.[29] Its substential N-methylation and its cyclic structure also contribute to its resistance against proteolytic degradation.[29,31] Due to its structural flexibility – it can be observed in six different conformations in polar solvents – and its hydrophobicity, it can be orally administered.[29] It exerts its function as a molecular glue by binding its primary target cyclophilin A (CypA), which acts as an accessory protein (AP), to form a complex. This complex inhibits calcineurin, which acts as a secondary target (protein of interest, POI). This in turn leads to the inhibition of lymphokine (interleukin-2, IL-2) transcription and consequently to a reduced function of effector T-cells leading to a suppressed immune response.[30,32] However, cyclosporine A does not exclusively bind to CypA. Instead, it has been shown to bind also other members of the Cyclophilin (Cyp) protein family due to their conserved peptidyl-prolyl cis-trans isomerase (PPI) domain, which actively binds to CsA. However, novel Cyp inhibitors have been discovered recently that target an adjacent S2 pocket, which would allow subtype-specific binding.[33] Other important immunosuppressive macrocyclic peptide drugs are rapamycin (sirolimus) and FK506 (tacrolimus). Both bind the FK506-binding protein 1A (FKBP1A, FKBP12) and other members of the FK-binding protein (FKBP) family as their primary targets and exert their function as a molecular glue, however, their secondarily bound protein of interest differs (Figure 3c). The FK506-FKBP1A complex, like the CsA-CypA complex, inhibits calcineurin, however with a tenfold lower concentration required.[30,32,34] In contrast, the rapamycin-FKBP1A complex binds to the mammalian target of rapamycin (mTOR) as its protein of interest and inhibits its function in the post-interleukin-2 receptor mTOR signal transduction pathway, which decreases T lymphocyte activation.[30,32,34]

To date, there have been around 120 different natural product databases and collections published. To make data on natural products more available, the containing information has been organised into in a COlleCtion of Open NatUral producTs (COCONUT) listing over 400,000 non-redundant natural products.[35,36] Collections as

such can serve as a starting point to obtain information about natural products of interest or for selecting candidates for screening projects.

1.2. Mass spectrometry-based proteomics

Biomolecules rarely carry out their function in isolation but rather by their interaction with other biomolecules organized into higher-order structures and networks.[37] Consequently, current research focusses on studying them under conditions close to their physiological conditions rather than in their purified state. To understand the complex interplay and its disruption in diseases, various methods such as genomics, transcriptomics, metabolomics, and proteomics are employed, often in combination with each other in multi-omics approaches. These approaches enable the construction of joint interaction networks of different biomolecules, integrating information from each method. The genome for instance is relatively static and the development of next-generation sequencing technology has made it possible to acquire large datasets in a short period of time, such as sequencing an entire genome in less than a day.[38] Genomic information provides valuable insights into gene function and dysfunction in human diseases, leading to advancements in disease diagnostics, personalized medicine and gene therapy. However, understanding how this genomic information is interpreted in a cell and how it translates into specific phenotypes is crucial. The same gene can give rise to different gene products (proteoforms) and the same genome can result in different proteomes, due to dynamic cellular events. Examples for such events are protein synthesis and degradation, alternative splicing and post-translational modification (PTM), which lead to different gene products.[39] It is estimated that the genome comprises around 20,000 coding genes, which can be translated into over a million different proteoforms, that differ accross individual cells, tissues and disease phenotypes.[40] Yet the differences between different proteoforms deriving from the same gene can significantly impact their structure, function, localisation, interaction and regulation.[37,41] Mass spectrometry (MS)-driven proteomics aims to bridge this gap and to provide additional layers of information, to comprehend biological processes and their disruptions in disease, as proteins play key roles in such.[1] Technological advancements in the past decade have led to the development of high-throughput MS-driven proteomic technologies, enabling the simultaneous identification and quantification of thousands of proteins and their modifications from complex biological

contexts in a single experiment.[37] These technologies have become powerful tools in biomarker identification, drug discovery, disease diagnostics and other fields.[1]

1.2.1. Shotgun proteomics

MS-driven proteomics can either study intact proteins (top-down proteomics)[42] or their their digested form as peptides (bottom-up proteomics).[37,43] Bottom-up proteomics allows for proteome-wide studies and different methods follow a common workflow (Figure 4a): proteins are isolated from their biological source, digested to peptides using a sequence-specific protease (typically trypsin), separated by liquid chromatography (LC) coupled to an on-line electrospray ionisation (ESI) source and analysed by a mass spectrometer. In bottom-up proteomics three main approaches are differentiated, each with its own advantages and suitabilities for different scientific questions: discovery or shotgun proteomics employing data-dependent acquisition (DDA), targeted proteomics and data-independent acquisition (DIA) proteomics.

In all these methods, peptides are separated in dependence of their hydrophobicity using reverse phase capillary chromatography, where hydrophilic peptides elute earlier, while hydrophobic peptides are retained on the column for longer. At the end of the column co-eluting peptides are vaporized and ionised by ESI. The resulting peptide ions are then detected in the mass analyser based on their mass-to-charge (m/z) ratios and a first mass spectrum is recorded, also referred to as full-scan, survey scan or MS^1 scan. This is followed by peptide fragmentation and acquisition of corresponding fragment-ion spectra (MS^2 scans), which differ depending on the acquisition methods mentioned earlier.

In DDA, the top *n* (e.g. top four) most abundant peptides from the full-scan are sequentially selected and fragmented, recording a second mass spectrum for each precursor ion (Figure 4b). This cycle is repeated for the next co-eluting peptide-ion package at the subsequent retention time point. In contrast, when using DIA, peptides are fragmented within a set of constant mass range windows independent of their MS^1 intensities. This is especially relevant for samples containing low abundance but biologically interesting peptides in the presence of peptides derived from highly abundant but biologically irrelevant proteins. In DDA, these peptides might not get selected for fragmentation and thus go unidentified. However, the complexity of MS^2 spectra in DIA still poses challenges for explicit protein identification. In targeted

methods, a predefined set of proteins relevant to the biological question is defined, and mass and retention time ranges of the corresponding peptides are used as fragmentation windows. In general the combination of acquisition on MS^1 and MS^2 level is referred to tandem mass spectrometry or MS/MS and when coupled directly to LC it is called LC-MS/MS. Mass spectrometers with high resolution mass analyzers, such as an orbitrap,[44] are known as high resolution LC-MS/MS instruments.

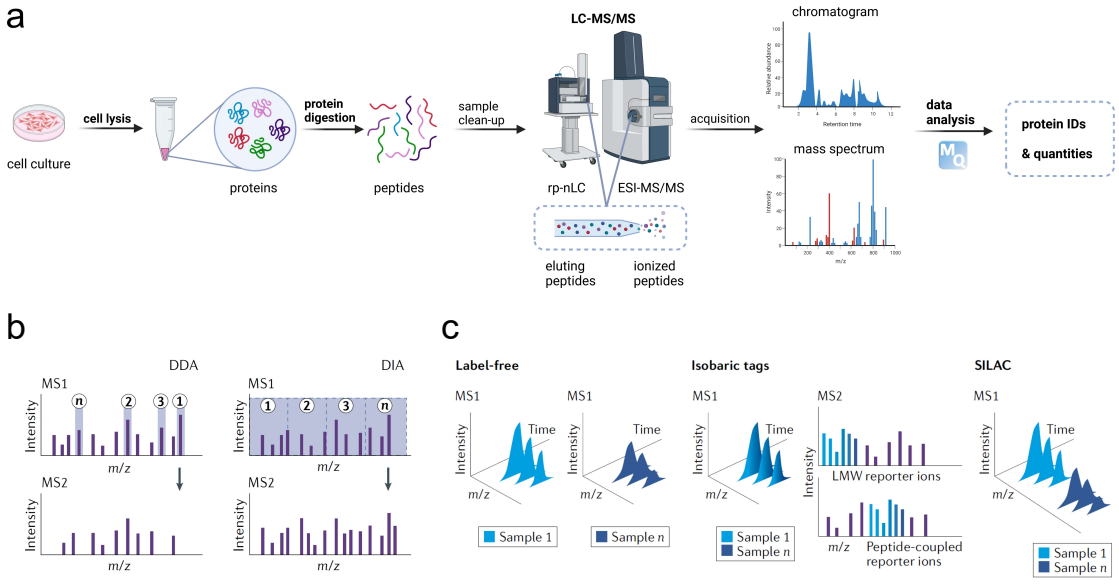


Figure 4. Bottom-up proteomics sample preparation and data acquisition. (a) Proteins are extracted from their biological source, such as cells in culture, and digested. Resulting peptides are desalted and separated by online reverse phase nano liquid chromatography (rp-nLC). Eluting peptides from the chromatography column are vaporized at the tip of the capillary and ionized in the electrospray ionisation (ESI) source of the tandem mass spectrometer (MS/MS). A chromatogram and mass spectra are acquired. Using the combined data, the proteins of the sample are identified and quantified in a computational platform such as MaxQuant (MQ). **(b)** Data-dependent acquisition (DDA) and data-independent acquisition (DIA). DIA: Acquisition of a full-scan (MS^1) and selection of a predefined number (*n*) of the most abundant peptide ions (1, 2, 3, …, n) followed by fragmentation scans (MS^2) for each of these selected precursor ions. In this schematic example *n* precursor ions are selected and fragmented within one cycle. This procedure is proceeded for each acquired full-scan. DIA: Fragmentation of peptides within a set of a predefined amount (*n*) of m/z-windows in the full-scan, resulting in more complex fragmentation spectra, when compared to DDA fragmentation spectra. **(c)** Commonly used quantification methods in high resolution LC-MS/MS: label-free quantification (LFQ), tandem mass tags (isobaric tags) and stable isotope labeling by amino acids in cell culture (SILAC). Figures (b) and (c) adapted from[11]

In proteomics experiments using a LC-MS/MS instrument three parameters are measured for each detected peptide ion species: the retention time in the LC system, the m/z and the intensity. The peaks in each of these MS^1 and MS^2 scans (mass spectra) can be considered three-dimensional objects, which can be detected and analysed in computational proteomics platforms such as MaxQuant.[45–47] To identify

the proteins present in the sample MaxQuant correlates the MS^2 fragment masses to in silico digested protein databases of the organism of interest. The quantification information depends on the chosen acquisition and quantification method (Figure 4c).

Absolute quantification, the aims is to ascertain the copy numbers or concentrations of a protein within a sample. In contrast, relative quantification methods involve ascertaining a quantitative ratio or relative change in protein concentrations by comparing intensities across different samples. Both approaches can utilize isotopic or chemical labels or work label-free. Stable isotope labeling by amino acids in cell culture (SILAC)[48] is a relative quantification method that compares isotope intensities at the MS^1 level. It requires incorporating stable isotopes through metabolic labeling at the cellular or organismic level. In contrast, in tandem mass tagging (TMT)[49] peptides from distinct samples are labeled with isobaric tags that have the same mass, but produce different fragment ions used for quantification at the MS^2 level. These tags with up to 18 channels significantly increase throughput and quantitative precision. Label-free quantification (LFQ), on the other hand, does not require sample labeling.[11,37,46] The LFQ algorithm in MaxQuant (MaxLFQ)[50] compares the intensities of MS^1 scan peptide features across different samples and calculates a normalized relative protein intensity. Given its experimental simplicity, compatibility with large-scale proteomic datasets, and suitability for limited sample availability, LFQ has a broad application range in various proteomic experiments.

1.2.2. Chemoproteomic target identification (target-ID) methods

As mentioned earlier, it is crucial to investigate the drug interaction profile of phenotypic screening hits to understand their modes of action, which led to the desired effects on the disease-phenotype of interest and might also imply undesired adverse effects. Chemical proteomics, an interdisciplinary field, develops and combines chemical tools and bioorthogonal techniques, often involving mass spectrometry, to screen bioactive substances in a proteome-wide manner. This allows to study the effect and interactions of bioactive compounds with native proteins in cell lysates or even in living cells, under conditions tailored to preserve the proteins' integrity including their structure, post-translational modifications (PTMs) and their protein-protein interaction networks. Chemoproteomic approaches can be used to study the cellular response to a bioactive compound (global proteome profiling), such as changes in protein abundance,

localization, interactions and PTMs, or for the identification of direct and indirect protein targets of the bioactive compound. Chemoproteomic target identification (target-ID) techniques can be classified into two main approaches: label-free approaches, where the effect of the unmodified compound on the structure and biophysical properties of directly interacting proteins is measured, and approaches that require the introduction of a linker trajectory (label) to attach the bioactive compound to a surface and/or to install further necessary functional groups in order to enrich and analyse the interacting sub-proteome (Figure 5). These methods are considered target-centric approaches as they screen one substance against an entire proteome as a complex target pool.

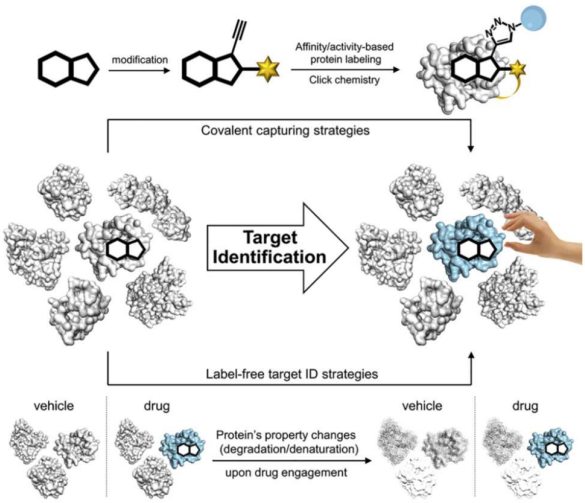


Figure 5. Chemoproteomic target-ID approaches. Top: basic principle of approaches that require the introduction of one or more functional groups (label) via side-chain modification on different sites on the molecule. One of these modifications implies a handle to immobilize the compound to a surface prior to or after the enrichment of interacting proteins. This handle can consist of a spacer with a terminal alkyne as displayed in this example, which can couple the compound to an azide-functionalized surface, such as magnetic beads (blue circle) using click chemistry. Additionally, introduced functional groups (yellow star) can be used to form a covalent bond between the compound and proteins in close proximity, allowing to also capture low affinity interactions. **Bottom:** basic principle of approaches that do not require the structural modification of the compound (label-free). The changes in the proteins' properties upon destabilizing conditions, such as degradation or denaturation, are measured and compared between a drug-treated sample and a vehicle control. Adapted from[2]

Label-free approaches are based on the assumption that drug-target binding stabilizes the protein structure against destabilizing factors such as proteolysis (Figure 6a), chemical denaturation (Figure 6b), or heating (Figure 6c). Based on this the stability of proteins is measured and compared between a drug treated and untreated (vehicle) proteome while being exposed to a stepwise increasing stability-modifying reagent or condition.

In limited proteolysis-small molecule mapping (LiP-SMap, Figure 6a),[51] the compound treated and untreated proteomes at native conditions are exposed to a nonspecific proteinase, such as proteinase K (PK). This proteinase cleaves proteins at their solvent accessible sites, whereas protein-compound binding sites are shielded from proteolysis and result in missing cleavage sites on the peptide level of the proteome. The resulting structure-specific peptides are further digested with a sequence-specific protease such as trypsin and analysed by LC-MS/MS. The LFQ data of both proteomes, drug treated and untreated, are compared and binding proteins can be identified by differences in the LiP patterns. However, this approach was developed to study proteomes of limited complexity and was therefore further developed by integrating a machine learning-based framework (LiP-Quant).[52] Using drug dose titrations, LiP-Quant identifies targets in complex eukaryotic proteomes, such as human cells, and delivers additional information on the target affinity and the predicted binding site at peptide level resolution.

In denaturation approaches, the protein structural stability is perturbed by the increase in guanidinium salt or a chaotropic agent such as urea (Figure 6b) or by the increase in temperature (Figure 6c), as it is the case in thermal proteome profiling (TPP)[53] or in proteome integral solubility alteration (PISA),[54] and measured in a proteome-wide manner using LC-MS/MS. Denatured proteins are more prone to be digested by a protease or oxidized by a reagent such as hydrogen peroxide than intact proteins. This concept is used in pulsed proteolysis (PP)[55] and in stability of proteins from rates of oxidation (SPROX).[56] The binding sites of a protein interacting with the corresponding compound are shielded from the perturbing agent and hence not oxidized or digested. Furthermore, upon thermal stress, proteins tend to irreversibly unfold, exposing their hydrophobic core and making them more susceptible to aggregate. This can be described by the melting temperature (T_m) - a temperature at which half of the proteins are denaturated - and is used in TPP, where proteins are precipitated by stepwise increase in temperature and soluble fractions of the samples are analysed by LC-MS/MS. Most commonly tandem mass tags (TMTs) are used in TPP for quantification, using one mass channel for each temperature point. A multipoint curve is constructed for each identified protein displaying the amount of intact (PP) or soluble (TPP) protein or its methionine oxidation level (SPROX) at each denaturant concentration or temperature for both, a compound treated (red) and untreated (black) proteome. Those

proteins showing a significant shift between these two curves and hence in their thermal, proteolytic and chemical stability upon target binding can be identified as potential targets.

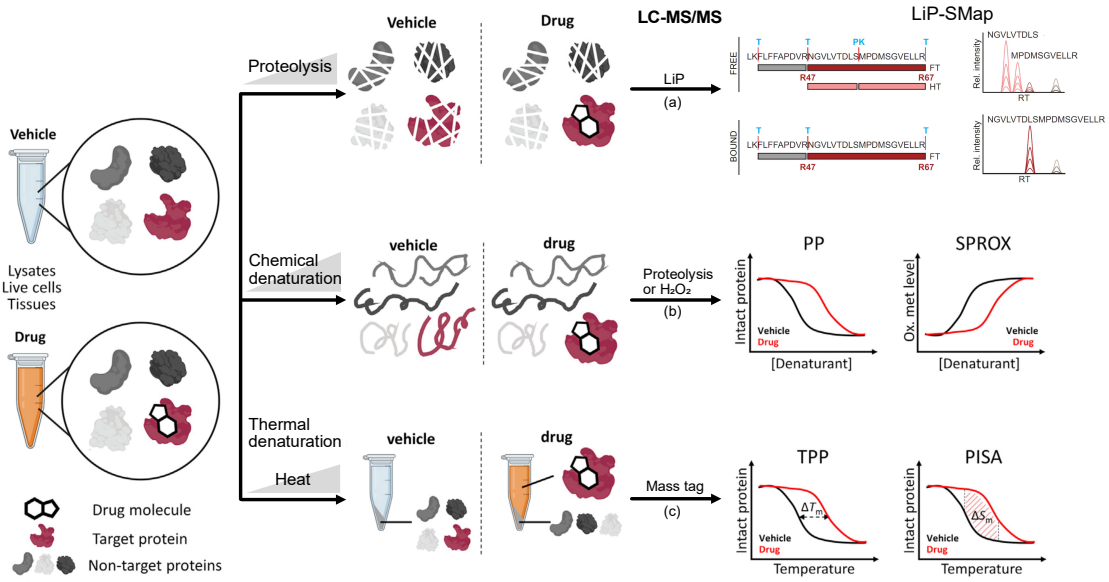


Figure 6. Examples of label-free chemoproteomic target-ID methods: (a) Limited proteolysis-small molecule mapping (LiP-SMap), **(b)** pulse proteolysis (PP) and stability of proteins from rates of oxidation (SPROX), (c) thermal proteome profiling (TPP) and proteome integral solubility alteration (PISA). Melting point shift indicated as ΔT_m and area under the melting curve as ΔS_m. Adapted from[2,51]

All of the above mentioned label-free methods have their own advantages and disadvantages and have been further developed since their original publication to address different challenges.[2,11,57–59] As an example, SPROX and LiP-SMap both deliver domain level binding information of the bioactive compound, whereas TPP allows to screen the underivatized drug even in living cells. However, they all share the common advantage that there is no need to synthetically modify the bioactive compound to screen its interaction landscape. On the other hand, they all require to analyse full-proteomes, which limits their throughput, and augments the risk of missing low-abundance targets.

1.2.3. Affinity Purification Mass Spectrometry (AP-MS) target identification

Unlike label-free methods, affinity- or activity-based target profiling involves engineering the bioactive substance of interest to capture single protein targets or entire sub-proteomes from full proteomes. The resulting LC-MS/MS samples of both affinity and activity enriched sub-proteomes have much lower complexity enabling higher throughput compared to label-free methods. Another main advantage is the

ability to capture low affinity interactions as well.[60] As a shortcoming these approaches require the introduction of one or more chemical modifications with a chemical and/or affinity function as mentioned earlier. Figure 7 illustrates the principles of affinity- and activity-based approaches along with some examples of affinity probe structures.

In the classical affinity-based approach, known as compound-centric chemical proteomics (CCCP), the compound of interest is modified with a handle to immobilize it onto a surface and proteins are enriched from a full proteome based on their affinity toward the modified compound (Figure 7a). This principle is also referred to as pull-down of the interacting proteins. The pull-down is followed by mild washing to reduce the amount of nonspecific background while preserving specific low-affinity interactions. The enriched proteins are then digested and analysed by LC-MS/MS. The compound can be directly attached to a surface, such as agarose or magnetic beads (Figure 7c, left). Alternatively, the compound can be modified with a reporter group (functional tag), such as biotin, which can be specifically captured after target enrichment using an avidin-modified surface. Terminal alkynes and azides are commonly used as bioorthogonal tags due to their low impact on compound properties. They are frequently employed for immobilizing the compound of interest via azide-alkyne Huisgen cycloaddition (‘click chemistry’) and have significantly gained in popularity for CCCP.[61] The immobilization can be performed after target enrichment due to the high chemoselectivity of the cycloaddition reaction or the biotin-avidin interaction (Figure 7c, right). However, all of these approaches are limited to study compound-protein interactions in lysed cells, as these (in most cases) non-covalent interactions would be disrupted during cell lysis and protein extraction when studying living cells.

To overcome this limitation, a workaround is to covalently capture the proteins that bind to the compound. This can be achieved by introducing a photoreactive group to the compound in addition to an alkyne-affinity handle (Figure 7d).[62] Photoreactive groups, also referred to as photo-affinity labels (PALs) or photocrosslinkers, are promiscuous binders that can form covalent bonds with a broad variety of functional groups in a UV light triggered reaction. These photocrosslinkers are discussed in more detail in chapter 1.3. The bifunctional compound can then be used to covalently enrich compound-binding proteins in living cells and analyse them using LC-MS/MS. This approach is also known as in situ CCCP and has the significant advantage that drug-

target interactions can be studied in the active proteome, closely reflecting true interactions. However, depending on the choice of the PAL-probe and the duration of UV light irradiation, this can lead to crosslinking of unspecific binders leading to false-positive hits. In some cases compounds are additionally equipped with a fluorophore such as rhodamine to detect compound-protein pairs in gel electrophoresis and analyse fluorescent bands using LC-MS/MS. Both approaches, CCCP and in situ CCCP, are also referred to as affinity purification (AP) approaches.

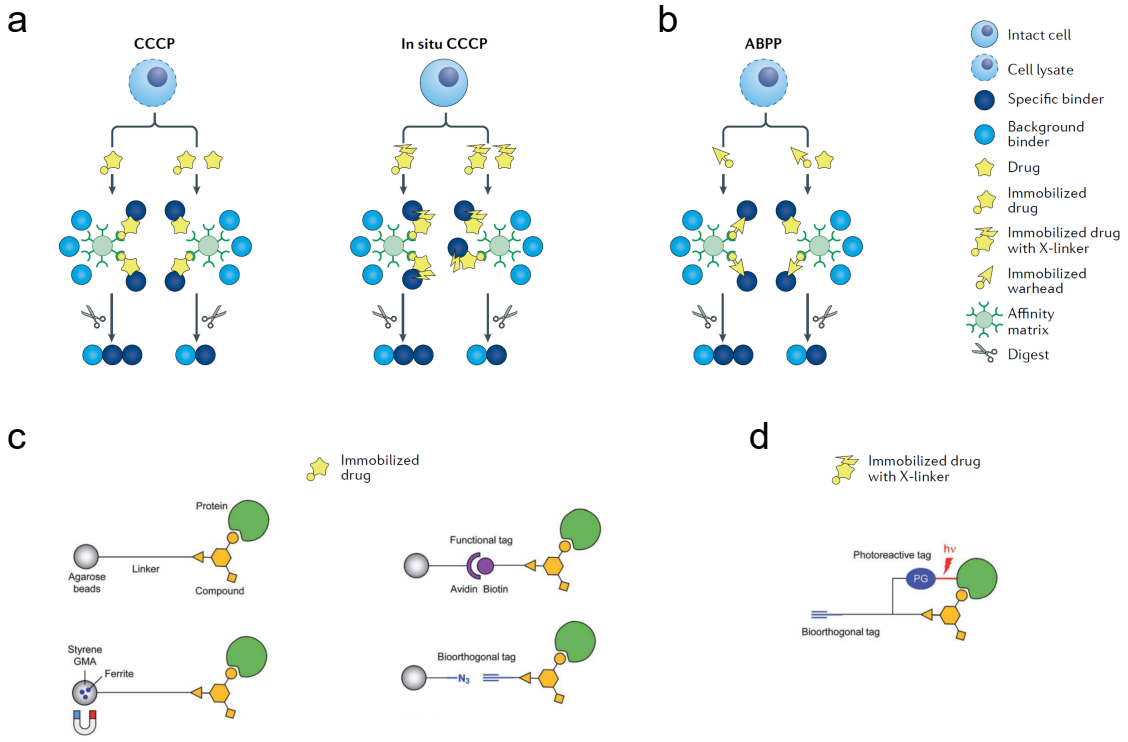


Figure 7. Chemoproteomic approaches using protein affinity or activity to study drug action. (a) Compound-centric chemoproteomics to enrich and identify proteins (blue circles) from cell lysates (CCCP) or living cells (in situ CCCP) using their affinity towards a derivate of a bioactive compound of interest (star). **(b)** Activity-based protein profiling (ABPP) to identify compound-binding proteins from cell lysates based on their activity towards an immobilized substructure (warhead) in concurrence with the free bioactive compound. **(c**) Examples of affinity probes for CCCP. (d) Example of an in situ CCCP affinity probe using a photoaffinity tag to covalently bind compound binding proteins. Adapted from[11,63]

In an alternative approach to CCCP or in situ CCCP, a surface material can be modified with the functional core (warhead) of a bioactive substance or with a mixture of broad-spectrum inhibitors to target a common reactive site or a conserved molecular recognition pocket (pharmacophore) of an entire protein class. In a competition assay, a cell extract is treated with varying concentrations of a bioactive compound or a vehicle control and incubated with the immobilized warhead (Figure 7b). The free

bioactive compound in the samples competes for the binding proteins of the warhead and partially elutes them, depending on the specificity of the interaction and the concentration of the competing compound. Specifically binding proteins are affected by the free compound, whereas nonspecific binders are expected to remain unaffected. The remaining proteins on the beads of the samples and the vehicle control can then be analysed by LC-MS/MS and specific binders can be identified by comparing their intensity. Dissociation constants can be obtained from the concentration-dependent data. This approach has been successfully applied to kinase targets, using affinity beads with broad specificity for several cellular kinases (kinobeads) competing against free kinase inhibitors of interest at different concentrations.[64–66] This strategy of dose-dependent competitive binding can also be applied in CCCP and in situ CCCP to gain additional information on the specificity and affinity of compound-protein interactions.

In an alternative activity-based protein profiling (ABPP) approach, the compound of interest is modified with an affinity tag, commonly a terminal alkyne for subsequent immobilization via cycloaddition to an azide-functionalized solid support, and with an electrophilic group that is reactive towards individual enzyme classes. These enzyme classes share a conserved reactive site or recognition pocket that contains a nucleophilic amino acid, such as a cysteine in cysteine hydrolases or serine in serine hydrolases. However, enzyme classes containing lysine, tyrosine, threonine, aspartate, glutamate, and methionine in their active sites can also be targeted. The electrophilic site-chain modification covalently attaches proteins binding to the compound of interest, which can then be pulled-down using the immobilization handle and analysed by LC-MS/MS. In contrast to promiscuous PAL labels, electrophilic labels are chemoselective.

In affinity-based approaches, despite washing the surface after affinity enrichment of compound-binding proteins, nonspecific background proteins are co-purified. Several strategies have been employed to address this issue. One approach is to compare enriched proteins from disease-related cell lysates with those from normal cell lysates.[67] Another strategy is to use inactive derivatives of the compound of interest as negative controls.[68,69] Additionally, multistep protocols including stringent washing have been implemented. However, especially for complex molecules derivatization requires laborious synthetic efforts and stringent washing can result in losing weak or

transient interactions, thereby increasing the number of false-negative results. Alternatively, affinity purification can be combined with quantitative mass spectrometry (qAP-MS).[3,70] A low stringency single-step protocol can preserve weak specific interactions, but it also leads to the enrichment of more unspecific weak interactions. These unspecific background binders typically bind to the surface material, such as beads, and are sometimes referred to as the 'beadome' or 'bead proteome'. This beadome is expected to remain consistent over affinity enrichments using the same surface material and protein pool (e.g. the same cell lysate). Instead of eliminating this consistent beadome, it can be used to normalize the quantitative data across different affinity enrichment (AE) experiments. To achieve this the proteomic data can be integrated in MaxQuant using the MaxLFQ algorithm and LFQ values can be employed to distinguish specific interactions from nonspecific contaminants. This can be accomplished by pairwise comparison of the normalized LFQ intensities of a replicate group with all other affinity enrichment samples of the other compounds studied in parallel. Statistical analysis, such as a Student's t-test, can be performed and the results can be displayed in a volcano plot.[70–72]

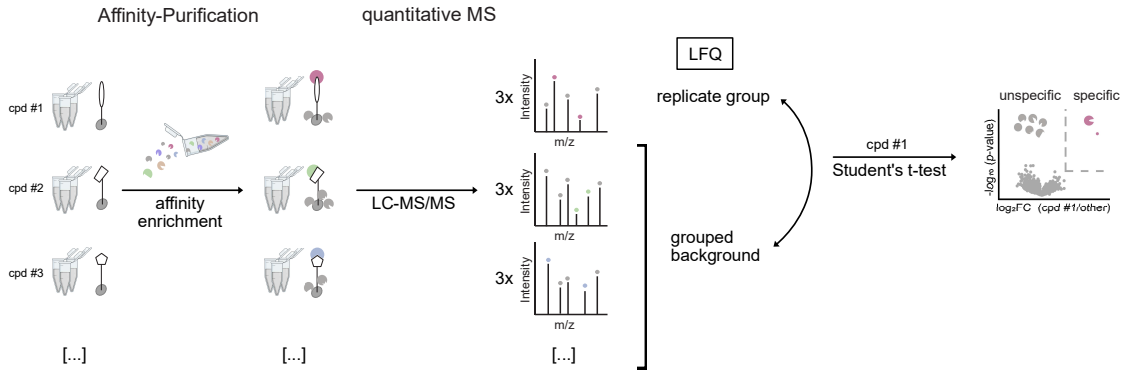


Figure 8. Using quantitative affinity purification mass spectrometry (qAP-MS) to identify specific compound-protein interactions. Triplicates of immobilized compounds (cpd #1-3) are used to enrich proteins from a full proteome. Specific interactions are coloured, whereas unspecific background binders are coloured in grey. Samples are analysed by LC-MS/MS and quantified. Obtained normalized label-free quantification (LFQ) values for the enriched proteins in each sample are grouped for the replicates of one compound, here compound 1 (cpd #1), and pairwise compared with the LFQ values of all other samples grouped as background in a Student's t-test. Results are presented as volcano plots displaying the $\log_2$ fold change (FC) against the Student's t-test derived $-\log_{10}$ p-value. Specific binders are expected with a high fold change and confidence (red). This procedure is repeated for each compound replicate group analogously.

This efficient and automated identification of specific interactors enables large-scale interaction screens. For such screens, cellulose provides a cost-effective, easy to handle, lightweight and printable solid support platform. For example, synthetic

peptides can be immobilized on cellulose arrays via SPOT synthesis[73,74] and used to screen for interacting proteins in whole cell extracts.[75–78] The high local concentration of the peptide ligands on the cellulose matrix and the mild washing conditions in such screens preserve even weak interactions, enabling high-sensitivity interaction screens.[70,72]

Despite these advantages, the required modification of the compound in both affinity- and activity-based approaches introduces an intrinsic bias in its chemical and steric properties. Therefore structure-activity relationship (SAR) studies need to be conducted to identify modification sites on the parent compound that produce a minimal or tolerable effect on its bioactivity upon derivatization. Performing SAR studies is a tedious and time-consuming effort and may unintentionally exclude additional target proteins.[3,18,79] In particular for large natural product derived molecules, synthetic access is often not available or feasible, which limits their suitability for affinity purification target-ID.

In summary, all the discussed chemoproteomic target-ID methods represent powerful tools for studying the target landscape of phenotypic screening hits. They share the common advantage that interactions can be studied under or close to physiological conditions and at endogenous protein expression levels and in the native state of proteins. However, the identified target candidates of both, label-requiring and label-free methods, require further validation to confirm their relationship with the phenotypic effect of the bioactive compound, thereby confirming them as true targets. This validation can be performed through cell culture assays involving knockdown, knockout, or overexpression of the potential target protein or through biophysical experiments with purified proteins to study direct target engagement.

Due to the strengths and weaknesses of each method, chemoproteomic approaches are often used in combination with one another in orthogonal combinatorial target-ID approaches. This can increase the coverage, success rate, and confidence of the target-ID. For example, targets identified through affinity-based approaches can be validated through subsequent label-free target identification, or vice versa. Multiple methods can also be employed in a one-pot approach.[80] The main difference between label-free and labeled approaches, is that immobilization-free methods, such as TPP and LiP-SMap, require long LC-MS/MS measurement times and are therefore limited in their throughput. However, they allow to study compounds without prior

derivatization, whereas affinity-based approaches offer a higher throughput, but require laborious functionalization and SAR studies for linker implementation.

Despite technological advancements in mass spectrometry instrumentation and methods as well as continuous development of novel labeled and label-free chemoproteomic methods, target-ID and MoA deconvolution of phenotypic screening hits remain the most time-consuming process in drug discovery.[2,4]

1.3. Photoaffinity labeling (PAL)

Undirected photocrosslinking, also called photoaffinity labeling (PAL), has emerged as a versatile tool for target-ID approaches. As previously mentioned, one application of PAL is for in situ CCCP approaches to covalently capture compound-protein interactions, however, the application range of PALs in proteomics is continuously growing. In contrast to other chemical reactions, where high chemo- and site-selectivity is desired, PAL approaches use promiscuous binding of photo-reactive groups as an advantage. Such groups are photolabile and become highly reactive against a broad range of different functional groups upon irraditation with ultraviolet (UV) light. This on the one hand allows the photoaffinity label to react with different types of molecules in its proximity and on the other hand it can react with different functional sites on one molecule, leading to a broad spectrum of PAL reaction products. This product distribution depends on the choice of the photocrosslinker, because even though photocrosslinkers do not exclusively react with only one particular functional group, they still show unique labeling preferences.[81–83]

1.3.1. Common photoaffinity labels (PALs) for proteomic applications

The prevalently used photoaffinity labels (PALs), also called photocrosslinkers, are aryl azides, benzophenones and diazirines (Figure 9), each of them with their own advantages and disadvantages and suitabilities for different applications. However, these groups are constantly further developed to achieve improved or altered properties, such as their photochemistry or size, and novel photocrosslinkers are implemented such as α-keto amides[84] or isoxazoles,[85] all leading to a broader application range of PAL for chemical biology applications.[86]

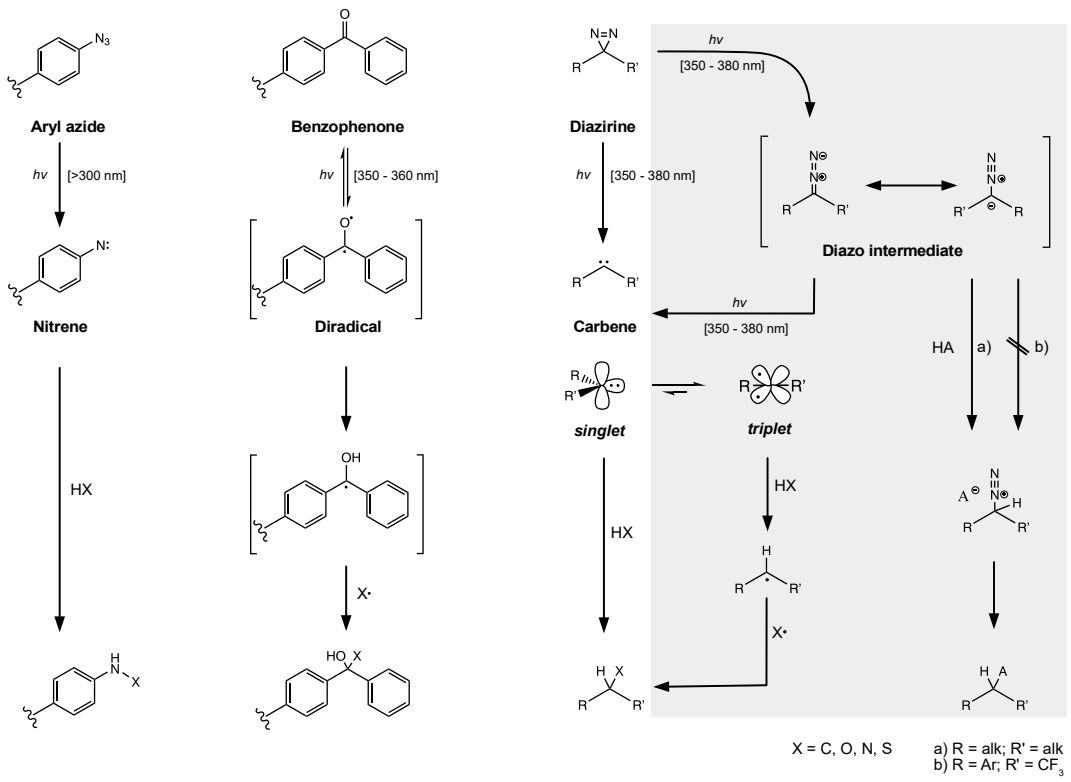


Figure 9. Overview of reactivity pathways of different photoaffinity labels (PALs). Left: Photolysis of aryl azides to nitrenes and their insertion into hydrogen-heteroatom (HX) bonds. **Centre:** Reversible photolysis of benzophenones and stepwise H abstraction and recombination of the formed diradical. **Right:** Photolysis of diazirines to reactive carbene intermediates and their insertion (singlet) into HX bonds. Side-reactions are marked with a grey frame. The formed singlet carbene is in equilibrium with a triplet carbene, which can also insert into HX bond, yet in a stepwise manner.[86] In concurrence to their photolysis to carbene intermediates, carbenes react in a side-reaction to diazo intermediates upon UV-irradiation. The diazo intermediate can convert into a carbene or a) in the case of alkyl diazirines undergo a side-reaction with acids (HA). Graph created using information from different sources.[81,86,87]

Aryl azides, for instance, when irradiated with UV light of rather short wavelengths below 300 nm form reactive nitrenes, which insert predominantly into heteroatom-hydrogen (X-H) bonds. The synthesis of an aryl azide PAL-probe is rather uncomplicated compared to other crosslinking groups, especially for compounds that already contain an aromatic ring. This not only reduces synthetic effort, but is also particularly advantageous for low molecular weight (MW) compounds. On the other hand, aryl azides are sensitive to reduction by thiols and their nitrene intermediates can undergo an undesired side-reaction and rearrange into benzazirines and dehydroazepines or ketenimines.[88] The high energy UV light required for the photoreaction might damage proteins at longer irradiation times - limiting their use for in situ CCCP applications. Another shortcoming is that the efficiency of the photoreaction is lower, when compared to diazirines or benzophenones.

Benzophenones, in contrast, are converted into their reactive diradical state by irradiation at higher wavelengths of the UV spectrum (350-360 nm). The formed singlet diradical quickly converts to a triplet diradical and can undergo X-H abstraction as well as radical recombination with substances in its proximity.[86] Unlike other photocrosslinkers, the photo-activation of benzophenone is reversible and quenching of the formed diradical by water or rearrangement is comparably much slower. Longer UV-irradiation times can hence lead to higher photoreaction yields. However, the efficiency of diradicals formed from benzophenones is lower when compared to the state of diazirines. Additionally, the formed diradicals show a clear preference toward methionine when crosslinked to proximal proteins.[89] Compared to other photocrosslinkers, benzophenone is rather bulky and its implementation to a compound structure - especially for small molecular weight compounds - can introduce a significant bias to its bioactivity.

The most frequently used and most versatile photocrosslinkers in chemoproteomics applications are diazirines.[90,91] Diazirines are rather compact photoaffinity labels and flexible concerning their substitution options. Compared to the other mentioned photocrosslinkers, they are relatively stable and show the most promiscuous binding behaviour. Upon irradiation with UV light between 350 nm and 380 nm (depending on the substituents) they form a reactive carbene or partially undergo isomerization to a linear diazo intermediate. The carbene intermediate, is a short-lived neutral molecule with six valence electrons that can be observed in a triplet or singlet spin state and which is highly reactive against a broad range of different functional groups.[92,93] The other intermediate of diazirine photolysis - the diazo intermediate - has an increased life-time and can transform into a carbene by photolysis or it can - depending on its substituents - stepwise insert into the functional group of a nearby molecule.[88] Labeling preferences towards nucleophilic and polar protic amino acids have been reported when crosslinked with neighbouring proteins in in situ CCCP experiments.[86] However, the efficiency and labeling preferences are highly dependent on the substituents of the diazirine group and their contributing electronic and steric effects.

Mainly aliphatic (alkyl) diazirines and aromatic (aryl trifluoromethyl) diazirines are differentiated, both having their own advantages and disadvantages. In aryl trifluoromethyl diazirines – also referred to as trifluoromethylphenyldiazirines (TPDs) - the α-position of the diazirine group binding carbon is substituted with an electron

withdrawing (trifluoromethyl) and an electron donating group (aryl). Due to the electronegativity of the trifluoromethyl (CF_3) group, electron density is pulled from the diazirine carbon, whereas the aryl group counteracts in pushing π electron density toward the diazirine carbon. For that reason, such diazirines are also referred to as 'push-and-pull' diazirines. The singlet spin state can be stabilized by the electronic effects of the described 'push-and-pull' structure, which in turn effects its reaction preferences. Singlet carbenes preferably insert into heteroatom-hydrogen (X-H) bonds and heteroatom-carbon (X-C) bonds rather than into carbon-hydrogen (C_{sp3}-H, C_{sp2}-H) bonds or non-polar (O_2, H_2) bonds.[94–96] This stabilisation further suppresses the rearrangement of carbenes to alkenes, an undesired side-reaction.[86,88] In addition to stabilization of the singlet carbene, the 'push-and-pull' substitution also stabilizes the diazo intermediate, which leads to a favoured reaction via the short lived and highly reactive carbene intermediate and hence facilitates the suppression of diazo-mediated side reactions.[86,92,97] This makes the photoreaction of TPDs and analogues highly dependent on proximity, very efficient and reactive toward a broad range of groups (Figure 10). As a shortcoming, their bulky structure and low crosslinking yields can be disturbing in some applications. In summary, TPDs are very promiscuous binders and are especially attractive for crosslinking larger compounds.[86]

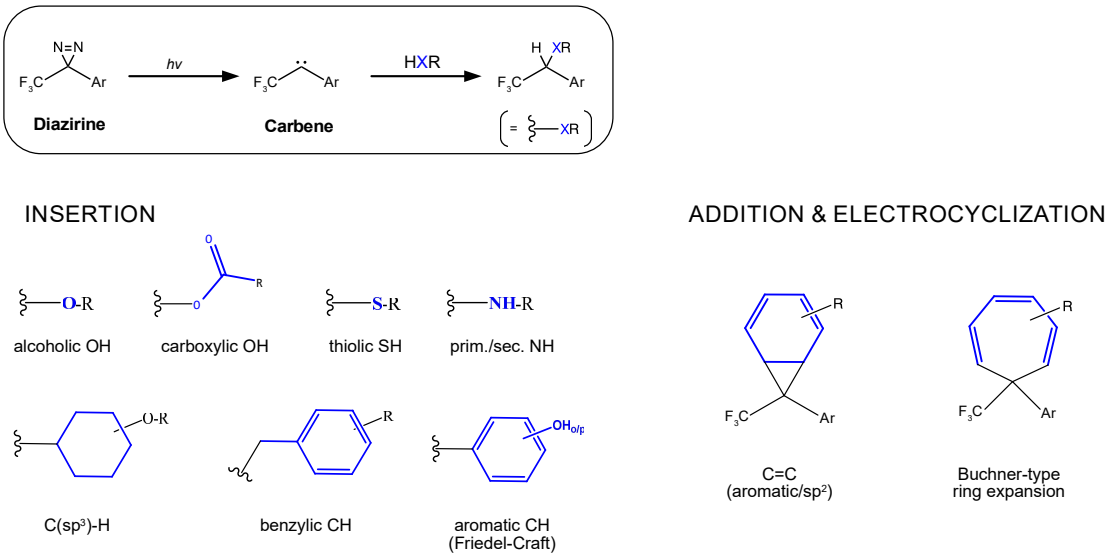


Figure 10. Examples of trifluoromethylphenyl diazirine (TPD) reactivities. Top: General reaction scheme of the diazirine photoreaction. **Bottom**: Examples of diazirine reaction products deriving form insertion, addition, and electrocyclization. Prim./sec. NH: primary or secondary amine, C=C: cyclopropanation of the aromatic C=C double bond, Buchner-type ring expansion: cyclopropanation of the aromatic C=C double bond and subsequent electrocyclic reaction. Graph created using information from[93]

In alkyl diazirines, the diazirine moiety is integrated into an alkyl chain or a terminal position of an alkyl chain. This can result in a very compact structure,[98] which is especially favourable for lower molecular weight bioactive compounds as it introduces a smaller sterical bias to the compound-protein interaction when compared to more bulky photocrosslinkers such as benzophenones. However, in contrast to TPDs, the diazo intermediate of alkyl diazirines is not stabilized and diazo-mediated side-reactions constitute a high proportion of the overall reaction. The preference of the diazo intermediate to react in a pH dependent manner and to acidic surfaces and carboxylic acids with high pKa, introduces a significant undesired bias to the promiscuity of the diazirine photoreaction.[81,99]

In summary, due to their favourable properties and versatility, diazirines have emerged as the photocrosslinkers of choice in many chemical biology applications.[86,90,91,98,100] In these applications often 'push-and-pull' diazirines are preferred due to the comparably higher carbene yields and lower degree of side-reactions caused by the diazo isomer when compared to aliphatic diazirines.[101]

1.3.2. Photoaffinity labeling in affinity purification target identification

As previously discussed in chapter 1.2.2, protein targets of bioactive substances can amongst others be identified by affinity purification (AP) approaches, which generally allow a high throughput. In a conventional AP-based target-ID approach, the substance of interest needs to be modified with a functional handle to immobilize it onto a surface (Figure 11a). However, this modification might introduce a bias toward the bioactivity of the compound: a risk that can be controlled by SAR studies. In a time-consuming and tedious effort a modification site on the compound is searched, which limits the effect on the compound's bioactivity to a tolerable level. On the other hand, derivatization of small molecules without loss of bioactivity remains a bottleneck for affinity-based target-ID despite modern synthesis techniques. Common challenges are small molecules lacking functional groups in a suitable position or limited synthetic access of large and complex molecules such as natural products.

To circumvent the limitation of compound derivatization and SAR studies, PAL has emerged as an attractive alternative to immobilize small molecules onto solid supports for affinity-based target-ID approaches.[102–105] To achieve this, a surface, typically either a glass slide or beads, is functionalized with a photocrosslinker (Figure 11b). In

most cases a trifluoromethylphenyl diazirine (TPD) is chosen due to its previously discussed advantages and attached to the surface using a hydrophilic spacer, such as polyethyleneglycol (PEG). Treating this photoreactive surface with a compound of interest and subsequently with UV light leads to photoimmobilization of the compound to the surface. The resulting photocrosslinked small molecule affinity matrix can then be used to identify binding proteins in an affinity enrichment (pull-down) assay. The promiscous binding of the carbene intermediate leads to various photoreaction products on spot of which a significant fraction is expected to be still bioactive. Hence, for this approach, the unmodified compound can be used and no previous SAR studies are needed.

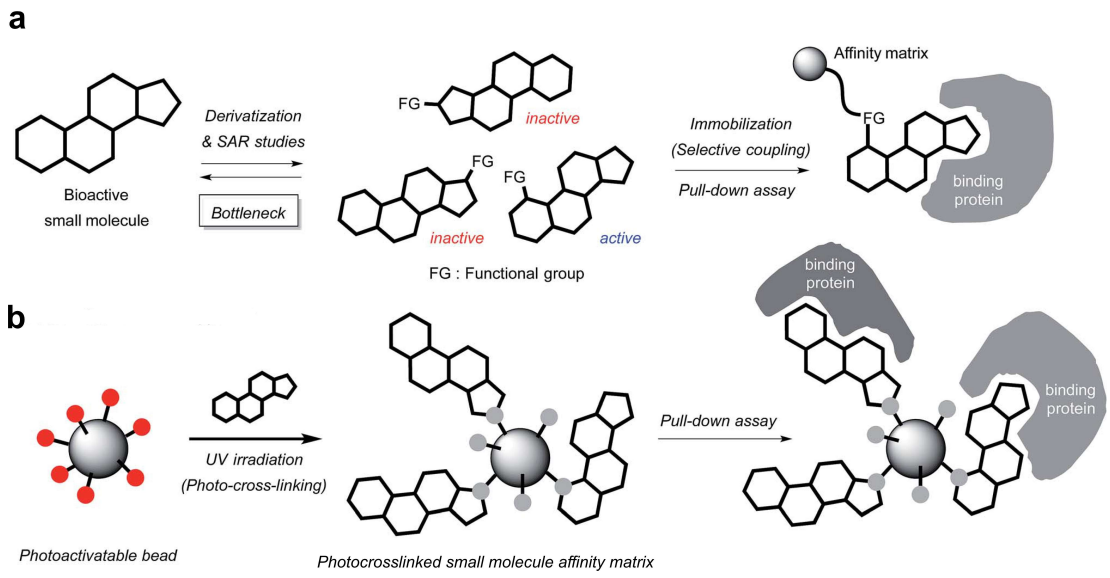


Figure 11. Affinity-based target-ID approaches using selective or unselective coupling. (a) Classical affinity-based target-ID approach: a bioactive small molecule is derivatized with a functional group (FG) at different molecular sites to immobilize it to a surface. The derivative with least affected bioactivity is chosen and selectively coupled to a surface (here a bead) via the introduced functional group. In a pull-down assay, binding proteins can be enriched and analysed. **(b)** Photocrosslinking affinity-based target-ID approach: a surface, such as a bead, is functionalized with a photocrosslinker. The resulting photoactivatable bead is incubated with the compound of interest and treated with UV light to induce the photocrosslinking reaction. The photocrosslinked small molecule affinity matrix is then used to pull-down binding proteins from a target pool. Modified from[93]

This approach was pioneered by Kanoh and co-workers in 2003, where they arrayed a few well studied compounds onto a photoactivatable glass slide and immobilized them in a parallel fashion to screen them against a single or a small set of fluorescently labelled target proteins, isolated or in a whole cell protein extract.[105] This setup on the one hand allowed to screen many compounds in parallel, but on the other hand, the

fluorescent read-out limited the amount of targets that could be screened in parallel (target-centric; multiple compounds, one candidate target protein) and implicated a possible bias to drug-target interactions.[105] This strategy was then further developed and applied to screen photocrosslinked small molecules against a complex target pool, such as a protein cell extract. However, these approaches require tedious experimental procedures to differentiate background binders from specific binders, which limits the throughput of compounds to be screened in parallel (drug-centric; one/few compound(s), multiple target proteins).[103,104] In other applications photoimmobilized drugs were used to screen against a phage display library.[83]

To further explore the reaction behaviour of TPDs with small molecules in such approaches, their photoreaction with different organic solvents or bioactive small molecules has been studied.[82,92,106] This was either performed with the TPD in solution or immobilized to a surface. In the latter approaches a photocleavable linker between the TPD and a surface material was used to detach and analyse the photocrosslinking products of the immobilized linker. This provided an additional layer of information to affinity-based target-ID methods using PAL and allowed researchers to control the photoimmobilization of the small molecules.

1.4. Surface chemistry analytical methods

Methods to analyse the composition of non-transparent surfaces, such as functionalized cellulose membranes, are X-ray photoelectron spectroscopy (XPS), attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR), and time-of-flight secondary ion mass spectrometry (ToF-SIMS).

1.4.1. X-ray photoelectron spectroscopy (XPS)

XPS is a quantitative non-destructive method to not only analyse the elemental composition of a surface, but also the binding state of the containing elements which allows to draw conclusions on present functional groups. The method relies on the photoelectric effect which describes the emission of electrons from a surface when irradiated with a light beam of sufficient energy (Figure 12a). The energy of photons of the light beam is absorbed by the atoms of the irradiated surface. If the acquired energy exceeds the binding energy of an electron in this atom, it can be liberated from its atomic bond. The excess energy contributes to the kinetic energy of the emitted

electron, which is detected and used to calculate the binding energy of the electron with knowledge of the photon energy from the X-ray source (Figure 12b). The binding energy in turn is specific for different chemical bonds and can be used to draw conclusions on functional groups present on the surface of interest. XPS is limited to the depth of 1-10 nm in conventional XPS and 50-200 nm in hard-XPS respectively as electrons in deeper layers of the material undergo significant energy loss by many inelastic collisions.[107] In a typical X-ray photoelectron survey spectrum (Figure 12c), the acquired electron intensities (y-axis) are displayed against their calculated binding energies (x-axis) with binding energy values decreasing from left to right. Peaks are assigned to photoelectrons originated from corresponding elements and specific orbitals according to the calculated binding energy. As an example, the O 1s peak corresponds to photoelectrons originated from electrons of the 1s orbital of oxygen. Following the survey scan, a scan of higher resolution in the binding energy area of interest is recorded, such as for O 1s photoelectrons. The spectral pattern of this high resolution scan is determined by the elements that are bound to the corresponding element and delivers information on the level of functional groups.

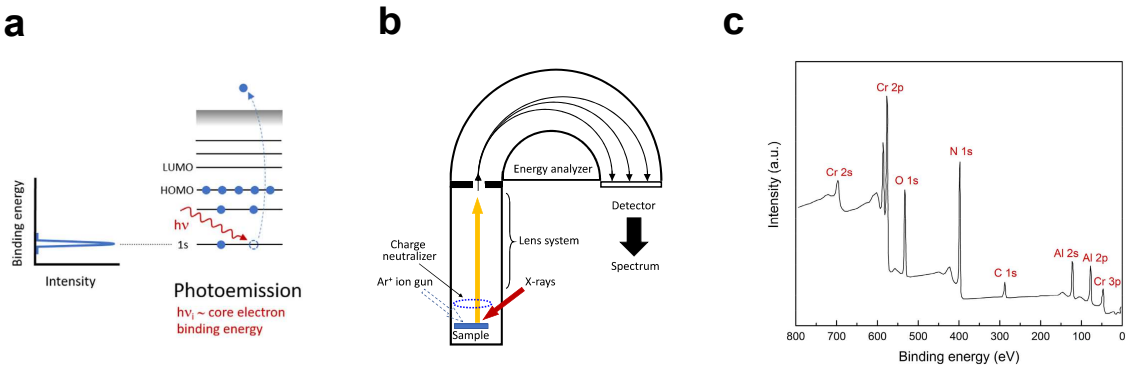


Figure 12. X-Ray photoelectron spectroscopy (XPS). (a) Photoelectric effect. **(b)** Schematic view of a photoelectron spectrometer. **(c)** Typical X-ray photoelectron survey spectrum. Modified from[107,108]

Elements except H and He can be detected with XPS with detection limits at 0.1% to 1.0%.[109] The main advantages of XPS are that it is a fast, non-destructive method that delivers qualitative as well as quantitative information about elemental composition and chemical bonding. The downsides of XPS are that it is a rather expensive and surface-sensitive method with a low probing depth, and the technique requires a high vacuum and trained personnel. Depending on the nature of the sample, the X-ray beam can be destructive and assignment of peaks can be challenging and require additional computational efforts. Further developments address the limitation of a high

vacuum and use hard X-rays and ambient pressure, which allows to perform in situ measurements. In summary, XPS is considered the most simple and direct method to obtain qualitative and quantitative chemical information on solid materials and their surfaces.[107]

1.4.2. Attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR)

ATR-FTIR is a non-invasive method in which an optically non-transparent solid material or liquid is irradiated with infrared (IR) light and the intensity of the reflected light is measured, delivering information about the absorbing material. If light from an IR light source is sent into an internal reflection element (IRE) with a high refractive index and above a certain angle – also referred to as the critical angle – it totally reflects off the inner surface as it propagates through the IRE, creating a standing wave, also referred to as an evanescent wave (Figure 13a). This wave penetrates into the sample, which is in direct contact with the IRE. By the interaction of the evanescent wave with the sample, the IR beam is attenuated before it reaches the detector. The resulting signal is detected and transformed into an IR spectrum (Figure 13b). The depth of sample penetration correlates with the wavenumber and the angle of incidence of the IR light and varies with the refractive index of the IRE and the sample.[110] The refractive index of the IRE must be higher than the refraction index of the sample being studied. Typical measures of penetration depths are between 0.5 µm and 5 µm.[111] The sensitivity of ATR-FTIR depends on the amount of reflections in the IRE, which can be varied by the angle of incidence. IRE is mostly a hemisphere or a prism consisting of a diamond, silicon, zinc selenide or germanium.[110,111]

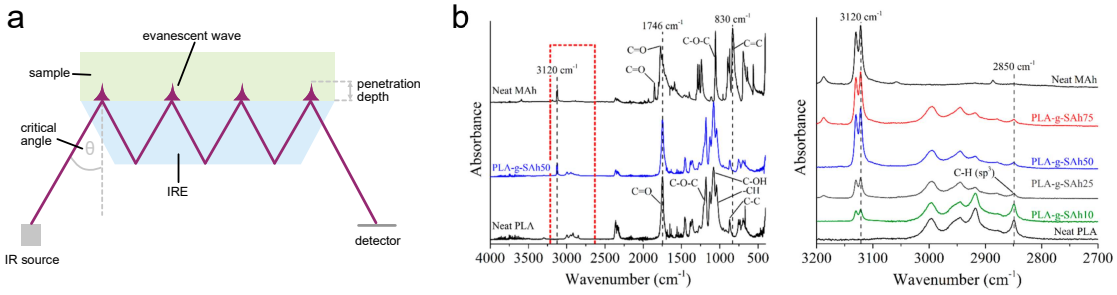


Figure 13. Attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR). (a) Schematic set up of ATR-FTIR. **(b)** Typical ATR-FTIR spectrum displaying the recorded absorbance of the IR light beam between wavenumbers of 4000 cm^{-1} and 400 cm^{-1} (left) and for a region of interest (red frame) between wavenumbers of 3200 cm^{-1} and 2700 cm^{-1} (right) for different polymer samples. Spectral bands were assigned to corresponding functional groups. Spectra taken from[112]

1.4.3. Time-of-flight secondary ion mass spectrometry (ToF-SIMS)

ToF-SIMS is a surface-sensitive method that is able to provide information on the molecular structure of a surface by detection of released secondary ions from a surface upon its exposure to a primary ion beam and under ultra-high vacuum conditions. If a a beam of primary ions in the energy range of keV bormbards a surface, these ions collide with atoms of the surface and transfer them a part of their energy and momentum (Figure 14a). Atoms or groups of atoms that acquire sufficient energy can overcome their binding energy, which leads to the breakage of molecular bonds and the release of molecular fragments, atoms, and ions from the surface. However on their trajectory through the surface material the primary ions significantly lose energy, which leads to the release of the secondary particles primarily from the first two to three atomic layers of a surface.[113,114] Additionally, as excited atoms in the surface collide with each other and exchange energies, the release of the secondary particles can take place in up to a 10 nm distance from the initial impact of the primary ion.[113] This process of multiple collisions, energy transfer and secondary particle release is also named sputting or sputtering.

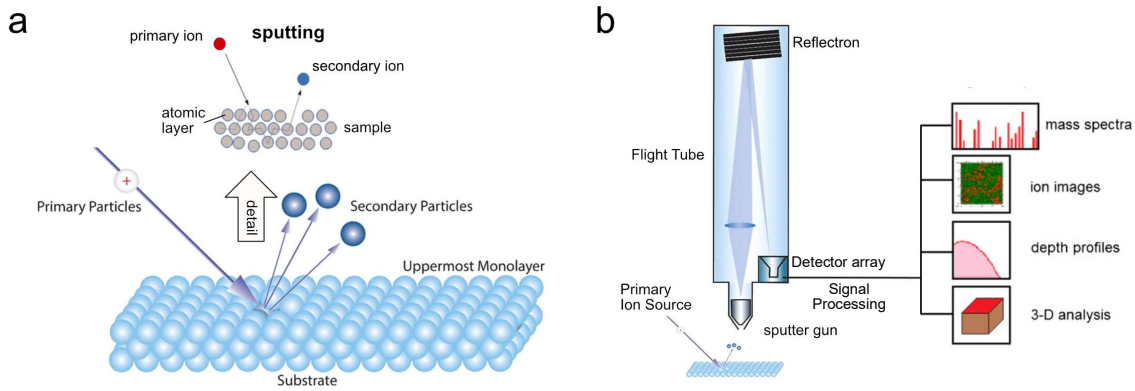


Figure 14. Time-of-flight secondary ion mass spectrometry (ToF-SIMS). (a) Schematic illustration of the sputting process. **(b)** Schematic set up of SIMS with a ToF mass analyzer. Modified from[113,114]

Only a small proportion of the emitted secondary particles is charged (10^{-6}-10^{-1}) and is accelerated at a fixed potential into the field-free drift tube of a time-of-flight (ToF) mass analyzer (Figure 14b).[114,115] The secondary ions at that point have the same kinetic energy and their detected time-of-flight in the drift tube is determined by their mass-to-charge ratio (m/z). ToF-SIMS is able to detect particles in the ppm to ppb ranges, which makes it on the one hand a very sensitive surface analysis method, but also

susceptible to detecting contamination on the surface.[114] In most cases no sample preparation is required and a small piece of a solid material is sufficient for analysis. However, data interpretation can result challenging and quantification of detected ions is limited and often performed using orthogonal surface analysis methods such as XPS.[114]

There are different operation modes of a ToF-SIMS instrument that can be used to perform a variety of different analysis on a surface sample and to obtain information in the form of high resolution mass spectra, ion images (chemical mapping), depth profiles, and 3D analysis (Figure 14b).[113] For the latter two operations, an additional sputter ion gun is required for the ToF-SIMS instrumentation. In the mass spectrometry mode, high resolution mass spectra are acquired using a very short primary ion pulse of several nanoseconds on a predefined spot on the sample (Figure 15). Sputtered secondary ions are then analysed in the ToF mass analyzer and the mass spectrum is recorded together with the coordinates of ion beam on the sample. This can be repeated for the adjacent pixels until a desired area has been scanned and analysed, resulting in a high resolution mass spectrum for each pixel. Peaks in these mass spectra represent secondary ions and peaks of interest can be selected to display their distribution over the entire analysed pixel area (chemical mapping). Such images are also referred to as chemical maps or ion images.[113]

ToF-SIMS can analyse surface materials in a wide mass range and when compared to other imaging surface analysis methods such as matrix-assisted laser desorption ionization time-of-flight (MALDI-ToF), it has a higher spatial resolution of 200 nm to 500 nm.

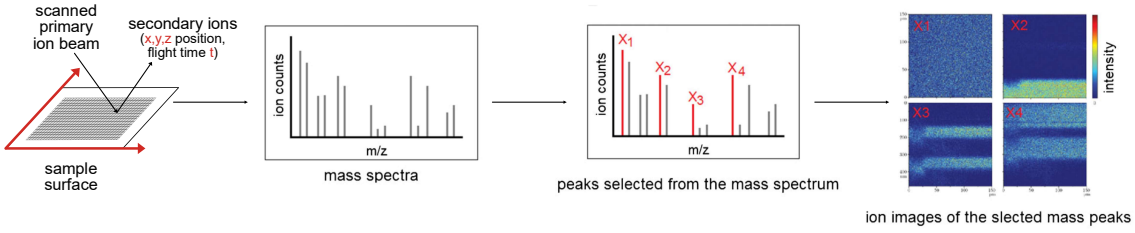


Figure 15. Mass spectrometry and chemical mapping operation modes of a ToF-SIMS instrument. Mass spectra are acquired for each pixel of a defined area on the surface (left). The data is combined to create ion images for selected peaks of a mass spectrum (right). Modified from[113]

2. Aim of the present work

As mentioned before, target deconvolution remains a key challenge in drug discovery. Protein targets of bioactive substances can be identified by affinity purification (AP) approaches, in which a molecule of interest is immobilized onto a surface and interacting proteins are enriched from a target pool using their binding affinity. Analysing the proteins enriched in an AP experiment with mass spectrometry and using the obtained quantitative data allows a high throughput and makes it possible to distinguish specific binders from promiscuous binders whilst preserving low affinity interactions.[70,75,78] However, in classical quantitative affinity purification mass spectrometry (qAP-MS) experiments that use drug-like structures as baits, these structures of interest need to be modified to immobilize them on a surface material. This in turn might bias their binding behaviour towards potential targets; a risk that can be controlled by structure-affinity-relationship (SAR) studies, which are tedious and time consuming, especially for large natural product analogues.[93] For the latter also their synthetic access is often crucially limited or even prohibitive and it might take up to years to accomplish their total synthesis or derivatization.

Undirected photocrosslinking as an alternative immobilization method in AP approaches can circumvent this issue as it uses the unmodified compound on the one hand and on the other hand leads to the attachment via several different molecular sites of the compound. However, current methods that use photocrosslinking for target identification are limited in their throughput due to the chosen method to identify the enriched proteins: approaches that used fluorescently labelled target proteins allowed the screening of many compounds in parallel, but were limited in the amount of targets that could be screened in parallel (target-centric) and the bulky fluorescent labels might have introduced a further bias towards the binding of the target;[105] approaches that screened against a complex unlabelled target pool (i.e. protein cell extract) required tedious experimental procedures to differentiate promiscuous binders that bound the substance of interest non-specifically from specific binders, which in turn reduced the amount of compounds that could be screened in parallel to a few (drug-centric).[103,104]

Our aim was to develop a novel target identification method that uses the high-throughput advantage of conventional qAP-MS (screening many compounds against many proteins) whilst minimizing the bias caused by the immobilization of the compound using PAL as an alternative immobilization approach. We hypothesised that

using PAL (to construct an interaction array of drug-like substances on cellulose membranes) together with qAP-MS (to identify specifically interacting proteins) would allow the screening of multiple compounds in parallel against a complex target pool such as an entire proteome in a large-scale, a less biased and a highly specific manner (Figure 16).

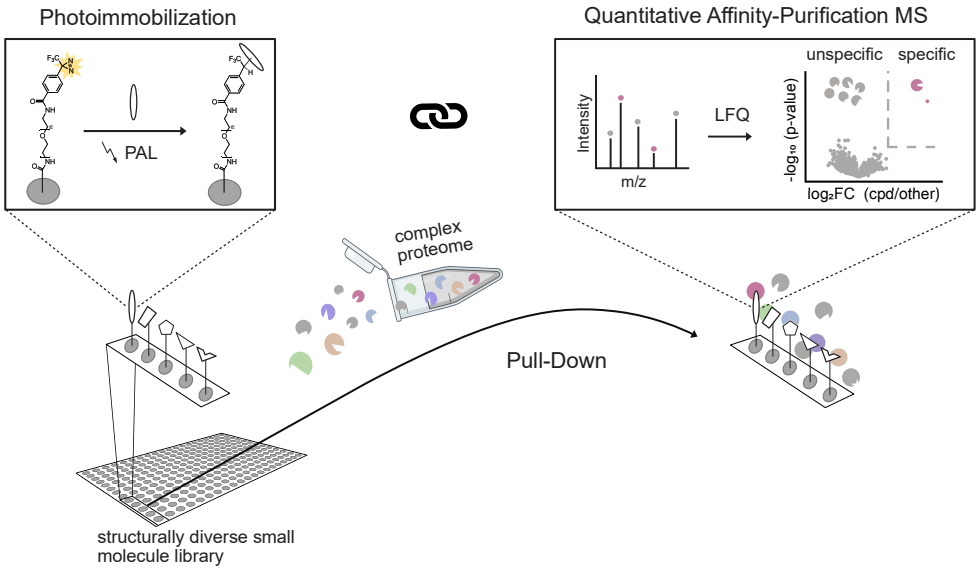


Figure 16. Combination of photoaffinity labeling (PAL) and quantitative affinity purification mass spectrometry (qAP-MS) to construct and screen small molecule cellulose arrays against complex target pools in large-scale.

Our main aims for developing such a method were as follows:

1 – A small molecule affinity matrix was constructed by production of a photoactivatable cellulose membrane and its undirected photoreaction with small molecules.

2 – The affinity matrix was used to identify specific protein interactions of small molecules by applying qAP-MS.

3 – The method was further developed, optimized and prepared for high-throughput target screening.

3. Results

3.1. Development of a compound interaction screen on a photoactivatable cellulose membrane (CISCM)

To combine the advantages of undirected photocrosslinking and qAP-MS, a compound interaction screen on a photoactivatable cellulose membrane (CISCM) was developed. CISCM consists of six steps. First, a cellulose membrane is functionalized with a photoreactive group, such as a trifluoromethylphenyl diazirine (TPD). In a second step this photoactivatable cellulose membrane is arrayed in triplicates with compounds of interest, followed by irradiation with UV light at 365 nm to initiate the undirected photoreaction between the compounds and the surface. The membranes are washed intensively in different organic solvents for hours to remove excess non-covalently attached compounds. In a fourth step, proteins interacting with the arrayed compounds are pulled-down from a complex target pool, such as a protein cell extract. After mild washing individual compound spots are excised and attached proteins are digested and analysed by high resolution LC-MS/MS.

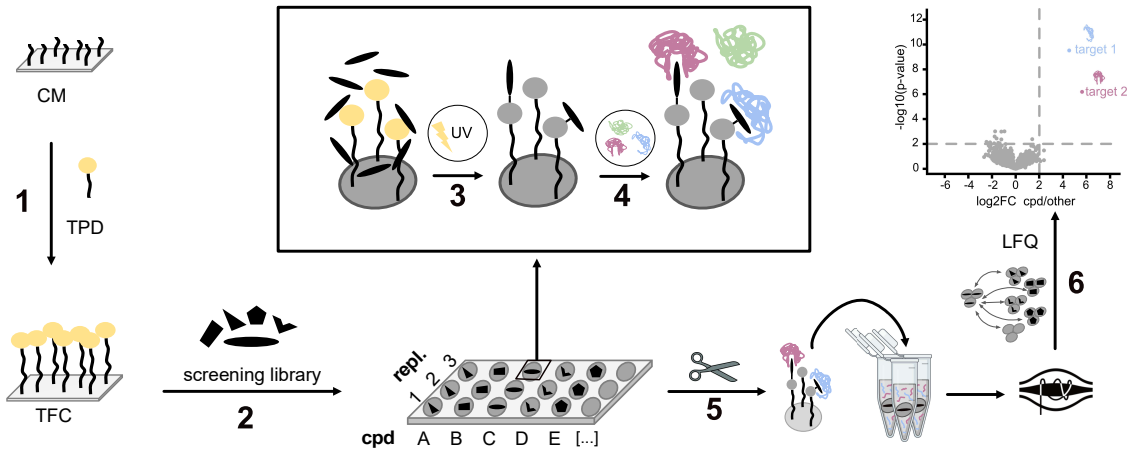


Figure 17 A compound interaction screen on a photoactivatable cellulose membrane (CISCM) performed in six steps: (1) Construction of a photoactivatable cellulose membrane (yellow: TPD warhead, black: Polyethylene glycol (PEG)-spacer), **(2)** spotting of a screening library onto this membrane, **(3)** covalent attachment of physisorbed compounds via undirected UV-crosslinking, **(4)** affinity enrichment of interacting proteins from a whole cell extract, **(5)** excision of each compound spot, protein digestion and LC-MS measurement, **(6)** target identification via quantitative analysis.

Some of the results presented in this thesis have been published in the Journal ChemMedChem in 2022:

Melder, F. T. I.; 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**, *17.* https://doi.org/10.1002/cmdc.202200346.

In the next sections, the preparation and analysis of a photoactivatable cellulose membrane will be shown. Next, the photoimmobilization of small molecules onto this membrane to construct a small molecule microarray (3.2) will be discussed and the application of the developed CISCM to a selection of well-studied natural compounds (3.3) will be presented. Last, further developments of the method (3.4) will be shown.

3.2. Generation of a photocrosslinked small molecule affinity matrix

3.2.1. Generation of a photoactivatable cellulose membrane using NHS-chemistry

For the undirected photoreaction we decided to use a trifluoromethylphenyl diazirine (TPD) derivate as PAL-probe of choice as it was previously used for AP-experiments.[104,105] This species has a better crosslinking efficiency compared to other photocrosslinkers[116,117] and a broader range of different functional groups that can be targeted on compounds of interest.[82] Its crosslinking product distribution also represents the chemical neighbourhood more accurately.[92] As a surface material we chose cellulose filter paper due to its beneficial properties for our application as well as its environmental and economic advantages. Cellulose is a highly abundant biopolymer, which is inexpensive, sustainable and environmentally friendly. Its properties and stability under various reaction conditions make it biocompatible and suitable for a broad range of applications.[118–120] In contrast to other surface supports like glass slides, cellulose consists of polysaccharide chains that are interconnected by hydrogen bonds and long ranging Van der Waals forces to form a porous network. This on the one hand allows a very dense functionalization of the surface. On the other hand, this makes functional groups on the surface more accessible than it would be the case on planar surfaces - a beneficial property for interaction screens with large biopolymers such as proteins. Functionalized cellulose has already successfully been used for qAP-MS applications. For example, in a protein interaction screen on peptide

matrix (PRISMA), synthetic peptides are immobilized onto cellulose membranes to screen for interacting proteins in a high-throughput manner. The high local bait density in this approach allows the capture of even low affinity interactions.[75,77,78]

Due to the strong network of non-covalent forces in the cellulose network, the synthetically accessible hydroxylic groups need to be activated before further functionalization is possible. To that end cellulose filter paper was incubated in a sodium hydroxide solution, to break hydrogen bonds, which leads to an increase of the reaction surface.[121]

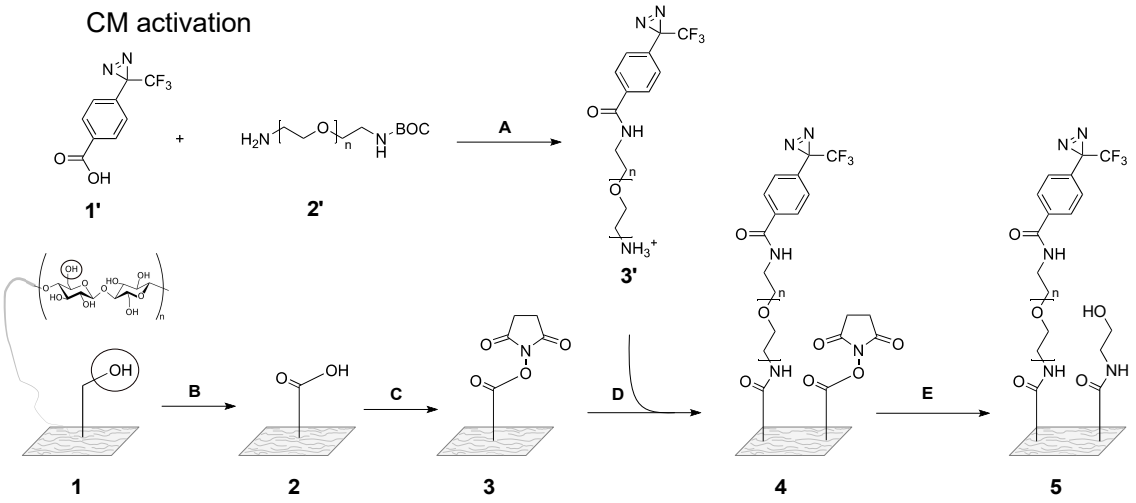


Figure 18. Construction of a photoactivatable cellulose membrane : (top) Functionalization of 4-[3-(trifluoromethyl)-3*H*-diazirin-3-yl]benzoic acid (TDBA) (1') with an amine-reactive PEG-spacer of variable length using *boc*-NH-PEG_n-$CH_2CH_2NH_2$ (2') to create a primary amine-containing trifluoromethylphenyl diazirine (TPD) (3'). (bottom) Stepwise activation of cellulose membranes (1) by 2,2,6,6-Tetramethylpiperidinyloxyl (TEMPO)- mediated oxidation to get oxidized cellulose (2) and activation with NHS and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) to obtain NHS-activated cellulose (NAC) (3). Immobilization of the photocrosslinker (3') on the activated cellulose (3) to form the TPD-functionalized cellulose (TFC) (4). Blocking of unreacted NHS-groups to obtain a blocked TFC membrane (5). Reaction conditions: (A) 1 eq. TDBA, 1.25 eq. *boc*-*N*-amido-PEG-amine, 0.35 eq. 4-Dimethylaminopyridine (DMAP), 30 min, THF; 1.75 eq. EDC, 18 h, RT, dark; (B) NaOH (w(NaOH) = 10%), 18 h, H_2O, RT; 0.39 mmol TEMPO, 14.1 mmol NaBr, 17.0 mmol NaOCl, 1 h, H_2O, pH 10; (C) 30 mmol NHS, 0.4 M EDC, sodium acetate, 1.5 h, H_2O, pH 5; (D) 10 mM amine-PEG TPD, 21 h, THF, RT; (E) 1-3 M ethanolamine (EA), 1-2 h, H_2O.

Membranes were then oxidized in a tetramethylpiperidinyloxyl (TEMPO)-mediated reaction (Figure 18, **B**) and activated with NHS and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (**C**) to obtain a NHS-activated cellulose (NAC) membrane (**3**).[122] To attach photoreactive species on these activated cellulose membranes a TPD warhead (**1'**) was equipped with a spacer of variable length and a terminal amine group (**A**) and immobilized via NHS-chemistry to obtain a TPD-

functionalized cellulose (TFC, **4**) membrane.[105] We chose a polyethylene glycol (PEG)-spacer due to its beneficial physical properties in aqueous screening environments, its flexibility and its low nonspecific protein binding.[63] Unreacted NHS was blocked with ethanolamine.

3.2.2. Confirmation of functionalization steps

To confirm the functionalization of the cellulose membrane (Figure 18), we teamed up with the Karlsruhe Institute of Technology (KIT, Germany) and monitored each of the functionalization steps with X-ray photoelectron spectroscopy (XPS), time-of-flight secondary ion mass spectrometry (TOF-SIMS) and attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR). These experiments were carried out and analysed by Vanessa Trouillet (XPS), Alexander Welle (TOF-SIMS) and Stefan Heißler (ATR-FTIR).

Using XPS, we were able to see significant changes in photoelectrons after oxidation, NHS-activation and UV-crosslinker immobilization of cellulose membranes (Figure 19). First a XPS survey scan across a wide energy range was performed to detect all photoelectrons emitted from a cellulose membrane at different functionalization stages (Figure 19a). Due to the high density of hydroxylic and acetal groups, all cellulose species showed the strongest signal for O 1s photoelectrons. Due to the high carbon content of cellulose, the signal for C 1s photoelectrons was the second intense. Only after functionalization of a cellulose membrane with NHS, a significant signal corresponding to N 1s photoelectrons appeared and after treatment with an amine-functionalized TPD-derivate (Figure 18, **3'**) a signal for F 1s photoelectrons appeared. Acquiring higher resolution spectra in selected energy ranges revealed more detail about the chemical context from which the photoelectrons originated and hence about the functional groups present in the corresponding cellulose species. Direct comparisons between subsequent functionalization steps are shown in Figure 19b for spectra in the binding energy regions, in which we expected the formation of a novel functional group. Upon oxidation a peak of weak intensity (2.2 atomic percent) at 289.2 eV binding energy appears in the high resolution C 1s spectra, when compared to unmodified cellulose, which corresponds to a newly formed carboxylic group. After further functionalization with EDC and NHS we detected N 1s photoelectrons at 402.0 eV binding energy (1.0 atomic percent) that originate from the succinimide

nitrogen of NHS bound to the oxygen as part of the ester group formed upon binding of the NHS hydroxyl group to the EDC-activated carboxylic acid of oxidized cellulose.[109] The other N 1s photoelectron signal, which is detected at 400.0 eV, can be explained by remaining EDC that contains N=C=N groups and amine (5.6 atomic percent). For cellulose treated with a PEG_4-TPD photocrosslinker (amine-PEG_4-linker) after NHS-activation we observed signal at 688.2 eV binding energy in the high resolution F 1s photoelectron spectrum that derives from covalently bound fluorine atoms of a trifluoromethyl group, such as present in the amine-PEG_4-linker.

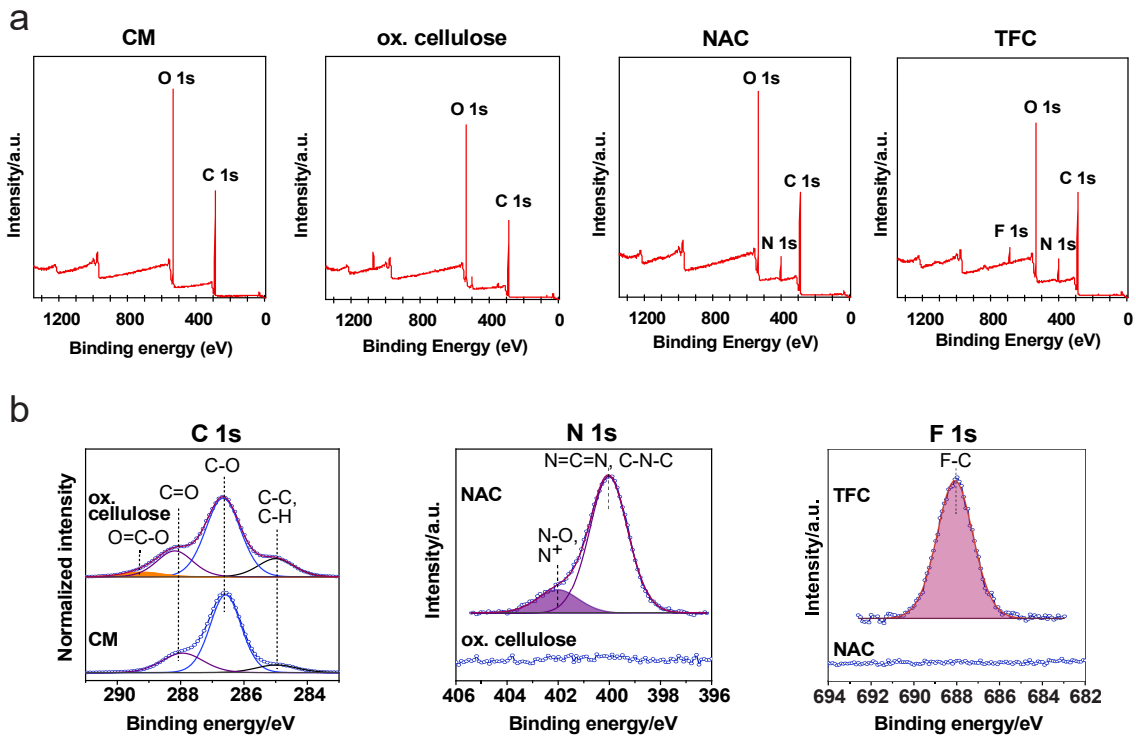


Figure 19. XPS spectra of stepwise cellulose membrane functionalization. (a) XPS survey scans of an unmodified cellulose membrane (CM), oxidized cellulose (ox. Cellulose), NHS-activated cellulose (NAC) and TPD-functionalized cellulose (TFC) across a broad range of binding energies. **(b)** Comparison of the C 1s, N 1s and F 1s photoelectron spectra in selected binding energy ranges of cellulose at different functionalization stages.

As a next step, our aim was to further test the functionalization with an orthogonal method such as ATR-FTIR (Figure 20). To that end, we first probed oxidized cellulose membranes and observed – in contrast to unmodified cellulose membranes as a reference - carboxyl vibrations at 1599 cm^{-1} (v C=O) in the fingerprint area, indicating the presence of a carboxylic group. For cellulose membranes treated with EDC and NHS we detected ATR-FTIR vibrations at 1705 cm^{-1} (v OCN), which corresponds to

an amide group, such as present in NHS. Further functionalization with an amine-PEG_4-linker led to a significant decrease of this signal. The decreased signal of amide vibrations at 1705 cm^{-1} completely vanished after blocking the membrane with ethanolamine. A second amide vibration at 1540 cm^{-1} was detected for both, NAC and TFC membranes. As expected, the additional functional group introduced by ethanolamine led to a broad signal from 1670-1540 cm^{-1}, instead of two separated peaks at 1650 cm^{-1} and 1539 cm^{-1}.

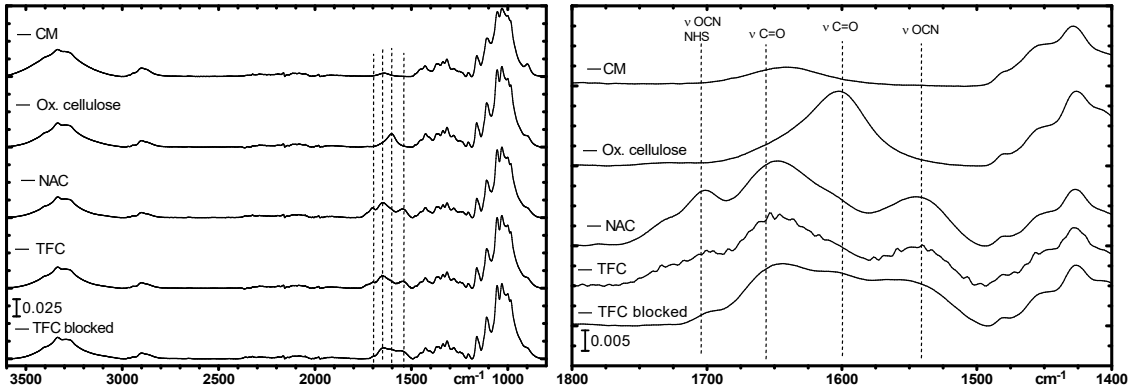


Figure 20. ATR-FTIR spectra of stepwise cellulose membrane functionalization. Spectra of different cellulose species in a broad wave number range (left). Interesting vibrations in the fingerprint area indicated with dashed lines. Magnified fingerprint area with assigned vibrations (right).

ATR-FTIR and XPS are direct and precise methods to study the surface chemistry of modified cellulose membranes. However, both methods require expert knowledge and specific equipment. Therefore, we aimed to develop a cheaper, faster and in-housed applicable strategy to routinely test the success of NHS-activation. As NHS-activated esters react with primary amines, we reasoned that NHS-activation could also be indirectly tested by the reaction of the NAC membranes with a primary amine-containing dye such as amine-flourescein, or an antibody coupled to the enzyme horseradish peroxidase (HRP). The binding can then be qualitatively detected in a western blot reader either by detecting the fluorescence of the dye or, for the latter, the chemiluminescence after reaction of HRP with its luminol substrate in the presence of an enhancer to obtain enhanced chemiluminescence (ECL).

Both approaches showed a strong signal for NAC when compared to previous functionalization steps or an unmodified cellulose membrane (Figure 21), even after short reaction times of only 30 minutes in the case of using 6-aminoflourescein. This signal could still be obtained after storing the NAC membranes for three months at low

vacuum and in the presence of a desiccant (Supporting Figure 1). NAC membranes that were blocked with ethanolamine before treatment with 6-aminoflourescein did not show fluorescence, nor did NAC membranes treated with flourescein (not containing a primary amine). The approach using a HRP-conjugated antibody required longer reaction times and washing steps and showed significant background binding of the antibody on unmodified cellulose. Therefore, we decided to use the covalent immobilization of 6-aminoflourescein on NAC membranes as a method of choice to test the success of NHS-activation.

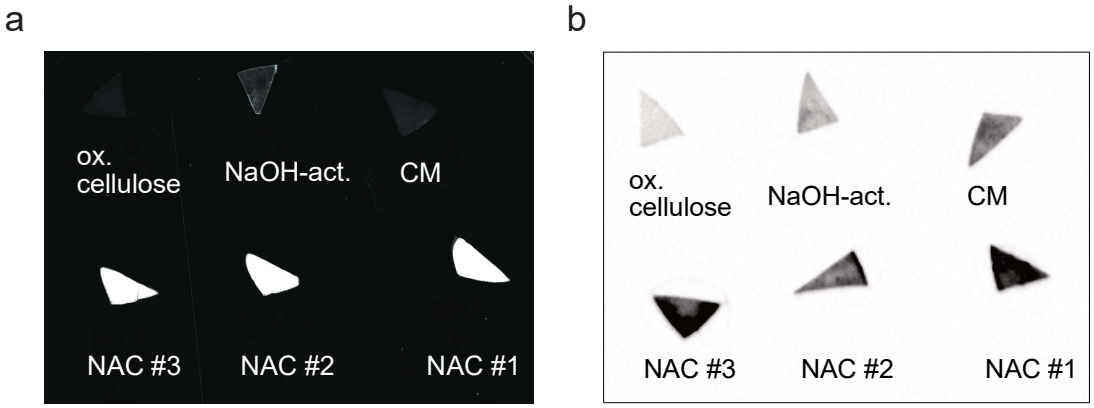


Figure 21. Control of NHS-activation on cellulose membranes using primary amine containing substances. Cellulose membranes at different functionalization stages of cellulose membranes after reacting with 6-aminoflourescein and washing with ethanol. Fluorescence appears as a white signal. **(b)** Cellulose membranes at different functionalization stages and after reaction with a HRP-conjugated antibody and intense washing in radioimmunoprecipitation assay (RIPA) buffer. Chemiluminescence appears as a black signal. Labels: unmodified cellulose membrane (CM), sodium hydroxide-activated cellulose (NaOH-act.), oxidized cellulose (ox. cellulose), and NHS-activated cellulose in triplicates (NAC#1-3).

3.2.3. Photoimmobilization of small molecules

Next, we tested whether we can covalently attach compounds to the constructed photoactivatable cellulose membrane using undirected photocrosslinking. To evaluate this, we used time-of-flight secondary ion mass spectrometry (ToF-SIMS), since it allows the detection of both intact molecules and corresponding fragments. As a well-studied test compound, we chose the large and polyfunctional natural structure analogue cyclosporine A (CsA).

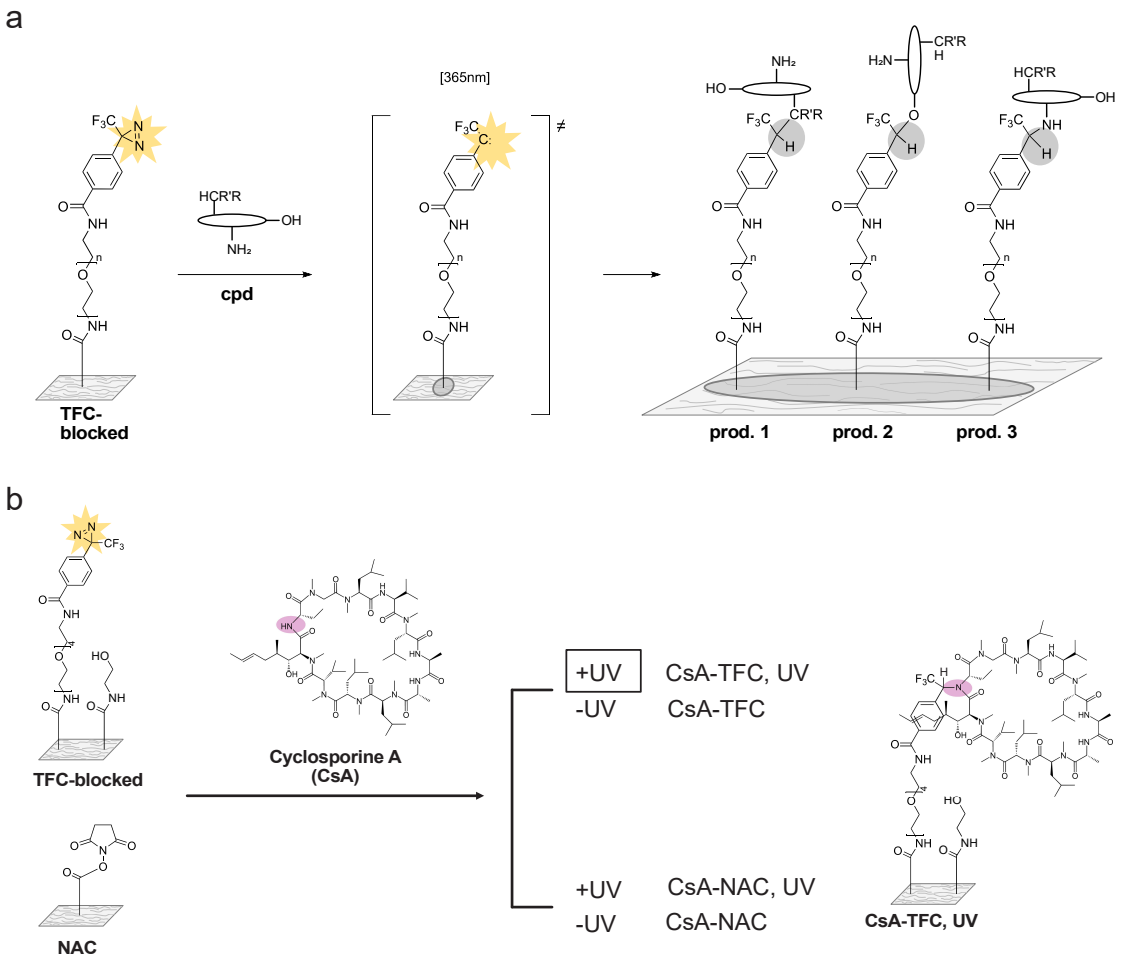


Figure 22. Photoimmobilization strategy and overview of ToF-SIMS samples and controls. (a) Spotting of generated TPD-functionalized cellulose (TFC) membrane with a compound (cpd) solution of interest (10 mM, DMSO), here represented by a fictional compound; evaporation of the solvent; formation of reactive carbene induced by irradiation with UV light at 365 nm; formation of different reaction products (prod. 1, prod. 2, prod. 3) of undirected photocrosslinking reaction between the photoactivated TFC and the dried compound of interest. **(b)** Overview of samples analysed with ToF-SIMS with 'CsA-TFC, UV' being the sample. One example of a possible crosslinking product of the sample is shown (right).

After spotting a solution of CsA (50 nmol) in quadruplicates on a photoactivatable cellulose (TFC) membrane and on a NHS-activated cellulose (NAC) membrane, the solvent was evaporated (Figure 22b). Two of each replicate membranes were irradiated with UV light at 365 nm (+UV) on both sides (spotted and unspotted), whereas the other two replicates were kept in the dark (-UV). For the spotted and irradiated TFC replicate samples (CsA-TFC, UV) the initiated photocrosslinking reaction with the different reactive sidechains of CsA was expected to yield a distribution of photoreaction products as schematically shown in Figure 22a on the example of a fictional compound and in Figure 22b for one possible photocrosslinking

product of CsA. To remove of non-covalently attached excess compound molecules, we washed the resulting four duplicate samples intensively in different organic solvents and analysed them in ToF-SIMS. As further controls, we prepared duplicates of CsA-spotted TFC membranes that were UV-irradiated without prior evaporation of the solvent and another two replicates that were UV-irradiated only on the unspotted side of the membrane.

First a drop cast sample of CsA was measured as a reference. This identified the molecular ion of CsA ($C_{62}H_{112}N_{11}O_{12}^+$), which correlates to the intact structure of the drug and hence indicates non-covalent attachment (Figure 23a). Interestingly this signal could only be detected in the reference, but neither in any of the four samples nor in the positive control (CsA-TFC, UV). Instead, we observed a corresponding fragment ion of CsA ($C_{60}H_{106}N_{11}O_{12}^+$) in the reference sample that could only be identified in the CsA-spotted and UV-treated TFC sample, but not in any of the three negative control samples (Figure 23b) or in CsA untreated samples. The presence of a CsA fragment ion only on a diazirine functionalized and UV-irradiated cellulose membrane (CsA-TFC, UV), but absence of the intact CsA molecular ion (M+) as observed in the reference, indicates that CsA was bound covalently to the TFC membranes. Therefore, CsA needed to be fragmented in those samples to be detected.

Next, we looked for fragments that are characteristic for cyclosporine A: amino acid fingerprints. We revealed an N-methylated leucine ion ($C_6H_{14}N^+$), which could only be identified in CsA-spotted and UV-treated TFC (CsA-TFC, UV), but not in any of the 3 negative control samples (CsA-TFC; CsA-NAC, UV; CsA-NAC, Figure 24a).

This fingerprint signal could also be identified when measuring the unspotted but also irradiated backside of this TFC membrane (Figure 24b). This was not surprising, as we used high molar amounts of compound and when spotting the solution of a dye, such as ponceau S, onto the front side of the cellulose membrane we could clearly see the compound spot also on the backside of the corresponding membrane. Interestingly, this signal could also be observed for both, front and backside of the membrane, when spotting the front side of TFC membrane and UV-irradiating the backside only. These results led us to the conclusion that UV-irradiation from one side could be sufficient to photoimmobilize CsA homogenously throughout the entire depth of the membrane. As previously suspected, if the compound spot was not dried prior to UV-irradiation, no

CsA fingerprint signal was observed. This indicates, that the photocrosslinker reacts predominantly with excess solvent molecules rather than with the spotted compound if spots are not dried prior to UV-irradiation. Using chemical mapping we could also show that the mentioned amino acid fingerprint signal was laterally homogenous across the cellulose membrane, indicating that CsA was photocrosslinked evenly throughout the sample (Figure 24c).

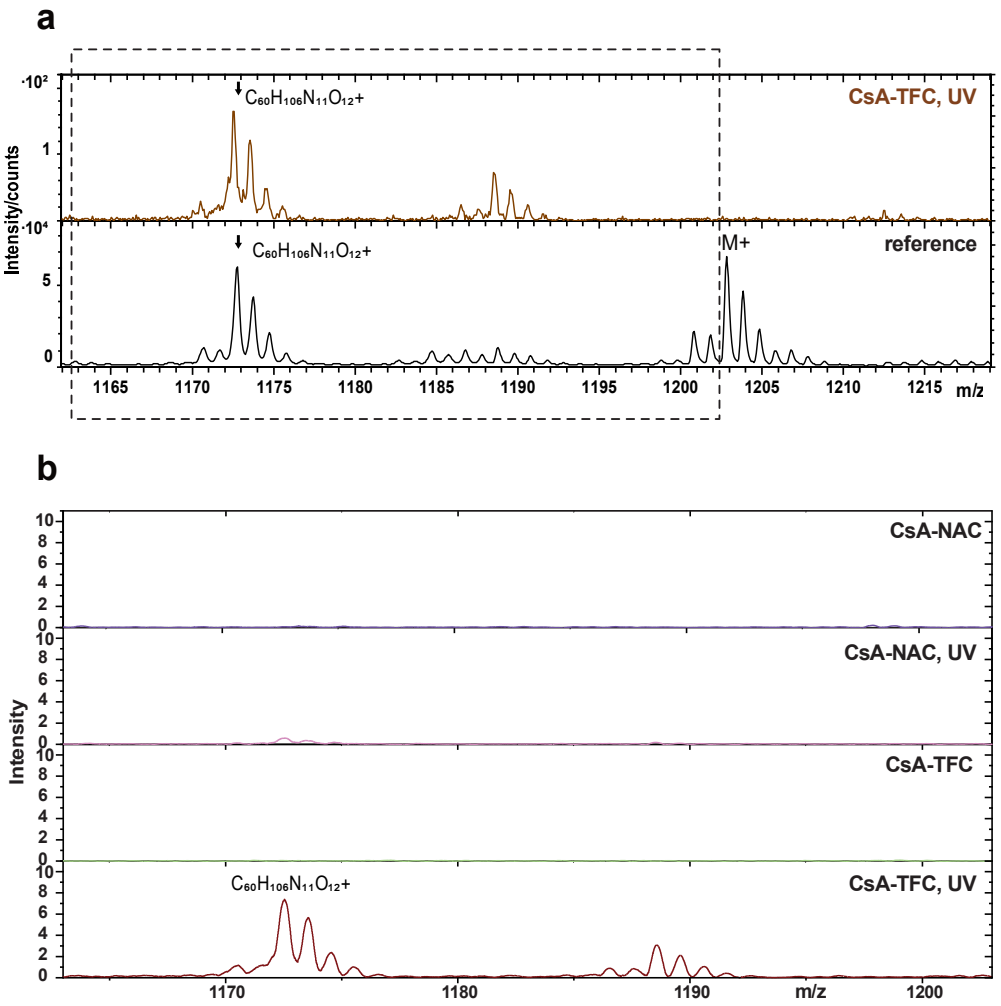


Figure 23. ToF-SIMS analysis of NAC and TFC membranes spotted with cyclosporine A (CsA), irradiated and not irradiated with UV light. (a) ToF-SIMS spectrum of a CsA drop cast (reference) compared to the spectrum of a CsA spotted and UV-irradiated TFC membrane (CsA-TFC, UV). Molecular ion in the reference indicated as M+ and a fragment occurring in both samples indicated as $C_{60}H_{106}N_{11}O_{12}^{+}$. **(b)** Smoothed (Savitzky-Golay) ToF-SIMS mass spectra of the four different CsA spotted cellulose membranes in mass range m/z 1163-1203 as highlighted in (a), where the fragment ion was detected for the positive control. CsA-TFC: CsA-spotted TFC membrane, CsA-NAC: CsA-spotted NAC, UV: UV-irradiated (365 nm).

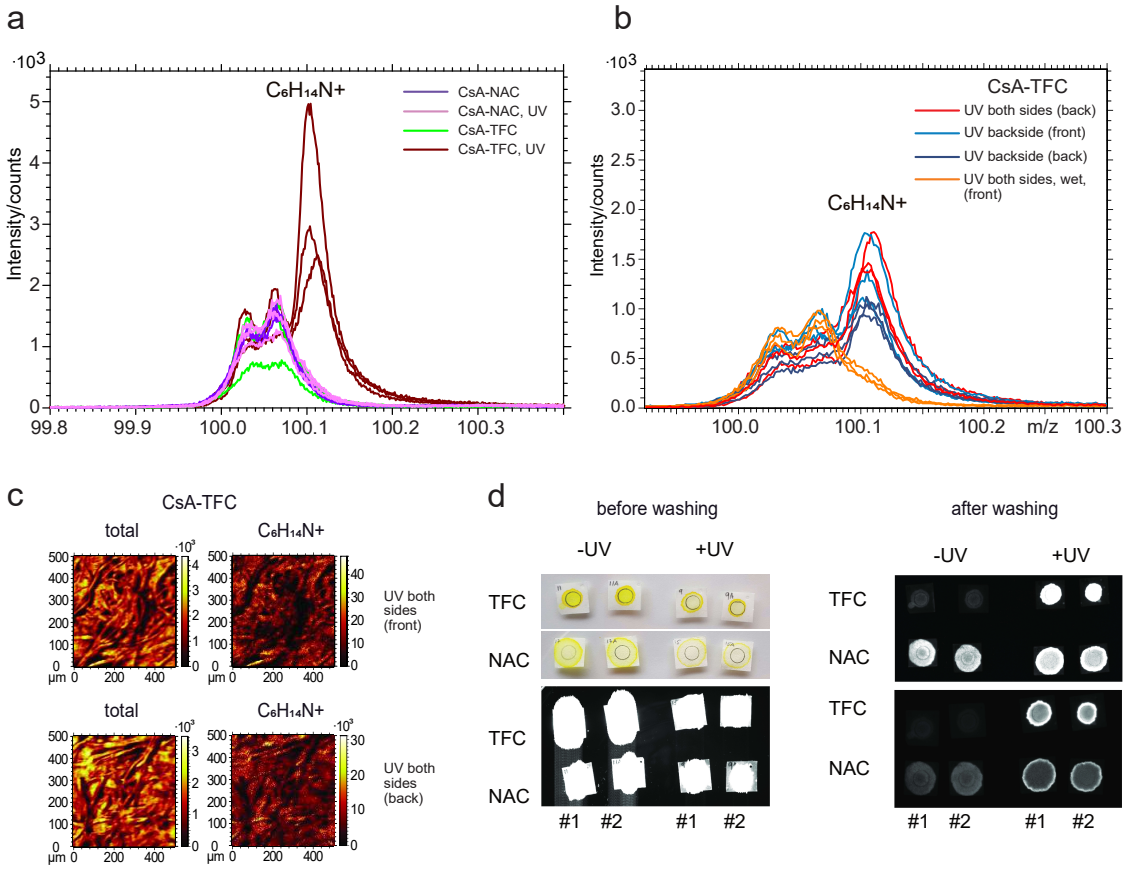


Figure 24. ToF-SIMS spectra of amino acid fingerprinting, chemical mapping and fluorescence imaging of a flourescein control. (a) ToF-SIMS spectrum in the mass range of the amino acid fingerprinting signal of N-methylated Leu ($C_6H_{14}N^+$, m/z 100) across different samples: CsA spotted NHS-activated cellulose (purple), CsA-spotted UV-treated NAC (pink), CsA spotted TPD-functionalized cellulose (green) and CsA spotted TFC treated with UV (brown). Spectra were acquired at three different lateral positions across the corresponding sample. **(b)** Amino acid fingerprint ToF-SIMS spectra of CsA spotted TFC differentially UV-treated and measured on front or backside respectively. **(c)** ToF-SIMS chemical mapping showing the lateral distribution of different CsA-fragment signals across CsA spotted TFC irradiated with UV. **(d)** Images of duplicates (#1, #2) of TFC and NAC membrane pieces spotted with a solution of flourescein with (+UV) or without (-UV) UV-irradiation and before (left) and after (right) washing the membranes in ethanol, DMF (right, top), THF, ethanol and water incubating in methanol overnight (right, bottom).

Similarly to the procedure mentioned in chapter 3.2.2, where we aimed to track the functionalization of cellulose in-house, we also aimed to control the UV-crosslinking in a cost- and effort effective and in-house applicable manner. To that end, we again used a fluorescent dye as an optical control compound – this time flourescein (without an amine-moiety). We reasoned that flourescein would enable us to detect even trace amounts of the dye on the membranes. As for cyclosporine A, we spotted blocked NAC and TFC membranes and irradiated one replicate with UV light (+UV), whereas the second one was not irradiated with UV light (-UV). Next, the membranes were washed

thoroughly. We compared the fluorescence of cellulose membranes before and after each washing step as shown for three selected time points in Figure 24d. As expected, before washing all membranes showed strong fluorescence. However, even though the membranes were spotted with a solution of flourescein in the same manner, the spots on the blocked NAC membranes had a greater diameter. This might be correlated with the different hydrophilicity of TFC and NAC membranes. The fluorescence signal decreased significantly for the spotted NAC membranes after 2 hours of washing and almost disappeared for the not irradiated TFC membrane. After further washing and incubation in methanol overnight, the positive fluorescein control (TFC, +UV) still showed strong fluorescence in contrast to the non-irradiated TFC membrane, for which the signal disappeared. For both irradiated and non-irradiated NAC membranes, we could only observe a low background fluorescence signal. This encouraged us to implement flourescein as an easy-to-perform internal UV-crosslinking control for each of the following experiments.

3.3. Application of CISCM to a selection of well-studied natural compound analogues

3.3.1. Target identification of cyclosporine A (CsA)

In the last section, we showed with the natural compound CsA, that we were able to photoimmobilize small molecules on a functionalized cellulose membrane. Next, we aimed to identify the known protein targets of photoimmobilized compounds using quantitative affinity purification mass spectrometry (qAP-MS). In addition to CsA, we included thalidomide as an example of a small molecular weight and compact structure with a - for its MoA - relevant stereocenter. Thalidomide was originally used as a sedative in the sixties and withdrawn as it caused severe skeletal birth defects in children and it was rediscovered in 1999 as a treatment in multiple myeloma.[123] We used the (S)-enantiomer of thalidomide due to its 10-fold higher binding affinity to the known target cereblon (CRBN).[124] As background for the quantitative analysis we included bromphenol blue, coomassie brilliant blue, ponceau S and 6-aminoflourescein.

Triplicates of blocked TFC membranes, equipped with a PEG_2-linker, were spotted with solutions of the mentioned compounds and the photoreaction was initiated by UV-irradiation. As controls, triplicates of unspotted non-functionalized cellulose

membranes and unspotted blocked TFC membranes were included. All membranes were washed thoroughly, dried and interacting proteins were enriched from a Jurkat cell extract. Cellulose spots corresponding to individual photocrosslinked compounds or controls were excised and enriched proteins digested. After standard shotgun proteomics sample preparation,[125] all 21 samples (2 compounds, 5 background samples, triplicates) were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer.

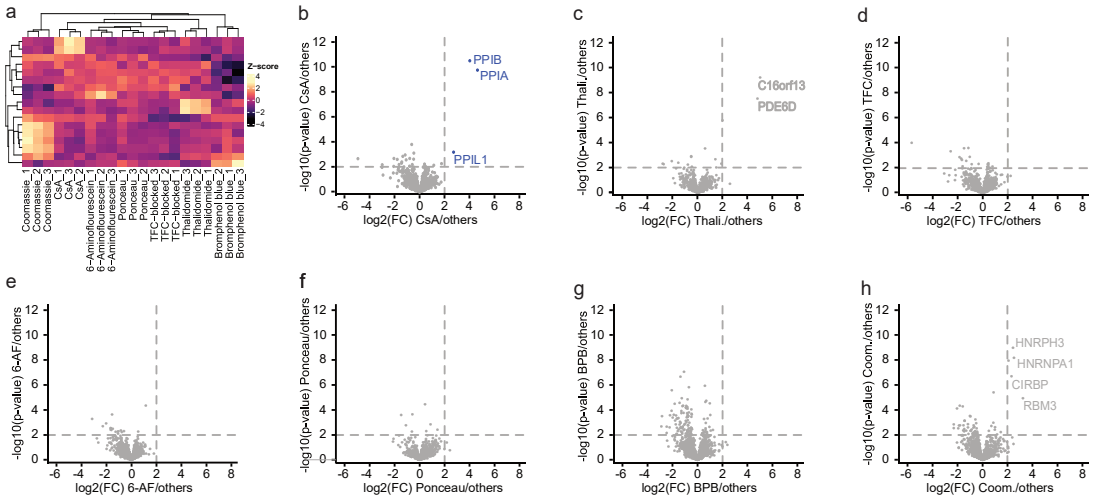


Figure 25. Differential protein identifications derived from AP-MS data. (a) Heatmap of Z-scores computed for ANOVA significant (FDR 5%, 250 randomizations) protein identifications of two immobilized compounds and five controls. Each column is an individual replicate of the total 21 samples and each row is the Z-scored. **(b-h)** Volcano plots displaying the $\log_2$ fold change (x-axis) against the Student's t-test derived $-\log_{10}$ p-value (y-axis) for pairwise comparison of grouped triplicates of (b) cyclosporine (CsA), (c) thalidomide (Thali.), (d) TPD-functionalized cellulose (TFC), (e) 6-aminoflourescein (6-AF), (f) ponceau S (Ponceau), (g) bromphenol blue (BPB) and (h) coomassie brilliant blue (Coom.), respectively, against all other samples (others). Proteins with t-test p-values < 0.01 and fold changes of at least four are labeled and known protein targets marked in blue.

MaxQuant[45] data analysis of raw spectra identified 2,509 protein groups (at 1% protein and peptide FDR) in all samples combined. To identify the proteins interacting specifically with a given immobilized compound, we used label-free quantification (LFQ) here and in all of the following chapters.[50] We first performed an analysis of variance (ANOVA, FDR 5%) to obtain differentially abundant proteins. Hierarchical clustering of ANOVA significant proteins revealed clustering of replicate samples (Figure 25a). To identify specific binders we compared protein abundances in the three replicates for a given compound or the empty TFC control to all other samples using the Student's t-test and presented the data as volcano plots (Figure 25b-h). As expected, most of the identified proteins did not show preferential binding and can thus

be considered nonspecific background proteins. We selected specific binders requiring t-test p-values < 0.01 and fold changes of at least 4. For cyclosporine A, this identified the cyclophilins PPIB, PPIA, and PPIL1 as specific interactors, corroborating previous mass spectrometry-driven chemoproteomic results (Figure 25b).[56,80,126–128] In contrast, for most of the dyes as well as for the blocked TFC membrane, no significant interactor was identified (Figure 25d-h). For the coomassie brilliant blue dye, four proteins, HNRH3, HNRNPA1, CIRBP and RBM3 were identified as significant. The isoelectric point prediction tool Expasy[129] revealed that HNRNPA1 (pI 9.17), CIRBP (pI 9.51), and RBM3 (pI 8.86) are all posit,ively charged at the given affinity purification conditions. This lead to the assumptions that those proteins show enhanced binding toward the negatively charged sulfonate groups of the dye due to electrostatic attraction. Therefore, these interactions were considered to be biochemically true but not biologically relevant. Also, the binding of HNRH3, a RNA-binding protein, could be explained by binding the negatively charged sulfonate groups of coomassie instead of the negatively charged phosphate groups of the RNA backbone. As expected our screen did not enrich any specific binders for the negative control (TFC membrane), which indicates that the background binding of the photocrosslinker with proteins of the target pool seems to be consistent. For the small drug thalidomide we failed to identify the known protein targets cereblon and DDB1 (Figure 25c), even though immobilization of immunomodulatory imide drugs (IMiDs) can enrich their direct target cereblon (CRBN) - a substrate receptor of the CRL4-CRBN-E3 ligase complex – and their indirect target DNA damage-binding protein 1 (DDB1).[130] Both targets, CRBN and DDB1 are expressed in Jurkat lysates that were used in this AP-MS experiment (Figure 26a).[131] Instead, we identified two other proteins as specific interactors: The hypothetical protein, which is predicted to be expressed from the open reading frame 13 of chromosome 16 (C16orf13), and the phosphodiesterase 6D (PDE6D). PDE6D plays a critical role in the RAS trafficking by controlling its localization upon binding. Degradation of PDE6D has emerged as a novel potential strategy to indirectly target RAS[132–134] and has been successful for other molecular glues (MG), also from the IMiDs family.[135]

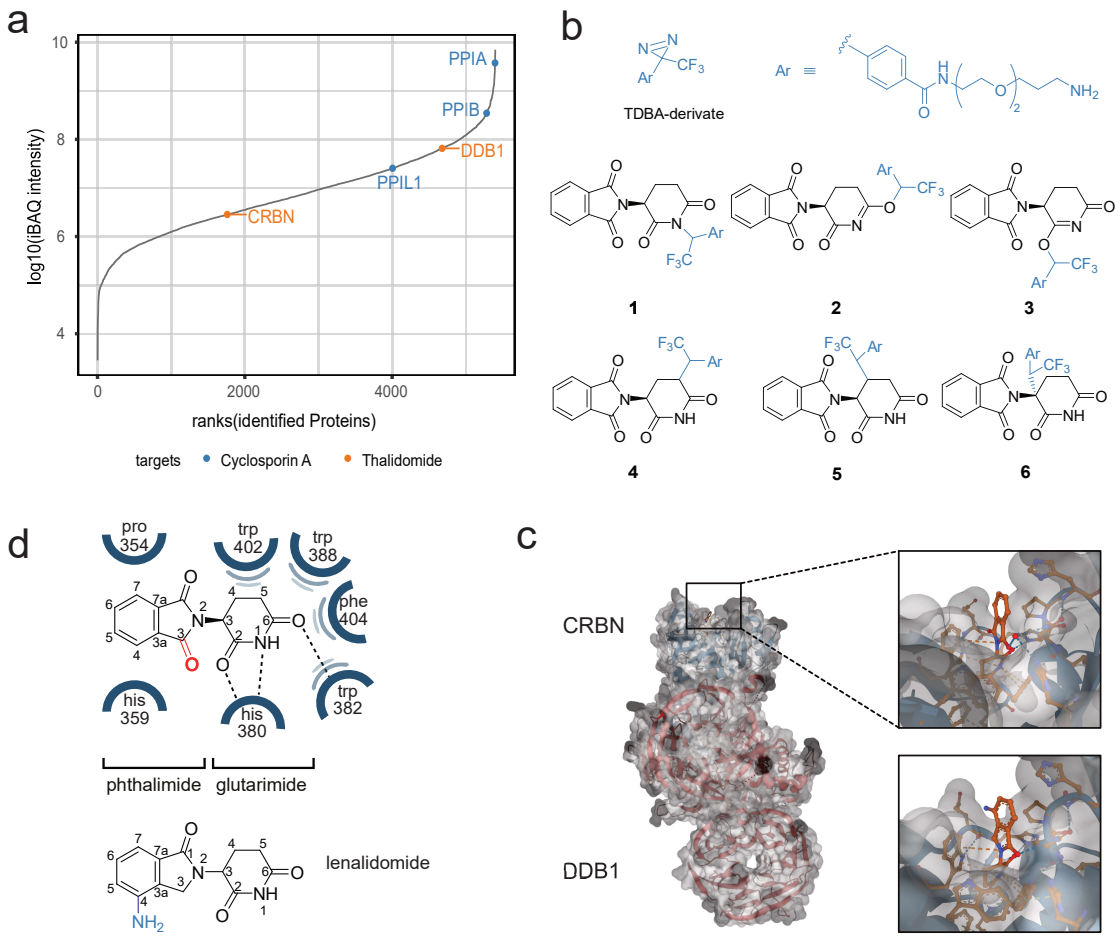


Figure 26. Thalidomide and its targets CRBN and DDB1. (a) Protein identifications of the used Jurkat proteome ranked by their $\log_{10}$ iBAQ intensities. Identified targets of CsA are highlighted in orange and expected targets of (S)-thalidomide are highlighted in blue. **(b)** Some possible photocrosslinking products of (S)-thalidomide by insertion (1-6). **(c)** Crystal structure of the G. gallus CRBN-DDB1 complex binding to (S)-thalidomide and magnified regions of the binding pockets (black boxes) with (S)-thalidomide (top, PDB code: 4CI1) and lenalidomide (bottom, PDB code: 4CI2). Solvent accessible surface area of CRBN is shown in grey. **(d)** Schematic representation of non-covalent interactions of the glutarimide ring with amino acid residues of the G. gallus CRBN binding pocket. Hydrogen bonds are indicated with dashed lines and hydrophobic interactions are indicated by blue semicircles.

We expect the photoreaction of (S)-thalidomide to occur preferably at the glutarimide ring - especially at its electron-rich amide group (**1**-**3**) – in the form of an N-H or O-H insertion (Figure 26b). When bound to CRBN, the glutarimide ring of (S)-thalidomide is buried in a shallow tritryptophan binding pocket (Figure 26c) and the binding is mediated by several non-covalent interactions of the glutarimide ring towards surrounding amino acids (Figure 26d).[124,136] The phthalimide substituent on the other

hand is designed to recruit substrates and hence protrudes out of the CRBN binding pocket: it remains accessible (Figure 26c) and its modification is not expected to influence the bindings toward CRBN. In contrast, any photoreaction at the glutarimide ring could perturb or even prevent the binding of CRBN. Lenalidomide (Revlimid®), another IMiD drug and a blockbuster for the treatment of multiple myeloma,[14,137] lacks one imide-oxygen and contains an additional aniline amine-group at position four of the solvent-exposed side (phthalimide) when compared to the structure of thalidomide (Figure 26d). Upon photocrosslinking we expect this additional amine-group to be more reactive towards the carbene than any other available functional group within the lenalidomide molecule. This in turn might shift the distribution of photocrosslinking products towards lenalidomide derivatives with an unmodified glutarimide moiety. We reasoned that by including lenalidomide in the next target screening we would be able to test whether the size of a drug – and hence its limited availability of functional groups – appears to be a general limitation of our method or if the case of thalidomide represents a special case of a tight structure affinity relationship.

3.3.2. Target identification of natural compound analogues

For one out of two test compounds and several controls, we have shown that we can photoimmobilize small molecules on a cellulose membrane and that we can use this compound (CsA) to pull down specific targets. However, as we were unable to identify the known targets of the second - rather small - test compound (S)-thalidomide, we aimed to further test the application range of our approach.

To that end we screened a library of seven well-studied model compounds (Figure 27a), because of their well characterized protein binding partners and their high structural complexity.[52,80,124,138–140] This library includes the previously discussed compounds (R,S)-lenalidomide and cyclosporine A (CsA) as well as the compounds sirolimus (rapamycin), methotrexate (MTX), vinblastine, paclitaxel and tacrolimus (FK506). The chosen compounds cover a broad molecular weight range when compared to all bioactive substances listed in ChEMBL[141] (Figure 27b), are structurally diverse with different levels of hydrophobicity (Figure 27c), and possess at least one stereocenter. Corresponding expected targets are sufficiently abundant in the used Jurkat cell extract (Figure 27d).

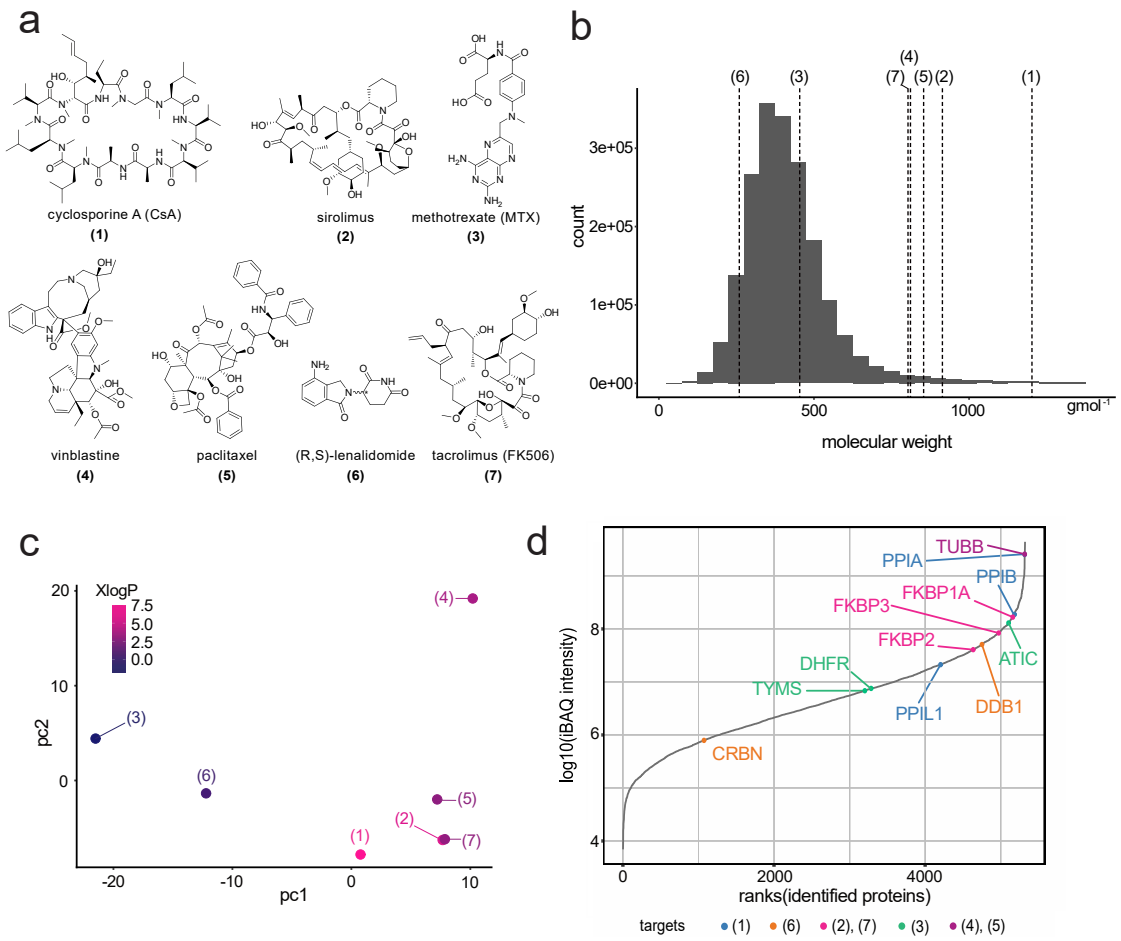


Figure 27. Selection of structurally diverse natural compound analogues and lenalidomide. (a) Structural formulas of the selected compounds. **(b)** Molecular weights of selected compounds (indicated by dashed lines) compared to the molecular weight distribution of all compounds listed in ChEMBL (excluding antibodies, enzymes, cells and structures with no listed molecular weight annotation): 1202.20 gmol^{-1} (1), 914.20 gmol^{-1} (2), 454.40 gmol^{-1} (3), 811.00 gmol^{-1} (4), 853.90 gmol^{-1} (5), 259.26 gmol^{-1} (6), 804.00 gmol^{-1} (7). **(c)** Principal component analysis on the chemical structure of the selected compounds. Spots are coloured according to the compounds' calculated XlogP value. **(d)** Protein identifications of the used Jurkat proteome ranked by their $\log_{10}$ iBAQ intensities. Identified unique and shared targets are highlighted in blue (1), orange (6), pink (2, 7), green (3) and violet (4, 5).

All compounds were spotted onto triplicates of TFC membranes and immobilized via photocrosslinking. Unmodified, oxidized and blocked TPD-functionalized cellulose without spotted compounds were used as controls. As an internal photocrosslinking control the compound-spotted TFC membrane triplicates were additionally spotted with fluorescein prior to UV-treatment (Figure 28a). After UV-treatment an empty area of these then photoinactive TFC membranes was spotted again with a solution of fluorescein, dried intensively and the membranes were washed thoroughly. We reasoned that the fluorescence of both, before and after UV-irradiation spotted fluorescein, correspond to photodependent and photoindependent interactions of the

compound with the TFC membrane and that a comparison of both can give evidence to the success of the photoreaction. Initially both spots showed a strong fluorescence signal, which - after intense washing - decreased for the photodependent interaction and vanished for the photoindependent interaction (Figure 28b). This led us to the conclusion that fluorescein, as a representative for all other studied compounds, was only detected in presence of an active TPD-linker in combination with UV light and we concluded that it was photoimmobilized.

For the interaction screen, cellulose membranes of samples and controls were incubated with Jurkat cell lysate and cellulose spots corresponding to individual photocrosslinked compounds and controls (excluding the fluorescein controls) were excised and processed for shotgun proteomic analyses using standard methods. All 30 samples (seven compounds and three controls, in triplicates) were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer.

Data analysis with MaxQuant[45] identified 3,275 protein groups (protein and peptide FDR of 1%) in all samples combined. Due to bad quality one replicate of the MTX affinity purification samples had to be excluded from the analysis. We first performed an analysis of variance (ANOVA, FDR 5%) with the LFQ data to obtain differentially abundant proteins and performed hierarchical clustering on ANOVA-significant proteins. This revealed clear differences between the control samples (unmodified, oxidized and TPD-functionalized cellulose) and the compound-spotted TFC samples (Figure 28c). To identify specific targets we compared protein abundances in the three replicates for a given compound to all other samples using the Student's t-test and presented the data as volcano plots (Figure 28d-j). We selected specific binders requiring t-test p-values < 0.01 and fold changes of at least 4. As expected, most of identified proteins did not show preferential binding. In accordance with the data presented in chapter 3.3.1 PPIB, PPIA and PPIL1 were identified as specific binders of CsA and an additional protein PPIF, another member of the PPI family, consonant with the results of previous proteomic efforts.[56,80,126–128]

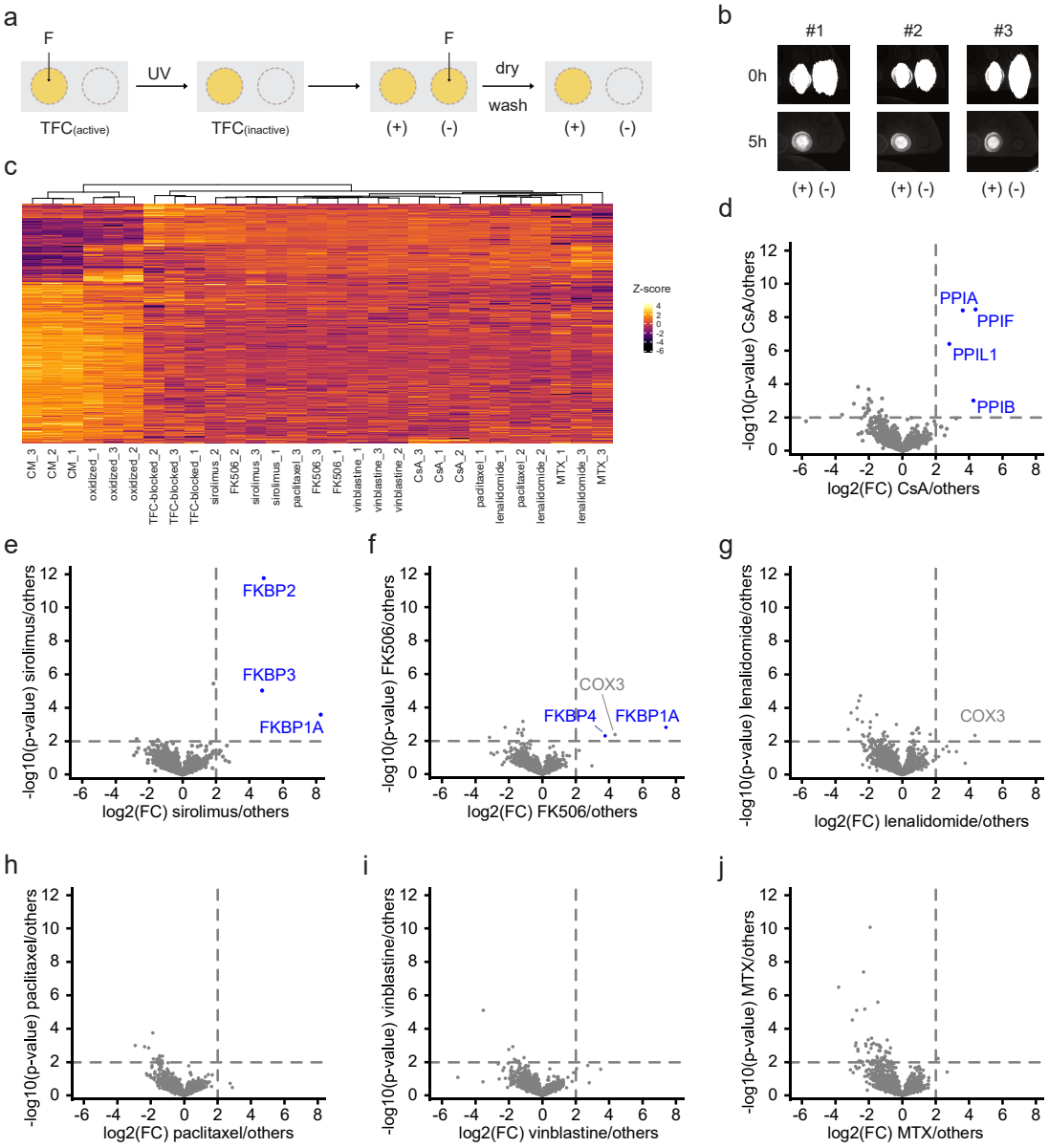


Figure 28. Target identification of a structurally diverse library. (a) Implemented photocrosslinking control with fluorescein (F) spotted onto triplicates of active TFC membranes (+) and by UV light (UV) inactivated TFC membranes (-).**(b)** Fluorescence signals of the fluorescein controls spots on the triplicate membranes (#1-3) before (0 h) and after (5 h) washing the membranes in the solvents ethanol, dimethylformamide, tetrahydrofurane, ethanol and water for one hour each. **(c)** Heatmap of Z-scores computed for ANOVA significant (FDR 5%, 250 randomizations) protein identifications of seven immobilized compounds and three controls. Each column is an individual replicate of the total 29 samples and each row is the Z-score. **(d-j)** Volcano plots displaying the $\log_2$ fold change against the Student's t-test derived $-\log_{10}$ p-value for pairwise comparison of grouped triplicates of (d) cyclosporine (CsA), (e) sirolimus, (f) tacrolimus (FK506), (g) lenalidomide, (h) paclitaxel, (i) vinblastine and (j)

methotrexate (MTX) against all other samples (others). Proteins with t-test p-values < 0.01 and fold changes of at least four are labeled and known protein targets marked in blue.

Similarly, for the other immunosuppressant drugs sirolimus (Figure 28e) and tacrolimus (Figure 28f) the main target FKBP1A as well as other members of the FK506-binding protein (FKBP) family could be identified as specific binders, which are known to bind the corresponding drugs and have previously been identified in chemoproteomic target identification experiments.[52,59,126,142] In contrast to our published data we additionally identified COX3 as specific binder for the drugs sirolimus and lenalidomide. This is due to the fact that in our publication we presented results based on the affinity enrichments of CsA, tacrolimus, sirolimus and lenalidomide, whereas for the data analysis presented in this dissertation the affinity enrichments of the drugs paclitaxel, vinblastine and methotrexate were additionally included, which changed the results of the analysis. COX3 was not identified in most of the added enrichments, which in turn led to a higher log_2-fold difference in the t-test for the enrichments in which it was present in all three replicates (sirolimus, lenalidomide). The interaction of thalidomide and structural analogues with the three isozymes of COX has been previously studied, yet the biological function of this interaction is still unclear.

Importantly, apart from COX3, all of the proteins identified as specific binders of the three natural products CsA, tacrolimus and sirolimus are previously known targets that have been identified in previous mass spectrometry-based proteomic methods. In contrast, as already observed for its structural analogue thalidomide, we also failed to detect the known targets for the small drug lenalidomide (Figure 28g), even though targeted immobilization of IMiDs via their phthalimide moiety can enrich their targets cereblon (direct binder) and DDB1 (indirect binder).[130,131] For paclitaxel, vinblastine and methotrexate (MTX) CISCM also failed to detect any specific binders, even though methotrexate immobilized on FG-beads can enrich its target DHFR.[139]

All the above mentioned compound binding targets were expressed in the Jurkat lysate we used in these experiments (Figure 27d).

3.4. Further development and optimization of the CISCM method

Using our novel CISCM method, we identified known and expected targets of three well-studied immunosuppressant compounds, which all have a cyclic peptide

structure, a high density of functional groups that can undergo a photoreaction with the used diazirine, and a spatially bulky structure. Until this point we were neither able to identify the target for the high molecular weight compounds vinblastine and paclitaxel nor the targets of the smaller or even fragment-like structures methotrexate and lenalidomide, even though the interactions with their corresponding targets could be preserved after side-chain modification of the drugs.[130,139]

The following chapter summarizes the variation of different parameters of the developed CISCM method, such as the spacer-arm length, the UV-irradiation times, the necessary molar amounts of compounds, as well as to upscale the method for future applications. Additionally, we further tested the photocrosslinking reaction and explored an alternative surface modification approach.

3.4.1. Dependence of CISCM target-ID on the PEG-spacer length

The length of the spacer and hence the amount of atoms separating the immobilized drug from the cellulose surface, is known to have a significant impact on the target binding.[143] We reasoned that increasing the spacer length to four instead of two PEG-units might improve the drug target interactions, especially for small molecules such as lenalidomide. A longer spacer was not considered in order to keep the protein binding towards the spacer itself to a minimum as this is known to be a significant source of false positives in AP experiments.[144] As CISCM with PEG_2-spacer length showed good performance for the compounds CsA and tacrolimus we included these as positive controls to monitor the performance after increasing the spacer length to four PEG-units. In addition to the previously studied small molecules (S)-thalidomide, lenalidomide, and methotrexate we added the fragment-like structure metformin. Metformin is used for treating of type 2 diabetes and subject of ongoing research in – to name a few - cancer, neurological disorders and aging, yet its mechanism of action is complex, controversial and highly debated.[145–148] It is known to inhibit fructose 1,5-bisphosphatase (FBP1), c-AMP response element-binding protein (CREB), glycerol-3-phosphatate dehydrogenase (GPD1) and the mitochondrial protein electron transfer flavoprotein-ubiquinone oxidoreductase (ETFDH) and to activate AMP-activated protein kinase (AMPK) elucidated by in vitro experiments.[145,149–151]

To study the effect of the spacer length on the target identification, TFC membranes were produced as previously discussed, however this time using a TPD conjugated to

a PEG_4-spacer arm instead of using a PEG_2-spacer. The mentioned compounds were spotted onto triplicates of these TFC membranes and immobilized via photocrosslinking. Unmodified, oxidized and NHS-activated cellulose without spotted compounds were used as controls. All cellulose membranes were incubated with HEK293T cell extract to enrich the interacting proteins. As before, cellulose spots corresponding to individual photocrosslinked compounds were excised, digested, and purified on StageTips.[125] Due to significant PEG-polymer contamination, the samples were additionally cleaned using a combination of strong cation exchange (SCX) and repeated subsequent desalting on StageTips. All 27 samples (six compounds and three controls, in triplicates) were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer.

As an internal photocrosslinking control we again spotted the TFC membranes with fluorescein, before and after their UV-irradiation and compared the fluorescence of the corresponding spotted compounds as described earlier (Figure 28a). As observed before (Figure 28b) both spots initially showed a strong fluorescence signal, which - after intense washing - decreased for the photodependent interaction and vanished for the photoindependent interaction (Supporting Figure 2).

Data analysis with MaxQuant[45] identified 3,259 protein groups (at 1% protein and peptide FDR) in all affinity enrichment samples combined. We first performed an analysis of variance (ANOVA, FDR 5%) with the LFQ data to obtain differentially abundant proteins. Hierarchical clustering of the ANOVA-significant proteins revealed clustering of most replicates (Figure 29a). Volcano plots were constructed from the LFQ data (Figure 29b-g) and as for the target identification using a PEG_2-spacer we selected specific binders requiring t-test p-values < 0.01 and fold changes of at least 4. In addition to the known targets PPIA, PPIB and PPIF this identified many other binders, predominantly 40S and 60S ribosomal proteins (Figure 29b, Table 1), which has previously been reported to be the case for a CsA target identification study.[80] For the drug tacrolimus the main target FKBP1A as well as two other members of the FKBP family, FKBP1B and FKBP4, could be identified. In contrast, we were unable to identify the known targets for any of the four small molecules methotrexate, (S)-thalidomide, lenalidomide, or metformin using the CISCM method with a longer spacer-arm (Figure 29d-g), even though their corresponding targets were present in the cell lysate we used

(Figure 29h). Hence, the variation of the spacer-arm length did not broaden the application range of the CISCM method.

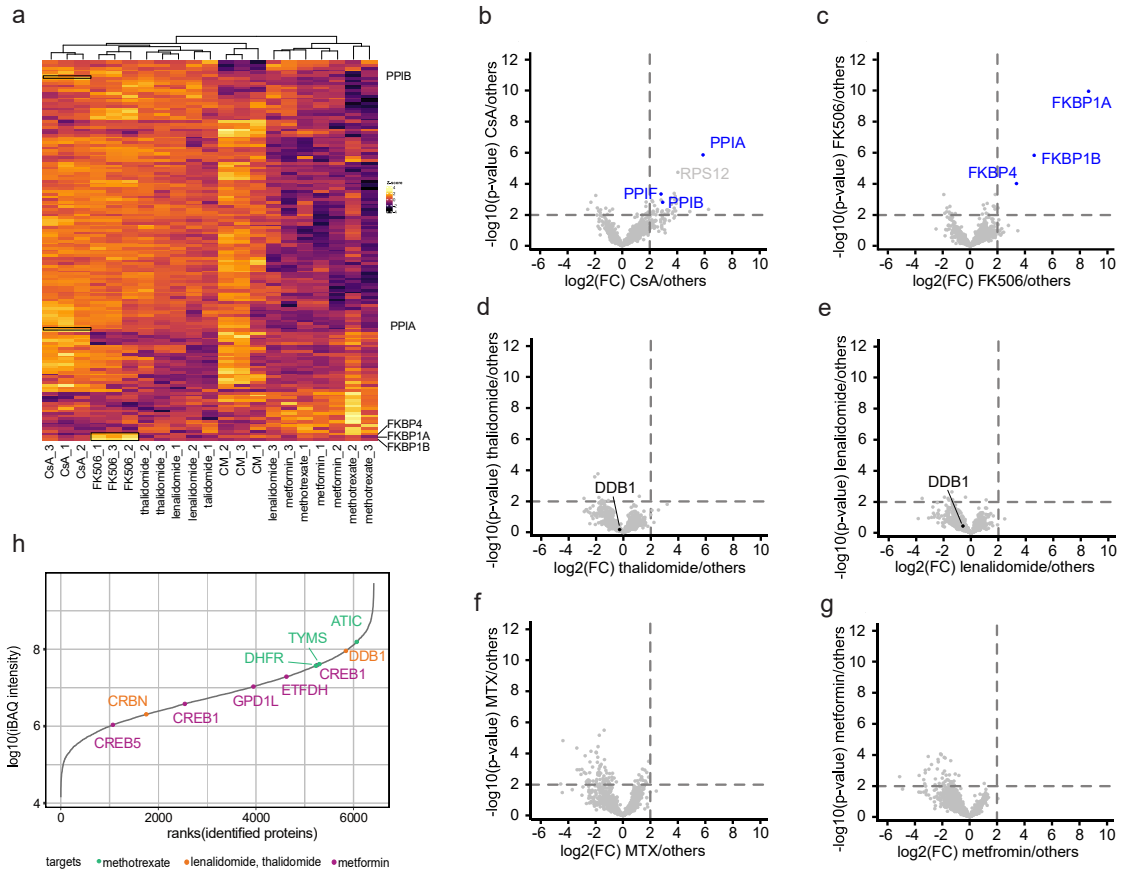


Figure 29. Target identification of a small test library after increasing the PEG-spacer length to four PEG-units. (a) Heatmap of Z-scores computed for ANOVA significant (FDR 5%, 250 randomizations) protein identifications of six immobilized compounds and one control. Each column is an individual replicate of the 21 samples and each row represents the Z-scored LFQ values. Known and identified targets of the positive controls are marked with a frame and labeled. **(b-g)** Volcano plots displaying the $\log_2$ fold change (x-axis) against the Student's t-test derived $-\log_{10}$ p-value (y-axis) for pairwise comparison of grouped triplicates of (a) cyclosporine (CsA), (b) tacrolimus (FK506), (c) methotrexate (MTX), (d) thalidomide, (e) lenalidomide and (f) metformin against all other samples (others). Proteins with t-test p-values < 0.01 and fold changes of at least four are labeled, known protein targets within that range are marked in blue and the ones outside that range –if identified - marked in black. For the sake of clarity proteins that are not part of the PPI protein family are not labeled, but listed in Table 1. **(h)** Ranked $\log_{10}$ iBAQ intensities of proteins identified in the for the affinity purification experiment used HEK293T proteome. Known and expected targets of the drugs with unsuccessful target-ID are labeled.

Table 1. Specific binders of cyclosporine A identified with a Student's t-test log_2 difference of at least 2 and a $-log_{10}$ p-value of at least 2.

MAJORITY PROTEIN IDS	PROTEIN NAMES	GENE NAMES	$-LOG_{10}$ STUDENT'S T-TEST P-VALUE	LOG_2 STUDENT'S T-TEST DIFFERENCE
A0A024R4M0;P46781	40S ribosomal protein S9	RPS9	2.097	3.307
A0A087WXM6;J3QQT2	60S ribosomal protein L17	RPL17;RPL17-C18orf32	3.193	3.846
A0A2R8YEM3;E7EQV9	Ribosomal protein L15;60S ribosomal protein L15	RPL15	2.719	3.922
H7BY10;K7EJV9	60S ribosomal protein L23a	RPL23A	2.210	4.906
C9JXB8;C9JNW5	60S ribosomal protein L24	RPL24	2.986	4.184
E7EPB3;P50914	60S ribosomal protein L14	RPL14	2.178	2.867
E9PKZ0;P62917	60S ribosomal protein L8	RPL8	2.254	2.924
I3L3P7;P62244	40S ribosomal protein S15a	RPS15A	3.091	3.541
K7EK07;Q71DI3	Histone H3;Histone H3.2	H3F3B;HIST2H3A	2.342	6.278
P18077;F8WBS5	60S ribosomal protein L35a	RPL35A	2.847	2.198
P18124;A8MUD9	60S ribosomal protein L7	RPL7	2.270	3.381
P19338;H7BY16	Nucleolin	NCL	2.964	2.829
P23284	Peptidyl-prolyl cis-trans isomerase B	PPIB	2.796	2.926
P25398	40S ribosomal protein S12	RPS12	4.734	4.049

P30405;R4GN99	Peptidyl-prolyl cis-trans isomerase F	PPIF	3.332	2.827
P39019;M0R2L9	40S ribosomal protein S19	RPS19	2.305	3.233
P49207	60S ribosomal protein L34	RPL34	3.368	3.793
P53985;Q5T8R5	Monocarboxylate transporter 1	SLC16A1	2.199	2.230
P60866;P60866-2	40S ribosomal protein S20	RPS20	2.038	3.075
P61353;K7ELC7	60S ribosomal protein L27	RPL27	2.079	3.647
P62081;A0A2R8Y623	40S ribosomal protein S7	RPS7	2.673	3.580
P62249;M0R3H0	40S ribosomal protein S16	RPS16	2.247	3.872
P62266;D6RD47	40S ribosomal protein S23	RPS23	2.583	2.343
P62851	40S ribosomal protein S25	RPS25	2.420	3.794
P62854;Q5JNZ5	40S ribosomal protein S26	RPS26;RPS26P11	2.980	3.798
P62899;H7C2W9	60S ribosomal protein L31	RPL31	2.350	2.559
P62937;F8WE65	Peptidyl-prolyl cis-trans isomerase A	PPIA	5.868	5.882
P62995-3;P62995	Transformer-2 protein homolog beta	TRA2B	2.064	2.212
P63244;D6RAC2	Guanine nucleotide-binding protein subunit beta-2-like 1	GNB2L1	2.215	2.546
Q02878	60S ribosomal protein L6	RPL6	2.198	3.697
Q14739;C9JXK0	Lamin-B receptor	LBR	2.003	2.451

Q7Z4V5-2;Q7Z4V5	Hepatoma-derived growth factor-related protein 2	HDGFRP2	2.330	3.279
Q8N5F7	NF-kappa-B-activating protein	NKAP	2.920	2.159
Q8TDD1;Q8TDD1-2	ATP-dependent RNA helicase DDX54	DDX54	2.061	2.294
Q96AG4	Leucine-rich repeat-containing protein 59	LRRC59	2.101	2.135
Q9Y5S9-2;Q9Y5S9	RNA-binding protein 8A	RBM8A	2.236	2.253

3.4.2. Dependence of CISCM target-ID on the presence of a PAL probe and of UV light

As the CISCM method showed good and reproducible performance with cyclic peptide drugs such as CsA, sirolimus (rapamycin) and tacrolimus (FK506), we wanted to study the dependency of target identification on the photocrosslinking reaction in more depth. As before we used NAC and TFC membranes and spotted six replicates for each membrane type with CsA. The half of these six replicates of CsA-NAC and CsA-TFC were then irradiated with UV light, whereas the remaining triplicates of each sample were not irradiated for control (Figure 30a).

As an optical control for the success of the photocrosslinking reaction we again used a qualitative fluorescence read-out of fluorescein. To that end fluorescein was spotted adjacent to CsA for each replicate. As expected, a strong fluorescence signal was observed in all samples directly after UV-irradiation and prior to any wash steps (Figure 30b). Samples were then washed in ethanol for one hour, which slightly decreased fluorescence in all samples. Even after further one hour washing steps with solvents DMF, THF, ethanol and methanol, a significant signal could still be observed in the negative controls without UV irradiation or diazirine. To further clean the membranes, we washed them for 53 hours in methanol which resulted in a significant decrease in the fluorescence signal in the not UV-irradiated fluorescein spotted samples. The fluorescence signal in the UV-irradiated replicates of diazirine containing TFC membranes remained constant, but surprisingly so did the fluorescence signal of UV-irradiated NAC membranes. This observation was not expected as no photoreactive group is present on the NAC membranes and consequently no photoreaction and no covalent attachment of fluorescein to the membranes is expected to occur and it was against previous observations presented earlier (chapter 3.2.3, Figure 22; chapter 3.3.2, Figure 28).

Protein interactors were then enriched from the HEK293T lysate, cellulose spots corresponding to one drug were excised and subjected to a digestion solution. Due to significant PEG-polymer contamination observed upon initial mass spectrometry analysis and to prepare the method for up-scaling to a larger compound library, we slightly modified the sample preparation steps. Bound proteins were eluted with a denaturating solution containing urea and thiourea followed by sp3 clean-up for sample

purification.[152] This adapted approach reduces polymer contamination significantly and allows processing of samples in a highly parallel and automated manner using a liquid handling platform, which enhances reproducibility whilst decreasing sample loss and risk of further contamination. The resulting 24 cleaned affinity enrichment peptide mixtures of both CsA and fluorescein in triplicates and at four different conditions were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer.

Data analysis in MaxQuant identified 1,374 protein groups (at 1% protein and peptide FDR) in all samples combined. We first performed an analysis of variance (ANOVA, FDR 5%) using LFQ intensities and ANOVA-significant proteins were clustered based on the z-scored protein interaction fold change to all other samples. This revealed two main clusters, one containing the CsA-spotted TFC membrane samples, and another one containing all other replicate samples (Figure 30c). Specific binders were selected requiring Student's t-test p-values smaller than 0.1 and $\log_2$-fold changes of at least 1.5 (Figure 30d). As before, most identified proteins did not bind preferentially. The known targets PPIF and PPIA (two isoforms) were specifically identified in the affinity enrichments of CsA-spotted TFC membranes (+D), surprisingly, independent of their UV-treatment. For the other two negative controls no known target of CsA could be specifically identified. Some binders passed the specificity cut-off for the UV-treated CsA-spotted NAC membrane. However, the binding of these proteins was suspected to not be related to CsA-binding and was therefore not further discussed. The results presented led us to two conclusions. First, elution of proteins interacting with compound-functionalized cellulose membranes using a denaturation mix of urea and thiourea and subsequent sample processing on sp3 beads can identify known interactors of the corresponding compound, however not with the same sensitivity. Second, specific drug-target interactions could be identified for both diazirine-modified (+D) cellulose membranes, independent of their treatment with UV light (+UV, -UV), which indicates that with the used conditions no UV-irradiation triggered photoreaction is necessary for a successful target-ID of CsA.

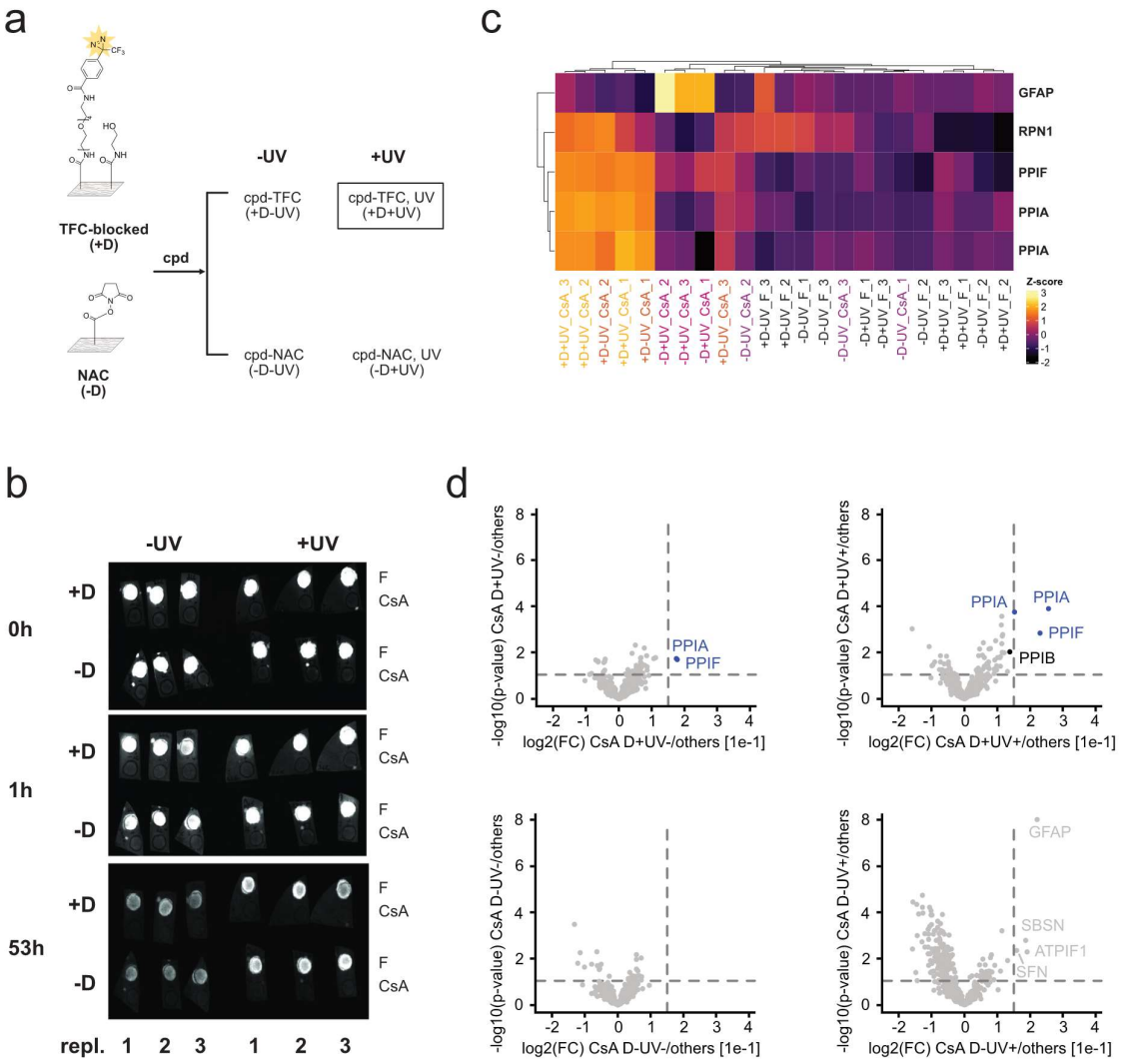


Figure 30. Sample overview, fluorescence control and qAP-MS results of CISCM in presence and absence of TPD and UV light. (a) Overview of samples and their naming: diazirine-containing (+D) and diazirine-free (-D) cellulose membrane, spotted compounds (cpd) CsA or fluorescein (F), UV-irradiation at 365 nm for two time 30 min (+UV) or storage in darkness (-UV). The positive control is indicated with a frame. **(b)** Fluorescence intensities of triplicates (replicates 1-3) of TFC (+D) and NAC (-D) membranes spotted with fluorescein (F) and CsA treated with UV light (+UV) or not UV-irradiated (-UV) and after washing in different solvents over time (0-53 h). **(c)** Heatmap of Z-scores computed for ANOVA significant (FDR 5%, 250 randomizations) protein identifications of two immobilized compounds at four different conditions. Each column is an individual replicate of the total 24 samples and each row is the Z-scored protein identifications. **(d)** Volcano plots displaying the $\log_2$ fold change (x-axis) against the Student's t-test derived $-\log_{10}$ p-value (y-axis) for pairwise comparison of grouped triplicates of cyclosporine (CsA) at four different conditions against all other samples (others). Proteins with t-test p-values < 0.1 and $\log_2$-fold changes of at least 1.5 are labeled, unknown interactors within that range are marked in grey whereas known protein targets within that range are marked in blue and the ones outside, but near that range marked in black.

3.4.3. Target identification of inkjet printed natural compounds

In chapter 3.4.2 we showed significant enrichment of known protein interactors of CsA in TFC membranes even in membranes not irradiated with UV, to which no covalent attachment via photoimmobilization was expected. We reasoned that non-covalent attachment of cyclic peptide compounds towards cellulose membranes might be sufficient to significantly enrich compound-target interactions. To this end we repeated the experiment and spotted compounds and controls onto non-functionalized cellulose filter membranes. To broaden the cellulose network and to increase the binding surface, cellulose membranes were incubated in sodium hydroxide, but not further functionalized.[121]

To first test to which extent compounds would be washed off during the affinity enrichment and following washes, we spotted two dyes – ponceau S and 6-aminoflurescein - on triplicates of NaOH-activated cellulose membranes and followed their signal at different stages of the protocol (Figure 31a). The spotted membrane was dried intensively and interacting proteins were enriched from a whole Jurkat cell lysate as before, however this time incubating the lysate and membrane for only 30 minutes followed by mild washing in a detergent-free lysis buffer to minimize the release of the compounds from the cellulose. Before washing, both dyes were clearly visible with a strong colour (before). Conditioning the membranes for five minutes in a detergent-free lysis buffer in order to prepare them for the affinity enrichment (after cond.) led to slightly reduced signal intensity for both dyes. After affinity enrichment (after AP) and subsequent washing (after washes) the colour intensity of the dyes as well as the fluorescence intensity of the 6-aminofluorescein control decreased significantly. These results were anticipated, as the compounds were expected to be bound to the non-functionalized cellulose membranes only by non-covalent forces.

To test whether the non-covalent attachment of compounds to cellulose would enable target enrichment three cyclic peptide drugs cyclosporine A (CsA), tacrolimus (FK506) and sirolimus (rapamycin), for which we previously achieved successful identification of protein interactors, were spotted and treated as described for the dye controls. After affinity enrichment and mild washing, cellulose spots corresponding to each of the three compounds were excised, interacting proteins were digested, and the resulting peptides were desalted. High resolution mass spectra of the resulting nine samples

(three compounds in triplicates) were acquired on a Q Exactive HFX mass spectrometer and analysed in MaxQuant, resulting in 2,157 protein groups (at 1% protein and peptide FDR) for all affinity enrichment samples combined.

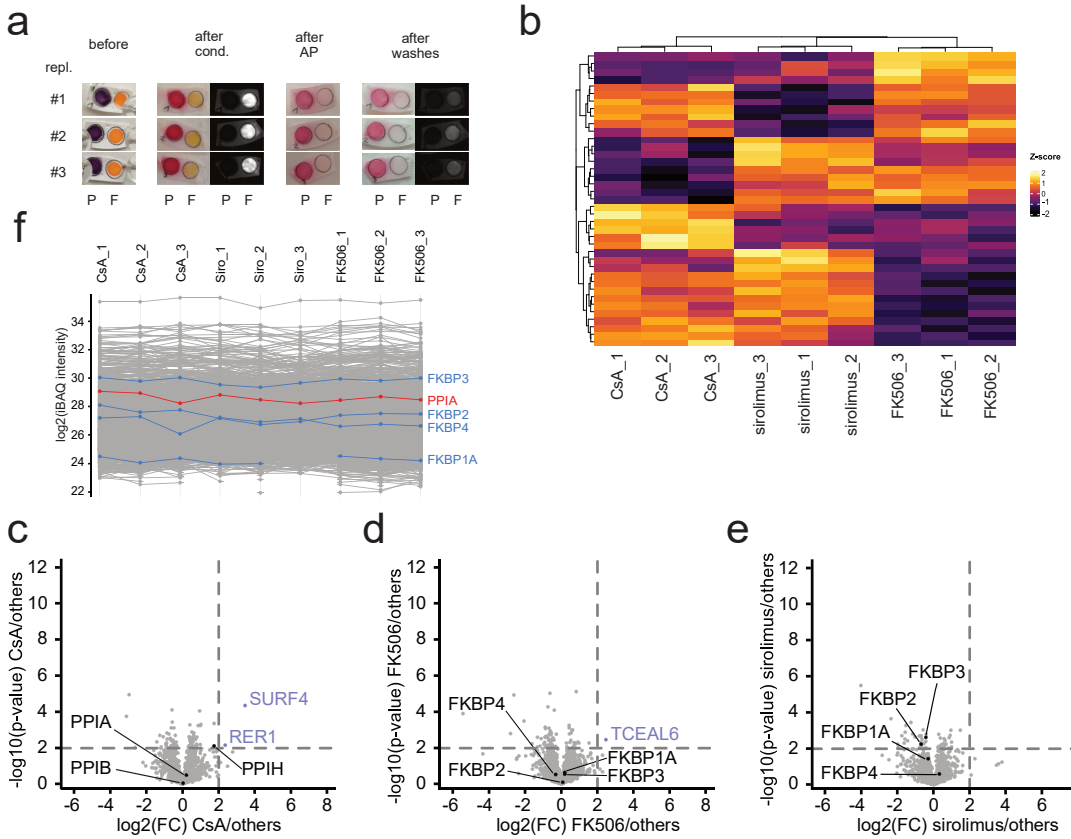


Figure 31. Target identification of inkjet printed natural compounds. (a) Camera and fluorescence images of inkjet printing controls ponceau S (P) and 6-aminofluorescein (F) in triplicates (repl. #1-3) before any washes (before), after conditioning in detergent-free lysis buffer for 5 min (after cond.), after affinity purification in a Jurkat cell lysate for 30 min at 4 °C (after AP) and after washing the membranes in detergent-free lysis buffer at 4 °C three times for 5 min each (after washes). **(b)** Heatmap of Z-scores computed for ANOVA significant (FDR 5%, 250 randomizations) protein identifications of the three compounds cyclosporine A (CsA), tacrolimus (FK506) and sirolimus. Each column is an individual replicate of the total 9 samples and each row is the Z-scored LFQ intensity of ANOVA significant protein identifications. **(c-e)** Volcano plots displaying the $\log_2$ fold change (x-axis) against the Student's t-test derived -$\log_{10}$ p-value (y-axis) for pairwise comparison of grouped triplicates of (b) cyclosporine (CsA), (c) tacrolimus (FK506), (d) sirolimus, respectively, against all other samples (others). Proteins with t-test p-values < 0.01 and fold changes of at least four are labeled in purple and known protein targets that did not pass this threshold are labeled and marked in black. (f) Profile plot displaying the $\log_2$ iBAQ values (y-axis) for comparison of protein intensities within each sample of the total 9 samples (x-axis). The $\log_2$ iBAQ values for the main known target of CsA (PPIA) are marked in red, whereas the $\log_2$ iBAQ values of the known targets of tacrolimus and sirolimus (FKBP1A, FKBP2, FKBP3 and FKBP4) are marked in blue.

To further analyse the significance of these proteins we first performed an analysis of variance (ANOVA, FDR 5%) using LFQ intensities (Figure 31b). Hierarchical clustering of ANOVA-significant protein identifications revealed clustering of replicates and separated the affinity enrichments of CsA from those of the structurally related drugs

sirolimus and tacrolimus, which further separated into two subclusters. The specificity of the interactions was analysed with a Student's t-test and proteins with a p-value < 0.01 and fold change of at least 4 were considered as specific binders, whereas all proteins not passing that specificity threshold were considered as background binders (Figure 31c-e). The proteins SURF4 and RER1 were identified as specific interactors of CsA, TCEAL6 of tacrolimus (FK506), and no specifically interacting proteins for the compound sirolimus, whereas none of the known targets passed the specificity cut-off although having been identified in the samples. To further corroborate this finding we used intensity-based absolute quantification[153] (iBAQ, Figure 31f). However, we did not observe higher iBAQ intensities for the known protein interactors in the affinity enrichment samples of the corresponding compound when compared to other proteins within the same sample.

The presented results indicate, that the presence of a photocrosslinker is necessary to successfully identify the molecular targets of cyclic peptide drugs, when these are attached to cellulose membranes, at least under the chosen conditions.

3.4.4. Alternative CISCM approach using click chemistry (click-CISCM)

Previous experiments using CISCM showed the identification of known and expected interacting protein partners of cyclic peptide compounds, but not of the small molecule compounds lenalidomide and methotrexate. Cellulose membranes contain a high amount of hydroxylic side-groups that are expected to be in the close proximity of the photocrosslinker due to dipole-dipole interactions. We hypothesised that performing the photoreaction on such a membrane might result in a high degree of side-reactions of the photocrosslinkers with the cellulose surface which can significantly lower the yield of the photoreaction and hence lead to a lower local density of immobilized compound. In addition, given that only a fraction of the photoimmobilized compound is crosslinked in a conformation that allows binding to proteins of biological relevance, the amount of immobilized and still available compound molecules might not be sufficient for a successful target identification. This might especially be relevant for compounds with limited functional groups for the photoreaction or that have a sensitive structure affinity relationship (SAR), as it is the case for the small molecule drugs lenalidomide and methotrexate.

We reasoned that photocrosslinking of the compounds in isolation and subsequent immobilization of the photocrosslinked constructs would allow both: isolation and analysis of individual reaction products to monitor and optimize the photocrosslinking reaction towards higher yields and prevention of a side-reaction of the photocrosslinker with the cellulose membrane. We decided that copper(I)-catalysed alkyne-azide cycloaddition (CuAAC, 'click chemistry') would provide a suitable approach to immobilize crosslinked constructs to azide-functionalized cellulose membranes, as it is a chemo- and site-specific reaction towards terminal alkyne moieties, which are rarely occurring in drug-like substances. Hence for the majority of photocrosslinked compounds no side-reactions between side chains of the crosslinked molecules and the azide-cellulose are expected. Furthermore CuAAC leads to almost complete conversions and is compatible with cellulose.[121,154,155]

In all previously described applications of the CISCM method NHS-chemistry was used to construct a photoactivatable cellulose (TFC) membrane (Figure 32a). This TFC membrane was then spotted with the compounds of interest, dried and irradiated with UV light to photoimmobilize the compounds and to construct a small molecule affinity matrix. Proteins interacting with this constructed small molecule affinity matrix were then enriched from a whole proteome and spots corresponding to each small molecule of interest were excised.

Using a click chemistry approach, we changed the order of the first two steps (Figure 32b): Each compound of interest was mixed with an alkyne-functionalized trifluoromethylphenyl diazirine (alkyne-TPD), dried and photocrosslinked separately. Next each crosslinking product pool corresponding to one compound of interest was resuspended and immobilized onto an excised spot of an azide-activated cellulose (AAC) membrane[121] via click chemistry and in a 96 well plate. After washing of each spot, HEK293T cell lysate was added to each well of the plate. Interacting proteins were enriched on each spot and digested. Resulting samples were digested, desalted and analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer (Figure 32c). The data was analysed as described for NHS-chemistry based CISCM screens and results of a Student's t-test using the LFQ data was displayed as volcano plots (Figure 32c).

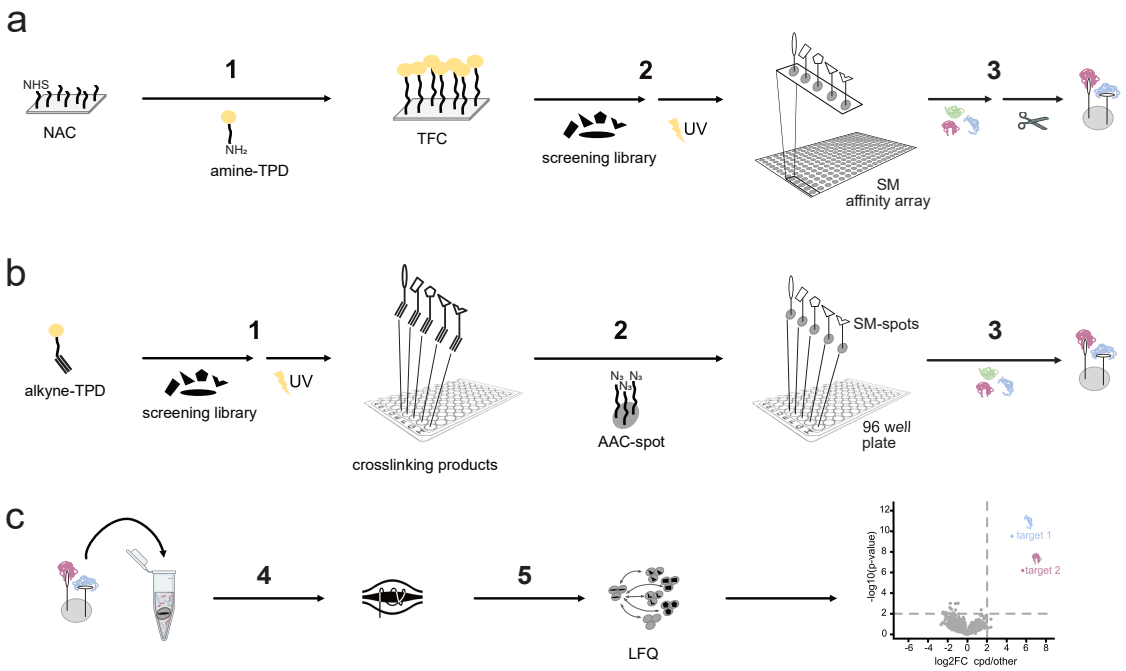


Figure 32. Comparison of CISCM approaches using NHS- or click chemistry: (a) Classical CISCM approach using NHS-activated cellulose (NAC) membranes. (1) Functionalization of NAC membranes with a trifluoromethylphenyl diazirine (TPD) via a nucleophile substitution to obtain a TPD-functionalized cellulose membrane (TFC). (2) Spotting of the TFC membrane with compounds of a screening library, photoimmobilization of small molecules in proximity to the TPD induced by UV light resulting in a small molecule (SM) affinity array. (3) Affinity enrichment from cell protein extract and excision of compound-cellulose spots. **(b)** Alternative CISCM approach using CuAAC to immobilize photocrosslinked compounds. (1) Photocrosslinking (UV) of an alkyne-modified TPD crosslinker with compounds of a screening library separately after evaporation of the solvent. (2) Covalent attachment of alkyne-functionalized photocrosslinking products on azide-activated cellulose-spot (AAC-spot) using CuAAC. (3) Affinity enrichment on small molecule spots (SM-spots) from cell protein extract in a 96 well format. **(c)** Common sample preparation and analysis. (4) Protein digestion and LC-MS measurement. (5) Target identification via quantitative analysis.

To control the success of both, the photoreaction and the click reaction, we again implemented fluorescein as an optical control in the manner that we photocrosslinked fluorescein with alkyne-TPD (Figure 32b, step 1). To explicitly control the success of the click reaction we additionally used 6-FAM-alkyne – an alkyne containing fluorescein derivate. To control non-covalent attachment of the dye to the cellulose membrane we used unmodified fluorescein. Triplicates of each of the three constructs, photocrosslinked fluorescein, 6-FAM-alkyne and fluorescein, were immersed separately in a reaction mixture containing $CuSO_4$ as well as sodium ascorbate and an excised spot of azide-activated cellulose was added to each replicate. After 24 hours the cellulose spots were washed thoroughly and the fluorescence of each spot was measured. For azide-cellulose spots, which were incubated with either the photocrosslinked fluorescein (F+D-alkyne) or the 6-FAM-alkyne (F-alkyne) at click

reaction conditions, a strong fluorescence could be observed (Figure 33a, Supporting Figure 3a). In contrast, the azide-cellulose spots incubated with unmodified fluorescein (F) in the click chemistry reaction mix, showed no fluorescence for two replicates and low fluorescence for one replicate. We concluded that the strong fluorescence observed in the positive controls indicates the success of both, the photo- and click reaction, and did not derive from non-covalently attached fluorescein.

To test this approach on its potential to identify drug targets, we applied it to CsA, as the target-ID of this compound has shown good and reproducible performance in the NHS-based CISCM approach. As an example for small molecules with a limited number of functional groups and a sensitive structure affinity relationship we additionally included both enantiomers of thalidomide. The click-CISCM protocol (Figure 32b, step 1-3) was performed with the three compounds in parallel to the fluorescein controls. To control unspecific background binding, we also included azide-activated cellulose (AAC) membranes for the affinity enrichment. The fluorescence of the fluorescein controls was measured, using the samples and the AAC spots as a background. As expected only for the fluorescein controls a fluorescence signal could be obtained (Supporting Figure 3a). The 12 samples (CsA, (R,S)-thalidomide, AAC, in triplicates) were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer.

In all affinity enrichment samples combined click-CISCM identified 4,779 protein groups after data analysis in MaxQuant (at 1% protein and peptide FDR). For the affinity purification of the immobilized drug CsA the known targets PPIA and PPIG could be identified as specific binders requiring Student's t-test p-values < 0.01 and fold changes of at least 2, among five other proteins with unknown affinity (Figure 33b, c). When compared to the affinity enrichment of the same drug but with a NHS-based CISCM approach and the same PEG-spacer length (four PEG units), the main target PPIA could be enriched with much higher specificity and a greater t-test difference when compared to the background (Figure 33b). However, the background sample size for the analysis of the NHS-based CISCM screen was greater than in the click-CISCM screen (Figure 29), which can also influence the results. For the S-enantiomer of thalidomide the protein PDE6D, which has also been identified in another experiment (Figure 25), could again be identified among several other binders. Another identified specific binder for S-thalidomide is the transcription factor YY1 which is – as

CRBN and DDB 1 - part of the nucleotide excision repair pathway.[156,157] Again, neither CRBN nor the indirectly binding protein DDB1 could be identified for any of the thalidomide enantiomers within the specificity criteria. We concluded that either photocrosslinking of small compounds in isolation did not improve the target identification or that – also taking the inferior results of the compound CsA when compared to NHS-based CISCM results into account - the click reaction parameters needed to be further improved. The affinity enrichment from an empty azide-activated cellulose control showed no strong background binding. We therefore didn't expect any significant influence of unreacted azide-groups after the click reaction to the target identification (Supporting Figure 3b).

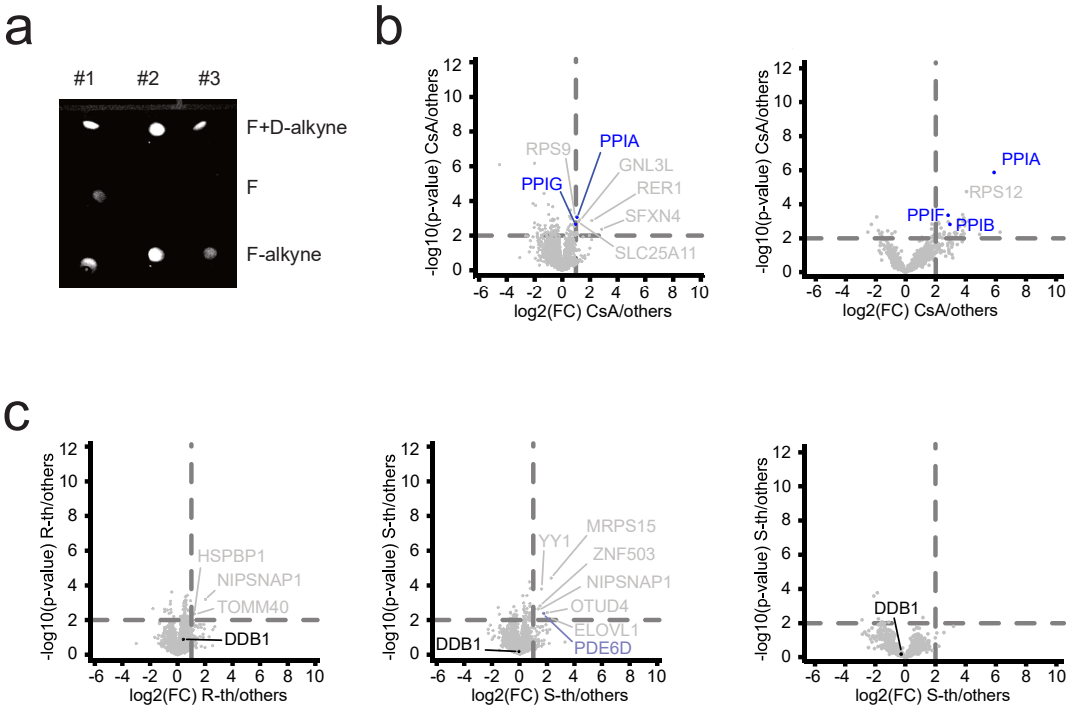


Figure 33. Target identification of a click chemistry-based variation of CISCM (click-CISCM). (a) Fluorescence signals of azide-cellulose spots after incubation at click reaction conditions in triplicates with photocrosslinking constructs of fluorescein (F+D-alkyne), with non-photocrosslinked and unmodified fluorescein (F) and with 6-FAM-alkyne (F-alkyne). **(b,c)** Volcano plots displaying the $\log_2$ fold change (x-axis) against the Student's t-test derived $-\log_{10}$ p-value (y-axis) for pairwise comparison of grouped triplicates of one drug-protein affinity enrichment against all other samples (others). Proteins with t-test p-values < 0.01 and fold changes of at least two (click-CISCM variation) or four (NHS-based CISCM) are labeled and known protein targets within that cut-off marked in blue and outside the cut-off (if identified) marked in black. Unknown specicific binders identified in other experiments of this thesis are marked in purple. (b) Volcano plots of CsA affinity enrichments using a click-variation of CISCM (left) and using normal CISCM (right). (c) Volcano plots of R-thalidomide (left) and S-thalidomide (center) affinity enrichments using a click-variation of CISCM and for comparison a volcano plot of an affinity enrichment of S-thalidomide (right) using NHS-based CISCM.

4. Discussion

One of the greatest bottlenecks in modern drug discovery remains the deconvolution of molecular mechanisms of action amongst phenotypic screening hits. A common approach involves elucidating a hit's molecular interactions to identify both, the targets that are responsible for a desired pharmacological effect as well as additional interactors that might cause undesired side-effects or that might alter the substance's efficiency (off-targets). For protein targets, existing proteomic approaches have enabled drug profiling in a nearly unbiased manner (label-free approaches). However, these methods typically require long LC-MS measurement times (i.e. thermal proteome profiling) and lack the throughput needed to screen larger compound libraries. Affinity-based approaches on the other hand allow a higher throughput, but require modification of the studied substance and hence entail a bias (conventional qAP-MS).

Here, we presented a scalable compound interaction screen on a photoactivatable cellulose membrane (CISCM), a qAP-MS approach, that is able to rapidly screen for the targets of natural compounds in a parallelized and comparably unbiased fashion. A key advantage of the approach - when compared to conventional qAP-MS approaches - is that it does not require the implementation of an immobilization handle and hence no laborious functionalization nor previous SAR studies, while still maintaining the scalability and speed of qAP-MS approaches. This is enabled by the use of diazirine-based undirected photoaffinity labeling (PAL), which allows immobilization of a broad range of functional groups.[82] Thus, while crosslinking via isocyanate or NHS potentially provides higher conversion for specific nucleophiles,[158] the diazirine-based approach employed here is more broadly applicable. Undirected photocrosslinking has been used before to assess the interaction of individual candidate proteins with drug libraries, yet we show that when combined with q-AP-MS it enables unbiased proteome-wide screening. Consequently, multiple protein target candidates can be studied in the complex context of their interaction network, which also opens the screening of the drug to yet unexpected target classes with the potential to expand the druggable proteome. To our knowledge, CISCM is the first chemoproteomic method that allows high-throughput target ID of multiple compounds and multiple candidate target proteins in parallel using PAL as immobilization method.

The main results of the development and application of CISCM will be discussed in the next section with a short glance at the further optimization potential and future applications.

4.1. Results summary and discussion

Following the protocol developed by Orelma and coworkers,[122] we successfully functionalised conventional cellulose filters with NHS in an easy in-house applicable manner and in only 2 days. We were able to confirm the activation of cellulose with XPS and ATR-FTIR (chapter 3.2, Figure 19, Figure 20). However, as this required expensive equipment and specially trained personnel on this complex technology, we employed a fluorescent assay (Figure 21). The NAC membranes were still active after storage at low vacuum conditions and in the presence of a desiccant even after 3 months (Supporting Figure 1). We equipped these NAC membranes with a trifluoromethylphenyl diazirine (TPD) using NHS-chemistry to obtain TPD-functionalized cellulose (TFC) membranes (chapter 3.2, Figure 19, Figure 20).

Using TPD-functionalized cellulose membranes we were able to photoimmobilize the natural compound cyclosporine A (chapter 3.2.3, Figure 23, Figure 24) and to identify its targets in a reproducible manner even after changing parameters of the protocol and using different backgrounds in the data analysis (chapter 3.3.1 - 3.4.2, Figure 25, Figure 28, Figure 29, Figure 30).

The developed and here presented CISCM screen also showed good and reproducible performance for the immunosuppressive natural compound analogues tacrolimus (FK506) and sirolimus (rapamycin), which share FK-binding proteins (FKBPs) as common targets, even when they were screened in the same experiment and hence in each other's background for pairwise comparisons (chapter 3.3.2, chapter 3.4.1, Figure 28, Figure 29). In contrast, we were unable to identify the shared target tubulin (TUBB) of the natural compound analogues vinblastine and paclitaxel, even though this target was sufficiently abundant in the cell lysate we used (Figure 27). Regarding their molecular weight in the range of the previously mentioned natural compound analogues for which CISCM showed good performance, and their abundance of functional groups reactive towards the used photocrosslinker we concluded that the unsuccessful target ID must have a different reason. Both vinblastine and paclitaxel bind their target TUBB at its dimer interface, which requires GTP to form.[159] As our

developed CISCM screen uses proteins extracted from lysed cells, where we expect GTP not to be produced, TUBB might solely be present in its monomeric form and hence not bind to the immobilized drug. Another possible explanation could be, that if the alpha-beta tubulin dimer interface is part of tubulin microtubule – as it is the case in living cells- this complex might be too sterically demanding to bind the immobilized drugs vinblastine and paclitaxel. Both possible explanations for missing tubulin in the CISCM for paclitaxel and vinblastine present potential limitations towards big target complexes and targets that reach their active binding form only under physiological conditions or metabolism. However, in future applications of CISCM to screen for tubulin binding substances, tubulin polymerization can be promoted by treating the cell lysate with Taxol and GTP.[160]

The developed CISCM method showed another limitation for small fragment-like drugs, such as metformin and methotrexate (Figure 28, Figure 29) or the IMiDs thalidomide and lenalidomide (Figure 25, Figure 28), which are known to have a sensitive SAR. The unselective nature of the photoreaction leads to a distribution of carbene insertion products, of which a fraction is expected to be inactive, leading to loss of target binding. This especially affects the target-ID of small molecules with a limited number of reactive attachment sites and molecules with a sensitive target binding affinity towards side-chain modifications.[161] The target-ID of those compounds remained unsuccessful even after increasing the spacer length which separates the drug from the surface (Figure 29) and after photocrosslinking the small molecules R- and S-thalidomide isolated from the cellulose membrane to prevent the photocrosslinker from a reaction with the abundant hydroxyl groups of the cellulose membrane (Figure 33). A longer spacer arm than four PEG-units was not considered as it might increase unspecific background binding. An alternative explanation for the unsuccessful target-ID of the IMiDs could be explained by the stronger interaction of CsA, FK506 and sirolimus towards their respective targets, when compared to the drug-target affinities of thalidomide and lenalidomide, even though quantitative affinity purification and mass spectrometry can in principle detect relatively weak interactions.[60] In addition, the choice of the photocrosslinker could also play a role. Even though photocrosslinkers don't exclusively react with only one particular functional group they still show unique labeling preferences.[81–83] Diazirines for instance insert into H-C and H-heteroatom bonds, however, their preferences towards these

groups depend on the diazirine's side-chains.[81] It has been shown that using a mixture of different diazirines and other PAL groups leads to a more diverse immobilization as the small molecules are attached in dependence of each photocrosslinker's site preferences leading to a more homogenous distribution of photocrosslinking products.[81,83] This in turn could increase the ratio of bioactive structures in the photocrosslinking product distribution, which in turn has previously been reported to improve target enrichment.[83,128] Additionally, this approach would be expected to boost immobilization efficiency as the different photocrosslinkers compensate for each other's reactivity profiles. In conclusion, using a mixture of different photocrosslinkers, such as aliphatic and aryl diazirines, benzophenones and aryl azides, might improve the overall performance of CISCM and overcome its current limitation to identify targets of small molecules, such as IMiDs, methotrexate or metformin.

An increased length of the PEG-spacer arm led to increased background binding for the target-ID of CsA, which corroborates observations of comparable studies, yet the overall performance of the target-IDs of the natural compounds CsA and FK506 did not significantly change (Figure 29). The photocrosslinking of the drug CsA isolated from the cellulose membrane and the subsequent immobilization of the photocrosslinking construct using click chemistry in contrast decreased the performance of the target-ID (Figure 33) when compared to the target-IDs of CsA using the NHS-based approach of CISCM. The known targets of these compounds were identified with a much smaller $\log_2$ fold changes. However, we assume that optimization of the chosen parameters for the click reaction could lead to comparable or even improved results when compared to the NHS-based approach of CISCM as it circumvents side-reactions of the photocrosslinker with the surface material and hence should lead to higher compound photocrosslinking yields. The usage of additives such as the ligand 3-[4-[[bis[[1-(3-hydroxypropyl)triazol-4-yl]methyl]amino]methyl]triazol-1-yl]propan-1-ol (THPTA) and fine tuning of the reactants' concentrations and molar ratios as well as the reaction temperature and time could significantly improve the results. As an additional advantage, the click-CISCM approach allows to monitor the photocrosslinking reaction products prior to their immobilization, giving an additional level of control and insight, especially for unsuccessful target-IDs.

Surprisingly, we were also able to identify the known targets PPIA and PPIF of the compound CsA using the CISCM method even if compound-spotted photoactivatable

cellulose membranes were not irradiated with UV light and no photoimmobilization of the compound is expected to occur (Figure 30). Interestingly, we could neither identify covalent nor physisorbed CsA in an analogously prepared sample analysed by ToF-SIMS (CsA-TFC, Figure 23b, Figure 24a). A fragment and fingerprint signal corresponding to CsA could only be identified if CsA-spotted cellulose membranes were treated with UV light and hence where we expect the compound to be photoimmobilized (CsA-TFC, UV). Even though ToF-SIMS is a very sensitive method detecting analytes in the range from ppm to ppb, it is limited to the penetration depth of the first two to three atomic layers.[113,114,162] A possible explanation why two known targets of but not the compound itself could be identified in the CsA-TFC sample might be that the high molar amount with which the compound was spotted, led to its complete penetration of the TFC membrane. This could be observed for the UV-treated CsA-TFC sample as a signal of the immobilized compound could be detected on both sides of the spotted TFC membrane even if the irradiation occured on its backside (Figure 24b). We therefore assume that the compound is present between front- and backside of the cellulose and hence also buried in its deep layers also before its covalent attachment by UV light. In the following washing steps of the CISCM protocol the non-covalently bound compound might only be washed away from the upper layers of the front and backside of the TFC cellulose, but remain present in the deeper layers, even after intense washing with different solvents. Its bulky structure and hydrophobic nature support this hypothesis. This might explain, why no signal for non-covalently bound CsA could be observed when analysing the CsA-TFC sample with ToF-SIMS. We expect this effect to be enhanced by the presence of the TPD-crosslinker, sterically and non-covalently hindering the release of the compound. This in turn might explain why we identified known targets of CsA in the presence of the TPD-crosslinker (Figure 30d, top), but not in its absence (Figure 30d, bottom) – independent of UV-irradiation. The compound's targets in contrast could have been released during the strong denaturing conditions of the shotgun proteomics sample preparation and identified by high resolution LC-MS/MS. One approach to further study the presence and absence of the drug CsA in the cellulose network and the nature of its binding to the cellulose structure could be to widen or break the cellulose structure of samples prepared analogously to such shown in Figure 30a. This could be achieved with highly concentrated bases and released structures could be analysed by LC-MS. However,

this might also affect the formed photocrosslinks with the drug and some drugs might not be stable under such conditions. Therefore employing a cleavable spacer between the cellulose and the photoreactive group instead of a PEG-spacer might offer a more straightforward solution. Using such a spacer in future applications of CISCM would offer the following additional advantages. Cleaving the spacer after the functionalization of NHS-activated cellulose with the diazirine moiety would allow to estimate the functionalization density when taking the expected cleavage yield into account. Even more important, after the photoreaction between the diazirine and the drugs, the photocrosslinking products could be cleaved from the cellulose membranes and analysed with LC-MS, nuclear magnetic resonance (NMR) or - if a fluorescent dye was photocrosslinked – with a spectrofluorometer to control the success of the photoreaction, to gain insights into different attachment sites of the photocrosslinker and even to quantify the yield of the photoreactions. An example for a cleavable linker is a disulphide linker, which can be cleaved using dithioreithol (DTT), resulting in higher cleavage yields when compared to photocleavable linkers. DTT is also used in standard shotgun proteomics workflow to cleave disulphide bonds in proteins and hence to assist the denaturation of protein structure to enhance the following tryptic digestion yield. The cleavage of drug-target complexes could hence be easily integrated to the shotgun proteomics workflow during denaturation and before digestion.

To increase the throughput for future CISCM applications, we chose automated sp3 on-bead digestion and sample clean-up. To that end, the enriched subproteomes were eluted from the drug-cellulose spots using a denaturation mix of concentrated urea and thiourea. However, especially for strong drug-target-interactions like the CsA-PPIA interactions, targets might partially remain on the drug-cellulose membrane. This is indicated by the much smaller fold-changes in the target-IDs of CsA (Figure 30) when compared to fold changes using digestion on cellulose spots along with StageTip purification (Figure 25b, Figure 28d, Figure 29b). Cleaving the entire drug-target complexes from the cellulose membrane instead of eluting the enriched targets could circumvent this current issue. Other ways of cleaving crosslinking products from cellulose membranes could be photoinduced cleavage[93,128] or using ammonia vapor.[163] Alternatively, to study the presence and quantities of CsA on the cellulose membrane a radiolabelled version of the drug – such as provided by American Labeled

Chemicals Inc. (MEbMT-b-3H, ART 1717) - could be used and detected with high sensitivity.

As discussed in the previous paragraph, we were able to identify known targets of the drug CsA, even if the drug was not photoimmobilized to cellulose membranes. Based on these results, we aimed to ascertain whether it would be possible to print drugs onto cellulose membranes in a non-covalent manner and to identify the corresponding targets. The non-covalent attachment of drugs to cellulose membranes has been previously reported in the context of novel drug dosage and delivery approaches and is referred to 'inkjet printing'.[164] We spotted the drugs CsA, sirolimus and FK506 onto non-functionalized cellulose membranes, dried them intensively and screened them against an entire proteome. The target-ID was unsuccessful for all three drugs. The compound spotted membranes were not washed prior to the target-ID in order to keep the non-covalent on-spot drug concentration as high as possible. Yet this might have led to the failure of the target-ID. The compounds are expected to be partially washed off the cellulose membranes during the incubation in cell lysate, even during short incubation times, and hence might have acted as in-solution concurrent opponent to the drug attached to the cellulose membrane in terms of the target binding. The release of a proportion of the drugs during lysate incubation might have also led to a distribution of the drugs across the cellulose membrane, leading to the enrichment of the corresponding targets across the membrane and hence across the drug background and not only on the original compound spot. This might explain why the protein identifications of expected targets in the subproteomes of the corresponding drug-spot – for example PPIA, PPIB and PPIH for the drug CsA - were not significantly or not at all different from the protein identifications of the background proteomes – in this case the protein-IDs of PPIA, PPIB and PPIH or the drug CsA in the subproteomes of FK506 and sirolimus. For future experiments it might be beneficial to extensively wash the dried drug spotted membranes prior to the target enrichment to prevent the two discussed limiting factors and to keep only the proportion of the drug that is buried in the deeper layers of the cellulose network. Additionally the cellulose membranes could be functionalized with NHS and subsequently with an amine-PEG-spacer containing a terminal aryl moiety to mimic the conditions of the unexpected target-ID presented in Figure 30 (+D-UV). Based on our current data it is not possible to estimate whether a

target-ID of inkjet printed drugs might be an attractive alternative to our here developed CISCM method.

4.2. Further outlook

As discussed in the previous chapters, the CISCM method described and developed here reliably detects targets of larger and complex compounds such as the natural structure analogues CsA, FK506 and sirolimus. Such structures, with a high mass and with properties beyond Lipinsky's rule of five have significantly gained importance in drug discovery as it can be observed by their increased number of drug approvals in the past years.[22,23] Amongst these, natural products or natural inspired analogues are the most important class as they comprise evolutionary optimized molecules with a broader structural diversity compared to purely synthetic structure pools. They further show high relevance for infectious diseases and cancer due to their original biological function. Due to its good performance for this important compound class, its simplicity and throughput, CISCM is perfectly suited to screen libraries of such structures, especially for structures of natural origin with confirmed or suspected bioactivity but yet unknown modes of action. There are numerous examples of screening libraries of natural or natural-like structures with known or unknown bioactivity and derived from different biological origins. The National Center for Complementary and Integrative Health provides a comprehensive overview for such libraries (https://www.nccih.nih.gov/grants/natural-product-libraries). Alternatively CISCM could be used to test drugs and structures that are in an advanced stage of drug discovery for their repurposing potential.[24] To give an example to such a library, we cherry-picked 46 natural compounds and derivatives (Figure 34) listed in the database ChEMBL, which are suspected to be applicable to alternate disease indications. This selection includes the compounds CsA, sirolimus and tacrolimus, for which CISCM has shown reliable performance. A Tanimoto similarity network[165,166] (Figure 34, Supporting Table 1, Supporting Table 2) of this handpicked library is shown to depict its chemical space. Structures are displayed as nodes and the connecting edges are weighted by the structural similarity based on their Simplified Molecular Input Line Entry Specification (SMILES). To that end, the software calculates a unique abstract molecule description for each molecule using a binary descriptor - substructure fragment fingerprint (FragFp) - to indicate whether or not a certain feature is present in the molecule. By comparing the descriptors of two molecules a similarity value, the

Tanimoto coefficient (Tc), can be calculated. If performed for every compound pair in the library, this results in a similarity matrix for each compound. Each compound is allocated to the most similar neighbour and compounds with a similarity greater than 80% are connected by edges and clustered.

The majority of the chosen compounds cluster into six main structure types: rapaloga, cyclic peptides, alkaloids, macrolactams, lignans and heterotetracycles; in addition to other structures. The degree of structural similarity within one cluster is not only determined by the amount of edges, to which one node is connected, but also by the degree of similarity as outlined by the size of each node: the more similar neighbours are, the bigger the node's size. Especially the rapaloga are known to be structurally very similar towards each other, whereas the structures of cyclic peptides differ more within the cluster. Across the network the structural similarity was assessed by comparing all structures to the reference molecule CsA as indicated by colour. This shows, that the library contains a large amount of structures somehow structurally related to CsA (green), which we expect to screen with comparably good performance using CISCM. However, the library also includes an amount of structurally different molecules (red), when compared to CsA, to further explore the application range of CISCM across different natural compound classes. An additional level of diversity is given by the known drug indication field indicated by the shape of each node, which ranges from antibiotic, immunosuppressive, and anticancer amongst others. The known drug indication field can differ within one cluster, suggesting the possibility of yet unknown modes of action and repurposing potential. Of the 46 structures, 35 structures are FDA approved for at least one prescription field. Of these, 23 are also under current investigation for novel treatments, and eleven structures of the library are not approved yet. The library was further chosen to cover a broad range of molecular masses, spanning from 364 $gmol^{-1}$ to 3080 $gmol^{-1}$ (Supporting Table 1).

This library hence contains a collection of well-studied natural compounds and derivatives covering a broad chemical space and application range. Screening such a library using CISCM could contribute to further understanding of the mechanisms of action of these substances and to study them for potential novel drug indication fields. Using CISCM, such a library could not only be screened against human targets of for example a cancer cell line, but also against bacterial protein extracts or any other protein target pool of interest.

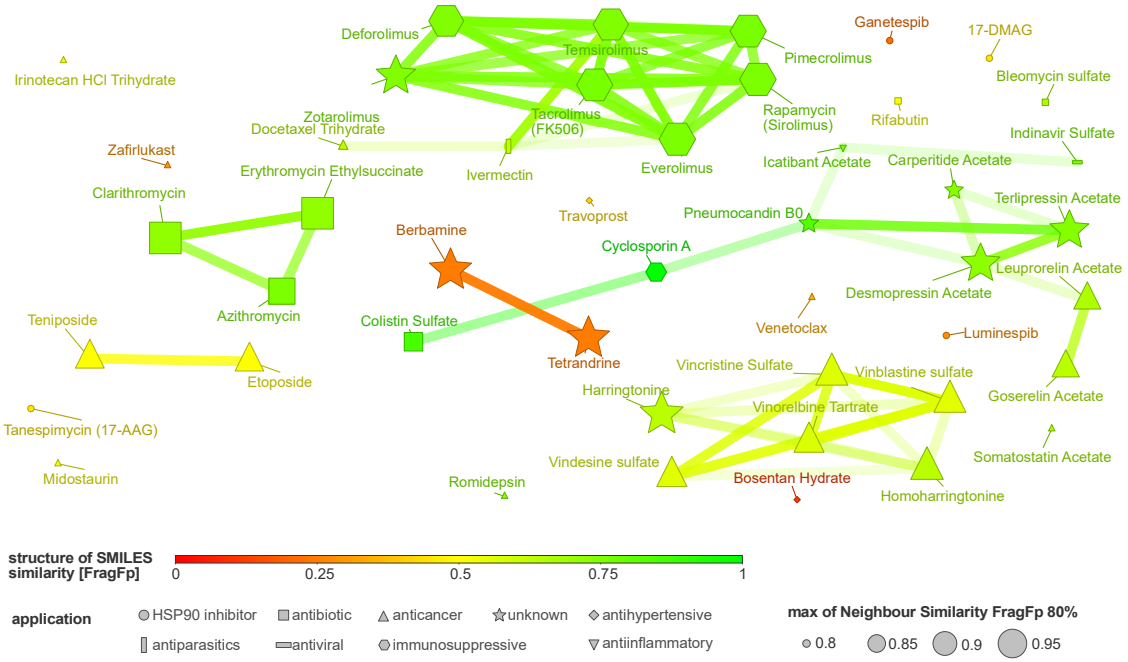


Figure 34. Tanimoto similarity network of a cherry-picked library of 46 compounds. Compounds are displayed as nodes and labeled with their corresponding names. Structurally similar compounds (similarity FragFp > 0.80) are connected by edges and clustered. The degree of similarity amongst direct neighbors within one cluster is indicated by the size of the node (neighbor similarity FragFP 80%), small nodes representing a similarity of 80% and the largest nodes representing a structural similarity of at least 95%. Structural similarity to the reference compound CsA is highlighted by color-grading (Similarity [FragFp]) with high values (green) referring to high structural similarity and low values (red) referring to low structural similarity). The treatment application field of each drug is indicated by the shape of the node.

However, the CISCM method developed here could also be used to screen much bigger libraries than the previously described cherry-picked selection. The technological advancement in both mass spectrometry driven proteomics instrumentation and automated sample processing and pipetting allows for upscaling CISCM in order to screen standard compound libraries containing several hundreds up to big libraries containing several thousands of substances. Modern chromatography systems, such as the Evosep system,[167,168] now allow to use very short gradients and to measure up to 300 mass spectrometry samples per day with high protein identifications. Modern pipetting robot systems such as the Agilent Bravo can furthermore process many samples in parallel for shotgun proteomic sample preparation and clean-up, which also favours the technical upscaling of the CISCM method to screen big compound libraries. A library, as the one described before, could hence be screened in about a week. Considering that NHS-activated cellulose membranes were stable over storage periods of 3 months and also the diazirine linker

could be synthesized in advance, this screening time could even be significantly further reduced. Also on the level of spotting the compounds of interest onto TPD-functionalized cellulose (TFC) membranes the throughput of CISCM could be upscaled. Both the throughput and the reproducibility could be improved by using an automated spotting system. The possibility to spot very small volumes, as it is only feasible with such systems, would allow a smaller compound spot size and hence a higher amount of compound spots per cellulose array, also contributing to the throughput potential.[74]

Other levels of CISCM that could be further optimized are the surface functionalization density and the efficiency of the crosslinking reaction. Instead of using NHS-chemistry based coupling to immobilize the photoreactive TPD-group an approach based on the SPOT synthesis could be applied. This approach has been developed to immobilize synthetic peptides onto cellulose supports via their amine-terminus using Fmoc-chemistry and reaching a high local concentration of peptides.[74,163] To further optimize the photocrosslinking reaction the parameters – such as UV-irradiation time and the stoichiometries of reactants - could be further adjusted, a mix of different photoreactive groups could be used or a variation of CISCM, where the crosslinking is taking place in absence of the cellulose membrane, could be further developed.

Additionally other surface materials could be explored such as silicium, gold wafers or glass slides, as these surfaces are used in numerous microarray applications and established for surface analysis, which would make it easier to monitor and quantify the surface functionalization amongst other advantages. Different materials and immobilizations strategies have been discussed broadly in the scientific literature.[169] Recent advancements in cellulose research provide transparent cellulose, which could also be considered as an attractive novel material, as this would offer the possibility to analyse the surface and its components with methods suitable for transparent materials, such as fluorescence microscopy, whilst maintaining the advantageous properties of cellulose.[169,170] This in turn could allow to quantify both the functionalization density of the TPD functionalized cellulose membrane (TFC) - when equipped with a fluorescent moiety – as well as the photocrosslinking reaction with a fluorescent drug or dye, which could not be performed in this study with non-transparent cellulose. Other suitable surfaces that would facilitate the characterization

and quantification of the functionalized surface would be gold nanoparticles or sepharose beads.[128,171]

4.3. Conclusion

In this thesis, a novel chemoproteomic drug target identification method has been developed. This method uses undirected photocrosslinking on cellulose membranes to create a compound cellulose array and quantitative mass spectrometry to screen this compound cellulose array against full proteomes and to identify corresponding targets.

A key advantage of this compound interaction screen on a photoactivatable cellulose membrane (CISCM) is that it does not require tedious functionalization and previous SAR studies for linker implementation – as it would be required for conventional qAP-MS experiments. This is beneficial especially for natural products, which are often complex and large molecules and for which the introduction of an immobilization handle remains a challenge, limiting their suitability for high-throughput target-ID via qAP-MS. This makes CISCM a powerful tool to study this important drug class, for which the molecular MoA and protein target classes often are yet to be uncovered. In the future, approaches combining parallel photoimmobilization similar to the MagicTag[83,172] or approaches using a mixture of aliphatic and aromatic diazirines,[128] could be used to maximize the bioactivity of immobilized small molecules. Compared to immobilization-free methods like TPP and LiP-SMap that require long LC-MS measurement times, the high throughput of CISCM enables multiplexed analysis of many compounds in parallel.

Current limitations of our method are due to the known limitations of diazirine-based photocrosslinking strategies and affinity purification mass spectrometry as follows. As an example, the approach does not seem to work efficiently for small fragment-like compounds like methotrexate, metformin and the IMiDs thalidomide and lenalidomide only exhibiting a limited number of reactive attachment sites. As CISCM uses lysed cells as potential target pool, it is also so far not possible to identify targets that reach their active binding form only under physiological conditions or metabolism, such as it is suspected for the protein tubulin, which is a target to vinblastine and paclitaxel.

However, given its simplicity, throughput potential and good performance for natural products, CISCM is an attractive method, complementing existing proteomic target

identification approaches in particular in the context of drug leads derived from natural products.[23,173] The application of CISCM to a bigger pool of chemically diverse natural compounds and analogues would give further insights towards its performance range. In the future, this method could be used to screen large collections of natural products and derivatives with yet unknown functional mechanisms, such as those derived from phenotypic screening.

5. Materials and Methods

Some parts of this section were taken from Melder, F. T. I. et al. *ChemMedChem* **2022**, *17.* https://doi.org/10.1002/cmdc.202200346.

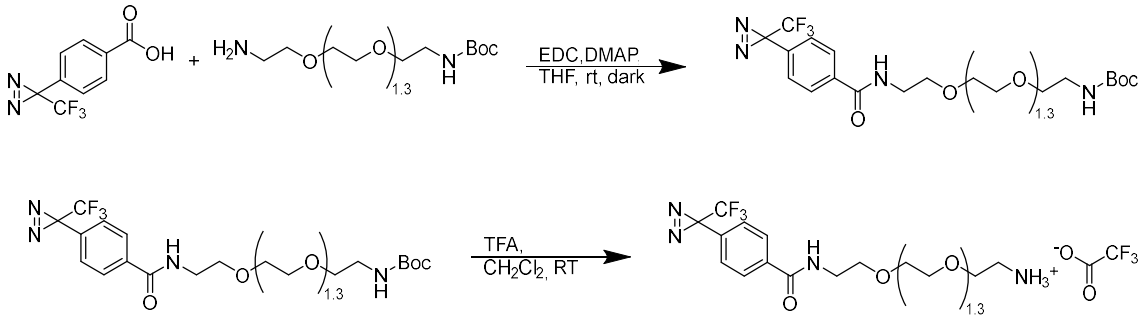
5.1. General information

Commercial chemicals from Sigma Aldrich, Chess, Merck, TCI, Acros and Roche were used as supplied. In all reactions described in chapter 5.2 deionized water was used. For HPLC of synthesis products milli-Q water was used. For LC-MS/MS sample preparation of affinity purification samples (AP-MS) LC-MS grade solvents were used. Nuclear Magnetic Resonance (NMR) mono-(1H, 13C and 19F) were recorded on a Bruker AVANCE 300 and Bruker AV 600 MHz spectrometers. All 13C-NMR-spectra were recorded with 1H-broad-band decoupling. All chemical shifts are reported in ppm relative to tetramethylsilane (δ = 0.00 ppm) and were calibrated with respect to their respective deuterated solvents. Mass spectrometry analyses on synthesized structures (chapter 5.2) were performed with two different spectrometers using the same column and method: Column: Thermo Accuore RP-MS; Particle Size: 2.6 µm; Dimension: 30 × 2.1 mm; Eluent A: Water with 0.1% TFA; Eluent B: Acetonitrile with 0.1% TFA; Gradient: 0.00 min 95% A, 0.2 min 95% A, 1.1 min 1% A, 2.5 min Stoptime, 1.3 min Posttime; Flow rate: 0.8 $mlmin^{-1}$; UV-detection: 220 nm, 254 nm, 300 nm. High resolution mass spectrometry analyses were carried out using the spectrometer Agilent Technologies 6220 Accurate Mass ToF LC/MS linked to Agilent Technologies HPLC 1200 Series, whereas liquid chromatography-mass spectrometry analysis was carried out employing Agilent Technologies 6120 Quadrupole LC/MS linked to Agilent Technologies HPLC 1290 Infinity (chapter 5.2). Thin Layer Chromatography was carried out on TLC-plates from Merck (Silicagel 60, fluorescence-indicator F254, layer thickness = 0.25 mm). For flash chromatography purifications, a Biotage Isolera One apparatus with RediSep®Rf Columns from Teledyne Isco, was used. Preparative HPLC: HPLC analyses were performed on a Shimadzu HPLC system, consisting of a CBM-20A controller, LC-20AP pump, SPD-20 a UV detector and a FRC-10 fraction collector. The separations were performed on a Macherey-Nagel VP250/21 Nucleodor 100-7 C18ec. Compounds were eluted at a flow rate of 30 $mlmin^{-1}$, using water (Solvent A) and acetonitrile (Solvent B) as a mixture of solvents. Magnetic Sp3 SpeedBeads from Cytiva were used (carboxylat-modified particles A

(45152105050250) and B (6515210505025)). Photocrosslinking was performed with a CL-1000L crosslinker (Analytik-Jena, emission at 365 nm, 5 x 8 W bulps) and with the 'user-defined UV time exposure setting' adjusted to 30 minutes. Affinity purification samples were analysed by high resolution LC-MS/MS on a Thermo Fisher Sientific Q Exactive HFX mass spectrometer connected to a Thermo Fisher Scientific EASY-nLC 1200 System using different gradient lengths as indicated in the corresponding sections (buffer A: 3% acetonitrile, 0.1% formic acid in LC-MS grade water; buffer B: 90% acetonitrile, 0.1% formic acid in LC-MS grade water) on a reverse phase column (C18, particle size: 1.9 µm) and data-dependent acquisition (DDA, Top20K). Graphics were created using BioRender (BioRender.com), Affinity Designer (version 2.1.1) and ChemDraw (version 18.0) software.

5.2. Synthesis of photocrosslinkers

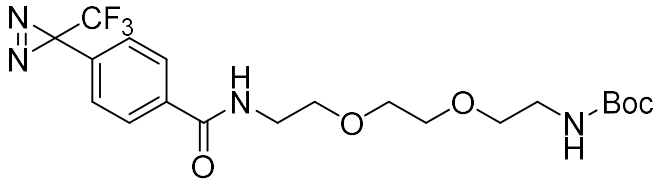
Three different photocrosslinkers were synthesized and analysed by Dr. Peter Lindemann (Nazaré lab, FMP Berlin) using a TBDA-warhead and PEG-spacers of different lengths and a terminal coupling group. Two of these photocrosslinkers employed amine-reactive PEG-spacers of two (amine-PEG_2-linker) or four (amine-PEG_4-linker) PEG-units respectively, and their synthesis followed the general procedures I and II. Another crosslinker was synthesized containing a PEG_4-spacer with a terminal alkyne (synthesis of propargyl-PEG_4-linker).



General procedure I: To a stirred solution of 4-[3-(trifluoromethyl)-3*H*-diazirin-3-yl]benzoic acid (TDBA) (1.00 eq.), *boc*-*N*-amido-PEG-amine (1.25 eq.), and 4-Dimethylaminopyridine (DMAP) (1.217 mmol; 0.35 eq.) in THF was added after 30 min. 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide (EDC) (1.75 eq.) was added and the reaction mixture was stirred for additional 18 hours under protection against light. After successful reaction the solvent was evaporated under reduced pressure.

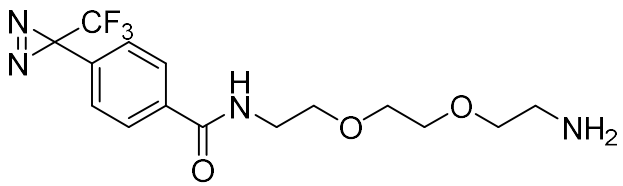
The crude product was purified by silica gel column flash chromatography using dichloromethane and methanol as solvents.

General procedure II: The *boc* protected amine was deprotected by treating the corresponding *boc*-*N*-amido-PEG-TDBA solved in dichloromethane with trifluoro acidic acid at room temperature (RT) for 90 minutes under protection from light. After complete reaction the mixture was quenched with ice water and extracted two times with dichloromethane. The combined organic phases have been dried over magnesia sulfate and the solvent was evaporated under reduced pressure. The crude product was purified with preparative HPLC.



Boc-PEG$_2$-linker: ***tert*-butyl (2-(2-(2-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamido)ethoxy)ethoxy) ethyl)carbamate:**[105] To a solution of 100 mg TDBA (0.434 mmol; 1.00 eq.), and 423 µl *N*-*boc*- 2,2′-(ethylenedioxy)diethylamine (1.78 mmol; 4.1 eq.) as corresponding *boc*-PEG-amine, and 18.6 mg DMAP (0.152 mmol; 0.35 eq.) in 6 ml THF, 146 mg EDC (0.760 mmol; 1.75 eq.) was added after 10 min. Following the general procedure I *tert*-butyl (2-(2-(2-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamido)ethoxy)ethoxy)ethyl)carbamate (138 mg, 0.43 mmol, 69%) was obtained as an colorless oil.

^{1}H NMR (600 MHz, DMSO-d_6) δ in ppm 8.66 (t, J = 5.6 Hz, 1H), 7.98 – 7.92 (m, 2H), 7.37 (d, J = 8.1 Hz, 2H), 6.75 – 6.70 (m, 1H), 3.55 – 3.47 (m, 6H), 3.41 (q, J = 5.8 Hz, 2H), 3.36 (t, J = 6.2 Hz, 2H), 3.04 (q, J = 6.0 Hz, 2H), 1.36 (s, 9H). ^{13}C NMR (151 MHz, DMSO) δ 165.2, 155.5, 149.3, 136.0, 130.1, 128.1, 126.4, 106.7, 77.5, 69.5, 69.4, 69.1, 68.7, 28.2. Calcd .mass for $C_{20}H_{28}F_3N_4O_5$, m/z 461.2006 [M+H]$^+$, found 461.2008.

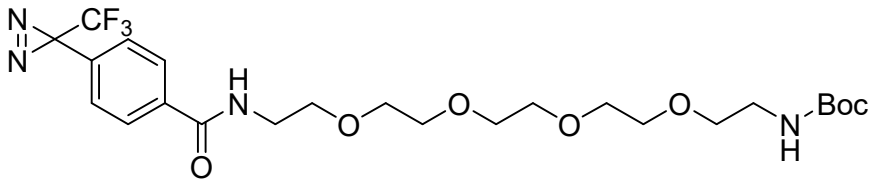


Amine-PEG$_2$-linker: ***N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide:**[105] Starting with 137 mg procedure *tert*-butyl (2-(2-(2-(4-

(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamido)ethoxy)ethoxy)ethyl)carbamate (0.298 mmol; 1.00 eq.) in 5 ml dichloromethane and using 745 µl TFA (9.67 mmol; 32.5 eq.), following general procedure II *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (80.6 mg, 0.224 mmol, 81%) was obtained as TFA salt and colorless oil.

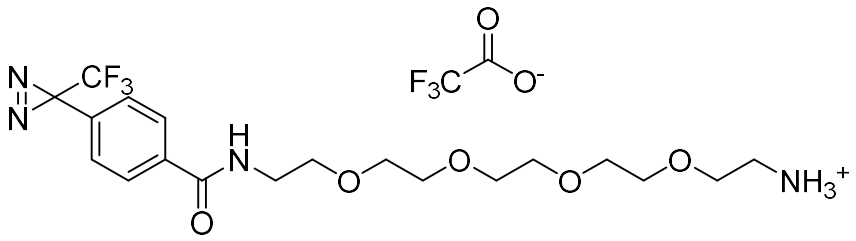
^{1}H NMR (600 MHz, DMSO-d_6) δ 8.68 (t, J = 5.7 Hz, 1H), 7.97 – 7.92 (m, 2H), 7.78 (s, 3H), 7.38 (d, J = 8.1 Hz, 2H), 3.57 (d, J = 4.5 Hz, 6H), 3.54 (t, J = 6.1 Hz, 2H), 3.43 (q, J = 6.0 Hz, 2H), 2.95 (q, J = 5.1 Hz, 2H). ^{13}C NMR (151 MHz, DMSO) δ 165.2, 135.9, 130.2, 128.1, 126.4, 121.74 (q, J = 274.6 Hz), 69.7, 69.4, 68.7, 66.6, 40.1, 38.6, 28.03 (q, J = 39.8 Hz).

Calcd. mass for $C_{15}H_{20}F_3N_4O_3$, m/z 361.1482 [M+H]$^+$, found 361.1484.



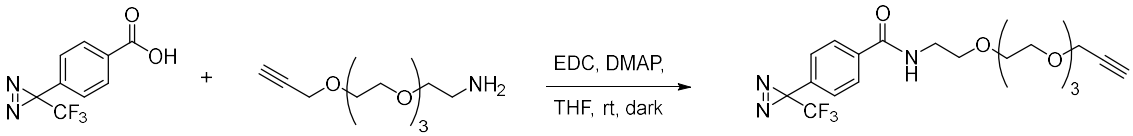
Boc-PEG$_4$-linker: ***tert*-butyl (1-oxo-1-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8,11,14-tetraoxa-2-azahexadecan-16-yl)carbamate:** To a solution of 800 mg TDBA (3.48 mmol; 1.00 eq.), and 1.38 ml *t-boc-N*-amido-PEG$_4$-amine (4,34 mmol; 2,50 eq.) as corresponding *boc*-PEG-amine, 148 mg DMAP (1.217 mmol; 0,35 eq.) in 50 ml THF, 1.17 g EDC (6.08 mmol; 1.75 eq.) was added after 60 min. Following the general procedure I 1.64 g *tert*-butyl (1-oxo-1-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8,11,14-tetraoxa-2-azahexadecan-16-yl)carbamate (2.99 mmol, 86%) was obtained as an colorless oil.

^{1}H NMR (300 MHz, DMSO) δ 8.69 (t, J = 5.5 Hz, 1H), 7.99 – 7.92 (m, 2H), 7.42 – 7.33 (m, 2H), 6.75 (t, J = 5.8 Hz, 1H), 3.58 – 3.29 (m, 18H), 3.04 (q, J = 6.0 Hz, 2H), 1.36 (s, 9H). ^{13}C NMR (75 MHz, DMSO) δ 165.2, 136.0, 130.2, 128.1, 126.4, 77.6, 69.8, 69.7, 69.6, 69.5, 69.2, 68.8, 54.9, 28.2. Calcd. mass for $C_{24}H_{35}F_3N_4O_7$, m/z 549.2531 [M+H]$^+$, found 549.2558.



Amine-PEG$_4$-linker: ***N*-(14-amino-3,6,9,12-tetraoxatetradecyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide:** Starting with 1.62 g *tert*-butyl (1-oxo-1-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8,11,14-tetraoxa-2-azahexadecan-16-yl)carbamate (2.95 mmol; 1.00 eq.) in 10 ml dichloromethane and using 7.39 ml TFA (18.9 mmol; 32.5 eq.), following general procedure II *N*-(14-amino-3,6,9,12-tetraoxatetradecyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (1.51 g, 2.69 mmol, 91%) was obtained as TFA salt and colorless oil.

^{1}H NMR (300 MHz, DMSO) *δ* 8.70 (t, *J* = 5.6 Hz, 1H), 8.00 – 7.90 (m, 2H), 7.78 (s, 3H), 7.43 – 7.34 (m, 2H), 3.62 – 3.46 (m, 15H), 3.42 (q, *J* = 5.6 Hz, 2H). 13C NMR (75 MHz, DMSO) δ 165.2, 157.9, 136.0, 130.2, 128.2, 126.4, 123.6, 69.8, 69.7, 69.7, 69.6, 68.8, 66.7. Calcd. mass for $C_{19}H_{28}F_3N_4O_5$, m/z 449.2006 [*M*+H]$^+$, found 449.2013.



Synthesis of propargyl-PEG$_4$-linker:[105] ***N*-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide:** To a solution of 25 mg TDBA (0.11 mmol; 1.00 eq.), and 52.8 mg 2-(2-(prop-2-yn-1-yloxy)ethoxy)ethan-1-amine (0.23 mmol; 2.10 eq.) as corresponding propargyl-PEG$_4$-amine, and 4.6 mg DMAP (0.04 mmol; 0.35 eq.) in 1.5 ml THF, 36.4 mg EDC (0.19 mmol; 1.75 eq.) was added and the reaction mixture was stirred for additional 72 hours under protection against light. After successful reaction the solvent was evaporated under reduced pressure. The crude product was purified with preparative HPLC and *N*-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (28.0 mg, 0.06 mmol, 58%) was obtained as an colorless oil.

^{1}H NMR (300 MHz, DMSO-d6) *δ* 8.69 (t, *J* = 5.7 Hz, 1H), 8.01 – 7.90 (m, 2H), 7.42-7.32 (m, 2H), 4.12 (d, *J* = 2.4 Hz, 2H), 3.55-3.50 (m, 10H), 3.48 (q, *J* = 1.5 Hz, 4H), 3.45 – 3.38 (m, 3H). ^{13}C NMR (75 MHz, DMSO) *δ* 165.2, 135.9, 130.2, 128.1, 126.4,

126.4, 80.3, 77.1, 69.8, 69.7, 69.6, 69.5, 68.8, 68.5, 57.5. m/z 444.4 $[M+H]^+$, found 444.1.

5.3. Preparation of NHS-activated cellulose (NAC) membranes

The main CISCM approach employed NHS-activation of cellulose membranes and immobilization of amine-PEG-linkers using NHS-based chemistry to obtain TPD-functionalized cellulose membranes (TFC membranes). For all procedures cellulose filter membranes from Whatman (type 50, 55 mm) were used.

For activation six cellulose membranes were incubated in a solution of sodium hydroxide in water (w(NaOH) = 10%) for 18 hours. Afterwards the membranes were rinsed with ethanol and stored under ethanol until further use for up to three days. The activated cellulose membranes were taken out of the ethanol and dried at 80 °C in an air stream. In a shortened version of this procedure, the membranes were rinsed in ethanol and then washed three times in ethanol (10 minutes) and dried with a hairdryer for 5 minutes. With oxidation and NHS-activation of the cellulose membranes performed on the same day, this enabled to obtain NAC membranes in 2 days.

The membranes were then oxidized in a 2,2,6,6-Tetramethylpiperidinyloxyl (TEMPO)-mediated method with sodium hypochlorite.[122] To achieve this, a fresh solution of 60.9 mg TEMPO (0.39 mmol; 0.2 eq.), 1.45 g NaBr (14.1 mmol; 0.83 eq.) in 292 ml water and 8.48 ml of $NaOCl_{aq}$ (17.0 mmol; c(NaOCl) = 2,00 M; $w(Cl_2) \approx 12\%$) was prepared. The solution was set to a pH-value of 10 with HCl (c = 1 M). To prevent overlapping of the membranes, the oxidation solution was distributed over 6 separate reaction vessels, containing one membrane each, and the pH-value re-adjusted to pH 10 if necessary. The oxidation was quenched after 60 minutes by rinsing the membranes with ethanol, continued with washing with water.

For NHS-activation a solution of 3.45 g NHS (n = 30 mmol; c = 0.10 mM and 23.0 g EDCl (c = 0.4 M) in 300 ml sodium acetate buffer (pH 5) was prepared.[122] The oxidized membranes were incubated for one and a half hours in the NHS-solution in separate reaction vessels. Afterwards the membranes were taken out and each one was washed with 400 ml of water. The NHS-activated cellulose membranes (NAC membranes) were dried between two paper filters for 18 hours.

5.4. Preparation of cell lysates and lysate proteome analysis

Both, Jurkat cells and HEK293T cells, were grown in label-free cell culture medium (RPMI-1640, 10% Fetal Bovine Serum and lysed with ice-cold lysis buffer (50 mM HEPES at pH 7.6, 150 mM NaCl, 1 mM EGTA, 1 mM $MgCl_2$, 0.5% Nonident P-40, 0.05% SDS, 0.25% Sodium-deoxycholate) subjected with protease inhibitor (cOmplete™, Mini, Roche) and benzonase® endonuclease (Emprove® Expert). Protein concentration of the lysate was measured using a detergent compatible protein assay (DC protein Assay, Bio-Rad) and adjusted to 4 $mgml^{-1}$. HEK293T cell culture was performed by Martha Hergeselle (Selbach lab, MDC Berlin).

To prepare samples for proteomic measurements, proteins in both lysates (Jurkat, HEK293T) were precipitated using methanol-chloroform-water (Wessel Flügge) precipitation,[174] resuspended in LC-MS grade denaturation buffer (6 M urea, 2 M thiourea, 10 mM HEPES, pH 8) and purified following the StageTip-procedure described in chapter 5.8 (reduction, alkylation, digestion, purification). The resulting peptide solutions of the StageTip-procedure were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer connected to a nLC1200 system using a long gradient and data-dependent acquisition (240 min, DDA: Top20, MS2-resolution: 7.5K, column: 1.9 µm).

The acquired spectra were analysed in MaxQuant[45] (MQ version 1.6.3.3) using a protein and peptide FDR of 1%, label-free-quantification (LFQ), match-between-runs, re-quantify and MQ standard parameters. The *in silico* digest was performed with Trypsin/P on the human Uniprot database (2018-04). Further data analysis was performed in R version 4.1.1. Reverse hits, potential contaminants and only identified by side were filtered out. The intensity based absolute quantification (iBAQ) data was log_{10}-transformed and missing values removed. The data of both lysate proteomes was then ranked according to their log_{10}(iBAQ) values and displayed as a rank plot.

5.5. Generation of photocrosslinked small molecule microarrays and affinity enrichments of interacting proteins

The samples corresponding to the data shown in the chapters 3.3.1, 3.3.2, 3.4.1, 3.4.2, 3.2.2 and 3.2.3 were prepared according to the following procedure with individual experimental parameters as stated in Table 2.

For the immobilization of the photocrosslinker triplicates of NAC and control membranes (see chapter 5.7), were incubated in a solution of an amine-PEG-linker (PEG-linker immobilization, in THF) at room temperature in the dark and under slight shaking of the solution. The resulting TPD-functionalized cellulose (TFC) and control membranes were then rinsed in THF, washed in water for 15 minutes and blocked in a solution of ethanolamine (blocking (EA)) in the dark. Then the blocked cellulose membranes were rinsed in water and dried.

Dried blocked TFC and control membranes were spotted with compound solutions (n(compounds)). For the data presented in chapter 3.3.1 CsA was dissolved in methanol and for all following CISCM screens compound solutions were prepared in DMSO. To control the spot size, volumes smaller or equals to 0.5 µl were used in combination with gel loader pipette tips. After evaporation of the solvent, the membranes were irradiated with UV light (365 nm, t(UV-irradiation)). Membranes were rinsed in ethanol, washed intensively in organic solvents (EtOH, DMF, THF, EtOH, water) for one hour each and dried. Samples of data shown in chapter 3.4.2 were additionally washed in methanol overnight.

Dried cellulose membranes containing the arrayed small molecules were then conditioned for 15 minutes in a LC-MS-grade wash buffer (50 mM HEPES, 150 mM NaCl, 1 mM EGTA at pH 8, 1 mM $MgCl_2$) at pH 7.6 and room temperature. Each membrane was then incubated with a cell lysate at 4 °C for two hours followed by washing with LC-MS-grade wash buffer three times for 5 minutes also at 4 °C. Membranes were dried, processed for shotgun proteomic analysis using standard methods and analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer as described in chapter 5.8 and analysed as described in chapter 5.9.

Table 2. Individual experimental details of NHS-based CISCM screens. In some screens, cellulose membranes were blocked with EA twice for one hour each (2*1 h). In other screens, such as the screen in chapter 3.3.2, membranes were irradiated from both sides, front side (fs) and backside (bs), for 30 minutes each side. Sample preparation data of chapters 3.2.2 and 3.2.3 were included for reasons of clarity, however a protocol is described elsewhere (chapter 5.10).

Chapter	PEG-linker immobilization	Blocking	n(cpds)	t(UV)	Cell lysate
3.3.1	Amine-PEG_2-linker, 1 mM, 20 h	1 M, 1 h	140 nmol	30 min	Jurkat
3.3.2	Amine-PEG_2-linker, 10 mM, 21 h	1 M, 1 h	50 nmol	30 min (fs), 30 min (bs)	Jurkat
3.4.1	Amine-PEG_4-linker, 10 mM, 24 h	1 M, 1 h	50 nmol	30 min (fs), 30 min (bs)	HEK293T
3.4.2	Amine-PEG_4-linker, 10 mM, 24 h	3 M, pH 9, 2*1 h	50 nmol	30 min (fs), 30 min (bs)	HEK293T
3.2.2, 3.2.3	Amine-PEG_4-linker, 10 mM, 24 h	3 M, pH 9, 2*1 h	50 nmol	30 min (fs), 30 min (bs)	None

5.6. Inkjet printing of natural compounds

Six pieces of cellulose membranes (Whatman type 50, 55 mm) were incubated in a solution of sodium hydroxide in water (w(NaOH) = 10%) overnight. Afterwards the membranes were washed three times with ethanol for 10 minutes each and dried with a hairdryer. Three of the activated cellulose membrane pieces were spotted with solutions of the compounds CsA, tacrolimus and sirolimus (10 mM, 25nmol) using gel loader pipette tips. As a visual control the remaining three cellulose membrane pieces were spotted with solutions of ponceau S and 6-aminofluorescein (10 mM, 25nmol). All membrane pieces were dried intensively with a hairdryer and conditioned for 5 minutes in a LC-MS-grade wash buffer (50 mM HEPES, 150 mM NaCl, 1 mM EGTA at pH 8, 1 mM $MgCl_2$) at pH 7.6 and room temperature. Each membrane was then incubated with a Jurkat cell lysate at 4 °C for 30 minutes followed by washing with LC-MS-grade wash buffer three times for 5 minutes also at 4 °C. Fluorescence of the

fluorescein spotted membranes was measured with a ChemiDoc (Bio-Rad) detector ('Fluorescein' setting) after spotting and after each incubation step. The compound-spotted cellulose membranes were dried, processed for shotgun proteomic analysis using standard methods (StageTip-procedure) and the resulting 9 samples were analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer (details see chapters 5.8). The resulting data was analysed as described in chapter 5.9. Additionally log_2-transformed intensity-based absolute quantification[153] (iBAQ) values of enriched proteins were displayed in a profile plot (Perseus version 1.6.7.0).

5.7. Implemented controls

For both, the photocrosslinking as well as the affinity purification, we implemented several controls. In the screens presented in the chapters 3.3.2 and 3.4.1 triplicates of blocked TFC membranes were produced and spotted with the compounds of interest as described in chapter 5.5. Each replicate of the compound-spotted TFC membranes was additionally spotted with fluorescein (50 nmol, DMSO) prior to UV-treatment using gel loader pipette tips and small volumes of below 0.5 µl. After UV-treatment of the membranes on their front- and backside (30 min each side), an empty area of these then photoinactive TFC membranes was spotted again with the fluorescein solution (50 nmol, DMSO) and dried (Figure 28a). The membranes were then washed thoroughly in different organic solvents for one hour each, as described earlier (chapter 5.5). The fluorescence of the controls was measured before washing and after each washing step with a ChemiDoc (Bio-Rad) detector in the 'Fluorescein' setting. These fluorescein controls were not analysed by high resolution LC-MS/MS.

In another CISCM screen we controlled the immobilization of compounds and subsequent target-ID in dependence of the presence of a photocrosslinker and of UV light (chapter 3.4.2). In addition to the prepared samples (details see chapter 5.5), six replicates of NAC membranes and triplicates of TFC membranes were prepared and spotted with a solution of CsA (50 nmol, DMSO). As fluorescent control, all membranes, including the sample membrane described earlier (chapter 5.5), were additionally spotted with a solution of fluorescein (50 nmol, DMSO). Triplicates of spotted NAC membranes were UV-irradiated from both sides (30 minutes each), whereas the remaining spotted NAC and TFC triplicates were not exposed to UV-irradiation (Figure 30). UV-irradiated and not irradiated membranes were then washed

as previously described (chapter 5.5) and additionally incubated in methanol overnight. The fluorescence of the controls was measured together with corresponding samples in a ChemiDoc (Bio-Rad) detector ('Fluorescein' setting) after the UV-irradiation and after each washing step. Resulting control samples were then treated analogously to the corresponding samples as described earlier (chapter 5.5, Table 2), including identical sample preparation and data acquisition (chapter 5.8) as well as data analysis (5.9). For label-free quantitative analysis all control samples were used, including fluorescein control samples to increase the background size.

We aimed to additionally control protein binding towards functionalized and non-functionalized cellulose membranes (chapters 3.3.1, 3.3.2, 3.4.1). To this end we included triplicates of cellulose membranes at different stages of functionalization (Table 3) to the CISCM screens. To create blocked NAC control membranes, NHS-activated cellulose (NAC) membranes were incubated in THF, instead of a solution of amine-PEG_4-linker, and blocked with ethanolamine (1 M, 1 h). None of these control membranes were spotted with compounds, but apart from this treated analogously to the samples as described earlier (chapter 5.5, Table 2), including identical sample preparation and data acquisition (chapter 5.8) as well as data analysis (5.9).

As background for the quantitative analysis of data presented in chapter 3.3.1, blocked and CsA-spotted triplicates of TFC membranes (chapter 5.5) were additionally spotted seven times with 0.2 µl of solutions of the dyes bromphenol blue (0.04%, EtOH), coomassie brilliant blue (0.01%, 10% phosphoric acid, 5% ethanol), ponceau S (10 mM, DMSO) and 6-aminofluorescein (10 mM, DMSO). Membranes were dried, irradiated with UV light (30 min) and further processed and analysed as described for CsA in the chapters 5.5, 5.8 and 5.9.

Table 3. Controls used in different CISCM screens presented in indicated chapters.

Chapter	dye	Fluorescence assay	AP-MS control membranes
3.3.1	Bromphenol blue, coomassie brilliant blue, ponceau S, 6-aminoflourescein	/	Non-functionalized cellulose, blocked TFC, dye-spotted blocked TFC
3.3.2	Fluorescein	Fluorescein-spotting on blocked TFC membranes before and after UV-irradiation	Non-functionalized cellulose, oxidized cellulose, blocked TFC
3.4.1	fluorescein	Fluorescein-spotting on blocked TFC membranes before and after UV-irradiation	Non-functionalized cellulose, oxidized cellulose, blocked NAC
3.4.2	fluorescein	Fluorescein-spotting on blocked NAC and blocked TFC membranes with and without UV-irradiation	Blocked NAC (compound-spotted), blocked TFC (compound-spotted, not UV-treated), blocked NAC and TFC (fluorescein-spotted)

5.8. LC-MS/MS sample preparation and data acquisition of affinity enrichments

Samples of data shown in chapters 3.3.1, 3.3.2, 3.4.1, 3.4.3 and 3.4.4 were digested and desalted using StageTip[125] purification (StageTip-procedure) and samples of the data shown in the chapter 3.4.2 were digested and purified using sp3 magnetic beads[152] (sp3-procedure) as described below.

StageTip-procedure: Cellulose spots corresponding to individual photocrosslinked compounds were excised using a paper puncher (d = 4 mm) and transferred into 96-

well plates containing LC-MS grade denaturation buffer (6 M urea, 2 M thiourea, 10 mM HEPES, pH 8). Each sample was then treated with dithiothreitol (10 mM, 50 mM ammoniumbicarbonate (ABC), LC-MS grade) for 30 minutes and afterwards with iodacetamide (55 mM, 50 mM ABC, LC-MS grade) for 20 minutes in the dark. Samples were then predigested with LysC (0.5 $\mu g \mu l^{-1}$, LC-MS grade) for one and a half hours, diluted four times with ABC (50 mM, LC-MS grade) and digested with trypsin (0.5 $\mu g \mu l^{-1}$, LC-MS grade) overnight. Digestion was stopped by reducing the pH (pH < 2.5) using a trifluoroacetic acid solution (10%, LC-MS grade water). Samples were then desalted using StageTip[125] purification (C18, reverse phase, Empore) and stored at 4 °C until further use.

Samples were eluted from StageTips using an elution buffer (0.1% formic acid, 50% acetonitrile in water, LC-MS grade), the solvent was evaporated and samples were resuspended in LC-buffer A (3% acetonitrile, 0.1% formic acid (FA) in water (LC-MS grade). The samples of the CISCM screen presented in chapter 3.4.1 were additionally purified by strong cation exchange chromatography (SCX)[175] and subsequent repetition of desalting by StageTip purification,[125] elution and resuspension in LC-buffer A.

Sp3-procedure: Cellulose spots corresponding to individual photocrosslinked compounds were excised using a paper puncher (d = 4 mm), transferred into 96-well plates and incubated in LC-MS grade denaturation buffer (6 M urea, 2 M thiourea, 10 mM HEPES, pH 8, 50 µl). After one hour the supernatants were transferred into another 96-well plate and stored on ice. Cellulose spots were incubated in another 50 µl LC-MS grade denaturation buffer for an additional hour. The following procedure was performed together with Mohamed Haji (Mertins lab, Proteomics Core Facility, MDC Berlin). Both supernatants were united and treated with dithiothreitol (10 µl, 100 mM, 50 mM ammoniumbicarbonate (ABC), LC-MS grade) at 37 °C for 30 minutes. The 96-well plate was then transferred to an automated liquid handling system (Agilent Bravo) for further sample processing. Iodacetamide (10 µl, 200 mM, 50 mM ABC, LC-MS grade) was pipetted to each well by the system and the plate was shaken for 30 minutes in the dark. Another 10 µl of dithiothreitol (100 mM, 50 mM ammoniumbicarbonate (ABC), LC-MS grade) were added to each well followed by the addition of magnetic sp3 beads (10 µl, 100 $\mu g \mu l^{-1}$, Sera-Mag SpeedBeads A and B, 1:1). Acetonitrile (200 µl, 70%) was added twice to reach a total volume of 540 µl in

each well. The beads were then washed three times with ethanol (200 µl each) and LC-MS grade ABC buffer (100 µl) was added together with trypsin (6 µl, 0.5 $mgml^{-1}$, LC-MS grade) and incubated overnight. The beads were then washed twice in LC-MS grade ABC buffer (100 µl) and the supernatants, containing the peptide sample, were united.

The resulting peptide solutions of StageTip-procedure, or the sp3-procedure respectively, were then analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer connected to a nLC1200 system using a short gradient and data-dependent acquisition (45 min, DDA: Top20, MS2-resolution: 15K, column: 1.9 µm).

5.9. Data analysis

The acquired raw files were analysed in MaxQuant[45] (MQ version 1.6.3.3) using a protein and peptide FDR of 1%, label-free-quantification (LFQ), match-between-runs, re-quantify and MQ standard parameters. For the in silico digest was performed with Trypsin/P on the human Uniprot database (2018-04). Reverse hits, potential contaminants and only identified by side were filtered out. The LFQ data was log_2-transformed and replicates grouped together. The data was then filtered on valid values (minimum 3 in at least one group) and missing values were imputed from normal distribution (width: 0.3, down shift: 1.8). The affinity enrichments of oxidized and NHS-activated cellulose membranes of data presented in chapter 3.4.1 were removed from further analysis, due to poor data quality.

Multiple sample testing was performed using LFQ intensities (ANOVA, permutation-based FDR: 5%, 250 randomizations) and ANOVA significant hits were Z-scored and clustered hierarchically. Protein abundances in the replicates for a given compound were compared to all other samples using the Student's t-test. Identified proteins within a defined specificity cut-off (Table 4) were identified as specific binders. Data filtering and statistical integration (ANOVA, Student's t-test) was performed in Perseus version 1.6.7.0. Hierarchical clustering was performed in R version 4.1.1. The mass spectrometry proteomics data presented in Figure 28 (chapter 3.3.2) for the compounds CsA, tacrolimus, sirolimus, lenalidomide, and controls (TFC-blocked, oxidized, cellulose) have been deposited to the ProteomeXchange Consortium via the PRIDE[2] partner repository with the dataset identifier PXD033050. For the here

presented data one replicate of methotrexate (MTX) needed to be removed from the analysis (Figure 28, chapter 3.3.2).

Table 4. Specificity cut-offs for different CISCM screens presented in the indicated chapters.

Chapter	p-value	Log_2(FC)
3.3.1, 3.3.2, 3.4.1, 3.4.3	<0.01	2
3.4.2	<0.1	1.5

5.10. Sample preparation and analysis of click-CISCM

In a variation of the CISCM approach three selected compounds were photocrosslinked with a propargyl-PEG_4-linker and resulting constructs covalently attached to cellulose membranes by azide-alkyne Huisgen cycloaddition ('click chemistry').

To achieve this a cellulose membrane (Whatman type 50, 55 mm) was incubated a solution of sodium hydroxide in water (w(NaOH) = 10%) for 24 hours and at room temperature. The membrane was then washed six times with ethanol and stored on ethanol until further use. Next, the activated cellulose membrane was tosylated and azide-functionalized.[121] To achieve this the membrane was shaken in 10 ml of a *p*-toluenesulfonyl chloride solution (464 mg, 2.44 mmol, pyridine) for 20 h at 40 °C. The tosylated cellulose was immersed in 20 ml DMF and sonicated three times. The membranes were stored in DMF until further use. Next, half of a tosylated cellulose membrane was shaken in 10 ml of a solution of sodium azide (NaN_3, 526 mg, 8.10 mmol) for 40 hours and at 60 °C. The azide-cellulose membrane was sonicated in 20 ml of water, acetone, ethanol and dichloromethane and dried.

A solution of propargyl-PEG_4-linker (6 eq., 29 mM, 'alkyne-TPD') was added to solutions of CsA (12 eq., 58 mM, EtOH), S-thalidomide (12 eq., 58 mM, DMF) or R-thalidomide (12 eq., 58 mM, DMF) in a glass-coated 96-well plate in triplicates, mixed thoroughly and dried in vacuo for one hour at 40 °C. Additionally, a solution of fluorescein (12 eq., acetone) was mixed in triplicates with propargyl-PEG_4-linker as a control sample and treated analogously. Samples and controls were UV-irradiated (30 minutes) and resulting photocrosslinking products were resuspended in a solution of DMSO/water (3/1), $CuSO_4$ (mol-10%) and sodium ascorbate (0.78 mmol).[121] To

control the click chemistry reaction, solutions of 6-FAM-alkyne (1 eq., 4.9 mM) and fluorescein (6 eq., 29.1 mM) in DMSO/water (3/1), $CuSO_4$ (mol-10%) and sodium ascorbate (0.78 µmol) were included in triplicates. Azide-cellulose spots were excised (d = 1 mm, 1 eq.) and added to the solutions for 24 hours at 60 °C while shaking. As an additional control, triplicates of azide-activated cellulose (AAC) spots were immersed in DMSO/water (3/1) for 24 hours at 60 °C. All cellulose spots were then washed twice in ethylenediaminetetraacetic acid (1 M), acetone and dichloromethane for 3 minutes.

The control cellulose spots, which were reacted with photocrosslinking products of fluorescein, with 6-FAM-alkyne or with fluorescein, were additionally washed in DMSO for one hour and then washed in dichloromethane for 10 minutes. Their fluorescence was measured with a ChemiDoc (Bio-Rad) detector ('Fluorescein' setting), using all other samples and the AAC control as a background.

Sample cellulose spots, corresponding to the immobilized compounds CsA, S-thalidomide and R-thalidomide, and the AAC control spots were conditioned for 10 minutes in a LC-MS-grade wash buffer (50 mM HEPES, 150 mM NaCl, 1 mM EGTA at pH 8, 1 mM $MgCl_2$) at pH 7.6 and room temperature. Each cellulose spot was then incubated with a HEK293T cell lysate at 4 °C for one hour followed by washing with LC-MS-grade wash buffer twice for 5 minutes also at 4 °C. Membrane spots were dried, processed for shotgun proteomic analysis using standard methods and analysed by high resolution LC-MS/MS on a Q Exactive HFX mass spectrometer as described in chapter 5.8 (StageTip-procedure).

The acquired raw files were analysed in MaxQuant[45] (MQ version 1.6.3.3) using a protein and peptide FDR of 1%, label-free-quantification (LFQ), match-between-runs, re-quantify and MQ standard parameters. For the *in silico* digest was performed with Trypsin/P on the human Uniprot database (2018-04). Reverse hits, potential contaminants and only identified by side were filtered out. The LFQ data was log_2-transformed and replicates grouped together. The data was then filtered on valid values (minimum 3 in at least one group) and missing values were imputed from normal distribution (width: 0.3, down shift: 1.8). Due to poor data quality one replicate of the CsA affinity enrichments was removed from the analysis. Protein abundances in the replicates for a given compound were compared to all other samples using the Student's t-test. For the analysis of (R)-thalidomide affinity enrichments the data of the

(S)-enantiomer was removed from the background and vice versa. Identified proteins with Student's t-test p-values < 0.01 and fold changes of at least 2 were identified as specific binders. Data filtering and statistical integration (Student's t-test) was performed in Perseus version 1.6.7.0.

5.11. Surface chemistry data acquisition and analysis

To evaluate the functionalization of cellulose membranes XPS (Figure 19) and ATR-FTIR (Figure 20) spectra were acquired after each functionalization step (Figure 18, 1, 2, 3, 4, 5). NHS-activation of the cellulose membranes was additionally controlled by using 6-aminofluorescein and horseradish peroxidase (HRP), respectively. To control the photocrosslinking on the TPD-functionalized cellulose (TFC) membranes, TOF-SIMS spectra (Figure 23, Figure 24) were acquired after photocrosslinking the compound cyclosporine A.

For preparation of TFC membranes (Figure 18, 4), NAC membranes (Figure 18, 3) were produced as described earlier and incubated in a solution of a photoactivatable amine-PEG_4-linker (Figure 18, 3', 10 mM, THF) for 24 hours in the dark. The membranes were then rinsed in THF, washed in THF (15 min, dark) and washed in water (15 min, dark). Subsequently the membranes were blocked in an ethanolamine solution (EA, 3 M, pH 9.0) twice for an hour and then rinsed and washed twice in water (5 min). The blocked TFC membranes were dried overnight at room temperature in the dark and shipped in falcon tubes.

XPS measurements were performed using a K-Alpha+ XPS spectrometer (ThermoFisher Scientific, East Grinstead, UK). The Thermo Avantage software was used for data acquisition and processing. All samples were analysed using a microfocused, monochromated Al Kα X-ray source (400 µm spot size). The K-Alpha+ charge compensation system was employed during analysis, using electrons of 8 eV energy, and low-energy argon ions to prevent any localized charge build-up. The spectra were fitted with one or more Voigt profiles (BE uncertainty: ± 0.2 eV) and Scofield sensitivity factors were applied for quantification.[176] All spectra were referenced to the C 1s peak (C-C, C-H) at 285.0 eV binding energy controlled by means of the well-known photoelectron peaks of metallic Cu, Ag, and Au, respectively. XPS measurements and data analysis were performed by Vanessa Trouillet (IAM-ESS and KNMFi, KIT).

ATR-FTIR spectra were recorded on a Bruker Tensor 27, equipped with a platinum ATR-Unit with diamond crystal. The spectra were recorded on a RT-DLaTGS detector with 32 scans and measured against air as background. To follow the functionalization of the cellulose ATR-measurements were performed for each single step. ATR-FTIR measurements were performed by Stefan Heißler (KNMFi, KIT) and analysed together with Peter Lindemann.

For the assessment of the photoimmobilization of small molecules ToF-SIMS measurements were performed using cyclosporine A (CsA) as test compound. To that end NAC membranes (Figure 18, 3) and TFC membranes (Figure 18, 4) were spotted in duplicates 10 times with 0.5 µl of a CsA solution (10 mM, DMSO) and a fluorescein control solution (10 mM, DMSO) and dried overnight. One replicate of each was then irradiated with UV light at 365 nm from both sides for 30 minutes each, whereas the second replicate was kept in the dark. All membranes were then rinsed in EtOH, washed in different organic solvents (EtOH, DMF, THF, EtOH, water) for one hour each and incubated in MeOH overnight in the dark. The membranes were dried and stored at room temperature until further use. The membranes were shipped in falcon tubes protected from light.

To control NHS-activation of cellulose membranes in a fluorescent assay, triplicates of NAC membrane pieces were conditioned in ethanol (EtOH) for 5 minutes and incubated in a solution of 6-aminofluorescein (10 mM, EtOH) for 30 minutes followed by washing three times for 5 minutes in ethanol. As controls non-functionalized cellulose (unmod.), non-functionalized cellulose preconditioned in sodiumhydroxide (NaOH-act.) and oxidized cellulose were treated equally. Fluorescence signals were acquired in a ChemiDoc (Bio-Rad) detector in the ‘Fluorescein’ setting.

In a chemiluminescence assay, anti-mouse horseradish peroxidase (HRP) antibody (anti-mouse IgG, enhanced chemiluminescence ECL™, GE Healthcare UK) was dialyzed, resuspended in phosphate buffered saline (1:2000) and incubated with triplicates of NAC membrane pieces for one hour at room temperature. As controls non-functionalized cellulose (unmod.), non-functionalized cellulose preconditioned in sodiumhydroxide (NaOH-act.) and oxidized cellulose were treated equally. Membranes were rinsed in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 1 M, pH 7.6) and incubated in HEPES buffer overnight at 4 °C. The membranes were then treated with Western Lightening® Plus-ECL (1:1, Perkin Elmer)

substrate. Chemiluminescence was measured in a ChemiDoc (Bio-Rad) detector in the 'Chemiluminescence' setting. Due to strong background signal the membranes were further washed in Tris-buffered saline (T-TBS, 20 mM Tris, 150 mM NaCl, 0.1% Tween 20) and washed in buffer (50 mM HEPES at pH 7.6, 150 mM NaCl, 1 mM EGTA, 1 mM $MgCl_2$, 0.5% Nonident P-40, 0.05% SDS, 0.25% Sodium-deoxycholate) for 36 hours. The chemiluminescence imaging was repeated as described before.

ToF-SIMS was performed on a TOF.SIMS5 instrument (ION-TOF GmbH, Münster, Germany) equipped with a Bi cluster primary ion source and a reflectron type time-of-flight analyser. Some samples were slightly outgassing, hence UHV base pressure during analysis was < 2×10^{-7} mbar. For high mass resolution the Bi source was operated in bunched mode providing short Bi_3^+ primary ion pulses at 25 keV energy, a lateral resolution of approx. 4 µm, a target current of 0.35 pA at 10 kHz repetition rate and 1.1 ns pulse length. For each sample three spots of 500×500 μm^2 were analysed, scanning 128×128 pixel with 75 scans (100 or 125 µs cycle time). Thereby the primary ion dose density was below the quasi static limit (2×10^{11} ionscm^{-2}). Charge compensation was performed by applying a 21 eV electron flood gun and tuning the reflectron accordingly. Peak broadening due to the roughness of the samples and slightly uneven charging was observed. Mass scale calibration was based on hydrocarbon signals. For pos. polarity spectra C^+, CH^+, CH_2^+, and $C_2H_3^+$ were used, for neg. polarity C^-, CH^-, CH_2^-, C_2^-, and C_3^-. Original data files, spectra and meta data are available on Radar4KIT. TOF-SIMS measurements and data analysis were performed by Dr. Alexander Welle (KNMFi, KIT).

5.12. Compound selection and related analysis

The molecular weight of seven selected compounds was compared to the molecular weights of all structures listed on CHEMBL[141] (access on 23.09.2020, https://www.ebi.ac.uk/chembl/g/#search_results/all), excluding antibodies, enzymes, cells and structures without molecular weight annotation. If the listed compound consisted on several monomeric structures the sum of the corresponding substructures was taken as molecular weight. A histogram (bin width: 50 $gmol^{-1}$) of the data was constructed in R version 4.1.1. A principle component analysis (PCA) was performed with the software DataWarrior[177] on the chemical structure of the

compounds. XlogP[178] values were calculated by pubchem XlogP 3.0 and a graph of the combined data was constructed in R version 4.1.1.

To construct a natural compound analogue screening library, 43 substances listed in ChEMBL were cherry-picked and the compounds CsA, tacrolimus, sirolimus were additionally included as controls. Information to these compounds is provided in Supporting Table 1 and Supporting Table 2. The compounds were then analysed in DataWarrior[177] as follows. Substructure fragment fingerprint (FragFp) descriptors were determined for each structure by using the corresponding Simplified Molecular Input Line Entry Specification (SMILES) and clustered based on the structural similarity and displayed as a Tanimoto similarity network. Known targets or drug indication fields were annotated and included in the network.

6. References

(1) Cui, M.; Cheng, C.; Zhang, L. High-Throughput Proteomics: A Methodological Mini-Review. *Lab. Investig.* **2022**, *102* (11), 1170–1181. https://doi.org/10.1038/s41374-022-00830-7.

(2) Ha, J.; Park, H.; Park, J.; Park, S. B. Recent Advances in Identifying Protein Targets in Drug Discovery. *Cell Chem. Biol.* **2021**, *28* (3), 394–423. https://doi.org/https://doi.org/10.1016/j.chembiol.2020.12.001.

(3) Ziegler, S.; Pries, V.; Hedberg, C.; Waldmann, H. Target Identification for Small Bioactive Molecules: Finding the Needle in the Haystack. *Angew. Chemie - Int. Ed.* **2013**, *52* (10), 2744–2792. https://doi.org/10.1002/anie.201208749.

(4) Moffat, J. G.; Vincent, F.; Lee, J. A.; Eder, J.; Prunotto, M. Opportunities and Challenges in Phenotypic Drug Discovery: An Industry Perspective. *Nat. Rev. Drug Discov.* **2017**, *16* (8), 531–543. https://doi.org/10.1038/nrd.2017.111.

(5) Hao, Y.-F.; Jiang, J.-G. Origin and Evolution of China Pharmacopoeia and Its Implication for Traditional Medicines. *Mini Rev. Med. Chem.* **2015**, *15* (7), 595–603.

(6) Kaufman, T. S.; Rúveda, E. A. The Quest for Quinine: Those Who Won the Battles and Those Who Won the War. *Angew. Chemie - Int. Ed.* **2005**, *44* (6), 854–885. https://doi.org/10.1002/anie.200400663.

(7) Renslo, A. R. Antimalarial Drug Discovery: From Quinine to the Dream of Eradication. *ACS Med. Chem. Lett.* **2013**, *4* (12), 1126–1128. https://doi.org/10.1021/ml4004414.

(8) Ehrlich, P.; Hata, S. *Die Experimentelle Chemotherapie Der Spirillosen*, 1st ed.; Springer, 1910. https://doi.org/https://doi.org/10.1007/978-3-642-64926-4.

(9) Bosch, F.; Rosich, L. The Contributions of Paul Ehrlich to Pharmacology: A Tribute on the Occasion of the Centenary of His Nobel Prize. *Pharmacology* **2008**, *82* (3), 171–179. https://doi.org/10.1159/000149583.

(10) Fischer, E. Einfluss Der Configuration Auf Die Wirkung Der Enzyme. *Berichte der Dtsch. Chem. Gesellschaft* **1894**, *27* (3), 2985–2993.

(11) 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.

(12) Radoux, C. J.; Vianello, F.; McGreig, J.; Desai, N.; Bradley, A. R. The Druggable Genome: Twenty Years Later. *Front. Bioinforma.* **2022**, *2*, 1–9. https://doi.org/10.3389/fbinf.2022.958378.

(13) Dai, L.; Li, Z.; Chen, D.; Jia, L.; Guo, J.; Zhao, T.; Nordlund, P. Target Identification and Validation of Natural Products with Label-Free Methodology: A Critical Review from 2005 to 2020. *Pharmacol. Ther.* **2020**, *216*. https://doi.org/10.1016/j.pharmthera.2020.107690.

(14) 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.

(15) Parvathaneni, V.; Kulkarni, N. S.; Muth, A.; Gupta, V. Drug Repurposing: A Promising Tool to Accelerate the Drug Discovery Process. *Drug Discov. Today* **2019**, *24* (10), 2076–2085. https://doi.org/10.1016/j.drudis.2019.06.014.

(16) Senger, M. R.; Fraga, C. A. M.; Dantas, R. F.; Silva, F. P. Filtering Promiscuous Compounds in Early Drug Discovery: Is It a Good Idea? *Drug Discov. Today* **2016**, *21* (6), 868–872. https://doi.org/https://doi.org/10.1016/j.drudis.2016.02.004.

(17) Merino, A.; Bronowska, A. K.; Jackson, D. B.; Cahill, D. J. Drug Profiling: Knowing Where It Hits. *Drug Discov. Today* **2010**, *15* (17–18), 749–756. https://doi.org/https://doi.org/10.1016/j.drudis.2010.06.006.

(18) Consortium, S. G.; Moustakim, M.; Felce, S. L.; Zaarour, N. *Target Identification Using Chemical Probes*, 1st ed.; Elsevier Inc., 2018; Vol. 610. https://doi.org/10.1016/bs.mie.2018.09.013.

(19) Paul, D.; Sanap, G.; Shenoy, S.; Kalyane, D.; Kalia, K.; Tekade, R. K. Artificial Intelligence in Drug Discovery and Development. *Drug Discov. Today* **2021**, *26* (1), 80–93.

(20) Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; Zhao, S. Applications of Machine Learning in Drug Discovery and Development. *Nat. Rev. Drug Discov.* **2019**, *18* (6), 463–477. https://doi.org/10.1038/s41573-019-0024-5.

(21) Schneider, G. Automating Drug Discovery. *Nat. Rev. Drug Discov.* **2018**, *17* (2), 97–113.

(22) 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.

(23) Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Orhan, I. E.; Banach, M.; Rollinger, J. M.; Barreca, D.; Weckwerth, W.; Bauer, R.; Bayer, E. A.; Majeed, M.; Bishayee, A.; Bochkov, V.; Bonn, G. K.; Braidy, N.; Bucar, F.; Cifuentes, A.; D'Onofrio, G.; Bodkin, M.; Diederich, M.; Dinkova-Kostova, A. T.; Efferth, T.; El Bairi, K.; Arkells, N.; Fan, T. P.; Fiebich, B. L.; Freissmuth, M.; Georgiev, M. I.; Gibbons, S.; Godfrey, K. M.; Gruber, C. W.; Heer, J.; Huber, L. A.; Ibanez, E.; Kijjoa, A.; Kiss, A. K.; Lu, A.; Macias, F. A.; Miller, M. J. S.; Mocan, A.; Müller, R.; Nicoletti, F.; Perry, G.; Pittalà, V.; Rastrelli, L.; Ristow, M.; Russo, G. L.; Silva, A. S.; Schuster, D.; Sheridan, H.; Skalicka-Woźniak, K.; Skaltsounis, L.; Sobarzo-Sánchez, E.; Bredt, D. S.; Stuppner, H.; Sureda, A.; Tzvetkov, N. T.; Vacca, R. A.; Aggarwal, B. B.; Battino, M.; Giampieri, F.; Wink, M.; Wolfender, J. L.; Xiao, J.; Yeung, A. W. K.; Lizard, G.; Popp, M. A.; Heinrich, M.; Berindan-Neagoe, I.; Stadler, M.; Daglia, M.; Verpoorte, R.; Supuran, C. T. Natural Products in Drug Discovery: Advances and Opportunities. *Nat. Rev. Drug Discov.* **2021**, *20* (3), 200–216. https://doi.org/10.1038/s41573-020-00114-z.

(24) Huang, B.; Zhang, Y. Teaching an Old Dog New Tricks: Drug Discovery by Repositioning Natural Products and Their Derivatives. *Drug Discov. Today* **2022**, *27* (7), 1936–1944. https://doi.org/https://doi.org/10.1016/j.drudis.2022.02.007.

(25) Huffman, B. J.; Shenvi, R. A. Natural Products in the “Marketplace”: Interfacing Synthesis and Biology. *J. Am. Chem. Soc.* **2019**, *141* (8), 3332–3346. https://doi.org/10.1021/jacs.8b11297.

(26) Mander, L.; Liu, H. W. Natural Products Drug Discovery. In *Comprehensive Natural Products II: Chemistry and Biology*; 2010; Vol. 1, pp 1–7451. https://doi.org/10.1016/C2009-1-28362-6.

(27) Atanasov, A. G.; Waltenberger, B.; Pferschy-Wenzig, E. M.; Linder, T.; Wawrosch, C.; Uhrin, P.; Temml, V.; Wang, L.; Schwaiger, S.; Heiss, E. H.; Rollinger, J. M.; Schuster, D.; Breuss, J. M.; Bochkov, V.; Mihovilovic, M. D.; Kopp, B.; Bauer, R.; Dirsch, V. M.; Stuppner, H. Discovery and Resupply of Pharmacologically Active Plant-Derived Natural Products: A Review. *Biotechnol. Adv.* **2015**, *33* (8), 1582–1614. https://doi.org/10.1016/j.biotechadv.2015.08.001.

(28) Davison, E. K.; Brimble, M. A. Natural Product Derived Privileged Scaffolds in Drug Discovery. *Curr. Opin. Chem. Biol.* **2019**, *52*, 1–8. https://doi.org/https://doi.org/10.1016/j.cbpa.2018.12.007.

(29) Vinogradov, A. A.; Yin, Y.; Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J. Am. Chem. Soc.* **2019**, *141* (10), 4167−4181. https://doi.org/10.1021/jacs.8b13178.

(30) Geiger, T. M.; Sch, S. C.; Dreizler, J. K.; Walz, M.; Hausch, F. Clues to Molecular Glues. *Curr. Res. Chem. Biol.* **2022**, *2*. https://doi.org/10.1016/j.crchbi.2021.100018.

(31) Muttenthaler, M.; King, G. F.; Adams, D. J.; Alewood, P. F. Trends in Peptide Drug Discovery. *Nat. Rev. Drug Discov.* **2021**, *20* (4), 309–325. https://doi.org/10.1038/s41573-020-00135-8.

(32) Shakya, A.; Sarma, R.; Ghimire, N.; Ghosh, S. K.; Bhat, H. R.; Rahman, O. Therapeutic Drug Monitoring and Toxicology of Immunosuppressant. In *Recent Advances in Therapeutic Drug Monitoring and Clinical Toxicology*; Kwabena, S., Yashwant, A., Eds.; Springer, 2022; pp 181–196. https://doi.org/10.1007/978-3-031-12398-6_12.

(33) Peterson, A. A.; Rangwala, A. M.; Thakur, M. K.; Ward, P. S.; Hung, C.; Outhwaite, I. R.; Chan, A. I.; Usanov, D. L.; Mootha, V. K.; Seeliger, M. A.; Liu, D. R. Discovery and Molecular Basis of Subtype-Selective Cyclophilin Inhibitors. *Nat. Chem. Biol.* **2022**, *18* (11), 1184–1195. https://doi.org/10.1038/s41589-022-01116-1.

(34) Soini, L.; Leysen, S.; Davis, J.; Ottmann, C. Molecular Glues to Stabilise Protein–Protein Interactions. *Curr. Opin. Chem. Biol.* **2022**, *69*, 1–6. https://doi.org/10.1016/j.cbpa.2022.102169.

(35) Sorokina, M.; Steinbeck, C. Review on Natural Products Databases: Where to Find Data in 2020. *J. Cheminform.* **2020**, *12*, 1–51. https://doi.org/10.1186/s13321-020-00424-9.

(36) Sorokina, M.; Merseburger, P.; Rajan, K.; Yirik, M. A.; Steinbeck, C. COCONUT Online: Collection of Open Natural Products Database. *J. Cheminform.* **2021**, *13* (1), 1–13. https://doi.org/10.1186/s13321-020-00478-9.

(37) Aebersold, R.; Mann, M. Mass-Spectrometric Exploration of Proteome Structure and Function. *Nature* **2016**, *537* (7620), 347–355. https://doi.org/10.1038/nature19949.

(38) Grada, A.; Weinbrecht, K. Next-Generation Sequencing : Methodology and Application. *J. Invest. Dermatol.* **2013**, *133* (8), 1–4. https://doi.org/10.1038/jid.2013.248.

(39) Bludau, I.; Aebersold, R. Proteomic and Interactomic Insights into the Molecular Basis of Cell Functional Diversity. *Nat. Rev. Mol. Cell Biol.* **2020**, *21* (6), 327–340. https://doi.org/10.1038/s41580-020-0231-2.

(40) Bludau, I.; Frank, M.; Dörig, C.; Cai, Y.; Heusel, M.; Rosenberger, G.; Picotti, P.; Collins, B. C.; Röst, H.; Aebersold, R. Systematic Detection of Functional Proteoform Groups from Bottom-up Proteomic Datasets. *Nat. Commun.* **2021**, *12* (1), 1–18. https://doi.org/10.1038/s41467-021-24030-x.

(41) Ong, S. E.; Mann, M. Mass Spectrometry–Based Proteomics Turns Quantitative. *Nat. Chem. Biol.* **2005**, *1* (5), 252–262. https://doi.org/10.1038/nchembio736.

(42) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L. Mapping Intact Protein Isoforms in Discovery Mode Using Top-down Proteomics. *Nature* **2011**, *480* (7376), 254–258. https://doi.org/10.1038/nature10575.

(43) Gillet, L. C.; Leitner, A.; Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu. Rev. Anal. Chem.* **2016**, *9*, 449–472. https://doi.org/10.1146/annurev-anchem-071015-041535.

(44) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Biochem.* **2013**, *85*, 5288–5296. https://doi.org/10.1021/ac4001223.

(45) 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.

(46) 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.

(47) Sinitcyn, P.; Daniel Rudolph, J.; Cox, J. Computational Methods for Understanding Mass Spectrometry–Based Shotgun Proteomics Data. *Annu. Rev. Biomed. Data Sci.* **2018**, *1*, 207–234. https://doi.org/10.1146/annurev-biodatasci-080917-013516.

(48) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as

a Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* **2002**, *1* (5), 376–386. https://doi.org/10.1074/mcp.M200025-MCP200.

(49) Thompson, A.; Schäfer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal. Chem.* **2003**, *75* (8), 1895–1904. https://doi.org/10.1021/ac0262560.

(50) 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.

(51) Piazza, I.; Kochanowski, K.; Cappelletti, V.; Fuhrer, T.; Noor, E.; Sauer, U.; Picotti, P. Resource A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication Resource A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell* **2018**, *172* (1), 358–361. https://doi.org/10.1016/j.cell.2017.12.006.

(52) Piazza, I.; Beaton, N.; Bruderer, R.; Knobloch, T.; Barbisan, C.; Chandat, L.; Sudau, A.; Siepe, I.; Rinner, O.; de Souza, N.; Picotti, P.; Reiter, L. A Machine Learning-Based Chemoproteomic Approach to Identify Drug Targets and Binding Sites in Complex Proteomes. *Nat. Commun.* **2020**, *11* (1), 1–13. https://doi.org/10.1038/s41467-020-18071-x.

(53) 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* **2014**, *346* (6205). https://doi.org/10.1126/science.1255784.

(54) Gaetani, M.; Sabatier, P.; Saei, A. A.; Beusch, C. M.; Yang, Z.; Lundström, S. L.; Zubarev, R. A. Proteome Integral Solubility Alteration: A High-Throughput Proteomics Assay for Target Deconvolution. *J. Proteome Res.* **2019**, *18* (11), 4027–4037. https://doi.org/10.1021/acs.jproteome.9b00500.

(55) Park, C.; Marqusee, S. Pulse Proteolysis : A Simple Method for Quantitative Determination of Protein Stability and Ligand Binding. *Nat. Methods* **2005**, *2* (3), 207–212. https://doi.org/10.1038/NMETH740.

(56) West, G. M.; Tucker, C. L.; Xu, T.; Park, S. K.; Han, X.; Yates, J. R.; Fitzgerald, M. C. Quantitative Proteomics Approach for Identifying Protein-Drug Interactions in Complex Mixtures Using Protein Stability Measurements. *Proc. Natl. Acad. Sci.* **2010**, *107* (20), 9078–9082. https://doi.org/10.1073/pnas.1000148107.

(57) Mateus, A.; Kurzawa, N.; Perrin, J.; Bergamini, G.; Savitski, M. M. Drug Target Identification in Tissues by Thermal Proteome Profiling. *Annu. Rev. Pharmacol. Toxicol.* **2022**, *62*, 465–482. https://doi.org/10.1146/annurev-pharmtox-052120-013205.

(58) Sun, J.; Prabhu, N.; Tang, J.; Yang, F.; Nordlund, P.; Dai, L. Recent Advances in Proteome-Wide Label-Free Target Deconvolution for Bioactive Small Molecules. *Med. Res. Rev.* **2021**, *41* (6), 2893–2926.

https://doi.org/10.1002/med.21788.

(59) Ruan, C.; Ning, W.; Liu, Z.; Zhang, X.; Fang, Z.; Li, Y.; Dang, Y.; Xue, Y.; Ye, M. Precipitate-Supported Thermal Proteome Profiling Coupled with Deep Learning for Comprehensive Screening of Drug Target Proteins. *ACS Chem. Biol.* **2022**, *17* (1), 252–262. https://doi.org/10.1021/acschembio.1c00936.

(60) Schulze, W. X.; Mann, M. A Novel Proteomic Screen for Peptide-Protein Interactions. *J. Biol. Chem.* **2004**, *279* (11), 10756–10764. https://doi.org/10.1074/jbc.M309909200.

(61) Bird, R. E.; Lemmel, S. A.; Yu, X.; Zhou, Q. A. Bioorthogonal Chemistry and Its Applications. *Bioconjug. Chem.* **2021**, *32* (12), 2457–2479. https://doi.org/10.1021/acs.bioconjchem.1c00461.

(62) Karaj, E.; Sindi, S. H.; Viranga Tillekeratne, L. M. Photoaffinity Labeling and Bioorthogonal Ligation: Two Critical Tools for Designing “Fish Hooks” to Scout for Target Proteins. *Bioorg. Med. Chem.* **2022**, *62*. https://doi.org/https://doi.org/10.1016/j.bmc.2022.116721.

(63) Kawatani, M.; Osada, H. Affinity-Based Target Identification for Bioactive Small Molecules. *Med. Chem. Commun.* **2014**, *5* (3), 277–287. https://doi.org/10.1039/c3md00276d.

(64) 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* **2017**, *358* (6367), 1–16. https://doi.org/10.1126/science.aan4368.

(65) Sharma, K.; Weber, C.; Bairlein, M.; Greff, Z.; Kéri, G.; Cox, J.; Olsen, J. V; Daub, H. Proteomics Strategy for Quantitative Protein Interaction Profiling in Cell Extracts. *Nat. Methods* **2009**, *6* (10), 741–744. https://doi.org/10.1038/nmeth.1373.

(66) Bantscheff, M.; Eberhard, D.; Abraham, Y.; Bastuck, S.; Boesche, M.; Hobson, S.; Sweetman, G.; Mathieson, T.; Perrin, J.; Raida, M.; Rau, C.; 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. Bio* **2007**, *25* (9), 1035–1044. https://doi.org/10.1038/nbt1328.

(67) Bassiouni, R.; Nemec, K. N.; Iketani, A.; Flores, O.; Showalter, A.; Khaled, A. S.; Vishnubhotla, P.; Jr, R. W. S.; Kaittanis, C.; Perez, J. M.; Khaled, A. R. Chaperonin Containing TCP-1 Protein Level in Breast Cancer Cells Predicts Therapeutic Application of a Cytotoxic Peptide. *Clin. Cancer Res.* **2016**, *22* (8), 4366–4379. https://doi.org/10.1158/1078-0432.CCR-15-2502.

(68) Won, S. J.; Eschweiler, J. D.; Majmudar, J. D.; Chong, F. S.; Hwang, S. Y.;

Ruotolo, B. T.; Martin, B. R. Affinity-Based Selectivity Profiling of an In-Class Selective Competitive Inhibitor of Acyl Protein Thioesterase 2. *ACS Med. Chem. Lett.* **2017**, *8* (2), 215–220. https://doi.org/10.1021/acsmedchemlett.6b00441.

(69) Kitamura, K.; Itoh, H.; Sakurai, K.; Dan, S.; Inoue, M. Target Identification of Yaku'amide B and Its Two Distinct Activities against Mitochondrial F0F1-ATP Synthase. *J. Am. Chem. Soc.* **2018**, *140* (38), 12189−12199. https://doi.org/10.1021/jacs.8b07339.

(70) Keilhauer, E. C.; Hein, M. Y.; Mann, M. Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). *Mol. Cell. Proteomics* **2015**, *14* (1), 120–135. https://doi.org/10.1074/mcp.M114.041012.

(71) Smits, A. H.; Vermeulen, M. Characterizing Protein–Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. *Trends Biotechnol.* **2016**, *34* (10), 825–834. https://doi.org/10.1016/j.tibtech.2016.02.014.

(72) Meyer, K.; Selbach, M. Quantitative Affinity Purification Mass Spectrometry: A Versatile Technology to Study Protein-Protein Interactions. *Front. Genet.* **2015**, *6*, 1–7. https://doi.org/10.3389/fgene.2015.00237.

(73) Frank, R. The SPOT-Synthesis Technique: Synthetic Peptide Arrays on Membrane Supports - Principles and Applications. *J. Immunol. Methods* **2002**, *267* (1), 13–26. https://doi.org/10.1016/S0022-1759(02)00137-0.

(74) Frank, R. Spot-Synthesis: An Easy Technique for the Positionally Addressable, Parallel Chemical Synthesis on a Membrane Support. *Tetrahedron* **1992**, *48* (42), 9217–9232. https://doi.org/10.1016/S0040-4020(01)85612-X.

(75) Meyer, K.; Selbach, M. Peptide-Based Interaction Proteomics. *Mol. Cell. Proteomics* **2020**, *19* (7), 1070–1075. https://doi.org/10.1074/mcp.R120.002034.

(76) Ramberger, E.; Suarez-Artiles, L.; Perez-Hernandez, D.; Haji, M.; Popp, O.; Reimer, U.; Leutz, A.; Dittmar, G.; Mertins, P. A Universal Peptide Matrix Interactomics Approach to Disclose Motif-Dependent Protein Binding. *Mol. Cell. Proteomics* **2021**, *20*, 1–12. https://doi.org/10.1016/j.mcpro.2021.100135.

(77) Ramberger, E.; Sapozhnikova, V.; Kowenz-Leutz, E.; Zimmermann, K.; Nicot, N.; Nazarov, P. V; Perez-Hernandez, D.; Reimer, U.; Mertins, P.; Dittmar, G. PRISMA and BioID Disclose a Motifs-Based Interactome of the Intrinsically Disordered Transcription Factor C/EBPα. *iScience* **2021**, *24* (6), 1–16. https://doi.org/10.1016/j.isci.2021.102686.

(78) Meyer, K.; Kirchner, M.; Uyar, B.; Cheng, J. Y.; Russo, G.; Hernandez-Miranda, L. R.; Szymborska, A.; Zauber, H.; Rudolph, I. M.; Willnow, T. E.; Akalin, A.; Haucke, V.; Gerhardt, H.; Birchmeier, C.; Kühn, R.; Krauss, M.; Diecke, S.; Pascual, J. M.; Selbach, M. Mutations in Disordered Regions Can Cause Disease by Creating Dileucine Motifs. *Cell* **2018**, *175* (1), 239–253. https://doi.org/10.1016/j.cell.2018.08.019.

(79) Wright, M. H.; Sieber, S. Chemical Proteomics Approaches for Identifying the Cellular Targets of Natural Products. *Nat. Prod. Rep.* **2016**, *33*, 681–708. https://doi.org/10.1039/c6np00001k.

(80) Cabrera, A.; Wiebelhaus, N.; Quan, B.; Ma, R.; Meng, H.; Fitzgerald, M. C. Comparative Analysis of Mass-Spectrometry-Based Proteomic Methods for Protein Target Discovery Using a One-Pot Approach. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (2), 217–226. https://doi.org/10.1021/jasms.9b00041.

(81) 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.

(82) Suzuki, T.; Okamura, T.; Tomohiro, T.; Iwabuchi, Y.; Kanoh, N. Third Generation Photo-Cross-Linked Small-Molecule Affinity Matrix: A Photoactivatable and Photocleavable System Enabling Quantitative Analysis of the Photo-Cross-Linked Small Molecules and Their Target Purification. *Bioconjug. Chem.* **2015**, *26* (3), 389–395. https://doi.org/10.1021/bc500559e.

(83) Dilly, S. J.; Bell, M. J.; Clark, A. J.; Marsh, A.; Napier, R. M.; Sergeant, M. J.; Thompson, J.; Taylor, P. C. A Photoimmobilisation Strategy That Maximises Exploration of Chemical Space in Small Molecule Affinity Selection and Target Discovery. *Chem. Commun.* **2007**, No. 27, 2808–2810. https://doi.org/10.1039/b704271j.

(84) Ota, E.; Usui, K.; Oonuma, K.; Koshino, H.; Nishiyama, S.; Hirai, G.; Sodeoka, M. Thienyl-Substituted α - Ketoamide: A Less Hydrophobic Reactive Group for Photo-Affinity Labeling. *ACS Chem. Biol.* **2018**, *13* (4), 876−880. https://doi.org/10.1021/acschembio.7b00988.

(85) Lougee, M. G.; Pagar, V. V.; Kim, H. J.; Pancoe, S. X.; Chia, W. K.; Mach, R. H.; Garcia, B. A.; Petersson, E. J. Harnessing the Intrinsic Photochemistry of Isoxazoles for the Development of Chemoproteomic Crosslinking Methods. *Chem. Commun.* **2022**, *58* (65), 9116–9119. https://doi.org/10.1039/D2CC02263J.

(86) West, A. V.; Woo, C. M. Photoaffinity Labeling Chemistries Used to Map Biomolecular Interactions. *Isr. J. Chem.* **2023**, *63*, 1–20. https://doi.org/10.1002/ijch.202200081.

(87) 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 Sci.* **2017**, *15*, 1–34. https://doi.org/10.1186/s12953-017-0123-3.

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

(89) Wittelsberger, A.; Thomas, B. E.; Mierke, D. F.; Rosenblatt, M. Methionine Acts as a '"Magnet"' in Photoaffinity Crosslinking Experiments. *FEBS Lett.* **2006**, *580*, 1872–1876. https://doi.org/10.1016/j.febslet.2006.02.050.

(90) 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.

(91) Kotzyba-Hibert, F.; Kapfer, I.; Goeldner, M. Recent Trends in Photoaffinity

Labeling. *Angew. Chemie Int. Ed. English* **1995**, *34* (12), 1296–1312. https://doi.org/10.1002/anie.199512961.

(92) Preston, G. W.; Wilson, A. J. Photo-Induced Covalent Cross-Linking for the Analysis of Biomolecular Interactions. *Chem. Soc. Rev.* **2013**, *42* (8), 3289–3301. https://doi.org/10.1039/c3cs35459h.

(93) Kanoh, N. Photo-Cross-Linked Small-Molecule Affinity Matrix as a Tool for Target Identification of Bioactive Small Molecules. *Nat. Prod. Rep.* **2016**, *33* (5), 709–718. https://doi.org/10.1039/c5np00117j.

(94) Kirmse, W. Carbenes and the O-H Bond. *Jounal of Cheminformatics* **1997**, *28* (10).

(95) Song, M.-G.; Sheridan, R. S. Regiochemical Substituent Switching of Spin States in Aryl(Trifluoromethyl)Carbenes. *J. Am. Chem. Soc.* **2011**, *133* (49), 19688–19690. https://doi.org/10.1021/ja209613u.

(96) Sander, W.; Bucher, G.; Wierlacher, S. Carbenes in Matrices: Spectroscopy, Structure, and Reactivity. *Chem. Rev.* **1993**, *93* (4), 1583–1621.

(97) Brunner, J.; Senn, H.; Richards, F. M. 3-Trifluoromethyl-3-Phenyldiazirine. A New Carbene Generating Group for Photolabeling Reagents. *J. Biol. Chem.* **1980**, *255* (8), 3313–3318. https://doi.org/10.1016/s0021-9258(19)85701-0.

(98) Conway, L. P.; Jadhav, A. M.; Homan, R. A.; Li, W.; Rubiano, J. S.; Hawkins, R.; Michael, R.; Parker, C. G. Evaluation of Fully-Functionalized Diazirine Tags for Chemical Proteomic Applications. *Chem. Sci.* **2021**, *12*, 7839–7847. https://doi.org/10.1039/d1sc01360b.

(99) Iacobucci, C.; Götze, M.; Piotrowski, C.; Arlt, C.; Rehkamp, A.; Ihling, C.; Hage, C.; Sinz, A. Carboxyl-Photo-Reactive MS-Cleavable Cross-Linkers: Unveiling a Hidden Aspect of Diazirine-Based Reagents. *Anal. Chem.* **2018**, *90* (4), 2805–2809. https://doi.org/10.1021/acs.analchem.7b04915.

(100) Hill, J. R.; Robertson, A. A. B. Fishing for Drug Targets: A Focus on Diazirine Photoaffinity Probe Synthesis. *J. Med. Chem.* **2018**, *61* (16), 6945–6963. https://doi.org/10.1021/acs.jmedchem.7b01561.

(101) Liu, M. T. H.; Choe, Y. K.; Kimura, M.; Kobayashi, K.; Nagase, S.; Wakahara, T.; Niino, Y.; Ishitsuka, M. O.; Maeda, Y.; Akasaka, T. Effect of Substituents on the Thermal Decomposition of Diazirines: Experimental and Computational Studies. *J. Org. Chem.* **2003**, *68* (19), 7471–7478. https://doi.org/10.1021/jo034949q.

(102) Kawamura, T.; Kondoh, Y.; Muroi, M.; Kawatani, M.; Osada, H. A Small Molecule That Induces Reactive Oxygen Species via Cellular Glutathione Depletion. *Biochem. J.* **2014**, *463* (1), 53–63. https://doi.org/10.1042/BJ20140669.

(103) Sasazawa, Y.; Kanagaki, S.; Tashiro, E.; Nogawa, T.; Muroi, M.; Kondoh, Y.; Osada, H.; Imoto, M. Xanthohumol Impairs Autophagosome Maturation through Direct Inhibition of Valosin-Containing Protein. *ACS Chem. Biol.* **2012**, *7* (5), 892–900. https://doi.org/10.1021/cb200492h.

(104) Kanoh, N.; Honda, K.; Simizu, S.; Muroi, M.; Osada, H. Photo-Cross-Linked Small-Molecule Affinity Matrix for Facilitating Forward and Reverse Chemical

Genetics. *Angew. Chemie* **2005**, *117* (23), 3625–3628. https://doi.org/10.1002/ange.200462370.

(105) Kanoh, N.; Kumashiro, S.; Simizu, S.; Kondoh, Y.; Hatakeyama, S.; Tashiro, H.; Osada, H. Immobilization of Natural Products on Glass Slides by Using a Photoaffinity Reaction and the Detection of Protein-Small-Molecule Interactions. *Angew. Chemie - Int. Ed.* **2003**, *42* (45), 5584–5587. https://doi.org/10.1002/anie.200352164.

(106) Kanoh, N.; Nakamura, T.; Honda, K.; Yamakoshi, H. Distribution of Photo-Cross-Linked Products from 3-Aryl-3-Trifluoro-Methyldiazirines and Alcohols. *Tetrahedron* **2008**, *64* (24), 5692–5698. https://doi.org/10.1016/j.tet.2008.04.031.

(107) Ketenoglu, D. A General Overview and Comparative Interpretation on Element-Specific X-Ray Spectroscopy Techniques: XPS, XAS, and XRS. *X-Ray Spectrom.* **2022**, *51* (5–6), 422–443. https://doi.org/10.1002/xrs.3299.

(108) Greczynski, G.; Hultman, L. X-Ray Photoelectron Spectroscopy: Towards Reliable Binding Energy Referencing. *Prog. Mater. Sci.* **2020**, *107*, 1–46. https://doi.org/10.1016/j.pmatsci.2019.100591.

(109) Goon, I. Y.; Lai, L. M. H.; Lim, M.; Amal, R.; Gooding, J. J. 'Dispersible Electrodes': A Solution to Slow Response Times of Sensitive Sensors. *Chem. Commun.* **2010**, *46* (46), 8821–8823. https://doi.org/10.1039/C0CC02690E.

(110) Liu, G. L.; Kazarian, S. G. Recent Advances and Applications to Cultural Heritage Using ATR-FTIR Spectroscopy and ATR-FTIR Spectroscopic Imaging. *Analyst* **2022**, *147* (9), 1777–1797. https://doi.org/10.1039/d2an00005a.

(111) Tiernan, H.; Byrne, B.; Kazarian, S. G. ATR-FTIR Spectroscopy and Spectroscopic Imaging for the Analysis of Biopharmaceuticals. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2020**, *241*. https://doi.org/10.1016/j.saa.2020.118636.

(112) Lopera-Valle, A.; Elias, A. Amine Responsive Poly (Lactic Acid)(PLA) and Succinic Anhydride (SAh) Graft-Polymer: Synthesis and Characterization. *Polymers* **2019**, *11* (9), 1–19. https://doi.org/10.3390/polym11091466.

(113) Fearn, S. Characterisation of Biological Material with ToF-SIMS: A Review. *Mater. Sci. Technol.* **2015**, *31* (2), 148–161. https://doi.org/10.1179/1743284714Y.0000000668.

(114) Prasad, A.; Salim, N. V; Mozetic, M.; Kailas, L.; Thomas, S. Time-of-Flight Secondary Ion Mass Spectrometric Analysis of Polymer Surfaces : A Review. *J. Appl. Polym. Sci.* **2022**, *139*, 1–20. https://doi.org/10.1002/app.52286.

(115) Vickerman, J. C.; Briggs, D. *ToF-SIMS: Surface Analysis by Mass Spectrometry*, 1st ed.; IM, 2001.

(116) Sakurai, K.; Ozawa, S.; Yamada, R.; Yasui, T.; Mizuno, S. Comparison of the Reactivity of Carbohydrate Photoaffinity Probes with Different Photoreactive Groups. *ChemBioChem* **2014**, *15* (10), 1399–1403. https://doi.org/10.1002/cbic.201402051.

(117) Dubinsky, L.; Krom, B. P.; Meijler, M. M. Diazirine Based Photoaffinity Labeling.

Bioorg. Med. Chem. **2012**, *20* (2), 554–570. https://doi.org/10.1016/j.bmc.2011.06.066.

(118) Klemm, D.; Heublein, B.; Fink, H.; Bohn, A. Cellulose : Fascinating Biopolymer and Sustainable Raw Material Angewandte. *Polym. Sci.* **2005**, *44* (22), 3358–3393. https://doi.org/10.1002/anie.200460587.

(119) Li, T.; Chen, C.; Brozena, A. H.; Zhu, J. Y.; Xu, L.; Driemeier, C.; Dai, J.; Rojas, O. J.; Isogai, A.; Wågberg, L.; Hu, L. Developing Fibrillated Cellulose as a Sustainable Technological Material. *Nature* **2021**, *590* (7844), 47–56. https://doi.org/10.1038/s41586-020-03167-7.

(120) Shaghaleh, H.; Xu, X.; Wang, S. Current Progress in Production of Biopolymeric Materials Based on Cellulose, Cellulose Nanofibers, and Cellulose Derivatives. *RSC Adv.* **2018**, *8* (2), 825–842. https://doi.org/10.1039/c7ra11157f.

(121) Nongbe, M. C.; Bretel, G.; Ekou, L.; Ekou, T.; Robitzer, M.; Le Grognec, E.; Felpin, F. X. Cellulose Paper Azide as a Molecular Platform for Versatile Click Ligations: Application to the Preparation of Hydrophobic Paper Surface. *Cellulose* **2018**, *25*, 1395–1411. https://doi.org/10.1007/s10570-017-1647-5.

(122) Orelma, H.; Filpponen, I.; Johansson, L. S.; Österberg, M.; Rojas, O. J.; Laine, J. Surface Functionalized Nanofibrillar Cellulose (NFC) Film as a Platform for Immunoassays and Diagnostics. *Biointerphases* **2012**, *7* (61), 1–12. https://doi.org/10.1007/s13758-012-0061-7.

(123) Singhal, S.; Mehta, J.; Desikan, R.; Ayers, D.; Roberson, P.; Eddlemon, P.; Munshi, N.; Anaissie, E.; Wilson, C.; Dhodapkar, M.; Zeldis, J.; Siegel, D.; Crowley, J.; Barlogie, B. Antitumor Activity of Thalidomide in Refractory Multiple Myeloma. *N. Engl. J. Med.* **1999**, *341* (21), 1565–1571. https://doi.org/10.1056/nejm199911183412102.

(124) Mori, T.; Ito, T.; Liu, S.; Ando, H.; Sakamoto, S.; Yamaguchi, Y.; Tokunaga, E.; Shibata, N.; Handa, H.; Hakoshima, T. Structural Basis of Thalidomide Enantiomer Binding to Cereblon. *Sci. Rep.* **2018**, *8* (1), 1–14. https://doi.org/10.1038/s41598-018-19202-7.

(125) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for Micro-Purification, Enrichment, Pre-Fractionation and Storage of Peptides for Proteomics Using StageTips. *Nat. Protoc.* **2007**, *2* (8), 1896–1906. https://doi.org/10.1038/nprot.2007.261.

(126) Harding, M. W.; Galat, A.; Uehling, D. E.; Schreiber, S. L. A Receptor for the Immunosuppressant FK506 Is a Cis-Trans Peptidyl-Prolyl Isomerase. *Nature* **1989**, *341* (6244), 758–760.

(127) Xu, C.; Zhang, J.; Huang, X.; Sun, J.; Xu, Y.; Tang, Y.; Wu, J.; Shi, Y.; Huang, Q.; Zhang, Q. Solution Structure of Human Peptidyl Prolyl Isomerase-like Protein 1 and Insights into Its Interaction with SKIP. *J. Biol. Chem.* **2006**, *281* (23), 15900–15908.

(128) 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. https://doi.org/10.1021/acschembio.2c00500.

(129) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31* (13), 3784–3788. https://doi.org/10.1093/nar/gkg563.

(130) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. Identification of a Primary Target of Thalidomide Teratogenicity. *Science* **2010**, *327* (5971), 1345–1350. https://doi.org/10.1126/science.1177319.

(131) Kim, K.; Lee, D. H.; Park, S.; Jo, S. H.; Ku, B.; Park, S. G.; Park, B. C.; Jeon, Y. U.; Ahn, S.; Kang, C. H.; Hwang, D.; Chae, S.; Ha, J. Du; Kim, S.; Hwang, J. Y.; Kim, J. H. Disordered Region of Cereblon Is Required for Efficient Degradation by Proteolysis-Targeting Chimera. *Sci. Rep.* **2019**, *9* (1), 1–14. https://doi.org/10.1038/s41598-019-56177-5.

(132) Papke, B.; Murarka, S.; Vogel, H. A.; Martín-Gago, P.; Kovacevic, M.; Truxius, D. C.; Fansa, E. K.; Ismail, S.; Zimmermann, G.; Heinelt, K.; Schultz-Fademrecht, C.; Al Saabi, A.; Baumann, M.; Nussbaumer, P.; Wittinghofer, A.; Waldmann, H.; Bastiaens, P. I. H. Identification of Pyrazolopyridazinones as PDEδ Inhibitors. *Nat. Commun.* **2016**, *7* (1–9). https://doi.org/10.1038/ncomms11360.

(133) Martín-Gago, P.; Fansa, E. K.; Klein, C. H.; Murarka, S.; Janning, P.; Schürmann, M.; Metz, M.; Ismail, S.; Schultz-Fademrecht, C.; Baumann, M.; Bastiaens, P. I. H.; Wittinghofer, A.; Waldmann, H. A PDE6δ-KRas Inhibitor Chemotype with up to Seven H-Bonds and Picomolar Affinity That Prevents Efficient Inhibitor Release by Arl2. *Angew. Chemie* **2017**, *129* (9), 2463–2468. https://doi.org/10.1002/ange.201610957.

(134) Zimmermann, G.; Papke, B.; Ismail, S.; Vartak, N.; Chandra, A.; Hoffmann, M.; Hahn, S. A.; Triola, G.; Wittinghofer, A.; Bastiaens, P. I. H.; Waldmann, H. Small Molecule Inhibition of the KRAS-PDEδ Interaction Impairs Oncogenic KRAS Signalling. *Nature* **2013**, *497* (7451), 638–642. https://doi.org/10.1038/nature12205.

(135) Teng, M.; Lu, W.; Donovan, K. A.; Sun, J.; Krupnick, N. M.; Nowak, R. P.; Li, Y. Der; Sperling, A. S.; Zhang, T.; Ebert, B. L.; Fischer, E. S.; Gray, N. S. Development of PDE6D and CK1α Degraders through Chemical Derivatization of FPFT-2216. *J. Med. Chem.* **2022**, *65* (1), 747–756. https://doi.org/10.1021/acs.jmedchem.1c01832.

(136) Fischer, E. S.; Böhm, K.; Lydeard, J. R.; Yang, H.; Stadler, M. B.; Cavadini, S.; Nagel, J.; Serluca, F.; Acker, V.; Lingaraju, G. M.; Tichkule, R. B.; Schebesta, M.; Forrester, W. C.; Schirle, M.; Hassiepen, U.; Ottl, J.; Hild, M.; Beckwith, R. E. J.; Harper, J. W.; Jenkins, J. L.; Thomä, N. H. Structure of the DDB1-CRBN E3 Ubiquitin Ligase in Complex with Thalidomide. *Nature* **2014**, *512* (1), 49–53. https://doi.org/10.1038/nature13527.

(137) Garber, K. The PROTAC Gold Rush. *Nat. Biotechnol.* **2022**, *40* (1), 12–16. https://doi.org/10.1038/s41587-021-01173-2.

(138) Lucke, C.; Weiwad, M. Insights into Immunophilin Structure and Function. *Curr. Med. Chem.* **2011**, *18* (35), 5333–5354.

https://doi.org/10.2174/092986711798194324.

(139) Sakamoto, S.; Kabe, Y.; Hatakeyama, M.; Yamaguchi, Y.; Handa, H. Development and Application of High-Performance Affinity Beads: Toward Chemical Biology and Drug Discovery. *Chem. Rec.* **2009**, *9* (1), 66–85. https://doi.org/10.1002/tcr.20170.

(140) Du, G.-H. *Natural Small Molecule Drugs from Plants*; Springer, 2018. https://doi.org/10.1007/978-981-10-8022-7.

(141) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: Towards Direct Deposition of Bioassay Data. *Nucleic Acids Res.* **2019**, *47* (DI), D930–D940. https://doi.org/10.1093/nar/gky1075.

(142) Brown, E. J.; Albers, M. W.; Bum Shin, T.; Ichikawa, K.; Keith, C. T.; Lane, W. S.; Schreiber, S. L. A Mammalian Protein Targeted by G1-Arresting Rapamycin-Receptor Complex. *Nature* **1994**, *369* (6483), 756–758. https://doi.org/10.1038/369756a0.

(143) Sato, S. I.; Kwon, Y.; Kamisuki, S.; Srivastava, N.; Mao, Q.; Kawazoe, Y.; Uesugi, M. Polyproline-Rod Approach to Isolating Protein Targets of Bioactive Small Molecules: Isolation of a New Target of Indomethacin. *J. Am. Chem. Soc.* **2007**, *129* (4), 873–880. https://doi.org/10.1021/ja0655643.

(144) Zheng, W.; Li, G.; Li, X. Affinity Purification in Target Identification: The Specificity Challenge. *Arch. Pharm. Res.* **2015**, *38*, 1661–1685. https://doi.org/10.1007/s12272-015-0635-2.

(145) Lamoia, T. E.; Shulman, G. I. Cellular and Molecular Mechanisms of Metformin Action. *Endocr. Rev.* **2021**, *42* (1), 77–96. https://doi.org/10.1210/endrev/bnaa023.

(146) Allende-Vega, N.; Marco Brualla, J.; Falvo, P.; Alexia, C.; Constantinides, M.; de Maudave, A. F.; Coenon, L.; Gitenay, D.; Mitola, G.; Massa, P.; Orecchioni, S.; Bertolini, F.; Marzo, I.; Anel, A.; Villalba, M. Metformin Sensitizes Leukemic Cells to Cytotoxic Lymphocytes by Increasing Expression of Intercellular Adhesion Molecule-1 (ICAM-1). *Sci. Rep.* **2022**, *12* (1), 1–12. https://doi.org/10.1038/s41598-022-05470-x.

(147) Li, N.; Zhou, T.; Fei, E. Actions of Metformin in the Brain: A New Perspective of Metformin Treatments in Related Neurological Disorders. *Int. J. Mol. Sci.* **2022**, *23* (15), 1–22. https://doi.org/10.3390/ijms23158281.

(148) Induri, S. N. R.; Kansara, P.; Thomas, S. C.; Xu, F.; Saxena, D.; Li, X. The Gut Microbiome, Metformin, and Aging. *Annu. Rev. Pharmacol. Toxicol.* **2021**, *62*, 85–108. https://doi.org/10.1146/annurev-pharmtox-051920-093829.

(149) Fontaine, E. Metformin-Induced Mitochondrial Complex I Inhibition: Facts, Uncertainties, and Consequences. *Front. Endocrinol. (Lausanne).* **2018**, *9*. https://doi.org/10.3389/fendo.2018.00753.

(150) Baur, J. A.; Birnbaum, M. J. Control of Gluconeogenesis by Metformin: Does

Redox Trump Energy Charge? *Cell Metab.* **2014**, *20* (2), 197–199. https://doi.org/10.1016/j.cmet.2014.07.013.

(151) Madiraju, A. K.; Erion, D. M.; Rahimi, Y.; Zhang, X.; Macdonald, J.; Jurczak, M.; Camporez, J.; Lee, H.; Cline, G. W. Metformin Suppresses Gluconeogenesis by Inhibiting Mitochondrial Glycerophosphate Dehydrogenase. *Nature* **2014**, *510* (7506), 542–546. https://doi.org/10.1038/nature13270.

(152) Hughes, C. S.; Foehr, S.; Garfield, D. A.; Furlong, E. E.; Steinmetz, L. M. Ultrasensitive Proteome Analysis Using Paramagnetic Bead Technology. *Mol. Syst.* **2014**, *10* (10), 1–14. https://doi.org/10.15252/msb.20145625.

(153) Schwanhäusser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global Quantification of Mammalian Gene Expression Control. *Nature* **2011**, *473* (7347), 337–342. https://doi.org/https://doi.org/10.1038/nature10098.

(154) Pohl, M.; Michaelis, N.; Meister, F.; Heinze, T. Biofunctional Surfaces Based on Dendronized Cellulose. *Biomacromolecules* **2009**, *10* (2), 382–389. https://doi.org/10.1021/bm801149u.

(155) Odinolfi, M. T.; Romanato, A.; Bergamaschi, G.; Strada, A.; Sola, L.; Girella, A.; Milanese, C.; Chiari, M.; Gori, A.; Cretich, M. Clickable Cellulosic Surfaces for Peptide-Based Bioassays. *Talanta* **2019**, *205*. https://doi.org/10.1016/j.talanta.2019.120152.

(156) Verheul, T. C. J.; van Hijfte, L.; Perenthaler, E.; Barakat, T. S. The Why of YY1: Mechanisms of Transcriptional Regulation by Yin Yang 1. *Front. Cell Dev. Biol.* **2020**, *8*, 1–9. https://doi.org/10.3389/fcell.2020.592164.

(157) Kobaisi, F.; Fayyad, N.; Rezvani, H. R.; Fayyad-Kazan, M.; Sulpice, E.; Badran, B.; Fayyad-Kazan, H.; Gidrol, X.; Rachidi, W. Signaling Pathways, Chemical and Biological Modulators of Nucleotide Excision Repair: The Faithful Shield against UV Genotoxicity. *Oxid. Med. Cell. Longev.* **2019**, *2019*, 1–18. https://doi.org/10.1155/2019/4654206.

(158) Bradner, J. E.; McPherson, O. M.; Koehler, A. N. A Method for the Covalent Capture and Screening of Diverse Small Molecules in a Microarray Format. *Nat. Protoc.* **2006**, *1* (5), 2344–2352. https://doi.org/10.1038/nprot.2006.282.

(159) Fontalba, A.; Paciuccil, R.; Avila, J.; Zabala, J. C. Incorporation of Tubulin Subunits into Dimers Requires GTP Hydrolysis. *J. Cell Sci.* **1993**, *106* (2), 627–632. https://doi.org/10.1242/jcs.106.2.627.

(160) Miller, L. M.; Xiao, H.; Burd, B.; Horwitz, S. B.; Angeletti, R. H.; Verdier-Pinard, P. Methods in Tubulin Proteomics. In *Method in Cell Biology*; 2010; Vol. 95, pp 105–126. https://doi.org/10.1016/S0091-679X(10)95007-3.

(161) Lin, Z.; Amako, Y.; Kabir, F.; Flaxman, H. A.; Budnik, B.; Woo, C. M. Development of Photolenalidomide for Cellular Target Identification. *J. Am. Chem. Soc.* **2022**, *144* (1), 606–614. https://doi.org/10.1021/jacs.1c11920.

(162) Noun, M.; Akoumeh, R.; Abbas, I. Cell and Tissue Imaging by TOF-SIMS and MALDI-TOF : An Overview for Biological and Pharmaceutical Analysis. *Microsc. Microanal.* **2022**, *28* (1), 1–26. https://doi.org/10.1017/S1431927621013593.

(163) Wenschuh, H.; Volkmer-Engert, R.; Schmidt, M.; Schulz, M.; Schneider-Mergener, J.; Reineke, U. Coherent Membrane Supports for Parallel Microsynthesis and Screening of Bioactive Peptides. *Pept. Sci.* **2000**, *55* (3), 188–206. https://doi.org/10.1002/1097-0282(2000)55:3<188::AID-BIP20>3.0.CO;2-T.

(164) Sandler, N.; Anni, M.; Ihalainen, P.; Kronberg, L.; Meierjohann, A.; Viitala, T.; Peltonen, J. Inkjet Printing of Drug Substances and Use of Porous Substrates-Towards Individualized Dosing. *J. Pharm. Sci.* **2011**, *100* (8), 3386–3395. https://doi.org/10.1002/jps.22526.

(165) Levandowsky, M.; Winter, D. Distance between Sets. *Nature* **1971**, *234* (5323), 34–35. https://doi.org/10.1038/234034a0.

(166) Lu, J.; Carlson, H. A. ChemTreeMap: An Interactive Map of Biochemical Similarity in Molecular Datasets. *Bioinformatics* **2016**, *32* (23), 3584–3592. https://doi.org/10.1093/bioinformatics/btw523.

(167) Bache, N.; Geyer, P. E.; Bekker-Jensen, D. B.; Hoerning, O.; Falkenby, L.; Treit, P. V.; Doll, S.; Paron, I.; Müller, J. B.; Meier, F.; Olsen, J. V.; Vorm, O.; Mann, M. A Novel LC System Embeds Analytes in Pre-Formed Gradients for Rapid, Ultra-Robust Proteomics. *Mol. Cell. Proteomics* **2018**, *17* (11), 2284–2296. https://doi.org/10.1074/mcp.TIR118.000853.

(168) Evosep. *Towards a Standardized Omics Platform with the 300 Samples per Day Method*; 2020. https://www.evosep.com/wp-content/uploads/2020/06/AN-005-20-06_300SPD.pdf.

(169) Nimse, S. B.; Song, K.; Sonawane, M. D.; Sayyed, D. R.; Kim, T. Immobilization Techniques for Microarray: Challenges and Applications. *Sensors* **2014**, *14* (12), 22208–22229. https://doi.org/10.3390/s141222208.

(170) Yin, Y.; Zhang, C.; Yu, W.; Kang, G.; Yang, Q.; Shi, Z.; Xiong, C. Transparent and Flexible Cellulose Dielectric Films with High Breakdown Strength and Energy Density. *Energy Storage Mater.* **2020**, *26*, 105–111. https://doi.org/10.1016/j.ensm.2019.12.034.

(171) Ghiassian, S.; Biesinger, M. C.; Workentin, M. S. Synthesis of Small Water-Soluble Diazirine-Functionalized Gold Nanoparticles and Their Photochemical Modification. *Can. J. Chem.* **2015**, *93* (1), 98–105. https://doi.org/10.1139/cjc-2014-0287.

(172) Dilly, S. J.; Clark, A. J.; Marsh, A.; Mitchell, D. A.; Cain, R.; Fishwick, C. W. G.; Taylor, P. C. A Chemical Genomics Approach to Drug Reprofiling in Oncology: Antipsychotic Drug Risperidone as a Potential Adenocarcinoma Treatment. *Cancer Lett.* **2017**, *393*, 16–21. https://doi.org/10.1016/j.canlet.2017.01.042.

(173) Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The Re-Emergence of Natural Products for Drug Discovery in the Genomics Era. *Nat. Rev. drug Discov.* **2015**, *14* (2), 111–129. https://doi.org/https://doi.org/10.1038/nrd4510.

(174) Wessel, D.; Flügge, U. I. A Method for the Quantitative Recovery of Protein in Dilute Solution in the Presence of Detergents and Lipids. *Anal. Biochem.* **1984**, *138* (1), 141–143. https://doi.org/10.1016/0003-2697(84)90782-6.

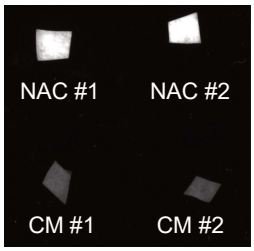
(175) Callesen, A. K.; Mohammed, S.; Bunkenborg, J.; Kruse, T. A.; Cold, S.; Mogensen, O.; DePont Christensen, R.; Vach, W.; Jørgensen, P. E.; Jensen, O. N. Serum Protein Profiling by Miniaturized Solid-Phase Extraction and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19* (12), 1578–1586. https://doi.org/10.1002/rcm.1960.

(176) Scofield, J. H. Hartree-Slater Subshell Photoionization Cross-Sections at 1254 and 1487 EV. *J. Electron Spectros. Relat. Phenomena* **1976**, *8* (2), 129–137. https://doi.org/10.1016/0368-2048(76)80015-1.

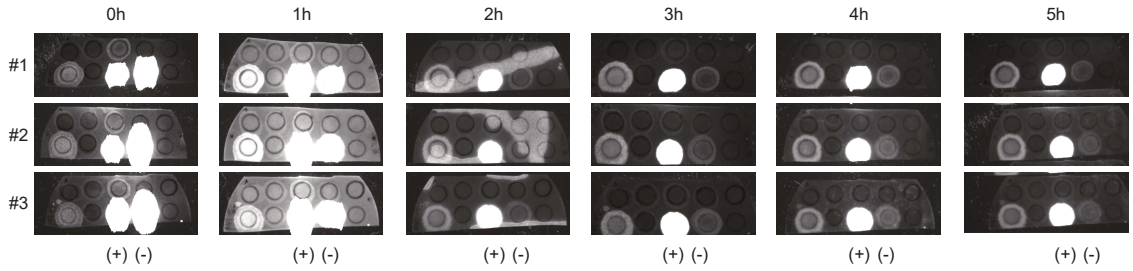
(177) Sander, T.; Freyss, J.; Von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program for Chemistry Aware Data Visualization and Analysis. *J. Chem. Inf. Model.* **2015**, *55* (2), 460–473. https://doi.org/10.1021/ci500588j.

(178) Wang, R.-X.; Fu, Y.; Lai, L. A New Method for Calculating Partition Coefficients of Organic Compounds. *Acta Phys. - Chim. Sin.* **1997**, *13* (01), 615–621. https://doi.org/10.3866/pku.whxb19970101.

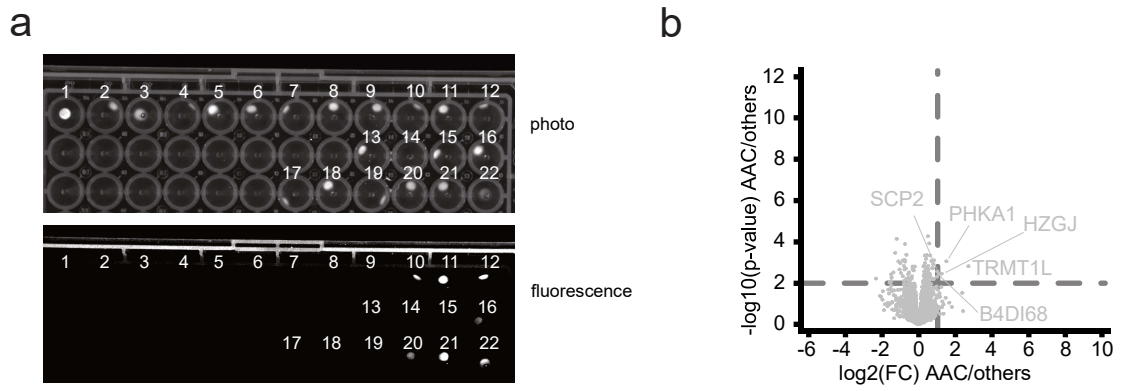
7. Supporting Material



Supporting Figure 1. Verification of NHS-activation on cellulose membranes after 3 months of storage and using 6-aminofluorescein. Duplicates of unmodified cellulose membrane (CM) pieces and NHS-activated cellulose (NAC) membrane pieces after incubation with 6-aminoflourescein for 30 minutes and washing with ethanol. Fluorescence appearing as a white signal.



Supporting Figure 2. Fluorescence intensity of implemented photocrosslinking control with fluorescein spotted onto triplicates of active TFC membranes (+) and by UV light inactivated TFC membranes (-) before (0h) and after washing in ethanol (1h), dimethylformamide (2h), tetrahydrofurane (3h), ethanol (4h) and water (5h) for one hour each.



Supporting Figure 3. Results of controls used in a click chemistry variation of CISCM. (a) Azide-cellulose spots after incubation at click reaction conditions in triplicates with photocrosslinking constructs of R-thalidomide (1-3), S-thalidomide (4-6), CsA (7-9, 13) and fluorescein (10-12), with non-photocrosslinked and unmodified fluorescein (14-16), with an empty control (17-19) and with 6-FAM-alkyne observed without (photo) and with (fluorescence) irradiation at absorption wavelength of fluorescein. **(b)** Volcano plot displaying the AP replicates of an azide-activated cellulose (AAC) spot against all other affinity enrichments.

Supporting Table 1. Properties of cherry-picked 46 compounds. Investigational: inv.

Name	MW	application	FDA-approval
Ganetespib	364.404	HSP90 inhibitor	inv.
Luminespib	451.521	HSP90 inhibitor	inv.
Tanespimycin (17-AAG)	452.593	HSP90 inhibitor	inv.
17-DMAG	453.625	HSP90 inhibitor	inv.
Travoprost	500.552	antihypertensive	approved
Harringtonine	531.6	unknown	inv.
Romidepsin (FK228)	540.704	anticancer	approved, inv.
Homoharringtonine	545.627	anticancer	approved, inv.
Bosentan Hydrate	551.622	antihypertensive	approved, inv.
Midostaurin (PKC412)	570.647	anticancer	approved, inv.
Zafirlukast	575.684	anticancer	approved, inv.
Irinotecan HCl Trihydrate	586.687	anticancer	approved, inv.
Etoposide	588.56	anticancer	approved
Berbamine	608.733	unknown	inv.
Indinavir Sulfate	613.8	antiviral	approved
Tetrandrine	622.759	unknown	inv.
Teniposide	656.659	anticancer	approved
Clarithromycin (Biaxin, Klacid)	747.96	antibiotic	approved
Azithromycin	748.992	antibiotic	approved
Vindesine sulfate	753.938	anticancer	approved, inv.
Vinorelbine Tartrate	778.944	anticancer	approved, inv.
Tacrolimus (FK506)	804.027	immunosuppressive	approved, inv.
Docetaxel Trihydrate	807.887	anticancer	approved, inv.
Pimecrolimus	810.462	immunosuppressive	approved, inv.
Vinblastine sulfate	810.986	anticancer	approved
Vincristine Sulfate	824.969	anticancer	approved, inv.
Rifabutin	847.015	antibiotic	approved, inv.
Erythromycin Ethylsuccinate	862.06	antibiotic	approved, inv.
Venetoclax	868.453	anticancer	approved, inv.
Ivermectin	875.102	antiparasitics	approved, inv.
Rapamycin (Sirolimus)	914.182	immunosuppressive	approved, inv.
Everolimus (RAD001)	958.235	immunosuppressive	approved
Zotarolimus (ABT-578)	966.222	unknown	inv.
Deforolimus (MK-8669)	990.217	immunosuppressive	inv.

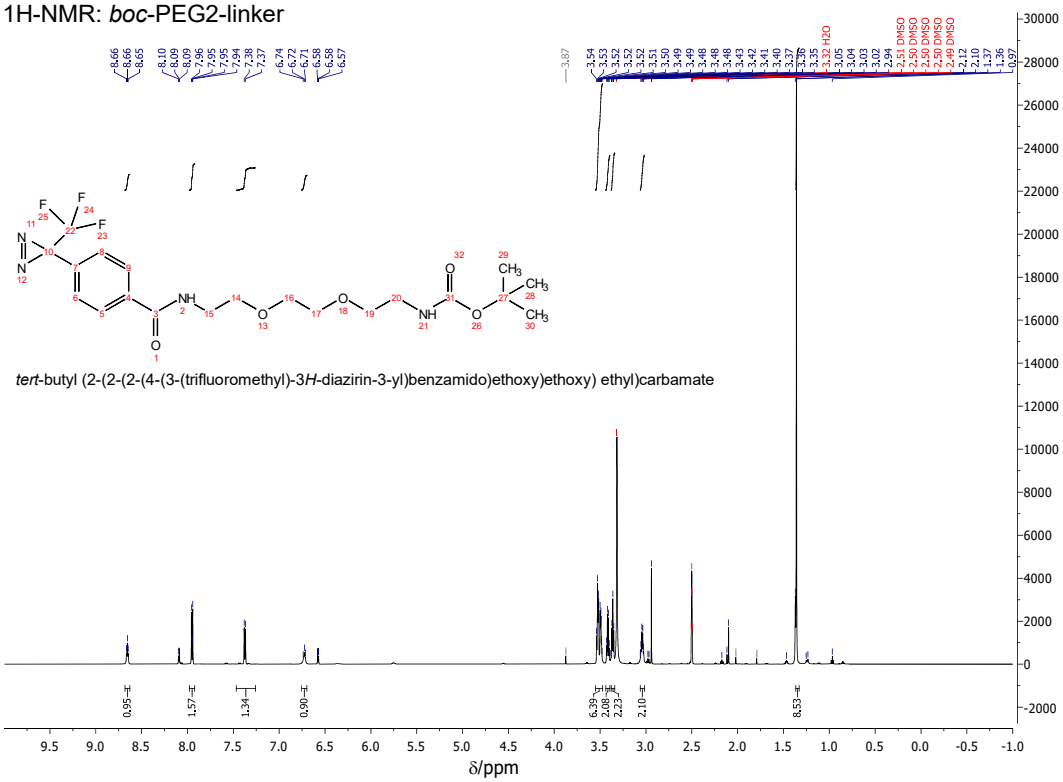
Temsirolimus	1030.3	immunosuppressive	approved
Pneumocandin B0	1065.22	unknown	inv.
Desmopressin Acetate	1069.23	unknown	approved
Colistin Sulfate	1155.45	antibiotic	approved
Cyclosporine A	1202.63	immunosuppressive	approved, inv.
Leuprorelin Acetate	1209.42	anticancer	approved, inv.
Terlipressin Acetate	1227.39	unknown	approved, inv.
Goserelin Acetate	1269.43	anticancer	approved
Icatibant Acetate	1304.54	antiinflammatory	approved, inv.
Bleomycin sulfate	1415.57	antibiotic	approved, inv.
Somatostatin Acetate	1637.9	anticancer	approved, inv.
Carperitide Acetate	3080.48	unknown	inv.

Supporting Table 2. Values of similarity analysis performed in DataWarrior for library of cherry-picked 46 compounds. N.: Neighbour

Name	N. Similarity FragFp 80%	N. Count	Neighbour Analysis X	N. Analysis Y
Ganetespib		0	0.54486877	0.7639063
Luminespib		0	0.6369065	-0.2484782
Tanespimycin (17-AAG)		0	-0.7664833	-0.49945953
17-DMAG		0	0.7032533	0.70325327
Travoprost		0	0.08998789	0.21505815
Harringtonine	0.9975 (max of 4)	4	0.21455482	-0.5188291
Romidepsin		0	-0.019804128	-0.8005699
Homoharringtonine	0.9975 (max of 5)	5	0.6011338	-0.705219
Bosentan Hydrate		0	0.39409122	-0.8123481
Midostaurin (PKC412)		0	-0.6864118	-0.6888331
Zafirlukast		0	-0.53869116	0.33464268
Irinotecan HCl Trihydrate		0	-0.6954823	0.6954823
Etoposide	0.95184	1	-0.39733782	-0.33743408
Berbamine	0.9887	1	-0.10144064	-0.024069957
Indinavir Sulfate	0.79692	1	0.8073966	0.34619337
Tetrandrine	0.9887	1	0.09855491	-0.2576313
Teniposide	0.95184	1	-0.65215755	-0.32756558
Clarithromycin	0.98592 (max of 2)	2	-0.5348252	0.08625671
Azithromycin	0.92256 (max of 2)	2	-0.36334506	-0.09635717
Vindesine sulfate	0.97045 (max of 4)	4	0.19778377	-0.73278135
Vinorelbine Tartrate	0.9684 (max of 5)	5	0.4197353	-0.61495537
Tacrolimus (FK506)	0.99425 (max of 6)	6	0.07820175	0.61181635
Docetaxel Trihydrate	0.80672	1	-0.29811615	0.401406

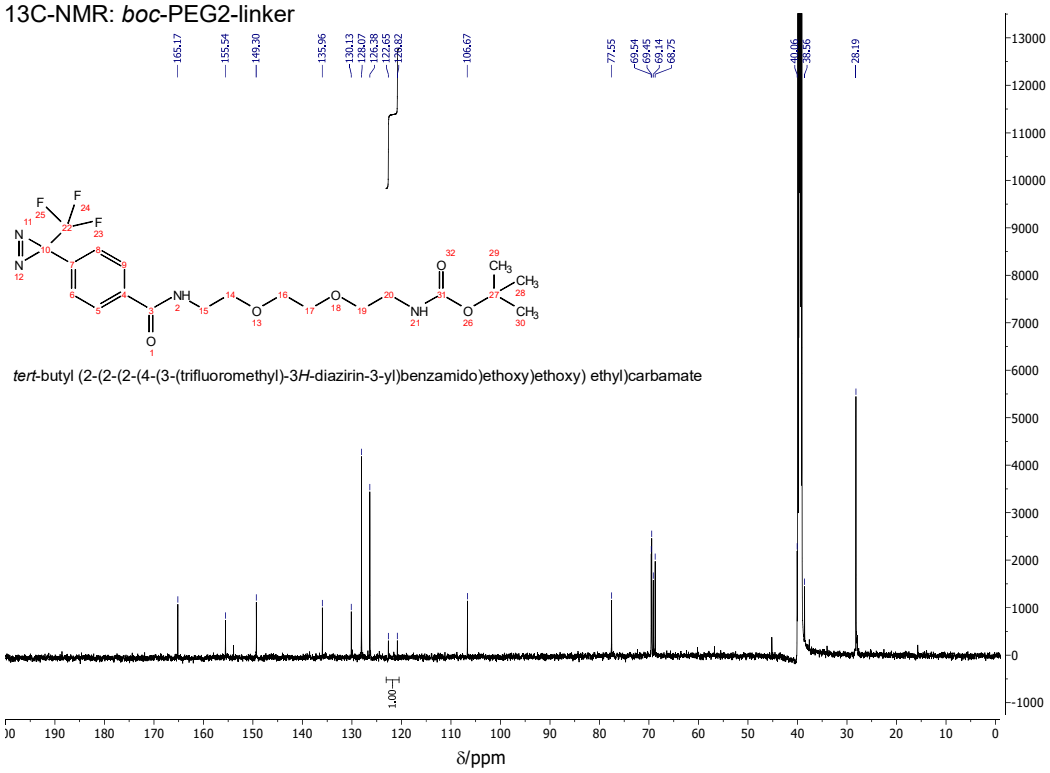
Pimecrolimus	0.99425 (max of 6)	6	0.34005412	0.7964692
Vinblastine sulfate	0.99093 (max of 5)	5	0.60964906	-0.47555658
Vincristine Sulfate	0.99093 (max of 5)	5	0.43688837	-0.3822108
Rifabutin		0	0.5564605	0.55670184
Erythromycin Ethylsuccinate	0.98592 (max of 2)	2	-0.300501	0.17106831
Venetoclax		0	0.41332817	-0.11807809
Ivermectin	0.80672 (max of 4)	4	-0.04650675	0.40109923
Rapamycin (Sirolimus)	1 (max of 7)	7	0.33635965	0.63065976
Everolimus (RAD001)	1 (max of 7)	7	0.21760438	0.42533424
Zotarolimus (ABT-578)	0.96409 (max of 6)	6	-0.17906275	0.64310867
Deforolimus (MK-8669)	0.98011 (max of 6)	6	-0.12221743	0.8316157
Temsirolimus	0.98571 (max of 7)	7	0.11578498	0.81937563
Pneumocandin B0	0.83234 (max of 4)	4	0.4337088	0.1357145
Desmopressin Acetate	0.96013 (max of 4)	4	0.65108436	-0.001473501
Colistin Sulfate	0.85616	1	-0.15772933	-0.27008805
Cyclosporine A	0.85616 (max of 2)	2	0.17792627	-0.0308373
Leuprorelin Acetate	0.93496 (max of 2)	2	0.8439436	-0.12657048
Terlipressin Acetate	0.96013 (max of 3)	3	0.80648047	0.11239565
Goserelin Acetate	0.93496	1	0.8149214	-0.35969618
Icatibant Acetate	0.79692 (max of 2)	2	0.44806662	0.39740655
Bleomycin sulfate		0	0.77284914	0.5521577
Somatostatin Acetate		0	0.77464503	-0.5690678
Carperitide Acetate	0.82831 (max of 2)	2	0.65311664	0.25013816

1H-NMR: *boc*-PEG2-linker



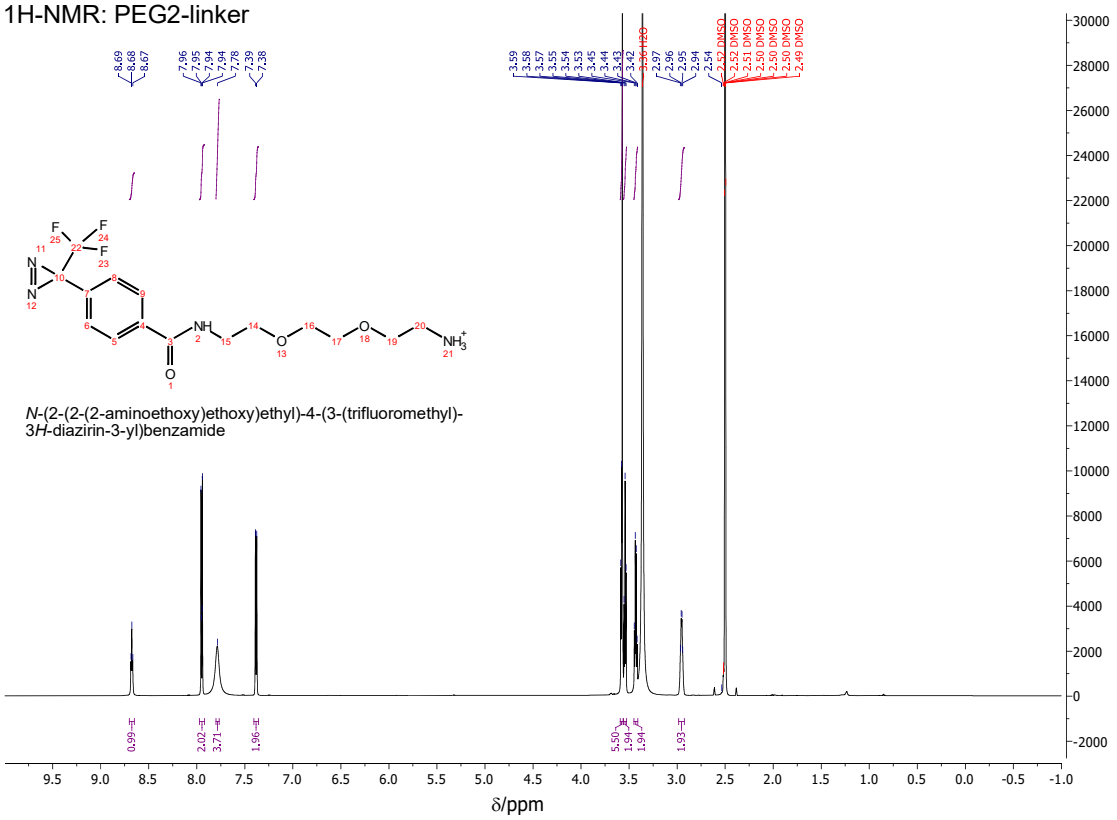
tert-butyl (2-(2-(2-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamido)ethoxy)ethoxy) ethyl)carbamate

13C-NMR: *boc*-PEG2-linker



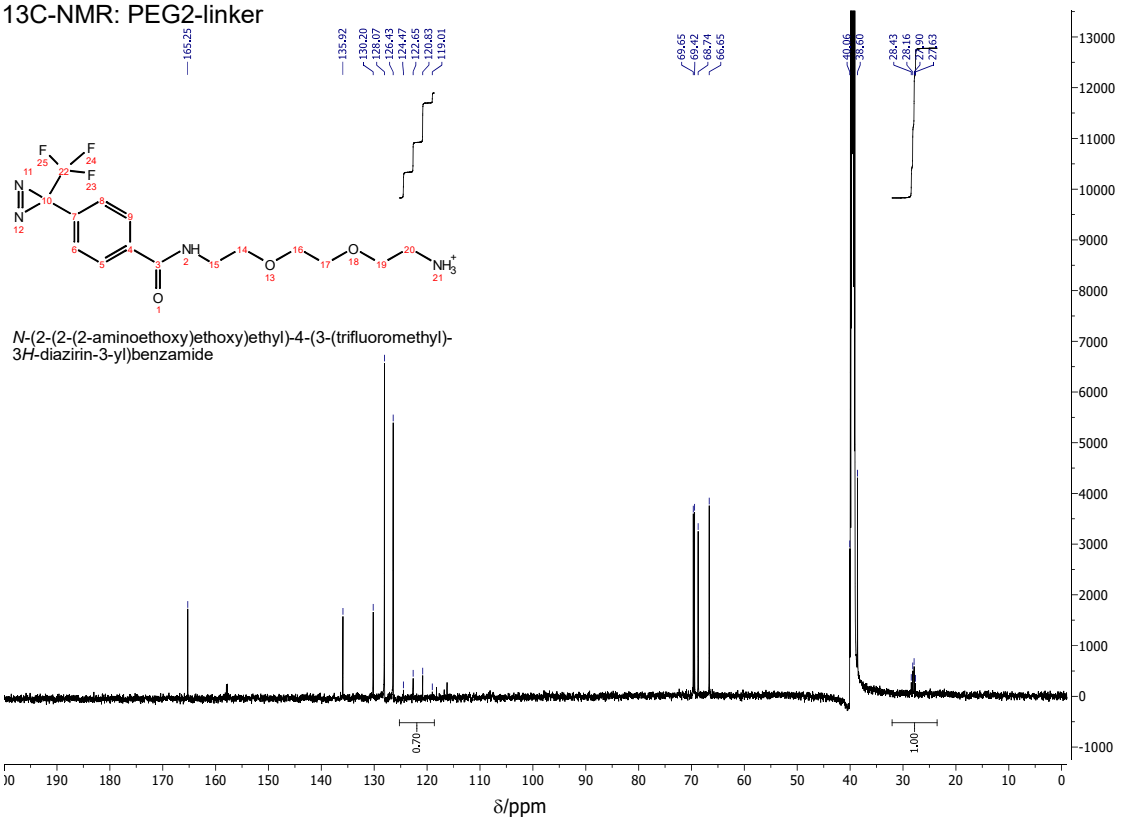
tert-butyl (2-(2-(2-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamido)ethoxy)ethoxy) ethyl)carbamate

1H-NMR: PEG2-linker



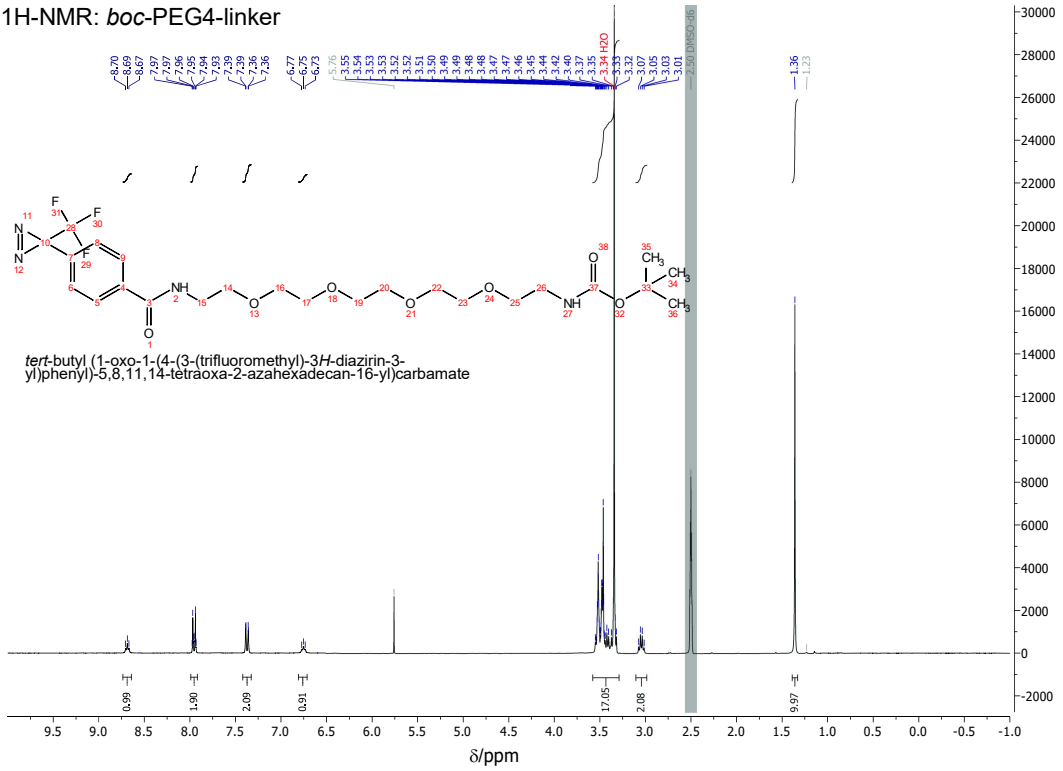
N-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide

13C-NMR: PEG2-linker



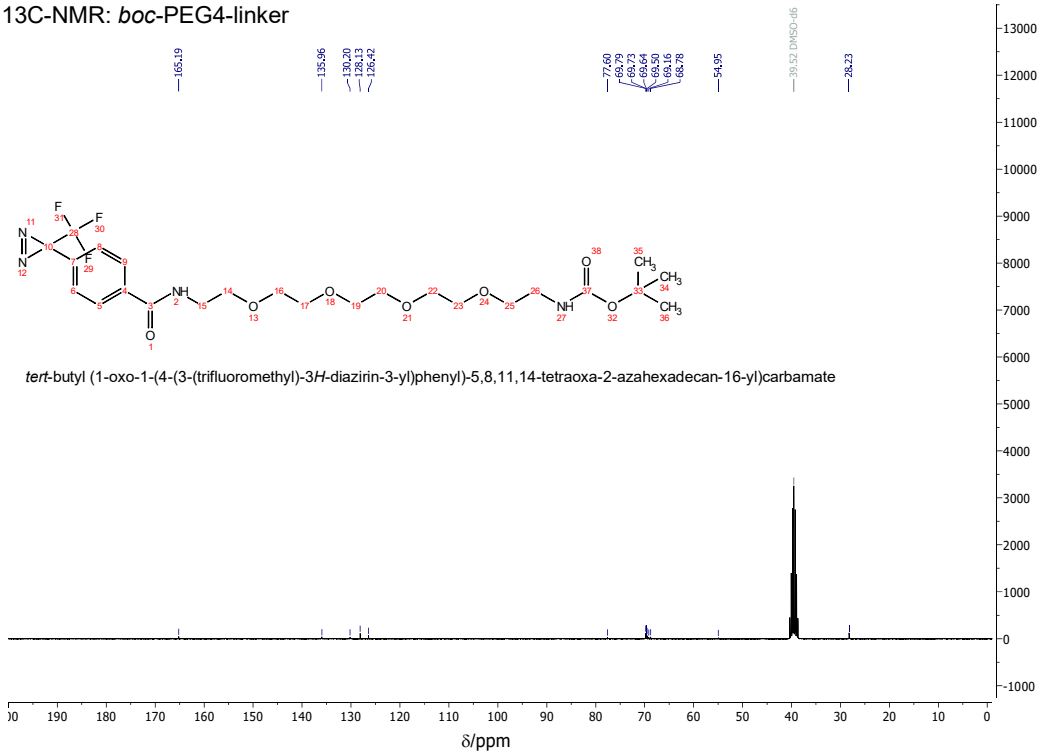
N-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide

1H-NMR: *boc*-PEG4-linker



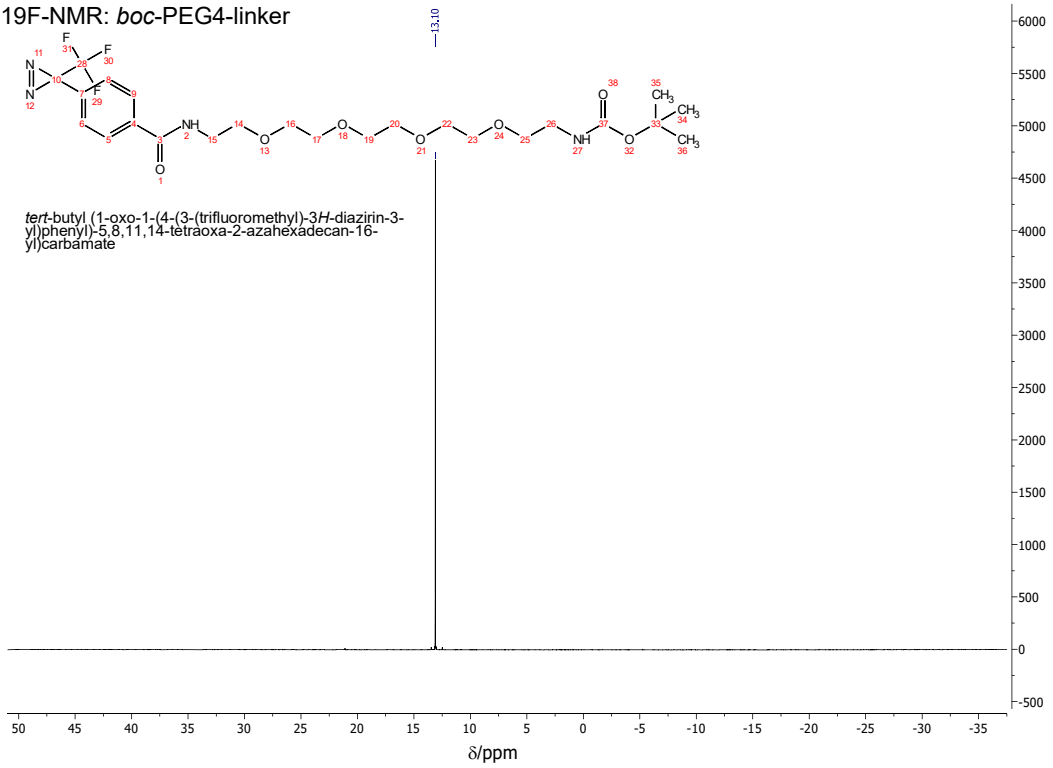
tert-butyl (1-oxo-1-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8,11,14-tetraoxa-2-azahexadecan-16-yl)carbamate

13C-NMR: *boc*-PEG4-linker

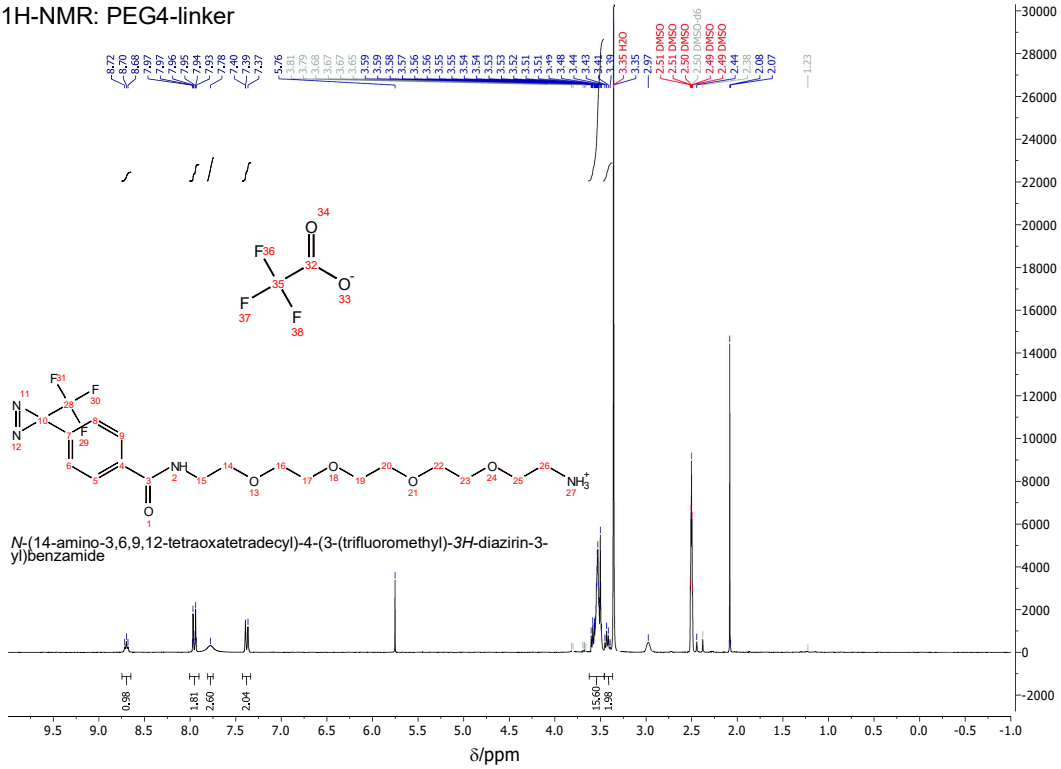


tert-butyl (1-oxo-1-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8,11,14-tetraoxa-2-azahexadecan-16-yl)carbamate

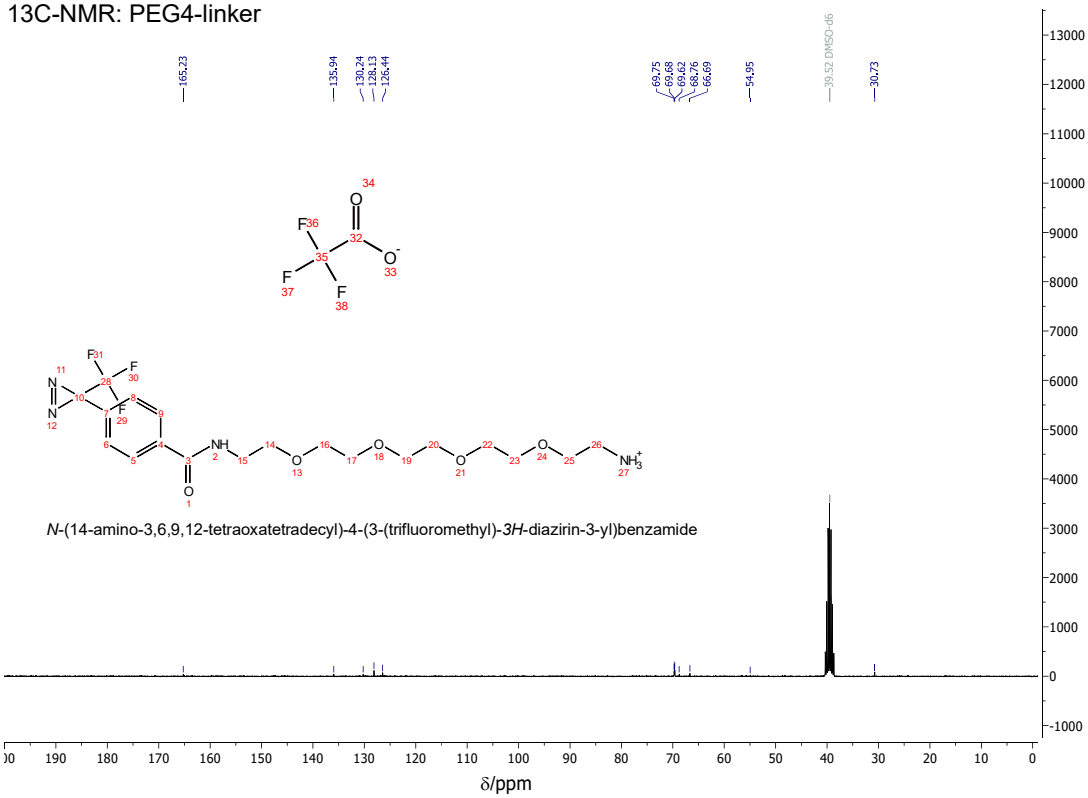
19F-NMR: *boc*-PEG4-linker



1H-NMR: PEG4-linker

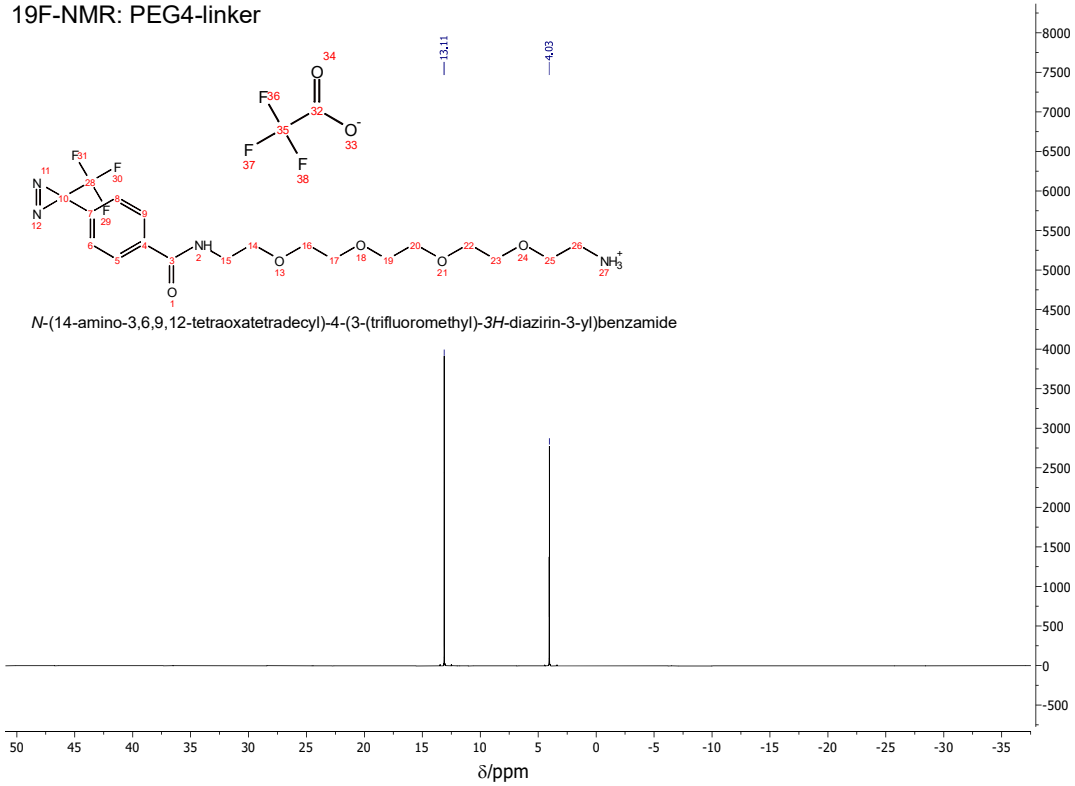


13C-NMR: PEG4-linker



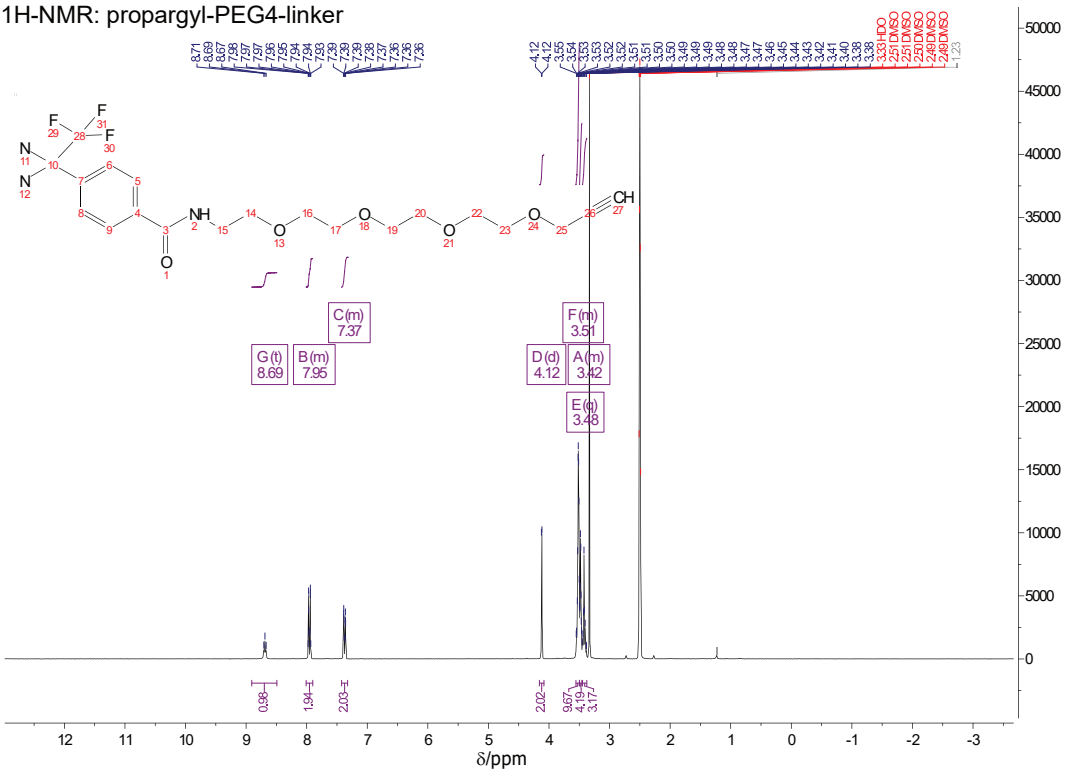
N-(14-amino-3,6,9,12-tetraoxatetradecyl)-4-(3-(trifluoromethyl)-*3H*-diazirin-3-yl)benzamide

19F-NMR: PEG4-linker



N-(14-amino-3,6,9,12-tetraoxatetradecyl)-4-(3-(trifluoromethyl)-*3H*-diazirin-3-yl)benzamide

1H-NMR: propargyl-PEG4-linker



13C-NMR: propargyl-PEG4-linker