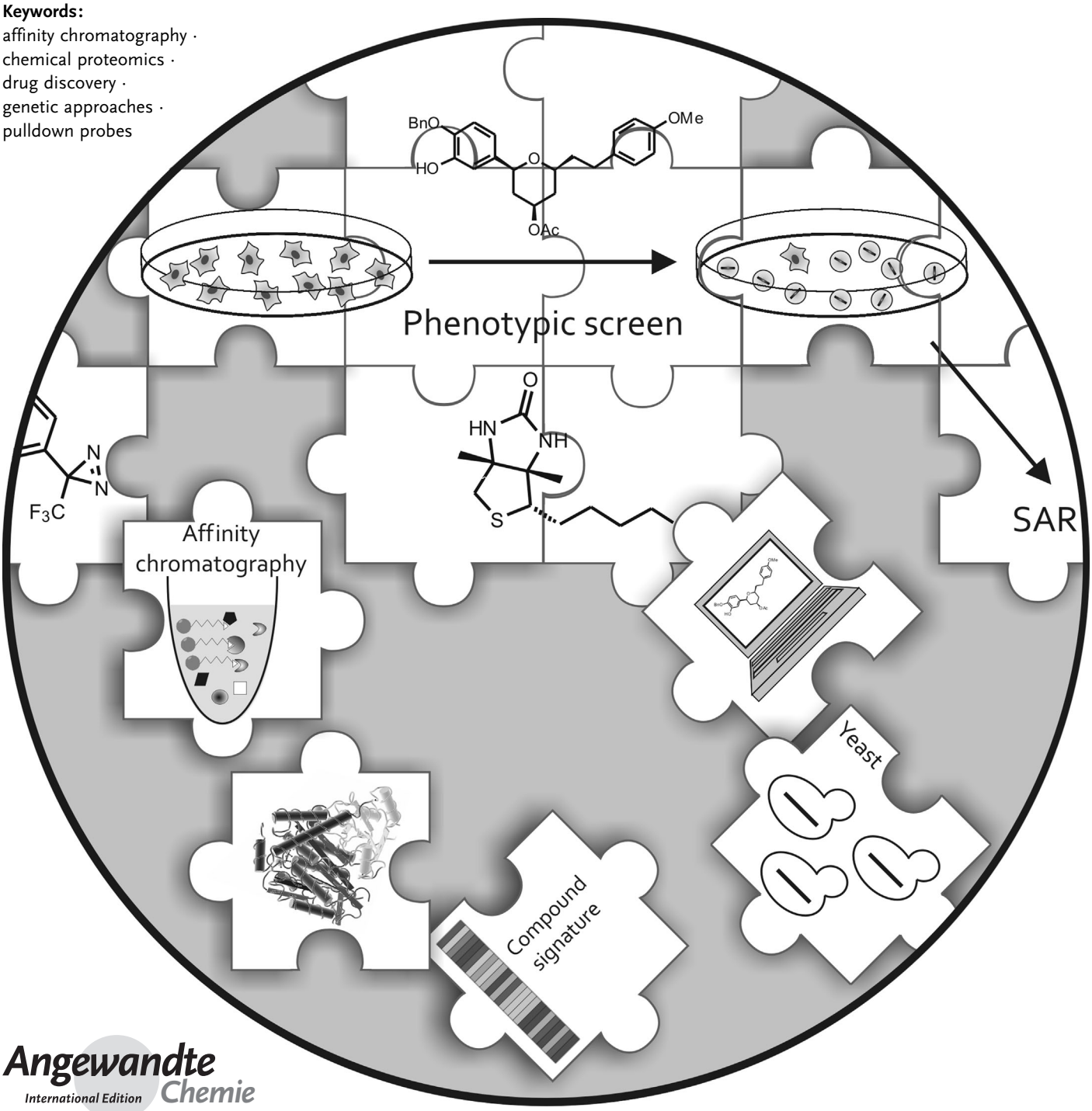
DOI: 10.1002/anie.201208749

Drug Discovery

Target Identification for Small Bioactive Molecules: Finding the Needle in the Haystack

Slava Ziegler, Verena Pries, Christian Hedberg,* and Herbert Waldmann**

Keywords:
affinity chromatography · chemical proteomics · drug discovery · genetic approaches · pulldown probes



Identification and confirmation of bioactive small-molecule targets is a crucial, often decisive step both in academic and pharmaceutical research. Through the development and availability of several new experimental techniques, target identification is, in principle, feasible, and the number of successful examples steadily grows. However, a generic methodology that can successfully be applied in the majority of the cases has not yet been established. Herein we summarize current methods for target identification of small molecules, primarily for a chemistry audience but also the biological community, for example, the chemist or biologist attempting to identify the target of a given bioactive compound. We describe the most frequently employed experimental approaches for target identification and provide several representative examples illustrating the state-of-the-art. Among the techniques currently available, protein affinity isolation using suitable small-molecule probes (pulldown) and subsequent mass spectrometric analysis of the isolated proteins appears to be most powerful and most frequently applied. To provide guidance for rapid entry into the field and based on our own experience we propose a typical workflow for target identification, which centers on the application of chemical proteomics as the key step to generate hypotheses for potential target proteins.

From the Contents

1. Introduction

The use of bioactive small molecules for the treatment of disease is an integral part of human culture. In past centuries these drugs often were applied either as single molecular entities or even as mixtures without detailed knowledge of their mode of action. For example, the active components of plant extracts used to treat various diseases often were unraveled only after decades and centuries. Since the 19th century, major advances in organic synthesis methodology and in compound isolation, including the ability to isolate natural products from complex mixtures, increasingly give access to potentially bioactive small molecules. This accessibility and the possibility to modify their structure through the power of organic synthesis have greatly promoted their use as research tools to elucidate biological phenomena, and in extension, to inspire medicinal chemistry and drug discovery. The use of small molecules offers several opportunities usually not shared by classical genetics methods: They act rapidly such that the consequences of treatment with them often can be recorded after only minutes or seconds. Their effects are tunable, for example, different concentrations of the compound or duration of treatment may give rise to graded phenotypes. Their influence is reversible and provides temporal control of protein function because the small molecules usually can be washed out or are metabolized. Small molecules can be used conditionally, that is, at any point in the time chosen for application. Thereby they allow proteins to be targeted whose mutation, knock-out, or knock-down would be lethal which may be of particular importance in the analysis of developmental processes. Notably, small molecules modulate but do not change biological systems. Whereas a gene knock-out will completely remove the protein and thereby change a given biological system, small molecules will not. In addition they may modulate only one or selected functions of the protein in question. As a result of these properties, the use of small molecules in the study of biology may offer alternatives to established biological methodology or provide entirely different opportunities to gain new insight. Consequently, the identification and synthesis of bioactive small molecules that specifically target proteins in the context of the cell or an organism are at the heart of chemical biology research.

However, while current genetics methods basically allow each gene of interest to be targeted, its chemical-biological counterpart (chemical genetics) is far from its ultimate goal of identifying a chemical modulator for each gene product of interest and their possibly different functions. By analogy to classical genetics, in chemical genetics a forward (or phenotype-based) and reverse (or target-based) approach can be employed.[1] Initially in reverse chemical genetics, modulators of given proteins are identified and then employed to study protein function in cellulo or in vivo. This approach was successfully applied for many protein classes, for example,

[*] Dr. S. Ziegler, Dr. V. Pries, Dr. C. Hedberg, Prof. Dr. H. Waldmann
Max-Planck-Institut für molekulare Physiologie
Abt. Chemische Biologie
Otto-Hahn-Strasse 11, 44227 Dortmund (Germany)
and
Technische Universität Dortmund, Fakultät Chemie, Lehrbereich Chemische Biologie
Otto-Hahn-Strasse 6, 44227 Dortmund (Germany)
E-mail: slava.ziegler@mpi-dortmund.mpg.de
christian.hedberg@mpi-dortmund.mpg.de
herbert.waldmann@mpi-dortmund.mpg.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201208749.

enzymes and receptors and for protein–protein interactions.[2] In forward chemical genetics first the modulation of living systems (cells or organisms) by compounds is recorded for example, by means of a reporter-gene activity, a fluorescence signal, or phenotypic changes detected by means of imaging methods, and only subsequently are the biological targets identified. Whereas reverse chemical genetics offers the opportunity to identify compounds with known protein targets of interest, the forward approach usually does not provide direct information about the target(s). Instead it calls for the development of powerful methods for the identification of one or a few cellular targets from the complex mixture of biomolecules present in cells. The human proteome comprises of approximately 20 000 proteins,[3] and in addition, nucleic acids, carbohydrates, or lipids may be the targets of bioactive small molecules. Thus, the search for cellular targets may well be regarded as a quest for a needle in a giant haystack.

In recent decades much progress was made in establishing methods for the identification of the cellular targets of small molecules. However, a generally applicable methodology that would allow a generic workflow to be defined that can successfully be applied in the majority of the cases—similar to workflows in the separation of individual small molecules (e.g. natural products) from complex mixtures—has not yet been established. This deficit may be due to the different chemical nature of small bioactive molecules (e.g. heterocycles vs. carbohydrates and carbocycles) and their varying affinity for targets, the different nature of the target proteins (e.g. membrane-bound vs. cytosolic proteins), and their varying abundance, and, in addition, a result of analytical methods not being sufficiently developed (e.g. very recent developments in mass spectrometry have significantly advanced target identification methodology). Because of this lack of methodology, in many cases target identification is the most challenging step and the most limiting bottleneck in a chemical biology research program. However, despite this limitation, phenotype-based approaches are increasingly being followed in academia and the pharmaceutical industry.[4] Cell-based screens monitor the influence of a compound on a complex living system in its entirety. Compounds identified in such assays have the proven ability to modulate such complex systems in the desired manner, a property not necessarily shared by hits obtained from target-based screens which need to be confirmed in cellular assays first.

Therefore, chemical–biology programs initiated with cell-based searches, for example, unbiased phenotypic screens promise to identify small molecules with various properties already encoded, most notably biological relevance[5] and the ability to modulate the complex biological system in the desired predefined manner. One of the main characteristics of target-based approaches is the focus on known (drug) targets. Sequencing of the human genome has revealed all the proteins coded. To date, only approximately 2% of all the predicted proteins have been targeted with small molecules and the estimated fraction of "druggable" proteins is approximately 15%.[6] Drug targets addressed today mostly represent selected protein classes, such as kinases, proteases, G-protein coupled receptors (GPCRs), ion channels, and receptors. There is an increasing demand for novel protein

Slava Ziegler studied biochemistry and obtained her PhD in 2004 under the guidance of Dr. O. Müller at the Max Planck Institute of Molecular Physiology on tumour genetics and Wnt signalling. Then she joined the group of Dr. R. Engers at the Institute of Pathology at the University of Düsseldorf and studied the mechanisms of tumour invasion and metastasis. Since July 2009 she has been leading a project group in the department Chemical Biology of Prof. H. Waldmann with focus on cell-based assays and identification of the targets of bioactive small molecules.

Christian Hedberg studied Organic chemistry at Uppsala University (1995–2000) and completed his PhD (2005) with Pher G. Andersson as part of the SSF program "SELCHEM". After a Postdoc with Herbert Waldmann (2005–2007) at the Max-Planck Institute for Molecular Physiology, Dortmund, he became a principle investigator there. His research interests include the development of small-molecule regulators of various biologically important proteins as well as investigations of rare posttranslational modifications.

Verena Pries studied Biology at the Ruhr-University Bochum, Germany, where she received her M.Sc. in 2007. She then joined the Chemical Biology group of Prof. H. Waldmann at the Max-Planck-Institute of Molecular Physiology in Dortmund, Germany. There she focused on target identification and validation of natural product inspired small molecules and obtained her PhD in 2011. Currently she is pursuing postdoctoral studies in the department for Developmental & Molecular Pathways at the Novartis Institutes for BioMedical Research in Basel, Switzerland.

Herbert Waldmann obtained his PhD in organic chemistry (1985) at the University of Mainz with Prof. H. Kunz after which he carried out postdoctoral research with Prof. G. Whitesides at Harvard University. He was appointed Professor of Organic Chemistry at the University of Bonn (1991), full Professor at the University of Karlsruhe (1993), and Director at the MPI of Molecular Physiology Dortmund and Professor at the University of Dortmund (1999). His research interests lie in the analysis of natural product structure for biology oriented synthesis (BIOS), the synthesis of lipidated peptides and proteins, and the development of small-molecule modulators.

targets that can be modulated by small molecules. The unbiased forward chemical genetics approach holds great promise in this regard because it identifies a desirable phenotype, bioactive hits, and potentially also novel targets simultaneously.

The identification of molecular targets is not only of interest to elucidate the mode of action of compounds, it may also be employed to identify so-called "off-targets" of known compounds and drugs, that is, additional target proteins which may give rise to undesired biological activity and toxicity. However, in contrast to this perception of undesirable "off-targets", the concept of "polypharmacology", that is, the notion that drugs may need to target several proteins simultaneously for efficacious disease modulation, as opposed to the concept of "one gene, one drug, one disease" is gaining increasing acceptance.[7] This change in concept might fuel a new wave of investigations and development in the pharmaceutical industry following this new idea.[7] The principle of polypharmacology is supported by observations from functional genomics studies in yeast, which revealed that many gene knock-outs have a minor or no phenotypic effect, for example, approximately 19% of the genes were essential, whereas the knock-out of only 15% showed an influence on yeast growth.[8] In addition, several knock-out studies in mice demonstrate that living organisms are robust,[8] a result of redundant protein functions and the existence of alternative signaling pathways. The need to overcome this robustness suggests that in drug discovery instead of searching for disease-causing genes and proteins, suitable perturbations of entire disease causing networks as a whole should be in focus. Simultaneous deletion or modulation of several genes or gene products might be required to overcome the robustness of a biological system. This concept is further supported by the existence of genetic interactions such as synthetic sickness or synthetic lethality.[9] Thus, polypharmacology may not be an undesired feature of biologically active compounds but rather may account for their activity and efficacy at reduced toxicity.[10]

By analogy to the identification of functional targets the quest for "off-targets" also calls for the development of powerful methodologies to identify the proteins bound and modulated by bioactive small molecules. To this end, in addition to experimental approaches, powerful computational methods for prediction of drug targets based on structural similarity are emerging.[7,10]

In this Review we give an overview of current methodology for target identification of small molecules. We primarily address a chemistry audience with the intention to provide a reference or an entry point, for instance for the interested chemist possibly faced with the request or inspired by the wish to identify the target of a given compound synthesized or identified in an ongoing research program, for example, a potent natural product. We first describe the most frequently applied experimental approaches for target identification and then provide several representative examples illustrating the state-of-the-art. They may serve as prototypic role model examples for successful chemical biology attempts to find the needle in the biological haystack of the cell. The Review is not meant to be comprehensive. In Table 1 (see Section 7: Appendix) and in the Supporting Information we provide a selection of bioactive small molecules, their targets and the approaches employed for their identification. For further examples the reader is referred to additional Review articles.[11]

2. Approaches to Target Identification

The currently most frequently employed approaches for target deconvolution can be divided into proteomics-, genetics-, and bioinformatics methods as well as comparative profiling-based approaches.[11a,c,12] In this Review we summarize and describe the most important of these individual approaches to target identification. In practice, techniques, such as activity-based proteome profiling (ABPP) and standard affinity pulldowns in connection with quantitative proteomics methods are far more widely used compared to elaborated biochemical and genetic techniques. Later in this Review, we discuss novel strategies to unravel the mode of action of bioactive compounds for which usually only a few examples have been described. Although these methods are discussed separately, we stress that they should be regarded as being complementary. Their combined application might generate a more complete picture of small-molecule–protein interactions in cells.[13]

2.1. Affinity Chromatography

Affinity-based proteomics ("pulldown") is one of the most widely applied methodologies to identify the targets of biologically active compounds (for a general outline of the method see Figure 1a). The pulldown probe is immobilized on a solid phase and exposed to a protein extract (e.g. cell lysate) to bind the target protein(s). Subsequently, proteins that bind non-specifically to the probe and the matrix are removed by stringent washing prior to release of the bound proteins by means of elution with the bioactive molecule or by heating. The target proteins then typically are identified by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis), tryptic digestion and identification of the resulting peptides by MS/MS (mass spectrometry/mass spectrometry) analysis.

2.1.1. Probe Synthesis

For a pulldown experiment the compound of interest needs to be modified to allow its immobilization on a matrix. To this end, a linker with a functional group or affinity function is attached to the compound such that target-protein binding is not impaired or only impaired to a tolerable extent. Hit identification is usually followed by investigation of the activity of structurally similar compounds which allow a structure–activity relationship (SAR) to be delineated and a site for attachment of the linker which tolerates substitutions and/or variation without significant loss in activity to be identified. For example, the attachment of a linker at two distinct positions in cGMP (Figure 2) led to the isolation of

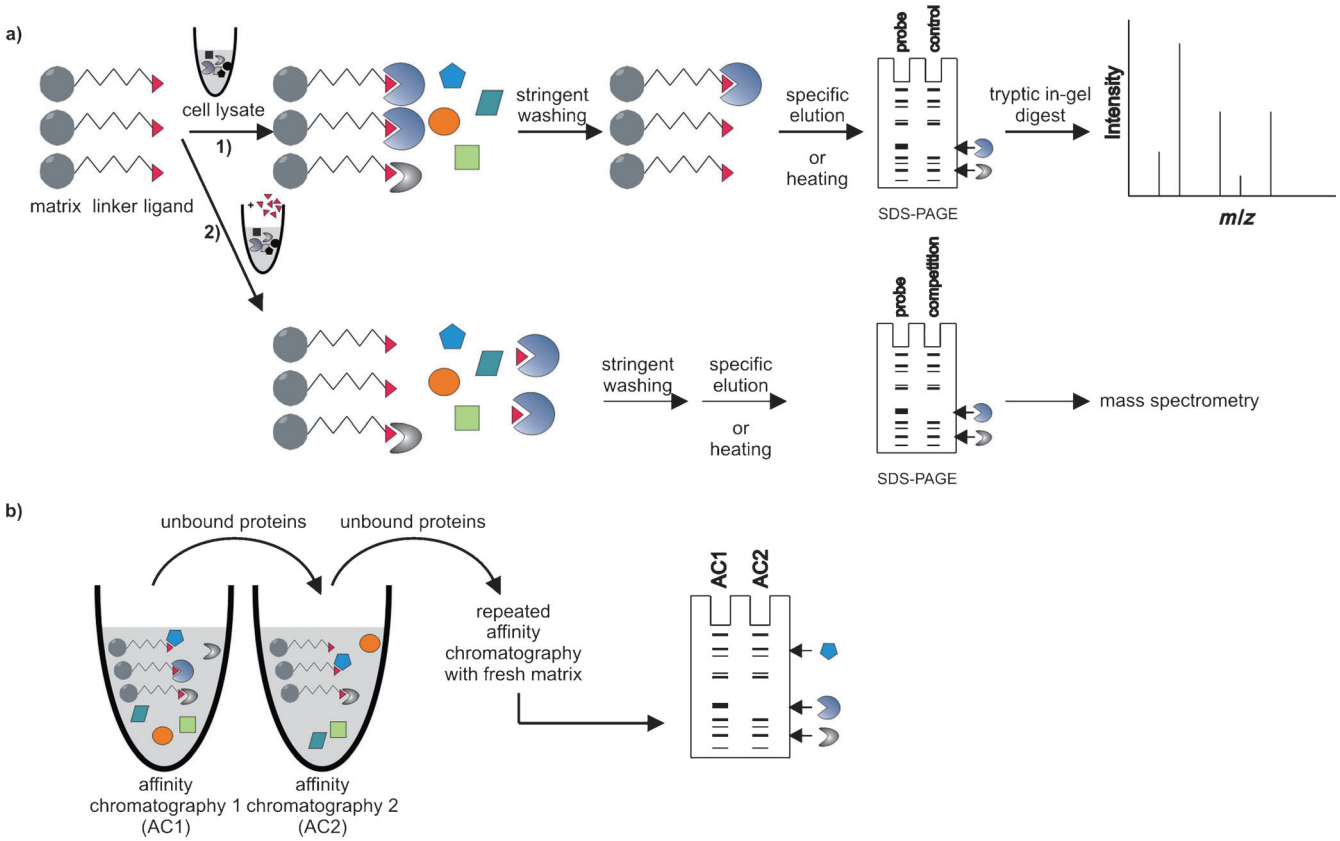


Figure 1. a) Principle of affinity-based proteomics for target identification 1) using control beads or 2) competition with unmodified compound. b) Serial affinity chromatography (see Section 2.1.8).

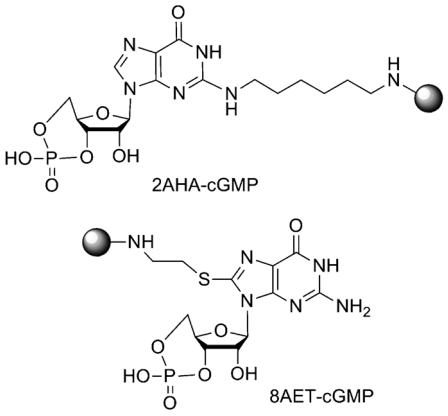


Figure 2. Structural formula of cGMP derivatives immobilized at different positions; gray sphere = bead.[14]

different proteins, which reflects the different nature and coupling positions of the compound.[14]

Frequently, for example, when working with natural products (NPs), structure–activity relationship studies are hampered by structural complexity of the hit or lack of total synthesis routes to access and modify the hit. In such cases functional groups in the hit compound may be employed to introduce modifications[15] which limit types of and sites of modification. Alternatives for target deconvolution of compounds without extensive structure–activity relationship studies are the use of immobilized photoreactive linkers which, upon irradiation, generate reactive intermediates to photo-crosslink small molecules to a solid surface (see also Section 2.1.9)[16] or label-free approaches (see Section 2.3).

2.1.2. *Linkers*

The attachment of a linker (spacer) to a small molecule to be employed in a pulldown experiment is obligatory not only due to the incorporation of a functional group required for immobilization but mainly to avoid steric interference of the matrix and the target proteins. Among the reported linkers are alkyl groups,[17] di- or triethylene glycol groups,[18] peptides,[19] the aminocaproylaminopentyloxy (ACAP) group,[20] divalent epoxides (ethylene glycol diglycidylether, EGDE),[21] and tartaric acid[22] (Figure 3a). The hydrophobic alkyl linkers may favor non-specific binding of proteins to the pulldown probe which will increase the complexity of the protein mixtures to be analyzed and thereby may hamper the detection and/or identification of the target proteins. Usually, a hydrophilic spacer is the linker of choice with di- or triethylene glycol ethers most often used. The rod-like polyproline linker was designed to overcome some of the drawbacks associated with the use of ethylene glycol linkers (shorter length and thus binding of bulky proteins might be hindered). It is more rigid, longer, and allows binding of bulky proteins.[19c] The tartaric acid linker was reported to decrease unspecific binding in comparison to oligoethylene glycol-

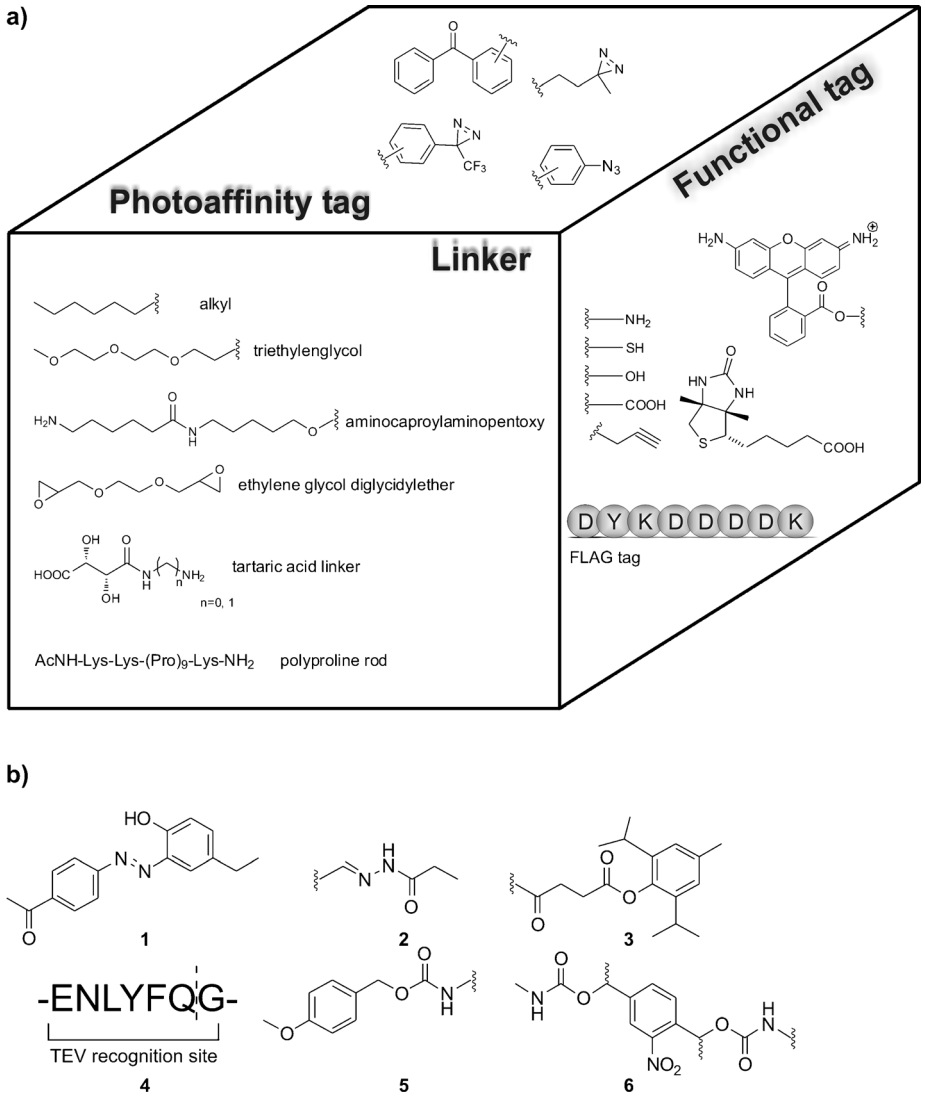


Figure 3. Toolbox for affinity probe-based target identification. a) For affinity probe-based target identification a linker group is attached to the bioactive compound of interest. This linker may additionally be equipped with a photoaffinity tag to covalently modify the target proteins and a functional tag to facilitate target protein identification. b) Groups for the design of cleavable linkers.

based probes.[22] To date, however, the polyproline and the tartaric acid linker were only employed in few target-identification studies.

In general, proteins that bind to a surface-immobilized probe are released from the resin either under harsh conditions (boiling in sodium dodecyl sulfate (SDS) sample buffer) or by elution with an excess of unmodified compound. The release of unspecifically bound proteins after boiling the samples increases the complexity of the mixtures to be analyzed by mass spectrometry (MS) and/or electrophoresis and can hamper the identification of the target proteins. Therefore, selectively cleavable linkers were designed based on diazobenzene (**1**; cleavage under reducing conditions, for example, sodium dithionite),[23] acylhydrazone (**2**; cleavage by hydrazides),[24] levulinoyl ester (**3**; cleaved by hydrazine),[25] peptide linkers (**4**; proteolytic cleavage site),[19b] acid-labile groups[26] (**5**) and the photolabile 1-(2-nitrophenyl)ethyl moiety[27] (**6**; Figure 3b). Cleavage allows the release of only those proteins that bind to the probe while proteins that unspecifically bind to the solid phase remain immobilized. The main disadvantage of thiol-based/acid-labile linkers is the requirement for a strong acid (e.g. trifluoroacetic acid, TFA) which is not compatible with a direct release of proteins from the resin since nonselective release can also occur. In addition, TFA needs to be removed prior to mass spectrometry analysis.[23a] Disulfide linkers are not compatible with buffers containing reducing reagents such as dithiothreitol (DTT). Peptide-based linkers may suffer from inefficient proteolytic cleavage and thus limited release of bound proteins.

2.1.3. *Functional Groups and Affinity Tags for Immobilization*

To ensure the immobilization of a pulldown probe either alone or in complex with bound proteins it should include either a functional group for covalent attachment to a resin, an affinity functionality, such as biotin which strongly binds to immobilized avidin, or a peptide tag or fluorophore (Figure 3a).

The pulldown probe can be covalently attached to an activated matrix (e.g. *N*-hydroxysuccinimide (NHS)- or epoxy-activated sepharose) by means of functional groups such as amines, alcohols, thiols, and carboxylic acids. Prior to exposure of the matrix-bound compound to a protein extract and binding of the target proteins, the non-modified groups are inactivated (e.g. by treatment with Tris buffer for NHS-esters) to prevent covalent binding of proteins to the beads. In many cases, incubation of the probe with cell lysate is sufficient to "fish out" the binding proteins, use of living cells is not required. However, Wirth et al. demonstrated that the pulldown probe of the natural product duocarmycin SA (Table 1, entry 1) only isolates a second target protein (aldehyde dehydrogenase, ALDH1A1) for the natural product after incubation with living cells, as opposed to treatment with cell lysates which emphasizes that lysis condition might compromise binding of the target protein to the ligand.[28]

Biotin is practical for affinity-based approaches owing to its strong interaction with the avian and amphibian protein avidin and the bacterial *Streptomyces avidinii* protein streptavidin. Furthermore, a deglycosylated version of native

avidin from white egg, termed neutravidin, is available. Neutravidin has a near-neutral isoelectric point (p*I*) and its use reduces non-specific binding.[29] Compared to avidin, streptavidin has several basic residues exchanged for neutral or acidic amino acids and is not glycosylated, which does not affect biotin binding.[30] The three proteins are homo-tetramers which bind four biotin molecules (one molecule per monomer). The affinity of the proteins to biotin is exceptionally high ($K_D \approx 10^{-14}$–10^{-15} M)[31] making it one of the strongest noncovalent interactions known. Biotin-(strept/neutr)avidin binding is widely used in many different protein and nucleic acid detection approaches, for example, in immunohistochemistry, enzyme-linked immunosorbent assays (ELISA) and protein purification. Biotin itself rarely interferes with the function of the labeled component (e.g. antibody, small molecule, nucleic acid). Furthermore, the strong interaction between avidin and biotin persists even under harsh conditions such extreme pH values and temperatures and denaturing environments, which allows its use in various biological approaches.

Affinity-based proteomics makes use of the biotin-streptavidin system to immobilize pulldown probes on a solid surface and enable the isolation of target proteins. Mostly, biotinylated compounds are immobilized to streptavidin-coated beads that are then exposed to a protein extract. However, biotinylated probes can also be first incubated with living cells to bind the target proteins prior to cell lysis and incubation with the streptavidin matrix, however, the attachment of biotin to a small molecule often reduces cell permeability. This approach might overcome the limitation of using cell lysate, since cell rupture might impair binding of a probe to its targets (see above) which may be particularly relevant for membrane proteins.[32] After the proteins are bound to the probe, the high affinity of biotin towards (strept)avidin allows stringent washing steps to remove unspecific binding proteins. The biotin moiety may also be incorporated in trifunctional probes (see Section 2.1.4) which allow the probe first to be covalently linked to the target protein (e.g. by means of photo-crosslinking) followed by isolation of the target proteins using streptavidin.[28,33] The high affinity of the biotin-streptavidin interaction requires very harsh conditions to release the bound proteins from the resin, generally, by heat-denaturation in SDS-containing buffer. An alternative is the elution of target proteins with an excess of unmodified compound. For covalent binders, the incorporation of a cleavable linker (see Section 2.1.2) offers an additional opportunity to release isolated proteins without denaturation. It should be considered that non-specific binding decreases from avidin to streptavidin with NeutrAvidin reported to give the lowest background.[29] Furthermore, endogenous biotinylated proteins might be enriched using avidin-based matrices, which will result in higher background.[34] Non-specific binding poses a serious problem in affinity-based proteomics (see Section 2.1.8). In the use of biotinylated probes the advantages and disadvantages have to be balanced. Biotinylated probes are valuable reagents also in various target confirmation approaches, for example, surface plasmon resonance (SPR), homogeneous time-resolved fluorescence (HTRF), imaging, and immuno-blotting experiments. Thus these probes may be used for target identification and target confirmation.

Instead of the biotin/avidin pair, antibodies can be applied to recognize tags in the affinity probes. The FLAG tag is a octapeptide with the sequence DYKDDDDK. The small size enables its encoding by a single oligonucleotide and the sequence was designed to ensure good antigenic properties (for the interaction with an anti-FLAG-tag antibody) while possessing maximal hydophillicity.[35] Moreover, this tag contains a protease cleavage site that allows its removal. Saxena et al. used a FLAG tag-modified probe of bisindolylmaleimide III (Bis-III, Table 1, entry 2), a derivative of a known protein kinase C (PKC) inhibitor.[36] After binding of the FLAG-tag-labeled probe to the target proteins, the complexes were isolated by means of a FLAG-tag antibody attached to an agarose resin. The isolated proteins were compared with targets identified after covalent attachment of Bis-III to a resin. The FLAG-tag-based approach identified several of the known Bis-III targets, such as PKC-α, GSK3β, CaMKII, adenosine kinase, but also some previously unidentified target proteins such as PKAC-α, prohibitin, and VDAC.[36] It is noteworthy that this approach was also compatible with release of the bound proteins by specific elution. The FLAG-tag used in this study is exchangeable with other known tags (e.g. HA,)[37] as well as with fluorophores, for which antibodies are available (e.g. 7-dimethylaminocoumarin, rhodamine, fluorescein).[38]

2.1.4. Trifunctional Probes

A combination of an affinity tag (biotin) with a fluorophore has been successfully applied for target isolation and subsequent detection of fluorescently labeled proteins in an SDS-PAGE in several cases. Prerequisite is the covalent attachment of the probe to the target proteins so that the probe remains attached to the proteins under the denaturating conditions of the SDS-PAGE. Trifunctional probes are widely used in particular for activity-based proteome profiling (ABPP).[28,33c,d,39] Such trifunctional probes (Figure 4a) usually contain an alkyne which can be coupled by copper(I)-catalyzed Huisgen [3+2] cycloaddition (Huisgen-Sharpless-Meldal click chemistry)[40] to either a biotin-azide or a fluorophore-azide (e.g. rhodamine) to allow for enrichment or visualization of the binding proteins on a gel.[33a] Coupling of both, biotin and rhodamine, to the compound containing an alkyne "handle", results in a trifunctional probe: the biotin tag facilitates enrichment of the target proteins, whereas the fluorophore allows detection of the target proteins on a gel. Fluorescent bands visible in the gel are then excised and subjected to mass spectrometry.[33b] This procedure markedly facilitates protein identification since only a few bands are analyzed instead of the entire lane or entire protein mixture. While this method is very well suited for reactive compounds, for noncovalent binders a photoaffinity group (e.g. benzophenone or alkyl-/aryl diazirine, see Figure 3a) needs to be incorporated in the probe.[33c] This strategy underlies the design and synthesis of capture compounds for capture-compound mass spectrometry (CCMS).[41] The capture compound contains a selectivity function for binding to the target

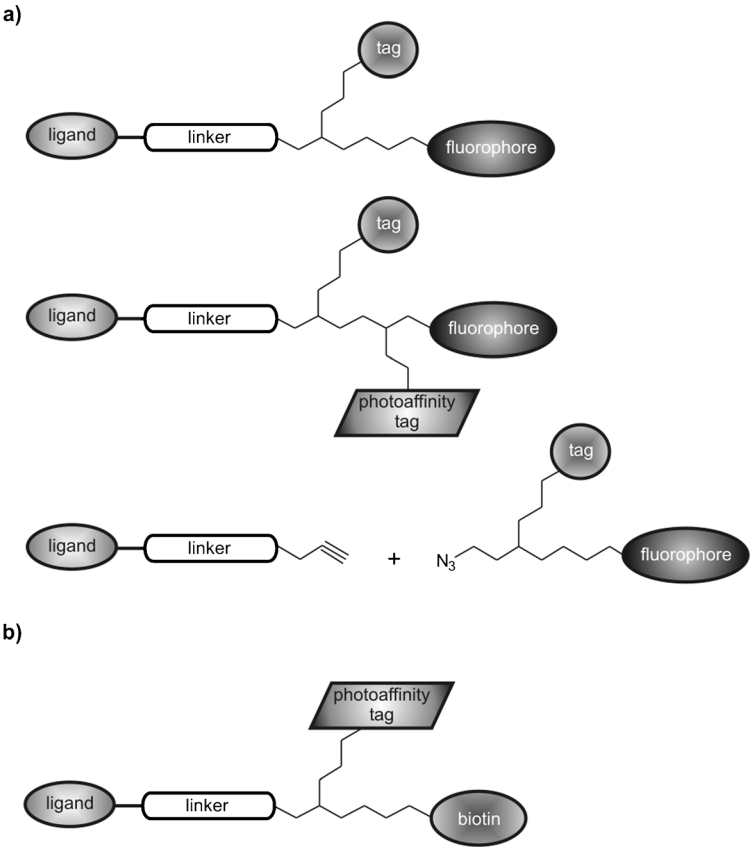


Figure 4. General structures of a) trifunctional probes and b) a probe for capture-compound mass spectrometry.

protein, a reactivity function (photoactivatable group) for covalent attachment to the target, and a sorting function (biotin) for enrichment (Figure 4b). CCMS was applied for the profiling of S-adenosyl-L-homocysteine (Table 1, entry 3) binding proteins, such as methyltransferases[42] and in the quest for off-targets of anti-Parkinson drugs.[43] Tolcapone and entacapone (Table 1, entry 4) target catechol-*O*-methyltransferase (COMT) and while tolcapone is hepatotoxic, entacapone is not. Using CCMS it was shown that besides COMT tolcapone targets proteins related to β-oxidation and oxidative phosphorylation. CCMS performed for the anticoagulant prodrug dabigatran[44] revealed that several proteins bind to the designed capture compound derived from dabigatran, among them the detoxifying oxidoreductase ribosyldihydronicotinamide dehydrogenase (NQO2).

2.1.5. Protein Extracts

For identification of target proteins by means of chemical proteomics, the pulldown probe is incubated with a protein extract usually originating from tissues or cultured cells as source of the proteome to be analyzed.[17–18] Generally, the proteome source is selected according to the activity of the compound, for example, a cell line sensitive to the compound of interest, to ensure the presence of the target proteins. Lysate preparation must not impair integrity of the target proteins and small molecule binding by excessive denaturation or residual enzymatic activity. Since mostly information on the target proteins will be missing, this issue can hardly be addressed in advance. However, if from lysate no target protein can be isolated intact cells should be considered.[28,38a,45] The use of cell lysates rather than living cells is required if the pulldown probe is membrane impermeable.

To minimize unspecific binding in affinity-based approaches lysate pre-fractionation may be considered.[46] Thereby sample complexity and thus background may be reduced, and use of a particular subcellular fraction will enrich the target protein and may provide better conditions for binding to the probe. Prerequisite to the fractionation is knowledge about the cellular localization of the binding proteins which often can be obtained by means of a fluorescently labeled compound.[47] For instance, pre-fractionation of HeLa cell extracts by means of anion exchange chromatography was carried out for the identification of ornithine δ-amino transferase (OAT) as a target of diazonamide (Table 1, entry 5).[46] Studies with BODIPY-labeled pladienolide B (Table 1, entry 6) revealed a nuclear localization of this natural product and the following target identification was performed with nuclear extracts.[48] Since substantial amounts of the protein extract are necessary, availability of the primary material could limit this approach. Nonetheless, fractionation should be considered if the target protein is, for example, membrane-associated, nuclear, or present in a particular chromatographic fraction. Care should be taken not to exclude insoluble fractions, since proteins of interest might well be bound to membrane components.

2.1.6. Protein Binding to the Affinity probe, Washing, and Elution

For binding of the pulldown probe to the target proteins the affinity compound may be immobilized on beads first (e.g. streptavidin-coated or NHS-activated matrix) and then incubated with lysate. For probes bearing a functional group for covalent immobilization on a matrix, this is the method of choice. Compounds with affinity tags (e.g. biotin or peptide tag) could be incubated with cells or protein lysate prior to immobilization on the solid phase. In particular cases immobilized compounds (e.g. nanopolymers, such as dendrimers) may also be cell permeable allowing incubation in situ to bind target proteins.[49] When employing tri-functional probes, for example, with an alkyne handle the compound is first incubated with living cells or cell lysates. In a second step, (if necessary after cell lysis) the linker that carries both an reactive tag and a fluorophore, is attached by means of copper(I)-catalyzed Huisgen [3+2] cycloaddition to the probe. Finally, the probe–protein complex is immobilized on a solid support.[28,33b]

Because of the high complexity of cell lysates, which usually contain numerous proteins, high abundant proteins might non-specifically bind either to the matrix, the linker or

the compound itself. To remove these proteins, extensive washing with increasing stringency during the washing steps is required. The degree of stringency depends on the affinity of the compound for the target proteins. Clearly, much harsher washing conditions can be applied when the compound is covalently bound to its target(s). In contrast, for noncovalent binders increasingly stringent washing steps can be controlled, for example, by comparison of the isolated proteins by means of SDS-PAGE and subsequent protein staining. To allow for enrichment of weakly bound proteins, mild washing conditions need to be found.

After removal of non-specific binders, proteins enriched on the solid phase are released either by denaturation through heating or by specific elution. Elution with denaturing agents such as SDS releases all proteins specifically bound to the pulldown probe as well as non-specifically to the linker or the bead surface. This procedure results in complex mixtures but may be the only way to liberate the isolated proteins when the compound is covalently bound to its target and the probe does not contain a cleavable linker. Cleavable linkers (see Section 2.1.2 Linkers) enable the specific release of proteins bound to ligands while leaving proteins that bind to the bead surface attached. For specific elution the bound proteins are incubated with an excess of unmodified active compound.[17] This step enables the release of proteins under mild conditions, and at the same time represents a second affinity step. Specific elution is compatible only with noncovalent binders and is limited by the solubility of the unmodified compound in aqueous buffers because usually high compound concentrations are required. For example, this approach was successfully applied to identify proteins that specifically bind to ADP, cAMP, GDP, and cGMP.[14] To this end, cGMP was immobilized by a 2-aminohexyl linker and was incubated with HEK293 lysate. Subsequently, a sequential elution with 10 mM ADP, 10 mM GDP, 5–10 mM cGMP, and 10 mM cAMP allowed isolation of the proteins that bind the respective nucleotide. Specific elution might be limited by the slow dissociation kinetics of the binding protein from its ligand.[50]

2.1.7. Control Experiments

To distinguish between proteins that bind specifically and non-specific binders control experiments are mandatory. The simplest experiment employs the same matrix material without the immobilized compound (bead only). This approach will reveal proteins that non-specifically bind to the matrix itself and was successfully employed in the identification of various target proteins.[15b,51] To identify proteins that bind non-specifically to the linker or the bioactive compound, either the linker alone[15b,18a,52] or even better an inactive analogue equipped with the same linker[18a,33b,53] needs to be immobilized on the matrix. An inactive compound should closely resemble the active compound in structure, size, hydrophobicity, and charge.[11b] Proteins that bind both to control beads and beads carrying the active probe are usually regarded as background. Although this approach proved to be powerful in several cases, it is noteworthy that this reasoning may be an oversimplification of the binding events during the pulldown experiment. Inactivity in cellular assays may be due to low solubility, low cell permeability, and low affinity for the target protein.[11c] These properties might prevent a phenotype of interest in a cell-based assay and therefore indicate lack of activity but the compound in question might still bind to the target proteins in cell lysates. The employment of such controls also may lead to identification of the same target proteins by both active and inactive probes and thereby to false negative results since these proteins will be regarded as non-specific binders. In theory, the ideal control for a chiral compound would be an inactive compound with the same physiochemical properties as the active probe but different stereochemistry, that is, an enantiomer. This strategy was successfully used for identification of nucleophosmin and Crm1 as targets of indoloquinolizines termed centrocountins (Table 1, entry 7, see also Section 5.4).[17] An enantiodifferentiating strategy for target identification was chosen also for jasmonate glucoside (Table 1, entry 8) which is a leaf-closing factor in plants and whose enantiomer is inactive.[54] A photoaffinity probe of jasmonate glucoside coupled to biotin was employed in pulldown experiments and the respective inactive enantiomer served as control. We stress, however, that it is not necessarily the case that the enantiomer of a given chiral compound does not bind at all to the same target protein, but rather its affinity may only be lower. Thus, an enantiomer may not be the ideal control compound. If the active small molecule is achiral, the control probe should be as closely related to it as possible. Furthermore, the attachment of the linker to a nonpermissive position (which does not tolerate modifications and leads to loss of activity) is another option for the design of control probes as reported for the neuritogenic compound TWS119 (Table 1, entry 9) in the identification of its target GSK3β[55] and for the anti-trypanosomal compound 4-[5-(4-phenoxyphenyl)-2*H*-pyrazol-3-yl]morpholine (Table 1, entry 10) and the identification of adenosine kinase as its target.[56]

An essential control strategy employs competition with binding of the protein to the immobilized ligand.[57] In this approach an excess of the bioactive compound is added to the protein mixture before pulldown. It binds to the target and thus prevents binding of the target protein to the (immobilized) affinity probe. Proteins identified in the standard pulldown but absent in the competition sample are considered as targets. The shortcomings of the competition approach are again the hydrophobicity and thus limited solubility of compounds in aqueous buffer. Competition studies should be combined with inspection of the isolated proteins on a SDS-PAGE, where the absence of the target proteins in the competition sample should be visible. To determine the most suitable conditions for competition, different concentrations of the unmodified compound should be applied.

Usually, proteins released from the solid support are separated by SDS-PAGE according to their molecular weight. After electrophoresis, proteins are stained using different methods. The widely used Coomassie Brillant Blue can detect proteins down to 10 ng and is compatible with mass spectrometry analysis.[58] More sensitive is the silver stain and staining with fluorescent dyes (detection of less than 1 ng

protein).[58] Whereas silver staining is not compatible with mass spectrometry without an additional destaining-step, many fluorescent dyes are, so that after protein visualization the bands of interest can be cut out from the gel and further processed. Sensitive protein staining allows the target proteins to be visualized on the gel, so that initial information is obtained, for example, the apparent molecular weight. By utilizing appropriate control probes, differences in the isolated proteins (amount, presence) can be detected after protein staining. For example, a protein band which is present on the gel only with the active but not with the inactive probe will be considered as potential target. In a competition experiment, the disappearance of a protein band (or decreasing amounts of it) in the presence of the competitor will be indicative of the target. For a competition experiment, protein staining after SDS-PAGE might help to find the optimal conditions prior to mass spectrometry analysis. It should be considered that even in the presence of a high excess of unmodified active compound, not all of the target protein will be completely released. Thus, the target protein might still bind to the immobilized probe which will lead to its detection by means of mass spectrometry. Simple comparison of the detected proteins in the presence or absence of a competitor will then lead to false negative results. Therefore, it is advisable to use quantitative proteomics to detect the relative abundance of the isolated proteins (see Section 2.1.11) as well as carrying out enough replicates to ensure statistical significance in the results.

2.1.8. Non-Specific Binding

In affinity-based proteomics, target identification often requires elimination of proteins that bind non-specifically to the probe and/or the matrix (background binding) to enable detection of the true targets. Non-specific binders increase the complexity of the mixtures to be analyzed by mass spectrometry and often are not easy to detect. High-affinity ligands will preferentially bind their physiological targets in cell lysate and interact with additional proteins with lower affinity. However, in pulldown experiments the protein extract is exposed to a high local concentration of the immobilized ligand which will enable binding of low-affinity targets as well, in particular if the target protein is present only in low concentration. This shortcoming can be overcome by decreasing the amount of the affinity probe on the matrix. Thus, high affinity of the probe for the target and sufficiently high target concentration often are indispensable for successful target identification. Preferably probes with nanomolar or low micromolar affinities should be used which may require substantial structure–activity relationship investigations. Since the origin of the bioactivity data is often a phenotypic screen, it might be difficult to assess the relative affinity between the small molecule of interest and its potential target(s). Nevertheless, it must not be forgotten that high cellular activity might originate from the polypharmacological properties of the compound in question which may result from the sum of low affinities towards multiple targets rather than high affinity for a single protein.[11c]

Of the successful target identifications reported, most employed high-affinity probes and/or the target proteins were present in relatively high concentrations. For low-affinity ligands the targets may be lost during the stringent washing steps, in particular if their abundance is low, thereby leading to identification of highly abundant background proteins. Frequently occurring background proteins may be identified by statistical analysis of pulldown experiments with different probes to identify the frequency of occurrence independent of the constitution of the probe ("black lists"[59]). In addition, the determination of proteins that (non-specifically) bind to commonly used matrices, so called "bead proteomes", could help to distinguish background binding.[60] Furthermore, quantitative assessment of human cell line proteomes and copy numbers in them[61] might help to rule out non-specific binders. Nonetheless, background proteins need to be considered for every probe since putative unspecific binders could prove to be the physiological target of the compound of interest (e.g. Tubulexin, Table 1, entry 11, see Section 5.5).[62]

Recently developed powerful mass-spectrometric techniques allow putative targets to be identified from complex mixtures immediately after release from the affinity resin.[63] However, separation by gel electrophoresis and staining of the proteins may still be advisable before mass spectrometry. Investigation of different pulldown conditions (e.g. probe design, amount of immobilized probe, concentration of the lysate, washing conditions, elution with free compound, competition) with subsequent staining of the proteins after SDS-PAGE can provide valuable information for background reduction. For example, linker variation or elution with an excess of unmodified compound and incorporation of a photoaffinity group to enable more stringent washing conditions may reduce unspecific binding. In addition, the protein mixture may be exposed to a control matrix to reduce or remove "sticky" proteins prior to the pulldown experiment with the active affinity compound ("pre-clearing").[64] Serial affinity chromatography provides a further option.[50] In this approach, probe-derivatized beads are exposed to a protein extract, the lysate is separated and added to a fresh affinity resin and these steps are repeated several times (Figure 1 b). It is expected that the first resin will enrich specifically binding proteins while non-specific binders would be captured to a similar extent by all resins. Serial affinity chromatography was successfully applied to the FK506/FKBP12, methotrexate/DHFR, and benzenesulfonamide/CA2 interactions and to the identification of clathrin as a target of bolinaquinone (Table 1, entry 12).[50,52]

2.1.9. Probe Immobilization by Photoactivation

Natural products provide a rich source of biologically active compounds. However, the identification of their targets is often hampered by lack of synthetic accessibility and, therefore, lack of conclusive structure–activity relationships. In cases were total synthesis of a natural product is not available, the natural products can be modified at positions, which allow chemical derivatization, giving raise to "semi-synthetic" derivatives. When substitutions are well tolerated, a linker for immobilization on a matrix can be incorporated.

An alternative method for immobilization of compounds without the need for modification is coupling to a solid surface by means of photoactivation.[16] In this method, beads are derivatized with oligoethylene glycol linkers containing aryldiazirine groups. Upon UV irradiation, highly reactive carbenes are generated which unspecifically react with the compounds to be immobilized (Figure 5). In a test study, all

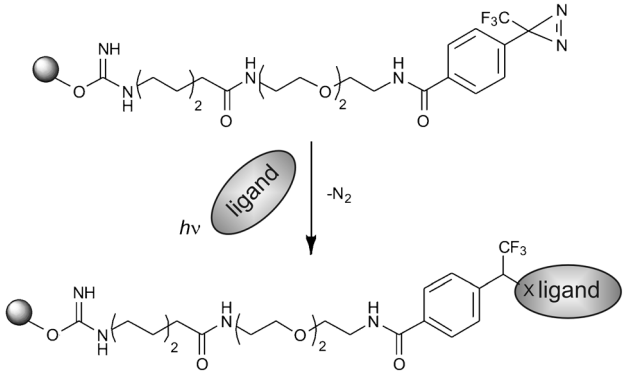


Figure 5. Compound immobilization by photoactivation.

possible C–H and C–O insertion products were generated in a relatively uniform distribution.[65] Thereby populations of differently oriented immobilized compounds are generated. Only insertions which target sites that tolerate modification without dramatic loss of activity will enable pulldown of the target proteins. This approach was successfully used to identify glyoxylase I as the target protein of the osteoclastogenesis inhibitor methyl gerfelin (Table 1, entry 13).[66] The method has enabled the generation of chemical arrays to screen for compounds that bind to a protein of interest.[16,67]

2.1.10. *Target Identification by Ligand-Directed Protein Labeling*

Transfer of a fluorescent tag from a small molecule to the target proteins upon ligand binding is a relatively novel strategy for target identification. To this end, the compound is modified to carry a reactive acylating or alkylating group (e.g. a phenyl ester or a phenylsulfonate) coupled to a fluorophore. In a first step the compound binds to the target protein before transfer of the reactive group to a proximal amino acid whereby the ligand, which acts as a leaving group, is displaced (Figure 6).[68] For instance, acyl-dye transfer was used to label proteins that bind to the natural product marinopyrrole A (Table 1 entry 14) with dimethylaminocoumarin.[68b] After incubation of live cells with the marinopyrrole A derivative, two fluorescent bands were observed after SDS-PAGE with the crude lysate. Co-immunoprecipitation using an anti-dye antibody and subsequent mass spectrometry led to the detection of three bands, two of which were identified as actin and the third as actinin which is tightly complexed with actin. Actin could be confirmed as target protein using microscopy, microcalorimetry, and an in vitro actin polymerization assay. In a similar setup, tosyl chemistry was employed to label proteins that bind to fusicoccin A (Table 1, entry 15, see Section 5.6).[69]

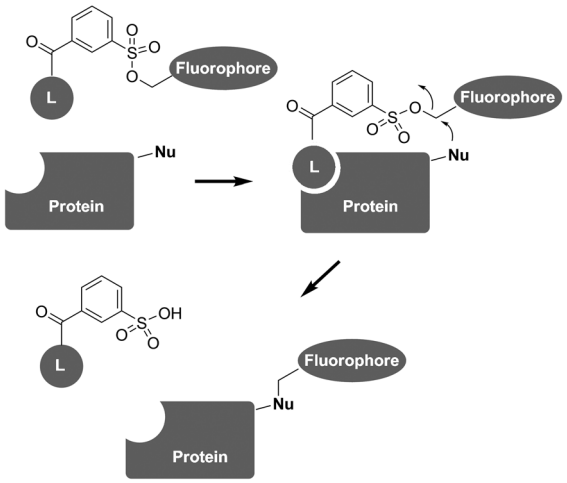


Figure 6. Ligand-directed protein labeling using tosyl chemistry.[68a]

2.1.11. *Quantitative Proteomics*

The limitations of comparisons between proteins identified with active and inactive probes can be overcome by relative quantification of proteins from different populations, that is, quantitative proteomics. For this purpose, usually stable isotope labels are incorporated in the proteins, although label-free quantification is possible as well.[70] Stable isotopes retain the physico-chemical properties of the proteins and thus do not interfere with binding. However, they differ in mass which enables their simultaneous detection and relative quantification. Isotope labeling can be performed metabolically or chemically. Mann and co-workers established stable isotope labeling by amino acids in cell culture (SILAC).[71] SILAC makes use of the requirement of mammalian cells for essential amino acids (e.g. methionine, arginine, leucine, lysine). Growing cells are cultured either under normal conditions (natural abundant amino acids, "light" amino acids) or supplemented with labeled essential amino acids (e.g. ^{13}C labeled L-arginine[72] and ^{13}C,^{15}N-labeled L-lysine,[73] so called "heavy" amino acids) which are fully incorporated into the cellular proteome after five rounds of cell divisions. Different lysates are used for the pulldown with the active (e.g. "heavy" lysate) and control probe ("light" lysate).

The eluted proteins from both experiments are combined, processed, and analyzed by mass spectrometry (Figure 7a). The use of isotopically labeled amino acids leads to a mass shift of 6 Da for ^{13}C-L-arginine-containing peptides and 8 Da for ^{13}C, ^{15}N-L-lysine-labeled peptides when comparing the mass spectra of "heavy" peptides with their "light" counterparts.[73] SILAC peptide pairs have addressable locations in mass and retention-time space—if a light peptide is detected at a certain mass and retention time, the heavy peptide is usually observed as well.[73] The relative quantity of the "heavy" to the "light" peptide can be derived from the ratio of ion intensities of the SILAC peptide pairs. Affinity chromatography using SILAC-labeled protein extracts allows for milder washing conditions allowing for isolation of weak binders and prioritization of targets by means of

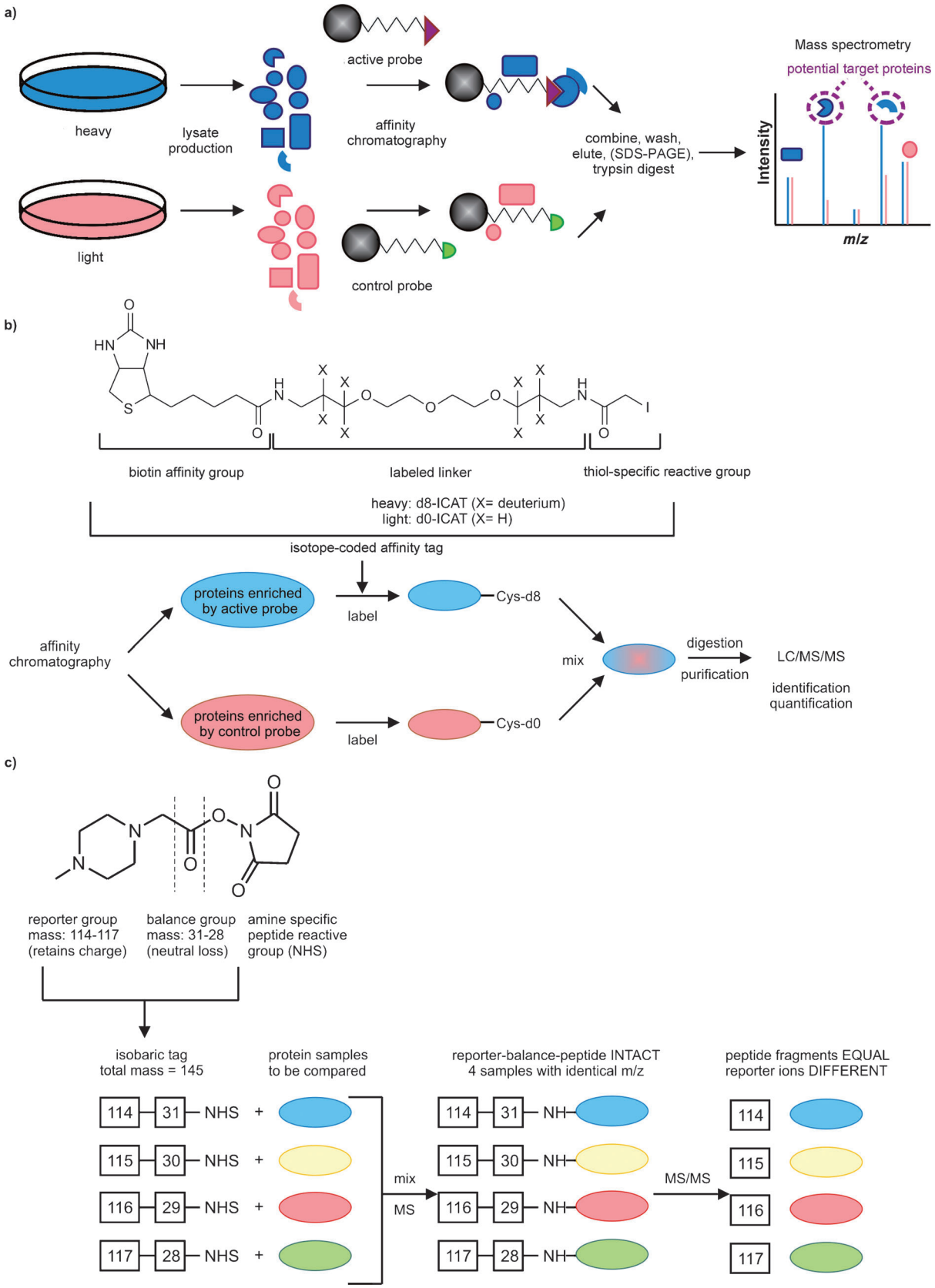


Figure 7. Quantitative proteomics approaches. a) Target identification employing the SILAC method.[73] b) Isotope-coded affinity tag (ICAT) strategy.[70] c) Quantification by using isobaric tags (iTRAQ).[74]

SILAC ratios.[73] Besides proof-of-concept for small molecules with known modes of action (kinase inhibitors, immunophilin ligands),[73] SILAC was used in the identification of tubulin and CSE1 L as target of the tetrahydropyran tubulexin A (Table 1, entry 11),[62] GSTP1 as target of piperlongumine (Table 1, entry 16),[75] epoxide hydrolase 1, and MT-ND1 as targets of a 5-benzoylindole derivative (Table 1, entry 17),[76] and for the analysis of imatinib targets[77] and the human kinome.[78] SILAC was also utilized to label embryonic stem cells.[79] Metabolic labeling was also reported for whole organisms such as mice,[80] *C. elegans* and *D. melanogaster*.[81]

Although powerful, SILAC is not applicable for labeling cells which cannot be maintained for sufficiently long time to ensure complete labeling of the proteome, for example, platelets, and for labeling harvested tissues and human body fluids. In such cases, chemical labeling is the alternative of choice. Several strategies are available, in particular the use of isotope-coded affinity tags (ICAT)[82] or isobaric tags for relative and absolute quantification (iTRAQ).[74] The ICAT reagent contains an affinity tag (e.g. biotin), a isotopically labeled linker (e.g. eightfold deuterated) and a thiol-specific reactive group (e.g. iodoacetamide) (Figure 7b).[82] The side chains of cysteine residues of the protein mixtures that are to be compared are derivatized either with the isotopically light or the heavy form of the ICAT reagent. After combining the samples and tryptic digest labeled peptides are isolated by avidin affinity chromatography, fractionated, and quantitatively analyzed by mass spectrometry. A second-generation ICAT reagent, containing a cleavable linker, was used for target identification of the antitumorigenic agent E7070 (*N*-(3-chloro-7-indolyl)-1,4-benzenedisulfonamide, Table 1, entry 18) and led to the isolation of malate dehydrogenase (MDH).[83]

ICAT significantly reduces the complexity of the mixtures to be analyzed and covers 96.1% of the human proteome.[84] Full coverage can be reached by labeling N- or C-terminal peptides.[84] iTRAQ includes isotopic labeling of amine groups.[74] The iTRAQ reagent consists of a reporter group (N-methylpiperazine), a mass balance group (carbonyl) and a group that reacts with peptides (NHS ester; Figure 7c).[74] It labels any amine group of the peptides from a protein mixture through an amide linkage. The reporter group incorporates combinations of ^{13}C, ^{15}N, and ^{18}O which results in tags with a mass of 114–117 Da. The balance group has a mass of 28 to 31 Da originating from isotopic enrichment at the carbonyl component to ensure identical total mass of the reagents (145 Da). Thus, peptides labeled with different iTRAQ reagents are isobaric and appear as one single MS peak after combining the labeled peptide mixtures. Collision-induced dissociation (CID) causes fragmentation of the tag amide bond and loss of the balance group while the reporter group produces MS/MS signature ions with *m/z* 114–117. The relative areas of the peaks are used for the quantification of peptide abundance in each sample. Using iTRAQ up to four samples can be labeled and analyzed simultaneously. iTRAQ was applied in the isolation of tankyrase as target of XAV939 (Table 1, entry 19, see also Section 5.10),[85] oxysterol binding protein as target of OSW-1 (Table 1, entry 20) and ORPphilins,[15a] and target profiling of ABL inhibitors using kinobeads and the competition approach (Table 1, entry 21).[86]

Besides the methods employing labeling of proteins or peptides, label-free quantification (LFQ) of proteins was performed in several studies.[87] For instance, integration of the complete mass spectrometric signal of each peptide can be used to quantify the same peptide in different LC-MS/MS runs[88] (for an extensive Review on current label-free quantification strategies see Ref. [89]) Label-free approaches are the least precise among the present techniques for mass spectrometric quantification because all variations between experiments are reflected in the data obtained.[86,89] However, LFQ is less costly and time-consuming (no additional steps to label proteins or peptides are required) and provides a higher dynamic range of quantification than stable isotope labeling.[86] Moreover, there is no limit to the number of probes that can be compared and mass spectral complexity is not increased.[86] To our knowledge, to date no target-identification studies employing LFQ have been reported. Nevertheless, label-free quantification should be considered as optional when designing an experiment for protein quantification.

2.1.12. *Tagged Libraries*

Limitations because of difficulties with linker incorporation after identification of bioactivity may be overcome by including the linker in the synthesis of compound libraries to be analyzed.[51c,90] Thereby the steps from hit identification to isolation of the target proteins may be made more efficient. Many of the reported examples are based on the triazine scaffold, which allows for versatile modifications (Figure 8a).[91] A triazine-based library including a triethyleneglycol linker was synthesized and screened in zebrafish embryos for brain and eye morphological changes, and from 1536 triazine derivatives one hit compound was identified (Table 1, entry 22). Removal of the linker (to yield encephalazine) did not influence the activity showing that neither the linker nor

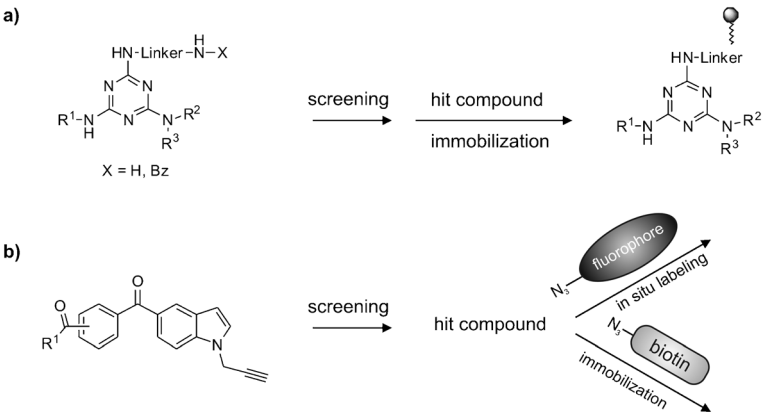


Figure 8. Tagged libraries for screening and target identification. a) Triazine-based library and b) 5-benzoylindole-based library.

the amino group is important for activity. The linker version of encephalazine was immobilized on an affinity matrix and exposed to protein extracts. Ribosomal accessory proteins were identified as the target proteins of encephalazine.[92] Similar approaches led to the identification of several hits in a screen for pigmentation inducers and the mitochondrial F1F0 ATPase as their target.[90,93] Tagged libraries with benzoyl capping were used in a screen for modulation of insulin signaling in *C. elegans*. An active compound (GAPDS, Table 1, entry 23) was identified and employed in an identification of GAPDH as the target protein following an affinity-matrix-based strategy.[94]

Cisar and Cravatt synthesized a library which also allows for the direct progression from a cell-based screen to target identification without additional chemical modifications of the hit compounds.[76] The library is based on 5-benzoylindole and 7-benzoyl benzo-1,4-diazepin-2,5-dione in which each compound has a photoreactive group for UV cross-linking to the target proteins and an alkyne handle for click chemistry (Figure 8b).[76] The library was screened for antiproliferative activity in MDA-MB-231 cells in normal or low-glucose medium. One compound (Table 1, entry 17) inhibited the proliferation under hypoglycemic conditions and could be directly employed in a SILAC-based affinity proteomics experiment to identify epoxide hydrolase 1 and MT-ND1 as targets, which are involved in mitochondrial respiration.

2.2. Two-Dimensional Gel Electrophoresis

Two dimensional gel electrophoresis (2DE) allows thousands of proteins to be separated simultaneously by isoelectric focusing according to their charge (1st dimension) and according to their size (2nd dimension). The resulting spots usually represent single proteins which can be isolated and identified. This technique was the first to enable proteome analysis[95] and can be used to identify targets of small bioactive molecules. A requirement is the covalent attachment of the compound to the target protein and its labeling (radiolabel or fluorophore) for detection. Jessen et al. employed 2DE for target identification of the apoptosis-inducing 3,5-diaryl-[1,2,4]-oxadiazoles MX-74420/MX-126374 (Table 1, entry 24).[96] A tritium-labeled photoaffinity derivative was cross-linked by UV irradiation to the target and a specifically labeled band of 50 kDa and with a p*I* of 5.3 was observed by SDS-PAGE. The appearance of this band was complemented by excess of unmodified active compound. Separation of the labeled lysate by means of 2DE revealed one single protein spot which was identified as TIP47 after tryptic digest analysis and mass spectrometry.

Park et al. developed a fluorescence difference approach in two-dimensional gel electrophoresis (FITGE) and employed it for the identification of the target of an anti-proliferative compound (Table 1, entry 25).[45] An active and a control pulldown probe were synthesized using a benzophenone as photoreactive group and an acetylene moiety for bioorthogonal click reaction. This probe set was used to label the proteomes in lysates or live cells after UV crosslinking and click reaction with Cy5- (active probe) or Cy3-azide (control). The labeled proteomes were mixed and subjected to 2DE. The merged images of the Cy5- and Cy3-fluorescence revealed only a few spots detectable in the Cy5 image but not in the Cy3 image and led to the identification and subsequent confirmation of tubulin as a target protein.

Although useful for target identification 2DE is time-consuming, labor-intensive, and limited by low resolution. Owing to the insufficient sensitivity of protein-staining procedures, mostly only the abundant proteins are detected. Furthermore, often hydrophobic, especially membrane proteins cannot be detected and proteome comparison is difficult because of gel-to-gel variations. Newer modifications include isoelectric point focusing over immobilized pH-gradients or cation-exchange columns prior to trypsination, and reversed-phase nano-HPLC, providing a similar sample simplification as 2DE, but with considerably higher sensitivity.[97]

2.3. Label-Free Approaches to Target Identification

2.3.1. Drug Affinity Responsive Target Stability (DARTS)

The drug affinity responsive target stability (DARTS) method was developed as a universally applicable strategy to investigate drug-target interactions.[98] DARTS is based on the finding that a target protein bound to a small molecule might be less susceptible to proteolysis than its unbound state.[99] A major advantage of DARTS is the use of unmodified compounds for target identification. Small-molecule binding to a target protein might induce local or global stabilization of the protein conformation and/or alter accessibility to proteolysis.[98] Compound binding may mask proteolytic cleavage sites, resulting in differences in proteolysis patterns between compound-bound and free state (Figure 9). The FKBP12-rapamycin/FK506 complex was protected against digestion by the protease subtilisin after binding of FKBP12 to rapamycin and FK506. Treatment with wortmannin was not effective, which demonstrates the selectivity of protease protection. Since neither rapamycin nor FK506 causes conformational changes in FKBP12, binding alone is sufficient to stabilize the

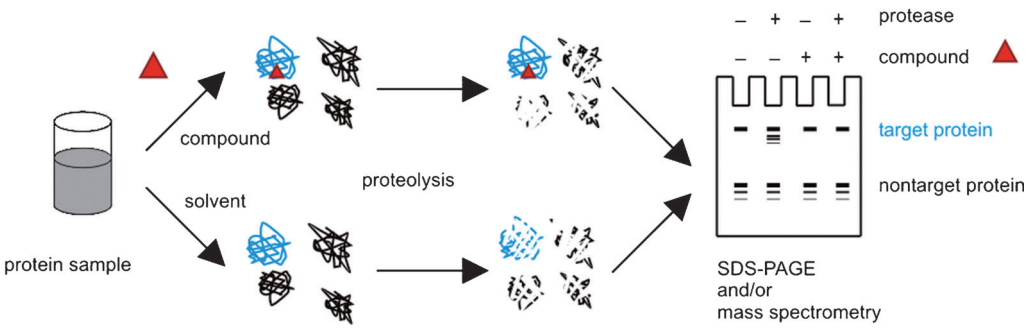


Figure 9. Target identification using drug-affinity responsive target stability (DARTS).[98]

protein in the proteolysis-resistant stage.[98] The applicability of DARTS to complex protein mixtures was demonstrated for the didemnin/B-EF1α, mTOR/rapamycin, COX-2/celecoxib, and SCF E3 ubiquitin ligase-inhibitor complexes.[98] Furthermore, DARTS was employed to identify the eukaryotic initiation factor 4A (eIF4A) as one target of resveratrol (Table 1, entry 26). DARTS might be limited by the affinity of the drug–target interaction, protein abundance and detection by mass spectrometry, and the susceptibility of the proteins to hydrolysis.

A related energetics-based methodology using pulse-proteolysis[99] was employed to identify ATP binding proteins in *E. coli*.[13] Protein samples were incubated with ATP and the urea-induced protein unfolding as compared to untreated samples was analyzed after a brief exposure to a protease. Ligand binding to its target proteins changes the stability of the protein, for example, thermodynamic stability, resulting in a decrease in unfolding rates and changes in the dynamics of the native protein.[13] The amount of the protein which remains after pulse proteolysis is compared by means of two-dimensional gel electrophoresis (2DE). As a proof-of-concept Liu et al., after 2DE, selected only 12 spots to identify proteins that bind to ATP, six of which were already reported to bind to ATP. For three of the remaining four proteins, an influence of ATP on their energetic properties was confirmed. Since this method suffers from the limitations for 2DE (see Section 2.2) Chang et al. used SDS-PAGE and liquid chromatography instead of 2DE for the detection of ATP binding proteins in *E. coli*.[100] Through the experimental ease of SDS-PAGE, different concentrations of urea could be tested which increases the chance of target identification because urea concentration modulates protein unfolding differently in the presence or absence of ligand.

West et al. used chemical denaturant-dependent oxidation rates of methionine (termed stability of proteins from rates of oxidation, SPROX) coupled to quantitative mass spectrometry to detect thermodynamic properties of protein un/refolding in the presence of a ligand.[101] To this end, a protein mixture is incubated with increasing concentrations of guanidinium hydrochloride as the denaturant in the presence or absence of a compound. The samples are treated with hydrogen peroxide so that the side chains of the methionine residues are selectively oxidized. The non-oxidized and oxidized proteins are quantified by means of isobaric tags. Transition midpoints are determined, that is, the assay conditions at which, depending on the concentration of denaturant, the non-oxidized methionine-containing peptides disappear and/or the oxidized peptides appear. Proteins are considered as targets if the transition midpoint between the ligand-bound and the ligand-free state for a peptide differ. By means of this method, cyclosporine A binding proteins were identified in a yeast lysate. Eight of these interactions were not reported before and one binding protein (calcineurin) was not detected. This method is limited by the requirement of having methionine residues in the ligand-binding site and the need for sufficiently high protein and ligand concentration to detect a transition midpoint shift.

2.3.2. Target Identification by Chromatographic Co-Elution (TICC)

Target identification by chromatographic co-elution (TICC) is based on the co-fractionation of ligand–target complexes during nondenaturing HPLC coupled with LC-MS to detect compound–protein interactions in complex protein mixtures with nearly native cellular proteins.[102] Ligand-target protein binding results in a characteristic shift in the chromatographic retention time profile of a compound. HPLC-MS/MS is then used to identify the binding protein. The methotrexate/DHFR, radiciciol/Hsp90, sordarine/eIf2, and TSA/HDAC1,2 pairs provided proof-of-principle. Chromatographic co-elution was employed in the identification of Erg6p, a cytoplasmic delta-sterol C-methyltransferase, as target of the antifungal natural product 4513-0042 (Table 1, entry 27) and in the identification of off-targets of the dopamine receptor agonist A77636.[102] This approach is also suitable to detect low-abundant target proteins as well as low-affinity (micromolar) interactions. One constraint is the need to separate unbound versus protein-bound compound which excludes covalent binders. The native conditions needed for TICC require conditions for sufficient solubilization of membrane proteins that will retain drug–target interactions. Furthermore, co-eluting proteins with similar retention properties might complicate the identification of the real target. Thus, this approach is limited primarily to hydrophobic noncovalent protein–ligand interactions and biological samples containing soluble native proteins.

2.3.3. Biochemical Suppression

In genetics, phenotypes caused by a mutation in a gene may be rescued by the overexpression of other genes. This phenomenon is used in high-copy suppressor screens.[103] By analogy, a biochemical suppression strategy was developed in which an inhibitor is used as a mutation and the addition of partially purified protein mixtures mimics protein overexpression, which should prevent the inhibition by the small molecule in a given assay.[103] Biochemical suppression is limited to soluble and fractionatable proteins, thus making it of limited use as a general technique. Peterson et al. employed the biochemical suppression of a compound-induced phenotype after addition of concentrated protein fractions to identify the target protein of the tetracyclic indole pirl1 (Table 1, entry 28).[103] pirl1 was identified in a screen for PIP_2-induced actin polymerization in *Xenopus* extracts. Untreated *Xenopus* extracts were fractionated, concentrated to mimic genetic protein overexpression, and added to the extracts treated with the compound. A fraction that suppressed pirl1 activity enhanced actin polymerization kinetics. Further chromatographic fractionation identified the Arp2/3 complex as a downstream component required for PIP_2 induced actin polymerization which prevents pirl1 inhibition. The cdc42/RhoGDI complex was isolated from a second fraction as a suppressor of the pirl1 phenotype. This approach is only applicable to the identification of target proteins if an activity amenable to an assay can be monitored and, to date, other examples for target identification are not published yet.

2.3.4. *Target Deconvolution by Comparison*

The increasing amount of data available for cells treated with small molecules, concerning activity fingerprints and signatures (e.g. gene expression and proteome profiling, phenotypic multiparameter analysis, cytotoxicity data) allows the collection of "fingerprints" or "signatures" of known compounds and novel biologically active compounds. These databases can be used to predict the mode of action and even the target proteins of an uncharacterized small molecule (Figure 10).

One possibility to profile known compounds is the screening for cytotoxicity or growth inhibition of cancer-cell-line panels. Growth inhibition data (e.g. GI_{50} values) derived from these studies can be used to create activity patterns of known drugs. Investigation of the NCI (National Cancer Institute)-60 cell-line panel was the first high-throughput cancer-cell-line screening program.[104] In addition to disease-oriented screening for drug discovery[105] using the COMPARE algorithm[106] for a query compound, the database of screened agents can be searched for similar profiles. Similarity in pattern often is related to similar mode of action, mode of resistance, or molecular structure[105] and creates a hypothesis that needs to be confirmed. This kind of analysis was performed for topoisomerase II inhibitors,[107] tubulin modulators,[108] and dihydroorotate dehydrogenase.[109] The activity patterns can be combined with additional information, for example, structural features, possible targets or modulators of activity in cells.[105] Besides the NCI60 platform, further cancer-cell-line panels were established.[104] The JFCR39 (Japanese Foundation for Cancer Research 39) system, which holds fewer cell lines than the NCI60 panel, was employed in several studies to predict the mechanism of action of uncharacterized small molecules.[110] Moreover, almost 1000 cancer cell lines are included in the Cancer Cell Line Encyclopedia (CCLE) which was released through a collaboration of academia and industry.[111] CCLE contains the mutational status of over 1600 genes, DNA gene copy numbers, mRNA expression levels, and the pharmacological profiles for 24 anticancer drugs.

Lamb et al. introduced connectivity maps as a generic tool to relate the action of drugs to genomic signatures, physiological processes or diseases, and to create drug signatures.[112] Using connectivity maps, the signature of a small molecule of interest can be compared to a database of compound signatures and thereby discover new connections. Lamb et al. chose mRNA expression profiling of compound treated cells as a genomic signature and selected 164 compounds representing a wide range of activities including compounds that target the same protein as well as different compounds for the same clinical indication. The breast cancer cell line MCF7 was used to generate genomic signatures. Then a "query signature" was compared to the reference expression profiles in the data set. Using the connectivity map, a high score was obtained for the HDAC inhibitors vorinostat and trichostatin A and their strong connectivity to valproic acid and HC toxin, which are known to inhibit HDAC activity but are structurally different. A high score was obtained also for estrogens such as E2 and genistein. The connectivity map showed a weak connectivity to 17α-estradiol which is consistent with the lower affinity for the estrogen receptor (ER). The approach also identified a negative connectivity for the ER-antagonists fulvestrant, tamoxifen, and raloxifen. Connectivity was also identified for antipsychotic drugs of different structural classes which act as dopamine receptor antagonists and do not directly influence gene expression. This method was then applied to the natural product gedunin (Table 1, entry 29) with unknown target. Gedunin was identified in a screen for inhibition of androgen receptor (AR) activation in prostate cancer cells.[112] High connectivity was found for geldanamycin and two derivatives thereof which are heat shock protein 90 (Hsp90) inhibitors. The involvement of Hsp90 in the regulation of AR stability was then confirmed which demonstrates the potential of the connectivity map to identify the targeted pathway of a poorly characterized compound. Since its introduction in 2006, the

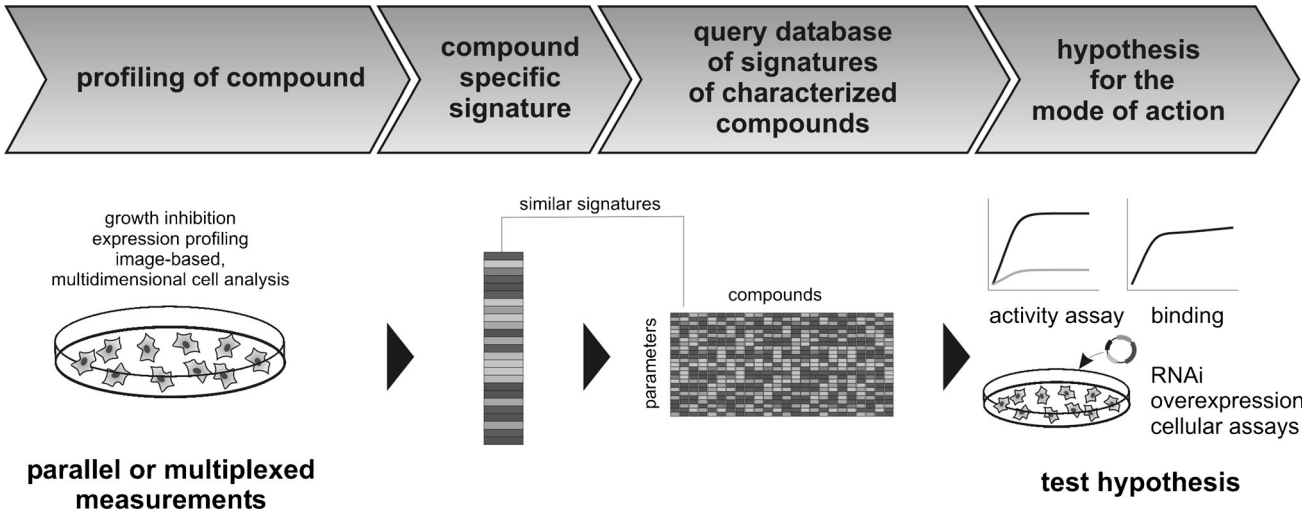


Figure 10. Target deconvolution by using compound profiling to create a signature of the compound and database query to create a hypothesis for the mode of action or the target protein.

connectivity-map database was employed for the identification of the targets or the involved pathways of uncharacterized compounds[113] or the identification of novel compounds to target particular physiological processes.[114]

A fingerprint or signature of a compound can also be generated by proteome profiling, for example, by means of two-dimensional difference gel electrophoresis (2D DIGE) which allows for the quantification of protein abundance.[115] Protein extracts are first covalently tagged with fluorescent *N*-hydroxysuccinimide derivatives (e.g. to attach Cy3 and Cy5) to label ε-amino groups of lysine residues. The proteomes to be compared are mixed and subjected to 2DE. Protein spots are detected after excitation at the appropriate wavelength for each fluorophore. Differences in protein abundance can be detected after merging of both images. This method allows two protein samples to be compared. Comparison of more samples is feasible when an internal standard is included in each 2DE to normalize the spots across all gels.[116] Muroi et al. established a proteome profiling system to elucidate mode of action by means of 2D DIGE.[117] The proteomic patterns of 19 known, well-characterized compounds were obtained by 2D DIGE and the proteome data obtained for query substances were compared with those of the known compounds. Iejimalides (Table 1, entry 30), which bind V-ATPase and affect actin polymerization, clustered into the same tree as other V-ATPase inhibitors, such as bafilomycin A1 and concamycin A, whereas they failed to cluster with actin inhibitors. Thus, proteome comparison using 2D DIGE may help elucidate the target or the mode of action of a small molecule of interest.

Perlman et al. employed data obtained by automated microscopy to create compound profiles.[118] For compounds with known mechanisms of action multidimensional single-cell phenotypic information was collected based on different probes (staining for DNA; tubulin, actin, phospho-ERK, phospho-p38, phospho-CREB, c-Fos, SC35, anillin, calmodulin) to cover a broad range of cellular mechanisms. For each cell, region, and probe, different measurements were performed (e.g. size, shape, intensity) to generate a set of descriptors. The population response to increasing concentrations of a given compound was used to create compound profiles. Similar profiles were obtained for compounds with different structures but common targets. Further proof-of-concept provided blinded alternate titrations of known compounds included in the set which clustered closely with their unblinded counterparts and with compounds with similar mode of action. In addition, one poorly characterized compound, austocystin, clustered with transcription and translation inhibitors and in vitro inhibition of transcription could be confirmed.[118]

Abassi et al. chose impedance measurements for dynamic monitoring of compound' influence on cells.[119] This method allows for detection of cellular effects of a small molecule over time whereas other compound-profiling strategies are based on end-point assays. An electrical sensor is employed that continuously tracks morphological changes (e.g. cell shape and movement) of adherent cells.[120] When cells are seeded in wells containing an electrode their attachment and spreading on the electrode restricts the current and causes changes in impedance which is extremely sensitive to external conditions such as temperature, pH value, and small molecules.[120] By screening 2000 FDA-approved drugs, natural products, and further bioactive compounds against the A549 non-small lung cancer cell line, Abassi et al. demonstrated that the impedance change can be predictive of the mechanism of action.[119] In addition, monitoring of impedance over time allowed the short- and long-term impact of compound treatment to be detected. The impedance-based time-dependent profiles of the screened compounds were used for cluster analysis. Good correlations were obtained for GPCRs, tubulin modulators, and inhibitors of DNA synthesis and this methodology predicted new mechanisms of action for characterized compounds.

These profile comparison methodologies help to create hypotheses for target proteins or mechanisms of action and therefore require experimental verification. clearly, this strategy for target deconvolution is restricted to known targets. Thus, when a query compound fails to cluster with characterized active molecules, novel modes of action should be considered.

2.4. Protein Microarrays

Proteins immobilized in an array can be exposed to small molecules to detect binding. Protein microarrays can be produced using technologies developed for the spotting of DNA. However, owing to possible denaturation, their generation is more complex than the production of DNA microarrays.[121] The main challenges in the development of protein microarrays are the development of immobilization strategies and methods which do not compromise protein structure and function (for an extensive Review see Jonkheijm et al.[121]), and the expression and purification of thousands of correctly folded and active proteins. Usual cellular expression and purification of large numbers of recombinant proteins is time-consuming and labor intensive.[122] Although cell-free systems could help to overcome these problems, current strategies suffer from lack of post-translational modifications which might be necessary for a given binding event. Nevertheless, protein microarrays should be considered as a target-identification tool because they may help to overcome problems of affinity-based proteomics, such as the preferential detection of high-abundant proteins and the challenging isolation of low-abundant or low-affinity proteins. On a microarray, proteins are displayed in equal amounts.[11a] Usually they are expressed as His- or GST-tagged proteins for reversible site-specific immobilization. Alternatively, proteins can be covalently attached to the array surface, for example, through their functional groups or by means of a chemical ligation reaction.[121] The analyzed small molecule is then exposed to a protein chip and its binding to the target proteins is detected. To this end, the compounds need to be labeled, for example, with fluorophores. Despite the potential usefulness of protein biochips, there are only a few examples for their application in interaction studies of proteins with small molecules (Table 1, entry 31).[123]

2.5. Computational approaches

Computational methods offer alternatives to experimental approaches to target identification. Keiser et al. used a statistics-based similarity ensemble approach (SEA)[124] to predict new targets of 878 FDA approved drugs and 2787 pharmaceutical compounds.[125] SEA compares protein targets considering similarities among their ligands. Subsequently a compound collection was screened against a panel of more than 1400 proteins. Each target was characterized only by its known ligands, and the two-dimensional structural similarity of each drug to each target's ligand set was calculated as an *E* value using SEA. The novelty of the predicted off-targets with strong *E* values was further analyzed against orthogonal databases and the literature. Known associations could be confirmed (Revanil (lisuride) as an α_2 adrenergic antagonist, Permax (pergolide) as 5-HT_{1D} receptor agonist). In addition 184 predictions were unprecedented. Cross-activity occurred among G-protein coupled receptors (GPCRs) which is well known.[126] Interestingly, for some drugs, the predicted new targets were unrelated to their known targets, both by sequence similarity and structure. The activity of these compounds against the predicted targets was then determined experimentally for 30 predictions, four of which could be confirmed (Table 1, entry 32–35).

This study was extended by Lounkine et al. who developed a guilt-by-association metric to create a drug-target adverse drug reactions (ADR) network.[7] Predicted new drug targets were linked to ADRs of those drugs for which they are the primary or well-known off-targets. The resulting drug-target-ADR network demonstrated that drugs often were associated with off-targets unrelated in sequence and structure to their primary target and provides a tool to evaluate candidate drugs according to their chemotype and the ADR related to it.

Further related studies showed the potential of computational methods for target and off-target analysis.[127] Therefore, this approach can be used for systematic prediction of targets and off-targets of bioactive compounds which can contribute to the better understanding of their mode of action. As for all comparative strategies, the limitations lie within the dependency of the method on known targets and their ligands.

2.6. Genetics Approaches for Target Identification

Yeast is often used as eukaryotic model organism for mammalian diseases and pathways, since at least 31% of the proteins encoded by yeast genes have human homologues and approximately 50% of the human genes implicated in heritable diseases are also found in yeast.[128] Its ease of manipulation and genetic tractability, the relative short life cycle, its inexpensive maintenance and growth as well as its stability in a haploid and diploid state, and the availability of the baker's yeast Saccharomyces cerevisiae genome since 1996 further fortify the employment of yeast as model organism.[129] The possibility to precisely insert DNA sequences at specific locations within the yeast genome by homologous recombination[130] led to the construction of a complete set of deletion mutants[131] thus enabling new genomic screens and large-scale comparative studies[132] which gave rise to successful strategies for target identification.[12b] One powerful approach established in *S. cerevisiae* is drug induced haploinsufficient profiling (HIP) which is based on the fact that lowering the dosage of a drug-target-encoding gene from two copies to one copy in a diploid yeast will lead to a heterozygote with increased sensitivity towards the drug (Figure 11 a).[133] An analogous approach to HIP is the homozygous deletion profiling (HOP) where both copies of non-essential genes are knocked out. In general HOP is not used for target-protein identification, but it can be applied to reveal functionally related or connected genes which buffer the drug-target pathway or are involved in chemical detoxification. HOP could identify genes related to the mechanism of action even if the target of the compound is not a protein. Pioneering experiments using individual heterozygous strains grown in the presence of sublethal concentrations of a compound that directly targets the gene product of the heterozygous locus demonstrated the feasibility of this approach for identifying molecular targets.[133] Each deleted gene was replaced by a KanMX deletion "cassette" flanked by two distinct 20-nucleotide sequences that serve as "molecular barcodes" which enable the unique identification of each gene disruption even in a pooled population.[131a] The relative sensitivities of the single strains could be quantified by using high-density oligonucleotide arrays carrying the barcode complements.[131]

The use of barcode strains enabled a first parallel approach analyzing the influence of 10 compounds on the whole yeast genome. For the anti-cancer agent methotrexate (Table 1, entry 36) not only the known target protein, dehydrofolate reductase (encoded by DFR1), could be identified but also genes acting upstream of DFR1 and genes involved in compound availability. However, for a second anticancer compound 5-fluorouracil (5-FU, Table 1, entry 37) the known target thymidylate synthase was not found but instead several genes involved in essential RNA processes were identified suggesting that 5-FU might be directly incorporated into RNA. In consequence this first genome-wide HIP assay revealed that the technique is powerful not only for identifying targets but also for clarifying modes of action and to gain information on transport proteins which are involved in drug resistance.[134] Several additional genome-wide profiling studies were performed making use of the heterozygous deletion strain collection. For tricyclic antidepressants (amitryptiline, imipramine, desipramine, clomipramine, chlorpromazine, and trifluopromazine) it was shown that they influence *NEO1*, a P-type ATPase and for the drug tunicamycin target (Table 1, entry 38) proteins involved in endoplasmic reticulum stress response and cell wall glycosylation were identified.[135] Furthermore it could be demonstrated that normal vacuolar and Golgi complex functions are responsible for the insusceptibility of the aminoglycoside antibiotic gentamicin (Table 1, entry 39) in eukaryotes[136] and that DNA damage and modulation of cytoskeleton dynamics are mechanisms of action of nitrogen-containing bisphosphonates (N-BPs), such as ibandronate,

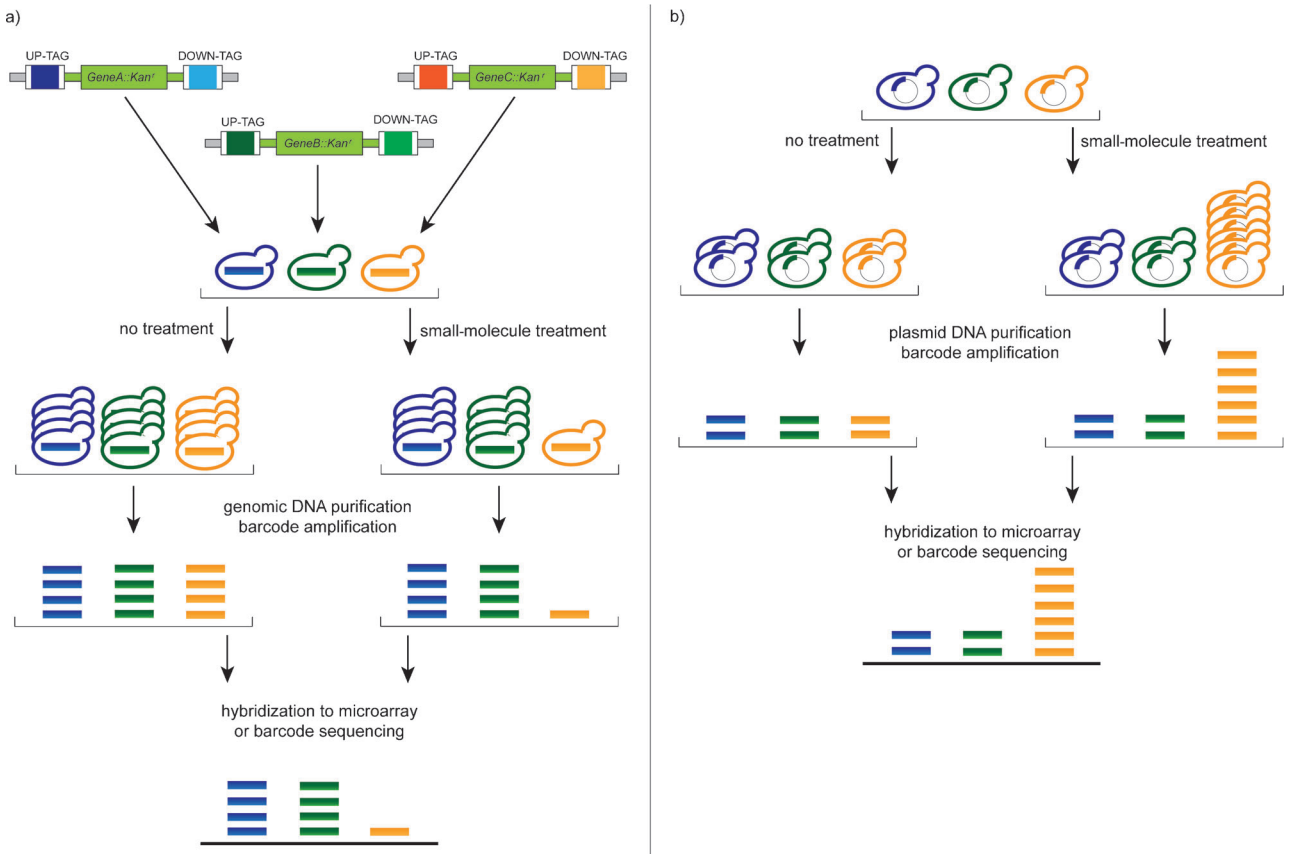


Figure 11. Yeast-based approaches to target identification. a) Principle of Haploinsufficiency profiling/homozygous profiling (HIP/HOP). b) Multicopy suppression profiling (MSP).[12b,132]

alendronate and risedronate (Table 1, entry 40).[137] Alternatively HIP can be used to search for haploproficient genes which lead to a fitness benefit instead of analyzing haploinsufficient genes. Huang et al. searched for haploproficient genes for cisplatin (Table 1, entry 41)[138] and doxorubicin (Table 1, entry 42)[139] treatment. Both drugs are major components of chemotherapy regimens but induce a high frequency of resistance. For cisplatin the approach identified several genes with distinct functions in nucleotide metabolism, mRNA catabolism, RNA Pol-II-dependent gene regulation, and vacuolar and membrane transport systems. For doxorubicin the small ubiquitin-related modifier pathway was shown to be an important factor for resistance.[138,139]

For the microtubule-targeting drug peloruside HIP could not identify the direct target protein. However a HOP approach revealed genes involved in regulation of mitosis and cell cycle, protein synthesis, transport, secretion, RNA processing, and steroid biosynthesis.[140]

The completion of a genome wide deletion collection of *Schizosaccharomyces pombe*[141] allowed to apply the principle of drug-induced haploinsufficiency in this fission yeast as well.[142] Thereby a more comprehensive screening for medically relevant compounds was enabled, since 454 genes in *S. pombe* are not conserved in *S. cerevisiae* but have a human orthologue. A library of 2815 gene-deleted strains was investigated concerning the mechanisms of the anti-tumor and adverse effects of bortezomib (Table 1, entry 43). Nineteen mutants with gene deletions showed synthetic lethality with bortezomib and could be clustered according to their function to the ubiquitin/proteasome pathway, to nuclear/chromatin proteins and nuclear transport, to vesicular traffic, to amino acid and vitamin metabolism, and to RNA metabolism. Of the 19 identified genes, 13 have human orthologues.

The creation of a sequence-verified organism and platform-independent TagModule collection compatible with Gateway cloning that can be readily adapted to any DNA tagging strategy[143] allowed the rapid generation of tagged mutants by transposon mutagenesis in diverse microorganisms. Hence the principle of drug induced haploinsufficiency could be applied to the pathogen *Candida albicans* and Sec7p was identified as target of brefeldin A (Table 1, entry 44).[143]

Multicopy suppression profiling (MSP) is a complementary gene dosage-based approach to HIP and HOP used to identify molecular targets (Figure 11b).[144] This method is based on the assumption that overexpression of a small molecule target results in increased tolerance towards this drug. Luesch and co-workers established a genome wide overexpression screen in *S. cerevisiae* and showed that strains overexpressing PKC1 and a subset of downstream kinases conferred resistance to the kinase inhibitor phenylaminopyrimidine (Table 1, entry 45). Furthermore they found proteins of two additional pathways genetically interacting with PKC1 signaling. A related approach by Butcher and coworkers was based on a pooled population of approximately 3900 yeast

strains each overexpressing a different protein. Resistant strains could be identified by a microarray technique. As proof of concept they confirmed TOR1 as target of rapamycin as well as several other genes involved in transcription, translation, and cell-cycle regulation.[145] Recently a similar approach was described for *S. pombe*.[146] A set of approximately 5000 different fission yeast ORFeome-expressing strains in a background lacking drug efflux pumps was used to analyze the mode of action of etoposide (Table 1, entry 46).

A powerful approach was developed by combining HIP and HOP with MSP on a single TAG microarray.[147] Examining the effect of increasing and decreasing gene dosage simultaneously improves the sensitivity and specificity of small-molecule target identification, as the probability that a gene product is a cognate target is higher when the deletion strain is sensitive and the overexpression strain is resistant to the drug. Hoepfner et al. made use of this principle in the identification of the target of the natural product cladosporin (Table 1, entry 47).[148] HIP revealed lysyl-tRNA synthetase as target protein that was confirmed by an overexpression experiment. Furthermore the examination of *S. cerevisiae* mutants that confer cladosporin resistance revealed non-synonymous mutations that were predicted to map near the binding pocket for ATP in lysyl-tRNA synthetase.

The molecular barcode technology was recently improved by switching from microarray-based analysis[149] to next-generation sequencing. Barcode analysis by sequencing ("bar-seq") surpassed barcode microarray hybridization in sensitivity, dynamic range, and detection limit.[150] However, most genetics-based approaches do not yield individual targets but rather pathways, as single deletions may influence a whole pathway and several deletions may show the same phenotype because of pathway changes. Thus, additional experiments are necessary to confirm the mode of action. Furthermore HIP and HOP depend on cell growth and thus are especially suitable for identifying target proteins important for oncology and antifungal applications. Typically, clear results from HIP/HOP studies are obtained for enzyme inhibitors or when the compound causes a loss of protein function. HIP will fail to identify the target if the coding gene is duplicated. In such cases, however, HOP might help to identify related or connected genes. Additionally, the method might be hampered by the drug target not being encoded in the yeast genome or the effect on the drug target might be rescued by a redundant pathway. Notably, the yeast cell wall may pose a significant hurdle for entering of small molecules into yeast. In addition, yeast and related model organisms have several drug efflux pumps which may efficiently reduce compound concentration. Such hindrances can be overcome by knocking out drug transporters[151] or enhancing drug accumulation in *S. cerevisiae* by repressing pleiotropic drug resistance genes.[152] Giaever et al. established a structure-based accumulation model (SAM) to overcome the general resistance of *Caenorhabditis elegans* to pharmacological perturbation.[153] This computer-based model can be applied to other organisms and can help to prioritize available molecules thus improving screen effectiveness.

The production of short-hairpin (shRNA) vector libraries has facilitated genome-wide screens using RNAi in mammalian cells.[154] Brummelkamp et al. introduced a collection of 23742 different pRETRO-SUPER vectors designed to target 7914 human genes for suppression by RNAi into MCF-7 cells to gain insight in the mechanism of action of nutlin-3 (Table 1, entry 48).[154] This large-scale shRNA barcode screen demonstrated the feasibility of this approach to reveal complete understanding of drug action. A similar approach was described by Burgess et al.[155] They chose shRNAs targeting a set of known cancer-relevant genes ("cancer 1000") to investigate resistance to doxorubicin in a well characterized mouse model of lymphoma. Carette et al. developed a haploid genetic screen in human cells by using insertional mutagenesis to generate null alleles in the 7 KBM7 chronic myeloid leukemia cell line haploid for all chromosomes except chromosome 8.[156] Application of this global gene-disruption strategy served to elucidate the entry pathways of tunicamycin[157] and the Ebola virus.[157] A comparable method was introduced by Elling et al.[158] They generated mammalian haploid embryonic stem cells from mouse blastocysts and mutagenized the haploid cells with retroviruses containing a reversible gene trap. Using this haploid embryonic stem-cell system a genome-wide screen for genes involved in ricin toxicity was performed. Luesch et al. uncovered the importance of FGFR-mediated signaling in modulating apratoxin A (Table 1, entry 49) activity by transiently transfecting a mammalian expression cDNA collection (27000) thereby enabling a genome-wide overexpression screen.[159]

2.7. Expression-Cloning-Based Methods

2.7.1. Three-Hybrid Systems

The yeast two-hybrid system is a powerful tool to identify protein–protein interactions.[160] By analogy the yeast three-hybrid system was established to detect interactions between small ligands and protein receptors.[161] This approach employs three hybrid molecules: the "fish", the "bait", and the "hook". The "bait" is a covalently linked heterodimer of two small-molecule ligands. If one ligand binds its receptor which is fused to a DNA-binding domain ("hook") and the other ligand binds to its receptor fused to a transcriptional activation domain ("fish") a reporter gene will be expressed which enables the selection of yeast cells with the relevant receptors (Figure 12a).

First pioneering experiments using a hybrid ligand of dexamethasone and the immunosuppressant FK506, the glucocorticoid receptor attached to a LexA DNA-binding domain and the FK506 binding protein attached to the transactivation domain of the bacterial protein B42 showed the feasibility of this approach.[161] Becker and co-workers studied the limitations of the yeast three-hybrid system focusing on the targets of ATP-competitive active site kinase inhibitors.[164] The screening of a purvalanol B (Table 1, entry 50) methotrexate fusion compound against a cDNA library and yeast cell arrays displaying selected polypeptide open reading frames identified known target proteins as well as novel targets. Furthermore it was demonstrated that the yeast three-hybrid approach is applicable to micromolar and nanomolar inhibitors.[164] Chidley

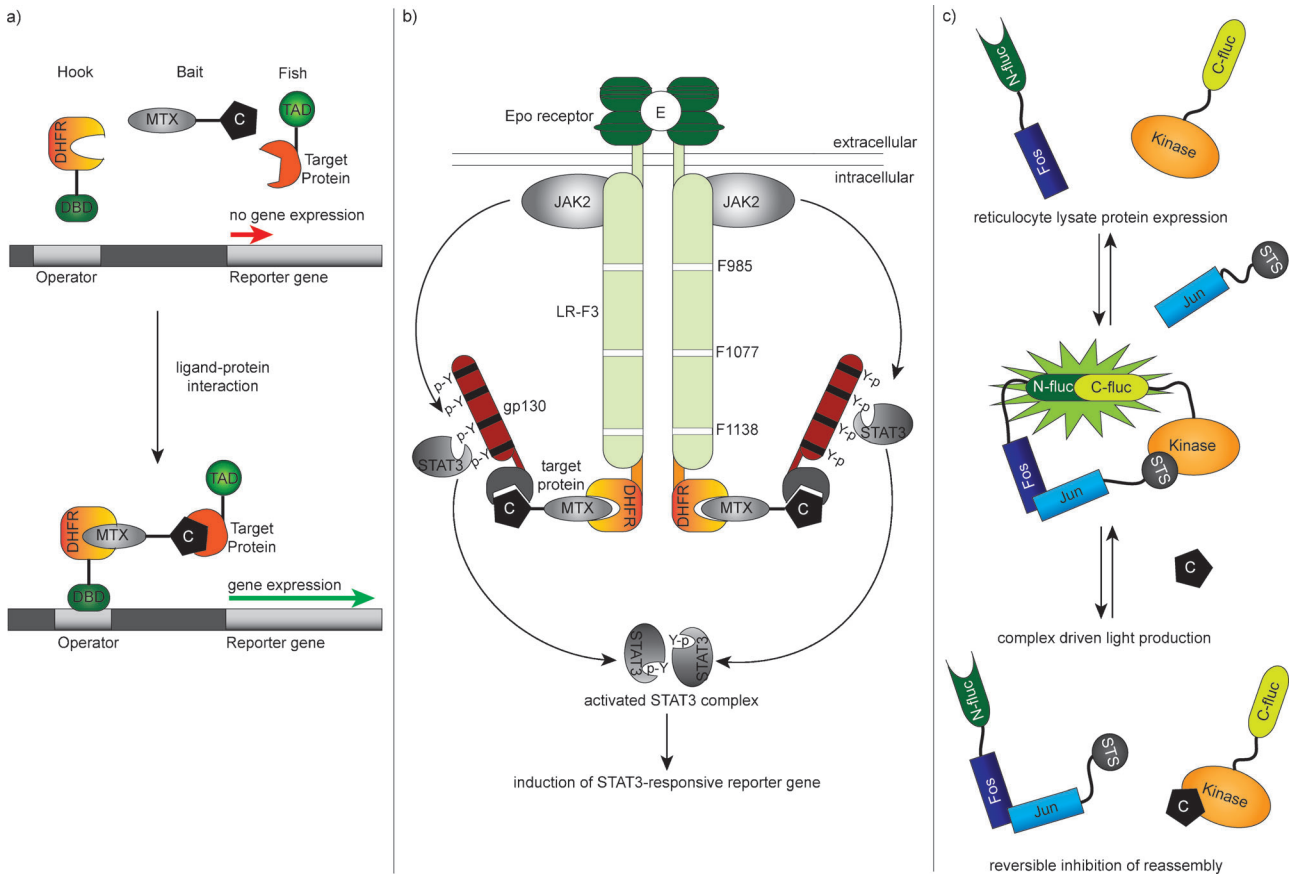


Figure 12. Three-hybrid systems. a) Yeast three-hybrid system.[161] b) Mammalian small-molecule protein interaction trap (MASPIT).[162] c) Coiled-coil enabled split-luciferase three-hybrid system.[163] DHFR: dihydrofolate reductase, DBD: DNA-binding domain, MTX: methotrexate, C: compound, TAD: transactivation domain, fluc: firefly luciferase, STS: staurosporine.

et al. developed a yeast three-hybrid method in which the compound of interest is coupled to O^6-benzylguanine (BG).[151] From this construct the compound of interest is transferred to the SNAP-tagged LexA DNA-binding domain inside the yeast cell. The BG derivative can be coupled to a variety of functional groups and enables the attachment to glutathione beads via a GST-SNAP fusion protein allowing the direct confirmation of identified target proteins by subsequent affinity chromatography. Notably, the sensitivity of the yeast strain was increased towards a broad range of compounds by deleting three genes coding for drug transporters to increase compound concentration in yeast cells. The method was used to identify the first nonkinase target of erlotinib, and to show that PDE6D binds atorvastatin (Table 1, entry 51). In addition, the anti-inflammatory drug sulfasalazine (Table 1, entry 52) was shown to bind to the enzyme sepiapterin reductase (SPR). In follow-up experiments it could be proven that inhibition of SPR by sulfasalazine led to reduced tetrahydrobiopterin levels and decreased NOS (nitric oxide synthase) activity in the gastrointestinal tract which may explain the role of sulfasalazine in the treatment of inflammatory bowel disease.

The yeast three-hybrid system has the advantage that cDNAs coding for the target proteins can be directly isolated and identified. However, if the target protein is not properly folded and posttranslationally modified or if it is an integral membrane protein, or part of a protein complex it will not be identified. Also the hybrid ligand must be permeable to yeast cells and it must retain its functional activity. Some of these limitations may be overcome by the mammalian small-molecule protein interaction trap (MASPIT, Figure 12b).[162] This three-hybrid system is based on the cytokine-receptor-associated JAK (Janus-activated kinase)-STAT (signal transducer and activator of transcription) signal transduction system. It consists of *Escherichia coli* DHFR fused to a chimeric cytokine receptor including the ligand binding domain of the erythropoietin receptor and the cytoplasmic domain of a mutated leptin receptor which has no functional STAT3 recruitment sites. The potential target proteins are expressed as fusions of gp130 which contains STAT3 binding sites. The DHFR protein provides an interaction site for a methotrexate–small-molecule hybrid. An interaction of the small molecule with its protein target restores an Epo-dependent activation of JAK2/STAT3 signaling and induces the expression of a STAT3-responsive reporter gene. As the interactions occur in the cytoplasm no nuclear translocation of fusion proteins or small molecules is necessary. Caligiuri et al. demonstrated the feasibility of this approach by using the FK506 analog AP1867, the ABL kinase inhibitor PD173955, the CDK2 inhibitors RGB-285961 and RGB-

286147, the CDK and GSK3 inhibitor RGB-285978, and the carbonic anhydrase inhibitor E7070.[162]

Another approach addressing at least the problem of cell permeability is the coiled-coil enabled split-luciferase three-hybrid system introduced by Jester et al. for profiling of protein kinase inhibitors.[163] Their approach employs a cell free translation system and consists of the N-terminal part of firefly luciferase attached to the Fos coiled-coil peptide and the C-terminal part of the luciferase attached to a kinase. Addition of the heteroconjugate composed of the pan-kinase inhibitor staurosporine conjugated to the Jun coiled-coil peptide lead to reassembly of the active luciferase. Kinase inhibitors can be identified by decrease of the luciferase signal caused by displacement of the modified staurosporine (Figure 12c). For a proof of principle the kinases PKA, AKT1, FGFR1/FLT2, and PIM1 were tested for the inhibition by 80 known kinase inhibitors. All four kinases showed distinct inhibitory profiles.

2.7.2. *Display Technologies*

Display systems physically link of a polypeptide's phenotype and its corresponding genotype.[165] Among the available display techniques, phage display is most widely adopted (Figure 13a). In phage display the genes encoding proteins of interest are cloned into bacteriophages to form fusion proteins with the phage coat protein which are expressed on the phage surface. Random sequences of peptides[166] or cDNA of human tissues can be displayed on phage particles.[167] By exposing the phages to and binding to an immobilized compound target proteins can be identified through binding. Target identification is achieved by amplification and sequencing of the phage DNA. The main advantage of this approach is the ability to amplify selected phage particles by transfection into *Escherichia coli* thus enabling additional rounds of target identification with phage populations enriched in phage particles displaying the target protein. This "biopanning" even allows the detection of normally only weakly expressed proteins. Sche et al. used biotinylated FK506 and identified the known target FKBP12 from a human brain cDNA library.[168] Further studies revealed hNopp140 as molecular target of doxorubicin,[169] and Bcl-2, NSC-1, and NFX1 as targets of paclitaxel (Table 1, entry 53),[170] Ca^{2+}/Calmodulin as target of the curcumin derivative HBC (Table 1, entry 54),[171] UQCRB in mitochondrial complex III as target protein for terpestacin (Table 1, entry 55)[172] and TACC3 as novel target for bisphenol A (Table 1, entry 56).[173] Piggott and Caruso could identify the human ribosomal protein S25 as the molecular target for the anticancer drug kahalalide F (Table 1, entry 57) using three different T7 phage-displayed human disease cDNA libraries.[174] By fingerprinting the selected clones with the frequent base cutter *Hin*f1 many clones could be screened very quickly, thereby facilitating the analysis of a far greater number of phages without the need for sequencing. For the target identification of paclitaxel[170b] and cyclosporin A[175] a photoaffinity coupling method was used in which a reactive carbene formed by UV irradiation immobilizes the compound nonspecifically.

Phage display offers several advantages. Most notably, problems due to protein abundance can be overcome by amplification of the target protein. The direct identification of

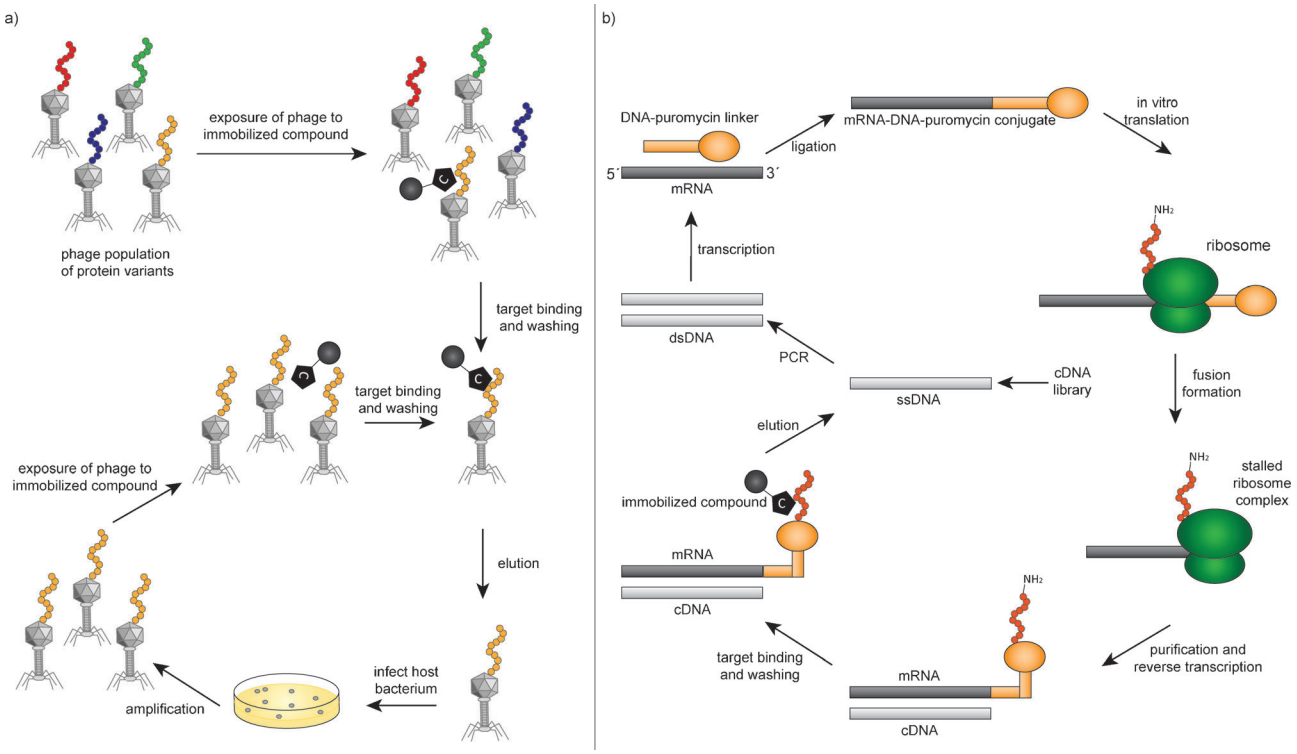


Figure 13. Display technologies for target identification. a) Principle of phage display. b) Principle of mRNA display.[11a]

target proteins from cDNA in the phage is possible and subdomains responsible for binding can be determined when the expressed cDNA is shorter than full length. Disadvantages are the lack of posttranslational modifications in proteins and the potentially improper folding of the expressed polypeptides. Furthermore, the display of membrane proteins may be problematic and the coding sequence of the target protein has to be in-frame with the coding sequence of the phage coat protein.

In mRNA display a newly translated mRNA molecule is covalently linked to its encoded polypeptide by puromycin. Puromycin is a nucleotide-amino acid chmiera that mimics the 3′ end of tyrosyl-tRNA (it mimics adenosine and tyrosine).[176]Puromycin has a non-hydrolysable amide bond and interferes with translation by premature peptide release. In vitro translation of these constructs results in protein–mRNA fusion molecules that are purified and reverse transcribed into cDNA templates for further amplification.[177] The mRNA-displayed protein library is then incubated with a small molecule immobilized on a solid support. Bound constructs are eluted and amplified by polymerase chain reaction (PCR) which results in an enriched drug-binding library that can be used for further selection (Figure 13b).

Although mRNA display in principle can be applied to every organism from which mRNA can be isolated and furthermore bypasses as an in vitro technique several limitations like problems in cloning and expression, and toxicity of the displayed peptides it was to date only applied in a proof of concept study. McPherson and co-workers constructed a library of mRNA-displayed proteins from human liver, kidney, and bone marrow transcripts, and selected against biotinylated FK506.[178] Recently, the successful target identification by mRNA display selection on a microfluidic chip was reported.[179] Nucleophosmin was identified as the target protein of the phthalimide derivative 2-(2,6-diisopropylphenyl)-5-amino-1*H*-isoindole-1,3-dione (TC11, Table 1, entry 58).

3. Target Confirmation

The methodologies for target identification serve to create a hypothesis for a potential mode of action that has to be demonstrated by confirmation experiments. If several proteins are target candidates they need to be prioritized according to their known functions and in relation to the compound-induced phenotype. For this prioritization the design and implementation of proper control experiments to distinguish non-specific binding is essential. Further it should be considered that identified proteins might not be direct targets but part of a protein complex containing the target. Protein interaction databases provide information on protein–protein interaction networks.[180]

Usually the identification of proteins that bind to bioactive small molecules is confirmed with specific antibodies in immuno-blotting experiments after affinity isolation of the target protein. The enrichment of the target protein is then compared with a control sample which can be obtained from a pulldown experiment with an inactive derivative[17–18,53a,55,62] or a competition experiment using an excess of unmodified compound.[17,47,51a,57,64]

Determination of the binding affinity of a small molecule for its putative target provides a strong evidence for target confirmation. Several methodologies, primarily used to examine protein–protein interactions, have been successfully employed for small-molecule–protein interaction studies, mainly surface plasmon resonance[18b,53a,b,83] and isothermal calorimetry (ITC).[15b,51b,53b,56] For surface plasmon resonance, the compound needs to be immobilized on the chip surface and thus is modified, which might decrease potency and might result in lower affinities compared to the unmodified small molecule. For ITC, unlabeled small molecules are used which allows to determine K_D (dissociation constant) values for the label-free compound. Further techniques to assess binding affinities include fluorescence polarization,[17] enzyme-linked immunosorbent assay (ELISA),[62] homogeneous time-resolved fluorescence (HTRF),[181] amplified luminescent proximity homogeneous assay (ALPHA),[182] and microscale thermophoresis (MST).[183] All of these methods require a purified target protein and mostly a modification of the active compound (e.g. with a fluorophore, biotin or a linker with functional group for immobilization). Dadvar et al. employed label-free compounds for UV/Vis absorption, fluorescence emission and ^{15}N NMR spectroscopy to confirm the interaction of PF-3717842 (Table 1, entry 59) with the phosphatidyl ethanolamine binding protein 2 (PEBP2).[184] Circular dichroism (CD) spectroscopy detected the conformational change in a protein's structure upon ligand binding.[15b] Also the thermal stability of a protein in unbound and ligand-bound state provides information on ligand binding since the small molecule might stabilize the protein's native confirmation. This will increase the thermal stability of the protein which can be detected by CD spectroscopy.[56] By means of the DARTS method also ligand-induced stabilization of a protein can be monitored.[98] Further methods like ultracentrifugation, microscale equilibrium dialysis[185] and electrochemical sensing[186] were reported for determination of protein–ligand interactions. In addition, the physical interaction of a small molecule with its target protein can also be detected in cells using Foerster resonance energy transfer (FRET)-based fluorescence life-time imaging microscopy (FLIM), which requires a fluorescently tagged protein and a compound labeled with a fluorophore.[17,187]

Binding of a small molecule to a protein not necessarily will modulate its functions. For that reason, further functional experiments are required for target confirmation. Clearly, when the target possesses enzymatic activity, the modulation of this activity should be assessed in an enzymatic assay.[28,56,188] Although an enzymatic activity assay can be employed to further optimize the activity for a given target, the resulting structure–reactivity relationship may significantly deviate from the data obtained with the initially used cell-based assay because of cell permeability, compound stability, cellular distribution and engagement in different pathways. If already chemical modulators are reported for the putative targets the phenotypes they induce should be compared with the query compound.[53c,57c] Related to this,

for the targets RNA interference (RNAi) and/or cDNA overexpression should be performed to analyze whether this might phenocopy the influence of the bioactive small molecule in the studied system. RNAi is widely used today to knock-down proteins of interest. For several identified target proteins the impact of small interfering RNA (siRNA) or short hairpin RNA (shRNA) might have been already reported and will help to prioritize given targets.[17,28] When RNAi for a gene encoding a putative target protein is not reported or present studies focused on different cellular characteristics, target confirmation should include this methodology as has been demonstrated for numerous target proteins.[47,53b,57c] Furthermore, usually less compound is required to produce a certain phenotype after a knock-down of the target gene.[57b] Knock-down efficiency must be quantified by either real-time PCR or by a specific antibody. Further, off-targets of the siRNA used should be excluded. It should be considered that a gene knock-down might not produce the same phenotype as a small molecule. Whereas RNAi will reduce the amount of the protein in the cells, small molecule will not. Target overexpression analysis could support the target hypothesis since it might abrogate the influence of the small molecule[51b,53a,189] or phenocopy the ligand activity if the compound acts as an activator.

Fluorophore-tagged small molecules can be employed to examine compound localization within cells and a potential co-localization with the target protein.[33a,47–48] However, the attached fluorophore may change the physicochemical properties of the compound and lead to a different distribution within the cells than the non-labeled small molecule. Mutation analysis or domain mapping of the protein could reveal the mode of ligand binding.[51b,53b] Co-crystallization of the small-molecule-target complex will not only reveal the binding site but also could explain the mechanism of modulation and suggest modification of the compound to improve its binding.[53b,66]

Altogether, target confirmation is as important as target identification. Confirmation of the target should not only include determination of the binding affinity of the ligand but should also be sought in the cellular context of the phenotypic screen. Combination of biophysical, biochemical, cell biological, and structural biology methodologies will contribute to the assembly of the puzzle and, in the end, yield an indicative picture of a compound's cellular activity.

4. Troubleshooting

Although many attempts might be undertaken to identify the molecular targets of an active small molecule, they may fail for several reasons. To identify target proteins using affinity chromatography, the design and the synthesis of proper probes, in case they are needed, is crucial. Structure–activity relationships should provide the information for determining the appropriate site to derivatize. If no structure–activity relationship is available, photo-crosslinking of the compound to a solid surface or attachment of the linker to a different position in the molecule or label-free methodologies should be considered. In order to distinguish non-specific binders, a control experiment should be performed employing either an inactive derivative or a competition strategy. To enrich low-abundant proteins, cell fractionation can be performed. When a simple comparison of isolated protein with the active and control probe cannot identify a target candidate, a quantitative proteomic analysis should be taken into consideration.

Several successful examples of target deconvolution are reported. However, often the proteins identified are highly abundant or have high affinity for the ligand. Moreover, the target proteins should be in a native conformation in the experimental setup to bind to the small molecule. The isolation of membrane proteins, especially of those spanning the membrane multiple times, might pose a big hurdle owing to the hydrophobic nature of these proteins and their low abundance. Moreover, membrane proteins usually need to be in their natural environment to exert their characteristic binding properties and to keep their three-dimensional structure.[32] A trifunctional probe which, in addition to the ligand, includes a protected hydrazine group to react with carbohydrates on glycoprotein receptors and a biotin functionality for affinity purification could help to overcome this problem.[32] It also should be considered, that the molecular targets of small molecules might not be proteins but rather other biomolecules, such as DNA or RNA, lipids, or carbohydrates.[190] In light of this fact, broad exploration of the activity profile of the compound could help to narrow down the possible targets. Different strategies for target identification are established. They should be regarded as complementary and need to be explored separately in each particular study.

5. Case studies

5.1. Target Identification of Adenanthin by Using Chemical Proteomics

Adenanthin (**7**, Figure 14), a diterpenoid isolated from the leaves of *Rabdosia adenantha*, induces differentiation of acute promyelocytic leukemia (APL) cells.[191] It induces APL-like cell differentiation, represses tumor growth in vivo, and prolongs the survival of mouse APL models that are sensitive and resistant to retinoic acid. APL is a unique subtype of acute myeloid leukemia (AML) that is genetically characterized by chromosome translocations involving the retinoic acid receptor. ATRA (*all-trans* retinoic acid) and arsenic trioxide, which are used for treatment of APL, are believed to target and lead to the destruction of the APL-specific fusion proteins. Unlike ATRA, however, adenanthin neither restored the APL-specific fusion protein-disrupted nuclear body, nor cleaved or degraded APL-specific fusion proteins, thus excluding the possibility that adenanthin targets fusion protein complexes to induce differentiation. These observations suggested a novel target to explain adenanthin's mode of action. After establishing a structure–activity relationship of adenanthin, a biotin-tagged adenanthin probe **8** and a negative control **9** were synthesized for target identification by means of a chemical proteomics approach (Figure 14).[192] It

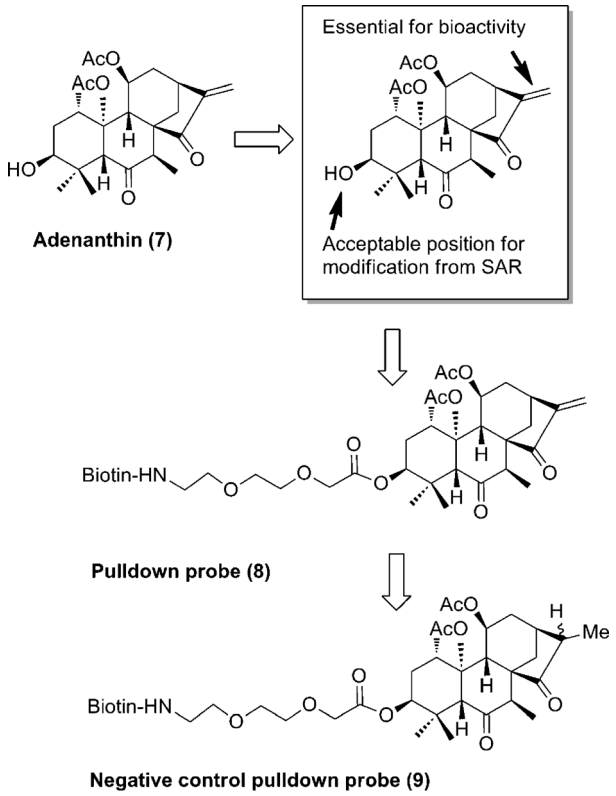


Figure 14. Structural formula of adenanthin (**7**) and adenanthin-based pulldown probes.

was assumed that the α,β-unsaturated moiety in adenanthin is a Michael acceptor that potentially captures nucleophiles, such as cysteines, at its targeted binding site and forms covalent adducts, thus allowing for efficient protein recovery by the pulldown probe. Moreover, reduction of the Michael-acceptor double bond of adenanthin to a single bond abolished its differentiation-inducing activity on NB4 cells, making it a suitable negative control. NB4 cell lysates were incubated with biotin–adenanthin or free biotin, and the bound proteins were isolated with streptavidin-coated agarose beads, followed by SDS-PAGE and silver staining. Only one band, with a molecular mass of approximately 23 kDa, was clearly precipitated by biotin–adenanthin, but not by free biotin. Formation of the band was completely prevented by high concentrations of unlabeled adenanthin, indicating a specific interaction. Mass spectrometry revealed that the adenanthin-bound protein is peroxiredoxin I and II (Prx I–II). Western blotting with selective antibodies against Prx I, Prx II, or the other four Prx isoforms, (Prxs III–VI) showed that NB4 cells expressed all six Prx isoforms, and biotin–adenanthin effectively pulled down Prx I and Prx II, but not Prxs IV–VI. Biotin–adenanthin also effectively bound in vitro the recombinant Prx I and Prx II protein, and this binding was competitively inhibited by higher concentrations of free unlabeled adenanthin.

Immunofluorescence staining with antibodies against Prx I or Prx II and streptavidin-fluorescein isothiocyanate (FITC) also demonstrated that biotin-adenanthin co-localized with Prx I and Prx II in the cytoplasm and nucleus of NB4 cells. Further investigations of MS-fragmentation patters revealed the selectivity for the cysteine residues modified in Prx I and II, respectively. Prx I and Prx II expression levels are upregulated in almost all AML subtypes, and in addition, Prx I and Prx II were reported to be elevated in a number of human cancers and were proposed as potential targets for anticancer drugs.

5.2. Activity-Based Proteome Profiling of Palmostatins

Recently, the ABPP-methodology was employed to elucidate the cellular targets of the Ras-modulating APT1 (acyl protein thioesterase 1) inhibitor palmostatin B (**13**)[187] and the related substrate analogue-based inhibitor palmostatin M (**10**) in live cells (Figure 15).[193] Mutated and consequently active H-Ras and N-Ras are major oncogenes responsible for early development of cellular tumourigenesis and enhanced survival signaling. It has been suggested that attenuation of H/N-Ras activity can be achieved by interfering with its localization at the plasma membrane by affecting its palmitoylation status. Palmostatin B has been shown to delocalize H-Ras in cells and inactivate APTs by reversible covalent modification of the enzyme active site through a β-lactone electrophile, thus acting as slowly hydrolyzed substrates, and therefore good starting candidates for the design of ABPP-probes.[33b] Palmostatin B (**13**) was equipped with an alkyne functionality at the lipid tail (**14**, Figure 15), and two versions of palmostatin M (**10**) were synthesized, in which the alkyne functionality was located either at the lipid tail (**11**) or at the polar head group (**12**) (Figure 15). When IC_{50} values of APT1 inhibition were determined for the three probes, no major differences resulting from the different regioisomeric alkyne labeling patterns were recorded. HeLa cells were incubated with the corresponding probes, washed and lysed and the resulting lysate was treated with Biotin–tetramethylrhodamine (TAMRA) azide in a copper(I)-catalyzed alkyne–azide cycloaddition, followed by enrichment with magnetic streptavidin beads (see Section 2.1.4). Subsequently, in-gel fluorescence scanning was employed to detect labeling events. Proteomic analysis after in-gel tryptic digestion identified APT1, APT2, and PPT1 as major targets. In quantitative western blotting of pulldown experiments, a significant difference in labeling efficiency between the regioisomeric alkyne derivatives (**11** vs. **12**) was noticed. This result suggested that the accessibility of the alkyne for the copper(I)-catalyzed alkyne-,azide cycloaddition when bound to the protein is of importance for the overall efficiency of the pulldown experiment. The proteomics investigation revealed that the isoenzyme APT2 also is targeted by the palmostatins, in addition to PPT1 and APT1. Biochemical inhibition assays of recombinant human APT2 revealed similar IC_{50} values for APT2 compared to APT1. The lysosomal depalmitoylating enzyme PPT1 has no importance for cell signaling, leaving APT1 and APT2 as the primary targets for the palmostatins. Notably, no other intracellular esterases (such as phospholipases A1, A2, C, and D) relevant to Ras signaling were identified as targets for the palmostatins, which confirms the selectivity of the probes. This study verified, for the first time,

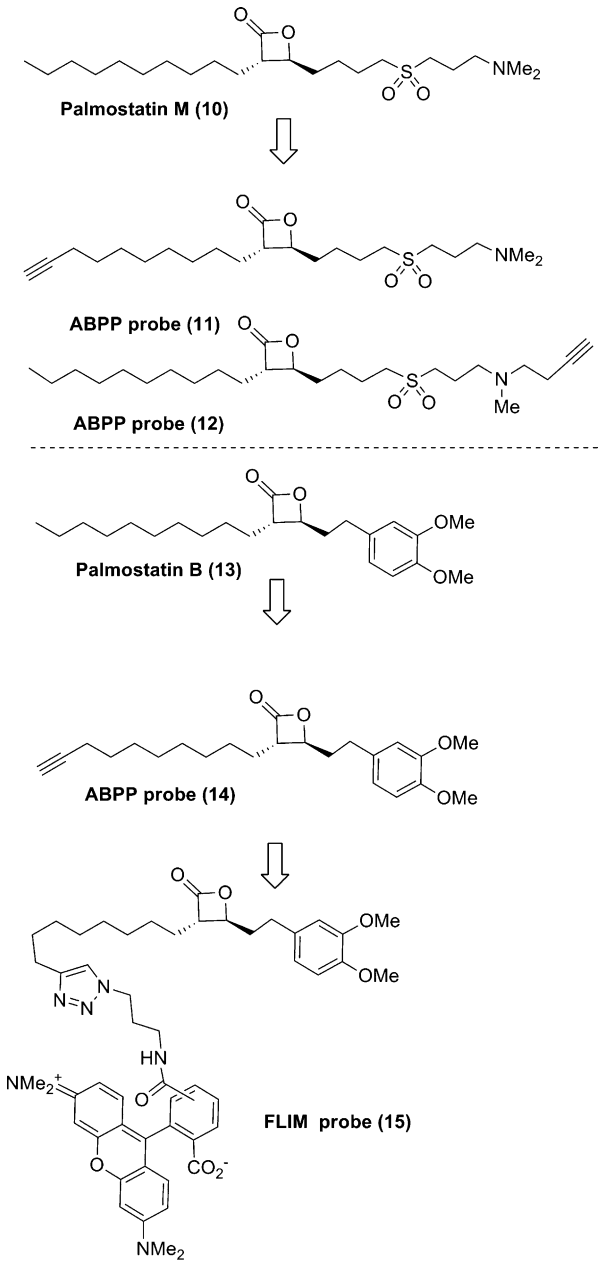


Figure 15. Structural formula of palmostatin M (**10**) and B (**13**) and related probes.

that APT2 is indeed an N-Ras depalmitoylating enzyme. The direct interaction between palmostatin B and APT1 in cells was investigated by fluorescence lifetime imaging (FLIM), in which cells were transfected with eGFP-tagged APT1 and subsequently treated with cell-permeable TAMRA-labeled palmostatin B (**15**, Figure 15).[187] It was found that the lifetime of the emitted fluorescence of the eGFP fluorophore of APT1 was significantly reduced, providing confirmation for a direct interaction through selective quenching by TAMRA. This approach exemplifies the complementary use of proteomics techniques with confocal imaging, allowing for direct correlation in live cells.

5.3. Affinity-Based Monoamine Oxidase Inhibitors

A major challenge in ABPP is to expand the set of probes beyond classical α,β-hydrolases, thus enabling the investigation of a larger set of the catalytically active proteome. In a recent paper, MAO-targeting substances were investigated using an ABPP-approach.[194] Flavine-dependent oxidases consist of a large family of proteins with diverse oxidative functions, such as halogenations, desaturations, and aminoxidations. Monoamine oxidases (MAOs) are flavin adenine dinucleotide (FAD)-containing enzymes, which catalyze the oxidative deamination of several important neurotransmitters, including serotonin, norepinephrine, and dopamine as well as xenobiotic amines. In MAOs, the flavine co-factor is covalently attached to the protein by a cysteine linkage and tertiary *N*-methyl, *N*-propargyl amines, such as deprenyl (**16**) and pargyline (**17**) (Figure 16a), are known to covalently modify the co-factor post oxidation by Michael-addition to the isoalloxazine ring of the FAD (Figure 16b).[195] Deprenyl and pargyline have been used in the cell biological and biochemical characterization of MAOs and are today used as approved drugs in clinical applications. A set of inhibitors (**18**–**19**) derived from pargyline and deprenyl was synthesized,

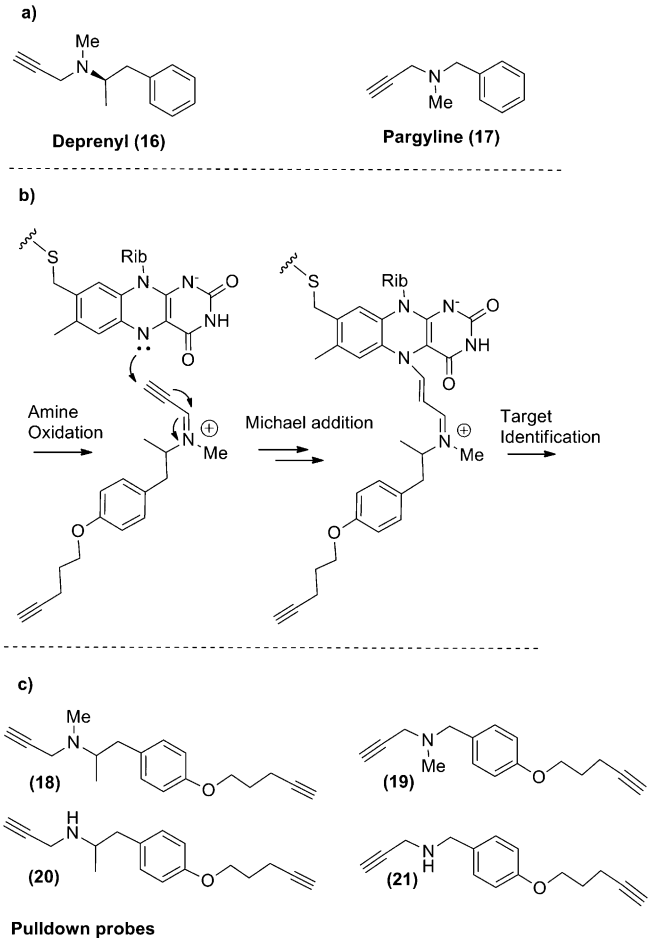


Figure 16. a) Structures of deprenyl (**16**) and pargyline (**17**). b) Labeling of monoamine oxidases. c) Structures of deprenyl and pargyline pulldown probes.

keeping the reactive propargylic amine intact and adding an alkyne tail at the aromatic moiety of the molecules, thus allowing for alkyne–azide cycloaddition visualization in cell lysates (Figure 16c). The probes were reasonably potent for MAO A (IC_{50} 0.93 μM for **17** and 0.04 μM for **19**) and MAO B (IC_{50} 0.11 μM for **16** and 1.76 μM for **18**) inhibitors in vitro. Control compounds **20**–**21**, carrying a free NH functionality did not react. Recombinant human MAO preparations were incubated with the probes, and subsequent attachment of a TAMRA-azide tag by copper(I)-catalyzed alkyne–azide cycloaddition. SDS-PAGE and in-gel fluorescence scanning were employed to detect labeling events. In initial experiments it was demonstrated that both isoforms of MAO (A and B) can be efficiently labeled as main targets by the probes (**18**–**19**) with slightly different isoform preferences, at concentrations as low as 100 nM. All potent probes identified in the initial screening include methyl-substituted tertiary amines, indicating that this feature might be important for efficient labeling, which is in good agreement with the established pharmacophore of known MAO inhibitors. In competition experiments it was shown that ABPP probes compete with MAO-specific inhibitors for the same binding site (flavin cofactor) in the enzyme active site, since pargyline is able to efficiently block MAO labeling by the ABPP-probe. Further, MAO-labeling was demonstrated in tissue homogenates employing deprenyl and pargyline-derived ABPP-probes. These results demonstrate that the developed ABPP system can serve as an effective chemical tool for profiling activity of isoforms of MAO in cell culture, as well as in tissue samples.

5.4. Target Identification of Centrocountins by using Chemical Proteomics

Investigation of a library of indoloquinolizines in a phenotypic screen for mitotic arrest identified compound (**22**) termed centrocountin 1 that caused the formation of multiple mitotic spindles in BSC-1 cells.[17] More detailed investigations identified the compound as a modulator of centrosome integrity. Centrocountin 1 caused formation of fragmented and supernumerary centrosomes, chromosome congression defects, multipolar mitotic spindles, acentrosomal spindle poles, and multipolar cell division. These effects suggested that centrocountin targets the centrosomal machinery. A focused subset of centrocountin derivatives was prepared and on the basis of a detailed structure–activity relationship, pulldown probe **24** and negative control probe **25** were synthesized for target identification by means of a chemical proteomics approach (Figure 17). Compound **24** was immobilized on sepharose beads by NHS-coupling and exposed to HeLa cell lysates. Enriched proteins were released by elution with a tenfold excess of unmodified centrocountin, followed by mass spectrometric identification of the bound proteins, revealing that probe **24**, but not control probe **25**, binds the nucleolar and centrosomal protein NPM and U2 small ribonuclear protein. Knock-down of NPM by RNAi induces fragmentation of centrosomes and impairs chromosome congression and mitotic spindle formation in HeLa cells.[196] NPM is involved in the regulation of centrosome duplication during mitosis and also promotes ribosome biogenesis. On the basis of these data, NPM was confirmed as a target protein. Reversible binding of NPM to **24** was further confirmed by immuno-blotting with an NPM-specific antibody and by concentration-dependent competition between immobilized and non-immobilized centrocountin. In addition, regulation of centrosome duplication by NPM includes complex formation with the nuclear export receptor Crm1.[197] Crm1 RNAi knock-down as well as Crm1 inhibition (by leptomycin B) led to similar defects in chromosome alignment and spindle assembly as observed upon treatment with centrocountin. Investigation of the affinity pulldown experiment by immuno-blotting with a Crm1-specific antibody indeed revealed binding of Crm1 to the immobilized centrocountin probe. Proof for the direct interaction between indoloquinolizine probe **23** and NPM, as well as Crm1, in HeLa cells was established by means of fluorescence lifetime imaging (FLIM) of the donors NPM-citrine and enhanced yellow

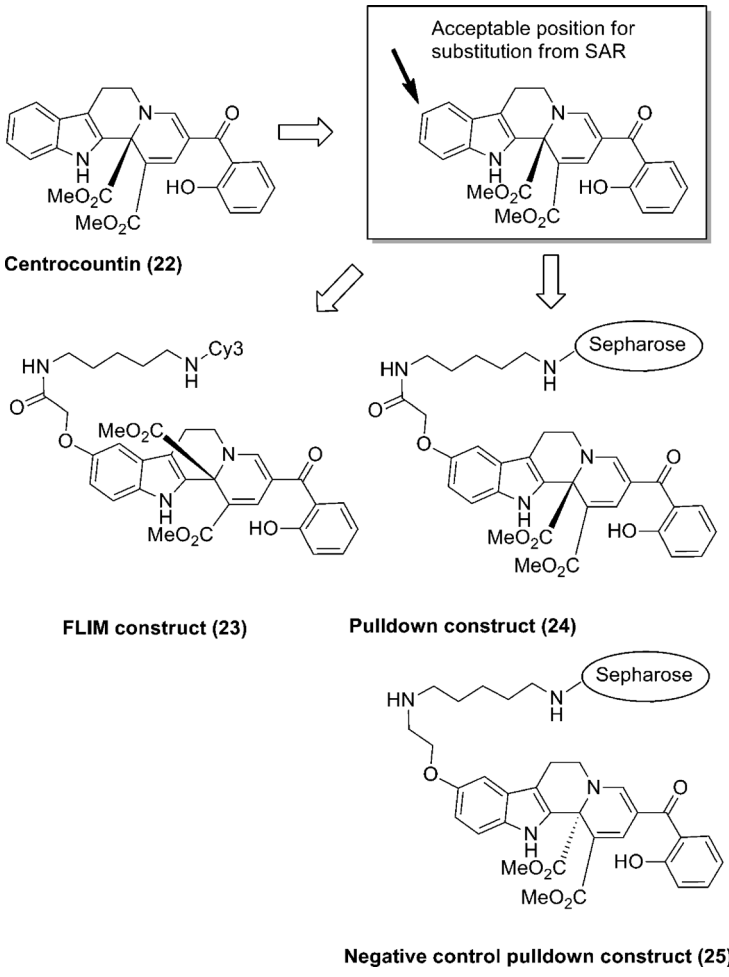


Figure 17. Structural formula of centrocountin (**22**) and probes for FLIM and affinity chromatography.

fluorescent protein (EYFP)-Crm1 after addition of the Cy3-labeled centrocountin (acceptor). These findings revealed that both NPM and Crm1 independently bind centrocountin and that centrocountin independently targets the centrosome-associated proteins NPM and Crm1 in cells. It thereby leads to impairment of centrosome and spindle integrity, chromosome congression defects, and cell-cycle arrest at the M-stage, thus ultimately resulting in apoptosis.

5.5. Target Identification of Tubulexins by Using SILAC

Tetrahydropyrans occur widely in nature and are endowed with pronounced biological activities. For instance the natural product centrolobine and closely related compounds have antibiotic and antioxidative activity. Evaluation of a library of tetrahydropyrans in a phenotypic screen monitoring changes associated with impaired mitosis revealed structurally novel modulators of mitosis termed tubulexins.[62] The most active compound tubulexin A (**26**, Figure 18) targets the chromosome segregation process and the vinca alkaloid binding site of α,β-tubulin. DNA staining and FACS analysis revealed that treatment with tubulexin A resulted in a virtually complete arrest of BSC-1, HeLa, and MCF-7 cells in the G2M phase. Tubulexin A arrested HeLa cells at concentrations as low as 2 μM, whereas tubulexin B (**27**; Figure 18) and tubulexin C (**28**) were less potent, and at least 10 μM was required for cell-cycle arrest in HeLa and BSC-1 cells. Induction of apoptosis by tubulexin A was confirmed by elevated activity of caspase-3 and caspase-7 in HeLa and BSC-1. Based on these results, tubulexin A was chosen for further characterization. From the tetrahydropyran library, a reasonable structure–activity relationship could be established, suggesting affinity probe **29** and negative control probe **30** for synthesis (Figure 18). The resulting probes were employed in affinity pulldown experiments using the quantitative SILAC (stable isotope labeling by amino acids in cell culture) approach to identify potential target proteins.[71] The biological activities of the selected biotinylated tubulexin probes **29** was characterized in relevant cellular assays, and was found to largely retain activity, compared to unmodified tubulexin A. The biotinylated probes **29** and **30** were immobilized on magnetic streptavidin-coated beads and incubated either with isotope-labeled or unlabeled HeLa cell lysates. After release from the streptavidin by heat denaturation, labeled and unlabeled pulldown samples were combined, and the proteins were separated by SDS-PAGE and identified by means of nano-HPLC-MS/MS analysis. Only proteins for which the binding ratio of quantified labeled SILAC peptides compared to unlabeled peptides was significantly altered in 75% of the replicates were considered potential targets. By means of this method, the chromosome segregation 1-like protein (CSE1L, CAS, exportin-2) and tubulin were identified as potential target proteins with relevance to mitosis. Binding of tubulexin A to CSE1L and tubulin was confirmed using Western blotting after the affinity pulldown experiment, thereby verifying the mass spectrometry result. The binding of CSE1L was reversible as shown by concentration dependent competition between free and immobilized tubulexin A. Investigation of in vitro polymerization of porcine tubulin in the absence of microtubule associated proteins by means of turbidity measurement revealed that tubulexin A inhibits tubulin polymerization. The majority of tubulin polymerization inhibiting molecules bind to the colchicine or the vinca alkaloid binding site. Although colchicine and the tubulexins both contain alkylated phenol groups, thus sharing some structural similarity, tubulexin A does not bind to the colchicine site. Instead, it successfully competes with a fluorescent-tagged analogue of vinblastine for binding to the vinca alkaloid binding site on tubulin in a concentration dependent manner. In addition, a synergistic effect of tubulexin A on tubulin and CSE1L was identified, thus suggesting tubulexin A to be a novel dual-action compound.

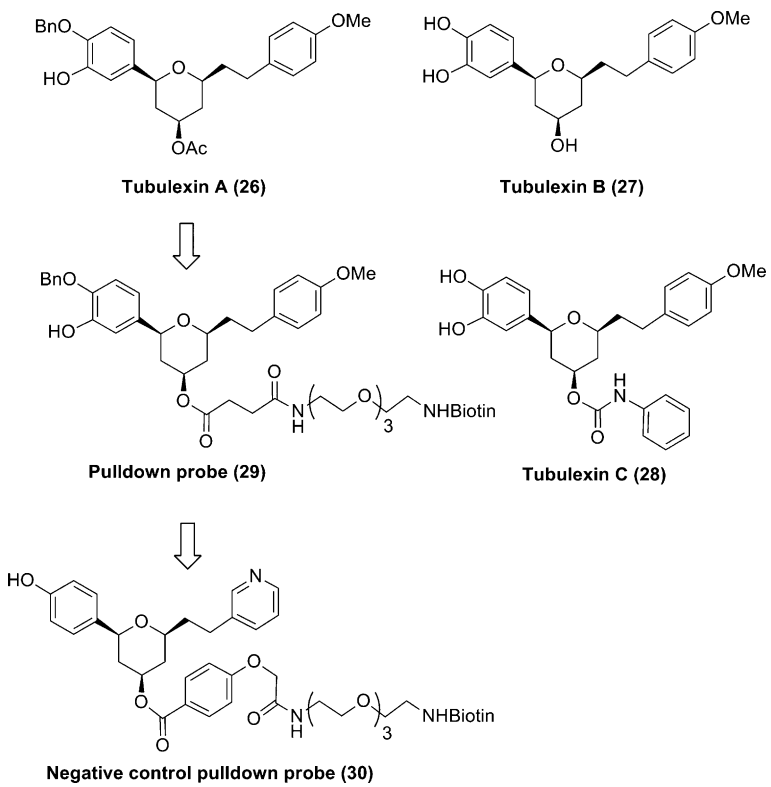


Figure 18. Structural formula of tubulexin A–C and probes for pulldown.

5.6. Ligand-Directed Protein Tagging of Fusiccocin Binding Proteins

Protein labeling by approaches based on ligand-directed covalent transfer of a tag to a target protein has been developed over the last years. Several different approaches have been explored, so far aryl sulfonates ("tosyl") is the prime and most successful approach. Fusicoccane A-derived probes were employed for profiling of 14-3-3 proteins in cancer cells.[69] 14-3-3 proteins

play a critical role in kinase-dependent signaling pathways through scaffolding protein–protein interactions (PPIs) with various phosphorylated ligands. To date, the 14-3-3 interactome has been poorly characterized. A ligand-dependent 14-3-3-detection technique would enable elucidation of 14-3-3-related intracellular signaling networks. A large number of intracellular ligand proteins that possess 14-3-3 consensus motifs have been identified, including proteins involved in signaling and cell-cycle control. The diterpene fusicoccin A[198] (**31**, Figure 19) activates the plant plasma membrane H+-ATPase through the formation of a ternary complex with plant 14-3-3, which leaves the host plant's stomata pores open, ultimately leading to wilting. Based on the crystal structure of 14-3-3ζ-bound fusicoccin A,[199] a functionalized probe was developed, where the complex 5-8-8-diterpene skeleton was kept intact and the linker carrying the reactive sulfonale ester was introduced in a spacer connected to the exocyclic sugar moiety (Figure 19). A spacer with an appropriate length was chosen to bring the sulfonylmethylene moiety near a potentially nucleophilic residue in 14-3-3ζ (His164). At the terminus of the transferable part of the reactive construct, a dansyl or a BODIPY fluorescent reporter group was introduced, thus resulting in probes **32** and **33**. In parallel, a control compound was synthesized, not carrying the selectivity-guiding fusicoccane core (**34**, Figure 19).

Labeling was evaluated by incubating the probes **32** and **33** with recombinant 14-3-3ζ in the presence or absence of the PMA2 phosphopeptide, followed by SDS-PAGE. Probes **32**/**33** labeled 14-3-3ζ in the presence of the peptide whereas no apparent labeling was detected when the peptide was absent during the experiment. This demonstrates that the formation of the ternary complex is the determining factor for the labeling reaction. The yield of the labeled protein increased with the concentration of the phosphopeptide until equimolarity with 14-3-3ζ was reached, thus supporting a 1:1:1 model for the formation of the ternary complex.

The selectivity of the reaction between **33** and 14-3-3σ, which has an Asn instead of a His at position 164, resulted in very weak labeling, suggesting that labeling predominantly involves His164 of 14-3-3ζ. This site specificity was further confirmed by using site-directed mutagenesis of 14-3-3ζ. Substitution of His for Ala at position 164 abolished labeling. A control experiment involving probe **34** did not result in labeling, thus further confirming that the fusicoccin anchor is essential for binding to 14-3-3. IThus, the fusiccocin A derived probes are capable of labeling 14-3-3ζ in a site-specific and ligand-dependent manner. It can be expected that ligand-directed protein labeling will become increasingly more important in the future.

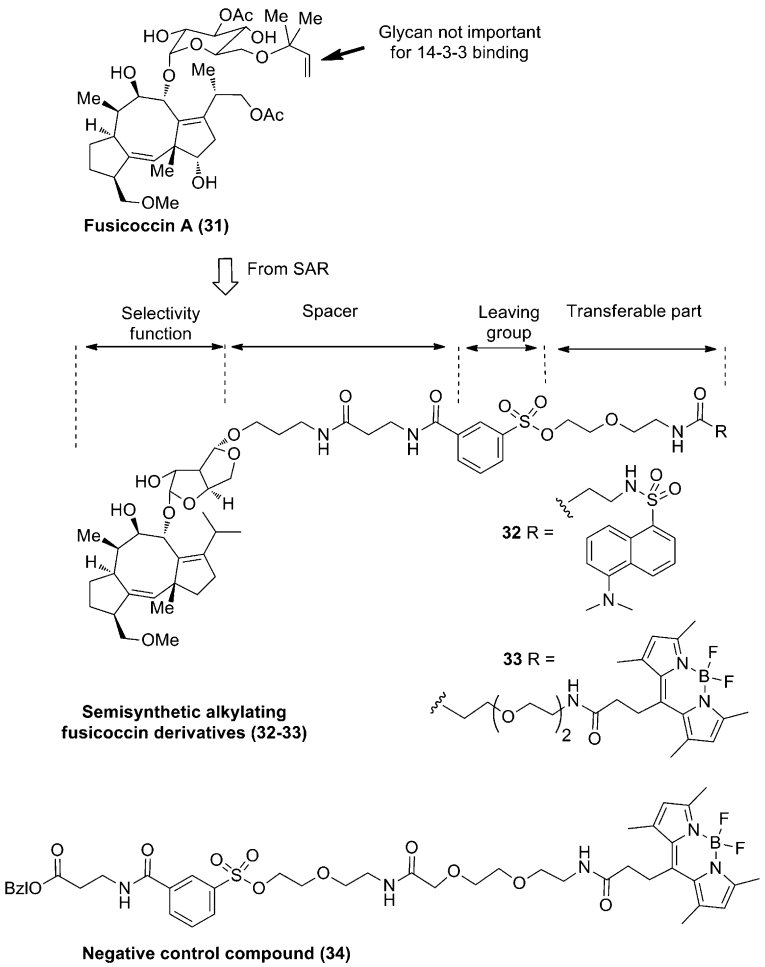


Figure 19. Structural formula of fusicoccin A (**31**) and probes for ligand-directed protein tagging.

5.7. ABPP-Fluorophosphonate Profiling

Target identification cannot only be carried out with specific probes, but also with unselective reporter probes. This approach has been especially successful for α,β-hydrolases in general and serine hydrolases in particular.[200] In a typical experiment, an inhibitor is added to cells or cell lysate, allowed to incubate for a given time, then a non-specific reporter fluorophosphonate (FP) carrying a fluorophore or reporter group (**35**, FP-rhodamine, **36**, FP-alkyne, **37**, FP-biotin, Figure 20) is added, which covalently binds to all α,β-hydrolases at their active sites. Subsequent in-gel imaging, comparing inhibitor treated and non-treated samples reveals a lack of gel-bands in the cases where the active site of a α,β-hydrolase was blocked. In contrast to the more reactive fluorophosphonates, carbamates can have high selectivity for individual serine hydrolases. These inhibitors have proven to be valuable research tools, and, in certain cases, have become approved drugs, for example, rivastigmine. Despite extensive screening efforts, potent and selective carbamate inhibitors have been identified for only a fraction of the mammalian serine hydrolases, pointing to the need for alternative chemical classes of serine hydrolase inhibitors. In a recent example, the classical fluorophosphonate reporter was replaced with "clickable" and tunable ureas (**38**, Figure 20), which were used for the proteome-wide evaluation

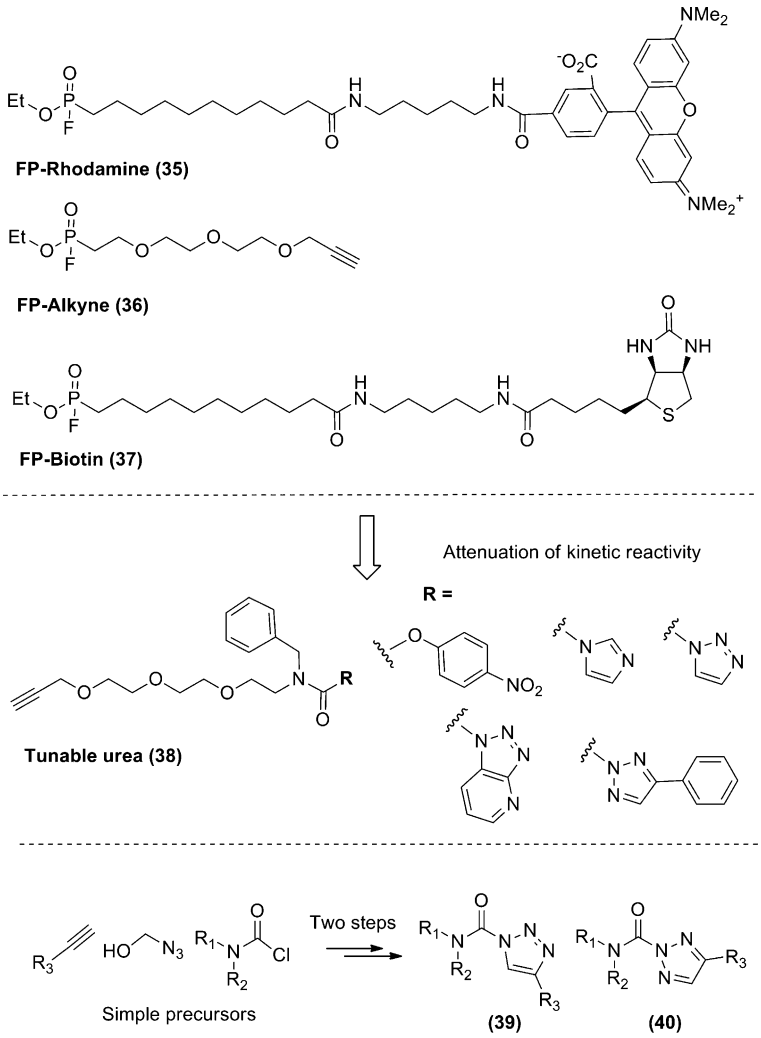


Figure 20. Structures of reporter probes for ABPP.

of a small library of carbamates based on a triazole urea chemotype.[201] Replacing the FP probes with ureas attenuates reactivity, thus allowing for the specific targeting of serine hydrolases over other α,β-hydrolases in the proteome.

The 1,2,3-triazole urea chemotype (**39**–**40**, Figure 20) displays unique reactivity for serine hydrolase inhibition. The properties are tunable across the enzyme class and specific selectivity for individual members can be achieved. A simple and efficient click-chemistry approach to create substituted triazole ureas was developed. In a two-step procedure, substituted alkynes were reacted with in situ-formed organoazide (azidomethanol) to yield 4-substituted triazoles, these were subsequently carbamoylated to give triazole urea products, typically as a 3:1 mixture of N^2- and N^1-carbamoylated regioisomers. Using this strategy, a 20-member library of 4-aryl and 4-alkyl triazole derivatives was prepared. The library was screened at 10 and 100 nM in mouse T-cell lysate, and several highly potent and selective inhibitors of several serine hydrolases were identified. A selected number of inhibitors were investigated in vivo in mice with excellent results. A good correlation between in vitro and in vivo activity for the 1,2,3-triazole ureas could be identified, thus suggesting usefulness of the compound class in vivo. Taken together, the 1,2,3-triazole urea chemotype stands out as a versatile scaffold for the development of potent and selective inhibitors for the poorly characterized serine hydrolases, which are abundant in the human proteome. One limitation of the fluorophosphonate and related detection techniques are that they are most suitable for detecting covalent inhibitors. In case of reversible, competitive inhibitors, the result of the global labeling will strongly depend on the incubation time, since an apparent off-rate of the noncovalent ligand determines the amount of protein labeling by the reporter group. A recent investigation has addressed this issue successfully by employing covalent reporter probes with tunable kinetic reactivity.[202]

5.8. Target Identification with Capture-Compound Mass Spectrometry (CCMS)

Target identification may not be limited to identification of single protein targets of small molecules. Often it extends to target profiling, where complete families of proteins binds with varying affinities (from high to low) to a small-molecule probe. If a classical pulldown experiment is carried out on such a multi-protein target family, most of the medium- and low-affinity binding partners will be lost during the washing process. A good example is compounds that specifically inhibiting a subset of kinases, which are of growing importance in cancer therapy. Development of improved therapies should ultimately also take low- and medium-affinity drug interactions into account, thus making improved target profiling technology important. Recently, a novel technology for the isolation of protein subfamilies based on trifunctional molecular probes called capture compound mass spectrometry (CCMS) has been developed (Figure 4b).[43] Capture compounds (CCs) carry three major functionalities: a selectivity function, which is the small-molecule ligand to be profiled, a photoactivatable crosslinking function, and a sorting function. In a typical experiment, a cell lysate is incubated with the capture compound for a certain time, allowing for equilibrium binding of the target proteins by the selectivity group, subsequent photolysis leads to the generation of a photoactivated intermediate, such as a nitrene or carbene, within the reactivity group, resulting in a covalent crosslink between the capture compound and the target proteins. By employing the sorting function (e.g. biotin) and streptavidin-coated magnetic beads as affinity matrix, the resulting protein complex can be pulled down. It has been demonstrated that a major advantage of this covalent cross-link between capture compounds and target proteins is that even weakly interacting proteins in the micromolar affinity range are irreversibly bound to the capture compound. The covalent nature of the

capture compound-target bond allows stringent washing conditions to be applied to reduce unspecific background.

The CCMS technique was used to characterize the staurosporine kinase binding profile.[203] First, from the structure–activity relationship of staurosporine (**41**) and the published crystal structures of the inhibitors in the respective target proteins, a suitable linking point for the trifunctional scaffold was identified, and the capture compound was synthesized (**42**) (Figure 21).

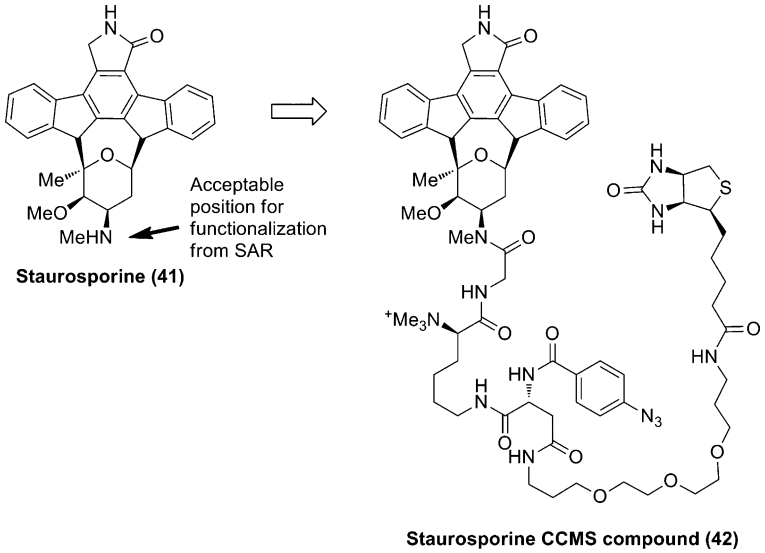


Figure 21. Structural formula of staurosporine (**41**) and staurosporine CCMS probe (**42**).

To demonstrate the validity of the CCMS approach, it was first validated that the staurosporine-derived capture compound formed a covalent bond with the purified recombinant human PKA-catalytic subunit and with proteins in HepG2 lysate. After incubation, photo-crosslinking and enrichment employing streptavidin beads, the captured proteins were enzymatically degraded using trypsin, the peptides were analyzed by LC-MS/MS and identified by database searches of the annotated peptides. To distinguish between specific target proteins and unspecific background proteins, competition control experiments were carried out in parallel to the capture assays by pre-incubation of the lysate with a large excess of free staurosporine. It was found that capture efficiency was satisfying in both cases. Carrying out the identical experiment in HepG2 lysate gave similar results, where one average capture experiment identified approximately 300 proteins. As staurosporine is an ATP-competitive kinase inhibitor, there may be promiscuity between proteins binding different nucleotides. Another possibility is that a very high affinity of some kinase-interacting proteins to the kinases or cross-linking because of the close proximity of the bound kinase may have led to the enrichment of some of these proteins. After data analysis, approximately 100 proteins were identified as kinases. Of the identified 100 kinases, 56 were serine/threonine kinases, while nine members of the tyrosine kinase family and four dual specificity kinases were also identified. Assignment of the affinity range of the captured kinases toward staurosporine revealed that three kinases have subnanomolar affinity, 12 kinases are in the range of 1–10 nM, 15 kinases are in the 10–100 nM range, and seven kinases have an affinity of 100 nM to 1 μM. Two kinases have an affinity of more than 10 μM. This shows that the majority of the kinases captured by staurosporine–CCMS had affinities in the nanomolar range, however, also kinases with affinities in the micromolar range were accessible.

5.9. Target Identification of Piperlongumine by Using SILAC

During a reporter-gene-based screening campaign for apoptosis-inducing small molecules, piperlongumine (**43**, Figure 22) was identified as a primary hit.[75] Piperlongumine is a natural product from the plant *Piper longum L*, from which extracts are known to have cytotoxic effects.[204] Upon evaluation of the effect of piperlongumine on healthy cells versus cancer cells, it was found that cell death in cancer cells was induced regardless of whether p53 was mutated or not, whereas healthy cells showed little or no sensitivity to the compound. It was concluded that piperlongumine has a cancer-specific cytotoxic effect, suggesting that the target of piperlongumine is a product of the malignant transformation. Next, piperlongumines apoptotic effect was investigated by quantitative immuno-blotting for wild-type (wt)

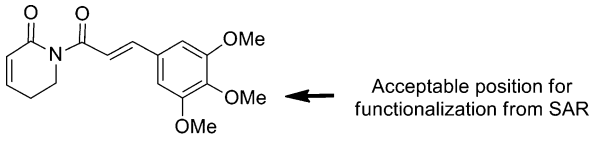


Figure 22. Structural formula of piperlongumine (**43**).

p53 expression, which revealed that p53 expression levels increased upon treatment with piperlongumine. Further, piperlongumine-treated cells showed attenuated levels of pro-survival proteins, suggesting that piperlongumine induces cell death or apoptosis in cancer cells by modulating the expression levels of proteins in the survival pathways. Piperlongumine was found to have suitable physico-chemical properties, as well as good oral uptake profile and reasonable plasma half-life for in vivo studies in mice. Piperlongumine-treatment showed a significant effect on blood-vessel development in tumor xenografts in mice, as well as in inhibiting secondary tumor formation. In a mammary gland cancer model in transgenic mice, piperlongumine was found to be more efficient than paclitaxel. To clarify piperlongumine's mode of action, affinity enrichment with SILAC and quantitative proteomics was used to identify the target proteins and their complexes. A suitable functionalized derivative was

prepared with guidance from the structure–activity relationship for piperlongumine. Twelve interaction partners of piperlongumine were identified in two separate cell types. Seven out of the twelve identified proteins were known to be involved in the cellular response to oxidative stress caused by elevated reactive oxygen species (ROS). Glutathione S-transferase 1 (GSTP1) had the highest score, followed by carbonyl reductase 1. Several of the proteins identified are known to be part of a common complex, suggesting that the affinity purification may have identified direct as well as indirect targets. The results indicate that, by binding to proteins known to regulate oxidative stress, piperlongumine modulates redox properties in the cell, leading to increase in ROS. Further, piperlongumine was shown to interact directly with purified recombinant GSTP1 and inhibit its activity and also that it can lead to a decrease in reduced glutathione (GSH) levels and an increase in oxidized glutathione (GSSG) levels in cancer cells. In contrary, piperlongumine treatment of normal cells did not increase GSSG levels. By measuring the concentration of individual ROS species with selective fluorescent probes, it was found that hydrogen peroxide and nitric oxide, but surprisingly not superoxide anion, were among the ROS species induced by piperlongumine in cancer cells. In contrast to the results in cancer cells, piperlongumine did not cause an increase in ROS levels in normal cells. This selective induction of ROS in cancer cells makes piperlongumine stand out from other small molecules that affect ROS levels. Taken together, it suggests a novel strategy for eliminating cancer cells by targeting the ROS stress-response pathway.

5.10. Identification of Tankyrase Inhibitors by Using iTRAQ

The evolutionarily conserved Wnt/β-catenin signal transduction pathway regulates numerous biological processes. One important regulatory function of the Wnt pathway is the strictly controlled proteolysis of the downstream effector β-catenin by the β-catenin destruction complex. In the absence of Wnt pathway activation, cytosolic β-catenin is kept in the phosphorylated state and targeted for degradation. On Wnt stimulation, the β-catenin destruction complex dissociates, leading to the accumulation of nuclear β-catenin and transcription of Wnt-responsive genes. Inappropriate activation of the Wnt pathway has been observed in many cancers, thus making it an attractive target for the development of Wnt modulators. In a high-throughput screen using a Wnt-responsive luciferase reporter assay, XAV939 (**44**, Figure 23) was identified as a small molecule inhibitor of the Wnt/β-catenin pathway.[85] Interestingly, XAV939 decreased β-catenin abundance, but significantly increased β-catenin phosphorylation, suggesting that XAV939 promotes the phosphorylation-dependent targeting of β-catenin for destruction by increasing the activity of the destruction complex. To find out how XAV939 increases the activity of the destruction complex, it was investigated if compound treatment altered the protein levels of known Wnt pathway components. The protein levels of axin1 and axin2 were strongly increased after XAV939 treatment, and a strong increase in axin-GSK3β complex

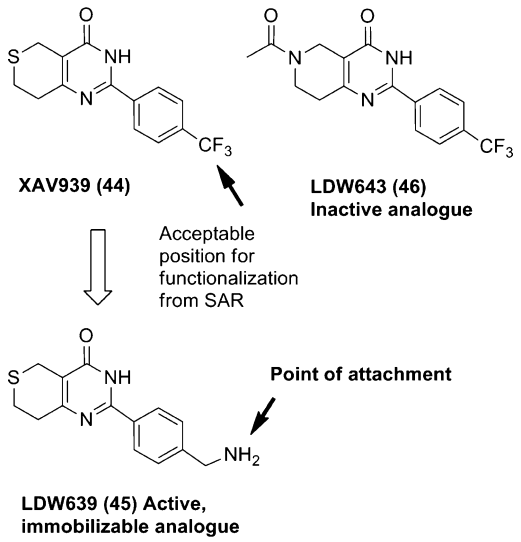


Figure 23. Structural formula of XAV939 (**44**) and related pulldown probes.

formation was observed, presumably because of increased axin protein levels. These findings suggested that XAV939 increases the concentration of the axin-GSK3ß complex, thereby promoting phosphorylation and degradation of β-catenin.

To identify the cellular target(s) for XAV939 which results in axin protein levels, a triple iTRAQ quantitative chemical proteomics approach was employed.[74] XAV939 was functionalized according to the structure–activity relationship, resulting in LDW639 (**45**, Figure 23), and subsequently immobilized to capture cellular proteins from cell lysates. HEK293-lysate was treated with an excess amount (20 mM) of XAV939 (**44**), the inactive analogue LDW643 (**46**, Figure 23) or DMSO. 18 proteins were significantly and specifically competed off with soluble XAV939 (**44**), including several poly(ADP-ribose) polymerases (PARP1, PARP2, TNKS1, TNKS2) as well as known PARP1 substrates (co-precipitated). In a completion experiment it was shown that XAV939 (**44**), but not LDW643 (**46**), blocks TNKS binding at 10 μM and blocks PARP1/2 binding at 1 mM. To determine loss-of-function phenotypes of the PARP family members copying the effect of XAV939, siRNA-mediated depletion was carried out. Co-depletion of TNKS1 and TNKS2 pheno-copied the effect of XAV939 by increasing the protein levels of axin1 and 2, whereas PARP1/2 co-knock-down showed no phenocopying effect.

The increase in axin protein levels in response to XAV939 treatment was suggested to be due to modulation of translation or protein stability. XAV939 treatment prolonged the half-life of endogenous axin 2 in cells, somehow protecting it from proteasome mediated degradation. In contrast, co-treatment of XAV939 with the proteasome inhibitor MG132 significantly diminished axin 1 and axin 2 polyubiquitination, suggesting that XAV939 may stabilize axin by preventing its polyubiquitination. Taken together the study suggests tankyrase to be an important player for axin stability, and identifies axin to be a limiting factor in the β-catenin

degradation complex, thus suggesting axin as a key player in the Wnt signaling cascade.

5.11. Target Identification of Thalidomide by Using Chemical Proteomics

The drug thalidomide (**47**, Figure 24) has been in clinical use as a mild sedative and is best known as the causative agent of severe birth defects during the 1960s. Thalidomide is still in clinical use for the treatment of leprosy and multiple

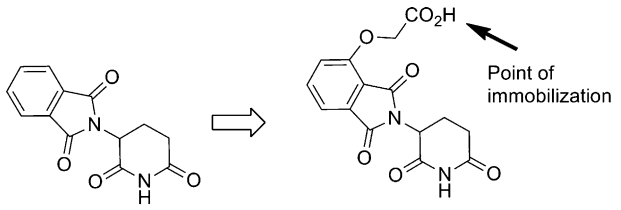


Figure 24. Structural formula of thalidomide (**47**) and FR259625 (**48**).

myeloma. Little is known about how the thalidomide-induced developmental defects are caused. Previous studies have suggested oxidative stress and its anti-angiogenic action as a possible cause of teratogenicity. To identify the origin of thalidomide's teratogenicity, a target profiling was recently carried out.[205] Thalidomide-induced birth defects display a distinct phenotype, which in particular includes severe deformation and shortage of limbs. The carboxylic acid derivative FR259625 (**48**, Figure 24) showed reproducible teratogenicity in zebrafish, similar to unmodified thalidomide. FR259625 was covalently conjugated to beads and incubated with human HeLa cell extracts. After extensive washing, bound proteins were eluted with free thalidomide, and subjected to SDS gel electrophoresis and subsequent silver staining. Two proteins were specifically eluted. When free thalidomide was added to extracts before incubation with the beads, the yields of these proteins were reduced, suggesting that these proteins specifically interact with thalidomide. The 127 and 55 kD proteins were subjected to trypsination and mass spectrometry and were identified as CRBN and damaged DNA binding protein 1 (DDB1). Further, protein identities were confirmed by immuno-blotting. Recombinant FLAG-tagged CRBN, but not (His)-tagged DDB1 bound to thalidomide beads, however, when co-added to the beads, both CRBN and DDB1 were bound, suggesting that DDB1 binds through an interaction with CRBN. The possible role of CRBN in thalidomide teratogenicity was investigated in animal models. Thalidomide is a teratogen in rabbits and chicken, but not in mice and rats. Zebrafish is an excellent model organism because of the rapid progress of embryo development, the transparency of the embryo, and knockdown of genes is straightforward. To examine the effects of thalidomide on zebrafish development, embryos were transferred to media containing different concentrations of thalidomide at 2 h post fertilization and allowed to develop for 3 days. It was apparent that in thalidomide-treated embryos, development of fins and otic vesicles was disturbed, whereas other aspects of development were not generally affected, compared to the untreated control. The phenotype could be partially rescued by injection of the corresponding zCRBN mRNA. To be taken into account is that thalidomide is rapidly hydrolyzed and metabolized to more than a dozen products in vitro and in vivo. Thalidomide and its products may have identical or different molecular targets, however, the current findings suggest that thalidomide exerts teratogenic effects by binding to CRBN and inhibiting the associated ubiquitin ligase activity. Since thalidomide is used for the treatment of multiple myeloma and leprosy, identification of its direct target may allow rational design of more effective thalidomide derivatives without teratogenicity.

6. Conclusion, Outlook, and a Suggested Workflow

Identification and confirmation of bioactive small-molecule targets is a crucial, often decisive step both in academic research and in pharmaceutical application. Knowledge of target proteins is required to arrive at valid conclusions and insights in chemical–biology research employing small-molecule perturbation of biological systems for analysis. By analogy, knowledge of the targets is advantageous in drug discovery to develop efficacious and safe drugs as well as for approval by the authorities.

The techniques and examples described in this Review show that target identification, in principle, is feasible, and, indeed the number of successful examples steadily grows. This progress has become possible through the development and availability of several new experimental techniques with varying applicability. At the outset of a target-identification project it often is unclear which technology will be best and it is left to trial-and-error or is determined by the general scientific expertise of and the infrastructure available to the individual research groups as to which approach is applied. A Review of the successful examples reported recently demonstrates that currently the application of affinity isolation employing suitable small-molecule probes (pulldown) and subsequent mass spectrometric analysis of the isolated proteins appears to be the most powerful and most frequently applied technology for target identification. This leading role is largely due to the rapid development of powerful mass spectrometry techniques and the availability of increasingly efficient mass spectrometers and analysis software.[206]

For target confirmation a variety of cell biological and biophysical techniques have opened up novel opportunities. The use of advanced microscopy techniques using FRET pairs and time-resolved measurements (fluorescence lifetime imaging (FLIM)), fluorescence anisotropy measurements, improvements in isothermal titration calorimetry, protein melting-point determination, and surface plasmon resonance techniques, as well as RNAi and target overexpression are of particular importance. It can be anticipated that improvements in these experimental techniques, and of course the development of novel methods will further enrich the toolbox for target identification and confirmation. We expect that an

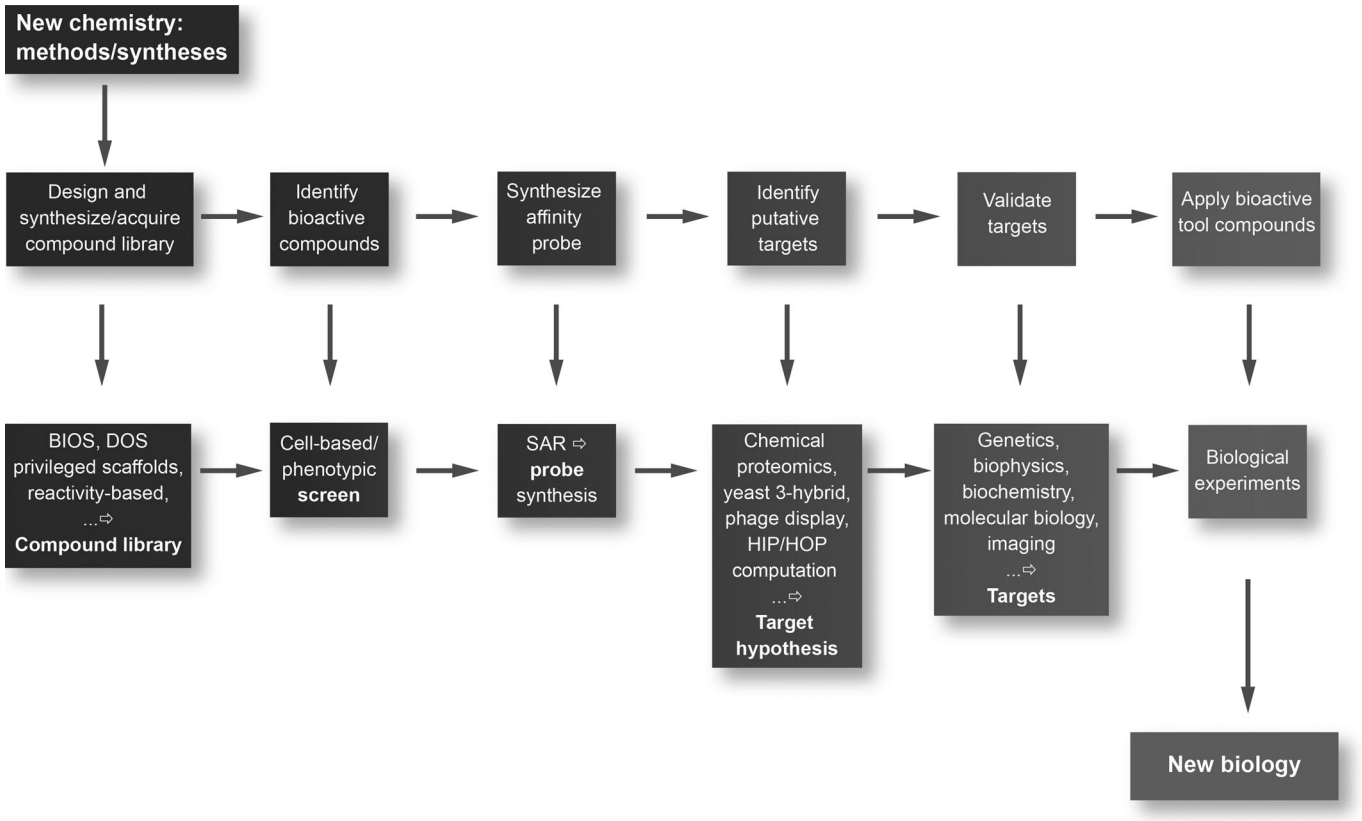


Figure 25. Workflow for design and synthesis of compound libraries, screening for bioactive small molecules, and target identification and confirmation.

important focus will be on the development of quantitative and label-free techniques. Prominent recent examples are the establishment of quantitative, mass spectrometry methods based on comparison of signal intensities recorded for individual peptides for target identification,[88,89] the introduction of microscale thermophoresis (which potentially can be label-free)[183b] and the drug-affinity responsive target stability (DARTS) approach[98] for target confirmation.

Despite the advances made and to be expected, target identification and confirmation for the foreseeable future will remain to be a difficult, labor-intensive, and time-consuming endeavor. Reasons are, for instance, that proteins may be expressed in only small amounts or they may be membrane bound, or stable only under the conditions given in cells and will unfold under the conditions of typical proteomics experiments. For both challenges currently no fully reliable methodology is available. Affinity pulldown experiments also usually result in the isolation of protein complexes which then calls for confirmation or deconfirmation of multiple proteins by means of biophysical, cell-biological, and genetic methods to identify the functional targets.

In particular, current methodology is not set up to reliably identify other biomolecules such as DNA, regulatory RNA, carbohydrates, or lipids as targets. In this context the identification of 3β-hydroxysterols as target of theonellamide F is a particularly impressive example.[190b]

Clearly, the development of novel powerful technology and in particular of reliable widely applicable, generic approaches is in high demand. The lack of generic approaches and workflows for target identification and confirmation is often the major obstacle for the chemist, biologist, or chemical biologist devoid of prior experience and expertise in the field but potentially in possession of a powerful compound active in a relevant cell-based screen.

To provide some guidance for rapid entry into the field and based on our own experience in the identification of the cellular targets of small bioactive compounds, below we sketch a workflow we typically apply in our related research projects (Figure 25). This workflow is not meant to be generic but rather represents one of a variety of possible experimental approaches. It centers on the application of chemical proteomics, that is, the use of affinity protein isolation as the key step to generate hypotheses concerning potential target proteins. We have also investigated phage-display technology but were less successful with this method. Investigation of haploinsufficiency profiling in yeast strains for hypothesis generation is currently ongoing.

Research into the identification of novel bioactive small molecules with potentially new mode of action and identification of their targets requires access to a compound library to be screened. Thus the initial step is library design and assembly, for which different approaches are established (see Ref. [5] and references therein). Our in-house library comprises approximately 150000 compounds synthesized and acquired according to the principles of biology oriented synthesis,[5] diversity,[207] drug-likeness,[208] coverage of estab-

lished drug-target classes, and occurrence in nature (natural products). In library design and synthesis, naturally the input from chemistry is most pronounced and important. Early biology input is strongest and most important in the selection, the set-up, and establishment of cell-based or in vivo screens to be carried out. Such screens should be as meaningful as possible and the recorded phenotypes or the pre-determined phenotypes to be induced by the small-molecule hits should ideally provide a link to existing knowledge to narrow down the number of potential target proteins to be identified subsequently. We stress that establishing a medium- or even high-throughput cell-based screen can be a major time-consuming task very often underestimated in chemical biology research.

Based on identified hits a structure-activity relationship for identified active compound classes is delineated which may result in further rounds of synthesis and biological investigation of focused compound collections. The resulting expanded structure–activity relationship then will allow sites within the active compounds to be identified for linker attachment (i.e. without loss or only with acceptable loss of bioactivity in the screen) and for the design of an inactive control probe. We note that identification of a truly inactive control probe for cell-based screens may be difficult. The chosen probe may not be sufficiently active to induce a desired phenotype, yet still it may bind the same target proteins albeit with lower affinity, thereby erroneously indicating unspecific interaction between active probe and putative target proteins. Notably, this may hold true for enantiomeric probe pairs for which it is assumed but probably not separately demonstrated that the enantiomer of a given bioactive compound should be inactive.

Active positive probe, and inactive negative probe is then applied simultaneously in isolation and potential target proteins identified. As mentioned above, we usually employ chemical proteomics for this step which includes mass spectrometric identification of putative target proteins. To this end, the probes typically are equipped with an affinity tag and immobilized, usually by employing the biotin/streptavidin pair, or they are covalently linked by means of amide bond formation to resin beads. As an additional complication, the protein band of interest may appear only as a weak signal when using cellular lysates rather than living, intact cells because of possible inactivation of the corresponding protein during cell rupture. Alternatively, phage-display techniques may be employed, and the use of yeast-3-hybrid technologies[151] and of yeast haploinsufficiency profiling[133] are very valid and powerful complementary approaches for this step. In the chemical proteomics approach, target proteins will ideally bind to the active probe only, but not to the inactive probe, and it should be possible to perform competition experiments either under the conditions of the pulldown experiment or subsequently for release from an affinity matrix. In addition, potential targets and successful competition may directly be monitored by comparison of gels obtained from the isolated protein mixture without mass spectrometric investigation. Thus, if a band is visible in the experiment with the positive probe, but not the negative probe, and it can be competed away with unmodified compound, it very likely represents a valid target.

The chemical proteomics experiment typically will yield a list of proteins that need to be validated or devalidated. This confirmation includes a variety of biochemical experiments (immuno-blotting with specific antibodies), enzyme activity assays, genetics techniques (RNAi experiments, induction of the same or a similar phenotype; induction of higher or lower sensitivity to the compound in question by knock-down or overexpression), and biophysical techniques monitoring direct interaction between the active compound (as opposed to an inactive control compound) and the target protein (e.g. fluorescence polarization, ELISA, isothermal titration calorimetry, thermophoresis, and surface plasmon resonance based assays).

A very powerful and assuring technique is the direct monitoring of the interaction between the probe and the target protein in cells, for example, by means of monitoring FRET pairs by fluorescence lifetime imaging (FLIM).

Finally, functional assay should be employed to confirm that the identified hit interferes with the biological functions of the putative target. To obtain reliable results in such functional assays it may be necessary to determine selectivity of the bioactive compound for the target in question, that is, selective inhibition of one or a few enzymes of a given protein class. For identification and confirmation of putative target proteins computational methods may advantageously complement the described experimental techniques, for example, by means of investigations of chemical similarity between identified hits and structurally similar compounds for which target proteins are already known.

Once the target is identified subsequent in vitro experiments may be employed to increase potency and selectivity of the hit, if possible including structural-biology investigations.

Following this workflow we have successfully identified protein–ligand pairs spanning different fields of biology. As stressed above the route is not generic, but prototypic at best, and the path through it still may be winding and riddled with multiple set-backs which need to be overcome by additional and alternative experiments. However, it may be used as an entry to the field, as a guideline for the first own endeavor. We sincerely hope that it may prove to be valid and applicable in many cases.

7. Appendix: Table 1

Table 1: Examples for bioactive small molecules for which cellular targets have been successfully identified.

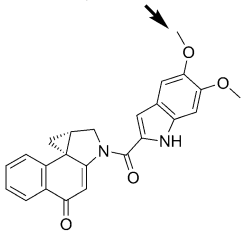
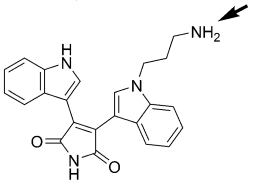
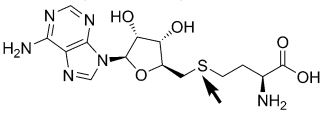
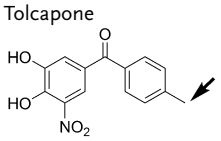
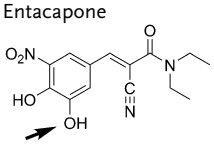
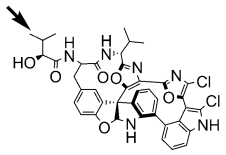
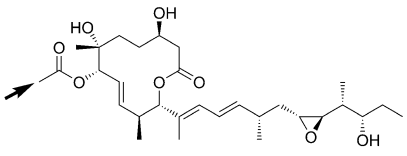
Entry	Compound	Target identification approach	Target(s)	Ref.
1	Duocarmycin SA derivative 	Chemical proteomics Trifunctional biotin–rhodamine probe	ADH1	[28]
2	Bisindolylmaleimide III 	Chemical proteomics FLAG-tagged probe	PKCα, GSK3β, CaMKIIΔ,γ, adenosine kinase, CDK2, NQO2, PKAC-α, prohibitin, VDAC	[36]
3	S-adenosyl-L-homocysteine 	Chemical proteomics CCMS[a]	Methyltransferases	[42d,c]
4	Tolcapone  Entacapone 	Chemical proteomics CCMS[a] (Caprotec) different linker attachment sites	COMT, Mitochondrial proteins binding to tolcapone	[43]
5	Diazonamide A 	Chemical proteomics prefractionation	Ornithine δ-amino transferase (OAT)	[46]
6	Pladienolide B 	Chemical proteomics prefractionation	SAP130 (SFb3 complex)	[48]

Table 1: (Continued)

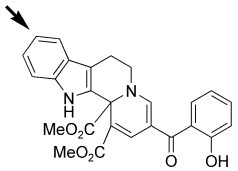
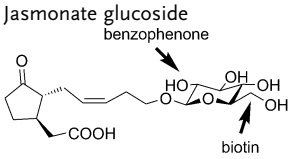
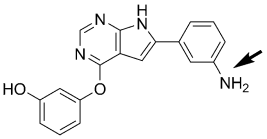
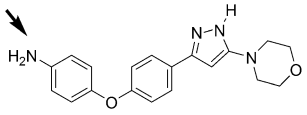
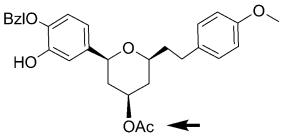
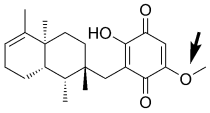
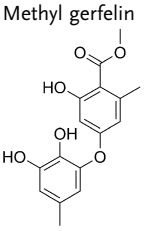
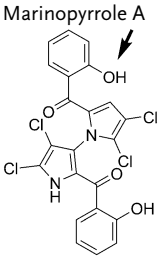
Entry	Compound	Target identification approach	Target(s)	Ref.
7	Centrocountin 	Chemical proteomics	NPM, Crm1	[17]
8	Jasmonate glucoside 	Chemical proteomics	Membrane protein	[54]
9	TWS119 	Chemical proteomics	GSK3β	[55]
10	4-[5-(4-phenoxyphenyl)-2*H*-pyrazol-3-yl]morpholine 	Chemical proteomics	Adenosine kinase (TbrAK)	[56]
11	Tubulexin A 	Chemical proteomics using SILAC[b]	CSE1L, tubulin	[62]
12	Bolinaquinone 	Serial affinity chromatography	Clathrin	[52]
13	Methyl gerfelin 	Chemical proteomics Immobilization by photocrosslinking	Glyoxalase 1 (GLO1)	[66]
14	Marinopyrrole A 	Chemical proteomics Acyl dye transfer	Actin	[68b]

Table 1: (Continued)

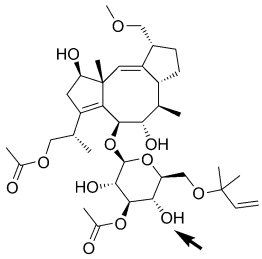
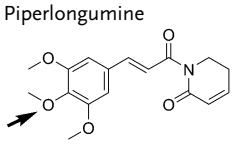
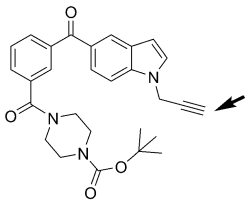
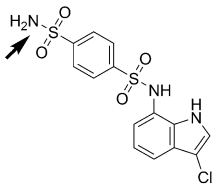
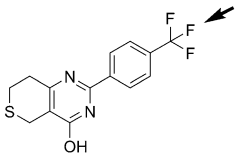
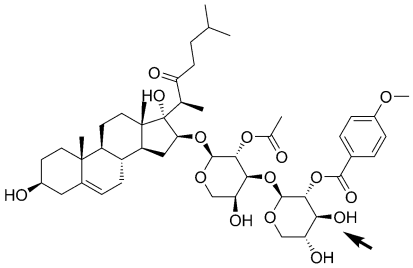
Entry	Compound	Target identification approach	Target(s)	Ref.
15	Fusicoccin A 	Chemical proteomics Tosyl chemistry	14-3-3	[69]
16	Piperlongumine 	Chemical proteomics using SILAC[b]	GSTP1	[75]
17		Chemical proteomics using SILAC[b] Tagged library	Epoxide hydrolase-1 MT-ND1	[76]
18	E7070 	2D-DIGE[c] ICAT[d]	MDH	[83]
19	XAV939 	Chemical proteomics iTRAQ[e]	Tankyrase 1, 2	[85]
20	OSW-1 	Chemical proteomics iTRAQ[e]	OSBP, ORP4L	[15a]

Table 1: (Continued)

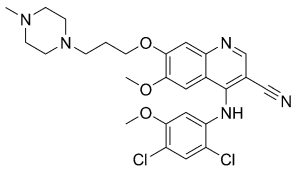
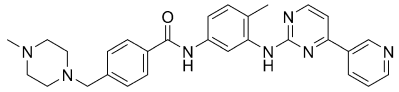
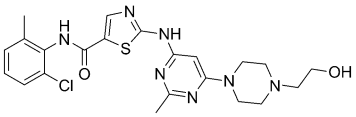
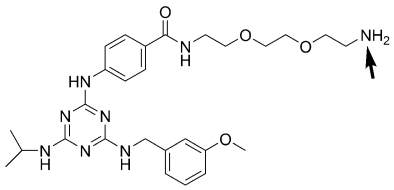
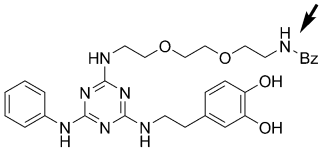
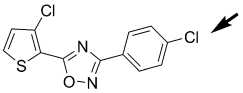
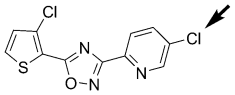
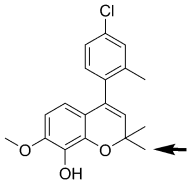
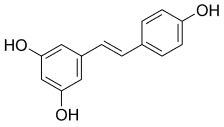
Entry	Compound	Target identification approach	Target(s)	Ref.
21	Bosutinib  Imatinib  Dasatinib 	Chemical proteomics Kinobeads, iTRAQ[e]	Different kinases	[86]
22		Chemical proteomics Tagged library	40S ribosomal subunit proteins S5, S13, and S18	[92]
23	GAPDS 	Chemical proteomics Tagged library	GAPDH	[94]
24	MX-74420  MX-126374 	2DE[f] ^{3}H-labeling	TIP47	[96]
25		FITGE[g] Cell lysates vs. live cells	Tubulin	[45]
26	Resveratrol 	DARTS[h]	eIF4A	[98]

Table 1: (Continued)

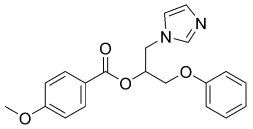
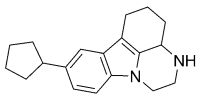
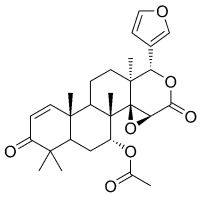
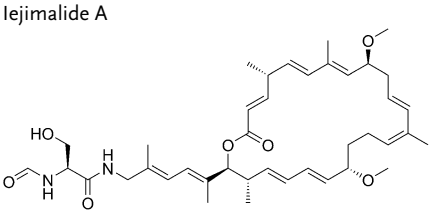
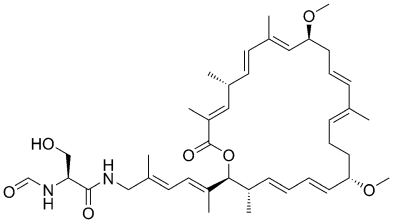
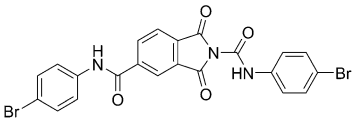
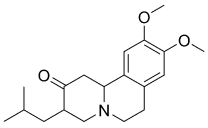
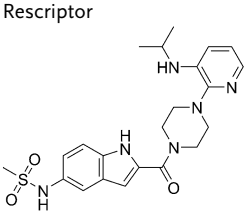
Entry	Compound	Target identification approach	Target(s)	Ref.
27	4513-0042 	TICC[i]	Erg6p	[102]
28	Pirl1 	Biochemical suppression	Cdc42/RhoGDI	[103]
29	Gedunin 	Comparison approach Connectivity map	HSP90	[112]
30	Iejimalide A  Iejimalide B 	Comparison approach 2D-DIGE[c]	V-ATPase	[117]
31	SMIR4 	Protein microarray	Ybr077cp	[123b]
32		Computation	α_2 adrenergic receptor	[125]
33	Rescriptor 	Computation	Histamine H_4 receptor	[125]

Table 1: (Continued)

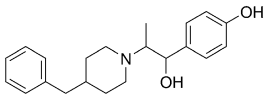
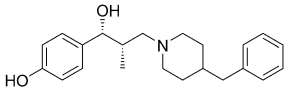
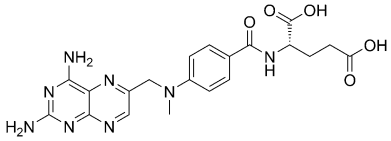
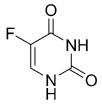
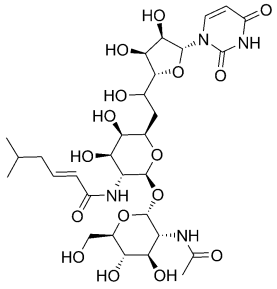
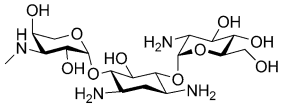
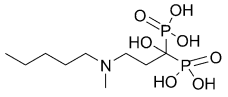
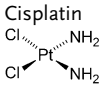
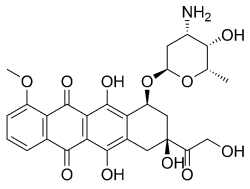
Entry	Compound	Target identification approach	Target(s)	Ref.
34	Validex 	Computation	μ-opioid receptor 5-HTT; serotonin transporter	[125]
35	Ro-25-6981 	Computation	5-HTT; serotonin transporter NET, noradrenalin transporter κ-opioid receptor	[125]
36	Methotrexate 	HIP/HOP[i]	DFR1, FOL1, FOL2	[134]
37	5-fluorouracil 	HIP/HOP[i]	CDC21, RRP6, RRP41, RRP44, RRP46, NOP4, MAK21, SSF1, YPR143W	[135]
38	Tunicamycin 	HIP/HOP[i]	ALG7p, HAC1, GFA1	[135]
39	Gentamicin 	HIP/HOP[i]	CAX4. GCS1, MNN9, SAC1, PEP3, PEP5, VPS15, VPS16, VPS33, VPS34	[136]
40	Ibandronate 	HIP/HOP[i]	TBCB, ASK/BDF4,	[137]
41	Cisplatin 	HIP/HOP[i]	FCY2, NMD2, NOT3, SKY1	[138]
42	Doxorubicin 	HIP/HOP[i]	SIZ1	[139]

Table 1: (Continued)

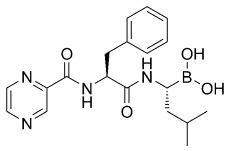
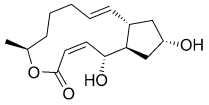
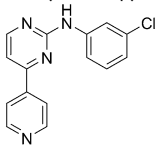
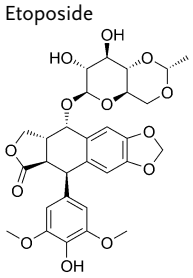
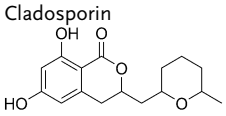
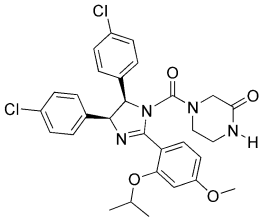
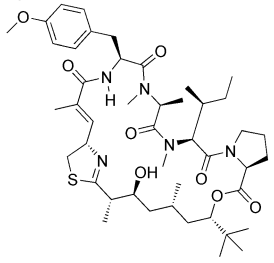
Entry	Compound	Target identification approach	Target(s)	Ref.
43	Bortezomib 	HIP[j] *S. pombe*	POF3, CHP1, SEC28, SNZ1, REX3, PKA1/GIT6	[142]
44	Brefeldin A 	HIP[j] *C. albicans*	Sec7p	[143]
45	Phenylaminopyrimidine 	MSP[k]	PKC1, CLG1, KSS1	[144]
46	Etoposide 	MSP[k] *S. pombe*	RHP51, RAD22	[146]
47	Cladosporin 	HIP[j]	Lysyl-tRNA synthase	[148]
48	Nutlin-3 	shRNA screen	TP53, 53BP1	[154]
49	Apratoxin A 	Mammalian overexpression	FGFR1	[159]

Table 1: (Continued)

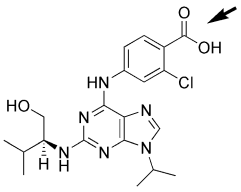
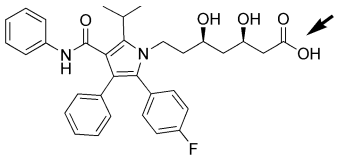
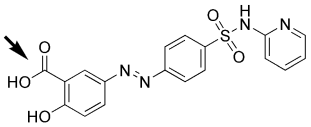
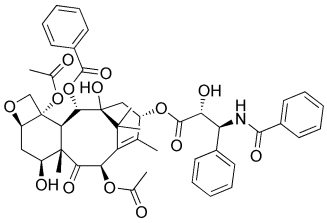
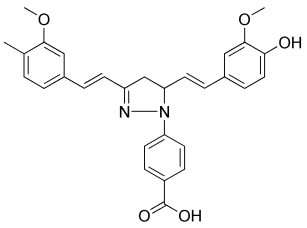
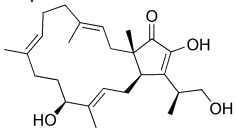
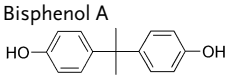
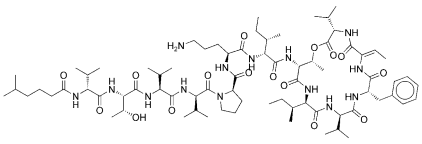
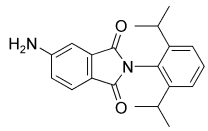
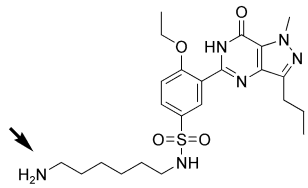
Entry	Compound	Target identification approach	Target(s)	Ref.
50	Purvanalol 	Yeast three-hybrid approach	Kinases	[164]
51	Atorvastatin 	Yeast three-hybrid approach	PDE6D	[151]
52	Sulfosalazine 	Yeast three-hybrid approach	SPR	[151]
53	Paclitaxel 	Phage display	Bcl-2, NSC-1, NFX1	[170]
54	HBC 	Phage display	Ca^{2+}/Calmodulin	[171]
55	Terpestacin 	Phage display	UQCRB	[172]
56	Bisphenol A 	Phage display	TACC3	[173]
57	Kahalalide F 	Phage display	RPS25	[174]

Table 1: (Continued)

Entry	Compound	Target identification approach	Target(s)	Ref.
58	TC11 	mRNA display	NPM	[179]
59	PF-3717842 	Chemical proteomics	PDEs PEBP1	[184]

[a] CCMS: capture compound mass spectrometry. [b] SILAC: stable isotope labeling by amino acids in cell culture. [c] 2D DIGE: two-dimensional difference gel electrophoresis. [d] ICAT: isotope-coded affinity tags. [e] iTRAQ: isobaric tags for relative and absolute quantification. [f] 2DE: two-dimensional gel electrophoresis. [g] FITGE: fluorescence difference approach in two-dimensional gel electrophoresis. [h] DARTS: drug affinity responsive target stability. [i] TICC: target identification by chromatographic co-elution. [j] HIP/HOP: haploinsufficient profiling/homozygous deletion profiling. [k] MSP: multicopy suppression profiling. The arrow indicates the site of modification.

Acknowledgements

Our research on target identification and confirmation of small bioactive molecules reflects the work of numerous former and present members of our research group, to whom we are more than grateful. Their names are found in the publications emanating from our group and cited in this Review. They have risen to the challenges posed by these truly interdisciplinary research projects and the need to embrace the methods and the different cultures of chemistry and biology. We are also grateful to our collaboration partners in various projects whose names are given in the author lists of our joint publications. Without their continued input and trusting collaboration many projects could not have been successfully realized. Our research was supported by the Max-Planck-Gesellschaft, the Deutsche Forschungsgemeinschaft, the Bundesministerium für Bildung und Forschung, the Alexander von Humboldt-Stiftung, the Volkswagen-Stiftung, the European Union (funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no268309), and the Land Nordrhein-Westfalen.

Received: October 31, 2012
Published online: February 18, 2013

[1] D. R. Spring, *Chem. Soc. Rev.* **2005**, *34*, 472 – 482.
[2] I. Gashaw, P. Ellinghaus, A. Sommer, K. Asadullah, *Drug Discovery Today* **2011**, *16*, 1037 – 1043.
[3] M. Clamp, B. Fry, M. Kamal, X. Xie, J. Cuff, M. F. Lin, M. Kellis, K. Lindblad-Toh, E. S. Lander, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19428 – 19433.
[4] a) J. Kotz, *J. SciBX* **2012**, *5*, 15; b) D. C. Swinney, J. Anthony, *Nat. Rev. Drug Discovery* **2011**, *10*, 507 – 519.
[5] a) R. S. Bon, H. Waldmann, *Acc. Chem. Res.* **2010**, *43*, 1103 – 1114; b) K. Kumar, H. Waldmann, *Angew. Chem.* **2009**, *121*, 3272 – 3290; *Angew. Chem. Int. Ed.* **2009**, *48*, 3224 – 3242; c) S. Wetzel, R. S. Bon, K. Kumar, H. Waldmann, *Angew. Chem.* **2011**, *123*, 10990 – 11018; *Angew. Chem. Int. Ed.* **2011**, *50*, 10800 – 10826; d) M. Kaiser, S. Wetzel, K. Kumar, H. Waldmann, *Cell. Mol. Life Sci.* **2008**, *65*, 1186 – 1201.
[6] S. J. Dixon, B. R. Stockwell, *Curr. Opin. Chem. Biol.* **2009**, *13*, 549 – 555.
[7] E. Lounkine, M. J. Keiser, S. Whitebread, D. Mikhailov, J. Hamon, J. L. Jenkins, P. Lavan, E. Weber, A. K. Doak, S. Cote, B. K. Shoichet, L. Urban, *Nature* **2012**, *486*, 361 – 367.
[8] A. L. Hopkins, *Nat. Chem. Biol.* **2008**, *4*, 682 – 690.
[9] D. A. Chan, A. J. Giaccia, *Nat. Rev. Drug Discovery* **2011**, *10*, 351 – 364.
[10] J. Mestres, E. Gregori-Puigjane, S. Valverde, R. V. Sole, *Mol. BioSyst.* **2009**, *5*, 1051 – 1057.
[11] a) G. C. Terstappen, C. Schluepen, R. Raggiaschi, G. Gaviraghi, *Nat. Rev. Drug Discovery* **2007**, *6*, 891 – 903; b) G. P. Tochtrop, R. W. King, *Comb. Chem. High Throughput Screening* **2004**, *7*, 677 – 688; c) S.-i. Sato, A. Murata, T. Shirakawa, M. Uesugi, *Chem. Biol.* **2010**, *17*, 616 – 623; d) R. K. Das, A. Samanta, K. Ghosh, D. Zhai, W. Xu, D. Su, C. Leong, C. Young-Tae, *IBC* **2011**, *3*, 3; e) A. M. Piggott, P. Karuso, *Comb. Chem. High Throughput Screening* **2004**, *7*, 607 – 630; f) B. J. Leslie, P. J. Hergenrother, *Chem. Soc. Rev.* **2008**, *37*, 1347 – 1360.
[12] a) B. Lomenick, R. W. Olsen, J. Huang, *ACS Chem. Biol.* **2011**, *6*, 34 – 46; b) F. Cong, A. K. Cheung, S. M. Huang, *Annu. Rev. Pharmacol. Toxicol.* **2012**, *52*, 57 – 78; c) T. Roemer, J. Davies, G. Giaever, C. Nislow, *Nat. Chem. Biol.* **2012**, *8*, 46 – 56.
[13] P. F. Liu, D. Kihara, C. Park, *J. Mol. Biol.* **2011**, *408*, 147 – 162.
[14] A. Scholten, M. K. Poh, T. A. B. van Veen, B. van Breukelen, M. A. Vos, A. J. R. Heck, *J. Proteome Res.* **2006**, *5*, 1435 – 1447.
[15] a) A. W. G. Burgett, T. B. Poulsen, K. Wangkanont, D. R. Anderson, C. Kikuchi, K. Shimada, S. Okubo, K. C. Fortner, Y. Mimaki, M. Kuroda, J. P. Murphy, D. J. Schwalb, E. C. Petrella, I. Cornella-Taracido, M. Schirle, J. A. Tallarico, M. D. Shair, *Nat. Chem. Biol.* **2011**, *7*, 639 – 647; b) N. S. Hegde, D. A. Sanders, R. Rodriguez, S. Balasubramanian, *Nat. Chem.* **2011**, *3*, 829 – 829.
[16] N. Kanoh, S. Kumashiro, S. Simizu, Y. Kondoh, S. Hatakeyama, H. Tashiro, H. Osada, *Angew. Chem.* **2003**, *115*, 5742 – 5745; *Angew. Chem. Int. Ed.* **2003**, *42*, 5584 – 5587.

[17] H. Dückert, V. Pries, V. Khedkar, S. Menninger, H. Bruss, A. W. Bird, Z. Maliga, A. Brockmeyer, P. Janning, A. Hyman, S. Grimme, M. Schurmann, H. Preut, K. Hubel, S. Ziegler, K. Kumar, H. Waldmann, *Nat. Chem. Biol.* **2012**, *8*, 179 – 184.

[18] a) S. Basu, B. Ellinger, S. Rizzo, C. Deraeve, M. Schurmann, H. Preut, H. D. Arndt, H. Waldmann, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6805 – 6810; b) T. Knoth, K. Warburg, C. Katzka, A. Rai, A. Wolf, A. Brockmeyer, P. Janning, T. F. Reubold, S. Eschenburg, D. J. Manstein, K. Hubel, M. Kaiser, H. Waldmann, *Angew. Chem.* **2009**, *121*, 7376 – 7381; *Angew. Chem. Int. Ed.* **2009**, *48*, 7240 – 7245.

[19] a) E. Weerapana, A. E. Speers, B. F. Cravatt, *Nat. Protoc.* **2007**, *2*, 1414 – 1425; b) A. E. Speers, B. F. Cravatt, *J. Am. Chem. Soc.* **2005**, *127*, 10018 – 10019; c) S. Sato, Y. Kwon, S. Kamisuki, N. Srivastava, Q. Mao, Y. Kawazoe, M. Uesugi, *J. Am. Chem. Soc.* **2007**, *129*, 873 – 880.

[20] F. Gug, N. Oumata, D. Tribouillard-Tanvier, C. Voisset, N. Desban, S. Bach, M. Blondel, H. Galons, *Bioconjugate Chem.* **2010**, *21*, 279 – 288.

[21] N. Shimizu, K. Sugimoto, J. Tang, T. Nishi, I. Sato, M. Hiramoto, S. Aizawa, M. Hatakeyama, R. Ohba, H. Hatori, T. Yoshikawa, F. Suzuki, A. Oomori, H. Tanaka, H. Kawaguchi, H. Watanabe, H. Handa, *Nat. Biotechnol.* **2000**, *18*, 877 – 881.

[22] T. Shiyama, M. Furuya, A. Yamazaki, T. Terada, A. Tanaka, *Bioorg. Med. Chem.* **2004**, *12*, 2831 – 2841.

[23] a) S. H. Verhelst, M. Fonovic, M. Bogyo, *Angew. Chem.* **2007**, *119*, 1306 – 1308; *Angew. Chem. Int. Ed.* **2007**, *46*, 1284 – 1286; b) M. Fonovic, S. H. L. Verhelst, M. T. Sorum, M. Bogyo, *Mol. Cell. Proteomics* **2007**, *6*, 1761 – 1770; c) Y. Y. Yang, M. Grammel, A. S. Raghavan, G. Charron, H. C. Hang, *Chem. Biol.* **2010**, *17*, 1212 – 1222; d) F. Landi, C. M. Johansson, D. J. Campopiano, A. N. Hulme, *Org. Biomol. Chem.* **2010**, *8*, 56 – 59.

[24] K. D. Park, R. Liu, H. Kohn, *Chem. Biol.* **2009**, *16*, 763 – 772.

[25] P. P. Geurink, B. I. Florea, N. Li, M. D. Witte, J. Verasdonck, C. L. Kuo, G. A. van der Marel, H. S. Overkleeft, *Angew. Chem.* **2010**, *122*, 6954 – 6957; *Angew. Chem. Int. Ed.* **2010**, *49*, 6802 – 6805.

[26] P. van der Veken, E. H. Dirksen, E. Ruijter, R. C. Elgersma, A. J. Heck, D. T. Rijkers, M. Slijper, R. M. Liskamp, *ChemBioChem* **2005**, *6*, 2271 – 2280.

[27] a) A. M. Piggott, P. Karuso, *Tetrahedron Lett.* **2005**, *46*, 8241 – 8244; b) T. Koopmans, F. J. Dekker, N. I. Martin, *RSC Adv.* **2012**, *2*, 2244 – 2246.

[28] T. Wirth, K. Schmuck, L. F. Tietze, S. A. Sieber, *Angew. Chem.* **2012**, *124*, 2928 – 2931; *Angew. Chem. Int. Ed.* **2012**, *51*, 2874 – 2877.

[29] A. T. Marttila, O. H. Laitinen, K. J. Airenne, T. Kulik, E. A. Bayer, M. Wilchek, M. S. Kulomaa, *FEBS Lett.* **2000**, *467*, 31 – 36.

[30] Y. Hiller, J. M. Gershoni, E. A. Bayer, M. Wilchek, *Biochem. J.* **1987**, *248*, 167 – 171.

[31] N. M. Green, *Biochem. J.* **1963**, *89*, 585 – 591.

[32] A. P. Frei, O. Y. Jeon, S. Kilcher, H. Moest, L. M. Henning, C. Jost, A. Pluckthun, J. Mercer, R. Aebersold, E. M. Carreira, B. Wollscheid, *Nat. Biotechnol.* **2012**, *30*, 997 – 1001.

[33] a) P.-Y. Yang, K. Liu, M. H. Ngai, M. J. Lear, M. R. Wenk, S. Q. Yao, *J. Am. Chem. Soc.* **2010**, *132*, 656 – 666; b) M. Rusch, T. J. Zimmermann, M. Burger, F. J. Dekker, K. Gormer, G. Triola, A. Brockmeyer, P. Janning, T. Bottcher, S. A. Sieber, I. R. Vetter, C. Hedberg, H. Waldmann, *Angew. Chem.* **2011**, *123*, 10012 – 10016; *Angew. Chem. Int. Ed.* **2011**, *50*, 9838 – 9842; c) M. H. Kunzmann, I. Staub, T. Böttcher, S. A. Sieber, *Biochemistry* **2011**, *50*, 910 – 916; d) T. Böttcher, S. A. Sieber, *J. Am. Chem. Soc.* **2010**, *132*, 6964 – 6972; e) J. Eirich, R. Orth, S. A. Sieber, *J. Am. Chem. Soc.* **2011**, *133*, 12144 – 12153.

[34] D. Kidd, Y. Liu, B. F. Cravatt, *Biochemistry* **2001**, *40*, 4005 – 4015.

[35] A. Einhauer, A. Jungbauer, *J. Biochem. Biophys. Methods* **2001**, *49*, 455 – 465.

[36] C. Saxena, E. Zhen, R. E. Higgs, J. E. Hale, *J. Proteome Res.* **2008**, *7*, 3490 – 3497.

[37] J. Arnau, C. Lauritzen, G. E. Petersen, J. Pedersen, *Protein Expression Purif.* **2006**, *48*, 1 – 13.

[38] a) C. C. Hughes, J. B. MacMillan, S. P. Gaudencio, W. Fenical, J. J. La Clair, *Angew. Chem.* **2009**, *121*, 742 – 746; *Angew. Chem. Int. Ed.* **2009**, *48*, 728 – 732; b) C. Saxena, T. M. Bonacci, K. L. Huss, L. J. Bloem, R. E. Higgs, J. E. Hale, *J. Proteome Res.* **2009**, *8*, 3951 – 3957.

[39] T. Böttcher, S. A. Sieber, *J. Am. Chem. Soc.* **2008**, *130*, 14400 – 14401.

[40] a) H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2001**, *113*, 2056 – 2075; *Angew. Chem. Int. Ed.* **2001**, *40*, 2004 – 2021; b) J. E. Hein, V. V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1302 – 1315.

[41] a) H. Köster, D. P. Little, P. Luan, R. Muller, S. M. Siddiqi, S. Marappan, P. Yip, *Assay Drug Dev. Technol.* **2007**, *5*, 381 – 390; b) T. Lenz, J. J. Fischer, M. Dreger, *J. Proteomics* **2011**, *75*, 100 – 115.

[42] a) T. Lenz, P. Poot, E. Weinhold, M. Dreger, *Methods Mol. Biol.* **2012**, *803*, 97 – 125; b) T. Lenz, P. Poot, O. Grabner, M. Glinski, E. Weinhold, M. Dreger, H. Koster, *J. Visualized Exp.* **2010**, *46*, e2264; c) L. Wirsing, K. Naumann, T. Vogt, *Anal. Biochem.* **2011**, *408*, 220 – 225; d) C. Dalhoff, M. Hueben, T. Lenz, P. Poot, E. Nordhoff, H. Koester, E. Weinhold, *ChemBioChem* **2010**, *11*, 256 – 265.

[43] J. J. Fischer, S. Michaelis, A. K. Schrey, O. G. Graebner, M. Glinski, M. Dreger, F. Kroll, H. Koester, *Toxicol. Sci.* **2010**, *113*, 243 – 253.

[44] S. Michaelis, A. Marais, A. K. Schrey, O. Y. Graebner, C. Schaudt, M. Sefkow, F. Kroll, M. Dreger, M. Glinski, H. Koester, R. Metternich, J. J. Fischer, *J. Med. Chem.* **2012**, *55*, 3934 – 3944.

[45] J. Park, S. Oh, S. B. Park, *Angew. Chem.* **2012**, *124*, 5543 – 5547; *Angew. Chem. Int. Ed.* **2012**, *51*, 5447 – 5451.

[46] G. Wang, L. Shang, A. W. Burgett, P. G. Harran, X. Wang, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2068 – 2073.

[47] J. E. Wulff, R. Siegrist, A. G. Myers, *J. Am. Chem. Soc.* **2007**, *129*, 14444 – 14451.

[48] Y. Kotake, K. Sagane, T. Owa, Y. Mimori-Kiyosue, H. Shimizu, M. Uesugi, Y. Ishihama, M. Iwata, Y. Mizui, *Nat. Chem. Biol.* **2007**, *3*, 570 – 575.

[49] L. Hu, A. Iliuk, J. Galan, M. Hans, W. A. Tao, *Angew. Chem.* **2011**, *123*, 4219 – 4222; *Angew. Chem. Int. Ed.* **2011**, *50*, 4133 – 4136, S4133/4131 – S4133/4113.

[50] K. Yamamoto, A. Yamazaki, M. Takeuchi, A. Tanaka, *Anal. Biochem.* **2006**, *352*, 15 – 23.

[51] a) D. Kaida, H. Motoyoshi, E. Tashiro, T. Nojima, M. Hagiwara, K. Ishigami, H. Watanabe, T. Kitahara, T. Yoshida, H. Nakajima, T. Tani, S. Horinouchi, M. Yoshida, *Nat. Chem. Biol.* **2007**, *3*, 576 – 583; b) W. Shi, X. Zhang, X. Jiang, H. Yuan, J. S. Lee, C. E. Barry III, H. Wang, W. Zhang, Y. Zhang, *Science* **2011**, *333*, 1630 – 1632; c) S. M. Khersonsky, D.-W. Jung, T.-W. Kang, D. P. Walsh, H.-S. Moon, H. Jo, E. M. Jacobson, V. Shetty, T. A. Neubert, Y.-T. Chang, *J. Am. Chem. Soc.* **2003**, *125*, 11804 – 11805; d) P. Dadvar, M. O'Flaherty, A. Scholten, K. Rumpel, A. J. Heck, *Mol. BioSyst.* **2009**, *5*, 472 – 482.

[52] L. Margarucci, M. C. Monti, B. Fontanella, R. Riccio, A. Casapullo, *Mol. BioSyst.* **2011**, *7*, 480 – 485.

[53] a) Q. Zhang, M. B. Major, S. Takanashi, N. D. Camp, N. Nishiya, E. C. Peters, M. H. Ginsberg, X. Jian, P. A. Randazzo, P. G. Schultz, R. T. Moon, S. Ding, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7444 – 7448; b) C.-w. Chung, H. Coste, J. H. White, O. Mirguet, J. Wilde, R. L. Gosmini, C. Delves, S. M. Magny, R.

Woodward, S. A. Hughes, E. V. Boursier, H. Flynn, A. M. Bouillot, P. Bamborough, J.-M. G. Brusq, F. J. Gellibert, E. J. Jones, A. M. Riou, P. Homes, S. L. Martin, I. J. Uings, J. Toum, C. A. Clement, A.-B. Boullay, R. L. Grimley, F. M. Blandel, R. K. Prinjha, K. Lee, J. Kirilovsky, E. Nicodeme, *J. Med. Chem.* **2011**, *54*, 3827 – 3838; c) S. M. Wignall, N. S. Gray, Y.-T. Chang, L. Juarez, R. Jacob, A. Burlingame, P. G. Schultz, R. Heald, *Chem. Biol.* **2004**, *11*, 135 – 146; d) A. V. Statsuk, R. Bai, J. L. Baryza, V. A. Verma, E. Hamel, P. A. Wender, S. A. Kozmin, *Nat. Chem. Biol.* **2005**, *1*, 383 – 388; e) J. R. Snyder, A. Hall, L. Ni-Komatsu, S. M. Khersonsky, Y. T. Chang, S. J. Orlow, *Chem. Biol.* **2005**, *12*, 477 – 484.

[54] Y. Nakamura, R. Miyatake, M. Ueda, *Angew. Chem.* **2008**, *120*, 7399 – 7402; *Angew. Chem. Int. Ed.* **2008**, *47*, 7289 – 7292.

[55] S. Ding, T. Y. H. Wu, A. Brinker, E. C. Peters, W. Hur, N. S. Gray, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7632 – 7637.

[56] S. Kuettel, M. Mosimann, P. Maser, M. Kaiser, R. Brun, L. Scapozza, R. Perozzo, *PLoS Neglected Trop. Dis.* **2009**, *3*, e506.

[57] a) T. Hirota, J. W. Lee, W. G. Lewis, E. E. Zhang, G. Breton, X. Z. Liu, M. Garcia, E. C. Peters, J. P. Etchegaray, D. Traver, P. G. Schultz, S. A. Kay, *Plos Biol.* **2010**, *8*, e1000559; b) T. Hirota, J. W. Lee, P. C. St John, M. Sawa, K. Iwaisako, T. Noguchi, P. Y. Pongsawakul, T. Sonntag, D. K. Welsh, D. A. Brenner, F. J. Doyle 3rd, P. G. Schultz, S. A. Kay, *Science* **2012**, *337*, 1094 – 1097; c) J. Lee, X. Wu, M. Pasca di Magliano, E. C. Peters, Y. Wang, J. Hong, M. Hebrok, S. Ding, C. Y. Cho, P. G. Schultz, *ChemBioChem* **2007**, *8*, 1916 – 1919.

[58] I. Miller, J. Crawford, E. Gianazza, *Proteomics* **2006**, *6*, 5385 – 5408.

[59] U. Rix, G. Superti-Furga, *Nat. Chem. Biol.* **2009**, *5*, 616 – 624.

[60] L. Trinkle-Mulcahy, S. Boulon, Y. W. Lam, R. Urcia, F. M. Boisvert, F. Vandermoere, N. A. Morrice, S. Swift, U. Rothbauer, H. Leonhardt, A. Lamond, *J. Cell Biol.* **2008**, *183*, 223 – 239.

[61] a) M. Beck, A. Schmidt, J. Malmstroem, M. Claassen, A. Ori, A. Szymborska, F. Herzog, O. Rinner, J. Ellenberg, R. Aebersold, *Mol. Syst. Biol.* **2011**, *7*, 549; b) N. Nagaraj, J. R. Wisniewski, T. Geiger, J. Cox, M. Kircher, J. Kelso, S. Paabo, M. Mann, *Mol. Syst. Biol.* **2011**, *7*, 548.

[62] T. Voigt, C. Gerding-Reimers, T. T. N. Tran, S. Bergmann, H. Lachance, B. Schölermann, A. Brockmeyer, P. Janning, S. Ziegler, H. Waldmann, *Angew. Chem.* **2013**, *125*, 428 – 432; *Angew. Chem. Int. Ed.* **2013**, *52*, 410 – 414.

[63] J. R. Wiśniewski, A. Zougman, N. Nagaraj, M. Mann, *Nat. Methods* **2009**, *6*, 359 – 362.

[64] R. R. Falsey, M. T. Marron, G. M. Gunaherath, N. Shirahatti, D. Mahadevan, A. A. Gunatilaka, L. Whitesell, *Nat. Chem. Biol.* **2006**, *2*, 33 – 38.

[65] N. Kanoh, T. Nakamura, K. Honda, H. Yamakoshi, Y. Iwabuchi, H. Osada, *Tetrahedron* **2008**, *64*, 5692 – 5698.

[66] M. Kawatani, H. Okumura, K. Honda, N. Kanoh, M. Muroi, N. Dohmae, M. Takami, M. Kitagawa, Y. Futamura, M. Imoto, H. Osada, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11691 – 11696.

[67] a) N. Watanabe, T. Sekine, M. Takagi, J. Iwasaki, N. Imamoto, H. Kawasaki, H. Osada, *J. Biol. Chem.* **2009**, *284*, 2344 – 2353; b) K. Hagiwara, T. Murakami, G. Xue, Y. Shimizu, E. Takeda, Y. Hashimoto, K. Honda, Y. Kondoh, H. Osada, Y. Tsunetsugu-Yokota, Y. Aida, *Biochem. Biophys. Res. Commun.* **2010**, *403*, 40 – 45; c) N. Kanoh, A. Asami, M. Kawatani, K. Honda, S. Kumashiro, H. Takayama, S. Simizu, T. Amemiya, Y. Kondoh, S. Hatakeyama, K. Tsuganezawa, R. Utata, A. Tanaka, S. Yokoyama, H. Tashiro, H. Osada, *Chem. Asian J.* **2006**, *1*, 789 – 797.

[68] a) S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura, I. Hamachi, *Nat. Chem. Biol.* **2009**, *5*, 341 – 343; b) C. C. Hughes, Y. L. Yang, W. T. Liu, P. C. Dorrestein, J. J. La Clair, W. Fenical, *J. Am. Chem. Soc.* **2009**, *131*, 12094 – 12096.

[69] M. Takahashi, A. Kawamura, N. Kato, T. Nishi, I. Hamachi, J. Ohkanda, *Angew. Chem.* **2012**, *124*, 524 – 527; *Angew. Chem. Int. Ed.* **2012**, *51*, 509 – 512.

[70] W. Yan, S. S. Chen, *Briefings Funct. Genomics Proteomics* **2005**, *4*, 27 – 38.

[71] S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **2002**, *1*, 376 – 386.

[72] S. E. Ong, I. Kratchmarova, M. Mann, *J. Proteome Res.* **2003**, *2*, 173 – 181.

[73] S.-E. Ong, M. Schenone, A. A. Margolin, X. Li, K. Do, M. K. Doud, D. R. Mani, L. Kuai, X. Wang, J. L. Wood, N. J. Tolliday, A. N. Koehler, L. A. Marcaurelle, T. R. Golub, R. J. Gould, S. L. Schreiber, S. A. Carr, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4617 – 4622.

[74] P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D. J. Pappin, *Mol. Cell. Proteomics* **2004**, *3*, 1154 – 1169.

[75] L. Raj, T. Ide, A. U. Gurkar, M. Foley, M. Schenone, X. Li, N. J. Tolliday, T. R. Golub, S. A. Carr, A. F. Shamji, A. M. Stern, A. Mandinova, S. L. Schreiber, S. W. Lee, *Nature* **2011**, *475*, 231 – 234.

[76] J. S. Cisar, B. F. Cravatt, *J. Am. Chem. Soc.* **2012**, *134*, 10385 – 10388.

[77] S. B. Breitkopf, F. S. Oppermann, G. Keri, M. Grammel, H. Daub, *J. Proteome Res.* **2010**, *9*, 6033 – 6043.

[78] a) F. S. Oppermann, F. Gnad, J. V. Olsen, R. Hornberger, Z. Greff, G. Keri, M. Mann, H. Daub, *Mol. Cell. Proteomics* **2009**, *8*, 1751 – 1764; b) H. Daub, J. V. Olsen, M. Bairlein, F. Gnad, F. S. Oppermann, R. Korner, Z. Greff, G. Keri, O. Stemmann, M. Mann, *Mol. Cell* **2008**, *31*, 438 – 448.

[79] T. A. Prokhorova, K. T. G. Rigbolt, P. T. Johansen, J. Henningsen, I. Kratchmarova, M. Kassem, B. Blagoev, *Mol. Cell. Proteomics* **2009**, *8*, 959 – 970.

[80] M. Krüger, M. Moser, S. Ussar, I. Thievessen, C. A. Luber, F. Forner, S. Schmidt, S. Zanivan, R. Fassler, M. Mann, *Cell* **2008**, *134*, 353 – 364.

[81] J. Krijgsveld, R. F. Ketting, T. Mahmoudi, J. Johansen, M. Artal-Sanz, C. P. Verrijzer, R. H. A. Plasterk, A. J. R. Heck, *Nat. Biotechnol.* **2003**, *21*, 927 – 931.

[82] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994 – 999.

[83] Y. Oda, T. Owa, T. Sato, B. Boucher, S. Daniels, H. Yamanaka, Y. Shinohara, A. Yokoi, J. Kuromitsu, T. Nagasu, *Anal. Chem.* **2003**, *75*, 2159 – 2165.

[84] H. Zhang, W. Yan, R. Aebersold, *Curr. Opin. Chem. Biol.* **2004**, *8*, 66 – 75.

[85] S. M. Huang, Y. M. Mishina, S. Liu, A. Cheung, F. Stegmeier, G. A. Michaud, O. Charlat, E. Wiellette, Y. Zhang, S. Wiessner, M. Hild, X. Shi, C. J. Wilson, C. Mickanin, V. Myer, A. Fazal, R. Tomlinson, F. Serluca, W. Shao, H. Cheng, M. Shultz, C. Rau, M. Schirle, J. Schlegl, S. Ghidelli, S. Fawell, C. Lu, D. Curtis, M. W. Kirschner, C. Lengauer, P. M. Finan, J. A. Tallarico, T. Bouwmeester, J. A. Porter, A. Bauer, F. Cong, *Nature* **2009**, *461*, 614 – 620.

[86] M. Bantscheff, D. Eberhard, Y. Abraham, S. Bastuck, M. Boesche, S. Hobson, T. Mathieson, J. Perrin, M. Raida, C. Rau, V. Reader, G. Sweetman, A. Bauer, T. Bouwmeester, C. Hopf, U. Kruse, G. Neubauer, N. Ramsden, J. Rick, B. Kuster, G. Drewes, *Nat. Biotechnol.* **2007**, *25*, 1035 – 1044.

[87] a) N. Nagaraj, M. Mann, *J. Proteome Res.* **2011**, *10*, 637 – 645; b) C. A. Luber, J. Cox, H. Lauterbach, B. Fancke, M. Selbach, J.

Tschopp, S. Akira, M. Wiegand, H. Hochrein, M. O'Keeffe, M. Mann, *Immunity* **2010**, *32*, 279 – 289.
[88] J. Cox, M. Mann, *Annu. Rev. Biochem.* **2011**, *80*, 273 – 299.
[89] M. Bantscheff, S. Lemeer, M. M. Savitski, B. Kuster, *Anal. Bioanal. Chem.* **2012**, *404*, 939 – 965.
[90] D. Williams, D. W. Jung, S. M. Khersonsky, N. Heidary, Y. T. Chang, S. J. Orlow, *Chem. Biol.* **2004**, *11*, 1251 – 1259.
[91] M. Uttamchandani, D. P. Walsh, S. M. Khersonsky, X. Huang, S. Q. Yao, Y. T. Chang, *J. Comb. Chem.* **2004**, *6*, 862 – 868.
[92] See Ref. [51c].
[93] D. W. Jung, D. Williams, S. M. Khersonsky, T. W. Kang, N. Heidary, Y. T. Chang, S. J. Orlow, *Mol. BioSyst.* **2005**, *1*, 85 – 92.
[94] J. Min, Y. K. Kim, P. G. Cipriani, M. Kang, S. M. Khersonsky, D. P. Walsh, J.-Y. Lee, S. Niessen, J. R. Yates III, K. Gunsalus, F. Piano, Y.-T. Chang, *Nat. Chem. Biol.* **2007**, *3*, 55 – 59.
[95] T. Rabilloud, M. Chevallet, S. Luche, C. Lelong, *J. Proteomics* **2010**, *73*, 2064 – 2077.
[96] K. A. Jessen, N. M. English, J. Yu Wang, S. Maliartchouk, S. P. Archer, L. Qiu, R. Brand, J. Kuemmerle, H. Z. Zhang, K. Gehlsen, J. Drewe, B. Tseng, S. X. Cai, S. Kasibhatla, *Mol. Cancer Ther.* **2005**, *4*, 761 – 771.
[97] A. S. Essader, B. J. Cargile, J. L. Bundy, J. L. Stephenson, Jr., *Proteomics* **2005**, *5*, 24 – 34.
[98] B. Lomenick, R. Hao, N. Jonai, R. M. Chin, M. Aghajan, S. Warburton, J. Wang, R. P. Wu, F. Gomez, J. A. Loo, J. A. Wohlschlegel, T. M. Vondriska, J. Pelletier, H. R. Herschman, J. Clardy, C. F. Clarke, J. Huang, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 21984 – 21989.
[99] C. Park, S. Marqusee, *Nat. Methods* **2005**, *2*, 207 – 212.
[100] Y. Chang, J. P. Schlebach, R. A. Verheul, C. Park, *Protein Sci.* **2012**, *21*, 1280 – 1287.
[101] G. M. West, C. L. Tucker, T. Xu, S. K. Park, X. Han, J. R. Yates 3rd, M. C. Fitzgerald, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9078 – 9082.
[102] J. N. Chan, D. Vuckovic, L. Sleno, J. B. Olsen, O. Pogoutse, P. Havugimana, J. A. Hewel, N. Bajaj, Y. Wang, M. F. Musteata, C. Nislow, A. Emili, *Mol. Cell. Proteomics* **2012**, *11*, M111 016642.
[103] J. R. Peterson, A. M. Lebensohn, H. E. Pelish, M. W. Kirschner, *Chem. Biol.* **2006**, *13*, 443 – 452.
[104] S. V. Sharma, D. A. Haber, J. Settleman, *Nat. Rev. Cancer* **2010**, *10*, 241 – 253.
[105] J. N. Weinstein, T. G. Myers, P. M. O'Connor, S. H. Friend, A. J. Fornace, Jr., K. W. Kohn, T. Fojo, S. E. Bates, L. V. Rubinstein, N. L. Anderson, J. K. Buolamwini, W. W. van Osdol, A. P. Monks, D. A. Scudiero, E. A. Sausville, D. W. Zaharevitz, B. Bunow, V. N. Viswanadhan, G. S. Johnson, R. E. Wittes, K. D. Paull, *Science* **1997**, *275*, 343 – 349.
[106] K. D. Paull, R. H. Shoemaker, L. Hodes, A. Monks, D. A. Scudiero, L. Rubinstein, J. Plowman, M. R. Boyd, *J. Natl. Cancer Inst.* **1989**, *81*, 1088 – 1092.
[107] F. Leteurtre, G. Kohlhagen, K. D. Paull, Y. Pommier, *J. Natl. Cancer Inst.* **1994**, *86*, 1239 – 1244.
[108] R. L. Bai, K. D. Paull, C. L. Herald, L. Malspeis, G. R. Pettit, E. Hamel, *J. Biol. Chem.* **1991**, *266*, 15882 – 15889.
[109] E. S. Cleaveland, A. Monks, A. Vaigro-Wolff, D. W. Zaharevitz, K. Paull, K. Ardalan, D. A. Cooney, H. Ford, Jr., *Biochem. Pharmacol.* **1995**, *49*, 947 – 954.
[110] a) T. Yamori, A. Matsunaga, S. Sato, K. Yamazaki, A. Komi, K. Ishizu, I. Mita, H. Edatsugi, Y. Matsuba, K. Takezawa, O. Nakanishi, H. Kohno, Y. Nakajima, H. Komatsu, T. Andoh, T. Tsuruo, *Cancer Res.* **1999**, *59*, 4042 – 4049; b) S. Yaguchi, Y. Fukui, I. Koshimizu, H. Yoshimi, T. Matsuno, H. Gouda, S. Hirono, K. Yamazaki, T. Yamori, *J. Natl. Cancer Inst.* **2006**, *98*, 545 – 556; c) Y. Ohashi, H. Iijima, N. Yamaotsu, K. Yamazaki, S. Sato, M. Okamura, K. Sugimoto, S. Dan, S. Hirono, T. Yamori, *J. Biol. Chem.* **2012**, *287*, 3885 – 3897.
[111] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehar, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jane-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, P. Aspesi, Jr., M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, L. A. Garraway, *Nature* **2012**, *483*, 603 – 607.
[112] J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J. P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, T. R. Golub, *Science* **2006**, *313*, 1929 – 1935.
[113] a) J. Gheeya, P. Johansson, Q. R. Chen, T. Dexheimer, B. Metaferia, Y. K. Song, J. S. Wei, J. He, Y. Pommier, J. Khan, *Cancer Lett.* **2010**, *293*, 124 – 131; b) Z. X. Zhang, T. Meng, J. X. He, M. Li, L. J. Tong, B. Xiong, L. P. Lin, J. K. Shen, Z. H. Miao, J. Ding, *Invest. New Drugs* **2010**, *28*, 715 – 728; c) F. M. Siu, D. L. Ma, Y. W. Cheung, C. N. Lok, K. Yan, Z. Q. Yang, M. S. Yang, S. X. Xu, B. C. B. Ko, Q. Y. He, C. M. Che, *Proteomics* **2008**, *8*, 3105 – 3117.
[114] a) Y. Ishimatsu-Tsuji, T. Soma, J. Kishimoto, *FASEB J.* **2010**, *24*, 1489 – 1496; b) T. Sanda, X. Li, A. Gutierrez, Y. Ahn, D. S. Neuberg, J. O'Neil, P. R. Strack, C. G. Winter, S. S. Winter, R. S. Larson, H. von Boehmer, A. T. Look, *Blood* **2010**, *115*, 1735 – 1745.
[115] G. Van den Bergh, L. Arckens, *Curr. Opin. Biotechnol.* **2004**, *15*, 38 – 43.
[116] A. Alban, S. O. David, L. Bjorkesten, C. Andersson, E. Sloge, S. Lewis, I. Currie, *Proteomics* **2003**, *3*, 36 – 44.
[117] M. Muroi, S. Kazami, K. Noda, H. Kondo, H. Takayama, M. Kawatani, T. Usui, H. Osada, *Chem. Biol.* **2010**, *17*, 460 – 470.
[118] Z. E. Perlman, M. D. Slack, Y. Feng, T. J. Mitchison, L. F. Wu, S. J. Altschuler, *Science* **2004**, *306*, 1194 – 1198.
[119] Y. A. Abassi, B. Xi, W. Zhang, P. Ye, S. L. Kirstein, M. R. Gaylord, S. C. Feinstein, X. Wang, X. Xu, *Chem. Biol.* **2009**, *16*, 712 – 723.
[120] I. Giaever, C. R. Keese, *Nature* **1993**, *366*, 591 – 592.
[121] P. Jonkheijm, D. Weinrich, H. Schroder, C. M. Niemeyer, H. Waldmann, *Angew. Chem.* **2008**, *120*, 9762 – 9792; *Angew. Chem. Int. Ed.* **2008**, *47*, 9618 – 9647.
[122] L. Berrade, A. E. Garcia, J. A. Camarero, *Pharm. Res.* **2011**, *28*, 1480 – 1499.
[123] a) H. Ge, *Nucleic Acids Res.* **2000**, *28*, 3e; b) J. Huang, H. Zhu, S. J. Haggarty, D. R. Spring, H. Hwang, F. Jin, M. Snyder, S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16594 – 16599; c) Y. Fang, A. G. Frutos, J. Lahiri, *J. Am. Chem. Soc.* **2002**, *124*, 2394 – 2395; d) C. Schnack, B. Hengerer, F. Gillardon, *Proteomics* **2008**, *8*, 1980 – 1986.
[124] a) M. J. Keiser, B. L. Roth, B. N. Armbruster, P. Ernsberger, J. J. Irwin, B. K. Shoichet, *Nat. Biotechnol.* **2007**, *25*, 197 – 206; b) SEA search tool can be accessed via http://sea.bkslab.org/.
[125] M. J. Keiser, V. Setola, J. J. Irwin, C. Laggner, A. I. Abbas, S. J. Hufeisen, N. H. Jensen, M. B. Kuijer, R. C. Matos, T. B. Tran, R. Whaley, R. A. Glennon, J. Hert, K. L. H. Thomas, D. D. Edwards, B. K. Shoichet, B. L. Roth, *Nature* **2009**, *462*, 175 – 181.
[126] J. U. Peters, P. Schnider, P. Mattei, M. Kansy, *ChemMedChem* **2009**, *4*, 680 – 686.
[127] a) E. Gregori-Puigjane, J. Mestres, *Comb. Chem. High Throughput Screening* **2008**, *11*, 669 – 676; b) A. Lagunin, A.

Stepanchikova, D. Filimonov, V. Poroikov, *Bioinformatics* **2000**, *16*, 747 – 748.
[128] J. Heitman, N. Movva, M. Hall, *Science* **1991**, *253*, 905 – 909.
[129] a) A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver, *Science* **1996**, *274*, 546 – 567; b) B. Suter, D. Auerbach, I. Stagljar, *Biotechniques* **2006**, *40*, 625 – 644.
[130] A. Wach, A. Brachat, R. Pöhlmann, P. Philippsen, *Yeast* **1994**, *10*, 1793 – 1808.
[131] a) E. A. Winzeler, D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, B. Andre, R. Bangham, R. Benito, J. D. Boeke, H. Bussey, A. M. Chu, C. Connelly, K. Davis, F. Dietrich, S. W. Dow, M. El Bakkoury, F. Foury, S. H. Friend, E. Gentalen, G. Giaever, J. H. Hegemann, T. Jones, M. Laub, H. Liao, N. Liebundguth, D. J. Lockhart, A. Lucau-Danila, M. Lussier, N. M'Rabet, P. Menard, M. Mittmann, C. Pai, C. Rebischung, J. L. Revuelta, L. Riles, C. J. Roberts, P. Ross-MacDonald, B. Scherens, M. Snyder, S. Sookhai-Mahadeo, R. K. Storms, S. Veronneau, M. Voet, G. Volckaert, T. R. Ward, R. Wysocki, G. S. Yen, K. Yu, K. Zimmermann, P. Philippsen, M. Johnston, R. W. Davis, *Science* **1999**, *285*, 901 – 906; b) G. Giaever, A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A. P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kotter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis, M. Johnston, *Nature* **2002**, *418*, 387 – 391.
[132] D. Delneri, *FEMS Yeast Res.* **2010**, *10*, 1083 – 1089.
[133] G. Giaever, D. D. Shoemaker, T. W. Jones, H. Liang, E. A. Winzeler, A. Astromoff, R. W. Davis, *Nat. Genet.* **1999**, *21*, 278 – 283.
[134] G. Giaever, P. Flaherty, J. Kumm, M. Proctor, C. Nislow, D. F. Jaramillo, A. M. Chu, M. I. Jordan, A. P. Arkin, R. W. Davis, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 793 – 798.
[135] P. Y. Lum, C. D. Armour, S. B. Stepaniants, G. Cavet, M. K. Wolf, J. S. Butler, J. C. Hinshaw, P. Garnier, G. D. Prestwich, A. Leonardson, P. Garrett-Engele, C. M. Rush, M. Bard, G. Schimmack, J. W. Phillips, C. J. Roberts, D. D. Shoemaker, *Cell* **2004**, *116*, 121 – 137.
[136] A. S. Blackburn, S. V. Avery, *Antimicrob. Agents Chemother.* **2003**, *47*, 676 – 681.
[137] N. Bivi, M. Romanello, R. Harrison, I. Clarke, D. Hoyle, L. Moro, F. Ortolani, A. Bonetti, F. Quadrifoglio, G. Tell, D. Delneri, *Genome Biol.* **2009**, *10*, R93.
[138] R.-Y. Huang, M. Eddy, M. Vujcic, D. Kowalski, *Cancer Res.* **2005**, *65*, 5890 – 5897.
[139] R.-Y. Huang, D. Kowalski, H. Minderman, N. Gandhi, E. S. Johnson, *Cancer Res.* **2007**, *67*, 765 – 772.
[140] A. Wilmes, R. Hanna, R. W. Heathcott, P. T. Northcote, P. H. Atkinson, D. S. Bellows, J. H. Miller, *Gene* **2012**, *497*, 140 – 146.
[141] D.-U. Kim, J. Hayles, D. Kim, V. Wood, H.-O. Park, M. Won, H.-S. Yoo, T. Duhig, M. Nam, G. Palmer, S. Han, L. Jeffery, S.-T. Baek, H. Lee, Y. S. Shim, M. Lee, L. Kim, K.-S. Heo, E. J. Noh, A.-R. Lee, Y.-J. Jang, K.-S. Chung, S.-J. Choi, J.-Y. Park, Y. Park, H. M. Kim, S.-K. Park, H.-J. Park, E.-J. Kang, H. B. Kim, H.-S. Kang, H.-M. Park, K. Kim, K. Song, K. B. Song, P. Nurse, K.-L. Hoe, *Nat. Biotechnol.* **2010**, *28*, 617 – 623.
[142] K. Takeda, A. Mori, M. Yanagida, *Plos One* **2011**, *6*, e22021.
[143] J. Oh, E. Fung, M. N. Price, P. S. Dehal, R. W. Davis, G. Giaever, C. Nislow, A. P. Arkin, A. Deutschbauer, *Nucleic Acids Res.* **2010**, *38*, e146.
[144] H. Luesch, *Mol. BioSyst.* **2006**, *2*, 609 – 620.
[145] R. A. Butcher, B. S. Bhullar, E. O. Perlstein, G. Marsischky, J. LaBaer, S. L. Schreiber, *Nat. Chem. Biol.* **2006**, *2*, 103 – 109.
[146] Y. Arita, S. Nishimura, A. Matsuyama, Y. Yashiroda, T. Usui, C. Boone, M. Yoshida, *Mol. BioSyst.* **2011**, *7*, 1463 – 1472.
[147] S. Hoon, A. M Smith, I. M. Wallace, S. Suresh, M. Miranda, E. Fung, M. Proctor, K. M. Shokat, C. Zhang, R. W. Davis, G. Giaever, R. P. St Onge, C. Nislow, *Nat. Chem. Biol.* **2008**, *4*, 498 – 506.
[148] D. Hoepfner, C. W. McNamara, C. S. Lim, C. Studer, R. Riedl, T. Aust, S. L. McCormack, D. M. Plouffe, S. Meister, S. Schuierer, U. Plikat, N. Hartmann, F. Staedtler, S. Cotesta, E. K. Schmitt, F. Petersen, F. Supek, R. J. Glynne, J. A. Tallarico, J. A. Porter, M. C. Fishman, C. Bodenreider, T. T. Diagana, N. R. Movva, E. A. Winzeler, *Cell Host Microbe* **2012**, *11*, 654 – 663.
[149] S. E. Pierce, R. W. Davis, C. Nislow, G. Giaever, *Nat. Protoc.* **2007**, *2*, 2958 – 2974.
[150] A. M. Smith, L. E. Heisler, J. Mellor, F. Kaper, M. J. Thompson, M. Chee, F. P. Roth, G. Giaever, C. Nislow, *Genome Res.* **2009**, *19*, 1836 – 1842.
[151] C. Chidley, H. Haruki, M. G. Pedersen, E. Muller, K. Johnsson, *Nat. Chem. Biol.* **2011**, *7*, 375 – 383.
[152] A. Stepanov, K. C. Nitiss, G. Neale, J. L. Nitiss, *Mol. Pharmacol.* **2008**, *74*, 423 – 431.
[153] A. R. Burns, I. M. Wallace, J. Wildenhain, M. Tyers, G. Giaever, G. D. Bader, C. Nislow, S. R. Cutler, P. J. Roy, *Nat. Chem. Biol.* **2010**, *6*, 549 – 557.
[154] R. Bernards, T. R. Brummelkamp, R. L. Beijersbergen, *Nat. Methods* **2006**, *3*, 701 – 706.
[155] D. J. Burgess, J. Doles, L. Zender, W. Xue, B. Ma, W. R. McCombie, G. J. Hannon, S. W. Lowe, M. T. Hemann, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 9053 – 9058.
[156] J. E. Carette, C. P. Guimaraes, M. Varadarajan, A. S. Park, I. Wuethrich, A. Godarova, M. Kotecki, B. H. Cochran, E. Spooner, H. L. Ploegh, T. R. Brummelkamp, *Science* **2009**, *326*, 1231 – 1235.
[157] J. H. Reiling, C. B. Clish, J. E. Carette, M. Varadarajan, T. R. Brummelkamp, D. M. Sabatini, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11756 – 11765.
[158] U. Elling, J. Taubenschmid, G. Wirnsberger, R. O'Malley, S.-P. Demers, Q. Vanhaelen, A. I. Shukalyuk, G. Schmauss, D. Schramek, F. Schnuetgen, H. von Melchner, J. R. Ecker, W. L. Stanford, J. Zuber, A. Stark, J. M. Penninger, *Cell Stem Cell* **2011**, *9*, 563 – 574.
[159] H. Luesch, S. K. Chanda, R. M. Raya, P. D. DeJesus, A. P. Orth, J. R. Walker, J. C. Izpisua Belmonte, P. G. Schultz, *Nat. Chem. Biol.* **2006**, *2*, 158 – 167.
[160] S. Fields, O.-k. Song, *Nature* **1989**, *340*, 245 – 246.
[161] E. J. Licitra, J. O. Liu, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12817 – 12821.
[162] M. Caligiuri, L. Molz, Q. Liu, F. Kaplan, J. P. Xu, J. Z. Majeti, R. Ramos-Kelsey, K. Murthi, S. Lievens, J. Tavernier, N. Kley, *Chem. Biol.* **2006**, *13*, 711 – 722.
[163] B. W. Jester, K. J. Cox, A. Gaj, C. D. Shomin, J. R. Porter, I. Ghosh, *J. Am. Chem. Soc.* **2010**, *132*, 11727 – 11735.
[164] F. Becker, K. Murthi, C. Smith, J. Come, N. Costa-Roldán, C. Kaufmann, U. Hanke, C. Degenhart, S. Baumann, W. Wallner, A. Huber, S. Dedier, S. Dill, D. Kinsman, M. Hediger, N. Bockovich, S. Meier-Ewert, A. F. Kluge, N. Kley, *Chem. Biol.* **2004**, *11*, 211 – 223.

[165] M. Paschke, *Appl. Microbiol. Biotechnol.* **2006**, *70*, 2 – 11.
[166] T. Clackson, J. A. Wells, *Trends Biotechnol.* **1994**, *12*, 173 – 184.
[167] R. Crameri, R. Jaussi, G. Menz, K. Blaser, *Eur. J. Biochem.* **1994**, *226*, 53 – 58.
[168] P. P. Sche, K. M. McKenzie, J. D. White, D. J. Austin, *Chem. Biol.* **1999**, *6*, 707 – 716.
[169] Y. Jin, J. Yu, Y. G. Yu, *Chem. Biol.* **2002**, *9*, 157 – 162.
[170] a) D. J. Rodi, R. W. Janes, H. J. Sanganee, R. A. Holton, B. A. Wallace, L. Makowski, *J. Mol. Biol.* **1999**, *285*, 197 – 203; b) S. Aoki, K. Morohashi, T. Sunoki, K. Kuramochi, S. Kobayashi, F. Sugawara, *Bioconjugate Chem.* **2007**, *18*, 1981 – 1986.
[171] J. S. Shim, J. Lee, H.-J. Park, S.-J. Park, H. J. Kwon, *Chem. Biol.* **2004**, *11*, 1455 – 1463.
[172] H. J. Jung, J. S. Shim, J. Lee, Y. M. Song, K. C. Park, S. H. Choi, N. D. Kim, J. H. Yoon, P. T. Mungai, P. T. Schumacker, H. J. Kwon, *J. Biol. Chem.* **2010**, *285*, 11584 – 11595.
[173] B. Van Dorst, J. Mehta, E. Rouah-Martin, V. Somers, W. De Coen, R. Blust, J. Robbens, *Toxicol. in Vitro* **2010**, *24*, 1435 – 1440.
[174] A. M. Piggott, P. Karuso, *ChemBioChem* **2008**, *9*, 524 – 530.
[175] K. Morohashi, H. Sahara, K. Watashi, K. Iwabata, T. Sunoki, K. Kuramochi, K. Takakusagi, H. Miyashita, N. Sato, A. Tanabe, K. Shimotohno, S. Kobayashi, K. Sakaguchi, F. Sugawara, *Plos One* **2011**, *6*, e18285.
[176] A. D. Keefe, *Current Protocols in Molecular Biology*, Wiley-VCH, Weinheim, **2001**.
[177] P. W. Hammond, J. Alpin, C. E. Rise, M. Wright, B. L. Kreider, *J. Biol. Chem.* **2001**, *276*, 20898 – 20906.
[178] M. McPherson, Y. Yang, P. W. Hammond, B. L. Kreider, *Chem. Biol.* **2002**, *9*, 691 – 698.
[179] H. Shiheido, F. Terada, N. Tabata, I. Hayakawa, N. Matsumura, H. Takashima, Y. Ogawa, W. Du, T. Yamada, M. Shoji, T. Sugai, N. Doi, S. Iijima, Y. Hattori, H. Yanagawa, *Plos One* **2012**, *7*, e38878.
[180] IntAct (http://www.ebi.ac.uk/intact/); STRING (http://string-db.org/), BioGRID (http://thebiogrid.org/); MINT (http://mint.bio.uniroma2.it/mint/); UniHI (http://www.unihi.org/); APID (http://bioinfow.dep.usal.es/apid/).
[181] G. Mathis, *Clin. Chem.* **1995**, *41*, 1391 – 1397.
[182] E. F. Ullman, H. Kirakossian, S. Singh, Z. P. Wu, B. R. Irvin, J. S. Pease, A. C. Switchenko, J. D. Irvine, A. Dafforn, C. N. Skold, et al., *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5426 – 5430.
[183] a) S. Duhr, D. Braun, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19678 – 19682; b) S. A. Seidel, C. J. Wienken, S. Geissler, M. Jerabek-Willemsen, S. Duhr, A. Reiter, D. Trauner, D. Braun, P. Baaske, *Angew. Chem.* **2012**, *124*, 10810 – 10814; *Angew. Chem. Int. Ed.* **2012**, *51*, 10656 – 10659.
[184] P. Dadvar, D. Kovanich, G. E. Folkers, K. Rumpel, R. Raijmakers, A. J. Heck, *ChemBioChem* **2009**, *10*, 2654 – 2662.
[185] T. Weidemann, J. M. Seifert, M. Hintersteiner, M. Auer, *J. Comb. Chem.* **2010**, *12*, 647 – 654.
[186] K. J. Cash, F. Ricci, K. W. Plaxco, *J. Am. Chem. Soc.* **2009**, *131*, 6955 – 6957.
[187] F. J. Dekker, O. Rocks, N. Vartak, S. Menninger, C. Hedberg, R. Balamurugan, S. Wetzel, S. Renner, M. Gerauer, B. Scholermann, M. Rusch, J. W. Kramer, D. Rauh, G. W. Coates, L. Brunsveld, P. I. Bastiaens, H. Waldmann, *Nat. Chem. Biol.* **2010**, *6*, 449 – 456.
[188] J. Yang, A. Shamji, S. Matchacheep, S. L. Schreiber, *Chem. Biol.* **2007**, *14*, 371 – 377.
[189] B. Chen, M. E. Dodge, W. Tang, J. Lu, Z. Ma, C.-W. Fan, S. Wei, W. Hao, J. Kilgore, N. S. Williams, M. G. Roth, J. F. Amatruda, C. Chen, L. Lum, *Nat. Chem. Biol.* **2009**, *5*, 100 – 107.
[190] a) S. Yamazoe, H. Shimogawa, S. Sato, J. D. Esko, M. Uesugi, *Chem. Biol.* **2009**, *16*, 773 – 782; b) S. Nishimura, Y. Arita, M. Honda, K. Iwamoto, A. Matsuyama, A. Shirai, H. Kawasaki, H. Kakeya, T. Kobayashi, S. Matsunaga, M. Yoshida, *Nat. Chem. Biol.* **2010**, *6*, 519 – 526.
[191] a) B. Jiang, H. Yang, M. L. Li, A. J. Hou, Q. B. Han, S. J. Wang, S. H. Li, H. D. Sun, *J. Nat. Prod.* **2002**, *65*, 1111 – 1116; b) Y. L. Wu, C. X. Liu, G. Q. Chen, *Blood* **2008**, *112*, 568 – 568.
[192] C. X. Liu, Q. Q. Yin, H. C. Zhou, Y. L. Wu, J. X. Pu, L. Xia, W. Liu, X. Huang, T. Jiang, M. X. Wu, L. C. He, Y. X. Zhao, X. L. Wang, W. L. Xiao, H. Z. Chen, Q. Zhao, A. W. Zhou, L. S. Wang, H. D. Sun, G. Q. Chen, *Nat. Chem. Biol.* **2012**, *8*, 486 – 493.
[193] C. Hedberg, F. J. Dekker, M. Rusch, S. Renner, S. Wetzel, N. Vartak, C. Gerding-Reimers, R. S. Bon, P. I. Bastiaens, H. Waldmann, *Angew. Chem.* **2011**, *123*, 10006 – 10011; *Angew. Chem. Int. Ed.* **2011**, *50*, 9832 – 9837.
[194] J. M. Krysiak, J. Kreuzer, P. Macheroux, A. Hermetter, S. A. Sieber, R. Breinbauer, *Angew. Chem.* **2012**, *124*, 7142 – 7147; *Angew. Chem. Int. Ed.* **2012**, *51*, 7035 – 7040.
[195] B. P. Zhou, D. A. Lewis, S. W. Kwan, C. W. Abell, *J. Biol. Chem.* **1995**, *270*, 23653 – 23660.
[196] M. A. Amin, S. Matsunaga, S. Uchiyama, K. Fukui, *FEBS Lett.* **2008**, *582*, 3839 – 3844.
[197] W. Wang, A. Budhu, M. Forgues, X. W. Wang, *Nat. Cell Biol.* **2005**, *7*, 823 – 830.
[198] A. Ballio, M. Mauri, E. B. Chain, P. Deleo, A. Tonolo, B. F. Erlanger, *Nature* **1964**, *203*, 297.
[199] M. Würtele, C. Jelich-Ottmann, A. Wittinghofer, C. Oecking, *EMBO J.* **2003**, *22*, 987 – 994.
[200] J. Z. Long, B. F. Cravatt, *Chem. Rev.* **2011**, *111*, 6022 – 6063.
[201] A. Adibekian, B. R. Martin, C. Wang, K. L. Hsu, D. A. Bachovchin, S. Niessen, H. Hoover, B. F. Cravatt, *Nat. Chem. Biol.* **2011**, *7*, 469 – 478.
[202] A. Adibekian, B. R. Martin, J. W. Chang, K. L. Hsu, K. Tsuboi, D. A. Bachovchin, A. E. Speers, S. J. Brown, T. Spicer, V. Fernandez-Vega, J. Ferguson, P. S. Hodder, H. Rosen, B. F. Cravatt, *J. Am. Chem. Soc.* **2012**, *134*, 10345 – 10348.
[203] J. J. Fischer, O. Y. Graebner Baessler, C. Dalhoff, S. Michaelis, A. K. Schrey, J. Ungewiss, K. Andrich, D. Jeske, F. Kroll, M. Glinski, M. Sefkow, M. Dreger, H. Koester, *J. Proteome Res.* **2010**, *9*, 806 – 817.
[204] D. P. Bezerra, G. C. G. Militao, F. O. de Castro, C. Pessoa, M. O. de Moraes, E. R. Silveira, M. A. S. Lima, F. J. M. Elmiro, L. V. Costa-Lotufo, *Toxicol. in Vitro* **2007**, *21*, 1 – 8.
[205] T. Ito, H. Ando, T. Suzuki, T. Ogura, K. Hotta, Y. Imamura, Y. Yamaguchi, H. Handa, *Science* **2010**, *327*, 1345 – 1350.
[206] J. Cox, M. Mann, *Nat. Biotechnol.* **2008**, *26*, 1367 – 1372.
[207] a) M. D. Burke, S. L. Schreiber, *Angew. Chem.* **2004**, *116*, 48 – 60; *Angew. Chem. Int. Ed.* **2004**, *43*, 46 – 58; b) G. L. Thomas, E. E. Wyatt, D. R. Spring, *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 700 – 712.
[208] a) M. M. Hann, T. I. Oprea, *Curr. Opin. Chem. Biol.* **2004**, *8*, 255 – 263; b) W. L- Scott, M. J. O'Donnell, *J. Comb. Chem.* **2009**, *11*, 3 – 13.