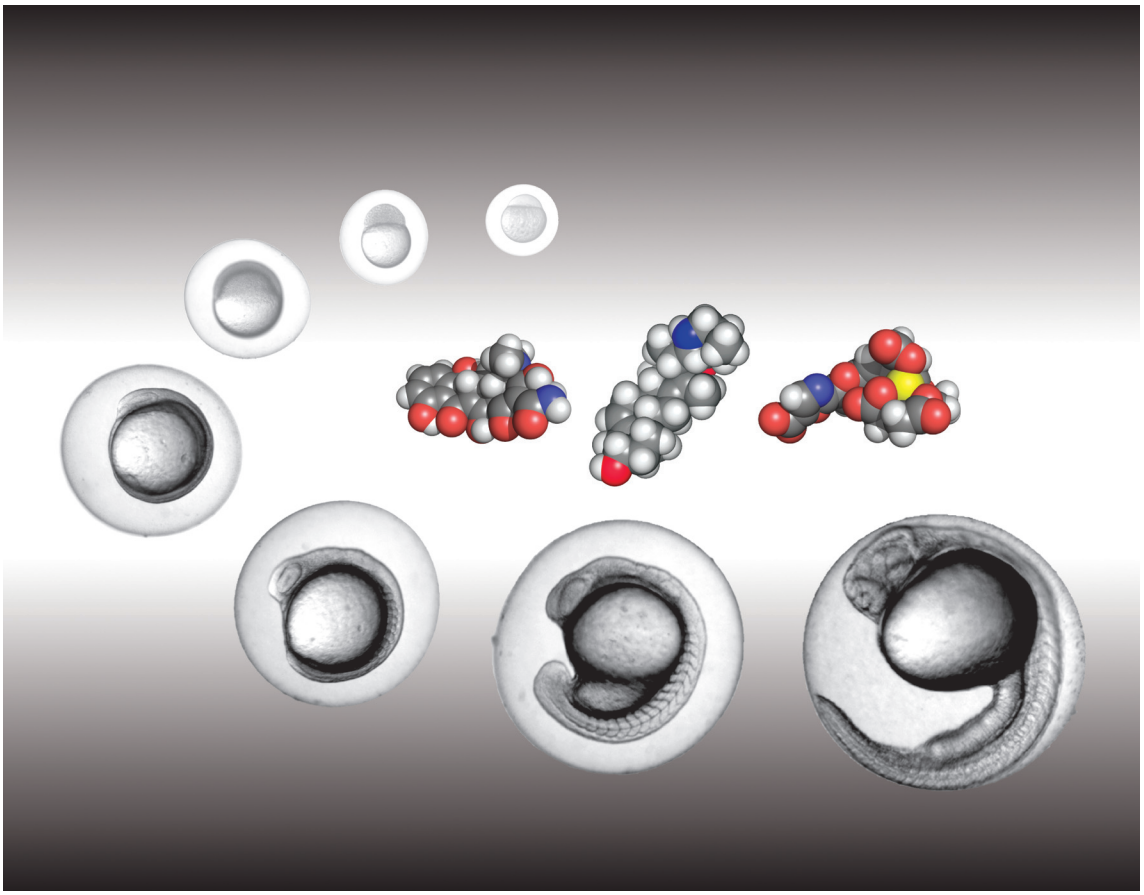
Chem Soc Rev

This article was published as part of the

2008 Chemistry–Biology Interface Issue

Reviewing research at the interface where chemistry meets biology

Please take a look at the full table of contents to access the other papers in this issue



Identification of the cellular targets of bioactive small organic molecules using affinity reagents†

Benjamin J. Leslie and Paul J. Hergenrother*

Received 11th April 2008
First published as an Advance Article on the web 23rd May 2008
DOI: 10.1039/b702942j

The elucidation of molecular targets of bioactive small organic molecules remains a significant challenge in modern biomedical research and drug discovery. This *tutorial review* summarizes strategies for the derivatization of bioactive small molecules and their use as affinity probes to identify cellular binding partners. Special emphasis is placed on logistical concerns as well as common problems encountered during such target identification experiments. The roadmap provided is a guide through the process of affinity probe selection, target identification, and downstream target validation.

1 Introduction

The ability of small organic molecules to induce phenotypic changes in a cell or organism by modulating protein function has made them invaluable as pharmaceuticals and as tools for the study of complex cellular processes. Many drugs in use today were discovered by phenotypic observation of the effect of a compound or extract on a cell or organism. In many cases, the drug's mechanism of action was not known until many years after its clinical implementation. Despite technological advances, identification of macromolecular targets of biologically active small molecules remains a central problem in medicinal chemistry and chemical biology.

Innovations in high-throughput screening and parallel synthesis have led to an increase in the number of new "lead" molecules and the rate at which large numbers of derivatives can be synthesized. However, access to large compound screening collections and enhanced throughput capabilities has yet to increase the number of drugs targeting new proteins and protein families. Today more than 50% of the 1357 FDA-approved drugs target just four of the predicted 16 000 protein families in the genome.[1] Unfortunately, there is still considerable risk and uncertainty associated with the discovery and validation of a novel drug target.

Thus, the determination of the protein targets of bioactive small molecules is a key step toward increasing the diversity of drug targets. Some of the most potent modulators of cellular functions are complex natural products. Frequently, extensive biological data are unavailable for these compounds, as challenges in chemical synthesis or isolation from natural sources limit their supply. This tutorial review is aimed at the researcher in possession of biologically active molecules, the targets of which could lead to novel strategies for the treatment of human disease. Highlighted herein are experiments using common reagents and equipment that can be performed to identify the protein binding partner of a bioactive small molecule. Although the strategies detailed in this review are applicable to the identification of targets from many classes of biomolecules, this tutorial will focus on the identification of target proteins, as they are currently without a doubt the most important macromolecular targets in drug

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana, IL 61801, USA. E-mail: hergenro@uiuc.edu
† Part of a thematic issue examining the interface of chemistry with biology.

Benjamin J. Leslie

Ben Leslie received a bachelor's degree in Chemistry from Colgate University. Currently Ben is a PhD candidate in the laboratories of Paul J. Hergenrother at the University of Illinois, Urbana-Champaign. His research involves the synthesis and discovery of novel anticancer small molecules and the use of these molecules to discover novel protein targets for the treatment of cancer.

Paul J. Hergenrother

Paul Hergenrother received a PhD with Prof. Stephen Martin at the U. of Texas-Austin (in 1999), then moved to the laboratory of Prof. Stuart Schreiber at Harvard U. as an American Cancer Society postdoctoral fellow. He began his independent career at U. of Illinois in 2001. The Hergenrother laboratory uses small molecules to define novel targets for the treatment of intractable diseases, including cancer, neurodegeneration, and drug-resistant bacteria.

This journal is © The Royal Society of Chemistry 2008

discovery and development.[1] The molecules featured in this tutorial are mainly "druglike" or natural product-like molecules, as the targets of structures of this type are most readily "druggable" given the current paradigm of molecular medicine. Such compounds, however, are just one of many classes of important molecules that modulate cellular functions. There are several excellent examples of the affinity purification of molecular targets of peptides,[2] phosphopeptides,[3] phospholipids and phosphoinositides,[4] glycopeptides,[5] and oligosaccharides.[6]

2 Overview of target identification strategies

Once a compound that induces a certain phenotypic response in a cell culture or *in vivo* experiment has been identified, several techniques may be employed to identify its protein target. This review focuses on methods where modified versions of lead compounds are synthesized, allowing the protein binder to be determined in a direct fashion. For target identification strategies based on technology-driven functional genetic or bioinformatic platforms, the reader is referred to several excellent reviews.[7–9] The central theme of the experiments described herein is that a binding event between the compound and its target is used to directly identify the target protein. A critical assumption is that the phenotype elicited by a small molecule arises from the cellular consequence of a covalent or non-covalent interaction between the small molecule and its target protein(s).

An overview of the various strategies is shown in Fig. 1; the choice of approach somewhat depends on whether it is suspected the compound binds covalently (which may be reversible) or non-covalently to its target. *Affinity matrices*, where the target molecule is covalently attached to a polymeric solid support, have been used for both irreversible and reversible binders. *Biotinylated* versions of the compound of interest are commonly synthesized and these probes can be utilized either in affinity matrix, cell lysate, or live cell experiments. *Radiolabel/imaging* probe molecules can be useful for target identification experiments, generally with covalent binders. Finally, *photoaffinity* probes can be used to artificially link a noncovalent compound to its target. The relative merits and shortcomings of these techniques are discussed below and representative examples for each category are shown in Fig. 1.

3 Target identification strategies employing affinity chromatography ("pulldowns")

The simplest and oldest of these techniques involves the tethering of modified versions of lead compounds to solid supports, followed by incubation with a lysate of a desired cell line, organ, or organism.[10–20] This method allows for affinity-based separation of the protein binding partner from all other proteins in the cell by simple filtration, washing, and elution. In most cases, the protein partners are then identified by modern degradative mass spectrometric techniques. The solid support, or resin, for these techniques is typically an agarose or sepharose-based polymer that is functionalized with reactive functional groups such as amines, thiols, or carboxylic acids. Activated resins functionalized with *N*-hydroxysuccinimidyl (NHS) esters are by far the most commonly used and will readily couple with amino and hydroxy functionalities; NHS resins are available from Bio-Rad, Pierce Biotechnology, Aldrich, and many other vendors.

In a typical affinity chromatographic purification, a cellular lysate or tissue homogenate is incubated with the affinity matrix under conditions designed to minimize protein degradation, typically at 4 °C in the presence of protease inhibitors. Incubation is usually performed in a so-called "batch" mode, where the suspension is actively agitated by stirring, rocking, or shaking to maximize protein contact with the resin. After incubation, the affinity resin is washed extensively with an aqueous buffer to elute any non-binding proteins from the resin, either by serially centrifuging the resin and aspirating the supernatant, or by washing the resin in fritted glass or plastic chromatography columns. At this point, the only proteins that remain are those which bind to the ligand or to some part of the affinity matrix. These proteins are then eluted from the affinity matrix under denaturing conditions, or by incubation with the free ligand, and resolved by electrophoresis on an SDS-PAGE gel. This purified target protein can then be identified by peptide microsequencing, immunohistochemical analysis, or, most commonly, by modern degradative mass spectrometric techniques. The protein targets of many biologically important natural products and small molecules have been discovered in this way. In addition to cultured human cells, affinity purifications have been performed using animal tissue or organ homogenates as a cheap, readily accessible source of protein.[10,11,21]

In a classic example, Schreiber and co-workers demonstrated that an agarose-supported derivative of the immunosuppressant FK-506, when incubated with the cellular lysate of bovine thymus or human spleen was able to selectively purify FKBP12, a *cis–trans* prolyl isomerase.[12] FKBP12 was competitively eluted from the matrix using soluble FK-506 (**1**, Fig. 2A), and was not observed to bind to a control resin which had been functionalized with ethanolamine instead of FK-506, indicating that FKBP12 was isolated on the affinity resin because it bound specifically to FK-506. Gels showing the use of resin **2** to affinity purify FKBP12 from bovine thymus and spleen lysates are shown in Fig. 2B.

This example illustrates the importance of control experiments in affinity-based target identification procedures, as it is common to isolate multiple proteins through affinity chromatography owing to nonspecific interactions between proteins and the affinity matrix. Although other solid supports have been used, including resins designed for use in solid-phase chemical synthesis,[13] sugar-based agarose and sepharose resins are the most popular, as their polar, uncharged surfaces reduce nonspecific binding of hydrophobic or charged proteins. A matrix control can be carried out in two ways. First, a "resin-only" control can be performed by coupling a blocking reagent such as ethanolamine to the resin of choice (see ref. 10, 14–17). Blocking agents are routinely added after coupling a compound to a solid support to quench all activated positions, minimizing the charge of the resin. Better still is the "inactive compound," or comparison, control, where a structurally similar, but biologically inactive version of the compound is used (see ref. 11, 14–16, 18, 19). Ideally this ligand

 This journal is © The Royal Society of Chemistry 2008

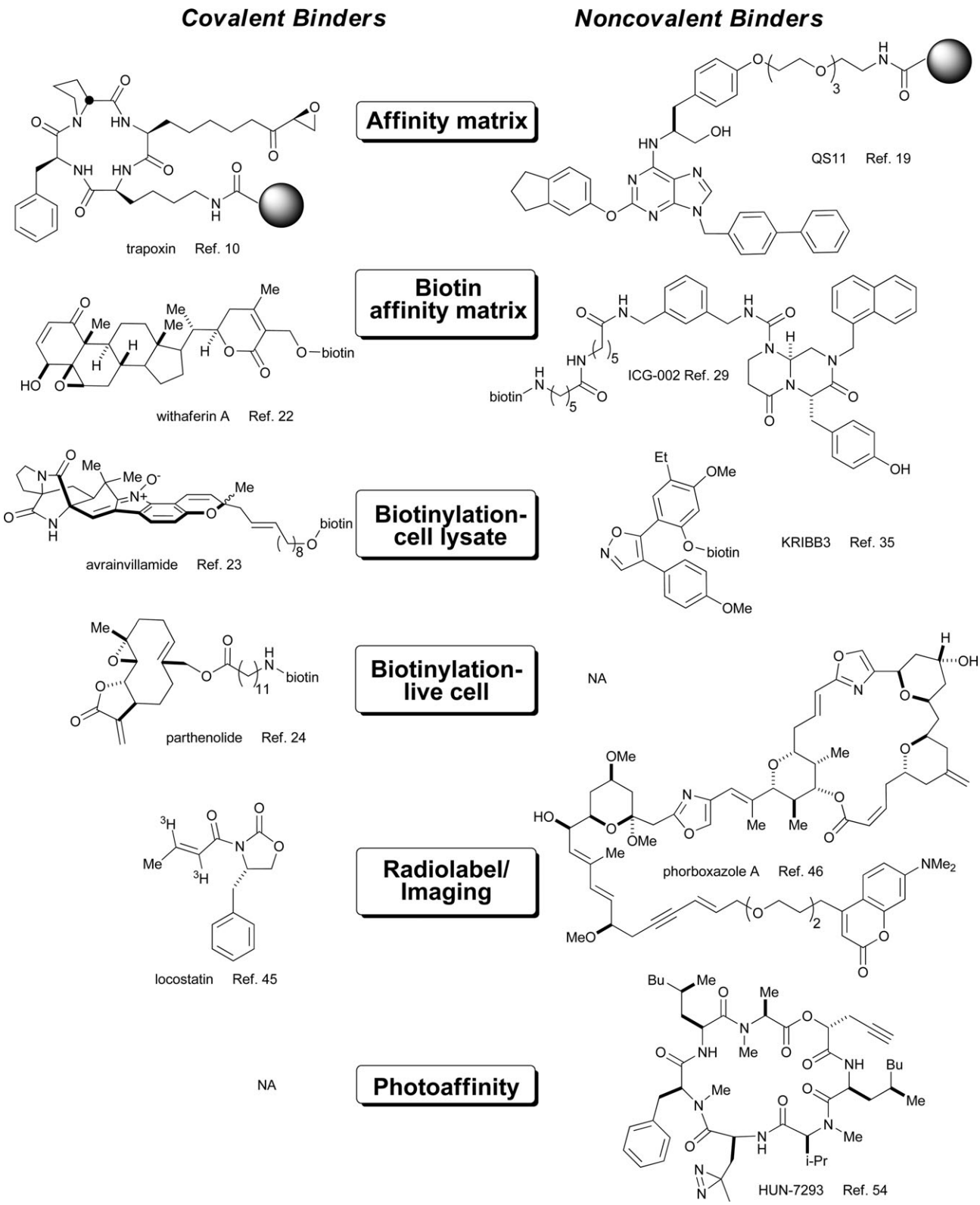


Fig. 1 An overview of the various strategies employed in affinity-based target identification experiments. Specific examples of probes used that bind their target covalently are on the left, and those that bind non-covalently are on the right. Affinity matrices are discussed in section 3, biotinylation strategies in section 4, radiolabel/imaging approaches in section 5, and photoaffinity probes in section 6.

would be a completely inactive enantiomer of the target compound, as was recently demonstrated by Meijer and co-workers in their discovery that affinity matrices made from the "selective cyclin-dependent kinase inhibitor" *R*-roscovitine

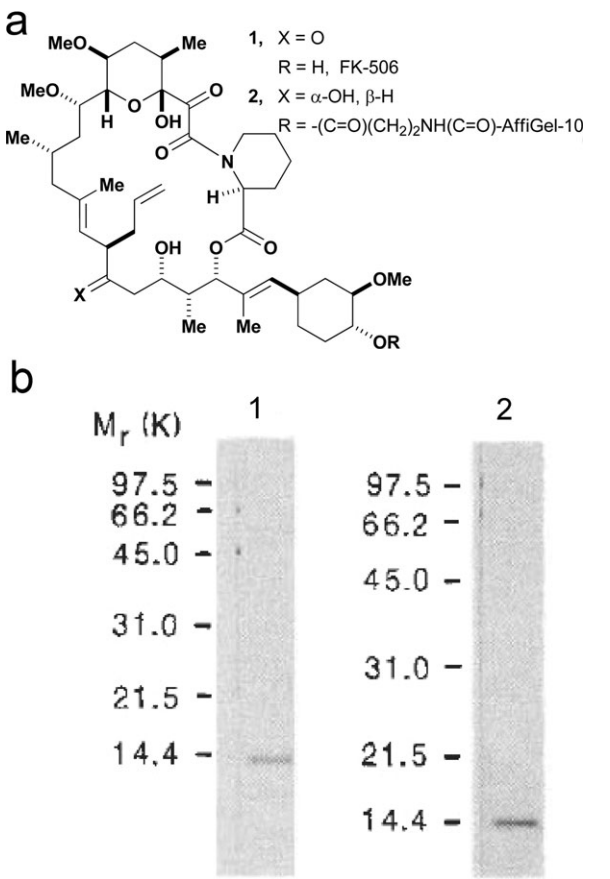


Fig. 2 FKBP12 is a cellular target of FK-506. (a) FK-506 (**1**) and an analogue-functionalized affinity matrix (**2**). (b) FK-506-binding protein 12 was isolated from bovine thymus (lane 1) and human spleen (lane 2) lysates by competitive displacement with **1** of proteins adsorbed onto affinity matrix **2**. Gels visualized by silver staining. Image reproduced with permission from Harding *et al.*[12] Copyright 1989 Nature Publishing Group.

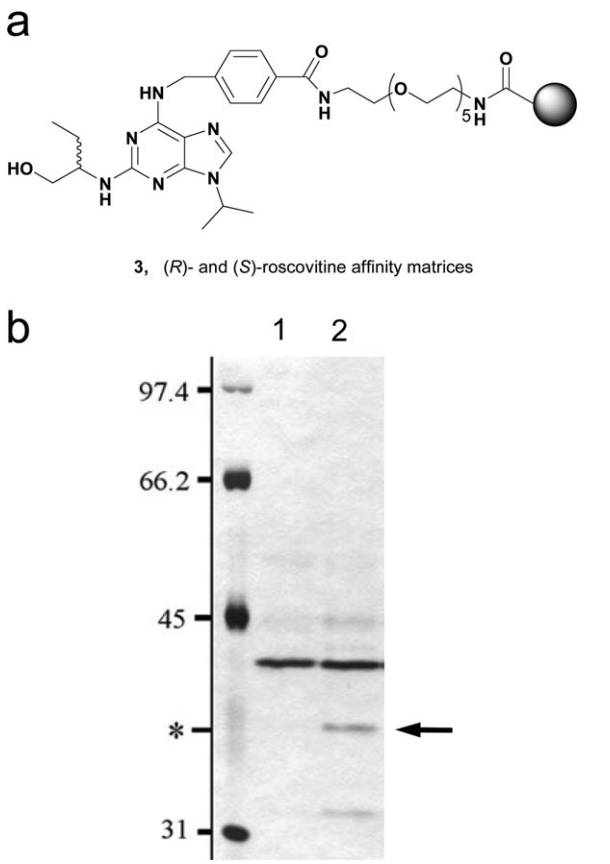


Fig. 3 Pyridoxal kinase is a cellular target of *R*-roscovitine. (a) Enantiomeric roscovitine affinity matrices (**3**). (b) Proteins isolated by affinity matrices functionalized with *S*-**3** (lane 1) and *R*-**3** (lane 2). In addition to known CDKs, *R*-**3** pulls down pyridoxal kinase (arrow). Gel visualized by silver staining. Image reproduced with permission from Bach *et al.*[11] Copyright 2005 American Society for Biochemistry and Molecular Biology.

also targets pyridoxal kinase (Fig. 3).[11] This control can exclude proteins which bind to the ligand on the matrix, but do so nonspecifically. An affinity purification using these control resins are run alongside the experiment, and the proteins eluted from these matrices can be discounted from the experimental sample.

Another important control experiment is a so-called ''competition'' control (see ref. 10, 11, 14, 18, 19, 22–27). In this experiment, the compound is added in solution to the protein sample at a concentration in excess of the ligand loading on the affinity matrix. This creates a competitive environment, where the protein target of a given small molecule is bound by the excess of soluble probe and cannot interact with the affinity matrix. Under the assumption that the affinity matrix has been constructed in such a way that it makes the same molecular contacts with the target protein as the compound itself, *the loss* of protein bands in an SDS-PAGE gel indicates a protein that was competed from the resin by a specific molecular interaction with the small molecule. Thus the target is identified by its presence in a standard solid-phase pulldown and its absence in a competition experiment. Alternatively, the competitor can be used as an agent to selectively elute bound proteins from the affinity matrix.[12,28,29]

One of the limitations of solid-phase affinity matrices is that their biological activity cannot be validated; there is no simple way to confirm that the chemical modification has not affected the molecule's ability to bind to the protein target. To counteract this problem, compounds are typically attached to affinity matrices through long, hydrophilic linkers derived from poly(ethylene glycol).[11,14–16,18–20] Then, biological activity of the linker-functionalized compound can be determined to confirm that the linker-modified compound will interact with the same proteins as the parent molecule. Additionally, the hydrophilic nature of the linker minimizes nonspecific binding and separates the probe from the surface of the resin, allowing for conformational flexibility of the probe so that it may assume a favorable binding pose, and preventing the steric bulk of the resin from blocking target protein binding.

One creative solution to the problem of attachment of compounds to solid supports has recently been put forth by Sugawara and co-workers (Fig. 4).[13] TentaGel resins were functionalized with a diazirine-based photoreactive group to generate resin **4** that, upon irradiation with UV light, nonselectively formed covalent bonds to small molecule inhibitors

 This journal is © The Royal Society of Chemistry 2008

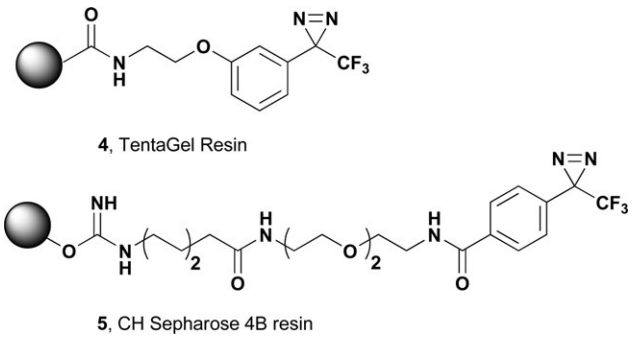


Fig. 4 Photocrosslinking resins used for covalent immobilization of unfunctionalized small molecule ligands.

of DNA polymerase β *via* C–H insertion. Although the exact site of C–H insertion on the small molecules was unknown, three different affinity matrices were generated with the ability to bind and purify pol β from a mixture of proteins, demonstrating the ability to create affinity matrices with no prior chemical modification of the target ligand. This technique had previously been reported by Osada and co-workers, who used sepharose resin **5** to immobilize cyclosporin A and FK-506 and isolate cyclophilin A and FKBP12, respectively, in proof-of-principle experiments.[20]

In order to rapidly identify the target of bioactive compounds, the concept of tagged library screening has been introduced to reduce time spent on structure–activity relationship (SAR) analysis (Fig. 5).[14,15,18] As a demonstration of this concept, a combinatorial library of 1536 multifunctionalized triazines was synthesized containing amino-triethylene glycol linkers and was screened for phenotypic effects in zebrafish. Active compounds which influence skin pigmentation (**6**) and brain/eye development (**7**) were subjected to Boc deprotection, and the resulting amine was then coupled to Affi-gel 10 (an NHS ester functionalized agarose resin) to generate affinity matrices with known activity, allowing for rapid identification of protein binding partners without further chemical synthesis.

Two critical factors in affinity chromatographic experiments are the affinity of a ligand for its target, and the ability to detect *specific* interactions over nonspecific ones. Logistical concerns regarding these factors have recently been highlighted by Burdine and Kodadek.[8] The affinity of a ligand for its target is a central concern, as the strength of this interaction has historically been the largest predictor of success in a target identification experiment. Protein binders must remain bound to affinity resins during the extensive washing required to remove nonspecific binding partners in order to be

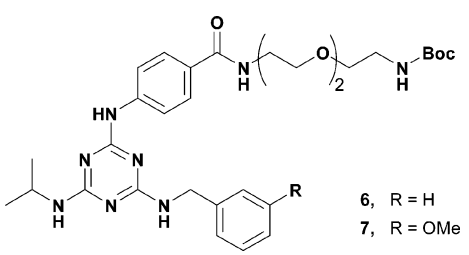


Fig. 5 Bioactive compounds from a tagged triazine library used directly in affinity purification experiments.

detected. Fortunately, natural products and druglike molecules often bind tightly to their targets. However two promising strategies have been proposed for target identification using low-affinity ligands. The first involves combinatorial assembly of two or more modest-affinity lead molecules in an effort to discover high-affinity bivalent noncompetitive ligands.[30,31] The second involves incorporation of a photoreactive group onto a lead scaffold to turn a weak-affinity ligand into an irreversible binder (see section 6 below).

Identification of protein targets is dependent on the ability to enrich the concentration of ligand binding partners over background nonspecific binding by attaining a balance between maximal target binding and minimal nonspecific interactions. As organic small molecules are typically uncharged and hydrophobic, high ligand loading on the resin can contribute to nonspecific interactions with the ligand. However, use of excessive amounts of affinity matrix may lead to increased nonspecific interactions with the resin. In practice it is advisable to construct a matrix with low ligand loading (modulating its concentration with a blocking agent such as ethanolamine), and then using the minimal amount of resin to isolate a detectable amount of protein. Selection of an analysis technique with a low limit of detection enables smaller amounts of protein to be used. For gel analysis, protein stains such as Sypro Ruby or Deep Purple have largely supplanted silver staining, as they can detect nanogram amounts of protein, and are more compatible with mass spectrometric analysis.

Since detection of a target is dependent on both the affinity and abundance of target proteins, it is imperative to maximize the ratio of total protein to ligand to ensure that high-affinity, but low-abundance binding partners can be detected over low-affinity, but high-abundance proteins. This means it is imperative to use as much protein as possible, as it should be assumed that the target is of low abundance until evidence suggests otherwise. Unfortunately, the costs associated with culturing large amounts of mammalian cells can be prohibitive. Two solutions to this limitation are to spike expensive mammalian cell lysates with *E. coli* lysates or to purchase large amounts of mammalian cells on ice from commercial suppliers.[8]

4 Target identification strategies employing biotinylated probes

Another common technique for target identification is the use of biotinylated probes (see ref. 21–25, 27–29, 32–40). Biotin, vitamin B_7, binds to the proteins avidin, streptavidin, and NeutrAvidin with dissociation constants of $\sim 10^{-15}$ M, making it one of the strongest noncovalent protein–ligand interactions known. Thus, avidin-functionalized resins, when incubated with a biotinylated probe, will create a noncovalent affinity matrix that is stable under all but the harshest conditions. Biotin and several derivatives are commercially available, allowing it to be easily appended to most common functional groups. Avidin's extensive glycosylation is responsible for its high charge (pI = 10) and propensity for nonspecific binding. The more expensive deglycosylated variants

streptavidin (pI = 5) and NeutrAvidin (pI = 6.3) are commonly used in order to minimize nonspecific binding.

One major advantage over the use of affinity matrices is that biotinylated probes are discrete small molecules that can be fully characterized. Since the binding of biotin to avidin is quantitative and essentially irreversible, the loading of a known amount of probe onto a solid support is possible. In contrast, when coupling affinity probes directly to commercially available solid-phase resins, analysis of coupling efficiency and chemical characterization of the matrix can be problematic, although some creative solutions have been put forward.[20]

Biotinylation strategies are not without their limitations, as biotinylation of a target molecule often significantly decreases its aqueous solubility. In addition, the solubility of biotin conjugates in standard organic solvents can also be poor; thus the introduction of biotin and further manipulation of biotin conjugates must be carried out in high-boiling solvents like DMF or DMSO. Consequently, biotinylation reactions are usually carried out on a small scale in the final step of a synthetic route.

Biotinylated probes can be utilized in various ways in target identification strategies. (1) They can be pre-complexed with a streptavidin resin, creating a non-covalent affinity matrix. This approach is directly analogous to the affinity resins employed in section 3, above. (2) They can be incubated with a cell lysate, and then the streptavidin resin can be added to pull out the biotinylated probe together with any bound protein. (3) They can be incubated with live cells, and after cell lysis, can be pulled out of solution with streptavidin resin, which will then pull out any bound protein. Biotinylated probes and their attached proteins may be eluted from avidin resins by 8 M guanidine–HCl, pH 1.5, or by boiling in a typical denaturing SDS-PAGE loading dye. Alternatives such as Invitrogen's CaptAvidin resin allow elution under more gentle conditions. Another option is the incorporation of peptide-based protease cleavage sites, to allow for selective ligand cleavage under gentle conditions.[32,33] One further advantage of biotinylated probes is that they also serve as imaging agents for immunoblots and microscopy through the use of labeled streptavidin or antibodies raised against biotin itself.[27]

4.1 Biotinylated probe pre-complexed with resin

Wang and co-workers recently used a biotinylated derivative of the antimitotic bicyclic natural product diazonamide A (**8**) to identify the mitochondrial enzyme ornithine δ-amino transferase (OAT) as its target (Fig. 6A).[34] In this experiment, biotinylated derivative **9** and *seco*-analog **10** were preincubated with avidin beads to generate noncovalent affinity matrices that were subsequently used to purify OAT from anion-exchange fractionated protein samples from HeLa (human cervical cancer) cell lysates, much like a solid-phase affinity matrix discussed in section 3 would be used. Importantly, because **9** is cell-permeable, the authors were able to first show that it induced the same mitotic spindle abnormalities as parent compound **8**, suggesting that the protein that diazonamide A interacts with inside the cell was indeed the same protein that was isolated by probe **9**. As shown in the gel

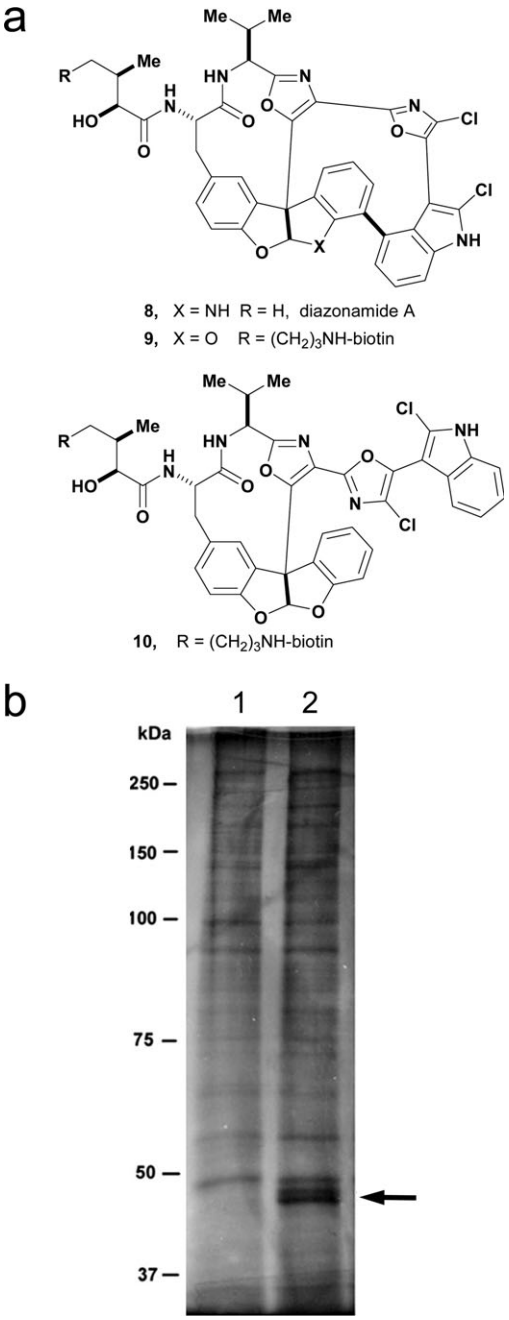


Fig. 6 OAT is a cellular target of diazonamide A. (a) Diazonamide A (**8**) and biotinylated derivatives (**9** and **10**). (b) Fractions of HeLa extract were incubated with **10** (lane 1) or **9** (lane 2) bound to avidin-agarose. Proteins eluted from the affinity matrix were separated *via* SDS-PAGE and analyzed by silver staining. OAT (arrow) binds specifically to **9**, but not inactive comparison variant **10**. Image reproduced with permission from Wang *et al.*[34]

in Fig. 6B, resins comprised of compound **9** (lane 2), but not of inactive analog **10** (lane 1), were able to purify OAT.

In order to decrease the nonspecific binding often observed when using avidin-based resins, a protein sample can be "precleared" of nonspecific binders by passing the cellular lysate though one of the control resins *before* incubation with the affinity matrix. This technique was recently used by Whitesell and co-workers to determine that the natural product

 This journal is © The Royal Society of Chemistry 2008

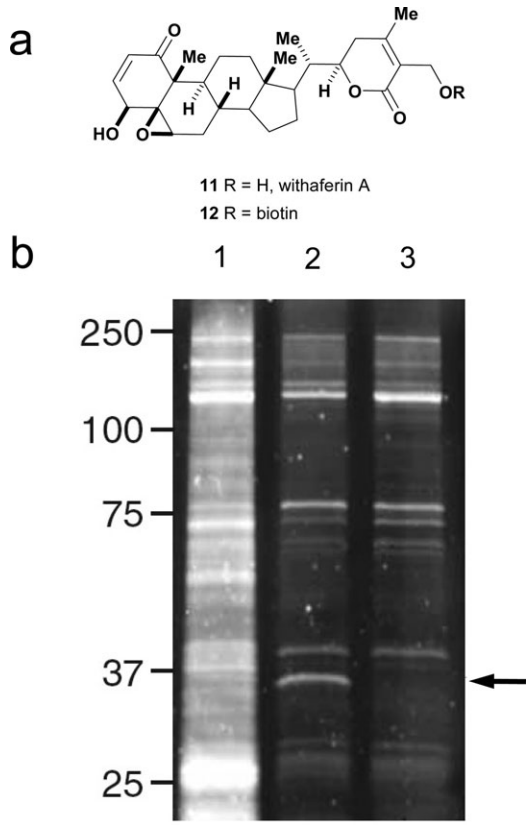


Fig. 7 Annexin II is a cellular target of withaferin A. (a) Withaferin A (**11**) and biotinyl–withaferin (**12**). (b) An affinity matrix obtained by binding **12** to NeutrAvidin beads was incubated with a whole-cell extract that had been treated with competitor **11** (lane 3) or DMSO (lane 2). Bound proteins were then eluted, run on a SDS-PAGE gel, and visualized with Sypro Ruby. The target, annexin II (arrow), is absent in the competitor lane. Lane 1 is the "preclear" sample, proteins which bind nonspecifically to unfunctionalized NeutrAvidin beads. Image reproduced with permission from Falsey *et al.*[22] Copyright 2006 Nature Publishing Group.

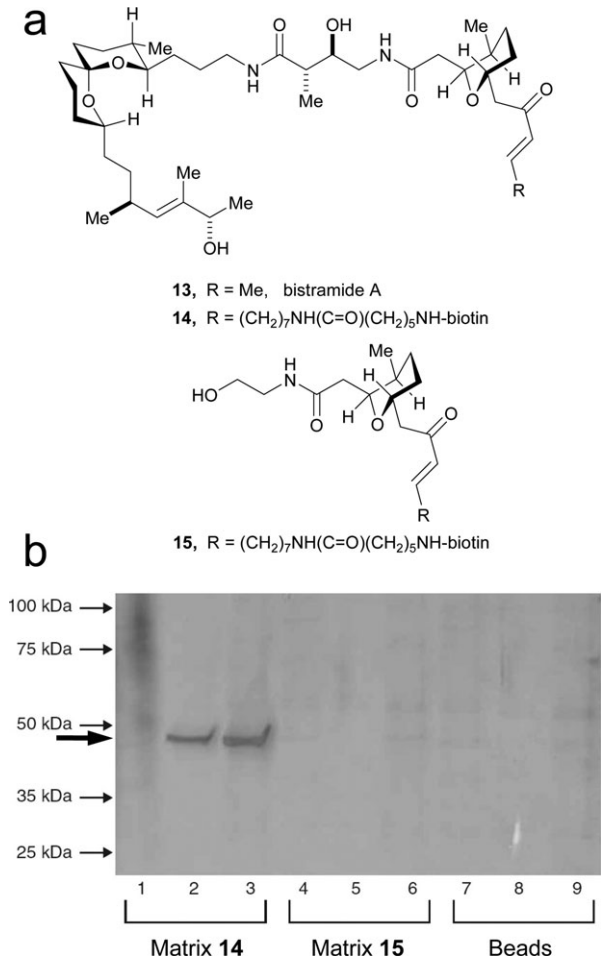


Fig. 8 Actin is a cellular target of bistramide A. (a) Bistramide A (**13**) and biotinylated analogs **14** and **15**. (b) Compounds **14** and **15** were preincubated with streptavidin beads, and incubated with A549 cell lysates. Matrices were eluted with lysis buffer (lanes 1, 4, 7), 82 μM bistramide A (lanes 2, 5, 8), and then 0.1% SDS (lanes 3, 6, 9). Elution fractions were run on poly(acrylamide) gels and silver stained. Actin (arrow) is isolated only by matrix **14** and is competitively displaced by **13** (lane 2). Image reproduced with permission from Statsuk *et al.*[28] Copyright 2005 Nature Publishing Group.

withaferin A (**11**) and its biotinylated probe **12** covalently bind to annexin II, which induces actin microfilament aggregation (Fig. 7A).[22] Sensitive protein detection with Sypro Ruby shows that many proteins isolated by probe **12** complexed with NeutrAvidin beads (Fig. 7B, lane 2) were still present when the experiment was performed in the presence of excess **11** (lane 3). These proteins bind to the resin, and not to the probe, as confirmed by their presence in the "preclear" lane (lane 1). The band that is present in lane 2, but absent in the presence of competitor (lane 3), is the target, determined to be annexin II (indicated by the arrow in Fig. 7B). This set of experiments indicates that even nonglycosylated variants of avidin can exhibit significant nonspecific binding.

A similar strategy was also used by Kozmin and co-workers to identify actin as the molecular target of the natural product bistramide A (**13**, Fig. 8).[28] Modification of their previously published convergent synthesis of bistramide A provided bioactive affinity probe **14** in seven synthetic steps. The authors were able to use a synthetic intermediate lacking spiroketal functionality as a comparison variant, inactive control compound **15** (Fig. 8A). Neither the comparison variant (Fig. 8B, lanes 4–6) nor the resin itself (lanes 7–9) isolated significant amounts of protein, whereas affinity matrix **14** bound solely and specifically to actin (lanes 2 and 3), which was eluted from the matrix with **13** or SDS.

4.2 Biotinylated probe incubated with cell lysate

It is less common to first incubate a biotinylated probe with a protein lysate and *then* add a streptavidin resin.[21,23] This protocol is non-ideal for several reasons. First, while the binding of biotin to streptavidin is rapid, pre-incubation of biotinylated compounds with streptavidin resins may be carried out overnight without risk of degradation, which ensures efficient coupling. Most importantly, pre-incubation allows the use of excess of the biotinylated probe (because the resin may be washed afterwards) to ensure that all biotin binding sites are saturated, preventing the binding of endogenous biotinylated proteins present in a cellular lysate. This creates an affinity matrix with maximal compound loading,

minimizing the amount of resin that must be used to pull down a detectable amount of protein, thus reducing the total amount of nonspecific binders isolated in the experiment.

4.3 Biotinylated probe incubated with live cells

An advantage of biotinylated probes over solid-phase affinity matrices is that they are often cell-permeable, meaning that it is possible for the probe to interact with a target protein in its native environment, inside a living cell or organism. Ultimately, the goal of an affinity purification experiment is to discover the protein that interacts with a compound inside the cell to produce a certain phenotype, not just any protein that binds to the compound *in vitro*. Consequently, experimental protocols that simplify affinity purification experiments, such as cellular lysis or subcellular fractionation, rely on the assumption that cellular lysis and subsequent manipulation do not disrupt the folding state or binding partners of the target protein needed for the compound to induce its cellular phenotype. As protein targets of many small molecules have been identified from cellular lysates, it certainly is the case that many proteins do retain their ability to bind compounds even after cellular lysis. However, the possibility remains that such simplifications are to blame in cases where affinity probes fail to isolate a protein binding partner.

In an interesting case, Mohan and co-workers reported the identification of the intermediate filament protein vimentin as a *second* covalent target of withaferin A, one year after Whitesell's identification of annexin II.[41] The authors note that, using a different biotinylated probe than Whitesell (which incorporated a longer aliphatic linker in between withaferin A and biotin), annexin II binding was not detected. This finding highlights two important points, namely that (1) small molecules may have more than one binding partner that may contribute to the same or different phenotypes, and (2) that the variables such as probe construction (linker length and composition) and experimental design (*in vitro vs.* live-cell labeling, cell type, *etc.*) may contribute to the preferential isolation of one target over another.

For live cell experiments, the biotinylated probe is incubated with cells for a defined period of time, cells are lysed, and an avidin resin is added in order to sequester the probe-bound target protein from the proteome. Decanting the supernatant leaves only the resin–probe complex and any proteins which bind to the probe. Myers and co-workers recently used this procedure to show that natural product (+)-avrainvillamide (**16**, Fig. 9A) covalently targets the oncoprotein nucleophosmin, a multifunctional protein overexpressed in many human cancers.[23] By switching one coupling partner in step 15 of their 17-step route to the natural product, more than 30 analogs of avrainvillamide were synthesized, including biotinylated analog **17**. Live T-47D cells were treated with **17** or a less potent, structurally simpler analog (not shown) for 90 min and then lysed and treated with streptavidin resin to purify nucleophosmin (Fig. 9B, compare lane 2 with lane 4). The greater potency of **16** over simpler analogs translates into more efficient and more specific binding in its associated probes.[23] After nucleophosmin purified from *in situ* isolation was identified by mass spectrometry, competition studies were carried out in cell

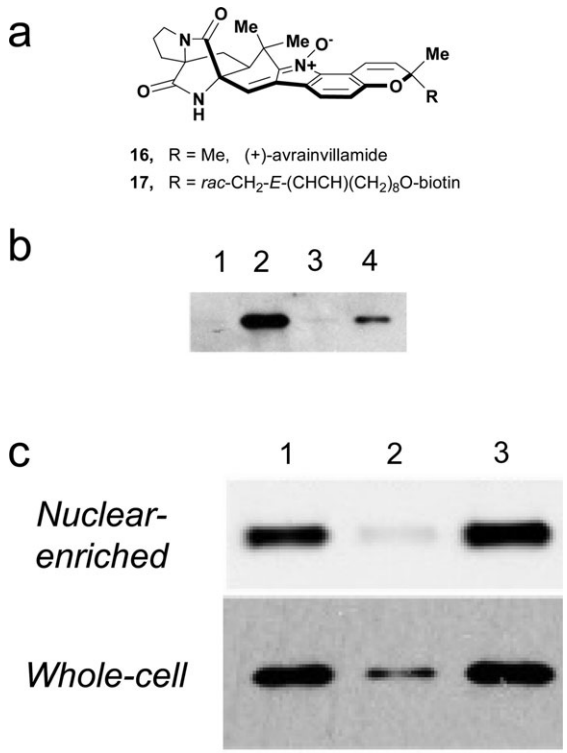


Fig. 9 Nucleophosmin is a cellular target of (+)-avrainvillamide. (a) Avrainvillamide (**16**) and its biotinylated derivative (**17**). (b) Western blot detection of nucleophosmin after addition of probes to live T-47D cells followed by lysis and affinity purification *via* streptavidin resin. Nucleophosmin is detected in the presence (lane 2), but not absence (lane 1) of 3 μM **17**. Attempted affinity purification from cells treated with 3 μM **16** plus 3 μM biotin 9-decenyl ester did not isolate nucleophosmin (lane 3). The complex molecular architecture of **16** aids in affinity purification, as the use of a structurally simpler and less potent analog (9 μM, not shown) is a less efficient affinity reagent for nucleophosmin (lane 4). (c) Western blot detection of nucleophosmin after affinity purification from lysates. A streptavidin affinity matrix made with **17** pulls down nucleophosmin from nuclear-enriched and whole-cell lysates of T-47D cells (lane 1). Competitive addition of **16** prevents affinity isolation (lane 2), whereas *ent*-**16** does not (lane 3). Image reproduced with permission from Wulff *et al.*[23] Copyright 2007 American Chemical Society.

lysates, showing by Western blotting that nucleophosmin binding to **17** was specific, as excess **16**, but not inactive *ent*-**16**, prevented nucleophosmin binding and isolation (Fig. 9C).

There are several other examples in the literature detailing the discovery of protein binding partners of biologically active small molecules by direct addition of probes to cells in culture. Target identification of the natural products fumagillin,[21] parthenolide,[24] and epoxomicin[25] were all carried out through incubation of biotinylated derivatives with human cell cultures, followed by lysis and purification of protein binding partners. However, in each of these cases, target purification was facilitated by the fact that the probe also covalently modified its target protein. Incubation of biotinylated derivatives with live cells in culture represents an important strategy for affinity purification of binding partners of covalent protein inhibitors. Interestingly, this strategy (without the use of photocrosslinking) has not been used to identify the protein

 This journal is © The Royal Society of Chemistry 2008

targets of noncovalent binders. It is possible that such a strategy would not be successful, as the conditions used for cell lysis could disrupt noncovalent interaction between the target protein and the probe.

5 Target identification strategies employing radiolabeled or fluorescent probes

In 2003, Cravatt and Sorensen *et al.* reported a "tag-free" identification of the target protein of (−)-FR182877, a natural product of interest for both its synthetic challenges and antitumor properties (**18**, Fig. 10A).[26] Using a technique pioneered in the field of proteomic profiling,[33,42] attachment of an azide generated tagged probe **19**. The azide functional group allowed fluorescent reporter tags to be appended using the ubiquitous Cu(I)-catalyzed variant of the azide–alkyne cycloaddition, avoiding side reactions observed when attempting to functionalize the scaffold with nucleophilic reagents. A sample of the mouse-heart proteome was incubated with rhodamine-tagged probe **20** and then resolved by SDS-PAGE; this strategy revealed that a 70 kDa protein was specifically labeled (Fig. 10B). The authors found that heat-denaturation of the proteome sample prior to incubation with the probe abolished labeling, suggesting that labeling only occurred when the probe was bound to the target protein in its native state. To identify the 70 kDa protein, probe **21** was used to purify the labeled protein from the proteomic mixture using avidin chromatography. This purified protein sample was resolved *via* SDS-PAGE and the rhodamine-tagged 70 kDa band was excised from the gel and analyzed by mass spectrometry to identify carboxylesterase-1 as a protein target of (−)-FR182877.

A similar strategy was employed by Crews and co-workers in the identification of the protein target of triptolide, a diterpene component of the Chinese medicinal herb *lei gong teng* (**22**, Fig. 11A).[43,44] A sample of the natural product was tritiated by a contract service and incubated with HeLa cells.

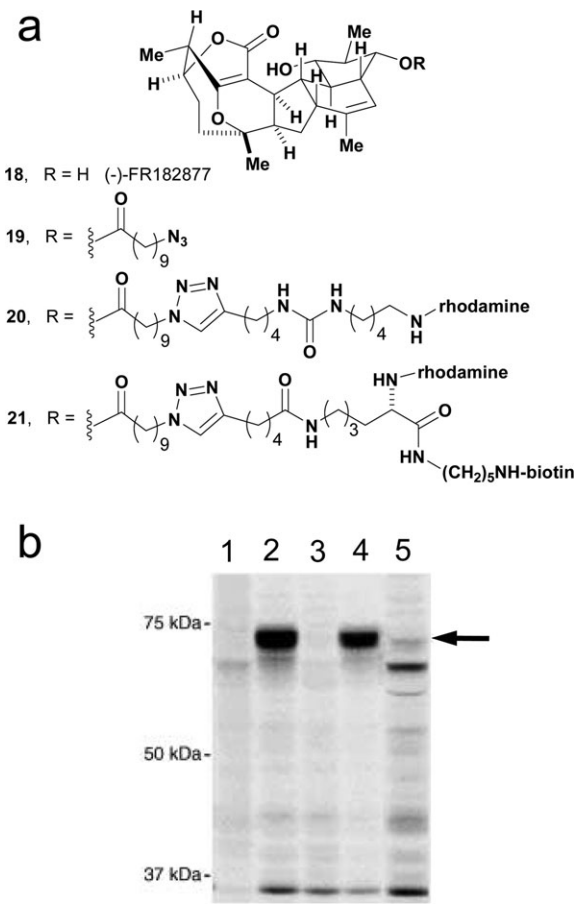


Fig. 10 Carboxylesterase-1 is a cellular target of (−)-FR182877. (a) (−)-FR182877 (**18**), azide-tagged analog **19** and fluorescent- and affinity-tagged derivatives (**20** and **21**). (b) Probe **20** efficiently labels carboxylesterase-1 under native conditions (lane 2), but not in a heat-treated proteome sample (lane 1). Labeling is prevented by competitive addition of **18** (lane 3), but not *ent*-**18** (lane 4). Labeling not seen with a rhodamine-tagged probe constructed from *ent*-**19** (lane 5, structure not shown). Image reproduced with permission from Adam *et al.*[26] Copyright 2003 Wiley-VCH Verlag GmbH & Co. KGaA.

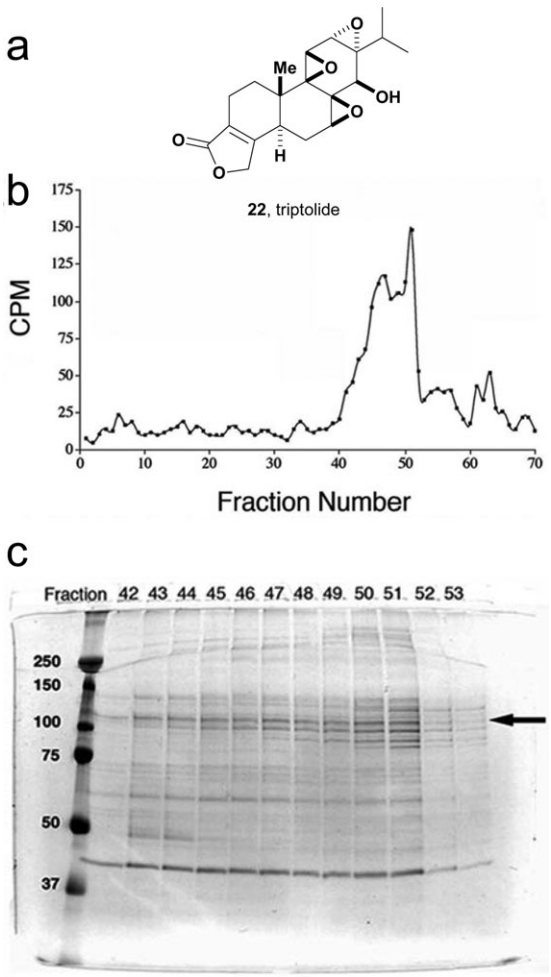


Fig. 11 PC2 is a cellular target of triptolide. (a) The diterpene triptolide (**22**). (b) [^{3}H]-Triptolide was incubated with HeLa cells and cell lysates were fractionated by anion exchange chromatography. The site of [^{3}H] incorporation was not determined. (c) Radioactive fractions from (b) were subjected to PAGE and stained with Coomassie Blue. PC2 (arrow) staining intensity correlates with the radioactivity of anion exchange fractions. Image reproduced with permission from Leuenroth *et al.*[44]

After cellular lysis, the proteome was fractionated *via* FPLC over an anion-exchange column and radioactive fractions were analyzed by SDS-PAGE and stained with Coomassie Blue (Fig. 11B and C). The intensity of Coomassie staining of a band at 110 kDa was proportional to the radioactive intensity of the fractions, indicating that [^{3}H]-triptolide bound specifically to this protein, which mass spectrometric analysis determined to be the calcium channel protein PC2.

These radiolabel/fluorescent probe target identification strategies are typically only useful when the small molecule covalently modifies its target, and thus is still linked to the protein under the denaturing conditions used for SDS-PAGE analysis.[45] One exception to this was the identification of the protein target of phorboxazole A using a fluorescent derivative of this compound in conjunction with native (non-denaturing) PAGE (Fig. 1).[46]

6 Target identification strategies employing photoaffinity probes

An ideal affinity probe should be easily characterized by routine spectroscopic techniques, and should be constructed so that it interacts with the target protein in live cells in the same way that the parent compound does. Additionally, it must retain sufficient affinity for the target protein such that it remains bound to the probe during elution of nonbinding proteins. Radiolabeled probes that covalently modify their protein targets probably come closest to this ideal.

But what if the compound of interest binds to its target in a non-covalent fashion? What if the compound has only a moderate affinity for its protein target? One of the shortcomings of the methods described in sections 3 and 4 is the difficulty in isolating less abundant or low affinity proteins. Sometimes, abundance problems can be solved simply by using more protein. Isolation of low-affinity proteins is inherently more problematic, however, because it is precisely the probe's affinity for its cellular target that is used to separate it from non-binders. It is perhaps no coincidence that the earliest examples of successful target identification *via* affinity purification involved molecules that bound abundant target proteins irreversibly, or with very high affinity.

One must not forget that the ultimate goal of a target ID experiment is not the discovery of the tightest-binding molecular target of a particular compound *per se*, but rather the elucidation of the protein whose inhibition or activation leads to the biologically interesting phenotype. Also, the knowledge of *all* proteins which bind to a compound can be important in predicting the off-target effects of a drug, as illustrated recently by Peters and Gray in a method to assay the binding partners of kinase inhibitors.[47]

In order to prevent the dissociation of weak binders during the course of an affinity purification experiment, crosslinking reagents can be used to irreversibly attach a small molecule to its target protein. The presence of chemically reactive electrophilic groups such as epoxides contributes to the exquisite potency of several classes of natural products, as these compounds often bind irreversibly to the target of interest. Taunton and co-workers' recent report of the structural bioinformatics-based design of human protein kinase inhibitors showed that targeting nonconserved nucleophilic functionality within the highly conserved human protein kinase active site with an electrophile produced highly potent and selective kinase inhibitors.[48] Use of this strategy in target identification, however, is not feasible, as proper placement of the reactive group such that it is in position to react with a nucleophilic residue on the target protein is unlikely when the target is unknown, and promiscuous reactivity of the electrophilic moiety with cellular nucleophiles would complicate analysis.

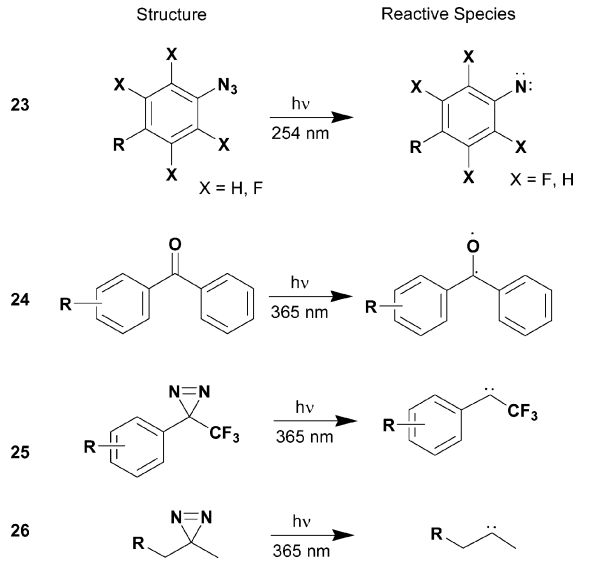


Fig. 12 Photoreactive groups commonly employed in photocrosslinking experiments. Aryl azides (**23**) are excited initially to singlet nitrenes using potentially damaging short-wave UV light. Irradiation of benzophenones (**24**) or diazirines (**25**, **26**) with longwave UV light sources forms radicals or carbenes, respectively.

To better accomplish this task, a compound can be functionalized with a *photoreactive* functional group.[49–52] In these experiments, the compound (containing a photoreactive group) is incubated with a lysate or whole cells to allow the probe to bind to its target protein. In the absence of UV light, the photoreactive group is stable, allowing the unactivated probe to come into binding contact with its target. Irradiation of the photoprobe at a specific wavelength will generate a reactive species that rapidly reacts with a variety of amino acid residues, forming a covalent linkage between the probe and its target protein. The three most common photophores used in target identification studies are aryl azides (**23**), benzophenones (**24**), and diazirines (**25**, **26**, Fig. 12).[49–52]

6.1 Aryl azides as photoprobes for target ID experiments

Although its crosslinking efficiency is generally regarded as poor (especially when lacking electron-withdrawing functionality on the ring), the azido functionality has been a popular choice due to its small size and ease of introduction. Jessen *et al.* recently reported the molecular target of a series of 3,5-diaryl-1,2,4-oxadiazoles that induce apoptosis and selectively inhibit tumor growth *in vivo* using this strategy (Fig. 13A).[53] SAR studies based on compound **27** indicated that functionality at the 4-chloro-position, required for activity, could be

 This journal is © The Royal Society of Chemistry 2008

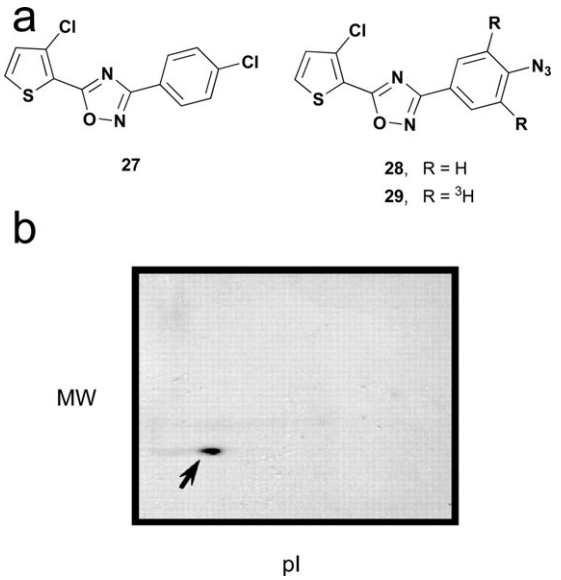


Fig. 13 TIP47 is a cellular target of antitumor 3,5-diaryl-1,2,4-oxadiazoles. (a) Active compounds **27** and **28** and tritiated photoprobe **29**. (b) Autoradiograph of a two-dimensional gel shows radiolabeling of a single target protein, TIP47, at MW 50 kDa and pI 5.3 (arrow). Image reproduced with permission from Jessen *et al.*[53] Copyright 2005 American Association for Cancer Research.

replaced with an azide photophore with no loss in activity (**28**). The 3,5-ditritium variant (**29**) of this compound was incubated with lysate from T-47D human breast cancer cells and irradiated with shortwave UV light. Two-dimensional electrophoretic separation of the photocrosslinked lysate followed by autoradiographic development revealed a single spot at MW 50 kDa and pI 5.3 (Fig. 13B). A duplicate gel of lysates not incubated with the tritiated probe was prepared and the analogous spot excised from the gel and analyzed by LCMS/MS. Twelve peptides matching tail-inducing protein 47 (TIP47), an insulin-like growth factor-II receptor binding protein, were identified by mass spectrometry, and thus this protein is believed to be the primary target leading to apoptotic induction by compound **27**.

6.2 Benzophenone and diazirines as photoprobes for target ID experiments

Fujii and co-workers used biotinylated photocrosslinking probes **31–33** to discover receptors for the leaf-movement factor potassium lespedezate (**30**), a ligand whose concentration governs circadian rhythmic leaf-movement in leguminous plants (Fig. 14A).[27] Three different probes were synthesized in order to examine the relationship between bioactivity and photocrosslinking efficiency of the probes, as systematic studies of such relationships are lacking. Probes were incubated with plasma membrane proteins from motor cells in plant leaves and irradiated with longwave UV light for 10 min to crosslink receptor proteins. After gel electrophoretic separation of the lysates, immunoblotting for probe-coupled (biotinylated) proteins identified a 210 kDa protein as a potential receptor protein. Probe **31**, which positions the photoactive group farthest away from the pharmacophore for increased bioactivity, showed no specific crosslinking (Fig. 14B, lanes 3

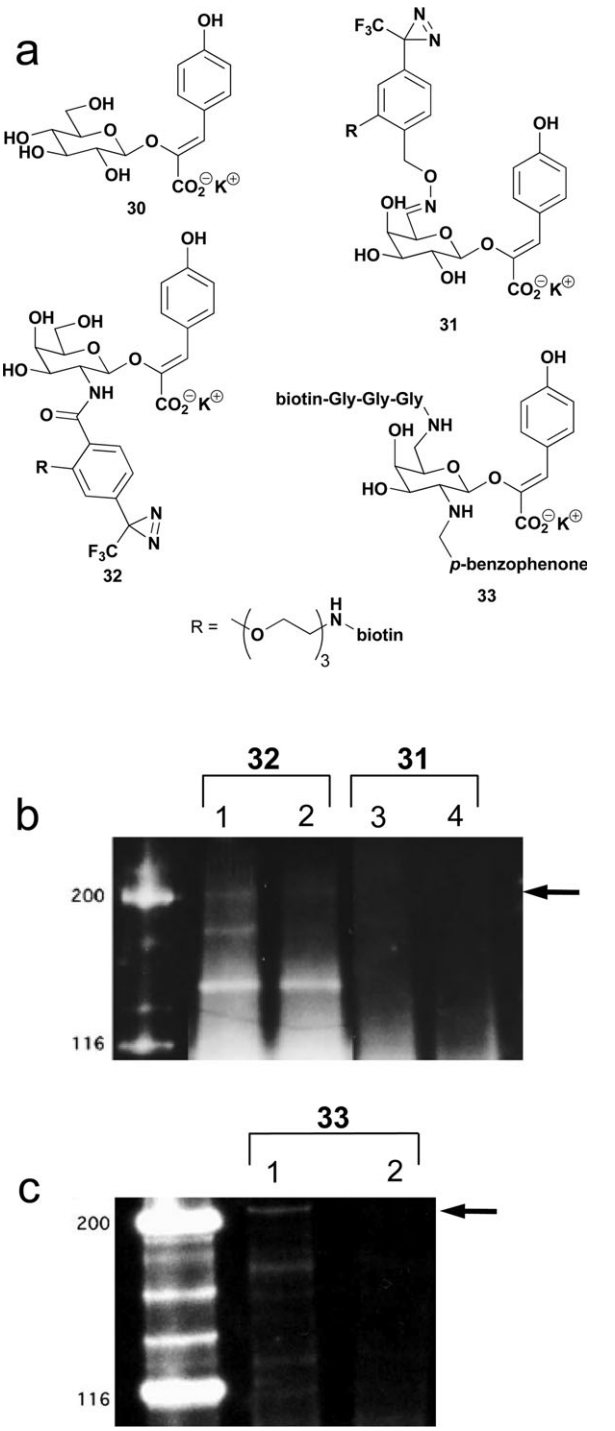


Fig. 14 Detection of receptor proteins of the leaf-movement factor potassium lespedezate. (a) Potassium lespedezate (**30**) and photoprobes (**31–33**). (b) Probe **32** incubated with membrane proteins of motor cells with (lane 2), and without excess **30** added (lane 1). Probe **31** incubated with membrane proteins of motor cells with (lane 4), and without excess **30** added (lane 3). (c) Probe **33** incubated with membrane proteins of motor cells with (lane 2), and without additional **30** added (lane 1). Western blots imaged by chemiluminescent detection of biotin. The authors report that probe **33** most efficiently labels the unknown 210 kDa receptor protein. Image reproduced with permission from Fujii *et al.*[27] Copyright 2005 Elsevier Ltd.

and 4). Diazirine **32** (Fig. 14B, lanes 1 and 2) and benzophenone **33** (Fig. 14C) both label an unknown 210 kDa protein, which is a potential receptor for **30**. The authors report that **33**, the least bioactive of the three photoprobes synthesized, produced the most efficient photolabeling, which they attribute to the proximity of the benzophenone photophore to the presumed binding pocket of **30**. However, in the absence of a binding model, it is unclear whether the difference in crosslinking efficiency between probes **32** and **33** is due to the proximity of these two groups to the binding pocket of the target, or difference in the photoreactivity of the photophores themselves. This highlights an important point, namely that there may be a fine balance between bioactivity and crosslinking efficiency, and that photophore selection may have a significant effect on crosslinking efficiency.

6.3 Photocrosslinking, followed by detection *via* click chemistry

The examples described in sections 6.1 and 6.2 rely on the compound of interest to have some sort of "tag" on it (*i.e.*, a radiolabel or biotin) that will allow its subsequent detection in electrophoresis experiments. As discussed above, introduction of biotin tags is often undesirable from a logistical standpoint, as large tags can decrease solubility and activity of a target ligand, and few chemistry labs are equipped to work with radioactive isotopes. Taunton and co-workers recently detailed the elucidation of the cellular target of HUN-7293 (**34**), a fungal cyclodepsipeptide known to block cotranslational translocation of vascular cell adhesion molecule (VCAM) and other proteins into the ER, using multifunctional photoprobe **35** (Fig. 15A).[54] They introduced an alkyne on a derivative of HUN-7293 as a bioorthogonal chemical handle through which they could rapidly tag their compound *after* photocrosslinking.

In order to maximize the efficiency of photocrosslinking and thus target identification, Taunton and co-workers substituted a leucine residue in the molecule with the diazirine photophore photo-leucine, and incorporated the alkyne tag. By making conservative structural modifications, **35** retains the potency of the parent compound, and places the photo-excitable group inside the scaffold of the natural product. Irradiation of a mixture of probe **35** with ER microsomal fractions, followed by azide–alkyne cycloaddition to a rhodamine–azide reporter selectively labeled Sec61, a 50 kDa ER membrane channel protein previously suspected to be a possible target (Fig. 15B). This interaction was competitively abolished by addition of **36**, a non-photoactivatable analog of **35**.

Several convenient multifunctional probe scaffolds are readily available from commercial sources and have been reported (Fig. 16).[33,36,55–59] Lysine, with its three functional groups, is available in many protecting group combinations and is a popular scaffold. Yao and co-workers have recently reported lysine-based modular probes of type **37** in their synthesis of affinity-based probes for profiling of aspartic proteases.[55,56] The unnatural amino acid 4-benzoylphenylalanine (Bpa, **38**) was used by Sewald and co-workers as a core scaffold in the synthesis of multifunctional probes used in one of the first reports of affinity-based tagging of protein families for proteomics using reversible inhibitors in combination with photocrosslinkers.[57] Bifunctional probes of type **39** are also popular.[58] They were used recently by Cremo and co-workers in crosslinking studies to generate a structural model of the regulatory domains of smooth muscle heavy meromyosin.[59] Based on these scaffolds, it is easy to imagine polyfunctional probes such as **40** and **41** where a large combination of purification, imaging, and bioorthogonally reactive tags may be appended as desired.

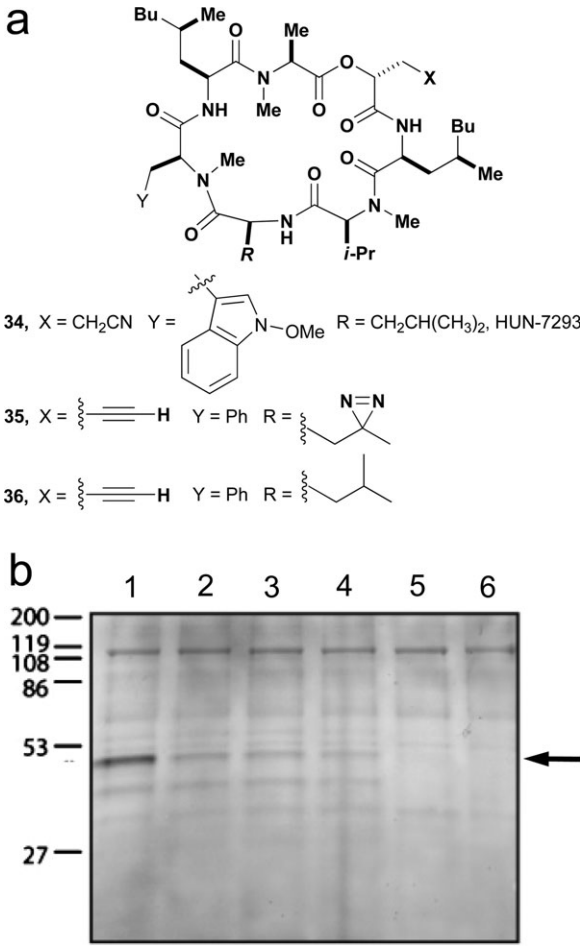


Fig. 15 Sec61α is a molecular target of HUN-7293. (a) HUN-7293 (**34**), its modular photoprobe (**35**), and a non-photoactivatable competitor (**36**). (b) In-gel fluorescence analysis of rhodamine-tagged photocrosslinked proteins. Sec61α crosslinking (lane 1, arrow) decreases with increasing amounts of **36** (lanes 2–4). Labeling was not observed in the absence of either UV irradiation (lane 5) or probe **35** (lane 6). Image reproduced with permission from MacKinnon *et al.*[54] Copyright 2007 American Chemical Society.

The relationship between photoprobe affinity and labeling efficiency was analyzed in depth by Salisbury and Cravatt in a recent paper detailing the optimization of "clickable" activity-based probes for the proteomic profiling of HDAC complexes.[60] No strong correlation was observed between photolabeling efficiency and inhibitory activity (and thus binding strength) of several potent probes. Their findings highlight the importance of the location and orientation of the crosslinking agent inside chiral binding pockets, as enantiomeric probes containing either L- or D-Bpa displayed disparate labeling

 This journal is © The Royal Society of Chemistry 2008

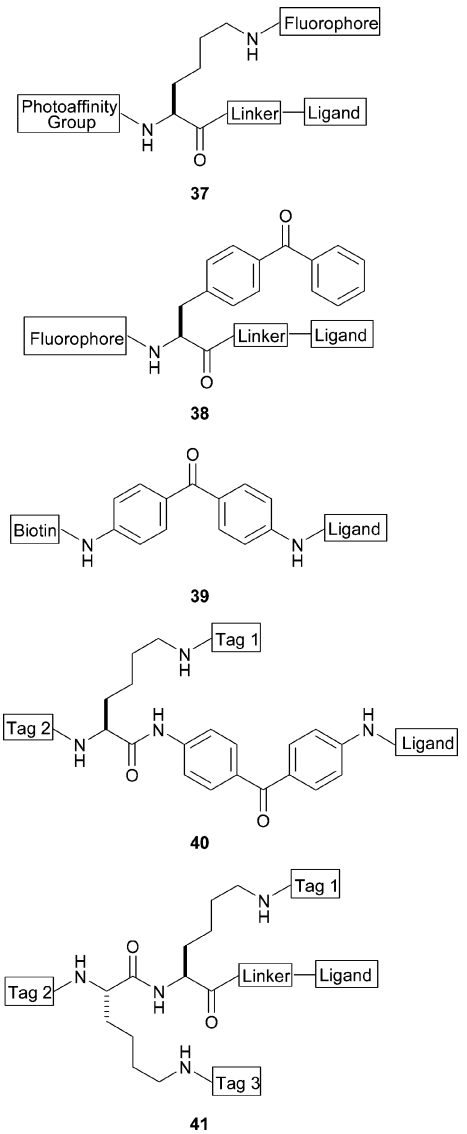


Fig. 16 Generalized scaffold designs for multifunctional affinity probes.

efficiencies. Furthermore, variation of the proximity of the crosslinking moiety to the inhibitor binding pocket was useful in this case, as longer linkers were also able to label known HDAC-associated proteins. Most interestingly, for one probe, SAHA-BPyne, the authors report superior labeling, both in terms of decreased background *and* increased sensitivity, in live cells (*in situ*) as opposed to *in vitro* photolabeling of lysates. This highlights the potential advantage of labeling putative target proteins inside the dynamic environment of the cell, as opposed to a simplified proteomic mixture.

7 Target identification and validation

In all cases discussed above, the affinity of an analog/derivative of a biologically active compound was used to selectively purify a protein from a complex mixture or tag it with a reporter such as a fluorophore or a radiolabel. Modern peptide mass fingerprinting has emerged as the primary means to identify target proteins from such experiments.[61] Typically the enzyme trypsin is used to digest isolated proteins, either in solution with affinity resins, or directly in SDS-PAGE gel slices after electrophoresis. Tryptic peptide fragments are separated *via* HPLC and directly injected into a mass spectrometer with an analyzer configuration. Masses of parent ions are recorded, and these peptides are sent to a collision chamber where they are degraded further. Mass spectral analysis of the resultant fragments is used to computationally piece together the amino acid sequence of the original tryptic peptide fragment. Comparison of these sequences to peptide sequence databases gives possible protein assignments with a confidence based on the quality of the mass spectral data. Today there are several commercially available MALDI and ESI mass spectrometers and software packages with this capability, and any large institution with a mass spectrometry facility should be capable of acquiring this type of data.

Once a putative target protein has been identified, several follow up experiments are possible to confirm that the compound of interest binds to the protein. The simplest experiments involve *in vitro* analysis of the purified protein with the compound. Direct binding can be analyzed by fluorescence anisotropy, isothermal titration calorimetry, or surface plasmon resonance. If the target is an enzyme, the compound or its analogs may be tested as potential inhibitors. Additionally, photoactivatable probes, when incubated with purified samples of target protein, can help map binding sites in the protein through photocrosslinking followed by peptide mass fingerprinting. Knowledge of a binding site enables site-directed mutagenesis studies to be performed, as deletion of key binding residues should alter affinity of the compound for the protein, and prevent affinity isolation.[23]

The most common cell culture validation experiments are designed to modulate the levels of the putative target protein. RNAi techniques capable of "knocking down" the expression level of the putative target protein can be used to confirm if the absence of the protein results in the same phenotype induced by chemical treatment with the compound of interest. In cases where a "knockout" cell line exists for the target protein, treatment with the compound should have no effect, as there should be no target for the compound to interact with. Conversely, artificially increasing the levels of the putative target protein by overexpression of the target from a transfected plasmid should in many cases diminish the cellular effect of the compound. Finally, analysis of the subcellular localization of a fluorescent version of the compound can be performed with confocal microscopy; if there is an antibody available for the putative protein target, cellular colocalization studies can be performed.

8 Conclusions and outlook

Although new technology-driven discovery platforms are emerging, the use of affinity-based probes in the identification of the molecular targets of bioactive small molecules is still as relevant in modern biomedical research as it was 20 years ago. Although difficult, the discovery of the molecular targets of biologically active compounds can unveil entirely new proteins for targeted drug discovery efforts, in addition to revealing basic information about the role of identified proteins in the cell.

Acknowledgements

We are grateful to the American Cancer Society (RSG-07-026-01-CDD) for supporting our work in the area of target identification of anti-cancer compounds. BJL thanks Eli Lilly and the ACS Division of Medicinal Chemistry for a pre-doctoral fellowship.

References

1. J. P. Overington, B. Al-Lazikani and A. L. Hopkins, *Nat. Rev. Drug Discovery*, 2006, **5**, 993–996.
2. M. Jürgens and M. Schrader, *Curr. Opin. Mol. Ther.*, 2002, **4**, 236–241.
3. T. Nühse, K. Yu and A. Salomon, *Curr. Protoc. Mol. Biol.*, 2007, 18.13.11–18.13.23.
4. G. D. Prestwich, *Chem. Biol.*, 2004, **11**, 619–637.
5. Y. Mechref and M. V. Novotny, *Chem. Rev.*, 2002, **102**, 321–370.
6. R. Johansson, L. C. Gunnarsson, M. Ohlin and S. Ohlson, *J. Mol. Recognit.*, 2006, **19**, 275–281.
7. G. C. Terstappen, C. Schlüpen, R. Raggiaschi and G. Gaviraghi, *Nat. Rev. Drug Discovery*, 2007, **6**, 891–903.
8. L. Burdine and T. Kodadek, *Chem. Biol.*, 2004, **11**, 593–597.
9. S. Wang, T. B. Sim, Y.-S. Kim and Y.-T. Chang, *Curr. Opin. Chem. Biol.*, 2004, **8**, 371–377.
10. J. Taunton, C. A. Hassig and S. L. Schreiber, *Science*, 1996, **272**, 408–411.
11. S. Bach, M. Knockaert, J. Reinhardt, O. Lozach, S. Schmitt, B. Baratte, M. Koken, S. P. Coburn, L. Tang, T. Jiang, D.-c. Liang, H. Galons, J.-F. Dierick, L. A. Pinna, F. Meggio, F. Totzke, C. Schächtele, A. S. Lerman, A. Carnero, Y. Wan, N. Gray and L. Meijer, *J. Biol. Chem.*, 2005, **280**, 31208–31219.
12. M. W. Harding, A. Galat, D. E. Uehling and S. L. Schreiber, *Nature*, 1989, **341**, 758–760.
13. K. Kuramochi, T. Haruyama, R. Takeguchi, T. Sunoki, M. Watanabe, M. Oshige, S. Kobayashi, K. Sakaguchi and F. Sugawara, *Bioconjugate Chem.*, 2005, **16**, 97–104.
14. D.-W. Jung, D. Williams, S. M. Khersonsky, T.-W. Kang, N. Heidary, Y.-T. Chang and S. J. Orlow, *Mol. BioSyst.*, 2005, **1**, 85–92.
15. J. R. Snyder, A. Hall, L. Ni-Komatsu, S. M. Khersonsky, Y.-T. Chang and S. J. Orlow, *Chem. Biol.*, 2005, **12**, 477–484.
16. S. Ding, T. Y. H. Wu, A. Brinker, E. C. Peters, W. Hur, N. S. Gray and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 7632–7637.
17. C. Xing, J. R. LaPorte, J. K. Barbay and A. G. Myers, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 5862–5866.
18. 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 and Y. T. Chang, *J. Am. Chem. Soc.*, 2003, **125**, 11804–11805.
19. Q. Zhang, M. B. Major, S. Takanashi, N. D. Camp, N. Nishiya, E. C. Peters, M. H. Ginsberg, P. G. Schultz, R. T. Moon and S. Ding, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 7444–7448.
20. N. Kanoh, K. Honda, S. Simizu, M. Muroi and H. Osada, *Angew. Chem., Int. Ed.*, 2005, **44**, 3559–3562.
21. N. Sin, L. Meng, M. Q. W. Wang, J. J. Wen, W. G. Bornmann and C. M. Crews, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 6099–6103.
22. R. R. Falsey, M. T. Marron, G. M. K. B. Gunaherath, N. Shirahatti, D. Mahadevan, A. A. L. Gunatilaka and L. Whitesell, *Nat. Chem. Biol.*, 2006, **2**, 33–38.
23. J. E. Wulff, R. Siegrist and A. G. Myers, *J. Am. Chem. Soc.*, 2007, **129**, 14444–14451.
24. B. H. B. Kwok, B. Koh, M. I. Ndubuisi, M. Elofsson and C. M. Crews, *Chem. Biol.*, 2001, **8**, 759–766.
25. L. Meng, R. Mohan, B. H. B. Kwok, M. Elofsson, N. Sin and C. M. Crews, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 10403–10408.
26. G. C. Adam, C. D. Vanderwal, E. J. Sorensen and B. F. Cravatt, *Angew. Chem., Int. Ed.*, 2003, **42**, 5480–5484.
27. T. Fujii, Y. Manabe, T. Sugimoto and M. Ueda, *Tetrahedron*, 2005, **61**, 7874–7893.
28. A. V. Statsuk, R. Bai, J. L. Baryza, V. A. Verma, E. Hamel, P. A. Wender and S. A. Kozmin, *Nat. Chem. Biol.*, 2005, **1**, 383–388.
29. K. H. Emami, C. Nguyen, H. Ma, D. H. Kim, K. W. Jeong, M. Eguchi, R. T. Moon, J.-L. Teo, S. W. Oh, H. Y. Kim, S. H. Moon, J. R. Ha and M. Kahn, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12682–12687.
30. D. J. Maly, I. C. Choong and J. A. Ellman, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 2419–2424.
31. M. M. Reddy, K. Bachhawat-Sikder and T. Kodadek, *Chem. Biol.*, 2004, **11**, 1127–1137.
32. S. I. Sato, Y. Kwon, S. Kamisuki, N. Srivastava, Q. Mao, Y. Kawazoe and M. Uesugi, *J. Am. Chem. Soc.*, 2007, **129**, 873–880.
33. E. Weerapana, A. E. Speers and B. F. Cravatt, *Nat. Protoc.*, 2007, **2**, 1414–1425.
34. G. Wang, L. Shang, A. W. G. Burgett, P. G. Harran and X. Wang, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 2068–2073.
35. K. D. Shin, M.-Y. Lee, D.-S. Shin, S. Lee, K.-H. Son, S. Koh, Y.-K. Paik, B.-M. Kwon and D. C. Han, *J. Biol. Chem.*, 2005, **280**, 41439–41448.
36. T. Kan, Y. Kita, Y. Morohashi, Y. Tominari, S. Hosoda, T. Tomita, H. Natsugari, T. Iwatsubo and T. Fukuyama, *Org. Lett.*, 2007, **9**, 2055–2058.
37. J. M. Nyangulu, M. M. Galka, A. Jadhav, Y. Gai, C. M. Graham, K. M. Nelson, A. J. Cutler, D. C. Taylor, G. M. Banowetz and S. R. Abrams, *J. Am. Chem. Soc.*, 2005, **127**, 1662–1664.
38. R. C. Clevenger, J. M. Raible, A. M. Peck and B. S. J. Blagg, *J. Org. Chem.*, 2004, **69**, 4375–4380.
39. N. Kitahata, T. Nakano, K. Kuchitsu, S. Yoshida and T. Asami, *Bioorg. Med. Chem.*, 2005, **13**, 3351–3358.
40. A. Reizelman, S. C. Wigchert, C. del-Bianco and B. Zwanenburg, *Org. Biomol. Chem.*, 2003, **1**, 950–959.
41. P. Bargagna-Mohan, A. Hamza, Y.-e. Kim, Y. Khuan Ho, N. Mor-Vaknin, N. Wendschlag, J. Liu, R. M. Evans, D. M. Markovitz, C.-G. Zhan, K. B. Kim and R. Mohan, *Chem. Biol.*, 2007, **14**, 623–634.
42. A. E. Speers and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 535–546.
43. S. J. Leuenroth and C. M. Crews, *Chem. Biol.*, 2005, **12**, 1259–1268.
44. S. J. Leuenroth, D. Okuhara, J. D. Shotwell, G. S. Markowitz, Z. Yu, S. Somlo and C. M. Crews, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 4389–4394.
45. S. Zhu, K. T. Mc Henry, W. S. Lane and G. Fenteany, *Chem. Biol.*, 2005, **12**, 981–991.
46. C. J. Forsyth, L. Ying, J. Chen and J. J. LaClair, *J. Am. Chem. Soc.*, 2006, **128**, 3858–3859.
47. E. C. Peters and N. S. Gray, *ACS Chem. Biol.*, 2007, **2**, 661–664.
48. M. S. Cohen, C. Zhang, K. M. Shokat and J. Taunton, *Science*, 2005, **308**, 1318–1321.
49. G. Dormán and G. D. Prestwich, *Trends Biotechnol.*, 2000, **18**, 64–77.
50. G. Dormán and G. D. Prestwich, *Biochemistry*, 1994, **33**, 5661–5673.
51. S. A. Fleming, *Tetrahedron*, 1995, **51**, 12479–12520.
52. F. Kotzyba-Hibert, I. Kapfer and M. Goeldner, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 1296–1312.
53. K. A. Jessen, N. M. English, J. Y. Wang, S. Maliartchouk, S. P. Archer, L. Qiu, R. Brand, J. Kuemmerle, H. Zhang, K. Gehlsen, J. Drewe, B. Tseng, S. X. Cai and S. R. Kasibhatla, *Mol. Cancer Ther.*, 2005, **4**, 761–771.
54. A. L. MacKinnon, J. L. Garrison, R. S. Hegde and J. Taunton, *J. Am. Chem. Soc.*, 2007, **129**, 14560–14561.
55. S. Chattopadhaya, E. W. S. Chan and S. Q. Yao, *Tetrahedron Lett.*, 2005, **46**, 4053–4056.
56. E. W. S. Chan, S. Chattopadhaya, R. C. Panicker, X. Y. Huang and S. Q. Yao, *J. Am. Chem. Soc.*, 2004, **126**, 14435–14446.
57. M. C. Hagenstein, J. H. Mussgnug, K. Lotte, R. Plessow, A. Brockhinke, O. Kruse and N. Sewald, *Angew. Chem., Int. Ed.*, 2003, **42**, 5635–5638.
58. C. M. Salisbury and B. F. Cravatt, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1171–1176.
59. J. L. Wahlstrom, M. A. Randall, Jr., J. D. Lawson, D. E. Lyons, W. F. Siems, G. J. Crouch, R. Barr, K. C. Facemyer and C. R. Cremo, *J. Biol. Chem.*, 2003, **278**, 5123–5131.
60. C. M. Salisbury and B. F. Cravatt, *J. Am. Chem. Soc.*, 2008, **130**, 2184–2194.
61. W. J. Henzel, C. Watanabe and J. T. Stults, *J. Am. Soc. Mass Spectrom.*, 2003, **14**, 931–942.

 This journal is © The Royal Society of Chemistry 2008