Phytochem Rev (2013) 12:895–914
DOI 10.1007/s11101-013-9312-6

Chemical genetics strategies for identification of molecular targets

José Manuel Botubol Ares · M[a] Jesús Durán-Peña Rosario Hernández-Galán · Isidro G. Collado

Received: 15 April 2013 / Accepted: 20 July 2013 / Published online: 9 August 2013

Abstract Chemical genetics is an emerging field that can be used to study the interactions of chemical compounds, including natural products, with proteins. Usually, the identification of molecular targets is the starting point for studying a drug's mechanism of action and this has been a crucial step in understanding many biological processes. While a great variety of target identification methods have been developed over the last several years, there are still many bioactive compounds whose target proteins have not yet been revealed because no routine protocols can be adopted. This review contains information concerning the most relevant principles of chemical genetics with special emphasis on the different genomic and proteomic approaches used in forward chemical genetics to identify the molecular targets of the bioactive compounds, the advantages and disadvantages of each and a detailed list of successful examples of molecular targets identified with these approaches.

Keywords Chemical biology · Forward chemical genetics · Small molecules · Target proteins

Abbreviations

AMP	Adenosine monophosphate
BSA	Bovine serum albumin
FITC	Fluorescein isothiocyanate
FKBP12	FK506-binding protein
HRP	Horseradish peroxidase
PEG	Polyethyleneglycol

Introduction

Despite advances made in combinatorial chemistry techniques (Lebl 1999) and high throughput screening (HTS) (Fox et al. 1999) by the pharmaceutical industry which have provided a greater number of novel chemical compounds and related biological data (Oprea 2002), the number of new drugs developed or under clinical trial has not increased proportionately (Bronson et al. 2011; Danishefsky 2010; Drews 2000; Gaudillière et al. 2001; Horrobin 2000).

Modern medicinal chemistry is focused on the study of protein-small molecule interactions and gene functionalities encoding protein synthesis (Stockwell 2004). Many of the medicinally relevant proteins have already been identified and in this connection the sequencing of the human genome (Drews 2000; Witkowski 2010) has helped scientists to identify

Electronic supplementary material The online version of this article (doi:10.1007/s11101-013-9312-6) contains supplementary material, which is available to authorized users.

J. M. B. Ares · M. J. Durán-Peña · R. Hernández-Galán · I. G. Collado (✉)
Department of Organic Chemistry, Faculty of Science, University of Cadiz, Polígono Río San Pedro s/n, 11510 Puerto Real, Cádiz, Spain
e-mail: isidro.gonzalez@uca.es

therapeutic targets and tackle the diseases facing humanity. However, elucidation of bioactive compound mechanisms of action is the major problem in chemical biology (Kwok et al. 2006) since most cellular targets of bioactive compounds have yet to be revealed. Natural products (NPs) are now increasingly being used as probes in the systematic search of inhibitors for key biochemical pathways and to delve further into the study of biological systems (Carlson 2010). The efficient identification of the chemical compounds that modulate protein functions both in vivo and in vitro is at the heart of research in chemical biology and medicinal chemistry.

Small molecules (SM) are essential as drugs in modern medicine and are valuable as probes to explore relevant biological processes. However, new approaches are needed to close the wide gap between the ability to study either single proteins or whole cellular processes. So, it is of interest to focus on studies designed to understand in more detail how SM disrupt particular signalling pathways and larger networks to yield distinct cellular phenotypes (Castoreno and Eggert 2011).

Chemical genetics

Chemical biology includes study and research of the interface between chemistry and biology (Altmann et al. 2009). In this context, chemical biology is aimed at studying the interactions of chemical compounds, including NPs, with proteins, to identify their role in biological processes (Schreiber 2003).

The above definition of chemical biology complements classical genetics where the focus is also on finding targets and signalling pathways. Chemical genetics can be defined simply as a genetic study using chemical tools (Schreiber 1998). Chemical genetics uses chemical compounds that may specifically activate or inhibit one or more target proteins (Chang 2009; Spring 2005). It offers several advantages over its classical counterpart and allows the study of unexplored biological space. For example, genes essential for survival or development cannot be studied using classical genetics; this can only be done using chemical genetics. Thus, the instantaneous effects of SM can be characterized using chemical genetics (Chang 2009; Walsh and Chang 2006). It also makes it possible to study mammals whereas classical

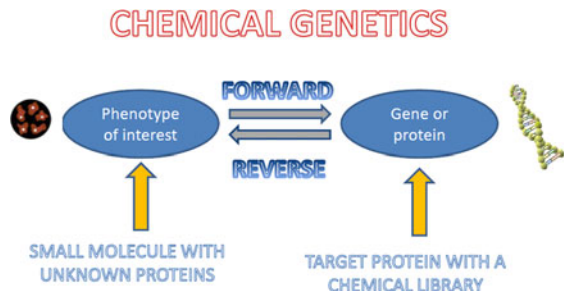


Fig. 1 Approaches in chemical genetics

genetic techniques are more complicated to apply due to their diploid genome, physical size and slow reproduction rate (Chang 2009). Other benefits of using chemical compounds are the temporal control and reversibility of the inhibition of protein function (Hübel et al. 2008).

As in classical genetics, two approaches can be taken to chemical genetics (Fig. 1) (Das et al. 2011): Forward chemical genetics (Chang 2009; Burdine and Kodadek 2004) and reverse chemical genetics (Blackwell and Zhao 2003; Neumann et al. 2003).

In forward chemical genetics, typical random mutagenesis is replaced by a screening of a library of typically not targeted SMs against multiple potential targets simultaneously (Chang 2009). Compounds that induce a phenotype of interest can be selected and then the target protein of this compound is identified. Forward chemical genetics require three components (Lokey 2003): (a) a collection or library of chemical compounds (chemical toolbox generation); (b) a biological assay with a quantifiable phenotypic output, usually performed using living cells or complex cellular extracts by means of the cyclobot method (Mendoza et al. 1999; Stockwell et al. 1999) and (c) a strategy to join an active compound to its biological target, otherwise known as target identification.

In reverse chemical genetics, a known target protein is screened using SM libraries to identify functional ligands that either stimulate or inhibit the target protein. Once a specific ligand that produces a change in the protein function is identified, it is introduced into a cell or organism and the resulting changes in the phenotype are studied (Chang 2009). Thus, chemical genetics work in reverse, i.e. from genotype to phenotype, while in forward chemical genetics the direction is from phenotype to genotype (Spring 2005).

Today, target identification is one of the greatest challenges in forward chemical genetics. Recent progress and difficulties in this field and different methods for the identification of biological targets will be discussed in depth in this review.

Approaches to the biological target identification in forward chemical genetics

Once a SM that induces a certain phenotypic response in a cell culture or in vivo experiment has been identified, the elucidation of it biological target is one of the greatest technical challenges in forward chemical genetics and phenotype-based drug discovery. The determination of these targets is particularly important in understanding the mode of action of potential SM and of various biological processes under study since a large number of drugs have multiple intracellular targets, some of which may be responsible for the undesirable side effects of the drugs. Therefore, identification of relevant targets can often lead to the discovery of more specific drugs with fewer side effects (Campillos et al. 2008). Furthermore, once a target protein is identified, its functions and cellular signalling pathways can be elucidated thus facilitating drug discovery research. However, no routine protocols can be adopted for the identification of molecular targets due to the great structural diversity of NPs.

Over the last several years a great variety of target identification methods have been developed, ranging from biochemical to genetic, which have received a number of excellent reviews (Das et al. 2011; Hübel et al. 2008; Leslie and Hergenrother 2008; Roti and Stegmaier 2012; Sleno and Emili 2008; Terstappen et al. 2007). We will classify these approaches as: affinity-based methods, genetic methods and other methods.

Affinity-based target identification methods

Traditionally, affinity-based methods play a major role in the identification of molecular targets for many biologically-active SM and NPs. These methods detect the direct binding of the SM to its target(s) and almost all successful affinity experiment have involved the combination of high affinity molecules with a high abundance of target proteins such as the isolation of FKBP12 (FK-506 binding protein-12) using the compound FK506 (Harding et al. 1989). However, it is possible to isolate a high abundance target protein with a low-affinity SM since, although interaction affinity is low, the abundance of the target proteins enable identification, a case in hand being the isolation of human GLO1 (glyoxalase 1) as a secondary target of the anti-inflammatory SM indomethacin (Sato et al. 2007).

We will classify the affinity-based target identification methods in the following groups: Matrix-based affinity methods: Affinity chromatography; Matrix-free affinity methods: Affinity by fluorescence, radioactivity, photoaffinity or immunoaffinity; Methods based on the stability of the target in response to its affinity for the drug (DARTS); Mass spectrometry-based affinity techniques and stability of proteins of rates of oxidation (SPROX).

Matrix-based affinity methods: affinity chromatography

Affinity chromatography is the oldest and most widely and successful used approach (Guiffant et al. 2007; Katayama and Oda 2007; Sakamoto et al. 2012; Sato et al. 2010). In this method a SM of interest is modified by introducing a suitable functional group (linker) through which it can be immobilized by attachment to a solid support (matrix) followed by the addition of a protein extract. Matrix-based methods should fulfil three requisites: (a) that the SM contains a derivatizable function, (b) that SM retains at least part of their activity after the derivatization and (c) that the matrix does not hinder the binding of the target protein to the SM (Lomenick et al. 2009). The latter two conditions cannot be predicted a priori and sometimes the SM cannot be provided or synthesized in sufficient amounts for the study.

First of all, it is necessary to know through a structure–activity relationship study which functional groups can be modified and used as points of attachment to the solid support through a linker. The linker type has a bearing on the success of obtaining the target proteins and is crucial for diminishing non-specific binding proteins. An optimal chemical linker should be mildly hydrophobic so as to prevent auto-aggregation and also must be long enough to prevent steric hindrance between the SM and the target protein (Fig. 2) (Sato et al. 2010).

Recent advances in this field to enhance the efficiency of affinity purification have been performed

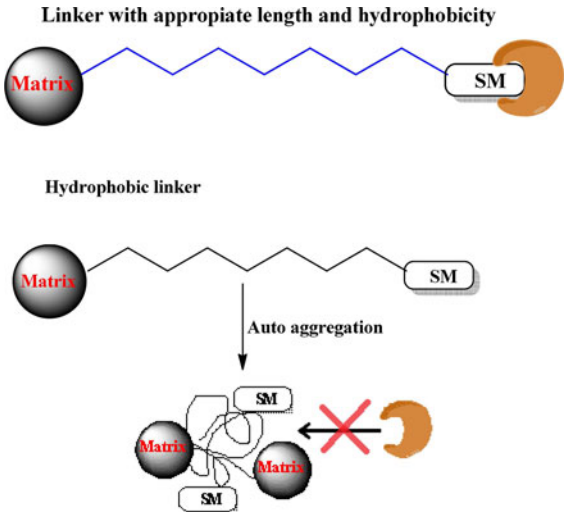


Fig. 2 Characteristics of an optimal chemical linker

and in this connection we would highlight the following linkers: (a) Hydrophilic linkers with PEG units (Bach et al. 2005; Furuya et al. 2006; Jung et al. 2005; Sato et al. 2007; Tamura et al. 2003; Zhang et al. 2007) or tartaric acid derivatives (Shiyama et al. 2004) to reduce non-specific binding proteins; (b) ACAP linkers (aminocaproylaminopentyloxy) to reduce auto-aggregation resulting from its optimal length, hydrophobicity and greater rigidity (Guiffant et al. 2007) and (c) elongation of the anterior linkers by insertion of a rigid polyproline helix for the isolation of low abundance or low affinity proteins. An example of this is the isolation of GLO1 as a new target of indomethacin (Sato et al. 2007).

SMs are immobilized to the matrix by means of complementary functional groups between the matrix and the linker. Traditionally, sugar-based affinity matrices such as agarose or Sepharose have been used in this approach, Affigel® matrices being one of the most popular affinity matrices; however, these matrices are unstable in organic solvents and also cause non-specific binding of proteins (Sakamoto et al. 2012). These problems can be solved by using resin-based affinity matrices due to their stability both in organic and aqueous media, physical and chemical stability and high ligand loading capacity. Matrices based on polymethacrylate derivatives have been developed as an alternative to Affigel® matrices, with special mention of the matrices known as Toyopearl.® Nevertheless, these matrices often show a high

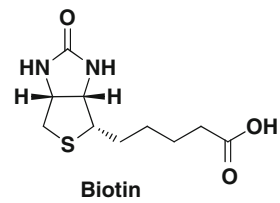


Fig. 3 Structure of biotin

number of non-specific binding proteins by comparison with Affigel® matrices (Sakamoto et al. 2012).

Alternatively, it is possible to immobilize SMs directly onto the resin using capturable molecules such as biotin (Fig. 3), which are easily trapped by avidin-agarose columns (McPherson et al. 2002). The strong non-covalent interaction between biotin and avidin results in the essentially non-reversible loading of the SM onto the resin (Hofmann and Kiso 1976; Hofmann et al. 1978). This approach has been used to identify the mitochondrial enzyme OAT (ornithine δ amino-transferase) as the target protein of the antimitotic marine diazonamide A (Wang et al. 2007).

Once the SM is immobilized on a suitable solid support it is incubated with a cell lysate or protein extract, usually by passing the extract through a column of immobilized material followed by extensive washing to remove non-specific binding proteins. Finally, the tightly bound proteins are eluted with an excess of free molecule or under strongly denaturing conditions using an electrophoresis buffer. These proteins are analysed by SDS-PAGE 1D (sodium dodecyl sulphate polyacrylamide gel electrophoresis), extracted from the gel and the protein bands are identified by mass spectrometry after partial tryptic digestion and databases are then searched for comparison of mass-sequencing of the digested peptides (Fig. 4).

On the other hand, as mentioned above, isolation and identification of target proteins by means of affinity chromatography is a difficult task because of the non-specific adsorption of background protein to the resin and linker. Therefore, several approaches have been used in order to differentiate between specific and non-specific interactions.

First of all, it is important to carry out well-designed negative control experiments. One approach is to use an inactive analogue of the SM with physical properties similar to the active SM (Sato et al. 2010; Snyder et al. 2005; Wang et al. 2007). In this regard, an

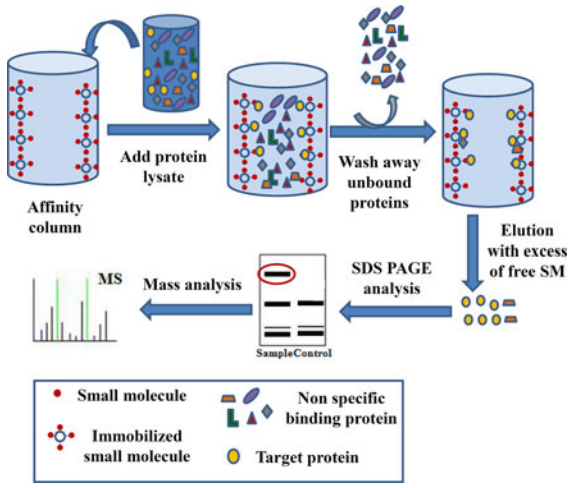


Fig. 4 Scheme of an experiment of affinity chromatography

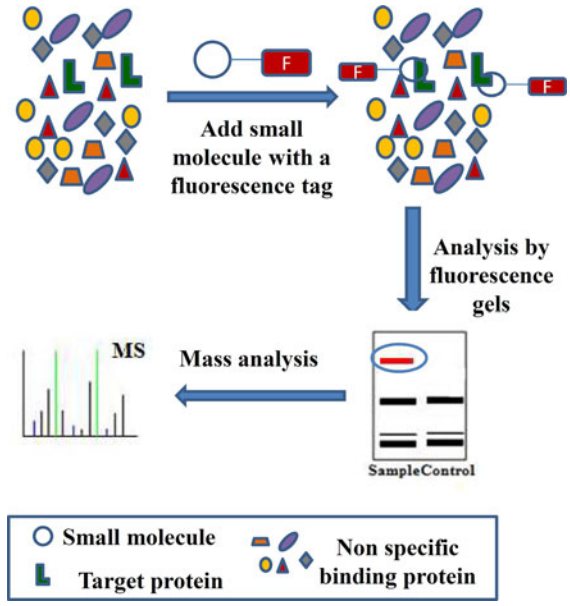


Fig. 5 Scheme of a fluorescence-based affinity experiment

inactive chiral isomer or *seco*-analogue of the active molecule would serve as an excellent negative control (Sato et al. 2010) by comparison of proteins eluted between the active molecule and the inactive analogue. An example is the identification of the target protein of diazonamide A using a *seco*-analogue of this SM (Wang et al. 2007).

Nevertheless, sometimes inactive structural analogues are not available for such study and different alternative strategies have been developed such as competitive elution experiments (Emami et al. 2004; Ito et al. 2010; Sleno and Emili 2008) and serial affinity chromatography (Yamamoto et al. 2006).

Competitive elution of bound proteins using an excess of free SM would permit selective elution under mild conditions; however, the SM's low water solubility is a hurdle that needs to be addressed in this approach. To this end, some hydrophilic organic solvents such as DMSO have been used to improve the water solubility of the SM (Sakamoto et al. 2012). One successful example of selective elution is the isolation of a cyclic AMP response element-binding protein as a molecular target of ICG-001, a SM that downregulates signalling by β-catenin/T cell factor (Emami et al. 2004).

Last, an alternate serial affinity chromatography approach has been reported where the protein extract is applied to the immobilized ligand matrix which is subsequently removed. The fresh matrix is then incubated with the same lysate (Yamamoto et al. 2006). The first matrix should capture most of the specific binding proteins due to their high affinity for the SM while the same amount of non-specific background proteins should be captured by both matrices. This approach has been used to identify the target proteins of FK506, benzenesulphonamide and methotrexate (Yamamoto et al. 2006) (see Table 1 supporting information).

A great number of small-molecules whose targets have been identified or confirmed by this approach are depicted in Table 1 of the supporting information, which highlights currently the success and widespread for this approach.

Matrix-free affinity methods

Matrix-free affinity methods rely on the incorporation of affinity tags to the SM of interest. An affinity tag can be defined as a compound that imparts an additional function to the SM to which it is attached. Like matrix-based affinity methods, the SM must be derivatizable and maintain at least part of its biological activity. The most important affinity tags are the fluorophores and photoreactive groups.

Fluorescence In this approach, the SM bound to proteins can be visualized in fluorescence gels and the fluorescence bands identified by mass spectrometry (Fig. 5). Furthermore, the introduction of a fluorophore

to a SM can reveal the subcellular location of the target proteins. To that end, appropriate lysates with an enrichment of the target protein can be used reducing the background levels in protein purification. An example is the isolation of SF3b as a molecular target of pladienolide B (Kotake et al. 2007).

One of the most frequently used and studied fluorophores is the jellyfish green fluorescent protein (GFP) (Tsien 1998; Zhang et al. 2002). However, the major drawback of GFP is its huge size (239 amino acids), sometimes diminishing the activity of the SM to which it binds.

An alternative is the use of styryl dyes (Garrett and Fattaey 1999; Sridhar et al. 2000), which span a broad range of fluorescent emission wavelengths, or the use of acylphenols as a transfer agent for immunoaffinity fluorescent tags (IAF) (Hughes et al. 2009).

Some examples of target proteins identified by fluorescence-based affinity are shown in Table 2 of the supporting information.

Photoaffinity This approach is commonly used when the SM has moderate or low affinity for its target protein(s) (Sadakane and Hatanaka 2006). In order to bind the SM irreversibly to its target protein a cross-linking reagent can be used to prevent the dissociation of weak binders (Leslie and Hergenrother 2008). In a photoaffinity experiment, a SM with a photoreactive functional tag and a reporter tag is added to a lysate or whole cell. The reporter tag is usually a radioactive isotope (MacKinnon et al. 2007) or biotin which enables the isolation and identification of target proteins. In the absence of UV light, the interaction between the SM and target protein takes place because the photoreactive group is stable while irradiation of the photoreactive group at a specific wavelength will generate a carbene that can produce the covalent attachment of the SM to the protein target (Leslie and Hergenrother 2008). Labelled protein mixtures are separated by denaturing-gel electrophoresis and detected with a phosphorimager. Specifically labelled proteins are then isolated and subjected to limited proteolytic digestion and identified by mass spectrometry (Fig. 6) (Dormán and Prestwich 2000).

A drawback of this approach is that the SM must retain biological activity after derivatization with a photoreactive group. The most frequently used photophores are benzophenones, (Dormán and Prestwich 1994) aryl azides (Fleming 1995; Kotzyba-Hibert et al. 1995) and diazirines (Brunner 1993). However, their large size can potentially interfere with SM/protein interactions. Recently, photo-leucine has been used as a photophore because of its smaller alkyl diazirine side chains (MacKinnon et al. 2007). In order to maximize the efficiency of photo cross-linking they incorporated an alkyne tag instead of biotin to perform an azide-alkyne cyclo-addition known as "click chemistry" (Cohen et al. 2007; Speers and Cravatt 2004).

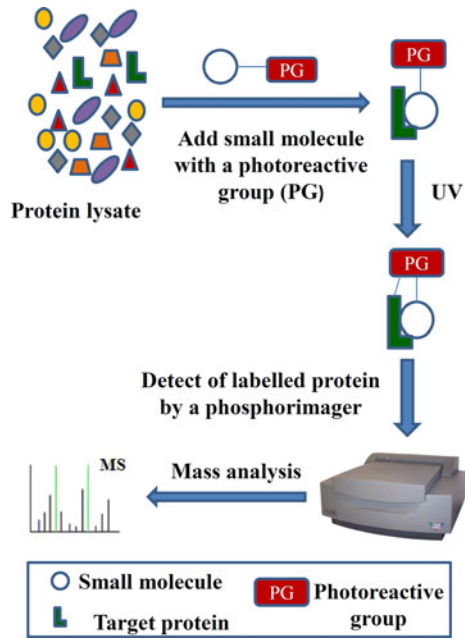


Fig. 6 Scheme of a photoaffinity experiment

Examples of successful target identification by this approach are shown in Table 3 of the supporting information.

Drug affinity responsive target stability (DARTS)

Most of affinity-based approaches described above require derivatization of the SM of interest and it is not always possible to maintain its biological activity. In order to overcome this obstacle, a target identification method known as DARTS has been recently developed where any SM library can be used without chemical derivatization or labelling (Lomenick et al. 2009, Lomenick et al. 2011; Nishiya et al. 2009).

DARTS is based on the notion that if a SM binds to its target protein, the protein-SM complex would be less susceptible to proteolysis. If the SM is not present,

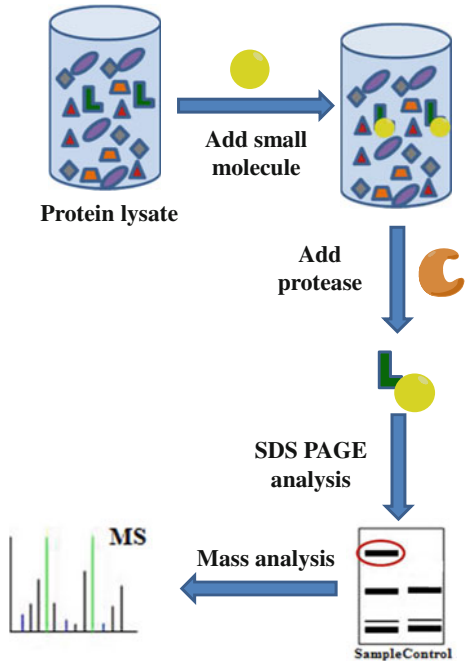


Fig. 7 General scheme of a DARTS experiment

a protease detects its recognition site(s) on the target surface initiating a proteolytic attack that leads to the complete digestion of the target protein.

In this approach, it is not necessary to know the structure of the SM, and NP extracts can be used (Lomenick et al. 2011). Furthermore, whereas other affinity-based approaches utilize positive enrichment by selectively assembling the target proteins of the SMs, DARTS uses negative enrichment by digesting non-target proteins while leaving behind the target proteins which are resistant to proteases. In a DARTS experiment, a SM incubated with a protein lysate and a control protein sample are each treated with varying amounts of protease and separated by 1D SDS-PAGE. Then, the bands whose abundance differs between the two samples are removed, treated with trypsin and analysed by mass spectrometry (Fig. 7) (Lomenick et al. 2011).

However, many target proteins can be missed in this approach because it is necessary to visualize the enrichment of the target protein in the gel (Lomenick et al. 2011). In order to solve this difficulty, different proteomic approaches have been performed to increase the sensitivity and throughput of this method, some of which can also be applied to other target identification methods. Special mention should be made of the development of 2D-PAGE (O'Farrell 1975), analysis of multiple samples by difference gel electrophoresis (DIGE) (Unlu et al. 1997) and advances in multidimensional protein identification technology (MudPIT) (Washburn et al. 2001).

Another drawback of this approach is the type of protease chosen which is crucial to maximize the digestion of all non-target proteins without compromising the protection of the target protein. Currently, protease mixtures are used and good results are being obtained with Pronase (Roche), a commercially available protease mixture capable of digesting both native and unfolded proteins (Lomenick et al. 2011).

If these drawbacks are solvent, DARTS could be an alternative to affinity chromatography in the coming years. Small-molecules whose targets have been identified or confirmed by this approach are summarized in Table 4 of the supporting information.

Stability of proteins from rates of oxidation (SPROX)

This affinity-based label-free approach is also based on thermodynamic measurements (West et al. 2010). In comparison with DARTS, the folding and thermodynamic stability of the target proteins are also studied but uses the ligand-induced changes in the methionine oxidation levels for target proteins as the readout (Lomenick et al. 2011). From an experimental standpoint, a cell lysate is studied in both the absence and presence of SM. Both samples are treated with increasing concentrations of guanidinium hydrochloride, a chemical denaturant, and the same amount of H_2O_2 in order to selectively oxidize the thioether groups in the side chain of methionine residues. Finally, a quantitative proteomic analysis is performed and the target proteins are identified by a change in protein stability since both samples are identical (West et al. 2010).

On the other hand, a disadvantage of SPROX is the fact that only the most abundant proteins in each sample can be identified and accurately quantified because there is no a target protein enrichment mechanism (Lomenick et al. 2011). But the biggest drawback of SPROX is that the target proteins must have a methionine-containing peptide and not all methionine residues have different oxidation rates whereby to determine thermodynamic changes (Roti and Stegmaier 2012).

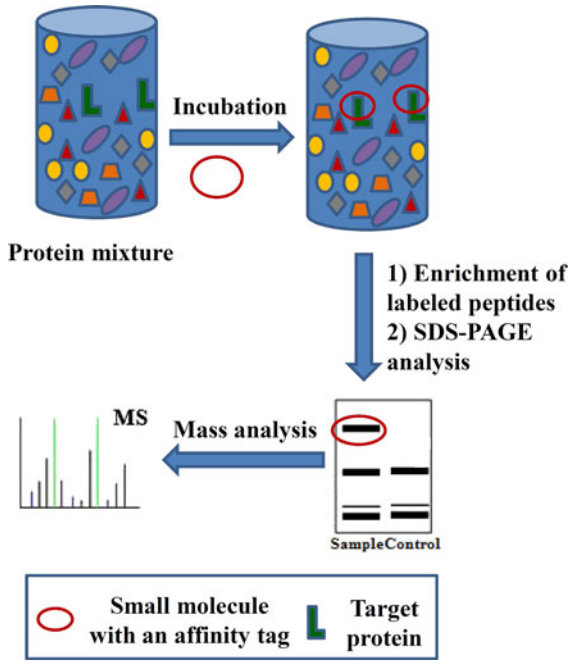


Fig. 8 General scheme of an ABPP experiment

Some examples of target proteins identified by means of this approach are shown in Table 5 of supporting information.

Activity-based protein profiling (ABPP)

This approach is based on the response of reactive functional groups (mainly electrophiles) of SMs with catalytic residues in the enzyme's active site (Liu et al. 1999; Sadaghiani et al. 2007). The SMs are comprised of a reactive group, a linker and a tag. Radioactive groups, fluorophores such as rhodamine and biotin can be used as tags to visualize and purify labelled proteins (Cravatt et al. 2008). Covalent modification can occur directly using a highly nucleophilic active site residue or by incorporation of a photoreactive group (i.e. benzophenone) followed by UV irradiation (Sleno and Emili 2008).

The protein targets for bioactive SMs are identified by comparing the labelling reduction when a compound is present (Fig. 8) (Sleno and Emili 2008). The addition of a SM, either in solution or in vivo, makes the inhibited target protein(s) less available to subsequent labelling.

The major advantage of ABPP is the fact that allow the targeted enzymes to be labelled for purification and

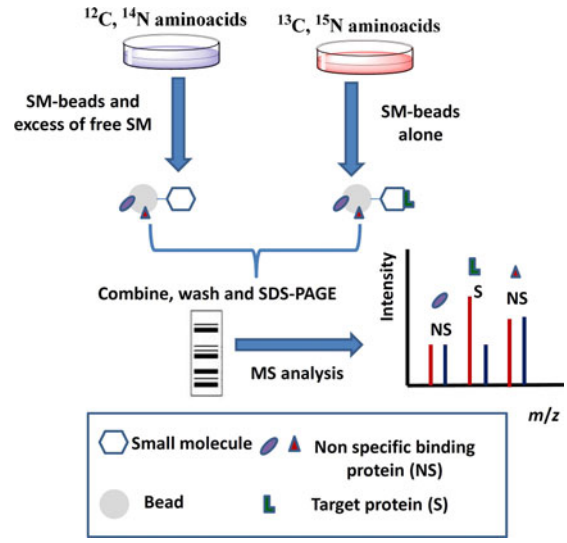


Fig. 9 Scheme of the SILAC technique. This figure is adapted from Ong et al. (2009)

analysis; However, it is necessary the derivatization of each SM which must contain reactive functional groups.

Small-molecule probes have been successfully used for the identification of proteases, lipases, cytochrome P-450s, glycosidases, kinases and phosphatases (Heal et al. 2011; Nomura et al. 2010; Simon and Cravatt 2010).

Mass spectrometry-based proteomic approaches

Quantitative proteomics have proven to be a powerful tool for discriminating between protein-SM specific interactions from background interactions in affinity experiments (Cheng et al. 2010). Nevertheless, the absolute intensity of a peptide in a mass spectrum depends, among other factors, on the type of ionization thus calling for the use of internal standards. The ideal internal standard should be a peptide of identical sequence but labelled with different stable isotopes. Hence, several strategies based on stable isotope labelling such as ICAT (isotope-coded affinity tag) (Gygi et al. 1999; Han et al. 2001; Oda et al. 2003), *i*TRAQ (isobaric relative and absolute tag for quantification) (DeSouza et al. 2005; Rix and Superti-Furga 2009; Ross et al. 2004) and SILAC (stable isotope labelling by amino acid in cell culture) (Ong et al. 2002, 2003, Ong et al. 2009; Yan and Chen 2005) have been developed for quantitative proteome analysis.

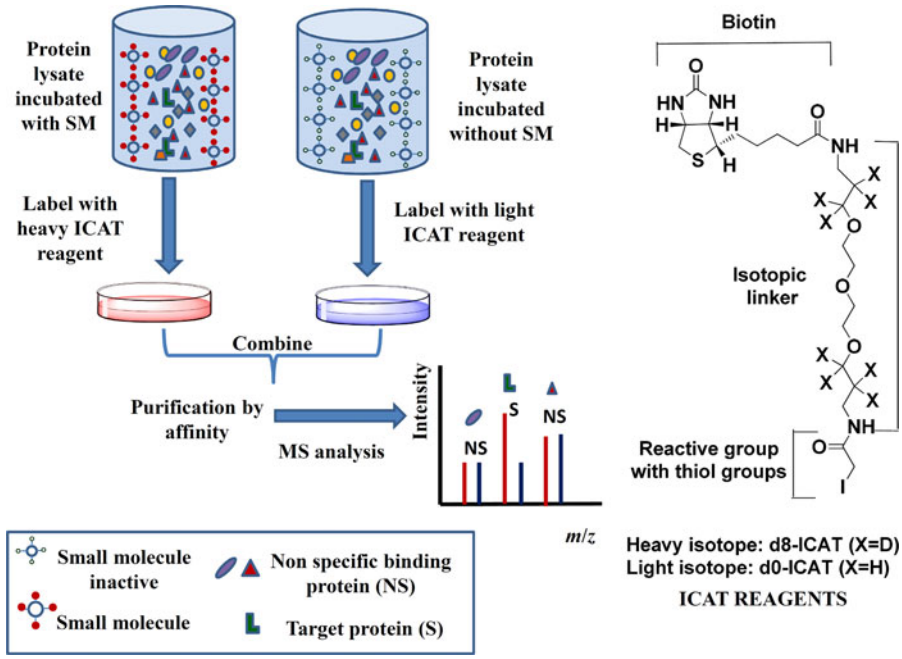


Fig. 10 Scheme of the ICAT technique

The SILAC approach exploits the ability of live cells to incorporate labelled amino acids through media supplementation. Proteins will incorporate either "light", natural isotope abundance forms, or the "heavy", ^{13}C and ^{15}N-bearing versions of arginine and lysine amino acids (Ong et al. 2002). Protein lysates from cultures are incubated either SM-loaded beads and soluble SM competitor or SM-bead alone. Then, the samples are combined, separated by SDS-PAGE and analyzed by mass spectrometry. Proteins interacting directly with the SM will be enriched in the heavy state over the light and will be identified by differential ratios. Non-specific interactions of proteins will be enriched equally in both states and will have the same isotopic ratio (Fig. 9) (Ong et al. 2009).

The molecular targets of several kinase inhibitors have been identified by SILAC (Ong et al. 2009) (see Table 6 supporting information).

A disadvantage of this method is that requires at least five populations to satisfactorily incorporate the isotopic labels and can only be used with biological samples grown in culture (Ong et al. 2003; Yan and Chen 2005). An alternative is the ICAT technique which is a gel-free approach enabling the quantification of proteins in any biological system (Roti and Stegmaier 2012). This approach uses a chemical reagent containing a biotin affinity tag for selective purification, a linker that incorporates stable isotopes and an iodoacetamide reactive group that specifically reacts with cysteinyl thiols (Yan and Chen 2004). In an ICAT experiment, a protein lysate incubated with the SM bound to an affinity matrix and the same lysate incubated with a molecule control are labelled in vitro with heavy (d8) and light (d0) ICAT reagents, respectively (Fig. 10). Then, both protein mixtures are combined, digested with trypsin into peptides and subjected to avidin affinity chromarography to enrich the labelled peptides that carry biotin tags in order to reduce the complexity of the mixtures. Finally, the labelled peptides are isolated, identified by their differences in the isotopic distribution and quantified using LC–MS.

Several target proteins of anticancer E7070 have been identified using this approach (Oda et al. 2003).

Although the ICAT approach drastically reduces the complexity of the mixture, only 96 % of human proteome possesses cysteine residues (Zhang et al. 2004). To overcome this difficulty, an amine group-based methodology known as *i*TRAQ has been developed by Ross et al. (2004). This in vitro gel-free approach utilizes isobaric reagents composed of a reporter group, a balance group and a reactive group that reacts specifically with primary amine groups of peptides (Fig. 11). The reporter group is a 4-methyl-piperazinyl tag with a mass ranging from 114 to 117 Da according to the different isotopic

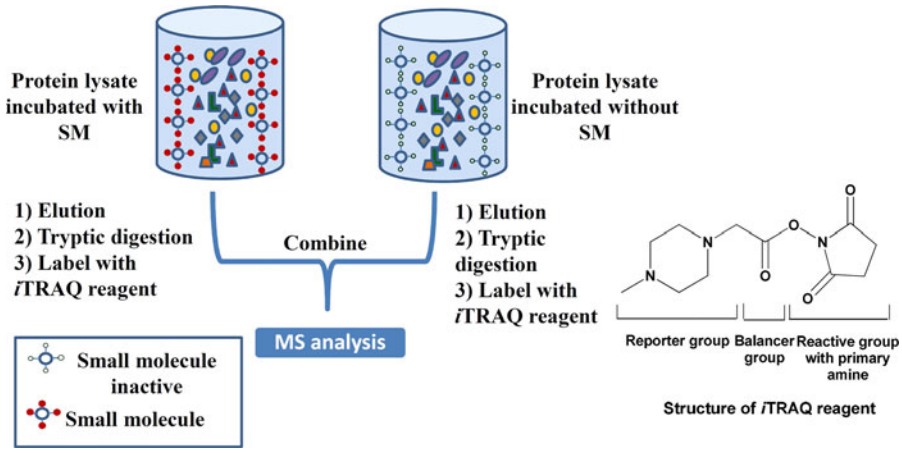


Fig. 11 Scheme of the *i*TRAQ approach

combinations of $^{12}C/^{13}C$ and $^{16}O/^{18}O$ in each individual reagent. A carbonyl group with a mass range between 28 and 31 Da is used as balancer group in order keep the combined mass of the reporter and balance groups constant at 145 Da for all four reagents. The reporter group is fragmented into ions from 114 to 117 Da in MS/MS experiments and the intensity of these fragments is used for quantification of up to four simultaneous samples (Yan and Chen 2005). In an *i*TRAQ experiment, a lysate incubated with the SM bound to an affinity matrix and the same lysate incubated with a molecule control are labelled in vitro each with a different *i*TRAQ reagent. Then, labelled peptides in both samples are combined and analyzed by LC–MS/MS and the target proteins are identified by their differences in the isotopic distribution (Fig. 11).

Recently, a kinobead matrix has been developed to evaluate the selectivity of several kinases by *i*TRAQ (see Table 7 of the supporting information) (Bantscheff et al. 2007).

Genetic methods to identify the molecular target

Genetic-based methods study the SM targets/pathways from the physiological responses or biochemical signatures produced by the SMs (Lomenick et al. 2009). We classify the genetic-based methods into the following groups: expression-cloning-based methods (Terstappen et al. 2007), microarray technologies (Sleno and Emili 2008), gene overexpression techniques (Luesch et al. 2005) and synthetic lethality approaches (Lum et al. 2004).

Expression-cloning-based methods

These approaches are based on the expression of target proteins from *c*DNA libraries. To solve the problems associated with the low abundance of target proteins, these methods include an affinity step to increase their quantity and facilitate their purification and isolation. Nevertheless, because the target proteins are expressed on the basis of fusion constructs of *c*DNA libraries, their properties might be different from those of their native counterparts and this could affect binding to the SM (Terstappen et al. 2007). Moreover, these methods are usually limited to yeast or other simple well-characterized model organisms.

The following approaches have been applied for the identification of target proteins: yeast and mammalian three hybrid systems, phage display, *m*RNA display and Drug westerns (Tanaka et al. 1999).

Yeast three hybrid (Y3H) and mammalian three hybrid systems The use of yeasts is widely used for the identification of target proteins, being *Saccharomyces cerevisiae* an organism model of simple eukaryotes given that its genome can easily be manipulated (Bjornsti 2002; Forsburg 2001).

Y3H is currently a promising approach for the identification of target proteins (Licitra and Liu 1996), which is comprised of three components: a synthetic hybrid ligand and two hybrid fusion proteins. One of the hybrid fusion proteins is composed of a DNA binding domain (LexA) fused generally to dihydrofolate reductase (DHFR), a target of methotrexate (MTX), which acts as a ligand binding domain (Abida et al. 2002; Baker et al. 2003; Lin et al. 2000). The

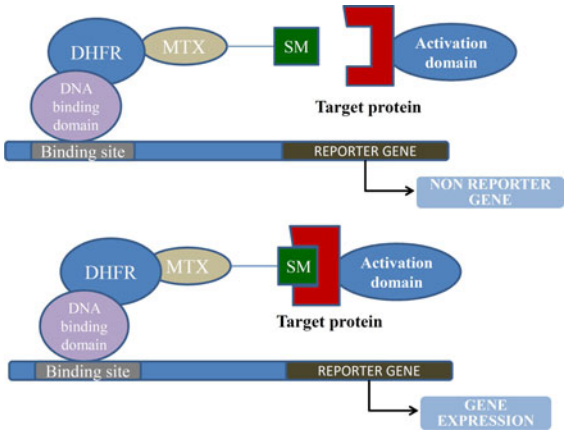


Fig. 12 Scheme of the Y3H approach

other fusion protein contains a transcriptional activation domain, usually derived from the yeast transcription factor GAL4 (Henthorn et al. 2002), fused to a cDNA library that includes the potential target proteins (Becker et al. 2004). The synthetic hybrid ligand is a SM that is covalently linked through a spacer to a ligand, usually MTX, in order to connect the first hybrid fusion protein to the second. From an experimental standpoint, a yeast strain that expresses DHFR is transformed with a library of *c*DNAs fused to a transcriptional activation domain. Then, the hybrid ligand is added to yeast and, if a binding protein is present in the library, the hybrid ligand will attach the transcriptional activation domain to the DNA activating expression of the reporter gene which is measured as the assay signal. Cells that express the reporter gene are selected and their DNA is extracted and sequenced in order to identify the potential target proteins by means of sequence-similarity searches (Fig. 12) (Becker et al. 2004).

Successful examples of identification of target proteins by Y3H are shown in Table 8 of the supporting information.

Unlike other expression cloning approaches the interaction between SMs and target proteins occur in living cells (Terstappen et al. 2007). Nevertheless, this approach can not be used directly in mammals due to differences in regard to yeast proteins. In order to solve this limitation, the MASPIT system (Mammalian Small Molecule Protein Interaction Trap) has been developed (Caligiuri et al. 2006). This is based on the associated JAK cytokine-receptor (janus-activated kinase)-STAT (signal transducer and activator of transcription) signal transduction system (Terstappen et al. 2007).

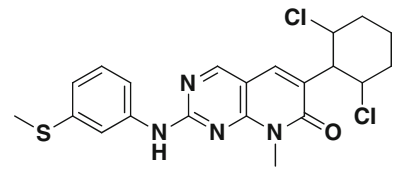


Fig. 13 Structure of PD173955

One successful example of MASPIT system in mammals is the identification of ephrin receptor tyrosine kinases and cyclin G-associated kinase of the ABL tyrosine kinase inhibitor PD173955 (Fig. 13) (Caligiuri et al. 2006).

mRNA display This approach utilizes *m*RNA display to identify proteins bound to a SM of interest (McPherson et al. 2002). Amplification by PCR (polymerase chain reaction) (Fig. 14 step 1) of a *c*DNA library with primers that introduce sequences permits in vitro transcription (Fig. 14 step 2). Then, the *m*RNAs generated are ligated to a puromycin-DNA linker (Fig. 14 step 3) and the in vitro translation is carried out to generate protein-*m*RNA fusion molecules (Fig. 14 step 4) (Liu et al. 2000; Terstappen et al. 2007). The complexes obtained are purified and reverse transcribed to generate a stable *c*DNA template (Fig. 14 step 5). Then, the biotin-labelled SMs of interest are immobilized on solid support and incubated with *m*RNA display molecules (Fig. 14 step 6) (Terstappen et al. 2007; Tochtrop and King 2004). Binding proteins are purified using avidin chromarography by elution with free biotin. After elution of the binding proteins (Fig. 14 step 7), the *c*DNAs are amplified by PCR (Fig. 14 step 1) resulting again in a *c*DNA library that is rich in target proteins. Finally, after several iteration processes, the *c*DNAs are purified, sequenced and the target protein identified by means of a DNA sequence similarity search (Hammond et al. 2001).

The major advantage of this approach is that the amplification process should permit the identification of low abundance targets by repetitive rounds of affinity selection; however, iterative rounds of amplification may promote the identification of short-length products rather than their full-length counterparts (Terstappen et al. 2007).

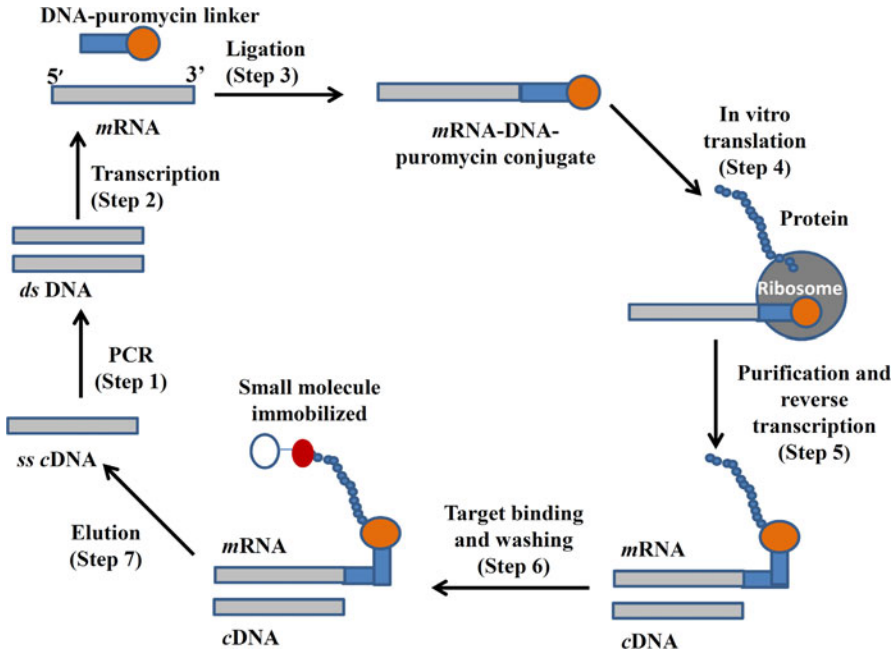


Fig. 14 General scheme of the *m*RNA approach. This figure is adapted from Tochtrop and King (2004)

To our knowledge, this method has only been successful with an immobilized conjugate of FK506-biotin to identify full-length FKBP12 (McPherson et al. 2002).

Phage display This approach has emerged as a valid approach to identify target proteins of SMs using the display of proteins on bacteriophages such as M13, T4 and T7, which enable the extraction of proteins from large libraries whereby a *c*DNA library is expressed with bacteriophage coat proteins (Kay et al. 1996; Sche et al. 1999). The phage library obtained is then exposed to the biotinylated immobilized SM resulting in the capture of bacteriophages that have affinity for the SM (Fig. 15 step 1) (King 1999). Next, an extensive wash is done to minimize non-specific interactions between some of the bacteriophage coat proteins and agarose column and the binding bacteriophages are eluted selectively with excess of biotin (Fig. 15 step 2) and transfected into bacterial host cells (Fig. 15 step 3) (usually *E. coli*) for amplification (Fig. 15 step 4) (Das et al. 2011; Terstappen et al. 2007). Phage populations obtained are re-exposed to immobilized SMs and are subjected to several rounds of affinity enrichment (Fig. 15 step 5) (Jung et al. 2009). Finally, a monoclonal phage

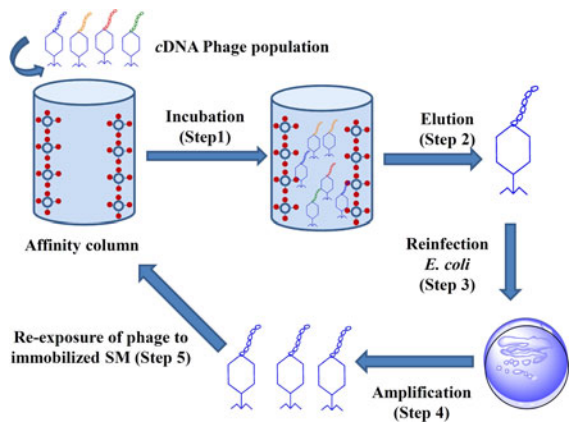


Fig. 15 Scheme of the phage display approach

population is obtained and the target proteins are identified by DNA sequence-similarity searches (Terstappen et al. 2007).

In general, low-abundance targets and targets of low-affinity SM can be identified by iterative amplification steps; however, in practice is usually complicated because clones that express high-affinity proteins will dominate the phage population. Another disadvantage of this approach is the fact that it is necessary to use a labelled-SM.

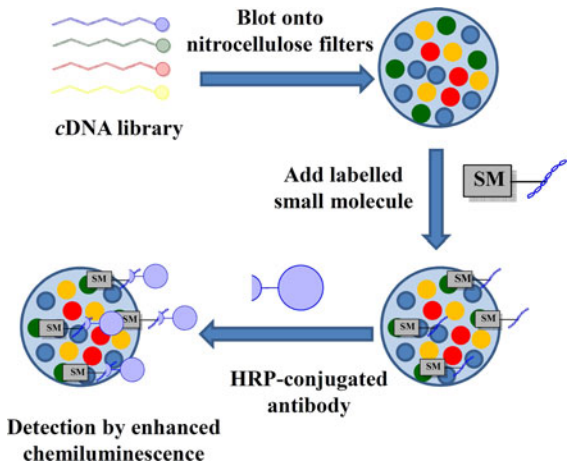
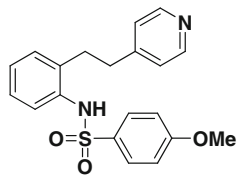


Fig. 16 Scheme of the Drug westerns approach

Fig. 17 Structure of HMN-154



Several successful examples of target protein identification using this approach are shown in Table 9 of the supporting information.

Drug westerns Tagged SMs are used in this approach to electrophoretically probe *c*DNA expression libraries (Tanaka et al. 1999). In this experiment, a plaque consisting of one member of a *c*DNA library is observed when bacteriophages are grown in a Petri dish together with the *c*DNA library (Chang 2009; Das et al. 2011). Proteins from the plaque are transferred to a nitrocellulose membrane where they are screened against labelled SMs with an antigen (BSA or FITC). Hits obtained from the plaques are detected by enhanced chemiluminescence, a single virus is isolated and purified and the target proteins are identified by DNA sequencing method (Fig. 16).

This method can facilitate immediate cloning of the genes encoding drug-binding proteins. However, as in other approaches, a limiting aspect of Drug westerns is the fact that a labelled SM is required.

An example of the application of Drug westerns method is the isolation of the transcription factor NF-YB and Thymosin β-10 as target proteins of the anticancer small molecule HMN-154 (Fig. 17) (Tanaka et al. 1999).

Microarray technologies

Miniaturized high-density arrays with immobilized proteins or *c*DNA (microarrays) can be used in high-throughput assays in order to identify the targets of the SMs. Protein microarrays and DNA microarrays have been developed for this purpose.

Protein microarrays In this approach the binding profile of a SM for an entire proteome can easily be obtained in a short period of time (Sleno and Emili 2008; Terstappen et al. 2007). A set of proteins are purified and immobilized by amino-terminal glutathione S-transferase (GST) tag on a glass microscope slide or some other derivatized surface (Heng et al. 2001; Heng and Snyder 2003). The microarray is then incubated with a labelled form of the SM, usually biotin. After thorough washing, bound targets are detected by adding a streptavidin fluorescently labelled conjugate and then the labelled proteins are identified by their positions on the microarray. As a loading control, the microarray is probed with GST-specific antibodies (Fig. 18) (Sleno and Emili 2008; Terstappen et al. 2007).

The major advantage of this method is that it allows for equal exposition of all proteins to the SM. Furthermore, proteins can be attached to chemically modified porous silicon and, after the addition of SMs, the bound substrates can be selectively monitored using techniques such as DIOS (direct ionization on silicon) (Shen et al. 2001) or MALDI-TOF.

Successful of this approach will be increased with the development of label-free binding detection compatible with the proteome array format.

On the other hand, this approach only reproduces the binding in vitro meaning that some proteins, whose affinity for the SM is due to post-translational modifications or their involvement in a complex formation with other proteins, might not be identified. Moreover, depending on the positioning of the protein on the chip, the SM might not bind to its target protein due to steric hindrance (Terstappen et al. 2007).

Some successful examples of this approach are shown in Table 10 of supporting information.

DNA microarrays This approach, known as "reverse transfected" cell microarrays, is based on the expression of defined *c*DNAs avoiding the use of individually purified proteins (Ziauddin and Sabatini

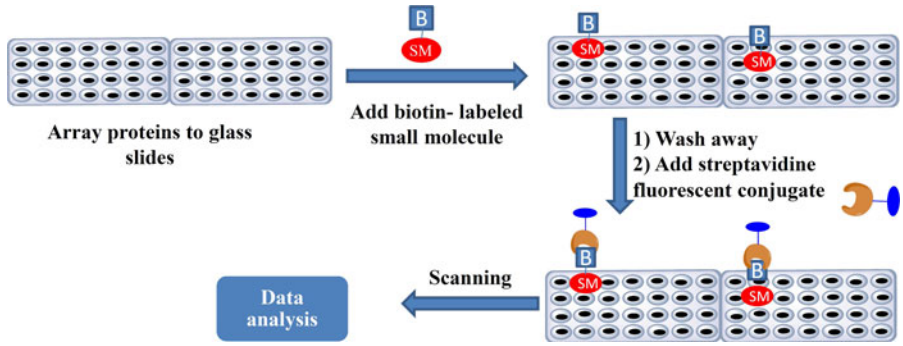


Fig. 18 Scheme of the protein microarrays approach. This figure is adapted from Terstappen et al. (2007)

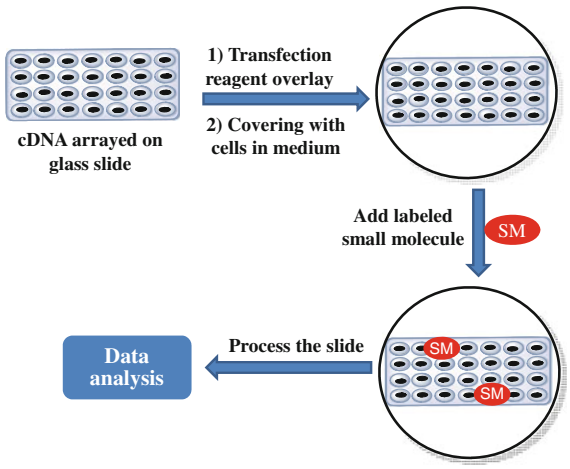


Fig. 19 Scheme of the DNA microarrays approach. This figure is adapted from Terstappen et al. (2007)

2001) since the glass slides are imprinted with sets of specific *c*DNAs in expression vectors, overlaid with a transfection reagent and covered with mammalian cells in a culture medium. Transfected cells will express *c*DNAs at a defined location on the chip and the target proteins will be revealed by the addition of a fluorescent or radioisotope-labelled SM to the chip and subsequent DNA sequence-similarity searches (Fig. 19) (Terstappen et al. 2007).

The major disadvantage of this approach is that the sensitivity of detection might be compromised for some cell types since transfection efficiencies can be relatively slow (Terstappen et al. 2007).

FKBP12 has been identified by DNA microarrays when radiolabelled FK506 is added to the culture medium of HEK293T cells (Ziauddin and Sabatini 2001).

Gene overexpression techniques

This technique is based on the principle that gene overexpression results in increased protein product, and if this protein is targeted by a SM, the cell should gain resistance to that SM (Luesch et al. 2005). This approach uses DNA libraries which are transformed by a yeast strain and the products obtained are grown on media containing the SM at a concentration that inhibits wild type growth. Target proteins are identified from drug-resistant transformants by means of sequencing of plasmid DNA (Fig. 20) (Bharucha and Kumar 2007).

However, the main disadvantage of this approach is that SMs often exhibit lower permeability in yeast cells than in mammalian cells.

Several successful examples of target protein identification using this approach are shown in Table 11 of the supporting information.

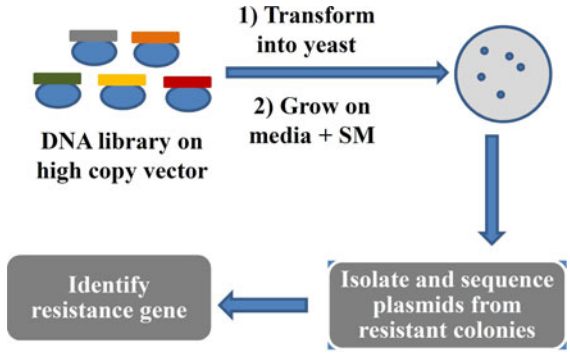


Fig. 20 General scheme of gene overexpression. This figure is adapted from Bharucha and Kumar (2007)

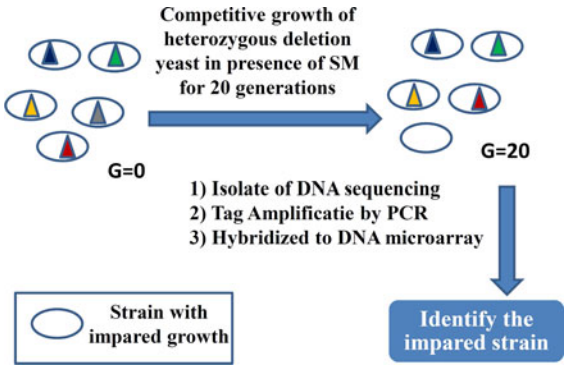


Fig. 21 Scheme representative of the synthetic lethality approach. This figure is adapted from Tochtrop and King (2004)

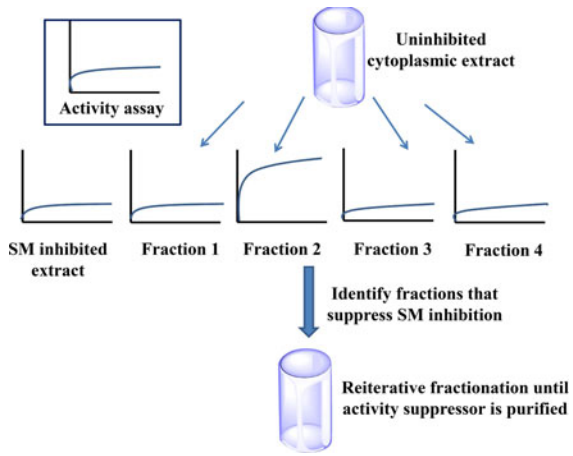


Fig. 22 Scheme of the biochemical suppression method. This figure is adapted from Peterson et al. (2006)

Synthetic lethality

This genetic approach uses collections of heterozygous mutants and studies their hypersensitivity to SMs (Giaever et al. 1999; Lum et al. 2004). The availability of the genome-wide collection of heterozygous yeast deletion strains is allowing the systematic application of this methodology.

It is often assumed that the reduction of the gene copy number of a target protein can result in sensitization to the small molecule of interest. Genomic DNA from cultures before (G = 0) and after twenty generations of growth (G = 20) are isolated, amplified by PCR and hybridized to DNA microarrays (Tochtrop and King 2004). Target proteins are identified by their ability to confer resistance to a strain when present at high copy (Fig. 21).

This approach has the advantage that a free-label SM can be used.

However, hypersensitivity may can due both direct and indirect mechanism, so other techniques will be required in order to confirm the molecular target of each SM (Tochtrop and King 2004). Moreover, only can be used SM that affect cell growth/viability.

Several successful examples of target identification using this approach are shown in Table 12 of the supporting information.

Other methods

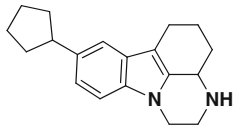
Despite the fact that genetic and affinity-based methods have successfully contributed to the identification of a large number of target proteins, these methods cannot be applied in a general manner due the nature of each ligand and target protein and it is necessary to investigate and develop new methods for the elucidation of biological targets. In this section, we will discuss the identification of targets using different basis to those described above.

Biochemical suppression

This method is based on the functional suppression of protein activity by chemical inhibitors by in vitro reactions regardless of the affinity that the SM has for its target protein (Peterson et al. 2006; Terstappen et al. 2007). From an experimental standpoint, an activity assay is used to measure the inhibition produced in a protein extract when a SM that inhibits the activity of interest is added. Fractions of uninhibited extract are added to the inhibited extract to identify the fraction that suppresses the SM's inhibitory activity. This process is repeated until the protein that suppresses the inhibition is purified and identified by gel electrophoresis and mass spectrometry (Fig. 22) (Terstappen et al. 2007).

The major advantages of this approach are that it does not require modification of the SM and small amounts of proteins can be identified because this method is not based on affinity. Furthermore, suppressors are introduced as native protein forms rather than individual gene products, allowing for identification not only of the target proteins but also of other pathway components (Terstappen et al. 2007).

Fig. 23 Structure of Pirl1



However, this method needs to use protein fractions which is a drawback for intact cells which would need to be permeabilized (Terstappen et al. 2007).

Although this approach might be a good alternative to affinity-based methods, to our knowledge, it has only been identified the Cdc42–RhoGDI (Rho-GDP dissociation inhibitor) complex as a direct target of pirl1 (Fig. 23) and the Arp 2/3 (actin and related proteins) complex as a downstream component of the actin assembly pathway that is capable of relieving upstream inhibition of Cdc42–RhoGDI when it is added at high concentrations (Peterson et al. 2006).

Target validation assays

Once a target protein is identified, the next step is to make certain that its modulation is associated with the phenotype observed in the assay where the SM was identified (Terstappen et al. 2007). There are several of approaches and tools available to discriminate between true interactions and false positives (Kramer and Cohen 2004) of which we would highlight bioinformatic analysis, if the three-dimensional protein structure is available (Macchiarulo et al. 2004; Mueller et al. 2007), study of the physical interaction of the SM and the putative target by surface plasmon resonance (Boozer et al. 2006; Elinder et al. 2011; O'Connell et al. 2010), fluorescence anisotropy (Zhang et al. 2011), isothermal titration calorimetry (Buurman et al. 2011; Zhou et al. 2008) or resonance acoustic profiling (Li et al. 2006) and the use of cell culture validation experiments such as RNA interference (RNAi) and overexpression of the target protein.

Concluding remarks and future prospects

NPs play an important role in the discovery of new drugs and agrochemicals owing to their great chemical diversity, high affinity and specificity for biological targets. Understanding the mechanism of action of bioactive compounds and identifying their molecular targets, are still the most important challenges facing chemical genetics. Despite the great variety of new methods that have been developed over the last several years, affinity chromatography is still the most widespread and successful method for identifying the target proteins of bioactive-small molecules. Proof of this is the disproportionately large number of target proteins identified by affinity chromatography highlighted in this review in comparison with the rest of the techniques.

Recent advances in mass-spectrometry proteomic methods such as SILAC, ICAT and iTRAQ, together with proteomic approaches such as 2D-PAGE, DIGE or MudPIT, have enabled a distinction to be drawn between specific and nonspecific binders and will help to identify low abundance targets with a low-affinity small molecule by affinity chromatography. Moreover, subcellular localization of the small molecule by means of fluorescence probe could be used for affinity purification in order to decrease sample complexity.

However, affinity chromatography is limited to SM that contain derivatizable functionalities and whose bioactivity is unaffected by modification, which is a serious hurdle. Future advances in new techniques should avoid the derivatization or labelling of SMs and should allow their application to protein extracts and membrane proteins.

In this context, DARTS has emerged as a promising alternative to circumvent the drawbacks outlined above. Nevertheless, new experimental conditions and new advances in degradomics must be developed in order to decrease the number of proteolysis conditions that need to be tested for each SM. Free-label protein microarrays could be used simultaneously with DARTS to study a whole proteome. Further identification of new target proteins using these approaches will prove whether they are suited for widespread use.

Acknowledgments This research was supported by grants from MICINN (AGL-2009-13359-C02-01) and from the Junta de Andalucia (P07-FQM-02689). Jose Manuel Botubol is grateful to the Junta de Andalucia for his research fellowship.

References

Abida WM, Crater BT, Althoff E et al (2002) Receptor-dependence of the transcription read-out in a small-molecule three-hybrid system. ChemBioChem 3:887–895

Altmann K-H, Buchner J, Kessler H et al (2009) The state of the art of chemical biology. ChemBioChem 10:16–29

Bach S, Knockaert M, Reinhardt J et al (2005) Roscovitine targets, protein kinases and pyridoxal kinase. J Biol Chem 280:31208–31219

Baker K, Sengupta D, Salazar-Jimenez G et al (2003) An optimized dexamethasone-methotrexate yeast 3-hybrid system for high-throughput screening of small molecule-protein interactions. Anal Biochem 315:134–137

Bantscheff M, Eberhard D, Abraham Y et al (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. Nat Biotechnol 25:1035–1044

Becker F, Murthi K, Smith C et al (2004) A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors. Chem Biol 11:211–223

Bharucha N, Kumar A (2007) Yeast genomics and drug target identification. Comb Chem High Throughput Screen 10:618–634

Bjornsti M-A (2002) Cancer therapeutics in yeast. Cancer Cell 2:267–273

Blackwell HE, Zhao YD (2003) Chemical genetic approaches to plant biology. Plant Physiol 133:448–455

Boozer C, Kim G, Cong S et al (2006) Looking towards label-free biomolecular interaction analysis in a highthroughput format: a review of new surface plasmon resonance technologies. Curr Opin Biotechnol 17:400–405

Bronson J, Dhar M, Ewing W et al (2011) To market, to market-2010. Annu Rep Med Chem 46:433–502

Brunner J (1993) New photolabeling and crosslinking methods. Annu Rev Biochem 62:483–514

Burdine L, Kodadek T (2004) Target identification in chemical genetics: the (often) missing link. Chem Biol 11:593–597

Buurman ET, Andrews B, Gao N et al (2011) In vitro validation of acetyltransferase activity of GlmU as an antibacterial target in *Haemophilus influenzae*. J Biol Chem 286:40734–40742

Caligiuri M, Molz L, Liu Q et al (2006) MASPIT: three-hybrid trap for quantitative proteome fingerprinting of small molecule-protein interactions in mammalian cells. Chem Biol 13:711–722

Campillos M, Kuhn M, Gavin A-C et al (2008) Drug target identification using side-effect similarity. Science 321:263–266

Carlson EE (2010) Natural products as chemical probes. ACS Chem Biol 5:639–653

Castoreno AB, Eggert US (2011) Small molecule probes of cellular pathways and networks. ACS Chem Biol 6:86–94

Chang YT (2009) Forward chemical genetics. Wiley Encyclopedia Chem Biol 2:94–111

Cheng K-W, Wong C-C, Wang M et al (2010) Identification and characterization of molecular targets of natural products by mass spectrometry. Mass Spectrom Rev 29:126–155

Cohen MS, Hadjivassiliou H, Taunton J (2007) A clickable inhibitor reveals context-dependent autoactivation of p90 RSK. Nat Chem Biol 3:156–160

Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. Annu Rev Biochem 77:383–414

Danishefsky SJ (2010) On the potential of natural products in the discovery of pharma leads: a case for reassessment. Nat Prod Rep 27:1114–1116

Das RK, Samanta A, Ghosh K et al (2011) Target identification: a challenging step in forward chemical genetics. IBC 3:1–16

DeSouza L, Diehl G, Rodrigues MJ et al (2005) Search for cancer markers from endometrial tissues using differentially labelled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. J Proteome Res 4:377–386

Dormán G, Prestwich GD (1994) Benzophenone photophores in biochemistry. Biochemistry 33:5661–5673

Dormán G, Prestwich GD (2000) Using photolabile ligands in drug discovery and development. Trends Biotechnol 18:64–77

Drews J (2000) Drug discovery: a historical perspective. Science 287:1960–1964

Elinder M, Geitmann M, Gossas T et al (2011) Experimental validation of a fragment library for lead discovery using SPR biosensor technology. J Biomol Screen 16:15–25

Emami KH, Nguyen C, Ma H et al (2004) A small molecule inhibitor of β-catenin/cyclic AMP response element-binding protein transcription. Proc Natl Acad Sci USA 101:12682–12687

Fleming SA (1995) Chemical reagents in photoaffinity labeling. Tetrahedron 51:12479–12520

Forsburg SL (2001) The art and design of genetic screens: yeast. Nat Rev Genet 2:659–668

Fox S, Farr-Jones S, Yund MA (1999) High throughput screening for drug discovery: continually transitioning into new technology. J Biomol Screen 4:183–186

Furuya M, Tsushima Y, Tani S et al (2006) Development of affinity chromatography using a bioactive peptide as a ligand. Bioorg Med Chem 14:5093–5098

Garrett MD, Fattaey A (1999) CDK inhibition and cancer therapy. Curr Opin Genet Dev 9:104–111

Gaudillière B, Bernardelli P, Berna P (2001) To market, to market-2000. Annu Rep Med Chem 36:293–318

Giaever G, Shoemaker DD, Jones TW et al (1999) Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. Nat Genet 21:278–283

Guiffant D, Tribouillard D, Gug F et al (2007) Identification of intracellular targets of small molecular weight chemical compounds using affinity chromatography. Biotechnol J 2:68–75

Gygi SP, Rist B, Gerber SA et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999

Hammond PW, Alpin J, Rise CE et al (2001) In vitro selection and characterization of Bcl-XL-binding proteins from a mix of tissue-specific mRNA display libraries. J Biol Chem 276:20898–20906

Han DK, Eng J, Zhou H et al (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol 19:946–951

Harding MW, Galat A, Uehling DE et al (1989) A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. Nature 341:758–760

Heal WP, Dang TH, Tate EW (2011) Activity-based probes: discovering new biology and new drug targets. Chem Soc Rev 40:246–257

Heng Z, Snyder M (2003) Protein chip technology. Curr Opin Chem Biol 7:55–63

Heng Z, Bilgin M, Bangham R et al (2001) Global analysis of protein activities using proteome chips. Science 293:2101–2105

Henthorn DC, Jaxa-Chamiec AA, Meldrum E (2002) A GAL4-based yeast three-hybrid system for the identification of small molecule-target protein interactions. Biochem Pharmacol 63:1619–1628

Hofmann K, Kiso Y (1976) An approach to the targeted attachment of peptides and proteins to solid supports. Proc Natl Acad Sci USA 73:3516–3518

Hofmann K, Finn FM, Kiso Y (1978) Avidin-biotin affinity columns. General methods for attaching biotin to peptides and proteins. J Am Chem Soc 100:3585–3590

Horrobin DF (2000) Innovation in the pharmaceutical industry. J R Soc Med 93:341–345

Hübel K, Lessmann T, Waldmann H (2008) Chemical biology-identification of small molecule modulators of cellular activity by natural product inspired synthesis. Chem Soc Rev 37:1361–1374

Hughes CC, Yang Y-L, Liu W-T et al (2009) Marinopyrrole a target elucidation by acyl dye transfer. J Am Chem Soc 131:12094–12096

Ito T, Ando H, Suzuki T et al (2010) Identification of a primary target of thalidomide teratogenicity. Science 327:1345–1350

Jung D-W, Williams D, Khersonsky SM et al (2005) Identification of the F1F0 mitochondrial ATPase as a target for modulating skin pigmentation by screening a tagged triazine library in zebrafish. Mol BioSyst 1:85–92

Jung JH, Shim JS, Park J et al (2009) Identification and validation of calmodulin as a binding protein of an anti-proliferative small molecule 3,4-dihydroisoquinolinium salt. Proteomics Clin Appl 3:423–432

Katayama H, Oda Y (2007) Chemical proteomics for drug discovery based on compound-immobilized affinity chromatography. J Chromatogr B 855:21–27

Kay BK, Adey NB, Sparks AB et al (1996) Construction of random peptide libraries in bacteriophage M13. In: Kay BK, Winter J, McCafferty J (eds) Phage display of peptides and proteins: a laboratory manual, San Diego

King RW (1999) Chemistry or biology: which comes first after the genome is sequenced? Chem Biol 6:R327–R333

Kotake Y, Sagane K, Owa T et al (2007) Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol 3:570–575

Kotzyba-Hibert F, Kapfer I, Goeldner M (1995) Recent trends in photoaffinity labeling. Angew Chem Int Ed Engl 34:1296–1312

Kramer R, Cohen D (2004) Functional genomics to new drug targets. Nat Rev Drug Discov 3:965–972

Kwok TCY, Ricker N, Fraser R et al (2006) A small-molecule screen in *C. elegans* yields a new calcium channel antagonist. Nature 441:91–95

Lebl M (1999) Parallel personal comments on "Classical" papers in combinatorial chemistry. J Comb Chem 1:3–24

Leslie BJ, Hergenrother PJ (2008) Identification of the cellular targets of bioactive small organic molecules using affinity reagents. Chem Soc Rev 37:1347–1360

Li X, Thompson KSJ, Godber JB et al (2006) Quantification of small molecule-receptor affinities and kinetics by acoustic profiling. Assay Drug Dev Technol 4:565–573

Licitra EJ, Liu JO (1996) A three-hybrid system for detecting small ligand-protein receptor interactions. Proc Natl Acad Sci USA 93:12817–12821

Lin H, Abida WM, Sauer RT et al (2000) Dexamethasone–Methotrexate: an efficient chemical inducer of protein dimerization in vivo. J Am Chem Soc 122:4247–4248

Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: the serine hydrolases. Proc Natl Acad Sci USA 96:14694–14699

Liu R, Barrick JE, Szostak JW et al (2000) Optimized synthesis of RNA-protein fusions for in vitro protein selection. Methods Enzymol 318:268–293

Lokey RS (2003) Forward chemical genetics: progress and obstacles on the path to a new pharmacopoeia. Curr Opin Chem Biol 7:91–96

Lomenick B, Hao R, Jonai N et al (2009) Target identification using drug affinity responsive target stability (DARTS). Proc Natl Acad Sci USA 106:21984–21989

Lomenick B, Olsen RW, Huang J (2011) Identification of direct protein targets of small molecules. ACS Chem Biol 6:34–46

Luesch H, Wu TYH, Ren P et al (2005) A genome-wide over-expression screen in yeast for small-molecule target identification. Chem Biol 12:55–63

Lum PY, Armour CD, Stepaniants SB et al (2004) Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. Cell 116:121–137

Macchiarulo A, Nobeli I, Thornton JM (2004) Ligand selectivity and competition between enzymes in silico. Nat Biotech 22:1039–1045

MacKinnon AL, Garrison JL, Hegde RS et al (2007) Photoleucine incorporation reveals the target of a cyclodepsipeptide inhibitor of cotranslational translocation. J Am Chem Soc 129:14560–14561

McPherson M, Yang Y, Hammond PW et al (2002) Drug receptor identification from multiple tissues using cellular-derived mRNA display libraries. Chem Biol 9:691–698

Mendoza LG, McQuary P, Mongan A et al (1999) High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA). Biotechniques 27:778,780, 782–786, 788

Mueller M, Martens L, Apweiler R (2007) Annotating the human proteome: beyond establishing a parts list. Biochim Biophys Acta 1774:175–191

Neumann G, Hatta M, Kawaoka Y (2003) Reverse genetics for the control of avian influenza. Avian Dis 47:882–887

Nishiya Y, Shibata K, Saito S et al (2009) Drug-target identification from total cellular lysate by drug-induced conformational changes. Anal Biochem 385:314–320

Nomura DK, Dix MM, Cravatt BF (2010) Activity-based protein profiling for biochemical pathway discovery in cáncer. Nat Rev Cancer 10:630–638

O'Connell DJ, Bauer MC, O'Brien J et al (2010) Integrated protein array screening and high throughput validation of 70 novel neural calmodulin-binding proteins. Mol Cell Proteomics 9:1118–1132

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

Oda S, Owa T, Sato T et al (2003) Quantitative chemical proteomics for identifying candidate drug targets. Anal Chem 75:2159–2165

Ong S-E, Blagoev B, Kratchmarova I et al (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386

Ong S-E, Foster LJ, Mann M (2003) Mass spectrometric-based approaches in quantitative proteomics. Methods 29:124–130

Ong S-E, Schenone M, Margolin AA et al (2009) Identifying the proteins to which small-molecule probes and drugs bind in cells. Proc Natl Acad Sci USA 106:4617–4622

Oprea TI (2002) Chemical space navigation in lead discovery. Curr Opin Chem Biol 6:384–389

Peterson JR, Lebensohn AM, Pelish HE et al (2006) Biochemical suppression of small-molecule inhibitors: a strategy to identify inhibitor targets and signaling pathway components. Chem Biol 13:443–452

Rix U, Superti-Furga G (2009) Target profiling of small molecules by chemical proteomics. Nat Chem Biol 5:616–624

Ross PL, Huang YN, Marchese JN et al (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169

Roti G, Stegmaier K (2012) Genetic and proteomic approaches to identify cancer drug targets. Br J Cancer 106:254–261

Sadaghiani AM, Verhelst SHL, Bogyo M (2007) Tagging and detection strategies for activity-based proteomics. Curr Opin Chem Biol 11:20–28

Sadakane Y, Hatanaka Y (2006) Photochemical fishing approaches for identifying target proteins and elucidating the structure of a ligand-binding region using carbene-generating photoreactive probes. Anal Sci 22:209–218

Sakamoto S, Hatakeyama M, Ito T et al (2012) Tools and methodologies capable of isolating and identifying a target molecule for a bioactive compound. Bioorg Med Chem 20:1990–2001

Sato S-I, Kwon Y, Kamisuki S et al (2007) Polyproline-rod approach to isolating protein targets of bioactive small molecules: isolation of a new target of indomethacin. J Am Chem Soc 129:873–880

Sato S-I, Murata A, Shirakawa T et al (2010) Biochemical target isolation for novices: affinity-based strategies. Chem Biol 17:616–622

Sche PP, McKenzie KM, White JD et al (1999) Display cloning: functional identification of natural product receptors using cDNA-phage display. Chem Biol 6:707–716

Schreiber SL (1998) Chemical genetics resulting from a passion for synthetic organic chemistry. Bioorg Med Chem 6:1127–1152

Schreiber SL (2003) The small-molecule approach to biology. Chem Eng News 81:51–61

Shen Z, Thomas JJ, Averbuj C et al (2001) Porous silicon as a versatile platform for laser desorption/ionization mass spectrometry. Anal Chem 73:612–619

Shiyama T, Furuya M, Yamazaki A et al (2004) Design and synthesis of novel hydrophilic spacers for the reduction of nonspecific binding proteins on affinity resins. Bioorg Med Chem 12:2831–2841

Simon GM, Cravatt BF (2010) Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. J Biol Chem 285:11051–11055

Sleno L, Emili A (2008) Protemic methods for drug target discovery. Curr Opin Chem Biol 12:46–54

Snyder JR, Hall A, Ni-Komatsu L et al (2005) Dissection of melanogenesis with small molecules identifies prohibitin as a regulator. Chem Biol 12:477–484

Speers AE, Cravatt BF (2004) Profiling enzyme activities in vivo using click chemistry methods. Chem Biol 11: 535–546

Spring DR (2005) Chemical genetics to chemical genomics: small molecules offer big insights. Chem Soc Rev 34: 472–482

Sridhar R, Hanson-Painton O, Cooper DR (2000) Protein kinases as therapeutic targets. Pharm Res 17:1345–1353

Stockwell BR (2004) Exploring biology with small organic molecules. Nature 432:846–854

Stockwell BR, Haggarty SJ, Schreiber SL (1999) High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. Chem Biol 6:71–83

Tamura T, Terada T, Tanaka A (2003) A quantitative analysis and chemical approach for the reduction of nonspecific binding proteins on affinity resins. Bioconjug Chem 14:1222–1230

Tanaka H, Ohshima N, Hidaka H (1999) Isolation of cDNAs encoding cellular drug-binding proteins using a novel expression cloning procedure: drug-western. Mol Pharmacol 55:356–363

Terstappen GC, Schlüpen C, Raggiaschi R et al (2007) Target deconvolution strategies in drug discovery. Nat Rev Drug Discov 6:891–903

Tochtrop GP, King RW (2004) Target identification strategies in chemical genetics. Comb Chem High Throughput Screening 7:677–688

Tsien RY (1998) The green fluorescent protein. Annu Rev Biochem 67:509–544

Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. Electrophoresis 18:2071–2077

Walsh DP, Chang Y-T (2006) Chemical genetics. Chem Rev 106:2476–2530

Wang G, Shang L, Burgett AWG et al (2007) Diazonamide toxins reveal an unexpected function for ornithine δ-amino transferase in mitotic cell division. Proc Natl Acad Sci USA 104:2068–2073

Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247

West GM, Tucker CL, Xu T et al (2010) Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. Proc Natl Acad Sci USA 107:9078–9082

Witkowski J (2010) Drawing the map of life: inside the human genome project by Victor K. McElheny. Nature 466: 921–922

Yamamoto K, Yamazaki A, Takeuchi M et al (2006) A versatile method of identifying specific binding proteins on affinity resins. Anal Biochem 352:15–23

Yan W, Chen SS (2005) Mass spectrometry-based quantitative proteomic profiling. Brief Funct Genomic Proteomic 4:27–38

Zhang J, Campbell RE, Ting AY et al (2002) Creating new fluorescent probes for cell biology. Nat Rev Mol Cell Biol 3:906–918

Zhang H, Yan W, Aebersold R (2004) Chemical probes and tandem mass spectrometry: a strategy for the quantitative analysis of proteomes and subproteomes. Curr Opin Chem Biol 8:66–75

Zhang Q, Major MB, Takanashi S et al (2007) Small-molecule synergist of the Wnt/β-catenin signaling pathway. Proc Natl Acad Sci USA 104:7444–7448

Zhang D, Lu M, Wang H (2011) Fluorescence anisotropy analysis for mapping aptamer-protein interaction at the single nucleotide level. J Am Chem Soc 133:9188–9191

Zhou X, Sun Q, Kini RM et al (2008) A universal method for fishing target proteins from mixtures of biomolecules using isothermal titration calorimetry. Protein Sci 17:1798–1804

Ziauddin J, Sabatini DM (2001) Microarrays of cells expressing defined cDNAs. Nature 411:107–110