Identification and validation of protein targets of bioactive small molecules

Denis V. Titov [a], Jun O. Liu [a,b,*]

[a] Department of Pharmacology, Johns Hopkins University School of Medicine, MD, United States
[b] Department of Oncology, Johns Hopkins University School of Medicine, MD, United States

ARTICLE INFO

ABSTRACT

Identification and validation of protein targets of bioactive small molecules is an important problem in chemical biology and drug discovery. Currently, no single method is satisfactory for this task. Here, we provide an overview of common methods for target identification and validation that historically were most successful. We have classified for the first time the existing methods into two distinct and complementary types, the 'top-down' and 'bottom-up' approaches. In a typical top-down approach, the cellular phenotype is used as a starting point and the molecular target is approached through systematic narrowing down of possibilities by taking advantage of the detailed existing knowledge of cellular pathways and processes. In contrast, the bottom-up approach entails the direct detection and identification of the molecular targets using affinity-based or genetic methods. A special emphasis is placed on target validation, including correlation analysis and genetic methods, as this area is often ignored despite its importance.

1. Introduction

Small molecules are widely used as tools to study and perturb biological systems. Bioactive small molecules provide an opportunity to quickly switch on or off the activity of proteins in both temporally and spatially controlled manner, which is unmatched by any other method. The main obstacle to a wider use of small molecules for studying biology is the difficulty of identifying compounds that can specifically alter the function of a given protein or protein family. Historically, many drugs and bioactive small molecules, particularly natural products, were discovered based on their desired or undesired physiological effects at the cellular or organismal level. Using *ad hoc* approaches, researchers have painstakingly identified the protein targets of many widely used drugs and bioactive natural products. This has led to a deeper understanding of the physiological functions of the newly identified protein targets and sometimes of the pathophysiology of diseases that the drugs were used to intervene. With the advent of contemporary chemical biology, the availability of high-throughput phenotypic screening platforms coupled with the significant increase in the number of chemical compounds in both public and private sectors has led to an emergence of a growing collection of chemical compounds that are known to interfere with various biological processes. In contrast to the discovery of biologically active compounds, however, target identification and validation has remained a major bottleneck that has prevented the wider use of

small molecules in biological research and impeded to certain extent drug development.

In this review, we attempt to provide an overview of some of the most widely used and effective methods for identification and validation of protein targets of bioactive small molecules. Due to space limitations, we will not be able to cover all the available methods. This review is divided into two parts. In the first part, we will discuss the common methods used to identify candidate target proteins. In the second part, we will cover methods commonly applied to validate physiological relevance of newly identified protein targets.

2. Target Identification

We have broadly divided target identification methods into two general classes, the bottom-up and the top-down approaches. For definition, we place the small molecule, its protein target and its cellular phenotype on a vertical line with the protein target at the bottom and the global cellular phenotype at the top (Fig. 1). The first approach consists of methods that allow for the direct identification of the protein target using genetics or affinity-based methods. We call this the 'bottom-up' approach because it starts at the 'bottom' with identification of the target protein before one goes 'up' to the next level to explain the phenotype through perturbation of the function of the target protein. The second approach consists of methods that allow identification of the protein target by exploiting the existing knowledge of a given cellular process that is perturbed by the small molecule. We call this the 'top-down approach' as it allows one to narrow 'down' the possible targets

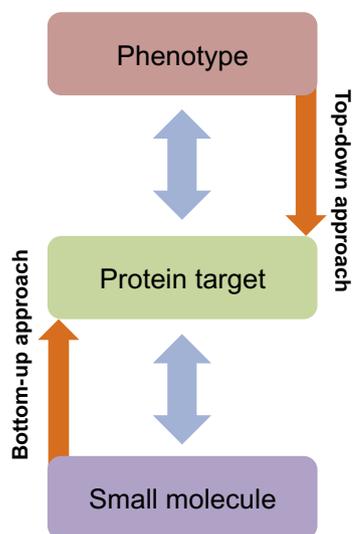* Corresponding author.
E-mail address: joliu@jhu.edu (J.O. Liu).

Figure 1. Schematic representation of top-down and bottom-up approaches to target identification.

based on a general understanding of what part of cellular and/or organismal physiology the small molecule affects and the proteins known to be involved in the relevant process.

2.1. Bottom-up approach

Up to now, the two most successful methods used for target identification are affinity purification and genetics. Genetics is the method of choice for target identification of antibiotics and antifungal drugs. Affinity purification has been successfully applied to drugs and small molecules that are active only in mammalian cells or whole organisms. Each method has its advantages and disadvantages. Affinity purification requires chemical modification of the compound under study that in turn requires prior knowledge on the structure/activity relationship of the compound. In contrast, genetic approaches employ unmodified compound. But genetic approaches require an appropriate cellular readout that is suitable for selection. Thus, not all biologically active small molecules can be readily studied using genetic methods (e.g. compounds that do not inhibit cell proliferation). In addition, genetic approaches may not lead to the identification of the direct target for small molecules, even though it may shed light on the pathways in which the target lie or with which the target protein may genetically interact. A high degree of uncertainty exists for both approaches and it is best to use both methods whenever possible to increase the chances of success.

2.1.1. Affinity-based methods

Affinity purification is one of the most widely used methods for target identification. It was successfully applied to a variety of small molecules and drugs. The basic premise of this approach is that a protein that binds a small molecule with the highest affinity is likely to mediate the biological activity of the compound. This approach comprises at least two distinct methods. In one method, a small molecule of interest is attached to a resin and then affinity chromatography is performed using an appropriate cell and tissue lysate in the presence or absence of excess free small molecule. Here, the goal is to identify proteins that bind to the affinity resin and are competed away in the presence of excess free compound. In the other method, a radioactively or fluorescently labeled small molecule is used to measure the binding activity of cell or tissue

lysates to facilitate the purification of a fraction that contains the majority of drug binding activity.

Classically, affinity chromatography approach is used with a small molecule immobilized on a resin such as agarose or sepharose or other types of beads (Fig. 2). Affinity resin is incubated with a cell lysate of choice, washed several times with a lysis buffer and proteins are eluted by either a solution of high concentrations of free small molecule or simply by boiling in Laemmli sample buffer. The eluted proteins are analyzed on an SDS–PAGE gel using silver staining. The identity of specific binding proteins that can be competed away by the free drug is determined by mass spectrometry. Success is dependent on a number of factors, including what cell or tissue type and what lysis buffer to use, how long to incubate the beads with the lysate and how stringent the washes should be. Unfortunately, there are no rules for arriving at the most optimal conditions and they are usually determined by trial and error. A prerequisite for a successful use of affinity chromatography approach is to attach the compound to the resin in such a way that it does not significantly abrogate its binding activity. Another important factor is that the target protein is abundant in the tissue or cell type that one uses. More recently, a variation of this method has emerged that takes advantage of the high affinity between biotin and streptavidin or avidin. Instead of directly immobilizing small molecules to solid phase, they are coupled to biotin. Such biotin-small molecule conjugates can be fully verified chemically and evaluated in relevant phenotypic assays to ensure that the chemical modifications allowed for retention of the biological activity of the small molecules. Some of the known examples of small molecules whose targets were identified using this approach include norepinephrine (β-adrenergic receptor),[1] FK506 (FKBP),[2] trapoxin (histone deacetylase),[3] TNP-470 (Methionine aminopeptidase 2),[4,5] thalidomide (cereblon or CRBN),[6] pateamine A (eukaryotic initiation factor or eIF4A),[7,8] methyl-gerfelin (glyoxalase I),[9] and spliceostatin A (SF3b involved in splicing).[10]

In conventional affinity chromatography experiments, the putative binding protein(s) is first detected on an SDS–PAGE gel upon silver staining before the protein band is excised for identification by mass spectrometry. The recent development of quantitative proteomics is changing the way target proteins are identified. Over the last decade, much effort has been made to develop methods for identification of protein–protein interaction on a genome wide scale.[11,12] A major goal was to better understand the functions of proteins present in various genomes. As a result, new methods for identifying specific interacting proteins that are cumulatively called quantitative proteomics were developed.[13–15] It started with efforts to identify protein interaction partners using affinity chromatography. But instead of small molecules, antibodies were attached to the resin to immunoprecipitate a specific protein along with its interacting partners that would be identified using mass spectrometry. With the development of more sensitive and more accurate mass spectrometers, it became possible to identify hundreds of proteins in a given experiment. Many of these proteins are not even visible on the gel. The problem is that the majority of those proteins bound non-specifically to the resin and since mass spectrometry is inherently non-quantitative it was difficult to determine which protein(s) bound specifically to the protein of interest. In order to solve this specificity problem, several groups developed methods to quantify the relative amounts of a particular protein bound to loaded and control resins. Two of the most popular quantitative mass spectrometry methods are iTRAQ and SILAC.[16,17] The theory behind both of those methods is that samples are labeled with different isotopes. The samples are mixed in a 1:1 ratio before the mixture is subject to MS analysis. The enrichment of any given protein can be detected by the ratio between peptides that have heavy vs light isotopes. After repeating the pull-down experiment, it is possible to obtain a list of specific
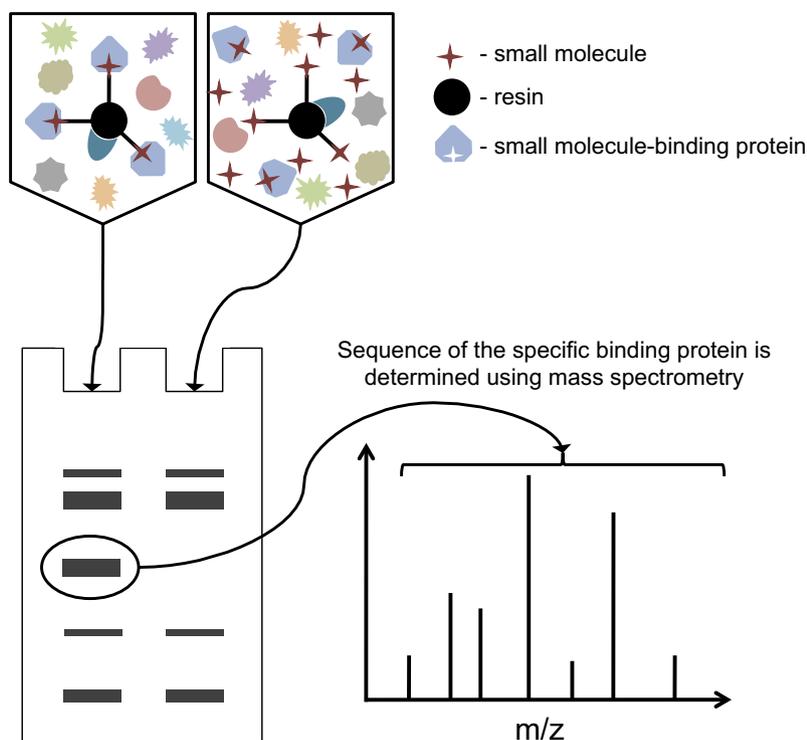
Figure 2. Schematic representation of protein target identification using affinity chromatography.

binding proteins based on that ratio for follow-up studies. Recently these techniques have been applied to target identification for small molecules and drugs.[18–21]

The second type of affinity-based methods relies on drug binding assays to guide target purification. The goal of this approach is to purify the binding protein (receptor) of a small molecule to homogeneity using drug binding assays to track the binding activity during protein chromatography and fractionation. A radioactively or fluorescently labeled small molecule such as a drug is usually incubated with a selected cell or tissue lysate and the protein–drug complex is separated from the free drug by various techniques including centrifugation, gel filtration, or absorption to charcoal.[22,23] Scintillation counting, spectrophotometry or other methods can then be applied to follow the protein fraction that binds the drug. This method is similar to that used for enzyme purification guided by enzymatic activity of a given protein or protein complex. When the putative binding protein is purified to near homogeneity, it can be subjected to identification by mass spectrometry. Binding specificity can be assessed by competition with excess unlabeled drug. K_d can also be determined by measuring the amount of drug bound to the protein at different drug and lysate/protein concentrations.[24] In fact, receptors for the majority of known drugs were identified using this method, including adrenergic receptors,[25–27] opioid receptors,[28,29] cyclophilins,[30] and FKBP12.[31] In many of these studies, affinity chromatography was used side by side with binding assays as one of the purification steps. Binding assays are essential for identification of appropriate lysis methods for optimal receptor solubilization and for identification of buffer conditions for maximal binding of the drug to its receptor. It is apparent that the optimal approach for target identification using these methods should consist of preparation of the extract enriched with target protein using binding assays followed by affinity chromatography in conjunction with quantitative mass spectrometry.

2.1.2. Genetic methods

The basic concept of genetic methods for target identification is to genetically perturb cells or whole organisms by mutagenesis, protein overexpression or knockdown and select for clones that are resistant to a particular small molecule. Identification of the gene(s) that caused the resistance can often lead to the target of a small molecule, though one needs to be cautious when analyzing genetic resistance since there are multiple ways resistance can arise. The most common phenotype that is exploited in genetic methods is cell viability. In this section we will first describe target identification using yeast genetics and then discuss examples of genetic methods applied to mammalian cells.

In the classic genetic method of target identification, bacterial or yeast cells are mutagenized and clones that are resistant to a small molecule of interest are selected by screening (Fig. 3). For yeast and bacterial cells it is often possible to select resistant clones that arise from spontaneous mutations. Once resistant clones are isolated, genomic DNA can be isolated and used to prepare a library. The genomic DNA library is used to transform wild type cells followed by selection on the small molecule. The genomic DNA that confers resistance of wild type cells to the drug can be isolated and sequenced, which will lead to the identification of the mutation on a specific gene that gave rise to resistance. Further studies will be required to assess whether the newly identified mutant gene product is the direct protein target for the small molecule. Occasionally, this approach can also be used in reverse whereby the DNA from wild type strain that confers sensitivity to the resistant clone is identified.

An illuminating example of target identification using yeast genetics is the identification of the target of the immunosuppressive and antiproliferative natural product rapamycin (TOR),[32] whose mammaliam counterpart (mTOR) was later identified using affinity chromatography.[33,34] The ability of rapamycin to arrest yeast cell cycle in G1 phase made it possible to apply yeast genetics to identify its target. Rapamycin, like another immunosuppressant FK506, works by first binding to its receptor FK506 binding protein (FKBP or FKBP proline rotamase 1 (FPR1) in yeast).[2,35,36] This binding is necessary but not sufficient for its antiproliferative activity since cells without FPR1 are viable and completely resistant to rapamycin. Yeast cells can spontaneously become resistant
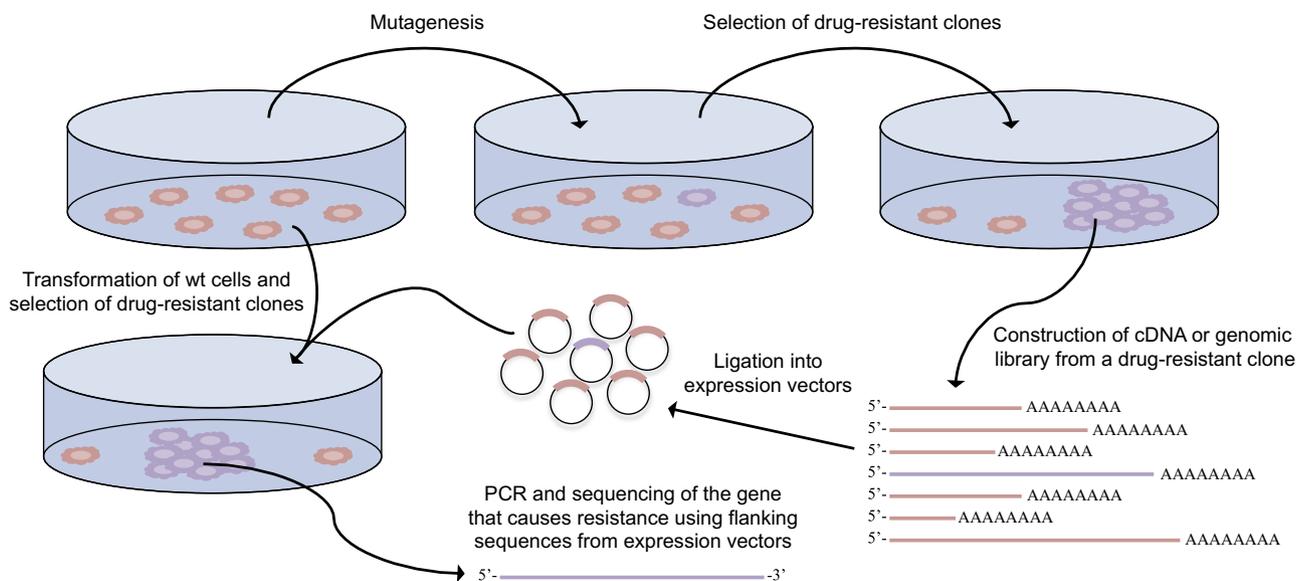
Figure 3. Schematic representation of protein target identification using genetic methods in yeast, bacterial or mammalian cells.

to rapamycin with a frequency of 10^{-7} so it was unnecessary to mutagenize cells in this case.[35] Heitman et al. isolated 18 independent spontaneous rapamycin-resistant strains. Using complementation analysis, they showed that 15 of 18 mutations were in FKBP gene and the other 3 mutations were in two genes that were named TOR1 and TOR2. Genomic libraries were prepared from rapamycin resistant strains as described above and the sequence of TOR1 and 2 genes was identified from the genomic fragments that caused rapamycin resistance in wild type strains.[37] Ensuing studies validated TOR1 and TOR2 as the targets of rapamycin.

In the last few years, genome-wide yeast knockout collections have become available where each gene is deleted in one strain and replaced by a DNA bar code to facilitate identification of the deletion mutants.[38,39] These knockout collections have been exploited to facilitate target identification as well. Several groups have used those collections to identify targets of some known and unknown drugs. There are two different approaches used. The first approach is to use a collection of haploinsufficiency diploid yeast strains in which one copy of a gene is deleted. The underlying principle for this approach is that cells that express half the amount of the target protein for a given small molecule will become hypersensitive to the molecule in comparison to the wild type strain. This approach was validated on a number of small molecules with known targets.[40] The key to success is to choose a concentration of the drug around or slightly lower than the IC_{50} so that the difference in sensitivity between wild type and deletion strains could be easily observed. Deletion strains can either be assayed separately using OD_{600} measurements in a plate reader[41] or they can be pooled and then the sensitive strains can be identified by the decrease in the bar code amount for a particular strain using microarrays.[42,43] Sphingolipid biosynthesis was identified as the target of a potential anticancer agent dihydromotuporamine C using these methods.[41] The second complementary approach is to use a haploid yeast deletion collection to screen for drug sensitivity.[44] The downside to this approach is that it is unlikely to yield a direct target of the drug since the complete deletion of the target protein is expected to produce non-viable yeast cells and will be absent from the collection. Nevertheless, this approach may yield information on genes that are functionally connected to the target of the small molecule under scrutiny. Similarly, gene overexpression screens in yeast can be used to identify targets of small molecules based on the expectation that cells with overexpressed target will become partially resistant to the small molecules.[45,46]

Given that many essential biological pathways are conserved between yeast and mammals, it is likely that many biologically active small molecules will be active in both. The homolog of the target that is identified in yeast can then be confirmed as a target in the corresponding organism where the small molecule of interest is also active. However, the use of yeast genetics for target identification does have its limitations. A number of mammalian cellular pathways are absent in yeast, which is evident from the much larger genome and proteome of mammals than those of yeast. Thus, yeast genetics will not be applicable to those small molecules that affect processes unique to mammalian cells. In addition, some small molecules cannot be used in yeast due to their low uptake and/or high efflux in comparison to mammalian cells.

Although less common, similar genetic approaches have been used with mammalian cells (Fig. 3). Mammalian cells can be readily mutagenized to select for clones that are resistant to a particular small molecule.[47–49] The advantage of using chemical mutagenesis is that dominant mutants can often be obtained using this method. Usually cells are treated with ethyl methanesulfonate (EMS), washed and allowed to proliferate for one cell cycle, as replication is required for mutations to incorporate into the genome. Then the cells are treated with a high concentration (10–1000 times higher then IC_{50}) of the small molecule of interest for days or even weeks to allow resistant cells to expand to form small colonies. Such drug-resistant colonies from different plates are then isolated and expanded in the presence of the drug. The most difficult part of this method is to isolate the gene that causes the resistance. One method is functional expression cloning.[50–52] First, a cDNA library is prepared from the resistant mutant cell line and cloned into a retroviral expression vector. Then the wild type cell line is transduced with the retroviral library and colonies that become resistant to the drug are isolated as described above. The gene that causes the resistance is incorporated into the genome of the wild type cell and can be amplified using PCR with primers complimentary to retrovirus-derived flanking regions and then sequenced to identify the mutant protein that confers resistance. Further development of sequencing technologies and wider availability of genome-wide mammalian ORF libraries will make genetic methods of target identification in mammalian cells even more efficient in the years to come.

A 'neo-genetic' approach for identifying novel protein–protein interactions, the yeast two-hybrid system,[53] was adapted to develop the yeast three-hybrid system for target identification.[54] This

method is based on the ability of a heterodimer between a small molecule of interest and another small molecule with a known high-affinity protein receptor to induce the formation of a ternary complex between two hybrid proteins to activate gene transcription. The original system was developed using the steroid hormone dexamethasone and its receptor, the rat glucocorticoid receptor, as the existing pair of small molecule/receptor. A limitation of the original version of the system is the requirement for high affinity between small molecule of interest and the target. Subsequently, a new ligand–receptor pair, methotrexate–dihydrofolate reductase, was developed that showed higher sensitivity.[55] That system was successfully used to discover targets for inhibitors of protein kinases.[56] Most recently, an even more sensitive SNAP-tag-based yeast three-hybrid system was developed that allowed for the identification of an enzyme involved in tetrahydrobiopterin biosynthesis as the target for sulfasalazine and its metabolites.[57]

2.2. Top-down approach

In contrast to the bottom-up approach discussed above, the top-down approach entails a path of going from a phenotype to specific protein target by exploiting the existing knowledge of a given cellular process. In this section we will describe a framework for the top-down approach along with some useful methods and examples from the literature where this approach was successfully applied. Top-down approach always starts with a phenotype that a compound of interest causes in cells or in organisms. All bioactive small molecules of course have some phenotype (e.g., cell proliferation inhibition). The goal is to go from the global phenotype in a step-wise fashion to eventually arrive at the protein target. A number of successful examples of the top-down target identification have been reported in the literature, including the identification of tubulin as the target of the anticancer drug taxol,[58,59] Sec61 complex as the target of a cyclodepsipeptide natural product that inhibits protein secretion,[60] and kinesin Eg5 as the target of a mitosis inhibitor named monastrol[61], to name a few. Described in more details below are three examples in this category. As our knowledge of biology increases, the top-down approach will become more and more effective.

The top-down approach was used successfully for the elucidation of the mechanism of action of a transcription inhibitor 5,6-dichloro-β-D-ribofuranosylbenzimidazole (DRB). In the late 1970s, DRB was shown to be a selective inhibitor of hnRNA and mRNA synthesis.[62–64] It inhibited hnRNA synthesis only by 60–70% but completely (>95%) inhibited cytoplasmic poly(A) containing RNA. After only several minutes of exposure to the drug, cells preferentially incorporated radioactively labeled nucleotides into long hnRNAs compared to untreated cells.[63] This was taken as an important clue that DRB is an inhibitor of transcription initiation rather than elongation. These in vivo observations were later confirmed using in vitro transcription assays where DRB was found to inhibit RNA polymerase II-driven transcription at concentrations similar to those used in vivo without affecting purified RNA polymerase II activity.[65–67] A more detailed investigation of the effects of DRB on in vitro transcription using pulse-chase assays showed that DRB inhibited the synthesis of full-length transcription only when added during the synthesis of the first 500 nucleotides or less and has no effect on elongating RNA polymerase II complex after it synthesized more than 500 nucleotides.[68,69] Furthermore, DRB had no effect on synthesis of RNA less than 100 nucleotides. This effect was explained by the observation that RNA polymerase II undergoes pausing soon after transcription initiation and the activity of positive transcription elongation factor (p-TEF) is required for the formation of productive elongation complex. This activity is inhibited by DRB. Now that the activity that is inhibited by DRB was identified, it was possible to fractionate cell extracts using this

activity as readout. pTEFb was purified this way and was shown to contain 2 subunits – Cdk9 and cyclin T.[70–72] DRB inhibited Cdk9 kinase activity at concentrations similar to those required for its effects on in vitro and in vivo transcription. Cdk9 was later shown to promote productive elongation by phosphorylating RNA polymerase II C-terminal domain and several negative regulators of elongation.[71,73] Thus, the top-down approach not only revealed the molecular target of DRB as Cdk9, but also led to identification of a new mechanism for transcription regulation.

Another bioactive small molecule whose mechanism was elucidated using the top-down approach is brefeldin A (BFA). In late 1980s, BFA was identified as a potent inhibitor of protein secretion.[74] Using careful pulse-chase and electron microscopy experiments, BFA was shown to inhibit transport of proteins from ER to Golgi and cause the disassembly of the Golgi complex.[74,75] Time course experiments showed that Golgi complex with its resident proteins is absorbed into the ER within 10 min after BFA treatment.[76,77] Disassembly of the Golgi complex proceeds through a dramatic intermediate state with long tubular projections spread throughout cytoplasm that is visible as soon as 2 min after BFA treatment. The absorption of Golgi complex into ER happens because BFA only inhibits transport of proteins from ER to Golgi but does not affect microtubule-dependent retrograde transport of proteins from Golgi to ER. Remarkably, all these effects are completely reversed within 20 min after BFA is washed out. BFA also causes the fusion of early endosomes and trans-Golgi network through mechanism similar to fusion of Golgi complex and ER.[78,79] The first clue to the biochemical mechanism of action of BFA came from a serendipitous observation that BFA promoted dissociation of a 110-kDa peripheral membrane protein from the Golgi apparatus within 30 s after treatment.[80] This process was inhibited by pre-incubation with GTPγS suggesting the participation of GTP binding proteins.[81] The 110-kDa protein was later found to be one of the major coat protein subunits of coated vesicles in Golgi and renamed β-COP.[82] BFA was shown to block the formation of non-clathrin coated vesicle from Golgi using a cell-free assay with purified Golgi.[83] This observation provided the first molecular understanding of BFA action and further studies focused on explaining the mechanism of vesicle formation inhibition by BFA and its dependence on GTP. At about the same time, several groups reported identification of ADP-ribosylation factor (ARF) small GTPase as a component of the coat of Golgi-derived vesicles that was required for β-COP binding to Golgi memberanes in GTP dependent manner.[84–86] Investigation of the effect of BFA on ARF showed that BFA inhibited ARF binding to Golgi membrane through inhibition of GDP–GTP exchange catalyzed by an unknown ARF GTP exchange factor (ARF GEF) present in Golgi fractions.[87,88] ARF GEF that is inhibited by BFA was purified using protein chromatography coupled with assay for BFA-inhibited ARF GEF activity.[89–91] It turned out that BFA is able to inhibit a number of ARF GEFs (BIG1, BIG2 and GBF1) that contain a common Sec7 domain. Independently, Sec7-containing yeast proteins GEA1p, GEA2p and Sec7p were later identified as targets of BFA using yeast genetics,[92,93] supporting the observations made in mammalian cells.

Most recently, the molecular target for triptolide, an anti-proliferative and immunosuppressive natural product isolated from a Chinese medicinal plant, was identified using the top-down approach as well (Fig. 4).[94] Since its initial isolation and identification in the early 1970s,[95] the mechanism of action of triptolide has attracted the attention of chemists and biologists. This interest was intensified when it was reported to possess potent anticancer activity through screening of the NCI 60 cancer cell lines. Using the 'bottom-up' approach with a [³H]-labeled triptolide, one group discovered a calcium channel polycystin-2 as a possible target[96] and another group detected the covalent binding of a 90-kD nuclear protein that bound to triptolide.[97] Unfortunately, polycystin-2

```
┌─────────────────────────────────────────┐
│  Transcription, translation, replication? │
└─────────────────────────────────────────┘
                    ▼
        ┌───────────────────────┐
        │     Transcription      │
        └───────────────────────┘
                    ▼
┌─────────────────────────────────────────┐
│ Pol I, Pol II or Pol III-dependent        │
│ transcription?                            │
└─────────────────────────────────────────┘
                    ▼
        ┌───────────────────────┐
        │    Pol II-dependent    │
        └───────────────────────┘
                    ▼
┌─────────────────────────────────────────┐
│ Pol II, TFIIA, TFIIB, TFIID, TFIIE        │
│ or TFIIH?                                 │
└─────────────────────────────────────────┘
                    ▼
              ┌──────────┐
              │  TFIIH   │
              └──────────┘
                    ▼
┌─────────────────────────────────────────┐
│ XPB, XPD, CDK7, p44 or other subunits?    │
└─────────────────────────────────────────┘
                    ▼
               ┌───────┐
               │  XBP  │
               └───────┘
```
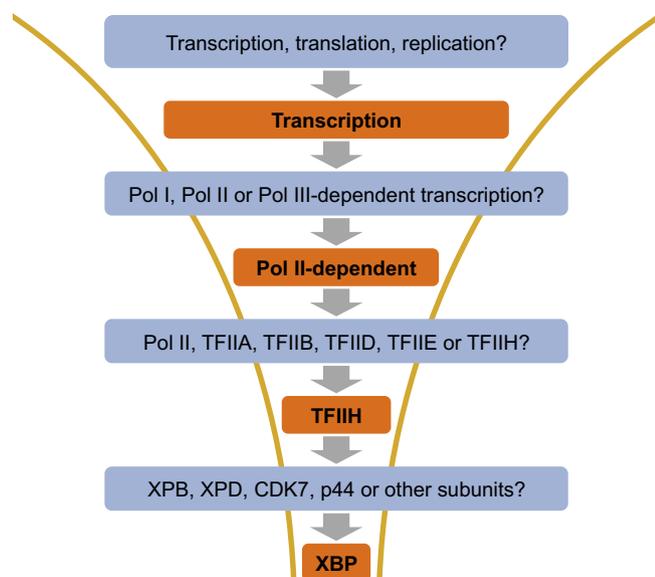
Figure 4. Top-down approach to identify XPB as the molecular target for triptolide.

cannot account for the general anti-proliferative activity of triptolide. Having detected the covalent binding of the 90-kD protein from cell lysate, it is theoretically possible to purify the putative triptolide target using [³H]-triptolide binding as an assay. But that did not happen. On a different front, the extensive synthetic efforts led to a large number of triptolide analogs and suggested that the C14 hydroxyl group is suitable for tethering an affinity probe. Affinity probes attached to the C14 OH group led to the isolation of dCTP pyrophosphatase, which does not seem to be relevant to the anti-proliferative activity of triptolide either.[98]

We took a top-down approach by beginning with a clue that triptolide may affect transcription.[99,100] We began by systematically examining the effects of triptolide on the synthesis of DNA, RNA and proteins in cancer cell lines. We quickly found that triptolide preferentially affected cellular RNA synthesis, corroborating with earlier findings. Fortunately, the mechanism and the key regulatory proteins of transcription initiation have been well characterized and there are a number of reconstituted in vitro transcription assays with defined protein components. Using those assays, we found that triptolide appeared to specifically affect TFIIH-mediated transcriptional initiation and nucleotide excision repair (NER). Four of the ten subunits of TFIIH possess defined enzymatic activities. A systematic examination of those activities narrowed the target to XPB, a DNA-dependent ATPase that is required for unwinding the double stranded DNA in the promoter region and around damaged nucleotides to allow RNAPII-dependent transcription initiation and NER to proceed. Subsequent experiments verified XPB as a physiologically relevant target for triptolide. It is interesting to note that for reasons that remain unknown, we are still unable to detect binding of affinity probes of triptolide by modifying triptolide at the C14 OH group to XPB both in vitro and in cell culture, highlighting the limitation of any given approach for target identification.

3. Target validation

Target validation is a crucial step in verifying the physiological relevance of newly identified small molecule targets. It is also a process during which new biological insights can be gained on the function of the target protein. Just because a small molecule binds to a particular protein or a mutation in a particular gene

causes resistance does not guarantee that the protein is the physiological target of the small molecule. Necessary and sufficient criteria for target validation include demonstration that a small molecule binds directly to the target protein and that expression of the functional target protein that is deficient in binding to the small molecule leads to the resistance of cells or organisms to the small molecule. In reality, very few targets meet all these criteria, underscoring the difficulty of target validation. We will briefly review two commonly used methods for target validation here.

3.1. Correlation between small molecule-target interaction and cellular activity

Correlation analysis is a most versatile and general method of target validation. It can be applied to virtually any target. The goal of the correlation analysis is to generate analogs of small molecule of interest that will have different affinity for the target protein and then determine if there is a good correlation between the affinity for the target protein and the potency of causing the corresponding phenotype. For example, if one analog is 10 times more potent at binding to a target protein then it should be 10 times more potent at causing the corresponding phenotype. If there is such a correlation, then it is strong evidence that this phenotype is caused by binding of the small molecule to the target protein. There are examples in the literature where correlation method is used for target validation but the correlation that is observed is 10 to 1 instead of 1 to 1, which in fact is the evidence against the putative protein as a relevant target of the small molecule. The key to a successful application of this method is to synthesize analogs of the bioactive small molecule that will have wide range of IC_{50} values spanning at least 3 orders of magnitude. It is also important to make compounds with IC_{50} values that are more or less equally distributed over those 3 orders of magnitude of activity and not concentrated around one particular value. Correlation analysis was successfully used for target validation of a number of bioactive small molecules including triptolide (XPB)[94] and fumagilin (MetAP2).[4] This method was used in 1970s to show that specific binding of 15 morphine analogs to homogenates of guinea pig intestine longitudinal muscle perfectly correlated (slope 0.98, $r = 0.97$) with their ability to inhibit electrically induced contraction of the guinea pig intestine.[101] This observation helped to solidify the proposition that binding to opioid receptors led to biological activity of morphine analogs and paved the way to the eventual isolation of these receptors.

3.2. Genetic approaches

If a bioactive small molecule causes a cellular phenotype by binding to a specific protein and (often) causes loss of its function, it is expected that (1) overexpression of the target protein will confer resistance to the small molecule; (2) deletion of the target protein will cause the same cellular phenotype as the treatment of cells with the small molecule and reduction of its expression will cause hypersensitivity to the small molecules; (3) mutants of the target protein that no longer bind to the small molecule but retain cellular activity of wild type protein will confer resistance to the small molecule. The genetic approaches of target validation entail experiments to address all three aforementioned aspects. Ideally, a complete validation of target should meet all three expectations.

The easiest and most widely used genetic method of target validation is to overexpress, knock down or less often, knock out the target protein. It is sometimes possible to observe resistance or sensitization by several orders of magnitude for the IC_{50} values for small molecules using those perturbations. Often, resistance or sensitization cannot be achieved because it is impossible to efficiently overexpress (e.g., protein is part of multi-subunit complex)

or knock out (e.g., protein is essential) the target protein. Another caveat of this approach is that overexpression or knockdown of any drug binding protein (not necessarily the physiological target) will lead to partial resistance or sensitivity to that drug, making it difficult to draw conclusions. Therefore, other methods of target validation should be used unless dramatic resistance or sensitization to the small molecule of interest is achieved.

Another genetic approach to target validation is to generate a mutant protein that is resistant to small molecule binding but retains its normal function. This method can lead to a very convincing validation of a physiological target and is comparable to correlation analysis in that regard. However, it is often difficult to generate a mutant target protein that is defective in small molecule binding while remaining active. This approach can be significantly facilitated if genetic methods of target identification were used and the mutations that perturb binding are already known.

A good example of the use of a genetic method for target validation is the identification of cereblon (CRBN) as the target of thalidomide's teratogenic activity.[6] Thalidomide was widely used as a sedative in the 1960s and was often prescribed to pregnant women. Before its teratogenic effect was discovered, more than 10,000 affected children were born from mothers taking thalidomide.[102,103] CRBN is a substrate receptor of E3 ubiquitin ligase complex that also contains DDB1, Cul4A and Roc1. Researchers generated Y384A and W386A double mutant of CRBN (CRBN$^{YW/AA}$) that was deficient for thalidomide binding by systematically mutating conserved residues in the C-terminal 104 amino acid long thalidomide-binding region of CRBN. To prove that CRBN is the physiological target of thalidomide's teratogenicity, chicks and zebrafish were used as models to assess thalidomide's teratogenicity. Overexpression of CRBN$^{YW/AA}$ but not wild type CRBN in chicks and zebrafish led to a complete resistance to teratogenic effect of thalidomide in those organisms. In addition, it was shown that knockdown of CRBN in zebrafish also led to similar defects as thalidomide treatment further validating it as a physiological target.

4. Conclusion

In this review, we highlighted a number of approaches for identification and validation of protein targets of bioactive small molecules. Each method has its own strengths and weaknesses. The choice of an appropriate subset of methods for target identification and validation will depend on the chemistry and the cellular phenotype associated with the small molecule. With the advent and improvement of target identification and validation methods, the bottle-neck of target identification in chemical biology will continue to be widened and eventually cease to exist, paving the way for small molecules to become more widely used in biology.

References and notes

1. Lefkowitz, R. J.; Haber, E.; O'Hara, D. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 2828.
2. Harding, M. W.; Galat, A.; Uehling, D. E.; Schreiber, S. L. *Nature* **1989**, *341*, 758.
3. Taunton, J.; Hassig, C. A.; Schreiber, S. L. *Science* **1996**, *272*, 408.
4. Griffith, E. C.; Su, Z.; Turk, B. E.; Chen, S.; Chang, Y. H.; Wu, Z.; Biemann, K.; Liu, J. O. *Chem. Biol.* **1997**, *4*, 461.
5. Sin, N.; Meng, L.; Wang, M. Q.; Wen, J. J.; Bornmann, W. G.; Crews, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6099.
6. Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. *Science* **2010**, *327*, 1345.
7. Low, W. K.; Dang, Y.; Schneider-Poetsch, T.; Shi, Z.; Choi, N. S.; Merrick, W. C.; Romo, D.; Liu, J. O. *Mol. Cell* **2005**, *20*, 709.
8. Bordeleau, M. E.; Matthews, J.; Wojnar, J. M.; Lindqvist, L.; Novac, O.; Jankowsky, E.; Sonenberg, N.; Northcote, P.; Teesdale-Spittle, P.; Pelletier, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10460.
9. Kawatani, M.; Okumura, H.; Honda, K.; Kanoh, N.; Muroi, M.; Dohmae, N.; Takami, M.; Kitagawa, M.; Futamura, Y.; Imoto, M.; Osada, H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11691.
10. Kaida, D.; Motoyoshi, H.; Tashiro, E.; Nojima, T.; Hagiwara, M.; Ishigami, K.; Watanabe, H.; Kitahara, T.; Yoshida, T.; Nakajima, H.; Tani, T.; Horinouchi, S.; Yoshida, M. *Nat. Chem. Biol.* **2007**, *3*, 576.
11. Uetz, P.; Giot, L.; Cagney, G.; Mansfield, T. A.; Judson, R. S.; Knight, J. R.; Lockshon, D.; Narayan, V.; Srinivasan, M.; Pochart, P.; Qureshi-Emili, A.; Li, Y.; Godwin, B.; Conover, D.; Kalbfleisch, T.; Vijayadamodar, G.; Yang, M.; Johnston, M.; Fields, S.; Rothberg, J. M. *Nature* **2000**, *403*, 623.
12. Gavin, A. C.; Bosche, M.; Krause, R.; Grandi, P.; Marzioch, M.; Bauer, A.; Schultz, J.; Rick, J. M.; Michon, A. M.; Cruciat, C. M.; Remor, M.; Hofert, C.; Schelder, M.; Brajenovic, M.; Ruffner, H.; Merino, A.; Klein, K.; Hudak, M.; Dickson, D.; Rudi, T.; Gnau, V.; Bauch, A.; Bastuck, S.; Huhse, B.; Leutwein, C.; Heurtier, M. A.; Copley, R. R.; Edelmann, A.; Querfurth, E.; Rybin, V.; Drewes, G.; Raida, M.; Bouwmeester, T.; Bork, P.; Seraphin, B.; Kuster, B.; Neubauer, G.; Superti-Furga, G. *Nature* **2002**, *415*, 141.
13. Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 49.
14. Schulze, W. X.; Usadel, B. *Annu. Rev. Plant Biol.* **2010**, *61*, 491.
15. Cox, J.; Mann, M. *Annu. Rev. Biochem.* **2011**, *80*, 273.
16. Miyagi, M.; Rao, K. C. *Mass Spectrom. Rev.* **2007**, *26*, 121.
17. Ong, S. E.; Kratchmarova, I.; Mann, M. *J. Proteome Res.* **2003**, *2*, 173.
18. Ong, S. E.; Schenone, M.; Margolin, A. A.; Li, X.; Do, K.; Doud, M. K.; Mani, D. R.; Kuai, L.; Wang, X.; Wood, J. L.; Tolliday, N. J.; Koehler, A. N.; Marcaurelle, L. A.; Golub, T. R.; Gould, R. J.; Schreiber, S. L.; Carr, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4617.
19. Fabian, M. A.; Biggs, W. H., 3rd; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lelias, J. M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. *Nat. Biotechnol.* **2005**, *23*, 329.
20. Bantscheff, M.; Eberhard, D.; Abraham, Y.; Bastuck, S.; Boesche, M.; Hobson, S.; Mathieson, T.; Perrin, J.; Raida, M.; Rau, C.; Reader, V.; Sweetman, G.; Bauer, A.; Bouwmeester, T.; Hopf, C.; Kruse, U.; Neubauer, G.; Ramsden, N.; Rick, J.; Kuster, B.; Drewes, G. *Nat. Biotechnol.* **2007**, *25*, 1035.
21. Huang, S. M.; Mishina, Y. M.; Liu, S.; Cheung, A.; Stegmeier, F.; Michaud, G. A.; Charlat, O.; Wiellette, E.; Zhang, Y.; Wiessner, S.; Hild, M.; Shi, X.; Wilson, C. J.; Mickanin, C.; Myer, V.; Fazal, A.; Tomlinson, R.; Serluca, F.; Shao, W.; Cheng, H.; Shultz, M.; Rau, C.; Schirle, M.; Schlegl, J.; Ghidelli, S.; Fawell, S.; Lu, C.; Curtis, D.; Kirschner, M. W.; Lengauer, C.; Finan, P. M.; Tallarico, J. A.; Bouwmeester, T.; Porter, J. A.; Bauer, A.; Cong, F. *Nature* **2009**, *461*, 614.
22. Qume, M. *Methods Mol. Biol.* **1999**, *106*, 3.
23. Hulme, E. C.; Trevethick, M. A. *Br. J. Pharmacol.* **2010**, *161*, 1219.
24. Hulme, E. C. *Methods Mol. Biol.* **1999**, *106*, 139.
25. Benovic, J. L.; Shorr, R. G.; Caron, M. G.; Lefkowitz, R. J. *Biochemistry* **1984**, *23*, 4510.
26. Shorr, R. G.; Strohsacker, M. W.; Lavin, T. N.; Lefkowitz, R. J.; Caron, M. G. *J. Biol. Chem.* **1982**, *257*, 12341.
27. Shorr, R. G.; Lefkowitz, R. J.; Caron, M. G. *J. Biol. Chem.* **1981**, *256*, 5820.
28. Cho, T. M.; Hasegawa, J.; Ge, B. L.; Loh, H. H. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 4138.
29. Nakayama, H.; Shikano, H.; Aoyama, T.; Amano, T.; Kanaoka, Y. *FEBS Lett.* **1986**, *208*, 278.
30. Handschumacher, R. E.; Harding, M. W.; Rice, J.; Drugge, R. J.; Speicher, D. W. *Science* **1984**, *226*, 544.
31. Siekierka, J. J.; Hung, S. H.; Poe, M.; Lin, C. S.; Sigal, N. H. *Nature* **1989**, *341*, 755.
32. Heitman, J.; Movva, N. R.; Hall, M. N. *Science* **1991**, *253*, 905.
33. Sabatini, D. M.; Erdjument-Bromage, H.; Lui, M.; Tempst, P.; Snyder, S. H. *Cell* **1994**, *78*, 35.
34. Brown, E. J.; Albers, M. W.; Shin, T. B.; Ichikawa, K.; Keith, C. T.; Lane, W. S.; Schreiber, S. L. *Nature* **1994**, *369*, 756.
35. Koltin, Y.; Faucette, L.; Bergsma, D. J.; Levy, M. A.; Cafferkey, R.; Koser, P. L.; Johnson, R. K.; Livi, G. P. *Mol. Cell. Biol.* **1991**, *11*, 1718.
36. Bierer, B. E.; Mattila, P. S.; Standaert, R. F.; Herzenberg, L. A.; Burakoff, S. J.; Crabtree, G.; Schreiber, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9231.
37. Kunz, J.; Henriquez, R.; Schneider, U.; Deuter-Reinhard, M.; Movva, N. R.; Hall, M. N. *Cell* **1993**, *73*, 585.
38. Winzeler, E. A.; Shoemaker, D. D.; Astromoff, A.; Liang, H.; Anderson, K.; Andre, B.; Bangham, R.; Benito, R.; Boeke, J. D.; Bussey, H.; Chu, A. M.; Connelly, C.; Davis, K.; Dietrich, F.; Dow, S. W.; El Bakkoury, M.; Foury, F.; Friend, S. H.; Gentalen, E.; Giaever, G.; Hegemann, J. H.; Jones, T.; Laub, M.; Liao, H.; Liebundguth, N.; Lockhart, D. J.; Lucau-Danila, A.; Lussier, M.; M'Rabet, N.; Menard, P.; Mittmann, M.; Pai, C.; Rebischung, C.; Revuelta, J. L.; Riles, L.; Roberts, C. J.; Ross-MacDonald, P.; Scherens, B.; Snyder, M.; Sookhai-Mahadeo, S.; Storms, R. K.; Veronneau, S.; Voet, M.; Volckaert, G.; Ward, T. R.; Wysocki, R.; Yen, G. S.; Yu, K.; Zimmermann, K.; Philippsen, P.; Johnston, M.; Davis, R. W. *Science* **1999**, *285*, 901.
39. Giaever, G.; Chu, A. M.; Ni, L.; Connelly, C.; Riles, L.; Veronneau, S.; Dow, S.; Lucau-Danila, A.; Anderson, K.; Andre, B.; Arkin, A. P.; Astromoff, A.; El-Bakkoury, M.; Bangham, R.; Benito, R.; Brachat, S.; Campanaro, S.; Curtiss, M.; Davis, K.; Deutschbauer, A.; Entian, K. D.; Flaherty, P.; Foury, F.; Garfinkel, D. J.; Gerstein, M.; Gotte, D.; Guldener, U.; Hegemann, J. H.; Hempel, S.; Herman, Z.; Jaramillo, D. F.; Kelly, D. E.; Kelly, S. L.; Kotter, P.; LaBonte, D.; Lamb, D. C.; Lan, N.; Liang, H.; Liao, H.; Liu, L.; Luo, C.; Lussier, M.; Mao, R.; Menard, P.; Ooi, S. L.; Revuelta, J. L.; Roberts, C. J.; Rose, M.; Ross-Macdonald, P.; Scherens, B.; Schimmack, G.; Shafer, B.; Shoemaker, D. D.; Sookhai-Mahadeo, S.; Storms, R. K.; Strathern, J. N.; Valle, G.; Voet, M.; Volckaert, G.; Wang, C. Y.; Ward, T. R.; Wilhelmy, J.; Winzeler, E. A.; Yang, Y.; Yen, G.; Youngman, E.; Yu, K.; Bussey,

H.; Boeke, J. D.; Snyder, M.; Philippsen, P.; Davis, R. W.; Johnston, M. *Nature* **2002**, *418*, 387.

40. Giaever, G.; Shoemaker, D. D.; Jones, T. W.; Liang, H.; Winzeler, E. A.; Astromoff, A.; Davis, R. W. *Nat. Genet.* **1999**, *21*, 278.
41. Baetz, K.; McHardy, L.; Gable, K.; Tarling, T.; Rebérioux, D.; Bryan, J.; Andersen, R. J.; Dunn, T.; Hieter, P.; Roberge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4525.
42. Giaever, G.; Flaherty, P.; Kumm, J.; Proctor, M.; Nislow, C.; Jaramillo, D. F.; Chu, A. M.; Jordan, M. I.; Arkin, A. P.; Davis, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 793.
43. Lum, P. Y.; Armour, C. D.; Stepaniants, S. B.; Cavet, G.; Wolf, M. K.; Butler, J. S.; Hinshaw, J. C.; Garnier, P.; Prestwich, G. D.; Leonardson, A.; Garrett-Engele, P.; Rush, C. M.; Bard, M.; Schimmack, G.; Phillips, J. W.; Roberts, C. J.; Shoemaker, D. D. *Cell* **2004**, *116*, 121.
44. Parsons, A. B.; Lopez, A.; Givoni, I. E.; Williams, D. E.; Gray, C. A.; Porter, J.; Chua, G.; Sopko, R.; Brost, R. L.; Ho, C. H.; Wang, J.; Ketela, T.; Brenner, C.; Brill, J. A.; Fernandez, G. E.; Lorenz, T. C.; Payne, G. S.; Ishihara, S.; Ohya, Y.; Andrews, B.; Hughes, T. R.; Frey, B. J.; Graham, T. R.; Andersen, R. J.; Boone, C. *Cell* **2006**, *126*, 611.
45. Luesch, H.; Wu, T. Y.; Ren, P.; Gray, N. S.; Schultz, P. G.; Supek, F. *Chem. Biol.* **2005**, *12*, 55.
46. Hoon, S.; Smith, A. M.; Wallace, I. M.; Suresh, S.; Miranda, M.; Fung, E.; Proctor, M.; Shokat, K. M.; Zhang, C.; Davis, R. W.; Giaever, G.; St Onge, R. P.; Nislow, C. *Nat. Chem. Biol.* **2008**, *4*, 498.
47. Bartolomei, M. S.; Corden, J. L. *Mol. Cell. Biol.* **1987**, *7*, 586.
48. Funanage, V. L. *Mol. Cell. Biol.* **1982**, *2*, 467.
49. Yan, J. P.; Colon, M. E.; Beebe, L. A.; Melancon, P. *J. Cell Biol.* **1994**, *126*, 65.
50. Kitamura, T.; Koshino, Y.; Shibata, F.; Oki, T.; Nakajima, H.; Nosaka, T.; Kumagai, H. *Exp. Hematol.* **2003**, *31*, 1007.
51. Whitehead, I.; Kirk, H.; Kay, R. *Mol. Cell. Biol.* **1995**, *15*, 704.
52. Shen, D. W.; Liang, X. J.; Suzuki, T.; Gottesman, M. M. *Mol. Pharmacol.* **2006**, *69*, 1383.
53. Fields, S.; Song, O. *Nature* **1989**, *340*, 245.
54. Licitra, E. J.; Liu, J. O. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12817.
55. Baker, K.; Bleczinski, C.; Lin, H.; Salazar-Jimenez, G.; Sengupta, D.; Krane, S.; Cornish, V. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16537.
56. Becker, F.; Murthi, K.; Smith, C.; Come, J.; Costa-Roldan, N.; Kaufmann, C.; Hanke, U.; Degenhart, C.; Baumann, S.; Wallner, W.; Huber, A.; Dedier, S.; Dill, S.; Kinsman, D.; Hediger, M.; Bockovich, N.; Meier-Ewert, S.; Kluge, A. F.; Kley, N. *Chem. Biol.* **2004**, *11*, 211.
57. Chidley, C.; Haruki, H.; Pedersen, M. G.; Muller, E.; Johnsson, K. *Nat. Chem. Biol.* **2011**, *7*, 375.
58. Schiff, P. B.; Fant, J.; Horwitz, S. B. *Nature* **1979**, *277*, 665.
59. Schiff, P. B.; Horwitz, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 1561.
60. Garrison, J. L.; Kunkel, E. J.; Hegde, R. S.; Taunton, J. *Nature* **2005**, *436*, 285.
61. Mayer, T. U.; Kapoor, T. M.; Haggarty, S. J.; King, R. W.; Schreiber, S. L.; Mitchison, T. J. *Science* **1999**, *286*, 971.
62. Tamm, I.; Hand, R.; Caliguiri, L. A. *J. Cell Biol.* **1976**, *69*, 229.
63. Sehgal, P. B.; Derman, E.; Molloy, G. R.; Tamm, I.; Darnell, J. E. *Science* **1976**, *194*, 431.
64. Sehgal, P. B.; Darnell, J. E., Jr.; Tamm, I. *Cell* **1976**, *9*, 473.
65. Zandomeni, R.; Mittleman, B.; Bunick, D.; Ackerman, S.; Weinmann, R. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 3167.
66. Zandomeni, R.; Bunick, D.; Ackerman, S.; Mittleman, B.; Weinmann, R. *J. Mol. Biol.* **1983**, *167*, 561.
67. Chodosh, L. A.; Fire, A.; Samuels, M.; Sharp, P. A. *J. Biol. Chem.* **1989**, *264*, 2250.
68. Marshall, N. F.; Price, D. H. *Mol. Cell. Biol.* **1992**, *12*, 2078.
69. Kephart, D. D.; Marshall, N. F.; Price, D. H. *Mol. Cell. Biol.* **1992**, *12*, 2067.
70. Marshall, N. F.; Price, D. H. *J. Biol. Chem.* **1995**, *270*, 12335.
71. Marshall, N. F.; Peng, J.; Xie, Z.; Price, D. H. *J. Biol. Chem.* **1996**, *271*, 27176.
72. Zhu, Y.; Pe'ery, T.; Peng, J.; Ramanathan, Y.; Marshall, N.; Marshall, T.; Amendt, B.; Mathews, M. B.; Price, D. H. *Genes Dev.* **1997**, *11*, 2622.
73. Wada, T.; Takagi, T.; Yamaguchi, Y.; Watanabe, D.; Handa, H. *EMBO J.* **1998**, *17*, 7395.
74. Misumi, Y.; Miki, K.; Takatsuki, A.; Tamura, G.; Ikehara, Y. *J. Biol. Chem.* **1986**, *261*, 11398.
75. Fujiwara, T.; Oda, K.; Yokota, S.; Takatsuki, A.; Ikehara, Y. *J. Biol. Chem.* **1988**, *263*, 18545.
76. Lippincott-Schwartz, J.; Donaldson, J. G.; Schweizer, A.; Berger, E. G.; Hauri, H. P.; Yuan, L. C.; Klausner, R. D. *Cell* **1990**, *60*, 821.
77. Lippincott-Schwartz, J.; Yuan, L. C.; Bonifacino, J. S.; Klausner, R. D. *Cell* **1989**, *56*, 801.
78. Wood, S. A.; Park, J. E.; Brown, W. J. *Cell* **1991**, *67*, 591.
79. Lippincott-Schwartz, J.; Yuan, L.; Tipper, C.; Amherdt, M.; Orci, L.; Klausner, R. D. *Cell* **1991**, *67*, 601.
80. Donaldson, J. G.; Lippincott-Schwartz, J.; Bloom, G. S.; Kreis, T. E.; Klausner, R. D. *J. Cell Biol.* **1990**, *111*, 2295.
81. Donaldson, J. G.; Lippincott-Schwartz, J.; Klausner, R. D. *J. Cell Biol.* **1991**, *112*, 579.
82. Serafini, T.; Stenbeck, G.; Brecht, A.; Lottspeich, F.; Orci, L.; Rothman, J. E.; Wieland, F. T. *Nature* **1991**, *349*, 215.
83. Orci, L.; Tagaya, M.; Amherdt, M.; Perrelet, A.; Donaldson, J. G.; Lippincott-Schwartz, J.; Klausner, R. D.; Rothman, J. E. *Cell* **1991**, *64*, 1183.
84. Stearns, T.; Willingham, M. C.; Botstein, D.; Kahn, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1238.
85. Serafini, T.; Orci, L.; Amherdt, M.; Brunner, M.; Kahn, R. A.; Rothman, J. E. *Cell* **1991**, *67*, 239.
86. Donaldson, J. G.; Cassel, D.; Kahn, R. A.; Klausner, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6408.
87. Donaldson, J. G.; Kahn, R. A.; Lippincott-Schwartz, J.; Klausner, R. D. *Science* **1991**, *254*, 1197.
88. Donaldson, J. G.; Finazzi, D.; Klausner, R. D. *Nature* **1992**, *360*, 350.
89. Morinaga, N.; Tsai, S. C.; Moss, J.; Vaughan, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12856.
90. Morinaga, N.; Moss, J.; Vaughan, M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12926.
91. Togawa, A.; Morinaga, N.; Ogasawara, M.; Moss, J.; Vaughan, M. *J. Biol. Chem.* **1999**, *274*, 12308.
92. Peyroche, A.; Paris, S.; Jackson, C. L. *Nature* **1996**, *384*, 479.
93. Peyroche, A.; Antonny, B.; Robineau, S.; Acker, J.; Cherfils, J.; Jackson, C. L. *Mol. Cell* **1999**, *3*, 275.
94. Titov, D. V.; Gilman, B.; He, Q. L.; Bhat, S.; Low, W. K.; Dang, Y.; Smeaton, M.; Demain, A. L.; Miller, P. S.; Kugel, J. F.; Goodrich, J. A.; Liu, J. O. *Nat. Chem. Biol.* **2011**, *7*, 182.
95. Kupchan, S. M.; Court, W. A.; Dailey, R. G., Jr.; Gilmore, C. J.; Bryan, R. F. *J. Am. Chem. Soc.* **1972**, *94*, 7194.
96. Leuenroth, S. J.; Okuhara, D.; Shotwell, J. D.; Markowitz, G. S.; Yu, Z.; Somlo, S.; Crews, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4389.
97. McCallum, C.; Kwon, S.; Leavitt, P.; Shen, D. M.; Liu, W.; Gurnett, A. *Immunobiology* **2007**, *212*, 549.
98. Corson, T. W.; Cavga, H.; Aberle, N.; Crews, C. M. *ChemBioChem* **2011**, *12*, 1767.
99. McCallum, C.; Kwon, S.; Leavitt, P.; Shoop, W.; Michael, B.; Felcetto, T.; Zaller, D.; O'Neill, E.; Frantz-Wattley, B.; Thompson, C.; Forrest, G.; Carballo-Jane, E.; Gurnett, A. *Therapy* **2005**, *2*, 261.
100. Vispe, S.; DeVries, L.; Creancier, L.; Besse, J.; Breand, S.; Hobson, D. J.; Svejstrup, J. Q.; Annereau, J. P.; Cussac, D.; Dumontet, C.; Guilbaud, N.; Barret, J. M.; Bailly, C. *Mol. Cancer Ther.* **2009**, *8*, 2780.
101. Creese, I.; Snyder, S. H. *J. Pharmacol. Exp. Ther.* **1975**, *194*, 205.
102. Speirs, A. L. *Lancet* **1962**, *1*, 303.
103. Taussig, H. B. *Sci. Am.* **1962**, *207*, 29.