Target Identification : A Challenging Step in Forward Chemical Genetics

8 authors, including:

Animesh Samanta
National University of Singapore
42 PUBLICATIONS   **1,076** CITATIONS

SEE PROFILE

Zhai Duanting
National University of Singapore
24 PUBLICATIONS   **800** CITATIONS

SEE PROFILE

Wang Wayne Xu
National University of Singapore
18 PUBLICATIONS   **546** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   UCNPs for bio imaging probe development View project

Project   Rapid detection View project

Target Identification: A Challenging Step in Forward Chemical Genetics

Raj Kumar Das[1], Animesh Samanta[1], Krishnakanta Ghosh[2], Duanting Zhai[2], Wang Xu[2], Dongdong Su[2], Cheryl Leong[3] and Young Tae Chang[1,2,4,*]

[1]Department of Chemistry, National University of Singapore, 117543, Singapore

[2]Department of Chemistry & MedChem Program of Life Sciences Institute, National University of Singapore, 117543, Singapore

[3]Graduate School for Integrative Sciences and Engineering, National University of Singapore, 119077, Singapore

[4]Singapore Bioimaging Consortium, Agency for Science, Technology and Research (A*STAR), 11 Biopolis Way 138667, Singapore

SYNOPSIS

Investigation of the genetic functions in complex biological systems is a challenging step in recent year. Hence, several valuable and interesting research projects have been developed with novel ideas to find out the unknown functions of genes or proteins. To validate the applicability of their novel ideas, various approaches are built up. To date, the most promising and commonly used approach for discovering the target proteins from biological system using small molecule is well known a forward chemical genetics which is considered to be more convenient than the classical genetics. Although, the forward chemical genetics consists of the three basic components, the target identification is the most challenging step to chemical biology researchers. Hence, the diverse target identification methods have been developed and adopted to disclose the small molecule bound protein. Herein, in this review, we briefly described the first two parts chemical toolbox and screening, and then the target identifications in forward chemical genetics are thoroughly described along with the illustrative real example case study. In the tabular form, the different biological active small molecules which are the successful examples of target identifications are accounted in this research review.
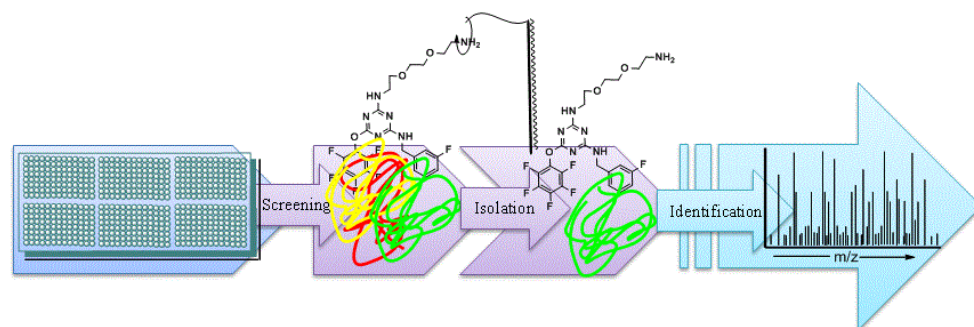
Keywords: chemical toolbox, small molecules, screening, cell imaging, phenotype change, genetic information, proteomes

Introduction

The completion of Human genome project[1], has led to the discovery of numerous novel genes. However in most cases, their functions in dynamic complex biological systems remain unknown. Therefore, illuminating the functions of these gene targets and the development of mechanistic pathways are poised to be the next challenging steps in the field of chemical genetics. Not surprisingly, the sheer amount of genetic information now available makes analysis a daunting task - a challenge that is currently being undertaken by the field of functional genomics[2]. The significant gap of functional genomics after identification of new pathways[3,4] and networks can be filled up by the emerging concept of chemical genetics[5]. Depending on the starting point of investigation, chemical genetics can be divided into two approaches[6]: forward chemical genetics[7] and reverse chemical genetics[8], which proceed "from effect to cause" and "from cause to effect", or "from phenotype to genotype" and "from genotype to phenotype"[9] respectively. Screening small-molecule libraries[10,11] for compounds that generate a phenotype of interest is an example of the forward chemical genetic approach whereas the reverse approach entails small molecules that target a single protein. Contrary to conventional genetics where gene knock-outs[12] or overexpression[13] are carried out, forward chemical genetics exploits the small-molecules to generate new phenotypes that are utilized in the explication of gene functions. Forward chemical genetics requires three basic components: a) a chemically diverse library of compounds; b) an assay, in which the library is screened for a cellular[14] or organismal[15]; phenotype and c) a method to trace an active compound to its biological target - otherwise known as target identification. It is this third component - target identification that remains one of the greatest challenges in chemical genetics.

Merit of Forward Chemical Genetics

In recent years, chemical genetics has increased in popularity due to the inducing capability of small molecules to modulate biological activity in reversible ways[16]. Although genetics can be applied for a better understanding of complex biological systems based on their capacity to modulate biological activity to a high degree of specificity, it does have its limitations. Direct induction of genes by sensitive mutation often raises the complicated situation to identify the effect of modulating gene products[17]. In genetics, delivery of gene is the most vital problem due to the issue of cell permeability. The generation of mutants one by one is also a tedious and time-consuming process. In mutation of genes, the importance of post translational modification or protein functions is not considered. Moreover, in mutation of genes, it is not feasible to confirm the reliable target for drugs and the over all process is also slow. In most cases, one protein has multiple functions, thus their modification (e.g. knockdown) does not lead to the expected effects. Moreover, other genetic methods, like antisense technology[18], mutagenesis or RNAi interference can be applicable at the level of the genes which cause some temporal or even permanent effects.

Keeping in mind these issues in genetics, small molecule induced chemical genetics especially forward chemical genetics has several advantages. For example, the biological effect of small molecules is typically rapid[19] and it allows for characterizing the instantaneous effects of small molecules. Moreover, all small molecules can be used to study the dynamic processes in a conditional manner[20]. For example, they can be applied at any time point in the experiment conditions with different concentration range to study critical genes at any developmental stage. Small molecules can also be used in multi cellular organisms[21] to see the phenotypic change *in vivo* systems. A successful forward genetic study offers the knowledge about a novel gene's function as well as a small molecule which has the potential to serve as an on/off switch to control biological processes[22]. Therefore, these small molecules can be useful biologic probes as well as potential new drug candidates.

Devices for Forward Chemical Genetics

1. Chemical toolbox

In the last few decades, various inventive chemical toolboxes have been developed for the study of numerous gene functions in system biology. The chemical genomics field is rapidly expanding and evolving to facilitate the discovery of different unknown functions of genes by the aid of chemical toolboxes[23]. Inspired by many successful attempts, research groups are racing to invent novel and effective functions of gene networks by applying diverse chemicals libraries. Identifying these new networks would significantly enhance the discovery of new drug molecules from this powerful chemical toolbox. This powerful technology, known as "chemical genomics" not only overcomes significant gaps in functional genomics but also has the potential for the discovery of novel drug-like small molecules. These small molecules which can be found from large libraries of compounds from different sources generally alter the functions of proteins after binding to the target proteins. The compounds are derived from natural plant[24] and animal sources subsequent extracted and synthesized into small molecules.

1a. The discovery of bioactive compounds

In system biology it is more interesting to choose the suitable novel chemical entities which are capable of modulating biological functions after binding to the target proteins. This unknown protein function is influenced by diverse drug-like small molecules which can be found from different commercial sources[25] such as ChemBridge Corporation, Maybridge Chemical, Thermo-Fisher Hit-Finder, ChemBridge MicroFormats, and Spectrum Collection etc.

1b. Combinatorial libraries

The large amount of chemical compounds can be synthesized in the laboratory by using basic scaffolds which may belong to different classes such as heterocycles, natural products, oligosaccharides, and fluorochromes. By applying the diversity-oriented design strategy[26], large collections of structurally diverse and complex compound sets are synthesized through parallel and mixture synthesis of small molecules by combinatorial chemistry. The approach is that a library with the same core scaffold, but with various diversity elements/branches directly attached around the core, may selectively respond to a broader range of analytes and thus show a greater likelihood to "hit" the target[27]. Numerous methods are available for creating such diversity, such as skeletal diversity[28,29], stereochemical diversity, and molecular property diversity[30,31].

1c. Natural product-like libraries

Natural products are considered as a vast source of chemical compounds by most industrial drug research organization where they widely utilized, explored and modified to construct better derivatives[32]. Many biologically validated drug developers would choose to generate libraries from natural product cores, which are in a way sensible and effective[33-35].

1d. Heterocyclic libraries

The majority of the heterocyclic compounds are widespread in various protein networks in the cellular context and built up for many cofactors of enzyme substrates, and also available as the main component of the DNA such as purine, pyrimidine bases. Hence, these are the highly desired structures for the development of new drug like small molecules[36]. In heterocyclic chemistry, cycloaddition and multicomponent reactions (MCRs) are mainly applied to construct a complex heterocyclic library[37] in DOS with different substitution. The number of hydrogen bond acceptors and donors in

heterocylic compounds are higher than normal chemicals and most of the time accomplishes the appropriate balance of hydrophilic and lipophilic characters which is suitable for bioavailability and membrane-permeability.

1e. Oligosaccharides libraries

The oligosaccharides are the main components of the glycoproteins and are known to be the extracellular segments of integral membrane proteins. These are attached with compatible amino acid chains in proteins or lipid moieties. The different structures and distribution of these oligosaccharides are responsible for the cell regulation and their several respective gene functions[38]. Diversity of the oligosaccharides part increases the chance of interactions between the receptor and target analytes.

1f. Peptide libraries

The development of solid phase synthesis by Merrifield in 1963 and the discovery of the different functional groups protection methods led to the rapid design of different peptide libraries. Generally, the standard Fmoc protected aminoacids are used to make a peptides library by using either modern microwave assisted coupling with low cost coupling reagents or even utilizing the more sophisticated mechanical peptide synthesizer in a very short time frame. After completing the reactions, all peptides are cleaved and purified by HPLC. This peptide library can be further functionalized by different linkers to incorporate the fluorochromes for specific applications for *in vitro* cell analysis or *in vivo* experiments[39].

2. High-Throughput Screenings

Phenotypic screenings are one of the most vital steps in forward chemical genetics study where high throughput screening is widely utilized to explore new biological active compounds. The active compounds show phenotype change mainly by changing protein (small molecule bind protein) function. The (high throughput) screening with small molecules is carried out mostly by use of model organism, mammalian cells or cell free systems. The most popular model organisms reported for phenotypic screening are yeast, plants, zebrafish, *Drosophila* and *C. Elegans*. Two types of yeast[40] (budding and fission yeast) are used. Yeast is applied for screening because of its easy growth, high genetic conservation with human and known genome sequence. Plants[41] are also commonly used (e.g. *Arabidopsis thaliana*) as the entire genome has been sequenced, and small molecules are readily up taken by plant roots. *Drosophila*[42] is sometimes used for screening for its short life cycle, known genome sequence and RNAi. Zebrafish[43] are widely used because it is vertebrate (having brain, heart, bone) and hence comparable to the human vertebrate system. They are also prolific reproducers, transparent and thus easy to visualize when observing phenotypic change. The use of *C. Elegans*[44] for the screening is also commonly reported. Its short life span, small and transparent appearance, known genome sequence and RNAi make it a useful model organism.

Recent technological advancements in liquid handling and robotics have also facilitated the development of high throughput screening of several individual compounds in a very short time. The screening is executed by exploiting living cells (mammalian cell), or complex cellular extracts. The cell phenotype can be simply and swiftly determined by the recent use of ELISA[45] (cytoblot) method. The compounds treated cells are appended and stained by the antibody to an epitope which speculates the phenotype of interest. Secondary antibody conjugated to horseradish peroxidase is normally put in to find the signal as a luminescence.

'Screening by imaging'[46] has recently appeared as an advanced technique whereby an automated microscope is employed to observe and record cellular phenotype change in response to compound addition. The screening is carried out in multi well plates and the generated data is analysed by the software. Fluorescence spectroscopy[47] and transmitted light microscopy[48] are also routinely used to perform the phenotypic screen.

3. Target Identification Approaches

Drug discovery research requires wide knowledge of disease-related proteins and their functions and the forward chemical genetics approach robustly contributes to this field by identifying the target proteins which bind to the small molecules. Once the proteins are identified, their functions and cellular signalling pathways can be elucidated thereby facilitating drug discovery research. However, exisiting methods for target identification such as affinity matrix have met with limited success. Many researchers encounter difficulty when investigating protein targets as most of the time the protein of interest is expressed at levels that are too low in biological systems for effective detection. Low binding affinity to the small molecules as well as non-specific binding also make up some of the many problems that plague the field of target identification research. To circumvent the problem, many researchers have come up with advanced methods through the use of small molecules to identify and discover new protein targets.

3.1. Affinity matrix approaches

One of the most powerful techniques to identify a target protein is the affinity matrix approach[26]. The affinity matrix can be achieved by modifying the hit compound on a bead or by attaching a tagged molecule (photoaffinity, chemical affinity, biotin or fluorescence). In each case, the binding affinity of the proteins with small molecules is utilized to find the target proteins[49]. After binding the small molecules with the proteins, the respective complexes are fished out by immobilizing small molecule[50] and can be identified by gel separation (SDS-PAGE or fluorescence gel) followed by MS-MS (tandem mass-spectrometry).

3.1a. On-bead affinity matrix

In this approach, structure-affinity relationship (SAR) has to be first studied for a small-molecule of interest in order to find out the proper position for suitable linker modification. Second a solid-phase matrix (agarose beads) has to be attached covalently to the small molecules in a specific site which does not affect the original activity of interest of the small molecule[17]. Furthermore, the resin bound small molecules are exposed to cell extract to separate the target protein. SDS-PAGE gel chromatography is applied to analyze the proteins mixtures which are immobilized through the column. The specific target proteins are isolated by competition assay that assists to exclude the non specific binders. The isolated proteins are identified by mass spectrometry after partial tryptic digestion and database searched for mass-sequencing of the digested peptides. The result would be the discovery of target proteins or genes.
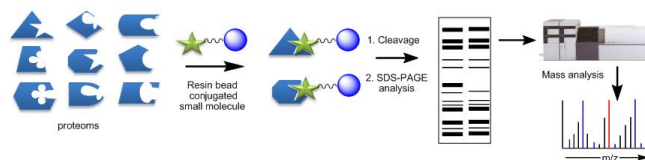

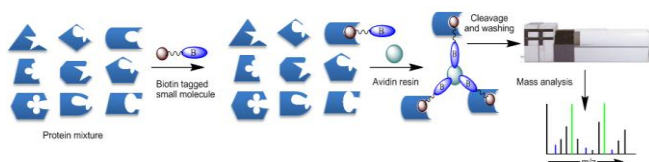**Figure 1.** Schematic representation of affinity matrix: on bead for target identification.

[Case study]

The high throughput screening of library molecules for finding the brain/eye morphological changes in the zebrafish embryo assay revealed that the encephalazine can inhibit the brain/eye development up to the different time point (1,8, and 1K cell stages). After the SAR study, the encephalaizine compound was attached with the agarose beads and then used for finding the target protein.

Affinity matrix based pull down experiment followed by 14% SDS-PAGE resolved and silver staining identified the two strong bands (23 and 18 kDa) to be ribosomal protein[49] (S5, S13, S18, and L28) which was further confirmed by LC-MS/MS .

encephalaizine
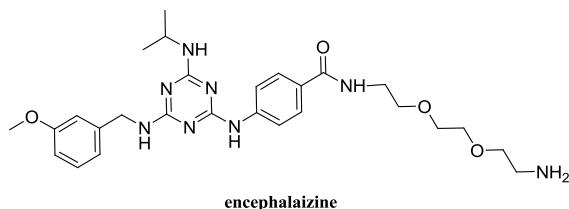
3.1b. Biotin tags in affinity matrix

The strategy of finding target proteins using biotin based affinity matrix is regularly used in chemical biology. After rigorous SARs study, the biologically active biotin attached small molecules are developed which are then treated to the cell extracts containing proteomes[51]. Next, the small molecule-biotin-protein complex can be fished out by using avidin/streptavidin resin beads[52] and then the corresponding target protein (after cleaved from the resin) can be identified by gel running followed by mass spectrometry analysis.


Figure 2. Schematic representation of affinity matrix: biotin tagged target identification method.

[Case study]

Image based high throughput screening of ESCs (R1 cell line) identified the small molecule straupimide which interacts with the NME2 protein and inhibits its nuclear localization, subsequently, the efficiency of the ESCs (human and mouse) differentiation increased**.** Stauprimide promotes ESCs differentiation toward the definitive endoderm fate which was confirmed by cell image experiment using various markers (definitive endoderm specific, mesoderm specific, visceral/parietal endoderm specific) and with the RT-PCR experiment using hepatocytes and pancreatic precursors cells (Afp, albumin, Cyp7A1, Pdx1, Ngn3 cells). The results obtained from both the experiments indicate that the stauprimide promotes the definitive endoderm mediated ESCs differentiation. Using biotin tagged affinity based approach to precipitate out the target protein followed by mass spectrometry confirmed the NME2 protein as a cellular target of the stauprimide small molecule[53]. The *in vitro* experiment and gene knock down experiment (KD1 and KD2 genes sequences), further confirmed that the NME2 protein is the cellular target of the straupimide small molecule.

Stauprimide
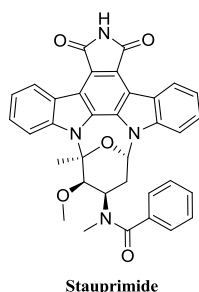
3.1c. Fluorescent tags in affinity matrix

Sometimes, affinity matrix with fluorescence tag can be used to visualize a target protein quite easily. The approach is quite similar to the biotin tag approach. In the case of fluorescence tag approach, small molecule bound proteins can be recognized by running fluorescence gels. The visualized fluorescence bands are excised and then analyzed by mass spectrometry to identify target proteins.
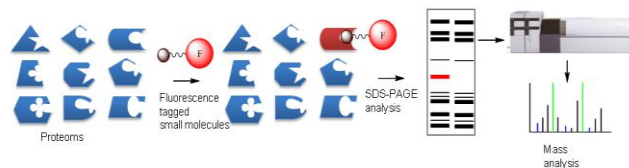

Figure 3. Schematic representation of affinity matrix: fluorescence tagged target identification method.

[Case study]

Cell (MDA-MB-231 human breast cancer cell) based screening of 50-membered natural product-like library revealed the compound MJE3, which can inhibit breast cancer cell proliferation (IC50 value of 19 M) and then activity based protein profiling (ABPP) which is quite similar like affinity matrix contained a reactive group, a binding group and an analytical tag. In this study a fluorescence tag was employed for successful identification of the target protein. After incubating the MJE3 in the MDA-MB-231 cells, the click reaction was carried out between the MJE3 labeled protein with a trifunctional biotin/rhodamine-azide reporter tag and the probe labeled protein was separated in SDS-PAGE and then visualized in gel fluorescence scanning[25]. The probe labeled protein was subsequently purified using avidin agarose beads. The gel band (26-kDa) was excised and then after the trypsin digestion the resulting peptide mixtures were analyzed and the MS data based search identified the 26-kDa protein as brain-type phosphorglycerate mutase 1 (PGAM1).

MJE3
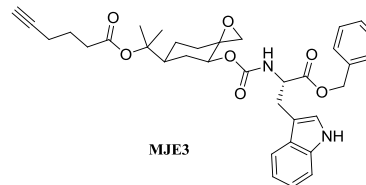
3.1d. Photoaffinity tags in affinity matrix

In this method, one photoaffinity moiety along with a reporter tag is added to the initial molecule (hit compound), making it highly reactive upon UV irradiation, releasing a reactive carbene functional group that can covalently attach to the specific target proteins. Subsequently, this binding protein can be fished out from the complex by a biological cell assay[54]. The reporter tag is usually a radioactive isotope[55] or chemical affinity group such as biotin, which allows for isolation and identification[8]. Mainly stable isotopes that can give unique isotopic pattern have been used to selectively identify the binding proteins even in very complex mixtures by mass spectrometry.
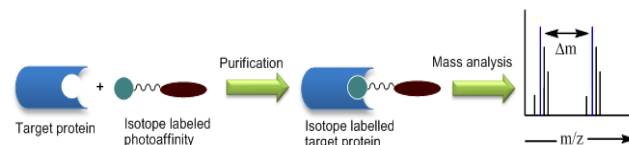

Figure 4. Schematic representation of the affinity matrix: photoaffinity based target identification method.

[Case study]

Image based high throughput phenotypic screening of neuropathiazols compounds identified some compounds which induce neuronal differentiation of cultured rat hippocampal NPCs. Synthesis of several analogs of the original neuro-pathiazol structure and a focused structure-activity relationship (SAR) study afforded a molecule (KHS101) of best activity towards the neuronal differentiation. The RT-PCR experiment and image based experiment revealed that the treatment of KHS101 significantly suppresses astrogenesis and at the same time it increases the neurogenesis. Protein-KHS101-BPcomplex was detected after

photocrosslinking and biotin-tag labeling by using two dimensional SDS/PAGE and western blotting of NPC cell lysates. Finally, mass spectrometry analysis revealed the 80-kDa protein to be TACC3. Gene knock down experiments and *in vivo* imaging studies confirmed that the TACC3 protein is the cellular target of the KHS101 small molecule[27,56].

KHS101
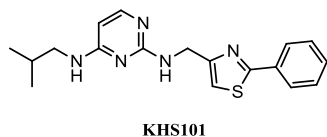
3.2. Drug western approaches

In the Drug westerns method[57], bacteriophages are grown in a petridish. Then the lysis caused by a viral infection can lead to a clearing, consisting of one member of the library. The proteins from the plaque are moved to the nitrocellulose and then these are screened against the tagged small molecules. Hit obtained from the plaques are isolated and then each single virus is purified and followed by the application of the DNA sequencing method, the each target protein is identified.
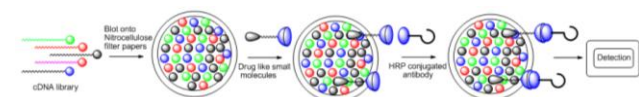

Figure 5. Schematic representation of the drug western techniques for target identifiaction.

[Case study]

Cell based screening of sulfonamide drug molecules, followed by SARs study found the small molecule HMN-154 which showed the potential anticancer activity towards the various cancer cell lines. The HMN-154 coupled BSA tagged was employed to identify the target protein using the drug-western method. BSA-conjugated HMN-154 was incubated in nitrocellulose membrane where the protein was expressed by the lTriplEx cDNA library. The positive clone, expressing HMN-154-binding protein was detected[57] and then the binding protein was identified after the DNA sequencing study to be transcription factor inhibitor NF-YB. The identified protein was further confirmed by an *in vitro* inhibition assay and immune-precipitation experiments.

HMN-154
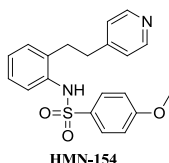
3.3. Three-hybrid system approaches

So far, two types of three hybrid systems have been developed in order to identify target proteins. Yeast three hybrid systems are performed in yeast cells, and mammalian three hybrid systems are carried out mammalian but both the approaches are quite similar. Initially, both the systems were used to study the protein- protein interaction, but later the applications were widened to study the small molecule-protein interactions too.

3.3a. Yeast three-hybrid systems

The yeast three-hybrid system (Y3H) is evolved from yeast two-hybrid screens and incorporates a dimer of organic small molecules into the screening spot. It comprises of three components: a synthetic hybrid ligand and two hybrid fusion proteins. The synthetic hybrid ligand is formed by covalently linking two small molecule ligands (hetero/homodimer) which induces cellular proteins in different cells. The synthetic hybrid ligand remains attached to one of the two hybrid fusion proteins. One of the small molecules in hybrid ligand brings the third hybrid protein to close constituting a functional transcription factor that drives expression of a reporter gene[58]. Whereas first ligand of the dimer binds to the receptor of DNA-binding domain, a second ligand binds to the receptor of transactivation domain allowing for the selection of yeast cells that

harbor the relevant receptors. The two functions (ligand–receptor interaction) of the system are screened by the small molecule dimers with a known activation domain or using a known small molecule to identify target proteins, such as dexamethasone-glucocorticoid receptor to find interacting proteins. The advantages of Y3H system is that it is carried out *in vivo* thus the phenotype and genotype are closely linked. On other hand it is restricted in a unicellular system.
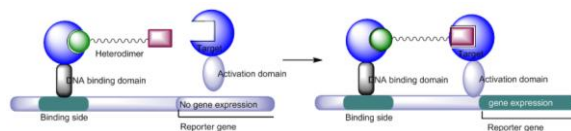

Figure 6. Schematic diagram of Y3H system for target identification.

[Case study]

Cell based CDK (Cyclin Dependent Kinase) inhibition assays and a purine library screen, followed by SARs study discovered a small molecule named Purvalanol B which can inhibit cyclin dependent kinase activity in human leukemic cell at a nanomolar concentration range. In order to identify the target protein, Yeast three hybrid cDNA libraries screening employing small molecule Purvalanol B-MFC (methotrexate fusion compound coupled with the Purvalanol B) was pursued by utilizing the following steps. First, the lexA-DHFR expressing yeast cells were transformed with the choice of cDNA library. Next, transcriptional expression of the HIS3 auxotrophic marker was selected in the presence of an MFC. Then, the positive colonies were picked up and arrayed robotically; following this the compound dependence to the HIS3- reporter expression was reconfirmed. Finally, plasmid isolation, sequencing, retransformation of yeast expressing LexA-DHFR with purified plasmids, robotic arraying of the transformed yeast cells and reconfirmation of specific HIS3-reporter activation by the test MFC using a series of genetic or compound-based counter screens identified small molecule binding target proteins[58] (CDK1, CDK5, CDK6, CLK3, EPHB2, FLT4, FYN, PAK4, PCTK1, PCTK2, RSK3 and YES), many of which were further confirmed by using the affinity matrix pull down experiment and secondary enzyme assays.

Purvalanol B
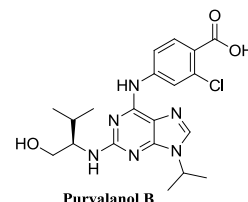
3.3b. Mammalian three-hybrid systems

The mammalian three-hybrid system is similar to the yeast three hybrid systems where mammalian cells are used instead of yeast. Initially, S. Eyckerman, et al developed the MAPPIT[59] explain acronym system in order to identify protein-protein interactions. Later C. Maureen, et al expanded this concept to the MASPIT[60] system to investigate small molecule-protein interaction and target identification.

[Case study]

After the synthesis of the several small molecules of Pyrido [2,3-d] pyrimidine, followed by their application in the kinase inhibition assay and SARs, studies revealed PD17395 as a SRC kinase inhibitor which can inhibit several ephrin receptor tyrosine kinases. In order to identify the target protein, the MASPIT system was employed in mammalian cells. For target identification, the following pathways were adapted as l in Y3H. First, a cDNA library was built up from HEK293 mRNA in a retroviral vector and then the IL5R reporter cells were infected with the retroviral library and subjected to various cycles of enrichment for MFC-dependent IL5R-positive cells, followed by flow cytometric single-cell sorting into 96-wellmicro-titer plates. Individual cell populations were subsequently screened for MFC and Epo-dependent reporter activation by fluorescence activated cell sorting (FACS). The cDNA library screen

with RGB-286649 and the MFC incorporating the ABL tyrosine kinase inhibitor PD173955 discovered a number of different tyrosine kinases as well as one Ser/Thr kinase (Cyclin G-associated kinase, ephrin receptor tyrosine kinases, FGFR1 and SRC kinases FYN and LYN)[60]. These proteins were further confirmed with the *in vitro* enzymatic assay (competition assay).

PD17395
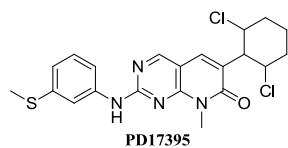
3.4. Phage display approaches

In recent years, phage display technology has emerged as a popular method to connect proteins or polypeptides with genes in bacteriophages to study the protein-protein, protein-DNA and protein-peptide interactions. This technology utilized the display of proteins or peptides on bacteriophages such as filamentous, M13, T4, T7, phage and it enables the extraction of proteins from a large collection of variants by immobilizing relevant DNA or proteins. The resulting phage particles containing genes and encoded proteins provide a connection between phenotype and genotype that enables large libraries of proteins to be screened and further amplified. In the process of screening, the bacteriophage displayed protein bound to the target will remain in the washing step and other non interacting proteins are removed. The recovered proteins are identified by sequencing and more phages can be produced by bacterial infection to enrich the interacting proteins for further selection[61].
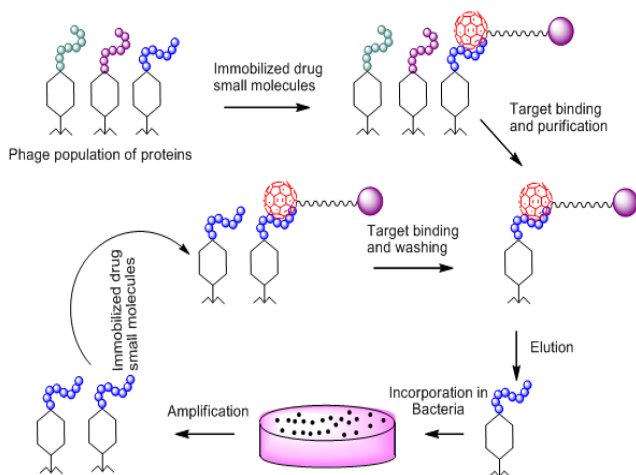

Figure 7. Schematic diagram of phase display method for identifying target protein.

[Case study]

Screening of a library of 1200 compounds revealed a candidate, IHY-153 that can effectively inhibit the proliferation of several human cancer cell lines, including human Cervical Cancer cell line (HeLa), human liver carcinoma cell line (HepG2), human fibrosarcoma cell line (HT 1080) and human colon carcinoma cell line (HCT116). Among them, HCT116 colon cancer cells are the most sensitive to IHY-153. Investigation of the effect of IHY-153 on HCT116 cell cycle progression by flow cytometry indicates that IHY-153 inhibits cell-cycle in a dose-dependent manner and induces cell cycle arrest at G0-G1 phase. Bacteriophage display biopanning approach was applied in the target identification. T7 phage particles expressing human cDNA libraries were added into the BH1-immobilized wells. After incubation and washing, bound phage particles were eluted with 1HY-153 and eluted phage particles were amplified after infection into *E. coli* strain BLT5615 and used for a second round of biopanning. After the fourth biopanning, eluted phages were

infected into LB agar *E. coli* and plaques formed were isolated. Amplified phage lysates from the isolated plaques were used as PCR templates and obtained sequences were compared with database. The phage encoding region matched 100% with human calmodulin (CaM). The specificity of IHY-153 towards CaM was also tested via phage binding assay and phage library. Necessity of Ca^{2+} for CaM binding with IHY-153 is tested and results demonstrate that Ca^{2+} is required for IHY-153 binding to CaM[61].

1HY-153
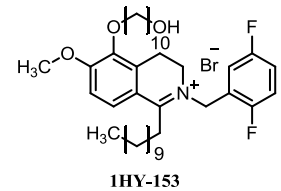
3.5. mRNA display approaches

mRNA display is an well known *in vitro* technique and has recently been used in studying protein-protein interaction[62]. The technique was initially developed to amplify the number of peptides[63] displayed in the phage display method. After the cDNA library amplification by PCR, the puromycin–DNA linker is ligated to the produced mRNA and using this mRNA-DNA complex[64] *in vitro* translation is carried out. The translation leads to generation of mRNA-protein fusion molecules which are then purified and reverse transcribed to generate the cDNA template which can be utilized to construct further amplification. Then the small molecules of interest are immobilized on solid support and are incubated with mRNA display molecules. The unbound protein-mRNA complexes are washed out and bound complexes are eluted and the bound genes amplified using PCR technique. Finally, after the several iteration processes, the cDNAs are purified, cloned, sequenced and put thorough a sequence similarity search to identify the target protein.
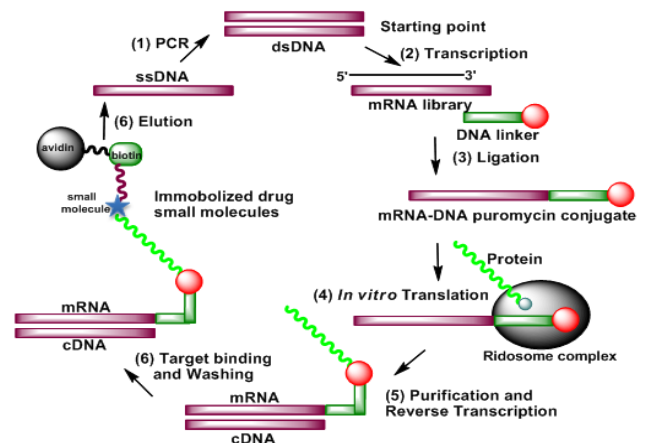

Figure 8. Schematic representation of mRNA display techniques for identifying target protein.

[Case study]

Tacrolimus (FK-506)[65] is a natural product which is established to be an immune-osuppressive drug. Molecules bound to the target protein were identified by using FK-506 biotinylated compounds in mRNA display method. After making the biotin tag with the FK-506, several steps are carried out in order to identify the molecular target in the mRNA display method.

In the beginning, the PCR-amplified cDNA was generated by applying primers that commence the engineered sequences necessary for transcription, ligation of the corresponding mRNA, *in vitro* translation of the mRNA-puromycin linker conjugate, and

epitope-based purification of the mRNA-protein fusion, Then this engineered PCR product was transcribed to construct mRNA which was then hybridized to a poly-dA-containing linker carrying a 5'-psoralenmoiety and a 3'- terminus containing the transfer-RNA mimic puromycin (Pu). Next a covalent crosslink between the mRNA and DNA-puromycin linker was introduced upon UV irradiation and this conjugate (mRNA-DNA-puromycin) was employed as a template for *in vitro* translation where ribosome could translate the open reading frame and stop at the mRNA-DNA junction. The dearth of a stop codon hindered the action of release factors and permitted the conjugate puromycin to enter the A-site of the ribosome and then the peptidyl transferase subunit catalyzes amide bond formation between an amine group on the puromycin and the carboxyl terminus of the mature protein to give an mRNA-protein fusion. Next, the fusion was purified (dissociation from the ribosome happened in presence of poly dA linker with oligo-dT cellulose) and the cDNA strand was made by reverse transcription of the fusion that protected the mRNA against degradation and served as a template for future PCR. After the initial random priming of cellular mRNA, the library of mRNA display molecules were then incubated with an immobilized drug (immobilized on streptavidin bead) or small molecule, and unbound material was eliminated by washing. The bound fusion protein was then eluted either specifically using excess of drug or nonspecifically using KOH. The identified target protein was found to be **FKBP12**[51]. This was further confirmed with the *in vitro* protein binding assay.

FK-506
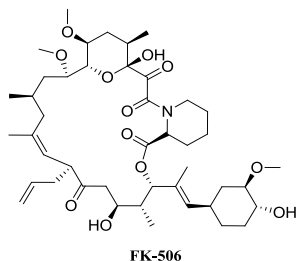
3.6. Protein microarray approaches

The high throughput analysis of interactions between the target proteins and small molecules can be performed by using protein chip technology which is known as protein microarray[66]. The high density format is used in this approach. Hence, only few days are required to study the binding profile of a certain small molecule to an entire proteome. The proteins of interest for analysis are purified and consequently immobilized on the glass microscope slide or derivatized on another surface[67,68]. The fluorescent tag or radioisotope-labeled small molecules are incubated in the array and then the positions where small molecule binds protein on the array are determined and the target protein subsequently identified.
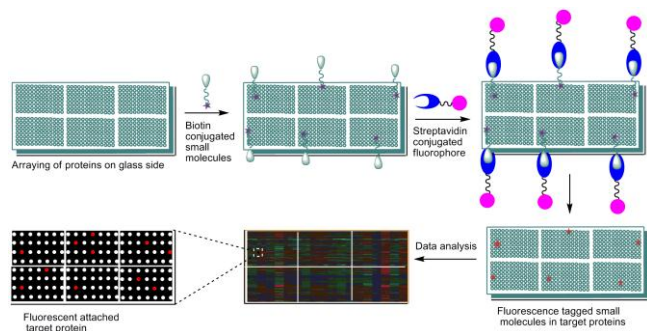
Figure 9. Schematic representation of protein microarray method for target identification.

[Case study]

Yeast growth inhibition studies using the chemical genetic screen of small molecules (16,320 compounds) identified SMIR3 and SMIR4 which can fully suppress rapamycin's anti-proliferative effects in yeast. Both the molecules were then biotinylated (to preserve their bioactivity) to identify target protein.

Biotin-small molecules (SMIR3-biotin and SMIR4-biotin) were probed on the proteome chip which consisted of almost the entire yeast proteome. After adding Cy3 labeled streptavidin on the proteome chip, 8 different candidate proteins binding to SMIR3-biotin and 30 different candidate proteins binding to SMIR4-bioitin were identified. This was followed by an *in vivo* study of rapamycin sensitivity and the ability of SMIR to suppress the rapamycins's effect towards the yeast cell using the yeast deletion strain of each candidate protein (found by protein chip) discovered YBR077CP(Nir1P) protein which is responsible for SMIR4 to suppress rapamycins's effect[69].

SMIR4

3.7. Drug affinity responsive target stability (DARTS)

A few successful cases have been reported based on the affinity matrix method. Introducing an affinity tags on to effective drug molecules to identify the target proteins is a major challenge due to the present limitations. A new technique, drug affinity responsive target stability (DARTS) may have great prospective to discover target proteins by stabilizing the target protein upon the binding of drug like mall molecules[70,71]. This modern methodology represents a target identification strategy without the modification of small molecules and relies on drug–protein binding. After binding with drug molecules, target proteins are less susceptible towards proteolysis thus helping to readily identify target proteins even in complex biological systems. Generally the ligand-bound states are thermodynamically favourable structures among multiple dynamic conformations due to hydrogen bonding, hydrophobic and/or electrostatic interactions between the protein and the small molecules. Hence, the target proteins are stabilized by restricting the multiple conformations.
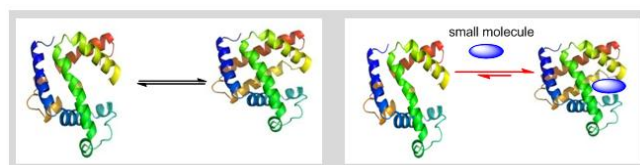
Figure 10. Schematic representation of DARTS principle.

To date the affinity matrix as well as affinity chromatography methods are well studied by different research groups and its major limitations are well defined. In spite of several good examples having been reported, the major drawbacks (e.g. nonspecific binding to nontarget proteins, time consuming SAR study and impossible to incorporate the affinity tag for some drug like small molecules) limit its application. In general, nonspecific binding proteins can be eliminated by simple washing for multiple times but in the same time interested binding proteins may be eliminated if the interaction between target proteins and small molecules is too weak. Therefore, compared to affinity matrix, DARTs presents a unique alternative path that allows for the detection of the target proteins by small molecules which requires neither further chemical derivatization nor extensive washing. Furthermore, this approach allows all bioactive small molecules, different source of extensive chemical library with diverse structural diversity including natural products too. DARTS apply to find target proteins from complex biological samples by digesting away nontarget proteins that are

susceptible to protease hydrolysis and vice versa.

[Case study]

Resveratrol, a compound that can be found in grapes and red wine is known to be responsible for various health benefits. However, its direct molecular target protein has not been reported so far due to low specific binding affinity of small molecules towards proteins. A potential requirement of the polyphenol groups in resveratrol for its activity has discouraged generation of affinity reagents for target identification. To overcome this limitation, Lomenick et al, came up with a novel idea considering the less susceptibility of the target protein to drug binding named as drug affinity responsive target stability (DARTS). As a proof of principle small molecule-protein complexes (mTOR-rapamycin and COX-2–celecoxib) are identified by DARTS method. Eventually, a target protein eIF4A was identified by applying new tools (DARTS). DARTS with resveratrol dosed yeast cell lysates revealed two silver stained bands between the 15- and 20-kDa MW markers that were more intense in the resveratrol treated lysate post proteolysis compared with the control. Mass spectrometry analysis identified the target protein as eIF4A[70]. Further gene mutation of target protein clearly depicts that eIF4A is the vital protein for various health benefits.

Resveratrol
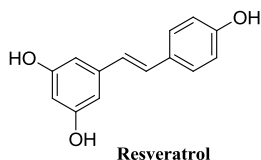
3.8. *In vitro* inhibition assay: guess and check method

Sometimes, the target protein could be determined by using a simple *in vitro* assay. After the detection of a hit compound by looking at the change in the cell or organismal phenotype or by looking at the cell images, expert researchers can predict some relative target proteins responsible for the respective alternations. This knowledge based approach leads one to guess and then to check some *in vitro* assay with some presumed proteins thereby leading to the identification of target proteins. This is a frequently used method[36] to find out the target proteins.

[Case study]

Optical density based high throughput screening of conditionally essential enzymes identified a small molecule which can prevent the growth of a wild type bacterial strain but do not affect the growth of a mutant strain incapable of initiating polymer synthesis. This approach led to the discovery of 1835F03[72], a molecule that can inhibit the wall teichoic acid (WTA) biosynthesis in *Staphylococcus aureus*. Using radiometric *in vitro* assays to test the inhibitory effects of 1835F03 on Tar B,D,F and L, and using overexpression and resistant mutants analysis showed that the 1835F03 compound can potentially inhibit the function of the TarG, which can exports WTAs to the cell surface, with an MIC of 1.3 µg/mL. The assessment of the antibacterial properties of 1835F03 compound can clearly demonstrate that its activity was fully bacteriostatic.

1835F03
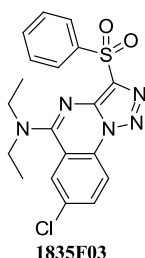
3.9. Magnetism-based Interaction Capture (MAGIC)

Magnetism-based interaction capture (MAGIC)[73,74] is an *in vivo* target identification approach in mammalian cells. The compound specific for target proteins is bound to cDNA library coupled with EGFP motif. When proteins are attached to the compounds, they will be separated from matrix under magnetic field. Hence through this magnetism-based approach the target protein can be extracted out and identified.
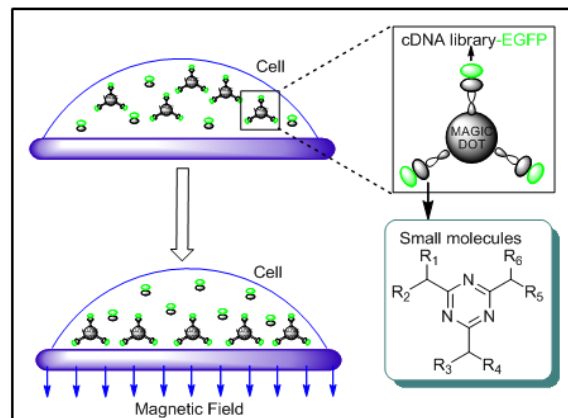
[Case study]

Figure 11. Schematic representation of MAGIC method for target identification.

The magnetism-based interaction capture (MAGIC) method was first developed by the Korean scientist (Jaejoon Won and his coworkers) in science journal[73] to identify target protein in living cell. High throughput screening revealed the ATM protein as the cellular target of the small molecule CGK733[74]. However, both the journals were retracted for data fabrications and results misrepresentation. Hence, there is no real example of target protein identification in this approach, although this method may be a good potential to find out the target protein in complex system.

In above study, we have demonstrated different useful strategies for the target protein identifications in complex biological systems. Moreover, we also highlighted the potential application of these strategies in forward chemical genetics with specific real examples. In the below table we have summarized several examples in a short and succinct way for better understanding (Table 1).
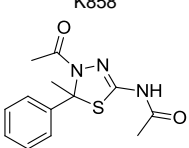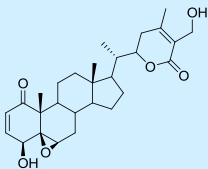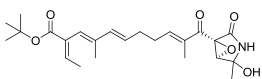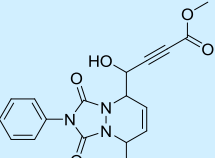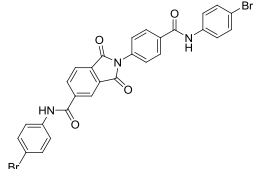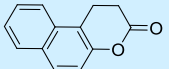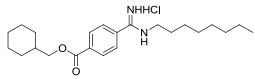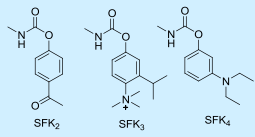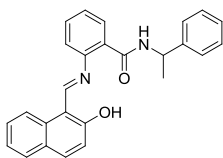
Conclusion and Prospects
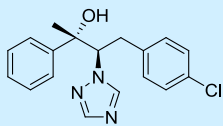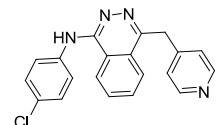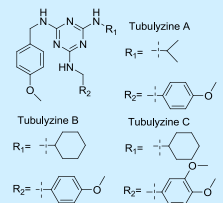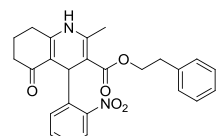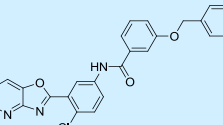
In this review, we described all the three components of forward chemical genetics. The approach consists of the collection of chemical compounds from various sources, screening in a high throughput format and then the daunting task of target identification. The first two parts were covered briefly and then we mainly focused on the several target identification methods with successful story of discovering small molecules bound prot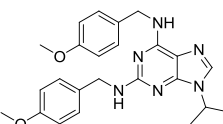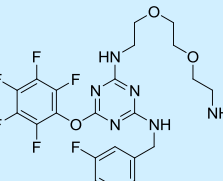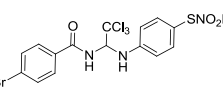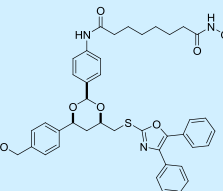eins. Our objective in this review is to bring most of the target identification methods which are successful for identifying target proteins into one frame. Previously most routinely used methods like affinity matrix and other protocols sometimes show hurdle to discover target proteins. To resolve the problem, more technically advanced target identification methods have been steadily developed in the recent years. Our review has included both the previous and recent target identification protocols along with illustrations of several real examples which were accumulated from several research works. Hence, chemical biology researchers can easily follow as well as apply any target identification method in their respective drug discovery research. We prognosticate that this review would be supportive for the researcher to employ forward chemical genetics in accelerating the drug discovery research.
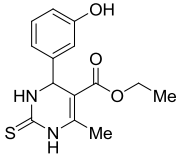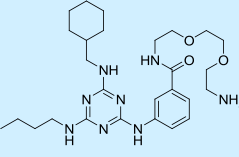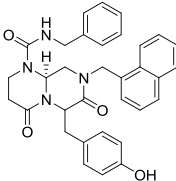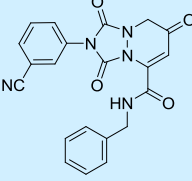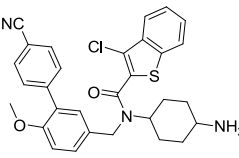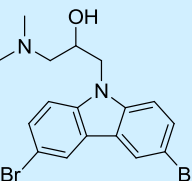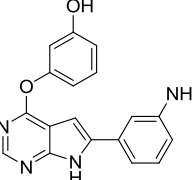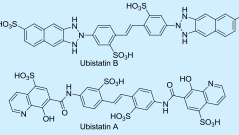
Acknowledgements
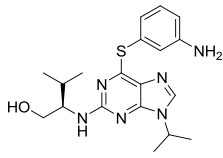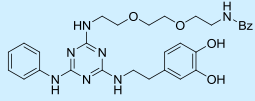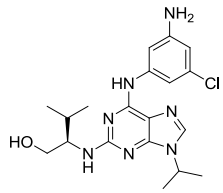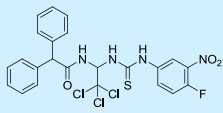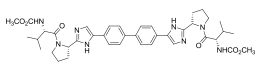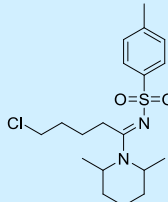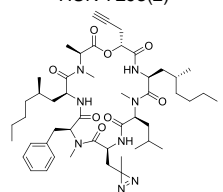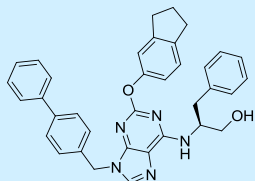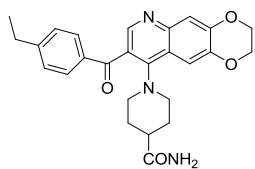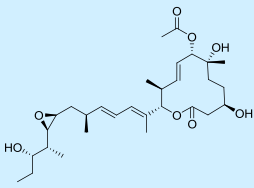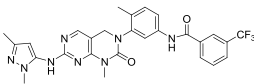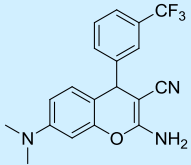
Table 1. Representative examples of successful target identification approaches

Small molecule	Library size	Activity	Target protein	Target identification methods	Reference
K858	Single molecule	Inhibit the ATPase activity of Eg5	Eg5	Guess and check method	75
Withaferin A(WFA)	Not specified	Antitumor and antiangiogenesis activity *in vivo*	Type (III) intermediate filament (IF) protein, vimentin.	Affinity matrix: biotin tag	33
Epolactaene	Singles molecule	Inhibited mammalian DNA polymerases and human DNA topoisomerase II *in vitro*	Hsp (heatshock protein) 60	Affinity matrix: biotin tag	76
MOL 249	Not specified	NF-κB inhibition	Thioredoxin	Guess and check method	77
Small molecule inhibitors of rapamycin (SMIR4)	16,320	Inhibition of rapamycin activity	Tep1p and Ybr077cp (Nir1p)	Protein microarrays	69
Splitomicin	6,000	Yeast growth inhibition	Sir2p	Guess and check method	78
SFK1	100,000	Suppressor of FK506	Por1p	Affinity matrix: on bead	79
SFK$_{2,3,4}$	100,000	Suppression of FK506 activity	Ald6p (aldehyde dehydrogenase)	Guess and check method	80
Sirtinol	1,600	Auxin signal increase	Sirtuin	Guess and check method	81,82

Compound		Quantity	Biological effect	Target	Method	Ref.
Brassinazole		Single compound	BR biosynthesis inhibition in plant	DWF4 (cytochrome P450)	Guess and check method	83
PTK787/ZK222584		Single compound	Block blood vessel formation	VEGFR inhibitor	Guess and check method	84
Tubulyzine A,B,C		>100	Delaying the epiboly development in zebrafish	Tubulin	Guess and check method	85
MP576		50,240	Inhibition of SARS-CoV	SARS-CoV M^{Pro}	Guess and check method	86
HE602		50,240	Inhibition of SARS-CoV	SARS-CoV helicase	Guess and check method	86
Myoseverin		Single compound	Fission of multinucleated myotubes into mononucleated fragments	Tubulin	Affinity matrix: biotin tag	87,88
Melanogenin		1,170	Pigmenting melanocyte	Prohibitin	Affinity matrix: on bead	26
Syntab A		16,320	Mitosis perturbation	Tubulin	Guess and check method	89
Tubacin		7,392	Inhibits β-tubulin acetylation	HDAC6	Guess and check method and drug western	90

Compound	Structure	Value	Effect	Target	Method	Ref
Monastrol		16,320	Block mitosis endering	Mitotic kinesin Eg5	Guess and check method	91
MP C11		1,536	Induce pigmentation in albino murine melanocytes	F1F0-ATP synthase	Affinity matrix: on bead	92
ICG-001		5,000	Down regulation of β-catenin	Cyclic AMP response element-binding protein (CBP)	Affinity matrix: biotin tag	93
PNRI-299		12	Activator protein 1 (AP-1) inhibitor	Redox factor 1 (Ref-1)	Affinity matrix: biotin tag	94
Hh-Ag 1.1		140,000	Inhibition of Smo	Smo	Guess and check method	95
Wiskostatin		16,320	Prevent net polymerisation of actin Inhibits activation of Arp2/3 complex	N-WASP	Guess and check method	96,97
TWS119		Not specified	Induce differentiation of pluripotent EC and ES cells to neurons	Glycogen synthase kinase-3 (GSK-3)	Affinity matrix: on bead	22
Ubistatin A, B		109,113	Inhibition of mitotic entry	Ubiquitin	Guess and check method	98

Diminutol	1,561	Inhibitor of mitotic spindle assembly	NQO1 (an NADP-dependent oxidoreductase)	Affinity matrix: on bead	99
GAPDS	1,120	Regulates DAF-2 insulin signaling in *C. Elegans*	GAPDH	Affinity matrix: on bead	100
Aminopurvalanol	150	Cell cycle arrest at G2/M	CDK1	Affinity matrix: on bead	87
CGK733	20,000	Selectively block the ATM and ATR kinase checkpoint signaling pathway	Ataxia telangiectasia mutated (ATM) protein	MAGIC approach	74
BMS-790052	>1million	HCV replication inhibitor	HCV NS5A	Affinity matrix: biotin tag	101
Sulfonyl Amidine	64	Anti-resorptive activities to inhibit the osteoclastogenesis	Prohibitin	Affinity matrix: on bead	102,103
HUN-7293(2)	Single molecule	Inhibitor of vascular cell adhesion molecule (VCAM) expression	Sec61α	Affinity matrix: photoaffinity	55
QS11	100,000	Synergizes with Wnt-3a ligand in the activation of Wnt/catenin signal transduction and inhibits ARFGAP1	ARFGAP1	Affinity matrix: on bead	104
Quinostatin	20,000	Antagonist of mTOR signaling	Class Ia PI3Ks	Affinity matrix: on bead	36

Pladienolide	Single molecule	Antitumor agent	SF3b	3H-labeled, fluorescence-tagged and photoaffinity/Biot in (PB)-tagged	52
SC1	50,000	Self-renewal of ES cells	ERK1- and RasGAP	Affinity matrix: on bead	105
Chromeceptin	10,000	Cellular attenuators of the IGF signals, through activation of STAT6	MFP-2	Affinity matrix: biotin tag	66,106

References

1. Witkowski, J. (2010). Long view of the Human Genome Project. *Nature* 466, 921-922.
2. Feder, M.E., and Mitchell-Olds, T. (2003). Evolutionary and ecological functional genomics. *Nat Rev Genet* 4, 651-657.
3. Huttenhower, C., Haley, E.M., Hibbs, M.A., Dumeaux, V., Barrett, D.R., Coller, H.A., and Troyanskaya, O.G. (2009). Exploring the human genome with functional maps. *Genome Res* 19, 1093-1106.
4. Ho, C.H., Piotrowski, J., Dixon, S.J., Baryshnikova, A., Costanzo, M., and Boone, C. (2010). Combining functional genomics and chemical biology to identify targets of bioactive compounds. *Curr Opin Chem Biol* 15, 1-13.
5. Stockwell, B.R. (2000). Chemical genetics: ligand-based discovery of gene function. *Nat Rev Genet* 1, 116-125.
6. De Rybel, B., Audenaert, D., Beeckman, T., and Kepinski, S. (2009). The past, present, and future of chemical biology in auxin research. *ACS Chem Biol* 4, 987-998.
7. Chang, Y.T. (2008). Forward Chemical Genetics. *Willey Encyclopedia Chem Bio*, 1-8.
8. Neumann, G., Hatta, M., and Kawaoka, Y. (2003). Reverse genetics for the control of avian influenza. *Avian Dis* 47, 882-887.
9. Zheng, X.F., and Chan, T.F. (2002). Chemical genomics in the global study of protein functions. *Drug Discov Today* 7, 197-205.
10. Stockwell, B.R., Haggarty, S.J., and Schreiber, S.L. (1999). High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem Biol* 6, 71-83.
11. Stockwell, B.R. (2004). Exploring biology with small organic molecules. *Nature* 432, 846-854.
12. Egener, T., Granado, J., Guitton, M.C., Hohe, A., Holtorf, H., Lucht, J.M., Rensing, S.A., Schlink, K., Schulte, J., Schween, G., et al. (2002). High frequency of phenotypic deviations in Physcomitrella patens plants transformed with a gene-disruption library. *BMC Plant Biol* 2, 6.
13. Namy, O., Hatin, I., Stahl, G., Liu, H., Barnay, S., Bidou, L., and Rousset, J.P. (2002). Gene overexpression as a tool for identifying new trans-acting factors involved in translation termination in *Saccharomyces cerevisiae*. *Genetics* 161, 585-594.
14. Koivisto, U.M., Hubbard, A.L., and Mellman, I. (2001). A novel cellular phenotype for familial hypercholesterolemia due to a defect in polarized targeting of LDL receptor. *Cell* 105, 575-585.
15. Ballard, J.W., and Melvin, R.G. (2010). Linking the mitochondrial genotype to the organismal phenotype. *Mol Ecol* 19, 1523-1539.
16. Spring, D.R. (2005). Chemical genetics to chemical genomics: small molecules offer big insights. *Chem Soc Rev* 34, 472-482.
17. Kim, Y.K., and Chang, Y.T. (2007). Tagged library approach facilitates forward chemical genetics. *Mol Biosyst* 3, 392-397.
18. Scherer, L.J., and Rossi, J.J. (2003). Approaches for the sequence-specific knockdown of mRNA. *Nat Biotechnol* 21, 1457-1465.
19. Lokey, R.S. (2003). Forward chemical genetics: progress and obstacles on the path to a new pharmacopoeia. *Curr Opin Chem Biol* 7, 91-96.
20. Kidera, A., and Go, N. (1990). Refinement of protein dynamic structure: normal mode refinement. *Proc Natl Acad Sci U S A* 87, 3718-3722.
21. Shapiro, J.A. (1998). Thinking About Bacterial Populations as Multicellular Organisms. *Annu Rev Microbiol* 52, 81-104.
22. Ding, S., Wu, T.Y., Brinker, A., Peters, E.C., Hur, W., Gray, N.S., and Schultz, P.G. (2003). Synthetic small molecules that control stem cell fate. *Proc Natl Acad Sci U S A* 100, 7632-7637.
23. Boshoff, H.I., and Dowd, C.S. (2007). Chemical genetics: an evolving toolbox for target identification and lead optimization. *Prog Drug Res* 64, 49, 51-77.
24. Raikhel, N., and Pirrung, M. (2005). Adding precision tools to the plant biologists' toolbox with chemical genomics. *Plant Physiol* 138, 563-564.
25. Evans, M.J., Saghatelian, A., Sorensen, E.J., and Cravatt, B.F. (2005). Target discovery in small-molecule cell-based screens by *in situ* proteome reactivity profiling. *Nat Biotechnol* 23, 1303-1307.
26. Snyder, J.R., Hall, A., Ni-Komatsu, L., Khersonsky, S.M., Chang, Y.T., and Orlow, S.J. (2005). Dissection of melanogenesis with small molecules identifies prohibitin as a regulator. *Chem Biol* 12, 477-484.
27. Warashina, M., Min, K.H., Kuwabara, T., Huynh, A., Gage, F.H., Schultz, P.G., and Ding, S. (2006). A synthetic small molecule that induces neuronal differentiation of adult hippocampal neural

progenitor cells. *Angew Chem Int Ed Engl* 45, 591-593.

28. Wang, S., and Chang, Y.T. (2008). Discovery of heparin chemosensors through diversity oriented fluorescence library approach. *Chem Commun (Camb)*, 1173-1175.

29. Wang, S., Kim, Y.K., and Chang, Y.T. (2008). Diversity-oriented fluorescence library approach (DOFLA) to the discovery of chymotrypsin sensor. *J Comb Chem* 10, 460-465.

30. Vendrell, M., Lee, J.S., and Chang, Y.T. (2010). Diversity-oriented fluorescence library approaches for probe discovery and development. *Curr Opin Chem Biol* 14, 383-389.

31. Lee, J.S., Kim, Y.K., Vendrell, M., and Chang, Y.T. (2009). Diversity-oriented fluorescence library approach for the discovery of sensors and probes. *Mol Biosyst* 5, 411-421.

32. Carlson, E.E. (2010). Natural products as chemical probes. *ACS Chem Biol* 5, 639-653.

33. Bargagna-Mohan, P., Hamza, A., Kim, Y.E., Khuan Abby Ho, Y., Mor-Vaknin, N., Wendschlag, N., Liu, J., Evans, R.M., Markovitz, D.M., Zhan, C.G., et al. (2007). The tumor inhibitor and antiangiogenic agent withaferin A targets the intermediate filament protein vimentin. *Chem Biol* 14, 623-634.

34. Crews, C.M., Yeh, J., Mohan, R., Meng, L., Kim, K., Splittgerber, U., Kwok, B. H. B., and Elofsson, M. (2000). Natural products as molecular probes: A chemical genetic approach to pharmaceutical target validation. *Abstr Pap Am Chem S* 219, U480-U480.

35. Crews, C.M. (2006). Chemical genetics: Using natural products as probes for cell biology. *Abstr Pap Am Chem S* 231.

36. Yang, J., Shamji, A., Matchacheep, S., and Schreiber, S.L. (2007). Identification of a small-molecule inhibitor of class Ia PI3Ks with cell-based screening. *Chem Biol* 14, 371-377.

37. Zhang, W., Lu, Y., Hiu-Tung Chen, C., Zeng, L., and Kassel, D.B. (2006). Fluorous mixture synthesis of two libraries with hydantoin-, and benzodiazepinedione-fused heterocyclic scaffolds. *J Comb Chem* 8, 687-695.

38. Miyazaki, K., Hirase, T., Kojima, Y., and Flint, H.J. (2005). Medium- to large-sized xylo-oligosaccharides are responsible for xylanase induction in *Prevotella bryantii* B14. *Microbiology* 151, 4121-4125.

39. Kelly, K., Alencar, H., Funovics, M., Mahmood, U., and Weissleder, R. (2004). Detection of invasive colon cancer using a novel, targeted, library-derived fluorescent peptide. *Cancer Res* 64, 6247-6251.

40. Botstein, D., Chervitz, S.A., and Cherry, J.M. (1997). Yeast as a model organism. *Science* 277, 1259-1260.

41. Blackwell, H.E., and Zhao, Y. (2003). Chemical genetic approaches to plant biology. *Plant Physiol* 133, 448-455.

42. Rorth, P., Szabo, K., Bailey, A., Laverty, T., Rehm, J., Rubin, G.M., Weigmann, K., Milan, M., Benes, V., Ansorge, W., et al. (1998). Systematic gain-of-function genetics in *Drosophila*. *Development* 125, 1049-1057.

43. Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000). The genome sequence of *Drosophila* melanogaster. *Science* 287, 2185-2195.

44. Lacchini, A.H., Everington, M.L., Augousti , A.T., and Walker, A.J. (2007). Use of *C. Elegans* as a model organism for sensing the effects of ELF-EMFs. *J. Phys.: Conf. Ser.* 76, 012027.

45. Mendoza, L.G., McQuary, P., Mongan, A., Gangadharan, R., Brignac, S., and Eggers, M. (1999). High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA). *Biotechniques* 27, 778-780, 782-776, 788.

46. Eggert, U.S., and Mitchison, T.J. (2006). Small molecule screening by imaging. *Curr Opin Chem Biol* 10, 232-237.

47. Guy, R.L., Gonias, L.A., and Stein, M.A. (2000). A fluorescence microscopy based genetic screen to identify mutants altered for interactions with host cells. *J Microbiol Methods* 42, 129-138.

48. Taylor, D.L. (2010). A personal perspective on high-content screening (HCS): from the beginning. *J Biomol Screen* 15, 720-725.

49. Khersonsky, S.M., Jung, D.W., Kang, T.W., Walsh, D.P., Moon, H.S., Jo, H., Jacobson, E.M., Shetty, V., Neubert, T.A., and Chang, Y.T. (2003). Facilitated forward chemical genetics using a tagged triazine library and zebrafish embryo screening. *J Am Chem Soc* 125, 11804-11805.

50. Ahn, Y.H., and Chang, Y.T. (2007). Tagged small molecule library approach for facilitated chemical genetics. *Acc Chem Res* 40, 1025-1033.

51. McPherson, M., Yang, Y., Hammond, P.W., and Kreider, B.L. (2002). Drug receptor identification from multiple tissues using cellular-derived mRNA display libraries. *Chem Biol* 9, 691-698.

52. Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat Chem Biol* 3, 570-575.

53. Zhu, S., Wurdak, H., Wang, J., Lyssiotis, C.A., Peters, E.C., Cho, C.Y., Wu, X., and Schultz, P.G. (2009). A small molecule primes embryonic stem cells for differentiation. *Cell Stem Cell* 4, 416-426.

54. Webb, Y., Zhou, X., Ngo, L., Cornish, V., Stahl, J., Erdjument-Bromage, H., Tempst, P., Rifkind, R.A., Marks, P.A., Breslow, R., et al. (1999). Photoaffinity labeling and mass spectrometry identify ribosomal protein S3 as a potential target for hybrid polar cytodifferentiation agents. *J Biol Chem* 274, 14280-14287.

55. MacKinnon, A.L., Garrison, J.L., Hegde, R.S., and Taunton, J. (2007). Photo-leucine incorporation reveals the target of a cyclodepsipeptide inhibitor of cotranslational translocation. *J Am Chem Soc* 129, 14560-14561.

56. Wurdak, H., Zhu, S., Min, K.H., Aimone, L., Lairson, L.L., Watson, J., Chopiuk, G., Demas, J., Charette, B., Halder, R., et al. (2010). A small molecule accelerates neuronal differentiation in the adult rat. *Proc Natl Acad Sci U S A* 107, 16542-16547.

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

58. Becker, F., Murthi, K., Smith, C., Come, J., Costa-Roldan, N., Kaufmann, C., Hanke, U., Degenhart, C., Baumann, S., Wallner, W., et al. (2004). A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors. *Chem Biol* 11, 211-223.

59. Eyckerman, S., Verhee, A., der Heyden, J.V., Lemmens, I., Ostade, X.V., Vandekerckhove, J., and Tavernier, J. (2001). Design and application of a cytokine-receptor-based interaction trap. *Nat Cell Biol* 3, 1114-1119.

60. Caligiuri, M., Molz, L., Liu, Q., Kaplan, F., Xu, J.P., Majeti, J.Z., Ramos-Kelsey, R., Murthi, K., Lievens, S., Tavernier, J., et al. (2006). MASPIT: three-hybrid trap for quantitative proteome fingerprinting of small molecule-protein interactions in mammalian cells. *Chem Biol* 13, 711-722.

61. Jung, J.H., Shim, J.S., Park, J., Ha, H.J., Kim, J.H., Kim, J.G., Kim, N.D., Yoon, J.H., and Kwon, J. (2009). *Proteomics Clin. Appl.* 3, 423-432.

62. Hammond, P.W., Alpin, J., Rise, C.E., Wright, M., and Kreider, B.L. (2001). *In vitro* selection and characterization of Bcl-X(L)-binding proteins from a mix of tissue-specific mRNA display libraries. *J Biol Chem* 276, 20898-20906.

63. Lipovsek, D., and Pluckthun, A. (2004). *In-vitro* protein evolution by ribosome display and mRNA display. *J Immunol Methods* 290, 51-67.

64. Liu, R., Barrick, J.E., Szostak, J.W., and Roberts, R.W. (2000). Optimized synthesis of RNA-protein fusions for *in vitro* protein selection. *Methods Enzymol* 318, 268-293.

65. Kino, T., Hatanaka, H., Hashimoto, M., Nishiyama, M., Goto, T., Okuhara, M., Kohsaka, M., Aoki, H., and Imanaka, H. (1987). FK-506, a novel immunosuppressant isolated from a Streptomyces. I. Fermentation, isolation, and physico-chemical and biological characteristics. *J Antibiot (Tokyo)* 40, 1249-1255.

66. MacBeath, G., and Schreiber, S.L. (2000). Printing proteins as

microarrays for high-throughput function determination. *Science* 289, 1760-1763.

67. Zhu, H., and Snyder, M. (2003). Protein chip technology. *Curr Opin Chem Biol* 7, 55-63.

68. Astle, J.M., Simpson, L.S., Huang, Y., Reddy, M.M., Wilson, R., Connell, S., Wilson, J., and Kodadek, T. (2010). Seamless bead to microarray screening: rapid identification of the highest affinity protein ligands from large combinatorial libraries. *Chem Biol* 17, 38-45.

69. Huang, J., Zhu, H., Haggarty, S.J., Spring, D.R., Hwang, H., Jin, F., Snyder, M., and Schreiber, S.L. (2004). Finding new components of the target of rapamycin (TOR) signaling network through chemical genetics and proteome chips. *Proc Natl Acad Sci U S A* 101, 16594-16599.

70. Lomenick, B., Hao, R., Jonai, N., Chin, R.M., Aghajan, M., Warburton, S., Wang, J., Wu, R.P., Gomez, F., Loo, J.A., et al. (2009). Target identification using drug affinity responsive target stability (DARTS). *Proc Natl Acad Sci U S A* 106, 21984-21989.

71. Lomenick, B., Olsen, R.W., and Huang, J. (2011). Identification of direct protein targets of small molecules. *ACS Chem Biol* 6, 34-46.

72. Swoboda, J.G., Meredith, T.C., Campbell, J., Brown, S., Suzuki, T., Bollenbach, T., Malhowski, A.J., Kishony, R., Gilmore, M.S., and Walker, S. (2009). Discovery of a small molecule that blocks wall teichoic acid biosynthesis in *Staphylococcus aureus*. *ACS Chem Biol* 4, 875-883.

73. Won, J., Kim, M., Yi, Y.W., Kim, Y.H., Jung, N., and Kim, T.K. (2005). A magnetic nanoprobe technology for detecting molecular interactions in live cells. *Science* 309, 121-125.

74. Won, J., Kim, M., Kim, N., Ahn, J.H., Lee, W.G., Kim, S.S., Chang, K.Y., Yi, Y.W., and Kim, T.K. (2006). Small molecule-based reversible reprogramming of cellular lifespan. *Nat Chem Biol* 2, 369-374.

75. Nakai, R., Iida, S., Takahashi, T., Tsujita, T., Okamoto, S., Takada, C., Akasaka, K., Ichikawa, S., Ishida, H., Kusaka, H., et al. (2009). K858, a novel inhibitor of mitotic kinesin Eg5 and antitumor agent, induces cell death in cancer cells. *Cancer Res* 69, 3901-3909.

76. Nagumo, Y., Kakeya, H., Shoji, M., Hayashi, Y., Dohmae, N., and Osada, H. (2005). Epolactaene binds human Hsp60 Cys442 resulting in the inhibition of chaperone activity. *Biochem J* 387, 835-840.

77. Misra-press, A., McMillan, M., Cudaback, E., Qabar, M., Ruan, F., Nguyen, M., Vaisar, T., Nakanishi, H., and Kahn, M. (2002). Identification of a Novel Inhibitor of the NF-kB Pathway. *Curr. Med. Chem. – Anti-Inflammatory & Anti-Allergy Agents* 1, 29-39.

78. Bedalov, A., Gatbonton, T., Irvine, W.P., Gottschling, D.E., and Simon, J.A. (2001). Identification of a small molecule inhibitor of Sir2p. *Proc Natl Acad Sci U S A* 98, 15113-15118.

79. Butcher, R.A., and Schreiber, S.L. (2003). A small molecule suppressor of FK506 that targets the mitochondria and modulates ionic balance in *Saccharomyces cerevisiae*. *Chem Biol* 10, 521-531.

80. Butcher, R.A., and Schreiber, S.L. (2004). Identification of Ald6p as the target of a class of small-molecule suppressors of FK506 and their use in network dissection. *Proc Natl Acad Sci U S A* 101, 7868-7873.

81. Grozinger, C.M., Chao, E.D., Blackwell, H.E., Moazed, D., and Schreiber, S.L. (2001). Identification of a class of small molecule inhibitors of the sirtuin family of NAD-dependent deacetylases by phenotypic screening. *J Biol Chem* 276, 38837-38843.

82. Zhao, Y., Dai, X., Blackwell, H.E., Schreiber, S.L., and Chory, J. (2003). SIR1, an upstream component in auxin signaling identified by chemical genetics. *Science* 301, 1107-1110.

83. Asami, T., Mizutani, M., Fujioka, S., Goda, H., Min, Y.K., Shimada, Y., Nakano, T., Takatsuto, S., Matsuyama, T., Nagata, N., et al. (2001). Selective interaction of triazole derivatives with DWF4, a cytochrome P450 monooxygenase of the brassinosteroid biosynthetic pathway, correlates with brassinosteroid deficiency in planta. *J Biol Chem* 276, 25687-25691.

84. Drevs, J., Hofmann, I., Hugenschmidt, H., Wittig, C., Madjar, H., Muller, M., Wood, J., Martiny-Baron, G., Unger, C., and Marme, D. (2000). Effects of PTK787/ZK 222584, a specific inhibitor of vascular endothelial growth factor receptor tyrosine kinases, on primary tumor, metastasis, vessel density, and blood flow in a murine renal cell carcinoma model. *Cancer Res* 60, 4819-4824.

85. Moon, H.S., Jacobson, E.M., Khersonsky, S.M., Luzung, M.R., Walsh, D.P., Xiong, W., Lee, J.W., Parikh, P.B., Lam, J.C., Kang, T.W., et al. (2002). A novel microtubule destabilizing entity from orthogonal synthesis of triazine library and zebrafish embryo screening. *J Am Chem Soc* 124, 11608-11609.

86. Kao, R.Y., Tsui, W.H., Lee, T.S., Tanner, J.A., Watt, R.M., Huang, J.D., Hu, L., Chen, G., Chen, Z., Zhang, L., et al. (2004). Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics. *Chem Biol* 11, 1293-1299.

87. Rosania, G.R., Chang, Y.T., Perez, O., Sutherlin, D., Dong, H., Lockhart, D.J., and Schultz, P.G. (2000). Myoseverin, a microtubule-binding molecule with novel cellular effects. *Nat Biotechnol* 18, 304-308.

88. Perez, O.D., Chang, Y.T., Rosania, G., Sutherlin, D., and Schultz, P.G. (2002). Inhibition and reversal of myogenic differentiation by purine-based microtubule assembly inhibitors. *Chem Biol* 9, 475-483.

89. Haggarty, S.J., Mayer, T.U., Miyamoto, D.T., Fathi, R., King, R.W., Mitchison, T.J., and Schreiber, S.L. (2000). Dissecting cellular processes using small molecules: identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. *Chem Biol* 7, 275-286.

90. Haggarty, S.J., Koeller, K.M., Wong, J.C., Grozinger, C.M., and Schreiber, S.L. (2003). Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation. *Proc Natl Acad Sci U S A* 100, 4389-4394.

91. Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., and Mitchison, T.J. (1999). Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 286, 971-974.

92. Williams, D., Jung, D.W., Khersonsky, S.M., Heidary, N., Chang, Y.T., and Orlow, S.J. (2004). Identification of compounds that bind mitochondrial F1F0 ATPase by screening a triazine library for correction of albinism. *Chem Biol* 11, 1251-1259.

93. Emami, K.H., Nguyen, C., Ma, H., Kim, D.H., Jeong, K.W., Eguchi, M., Moon, R.T., Teo, J.L., Kim, H.Y., Moon, S.H., et al. (2004). A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. *Proc Natl Acad Sci U S A* 101, 12682-12687.

94. Chong, T., McMillan, M., Teo, J.L., Henderson, J.W.R., and Kahn, M. (2004). Chemogenomic Investigation of AP-1 Transcriptional Regulation of LTC4 Synthase Expression. *Lett Drug Des Discovery* 1, 211-216.

95. Jonathan, B.W. (2002)). An approach using 'chemical genetics' has identified small-molecule agonists of the Hedgehog signaling pathway that may lead the way to drugs for chronic degenerative diseases. *Journal of Biology* 1, 1-7.

96. Peterson, J.R., Lokey, R.S., Mitchison, T.J., and Kirschner, M.W. (2001). A chemical inhibitor of N-WASP reveals a new mechanism for targeting protein interactions. *Proc Natl Acad Sci U S A* 98, 10624-10629.

97. Peterson, J.R., Bickford, L.C., Morgan, D., Kim, A.S., Ouerfelli, O., Kirschner, M.W., and Rosen, M.K. (2004). Chemical inhibition of N-WASP by stabilization of a native autoinhibited conformation. *Nat Struct Mol Biol* 11, 747-755.

98. Verma, R., Peters, N.R., D'Onofrio, M., Tochtrop, G.P., Sakamoto, K.M., Varadan, R., Zhang, M., Coffino, P., Fushman, D., Deshaies, R.J., et al. (2004). Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. *Science* 306, 117-120.

99. Wignall, S.M., Gray, N.S., Chang, Y.T., Juarez, L., Jacob, R., Burlingame, A., Schultz, P.G., and Heald, R. (2004). Identification of a novel protein regulating microtubule stability through a chemical approach. *Chem Biol* 11, 135-146.

100. Min, J., Kyung Kim, Y., Cipriani, P.G., Kang, M., Khersonsky, S.M., Walsh, D.P., Lee, J.Y., Niessen, S., Yates, J.R., 3rd, Gunsalus, K., et al. (2007). Forward chemical genetic approach identifies new role for GAPDH in insulin signaling. *Nat Chem Biol* 3, 55-59.

101. Gao, M., Nettles, R.E., Belema, M., Snyder, L.B., Nguyen, V.N., Fridell, R.A., Serrano-Wu, M.H., Langley, D.R., Sun, J.H., O'Boyle, D.R., 2nd, et al. (2010). Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature* 465, 96-100.

102. Lee, M.Y., Kim, M.H., Kim, J., Kim, S.H., Kim, B.T., Jeong, I.H., Chang, S., and Chang, S.Y. (2010). Synthesis and SAR of sulfonyl- and phosphoryl amidine compounds as anti-resorptive agents. *Bioorg Med Chem Lett* 20, 541-545.

103. Chang, S.Y., Bae, S.J., Lee, M.Y., Baek, S.H., Chang, S., and Kim, S.H. (2011). Chemical affinity matrix-based identification of prohibitin as a binding protein to anti-resorptive sulfonyl amidine compounds. *Bioorg Med Chem Lett* 21, 727-729.

104. Zhang, Q., Major, M.B., Takanashi, S., Camp, N.D., Nishiya, N., Peters, E.C., Ginsberg, M.H., Jian, X., Randazzo, P.A., Schultz, P.G., et al. (2007). Small-molecule synergist of the Wnt/beta-catenin signaling pathway. *Proc Natl Acad Sci U S A* 104, 7444-7448.

105. Chen, S., Do, J.T., Zhang, Q., Yao, S., Yan, F., Peters, E.C., Scholer, H.R., Schultz, P.G., and Ding, S. (2006). Self-renewal of embryonic stem cells by a small molecule. *Proc Natl Acad Sci U S A* 103, 17266-17271.

106. Choi, Y., Shimogawa, H., Murakami, K., Ramdas, L., Zhang, W., Qin, J., and Uesugi, M. (2006). Chemical genetic identification of the IGF-linked pathway that is mediated by STAT6 and MFP2. *Chem Biol* 13, 241-249.