Phenotypes and targets-based chemical biology investigations in cancers

Guo-Qiang Chen*, Ying Xu, Shao-Ming Shen, Jian Zhang

Affiliation notes

Department of Pathophysiology, Key Laboratory of Cell Differentiation and Apoptosis of Chinese Ministry of Education, Shanghai Jiao Tong University School of Medicine (SJTU-SM), Shanghai 200025, China.

*Corresponding author: chengq@shsmu.edu.cn

Abstract

Chemical biology has been attracting wide interests because eminent roles of chemical methods and techniques in helping to decipher and manipulate biological systems. Although chemical biology involves a wide field, this review prefers to focus on chemical biology aiming to use exogenous chemical probes to interrogate, modify, and manipulate biological processes at the cellular and organismal levels in a highly controlled and dynamic manner. In this direction, many advances have been achieved for cancer biology and therapeutics from target identification and validation based on active anti-cancer compounds (forward approach) to discoveries of anti-cancer molecules based on some important targets including protein-protein interaction (reverse approaches). Herein I attempts to summarize some recent progresses mainly from China through applying chemical biology approaches to explore molecular mechanisms of carcinogenesis. Additionally, we also outline several new strategies for chemistry to probe cellular activities such as proximity-
dependent labeling methods for identifying protein-protein interactions, genetically
encoded sensors, and light activating or repressing gene expression system.

Keywords

Chemical Biology, chemical compounds, target, cancer, leukemia

INTRODUCTION

Chemistry has been presenting eminent roles in helping to decipher and
manipulate life activities, and the interdependency of chemistry with biology,
pharmacology and medicine has been shown to be of great synergistic values[1,2].
For example, discoveries of potent chemical probes JQ1[3] and I-BET (Inhibitor for
bromodomain and extra terminal) [4] are triggering the revolutionary progresses in
our understanding of bromodomain biology and pharmacology[5,6]. As a thriving
interdisciplinary scientific area, hence, chemical biology has been attracting wide
interests[7].

Different from classical biochemistry that focuses on the understanding of
endogenous chemical processes in living systems, chemical biology employs
methods and techniques of chemistry to investigate biological phenomena.
Especially, chemical biology aims to use exogenous chemical probes to interrogate,
modify, and manipulate biological processes at the cellular and organismal levels in
a highly controlled, reversible and dynamic manner. By analogy to classical
genetics, chemical biology can also employ the forward and reverse approaches
(Fig. 1). Forward or phenotype-based chemical biology begins from screening
compounds to trigger interesting phenotypes of cells or organisms, and the
biological target(s) of the interesting compound is identified. Reciprocally, the
reverse or target-based chemical biology usually starts with known targets, which has been validated to play a critical role in a particular signaling pathway, biological activity or disease of interest.

Chemical biology has also being appreciated by scientists in China. The National Climbing Program for basic research program initiated in the late 1980s, the precedent of the National Basic Research (973) Program of China, funded a project entitled “chemical studies in biological processes” [8]. From then on, scientists in China made greater contribution to chemical biology research, especially forward chemical biology based on screening of active compounds, such as application and target identification of all-trans retinoic acid (ATRA) and arsenic trioxide (As$_2$O$_3$) in the treatment of acute promyelocytic leukemia (APL), a unique subtype of acute myeloid leukemia (AML)[9,10], and Chemical induction of pluripotent stem cells from mouse somatic cells [11-13].

To push the development of chemical biology as a multidisciplinary research priority area, the National Natural Science Foundation of China (NSFC) launched a major research program on chemical biology in 2005, named “investigations on signal transduction processes utilizing small chemical probes”, which focused on development of new techniques and methods to detect the information of signaling processes, explore chemical compounds-based signaling mechanisms of cellular functions, and discover targets and lead compounds based on signal transduction processes[8]. Following the program, a new plan entitled “dynamic modification of biological macromolecules and chemical interference” was initiated by NSFC in 2017.
Due to extremely high morbidity and mortality worldwide, cancers are always the main focus of medical research. In the past two decades, increasing numbers of potential therapeutic targets for cancers were generated by an exponential growth in genomic information with the support of data from various kinds of other omics such as epigenomics, transcriptomics, proteomics, metabolomics as well as synthetic lethal screens based on RNA interference (RNAi), CRISPR/Cas systems and classical hypothesis-driven approaches[14-16]. However, it is estimated that approved drugs are available for only 5% of the 500 or so cancer-causing gene products [17,18]. Available chemical modulators (inhibitors for oncogenic driver and activators for dysfunctional tumor-suppressor proteins) only applied for less than 10% of the cancer proteome. Moreover, at least 10% of pathogenic cancer driver genes are considered to be druggable with current technologies, but they are not yet chemically explored. Therefore, there is an urgent medical need to extend the chemical targeting of the cancers, which will be essential to help understand the functioning of cancer-related genes and pathogenic networks and to develop personalized, precision therapeutic strategies for cancer patients. This review tried to highlight some but not all examples of recent progresses, mainly from China, in the discovery of chemical tools and application of chemical biology approaches in cancers.

CHEMICAL LIBRARY

The chemical or drug library is a collection of stored chemicals usually used in high-throughput screening systems for identifying chemical probes of disease-related targets and potential starting points for drug discovery. Each chemical library should have associated information collected in a form of database, such as chemical structure, purity, quantity and known potential biological activities. A
rational chemical library with wide chemical structure space will increase the chance to find a “hit” in the high-throughput screening. Considering that many active chemicals identified through high-throughput screening cannot be druggable, drug-like properties such as solubility and bioavailability are important to chemicals collected in a drug library.

Chemical compounds include natural and synthetic compounds. Natural compounds or products derived from plants such as traditional Chinese medicine (TCM) herbs, minerals, microorganisms and animals have been used as valuable sources for clinical drugs or chemical probes to find mechanisms for life activities. Although natural compounds contain common disadvantages such as access and supply, processing complexities of natural product chemistry and inherent development delays, and entangled interests regarding intellectual property rights, they have been attracting significant attention for the development of novel chemotherapeutics to cancers because of their remarkable efficacy and generally low toxicity. As estimated[19], more than half of anti-cancer drugs approved by the Food and Drug Administration of American (FDA) between 1981 and 2014 originated from natural products and/or their derivatives.

Natural products alone cannot build a huge library with diverse structures. The synthesis of natural products has been one of the mainstays of organic chemistry over a century, and great progresses in the synthesis of natural product analogues have been achieved[1]. Combinatorial chemistry allows for the synthesis of vast numbers of compounds. Diversity-oriented synthesis (DOS) can provide an efficient manner to generate such a library[20]. By using the DOS library, many active molecules have been discovered to modulate protein-protein interactions (PPI), transcription factor activity and multidrug resistance, as reviewed[21]. Moreover,
the synthetic approach enables the mass production of some rare natural products, as well as the optimized version of primary active compounds identified in a high-throughput screening.

A chemogenomic library is a relatively small library containing hundreds-to-thousands (rather than millions) of selective small molecules with known or potential targets or functions, of which majority target G-protein-coupled receptors (GPCRs), kinases and ion channels, the most common molecular targets for drug discovery [22]. Using such a library can greatly increase the opportunity to repurpose a drug that acts on a novel pathway or target[23]. It should be pointed out that dark chemical matter (DCM), those small molecules in a screening collection that have never shown biological activity despite having been exhaustively tested in high-throughput screening, was also reported to occasionally result in potent hits with unique activity and clean safety profiles, which makes them valuable starting points for lead optimization efforts[7,24].

Here, we want to introduce a small compound library from Wu Q’s group in Xiamen University, which specifically targets to the orphan nuclear receptor Nur77 (also known as NR4A1, NGFIB, TIS1, NAK-1, TR3, or N10), a member of the nuclear receptor NR4A subfamily. Nur77 is an immediate early gene-encoded unique transcription factor that is rapidly and transiently induced in response to changes in the extracellular environments. In addition to its transcriptional function, Nur77 also presents a non-genomic signaling function through its physical interactions with various signaling proteins, such as p53, hypoxia inducible factor (HIF)-1α, protein arginine N-methyltransferase 1 (PRMT1), protein kinase C (PKC), retinoid X Receptor (RXR) and Wnt (wingless)/β-catenin signaling, thereby
modulating a wild range of important biological functions, including cell cycle progression, apoptosis, autophagy, inflammation, metabolism, and energy homeostasis. Nur77 expression and signaling transduction are regulated in many cancers, which provide an important molecule target for drug screening[25]. Wu Q’s group identified cytosporone-B (Csn-B, Fig. 2) extracted from a mangrove endophytic fungus as the first naturally occurring agonist for Nur77[26]. Based on the Csn-B structure, they designed and synthesized more than 300 derivatives to construct a unique library of small-molecule compounds that specifically target Nur77. Thereafter, they extensively explored the molecular mechanisms of Nur77 in the regulations of glucose metabolism, autophagy, inflammation and carcinogenesis. For examples, they reported that the chemical compound TMPA can bind to the ligand binding domain (LBD) of Nur77, causing Nur77 conformational change and resulting in the disruption of association of Nur77 with liver kinase B1 (LKB1), the latter making an important role in governing energy homeostasis by regulating the activity of the energy sensor the AMP-activated protein kinase (AMPK). As a result, LKB1 is released to the cytoplasm to phosphorylate and activate AMPK, finally downregulating glucose level in diabetes mice (Fig. 2)[27]. They also demonstrated that the compound THPN from their compound library triggers Nur77 into the mitochondria by Nix interaction, where Nur77 is located in the mitochondrial inner membrane and interacts with ANT1, and then causes the opening of mitochondrial permeability transition pore and the depolarization of mitochondrial membrane, eventually leading to the irreversible autophagic death of melanoma cells. In a mouse model of spontaneous skin melanoma, they confirmed THPN-induced autophagic cell death in a Nur77-dependent manner, which further inhibits the development and metastasis of melanoma (Fig. 2)[28]. This work not only
elucidated the new mechanism by which Nur77 participates in the autophagic cell death induction through mitochondrial signaling pathway, but also demonstrated that inducing autophagic death of melanoma cells by THPN can overcome the resistance of melanoma cells to drugs-inducing apoptosis. Additionally, the small-molecule compound PDNPA obviously inhibits the transcriptional activity of NF-κB in a Nur77-dependent manner (Fig. 2), thus effectively activating the Nur77 anti-inflammatory function [29].

PHENOTYPE-BASED/ FORWARD CHEMICAL BIOLOGY

Following the chemical library and active compound screening, the forward chemical biology calls for the development of powerful methods for the identification of one or a few cellular targets from the complex mixture of biomolecules present in cells, including proteins, nucleic acids, carbohydrates, or lipids[30-33]. Thus, the search for cellular targets may well be regarded as a quest for a needle in a giant haystack.

Cell-based screening

With the advantage of high throughput and low cost, cell-based screening assay has been widely used in chemical biologic investigation and drug discovery. Cells most used in such an assay are different cancer cell lines or primary cells directly derived from patients. Also, the merging in vitro 3D tissues called organoids are presenting great potentials in compound screen and drug discovery[34]. The phenotypes used in the cell-based assay can refer to the hallmarks of cancer, and cytotoxicity, cell cycle, differentiation, invasion and, migration and drug resistance are main test indicators used in the cell-based screening assay. For high-throughput screening, however, it is better and appropriate that phenotypic changes are
recorded by the reporter-gene activity, fluorescence signal or imaging methods. For examples, rapid tumor growth leads to excessive oxygen demand and a hypoxic environment in most solid tumors, which adapt to the hypoxia environment by upregulating the transcription of target genes that regulate cell proliferation, angiogenesis, energy metabolism, apoptosis resistance, and metastasis. These processes are generally associated with the transcriptional factor HIF-1, which is upregulated in the tumor and is a promising target for cancer chemotherapy[35,36]. Though no HIF-1 inhibitor is clinically available to date, a lot of effort has been applied during the last decade in search of potent HIF-1 inhibitors, as recently reviewed[37].

Here I give examples for a reporter-gene activity-guiding phenotypic screening based on Wnt/β-catenin signaling. The signaling is a highly conserved pathway in organism evolution and regulates many biological processes, and its aberrant activation is closely related to tumor progression [38]. For example, the pathway is frequently activated in colorectal cancer (CRC) as a result of the mutation of adenomatous polyposis coli (APC). In a screening of FDA-approved drugs using the Top-flash (Wnt/β-catenin pathway-responsive firefly luciferase) assay, nitazoxanide (NTZ, Fig. 1), a clinically approved antiparasitic drug, was shown to block Wnt/β-catenin signaling in vitro and in a murine model for familial adenomatous polyposis, Apc^min/+ mice, which spontaneously generate tumor. They also showed that NTZ promotes β-catenin degradation independent upon glycogen synthase kinase 3β (GSK3β) or APC[39]. With the similar screening system, Li L’s group in Shanghai Institute of Biological Science of Chinese Academy of Science identified 15-oxospiramilactone (NC043), a diterpenoid derivative, to inhibit Wnt3a or LiCl-stimulated Top-flash reporter activity in HEK293T cells and growth of CRC cells,
SW480 and Caco-2, together with decrease in the mRNA and/or protein expression of Wnt target genes Axin2 and cyclin D1. Further, NC043 did not affect the cytosol-nuclear distribution and protein level of soluble β-catenin, but decreased β-catenin/TCF4 association in SW480 cells[40]. More recently, they showed that NC043 directly binds to CARF (the collaborator of ARF) through forming a covalent bond with the Cys\textsuperscript{516} residue of CARF, and disrupts the CARF-Dishevelled (Dvl) interaction, thereby promoting Wnt signaling activation[21]. This group also screened a synthetic chemical library of lycorine derivatives, and identified small molecule compound HLY78 as an activator of the Wnt/β-catenin signaling pathway in a Wnt ligand-dependent manner. Shortly, HLY78 targets the DIX domain of Axin and potentiates the Axin-LRP6 (low-density lipoprotein receptor-related protein 6) association, thus promoting LRP6 phosphorylation and Wnt signaling transduction[41].

**Animal-based screening**

Comparing with cell model, animal model has the advantage of high relevance with disease, because it can mimic pathophysiological features similar to that of patients. In spite of the disadvantage with low-throughput, high cost and time-consuming, animal-based screening is indispensable for the pre-clinic research of efficacy, side effect and toxicity of a novel potential drug. Small animals such as zebrafish, drosophila, xenopus tadpole are mostly used for the animal-based screening. Using a *drosophila* Ras-driven tumor model with a large-scale screening with 2,000 compounds, the glutamine analogue acivicin was finally identified as a potent and specific inhibitor of *Drosophila* tumor formation[42]. After screening with 26,400 molecules in a T-cell reporting zebrafish model, a novel molecule named lenaldekar was identified as a selective toxicity against leukemia[43]. The
compound has similar effects on the human leukemia samples and xenograft mice, indicating a high relevance between zebrafish and human. Kalin et al[44] employed xenopus tadpole embryos as an in vivo model to identify novel compounds involved in angiogenesis and lymphangiogenesis through a simple phenotypic read-out (edema formation or larval lethality), and identified 32 interfering with blood vascular and/or lymphatic development in xenopus from 1280 bioactive compounds.

**Target identification**

Phenotypic screening is often the most straightforward way to discover relevant bioactive compounds usually with unknown molecular targets. Some researchers may argue that target identification for hits from phenotypic screening is unnecessary if the relevant readout from the phenotypic assay is fairly reliable. However, target identification and validation of bioactive small molecules are the essential and often decisive step for understanding action mechanism of compound and for further compound optimization[30,45]. Through the development and availability of several new experimental techniques, in principle, target identification is feasible and the number of successful examples steadily grow as reviewed[46].

Bioactive compounds may interact with off-target proteins, resulting in undesired biological activity or toxicity[47]. We can optimize the compound with more powerful activity and less toxicity by knowing both therapeutic targets and off-targets. Recent technological advances in genomics, proteomics and bioinformatics have accelerated the process of target identification. Proteomics and genomics-based approaches provide with powerful tools for the all-round identification[48]. Quantitative mass spectrometry (MS) techniques greatly
enhanced the sensitivity of target protein detection[49]. Protein microarrays have greatly simplified the process of target identification[50]. Up to now, a number of technologies have been explored to identify targets from phenotypic screens.

Among the techniques currently available in small molecule target identification, protein affinity isolation using suitable small-molecule probes (pulldown) and subsequent MS analysis of the isolated proteins appears to be most powerfully and most extensively used, aspiring to identify the full protein-binding spectrum of a compound. After a structure activity relationship (SAR) study, an active compound is conjugated with a specific affinity tag such as polyethylene glycol and biotin at an appropriate position without affecting the compound’s activity. Then the compound-tag is incubated with cells or total cell lysates, followed up by the capture of target proteins using a specific solid matrix. Lastly, target proteins are revealed by MS. Although affinity-based approaches can provide an effective method for target identification, it is a challenge to find the appreciate position of active compound for the conjugation of an affinity tag without affecting the compound activity. Besides, the affinity tag somehow alters the structure of the original compound, resulting in false positive and negative targets.

Protein microarrays also provide a high throughput and high sensitive method for the target identification of small molecules. As for this, desired proteins (up to thousands) are immobilized on a treated glass microscope slide to generate a protein microarray[51]. Small molecules need to be labeled with a reporting tag (fluorescence, biotin or isotope), and then they are incubated with the protein microarray, followed up by washing and signal developing. Huang et al[52] identified small-molecule inhibitors and enhancers of rapamycin through a yeast proteome microarray. By using a protein-domain microarray of human methyllysine
effector molecules and biotin-labeled UNC1251 analogs, Bae et al.[53] identified EML405 as an inhibitor of tudor-domain-containing protein Spindlin 1 (SPIN1). Regardless of the type of labeling, protein microarrays owe the same disadvantage of affinity-based approaches, that is, the labeled small molecule may interfere with the small molecule-protein interaction.

Label-free approaches are relatively simple and direct approaches which do not require any chemical modification of an active compound. This kind of strategy relies on the principle that the protein becomes susceptible to proteolysis once it binds to a small molecule[32,54]. The drug affinity responsive target stability (DARTS) is one of such label-free approaches with the principle that small-molecule binding proteins are protected and enriched during proteolysis, and it has been used to successfully identify cellular targets for several active compounds [31]. After the above-mentioned finding of NTZ for its inhibition of Wnt signaling, for instance, Qu et al.[39] profiled proteins that bind directly using DARTS and found that NTZ directly interacts with peptidyl arginine deiminase 2 (PAD2), an enzyme that catalyzes the conversion of the protein arginine residue to citrulline (a post-translational modification called deamination or citrullination) and citrullinates β-catenin to promotes β-catenin degradation.

A detailed example for target identification with the needle in the haystack

Leukemia is an aggressive and heterogeneous disorder of malignant hematopoiesis that occurs worldwide. In the past thirty years, many important advances had been achieving in the biological, molecular and cytogenetic aspects of leukemia. It has been widely understood that various kinds of leukemias present specific cytogenetic alterations, especially chromosome translocations, which
generate abnormal oncogenic fusion proteins. These alterations disrupt the normal signaling of hematopoietic development and cause uncontrolled proliferation, blocked differentiation and/or damaged apoptosis of malignant hematopoietic cells. In the past years, some natural compounds from TCM and synthetic small compounds were investigated for their anti-leukemia activity[55-57]. Adenanthin, a kind of ent-kaurene diterpenoids, was originally isolated from the leaves of *Rabdosia adenantha*. The previous investigations demonstrated that diterpenoids have a wide spectrum of biological activities such as anti-tumor, anti-inflammatory and significant cardiovascular effects. Up to now, more than 600 diterpenoids have been found in China. Previously, we also identified pharicin B, a novel natural ent-kaurene diterpenoid derived from *Isodon pharicus* leaves, to rapidly stabilize RARα (*retinoic acid receptor α*) protein in various subtypes of AML cells, especially APL, and thus to present a synergistic or additive differentiation-enhancing effect in combination with ATRA[58]. We also found that pharicin A, another new ent-kaurene diterpenoid, induces mitotic arrest in leukemia and solid tumor-derived cells, which is associated with unaligned chromosomes, aberrant BubR1 localization and deregulated spindle checkpoint activation[59]. With the encouragement from these works, we used a cell-based phenotypic assay to screen up to 400 natural ent-kaurene diterpenoids which were provided and isolated by Sun HD’s group in Kunming Institute of Botany of Chinese Academy of Sciences[60]. Among these diterpenoids, adenanthin was shown to decrease viability of APL cells at a concentration of more than 4 μmol, while it also induce the differentiation of APL cell lines and primary leukemic blasts from APL and non-APL AML patients[61,62]. Further investigations in ATRA-sensitive and ATRA-resistant APL transgenic mice models revealed that intravenous administration of adenanthin (5
mg kg\(^{-1}\) body weight, each day for five consecutive days a week) significantly induces differentiation and tumor regression, and prolongs the survival of both kind of leukemic mice\(^{[61]}\). Moreover, adenanthin also markedly eliminates APL-initiating progenitor cells (CD34\(^{+}\), c-kit\(^{+}\), FcγRIII/II\(^{+}\), and Gr1\(^{int}\)) in ATRA-sensitive leukemic mice. All these results strongly suggest that adenanthin has potential therapeutic efficacy on AML.

To further study the molecular mechanism of action (MMOA) of adenanthin, we tried to identify its potential protein target(s) via a chemical proteomic approach. After a SAR study of adenanthin, we synthesized a biotin-tagged adenanthin probe without affecting its differentiation-inducing activity. Thus, this biotin-tagged adenanthin probe was applied into the lysates of APL cell line NB4 cells, followed by precipitation with streptavidin-coated agarose beads. The bound proteins were run on an SDS-PAGE gel. On the gel with silver staining, only one detectable band at approximately 23 kD was clearly precipitated by biotin-adenanthin but not by free biotin, and the precipitated band could be competitively inhibited by higher concentrations of unlabeled adenanthin. Finally, mass spectrometry analysis revealed that the adenanthin-bound protein is peroxiredoxin (Prdx) I and Prdx II (Fig. 3)\(^{[61,62]}\).

Prdxs are a family of small nonseleno peroxidases that catalyze peroxide reduction of H\(_2\)O\(_2\), a reactive oxygen species which also play an important role as second messengers in cellular signaling pathways\(^{[63,64]}\). Prdxs have a conserved cysteine named the peroxidatic cysteine (C\(_P\)) that serves as the site of oxidation by peroxides. Peroxides oxidize the C\(_P\)-SH to C\(_P\)-SOH, which reacts with another cysteine typically named the resolving cysteine (C\(_R\)), forming a disulfide that is subsequently reduced by an appropriate electron donor. Depending on the location...
or absence of the Cᵣ, the mammalian Prdxs are classified into two-cysteine (Prdxs I–IV), atypical two-cysteine (Prdx V) and one-cysteine Prdx (Prdx VI) subfamilies. Thus, we further blotted the precipitates with antibodies against Prdx I–VI, all of which were expressed. The results showed that biotin-adenanthin effectively pulled Prdx I/II down but not Prdxs IV–VI, which was also confirmed by the in vitro recombinant Prdxs proteins. Next, we incubated the recombinant Prdx I protein with or without adenanthin followed by MS analysis to determine the adenanthin-modified specific residue, and the results revealed that adenanthin covalently modifies Cys¹⁷³, the Cᵣ of Prdx I. Remarkably, adenanthin also selectively bound the Cᵣ (Cys¹⁷²) of Prdx II. Accordingly, adenanthin effectively inhibited the peroxidase activity of recombinant Prdx II, and especially Prdx I.

To validate whether targeting Prdx I/II are directly associated with adenanthin-induced differentiation, we knocked down the expression of Prdx I or Prdx II in NB4 cells by specific small interfering RNAs, and found that knockdown of Prdx I or Prdx II induced NB4 cell differentiation. Our further studies showed that adenanthin treatment can moderately increase intracellular H₂O₂ level, and the elevated H₂O₂ activates signal-regulated protein kinases 1 and 2 (ERK1 and ERK2), consequently increasing expression of CCAAT/enhancer binding protein-β (C/EBPβ), which has been widely shown to drive AML cell differentiation (Fig. 3)[61].

More intriguingly, a recent work from Dr. Lei XG’s group in Peking University also showed that the natural product jungermannenone C (Fig. 3), a tetracyclic diterpenoid isolated from liverworts, induces AML cells to undergo differentiation through targeting Prdx I/II by selectively binding to the conserved cysteine residues and thus leading to cellular reactive oxygen species (ROS) accumulation (Fig.
More intriguingly, a recent report unveiled the biological function of the cyclin-dependent kinase 2 (CDK2)-Prdx II axis in blocking AML differentiation[66]. They showed that CDK2 undergoes ubiquitin-dependent proteasome degradation by specific E3 ubiquitin ligase KLHL6, which is accompanied by AML cell differentiation. Importantly, inhibiting CDK2 effectively induces granulocytic differentiation in AML cells and the differentiation blockade function of CDK2 may be achieved directly by maintaining the activity of Prdx II[66].

**TARGET-BASED SCREENING/REVERSE CHEMICAL BIOLOGY**

In target-based screening approach, the target of interest is usually recombinantly expressed in a purified system or in a cellular context, and then specific small chemicals with potential ability to modulate the activity of the target are screened through an appropriate *in vitro* assay. This approach is relatively easy to execute with less cost of time and money, comparing to phenotypic-based screening approach. But it still faces a big challenge, because an active compound identified *in vitro* assay may not work *in vivo*, due to the complex environment such as cell-membrane impenetrability of compound, post-translational modifications (PTMs) of target, undesired targets, compound metabolism of intact organisms. In spite of these, many activators or inhibitors of receptors and enzymes have been successfully identified through this screening approach. Especially, small molecule inhibitors of a number of recently identified protein targets offer new therapeutic options tailored to specific mutations or to counter resistance. Imatinib as the first tyrosine kinase inhibitor was discovered with high specificity for Bcr-Abl protein resulting from t(9, 22)-derived Philadelphia chromosome in chronic myeloid leukemia (CML). However, several Bcr-Abl-dependent and -independent mechanisms of
resistance to imatinib arose after becoming first-line therapy in CML patients. Consequently, new specific drugs, such as dasatinib, nilotinib, bosutinib, and ponatinib, were rationally designed and approved for clinic to override resistances [67]. On the basis of the finding that the NEDD8-activating enzyme subunit NAE1 is overexpressed in CML cells, the NAE1 inhibitor MLN4924 was also demonstrated to induce G2-M arrest and apoptosis in CML cells regardless of their T315I mutation status in Bcr-Abl, which offer a preclinical proof of concept for targeting protein neddylation as a novel therapeutic strategy to override mutations-derived imatinib resistance in CML[68]. With a c-Myc pathway-targeted screening of seven natural anticancer compounds, cryptotanshinone was identified as a dual inhibitor of pSTAT5 and pSTAT3 to effectively block IL-6-mediated STAT3 activation and reverse Bcr-Abl kinase-independent drug resistance in CML [69]. In the following, I like to provide several examples for targets-based discoveries of active compounds against cancers including leukemia.

**PDK1 inhibitor**

One of general hallmarks of malignancy is a unique metabolic profile of high aerobic glycolysis, a phenomenon of the enhanced conversion of glucose into lactate even in the presence of oxygen. The aerobic glycolysis, known as Warburg effect, confers a significant growth advantage of cancer cells by supplying essential ATP production, generating precursors for biosynthesis, and providing reducing equivalents for antioxidant defense. Recently, an increasingly recognized linkage between oncogenic proteins (for example, HIF-1 or its cooperation with dysregulated c-Myc) and pyruvate dehydrogenase kinase 1 (PDK1), a molecular switch that diminishes mitochondrial respiration and enhances aerobic glycolysis via phosphorylating and inactivating pyruvate dehydrogenase (PDH), and has
provided a glimpse into the molecular basis of the metabolic reprogramming in cancer cells. To discover new PDK1 inhibitors, Geng MY’s group in Shanghai Institute of Materia Medica of Chinese Academy of Sciences carried out a PDK1 enzymatic screen using an in-house small-molecule library composed of ∼600 commercially available known drugs, and thiram, an existing pesticide with anticancer activity, was found to be capable of remarkably inhibiting PDK1 activity. Based on thiram, their further chemical efforts led to discovery of a more potent new compound designated as bis(4-morpholinyl thiocarbonyl)-disulfide (JX06), which was identified as a selective covalent inhibitor of PDK1 in cancer cells[70]. JX06 forms a disulfide bond with the thiol group of a conserved cysteine residue (C240) based on recognition of a hydrophobic pocket adjacent to the ATP pocket of the PDK1 enzyme. With the covalent modification at C240, conformational changes at Arg286 through Van der Waals forces are induced, thereby hindering access of ATP to its binding pocket and in turn impairing PDK1 enzymatic activity. Notably, cells with a higher dependency on glycolysis were more sensitive to PDK1 inhibition[70], reflecting a metabolic shift that promotes cellular oxidative stress and apoptosis.

IDH1 inhibitor

Point mutations affecting isocitrate dehydrogenase 1 (IDH1) Arg132 (R132) and IDH2 Arg172 or Arg140 (R172 or R140) are driver mutations in AML and other cancers. A high-throughput biochemical screen targeting an IDH1 heterodimer composed of R132H mutant IDH1 and wild-type IDH1 identified a tetrahydropyrazolopyridine series of inhibitors, and additional structure optimization led to the identification of GSK321 as a highly potent inhibitor of mutant IDH1 enzymes, with IC50 values of 4.6 nM against R132H, 3.8 nM against R132C and 2.9
nM against R132G. GSK321 binds to an allosteric site and locks the IDH1 enzyme in a catalytically inactive conformation, thereby enabling inhibition of different clinically relevant IDH1 mutants. GSK321 treatment of primary AML cells with IDH1 mutant uniformly led to induction of granulocytic differentiation at the level of leukemic blasts and more immature stem-like cells, \textit{in vitro} and \textit{in vivo}, together with a decrease in intracellular oncometabolite 2-hydroxyglutarate (2HG) \cite{71}.

\textbf{SIRT6 activator}

One of PTMs is the lysine side chain $N^\epsilon$-acylation which can be reversed by deacylation of the resulting $N^\epsilon$-acyl-lysine. The sirtuins, which have seven sirtuin members (SIRT1-7), are able to catalyze the $N^\epsilon$-acyl-lysine deacylation reaction on histone and belong to a kind of acyltransferase which can catalyze the removal of the acyl group from lysine residues on histones and other non-histone protein substrates. The sirtuin-catalyzed deacylation reaction plays an important regulatory role in multiple crucial cellular processes such as transcription, DNA damage repair, genomic stability, cell cycle, apoptosis, inflammation, metabolism and caloric restriction. This reaction is also regarded as a current therapeutic target for human diseases such as cancers. More recently, a variety of chemical probes and modulators (inhibitors and activators) have been developed and some of them have been employed toward an enhanced mechanistic and functional understanding of the sirtuin-catalyzed deacylation reaction\cite{72}.

SIRT6, one of the most important sirtuin member which is widely expressed in almost all mammalian tissues, has been implicated in regulating several biological processes, including DNA repair, glucose/lipid metabolism, inflammation and aging. A series of studies have revealed that SIRT6 function as a tumor suppress
gene[73,74], but there are still no small-molecule activator specific to SIRT6 have been found to date. We hypothesized that SIRT6 can be activated via an allosteric mechanism. Thus, they adopted the allosite method[75] to predict allosteric sites for the activation of SIRT6, resulting in a pocket around Phe$^{82}$ and Phe$^{86}$ as a potential site. After a virtual docking with more than 5,000,000 compounds, top-ranked 20 compounds were selected out for experimentally validation. Two hits were confirmed to be able to increase the activity of SIRT6 deacetylation in the Fluor-de-Lys (FDL) assay. Subsequent optimization based on these two hits yielded two potent activators: MDL-800 and MDL-801, with EC$_{50}$ values of 10.3±0.3 μM and 5.7±0.023 μM, respectively. MDL-800 was proved to be a selective activator of SIRT6 among the 18 diverse histone deacetylase (HDAC) members. They also showed that MDL-800 can arrest the cell cycle of human hepatocellular carcinoma (HCC) cells and inhibit HCC in a xenograft mice model via enhancing the activity of SIRT6 deacetylation[76].

**SPOP inhibitor**

BTB domain-containing speckle-type POZ protein (SPOP), an adaptor of ubiquitin E3 ligase, plays important roles in development and tumorigenesis by mediating the ubiquitination of multiple substrates such as phosphatase and tensin homolog (PTEN) and dual specificity phosphatase 7 (DUSP7). The previous works showed that SPOP is overexpressed in virtually all clear-cell renal cell carcinoma (ccRCC), which accounts for about 75% of all RCC cases[77], and overexpressed SPOP accumulates in the cytoplasm of ccRCC cells, which is a nucleoprotein in normal cells. Furthermore, SPOP serves as a regulatory hub to promote ccRCC tumorigenesis through the ubiquitination and degradation of multiple regulators of cellular proliferation and apoptosis. These ideal characteristics make SPOP as a
potential antitumor target for treating ccRCC. Based on structural studies, Jiang HL’s group in Shanghai Institute of Materia Medica of Chinese Academy of Sciences performed computational screening through a hierarchical strategy combining pharmacophore modeling and molecular docking, and 109 compounds were selected from the SPECS database with \( \sim 200,000 \) drug-like structures[78].

They first performed a fluorescence polarization (FP) assay to measure the ability of small molecules to competitively inhibit SPOP-binding consensus (SBC1) peptide \textit{in vitro}. After compound 6a was identified as an initial hit capable of competing with SBC1 peptide binding to SPOP, additional synthetic optimization of the chemical core successfully yielded the more active lead compound 6b with a KD value of 35 \( \mu \text{M} \) on FP measurement. The \textit{in vitro} pull-down assay showed that compound 6b obviously disrupts the SPOP interaction with its substrate PTEN protein \textit{in vitro} with an IC\(_{50}\) of 2.8 \( \mu \text{M} \), and co-immunoprecipitation experiment showed that the inhibitor significantly disrupts SPOP binding to PTEN and DUSP7 in a dose-response manner. Accordingly, the compound efficiently inhibits the growth of six ccRCC cell lines and all primary ccRCC cells isolated from seven patients, with minimal effect on the growth of one non-tumor human proximal tubule epithelial cell line HK-2. Furthermore, a clear dose-dependent reduction in A498 tumor growth rate could be observed in mice treated with either compound 6b compared with the vehicle-injected control[78]. These results imply that SPOP would not classically be viewed as “undruggable,” and opens up an avenue that small-molecule-targeting cytoplasmic SPOP signaling might be more specific to ccRCC cancer cells, which would be a promising strategy to combat kidney cancer in future therapies.
PPI inhibitor

With an estimated 650,000 PPIs as part of the human interactome, PPIs are a critical means for the majority of proteins to exert and regulate their functions, and play crucial roles in various cellular processes and signaling pathways. The fast advance in MS technology allowed protein interactions to be elucidated in a systematic manner and tremendously enhanced the understanding of biological pathways and networks[79]. Dysregulation in PPIs are often found to be the primary cause of pathogenesis of some diseases, especially cancers. Thus, PPIs have become popular and attractive therapeutic targets. Although the development of PPI modulators (inhibitors or stabilizers) have been hindered because of the seemingly low druggability of PPI interfaces, extensive studies for PPI targets and modulators have been performed to understand these complex targets and identify distinct properties in its network, conformational structure, and ligand chemical space. Also, some PPI modulators has been in clinical trials or clinical use as summarized by a recent review[80], which have ascertained that these targets are tractable and can be modulated by small molecule compounds. Here, we provide three recent examples on the discoveries of target-based compounds against cancers.

MLL/menin interaction inhibitor

The mixed lineage leukemia (MLL) is a common target of chromosomal translocations found in patients with AML and ALL. Its fusion with one of over 50 different partner genes forms chimeric oncogenes encoding MLL fusion proteins in which the N-terminal 1400-amino-acid fragment of MLL is preserved and fused to distinct protein partners. The leukemogenic activity of MLL fusion proteins is critically dependent on their direct interaction with menin, a product of the multiple
endocrine neoplasia (MEN1) gene. Grembecka et al[81] screened a collection of 49,000 small molecules using a FP assay with a fluorescein-labeled MLL-derived peptide comprising the high-affinity menin binding motif (MBM1) to identify initial lead compounds that target menin and inhibit the menin-MLL interaction. They found that the most potent compound MI-1, which belongs to the thienopyrimidine class, reversibly inhibits the menin-MLL interaction with an IC$_{50}$ value of 1.9 μM. The compound effectively reverses MLL fusion protein–mediated leukemic transformation by downregulating the expression of target genes, and selectively blocks proliferation and induces both apoptosis and differentiation of leukemia cells harboring MLL translocations.

**CBFβ-SMMHC/RUNX1 interaction inhibitor**

AML with the chromosome inversion inv(16)(p13q22), a driver mutation that generates preleukemic progenitor cells that, upon acquisition of additional cooperating mutations, progress to leukemia, expresses the transcription factor fusion CBFβ-SMMHC (core binding factor β-smooth muscle myosin heavy chain), which cooperates with activating mutations in components of cytokine signaling pathways in leukemia transformation. CBFβ is a component of the heterodimeric transcription factor core binding factor, where it binds to RUNX proteins and enhances their affinity for DNA, and the resulting complex plays a key role in regulating hematopoiesis. CBFβ-SMMHC outcompetes CBFβ for binding to RUNX1, deregulates RUNX1 transcription factor activity in hematopoiesis, and induces AML. The CBFβ-SMMHC outcompetes wild-type CBFβ for binding to the transcription factor RUNX1, deregulates RUNX1 activity in hematopoiesis, and induces AML. Illendula et al[82] used a fluorescence resonance energy transfer (FRET) assay to
screen compounds that inhibit the binding of CBFβ-SMMHC to the RUNX1 Runt domain, and identified the active compound AI-4-57 with IC50 of 22 μM. Based on the lead compound, they found AI-10-49 to selectively bind to CBFβ-SMMHC and disrupts its binding to RUNX1, thus restoring RUNX1 transcriptional activity, displays favorable pharmacokinetics, and delays leukemia progression in mice. These data suggest that direct inhibition of the oncogenic CBFβ-SMMHC fusion protein may be an effective therapeutic approach for inv(16) AML, and they provide support for transcription factor targeted therapy in other cancers.

**APC-Asef interaction inhibitor**

Mutation and inactivation of APC, a widely accepted tumor suppressor gene highly mutated in CRC, is a key and early event almost uniquely observed in colorectal tumorigenesis. As a multi-domain protein, APC serves multiple functions through different binding partners. From the N terminus to the C terminus, there is an oligomerization domain, an armadillo repeat-domain, a 15- or 20-residue repeat domain, SAMP repeats domain, a basic domain and C-terminal domains. Alterations in the APC gene generate truncated gene products, leading to activation of the Wnt signaling pathway and deregulation of multiple other cellular processes. APC mutant proteins that retain at least the first 171 amino acids are able to bind to the oligomerization domain and may produce a dominant negative effect on APC protein. The armadillo repeat domain is the most conserved domain and has been shown to bind to IQ-motif-containing GTPase activation protein 1 (IQGAP1), PP2A, Asef (also known as Rho guanine nucleotide exchange factor 4), and KAP3. APC-Asef interaction can relieve the negative intramolecular regulation of Asef, which leads to aberrant migration in human colorectal cancer, suggesting that this interaction might be a potential target for the treatment of invasive migration in...
colorectal cancer. To identify potent inhibitors of APC-Asef interaction, we analyzed the direct binding interface on the basis of the previously determined structure of the APC-Asef complex, and a hot spot (181GGEQLAI187) in a flexible segment of Asef was used to design inhibitors (Fig. 4)[83]. Then a series of truncated peptides were synthesized and verified through a FP competitive assay. MAI-005 was identified as the best inhibitor in the first round screening, with a $K_i$ value of 44.62 ± 0.99 μM. After the first optimization with mutations at each position of MAI-005, three peptides were gotten with more potency ($K_i$ value is 3.12 ± 0.70 μM, 3.80 ± 0.72 μM, and 2.41 ± 0.88 μM, respectively). With the analysis of the structures of APC in complex with each of the three peptides, Arg549 at the center of the APC pocket was shown to have increased potential for interactions with the C termini of peptide inhibitors through salt-bridge or hydrogen bonds. Basing on that, they generated the second peptide library which contains restricted polar or acidic diversity at position 187, fully randomized diversity at position 186, and various capping groups on the terminus. MAI-150 was identified as the most potent peptide with a $K_i$ value of 0.12 ± 0.02 μM through the FP assay. For further optimization of MAI-150, we synthesized a panel of more than 50 peptidomimetic inhibitors, optimized for a capping group as well as the Leu185 and Tyr186 side chains. Finally, MAI-203 (Fig. 4) was identified as the most potent peptide with a $K_i$ value of 0.015 ± 0.001 μM through the FP assay. Next, it was examined whether MAIs inhibit APC-Asef interaction in vivo. An optimized 'GGGGG' linker was conjugated to the C terminus of the inhibitors (referred as MAITs relatively) to facilitate them across the cell membrane. Co-IP assay was carried out in HEK293T cells to corroborate the inhibitory effect of MAITs on APC-Asef interaction in vivo. Incubating cells with MAIT-203 and MAIT-150 reduced APC-Asef interaction in a
dose-dependent manner, while incubating with DMSO, 'GGGGG' linker, MAI-150, or MAI-203 showed no effect on APC-Asef interaction, confirming the fact that MAITs inhibit APC-Asef interaction in vivo. Through xCELLigence system Real Time Cell Analysis (RTCA), wound-healing assay and transwell assay, we showed that MAIT-203 at 10 μM significantly inhibited colorectal cell migration of two CRC cell lines (SW480 and HCT116). High concentrations (up to 100 μM) of MAIT-203 did not affect the morphology or growth of SW480 and HCT116 cells, indicating that the anti-migratory effect of MAIT-203 is worked through disrupting APC-Asef interaction rather than cell growth. Besides, RTCA invasion assay showed that MAIT-203 at 10 μM significantly repressed the invasion of SW480 and HCT116 cells, together with the results of migration assays, suggesting that MAIT-203 can inhibit the metastasis of colorectal cancer cells[83]. These results not only demonstrates the feasibility of exploiting the APC-Asef interaction as a target of metastatic colorectal cancer for drug discovery, but also establishes a new pharmacological paradigm for using peptide/peptidomimetic as an inhibitor of protein-protein interaction.

DEVELOPMENT OF CHEMICAL TOOLS

Much of our mechanistic knowledge about cellular processes has been gained by recreating a system in vitro or by engineering a cell or a model organism such as yeast, nematodes or mice. The chemical toolset to probe and validate such models in their real settings is a historical contribution of chemical biology to biological and biomedical sciences, as Chiosis addressed[84]. Therefore, exploring and developing new chemical toolset is very important for the application of chemical biology to manipulate biological systems. Here I like to outline several new strategies for chemistry to probe cellular activities.
Proximity-dependent labeling (PDL) methods for PPI identification

As mentioned above, PPIs play important roles in signaling pathways and cellular activities. PPIs can be either static in intact entities or be dynamically regulated. Previously, a variety of biochemical and/or high-throughput screening methods were developed to investigate the PPIs in vitro or in living cells, such as yeast two/three-hybridization, phage display, affinity pulldown coupled with MS characterization, protein array, FRET or bioluminescence resonance energy transfer and others. However, all these methods have some defects, such as presenting high false positive rate and failing to perform the real-time and in vivo PPI analysis. In the past few years, several groups have independently developed a class of methods termed proximity-dependent labeling (PDL) for PPIs mapping. Their basic strategy is that the protein of interest is genetically fused to a proximity-based labeling enzymes such as engineered biotin ligase or peroxidase, which are capable of covalently attaching the known reactive groups to the nearby proteins[85]. In the presence of biotin or biotin-containing substrates, the fused enzyme will activate and then release substrates to label proximal proteins. Interacting proteins that are in close proximity to the protein of interest are more likely to be labeled by the proximity enzyme [86]. Among these PDL methods, BioID is the earliest and the most widely used one. This technique harnesses a promiscuous biotin protein ligase BirA protein from E. coli (for the prototypic BioID) or its R118G mutant named BirA* to biotinylate proteins based on proximity. The ligase is fused to a protein of interest and expressed in cells, where it biotinylates proximal endogenous proteins (for detail, see reference [87]). The recent practices showed that BioID can identify weaker and/or transient interactions, is amenable to temporal regulation, and especially can be applied to insoluble proteins and to a variety of cell types from
diverse species. Like BioID, PDL with ascorbate peroxidase (APEX) is active in living cells and can catalyze biotin-phenol and H$_2$O$_2$ to generate biotin-phenolxyl radicals that covalently react with specific amino acids. APEX has also successfully applied to PPIs analyses in living cells, as recently summarized[85]. Recently, Wang H and Zhuang M’s groups in School of Life Science and Technology, ShanghaiTech University developed a new PDL system named PUP-IT (pupylation-based interaction tagging) to identify membrane protein interactions, in which a small protein tag Pup (a small bacterial protein that carries 64 amino acids with Gly-Gly-Gln at the C terminus) is applied to proteins that interact with a PafA, a gene that encodes Pup ligase)-fused protein, enabling transient and weak interactions to be enriched and detected by MS. With this approach to CD28, they identified multiple potential CD28-interacting proteins besides its known binding partners. They also showed that this method can identify the cell surface receptor and its ligand interaction[88].

**Genetically encoded sensors for NAD (P) H**

Real-time tracing of cell metabolism is a technical bottleneck for biomedical and bioengineering research. Traditional approaches to assess the cellular metabolic state, e.g. biochemical analysis, MS or nuclear magnetic resonance, are not effective for *in vivo*, real-time, spatiotemporal tracing of cellular metabolites. During the recent years, Yang Y and his co-workers in East China University of Science and Technology contribute significantly on the developing of genetically encoded sensors for nicotinamide adenine nucleotides(NADH)/NADH phosphate (NADPH), key metabolites for redox and energy metabolism. Reduced NADH and NADPH are the most important electron carries in cells. They participate in numerous metabolic redox reactions of organisms and are important parameters for cellular metabolic imbalances
and disease states. By rational design of fusion between fluorescent protein and NADH binding bacterial repressor Rex, they created a series of highly responsive and highly specific genetically encoded sensors for NADH and NADPH, which couple the emitting states of the fluorescent protein chromophore to the conformational changes induced by specific interaction between Rex protein and nicotinamide adenine nucleotide ligands. For instances, the NADH sensor Frex specifically detects free NADH level in the cells[89], SoNar reports NADH/NAD ratio[90] and iNap (a rational design mutant of SoNar) specifically detect NADPH[91]. These sensors can be targeted to various subcellular compartments or living tissues by genetic manipulation, which allows dynamic monitoring and imaging of cellular metabolic states in living cells or in vivo. Considering the central roles that NADH and NADPH play in cell metabolism and signaling, dynamic and quantitative monitoring of NADH and NADPH level in situ not only helps for the better understanding of substance and energy metabolism regulations and networks of, but also provides useful tools for biosynthetic of drugs, metabolic engineering, and drug discovery for metabolic diseases. These technological advances garner a lot of attention from the international peers, and are currently applied to various model organisms from bacteria, yeasts, fruitflies to mammals. Due to their favorite characteristics, these sensors are anticipated to revolutionize the study of redox biology, as many unresolved questions regarding the regulation and physiological role of different redox couples can be directly addressed.

**LightOn and LightOff gene expressions**

Regulation of gene expression is of vital importance for the study of various life activities such as cell metabolism. Gene expression systems induced by small molecular chemicals such as tetracycline are not able to precisely control gene
expression spatially and in a single cells. Based on the photosensitivity of the riboflavin moiety in light-oxygen-voltage-sensing domain, LightOn was developed as a light actable transcription activator and a robust and convenient single component light switchable gene expression system. LightOn enables spatiotemporal, quantitative and reversible expression regulation of functional genes such as insulin and Cre recombinase in vivo, and control the blood glucose levels and glucose metabolism in diabetic mice[92]. Further studies showed that integration of the system with tetracycline inducible gene expression system lead to gene expression only in the presence of inducer and light, which is significantly more stringent[93]. Based on similar strategies, the group also developed light activated repressor, the LightOff gene expression. Such a system achieved extremely stringent regulation of gene expression with induction ratio higher than 10,000 fold, far exceeding than that of chemical inducible gene expression system[94]. These novel gene expression system provide powerful tools not only for the analysis of complex biological systems, but also for accurate therapeutic strategies of accurate timing and dosage for important human diseases such as diabetes. For example, LightOn system offers an efficient way to control the proliferation and differentiation of neural progenitor cells by changing the light-exposure pattern, showing its applicability to the regeneration technology[95].

**Bioorthogonal cleavage reactions for gain-of-function study of proteins**

Cells have evolved a rich repertoire of enzymes to catalyze potent chemical modifications on diverse proteins that dictate virtually all signaling events. For example, there exist near 600 kinases in human cells to control phosphorylation and this enzymatic network becomes more complicated by the presence of many phosphatases that catalyze the removal of phosphorylation to reverse the signaling.
Mis-regulation of these enzymes is often linked with various diseases such as cancer and inflammation. Meanwhile, oncogenic mutations on these signaling proteins are the common features for cancer. Moreover, this process is further complicated by the feedback regulation of enzyme and signaling protein networks and little is known about this critical “fine-tuning” mechanism. *In vivo* manipulation of a given enzyme with high specificity and spatiotemporal resolution is highly desired yet exceedingly difficult. To address these challenges, the bioorthogonal cleavage reaction-based “chemical decaging” strategy has been developing and thriving in recent years as chemistry-enabled strategies to label and manipulate biomolecules without the interference on native biochemical processes. For instance, the critical roles of diverse PTMs on proteins have been increasingly appreciated, with many of the underlying mechanisms hardly addressable by traditional genetic-based approaches. An array of bioorthogonal chemistry tools have been developed to track, visualize and modify proteins as well as other biomolecules under living conditions, which significantly facilitated the study of these PTM events. However, almost all previous efforts have been centered on “bond-formation” reactions. In the opposite direction, Chen P’s group in Peking University started to recognize noticeable needs to develop bioorthogonal “bond-cleavage” tools that could be applied for spatial-temporal controlled rescue of intracellular proteins, surface glycans and even intact cells within a native cellular environment [96]. They showed that simple palladium species can effectively catalyze the depropargylation reaction with low toxicity in living cells, which can be employed as a biocompatible chemical decaging strategy to rescue intracellular protein activity[97]. They also employed the inverse electron demand Diels-Alder (iDA) reaction as a small molecule triggered bioorthogonal cleavage reaction. The classical iDA reaction has
been used as a “bioorthogonal triggered release” strategy via installation of a carbamate group next to the double bond on trans-cyclooctene (TCO), which readily reacted with tetrazines followed by rearrangement through electron shift, resulting in the deprotection of TCO group from protected amine moiety. His group elegantly demonstrated the application of this reaction as a chemical decaging strategy on essential lysine residues on various proteins[98].

By coupling these bioorthogonal cleavage reactions with the genetic code expansion methodology, they also created a mechanism-based kinase activation strategy[99]. The activity of each member of the near 600 kinases is precisely regulated by its native physiochemical inputs, which are often entangled within the complicated signaling networks, making it exceedingly difficult to manipulate a single kinase with high specificity and spatiotemporal resolution. Gain-of-function study of kinase is advantageous in probing the sufficiency of a specific kinase as opposed to the more widely adopted loss-of-function methods, but a general “activation” tool is lacking. Their kinase decaging strategy relies on replacing the kinase’s catalytic lysine by a chemical-caged lysine analogue TCOK via genetic code expansion to blockage its enzymatic activity. The subsequent addition of the bioorthogonal cleavage trigger Tetrazine (TZ) will unmask this lysine and thus rescue the corresponding native kinase. They have applied this strategy to specifically rescue a panel of kinases such as mitogen-activated protein kinase kinase-1/2, Focal adhesion kinase (FAK) and Src in living cells and further extended its utility to living animals[99]. Totally, these approaches offer a general tool to rescue the native sequence of desired protein machinery in situ such as PTM enzymes, metabolic enzymes, epigenetic regulators, which is uniquely positioned...
for gain-of-function, in contrast to more conventional loss-of-function, study of proteins within living systems.

**CONCLUSION**

The field of cancer biology has long benefited from the mechanistic insights provided by appropriately characterized chemical probes or tools of sufficient quality. These chemical probes enable temporal–spatial study of cellular pathways with confidence to interrogate complex biological systems. Numerous breakthroughs in biology have been enabled by the use of chemical probes or tools of sufficient quality, especially in combination with complementary biological reagents and molecular technologies[5]. Actually, the great reservoir of natural compounds derived from higher plants such as TCM herbs displays an important role in the discovery and development of new chemical probes and also therapeutic drugs. The diversity in the number of species of plants all over the world is enormous (from 215,000 to 500,000). However, only about 6% have been screened for biological activity, and 15% have been evaluated phytochemically. Therefore, we should continue to explore these small molecules, ideally of well-defined biological potency, selectivity, and cell permeability. Especially, China should provide financial and policy supports for collection of chemical compounds to nation-level libraries to be shared by all scientists. On the other hand, chemists and biologists should joined together to further explore and develop strategies of target identification and validation. In the era of big data, especially, the artificial intelligence technology also should be developed in various fields of the chemical probe or drug design and target identification [100,101]. Finally, there are challenges with the selection and use of chemical probes. The criteria for high-quality chemical probes often need to be more stringent than a drug used in patients,
especially in regard to selectivity, because chemical probes need to be much more selective to ask specific biological questions. Therefore, careful chemical optimization and biological testing must be carried out to minimize the risk of off-target effects in chemical probes.

ACKNOWLEDGMENTS

Due to number limitation, I do not cite all related references and described development of other important chemical tools.

FUNDING

Partial works in our laboratory were supported by National Natural Science Foundation of China (81230048, 81430061, 81502001 and 81721004).

REFERENCES

1 Altmann, KH, Buchner, J, Kessler, H et al. The state of the art of chemical biology. Chembiochem 2009; 10:16-29
2 Bucci, M, Goodman, C and Sheppard, TL. A decade of chemical biology. Nat Chem Biol 2010; 6:847-54
3 Filippakopoulos, P, Qi, J, Picaud, S et al. Selective inhibition of BET bromodomains. Nature 2010; 468:1067-73
4 Nicodeme, E, Jeffrey, KL, Schaefer, U et al. Suppression of inflammation by a synthetic histone mimic. Nature 2010; 468:1119-23
5 Blagg, J and Workman, P. Choose and Use Your Chemical Probe Wisely to Explore Cancer Biology. Cancer Cell 2017; 32:9-25
6 Tang, YJ, Gholamin, S, Schubert, S et al. Epigenetic targeting of Hedgehog pathway transcriptional output through BET bromodomain inhibition. Nature Medicine 2014; 20:732-40
7 Kostic, M, Crews, CM, Hertweck, C et al. Voices of Chemical Biology: Charting the Next Decade. Cell Chem Biol 2016; 23:199
8 Jiang, H, Wu, J, Zhang, L et al. Chemical biology in China takes on signal transduction. Nat Chem Biol 2008; 4:515-8
9 Zhang, XW, Yan, XJ, Zhou, ZR et al. Arsenic trioxide controls the fate of the PML-RARalpha oncoprotein by directly binding PML. Science 2010; 328:240-3
10 Li, K, Wang, F, Cao, WB et al. TRIB3 Promotes APL Progression through
Stabilization of the Oncoprotein PML-RARalpha and Inhibition of p53-Mediated Senescence. Cancer Cell 2017; 31:697-710

11 Hou, P, Li, Y, Zhang, X et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science 2013; 341:651-4

12 Yang, Y, Liu, B, Xu, J et al. Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. Cell 2017; 169:243-57

13 Li, X, Yang, Y, Yang, Z et al. Chemical reprogramming: the CiPSCs and the CiNs. National Science Review 2017; 4:7-10

14 Workman, P and Al-Lazikani, B. Drugging cancer genomes. Nat Rev Drug Discov 2013; 12:889-90

15 Blagg, J and Workman, P. Chemical biology approaches to target validation in cancer. Curr Opin Pharmacol 2014; 17:87-100

16 Shen, Q, Cheng, F, Song, H et al. Proteome-Scale Investigation of Protein Allosteric Regulation Perturbed by Somatic Mutations in 7,000 Cancer Genomes. Am J Hum Genet 2017; 100:5-20

17 Patel, MN, Halling-Brown, MD, Tym, JE et al. Objective assessment of cancer genes for drug discovery. Nat Rev Drug Discov 2013; 12:35-50

18 Guo, J, Yu, W, Su, H et al. Genomic landscape of gastric cancer: molecular classification and potential targets. Sci China Life Sci 2017; 60:126-37

19 Newman, DJ and Cragg, GM. Natural Products as Sources of New Drugs from 1981 to 2014. Journal of Natural Products 2016; 79:629-61

20 Galloway, WRJD, Isidro-Llobet, A and Spring, DR. Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. Nat Commun 2010; 1:80 doi: 10.1038/ncomms81

21 He, XL, Zhang, WJ, Yan, C et al. Chemical biology reveals CARF as a positive regulator of canonical Wnt signaling by promoting TCF/beta-catenin transcriptional activity. Cell Discovery 2017; 3:17003; doi:10.1038/celldisc.2017.3

22 Jones, LH and Bunnage, ME. Applications of chemogenomic library screening in drug discovery. Nat Rev Drug Discov 2017; 16:285-96

23 Liu, YR, He, Y, Yang, FF et al. A novel synthetic small molecule YF-452 inhibits tumor growth through antiangiogenesis by suppressing VEGF receptor 2 signaling. Sci China Life Sci 2017; 60:202-14

24 Wilson, MR, Zha, L and Balskus, EP. Natural product discovery from the human microbiome. Journal of Biological Chemistry 2017; 292:8546-52

25 To, SK, Zeng, JZ and Wong, AS. Nur77: a potential therapeutic target in cancer. Expert Opin Ther Targets 2012; 16:573-85

26 Zhan, Y, Du, X, Chen, H et al. Cytosporone B is an agonist for nuclear orphan receptor Nur77. Nat Chem Biol 2008; 4:548-56

27 Zhan, YY, Chen, Y, Zhang, Q et al. The orphan nuclear receptor Nur77 regulates LKB1 localization and activates AMPK. Nat Chem Biol 2012; 8:897-904

28 Wang, WJ, Wang, Y, Chen, HZ et al. Orphan nuclear receptor TR3 acts in autophagic cell death via mitochondrial signaling pathway. Nat Chem Biol
Li, L, Liu, Y, Chen, HZ et al. Impeding the interaction between Nur77 and p38 reduces LPS-induced inflammation. Nat Chem Biol 2015; 11:339-46
Schenone, M, Dancik, V, Wagner, BK et al. Target identification and mechanism of action in chemical biology and drug discovery. Nat Chem Biol 2013; 9:232-40
Lomenick, B, Hao, R, Jonai, N et al. Target identification using drug affinity responsive target stability (DARTS). P Natl Acad Sci USA 2009; 106:21984-89
Li, J, Xu, Y, Long, XD et al. Cbx4 Governs HIF-1 alpha to Potentiate Angiogenesis of Hepatocellular Carcinoma by Its SUMO E3 Ligase Activity. Cancer Cell 2014; 25:118-31
Huang, Y, Lin, D and Taniguchi, CM. Hypoxia inducible factor (HIF) in the tumor microenvironment: friend or foe? Sci China Life Sci 2017; 60:1114-24
Bhattarai, D, Xu, X and Lee, K. Hypoxia-inducible factor-1 (HIF-1) inhibitors from the last decade (2007 to 2016): A "structure-activity relationship" perspective. Med Res Rev 2018; 38:1404-42
Xia, XJ, Zuo, FF, Luo, MG et al. Role of TRIM33 in Wnt signaling during mesendoderm differentiation. Sci China Life Sci 2017; 60:1142-49
Qu, Y, Olsen, JR, Yuan, X et al. Small molecule promotes beta-catenin citrullination and inhibits Wnt signaling in cancer. Nat Chem Biol 2018; 14:94-101
Wang, W, Liu, HY, Wang, S et al. A diterpenoid derivative 15-oxospiramilactone inhibits Wnt/beta-catenin signaling and colon cancer cell tumorigenesis. Cell Res 2011; 21:730-40
Wang, S, Yin, JL, Chen, DZ et al. Small-molecule modulation of Wnt signaling via modulating the Axin-LRP5/6 interaction. Nat Chem Biol 2013; 9:579-85
Willoughby, LF, Schlosser, T, Manning, SA et al. An in vivo large-scale chemical screening platform using Drosophila for anti-cancer drug discovery. Disease Models & Mechanisms 2013; 6:521-29
Ridges, S, Heaton, WL, Joshi, D et al. Zebrafish screen identifies novel compound with selective toxicity against leukemia. Blood 2012; 119:5621-31
Kalin, RE, Banziger-Tobler, NE, Detmar, M et al. An in vivo chemical library screen in Xenopus tadpoles reveals novel pathways involved in angiogenesis and lymphangiogenesis. Blood 2009; 114:1110-22
Futamura, Y, Muroi, M and Osada, H. Target identification of small molecules
based on chemical biology approaches. Mol Biosyst 2013; 9:897-914

Ziegler, S, Pries, V, Hedberg, C et al. Target Identification for Small Bioactive Molecules: Finding the Needle in the Haystack. Angewandte Chemie-International Edition 2013; 52:2744-92

Ito, T, Ando, H, Suzuki, T et al. Identification of a primary target of thalidomide teratogenicity. Science 2010; 327:1345-50

Roemer, T, Davies, J, Giaever, G et al. Bugs, drugs and chemical genomics. Nat Chem Biol 2011; 8:46-56

Schirle, M, Bantscheff, M and Kuster, B. Mass spectrometry-based proteomics in preclinical drug discovery. Chem Biol 2012; 19:72-84

Berrade, L, Garcia, AE and Camarero, JA. Protein microarrays: novel developments and applications. Pharm Res 2011; 28:1480-99

Tao, SC, Chen, CS and Zhu, H. Applications of protein microarray technology. Comb Chem High Throughput Screen 2007; 10:706-18

Huang, J, Zhu, H, Haggarty, SJ et al. Finding new components of the target of rapamycin (TOR) signaling network through chemical genetics and proteome chips. P Natl Acad Sci USA 2004; 101:16594-99

Bae, N, Viviano, M, Su, XN et al. Developing Spindlin1 small-molecule inhibitors by using protein microarrays. Nat Chem Biol 2017; 13:750-56

Henzler-Wildman, K and Kern, D. Dynamic personalities of proteins. Nature 2007; 450:964-72

Xu, Y, Liu, J, Wu, Y et al. Natural products against hematological malignancies and identification of their targets. Sci China Life Sci 2015; 58:1191-201

Zhen, T, Wu, CF, Liu, P et al. Targeting of AML1-ETO in t(8;21) leukemia by oridonin generates a tumor suppressor-like protein. Sci Transl Med 2012; 4:127ra38

Wang, X, Lin, Q, Lv, F et al. LG-362B targets PML-RAR alpha and blocks ATRA resistance of acute promyelocytic leukemia. Leukemia 2016; 30:1465-74

Gu, ZM, Wu, YL, Zhou, MY et al. Pharicin B stabilizes retinoic acid receptor-alpha and presents synergistic differentiation induction with ATRA in myeloid leukemic cells. Blood 2010; 116:5289-97

Xu, HZ, Huang, Y, Wu, YL et al. Pharicin A, a novel natural ent-kaurene diterpenoid, induces mitotic arrest and mitotic catastrophe of cancer cells by interfering with BubR1 function. Cell Cycle 2010; 9:2897-907

Zhang, RH, Liu, ZK, Yang, DS et al. Phytochemistry and pharmacology of the genus Leonurus: The herb to benefit the mothers and more. Phytochemistry 2018; 147:167-83

Liu, CX, Yin, QQ, Zhou, HC et al. Adenanthin targets peroxiredoxin I and II to induce differentiation of leukemic cells. Nat Chem Biol 2012; 8:486-93

Liu, CX, Zhou, HC, Yin, QQ et al. Targeting peroxiredoxins against leukemia. Exp Cell Res 2013; 319:170-6

Schieber, M and Chandel, NS. ROS function in redox signaling and oxidative
Kang, SW, Lee, S and Lee, EK. ROS and energy metabolism in cancer cells: alliance for fast growth. Arch Pharm Res 2015; 38:338-45

Yue, Z, Xiao, X, Wu, J et al. ent-Jungermannenone C Triggers Reactive Oxygen Species-Dependent Cell Differentiation in Leukemia Cells. J Nat Prod 2018; 81:298-306

Ying, M, Shao, X, Jing, H et al. Ubiquitin-dependent degradation of CDK2 drives the therapeutic differentiation of AML by targeting PRDX2. Blood 2018; 131:2698-711

Rossari, F, Minutolo, F and Orciuolo, E. Past, present, and future of Bcr-Abl inhibitors: from chemical development to clinical efficacy. J Hematol Oncol 2018; 11:84. doi: 10.1186/s13045-018-0624-2.

Liu, C, Nie, D, Li, J et al. Antitumor Effects of Blocking Protein Neddylation in T315I-BCR-ABL Leukemia Cells and Leukemia Stem Cells. Cancer Res 2018; 78:1522-36

Dong, B, Liang, Z, Chen, Z et al. Cryptotanshinone suppresses key oncoproliferative and drugresistant pathways of chronic myeloid leukemia by targeting STAT5 and STAT3 phosphorylation. Sci China Life Sci 2018; 61:999-1009

Sun, W, Xie, Z, Liu, Y et al. JX06 Selectively Inhibits Pyruvate Dehydrogenase Kinase PDK1 by a Covalent Cysteine Modification. Cancer Res 2015; 75:4923-36

Okoye-Okafor, UC, Bartholdy, B, Cartier, J et al. New IDH1 mutant inhibitors for treatment of acute myeloid leukemia. Nat Chem Biol 2015; 11:878-86

Chen, B, Zang, W, Wang, J et al. The chemical biology of sirtuins. Chem Soc Rev 2015; 44:5246-64

Sebastian, C, Zwaans, BMM, Silberman, DM et al. The Histone Deacetylase SIRT6 Is a Tumor Suppressor that Controls Cancer Metabolism. Cell 2012; 151:1185-99

Kugel, S, Sebastian, C, Fitamant, J et al. SIRT6 Suppresses Pancreatic Cancer through Control of Lin28b. Cell 2016; 165:1401-15

Huang, WK, Lu, SY, Huang, ZM et al. Allosite: a method for predicting allosteric sites. Bioinformatics 2013; 29:2357-59

Huang, ZM, Zhao, ZX, Deng, W et al. Identification of a cellularly active SIRT6 allosteric activator. Nat Chem Biol 2018; in press:

Gao, YH, Wu, ZX, Xie, LQ et al. VHL deficiency augments anthracycline sensitivity of clear cell renal cell carcinomas by down-regulating ALDH2. Nat Commun 2017; 8:15337

Guo, ZQ, Zheng, T, Chen, B et al. Small-Molecule Targeting of E3 Ligase Adaptor SPOP in Kidney Cancer. Cancer Cell 2016; 30:474-84

Li, X, Wang, W and Chen, J. Recent progress in mass spectrometry proteomics for biomedical research. Sci China Life Sci 2017; 60:1093-113

Macalino, SJY, Basith, S, Clavio, NAB et al. Evolution of In Silico Strategies for Protein-Protein Interaction Drug Discovery. Molecules 2018; 23:pii:
Grembecka, J, He, S, Shi, A et al. Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. Nat Chem Biol 2012; 8:277-84

Illendula, A, Pulikkan, JA, Zong, H et al. Chemical biology. A small-molecule inhibitor of the aberrant transcription factor CBFbeta-SMMHC delays leukemia in mice. Science 2015; 347:779-84

Jiang, H, Deng, R, Yang, X et al. Peptidomimetic inhibitors of APC-Asef interaction block colorectal cancer migration. Nat Chem Biol 2017; 13:994-1001

Jiang, H, Hsieh-Wilson, L, Arruda, P et al. Voices of chemical biology. Nat Chem Biol 2015; 11:446-7

Zhu, H, Tamura, T and Hamachi, I. Chemical proteomics for subcellular proteome analysis. Curr Opin Chem Biol 2018; 48:1-7

Li, PP, Li, JJ, Wang, L et al. Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool. Proteomics 2017; 17: doi: 10.1002/pmic.201700002

Roux, KJ, Kim, DI, Burke, B et al. BioID: A Screen for Protein-Protein Interactions. Curr Protoc Protein Sci 2018; 91:19.23.1–19.23.15. doi: 10.1002/cpps.51

Liu, Q, Zheng, J, Sun, WP et al. A proximity-tagging system to identify membrane protein-protein interactions. Nat Methods 2018; 15:715-22

Zhao, Y, Jin, J, Hu, Q et al. Genetically encoded fluorescent sensors for intracellular NADH detection. Cell Metab 2015; 14:555-66

Zhao, Y, Hu, Q, Cheng, F et al. SoNar, a Highly Responsive NAD+/NADH Sensor, Allows High-Throughput Metabolic Screening of Anti-tumor Agents. Cell Metab 2015; 21:777-89

Tao, R, Zhao, Y, Chu, H et al. Genetically encoded fluorescent sensors reveal dynamic regulation of NADPH metabolism. Nat Methods 2017; 14:720-28

Wang, X, Chen, XJ and Yang, Y. Spatiotemporal control of gene expression by a light-switchable transgene system. Nat Methods 2012; 9:266-9

Chen, XJ, Li, T, Wang, X et al. Synthetic dual-input mammalian genetic circuits enable tunable and stringent transcription control by chemical and light. Nucleic Acids Research 2016; 44:2677-90

Chen, XJ, Liu, RM, Ma, ZC et al. An extraordinary stringent and sensitive light-switchable gene expression system for bacterial cells. Cell Res 2016; 26:854-57

Imayoshi, I, Isomura, A, Harima, Y et al. Oscillatory Control of Factors Determining Multipotency and Fate in Mouse Neural Progenitors. Science 2013; 342:1203-08

Fan, XY, Li, J and Chen, PR. Bioorthogonal chemistry in living animals. Natl Sci Rev 2017; 4:300-02

Li, J, Yu, J, Zhao, J et al. Palladium-triggered deprotection chemistry for protein activation in living cells. Nat Chem 2014; 6:352-61

Li, J, Jia, S and Chen, PR. Diels-Alder reaction-triggered bioorthogonal
protein decaging in living cells. Nat Chem Biol 2014; 10:1003-5

99 Zhang, G, Li, J, Xie, R et al. Bioorthogonal Chemical Activation of Kinases in Living Systems. ACS Cent Sci 2016; 2:325-31

100 Zhong, F, Xing, J, Li, X et al. Artificial intelligence in drug design. Sci China Life Sci 2018; 61:1191-204

101 Zhang, L, Tan, J, Han, D et al. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today 2017; 22:1680-85
Figure 1. General flow chart of forward/compounds-based and reverse/targets-based chemical biology approaches.
Figure 2. Chemical biology investigations based on Nur77-targeting compounds Csn-B and its derivatives. Csn-B specifically binds to the ligand-binding domain of Nur77 and subsequently activates its transactivational activity towards target genes including Nur77 itself. Then increased Nur77 is translocated to mitochondria to cause cytochrome C (cyto C) release, inducing apoptosis and retarded xenograft tumor growth. Other three derivatives of Csn-B, including TMPA (Ethyl 2-[2,3,4-trimethoxy-6-(1-octanoyl)phenyl]acetate), THPN (1-(3,4,5-trihydroxyphenyl)nonan-1-one) and PDNPA (n-Pentyl 2-[3,5-dihydroxy-2-(1-nonanoyl)phenyl]acetate), can exert different functions in decreasing glucose, inducing autophagy and inhibiting inflammation through modulating various signaling pathways.
Figure 3. Adenanthin and ent-Jungermannenone C target to Prdx I/II to induce leukemic cell differentiation. A flow chart for SAR analysis and target identification of adenanthin is shown in the dot-lined frame.

Figure 4. Design and optimization of small peptide to inhibit APC-Asef interaction.