

# ADVANCES IN CHEMICAL GENETICS

---

Inese Smukste and Brent R. Stockwell

*Department of Biological Sciences and Department of Chemistry, Columbia University, Fairchild Center, New York, New York 10027; email: stockwell@biology.columbia.edu*

**Key Words** small molecules, high-throughput screening, combinatorial chemistry, target identification

■ **Abstract** Chemical genetics is an emerging approach for studying biological systems using chemical tools. This strategy aims to reveal the macromolecules responsible for regulating biological systems; thus, the approach shares much in common with genetics. In both strategies, one must (a) develop an assay that reports on a biological process of interest, (b) perturb this process systematically (with mutations or small molecules), and (c) determine the target of each perturbation to reveal macromolecules (i.e., proteins and genes) regulating the process of interest. In this review, we discuss advances and challenges in this field that have emerged over the past four years. Several technologies have converged, raising the hope that it may be possible to systematically apply chemical probes to biological processes.

## INTRODUCTION

Chemical genetics is an approach for studying biological systems that has emerged over the past decade (100, 101). This strategy aims to reveal the specific macromolecules that are responsible for regulating biological systems using small organic molecules; this approach shares much in common with the science of genetics (88). In both approaches, one (a) develops an assay for a biological process, (b) perturbs this process systematically (with mutations or small molecules), and (c) determines the target of each perturbation to reveal proteins or genes regulating the process. In this review, we discuss advances and challenges in this field that emerged over the past four years.

Several technologies have converged over the past decade, raising the hope that it may be possible to systematically apply chemical probes to biological processes (90). First, rapid (i.e., high-throughput) screening has been developed, enabling the testing of many thousands of compounds in miniaturized phenotypic and protein-targeted assays. Recent advances and remaining problems in this area discussed herein include the development of new phenotypic assay formats and new screening methods.

Second, chemical genetics requires the ability to create thousands of small-molecule probes targeting many different proteins using efficient synthetic

chemistry, just as genetics requires the ability to create thousands of mutations in many different genes using efficient mutagens (89). The emergence of combinatorial chemistry in the past decade provided the opportunity to create thousands, or even millions, of distinct small molecules using synthetic chemistry (27, 29). Unfortunately, the range of compounds that can be synthesized in such a fashion is still limited; these procedures need to be refined to provide large numbers of useful chemical probes for screening efforts. Over the past four years, some advances have been reported in the complexity of compounds that can be synthesized in such a rapid fashion; these are described below.

Third, chemical genetics requires the ability to identify the molecular mechanisms of action of small molecules, just as genetics requires the ability to identify the mutated genes that have been selected in a mutagenesis screen (14). In the latter half of the twentieth century, several early successes at target identification for small molecules demonstrated that, in some cases, it is feasible to determine the detailed molecular mechanism of action for small-molecule probes (100). Advances in the past four years reinforce this notion, but also highlight the challenge in making this component of the chemical genetic process systematic and reliable.

Fourth, informatics and software tools have been developed to facilitate the analysis of large quantities of screening data (81). Such tools are needed to identify and to correct systematic errors in large data sets, to evaluate structure-activity relationships, and to manage compound inventories. With high-quality data sets, it is possible to perform global analyses to extract information about the compounds and systems under study. Herein, we review such tools and global analyses and highlight the need for attention to data quality.

Finally, numerous investigators have pursued reverse chemical genetic approaches, which are analogous to reverse genetic approaches. In this strategy, small-molecule inhibitors are created against a specific protein or protein family; subsequently, the phenotypic consequences of modulating this protein or protein family are determined. We review a number of such studies that were published over the past four years and what has been learned using this approach.

Thus, this review covers chemical libraries that have been created for genetic-like screens, screens and screening methods, approaches to determining the mechanism of action of small molecules, global analyses, and tools for reverse chemical genetics. This discussion of these related topics reveals both the challenges and promise of this interdisciplinary approach to studying biological systems.

## CHEMICAL LIBRARIES

Just as collections of mutated strains or organisms are essential for genetic studies, collections of small molecules are essential for chemical genetic studies. Small molecules vary in quality and usefulness for perturbing biological systems. Useful compounds have a high affinity for a single target protein, are potent (i.e., effective at a low concentration), display high intrinsic activity (i.e., completely inhibit or activate their target protein), and are specific (i.e., don't affect other proteins). An

example of such a useful compound is the natural product rapamycin, which is an inhibitor of the mTOR protein (mammalian target of rapamycin) (11, 83). A major challenge for the field of chemical genetics is to discover the structural elements in small molecules that enable these useful attributes.

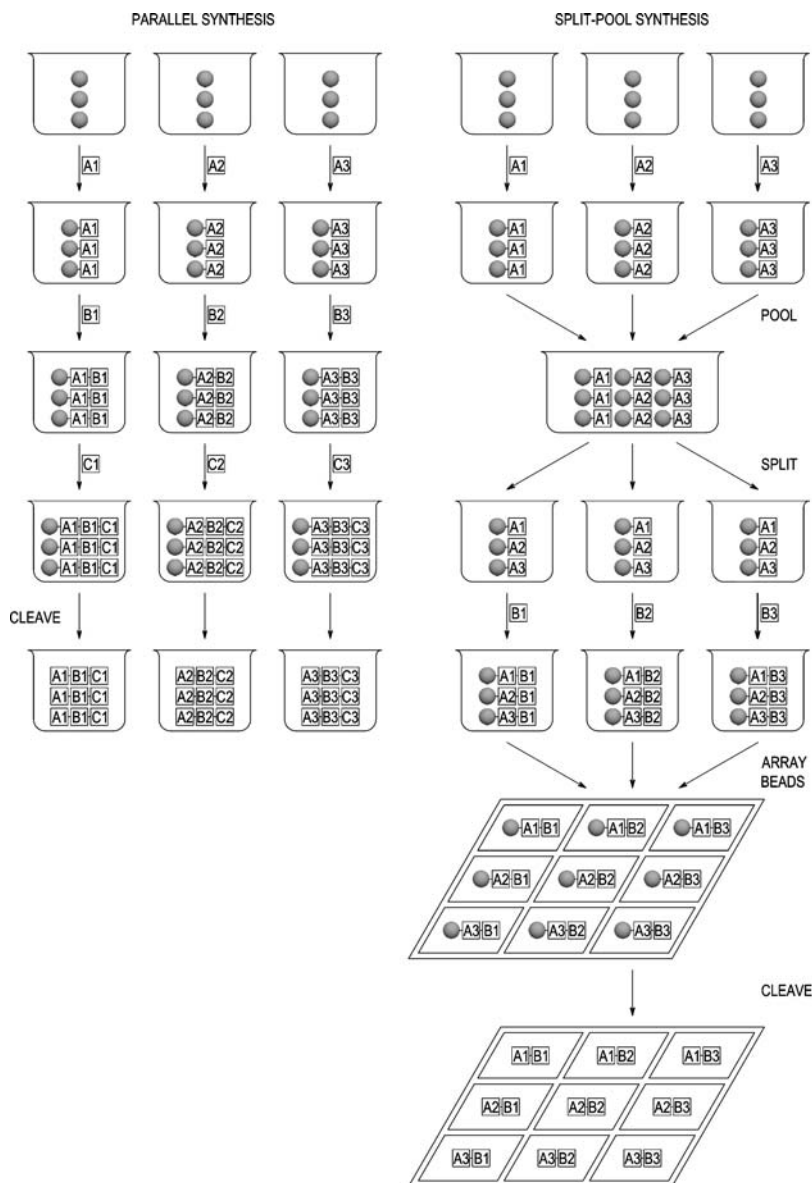
The first chemical libraries were painstakingly assembled by pharmaceutical companies by adding compounds that had been synthesized for internal drug discovery projects one by one. Many such compounds were analogs (structural relatives) of existing biologically active compounds—either existing drugs or natural products (94). Only with the development of high-throughput chemical synthesis, including the use of solid-phase resins, did collections of chemicals become available to investigators in academia and smaller companies (35). A solid-phase resin is essentially a group of small plastic beads on which a peptide or small molecule is created; this allows easy separation of the growing peptide or compound from undesired leftover material in the reaction solution (Figure 1).

Merrifield (65) first developed solid-phase synthesis in the context of peptides; this technology enabled the synthesis of libraries of peptides (41, 51). However, syntheses in which peptides are created combinatorially (rather than side by side, as in parallel synthesis) became possible with the introduction of the split-pool method, which allows the synthesis of tens of thousands of synthetic peptides simultaneously (37, 59). In parallel synthesis, each compound or peptide is synthesized in a separate small container (e.g., test tube or well of a multi-well plate). In split-pool synthesis, a mixture of beads is initially split in several containers and the first amino acid is attached to the beads; subsequently, all the beads are pooled together and resplit into a new set of containers, where the second amino acid is attached to the first amino acid (Figure 2). Using this process, it is possible to synthesize millions of chemicals using only a few hundred reactions (106, 107).

Early efforts to generate compounds using combinatorial methods focused on synthesizing biopolymers such as peptides and polynucleotides. In the past decade, methods for the synthesis of chemical libraries of nonpeptidic, low molecular weight compounds have been developed. Comprehensive reviews of combinatorial chemistry are detailed elsewhere (27–33).

Although combinatorial libraries have provided many new chemical entities, until recently most combinatorial libraries contained simple compounds with few or no stereogenic centers (e.g., carbon atoms with four different groups that create the possibility of mirror images existing as distinct compounds) and a limited variety of functional groups. Clustered in a limited region of chemical space, these compounds were less likely to be active in biological systems than compounds isolated from natural sources (80). There has been a recent trend toward creating structurally complex, natural product-like compound libraries that may provide higher-quality collections for screening (2, 10, 66, 71, 89, 90).

Although natural products cover a larger region of chemical diversity space, they are limited in their chemical diversity by the building blocks available to living organisms—primarily short chain fatty acids, amino acids, five carbon terpenoid precursors, and sugars. Natural products are thus synthesized through a



handful of biochemical pathways creating terpenes, polyketides, cyclic peptides, and alkaloids (63). Has nature already defined the optimal regions of chemical space for probing biological systems, or has nature created the best compounds with the limited chemical matter that was available? If the latter is true, should we then further diversify into unexplored regions of chemical space?

To address this question, the new field of diversity-oriented synthesis (DOS) aims to provide structurally complex and diverse chemical collections. DOS seeks to expand the accessible regions of chemical space by creating compound collections that cover currently poorly populated areas (16, 89, 90). The first such library was created based on the shikimic acid scaffold, which was further modified in a split-pool synthesis to provide more than 2 million complex compounds (106, 107).

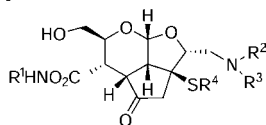
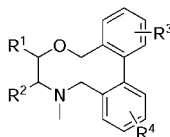
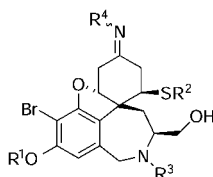
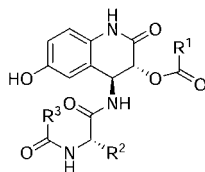
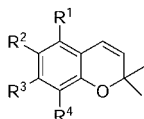
Highly functionalized (i.e., having many different chemical groups) chiral, medium, and large rings are often encountered in natural products; such structures provide rigid scaffolds for displaying chemical groups. Although these compounds have potential for providing biologically active molecules, their syntheses present a formidable challenge even using the solution phase; until recently such compounds were out of reach of combinatorial methods. However, several noteworthy examples of complex molecule synthesis were recently reported (Figure 2): Ferrier & Pauson-Khand reactions (56) were used to prepare a library of 2500 tricyclic and tetracyclic small molecules. Spring and colleagues (96, 97) developed methodology for the stereoselective synthesis of biaryl-containing medium-size rings. This library represents the first large-scale, solid-phase atropidiastereoselective synthesis of 9-, 10-, and 11-membered rings. Recently, a synthesis of macrolactones was reported (86). Thus, libraries of complex small molecules have been derived around well validated bioactive polycyclic natural product scaffolds, or so-called privileged structures.

Typically, libraries of diverse small molecules are obtained by appending various building blocks (akin to amino acids in peptides) to a common scaffold (27–33); however, creating libraries with diverse skeletons is more challenging. Among the

---

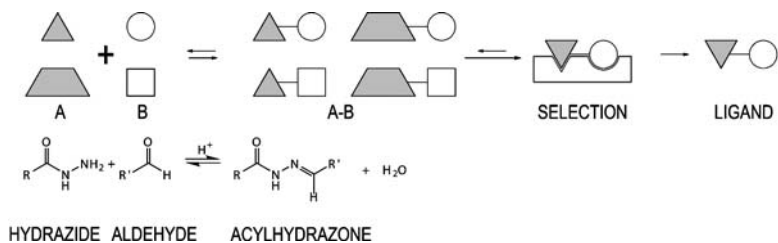
←

**Figure 1** Comparison of parallel and split-pool syntheses. Spheres represent polymer beads; A, B, and C represent different building blocks (e.g., amino acids): 1. In a solid-phase, parallel synthesis procedure, compounds are constructed on a solid support such as polymer beads, which enables facile purification of the products via filtration and use of excess of reagents to ensure a high yield of the reaction product. Each compound is synthesized in a separate container, and after cleaving the linkage to the beads, pure compounds are obtained. Automation enables convenient parallel synthesis using a 96-well format. 2. Split-pool synthesis generates many compounds using a small number of synthetic reactions. In each reaction step, polymer beads are pooled together in several containers and a particular building block is added to each set of beads. After the reaction is complete, all beads are pooled together and split into separate flasks for the next reaction. This figure shows a two-step synthesis of nine compounds using three different reagents in each step ( $3 \times 3 = 9$ ). After the synthesis, the beads can be pooled together, providing a mixture of compounds upon cleavage from the beads. Alternatively, in the “one-bead, one-compound” approach developed by the Schreiber group, each bead is placed in a separate well of a 384-well plate and upon cleavage pure compounds are obtained (7, 24, 105).

**Library 1****Library 2****Library 3****Library 4****Library 5**

**Figure 2** Examples of recent combinatorial library synthesis. Library 1: Diversity-oriented synthesis (DOS) library, seven stereogenic centers, three to four rings, 2500 members (56); Library 2: DOS library, three to four rings, 1412 members (97); Library 3: combinatorial library based on the natural product Galanthamine, 2527 members (71); Library 4: natural-product-like library based on tetrahydroquinoline scaffold, three stereogenic centers, two rings, 27 members (25); Library 5: combinatorial library based on benzopyran scaffold, 10,000 members, DOS library (67).

first attempts to introduce skeletal diversity was a synthesis of kinase inhibitors that incorporated heterocyclic scaffolds as a diversity element (26). In another case, 29,400 polycyclic compounds with 10 distinct skeletons were prepared via a branching DOS pathway (58). Burke and colleagues (15, 16) prepared a collection of ~1260 compounds representing a combinatorial matrix of six molecular



**Figure 3** Use of dynamic combinatorial libraries to discover ligands for proteins. Dynamic combinatorial libraries (DCLs) consist of a set of reversibly interchanging components. These libraries are created by selecting building blocks and reaction conditions that allow dynamic and reversible connections between components. This figure shows construction of a dynamic acylhydrazone library ( $A-B$ ) from a pool of aldehydes ( $A$ ) and hydrazides ( $B$ ) (78). Treating the dynamic combinatorial library with a target protein stabilizes molecules that bind to the protein and enables identification of preferred ligand structures.

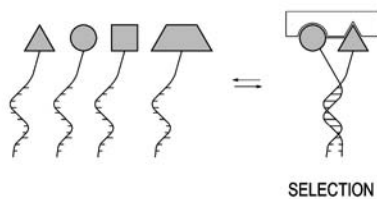
skeletons, each derivatized with a matrix of 105 building blocks in both enantiomeric and diastereomeric forms.

Contrary to the above-listed methods, where defined chemicals are produced, dynamic combinatorial libraries (DCLs) are composed of reversibly interchanging building blocks (76, 78). These libraries are created by selecting chemical fragments and reaction conditions that allow dynamic and reversible connections to be made between the fragments, forming a dynamic mixture in solution (Figure 3). Unfortunately, the method is considerably limited in the types of compatible fragments and reaction conditions that are available. Recent examples include creation of a small carbohydrate library against the plant lectin concanavalin A using disulfide interchange (77) or reversible acylhydrazone formation (78). In addition, a potent bis-cationic heterocyclic inhibitor of the *Bacillus subtilis* HPr kinase/phosphatase was developed using a dynamic combinatorial library derived from 16 hydrazide and 5 mono- and di-aldehyde building blocks (13).

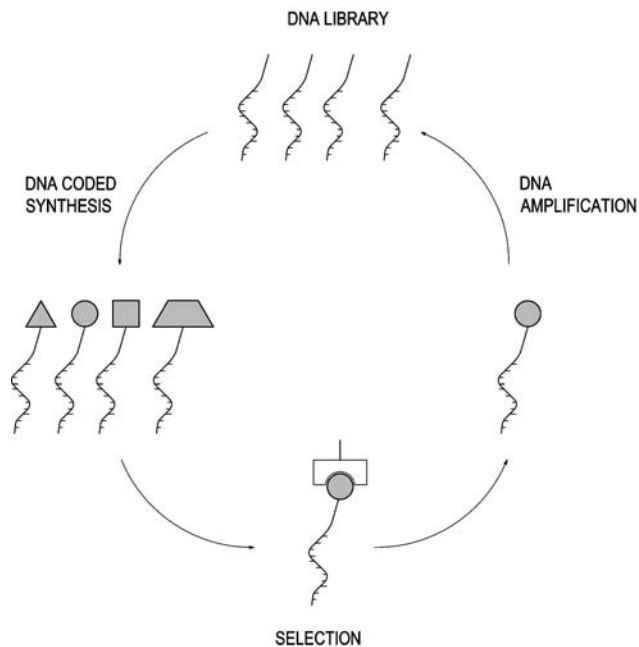
Testing small molecules in biological assays generally requires spatial separation of each test compound, which even with current screening technologies limits the number of small molecules that can be analyzed. Another approach involves testing mixtures and labeling each small molecule with a unique tag that allows isolation of the active compound from a complex mixture. Techniques developed for DNA sequence detection and amplification enable efficient use of DNA tags for such a purpose (Figure 4). To explore this avenue, Melkko and colleagues (64) developed encoded self-assembling chemical libraries (ESACs), in which oligonucleotide sequences carrying small molecules induce self-assembly of the library and are used to identify individual compounds.

Another DNA-based synthesis procedure uses DNA display. These libraries contain single-stranded DNA “genes” with defined sets of “codon” sequences

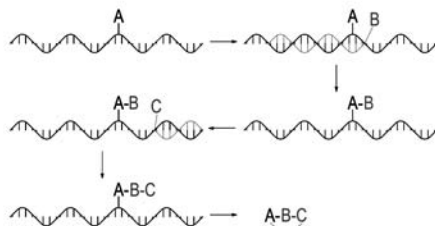
1. ENCODED SELF-ASSEMBLING CHEMICAL LIBRARIES



2. DNA DISPLAY LIBRARIES



3. DNA-TEMPLATED ORGANIC SYNTHESIS





appended with a polyethylene glycol linker. These codons serve three purposes: (a) they specify a reaction sequence for the construction of a small molecule by splitting single-strand DNA into subpools for each reaction step, (b) the DNA sequence functions as an extremely sensitive tag for the constructed molecule, and (c) noncoding regions enable genetic recombination of support sequences by the polymerase chain reaction (PCR) (47–49).

An alternative methodology was applied in DNA-templated organic synthesis (DTS), in which starting materials are linked to 20–50 nucleotide-long DNA templates and combined in very dilute solutions with reagents linked to complementary DNA sequences. DNA sequence pairing increases the effective molarity of the reagents and promotes the desired DNA-directed reaction (38–40, 61). Each synthesized molecule contains a unique DNA tag that is compatible with PCR amplification and, therefore, only a small amount of template is needed for PCR amplification and selection.

## NEW PHENOTYPIC ASSAYS

Once a set of small molecules is assembled, the compounds must be tested in an appropriate biological assay. Forward chemical genetics, like forward genetics, uses phenotypic screening to identify the biological macromolecules that regulate a specific process of interest. Rapid, miniaturized phenotypic screens must then be developed to enable testing of many compounds for their ability to affect a biological process of interest.

In recent years, phenotype-based chemical screens have been extended from single-cell organisms, such as bacteria and fungi, to multicellular organisms such as zebrafish and plants (8, 20). Examples of increasingly complex screens have

---

← **Figure 4** DNA-based chemical libraries. 1. Encoded self-assembling chemical libraries (ESACs) comprise small molecules tethered to oligonucleotides that are responsible for the self-assembly of the library and for the subsequent ability to identify each small molecule. Oligonucleotides can form DNA duplexes (shown in the figure) or DNA triplexes, bringing two or three chemical groups into close proximity. Combinations of small molecules that exhibit synergistic binding to the target protein are isolated by affinity chromatography. 2. DNA display libraries contain single-stranded DNA “genes” with defined sets of “codon” sequences appended with a polyethylene glycol linker. These codons direct the synthesis of a small molecule by splitting ssDNA into appropriate subpools for each reaction step. During the binding assay, the DNA sequence is used to identify and amplify the small molecule that binds to the target. 3. In DNA-templated organic synthesis (DTS), starting materials (*A*, *B*, *C*) are linked to 20–50 nucleotide-long DNA templates and combined in dilute solutions. Complementary DNA sequence pairing increases the effective molarity of the reagents by bringing two reagents into close proximity, and thereby promotes the reaction.

been reported (22). For example, gene expression–based high-throughput screens (GE-HTSs) exploit a gene expression signature, such as the changes in gene expression that occur during cell differentiation, as a readout for cellular states. GE-HTS was used to identify compounds that induce terminal differentiation of acute myelogenous leukemia cells (AML) (98). The mechanisms underlying aborted differentiation in AML are not well characterized, preventing small-molecule screens against well-defined targets. Instead, a collection of 1739 compounds was screened in a cell-based assay using AML tumor cells. A gene expression signature was created that reliably indicated whether these tumor cells had differentiated into neutrophil-like cells. This gene expression signature was monitored in a 384-well format and analyzed after treating these cells with each of the 1789 compounds. The signature was detected using semiquantitative reverse transcription-PCR with single-base extension mass spectrometry. The screen revealed eight novel compounds that induce neutrophil differentiation.

Of increasing interest are genotype-selective antitumor agents that become lethal to cells only in the presence of specific oncoproteins or in the absence of specific tumor suppressor proteins. In a search of small molecules targeting oncogene-expressing tumorigenic cells, Dolma et al. (34) used synthetic lethal screening of tumor cells with defined genetic changes. Normal human fibroblasts were created expressing various combinations of hTERT (the protein component of telomerase), the SV40 (simian virus-40) large and small T oncoproteins, HPV16 (human papillomavirus type 16) E6 and E7 oncoproteins, and oncogenic RAS. These investigators discovered both known and novel compounds with previously unreported genetic selectivities, including a novel compound from a combinatorial library that induces RAS and small T-dependent, nonapoptotic cell death.

Mutagenesis followed by phenotypic screening has been extensively used in vertebrate development studies (43). However, such studies can be limited by the irreversible nature of genetic alterations and by lethal mutations. A chemical approach provides the advantage of a temporal control of gene function. Compounds identified in such screens offer tools for analyzing developmental processes. For example, a phenotype-based screen of zebrafish embryos was used to identify small molecules capable of specifically modulating various aspects of vertebrate ontogeny, including development of the central nervous system, the cardiovascular system, the neural crest, and the ear (72–74). In a medium-throughput screen, synchronously fertilized zebrafish eggs were treated with small molecules in a 96-well format and visually examined under a dissecting microscope. One of the small molecules, concentramide, which disrupts heart patterning, exhibited a similar effect to a genetic mutation of the *heart and soul* gene.

In another such study, a forward chemical genetic screen was performed using a tagged triazine library and zebrafish embryo screening (54). One triazine generated a significant phenotype change in zebrafish brain/eye morphology. This compound was linked to a solid-phase resin and used to identify two proteins: the 40S ribosomal subunit proteins S5 and S18.

## FORWARD CHEMICAL GENETIC SCREENS AND TECHNOLOGIES

Schreiber and colleagues have developed an efficient approach to split-pool, diversity-oriented synthesis using high-capacity macrobeads as individual microreactors (7, 24). This diversity-oriented synthesis platform is based on 500–600- $\mu\text{m}$  polystyrene macrobeads (i.e., large beads compared to the 80–200- $\mu\text{m}$  beads typically used in solid-phase synthesis) that provide robust support and can yield more than 50 nanomoles of compound per bead, which is sufficient for more than 100 screens (105). An encoding strategy is used to record the reaction history for each compound, where chloroaromatic, diazoketone tags are introduced via a rhodium-catalyzed carbene insertion into the aromatic rings of the resin. These tags are released under oxidative conditions with ceric ammonium nitrate, silylated and detected by gas chromatography. This methodology was validated by enantioselective synthesis of an encoded 4320-member library of structurally complex dihydropyranocarboxamides. Clemons and colleagues (24) further developed the platform by optimizing bead array, automating compound cleavage, and robotically implementing library formatting. The one-bead, one-stock solution platform was also used in the split-pool synthesis of 1890 different 1,3-dioxanes, which were evaluated in five phenotypic assays and more than 50 protein-binding experiments (57, 99).

A recent application of this technology involved the Sir2 family of proteins. The Sir2 family of proteins (sirtuins) function in transcription regulation, cell cycle regulation, DNA damage repair, and aging. Sir2p deacetylates histones in a nicotinamide adenine dinucleotide (NAD)-dependent manner, which accounts for its transcriptional silencing activity. A phenotype-based screen revealed three compounds that inhibit transcriptional repression mediated by Sir2p (42). The primary screen identified compounds that inhibit Sir2p-mediated silencing of a URA3 reporter gene and consequently inhibit the yeast strain growth in the presence of 5-fluoroorotic acid. Preliminary studies of one of these inhibitors suggest that sirtuins do not regulate global histone acetylation levels in mammalian cells and that they may be involved in body-axis formation during plant development. Thus, the approach provided novel inhibitors of the transcriptional regulator Sir2p and allowed the study of Sir2p function.

It is possible to use chemical genetics to study the mammalian cell cycle: a chemical genetic modifier screen was performed using a cyto blot (antibody-based) assay to identify small molecules that suppress the G2-phase arrest caused by topoisomerase II inhibitor ICRF-193 (44). The specific suppressors identified in the screen affected cell cycle progression, microtubule stability, and nucleocytoplasmic transport of cyclin B1, but did not affect the chromatin deacetylation checkpoint induced by trichostatin A (TSA). Such cyto blot assays can be used to detect translational and posttranslational changes that mark desired phenotypes (103). In addition, a cyto blot-based screen was used to identify small-molecule suppressors of TSA-induced cell cycle arrest (55).

Another strategy involves screening for compounds that alter the subcellular localization of a protein of interest. For example, Kau et al. (52) used a cell-based small-molecule screen to identify compounds that inhibit FOXO1a nuclear export in PTEN-null tumor cells. This is one of the first examples in which small molecules reversed the consequences of losing a tumor suppressor protein (95): Loss of PTEN was reversed by affecting downstream target proteins.

Chemical genetic screens related to cellular uptake and trafficking have been performed (12, 68, 91). For example, a fluorescence-based assay, which uses calcein fluorescence quenching as a measure of iron uptake, was used to screen mammalian cells for inhibitors of iron uptake (12). Calcein fluorescence was also used to determine cell viability in the absence of iron. This screen identified 10 small-molecule inhibitors of iron transport. Two of these compounds blocked uptake of iron mediated by transferrin, five compounds exhibited cytotoxicity at higher concentration, and three compounds presented interesting probes for dissecting iron uptake mechanisms in mammalian cells.

Finally, combination chemical screening was developed using several different phenotypic assays (9). Borisy et al. developed high-throughput phenotypic screens compatible with testing hundreds of thousands of compound combinations at multiple doses. Numerous combinations were reported that inhibit fungal cell proliferation, inhibit tumor cell proliferation, and suppress production of the proinflammatory cytokine TNF alpha.

## REVERSE CHEMICAL GENETICS: ORTHOGONAL TARGETING OF PROTEINS

Although small molecules can be used to probe protein function, their major disadvantage is lack of specificity. This shortfall becomes particularly apparent in targeting classes of proteins, such as kinases or G-protein coupled receptors (GPCRs), that share highly conserved active sites. Orthogonal chemical genetics relies on a combination of protein engineering and synthesis to explore the functions of individual proteins within such families (3–6, 19, 85, 93). This approach involves modifying a small-molecule ligand to eliminate its binding affinity for its native target (to render it orthogonal) and subsequent mutation of a protein to allow it to recognize the orthogonal ligand. This orthogonality approach has been used to dissect the function of individual GPCRs (79, 85).

Shokat and colleagues have extensively used this approach to study protein kinases. The protein kinases share highly conserved ATP-binding sites; therefore, it is difficult to develop selective kinase inhibitors and dissect their individual functions. Substituting alanine in place of a conserved threonine within an ATP-binding pocket (the “gatekeeper position”) sensitizes kinases toward the ATP competitive inhibitor 1-Naphthyl-PP1 (NaPP1) (36).

The role of cyclin-dependent kinases CDK7 and CDK8 in transcription, specifically in preinitiation complex (PIC) dissociation, was investigated using this

methodology (62). The *Saccharomyces cerevisiae* orthologs of these kinases, Kin28 and Srb10, were engineered to respond to the specific inhibitor Na-PP1. It was determined that Kin28 and Srb10 both have a positive role in promoting transcription by DNA polymerase II, with Kin28 being the dominant kinase.

In a related study, the oncoprotein v-Src was modified to recognize a non-natural phosphate donor substrate N6-(Benzyl)-ATP (A\*TP), which is poorly accepted by wild-type kinases (92). Consequently, only v-Src could transfer radiolabeled A\*TP onto its substrates. A screen for radiolabeled proteins in NIH 3T3 cells in this system identified novel candidate substrates of v-Src: Cofilin, Dok-1, and Calumenin.

Similarly, an engineered Y119G myosin Vb mutant that hydrolyzes ATP, but is selectively sensitized to an N6-substituted ADP analog, was used to identify myosin Vb's physiological functions (75). Mutant myosin Vb was selectively inhibited by N6-(2-phenylethyl)-ADP (PE-ADP), causing it to remain bound to actin, which resulted in decreased levels of transferrin in the cytoplasm and increased levels of plasma-membrane transferrin receptor, suggesting myosin Vb functions in trafficking between peripheral and pericentrosomal compartments. An engineered Y61G myosin-1c was shown to be susceptible to inhibition by N6-modified ADP analogs, such as N6-(2-methylbutyl)ADP (NMB-ADP). Such analogs were used to dissect myosin-1c function in the slow adaptation by hair cells (the sensory cells of the inner ear) (50).

In another example, this protein engineering approach was used to study function of a 1Na-PP1-sensitive mutant of Apg1 kinase (1). It was known that Apg1 is required for the Cvt (cytoplasm-to-vacuole) trafficking pathway and autophagy; however, this study determined that Apg1 kinase activity was required only for Cvt trafficking of aminopeptidase I, but not for import via autophagy.

When a rational bump/hole ligand/receptor combination cannot be designed, it is possible to use selection methods to discover effective bump/hole combinations. For example, Clemons et al. (23) reported the synthesis of calcineurin-resistant derivatives of FK506 and the selection of compensatory calcineurin-derived receptors. These reagents provide small molecule–protein combinations orthogonal to existing chemical dimerizer systems.

A similar ability to regulate individual gene products at the posttranslational level could be achieved by causing selective protein degradation via the ubiquitin-proteasome pathway. Toward this end, Schneekloth and colleagues designed small molecules capable of inducing the degradation of proteins in vivo (87). They showed that PROteolysis TARgeting Chimeric molecules (PROTACS), comprising a ligand for a target protein, a linker, and a ligand for an E3 ligase, selectively facilitated degradation of green fluorescent protein (GFP)-labeled FKBP12 (the 12-kilodalton form of the FK506 binding protein) and GFP-labeled androgen receptor (AR). This PROTAC-mediated approach could provide a general strategy for conditional “chemical knockouts” if high-affinity ligands are developed for proteins of interest and for E3 ligases sufficient for ubiquitination of each protein.

## MECHANISM OF ACTION STUDIES

One of the most difficult challenges for chemical genetic studies is determining the mechanism of action of compounds discovered in phenotypic screens. There are two basic approaches to determining a compound's mechanism of action: hypothesis-based and discovery-based experiments. Within each category, there are several variant strategies that can be employed.

### Hypothesis-Based Mechanism of Action Studies: Pathway Based

One approach to determining mechanism of action uses hypothesis-based experiments based on knowledge of the pathway or phenotype being studied. For example, in one recent study, Verma et al. (109) screened for compounds that block cell cycle progression and discovered a class of compounds, termed ubistatins, that prevent the degradation of ubiquitylated substrates. They found that ubistatins act by binding to the ubiquitin-ubiquitin interface of lysine-48-linked polyubiquitin chains. By using existing mechanistic knowledge of the process they were studying, these investigators discovered the precise mechanism of action of a class of compounds.

A second example of this pathway-based approach is Grozinger et al.'s (42) discovery of small-molecule inhibitors of Sir2p, described previously. They discovered three compounds, including the novel compound "Sirtinol," that inhibit Sir2p-mediated silencing in cells and that directly inhibit Sir2p enzymatic activity *in vitro*. Thus, by knowing that Sir2p is a crucial mediator of chromatin silencing, these investigators determined that these compounds directly inhibit this enzyme.

In a third example of this pathway-based approach, Kau et al. (52) discovered the target of several compounds that prevent nuclear export of the FOXO1a transcription factor. Because the nuclear export receptor CRM1 is the target of leptomycin B, an inhibitor of nuclear export, these investigators tested whether any of the compounds they discovered act by inhibiting CRM1. They found 19 compounds that block nuclear export by reacting with cysteine-528 on CRM1. Most of these compounds had an obvious cysteine-reactive, electrophilic group that would likely react with thiols, and therefore cysteine-528 on CRM1.

Finally, in a fourth example, a screen of 1,3-dioxanes for inhibitors of deacetylases yielded the small-molecule "tubacin," which inhibits tubulin deacetylation in mammalian cells (46). Tubacin was used to identify its cellular target histone deacetylase 6 (HDAC6). Tubacin treatment did not affect the stability of microtubules, but decreased cell motility.

### Hypothesis-Based Mechanism of Action Studies: Annotation Based

A second strategy for determining the mechanism of action of compounds derived from a phenotypic screen is to use a previously annotated compound library. Using

such a strategy, it is possible to link known proteins and compounds to new phenotypes.

For example, Root et al. (80) reported the assembly of a library of biologically active compounds for phenotypic screens. This library comprises approved drugs, controlled substances, and natural products with known targets and mechanisms. The compounds in this library were annotated with more than 12,000 potential mechanisms of action using automated indexing with the Medline database. Such annotated libraries can be useful for identifying known proteins that regulate phenotypes. Root et al. showed that a subset of 85 compounds from this library of 2000+ compounds inhibited proliferation of a tumor cell line and that it was possible to determine the mechanism of action of these 85 compounds by analyzing the associated annotation. This analysis revealed many known methods for inhibiting tumor cell proliferation, such as inhibiting tubulin and binding to topoisomerases, as well as several novel mechanisms, such as potassium ion chelation. Annotated compound libraries can be useful for accelerating some mechanism of action studies.

## Discovery-Based Mechanism of Action Studies: Biochemical Approaches

A second approach to determining mechanism of action uses discovery-based experiments that require no prior knowledge of the pathway or phenotype being studied. Such a discovery-based strategy can use either a biochemical or a genetic approach.

In an example of discovery-based target identification using a biochemical approach, Khersonsky et al. (54) synthesized an affinity matrix based on the structure of the active triazine discovered in a screen for morphological alterations in the brain or eye of the zebrafish *Danio rerio*. This affinity matrix, but not a related control matrix, was used to purify several ribosomal proteins that had been previously implicated as regulators of head and eye morphology using zebrafish genetic screens. Although these targets have not yet been directly validated as the relevant binding proteins for this active triazine, the fact that previous genetic screens implicated these same proteins in eye and head development offers some degree of validation for this approach.

In a second example, Williams et al. (111) screened the same triazine library to identify compounds that induce pigmentation in albino melanocytes. They found six compounds that induce pigmentation by correcting the mistrafficking of tyrosinase that causes the albinism in these cells. Affinity purification experiments with these compounds, and subsequent validation studies, revealed a cellular target to be the mitochondrial F1FO-ATP synthase. In this case a similar result would have been found by screening an annotated compound library.

In a third example, Wan et al. (110) immobilized on a solid-phase resin derivatives of hymenialdisine, which is a natural product kinase inhibitor that inhibits proliferation of leukemia cells. This resin was used to identify new kinase targets

of hymenialdisine, including p90RSK. Subsequent *in vitro* assays confirmed that hymenialdisine directly inhibits p90RSK and several other kinases.

A fourth example of this biochemical approach, reported by Oda et al. (69), involved identifying protein targets of E7070, which is a sulfonamide that perturbs the cell cycle at the G1-S transition. In this case, immobilizing E7070 and an inactive analog (as a negative control matrix) on a solid support resulted in the identification of 285 candidate binding proteins, almost all of which bind to both the E7070 matrix and the inactive analog matrix. To improve their ability to detect E7070 matrix-specific binding proteins, Oda et al. eluted all proteins from both matrices and labeled the E7070 matrix-derived proteins with Cy5 and the control matrix-derived proteins with Cy3. These pools were then mixed and the ratio of each protein in the two pools was determined using 2D gel electrophoresis and fluorescence imaging. This approach revealed that malate dehydrogenase (MDH) specifically bound the E7070 matrix, but not the control matrix; subsequently, the interaction between MDH and E7070 was confirmed using surface plasmon resonance. Thus, differential labeling of resin-binding proteins can be used to increase the sensitivity of protein target detection.

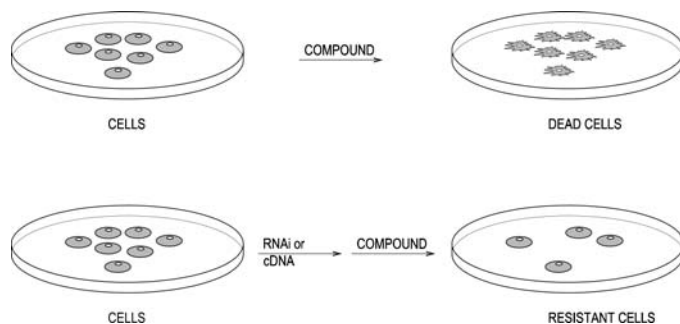
A variation of this biochemical, discovery-based approach involves looking for changes in the level or state of proteins in cells treated with a compound. Towbin et al. (108) used this proteomic-based study to discover that LAF389, a marine natural product analog that inhibits tumor growth, causes a change in the mobility of 14-3-3 $\gamma$  on 2D gels. This change in mobility is due to retention of the aminoterminal methionine, which suggests that LAF389 inhibits methionine aminopeptidases (MetAPs). These authors confirmed that both known MetAPs are inhibited by LAF389. Thus, by observing changes in posttranslational modifications of proteins, it is possible to identify targets of a compound of interest by knowing which enzymes are responsible for those posttranslational changes.

## Discovery-Based Mechanism of Action Studies: Genetic Approaches

A genetic approach to determining the mechanism of action of a compound may entail screening for genetic suppressors or enhancers of the compound (Figure 5). This strategy may reveal the direct target of the compound or related components of the relevant pathway.

In one recent example, Zhao et al. (112) used a genetic approach to determine how the small-molecule sirtinol activates the auxin developmental program in the plant *Arabidopsis*. These investigators screened 60,000 ethylmethane sulfonate mutagenized M2 *Arabidopsis* seeds and identified a single mutant (named *sir1*) that was resistant to sirtinol. Cloning *sir1* revealed that it encodes an E1-like ubiquitin activating enzyme and a Rhodanese-like domain homologous to *cis-trans* peptidyl-prolyl isomerases. The SIR1 protein may be a direct target of sirtinol in *Arabidopsis*, or it may be simply involved in mediating the response to sirtinol. Zhao et al. showed that in either case, *sir1* negatively regulates auxin signaling.





**Figure 5** Use of genetic tools to elucidate the mechanism of action of small molecules. Genetic suppressor and enhancer screens can be used to identify proteins and genes that are involved in the response of cells to a particular compound. For example, a compound that causes cell death (*top*) on its own can be subjected to a genetic suppressor screen that identifies cDNA or siRNA reagents that prevent cell death caused by the compound (*bottom*). Such suppressor reagents are likely to encode proteins involved in the response to the compound, including the direct protein target of the compound (34).

A second, similar genetic screen revealed another genetic suppressor of sirtinol, named AtCAND1, a HEAT-repeat protein that functions in auxin signaling (21). Thus, this genetic screening approach revealed components of the sirtinol-induced and auxin-induced responses in *Arabidopsis*.

In a second example of a genetic approach to mechanism of action, Butcher & Schreiber (17) used 4700 yeast haploid deletion strains to identify targets of small-molecule suppressors of FK506. The natural product FK506 inhibits cell growth under  $\text{Na}^+/\text{Li}^+$  salt stress conditions in the yeast *Saccharomyces cerevisiae* by inhibiting the phosphatase calcineurin. Butcher & Schreiber screened for either small-molecule suppressors of FK506 (SFKs) or yeast gene deletions that allow growth in the presence of FK506 and salt stress conditions. These investigators reasoned that the deletions that cause resistance to FK506/salt stress might be in genes that encode protein targets of the SFKs. They found that one of the deletions that caused resistance to FK506/salt stress, in the *ALD6* gene, encodes a protein that is directly inhibited by several SFKs (18). Thus, screening for deletions that phenocopy a small molecule can be a means to discovering the target of that small molecule.

In a similar approach, Parsons et al. (70, 102) used a haploid deletion panel to detect gene deletions that render yeast cells more sensitive to the microtubule depolymerizer benomyl. They termed the list of gene deletions that make cells more or less sensitive to a compound a “fitness profile.” For example, deleting genes involved in mitosis rendered yeast cells more sensitive to benomyl. Among these genes was *TUB3*, a component of microtubules, the direct target of benomyl.

A similar strategy was taken by Lesage et al. (60) for analyzing the mechanism of action of the glucan synthase inhibitor caspofungin. Deletion of *FKS1*, which encodes the direct target of caspofungin, led to increased sensitivity to this compound. Thus, in some cases, deleting the gene that encodes the protein target of a compound will render cells more sensitive to that compound.

Parsons et al. (70) found that the fitness profile for a compound is often similar to the fitness profile for a yeast strain with a deletion in the gene encoding the compound's target. This latter profile is obtained by testing the relative growth rates of yeast strains containing a deletion in the target as well as each of the other 4700 gene deletions, in 4700 separate experiments. For example, the fitness profile of benomyl is similar to the fitness profile of a *TUB2* deletion, another component of microtubules. At present, this strategy is limited to the analysis of yeast genes, although one can imagine extending the approach to mammalian systems using genome-scale RNA interference.

## GLOBAL ANALYSES

There are two approaches to analyzing large data sets: screens and global analyses (81, 84). Both approaches involve collecting a large amount of data on the effects of specific compounds (104). However, the goals of the two approaches differ: screens seek to identify active reagents that can be pursued in subsequent experiments, whereas global analyses seek to draw conclusions regarding all of the reagents tested in a screen. A high rate of false negatives and false positives can be tolerated in a screen because as long as a few true positives can ultimately be confirmed, the screen is successful. Unfortunately, the same is not true for global analyses, which require low false positive and false negative rates to meaningfully interpret the data.

Both types of high-throughput approaches generate immense quantities of data that require sophisticated data analysis tools. Kelley et al. (53) created a freely available software tool, SLIMS (Small Laboratory Information Management System), that facilitates analysis of large-scale screening data. In this system, raw data from high-throughput assays can be normalized and systematic spatial errors are automatically identified and corrected (82). If annotated compounds are used, published literature associated with active compounds can be automatically retrieved from Medline and processed to yield potential mechanisms of actions. Thus, SLIMS provides a framework for analyzing high-throughput assay data both as a laboratory information management system and as a platform for experimental analysis.

Several groups have begun performing global analyses. Haggarty et al. (45) collected a matrix of data using a set of 7392 different 1,3-dioxanes with appended metal chelation moieties in a set of protein acetylation assays. They calculated the Euclidian distance between both compounds and assays by treating each point in the matrix as a point in multidimensional assay data space. This analysis allowed

these authors to determine which compounds were functionally similar (as opposed to structurally similar) and which assays were functionally similar, in that they yielded a similar set of active compounds. Thus, obtaining data on the same set of compounds in multiple assays in a systematic fashion enables meaningful comparisons between both the compounds and the assays.

## CONCLUSION

An increasing number of investigators are turning to chemical approaches to probe signaling pathways and phenotypes of interest. Over the past four years, it has become clear that procuring or creating tens of thousands of compounds is feasible. Similarly, many different high-throughput assays can be developed that report on cellular and molecular processes in mammalian systems. Thus, compound procurement and screening no longer represent rate-limiting steps in the chemical genetic process. Nonetheless, it would be advantageous to create synthetic routes to more complex and diverse compound libraries, which might increase the potency and selectivity of compounds that emerge from high-throughput screens. In addition, more complex phenotypic readouts, such as those obtained with automated imaging, would increase the range of processes that can be studied using this approach. We can expect the fields of high-throughput chemical synthesis and high-throughput screening to continue to advance toward these goals over the next several years.

Mechanism of action studies represent the most intractable part of the chemical genetic process. We have seen that there are a wide variety of strategies that may be employed in the study of a particular compound of interest: genetic, biochemical, and hypothesis-based approaches can all be fruitful under some circumstances. Nonetheless, all of these approaches have disadvantages that reduce their generality: biochemical approaches require high-affinity compounds, genetic approaches require comprehensive collections of genetic reagents, and hypothesis-based experiments cannot always be applied to a case of interest. Thus, significant efforts need to be directed toward creating more generally applicable approaches for mechanism-of-action studies. If systematic, algorithmic methods were available, the chemical genetic approach would become an increasingly attractive one for studying biological systems that are not compatible with traditional genetic approaches.

## ABBREVIATIONS

AAK1: AP2 associated kinase 1  
ADP: adenosine diphosphate  
AML: acute myelogenous leukemia  
AR: androgen receptor

- ATP: adenosine triphosphate  
C175A: cysteine 175 substitution with alanine  
CAN: ceric ammonium nitrate  
CDK7: cyclin-dependent protein kinase 7  
CDK8: cyclin-dependent protein kinase 8  
CRM1: major karyopherin, involved in exporting proteins, RNAs, and ribosomal subunits from the nucleus of *Saccharomyces cerevisiae*  
DNA: deoxyribonucleic acid  
DOS: diversity-oriented synthesis  
DCL: dynamic combinatorial libraries  
E7070: a sulfonamide that perturbs the cell cycle at the G1-S transition  
ESAC: encoded self-assembling chemical libraries  
FKBP12: FK506 binding protein, 12 kDa  
FKS1: catalytic subunit of 1,3-beta-D-glucan synthase, target of caspofungin  
FOXO1a: forkhead box O1A of the forkhead family of transcription factors  
GAK: cyclin G associated kinase  
GE-HTS: gene expression-based high-throughput screening  
GFP: green fluorescent protein  
GPCRs: G-protein coupled receptors  
GPD1: glycerol-3-phosphate dehydrogenase 1  
HDAC: histone deacetylase  
HPV 16: human papillomavirus type 16  
hTERT: human telomerase reverse transcriptase  
LAF389: a marine natural product analog that inhibits tumor growth  
M108G: methionine 108 substitution with glycine  
MDH: malate dehydrogenase  
MetAPs: methionine aminopeptidases  
mTOR: mammalian target of rapamycin  
 $\mu\text{m}$ : micrometers  
NaPP1: ATP competitive inhibitor 1-Naphthyl-PP1  
p90RSK: ribosomal protein S6 kinase, 90kD, polypeptide 2  
PCR: polymerase chain reaction  
PE-ADP: N6-(2-phenylethyl)-ADP  
PIC: preinitiation complex  
Pol II: component of RNA polymerase II holoenzyme  
PS: polystyrene  
PTEN: phosphatase and tensin homolog  
RNA: ribonucleic acid  
SAR: structure-activity relationship  
SFks: small-molecule suppressors of FK506  
Sir2p: transcription protein, regulator of silencing at HML, HMR, telomeres, and rDNA  
SV40: simian virus-40  
TG-3 mAb: monoclonal antibody that recognizes phosphorylated nucleolin

TG: triethyleneglycol  
 TSA: trichostatin A  
 TUB2: beta-tubulin  
 URA3: orotidine-5'-phosphate decarboxylase (*Saccharomyces cerevisiae*)  
 YGGFL: pentapeptide: tyrosyl-glycyl-glycyl-phenylalanyl-leucine  
 Y119G: tyrosine 119 substitution with glycine

## ACKNOWLEDGMENTS

Brent R. Stockwell is supported in part by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund.

**The Annual Review of Genomics and Human Genetics is online at  
<http://genom.annualreviews.org>**

## LITERATURE CITED

1. Abeliovich H, Zhang C, Dunn WA Jr, Shokat KM, Klionsky DJ. 2003. Chemical genetic analysis of Apg1 reveals a non-kinase role in the induction of autophagy. *Mol. Biol. Cell.* 14:477–90
2. Arya P, Joseph R, Chou DT. 2002. Toward high-throughput synthesis of complex natural product-like compounds in the genomics and proteomics age. *Chem. Biol.* 9:145–56
3. Belshaw PJ, Schoepfer J, Liu K, Morrison K, Schreiber SL. 1995. Rational design of orthogonal receptor-ligand combinations. *Angew. Chem. Int. Ed. Eng.* 34:2129–32
4. Bishop A, Buzko O, Heyeck-Dumas S, Jung I, Kraybill B, et al. 2000. Unnatural ligands for engineered proteins: new tools for chemical genetics. *Annu. Rev. Biophys. Biomol. Struct.* 29:577–606
5. Bishop AC, Buzko O, Shokat KM. 2001. Magic bullets for protein kinases. *Trends Cell. Biol.* 11:167–72
6. Bishop AC, Ubersax JA, Petsch DT, Matheos DP, Gray NS, et al. 2000. A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* 407:395–401
7. Blackwell HE, Perez L, Stavenger RA, Tallarico JA, Cope Eatough E, et al. 2001. A one-bead, one-stock solution approach to chemical genetics: part 1. *Chem. Biol.* 8:1167–82
8. Blackwell HE, Zhao Y. 2003. Chemical genetic approaches to plant biology. *Plant Physiol.* 133:448–55
9. Borisy AA, Elliott PJ, Hurst NW, Lee MS, Lehar J, et al. 2003. Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. USA* 100:7977–82
10. Breinbauer R, Manger M, Scheck M, Waldmann H. 2002. Natural product guided compound library development. *Curr. Med. Chem.* 9:2129–45
11. Brown EJ, Albers MW, Shin TB, Ichikawa K, Keith CT, et al. 1994. A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature* 369:756–58
12. Brown JX, Buckett PD, Wessling-Resnick M. 2004. Identification of small molecule inhibitors that distinguish between non-transferrin bound iron uptake and transferrin-mediated iron transport. *Chem. Biol.* 11:407–16
13. Bunyapaiboonsri T, Ramstrom H, Ramstrom O, Haiech J, Lehn JM. 2003. Generation of bis-cationic heterocyclic inhibitors of *Bacillus subtilis* HPr kinase/phosphatase from a ditopic dynamic

- combinatorial library. *J. Med. Chem.* 46: 5803–11
14. Burdine L, Kodadek T. 2004. Target identification in chemical genetics: the (often) missing link. *Chem. Biol.* 11:593–97
  15. Burke MD, Berger EM, Schreiber SL. 2003. Generating diverse skeletons of small molecules combinatorially. *Science* 302:613–18
  16. Burke MD, Schreiber SL. 2004. A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed. Engl.* 43: 46–58
  17. Butcher RA, Schreiber SL. 2003. A small molecule suppressor of FK506 that targets the mitochondria and modulates ionic balance in *Saccharomyces cerevisiae*. *Chem. Biol.* 10:521–31
  18. Butcher RA, Schreiber SL. 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. USA* 101:7868–73
  19. Carroll AS, Bishop AC, DeRisi JL, Shokat KM, O'Shea EK. 2001. Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. *Proc. Natl. Acad. Sci. USA* 98:12578–83
  20. Carroll PM, Dougherty B, Ross-Macdonald P, Browman K, FitzGerald K. 2003. Model systems in drug discovery: chemical genetics meets genomics. *Pharmacol. Ther.* 99:183–220
  21. Cheng Y, Dai X, Zhao Y. 2004. AtCAND1, a HEAT-repeat protein that participates in auxin signaling in Arabidopsis. *Plant Physiol.* 135:1020–26
  22. Clemons PA. 2004. Complex phenotypic assays in high-throughput screening. *Curr. Opin. Chem. Biol.* 8:334–38
  23. Clemons PA, Gladstone BG, Seth A, Chao ED, Foley MA, Schreiber SL. 2002. Synthesis of calcineurin-resistant derivatives of FK506 and selection of compensatory receptors. *Chem. Biol.* 9:49–61
  24. Clemons PA, Koehler AN, Wagner BK, Sprigings TG, Spring DR, et al. 2001. A one-bead, one-stock solution approach to chemical genetics: Part 2. *Chem. Biol.* 8:1183–95
  25. Couve-Bonnaire S, Chou DT, Gan Z, Arya P. 2004. A solid-phase, library synthesis of natural-product-like derivatives from an enantiomerically pure tetrahydroquinoline scaffold. *J. Comb. Chem.* 6:73–77
  26. Ding S, Gray NS, Wu X, Ding Q, Schultz PG. 2002. A combinatorial scaffold approach toward kinase-directed heterocycle libraries. *J. Am. Chem. Soc.* 124:1594–96
  27. Dolle RE. 1998. Comprehensive survey of combinatorial libraries with undisclosed biological activity: 1992–1997. *Mol. Divers.* 4:233–56
  28. Dolle RE. 2000. Comprehensive survey of combinatorial library synthesis: 1999. *J. Comb. Chem.* 2:383–433
  29. Dolle RE. 2001. Comprehensive survey of combinatorial library synthesis: 2000. *J. Comb. Chem.* 3:477–517
  30. Dolle RE. 2002. Comprehensive survey of combinatorial library synthesis: 2001. *J. Comb. Chem.* 4:369–418
  31. Dolle RE. 2003. Comprehensive survey of combinatorial library synthesis: 2002. *J. Comb. Chem.* 5:693–753
  32. Dolle RE. 2004. Comprehensive survey of combinatorial library synthesis: 2003. *J. Comb. Chem.* 6:623–79
  33. Dolle RE, Nelson KH Jr. 1999. Comprehensive survey of combinatorial library synthesis: 1998. *J. Comb. Chem.* 1:235–82
  34. Dolma S, Lessnick SL, Hahn WC, Stockwell BR. 2003. Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells. *Cancer Cell* 3:285–96
  35. Ellman JA. 2000. Combinatorial methods to engineer small molecules for functional genomics. *Ernst. Schering. Res. Found. Workshop* 32:183–204
  36. Fan QW, Specht KM, Zhang C,

- Goldenberg DD, Shokat KM, Weiss WA. 2003. Combinatorial efficacy achieved through two-point blockade within a signaling pathway—a chemical genetic approach. *Cancer Res.* 63:8930–38
37. Furka A, Sebestyen F, Asgedom M, Dido G. 1991. General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* 37:487–93
38. Gartner ZJ, Grubina R, Calderone CT, Liu DR. 2003. Two enabling architectures for DNA-templated organic synthesis. *Angew. Chem. Int. Ed. Engl.* 42:1370–75
39. Gartner ZJ, Kanan MW, Liu DR. 2002. Multistep small-molecule synthesis programmed by DNA templates. *J. Am. Chem. Soc.* 124:10304–6
40. Gartner ZJ, Tse BN, Grubina R, Doyon JB, Snyder TM, Liu DR. 2004. DNA-templated organic synthesis and selection of a library of macrocycles. *Science* 305:1601–5
41. Geysen HM, Rodda SJ, Mason TJ. 1986. The delineation of peptides able to mimic assembled epitopes. *Ciba Found. Symp.* 119:130–49
42. Grozinger CM, Chao ED, Blackwell HE, Moazed D, Schreiber SL. 2001. Identification of a class of small molecule inhibitors of the sirtuin family of NAD-dependent deacetylases by phenotypic screening. *J. Biol. Chem.* 276:38837–43
43. Haffter P, Nusslein-Volhard C. 1996. Large scale genetics in a small vertebrate, the zebrafish. *Int. J. Dev. Biol.* 40:221–27
44. Haggarty SJ, Koeller KM, Kau TR, Silver PA, Roberge M, Schreiber SL. 2003. Small molecule modulation of the human chromatid decatenation checkpoint. *Chem. Biol.* 10:1267–79
45. Haggarty SJ, Koeller KM, Wong JC, Butcher RA, Schreiber SL. 2003. Multidimensional chemical genetic analysis of diversity-oriented synthesis-derived deacetylase inhibitors using cell-based assays. *Chem. Biol.* 10:383–96
46. Haggarty SJ, Koeller KM, Wong JC, Grozinger CM, Schreiber SL. 2003. Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation. *Proc. Natl. Acad. Sci. USA* 100:4389–94
47. Halpin DR, Harbury PB. 2004. DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol.* 2:E173
48. Halpin DR, Harbury PB. 2004. DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol.* 2:E174
49. Halpin DR, Lee JA, Wrenn SJ, Harbury PB. 2004. DNA display III. Solid-phase organic synthesis on unprotected DNA. *PLoS Biol.* 2:E175
50. Holt JR, Gillespie SK, Provance DW, Shah K, Shokat KM, et al. 2002. A chemical-genetic strategy implicates myosin-1c in adaptation by hair cells. *Cell* 108:371–81
51. Houghton RL, Benson DR, Reynolds LD, McNeill PD, Sleath PR, et al. 1999. A multi-epitope synthetic peptide and recombinant protein for the detection of antibodies to *Trypanosoma cruzi* in radioimmunoprecipitation-confirmed and consensus-positive sera. *J. Infect. Dis.* 179:1226–34
52. Kau TR, Schroeder F, Ramaswamy S, Wojciechowski CL, Zhao JJ, et al. 2003. A chemical genetic screen identifies inhibitors of regulated nuclear export of a Forkhead transcription factor in PTEN-deficient tumor cells. *Cancer Cell* 4:463–76
53. Kelley BP, Lunn MR, Root DE, Flaherty SP, Martino AM, Stockwell BR. 2004. A flexible data analysis tool for chemical genetic screens. *Chem. Biol.* 11(11):1495–503
54. Khersonsky SM, Jung DW, Kang TW, Walsh DP, Moon HS, et al. 2003. Facilitated forward chemical genetics using a tagged triazine library and zebrafish embryo screening. *J. Am. Chem. Soc.* 125:11804–5

55. Koeller KM, Haggarty SJ, Perkins BD, Leykin I, Wong JC, et al. 2003. Chemical genetic modifier screens. Small molecule trichostatin suppressors as probes of intracellular histone and tubulin acetylation. *Chem. Biol.* 10:397–410
56. Kubota H, Lim J, Depew KM, Schreiber SL. 2002. Pathway development and pilot library realization in diversity-oriented synthesis: exploring Ferrier and Pauson-Khand reactions on a glycol template. *Chem. Biol.* 9:265–76
57. Kuruvilla FG, Shamji AF, Sternson SM, Hergenrother PJ, Schreiber SL. 2002. Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays. *Nature* 416:653–57
58. Kwon O, Park SB, Schreiber SL. 2002. Skeletal diversity via a branched pathway: efficient synthesis of 29 400 discrete, polycyclic compounds and their arraying into stock solutions. *J. Am. Chem. Soc.* 124:13402–4
59. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ. 1991. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 354:82–84
60. Lesage G, Sdicu AM, Menard P, Shapiro J, Hussein S, Bussey H. 2004. Analysis of beta-1,3-glucan assembly in *Saccharomyces cerevisiae* using a synthetic interaction network and altered sensitivity to caspofungin. *Genetics* 167:35–49
61. Liu DR. 2004. Translating DNA into synthetic molecules. *PLoS Biol.* 2:E223
62. Liu Y, Kung C, Fishburn J, Ansari AZ, Shokat KM, Hahn S. 2004. Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol. Cell. Biol.* 24:1721–35
63. Mann J, Davidson RS, Hobbs JB, Bantorpe DV, Harborne JB. 1994. *Natural Products. Their Chemistry and Biological Significance*. Essex: Longman. 455 pp.
64. Melkko S, Scheuermann J, Dumelin CE, Neri D. 2004. Encoded self-assembling chemical libraries. *Nat. Biotechnol.* 22:568–74
65. Merrifield B. 1997. Concept and early development of solid-phase peptide synthesis. *Methods Enzymol.* 289:3–13
66. Myers AG, Lanman BA. 2002. A solid-supported, enantioselective synthesis suitable for the rapid preparation of large numbers of diverse structural analogues of (-)-saframycin A. *J. Am. Chem. Soc.* 124:12969–71
67. Nicolaou KC, Pfefferkorn JA, Mitchell HJ, Roecker AJ, Barluenga S, et al. 2000. Natural product-like combinatorial libraries based on privileged structures. 2. Construction of a 10,000-membered benzopyran library by directed split-and-pool chemistry using NanoKans and optical encoding. *J. Am. Chem. Soc.* 122:9954–67
68. Nieland TJ, Feng Y, Brown JX, Chuang TD, Buckett PD, et al. 2004. Chemical genetic screening identifies sulfonamides that raise organellar pH and interfere with membrane traffic. *Traffic* 5:478–92
69. Oda Y, Owa T, Sato T, Boucher B, Daniels S, et al. 2003. Quantitative chemical proteomics for identifying candidate drug targets. *Anal. Chem.* 75:2159–65
70. Parsons AB, Brost RL, Ding H, Li Z, Zhang C, et al. 2003. Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat. Biotechnol.* 22(1):62–69
71. Pelish HE, Westwood NJ, Feng Y, Kirchhausen T, Shair MD. 2001. Use of biomimetic diversity-oriented synthesis to discover galanthamine-like molecules with biological properties beyond those of the natural product. *J. Am. Chem. Soc.* 123:6740–41
72. Peterson RT, Link BA, Dowling JE, Schreiber SL. 2000. Small molecule developmental screens reveal the logic and timing of vertebrate development. *Proc. Natl. Acad. Sci. USA* 97:12965–69
73. Peterson RT, Mably JD, Chen JN, Fishman MC. 2001. Convergence of distinct



- pathways to heart patterning revealed by the small molecule concentramide and the mutation heart-and-soul. *Curr. Biol.* 11:1481–91
74. Peterson RT, Shaw SY, Peterson TA, Milan DJ, Zhong TP, et al. 2004. Chemical suppression of a genetic mutation in a zebrafish model of aortic coarctation. *Nat. Biotechnol.* 22:595–99
75. Provance DW Jr, Gourley CR, Silan CM, Cameron LC, Shokat KM, et al. 2004. Chemical-genetic inhibition of a sensitized mutant myosin Vb demonstrates a role in peripheral-pericentriolar membrane traffic. *Proc. Natl. Acad. Sci. USA* 101:1868–73
76. Ramstrom O, Bunyapaboonsri T, Lohmann S, Lehn JM. 2002. Chemical biology of dynamic combinatorial libraries. *Biochim. Biophys. Acta* 1572:178–86
77. Ramstrom O, Lehn JM. 2000. In situ generation and screening of a dynamic combinatorial carbohydrate library against concanavalin A. *ChemBioChem* 1:41–48
78. Ramstrom O, Lohmann S, Bunyapaboonsri T, Lehn JM. 2004. Dynamic combinatorial carbohydrate libraries: probing the binding site of the concanavalin A lectin. *Chemistry* 10:1711–15
79. Redfern CH, Coward P, Degtyarev MY, Lee EK, Kwa AT, et al. 1999. Conditional expression and signaling of a specifically designed Gi-coupled receptor in transgenic mice. *Nat. Biotechnol.* 17:165–69
80. Root DE, Flaherty SP, Kelley BP, Stockwell BR. 2003. Biological mechanism profiling using an annotated compound library. *Chem. Biol.* 10:881–92
81. Root DE, Kelley BP, Stockwell BR. 2002. Global analysis of large-scale chemical and biological experiments. *Curr. Opin. Drug Discov. Dev.* 5:355–60
82. Root DE, Kelley BP, Stockwell BR. 2003. Detecting spatial patterns in biological array experiments. *J. Biomol. Screen* 8:393–98
83. Sabatini DM, Erdjument-Bromage H, Lui M, Tempst P, Snyder SH. 1994. RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell* 78:35–43
84. Savchuk NP, Balakin KV, Tkachenko SE. 2004. Exploring the chemogenomic knowledge space with annotated chemical libraries. *Curr. Opin. Chem. Biol.* 8:412–17
85. Searce-Levie K, Coward P, Redfern CH, Conklin BR. 2001. Engineering receptors activated solely by synthetic ligands (RASSLs). *Trends Pharmacol. Sci.* 22:414–20
86. Schmidt DR, Kwon O, Schreiber SL. 2004. Macrolactones in diversity-oriented synthesis: preparation of a pilot library and exploration of factors controlling macrocyclization. *J. Comb. Chem.* 6:286–92
87. Schneekloth JS Jr, Fonseca FN, Koldobskiy M, Mandal A, Deshaies R, et al. 2004. Chemical genetic control of protein levels: selective in vivo targeted degradation. *J. Am. Chem. Soc.* 126:3748–54
88. Schreiber SL. 1998. Chemical genetics resulting from a passion for synthetic organic chemistry. *Bioorg. Med. Chem.* 6:1127–52
89. Schreiber SL. 2000. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287:1964–69
90. Schreiber SL. 2003. The small-molecule approach to biology: Chemical genetics and diversity-oriented organic synthesis make possible the systematic exploration of biology. *Chem. Eng. News* 81:51–61
91. Sekiya-Kawasaki M, Groen AC, Cope MJ, Kaksonen M, Watson HA, et al. 2003. Dynamic phosphoregulation of the cortical actin cytoskeleton and endocytic machinery revealed by real-time chemical genetic analysis. *J. Cell. Biol.* 162:765–72
92. Shah K, Shokat KM. 2002. A chemical genetic screen for direct v-Src substrates reveals ordered assembly of a retrograde signaling pathway. *Chem. Biol.* 9:35–47

93. Shokat K, Velleca M. 2002. Novel chemical genetic approaches to the discovery of signal transduction inhibitors. *Drug Discov. Today* 7:872–79
94. Shu YZ. 1998. Recent natural products based drug development: a pharmaceutical industry perspective. *J. Nat. Prod.* 61: 1053–71
95. Smukste I, Stockwell BR. 2003. Restoring functions of tumor suppressors with small molecules. *Cancer Cell* 4:419–20
96. Spring DR. 2003. Diversity-oriented synthesis; a challenge for synthetic chemists. *Org. Biomol. Chem.* 1:3867–70
97. Spring DR, Krishnan S, Blackwell HE, Schreiber SL. 2002. Diversity-oriented synthesis of biaryl-containing medium rings using a one bead/one stock solution platform. *J. Am. Chem. Soc.* 124:1354–63
98. Stegmaier K, Ross KN, Colavito SA, O'Malley S, Stockwell BR, Golub TR. 2004. Gene expression-based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nat. Genet.* 36:257–63
99. Sternson SM, Louca JB, Wong JC, Schreiber SL. 2001. Split-pool synthesis of 1,3-dioxanes leading to arrayed stock solutions of single compounds sufficient for multiple phenotypic and protein-binding assays. *J. Am. Chem. Soc.* 123: 1740–47
100. Stockwell BR. 2000. Chemical genetics: ligand-based discovery of gene function. *Nat. Rev. Genet.* 1:116–25
101. Stockwell BR. 2000. Frontiers in chemical genetics. *Trends Biotechnol.* 18:449–55
102. Stockwell BR. 2004. The biological magic behind the bullets. *Nat. Biotechnol.* 22:37–38
103. Stockwell BR, Haggarty SJ, Schreiber SL. 1999. High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem. Biol.* 6:71–83
104. Strausberg RL, Schreiber SL. 2003. From knowing to controlling: a path from genomics to drugs using small molecule probes. *Science* 300:294–95
105. Tallarico JA, Depew KM, Pelish HE, Westwood NJ, Lindsley CW, et al. 2001. An alkylsilyl-tethered, high-capacity solid support amenable to diversity-oriented synthesis for one-bead, one-stock solution chemical genetics. *J. Comb. Chem.* 3:312–18
106. Tan DS, Foley MA, Shair MD, Schreiber SL. 1998. Stereoselective synthesis of over two million compounds having structural features both reminiscent of natural products and compatible with miniaturized cell-based assays. *J. Am. Chem. Soc.* 120:8565–66
107. Tan DS, Foley MA, Stockwell BR, Shair MD, Schreiber SL. 1999. Synthesis and preliminary evaluation of a library of polycyclic small molecules for use in chemical genetic assays. *J. Am. Chem. Soc.* 121:9073–87
108. Towbin H, Bair KW, DeCaprio JA, Eck MJ, Kim S, et al. 2003. Proteomics-based target identification: bengamides as a new class of methionine aminopeptidase inhibitors. *J. Biol. Chem.* 278:52964–71
109. Verma R, Peters NR, D'Onofrio M, Tochtrop GP, Sakamoto KM, et al. 2004. Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. *Science* 306:117–20
110. Wan Y, Hur W, Cho CY, Liu Y, Adrian FJ, et al. 2004. Synthesis and target identification of hymenialdisine analogs. *Chem. Biol.* 11:247–59
111. Williams D, Jung DW, Khersonsky SM, Heidary N, Chang YT, Orlow SJ. 2004. Identification of compounds that bind mitochondrial F1F0 ATPase by screening a triazine library for correction of albinism. *Chem. Biol.* 11:1251–59
112. Zhao Y, Dai X, Blackwell HE, Schreiber SL, Chory J. 2003. SIR1, an upstream component in auxin signaling identified by chemical genetics. *Science* 301:1107–10