Part I
Concept of Screening

1
Chemical Genetics: Use of High-throughput Screening to Identify Small-molecule Modulators of Proteins Involved in Cellular Pathways with the Aim of Uncovering Protein Function

Sa L. Chiang

1.1
Introduction

Understanding cellular pathways and their molecular mechanisms is one of the longstanding goals of scientific research. However, the tools that researchers utilize to study biological processes become progressively more sophisticated as technology and our knowledge of biology advance. This chapter discusses one of the most recent and exciting developments in this area: chemical genetics and the use of small, bioactive molecules to characterize the components of cellular pathways and their functions.

1.2
Classical and Chemical Genetics

All genetic approaches depend on the ability to perturb gene function and to correlate phenotypic changes with specific changes in gene function. Classical genetics relies on physical perturbation of a gene, through methods such as irradiation, chemical mutagenesis or insertional mutagenesis (e.g. via the use of transposons). There are also numerous molecular biology techniques for creating directed mutations that allow highly specific modifications at the level of the gene or even the nucleotide in a variety of experimental systems. These methods are well established and readers should consult a standard genetics text for the relevant descriptions.

A common feature of all classical genetic methods is that they cause a permanent change in the structure of a gene. Therefore, except for the few situations noted below, the phenotypes arising from classical genetic mutations are irreversible. A notable disadvantage of this irreversibility is that it hinders the study of genes that are essential for viability. Irreversibility also makes it difficult to study the effects of

High-Throughput Screening in Drug Discovery. Edited by Jörg Hüser

ISBN: 3-527-31283-8

temporal variations in gene expression or protein function. Placing genes under the control of inducible promoters has gone some way towards solving these problems, but inducible promoters, of course, act at the level of transcription and the researcher may not obtain sufficient control over the activity of the encoded protein. Temperature-sensitive mutations may provide some control over protein activity, but they are not easy to construct and generally cannot be used in animal models. Pleiotropic effects caused by temperature shifts may also complicate analysis.

In the last decade, researchers have increasingly explored the use of low molecular weight chemical entities to modulate and characterize protein function. These methods are generally analogous to classical genetic approaches and have accordingly been termed *chemical genetic* strategies [1, 2]. Whereas classical genetics uses physical modification of a gene to perturb protein function, chemical genetics employs specific, biologically active small-molecule modulators to perturb protein function.

It is relatively simple to imagine mechanisms by which small molecules may modulate protein function. Enzymatic activity can be affected by the binding of small molecules to active or allosteric sites and protein–ligand interactions may be disrupted by small molecules that interfere with binding between interaction partners. Alternatively, interaction partners could be brought together more effectively by ligands that bind to both. Interactions between small molecules and their targets may be reversible or irreversible and protein function can be either diminished or enhanced, depending on the situation.

One advantage of small molecules is that they can be used in biological systems where there is little or no ability for classical genetic manipulation. Moreover, since many small molecules will not interact irreversibly with their targets, chemical genetics is expected to provide enhanced opportunities to create conditional phenotypes. Theoretically, a high degree of control over protein function could be afforded by simply adding or washing away a small-molecule modulator. Indeed, several widely used inducible promoter systems (lac, arabinose, tetracycline) employ small-molecule inducers. In addition, conditional phenotypes induced by small molecules can potentially be studied in animal models, where temperature-dependent phenotypes generally cannot be used.

In actuality, the ability of small molecules to induce and reverse phenotypes will depend on factors such as binding kinetics and the physical accessibility of the target. More important, most protein targets have no known small-molecule modulators of their activity. Obtaining a specific and potent small-molecule modulator for a chosen target often requires structure-based drug design or full-scale high-throughput screening (HTS) and medicinal chemistry optimization may also be necessary. Owing to this "front-end" effort, directed perturbations in many systems are currently more difficult to achieve with chemical genetics than with classical genetics.

1.2.1 Forward and Reverse Screens

Many classical genetic screens begin with mutagenesis of organisms or populations of cells, followed by attempts to associate the resulting phenotypic changes with specific genes. These approaches are termed *forward genetic* strategies and can be generally characterized as starting with phenotypes and progressing towards the identification of the genes that are responsible for those phenotypes. With the advent of molecular biology, directed mutagenesis techniques became available and these advances allowed *reverse genetic* studies, which begin with the introduction of programmed or directed mutations into a known target gene, followed by analysis of the resulting phenotypes to obtain information about the function of that gene.

Both forward and reverse genetic strategies have also been employed for chemical screens. In a forward chemical genetic screen, chemical libraries are screened for compounds that produce a phenotype of interest, typically in a cell-based assay. An example of forward chemical genetics is the strategy that was employed to identify inhibitors of mitotic spindle bipolarity [3]. Monastrol was identified through a chemical genetic screen employing a whole-cell immunodetection assay to screen a library of 16 320 compounds for those that increase nucleolin phosphorylation, a phenotype predicted for cells experiencing mitotic arrest. A total of 139 compounds found to increase nucleolin phosphorylation were subjected to further analysis, resulting in the identification of five compounds that affect the structure of the mitotic spindle. One of these induced the formation of a monoastral microtubule array and was accordingly named monastrol. Monastrol has since been employed in multiple studies as a tool for investigating the process of cell division.

Conversely, in a reverse chemical genetic screen, small-molecule libraries might be screened for compounds that bind to a purified protein target, modulate the activity of the target or affect the target's ability to interact with other proteins. Such compounds could then be used in cell-based assays to characterize the function of the target protein in cellular pathways. For instance, a luminescence-based reverse chemical genetic screen was employed to identify inhibitors of rabbit muscle myosin II subfragment (S1) actin-stimulated ATPase activity [4]. The most potent compound identified (*N*-benzyl-*p*-toluenesulfonamide; BTS) inhibited S1 ATPase activity with an IC_{50} of ~5 μM and BTS also inhibited the activity of skeletal muscle myosin in a gliding-filament assay. Subsequent studies demonstrated that BTS inhibits the binding of myosin–ADP to actin and also affects various properties of rabbit and frog muscle preparations.

A similar reverse chemical genetic screen identified an inhibitor (blebbistatin) of nonmuscle myosin II [5]. Blebbistatin was found to inhibit a variety of activities in whole vertebrate cells, including directed cell migration and cytokinesis. The use of blebbistatin and additional drugs (including monastrol) to manipulate the mitotic process in whole cells led to the discovery that ubiquitin-dependent proteolysis is required for exit from the cytokinetic phase of the cell cycle. Blebbistatin is thus

another example of how small molecules identified via HTS can be used to study the function of both individual proteins and major cellular processes.

The above points are summarized in Table 1.1.

Table 1.1 Attributes of classical genetic, chemical genetic, and RNAi approaches.

	Classical genetics	Chemical genetics	RNAi
Nature of perturbation	Permanent genetic change; true null is usually possible; heritable	Transient or irreversible, depending on situation; not heritable	Transient; generally not heritable
Directed perturbation	Possible in many systems, can be accomplished at the single-nucleotide level	Requires identification of a specific effector; may require synthetic chemistry effort to optimize	Yes; some occasional off-target effects
Conditional perturbation	Possible in some situations (temperature-sensitive alleles, inducible expression)	Possible, depending on specific situation	Yes, although difficult to control temporally
Target identification	Often simple, depending on situation	Often difficult	Target is known

1.3 Identifying Bioactive Molecules

Biologically active small molecules have often been discovered by testing a single compound at a time, but such an approach is obviously highly inefficient and cost-intensive in terms of both reagents and personnel time. As a result, HTS technologies have been developed to screen large numbers of compounds simultaneously, typically through the miniaturization and automation of assay protocols. What constitutes high throughput will vary depending on technical considerations and the screener's economic situation, but generally, the number of compounds involved may range from tens of thousands to several million. With current technologies, throughput is high enough that screening this number of compounds can be accomplished within several weeks, if not within several days.

A large number of screening technologies are available today for identifying bioactive small molecules; in fact, there are too many methods to be discussed adequately in this chapter. A very general discussion is provided here, but the reader is referred to other chapters in this volume for greater detail.

High-throughput screening assays can be divided into two main classes: "pure protein" and "cell-based". Pure protein screens generally have optical assay readouts that monitor enzymatic or binding activity. For instance, fluorescence polar-

ization or FRET techniques are commonly used to screen for compounds that affect binding between protein partners. In pure protein assays, every compound screened should have equal access to the target. However, the membrane permeability characteristics of any active molecules that are identified may subsequently pose a major concern if the target is intracellular.

Readouts for cell-based screens may rely on reporter gene systems (e.g. luciferase, β-lactamase), cell density, cell viability or cell morphology. Screen readouts can also be divided into two broad classes: uniform well readout acquired via plate readers and images acquired via automated microscopy. Screens involving image-based readouts are usually technically more difficult and their computational analysis more challenging than plate-reader screens. However, image-based readouts can provide a vast amount of information and microscopy screens are therefore often referred to as *high-content screens* [6–8]. The development of equipment and data analysis techniques for automated imaging screens is an area of active research, but application of early technologies has already yielded promising results [9–15].

1.4 Target Identification

When biologically active small molecules are identified through chemical screens, particularly forward genetic screens, a substantial amount of work is often required to identify the molecular target. Ignorance of the target does not preclude clinical or research use of the molecule; indeed, many clinical agents have been used effectively even when their targets were not known (e.g. aspirin, nitroglycerin, fumagillin, epoxomycin). Nevertheless, defining the molecular mechanism of action is vital to understanding the biological principles involved and also for potentially creating more potent or specific molecules through synthetic chemistry approaches.

Target identification is currently a major area of research in chemical genetics and Tochtrop and King have recently provided an excellent discussion of this topic [16]. This section will therefore emphasize HTS-related examples and also discuss some recent work not included in that review.

1.4.1 Hypothesis-driven Target Identification

For well-characterized cellular pathways, it is sometimes possible to deduce the target of a small molecule by comparing data from across the field. For instance, characteristic phenotypes induced by the small molecule may permit the assignment of the target to a previously identified complementation group. Subsequent hypothesis-driven testing of potential targets can then be undertaken.

This deductive approach was used successfully to define the target of monastrol, a small-molecule inhibitor of mitotic spindle bipolarity [3]. As noted above, mon-

astrol was identified in a forward chemical genetic screen for compounds that affect the structure of the mitotic spindle. Monastrol induces the formation of a monoastral microtubule array and, since previous studies had noted that antibody inhibition of the Eg5 kinesin also causes monoaster formation, it was postulated that Eg5 might be a target of monastrol. Subsequent work demonstrated that monastrol reversibly inhibits Eg5-driven microtubule motility *in vitro* but is not a general inhibitor of motor proteins.

Hypothesis-driven target identification was also used in a chemical screen for inhibitors of SARS coronavirus (SARS-CoV) replication [17]. In this study, a library of 50 240 compounds was screened by imaging for compounds that inhibit the cytopathic effect (CPE) of SARS-CoV towards Vero cells. A total of 104 compounds demonstrating effective inhibition of CPE and viral plaque formation were then tested *in vitro* against two protein targets known to affect SARS-CoV replication (M^{pro} protease and the NTPase/helicase), and also in a pseudotype virus assay for inhibition of S protein–ACE2-mediated entry of SARS-CoV into 293T cells. Two inhibitors of M^{pro}, seven inhibitors of Hel and 18 inhibitors of viral entry were identified and each of these three classes contained at least one inhibitor active in the low micromolar range. The authors subsequently assayed the 104 compounds against other common RNA viruses and found that most were specifically active against SARS-CoV, and approximately 3% were active against all viruses tested.

1.4.2 Affinity-based Target Identification

Affinity-based methods such as affinity labeling, affinity chromatography and crosslinking are also commonly used target identification strategies. Since these approaches often require synthetic modification of the molecule (e.g. addition of linkers or immunoreactive epitopes), care must be taken that the modifications do not interfere with the molecule–target interaction. Nonspecific interactions may also complicate the analysis and controls must be designed to address this issue. Despite such limitations, however, these strategies are of tremendous utility in target identification.

With affinity labeling, molecules may be radioactively or chemically labeled; they may be synthetically modified with reactive groups to promote covalent attachment to the target or tagged with specific moieties to facilitate detection. Standard protein fractionation and detection techniques can then be used to identify proteins from crude extracts that are specifically labeled by the molecule. Examples of drugs whose cellular targets were determined by such methods include acetylcholine, the anti-angiogenic agent fumagillin [18] and the antifungal lipopetide echinocandin [19].

Affinity purification of putative targets from cellular extracts is accomplished with molecules immobilized on a solid support used either as a column matrix or as a bead slurry. Small molecules for which cellular targets were identified through affinity purification include the immunosuppressant FK506 [20], the kinase inhibitor purvalanol [21] and the anti-inflammatory compound SB 203580 [22].

Additional studies involving affinity chromatography were instrumental in further characterizing the biochemistry of FK506 and purifying the target of the structurally related compound rapamycin [23].

A natural technological development of affinity-based target identification involves probing proteome chips with small molecules [24]. A recent application of this approach [25] is discussed in Section 1.4.4. Conversely, it is also possible to probe small-molecule microarrays with a known protein to identify specific interactions [26].

1.4.3 Genomic Methods of Target Identification

Yeast three-hybrid system The yeast two-hybrid system has been widely used to study protein–protein interactions and the approach was adapted to create a yeast three-hybrid system for identifying protein–ligand interactions [27]. The method relies on the use of a hybrid "bait" ligand consisting of the query molecule linked to a known ligand. A protein fusion between the known ligand's receptor and a DNA-binding domain serves as the "hook", while the "fish" is a protein fusion between a transactivation domain and the target protein. If the target protein interacts with the query molecule, the "fish" and "hook" will be brought together by the "bait" and transcription from a reporter gene is activated. Hence it should be possible to identify the target of the query molecule by cloning a library into the "fish" domain and screening or selecting for activity of the reporter.

The three-hybrid proof-of-principle study was conducted with a dexamethasone–FK506 hybrid ligand and a Jurkat cDNA library. This experiment successfully identified two variants of human FKBP12 as targets of FK506. Another proof-of-principle study employed a dexamethasone–methotrexate hybrid ligand to screen a mouse cDNA library and identified dihydrofolate reductase as a target of methotrexate [28].

A three-hybrid approach was recently used to identify targets of various kinase inhibitors [29]. For each inhibitor, a hybrid ligand was synthesized by attachment to methotrexate and the hook protein was a LexA–DHFR fusion. A three-hybrid screen of a human cDNA library with a purvalanol B–methotrexate ligand identified several (but not all) known purvalanol targets and also several new candidate targets, and all identified targets were kinases. Affinity chromatography and enzymatic assays confirmed 12 of 16 novel candidate targets identified in the cDNA screens.

Induced haploinsufficiency One genomic approach to target identification relies on the premise that altering the gene dosage of the target will affect sensitivity to the small molecule. For example, inactivating one copy of the target gene in a diploid organism would in many cases be expected to increase sensitivity. This method of identifying drug targets through *induced haploinsufficiency* was established by Giaever et al. [30], who constructed and screened a set of 233 heterozygous yeast deletion mutants for those demonstrating increased sensitivity to

known drugs. Each mutant was chromosomally tagged with a unique oligonucleotide and the mutants were pooled and grown in the presence of tunicamycin, at a level of drug that is sublethal for wild type. The relative number of each mutant in the pool was monitored at various time points by polymerase chain reaction (PCR) amplification and fluorescence labeling of all tags in the pool, followed by hybridization of the PCR-generated probes to an oligonucleotide microarray. The fluorescence intensity generated at each spot on the array permitted quantitation of the relative abundance of each corresponding heterozygote in the pool. Mutants unaffected by tunicamycin showed no reduction in signal over time, whereas tunicamycin-sensitive heterozygotes showed varying decreases in signal. In this study, one known and two new loci were identified and confirmed as involved in tunicamycin resistance.

A subsequent study extended this approach to test 78 compounds against a pool of 3503 yeast heterozygotes, representing over half the yeast genome [31]. Most of the compounds tested had known activities and many also had known targets. Of 20 compounds with known targets, in most cases this method correctly identified the target or members of the target complex. Targets were also identified for a number of compounds with previously unknown targets in yeast.

Recently, a complementary approach of using gene overexpression to identify small-molecule targets has been pioneered [32]. A plasmid-based yeast genomic library was introduced into yeast and 7296 individual transformants were arrayed in 384-well plates. The arrayed library was then replicated on to solid agar containing an inhibitor. The plasmid inserts from resistant strains were then recovered and sequenced. Pkc1 was identified via this method as a target for the inhibitor and this was subsequently confirmed by affinity chromatography and genetic and biochemical assays.

Expression profiling Expression profiling was recently employed in identifying the targets of a class of small-molecule antagonists of FK506 in yeast cells subjected to salt stress [33]. These molecules (termed SFKs for suppressors of FK506) were identified in a chemical genetic screen for molecules that rescue yeast growth in the presence of high salt and FK506 [34]. Expression profiling results obtained with SFK2-treated yeast suggested that the Ald6 p pathway was a target of SFK2 and haploinsufficency screening supported this hypothesis. Overexpression of ALD6 was found to suppress the effects of SFK2–SFK4 on growth and the ability of SFKs to inhibit Ald6 p *in vitro* was subsequently demonstrated.

In addition to using DNA microarray technologies in conjunction with haploinsufficiency studies, it is likely that gene expression profiling will be used increasingly as a primary means of identifying the targets of bioactive small molecules. Many research groups have used expression profiling to identify characteristic patterns of gene expression ("fingerprints") that are associated with certain disease states or biological pathways. If the patterns are sufficiently unique, then it is sometimes possible to assign previously uncharacterized mutants to specific cellular pathways or complementation groups based on their expression profiles. Analogously, by profiling cells grown with and without a small-molecule modu-

lator, it may be possible to identify genes or pathways that are affected by the small molecule.

The utility of expression profiling as a target identification method has been explored in a study involving 300 full-genome expression profiles in yeast generated with 276 deletion mutants, 11 regulatable alleles of essential genes and treatment with 13 compounds with known targets [35]. Of the 276 deletion mutants, 69 contained deletions of open reading frames of unknown function. To test the predictive value of the expression profiles, the profile associated with the well-defined ergosterol pathway was used to assign function to an uncharacterized ORF (YER044 c) and to identify the target of dyclonine. Subsequent biochemical and genetic experiments confirmed that both YER044 c and dyclonine affect sterol biosynthesis. The expression profiles associated with disruptions in cell wall function, mitochondrial function or protein synthesis were also used to assign function to several uncharacterized ORFs.

In combination with RNAi RNA interference (RNAi) refers to sequence-specific gene silencing triggered by the presence of double-stranded RNA within cells [36, 37]. This phenomenon has been recognized in multiple organisms, including plants, *Drosophila*, mice and humans, and researchers have rapidly employed RNAi as a tool for genome-wide studies in a variety of organisms. RNAi-based screens are logistically very similar to small-molecule screens and provide a powerful complement to both classical and chemical genetic methods.

Recently, parallel chemical genetic and RNAi screens were performed to identify cytokinesis inhibitors and their targets [38]. Over 51 000 small molecules and 19 470 dsRNAs (representing >90% of the annotated *Drosophila* genome) were screened by imaging in *Drosophila* Kc_{167} cells and assessed for the ability to increase the frequency of binucleate cells. This study resulted in the identification of 50 small-molecule cytokinesis inhibitors and 214 genes that are important for cytokinesis. Analysis of compounds and dsRNAs that induced similar phenotypes led to the finding that one of the small-molecule inhibitors (binucleine 2) affects the Aurora B pathway.

1.4.4 Proteomic Methods

A proteomic approach to target identification was recently reported by Huang et al. [25], who used a chemical genetic screen to isolate small-molecule enhancers and inhibitors of rapamycin's effect on yeast growth. Biotinylated inhibitor molecules were used to probe yeast proteome chips, followed by detection with fluorescently labeled streptavidin. This resulted in the identification of several putative target proteins, the role of which in rapamycin sensitivity was assessed using mutants bearing deletions in those genes. The deletion of one of these genes (YBR077C) resulted in rapamycin hypersensitivity. Deletion of another gene (Tep1 p) had no

effect on rapamycin sensitivity, but deleting the Apl2 p gene, which encodes a previously identified interaction partner of Tep1 p, did result in rapamycin hypersensitivity.

A proteomic method based on two-dimensional gel electrophoresis was recently used to identify the target of the synthetic bengamide analog LAF389 [39]. Cells were grown in the presence and absence of a natural bengamide, bengamide E, and proteins isolated from both populations were separated on 2D gels. One protein, a 14-3-3 isoform, was noted to display a bengamide-dependent change in charge. A similar effect was observed with LAF389 and the affected isoform was identified as 14-3-3 γ. Analysis of the altered 14-3-3 γ and additional studies established that LAF389 directly inhibits methionine aminopeptidases, resulting in retention of the initiator methionine in 14-3-3 γ.

1.5
Discovery for Basic Research Versus Pharmacotherapy Goals

The required characteristics of a small-molecule research tool are much less stringent than those for a lead molecule destined for clinical development. First, since a research tool will be used *in vitro* or in animal models, safety and regulatory issues are limited. ADMET constraints are not an issue for *in vitro* research and they are obviously more easily met for experimental animals than human patients, in whom adverse effects are tolerated only under extraordinary circumstances. This "experiment versus patient" issue also affects specificity and efficacy requirements, as well as the motivation to undertake subsequent structure–activity relationship (SAR) studies. For a research tool, the parameters that define adequate specificity and efficacy considerations will vary depending on whether the molecule is to be used in enzymatic, cell-based or animal studies and on the experiment's time course and sensitivity of detection. Minor or short-lived effects can be extremely useful in the basic research setting, but they will generally not suffice for a therapeutic agent and substantial SAR studies may be needed to optimize molecules intended as therapeutic leads. Molecules used as research tools often do not undergo an intensive SAR effort, as increased specificity or efficacy may not be needed.

Molecules that have been discovered through primary screens and used for basic research typically have IC_{50} values ranging from the low micromolar to approximately 50 μM and the upper end of this range likely reflects the compound concentrations employed during primary screens. Two such molecules that have been used in multiple basic research studies without further structural modification are monastrol and blebbistatin, which have IC_{50} values of 14 and 2 μM, respectively [3, 5]. In contrast, nanomolar IC_{50}s are considered desirable for therapeutic agents.

1.6
Chemical Genetic Screens in the Academic Setting

During the last decade, HTS methods have been actively pursued by the biotechnology and pharmaceutical industries to accelerate the identification of lead compounds for drug discovery. Owing to the financial resources needed to support even a modest HTS effort, these methods were largely unavailable to academic investigators. Therefore, to promote chemical genetics as an academic discipline, the Institute of Chemistry and Cell Biology/Initiative for Chemical Genetics (ICCB/ICG) at Harvard Medical School established one of the first academic HTS facilities, using an organizational model that relies on a single screening facility used by multiple investigators. This model (termed Investigator-initiated Screening) has been highly successful, facilitating screening projects for more than 80 different research groups from throughout the USA and abroad. In contrast to industry efforts towards drug discovery, the interest of the ICCB/ICG was primarily the identification of bioactive molecules for use as research tools. It should noted, however, that bioactive molecules discovered in academic screens may also show promise as pharmaceutical leads and that "high-risk" lead discovery (e.g. for targets that are not yet validated or that are considered economically nonviable for the private sector) may be better suited to academia than industry.

Under the Investigator-initiated Screening program, individual researchers propose and carry out the majority of the work for their own screening projects, including assay development, reagent dispensing, assay readout and subsequent data analysis. The screening facility staff assist by maintaining and providing access to compound collections and screening robots, training screeners for independent operation of some machines and providing some informatics and chemistry support. Data from all screens conducted at the ICCB/ICG were deposited in a nonpublic database and comparison of information across screens was used to help screeners eliminate uninteresting screening positives from further analysis. For example, fluorescent compounds that score as positive in fluorescence-based screens are most likely irrelevant to the actual screen target and compounds that score as positive in multiple cell-based screens may have nonspecific effects.

Most academic screens have been performed in 384-well format. Through spring 2005, the ICCB/ICG screening facility performed an average of 12–15 screening sessions per week, screening 14 080 compound wells (20 plates in duplicate) in a typical session. A typical investigator-initiated screening project screened 50 000–100 000 compound wells in duplicate. At this capacity, the facility could comfortably initiate 40 new screens each year.

Although academic investigators share some goals with their industry counterparts, there are certain key differences. For example, most industrial screening programs target a relatively small number of disease-relevant pathways and proteins, whereas academic investigators may wish to use chemical genetics to study a wide variety of biological pathways in diverse organisms. This scientific heterogeneity among academic groups is somewhat unpredictable and tends to result in a

greater variety of assay protocols than typically encountered in an industry setting. For this reason, the ICCB/ICG screening facility was specifically designed to have the flexibility to accommodate many different types of assays, generally by employing screening instruments that can be used for as wide a variety of assay types as possible and that can work either in stand-alone mode or in custom configurations with other pieces of equipment for automation of sequential steps in a screening protocol.

Compound selection is also affected by the diverse nature of academic screens. Computational methods are employed in both academia and industry during the compound selection process to identify compounds that are "drug-like" or "lead-like" and to eliminate molecules that are unstable, toxic or otherwise unsuitable for screening. However, since academic screens may address any area of biology, the ICCB/ICG also chose to acquire compound libraries of maximal diversity and to avoid "targeted libraries" that are predicted to contain a relatively high proportion of compounds likely to act on specific pathways (e.g. kinase-targeted libraries). The ICCB/ICG libraries also contained compounds donated by chemists in exchange for data acquired by screening those compounds, thus extending the collaborative model to the wider community of synthetic chemists.

Finally, the many different types of primary screen data (including both numerical and imaging data) acquired under this organizational model present a particular challenge for data analysis and archiving purposes. Each dataset will have a unique definition of screening positives and may need highly individualized analysis techniques. Any cross-screen data comparison in this situation also requires a fairly detailed understanding of each screen being considered.

1.7 Conclusions

The use of HTS and small molecules to study cellular processes has recently begun to gain momentum. The substantial costs and logistic issues associated with establishing even a modest screening effort has deterred most academic researchers (and potentially also small biotech firms) from pursuing HTS, but it is hoped that more and more investigators will soon be able to avail themselves of this exciting new technology. Collaborative efforts such as the ICCB/ICG and the multiple screening centers currently being established by the National Institutes of Health will be vital in this respect and commercial efforts could play also a large role in putting HTS within the reach of smaller groups. A similar "popularization" of DNA microarray technologies was achieved over the last decade using both private and public sector resources and microarray experiments are now common in both academic and industry laboratories. Chemical genetics and HTS admittedly require a higher level of investment, but the potential gains in both basic science and clinical medicine are immense.

References

1 T.J. Mitchison, *Chem. Biol.* **1994**, *1*, 3–6.
2 S.L. Schreiber, *Bioorg. Med. Chem.* **1998**, 6, 1127–1152.
3 T.U. Mayer, et al., *Science* **1999**, *286*, 971–974.
4 A. Cheung, et al., *Nat. Cell. Biol.* **2002**, *4*, 83–88.
5 A. F. Straight, et al., *Science* **2003**, *299*, 1743–1747.
6 V.C. Abraham, D. L. Taylor, J. R. Haskins, *Trends. Biotechnol.* **2004**, *22*, 15–22.
7 R.A. Blake, *Curr. Opin. Pharmacol.* **2001**, *1*, 533–539.
8 J.C. Yarrow, et al., *Comb. Chem. High. Throughput Screen.* **2003**, *6*, 279–286.
9 Y. Feng, et al., *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6469–6474.
10 T.J. Mitchison, *ChemBiochem* **2005**, *6*, 33–39.
11 T.J. Nieland, et al., *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15422–15427.
12 Z.E. Perlman, T. J. Mitchison, T. U. Mayer, *ChemBiochem.* **2005**, 6, 218.
13 Z.E. Perlman, et al., *Science* **2004**, *306*, 1194–1198.
14 N. Venkatesh, et al., *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8969–8974.
15 J.C. Yarrow, et al., *Chem. Biol.* **2005**, *12*, 385–395.
16 G.P. Tochtrop, R. W. King, *Comb. Chem. High. Throughput Screen.* **2004**, *7*, 677–688.
17 R.Y. Kao, et al., *Chem. Biol.* **2004**, *11*, 1293–1299.
18 N. Sin, et al., *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6099–6103.
19 J.A. Radding, S. A. Heidler, W. W. Turner, *Antimicrob. Agents Chemother.* **1998**, *42*, 1187–1194.
20 M.W. Harding, et al., *Nature* **1989**, *341*, 758–760.
21 M. Knockaert, et al., *Chem. Biol.* **2000**, *7*, 411–22.
22 K. Godl, et al., *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15434–15439.
23 E.J. Brown, et al., *Nature* **1994**, *369*, 756–758.
24 G. MacBeath, S. L. Schreiber, *Science* **2000**, *289*, 1760–1763.
25 J. Huang, et al., *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16594–16599.
26 A.N. Koehler, A. F. Shamji, S. L. Schreiber, *J. Am. Chem. Soc.* **2003**, *125*, 8420–8421.
27 E.J. Licitra, J. O. Liu, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12817–12821.
28 D.C. Henthorn, A. A. Jaxa-Chamiec, E. Meldrum, *Biochem. Pharmacol.* **2002**, *63*, 1619–1628.
29 F. Becker, et al., *Chem. Biol.* **2004**, *11*, 211–223.
30 G. Giaever, et al., *Nat. Genet.* **1999**, *21*, 278–283.
31 P.Y. Lum, et al., *Cell* **2004**, *116*, 121–137.
32 H. Luesch, et al., *Chem. Biol.* **2005**, *12*, 55–63.
33 R.A. Butcher and S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7868–7873.
34 R.A. Butcher, S. L. Schreiber, *Chem. Biol.* **2003**, *10*, 521–531.
35 T.R. Hughes, et al., *Cell* **2000**, *102*, 109–126.
36 G.J. Hannon, J. J. Rossi, *Nature* **2004**, *431*, 371–378.
37 C.C. Mello, D. Conte Jr., *Nature* **2004**, *431*, 338–342.
38 U.S. Eggert, et al., *Public Library of Science Biology* **2004**, *2*, e379.
39 H. Towbin, et al., *J. Biol. Chem.* **2003**, *278*, 52964–52971.