Small Molecules, Big Players: the National Cancer Institute's Initiative for Chemical Genetics

Nicola Tolliday,[1] Paul A. Clemons,[1] Paul Ferraiolo,[1] Angela N. Koehler,[1] Timothy A. Lewis,[1] Xiaohua Li,[1] Stuart L. Schreiber,[1,2] Daniela S. Gerhard,[3] and Scott Eliasof[1]

[1]Broad Institute of Harvard and MIT; [2]Howard Hughes Medical Institute, Cambridge, Massachusetts; and [3]Office of Cancer Genomics, National Cancer Institute, Bethesda, Maryland

Abstract

In 2002, the National Cancer Institute created the Initiative for Chemical Genetics (ICG), to enable public research using small molecules to accelerate the discovery of cancer-relevant small-molecule probes. The ICG is a public-access research facility consisting of a tightly integrated team of synthetic and analytical chemists, assay developers, high-throughput screening and automation engineers, computational scientists, and software developers. The ICG seeks to facilitate the cross-fertilization of synthetic chemistry and cancer biology by creating a research environment in which new scientific collaborations are possible. To date, the ICG has interacted with 76 biology laboratories from 39 institutions and more than a dozen organic synthetic chemistry laboratories around the country and in Canada. All chemistry and screening data are deposited into the ChemBank web site (http://chembank.broad.harvard.edu/) and are available to the entire research community within a year of generation. ChemBank is both a data repository and a data analysis environment, facilitating the exploration of chemical and biological information across many different assays and small molecules. This report outlines how the ICG functions, how researchers can take advantage of its screening, chemistry and informatic capabilities, and provides a brief summary of some of the many important research findings. (Cancer Res 2006; 66(18): 8935-42)

Small Molecules and Chemical Genetics

Many of life's critical processes are carried out by small molecules and so they can arguably be considered another essential component of the "central dogma" of molecular biology (1). Biologically relevant small molecules may be the product of evolution or they may be synthesized by organic chemists who seek to develop drugs or chemical probes for the study of dynamic processes. Such probes have been critical for characterizing and dissecting many of the cellular circuits and signaling pathways known today (e.g., the mammalian target of rapamycin and Ras pathways; refs. 2, 3). However, new small molecules are needed to function as probes to further dissect the known pathways, to discover novel pathways and interactions, to validate potential drug targets, and to function as drugs to treat unmet medical needs, including more targeted therapies.

The use of small molecules to uncover molecular mechanisms, a field of inquiry termed chemical genetics (4), is complementary to techniques developed by the molecular genetics community (e.g., gene knockouts and RNA interference). Small molecules can offer excellent temporal control, turning processes on and off with the addition or removal of a compound, and even spatial control can be accomplished on a cellular level with caged compounds or in animal studies following local administration. Small molecules can also alter protein translation and transcription in more subtle ways than simply turning protein expression up or down, such as modulating one of its several functions. For example, the histone deacetylase, HDAC6, possesses two distinct and active catalytic domains, but only tubulin deacetylation activity is selectively inhibited by the small molecule tubacin (5). Genetic knockdown experiments would not be capable of decoupling these two activities.

The special properties of the small molecules described above are important in drug development. Once a small molecule is discovered that selectively and potently alters a target with potential therapeutic importance, it can be modified by medicinal chemists to overcome the pharmacodynamic, pharmacokinetic, toxicity, and efficacy hurdles necessary for bringing a drug to market. The pharmaceutical industry has successfully exploited the ability of small molecules to serve as drugs by focusing on a relatively small number of biological targets, a subset of the so-called "druggable genome" (6). High-throughput screening (HTS), a mainstay of the drug industry, is well suited for this process because it is designed to handle hundreds of thousands of biological measurements per day.

Chemical genetics has also used HTS, but with the related aim of investigating cellular pathways and identifying key proteins underlying cellular processes using small molecules as the sources of perturbation (7). HTS automation is exploited to systematize chemical genetic screens across a wide swath of both chemical and biological space. Often, complex measurements of cellular phenotypes are used without *a priori* knowledge of the relevant targets (referred to as phenotypic screening). The biological activities of both "druggable" and "nondruggable" targets are relevant for chemical genetic studies, although the latter offer new challenges for assay development and new opportunities for the identification of potent and selective modulators. Therefore, this approach requires new chemistry that does not necessarily adhere to the traditional prescripts for what is "drug-like" and that has not already been explored in-depth by the pharmaceutical industry.

The History and Mission of the Initiative for Chemical Genetics

In 2002, the National Cancer Institute (NCI) created the Initiative for Chemical Genetics (ICG) out of its commitment to improve the pathway to cancer drug discovery. The ICG's mission is to enable academic research using small molecules to explore biology with

Requests for reprints: Scott Eliasof, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA 02142. Phone: 617-324-9603; Fax: 617-324-9601; E-mail: seliasof@broad.harvard.edu.

doi:10.1158/0008-5472.CAN-06-2552

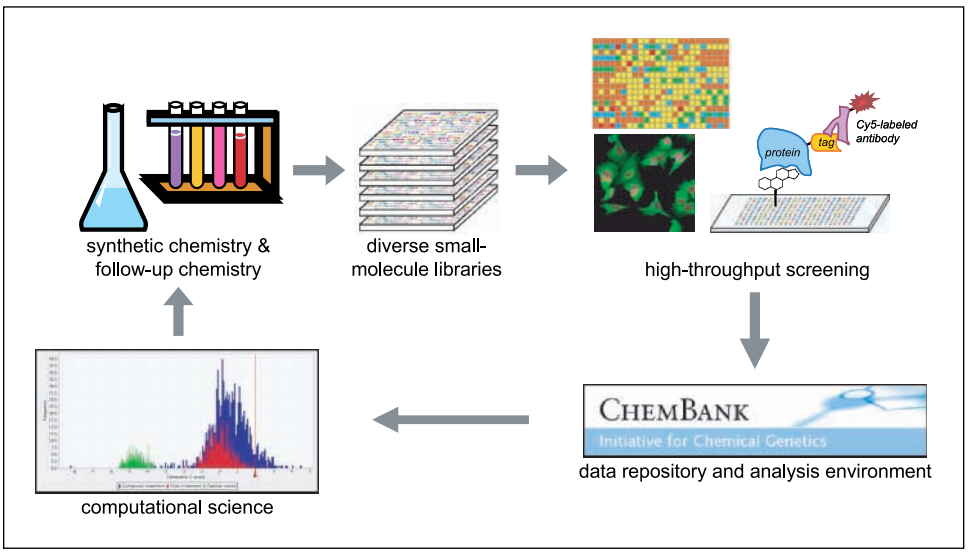

Figure 1. Flow of information at the ICG. Novel molecules are synthesized and screened, screening and chemistry data are imported into ChemBank, and data analysis guides further chemistry and biological experimentation.

an emphasis on cancer biology. The ICG is located at the Broad Institute of Harvard and Massachusetts Institute of Technology in Cambridge, MA and includes teams of synthetic and analytical chemists, assay developers, screening and automation engineers, computational scientists, and software developers. The activities of the ICG are tightly integrated into a single, cohesive unit (Fig. 1).

In creating the ICG, the NCI was able to leverage existing chemical genetics infrastructures, including the Harvard Institute of Chemistry and Cell Biology, founded by one of us (S.L.S.) and Timothy Mitchison in 1997. The ICG fosters wide-reaching collaborations between chemists, biologists, and computational scientists. The ICG Screening Facility is available to the entire academic research community, providing scientists with access to screening automation and novel chemistry. Chemists anywhere who supply compounds can observe the annotation of their biological activity via the World Wide Web.

More recently, the ICG has increasingly focused on supporting the basic cancer research developed at the NCI Cancer Centers. The ICG is prioritizing screens coming from cancer center scientists and, when possible, directly assisting cancer scientists at the assay development stage. This mission-oriented focus of the ICG is one important way in which it differs from the NIH Molecular Libraries and Imaging Initiative (MLII).[4] The MLII, following the model of the ICG, is facilitating HTS, follow-up chemistry, and data sharing (via the publicly accessible database, PubChem)[5] for the NIH-funded community, with the aim of discovering new chemical probes. Although both the ICG and the MLII will use many common technologies, the ICG intends to do more than make probes. As highlighted below, the ICG is seeking to maximize its effect on the discovery of novel cancer therapeutics.

Collaborating with the ICG is different from collaborating with the pharmaceutical industry. First, as described below, all data generated at the ICG are available to the collaborating cancer scientists, thereby enabling them to compare their screening data with all other assays done at the ICG. Access to all screening data ensures that promiscuous screening "hits" can be eliminated and allows for the integration of multiple data sources that can yield new insights into how disparate biological pathways interconnect. In contrast, the pharmaceutical industry appropriately keeps their screening data guarded and can generally only compare data across screens that are done within a single company. Second, as described below, the ICG is an academic research center, developing important innovations in the areas of synthetic chemistry, screening technologies as well as computational methodologies, integrating the new technologies together, and publishing the findings.

Capabilities and Technology Development at the ICG

Advances made by ICG scientists in the areas of synthetic chemistry, screening, and informatics have greatly enriched the screening experience at the ICG.

Available chemistry. The ICG small molecule screening collection has benefited from a close relationship with the Harvard University Department of Chemistry and Chemical Biology and the Broad Institute of Harvard and MIT, where research on diversity-oriented synthesis (DOS) has been done. DOS chemistry applies diversity-based reaction planning and combinatorial chemistry to produce libraries of molecules with stereochemically and structurally diverse skeletons with high-appending potential that populate chemical space in an unbiased way (8). Importantly, these DOS libraries consist of complex, three-dimensional molecules that feature multiple stereocenters, resembling natural products (Fig. 2). The diversity of molecular shape that is possible to obtain with DOS compounds and natural products is necessary for maximizing the range of biological targets that can be modulated by small molecules (9).

Compound libraries available for screening within the ICG also contain commercially available simple, flat small molecules that follow the Lipinski Rule of Five, similar to those screened by the pharmaceutical industry (10). Although the strict Lipinski Rule of Five filter may increase the probability of discovering orally active drugs, it is too limiting for finding the widest variety of chemical probes. Additionally, exclusive focus on Lipinski Rules may also be limiting in the quest for new therapeutic opportunities.

[4] http://nihroadmap.nih.gov/molecularlibraries/.
[5] http://pubchem.ncbi.nlm.nih.gov/.

Natural products, which can be quite complex, generally do not fall within the Lipinski Rule of Five. However, >25% of drugs approved by the Food and Drug Administration between 1981 and 2002 were natural products or natural product–derived. Furthermore, in the areas of cancer and infectious disease, >60% and 75% of drugs are of natural origin, respectively (11). The ICG has access to natural product collections through a collaboration with Jon Clardy at the Harvard Medical School. These collections comprise prefractionated extracts from cyanobacteria, marine organisms, and lichens (collected with support from the NCI National Cooperative Drug Discovery Group) and these are of high interest in cancer-relevant screening projects.

In addition, the ICG is actively screening small molecules synthesized by synthetic chemistry laboratories from around North America, including: McGill University, University of Alberta, University of California Los Angeles, University of California Irvine, Utah State, Northwestern University, Harvard University, and three National Institute of General Medical Sciences Centers of Excellence in Chemical Methodologies and Library Development, located at the University of Pittsburgh, Boston University, and the Broad Institute, respectively. Like the DOS molecules, these small molecule collections are the products of unique synthetic approaches and are not available commercially.

Screening technologies. The screening capabilities of the ICG are described in Table 1. Briefly, three types of screening technologies are available: plate readers and automated microscopes (high-content imaging systems) for assays done with small molecules in solution and protein-binding assays using small-molecule microarrays, in which small molecules are immobilized at high density on a solid surface (12).

To date, the majority of screens done within the ICG are phenotypic assays, typically done on cells or whole model organisms, in which the response of a small molecule is the result of a modification to an entire biological system. Most screens use plate readers because these screens are simple and fast. Plate readers can

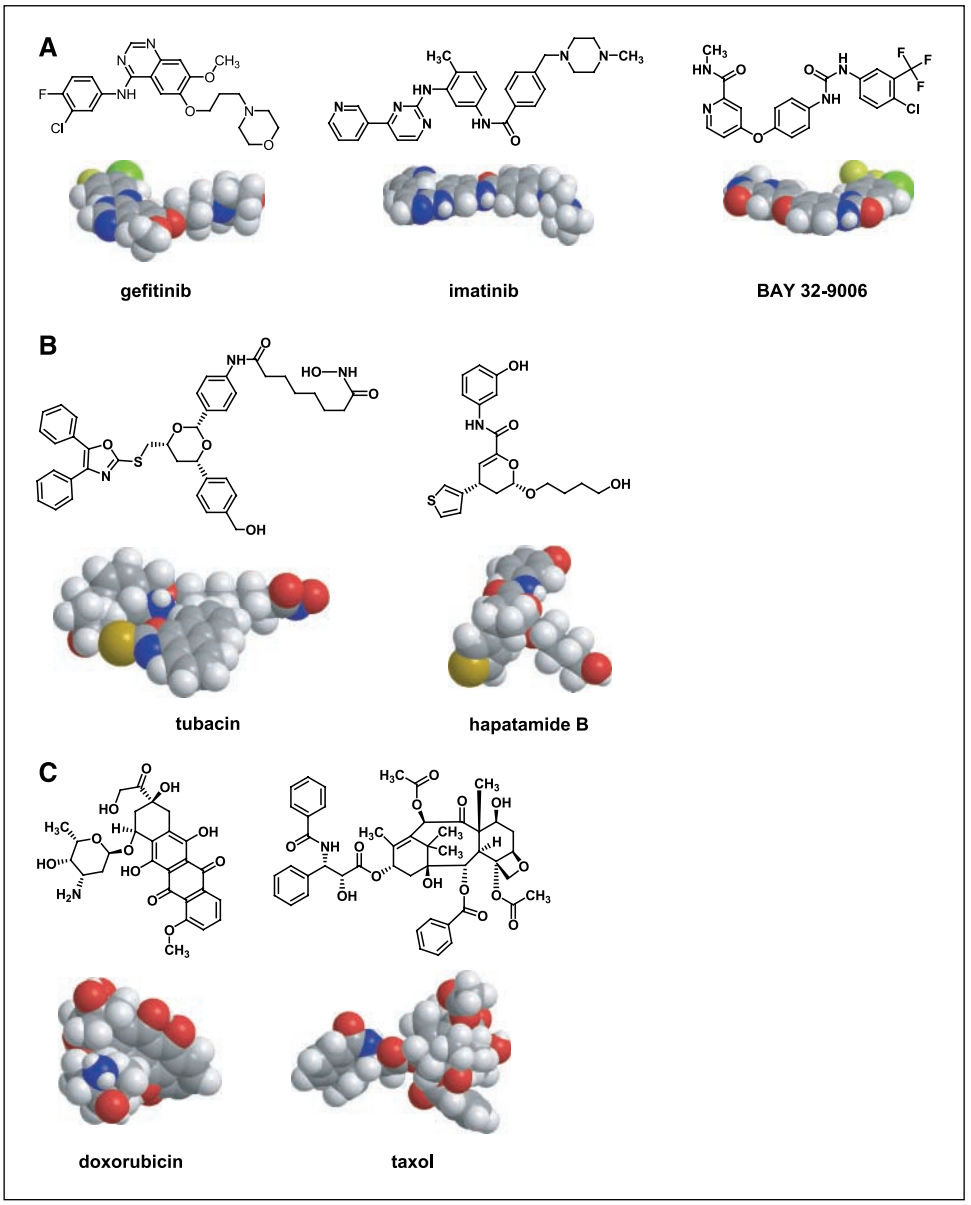

Figure 2. The three classes of small molecules available for screening at the ICG. Molecules synthesized by DOS (*B*) are more three-dimensional in shape than compounds typically synthesized by pharmaceutical companies (*A*), and more closely resemble more structurally complex and synthetically difficult natural products (*C*).

Table 1. Screening modalities available at the ICG

- Plate reader assays
  - Fluorescence
    - Fluorescence intensity assay
    - Fluorescence polarization assay
    - Fluorescence resonance energy transfer assay
  - Absorbance
    - Optical density assay
    - Colorimetric assay
  - Luminescence
    - Chemiluminescent cytoblot assay
    - Chemiluminescent ELISA assay
    - Cell viability assay
- High-content screening
  - Nuclear foci formation assay
  - Cell morphology assay
  - Protein translocation assay
- Small molecule microarrays
  - Purified protein binding assay
  - Cell lysate binding assay

handle multiple detection formats (luminescence, absorbance, and fluorescence) but are limited because the signal is averaged over a well of a microtiter plate. In contrast, high-content screening assays are significantly slower but can produce detailed measurements at the level of an individual cell or subcellular compartment. For example, image analysis software can be trained to monitor the translocation of a fluorescently tagged protein from the cell nucleus to the cytoplasm, or to detect changes in cell morphology. Because such measurements are done at the cellular level rather than averaged over a well, the signal-to-noise ratio is considerably higher; each cell, in essence, serves as its own set of data points.

The ICG also uses small-molecule microarrays. Up to 10,000 small molecules can be covalently attached to a functionalized glass microscope slide through an optimized-length polyethylene glycol linker (12). Recently, the ICG adapted this technology to attach a wide variety of molecules, including Food and Drug Administration–approved drugs, natural products, and DOS molecules (13). The "printed" slides are incubated with purified protein, washed, and the bound purified protein is visualized with a fluor-labeled antibody (typically directed to an epitope tag on the protein or to the protein itself) using a standard microarray scanner. In addition, proteins residing in clarified cellular lysates can also be detected using this technology. Binding strength and kinetic variables are determined using surface plasmon resonance secondary assays, using a Biacore detector (Biacore International AB, Uppsala, Sweden).

Computational tools. Central to the ICG's activities is the community-based resource and research environment, ChemBank.[6] ChemBank is a web-based database and application that currently stores information on both small molecules that are screened by the ICG and biological assays that have been done at the ICG. These data are available to the scientific community through an easy-to-access, publicly accessible web interface (see below). In addition, the small molecules screened at the ICG are available from PubChem, with links back to the appropriate molecule and assay data in ChemBank. However, ChemBank is more than a simple data repository; it offers complex search and visualization tools designed to correlate and navigate chemical and biological space.

The creation of ChemBank is the result of a close interaction between a wide variety of experts including synthetic and analytical chemists, computational scientists, biologists, biotechnicians, and software engineers. The goal was to create a resource where chemists and biologist can cross-fertilize their thinking by providing mutual feedback on the relationship between chemical space and biological space. ChemBank makes it possible for chemists to explore such issues as what characteristics make a compound more or less likely to have biological activity and whether it is possible to identify classes of chemical structures that are unusually effective in performing challenging tasks like preventing or promoting protein-protein interactions. Biologists can learn more about targets or signaling pathways that play a key role in a phenotypic screen, for example, by comparing the behavior of known, biologically active ("bioactive") small molecules with the behavior of novel small molecules across multiple assays. In this way, ChemBank can be used to predict the identity of an unknown target of a small molecule (14).

In order to satisfy the conflicting needs of enabling scientific research and ensuring that experiments are properly developed by the originating scientist free of immediate competition from the entire research community, we have instituted a time delay of 1 year between screening data collection and public data dissemination. This delay ensures that the primary screening data are confirmed in subsequent experiments, so that the public data are quality controlled. All chemists and biologists who participate in the ICG are free to examine each other's data because all are required to sign a data sharing agreement (DSA). The DSA creates a community of hundreds of collaborating scientists by establishing the ground rules for interacting with each other and with ChemBank. The DSA ensures that deposition of data into ChemBank does not constitute "publication" for patent or future publication purposes and that no precommitments of intellectual property exist for any small molecules or screening projects. More importantly, the DSA explicitly reminds contributing scientists that their data are shared with the other contributing scientists and that they may not publish each other's findings. For example, scientists may wish to perform cross-sectional analyses, searching for the performance of a screener's hits in other assays. In this way, all contributing scientists have access to and may learn from the very latest data, without having to wait for the 1 year data delay. Thus, there are two versions of ChemBank: the public version, currently containing 36 screening projects (441 distinct assays) that have surpassed the delayed release, and the version available only to those who participate and sign the DSA, currently containing 137 screening projects (1,307 distinct assays). As time progresses, screens will move from the restricted version to the public version. A current copy of the DSA can be found at http://www.broad.harvard.edu/chembio/sci/screen/facil/DataSharingAgreement.pdf, and we welcome its use as a model for similar projects.

How the Cancer Research Community Can Engage with the ICG

A scientist wishing to benefit from the ICG begins by contacting the ICG to discuss their ideas confidentially for a high-throughput screen. In most cases, the ICG operates in a "staff-assisted" fashion: the screening facility personnel assist efforts by operating and

[6] http://chembank.broad.harvard.edu/.

maintaining the screening robots and the collaborating scientists perform their own assays. This ensures that the assay is done optimally because the investigators are the experts in their particular field of biology. In special cases, the ICG can help with assay development; however, only limited resources are available for this support. Once it has been established that an assay is sufficiently developed for a high-throughput screen, the investigator completes the screening application found on the ICG web site[7] and agrees to the DSA described above. The application is reviewed by a screening review committee, an impartial committee of ICG scientists that is tasked with making determinations regarding the "readiness" of a particular screening project (to ensure optimal use of the screening facility), and the potential for overlap with existing screening projects (to avoid multiple laboratories wishing to perform an identical screen). All applications are kept strictly confidential. The committee also provides advice on screening projects, including suggestions for improvement of assay design (i.e., methodology) and experimental design (i.e., experimental conditions). Typically, a screen is done in 384-well microtiter plates, in duplicate, using a collection of 25,000 to 50,000 small molecules as described above. A screening campaign is usually done in ~2 weeks. Investigators are not charged a fee for access to the facility or for screening of the libraries, but pay the consumable costs (typically 3-7¢ per well).

The output file is both imported into ChemBank and given to the screener, together with molecular structures and can be subsequently analyzed in ChemBank or other data analysis environments, such as Spotfire (Spotfire, Inc., Somerville, MA). The data are processed using a standard error model, i.e., a set of analytic techniques designed to separate maximally signal from noise.[8] Importantly, these techniques also standardize the data to render it neutral with respect to the screening technology used to produce it, thereby enabling comparisons of assay data across all screening modalities.

Although a well-designed and implemented high-throughput screen is important to ensure high-quality assay data, the most essential determinant of a successful screen is whether or not there are well-designed plans for follow-up experimentation and a commitment to carry these through. The experiments that need to be carried out include: (*a*) a retest of the compounds to eliminate false-positives, (*b*) varying compound concentration to prioritize screening hits and begin structure-activity relationship (SAR) studies, (*c*) assays designed to determine specificity, (*d*) assays that are increasingly relevant for cancer therapeutic development, such as animal models, and (*e*) determining the small molecule target, if the original assay was a phenotypic screen. The ICG can assist in this process with analysis of the screening hits, identification of related molecules that are commercially available (so-called "SAR by purchase") and follow-up chemistry if the screeners request it and there is an adequate staff to carry out the tasks.

Scientific Direction of the ICG

Although the ICG performs an important service to the cancer community, it is not simply a "service" or "core" facility. ICG scientists innovate to improve the screening experience and advance the field of chemical genetics and its effect on cancer therapeutics. For example, one major focus of the ICG is multidimensional screening. HTS is traditionally done as a one-dimensional experiment, where different small molecules are added to assay wells containing a single cell type in a single state, and a single type of assay measurement is made at a single time point using a single dose of compound. However, it is becoming increasingly useful to view screening experiments as multidimensional. Cell-based assays, for example, can be viewed as experiments that can be done along multiple dimensions. The effects of varying small molecule structure, the duration of time that a cell is exposed to a small molecule, the concentration of the small molecule being applied, the cellular properties being measured (e.g., cell viability or proliferation), and the properties of the cell being studied (e.g., a cell's origin or tumorigenic state) can all be assessed and simultaneously considered. These multidimensional screens can provide insights not accessible via their simpler counterparts. For example, through the NCI's implementation of the "NCI-60" screening system, a matrix of data is available on the Internet that comprises 60 distinct cell lines exposed to a large collection of small molecules (15). As a result, signatures of small molecules derived from these screens have been used to infer the protein targets of small molecules by matching response profiles to compounds of known biological activity (16). The ICG is developing analytic techniques to view and explore such multidimensional screening data and will make these tools available through ChemBank once they are developed.

Another major undertaking at the ICG is the systematic exploration of both chemistry and biology space. Through ChemBank, the ICG is creating a common, open-source research environment that will enable several new collaborations between chemists, biologists, and computational scientists. Because data from multiple assay technologies are processed using a standardized error model, ChemBank enables scientists to benefit maximally from the near-infinite number of possible correlations of the data. Simply put, we believe that the whole will be greater than the sum of its parts. Ultimately, the goal of ChemBank is to: (*a*) guide chemists synthesizing novel compounds or libraries, (*b*) assist biologists searching for small molecules that perturb specific biological pathways, (*c*) develop new computational methodologies for exploring chemical genetic data, and (*d*) catalyze the process by which drug hunters discover new and effective medicines. The ICG is also working closely with the NCI's cancer Biomedical Informatics Grid (caBIG)[9] with the ultimate goal of linking ChemBank data with cancer-relevant research data and tools being developed elsewhere. This effort will enable scientists to connect ICG's data with >800 cancer scientists from >80 organizations, who are working collaboratively on >70 different projects.

The use of a common data repository and analysis environment also enables the ability to perform meta-analyses across many millions of data points. For example, the ICG's Forma Project ("forma" is Latin for shape or form), seeks to examine the effect of molecular shape on biological outcome. Small molecules from a variety of sources (DOS compounds, commercially available small molecules, and natural products) will be screened using a variety of assay technologies (plate readers, automated microscopes, and small-molecule microarrays), using cell-based phenotypic assays, protein binding assays, and biochemical assays. Importantly, the ICG will analyze small molecules that are not selected on the basis of known biological activity or structural biasing elements, so that

[7] http://www.broad.harvard.edu/chembio/sci/screen/facil/ScreenApplication.pdf.
[8] Manuscript in preparation.
[9] https://cabig.nci.nih.gov

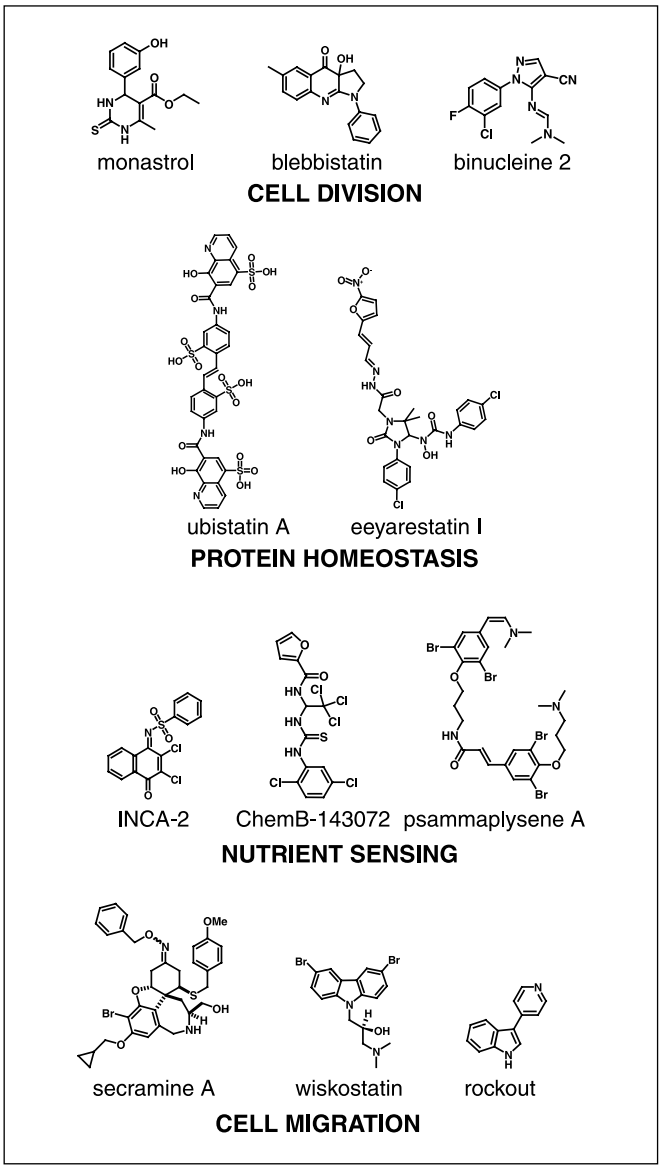

Figure 3. Cancer-relevant small molecules discovered at the ICG.

the study is unbiased. The expectations are that shape will be an important determinant of biological activity of small molecules, for example, when comparing enzymes, which tend to have binding pockets, to other proteins that possess more flat surfaces.

ICG Achievements

The success of the ICG can be measured in a number of ways, but its effect on science, and on cancer in particular, is the ultimate ruler. To date, 76 different research groups representing 39 different institutions have participated in >130 screening projects (~1,300 assays) at the ICG and more than a dozen organic synthesis laboratories have provided compounds. Results from most of the assays are currently in ChemBank. This work has resulted in 62 primary publications in the past 4 years,[10] describing 229 skeletally distinct chemical probes (simple R-group substitutions are not considered distinct skeletons). These screening efforts have led to the discovery of novel cellular processes thought to play key roles in cancer and novel modulators of existing pathways (Fig. 3). Many of these processes are active areas of study in the development of new cancer therapeutics. Specifically, small molecules have been found that modulate angiogenesis, metastasis, protein homeostasis, cell cycle, growth factor pathways, cell survival, and nutrient growth signaling. Although most of the small molecules shown in Fig. 3 will not become drugs, they will likely lead to new insights into the clinical development of new cancer therapeutics and pave the way for future drugs.

For example, among those chemotherapeutic agents that work by inhibiting mitosis, many are effective anticancer agents because they interfere with microtubule function responsible for cell proliferation (17). However, microtubule disruption affects many important cellular processes including intracellular transport, maintenance of cell shape, cell signaling, and cell motility; therefore, the current antimitotic drugs suffer from serious side effects, including peripheral neuropathy (18). Work at the ICG has uncovered several new small molecules that disrupt additional key steps in cell division. The small molecules monastrol, blebbistatin, and binucleine 2 were discovered using cell-based phenotypic high-content screens (19), screens for non–muscle myosin II ATPase activity (20) and parallel RNAi and chemical genetic screens (21), respectively. As novel tools and lead compounds for biological study and pharmaceutical development, these compounds have the potential to evolve into chemotherapeutic agents with more tolerable side effects.

Tumor cells often exhibit excess protein production, and thus, small molecules that affect protein homeostasis are another major area of research that has led to new cancer therapeutics. For example, 17-allylamino-17-demethoxygeldanamycin, an inhibitor of the chaperone protein HSP90, is currently under clinical investigation in both solid tumors and hematologic malignancies (22), and bortezomib, an inhibitor of the proteasome (a protein-degrading organelle), is approved for use in treating multiple myeloma (23). The ICG has made several new discoveries in this area that could lead to new drug candidates or new drug targets. The HDAC6-specific inhibitor tubacin (Fig. 2) has been shown to work synergistically with bortezomib to induce significant antitumor activity in multiple myeloma cells by inhibiting a parallel protein degradation mechanism known as the aggresome pathway (24). In a screen for compounds that stabilize cyclin B in synchronized *Xenopus laevis* cell extracts, a novel class of protein degradation modulators (ubistatins) were discovered that disrupt the binding of ubiquitinated substrates to the proteasome (25). A screen for small molecules that interfere with the degradation of the class I MHC heavy chain and its endogenous unmodified class I MHC heavy chain counterpart identified the closely related small molecules eeyarestatin I and II that interfere with a step preceding proteasomal degradation of these MHC heavy chains (26). The target of eeyarestatin is not yet known. Finally, a series of small molecules have been discovered that either activate or inactivate autophagy in cells.[11]

Additional examples of cancer-related discoveries derived from the interrogation of the nutrient-sensing network involving a number of anticancer drug targets, including the mammalian target

[10] http://www.broad.harvard.edu/chembio/icg/pubs.html.

[11] Perlstein, et al., submitted for publication.

of rapamycin, the protein phosphatase calcineurin and its substrate nuclear factor of activated T cells (NFAT), as well as the upstream phosphatidylinositol-3′-kinase/Akt signaling pathway. Providing evidence for the therapeutic relevance of nutrient sensing, rapamycin, and several analogues currently under clinical investigation seem to be well tolerated and may delay disease progression in some patients with renal cell carcinoma, breast cancer, or mantle cell lymphoma (27, 28). Small-molecule screens done at the ICG have uncovered novel compounds working at various points along this signaling network. For example, a fluorescence polarization assay was used to identify a series of inhibitors of calcineurin-NFAT signaling [inhibitors of NFAT-calcineurin association (INCAs)] by blocking the targeting of calcineurin to its substrate NFAT (29). Cell-based assays used to monitor NFAT dynamics in nonlymphoid cells led to the discovery of small molecules that inhibit translocation of NFAT to the nucleus, including ChemB-143072 (30). The psammaplysenes—natural products derived from the marine sponge, *Psammaplysilla*—were discovered in a screen to compensate for lost tumor suppressor functionality in cells deficient in phosphatidylinositol phosphate 3′-phosphatase (a phosphatase that works in opposition to phosphatidylinositol-3′-kinase). These natural products induce the relocalization of the transcription factor FOXO1a to the nucleus (31).

Tumor cell metastasis is a complex, desirable but classically inaccessible area for cancer therapeutic discovery. Among several key pathways that contribute to cell migration, the role of Rho GTPases, including RhoA, RhoB, Rac1, and Cdc42, seem quite promising as anticancer targets (32). Screens done at the ICG have credentialed several relevant cancer targets, using a chemical genetic approach. Phenotypic screening led to the discovery of secramine, a small molecule that inhibits activation of the Rho GTPase effector Cdc42 and Cdc42-dependent processes in cancer cells, including the inhibition of actin polymerization (33). A high-throughput screen for inhibitors of actin polymerization in *X. laevis* egg extracts led to the discovery of wiskostatin (34), an inhibitor of a Cdc42 effector, known as neural Wiskott-Aldrich syndrome protein. Screening for inhibitors of cell migration using an image-based wound healing assay yielded several noteworthy small molecules, including the Rho-kinase inhibitor Rockout (35). Rockout inhibited cell blebbing and disassembly of stress fibers in the human melanoma cell line M2. Although Rockout is less potent than existing ROCK inhibitors, its specificity is comparable and it offers a new structural class of Rho-kinase inhibitors that medicinal chemists can further optimize to develop therapeutic leads.

The small molecules identified by the ICG will require significant follow-up biological experimentation and medicinal chemistry before any would yield a clinical candidate. However, by identifying small molecules that can alter specific cancer-relevant pathways, it may one day be possible to fine-tune cancer treatment, akin to the "target and control" paradigm as described by von Eschenbach (36).

In summary, the ICG functions as a public screening center for cancer scientists seeking small-molecule probes, as an innovator of new screening methodologies, novel small molecules and new analysis techniques, and most importantly, as a discovery engine for both novel cancer targets and small-molecule leads for further preclinical development. In addition, the new database and analytic tool, ChemBank, is available for public data-mining once the screening data are released.

Acknowledgments

Received 7/11/2006; accepted 7/27/2006.

Grant support: National Cancer Institute's Initiative for Chemical Genetics under contract N01-CO-12400 (to Science Applications International Corporation, Frederick, Inc.).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Dr. Jeffery Derge (Science Applications International Corporation) for his continued support and guidance, and Dr. James Bradner for his thoughts regarding this manuscript.

References

1. Schreiber SL. Small molecules: the missing link in the central dogma. Nat Chem Biol 2005;1:64–6.
2. Hardwick JS, Kuruvilla FG, Tong JK, Shamji AF, Schreiber SL. Rapamycin-modulated transcription defines the subset of nutrient-sensitive signaling pathways directly controlled by the Tor proteins. Proc Natl Acad Sci U S A 1999;96:14866–70.
3. Muller O, Gourzoulidou E, Carpintero M, et al. Identification of potent Ras signaling inhibitors by pathway-selective phenotype-based screening. Angew Chem Int Ed Engl 2004;43:450–4.
4. Schreiber SL. The small-molecule approach to biology: chemical genetics and diversity-oriented organic synthesis make possible the systematic exploration of biology. Chem Eng News 2003;81:51–61.
5. Haggarty SJ, Koeller KM, Wong JC, Grozinger CM, Schreiber SL. Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation. Proc Natl Acad Sci U S A 2003;100:4389–94.
6. Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov 2002;1:727–30.
7. Stockwell BR. Exploring biology with small organic molecules. Nature 2004;432:846–54.
8. Burke MD, Schreiber SL. A planning strategy for diversity-oriented synthesis. Angew Chem Int Ed Engl 2004;43:46–58.
9. Sauer WH, Schwarz MK. Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. J Chem Inf Comput Sci 2003;43:987–1003.
10. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 2000;44:235–49.
11. Newman DJ, Cragg GM, Snader KM. Natural products as sources of new drugs over the period 1981–2002. J Nat Prod 2003;66:1022–37.
12. MacBeath G, Koehler AN, Schreiber SL. Printing small molecules as microarrays and detecting protein-ligand interactions en masse. J Am Chem Soc 1999;121:7967–8.
13. Bradner JE, McPherson OM, Mazitschek R, et al. A robust small-molecule microarray platform for screening cell lysates. Chem Biol 2006;13:493–504.
14. Perlman ZE, Slack MD, Feng Y, Mitchison TJ, Wu LF, Altschuler SJ. Multidimensional drug profiling by automated microscopy. Science 2004;306:1194–8.
15. Monks A, Scudiero D, Skehan P, et al. Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor cell lines. J Natl Cancer Inst 1991;83:757–66.
16. Shi LM, Fan Y, Lee JK, et al. Mining and visualizing large anticancer drug discovery databases. J Chem Inf Comput Sci 2000;40:367–79.
17. Lampson MA, Kapoor TM. Unraveling cell division mechanisms with small-molecule inhibitors. Nat Chem Biol 2006;2:19–27.
18. Quasthoff S, Hartung HP. Chemotherapy-induced peripheral neuropathy. J Neurol 2002;249:9–17.
19. Mayer TU, Kapoor TM, Haggarty SJ, King RW, Schreiber SL, Mitchison TJ. Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. Science 1999;286:971–4.
20. Straight AF, Cheung A, Limouze J, et al. Dissecting temporal and spatial control of cytokinesis with a myosin II Inhibitor. Science 2003;299:1743–7.
21. Eggert US, Kiger AA, Richter C, et al. Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets. PLoS Biol 2004;2:e379.
22. Dai C, Whitesell L. HSP90: a rising star on the horizon of anticancer targets. Future Oncol 2005;1:529–40.
23. Burger AM, Seth AK. The ubiquitin-mediated protein degradation pathway in cancer: therapeutic implications. Eur J Cancer 2004;40:2217–29.
24. Hideshima T, Bradner JE, Wong J, et al. Small-molecule inhibition of proteasome and aggresome function induces synergistic antitumor activity in multiple myeloma. Proc Natl Acad Sci U S A 2005;102:8567–72.
25. Verma R, Peters NR, D'Onofrio M, et al. Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. Science 2004;306:117–20.
26. Fiebiger E, Hirsch C, Vyas JM, Gordon E, Ploegh HL, Tortorella D. Dissection of the dislocation pathway for type I membrane proteins with a new small molecule inhibitor, eeyarestatin. Mol Biol Cell 2004;15:1635–46.
27. Dancey JE. Inhibitors of the mammalian target of rapamycin. Expert Opin Investig Drugs 2005;14:313–28.
28. Harding MW. Immunophilins, mTOR, and pharmacodynamic strategies for a targeted cancer therapy. Clin Cancer Res 2003;9:2882–6.

29. Roehrl MH, Kang S, Aramburu J, Wagner G, Rao A, Hogan PG. Selective inhibition of calcineurin-NFAT signaling by blocking protein-protein interaction with small organic molecules. Proc Natl Acad Sci U S A 2004; 101:7554–9.
30. Venkatesh N, Feng Y, DeDecker B, et al. Chemical genetics to identify NFAT inhibitors: potential of targeting calcium mobilization in immunosuppression. Proc Natl Acad Sci U S A 2004;101: 8969–74.
31. Schroeder FC, Kau TR, Silver PA, Clardy J. The psammaplysenes, specific inhibitors of FOXO1a nuclear export. J Nat Prod 2005;68:574–6.
32. Aznar S, Fernandez-Valeron P, Espina C, Lacal JC. Rho GTPases: potential candidates for anticancer therapy. Cancer Lett 2004;206:181–91.
33. Pelish HE, Peterson JR, Salvarezza SB, et al. Secramine inhibits Cdc42-dependent functions in cells and Cdc42 activation *in vitro*. Nat Chem Biol 2006;2: 39–46.
34. Peterson JR, Bickford LC, Morgan D, et al. Chemical inhibition of N-WASP by stabilization of a native autoinhibited conformation. Nat Struct Mol Biol 2004;11:747–55.
35. Yarrow JC, Totsukawa G, Charras GT, Mitchison TJ. Screening for cell migration inhibitors via automated microscopy reveals a Rho-kinase inhibitor. Chem Biol 2005;12:385–95.
36. von Eschenbach A. Collaboration to play key role in NCI's future, director says. Interview by Katherine Arnold. J Natl Cancer Inst 2002;94:790–2.

Small Molecules, Big Players: the National Cancer Institute's Initiative for Chemical Genetics

Nicola Tolliday, Paul A. Clemons, Paul Ferraiolo, et al.

Cancer Res 2006;66:8935-8942.