Chemistry & Biology, Vol. 10, 881–892, September, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2003.08.009

Biological Mechanism Profiling Using an Annotated Compound Library

David E. Root, Stephen P. Flaherty, Brian P. Kelley, and Brent R. Stockwell*
Whitehead Institute for Biomedical Research
9 Cambridge Center
Cambridge, Massachusetts 02142

Summary

We present a method for testing many biological mechanisms in cellular assays using an annotated library of 2036 small organic molecules. This annotated compound library represents a large-scale collection of compounds with diverse, experimentally confirmed biological mechanisms and effects. We found that this chemical library is (1) more structurally diverse than conventional, commercially available libraries, (2) enriched in active compounds in a tumor cell viability assay, and (3) capable of generating hypotheses regarding biological mechanisms underlying cellular processes. We elucidated biological mechanisms relevant to the antiproliferative activity of 85 compounds from this library that were selected using a high-throughput cell viability screen. We developed a novel automated scoring system for identifying statistically enriched mechanisms among such a subset of compounds. This scoring system can identify both previously known and potentially novel antiproliferative mechanisms.

Introduction

Small organic molecules have been used to mimic genetic mutations for the study of biological systems in an approach called chemical genetics [1–4]. This strategy generally requires screening tens of thousands of organic compounds in cellular assays to identify a small number of active compounds and then requires significant additional effort to illuminate the mechanistic basis for each compound's activity [3]. Just as genetic approaches have benefited from the use of large collections of annotated genes and genetic probes [5, 6], chemical genetic approaches might benefit from the use of annotated compounds. We set out to identify, collect, and assemble into a screenable format thousands of small molecules with experimentally verified biological mechanisms and activities.

Our goal was to use such an annotated compound library (ACL) to capture information that previously would have required many target-identification experiments. We chose to focus, in our initial study, on aspects of viability and proliferation unique to tumor cells. Proliferation rate is one of many cellular alterations associated with cancer that can be used as the basis for a chemical genetic screen; others include activation of angiogenesis and metastatic potential, epithelial to mesenchymal transition, and altered adhesiveness and growth rates [7]. In this study, we developed a method for integrating existing biological information related to a large number of compounds to study the biological mechanisms sufficient for killing or arresting the growth of a human cancer cell line derived from a clinical specimen.

The National Cancer Institute has assembled a large collection of compounds with either potential or demonstrated activity in cancer and viral-related assays. Our approach is distinct from the NCI's approach in that we obtained and tested compounds with a more diverse range of biological activities, including neurologically active compounds, antidiabetic compounds, controlled substances, and compounds with many other activities unrelated to cancer and infectious diseases. Our hypothesis was that an annotated compound library covering such a wide range of activities would be useful in many different cell-based assays. This collection is complementary to the NCI collection, and we have recently begun to compile a composite set comprising both our collection and the NCI diversity set of 1990 compounds (vide infra).

Many previous studies [8–14] have identified specific compounds that prevent the growth of tumor cells, a number of which have been developed into clinically effective anticancer drugs. In such cases, much effort was required to identify the specific proteins targeted by each of these antitumor agents. We hypothesized that an ACL composed of well-studied organic compounds with diverse biological mechanisms *not limited to those known to affect tumor cell viability* would allow for rapid elucidation of new biological targets and mechanisms relevant to tumor cell viability, and eventually other cellular processes of interest. In our view, such a result could be achieved by testing annotated compounds for their effects on tumor cell proliferation, identifying the subset of these compounds that prevent tumor cell proliferation, and determining the biological mechanisms that are statistically enriched among this group of compounds relative to the collection as a whole. We report herein the successful execution of this biological mechanism profiling strategy.

Results and Discussion

We assembled a collection of 2036 biologically active compounds and assigned each compound one of 169 broad biological mechanism descriptors ("primary descriptors"). Ninety-nine of these primary descriptors are represented by three or more compounds (Table 1), illustrating the wide variety of biological mechanisms probed by the compounds present in the library.

All of the compounds in our ACL have biological activity, but only a small percentage of compounds in conventional compound libraries are expected or reported to have biological activity. We sought to identify differences in distributions of molecular descriptors between

*Correspondence: stockwell@wi.mit.edu

Table 1. Representative Classes of Biologically Active Compounds and the Number of Compounds in Each Class

ACE inhibitor	6	Cl^- ion signaling modulator	3	muscle relaxant	6
actin ligand	5	cyclooxygenase inhibitor	29	Na^+ ion signaling	18
adenosine receptor ligand	78	dehydrogenase modulator	5	neuronal function	9
adenylate cyclase ligand	15	detergent	3	nitric oxide signaling	45
adrenergic receptor ligand	105	diacylglycerol kinase modulator	3	opioid receptor ligand	32
aa metabolism	12	diuretic	4	P450 modulator	10
aminopeptidase inhibitor	7	DNA binding agent	66	PARP	5
amphetamine	7	DNA gyrase ligand	5	phosphatase ligand	4
analgesic	11	dopaminergic	111	phosphodiesterase	14
anesthetic	12	dye (cell staining)	9	phospholipase	13
antibacterial agent	46	enkephalinase ligand	4	PI3K inhibitor	3
antibiotic (general)	32	folate metabolism modulator	9	platelet aggregation inhibitor	6
anticonvulsant	3	G protein modulator	3	prostaglandin	16
antifungal	11	GABA receptor ligand	45	protease inhibitor	28
antihypertensive agent	5	glucose modulator	5	protein kinase inhibitor	62
antiinflammatory agent	6	glutamatergic agent	88	protein synthesis inhibitor	8
antimalarial agent	4	glutathione modulator	4	ribosome activity modulator	8
antioxidant	15	glycoprotein synthesis	3	RNA biosynthesis inhibitor	5
antiviral agent	9	glycosidase inhibitor	9	serotonergic agent	71
aromatase inhibitor	4	gonadotropin	6	steroid	67
ATPase inhibitor	10	guanylate cyclase	4	taste receptor ligand	5
benzodiazepine	5	histamine modulator	37	thymidylate synthase inhibitor	8
Ca++ ion signaling modulator	45	imidazoline modulator	4	TNF modulator	5
calmodulin inhibitor	24	ion signaling (general)	36	topoisomerase ligand	7
cannabinoid receptor ligand	6	K^+ ion signaling modulator	17	tryptophan hydroxylase ligand	7
caspase inhibitor	5	L-aa decarboxylase inhibitor	8	tubulin ligand	20
ceramide signaling modulator	7	lectin ligand	3	tyrosine hydrolase ligand	6
cGMP modulator	3	lipid biosynthesis modulator	9	tyrosine kinase ligand	25
ion chelator	7	lipoxygenase ligand	5	vanilloid receptor ligand	3
cholecystokinin modulator	6	melatonin modulator	3	vasodilator	8
cholesterol biosynthesis modulator	5	metal ion modulator	3	venom	7
cholinergic agent	79	mitochondrial modulator	3	vitamin	18
cholinesterase inhibitor	16	monoamine oxidase inhibitor	15	xanthine oxidase inhibitor	4

Abbreviations: amino acid, aa; angiotensin converting enzyme, ACE; adenosine triphosphate, ATP; cyclic guanosine monophosphate, cGMP; deoxyribonucleic acid, DNA; gamma amino butryic acid, GABA; cytochrome P450, P450; poly(ADP ribose) polymerase, PARP; phosphatidylinositol-3-kinase, PI3K; and tumor necrosis factor, TNF.

these two types of libraries. We compared the distributions of 138 molecular descriptor values for the 1613 compounds in the ACL with associated electronic structure files versus 20,000 compounds in a synthetic combinatorial library (Comgenex) and 29,996 compounds in a commercially available synthetic compound collection (Chembridge). The two latter libraries are typical of those used in high-throughput, chemical genetic screens [15–18]. As seen in Figure 1A, the middle 80% of descriptor values for the ACL has a wider range than either the Chembridge or Comgenex libraries for most descriptors (average ratio of ranges, ACL/Chembridge = 1.8), whereas the Chembridge and Comgenex libraries have similar ranges for most descriptors (average ratio of ranges, Comgenex/Chembridge = 1.1). The middle 98% of ACL descriptor values exhibits an even greater spread relative to the corresponding range for the Chembridge and Comgenex libraries (Figure 1B). We used a self-organizing map (SOM) [19] to cluster compounds from the three libraries based on their descriptor values and to illustrate that the ACL spans a different region of descriptor space and is more diverse than either commercial library (see Supplemental Figure S1 at http://www.chembiol.com/cgi/content/full/10/9/881/DC1). This SOM illustrates that the Comgenex library is largely composed of molecules in small and densely packed regions of chemical space, likely due to the combinatorial nature of the library design.

The small number of compounds in the ACL covers a significantly larger range of descriptor values than the 20,000 compounds and 29,996 compounds in the Comgenex and Chembridge libraries, respectively, (Figure 1) despite the fact that a major criterion in the design of the latter two libraries was structural diversity [18, 20]. The ACL was designed to have broad *functional* diversity, and it incorporates compounds affecting a wide range of biological mechanisms. The observation that the ACL is structurally as well as functionally diverse is intriguing but implies no causal relationship between structural and functional diversity. However, these results do demonstrate that some commercial libraries fail to test large biologically relevant descriptor ranges, and, while there is no evidence that structural diversity per se will yield a higher hit rate, a more structurally diverse library can yield a more structurally diverse set of active compounds.

We anticipated that compounds with previously described biological activity would have a greater proba-

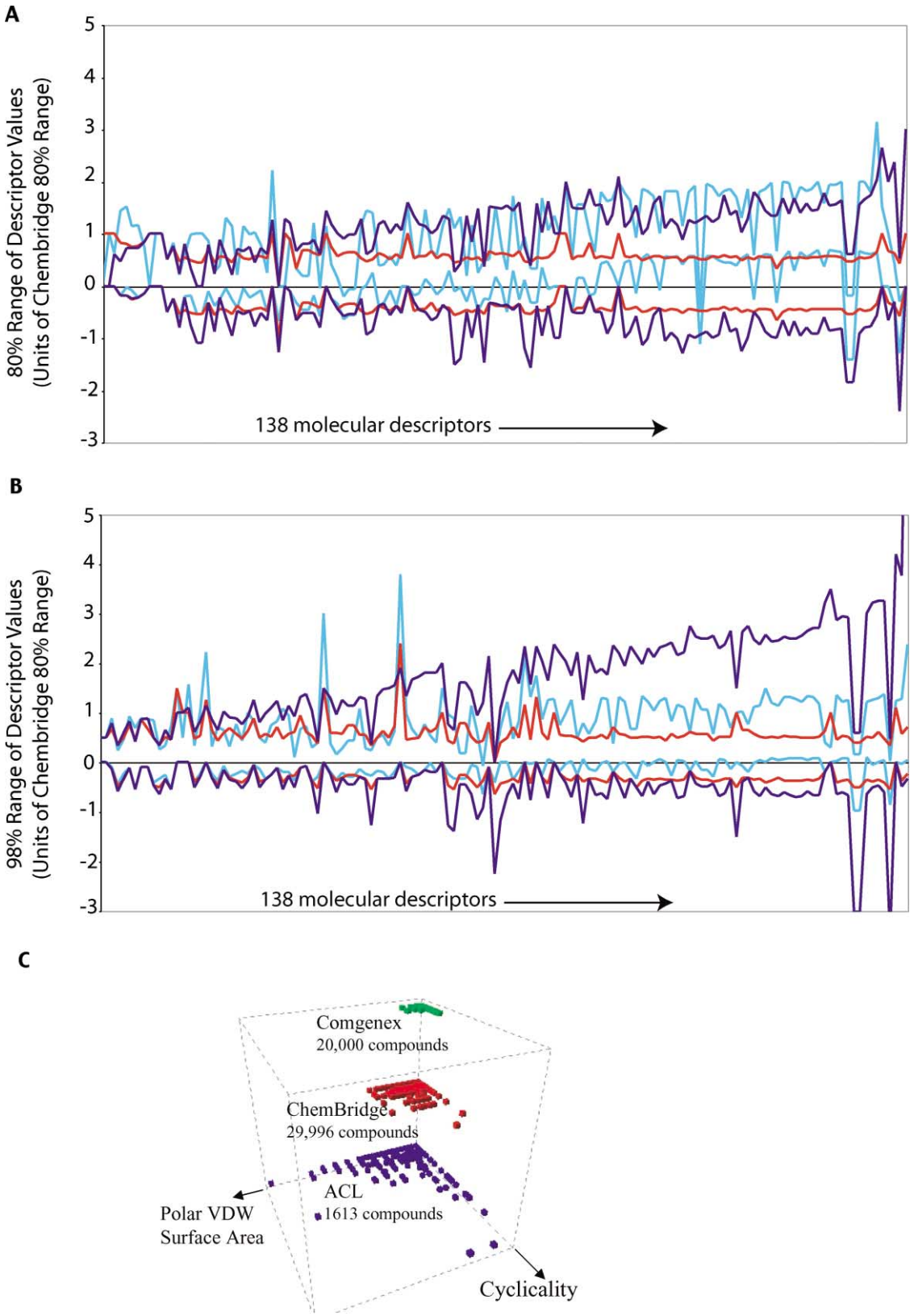

Figure 1. An Annotated Library of Biologically Active Compounds Is More Diverse Than Conventional Compound Libraries

(A and B) The plots depict the upper and lower bounds for values of 138 molecular descriptors for the middle n% of descriptor values in three libraries (ACL, Chembridge [CB], and Comgenex [CGX]), where n = 80 (A) and 98 (B). In (A), the upper and lower red lines indicate the 90th and 10th percentile values of each molecular descriptor for the CB library. Similarly, the dark blue and light blue lines indicate the 90th and 10th percentile values for the ACL and CGX libraries, respectively. The descriptors were sorted in order of increasing range (upper bound minus lower bound) in the ACL library, resulting in an increasing gap between dark blue lines.

(C) Scatter plot of two molecular descriptors for all three libraries. The values of polar van der Waals surface area and cyclicality, as defined by Petitjean [49], were plotted for 1613 ACL compounds (blue), 29,996 CB compounds (red) and 20,000 CGX compounds (green).

bility than random compounds of being active in new cellular assays because their biological molecular mechanisms might be operative in a new context. To test this hypothesis, we evaluated the ability of the ACL and the Comgenex libraries to selectively inhibit the proliferation of an engineered human tumor cell line [21]. This cell line was produced by introduction of genetic elements encoding the simian virus 40 large T (LT) and small T (ST) oncoproteins, the telomerase catalytic subunit (hTERT), and an oncogenic allele of RAS into primary human fibroblasts [22, 23]. When all four of these proteins are expressed together, these cells (BJELR) acquire characteristics of tumor cells, such as growth in soft agar and tumor formation in mice [22, 23]. We tested the effects on proliferation of BJELR and BJ cells of the ACL and Comgenex compounds at 4 μg/ml for 48 hr. By measuring the staining of these cells with calcein acetoxymethyl ester (calcein AM) [24], a viability dye, we found that 2.5% of the ACL compounds inhibited proliferation of engineered BJELR tumor cells by at least 80%, whereas only 0.69% of the Comgenex compounds did so, indicating a 4-fold enrichment for such antitumor agents in the ACL. More importantly, when we tested these compounds for tumor cell selectivity versus primary cells (i.e., the ratio of concentrations required to achieve 50% reduction of calcein staining in tumor cells and primary cells), we found that 1% of the ACL compounds were at least 4-fold selective for killing of tumor cells versus primary cells, but only 0.01% of the Comgenex compounds met this selectivity threshold. Since primary cells are normal, nontumorigenic cells freshly derived from human tissue, selectivity for tumor cells compared to these cells indicates specificity for mechanisms particular to tumor cells. Thus, there was enrichment of compounds with tumor cell-selective killing in the ACL relative to a combinatorial library [21]. It should be noted that although we selected compounds for inclusion in the ACL based on reported biological activity, none of the compounds had previously been tested in this assay system, i.e., for their ability to kill engineered tumor cells selectively relative to their isogenic primary cell counterparts.

In a second experiment with the ACL, we sought to test its ability to uncover mechanisms associated with particular cellular processes, in this case tumor cell proliferation and viability. We treated A549 lung carcinoma cells [25, 26] with each ACL compound in triplicate for 1, 2, and 3 days to identify both rapidly and slowly acting antiproliferative and cytotoxic mechanisms (Figure 2A). Of the 2036 compounds in our collection, 85 caused a 50% or more reduction in staining of A549 cells with the viability dye calcein AM [24] (Table 2). A number of these compounds (37) have not been previously tested in the NCI's multiyear comprehensive compound screening effort [10], and 25 of these compounds were not in the NCI collection of 249,071 compounds (see Supplemental Table S1 at http://www.chembiol.com/cgi/content/full/10/9/881/DC1).

The abundance of active compounds and large amount of associated literature-based information makes it difficult to follow up on each active compound with a manual search of the literature. In a traditional chemical genetic approach, the 85 active compounds would serve as a starting point for 85 separate time-intensive, target-identification projects. Instead, we integrated and analyzed the existing literature on these 85 compounds to extract biological mechanisms enriched in this group of compounds relative to the entire ACL. The results of such an analysis can be used to guide and prioritize validation experiments.

To facilitate analysis of existing literature, we developed automated algorithms for both mechanistic annotation of compounds and identification of enriched mechanisms among a subset of compounds. One obstacle to making full use of the existing literature is that most published reports associate a single biological mechanism with any given small molecule, despite the fact that within cells many diverse molecular changes are observed after treatment with each compound. Compound annotation that accommodates multiple effects associated with each compound is critical for correctly inferring mechanisms of action. We developed a comprehensive vector-based strategy for compound annotation that is compatible with multiple mechanisms for each compound. We generated a vector for each compound with a quantitative score for each of 12,755 different biological mechanisms, comprising the 169 primary descriptors, 200 Medline medical subject heading terms related to pharmacology, and more than 12,000 human gene names (our "12K annotation"). For each compound in the ACL, we scored the relevance of each biological mechanism term by performing a search of the more than 11 million records of the Medline biomedical literature database for the number of co-occurrences of the biological mechanism term and the given compound name in the title, abstract, or keyword fields (Figure 3).

To determine the reliability of this 12K annotation, we compared it with our manual annotation. In a randomly chosen subset of 235 compounds, 210 (89%) had manually assigned primary descriptors that were the same as one of the top three ranked mechanism descriptors assigned by the automated annotation method (Figure 3A), indicating reasonable agreement between manual and automated annotation methods. The automated annotation can be updated regularly to incorporate the most recent literature and to add literature annotation for new ACL compounds. Future studies may focus on incorporating additional biological mechanism terms and databases and on implementing natural language processing systems [27] to improve further the accuracy of automated annotation. Recently, a similar strategy has been applied to the related problem of gene network analysis [28].

We used our 12K annotation to identify, in an unbiased fashion, 28 mechanisms that were statistically overrepresented among the 85 active compounds identified in our antitumor screening procedure versus the parent library (Table 3; see Supplemental Data at *Chemistry & Biology*'s website for details). Our website (http://staffa.wi.mit.edu/stockwell/) allows entry of a set of compound names from the ACL and returns a list of enriched mechanisms in this subset of the ACL. We refer to this procedure for identifying enriched biological mechanisms as Global Mechanism Extraction. The list of enriched biological mechanisms consists of general antiproliferative

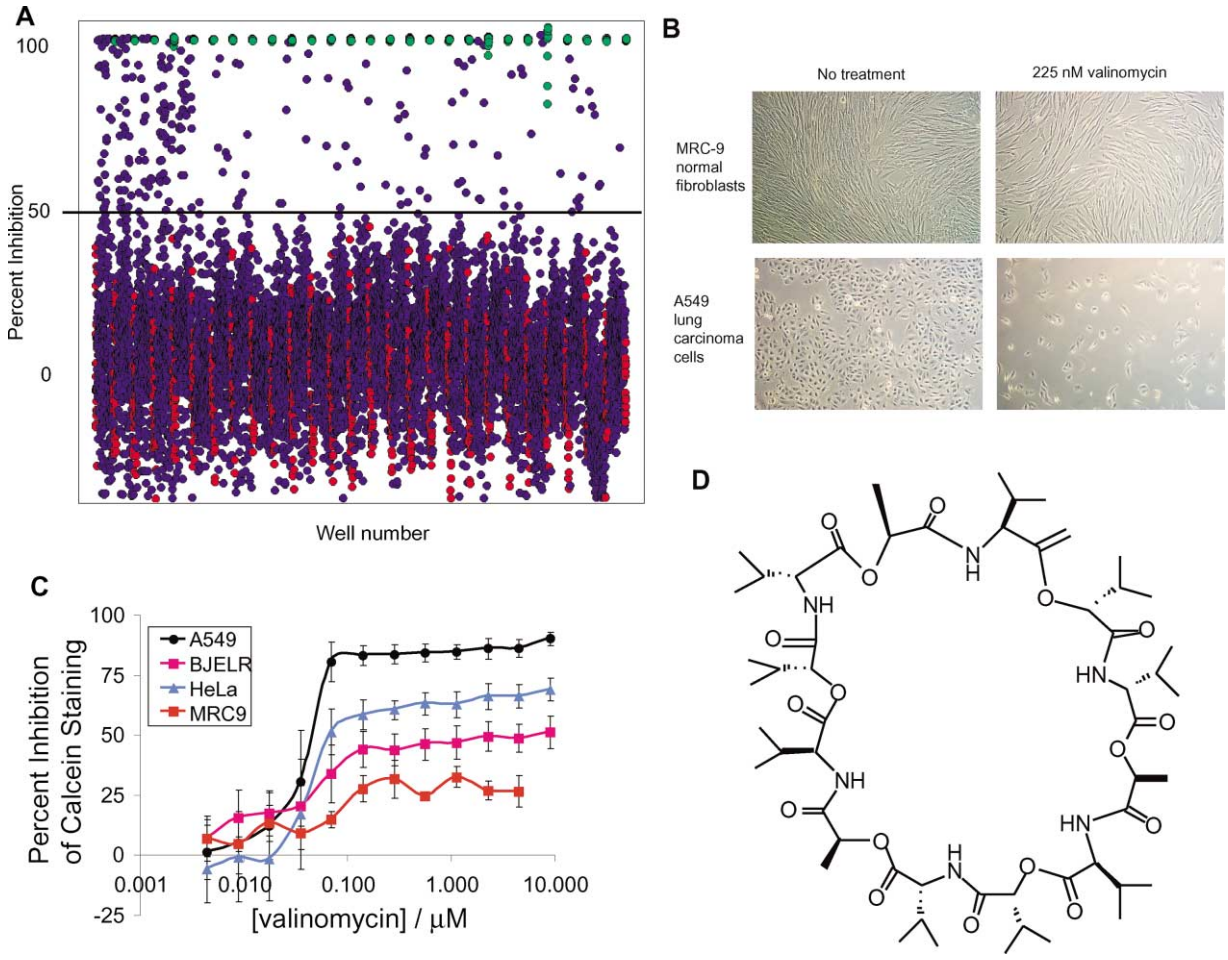

Figure 2. 85 ACL Compounds Inhibit A549 Cell Viability and Proliferation, and a Subset of Compounds Including Valinomycin Selectively Kills Lung Carcinoma Cells but Not Normal Lung Fibroblasts

(A) Effect of all 2036 compounds in the ACL on viability of A549 human lung carcinoma cells. Mean percent inhibition (of three replicates) of calcein AM signal for each compound at each time point is plotted (blue circles), along with untreated negative control wells (red circles) and positive control wells lacking cells (green circles). The horizontal bar indicates 50% inhibition of signal and separates active compounds from all negative controls.

(B) Valinomycin kills lung carcinoma cells (A549) but not normal lung fibroblasts (MRC9).

(C) Dose response for valinomycin in multiple cell types.

(D) Structure of valinomycin.

terms, clinically validated anticancer mechanisms [12, 29–34], other cancer or cell death-related mechanisms, as well as several mechanisms with no obvious or previously recognized relationship to cell death. The identification of known anticancer mechanisms confirms the internal consistency of this automated procedure, and the identification of unanticipated mechanisms offers the potential for finding novel associations.

Among the novel antiproliferative mechanisms generated by this screen, the biological mechanism term "ionophores" was determined to be enriched among the hit compounds. The hit compounds narasin, nonactin, thiomuscimol, and valinomycin (Figures 2B and 2C) appeared in four Medline records along with the term "ionophores"[35–38]. In contrast, when 85 randomly selected compounds from the ACL were similarly analyzed in 1000 repeated trials, the term ionophore was rarely associated (i.e., the term is not promiscuous). This suggests a specific hypothesis, namely that ionophores or a subset of ionophores are capable of killing human A549 lung cancer cells. These and others compounds in the set of 85 that inhibits the growth of A549 tumor cells would not have been selected a priori as likely to have such activity in tumor cells. For example, the "ionophore" thiomuscimol is primarily used as a GABA receptor agonist, while narasin and nonactin are primarily used as antiinfective agents; none of these compounds are described in the literature as having strong antiproliferative activity in human tumor cells. To quantify this point directly, we determined the number of the 85 active compounds that were found in at least one Medline record with tumor or cell death-related biological terms (tumor, tumors, carcinoma, sarcoma, adenocarcinoma, squamous cell carcinoma, small cell carcinoma, cell division, antineoplastic agents, cultured tumor cells, phytogenic antineoplastic agents, apopto-

Table 2. Compounds that Reduce Viability of Lung Carcinoma Cells by >50%

Compound Name	24 hr	48 hr	72 hr	Compound Name	24 hr	48 hr	72 hr	Compound Name	24 hr	48 hr	72 hr
(+/−)-Dropropizine			59 ± 2	Cytochalasin E		94 ± 1	82 ± 4	Parthenolide			86 ± 5
(−)-Indolactam V	52 ± 2			Deferoxamine mesylate			55 ± 4	Phenylarsine oxide	100 ± 0	100 ± 0	100 ± 0
(2S, 3R)-3-Amino-2-hydroxy-4-(4-nitrophenyl)butanoyl-L-leucine		71 ± 3	80 ± 5	Digoxin	67 ± 6	100 ± 2	94 ± 2	Phomopsin A	60 ± 1	55 ± 1	77 ± 3
(D,L)-Glutamic acid		52 ± 8	71 ± 15	Dihydroouabain	69 ± 7	59 ± 22	80 ± 3	Pinacidil	55 ± 4	89 ± 7	53 ± 28
(±)-Pindobind			69 ± 8	Diphenyleneiodonium chloride		50 ± 6	77 ± 2	Podophyllotoxin	71 ± 9	84 ± 4	90 ± 8
1-Aminocyclohexane-trans-1,3-dicarboxylic acid			54 ± 5	Dipyridamole			52 ± 13	Proscillaridin A	84 ± 4	97 ± 1	100 ± 0
11-nor-delta8-tetrahydro-cannabinol-9- carboxylic acid			52 ± 12	Echinomycin	60 ± 1	89 ± 1	97 ± 2	Quinacrine		69 ± 8	91 ± 2
2-Amino-6-mercaptopurine		68 ± 18	72 ± 18	Emetine	62 ± 2	101 ± 5	92 ± 1	Sangivamycin		83 ± 4	79 ± 2
2-Amino-8-hydroxyquinoline		100 ± 0	99 ± 2	Enniatin	68 ± 5	75 ± 9	92 ± 2	Sanguinarine		55 ± 12	72 ± 4
2-Methoxyestradiol	51 ± 2			Epitestosterone			52 ± 40	Streptonigrin	100 ± 0	100 ± 0	100 ± 0
3′-Azido-3′-deoxythymidine		93 ± 3	95 ± 2	Fumagillin			59 ± 8	Sulfacetamide		60 ± 6	59 ± 19
3-Phenylpropargylamine		67 ± 3	79 ± 4	Hydroxytolbutamide		51 ± 10	59 ± 22	Sulfamerazine	56 ± 7	72 ± 5	55 ± 8
4,5,6,7-Tetrahydro-isoxazolo[4,5-c]pyridin-3-ol			57 ± 6	L-Erythro-sphingosine			60 ± 46	Sulfamethizole		61 ± 8	61 ± 13
5- Hydroxyuridine		80 ± 7	69 ± 14	Menadione	77 ± 15	61 ± 11	67 ± 17	Tetraethylthiuram disulfide		100 ± 1	100 ± 0
5-Ethyl-2′-deoxyuridine		95 ± 2	95 ± 3	Methapyrilene		73 ± 12	66 ± 2	Thiabendazole			60 ± 24
5-Fluorouracil			55 ± 12	Methotrexate			52 ± 4	Thiomuscimol			81 ± 7
5-Thio-D-glucose			56 ± 9	Methyl (5-[2-thienylcarbonyl]-1H-benzimidazol-2-yl)carbamate	58 ± 2	65 ± 1	78 ± 5	Trazodone		54 ± 4	
9-Aminoacridine		94 ± 5	92 ± 5	Mianserin		79 ± 16	72 ± 37	Trimethoben-zamide	59 ± 10	83 ± 14	93 ± 4
Acetyl-beta-methylcholine chloride			62 ± 20	Mitoxantrone		88 ± 13	95 ± 3	Trimethyl-(m-aminophenyl)-ammonium chloride		74 ± 8	72 ± 21
Actinomycin C		78 ± 18	89 ± 4	Mycophenolic acid		57 ± 3	80 ± 5	Tunicamycin		79 ± 4	91 ± 7
Actinomycin D		81 ± 4	86 ± 2	N-Oleoylethanolamine		80 ± 5	73 ± 12	Tyrphostin 23			80 ± 19
Adenosine			53 ± 2	N-p-tosyl-L-valine chloromethyl ketone		61 ± 12	77 ± 6	Tyrphostin 46			72 ± 12
Adenosine 5′-O-(3-thiotri-phosphate)			69 ± 26	N6-Benzyladenosine		65 ± 1	64 ± 7	Tyrphostin A9			53 ± 9
Camptothecin		90 ± 3	92 ± 7	Narasin			57 ± 1	Valinomycin	65 ± 1	81 ± 2	92 ± 3
Cantharidin	89 ± 7	96 ± 2	99 ± 0	Nigericin			62 ± 22	Verrucarin A	73 ± 15	99 ± 2	95 ± 3
Chelerythrine chloride		70 ± 50	70 ± 41	Nonactin	65 ± 8	73 ± 4	90 ± 3	Vinblastine		79 ± 4	82 ± 1
Colchicine		84 ± 7	81 ± 14	Oligomycin		80 ± 4	73 ± 6	Vincristine		85 ± 6	81 ± 6
Crystal Violet	69 ± 6	89 ± 2	94 ± 1	Ouabain	78 ± 5	93 ± 2	99 ± 0	Willardiine		97 ± 1	93 ± 3
				Paclitaxel		75 ± 12	76 ± 7				

Percent inhibition of fluorescence intensity is reported. Error bars indicate one standard deviation.

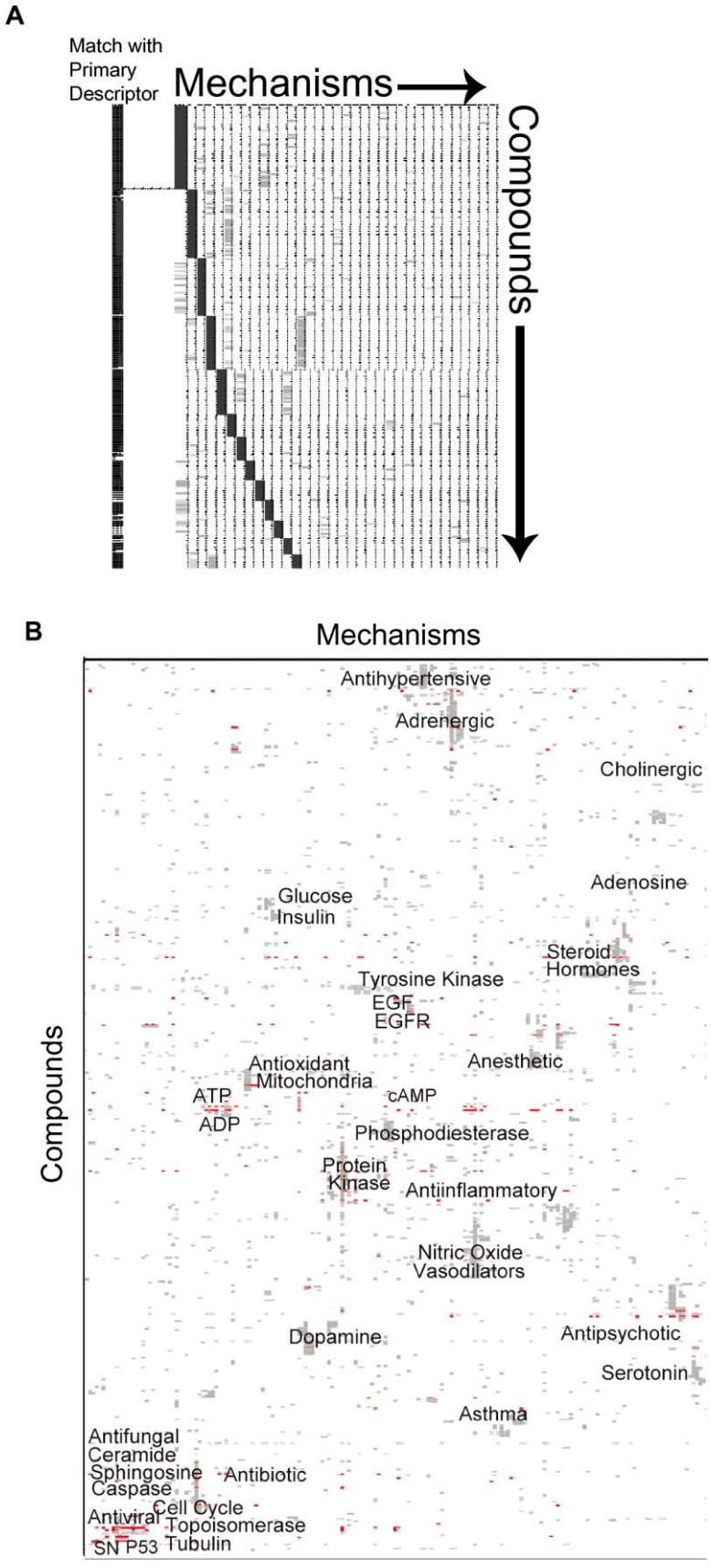

Figure 3. Automated Vector Annotation of Biologically Active Compounds

(A) The table illustrates the ranking of 36 mechanism descriptors for 235 compounds. Each row in the table represents a compound, and each column represents a mechanism. The top ranked mechanism for each compound was colored black, the second ranked mechanism was colored dark gray, the third ranked mechanism was colored light gray, and all other mechanisms were colored white. Compound vectors were sorted in the table by their top ranked mechanisms. The "Match with Primary Descriptor" column was shaded black for a compound vector if the manually assigned primary descriptor for the compound appeared as one of the top three automatically ranked mechanisms.

(B) To illustrate the distribution of compounds and mechanisms in the ACL, we highlighted in dark gray (#1), medium gray (#2), and light gray (#3) the top scoring three mechanisms for each compound and performed a hierarchical clustering of both compound and mechanism vectors. Large clusters of compounds with the same or related mechanisms are labeled with the name of a representative mechanism to illustrate islands of relatedness in this plot of compound/mechanism space. Non-zero entries in the rows for 85 compounds that inhibited proliferation of A549 lung carcinoma cells are highlighted in red, illustrating the mechanisms that block proliferation of these cells.

Table 3. Mechanisms for Reducing A549 Lung Carcinoma Cell Viability Identified in Global Mechanism Extraction

	Mechanism	Score	No. of Hits	No. in ACL	No. of Records Hits	No. of Records ACL
General antiproliferative terms	antineoplastic agents	0	15	118	27	118
	growth inhibitors	0	2	14	2	7
	cell cycle	0	4	50	5	58
	immunosuppressive agents	0	4	38	20	43
Clinically validated mechanisms	topoisomerase I	0	4	10	3	4
	topoisomerase	0	5	15	4	8
	tubulin	0	5	18	4	12
	antimetabolites	0	4	34	7	32
	protein kinase	0	13	193	12	180
	DNA	0	24	305	33	612
Cancer/death-related mechanisms	caspase	0	5	44	5	32
	TNF	0	2	17	2	14
	BRCA1	0	2	9	2	6
	p53	0	4	33	3	48
	oxidative phosphorylation	0	2	14	3	8
	mitochondria	0	4	71	8	54
Other mechanisms	ATPase	0	4	43	4	34
	ionophores	0	4	28	4	23
	antidepressive agents	0	3	55	4	50
	potassium channel	0	5	50	6	38
	protein synthesis	0	3	32	3	30
	actin	0	3	26	2	28
	dehydrogenase	0	6	121	11	88
	PDGF	0	2	12	2	7
	ATP	0	4	76	10	67
	oxidants	0	2	27	3	17
	antibiotics	3.E-04	9	137	8	154
	cardiotonic agents	4.E-04	3	14	3	14

85 active compounds that decreased viability of A549 lung carcinoma cells were compared to all 2036 compounds in the ACL using a procedure we termed Global Mechanism Extraction. Biological mechanism terms with statistical enrichment among these 85 compounds (score $<.0004$, $p < 0.02$) are listed. For each biological mechanism term, the table lists the "score," which is the product of (1) the percentile score for the number of Medline records reporting that biological mechanism in this group of 85 hit compounds relative to the distribution obtained for that mechanism when 85 randomly selected compounds are subjected to the same algorithm, and (2) the percentile score for the number of compounds associated with each biological mechanism relative to the distribution obtained for the mechanism when 85 randomly selected compounds are subjected to the same algorithm. Also listed is the actual number of hit compounds (No. of Hits) and library compounds (No. in ACL) that are associated with each mechanism and the number of Medline records containing both the mechanism term and the name of a hit compound (No. of Records Hits) and the number of Medline records containing both the term and the name of any library compound (No. of Records ACL). For more detail, see Supplemental Data at *Chemistry & Biology*'s website.

sis, apoptotic, antineoplastic combined chemotherapy protocol, lung neoplasms, breast neoplasms, cell cycle, tumor necrosis factor, cytotoxin, cytotoxic, and cytotoxicity). We found that for 58 of the 85 hit compounds there was at least one Medline record listing one of these terms and the compound name. In other words, one-third of the compounds (27 out of 85) would not have been selected a priori as antitumor agents using this broad definition of previously identified tumor-related activity.

The ACL can accelerate the process of identifying promising candidate mechanisms underlying a biological process, but subsequent experimental confirmation of these proposed mechanisms remains a potentially difficult and ad hoc procedure. The ACL has revealed mechanisms underlying phenotypic effects in several studies we have undertaken, such as mechanisms that enable selective killing of BJELR engineered tumor cells but not isogenic BJ primary cells [21]. We compared the expression levels in these two cell lines of topoisomerase I (TOP1), the putative target of camptothecin [11, 39–44], one of our BJELR-selective compounds, and found that BJELR cells upregulate TOP1 relative to BJ cells. As camptothecin's putative mechanism of action in other cells involves introduction of double-strand DNA breaks in a TOP1-dependent manner [39, 45–47], upregulation of TOP1 explains the increased sensitivity of BJELR cells to camptothecin. In support of this interpretation, we found that genetic inactivation of TOP1 with a small interfering RNA (siRNA) in BJELR cells confers partial resistance to camptothecin [21]. Thus, in contrast to screens with conventional libraries, screens with the ACL generate hypotheses regarding biological mechanisms underlying cellular phenotypes; in some cases these hypotheses can be validated directly.

After identifying the 85 compounds that inhibit proliferation of A549 cells, we sought to determine whether any of these compounds were A549 selective rather than being generally cytotoxic toward many cell types. From the group of 85 active compounds, we selected 30 compounds representing 20 biological mechanisms for retesting in a 20-point, 2-fold dilution series in triplicate in both A549 lung carcinoma cells and in MRC9 normal human lung fibroblasts [48]. We calculated the ratio of the concentrations required in the two cell lines for 50% inhibition of calcein AM viability dye signal, yielding an A549 cell selectivity value for each compound (Table 4). Compounds that were more potent in

Table 4. Mechanisms for Reducing Calcein Staining in Lung Carcinoma Cells

Proposed Mechanism	Mean Tumor Selectivity	Compound(s)
ionophore	>128	valinomycin
tubulin binding	15.5	paclitaxel, phomopsin A, podophyllotoxin, vinblastine
inhibition of glycosylation	6.4	tunicamycin
DNA synthesis	5.4	actinomycin D, echinomycin, streptonigrin, sangivamycin
phosphatases	3	phenylarsine oxide
proteases	2	(2S, 3R)-3-Amino-2-hydroxy-4-(4-nitrphenyl)butanoyl-L-leucine
protein synthesis	2	emetine, verrucarin A
ceramidase	1.6	N-oleoylethanolamine
dehydrogenase	1.5	tetraethylthiuram disulfide
nitric oxide signaling	1.4	N-p-tosyl-L-valine chloromethyl ketone
MAO	1.1	quinacrine
NFKB	1.1	parthenolide
topoisomerase II	1.1	mitoxantrone
PKC	1	chelerythrine chloride
(Na+, K+)-ATPase	0.8	digoxin, dihydroouabain, proscillaridin A
adenosine deaminase	0.8	N6-benzyladenosine
PP2A	0.8	cantharidin
metal chelation	0.53	2-amino-8-hydroxyquinoline
actin	0.5	cytochalasin E
topoisomerase I	0.04	camptothecin

A549 cells compared to normal MRC9 cells received a selectivity ratio greater than one. Several compounds, including those that act by binding to DNA or small ions (i.e., with ionophore activity such as valinomycin; Figures 2B and 2C) or by inhibiting protein glycosylation, exhibited A549 cell selectivity. These results demonstrate that A549 lung carcinoma cells (derived from lung epithelial cells) are more sensitive than MRC9 cells (lung fibroblasts) to killing by these compounds. This may reflect a selectivity of these compounds and mechanisms for A549 cells in particular, or it may reflect the fact that these compounds and mechanisms are selective for lung epithelial cells over lung fibroblasts. These results illustrate that although many compounds kill mammalian cells nonspecifically, it is possible to identify cell-type selective compounds. In other studies, we have extended this concept to search for even more selective compounds, namely those that selectively kill cells expressing one or more oncoproteins [21].

We recognized that valinomycin's selectivity for A549 cells relative to MRC9 cells (Table 4) could either be due to the reported ionophore activity of valinomycin or due to a novel mechanism. To distinguish between these alternatives, we identified and tested five additional ionophores (narasin, nigericin, salinomycin, enniatin, and ammonium ionophore) for their activity in both A549 and MRC9 cells (see Supplemental Figure S2 at *Chemistry & Biology*'s website). We tested each compound in six replicates in a dilution series in both cell lines using the calcein AM viability assay, and determined that these five ionophores displayed A549-selective lethality with no detectable activity in MRC9 cells, although the maximum level of activity and the potency in A549 cells varied among the compounds. Since we observed that identification of such A549-selective activity is uncommon (Table 4), our finding that five additional ionophores display A549 selectivity implies that valinomycin's selectivity is due to its ionophore activity.

In summary, we have created a method for evaluating numerous biological mechanisms that may underlie a specific cellular phenotype of interest. The ability of this analytic method to evaluate multiple mechanisms associated with a particular compound makes it a useful tool in interpreting the effects of compounds on phenotypes. Future annotation strategies could be broadened to integrate systematic experimentation into the annotation strategy. For example, such annotation could include the pattern of transcriptional changes induced by a compound or the pattern of protein binding exhibited by a compound. In such cases, this methodology could be applied to sift through large numbers of selected compounds to determine the relevant biological mechanisms shared by those compounds. Moreover, the ACL and conventional synthetic libraries provide complementary approaches for an initial phenotype-based screen; compounds from conventional libraries can be incorporated into the ACL once they have verified biological activity. We envision an expanding collection of annotated compounds that define, with increasing precision, the biological molecular mechanisms that are involved in regulating cellular processes.

Significance

Annotated compound libraries are capable of generating hypotheses regarding underlying biological mechanisms, in contrast to traditional compound libraries. Just as in DNA microarray-based transcription profiling experiments and proteomic experiments, these hypotheses must ultimately be tested and validated using conventional biological methods. It is, however, possible to group compounds from a primary screen on the basis of their mechanistic similarity to guide and prioritize validation studies. We demonstrate a strategy for screening an annotated compound library in cellular assays and then generating mechanistic hypotheses; this strategy can likely be extended to other cellular phenotypes. For example, we are using this strategy to identify biological mechanisms underlying Huntington's disease and spinal muscular atro-

phy using cell-based models. We propose that such annotated compound libraries will prove to be an increasingly valuable resource for both chemists and biologists interested in performing chemical genetic screens.

Experimental Procedures

Selection and Storage of Compounds for the ACL

We assembled a collection of 2036 biologically active compounds available from Sigma-Aldrich Co. by identifying all compounds that were either (a) approved drugs or (b) had some associated annotation reporting biological activity. In cases where many close analogs or salt forms of a compound were available, three representative members of the compound class were chosen. We created a list of structurally distinct FDA-approved small molecule drugs from the electronic Orange Book [50], which is maintained by the FDA and lists all drug products that have been approved for use in the United States since 1938. As of December 1999, there were 19,299 products approved by CDER (the Center for Drug Evaluation and Research within the FDA), including over-the-counter medications, prescription drugs, and discontinued products. By eliminating duplicate products with the same active ingredients, we reduced this list to 1300 distinct FDA-approved drugs and purchased the 514 that were available from Sigma-Aldrich Co. All compounds were dissolved in dimethylsulfoxide (DMSO) at a concentration of 4 mg/m (corresponding to a concentration of 10 mM for a molecular weight of 400 g mol^{-1}) and dispensed into 384-well polypropylene plates. We created an electronic archive containing both chemical and biological information for each compound, including the chemical structure, well position, plate number, FDA-approval status, common name, and alternative names (see Supplemental Data at http://www.chembiol.com/cgi/content/full/10/9/881/DC1). We assigned one of 169 primary biological descriptor terms to each compound to categorize the general mechanism of action. These 169 primary descriptors were created manually by reading literature associated with each compound and creating one term that best captures that consensus literature view on the compound's primary mode of action. The 99 primary descriptors associated with three or more compounds are shown in Table 1 with the number of compounds in the primary descriptor category indicated.

Calculation of Molecular Descriptor Values

To compare the molecular descriptor ranges of the ACL with conventional synthetic libraries, we used the Molecular Operating Environment (MOE) software package (Chemical Computing Group) to electronically (1) remove counter ions by retaining only the largest covalently bound fragment, (2) adjust the protonation state to that of the predominant species that exists at pH 7.0, (3) set atom ionization to formal charge, and (4) add explicit hydrogens. We then removed duplicate compounds and compounds with no associated structural information from each library, reducing the ACL list from 2036 compounds to 1613 compounds with unique, available structures. The ChemBridge (CB) library list was reduced from 30,000 to 29,996 compounds; all 20,000 compounds of Comgenex (CGX) library were retained in this procedure. Using MOE, we calculated 138 molecular descriptors for each compound, including both 1D (atom counts) and 2D (bond connectivity) descriptors. For each library, the following percentiles were computed for each descriptor value distribution: 1, 5, 10, 50, 90, 95, and 99. Different descriptors have vastly disparate ranges; to facilitate comparisons over all the descriptors, each descriptor was normalized to its range in one of the libraries (CB) and also scaled to the CB median. This was accomplished as follows: the unit for descriptor normalization was the 80% range of the CB library (CB_range80), i.e., CB 90th percentile (CB90) minus CB 10th percentile (CB10). Percentile values for each library were normalized and scaled for each descriptor *i* as in the following example: $ACL90_i$ (normalized) = $ACL90'_i = (ACL90_i - CB\ median_i) / CB\_range80_i$ (see Figure 1). For example, if the ACL10′ and ACL90′ for a descriptor are −1 and +1, respectively, then ACL_range80′ = 2, the ACL has twice the range of the 80% range of the CB library, and the tails of the distribution are symmetrical about the CB median for this particular descriptor.

Tumor Cell Viability Assay

A549 lung carcinoma cells or BJELR engineered tumor cells were seeded in black, clear-bottom, tissue culture-treated 384-well plates (Costar #3712, VWR#29444-078), treated with each compound in triplicate at 4 μg/ml for 1, 2, and 3 days, washed with phosphate-buffered saline, and incubated for 4 hr with 0.7 μg/ml calcein acetoxylmethyl ester (calcein AM; Molecular Probes). Calcein AM is a nonfluorescent compound that freely diffuses into cells and is cleaved by intracellular esterases, generating the fluorescent anionic calcein, which is impermeable to live cell plasma membranes and is therefore retained only in live cells. Total fluorescence intensity of each well was measured on a Packard Fusion plate reader with a 485 nm excitation filter (20 nm bandpass) and a 530 nm emission filter (25 nm bandpass), and converted to percent inhibition of cell viability by subtracting the instrument background and dividing by the average signal from untreated control cells. For dose-response testing, cells (A549 human lung carcinoma cells, HeLa cervical carconoma cells, BJELR engineered tumor cells, normal BJ fibroblasts, or normal MRC9 fibroblasts) were seeded at a density of 3000 cells/well in 384-well plates, incubated for 2 days at 37°C with 5% CO_2, then processed as described above. For tumor selectivity determinations, compounds were tested in triplicate for their effect on viability of A549 lung carcinoma cells and MRC9 normal lung fibroblasts in 384-well plates in a 20 point, 2-fold dilution series at a maximum concentration of 4 μg/ml. The ratio of concentrations required for 50% inhibition of viability (IC_{50}) was calculated (MRC9/A549) and the geometric mean determined when multiple compounds acting by the same mechanism were tested (see Table 4).

Microscopic Examination of Cell Death

Human A549 lung carcinoma cells and MRC-9 primary human lung fibroblasts were seeded in 2 ml of medium in each well of six-well plates at a concentration of 100,000 cells/ml. Cells were either not treated or treated with 225 nM valinomycin, incubated for 2 days at 37°C in the presence of 5% CO_2, and photographed using phase contrast microscopy to confirm visually selective cell death in A549 cells.

Annotation of Compounds with Biological Mechanisms

We obtained a local copy of the Medline biomedical literature database and indexed both the 2,036 compound names in the ACL and 12,765 biological mechanism terms. We identified the number of times a compound name and a biological mechanism term both occurred in an abstract, keyword, or title of a record and recorded this number as the score for that compound name and biological mechanism pair. A searchable website containing this data is available at http://staffa.wi.mit.edu/stockwell/. Biological mechanisms were scored for each compound by counting the number of abstracts in Medline that contain both the compound name and a given biological mechanism.

Global Mechanism Extraction

Using the Medline database and a selection of hit compounds, a single score was generated for each mechanism term that ranked our confidence that the mechanism term was associated with the hit compounds in the context of the assay. Each score was computed independently across all mechanism terms that co-occurred with any hit compound name. In computing this score, we were interested in answering two questions. First, how many of the *hit compounds* were associated with a given mechanism term in the literature, and second, how many *Medline records* had both the mechanism term and one of our hit compound names? Having many compounds associated with a given mechanism term adds confidence that the mechanism is actually relevant in the context of the assay. Similarly, having many records in Medline containing our hit compounds and a particular mechanism term increases our confidence that this biological mechanism is justifiably associated with the hit compounds. Finally, we needed to remove the underlying biases of our compound library and Medline. For example, one of the most often used terms in Medline is "DNA." We expect to see more associations

between our hit compounds and this term than associations with other terms. To overcome this problem, we generated distributions by randomly selecting a set of 85 compounds and extracting the collection of Medline records that refer to at least one of these compounds. For each of 1000 iterations, we recorded, for each relevant mechanism term, the *number of compounds* in the randomly selected set with the mechanism (DIST1) and the *number of Medline records* in the extracted collection with the mechanism (DIST2). DIST1 is the expectation of observing, given a random selection of compounds, the number of compounds associated with a given mechanism. DIST2 is the expectation of observing, given the collection of Medline records extracted using a random set of compounds, the number of records with a given mechanism. Given these two distributions, DIST1 and DIST2, we computed, given a selection of hit compounds, (1) the probability of finding at least the number of compounds observed associated with a mechanism term and (2) the probability of finding at least the number of observed Medline records associated with both the hit compounds and a given mechanism term. These two probabilities were multiplied together to form the final score. A score of 0 indicates that the number of compounds or mechanism terms obtained for the 85 A549 hits was greater than the value obtained for that mechanism in all of the 1000 trials with random compounds. The full list of mechanisms and scores is available at http://staffa.wi.mit.edu/stockwell/. Compound names were stripped of stereochemical notations, salt designations, and solvent terms. An occurrence in a Medline record indicated an exact text match of the compound name and the biological mechanism term. All biological mechanism terms that occurred at least once with one selected compound name were listed in a table that includes the mechanism term, the compounds found with that mechanism, and a link to the Medline abstracts. These statistics are useful primarily as a guide toward expert analysis of the literature results. In practice, we have found that ordering the abstracts in this manner has simplified searching Medline for the biological mechanisms by which selected compounds operate.

Acknowledgments

This research of B.R.S. was funded in part by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund and by the National Cancer Institute (1R01CA97061-01).

Received: February 18, 2003
Revised: August 4, 2003
Accepted: August 4, 2003
Published online: August 28, 2003

References

1. Mitchison, T.J. (1994). Towards a pharmacological genetics. Chem. Biol. *1*, 3–6.
2. Schreiber, S.L. (1998). Chemical genetics resulting from a passion for synthetic organic chemistry. Bioorg. Med. Chem. *6*, 1127–1152.
3. Stockwell, B.R. (2000). Chemical genetics: ligand-based discovery of gene function. Nat. Rev. Genet. *1*, 116–125.
4. Stockwell, B.R. (2000). Frontiers in chemical genetics. Trends Biotechnol. *18*, 449–455.
5. Brown, P.O., and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. Nat. Genet. *21*, 33–37.
6. Maeda, I., Kohara, Y., Yamamoto, M., and Sugimoto, A. (2001). Large-scale analysis of gene function in Caenorhabditis elegans by high-throughput RNAi. Curr. Biol. *11*, 171–176.
7. Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.
8. Weinstein, J.N., and Buolamwini, J.K. (2000). Molecular targets in cancer drug discovery: cell-based profiling. Curr. Pharm. Des. *6*, 473–483.
9. Weinstein, J.N., Myers, T.G., O'Connor, P.M., Friend, S.H., Fornace, A.J., Jr, Kohn, K.W., Fojo, T., Bates, S.E., Rubinstein, L.V., Anderson, N.L., et al. (1997). An information-intensive approach to the molecular pharmacology of cancer. Science *275*, 343–349.
10. Shi, L.M., Fan, Y., Lee, J.K., Waltham, M., Andrews, D.T., Scherf, U., Paull, K.D., and Weinstein, J.N. (2000). Mining and visualizing large anticancer drug discovery databases. J. Chem. Inf. Comput. Sci. *40*, 367–379.
11. Liu, L.F., Desai, S.D., Li, T.K., Mao, Y., Sun, M., and Sim, S.P. (2000). Mechanism of action of camptothecin. Ann. N Y Acad. Sci. *922*, 1–10.
12. Long, B.H., and Stringfellow, D.A. (1988). Inhibitors of topoisomerase II: structure-activity relationships and mechanism of action of podophyllin congeners. Adv. Enzyme Regul. *27*, 223–256.
13. Miller, M.L., and Ojima, I. (2001). Chemistry and chemical biology of taxane anticancer agents. Chem. Rec. *1*, 195–211.
14. Zalacain, M., Zaera, E., Vazquez, D., and Jimenez, A. (1982). The mode of action of the antitumor drug bouvardin, an inhibitor of protein synthesis in eukaryotic cells. FEBS Lett. *148*, 95–97.
15. Stockwell, B.R., Hardwick, J.S., Tong, J.K., and Schreiber, S.L. (1999). Chemical genetic and genomic approaches reveal a role for copper in specific gene activation. J. Am. Chem. Soc. *121*, 10662–10663.
16. Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., and Mitchison, T.J. (1999). Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. Science *286*, 971–974.
17. Peterson, R.T., Link, B.A., Dowling, J.E., and Schreiber, S.L. (2000). Small molecule developmental screens reveal the logic and timing of vertebrate development. Proc. Natl. Acad. Sci. USA *97*, 12965–12969.
18. Darvas, F., Dorman, G., and Papp, A. (2000). Diversity measures for enhancing ADME admissibility of combinatorial libraries. J. Chem. Inf. Comput. Sci. *40*, 314–322.
19. Kohonen, T. (2001). Self-Organizing Maps, Volume 30 (Berlin: Springer).
20. Markham, P.N., Westhaus, E., Klyachko, K., Johnson, M.E., and Neyfakh, A.A. (1999). Multiple novel inhibitors of the NorA multidrug transporter of Staphylococcus aureus. Antimicrob. Agents Chemother. *43*, 2404–2408.
21. Dolma, S., Lessnick, S.L., Hahn, W.C., and Stockwell, B.R. (2003). Identification of genotype-selective anti-tumor agents using synthetic lethal chemical screening in engineered human tumor cells. Cancer Cell *3*, 285–296.
22. Hahn, W.C., Counter, C.M., Lundberg, A.S., Beijersbergen, R.L., Brooks, M.W., and Weinberg, R.A. (1999). Creation of human tumour cells with defined genetic elements. Nature *400*, 464–468.
23. Hahn, W.C., Dessain, S.K., Brooks, M.W., King, J.E., Elenbaas, B., Sabatini, D.M., DeCaprio, J.A., and Weinberg, R.A. (2002). Enumeration of the simian virus 40 early region elements necessary for human cell transformation. Mol. Cell. Biol. *22*, 2111–2123.
24. Wang, X.M., Terasaki, P.I., Rankin, G.W.J., Chia, D., Zhong, H.P., and Hardy, S. (1993). A new microcellular cytotoxicity test based on calcein AM release. Hum. Immunol. *37*, 264–270.
25. Lieber, M., Smith, B., Szakal, A., Nelson-Rees, W., and Todaro, G. (1976). A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. Int. J. Cancer *17*, 62–70.
26. Giard, D.J., Aaronson, S.A., Todaro, G.J., Arnstein, P., Kersey, J.H., Dosik, H., and Parks, W.P. (1973). In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J. Natl. Cancer Inst. *51*, 1417–1423.
27. Rindflesch, T.C., Tanabe, L., Weinstein, J.N., and Hunter, L. (2000). EDGAR: extraction of drugs, genes and relations from the biomedical literature. Pac. Symp. Biocomput. 517–528.
28. Jenssen, T.K., Laegreid, A., Komorowski, J., and Hovig, E. (2001). A literature network of human genes for high-throughput analysis of gene expression. Nat. Genet. *28*, 21–28.
29. Plunkett, W., Gandhi, V., Huang, P., Robertson, L.E., Yang, L.Y., Gregoire, V., Estey, E., and Keating, M.J. (1993). Fludarabine: pharmacokinetics, mechanisms of action, and rationales for combination therapies. Semin. Oncol. *20*, 2–12.
30. DeVita, V.T., Jr. (1997). In Cancer: Principles & Practice of Oncology, V.T. DeVita, Jr., S. Hellman, and S.A. Rosenberg, eds. (Philadelphia, PA: Lippincott-Raven), pp. 333–347.
31. Denny, W.A., and Wakelin, L.P. (1987). Mode and kinetics of

DNA binding of the anti-tumour agent bisantrene. Anticancer Drug Des. *2*, 71–77.
32. Druker, B.J. (2002). STI571 (Gleevec) as a paradigm for cancer therapy. Trends Mol. Med. *8*, S14–S18.
33. Iwahana, M., Ochi, Y., and Ejima, A. (2000). Antiproliferative activity and mechanism of action of DZ-3358, a novel pyrimidinyl pyrazole derivative. Anticancer Res. *20*, 785–792.
34. Di Leo, A., and Piccart, M.J. (1999). Paclitaxel activity, dose, and schedule: data from phase III trials in metastatic breast cancer. Semin. Oncol. *26*, 27–32.
35. Kaldhusdal, M., Schneitz, C., Hofshagen, M., and Skjerve, E. (2001). Reduced incidence of Clostridium perfringens-associated lesions and improved performance in broiler chickens treated with normal intestinal bacteria from adult fowl. Avian Dis. *45*, 149–156.
36. Bala, S., Kombrabail, M.H., and Prabhananda, B.S. (2001). Effect of phloretin on ionophore mediated electroneutral transmembrane translocations of H(+), K(+) and Na(+) in phospholipid vesicles. Biochim. Biophys. Acta *1510*, 258–269.
37. Rabe, H., Picard, R., Uusi-Oukari, M., Hevers, W., Luddens, H., and Korpi, E.R. (2000). Coupling between agonist and chloride ionophore sites of the GABA(A) receptor: agonist/antagonist efficacy of 4-PIOL. Eur. J. Pharmacol. *409*, 233–242.
38. Yeung, C.H., and Cooper, T.G. (2001). Effects of the ion-channel blocker quinine on human sperm volume, kinematics and mucus penetration, and the involvement of potassium channels. Mol. Hum. Reprod. *7*, 819–828.
39. Madden, K.R., and Champoux, J.J. (1992). Overexpression of human topoisomerase I in baby hamster kidney cells: hypersensitivity of clonal isolates to camptothecin. Cancer Res. *52*, 525–532.
40. Eng, W.K., Faucette, L., Johnson, R.K., and Sternglanz, R. (1988). Evidence that DNA topoisomerase I is necessary for the cytotoxic effects of camptothecin. Mol. Pharmacol. *34*, 755–760.
41. Hsiang, Y.H., and Liu, L.F. (1988). Identification of mammalian DNA topoisomerase I as an intracellular target of the anticancer drug camptothecin. Cancer Res. *48*, 1722–1726.
42. Champoux, J.J. (2000). Structure-based analysis of the effects of camptothecin on the activities of human topoisomerase I. Ann. N Y Acad. Sci. *922*, 56–64.
43. Bjornsti, M.A., Benedetti, P., Viglianti, G.A., and Wang, J.C. (1989). Expression of human DNA topoisomerase I in yeast cells lacking yeast DNA topoisomerase I: restoration of sensitivity of the cells to the antitumor drug camptothecin. Cancer Res. *49*, 6318–6323.
44. Andoh, T., Ishii, K., Suzuki, Y., Ikegami, Y., Kusunoki, Y., Takemoto, Y., and Okada, K. (1987). Characterization of a mammalian mutant with a camptothecin-resistant DNA topoisomerase I. Proc. Natl. Acad. Sci. USA *84*, 5565–5569.
45. D'Arpa, P., Beardmore, C., and Liu, L.F. (1990). Involvement of nucleic acid synthesis in cell killing mechanisms of topoisomerase poisons. Cancer Res. *50*, 6919–6924.
46. Hsiang, Y.H., Lihou, M.G., and Liu, L.F. (1989). Arrest of replication forks by drug-stabilized topoisomerase I-DNA cleavable complexes as a mechanism of cell killing by camptothecin. Cancer Res. *49*, 5077–5082.
47. Traganos, F., Seiter, K., Feldman, E., Halicka, H.D., and Darzynkiewicz, Z. (1996). Induction of apoptosis by camptothecin and topotecan. Ann. N Y Acad. Sci. *803*, 101–110.
48. Jacobs, J.P., Garrett, A.J., and Merton, R. (1979). Characteristics of a serially propagated human diploid cell designated MRC-9. J. Biol. Stand. *7*, 113–122.
49. Petijean, M. (1992). Applications of the radius-diameter diagram to the classification of topological and geometric shapes of chemical compounds. J. Chem. Inf. Comput. Sci. *32*, 331–337.
50. Knoben, J.E., Scott, G.R., and Tonelli, R.J. (1990). An overview of the FDA publication Approved Drug Products with Therapeutic Equivalence Evaluations. Am. J. Hosp. Pharm. *47*, 2696–2700.