An efficient rapid system for profiling the cellular activities of molecular libraries

Jonathan S. Melnick,*‡ Jeff Janes†‡, Sungjoon Kim‡§, Jim Y. Chang‡, Daniel G. Sipes‡, Drew Gunderson‡, Laura Jarnes‡, Jason T. Matzen‡, Michael E. Garcia‡, Tami L. Hood‡, Ronak Beigi‡, Gang Xia‡, Richard A. Harig‡, Hayk Asatryan‡, S. Frank Yan‡, Yingyao Zhou‡, Xiang-Ju Gu‡, Alham Saadat‡, Vicki Zhou‡, Frederick J. King‡, Christopher M. Shaw‡, Andrew I. Su‡, Robert Downs‡, Nathanael S. Gray‡, Peter G. Schultz‡§, Markus Warmuth‡§, and Jeremy S. Caldwell‡§

*Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, CA 92121; ‡The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037

Contributed by Peter G. Schultz, December 29, 2005

Rapid quantitative methods for characterizing small molecules, peptides, proteins, or RNAs in a broad array of cellular assays would allow one to discover new biological activities associated with these molecules and also provide a more comprehensive profile of drug candidates early in the drug development process. Here we describe a robotic system, termed the automated compound profiler, capable of both propagating a large number of cell lines in parallel and assaying large collections of molecules simultaneously against a matrix of cellular assays in a highly reproducible manner. To illustrate its utility, we have characterized a set of 1,400 kinase inhibitors in a panel of 35 activated tyrosine-kinase-dependent cellular assays in dose–response format in a single experiment. Analysis of the resulting multidimensional dataset revealed subclusters of both inhibitors and kinases with closely correlated activities. The approach also identified activities for the p38 inhibitor BIRB796 and the dual src/abl inhibitor BMS-354825 and exposed the expected side activities for Glivec/STI571, including cellular inhibition of c-kit and platelet-derived growth factor receptor. This methodology provides a powerful tool for unraveling the cellular biology and molecular pharmacology of both naturally occurring and synthetic chemical diversity.

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

Abbreviations: ACP, automated compound profiling; SAR, structure–activity relationship.

1J.S.M., J.J., and S.K. contributed equally to this work.
2To whom correspondence may be addressed. E-mail: schultz@scripps.edu, mwarmuth@gnf.org, or jcalwell@gnf.org.

© 2006 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0511292103

PNAS | February 28, 2006 | vol. 103 | no. 9 | 3153–3158

CELL BIOLOGY
kinase oncogenes, e.g., Bcr-Abl, Flt3, and NPM-Alk (13–15). This rendered IL-3-independent by transformation with known tyrosine kinases (Abl, NTRK3, PDGFR, and Jak2) (9–12), suggesting that the fusion partner for tyrosine kinases is ETV6

Several chimeric tyrosine kinases, including Bcr-Abl, NPM-Alk, and ETV6-NTRK3, have been described and shown to be causative leading to the juxtaposition of a fusion partner to a kinase (8).

In vitro kinase-directed heterocycles. Although parallel interrogations of the kinome against chemical space have been performed by using in vitro biochemical and phage-display assays, the cell-based format described here assays the physiological conformation of the kinase in the presence of other cellular components, cell permeability, and nonselective cellular toxicity (6, 7). To profile tyrosine kinases in a cell-based format, we made use of the well established fact that kinases can be constitutively activated by genomic rearrangements leading to the juxtaposition of a fusion partner to a kinase (8). Several chimeric tyrosine kinases, including Bcr-Abl, NPM-Alk, and ETV6-NTRK3, have been described and shown to be causative to human cancer and hematopoietic malignancies. Frequently found fusion partners include structural proteins as well as transcription factors or genes of unknown functions. The most frequent fusion partner for tyrosine kinases is ETV6/Tel, a gene that has been found in chimeras with both cytotoxic and receptor tyrosine kinases (Abl, NTRK3, PDGFR, and Jak2) (9–12), suggesting that fusion to ETV6/Tel might be a generally applicable strategy to activate tyrosine kinases.

To generate a cDNA library of kinases fused to ETV6/Tel, a retroviral plasmid based on a pMSCV backbone was constructed that allows rapid in-frame cloning of a kinase domain upstream of a Tel cassette and downstream of a Myc tag using a Gateway cassette system (Fig. 2a). A total of 81 Tel fusion vectors representing individual kinases, as well as the control non-Tel fusion Bcr-Abl, were established and used to transform Ba/F3 cells as described in Material and Methods. Ba/F3 is a hematopoietic cell line that depends on IL-3 for proliferation and survival but is rendered IL-3-independent by transformation with known tyrosine kinase oncogenes, e.g., Bcr-Abl, Flt3, and NPM-Alk (13–15). This cell line model has frequently been used to demonstrate the differential cytotoxicity of kinase inhibitors against transformed IL-3-independent cells versus parental Ba/F3 cells grown in the presence of IL-3 (16–18).

After selection for puromycin and IL-3 withdrawal, 35 of the 81 fusion kinase constructs gave rise to unique cell populations exhibiting growth factor independence. Fig. 2b summarizes the identity of the kinases found to induce growth factor independence in Ba/F3 cells. All cells were further validated as described (Materials and Methods) by Western blot and RT-PCR/sequencing of the respective fusion and by using sets of reference compounds for each individual kinase. In the latter case, two independently generated cultures of IL-3-independent cells were required to show similar responses to reference compounds. To determine whether kinase inhibitor activity was accurately reflected by the Tel-kinase-transformed Ba/F3 cells, a collection of ~80 tyrosine kinase inhibitors with known activities and selectivities was tested in dose–response format against the cell panel. For example, the Sugen c-Met inhibitor PHA-665752 inhibited the growth of Ba/F3/Tel-Met with an IC50 of 80 nM and showed no toxicity toward the rest of the panel, in agreement with published results (19).

Likewise, the kinase-insert domain receptor (KDR) inhibitor AAL993 (20) shows selective nanomolar potency against Ba/F3/Tel-KDR and the homologous VEGF receptor family member FLT4. Known BCR-ABL inhibitors, STI571 and AMN107, were also tested against the Ba/F3/Tel-TK panel and a panel of recombinant enzymes. As shown in Fig. 2c, both STI571 and AMN107 inhibited Bcr-Abl, c-Kit, and PDGF-RA/B with IC50s between 1 and 100 nM as expected [data are shown as percent activity of the enzyme versus control at a single concentration of 10 μM and was compared to the inhibitory concentrations (IC50) generated by Ba/F3 cell proliferation assays]. In addition, AMN107 also inhibited the proliferation of Ba/F3 cells expressing Tel fusions of the kinase inhibitors EphB1, B2, and B4, suggesting that the increased potency of AMN107 compared with STI571 results in somewhat decreased selectivity. Interestingly, several discrepancies between the cellular and enzymatic assays were found. For example, although very potent on both PDGF-RA/B in cellular assays, both STI571 and AMN107 only partially inhibited PDGF-RA/B in the enzymatic assays. A similar observation was made for EphB4 and AMN107. This is best explained by the fact that both inhibitors were shown to bind the inactive conformations of kinases, which might constitute only a fraction of the enzyme preparations used. Finally, some of the activities seen in the enzymatic assays could not be reproduced using cellular assays. IC50 measurements for these enzymes would be required to draw more detailed conclusions.

Kinase Profiling Experiment. A chemical library totaling 1,400 unique small molecules targeting tyrosine kinases was selected that includes purines, pyrimidines, benzimidazoles, and quinazolines (21). The library was supplemented with an additional 10 known tyrosine kinase inhibitors as controls and plated into 1,536-well microtiter plates prediluted into dose–response format ranging from 10 μM down to 3 nM. Each individual population of Ba/F3-TK cells was seeded on the automated tissue culture station, propagated for two to three passages; checked for cell viability, growth rates, and cell density; and then plated into 1,536-well assay plates on the robotic platform. The preplated compounds were transferred to assay plates in register (1:1) by using a 1,536-pin transfer device, incubated overnight, then assayed for cell viability as described in Materials and Methods. The experiment was repeated in triplicate resulting in ~1.5 million data points. To compare the data quality generated on the robotic system with data generated manually with workstations, a subset (936) of the compounds was tested against 20 of the Ba/F3-TK cells in 384-well assay plate format in duplicate. Analysis of the resulting datasets showed that for both the ACP and the manual method, >95% of the data had coefficients of variation of <10%. By defining outliers as those
IC₅₀s with >3-fold or less than one-third of the average IC₅₀, only two outliers of 936 data points were observed in the ACP dataset relative to the manual dataset.

**Global Profile Data Analysis.** An examination of the 1,400 compounds tested demonstrated that only 30 (2.1%) were toxic to wild-type IL-3-dependent Ba/F₃ cells. Interestingly, 282 (20.1%) compounds did not affect the activity of any TK-dependent cell line; however, every kinase on the panel was inhibited by at least one compound. Finally, 83 (5.9%) compounds selectively inhibited a single kinase in the panel.

Typically, as the potency of a compound increases, parallel gains in selectivity occur (22, 23). Regression analysis was used to determine whether the profiling data are consistent with this premise. Compounds were classified according to specificity by counting the number of the 36 assays in which each compound displayed a 50% growth inhibition (ГI₅₀) < 10 μM, giving a “non-specificity count.” Every test point for which the ГI₅₀ was <10 μM was plotted with the negative log of the ГI₅₀ on the ordinate and the non-specificity count of that compound on the abscissa. Although the global dataset of 935 nontoxic compounds was uninformative, inspection of clusters of structurally related compounds revealed 9 of 14 classes that showed a modest correlation between increases in potency and selectivity (Fig. 3a).

Next, we asked whether chemical similarity was a predictor of biological activity within this dataset. For each pair of compounds, the chemical similarity was computed by the Tanimoto similarity of 512-bit Daylight fingerprints. The similarity in biological response between two compounds was calculated as the ordinary Pearson correlation of the vectors composed of the log₁₀ GI₅₀ values across the 36 assays. The plot of these two metrics displays a very strong relationship between the similarity in chemical structure and similarity in biological activity (Fig. 3b).

**Fig. 2.** Ba/F₃ construction. (a) Retroviral construct for Ba/F₃-TK library. pTK is based on the pMSCV (Clontech) backbone. The first 1 kb of ETV6, with an artificial myristoylation sequence (mTel), is located upstream of the Gateway reading frame B cassette (Invitrogen), which is followed by the Myc tag sequence EKQISEEDL and a stop codon. This is linked by the HCV IRES to puromycin in which the stop codon has been replaced by the luciferase coding sequence from pLucN2 (Clontech). The expressed protein, a functional dimer, is detailed to the left. Dimerization of the mTel domains bring the two fused kinase domains into proximity, leading to autophosphorylation and activation. (b) The tyrosine kinase. The tyrosine kinase-dependent Ba/F₃ lines used in this screen are highlighted in green boxes against a phylogeny of all 90 tyrosine kinases reproduced courtesy of Cell Signaling Technology (Danvers, MA). (c) Cellular and biochemical assays. The Bcr-Abl-targeted kinase inhibitors Gilvec and AMN107 were tested against the kinase cell line panel (Ba/F₃) and the purified enzyme (in vitro). Data for Ba/F₃ are listed as the IC₅₀ in μM, whereas the in vitro data are listed as the percent enzymatic activity remaining at 10 μM compound. The boxes are colored in green (potent inhibition), black (mild inhibition), or red (little to no inhibition).
between chemical and biological similarity is strong, it is clearly nonlinear and noisy. One source of this nonlinearity is the folded nature of Daylight fingerprint bitmaps, which causes the similarity for unrelated compounds to cluster around a Tanimoto coefficient of 0.5. The most interesting outliers are those that have a high chemical but a low biological similarity (enclosed by the triangle in Fig. 3b). An inspection of the biological profiles of these outliers reveals three general classifications. There are outliers in which the biological profile vector has low variance for one or both compounds in the pair, usually because the compound has little or no activity in all of the kinase assays. Such low variance causes correlation-based distance measures to be brittle, responding dramatically to slight changes in the measured GI50 for a single assay. Another group of outliers are compound pairs in which a small structural change leads to a slight general cytotoxicity. Because this cytotoxicity is reflected in the GI50 for all 36 kinase assays, the cumulative effect is to produce large differences in biological profile. Finally, there are a smaller number of outliers that appear to be genuine exceptions to the SAR hypothesis, in which small changes of chemical structure lead to large changes in biological profile. These three categories are comined in Fig. 3b, and inspection of the individual profiles (not shown, but analogous to those in Figs. 2c and 4b) is necessary to distinguish them.

An SAR dendrogram was created to relate kinase similarity as a function of compound activity. Distances between kinase pairs in “profile space” were calculated as the Euclidean distance between the vectors composed of the pGI50 values of the 935 nontoxic compounds, with inactive compounds being assigned a surrogate GI50 of 10 μM. This distance matrix was subjected to agglomerative clustering by using the complete linkage method. The data are represented as a dendrogram (Fig. 5, which is published as supporting information on the PNAS web site). As expected, highly homologous kinases are most often inhibited comparably by small molecules. The most notable exception is the close proximity of Flt3 to the Trk kinases in the SAR dendrogram versus the sequence-based tree. This is likely because these kinases share a critical “gatekeeper” phenylalanine side chain in the ATP-binding center (see Fig. 2b). In general, the SAR dendrograms can be used as tools to guide knowledge-based target selection and analyze multiparameter datasets obtained from compound profiling. Targets in close proximity on the dendrogram are more likely to exhibit comparable inhibition by a particular group of inhibitors than targets that are more distant. The dendrogram can also be used to select and pursue targets that are inhibited by similar chemotypes.

Previously Uncharacterized Activities. The data were next scanned to identify whether any known drugs and well characterized compounds from the chemical collection displayed any previously unknown activities. The data were configured into a heat map by hierarchical clustering for ease of analysis (Fig. 4a). The data were examined to find targets for existing molecules and to understand the activity of molecules on related targets. For Gleevec, activity against Bcr-Abl, platelet-derived growth factor, and c-kit were all manifest (see Fig. 4b). Interesting cross-activities for other less well characterized kinase inhibitors currently under clinical investigation were also discovered. For example, BIRB796, a kinase inhibitor for all p38 kinase isomers (24, 25), was also shown to inhibit Tie2 and to a lesser extent Tie1 (Fig. 4b). These activities were confirmed to be target-related and not dependent on inhibition of p38 by Western blot of both Tel-Tie1 and Tel-Tie2 immunoprecipitated from Ba/F3 cells and probed with antiphosphotyrosine antibodies (data not shown). Tie2 has recently been implicated in angiogenesis as well as stem cell quiescence and mobilization during chemotherapy (24–27). Therefore, expansion to Tie2 as a target could potentially increase the utility of this compound.

Although BIRB796 displayed high selectivity versus other tyrosine kinases in our experiment, BMS-354825, a dual inhibitor of Src and Abl (28), was shown to inhibit multiple other kinases, including several Ephrin receptors (B1, B2, and B4). Ephrin receptors have been implicated in both tumor angiogenesis and growth and survival of tumor cells (29–31). BMS-354825 is currently being tested in humans for its potential to overcome Gleevec-resistant Bcr-Abl-positive chronic myelogenous and acute lymphoblastic leukemia (32). Our results indicate the potential for expansion into other tumor types. In summary, these data demonstrate the power of the aforementioned ACP in identifying novel indications for known drugs and drug candidates, a fact that will
become increasingly important given that market sizes for molecularly targeted therapies based on genetic lesions might generally be rather small.

**Conclusion**

Herein, we have profiled a collection of 1,400 small-molecule kinase inhibitors in a dose–response format against an array of 35 cell-based tyrosine kinase assays in a single experiment using a highly efficient profiling technology. The ACP provides a mechanism for systematic 2D combinatorial screening of chemical space against biological space. Automated kinase profiling expands SAR from one target against a set of compounds to many targets against many compounds, thus providing a more comprehensive dataset. In the case of the kinome, this type of information facilitates the process of “kinase hopping” to determine which scaffolds are most likely to have activity on new targets of interest.

Opportunistic cellular profiling is the preclinical corollary to serendipitous clinical profiling, which led to discoveries such as Viagra’s use in erectile dysfunction (despite its originally intended use in heart disease) or the demonstration that cholesterol synthesis...
inhibitors, statins, reduce CD69 T cell antigen levels in T cells, which may extend the statin’s benefits to immune regulation (33–35). As demonstrated here, the ACP experiment identified PDGFR and c-kit as side activities for Glivec, which are under investigation for alternative treatments of asthma and gastrointestinal stromal disorder (GIST). Similarly, the activities identified for the p38 kinase inhibitor BIRB796 and dual src/abl inhibitor BMS-354825 may prove useful as tools to validate Tie2 and the Ephrins as drug targets in angiogenesis.

ACP profiling of molecular libraries against diverse cellular assays can be applied to many other problems as well. For example, it may be possible to identify novel ligands for whole panels of orphan G-protein-coupled receptors by profiling collections of diverse lipid, metabolite, and neuropeptide hormone libraries. It may also be possible to identify combinations of drugs that act synergistically against panels of patient-derived tumor cell lines. For pharmacogenomics, disease-associated SNPs identified by haplotype mapping can be engineered into SNP-dependent cellular assays and profiled against panels of preclinical drug candidates to prospectively match patient variants with treatment. The configuration of the ACP also allows screens for ligands with enhanced potency, selectivity, or expression levels evolved from evolved protein libraries.

Materials and Methods

Automated Cellular Assays. The assay portion of the ACP consists of the same control software as the automated tissue culture station, a noncontact 1,536-well reagent dispenser (dispense range of 250 nl to 5 μl at 5% coefficient of variation (CV), dispense time of <1 min/plate at 5 μl), a 1,536-pin transfer device (validated at 9 nl and >10% CVs), room-temperature hotels, the same incubators as described above, and a Perkin–Elmer ViewLux plate reader. To request cells for an assay, the operator inputs the desired final number of cells per well. The automated tissue culture station calculates the correct number of daughter flasks for pooling (if the number of cells required is unavailable, the cell propagation protocol is enacted). When daughter flasks have incubated for a defined amount of time, they are pooled in a matrix fashion into several empty recipient flasks. An equal volume from each daughter cell flask will be deposited into each recipient flask. A sample from one pooled flask will be counted (only one sample is necessary, because all of the flasks have been adjusted to the same cell density by the pooling process). All of the flasks will have their volume adjusted to the desired cell density for plating. This pooling function, although complex, obviates the requirement for a common pooling container, thus decreasing the possibility of contamination. The flasks are then moved to the cell dispenser, where cells are pulled from flasks and dispensed into 1,536-well assay plates, typically 5 μl/well. The plates of cells are placed into an incubator and the empty flasks discarded. For the assay described in this paper, compounds are introduced to the system in 1,536-well plates through the room-temperature hotels. Cell plates are requested at 10^5 cells/ml, 5 μl/well, in 1,536-well plates, for each of the 35 Ba/F3 cell lines. The cell requests are processed as described above. After cell plating, preplated compounds in single-dose or dose–response format are transferred from the compound plates to the assay plates by using the 1,536 pintool. The assay plates are incubated in the environmentally controlled (37°C, 95% humidity, 5% CO2) incubators for 48 h (a custom lid design minimizes edge effects from evaporation and enables 5-μl cell-based assays to be incubated for up to 5 days). After incubation, 5 μl/well Cell Titer Glo (Promega) is added with well Cell Titer Glo (Promega) is added with

Supporting Text. Protocols detailing the automatic propagation of cell lines using the ACP, as well as the construction of the Ba/F3/Tel-kinase library, are presented in Supporting Text, which is published as supporting information on the PNAS web site.