Compound Management for Quantitative High-Throughput Screening

*Adam Yasgar, Paul Shinn, Ajit Jadhav, Douglas Auld, Sam Michael, Wei Zheng, Christopher P. Austin, James Inglese, and Anton Simeonov**
NIH Chemical Genomics Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD

Keywords:
screening,
qHTS,
inter-plate titrations,
serial dilution,
concentration–response curve,
dose–response curve,
compound registration,
liquid handling,
automation,
cherry picking,
follow-up

An efficient and versatile Compound Management operation is essential for the success of all downstream processes in high-throughput screening (HTS) and small molecule lead development. Staff, equipment, and processes need to be not only reliable, but remain flexible and prepared to incorporate paradigm changes. In the present report, we describe a system and associated processes that enable handling of compounds for both screening and follow-up purposes at the NIH Chemical Genomics Center (NCGC), a recently established HTS and probe development center within the Molecular Libraries Initiative of the NIH Roadmap. Our screening process, termed quantitative HTS (qHTS), involves assaying the complete compound library, currently containing >200,000 members, at a series of dilutions to construct a full concentration–response profile. As such, Compound Management at the NCGC has been uniquely tasked to prepare, store, register, and track a vertically developed plate dilution series (i.e., inter-plate titrations) in the 384-well format. These are compressed into a series of 1536-well plates and are registered to track all subsequent plate storage. Here, we present details on the selection of equipment to enable automated, reliable, and parallel compound manipulation in 384- and 1536-well formats, protocols for preparation of inter-plate dilution series for qHTS, as well as qHTS-specific processes and issues. (JALA 2008;13:79–89)

*Correspondence: Anton Simeonov, Ph.D., NIH Chemical Genomics Center, 9800 Medical Center Dr., Bethesda, MD 20892, USA; Phone: +1.301.217.5721; Fax: +1.301.217.5736; E-mail: asimeono@mail.nih.gov

1535-5535/$32.00

Copyright © 2008 by The Association for Laboratory Automation

doi:10.1016/j.jala.2007.12.004

Introduction

In high-throughput screening (HTS), the historical practice of using a single compound concentration for the primary screen has been associated with a high proportion of false positives.[1] To minimize the number of false positives, we recently introduced quantitative HTS (qHTS) where concentration–response profiles, as opposed to single-concentration data points, are generated at the primary-screen level for each library compound.[2–6] To enable qHTS, compound libraries have to be prepared as a dilution series. Two conceptually different schemes exist for generating such titrated sample sets. The first is the intra-plate method, which places all concentration points for a given compound within the same plate. The second is the "vertical" or inter-plate method, placing all concentration points for a given compound on different plates. Under the vertical, or inter-plate dilution method selected by the NIH Chemical Genomics Center (NCGC) for its qHTS operations, the first plate contains the highest concentration of a set of compounds, or sublibraries, whereas subsequent plates contain the same compounds in the same well locations, but at successive lower concentrations. In this manner, a dilution series of plates can be generated where multiple copies of the same library are made with each copy differing only in the concentration of library members

and where the number of copies is determined by the number of titration points desired for the dilution series. This approach to concentration-series plating is superior to the intra-plate method because it offers increased flexibility in plate usage for screening a wide variety of assay systems, and ease of plate preparation. In vertically developed multiconcentration-compound collections, the user is free to choose between screening all concentrations and selecting dilution plates in a random-access manner. For example, by screening only selected plates, the concentration—response curve can be constructed to favor a certain concentration range, as necessitated by the assay biology and/or reagent costs.

Compound Management at NCGC operates in a unique environment with respect to its innovative screening process and the input and output of material. First, the application of qHTS means that each compound does not reside in a single well, but is represented by (at least) seven samples. Thus, a 200,000-compound collection equates to approximately 1.4 million wells, a size comparable to a large pharmaceutical company.[7] Second, screening libraries are not supplied from one centralized vendor source but instead arrive via widely different mechanisms. Third, Compound Management is responsible not only for the preparation of uniform screening libraries for qHTS but also for reacquiring samples from a multitude of suppliers and collaborators for the purposes of follow-up experiments. Herein, we present the equipment and processes implemented at NCGC to acquire, register, and plate compounds to support qHTS-based screening and follow-up.

Materials and Methods

Compound Receipt and Initial Processing

The processing of an incoming compound involves both physical and virtual components (Fig. 1). Before any physical manipulation of the compounds can occur, they are registered using ActivityBase (IDBS, Guildford, UK) to auto-generate unique identifiers. A structure data (SD) file containing the minimum compound structure, source of compound, and the source sample identifier is typically used on initial registration. ActivityBase is used to manage salts and solvates table[8] for salt stripping or the addition of any new salts that may be encountered during registration. Batch numbering of samples is automatically incremented by ActivityBase for stereoselective and duplicated compounds. Additional information on compounds such as analytical data is added to the database once registration is complete. Samples containing unknown structure or large mixtures of compounds are registered as samples and then tracked throughout all subsequent processes using unique identifiers. This allows screening in qHTS mode for any type of sample ranging from purified compounds to large mixtures of combinatorial libraries. An individual compound or sample can be present in many different physical locations, starting with the initial container provided by the supplier to a microtiter plate present on the online screening system. Every location of a particular sample is stored and tracked in ActivityBase.

Compounds are received either in solid state contained in individual vials or 96-well tube racks, or as solutions in vials, tubes, or plates (Figs. 1 and 2). In the rare cases of special packaging such as ampoules or septum-capped vials, the respective samples are solubilized through manual methods. A barcode label with an NCGC ID is affixed to the corresponding vial before solubilization. In general, the entire sample is dissolved and transferred to a tube, or several tubes, if the quantity is less than 10 mg. Whenever possible, samples are requested to be sent in 2D-barcoded Matrix tubes (1.4 mL size; Matrix/Thermo Scientific, Hudson, NH) and racks, to maintain uniform processing of containers. The volume of DMSO needed to produce a 10-mM solution is then pipetted into the containers. The containers are recapped, inverted, and vortexed to capture any powder that might have adhered to the cap or side of the tube, and centrifuged to 1000 rpm (or $145 \times g$, Eppendorf 5810R) for 1 min. Although most powder samples dissolved after brief vortexing, visual inspection of the tubes was implemented to identify and segregate mixtures containing undissolved material. Tubes containing such mixtures were subjected to sonication treatment for up to 10 min to complete the dissolution process.

Solutions from vials and other types of containers are pipette-transferred into the same type of 2D-barcoded 1.4-mL Matrix tubes. If the solution volume exceeds 750 μL, the excess is transferred to another 2D-barcoded tube and the multiple instances of the compound in tube format are recorded. The tracking of the 2D barcodes is performed using a BioMicrolab 2D Scanner or BioMicrolab XL20 Tube Handler (BioMicroLab, Concord, CA). Each 1.4-mL tube is then placed into a 96-tube rack, with four of these racks representing one quadrant of a 1536-well plate, or a total of 16 racks corresponding to one 1536-well plate (Fig. 3). The NCGC ID, plate and position in the Matrix rack, and the 2D barcode are recorded using a Pocket Excel spreadsheet on a Symbol MC 50 PocketPC (Motorola, Holtsville, NY) equipped with an integrated barcode scanner.

Compound libraries shipped directly as solutions in 96- or 384-well plates are centrifuged and the plates are heat sealed (using PlateLoc Thermal Plate Sealer equipped with BenchCel 2 × stacker system, Velocity 11, Palo Alto, CA) if necessary and stored either at room temperature or at −80 °C depending on the recommendation of the sender and timing of subsequent operations.[9–14] In the event that the 384-well plates were shipped with empty wells or columns (i.e., partially filled plates), they are condensed using Tecan Freedom Evo. The opposite process could also be performed, where in the event that a 384-well plate is received with all of its wells filled, samples contained in columns 1 and 2 can be transferred to a different plate.

Sample Compression into 384-Well Plates

Compounds in 96-tube racks are compressed into 384-well Greiner Bio-One or Matrix polypropylene plates via interleaved quadrant transfer using an Evolution P^3 (EP^3) system

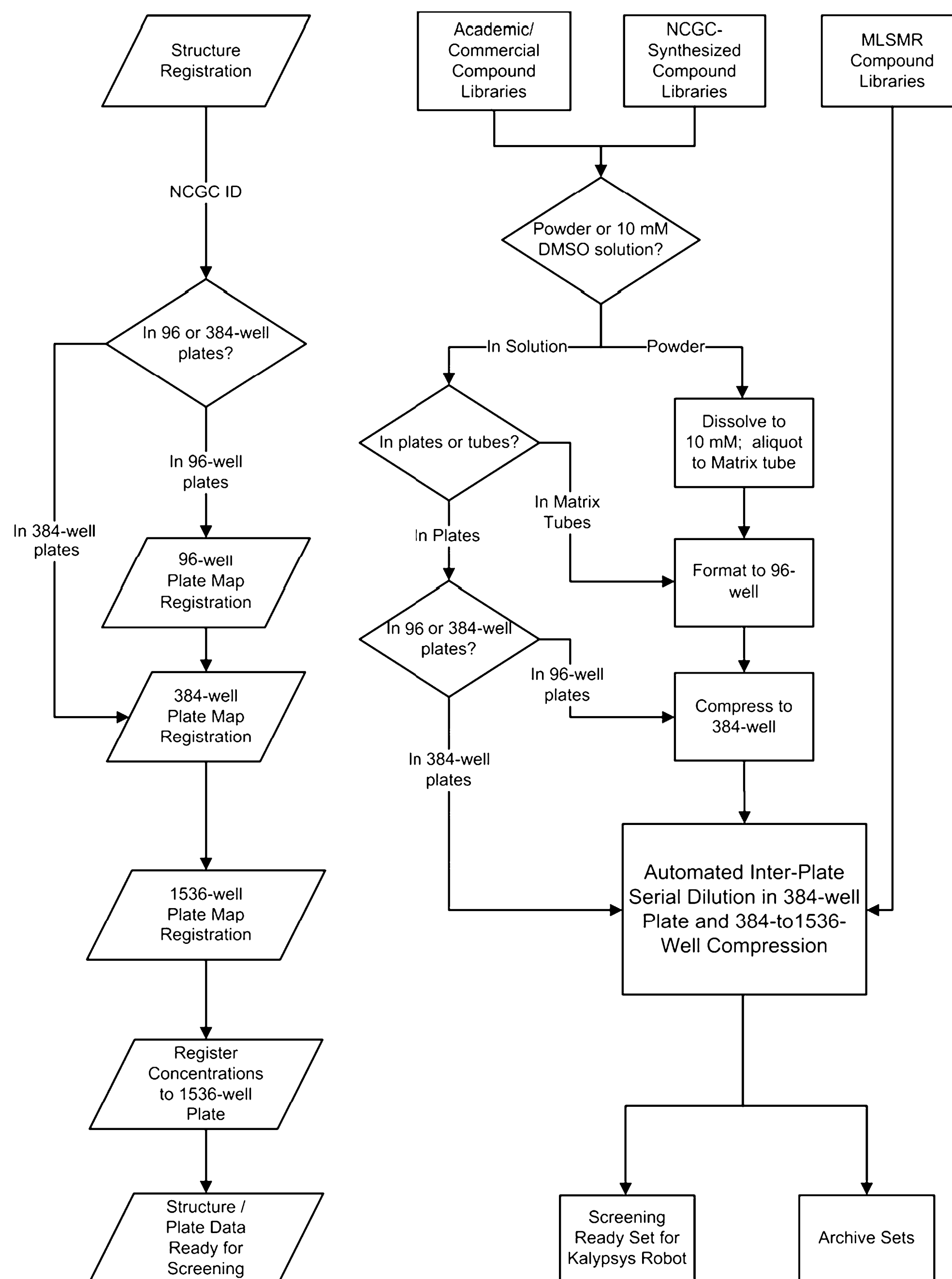

Figure 1. Compound management workflow. (*Left*) Diagram of the NIH Chemical Genomics Center's virtual process for structure and plate-map registration. (*Right*) Diagram illustrating the physical steps involved in preparation of 1536-well plate sets for quantitative HTS.

equipped with a 96-tip head and PlateStak Automated Microplate Handler (PerkinElmer Life and Analytical Sciences, Shelton, CT). The EP^3 system used here consisted of four plate stackers, 16 deck positions, and one dispense head capable of addressing all deck positions. The front left stacker (1A), front right stacker (1B), back left stacker (13A), and back right stacker (13B) are the locations for plate loading (Fig. 4). For each 96-tube rack, column 1 is left empty for control placement resulting in blank columns 1 and 2 of the final 384-well plate, the latter propagating to four empty columns in the final 1536-well plate.

The process for compression from 96-tube racks to 384-well plates involves the following steps. First, four 96-tube racks are centrifuged at 1000 rpm for 1 min. These are then placed into stacker 1A, with 96-tip boxes located in grid positions 10, 12, 14, 16, and a 384-well plate on grid position 7. The plate formatting starts with the first tube rack downstacked from 1A once its barcode is scanned; simultaneously, tips are automatically loaded onto the dispense head. The samples in the first tube rack are mixed three times by aspirating and dispensing 20 μL of solution 1 and 4 mm above the bottom of the tubes, respectively.[15] This is followed by

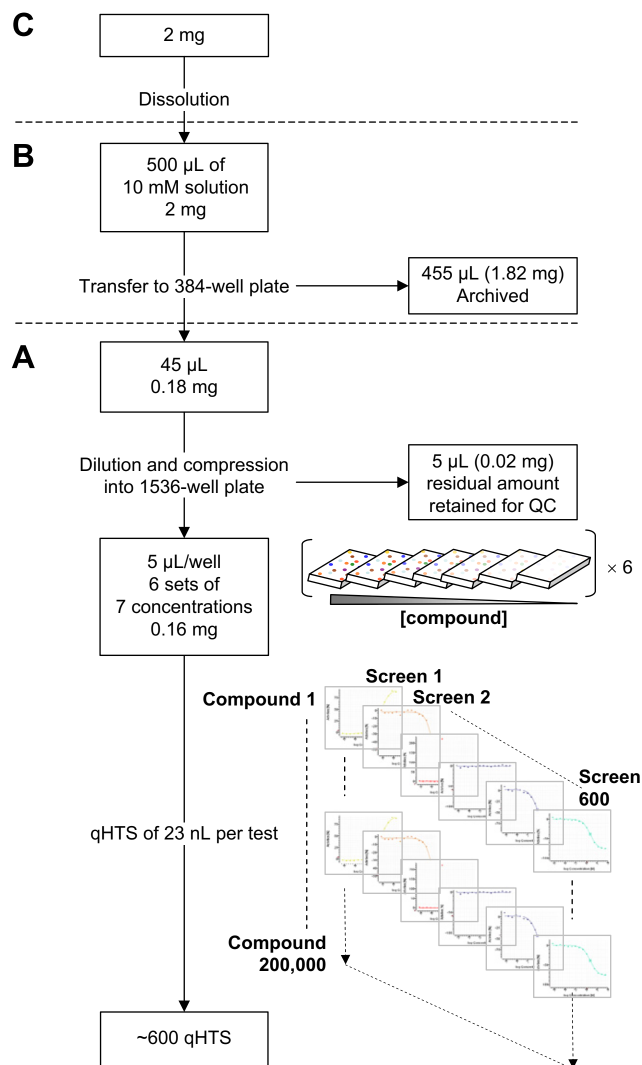

Figure 2. Compound budgeting. Shown is the utilization of a typical compound sample (45 µL of 10 mM DMSO solution equivalent to 0.18 mg) committed to a plating job for quantitative HTS (A). Of that amount, approximately 0.02 mg, or 11% remains stored and available for QC and other uses, whereas the majority is spread across six sets of seven concentrations, or a total of 42 plates, and used in several hundred screens, yielding a full concentration—response profile from each assay. Compound samples received in solution of larger volume (B) or powder (C) forms are processed such that only 45 µL of top-concentration solution are used in the above plating job, whereas the remainder is stored for future use.

a 55 µL transfer into quadrant 1 of the 384-well plate. Although a fresh set of tips is being loaded, rack 1 is upstacked to 1B, rack 2 is downstacked, barcode scanned, aspirated, dispensed, with all the steps repeated until all the remaining racks are completed. Depending on the volume of solution contained in the tubes, multiple copies of these 384-well top-concentration plates can be made and stored. The same process was performed for a total of 16 racks of 96 tubes until four 384-well top-concentration plates were prepared,

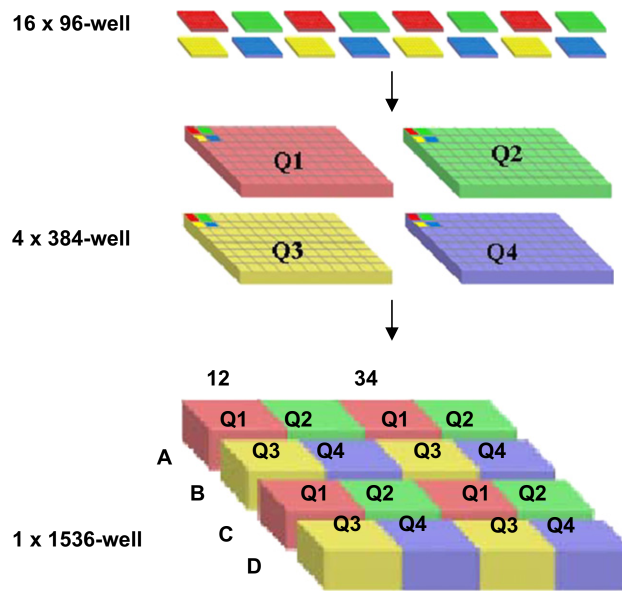

Figure 3. Interleaved quadrant compression. Definition of four-well quadrants (Q1, Q2, Q3, and Q4) within 384- and 1536-well plates.

which could then be used for the downstream dilution and 384- to 1536-compression process.

Inter-Plate Titration Series Preparation

The inter-plate titration plating into 384-well plates and subsequent compression into 1536-well plates can be executed as one fully automated protocol using an EP^3 system with integrated two plate stack units (for transport and presentation of *both* 384- and 1536-well plates), barcode

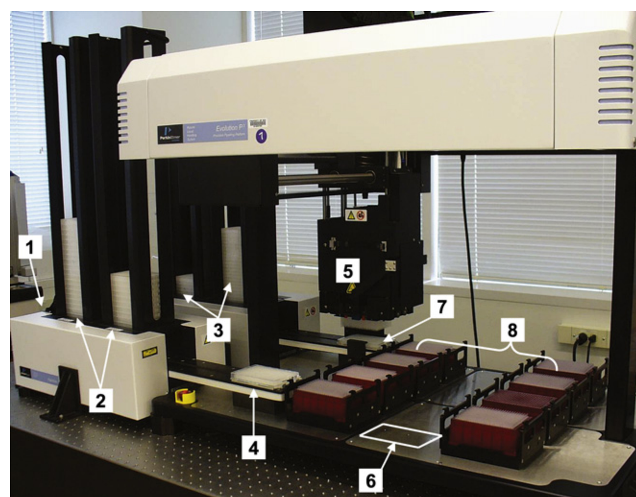

Figure 4. Layout of the Evolution P^3 dispenser used in library preparation. Indicated on the photograph are (1) Barcode scanner, (2) 384-well plate stackers 1A (left) and 1B (right), (3) 1536-well plate stackers 13A (left) and 13B (right), (4) 384-well plate stack shuttle, (5) 384-tip dispense head, (6) Example of a grid position, (7) 1536-well plate stack shuttle, and (8) 384-tip boxes.

scanners, and 384-tip dispense head (Fig. 4). The plates that will contain the lower concentrations are prefilled with 40 μL DMSO (Matrix WellMate Stacker Base) and all 28 plates, consisting of the four top-concentration plates and 24 future lower-concentration plates corresponding to the four quadrants of the 1536-well plate, are arranged by quadrants in the front stacker of the EP^3 system. The opposite stacker then contains 42 1536-well empty destination plates for creation of the six identical copies of the seven-point dilution series. The solution in the source (higher-concentration) plate is downstacked from 1A and mixed by five 20 μL aspiration and dispense cycles. Then, 10 μL of solution is aspirated from the higher-concentration plate and the plate is upstacked to 1B, the solution is dispensed into the 40 μL DMSO contained in the next concentration plate (downstacked from 1A) to achieve a five-fold dilution. The steps are then repeated until all seven 384-well plates are prepared. Tip changing occurs only once (after dispensing into the fourth-concentration plate) for each seven-plate dilution set to minimize the carryover of highest-concentration solutions to the lowest-concentration plates.

On conclusion of the serial dilution of the first quadrant, the lowest-concentration 384-well plate is downstacked from 1B, 30 μL of solution is aspirated from each well, and the plate upstacked to 1A. Simultaneously, the future first replicate of the lowest-concentration 1536-well plate is then downstacked from 13A, 5 μL of solution is dispensed into the first quadrant (Fig. 3), and the plate upstacked to 13B. The downstacking, dispense of 5 μL, and upstacking steps are then repeated until all six replicates are made. This subroutine is similarly executed for the subsequent higher-concentration sets. After the completion of a quadrant, there is a tip change, simultaneous reshuffling of the 1536-well plates from stacker 13B to 13A, and partial return of 384-well plates from stacker 1A to 1B to expose the next quadrant's source plates. The remaining three quadrants are processed as described above. On conclusion, the 1536-well plates are centrifuged at 1000 rpm for 1 min, and the five sets are heat sealed for long-term storage at −80 °C, whereas the remaining set is fitted with compound lids and made available for immediate use in screening (Fig. 2).

Plate Registration

Each set of plates is tracked using a unique ''platejob'' identifier. The plate job identifier is assigned at the time of initial reformatting on the EP^3, and used for registration and to enable tracking of qHTS inter-plate concentrations. A qHTS plate job typically consists of six copies of an inter-plate titration series, with each plate containing the same compound layout. Once a plate reformatting is completed on the EP^3, its detailed log file is processed using a macro that was designed to extract the proper concentration order of the 1536- and 384-well plates used in the physical formatting.

All plates, starting with the initial 96- and 384-well source plates and down to the individual 1536-well compound titration sets, are registered using ActivityBase batch processing tools. The processed EP^3 log file information is used to register the various copies of 384-well plate titration series. The 384-well plates' quadrant mapping into their respective 1536-well plates is performed using ActivityBase's plate reformatting tool. An in-house database is used for storing the relationships between plates, wells, compounds, and their concentration mappings. This database is updated daily by joining two sets of information: the ActivityBase plate-compound tracking information and the concentrations assigned from the EP^3 plate job processing macro.

The in-house database is an Oracle Enterprise db (Oracle, Redwood Shores, CA) that is designed specifically for qHTS data management. To address compound and plate tracking, the db schema includes tables that are created or updated through a cron job that is executed automatically on a daily basis. The cron job executes a structure query language (SQL) script using Oracle SQL*Plus in batch command line mode. First, a materialized table view is created to store up-to-date ActivityBase data on plates, wells, compounds, batches, and compound supplier information. For optimal database performance, indexes are created on plate, row, column, sample batch identifier, parent compound identifier, and supplier data. Second, a separate table is updated to insert new plate and well concentration information from the compound plating jobs. During processing of screening data, these two tables are used jointly to create qHTS concentration mappings that convert plate-well driven data into titration response activities centered on individual samples.

Follow-Up Compound Processing

For compounds within the Molecular Libraries Small Molecule Repository (MLSMR) collection, an order is placed to the MLSMR contractor online for 23 μL of 10 mM solutions to be delivered in a Matrix 384-well plate. For follow-up samples from other libraries, compound vendors such as Sigma-Aldrich, or chemical procurement companies such as ChemNavigator, are used to procure the appropriate amount of powder needed to prepare at least 250 μL of a 10-mM solution.

Follow-up confirmation samples are arrayed as 24-point titrations at 1:2 dilution steps developed within the same plate. The 24-point intra-plate titration is first prepared in 384-well plate format by dispensing 20 μL of 10 mM solution in column 1, rows 1−15 of a 384-well plate (row 16 is left empty for control placement). For diluent, 10 μL of DMSO is dispensed into columns 2−24. After centrifuging the plate, EP^3 equipped with 384-tip head loaded with 16 tips in column 1 is used to perform twofold dilutions across the plate using the same mixing, aspirating, and dispensing steps described in previous section. Tip changes are performed after every sixth dilution point, for a total of four tip changes per intra-plate dilution. The plate is centrifuged again and placed back on the EP^3 deck. Then, 7 μL is aspirated from all wells, and the contents are transferred in two 3.5-μL replicate dispenses into Q1 and Q3 of a 1536-well plate (Fig. 3). Similarly,

Q2 and Q4 are filled with the dilution series of additional compounds, sourced from another 384-well plate, for a total of up to 30 compounds per plate. The plate is centrifuged, heat sealed, and made available for confirmation studies.

Results and discussion

Library Preparation for qHTS

qHTS is a novel screening paradigm that assays all compounds in the screening collection at multiple concentrations. Concentration–response curves obtained from qHTS allow for reliable activity assignment and enable the derivation of structure–activity relationships from the primary-screen data.[3,5,6] By comparison, triaging of hits from single-concentration-based screens is typically inefficient; because the total number of samples screened (typically > 1 million) is extremely large, even superior assays with false positive rates of 0.1% (99.9% accuracy) can produce more false positive than true positive hits. This leads to erroneous assignment of activity and triggers resource-intensive follow-up activities on samples that show a poor confirmation of activity on retesting.[16] In qHTS, single "hits" that do not fit to a concentration–response curve are easily identified and discarded as inactive (Fig. 5). Therefore, qHTS resolves the active from inactive compounds with much greater accuracy and sensitivity than traditional HTS. This information can then be used to rank compounds by potency and efficacy, which enables SAR analysis using the primary data and other queries of the primary-screening database to address, for example, the selectivity profile of the compounds across other qHTS assays.

In addition, qHTS is operationally a more robust process. Inter-plate titrations allow risk of equipment malfunction to be spread over several plates and thus make the screen refractory to single-plate losses. During the curve-fitting step of data analysis, occasional assay plates exhibiting high levels of data variation and/or unexplained positional drifts can simply be masked from further consideration without the need for such plates to be scheduled for rescreening.[3]

Though qHTS can be performed with compounds in 384- or 1536-well plates, we prefer the 1536-well format given the

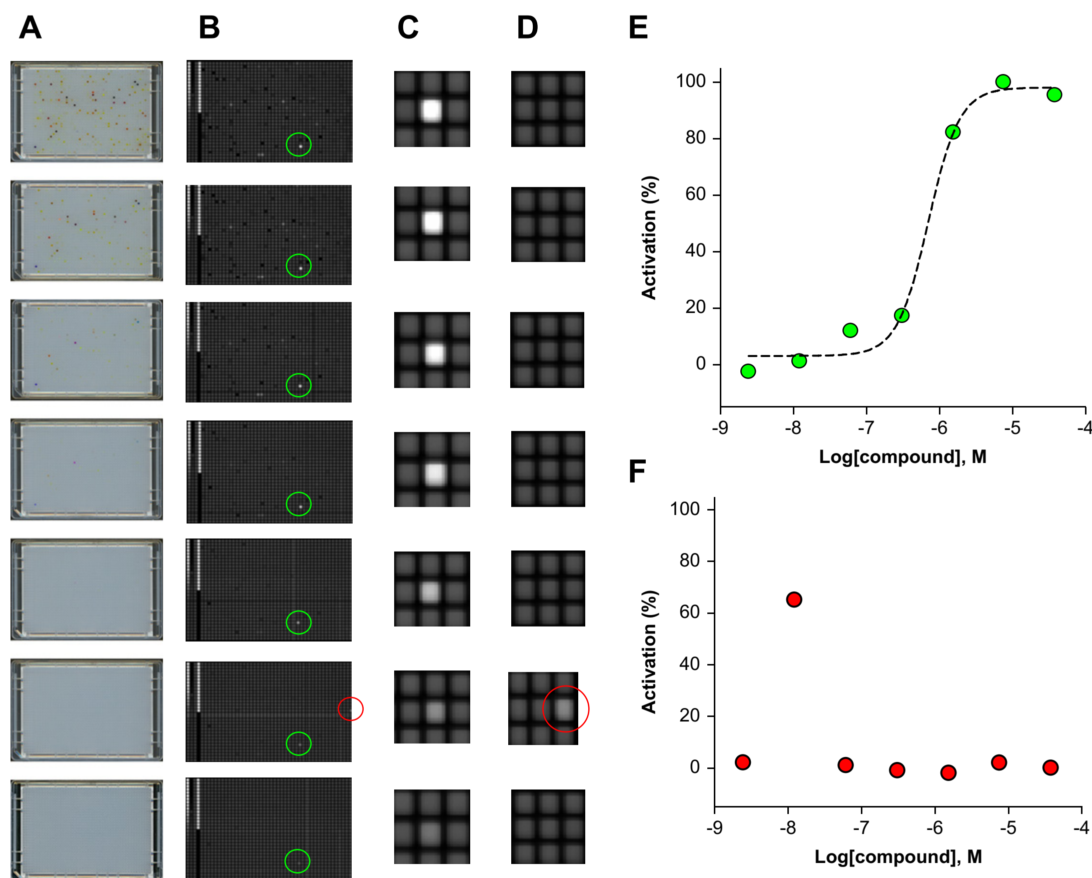

Figure 5. Inter-plate series in qHTS. A) A library set of seven vertically titrated 1536-well compound plates arranged from high to low concentration, top to bottom. B) Charge-coupled device (CCD)-based plate-reader images showing an assay response (wells circled in green) to increasing compound concentration in a vertically titrated concentration plate series. A single high-response well (circled in red) is observed that does not show a concentration–response curve (e.g., a false positive). C) Magnified images of wells exhibiting concentration–response curve. D) Magnified images of wells in the same region as the false positive well. E) and F) Concentration–response curves derived from the samples indicated in C) and D), respectively. Curve-fitting software recognizes data from D) as a flat response with a single-point outlier and classifies this as inactive, thus eliminating the sample from unnecessary follow-up analysis.

cost savings in reagents, and have found that with precise liquid handling and strict attention to automation protocols, assays that cannot be miniaturized to 1536-well format are very rare; of the first 100 projects undertaken at the NCGC, only one could not be run in 1536-well format.

To create the compound titrations for qHTS, we start with 45–60 μL of highest concentration-compound solution per well in a 384-well plate, the upper end of this range being driven by well capacity considerations and the lower end being dictated by the minimum volume needed to execute the dilution protocol. Given the fact that the majority of vendor-supplied compound collections, when shipped in solution form, are maintained as 10-mM stock solutions in DMSO, this value served as the highest-concentration point in preparing our libraries for qHTS. Pintool transfer[17] of 23 nL of 10-mM solution into an assay volume of 3–9 μL produces a final assay concentration of 25–77 μM. This concentration range is modestly higher than that generally used in single-concentration screening, where final compound concentrations are between 10 and 20 μM. Where the biological question requires testing compounds at concentrations approaching or exceeding 100 μM, this is easily achieved by performing two successive pintool transfers.[18] Though compound precipitation typically makes interpretation of such high concentration screening data difficult, qHTS-generated concentration response curves aid this interpretation, because such compounds typically show partial efficacy responses.

In developing our dilution series, we strived to maintain balance among several factors. First, we needed to include enough concentration points to allow for robust curve fitting. Second, the concentrations tested needed to cover a meaningful range, applicable to a broad range of assay detection formats and target types. Third, the number of concentration points had to be kept to a minimum due to typical limitations associated with creation, storage, and handling resources. Last, the dilution factor needed to be such that the library preparation could be executed in a simple, precise, and reliable manner using generally available liquid-dispensing equipment. Our chosen dilution series, consisting of seven points separated by five-fold steps met the above criteria. Depending on the final assay volume, the lowest concentration tested is between 1 and 3 nM, and the highest concentration tested between 25 and 77 μM, thus yielding a concentration series that spans a range of over four logs and covers the crucial low-micromolar and nanomolar ranges most frequently tested in search of small molecule effectors. Seven points are largely adequate for fitting a Hill equation[19] through the data, and according to a recent report may even be redundant.[20] However, seven points may be preferred for cell-based assays that typically show more variation in assay noise and activity is often associated with complex-concentration response relationships such as bell-shaped curves.[2] The five-fold dilution step is generous enough to permit the above concentration coverage with just seven points, and has an interval that provides for more than a single point per log concentration, thus ensuring that enough points are available to map the response between asymptotes and accurately calculate the Hill slope. Although "front-loading" of the library with multiple concentrations increases the number of compound plates to be stored and handled, this does not mean that the entire set of seven plates has to be screened against every assay. On the contrary, the inter-plate manner of dilution series preparation allows the user to choose the number and type of concentrations tested depending on the assay system without having to run extensive pilot studies (see section Process Robustness and Flexibility). Additionally, as noted elsewhere, resources invested in library preparation and testing in concentration–response format are largely recouped by decreasing the load on cherry picking and confirmation activities.[2,3,5,6]

The volume of 45 μL per well (Fig. 2), while not representing the absolute minimum required, is optimal for execution of the dilution protocol described here. In this case, 10 μL of solution is used to make the next concentration point (by mixing with 40 μL DMSO) and the remaining 35 μL serve as source for the highest-concentration 1536-well plates. Thus, the preparation of six sets of plates containing 5 μL of solution per well consumes 30 μL, leaving approximately 5 μL in the well. This volume serves three purposes. First, it represents an amount needed to ensure that the 30 μL aspirated does not contain air bubbles due to minor variations in meniscus and depth of tip approach. Second, this sample can be used for retrospective QC analysis to investigate compound stability and to resolve library plating or registration issues (see example in Process Robustness and Flexibility). Third, it represents an archive amount, which can be accessed to perform limited follow-up tests should a resupply of the compound become difficult.

Production Metrics and Compound Budget

Although inter-plate titrations are prepared starting from solutions in 384-well plates, only the MLSMR collection, designed by NIH and presently handled by the NIH-selected contractor Biofocus DPI, is delivered in this ready-to-plate state (Figs. 1 and 2A). Compound acquisitions are categorized in the following three ways (Fig. 6): External (Exploratory Collection, compound libraries from vendors or academic collaborators), Internal (Exploratory Collection, probe analogues), and the NIH MLSMR. Although the MLSMR library currently represents the greatest fraction of compounds screened at NCGC, additional collections are continuously being acquired from commercial and academic sources both to increase the diversity of the screening set and to better address the previously uncharacterized targets which are frequently screened at the center.[21] Non-MLSMR libraries are delivered in various formats including but not limited to solutions in regular 96-well plates (for example, the NCI Diversity Set), solutions in deep-well plates (the LOPAC and Prestwick libraries), 0.5 mg solid per well (University of Pittsburgh CMLD), and 2 mg of solid per tube (NCGC Synthetic Chemistry Group) (Fig. 2B, C). Collections arriving in solid form are first dissolved and then formatted in

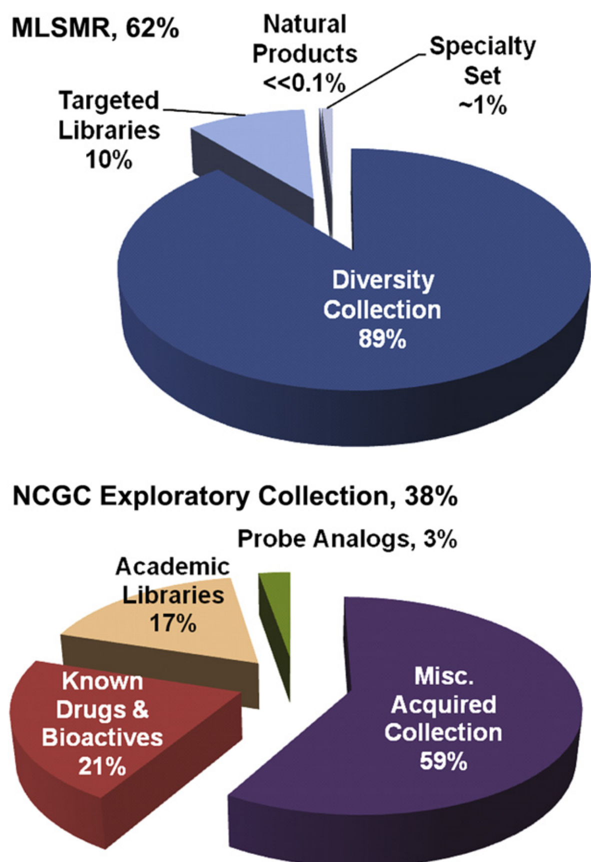

Figure 6. NIH Chemical Genomics Center Compound Library Composition. Reflected are a total of 222,183 compounds as of September 2007.

96-tube arrays and 384-well plates, whereas solution samples arriving from the majority of library vendors (such as Sigma-Aldrich, Prestwick Chemical, and Microsource) in 96-tube or deep-well format are compressed directly into 384-well plates. Thus, libraries entering our center at these upstream points require additional processing effort to be brought to the common 45 μL per well, 384-well plate state needed for qHTS library preparation (Fig. 1). Additionally, some specialized libraries are processed separately to yield a 15-point dilution series needed to generate dense data coverage for select projects (data not shown).[18]

Using the above-described processes, we have performed over 400 reformatting jobs to date, which cover the process from receipt of 96-tube racks or 96-well plates to the generation and disposition of series of 1536-well plates of seven or more concentrations. Currently, six 384- to 1536 dilution and compression jobs are performed per day, and at six final replicates being generated per library series, a total of 252 1536-well plates (or nearly 400,000 samples) are produced daily. To date, 4872 1536-well plates have been prepared, equivalent to over 7 million total samples. For a 100,000-member library to be processed from a receipt of 384-well source plates to a set of 1536-well inter-plate titrated qHTS plates, it takes one full-time equivalent (FTE) approximately 12 working days. Once a 1536-well library plate (containing 5 μL of compound solution per well) is prepared, depending on the storage conditions, approximately 100 screens can be run before the sample volume is depleted by ~50%. Taken as a whole, our plating process uses 45 μL of 10 mM starting solution, or 0.18 mg of compound (assuming a molecular weight of 400), as an input for each library member. As an output, six complete copies of the library titration series are produced and those replicate sets are typically used in screening over a period of approximately 2½ years, with each set residing on our robotic system for approximately 5 months, a period chosen to both maximize compound usage and minimize sample degradation.[7,9,11,14,22,23] The sets of plates which are heat sealed and stored at −80 °C remain stable for a relatively long period of time. The screening set of plates is stored lidded at room temperature and is continuously used in HTS for approximately 5 months, during which period we have observed only minimal differential evaporation of the edge wells (approximately 20% at the end of the entire period, data not shown). Depending on their frequency of utilization, the six replicate sets resulting from one plating job can be used to generate concentration—response profiles for every compound in up to 600 separate qHTS assays (Fig. 2).

Dispense Equipment Considerations

When evaluating instrumentation to perform the dilution and reformatting of the compound library, the most important aspect for our center was the error recovery capability of the liquid dispenser. The main rationale behind such prioritization was that while the majority of our compound libraries were generally easy to reacquire, numerous distinct subsets originated from academic laboratories where the compounds had been synthesized in limited quantities. Because these libraries often explore chemical space not addressed by commercial sources, and can be difficult to resynthesize on short notice, it was imperative that we ensure efficient utilization of these libraries. In this regard, a key feature of the EP^3 system is the ability for the operator to assume manual control over the stackers and the dispense head, should an error or crash occur. In case of such an error, one is able to terminate the run without resetting the instrument. The user is then able to manually control the position of the dispense head, the transport of an empty plate, and dispensing of compound solution remaining within the tips, to salvage the plating operation without loss of compound solution.

The task of accurately positioning the 384-tip head over the 1536-well plate presents special issues due to our requirement to use polypropylene compound storage plates that are customized for our robotic plate handling. These plates have round wells for use by the 1536-based pintool station. To maximize the height-to-volume ratio and thus enable multiple dips for repeated screening, the wells of these plates have smaller diameter than that of typical 1536-well plates and as such the array of wells becomes more difficult to address by

384-tip-head pipetting systems. Even with good initial X/Y position optimization, problems can occur when the tips are inserted approximately halfway into the wells, causing one or more of them to come in contact with the walls of the well even with a minimal (0.1 mm) shift of the head position. If enough tips become similarly lodged, the cumulative frictional force becomes large enough for the plate to be lifted with the dispense head upon its retraction, resulting in dropped plate or crashed head. In this regard, we found it advantageous to use the ''liquid tracking'' feature, in which the head retracts while the solution is dispensed, thereby greatly reducing the probability of solution spillage and tip lodging.

The ability of the EP^3 to perform parallel plate processing results in significant time and material savings. For example, as one quadrant is finishing, the system begins the dilution process for the next quadrant by loading a fresh set of tips and reshuffling the 1536-well plates. The EP^3's parallel processes decrease the duration of the plating job and also allow the plating to become ''setup and walk away,'' with no operator intervention needed after initiating the software protocol. Additionally, because we found that the instrument performed each dilution step cleanly, as evidenced by the absence of residual droplets inside the tips after dispense (data not shown), we were able to reduce the tip usage, and thereby operational costs, by limiting the tip sets needed per quadrant to a total of two.

Process Robustness and Flexibility

Although compound management may be thought of as a purely ''physical'' process, the ''virtual'' or database, aspect is equally important. Compound and plate management in qHTS has particular advantages in error recovery throughout the virtual processes of compound and plate registration and data processing. In qHTS, compound activity assignment is

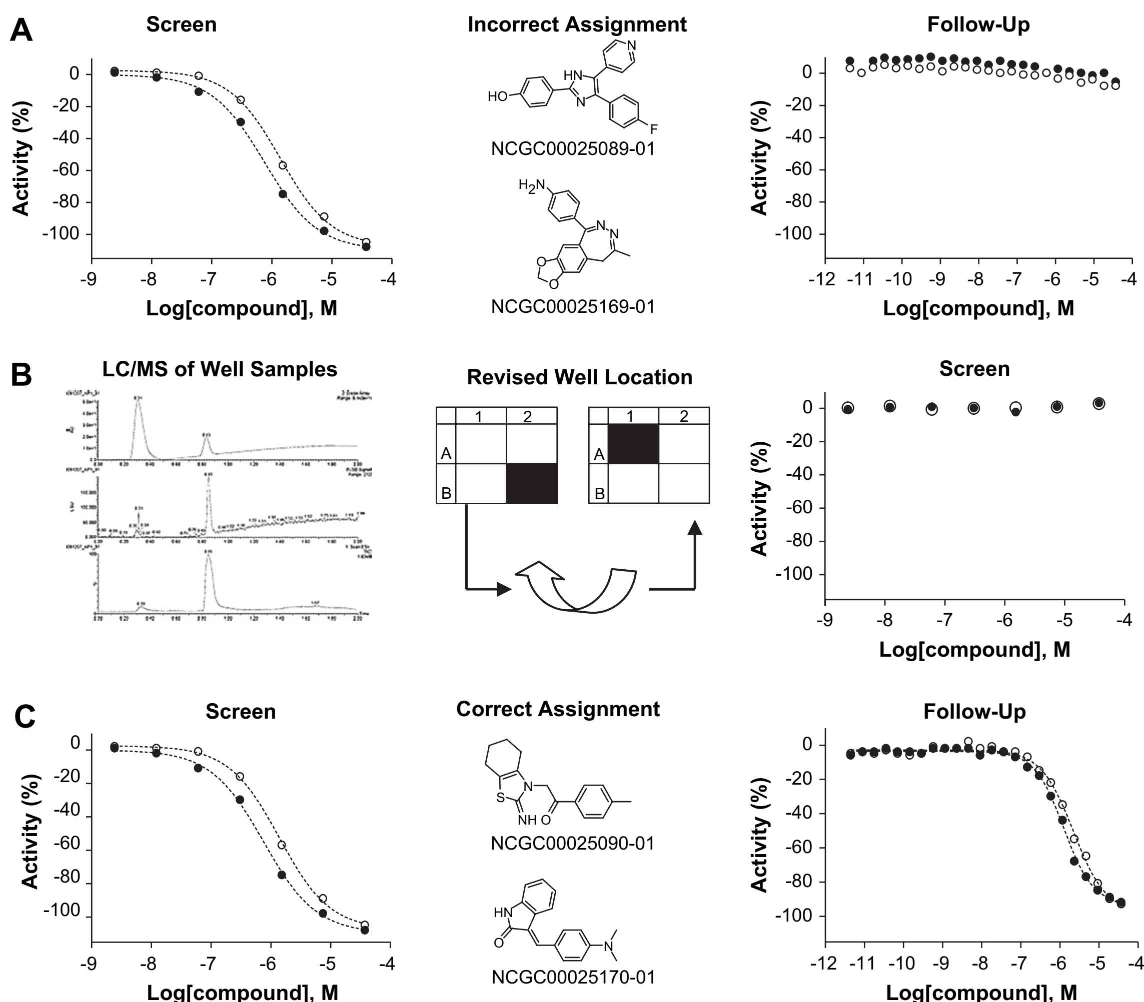

Figure 7. Registration error recovery. A) Two active compounds were identified in the quantitative HTS (qHTS) assay but showed to be inactive upon follow-up testing. B) Solutions sampled out of the 1536-well plate used in the qHTS were analyzed by LC/MS and it was identified that one of the 384-well quadrants was registered incorrectly. The plate map was revised by reregistering the quadrant. The compounds corresponding to the reassigned IDs and the data showed to be inactive, consistent with the follow-up testing. C) The compounds reassigned to the active wells were purchased and the activity was confirmed in the follow-up testing.

based on the cumulative data from the entire concentration series tested, and as such provides a more robust basis for selection of actives for follow-up. Such confidence of activity assignment is supported by a comprehensive record of compound provenance throughout the compound management process, allowing tracking of potential sources of error in the registration or plating processes should a follow-up compound fail to confirm. Because there is complete traceability of all compounds in all plates, if errors are detected compound registration may be corrected, in turn correcting the associated plate map and leading the investigator to the correct follow-up compound. This procedure creates a "closed loop system," allowing any errors to be corrected at the registration level, with the new registration acting as the input to the follow-up process and data (Fig. 1). Thus, compound registration is not static, and if corrections are made, they are propagated throughout the system. In this regard, this process should not be viewed as "re-registering" a compound ID, but rather as a plate-map correction. A major consideration in the design of the db schema was the ability to scale across potentially hundreds of millions of plate, well, compound, concentration combinations. Database performance and scalability have been achieved by balancing various factors in table design, use of Oracle's partitioning features, and effective use of materialized views. Although Oracle was used for the database technology, much of the SQL is generic and any widely used database such as Postgres (www.postgresql.org) or MySQL (www.mysql.com) can be used to manage qHTS data.

Registration errors can occur via incorrect compound information provided for a sample well or through a SD file uploading error, which creates an erroneous plate map. We encountered an error of the latter type during the analysis of follow-up experiments for compounds identified from a luciferase-based screen. Two active compounds identified from the screen (Fig. 7A, left panel) were purchased, but showed no activity in the follow-up assay (Fig. 7A, right panel). To resolve the discrepancy, the follow-up solutions were subjected to LC/MS analysis to confirm purchased compounds' identity; these showed molecular weights of 331 and 293. We then similarly analyzed the active wells from the *original* 384- and 1536-well library plates using the archived volume mentioned above, and found compounds of molecular weights of 287 and 264, an obvious contradiction with the purchased compounds that were assayed in the follow-up experiments. Examination of the log file indicated that one of the 384-well plate maps was inverted in the database (Fig. 7B, center panel). After updating the plate map, reanalysis of the data showed that the follow-up compounds had in fact been inactive in the primary qHTS (Fig. 7B, right panel), in concurrence with the follow-up results. After plate-map revision, the correct compounds associated with the active wells were purchased and the follow-up experiments confirmed the activity of observed in the qHTS (Fig. 7C). This example illustrates how the small residual volumes in all intermediate and final plates are not wasted but serve as an essential sample sources for *retrospective* library analysis, QC, and error recovery.

As stated earlier, inter-plate, or vertical, preparation of dilution series affords additional flexibility during screen design. In addition to making the screening process less susceptible to single-plate losses due to crashes and/or equipment malfunction, concentration series developed across multiple compound plates allow the implementation of asymmetric concentration–response screening, or in the extreme case, single-concentration testing. In a recently implemented screening assay, we needed to modify the dilution series to better reflect the underlying biochemistry of the assay system. The system under study was a protein–ligand interaction and we used a fluorescence polarization assay in search of inhibitors. However, the relatively low affinity of the interaction, characterized by a K_d of approximately 15 μM, necessitated the use of high concentration of protein and probe. This in turn not only made it unnecessary to test compounds at many of the lowest library concentrations but also prompted us to explore a higher than normal maximal concentration.

To bias the concentration–response testing toward higher compound concentrations, we used the built-in flexibility of our vertically prepared titrations. To deemphasize the lower concentrations, we simply skipped three of the lowest-concentration dilution plates when scheduling the screen on the robotic system, while keeping the top-concentration plates, as shown in Fig. 8. In addition, we added one concentration higher than that routinely tested by pin transferring two successive 23-nL aliquots from the highest-concentration library plate into the same assay plate, thus generating a new concentration point equal to double the previous value

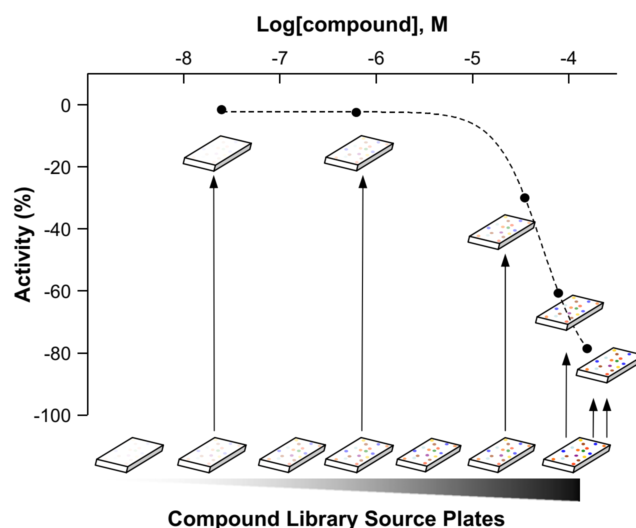

Figure 8. Implementation of asymmetric titration. Arrows represent the pin transfer of compound from the source plates available for testing (below) to the assay plates used in quantitative HTS (above). In this example, the highest-concentration point is customized to contain a twofold higher compound concentration achieved by a double pin transfer (dual arrows) of solution out of the library plate containing the highest available concentration.

(Fig. 8). By exploiting such flexibility, we not only shifted and appropriately biased the concentration–response profile, but also conserved precious protein reagent by dropping unnecessary low-concentration points.

Summary

In this report, we have presented the general strategy and specific processes for preparing compound libraries for use in qHTS. Our processes and equipment selections were driven by the need to handle the diverse streams of incoming compounds, follow-up tests, and the requirement to create an efficient plating process to support the qHTS paradigm. We find that inter-plate (vertical) titrations are appropriate for the screening phase of a project where library sets remain constant over extended periods of time and flexibility of concentration-range selection is important. The compound management schema described here enables the efficient, flexible, and cost-effective generation of concentration–response information on compound collections containing >200,000 members.

Acknowledgments

This research was supported by the NIH Roadmap for Medical Research and the Intramural Research Program of the NHGRI, NIH. We thank Fonda Newcomb and Mark McKown for expert help with EP^3 set-up and initial parameter optimization.

References

1. Malo, N.; Hanley, J. A.; Cerquozzi, S.; Pelletier, J.; Nadon, R. Statistical practice in high-throughput screening data analysis. *Nat. Biotechnol.* **2006,** *24(2)*, 167–175.
2. Auld, D.; Inglese, J.; Jadhav, A.; Austin, C.; Sittampalam, G.; Montrose-Rafizadeh, C.; Mcgee, J.; Iversen, P. HTS technologies to facilitate chemical genomics. *Eur. Pharm. Rev.* **2007,** *2*, 53–63.
3. Inglese, J.; Auld, D. S.; Jadhav, A.; Johnson, R. L.; Simeonov, A.; Yasgar, A.; Zheng, W.; Austin, C. P. Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103(31)*, 11473–11478.
4. Inglese, J.; Shamu, C. E.; Guy, R. K. Reporting data from high-throughput screening of small-molecule libraries. *Nat. Chem. Biol.* **2007,** *3(8)*, 438.
5. Simeonov, A.; Jadhav, A.; Sayed, A.; Wang, Y.; Nelson, M.; Thomas, C.; Inglese, J.; Williams, D.; Austin, C. Quantitative High-throughput Screen Identifies Inhibitors of the *Schistosoma mansoni* Redox Cascade. *PLoS Negl Trop Dis* **2008,** *2(1)*, e127.
6. Zheng, W.; Padia, J.; Urban, D. J.; Jadhav, A.; Goker-Alpan, O.; Simeonov, A.; Goldin, E.; Auld, D. S.; LaMarca, M. E.; Inglese, J.; Austin, C. P.; Sidransky, E. Three classes of gluococerebrosidase inhibitors identified by quantitative high-throughput screening as chemical chaperone leads for gaucher disease. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104(32)*, 13192–13197.
7. Archer, J. R. History, evolution, and trends in compound management for high throughput screening. *Assay Drug Dev. Technol.* **2004,** *2(6)*, 675–681.
8. http://ncgc.nih.gov/pub/compoundmanagement/salttable.pdf (accessed January 9, 2008).
9. Bowes, S.; Sun, D.; Kaffashan, A.; Zeng, C.; Chuaqui, C.; Hronowski, X.; Buko, A.; Zhang, X.; Josiah, S. Quality assessment and analysis of Biogen Idec compound library. *J. Biomol. Screen.* **2006,** *11(7)*, 828–835.
10. Catalan, J.; Diaz, C.; Garcia-Blanco, F. Characterization of binary solvent mixtures of DMSO with water and other cosolvents. *J. Org. Chem.* **2001,** *66(17)*, 5846–5852.
11. Cheng, X.; Hochlowski, J.; Tang, H.; Hepp, D.; Beckner, C.; Kantor, S.; Schmitt, R. Studies on repository compound stability in DMSO under various conditions. *J. Biomol. Screen* **2003,** *8(3)*, 292–304.
12. Kozikowski, B. A.; Burt, T. M.; Tirey, D. A.; Williams, L. E.; Kuzmak, B. R.; Stanton, D. T.; Morand, K. L.; Nelson, S. L. The effect of freeze/thaw cycles on the stability of compounds in DMSO. *J. Biomol. Screen.* **2003,** *8(2)*, 210–215.
13. Kozikowski, B. A.; Burt, T. M.; Tirey, D. A.; Williams, L. E.; Kuzmak, B. R.; Stanton, D. T.; Morand, K. L.; Nelson, S. L. The effect of room-temperature storage on the stability of compounds in DMSO. *J. Biomol. Screen.* **2003,** *8(2)*, 205–209.
14. Rasmussen, D. H.; MacKenzie, A. P. Phase diagram for the system water-dimethylsulphoxide. *Nature* **1968,** *220(5174)*, 1315–1317.
15. Walling, L.; Carramanzana, N.; Schulz, C.; Romig, T.; Johnson, M. Mixing in 384-well plates: issues, measurements, and solutions. *Assay Drug Dev. Technol.* **2007,** *5(2)*, 265–275.
16. Fogel, P.; Collette, P.; Dupront, A.; Garyantes, T.; Guedin, D. The confirmation rate of primary hits: a predictive model. *J. Biomol. Screen.* **2002,** *7(3)*, 175–190.
17. Cleveland, P. H.; Koutz, P. J. Nanoliter dispensing for uHTS using pin tools. *Assay Drug Dev. Technol.* **2005,** *3(2)*, 213–225.
18. Xia, M.; Huang, R.; Witt, K.; Southall, N.; Fostel, J.; Cho, M. -H.; Jadhav, A.; Smith, C.; Inglese, J.; Portier, C.; Tice, R.; Austin, C. Compound cytotoxicity profiling using quantitative high-throughput screening. *Environ. Health Perspect.* 2007, in press.
19. Hill, A. V. The possible effects of the aggregation of the molecule of haemoglobin on its dissociation curves. *J. Physiol. (London)* **1910,** *40*, 4–7.
20. Turner, R. J.; Charlton, S. J. Assessing the minimum number of data points required for accurate IC50 determination. *Assay Drug Dev. Technol.* **2005,** *3(5)*, 525–531.
21. Verheij, H. J. Leadlikeness and structural diversity of synthetic screening libraries. *Mol. Divers.* **2006,** *10(3)*, 377–388.
22. Chan, J. A.; Hueso-Rodriguez, J. A. Compound library management. *Methods Mol. Biol.* **2002,** *190*, 117–127.
23. Oldenburg, K.; Pooler, D.; Scudder, K.; Lipinski, C.; Kelly, M. High throughput sonication: evaluation for compound solubilization. *Comb. Chem. HTS* **2005,** *8(6)*, 499–512.