USING HIGH THROUGHPUT SCREENING TO ACQUIRE PROMISING DRUG CANDIDATES AGAINST *Mycobacterium tuberculosis*

A Thesis

by

STEVEN KENNETH LAIHING

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2011

Major Subject: Biochemistry

Using High Throughput Screening to Acquire Promising Drug Candidates Against

Mycobacterium tuberculosis

USING HIGH THROUGHPUT SCREENING TO ACQUIRE PROMISING DRUG CANDIDATES AGAINST *Mycobacterium tuberculosis*

A Thesis

by

STEVEN KENNETH LAIHING

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	James Sacchettini
Committee Members,	William Park
	Thomas Ioerger
Head of Department,	Gregory Reinhart

August 2011

Major Subject: Biochemistry

ABSTRACT

Using High Throughput Screening to Acquire Promising Drug Candidates Against *Mycobacterium tuberculosis*. (August 2011)

Steven Kenneth LaiHing, B.S., Oakwood University

Chair of Advisory Committee: Dr. James Sacchettini

Mycobacterium tuberculosis currently affects 1/3 of the world's population. Over the past 20 years tuberculosis has become more resistant to all front line drugs used against it. Because of this, the threat of Multi Drug Resistant (MDR-TB) and Extensive Drug Resistant (XDR-TB) strains has grown greater and emerges as a world health issue. Modern travel has greatly facilitated the spread of these resistant strains. For this reason, more frontline drugs are urgently needed in the fight against TB infection. High Throughput Screening can be used to both find and analyze promising drug candidates. Using automation, thousands of compounds can be tested against an attenuated strain of Tuberculosis and separate the promising compounds from the ineffective ones. We have found a select subset of candidates from our custom built ~52,000 compound diversity library which show potent inhibitory effects against our mc^2-7000 attenuated TB strain. These compounds have IC_{50s} ranging from 1.98 μM to 11.3 μM and should be considered for future development as drugs against TB. Among the active compounds, we have found enrichment for hydrazines, as well as representation of several chemi-classes including quinolones. To determine possible toxicity issues, we have also vetted

these compounds against a strain of human lymphoma; all of our promising compounds meet the threshold for non-cytoxicity.

DEDICATION

I dedicate this to my parents, Dr. Kenneth LaiHing and Jean LaiHing, as well as my fiancée, Ebony Hammond. It is also dedicated to all of my family and friends who supported me.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Sacchettini, as well as my committee members, Dr. Ioerger and Dr. Park, for their guidance and support throughout this research.

Thanks also go to my friends and colleagues and the Biochemistry Department faculty and staff for making my time at Texas A&M University a great experience. Lastly, thanks to Dr. Dwight Baker for his guidance, advice, and expertise in my project field.

NOMENCLATURE

HTS	High Throughput Screening
TB	Tuberculosis
CDD	Collaborative Drug Discovery
MDR	Multi Drug Resistant
XDR	Extensive Drug Resistant
INH	Isoniazid
MOA	Mode of Action

TABLE OF CONTENTS

1. INTRODUCTION: THE NEED FOR HIGH THROUGHPUT SCREENING

*Mycobacterium tuberculos*is (TB) has plagued mankind for centuries. Currently TB infects a new person every second, and greatly afflicts those immunocompromised patients (World Health Organization). On the continent of Africa, TB claims the lives of 70% of those who suffer from AIDS. Tuberculosis also greatly affects victims by its latent infection attributes, i.e. its ability to stay dormant in the body for many years before becoming symptomatic (1).

The problems facing TB treatments are symptomatic of the larger issues facing medicine with the resistance of popular drugs and treatment of disease. According to the Centers for Disease Control, TB will be resistant to all current frontline antibiotics by 2020. During the increasingly encountered MDR-TB, resistance is observed to at least isoniazid (INH) and rifampicin (RIF), the two most powerful frontline drugs. This threat of resistance has demanded that research and development expeditiously discover and bring new drugs to market to replace or supplement our current treatments. According to a 2007 Nature Review in Drug Discovery by Payne *et.al* it takes approximately 13.5 years for a drug to make it fully through the Food and Drug Administration approval process (2), making the issue of drug development especially urgent. The potential impact of bacterial resistance to current drugs poses a serious threat to world health.

This thesis follows the style of *Journal of Biological Chemistry.*

Adding to the urgency of drug development is the fact that the current compounds in the development pipeline are derivatives of chemical classes for which there are known resistance mechanisms. Developing new, chemically unrelated compounds requires the ability to rapidly evaluate compounds across a range of chemical space. High throughput screening (HTS) is a powerful tool for quickly assessing the effectiveness of many compounds at once (2). It is widely used in industry and is gaining acceptance in academia due to its multiple applications, rapidness, and minimization of human error in experiments, reproducibility, and quality of data. Most labs which feature a HTS program use automation for their screening purposes. While the cost of entry to a robotics based HTS program may be high, a lab that makes the required investment can progress much faster in the drug discovery process. HTS requires more equipment (like robotics) than enzyme based assays, but carries with it several advantages. With an enzyme or target based approach, promising inhibitors may bind and show efficacy, but in actuality may not be permeable to the cell membrane. This does not translate well into a promising drug. With activity based or a HTS approach, by definition the compounds which are found to inhibit are both active and cell wall permeable. This can prove to be greatly advantageous, as some believe it is an easier process to identify a target for an already effective inhibitor versus modifying a compound candidate to confer permeability. It is not enough however, to simply gather thousands of compounds together and screen them against a specific strain. The real strength of the HTS program is in the formation of a diverse library and the generation of high quality, reproducible data. Previous whole cell screens have met with varying degrees of success. For

instance, the widely known 2009 screen run by Southern Research Institute (SRI) featured a library of 100,997 compounds from ChemBridge Corporation (3) . Their activity based assay discovered hundreds of promising compounds by sequentially screening in lessening concentrations until even the most promising compounds were ineffective, in order to generate IC_{90} values. Out of their 100,000+ compound library, they discovered a highly active subgroup of 293 compounds with high IC_{90} values (<0.1-1.6 μg/mL). They also discovered multiple representations of different chemi-classes in their inhibitor subset, including aminothiazoles, hydroxyquinolones to name a few. In many ways our HTS builds upon the screening foundation laid out by the SRI screen. Our screen, like theirs, began with the formation of a diversity selected library of drug like compounds. We also used Lipinski's rules as a guide in our selection criteria. In contrast, the previously mentioned GlaxoSmithKline (GSK) HTS program of the mid to late 1990s spent approximately 1 million USD per screen with 70 target based screens performed, yielding only 5 tractable hits which were later discarded from active study (4). Their library featured an over representation of certain scaffolds such as nucleoside analogs. Their approach in searching only for compounds against specific targets caused them to possibly overlook otherwise active compounds with unknown modes of action.

Unlike most of the other labs performing screens against MTB, the Sacchettini HTS program uses a vaccine strain of TB, mc^2-7000 instead of wild type H37Rv. This doubly attenuated pantothenic knock out strain ($\triangle$ *panCD)* removes genes required for pantothenate synthesis (5) . Pantothenic acid is used in the synthesis of coenzyme A (CoA). Coenzyme A acts as an acyl group carrier to form acetyl-CoA and other related

compounds; this is a way to transport carbon atoms within the cell. Mc^2-7000 also allows us to perform all of our experiments at a BL2 level due to the region of difference one (RD1) deletion which encodes several of the strain's virulence factors (6). RD1 locus is absent in all BCG vaccine strains of TB, yet is present in all *M. tuberculosis* strains. The *panCD* deletion was added as a backup safety step due to the significant increase in virulence of vaccine strains which only had the RD1 deletion. It is an acknowledged concession that compounds which act upon the virulence pathway(s) will probably be overlooked and undetected in our screen because of our model strain. While this is less than favorable, the ability to test for all the compounds which are not linked to TB's virulence pathways is still a valuable opportunity. It is also a superior model than screening compounds against *M. Smegmatis*. Understandably this might bring up questions as to the results acquired using this strain. It has been shown that the differences between mc^2-7000 and wild type (WT) H37Rv are less than between WT and the much used BCG vaccine strain in mouse organ CFU studies (5). We find that, properly supplemented with pantothenic acid to account for the *panCD* deletion, our model strain allows us to adequately test whole cell compounds in a high throughput method and determine whether each candidate has any inhibitory effects against TB. Testing our promising hit subset independently against WT H37Rv would be useful to assess the real world benefits of our candidates. Information from whole cell assays combined with target identification and crystal structure inhibitor complexes will help in the structure activity (SAR) relationship of our validated hit compounds in a rational drug design strategy.

2. EXPERIMENTAL METHODS AND MATERIALS

Our HTS uses the mc^2-7000 strain of TB (5), courtesy of the Bill Jacobs lab. The glycerol stocks of the strain are kept at -80 ° C and a new starter culture is set up every week, which is kept at 37 ° C on a shaker in an incubator. Mc^2-7000 is not virulent, can be grown in a BL2 facility, and requires pantothenic acid to grow. The diversity library was custom built by the Sacchettini lab in order to acquire promising drug candidates. The diversity library consists of compounds from multiple vendors (Chembridge, Chem-Div, and Enamine) (Fig.2). The compounds were chosen specifically for their conformance with a preset criteria based on Lipinski's rules for selecting drug like compounds (7) as well as their acceptance based on our Tanimoto scoring cutoff (8). It is important to note here that Lipinski's rules were never meant for natural products or their analogues, which many of the current TB drugs on the market are derived from. In our case the Sac-1 diversity library features no natural products. Tanimoto scoring is a similarity calculation which can be run to avoid duplicate representation of compound family scaffolds being selected for a library; it can be used to enhance library diversity. Tanimoto scoring (also known as Jaccard Similarity and distance index) allows selection from a scale of -.5 (all compounds) to 1 (fewest compounds). Our Tanimoto selection criterion was 0.7. The compounds were shipped to us as individual powders in a 96 well plate format; we converted 4 of these 96 well "mother plates" into one 384 well daughter plates (162 of them total), solubilized with 100% DMSO. These plates were then tested against the mc^2-7000 strain (Fig. 4).

2.1 Preparation of Starter Culture

Briefly, a 10 mL starter culture with mc^2-7000 was grown for ~4 days. The starter culture consists of 10mL Difco 7H9 (pH ~7.4), 100 µL dextrose, 1 mL OADC (Middlebrook, 212351), 85 µL NaCl (10% solution) 25 µL tween, and 10 µL each of malachite green and 1 mM pantothenic acid (50 mg/mL) mixed in a 20 mL capacity ink bottle. As a precaution, every starter culture had a sample taken from it and swabbed onto a Luria Broth (LB) plate with no antibiotic. This LB plate is placed into an incubator at 37 degrees centigrade and is allowed to sit overnight. TB grows much slower than other bacteria, so if any growth is observed on the LB contamination check plate after only a day, the starter is contaminated and disposed of by filling the culture bottle with Environ™ Vesphene™, a tubercidal agent.

2.2 Preparation of Assay Plates

Pending verification of non-contamination, 300 µL of this grown starter culture (OD ~0.6) is added to a 20 mL plate assay solution with the exact same concentrations of supplements, sans OADC. Our assay requires 20 mL of 7H9 media per plate, be it 96 well (200 µL/well) or the more commonly used 384 well (50 µL/well). Therefore, when setting up an assay, it might be 500 mL of media being made for the run (20 plates worth).

The screening program features a group of robots that create the 320 compound daughter plates, the assay test plates, as well as stain, analyze, and re-plate promising compounds for further testing. Chief among these robots is the Cy-Bio Vario, featuring 96 and 384 well interchangeable heads and a dual rack system for storing and moving

plates as they are needed during the assay. The Vario system is enclosed in a HEPA™ filtered hood, enabling its use in testing many different strains. One μL of each compound was added to a plate well with 49 μL of cell assay solution. Minimum (Rifampicin) and maximum (no drug) growth control wells were placed in each plate (Fig. 3). The assay plates were then sealed and stored on a platform shaker in an upright incubator (Forma Scientific Incubator model 3282) set for 37 ° C. The plates were grown for 6 days, and on the 6th day were stained with 1.25 μL/well of Cell Titer-Blue™ and allowed to incubate for an additional 24 hours. Cell Titer-Blue™ (or Resazurin) is an oxidation reduction indicator used in whole cell viability assays (9). It is reduced from its normal blue appearance to a pinkish red when test cells are actively respiring in a well.

2.3 Acquiring and Analyzing Assay Data

On the 7^{th} day, the plates were read using a POLARstar Omega spectrophotometer running an absorbance program set to calculate a differential reading at 573nm and 605nm. The resulting output number, if greater than zero, was designated as an ineffective compound against TB, showing growth, (pink). If the resulting number was less than zero, the compound was inhibitory towards TB (blue). Data was taken from the Omega spectrophotometer as a Comma Delimited (CSV) file and used for sorting and calculating plate specific percent controls, as well as mean (Fig. 5.) The data was then formatted and uploaded onto the Collaborative Drug Design (CDD) website, where further calculations could be performed, including IC_{50}, and heat mapping. The CDD website offers our lab a secure cloud environment for storage and analysis of our data,

while also giving us the ability to share data and results with colleagues not only in our lab but in other collaborating labs as well. It is a powerful tool for our screening program (2).

2.4 Statistical Analysis

After recovering the data readouts from the spectrophotometer, it was analyzed and its quality determined. Due to the highly variable nature of the whole cell assay, several different methods of analyzing the data were taken into consideration. The first way which the data was analyzed was to perform a calculation on each plate using the *percent of control method* for each assay well. In essence, we compared the inhibitory properties of each drug to the known inhibitory properties of our max control (no drug present) and our minimum control (rifampicin present). The following equation was then used to calculate the percent inhibition:

$$1 - \frac{(\text{Sample} - \text{minimum})}{(\text{Maximum} - \text{minimum})} \times 100 = \%\ \text{inhibition} \qquad \text{(Eq. 1)}$$

This method for analyzing the data would have been satisfactory for all of the data run if it were not for the fact that the control wells behaved differently not only from run to run but occasionally from plate to plate. This caused statistical issues with both the

reproducibility and the similarity of the data (*Note: we were able to use Equation 1 on our potent hit data, described below*). We then used an alternative method for analyzing the primary screen data, namely the *population- mean method*. The population-mean method takes the individual sample and divides it by the mean of all the samples on the plate, with the controls removed.

The formula:

$$1 - \frac{\text{(Sample)}}{\text{(Mean of samples)}} \times 100 \qquad \text{(Eq. 2)}$$

As part of the statistical analysis, the plate data was checked for variability between different concentrations. The purpose of this was to ensure that the data generated from the screens could be compared with itself for consistency and would support legitimate conclusions about the data results. To accomplish this, the software Pipeline Pilot was used to compare different concentration runs such as 10 μM and 5 μM (Figure 8). The more steep the output graph's slope (0.90 in Figure 8) the higher the consistency of the data. If comparisons were made between two sets of identical compounds at the same concentration, the ideal slope would be 1, showing complete agreement and reproducibility between the two runs. Since the previously stated comparison was made between identical compounds at two different concentrations, the .90 slope is acceptable,

since it is expected that a certain small amount of compounds would drop out after lowering the concentration. In order to ensure the highest quality of data, plates with the most potent hits from the Sac-1 library were run three times (3 runs) each in triplicate. They were run on different days, with independently made supplements as well as different source starter cultures. This data was used to generate the IC_{50} curves seen in Figure 1. The percent control data was plotted in Excel for the curves (using Equation 1) and the concentration data was inputted into Pipeline Pilot to generate the IC_{50} value. The minimum control used was Rifampicin and the max control was the wells with no compound present. Error bars were calculated by using the Standard Deviation of Population (Stdev.P) function in Excel 2010 for each plate.

2.5 Assay Optimization

The original 20 μM screen was performed with an incubation time totaling 5 days. On the 5^{th} day they were stained, with absorbance readings taken on the 6^{th} day. As the data was examined it was discovered that the dynamic range on each plate was statistically too small (Fig. 5 Cont.). The inhibitor cutoff or minimum was observed to be slightly less than zero while the max was only around OD 0.2. In response to this, in the secondary screening portion of the HTS program the time of incubation was lengthened to allow for more development and growth of the mc^2-7000 culture in each well. Once stained and analyzed, it was observed that the range had nearly doubled to OD ~0.4 (Fig. 6). This proved to be a much more desirous dynamic range for the statistical analysis and became part of the standard operating procedure. The secondary screening also included an added supplement of malachite green to combat any contamination that may arise.

Malachite green targets *E.coli* and other harmful bacteria in whole cell cultures and proved to be quite valuable in limiting losses and wastes of materials due to contamination down time (10).

During the secondary phase of screening a robot mixing step was added (validated by bromophenol blue). This step occurs immediately after the compounds are dispensed into the cell media solution. It is a four step mix for each well with the plate shifting as 20 uL are aspirated and dispensed in each corner of the well. It ensures a full exposure of the compound to the test strain.

3. RESULTS

The preliminary 20 μM screen yielded many hits which showed activity against mc^2-7000 (Fig. 7). These hits were re-plated using an EP-Motion robot which cherry picked the promising leads from the daughter plates into a fresh plate containing only hits. Compounds were then tested again at lower concentrations and at each decreasing level of concentration the hits that did not inhibit mc^2-7000 growth were removed from further testing. At 10 μM we detected ~700 compounds which had activity against TB. At 5 μM we detected ~350 active compounds. At 1.25 μM we detected 174 compounds, and at 0.625 μM we detected 107 compounds. These compounds were then re-plated and tested in another dilution screen beginning with 10 μM and continuing serially downward to 9 nM. At the two lowest concentrations, no hits showed activity. To ensure good quality data, the most potent hits (21 compounds) were tested in 3 runs (triplicate for each run). IC_{50}s were calculated using Pipeline Pilot (11) and graphs were generated in Excel 2010 (Fig. 1). Another subset of less potent compounds is presented in Figure 10. Percent of Control values were determined using *Equation 1*. These promising compounds also were run against two cytotoxicity screens; the WT 2561 strain of the human lymphoma cell line (12) (at a concentration of 5 μM) and a human fibroblasts screen. All of the top compounds were beneath the lethal threshold established in both toxicity screens. The top ranked compound, CD:0325-0400 (2,2'-(E)-diazene-1,2-diyldiphenol) (Figure 9) has been used in previous research (13) to validate what is called a "molecular keypad lock." It functions as a mono-phosphate chemo-sensor. Many other compounds which had inhibitory properties against TB have other known

properties or patents, such as the second best compound by IC_{50} CB:5538347 (1-methyl-2-[3-(1-methyl-1H-benzimidazol-2-yl)triaz-1-en-1-yl]-1H-benzimidazole) has patents pending or approved. Many of our compounds might have broad based implications against other diseases. Using a clustering algorithm, we determined the overall abundance of certain chemi-classes in our library. We observed that one of our top hits, EN-T5387803 was a quinolone (Fig. 10). This is surprising considering the extremely low abundance in the general library population. Only 227 quinolones are present in the entire 51K Sac-1 library (~0.45%). Quinolones (and fluoroquinolones) are chemotherapeutic drugs, which kill bacteria by interfering with DNA replication, mainly by inhibition of DNA gyrase. DNA gyrase belongs to a class of enzymes known as topoisomerases enzymes (14) . These enzymes are highly essential for the catalyzation required for DNA replication, transcription, and recombination in prokaryotic and eukaryotic cells. This differs from other drugs such as penicillin and cephalosporin which inhibit cell wall biosynthesis. Ciprofloxacin is an example of an antibiotic in the fluoroquinolone class (15). Quinolones are still being investigated for their inhibitory abilities; novel synthetic quinolones have recently been reported which inhibit *M. tuberculosis* and rapid-growing mycobacteria as well as methicillin-resistant *S. aureus*, while showing no inhibition for *E. coli and P. aeruginosa*, suggesting a specific uracil binding site (16). Of the top compounds, we observed an enrichment of hydrazines. Considering only 5,528 compounds out of the total 51K library fit this class (~10%), it is notable that over 50% of our top hits were hydrazines. Isoniazid (INH), a front line anti-tuberculosis drug is a hydrazine, so further testing (perhaps against an INH resistant

strain) should be performed on these hydrazines to observe whether they possess the same mode of action. Some current frontline drugs such as INH require bioactivation to be effective. These types of drugs are known as prodrugs. For example, INH must be activated by a bacterial catalase-peroxidase enzyme called KatG. KatG couples the isonicotinic acyl with nicotinamide adenine dinucleotide (NADH) to form isonicotinic acyl-NADH complex. This complex tightly binds to the enoyl-acyl carrier protein reductase InhA, which blocks the natural enoyl-AcpM substrate and the action of fatty acid synthase. This process inhibits the synthesis of mycolic acid, required for the mycobacterial cell wall (17) . INH is most often given in a regimen combination with other frontline or second line drugs dependent on type of TB. Drug regimens which use only single drugs result in rapid resistance development and treatment failure.

Out of our top compounds, four (CB:5109296, EN:T5996673, CD:8010-0479 and EN:T6076227) have the scaffolding required to be prodrugs. These 4 compounds have the necessary hydrolyzing groups to degrade and demonstrate their efficacy *in vitro*. Recent research illustrates that esters (lipophilic specifically) have shown more activity (in concentrations 10 fold lower) against TB than current drugs including pyrazinamide, while still easily activated by mycobacterial enzymes (18). There is still much discussion in the drug development community as to whether prodrugs are desirable candidates. Many frontline drugs are prodrugs, such as INH. The major argument put forth against prodrugs is that it is easier to develop resistance to them versus other conventional classes. It has been previously reported that INH encounters resistance on the order of 1/1,000,000 organisms, while Rifampicin encounters resistance in 1/100,000,000 (19).

Also, the overwhelming majority of mutations that confer resistance to drugs requiring activation affect the activation process, not the target (19). This would infer that prodrugs would be more affected than other drug classes with resistance problems.

Overall many of the known chemi-classes of current use tuberculosis drugs are represented in our candidate subset. While seeing similarities between candidate and frontline drug may give us a hint as to target and mode of action, the presence of previously unseen chemi-classes and emerging compound families gives us insight into novel strategies of attacking TB.

4. DISCUSSION AND CONCLUSIONS

The screening program yielded several promising candidates that can continue to mouse studies or direct application studies against *Mtb* H37Rv. Target identification using mutagenesis or further screening studies is a natural progression for these compounds based on the data acquired from this screening program. Previous screening approaches have been a process in which the goal was to choose a target and then find compounds which were active against it. Our HTS program chose in essence a strategy that was the reverse of that. The goal of the Sacchettini HTS program was to first and foremost find promising compound candidates which not only were effective against TB, but also were specifically targeting a portion of the TB genome. The hope is that when the active compounds are matched with their identified targets, we will be able to structurally optimize and design new pharmaceuticals which will be less prone to resistance (or lengthens the time before resistance develops), less toxic, and efficient at lower clinical doses. Another opportunity for our promising compounds is to pair them with known TB treatments to determine whether they work in synergy. The future of effective TB treatment may very well be the reapplication of a compound that is ineffective due to resistance, coupled with a new compound that allows efficacy to return. Our secondary screen with human lymphoma cells and fibroblasts allowed us to remove compounds which were killing the bacteria due to toxicity. The remaining compounds are promising compounds (due to their non-toxicity, combined with their reproducible effectiveness against our TB strain) and are desirable candidates for further study.

It is a fact that Tuberculosis greatly affects the poorer regions of our planet. The people who suffer worst from TB are those who have very little money, little to no access to pure drinking water, and live in areas with poor public health programs (20). For these reasons, many drug companies have pulled out of anti-TB drug development. Simply put, it is more rewarding financially for drug companies to target diseases and conditions whose sufferers (and their insurance) can afford much greater prices for their care, versus poor people who cannot afford medicine. This unfortunate truth demands that the scientific community come up with novel solutions and more efficient ways to develop treatments for these "poor" diseases. High Throughput Screening allows an academic lab to function and analyze large batches of compounds for leads much like a pharmaceutical company, but is freer of the financial considerations of a large company. By using HTS it is conceivable that the academic research of today could lead to a base for a new anti-tubercular treatment in the future.

REFERENCES

1. de Steenwinkel, J. E. M., de Knegt, G. J., ten Kate, M. T., van Belkum, A., Verbrugh, H. A., Kremer, K., van Soolingen, D., and Bakker-Woudenberg, I. A. J. M. (2010) *Journal of Antimicrobial Chemotherapy* **65**, 2582-2589

2. Ekins, S., Bradford, J., Dole, K., Spektor, A., Gregory, K., Blondeau, D., Hohman, M., and Bunin, B. A. (2010) *Molecular BioSystems* **6**, 840-851

3. Ananthan, S., Faaleolea, E. R., Goldman, R. C., Hobrath, J. V., Kwong, C. D., Laughon, B. E., Maddry, J. A., Mehta, A., Rasmussen, L., Reynolds, R. C., Secrist Iii, J. A., Shindo, N., Showe, D. N., Sosa, M. I., Suling, W. J., and White, E. L. (2009) *Tuberculosis* **89**, 334-353

4. Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007) *Nat Rev Drug Discov* **6**, 29-40

5. Sambandamurthy, V. K., Derrick, S. C., Hsu, T., Chen, B., Larsen, M. H., Jalapathy, K. V., Chen, M., Kim, J., Porcelli, S. A., Chan, J., Morris, S. L., and Jacobs, J. W. R. (2006) *Vaccine* **24**, 6309-6320

6. Junqueira-Kipnis, A. P., Basaraba, R. J., Gruppo, V., Palanisamy, G., Turner, O. C., Hsu, T., Jacobs, W. R., Fulton, S. A., Reba, S. M., Boom, W. H., and Orme, I. M. (2006) *Immunology* **119**, 224-231

7. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001) *Advanced Drug Delivery Reviews* **46**, 3-26

8. Searls, D. B. (2010) *PLoS Comput Biol* **6**, e1000809

9. Czekanska, E. M. (2011) Assessment of Cell Proliferation with Resazurin-Based Fluorescent Dye. in *Mammalian Cell Viability* (Stoddart, M. J. ed.), pp 27-32, Humana Press, New York

10. Brovko, L., Meyer, A., Tiwana, A. S., Chen, W., Liu, H., Filipe, C. D. M., Griffiths, M.W., (2009) *Journal of Food Protection* **72**, 1020-1024

11. Altintas, I., Berkley, C., Jaeger, E., Jones, M., Ludascher, B., and Mock, S. (2004) Kepler: an extensible system for design and execution of scientific workflows. in Proceedings of *16th International Conference on Scientific and Statistical Database Management, 2004.*

12. Maxwell, S. A., Li, Z., Jaya, D., Ballard, S., Ferrell, J., and Fu, H. (2009) *Journal of Biological Chemistry* **284**, 22379-22389

13. Wang, J., and Ha, C.-S. (2010) *Analyst* **135**, 1214-1218

14. Appelbaum, P. C., and Hunter, P. A. (2000) *International Journal of Antimicrobial Agents* **16**, 5-15

15. Bolon, M. K. (2011) *The Medical Clinics of North America* **95**, 793-817

16. Guzman, J. D., Wube, A., Evangelopoulos, D., Gupta, A., Hüfner, A., Basavannacharya, C., Rahman, M. M., Thomaschitz, C., Bauer, R., McHugh, T. D., Nobeli, I., Prieto, J. M., Gibbons, S., Bucar, F., and Bhakta, S. (2011) *Journal of Antimicrobial Chemotherapy,* **66,** 1766-1772

17. Miesel, L., Rozwarski, D. A., Sacchettini, J. C., and Jacobs Jr, W. R. (1998) *Novartis Found Symp* **217**, 209-220; discussion 220-201

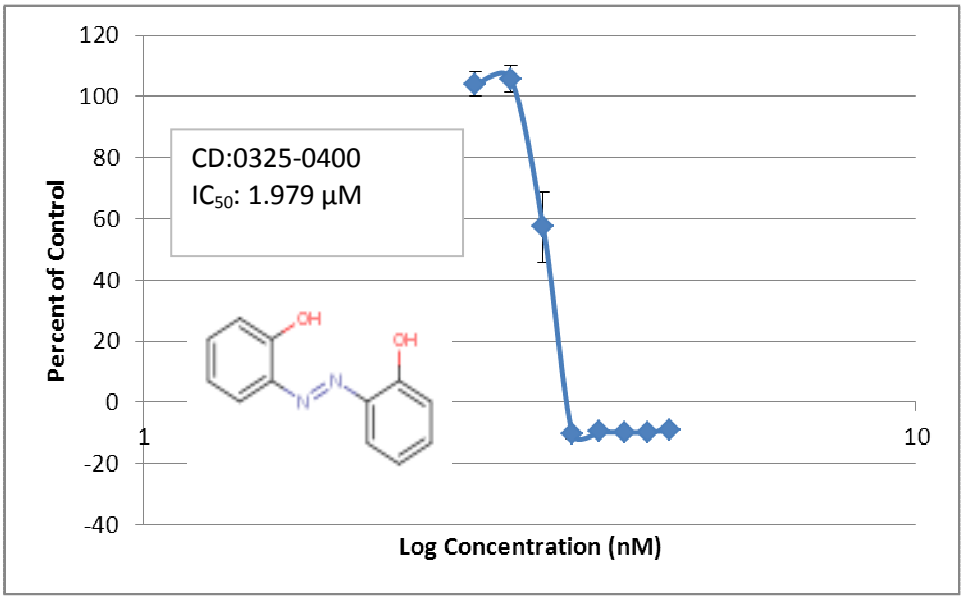
18. Simões, M. F., Valente, E., Gómez, M. J. R., Anes, E., and Constantino, L. (2009) *European Journal of Pharmaceutical Sciences* **37**, 257-263

19. Barry, C. E., Slayden, R. A., Sampson, A. E., and Lee, R. E. (2000) *Biochemical Pharmacology* **59**, 221-231

20. Maher, D. (2010) *Tropical Medicine & International Health* **15**, 11-17

APPENDIX

FIGURES



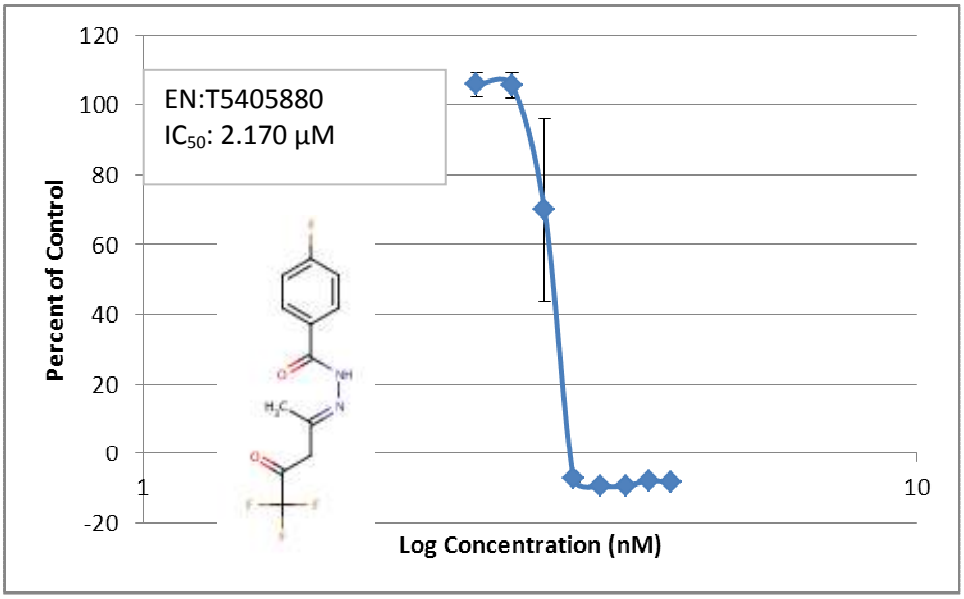
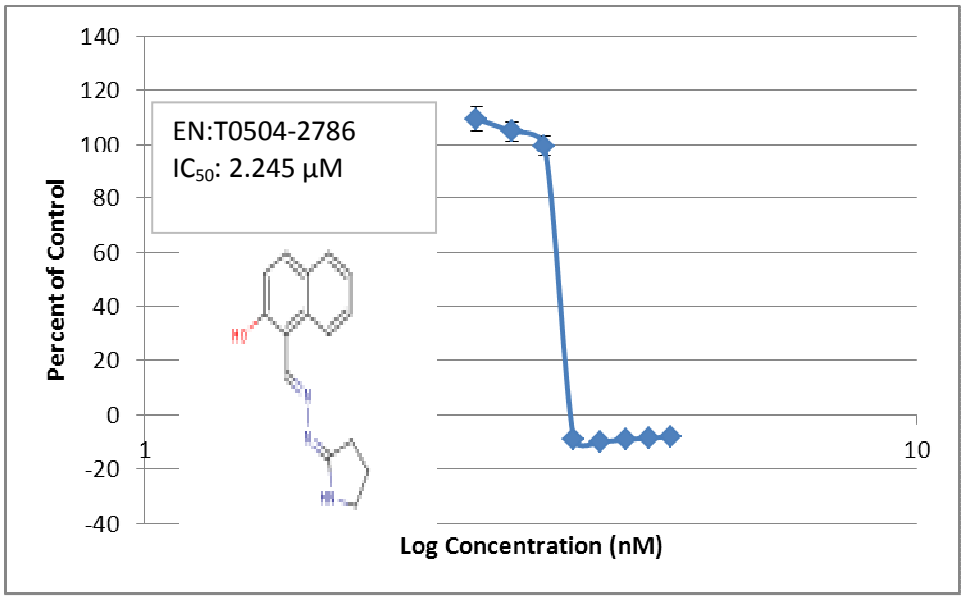


Figure 1 Top Hits by IC_{50}



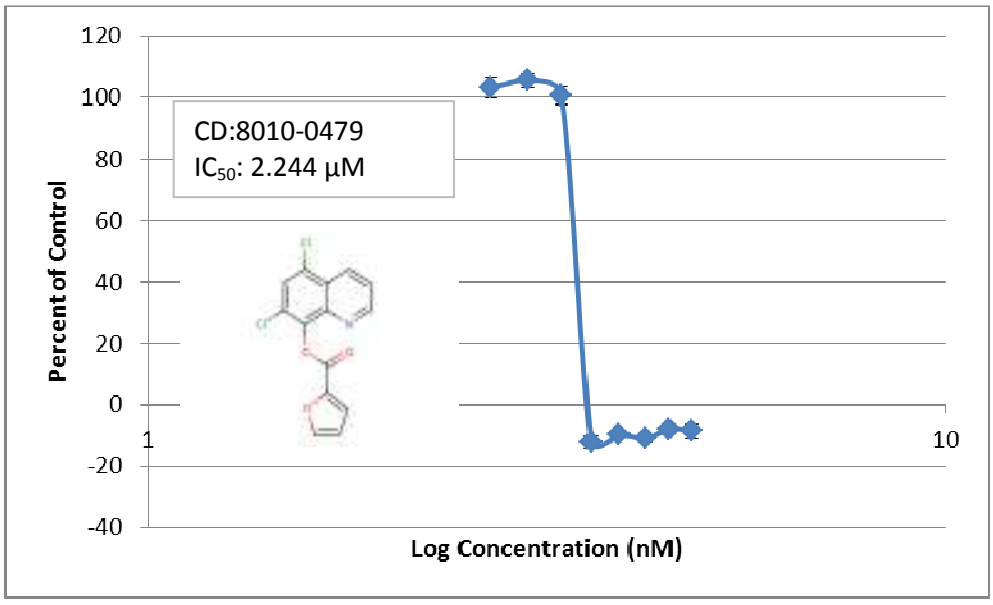
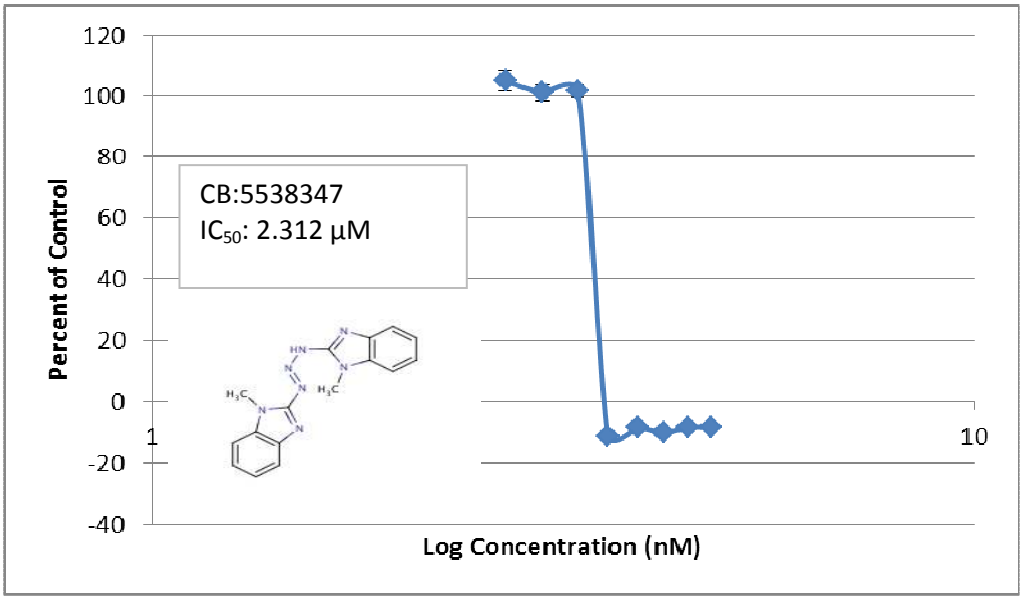


Figure 1. continued.



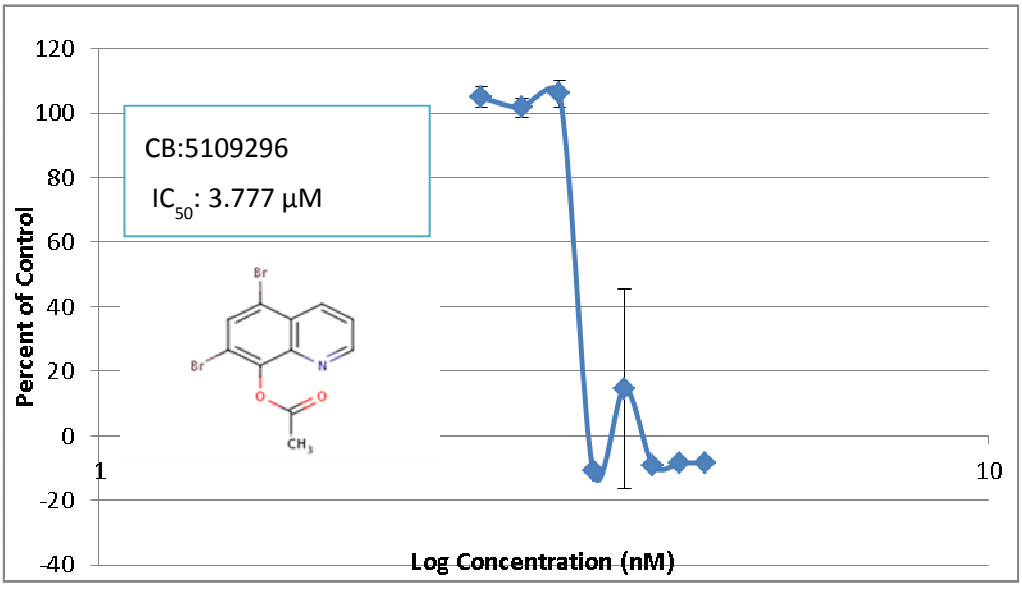
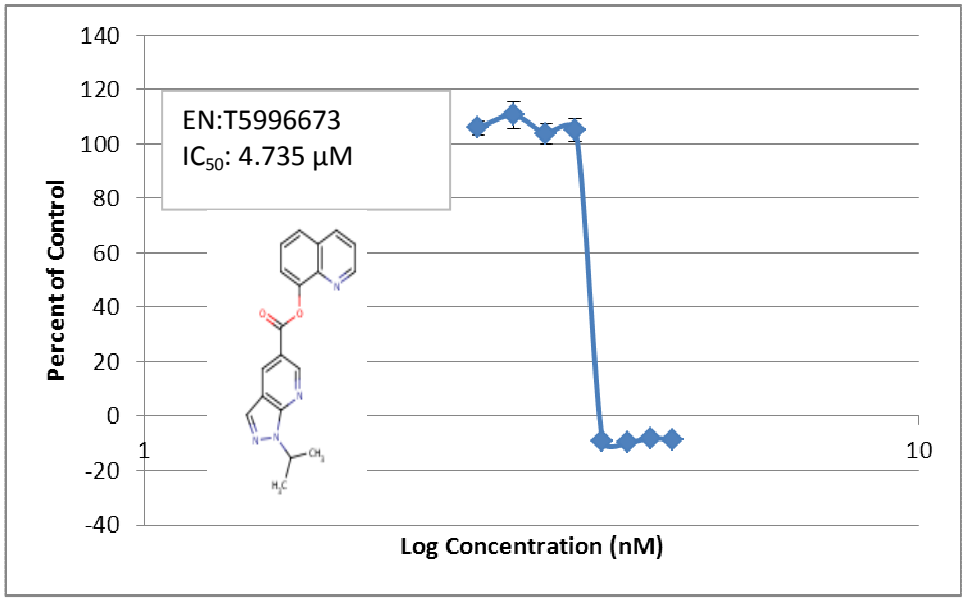


Figure 1. continued.



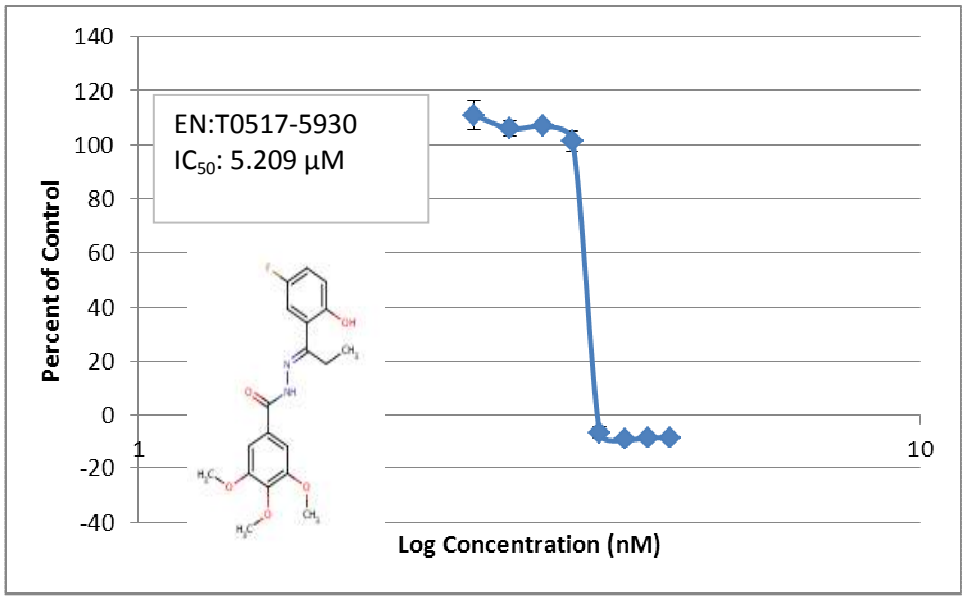
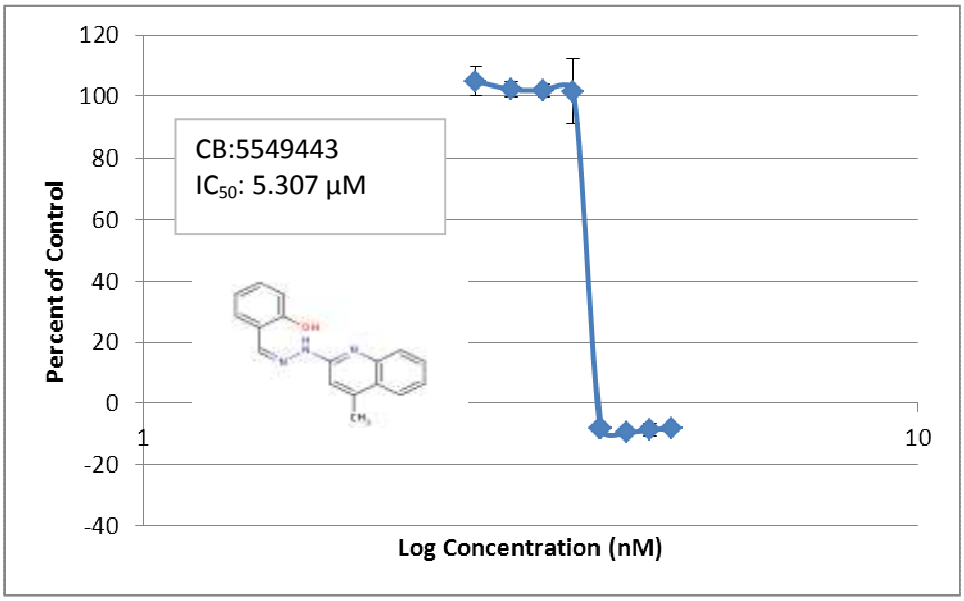


Figure 1. continued.



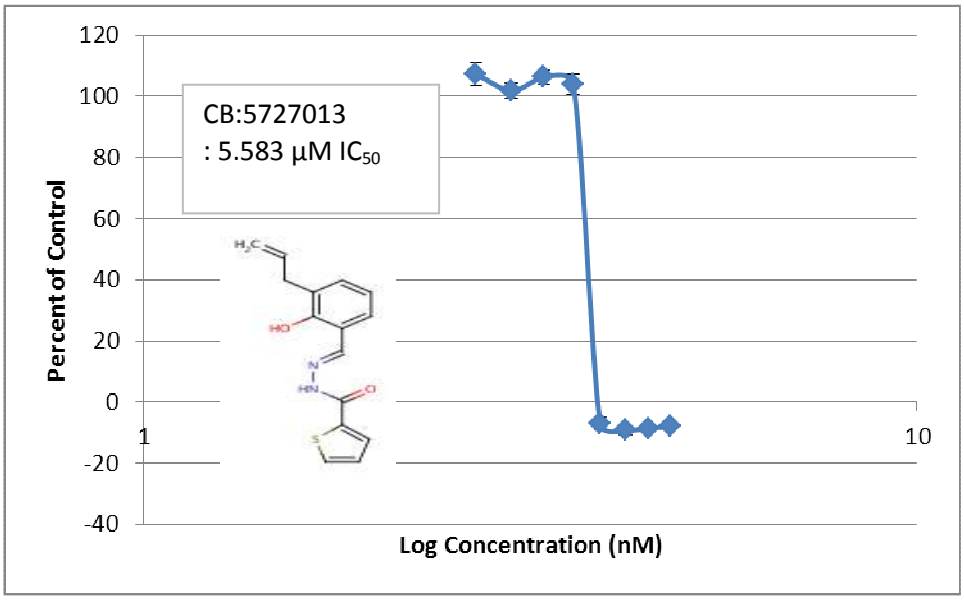
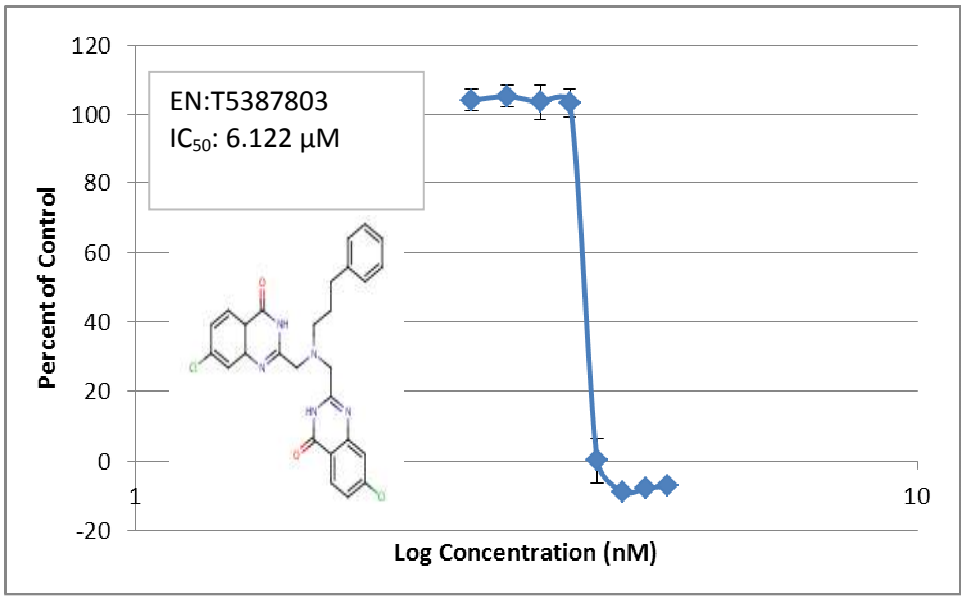


Figure 1. continued.



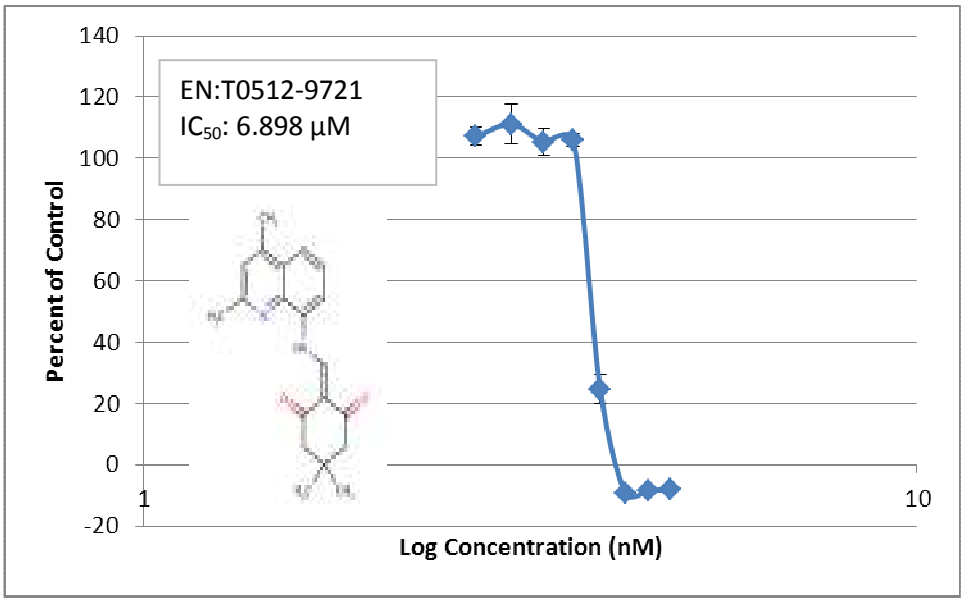
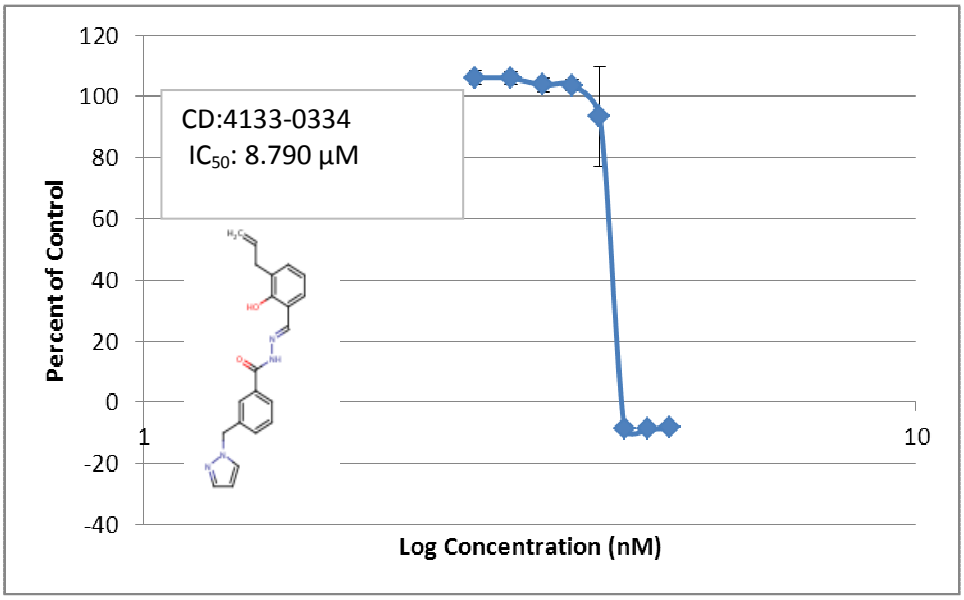


Figure 1. continued.



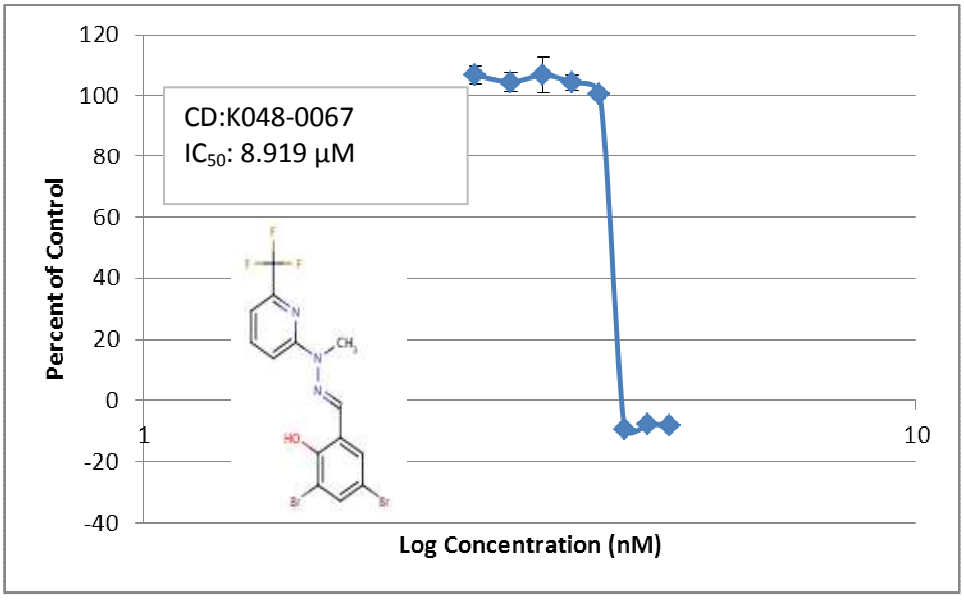
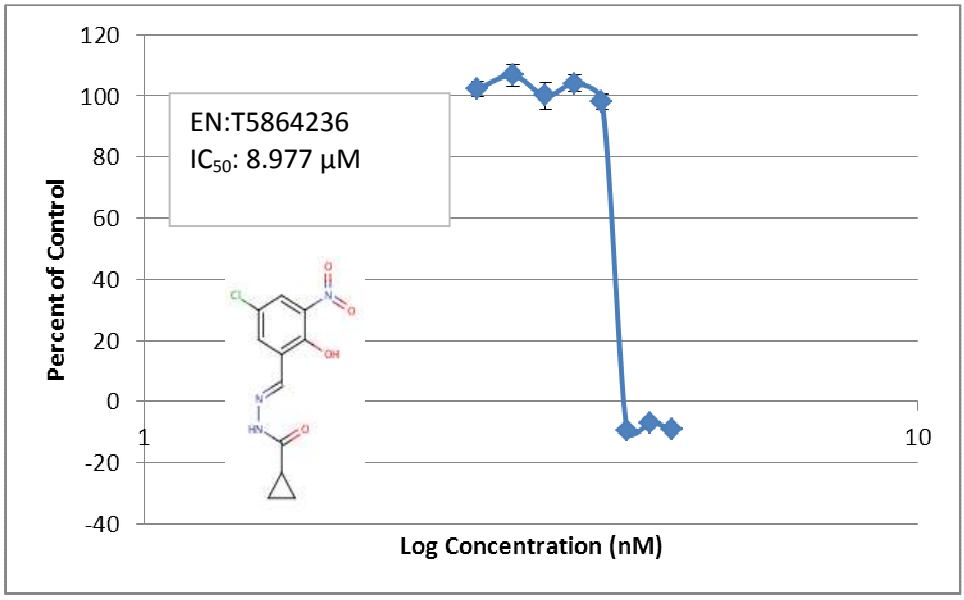


Figure 1. continued.



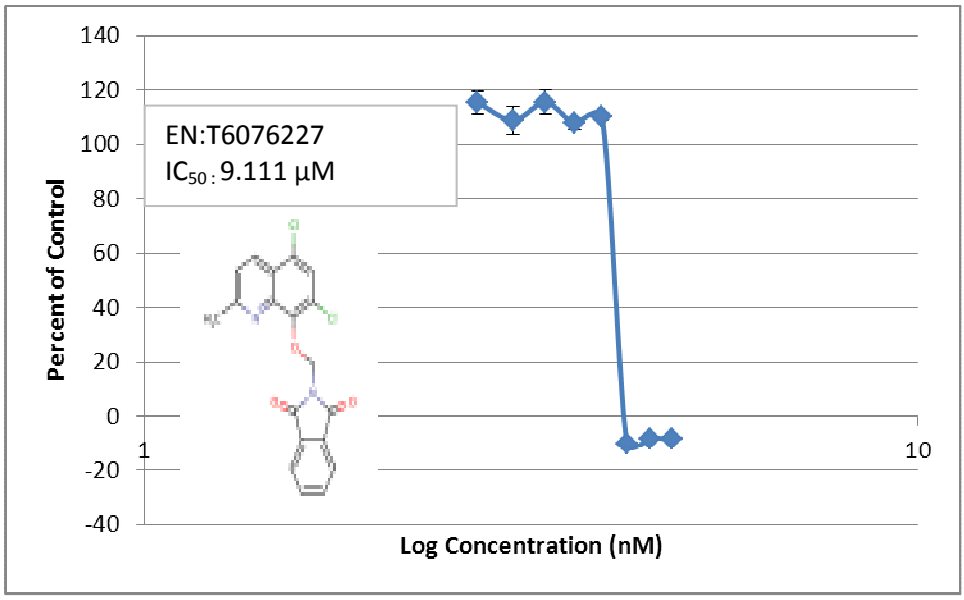
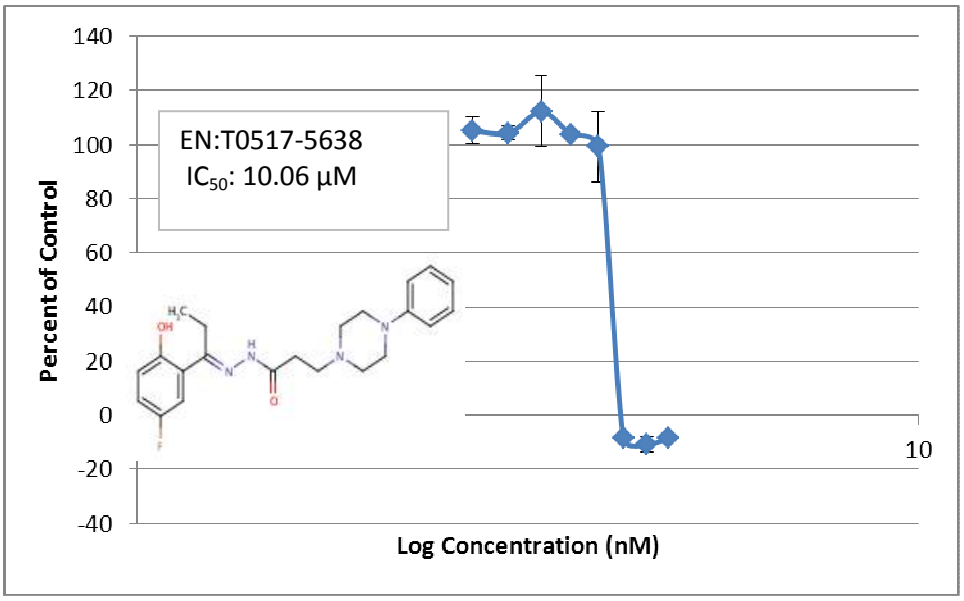


Figure 1. continued.



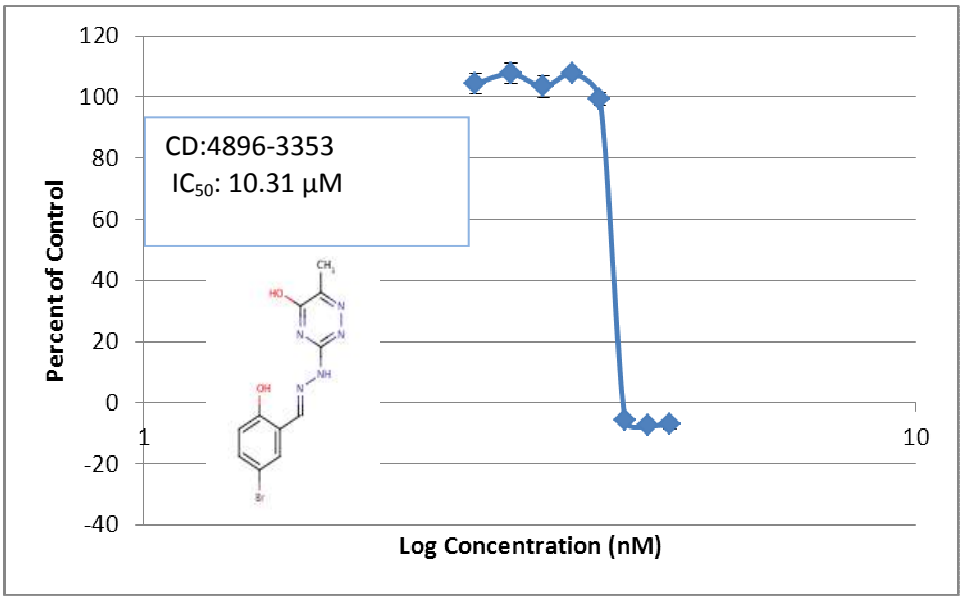
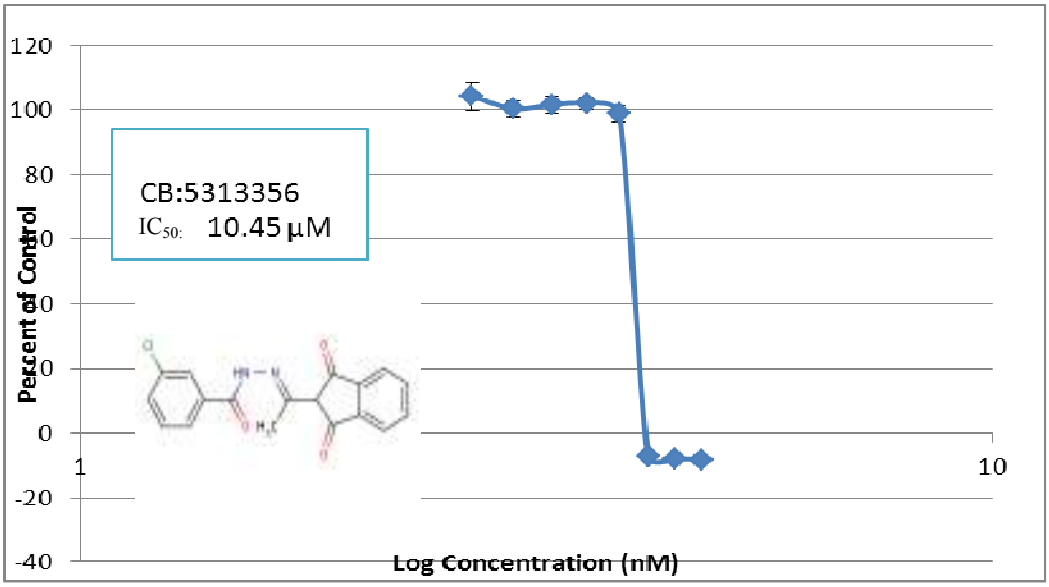


Figure 1. continued.



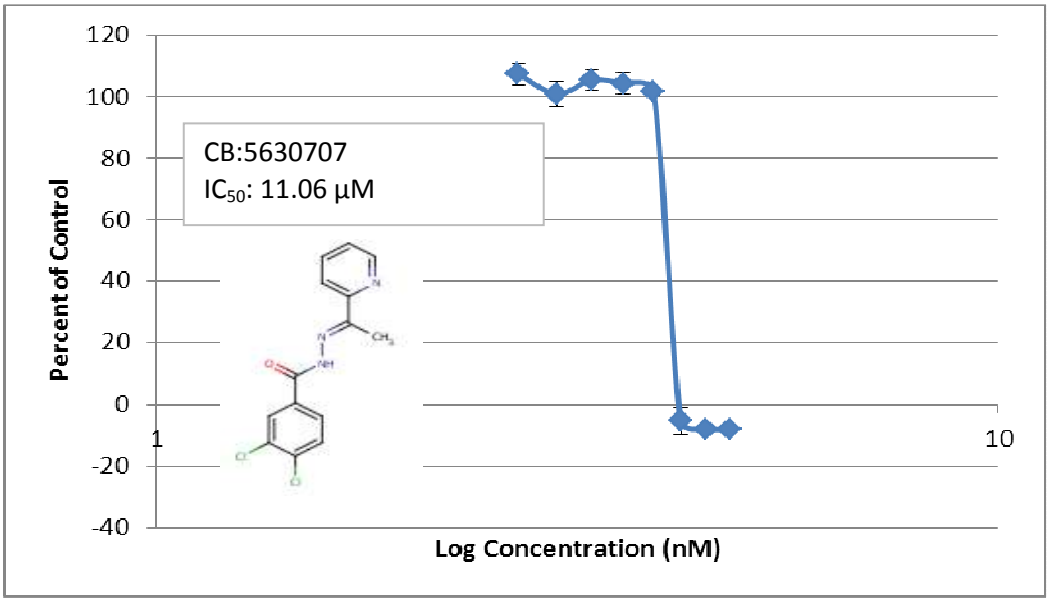


Figure 1. continued.

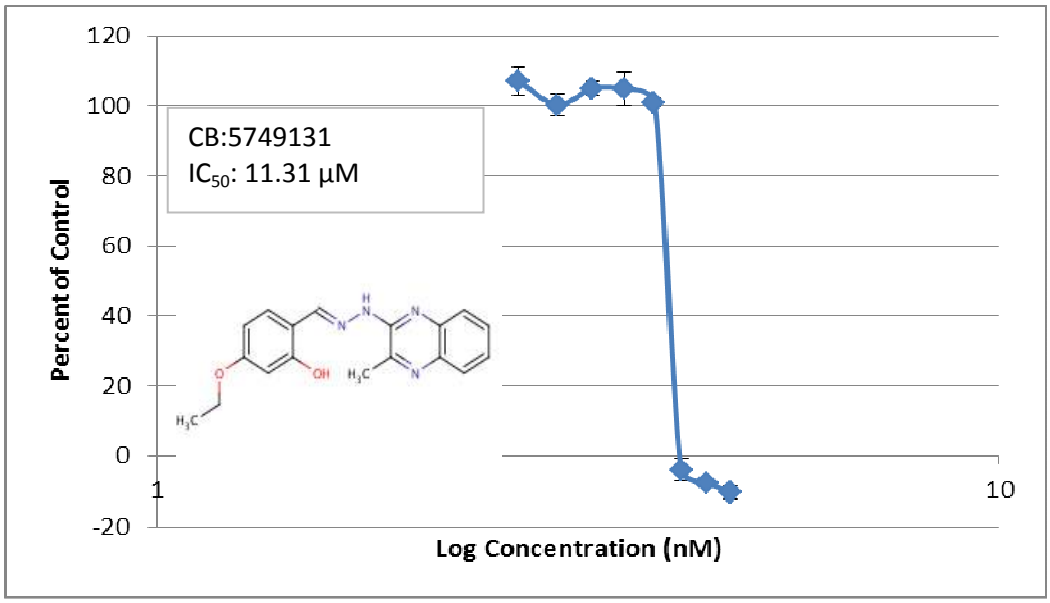


Figure 1. continued.

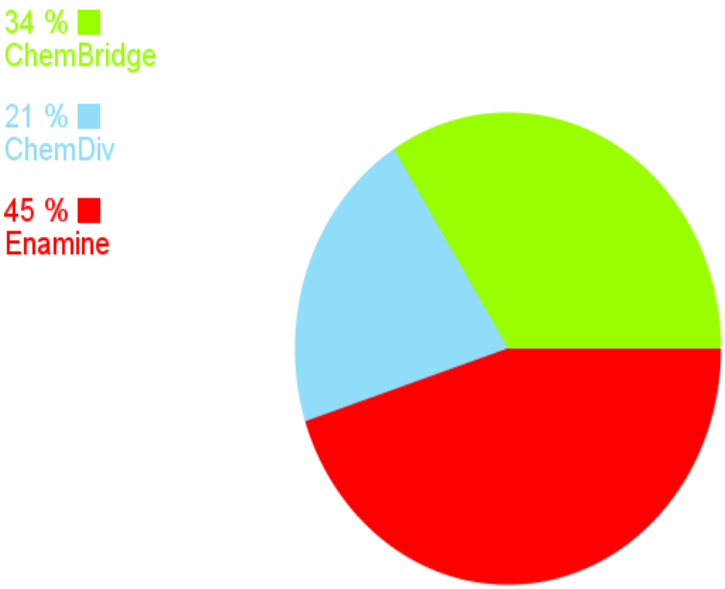


Figure 2. Sac-1 Diversity Library total compounds by vendor.

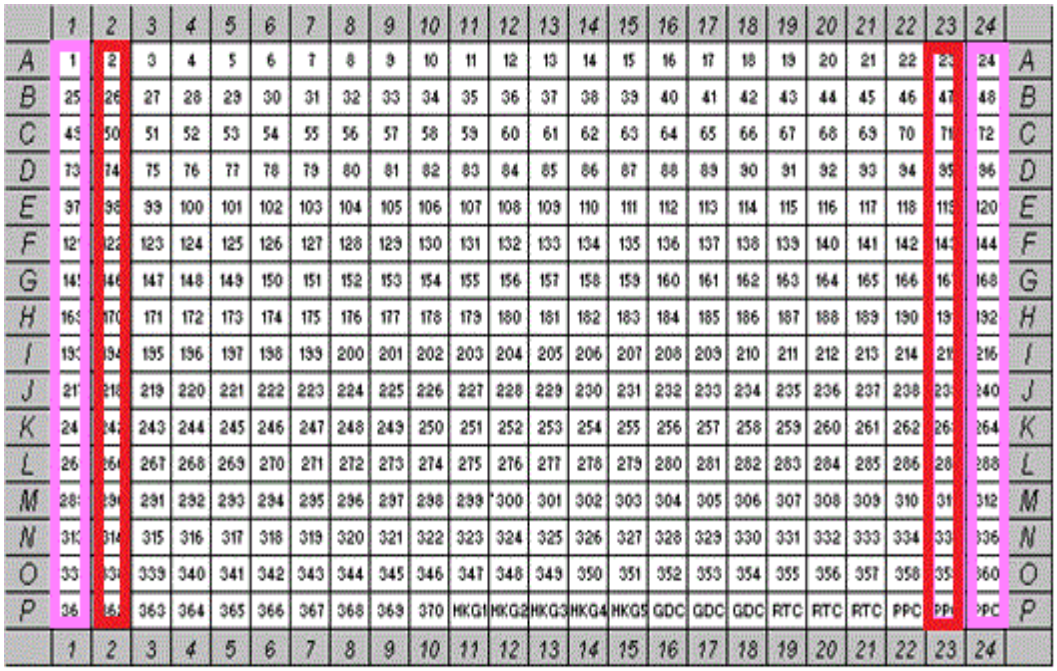


Figure 3. General layout of 384 well HTS Plate.

The pink columns *(1 and 24)* hold the "max" controls, which are wells that contain no drugs. After optimization we were pleased to find that the OD was ~ 0.4. The red columns *(2 and 23)* hold the minimum controls, normally Rifampicin.

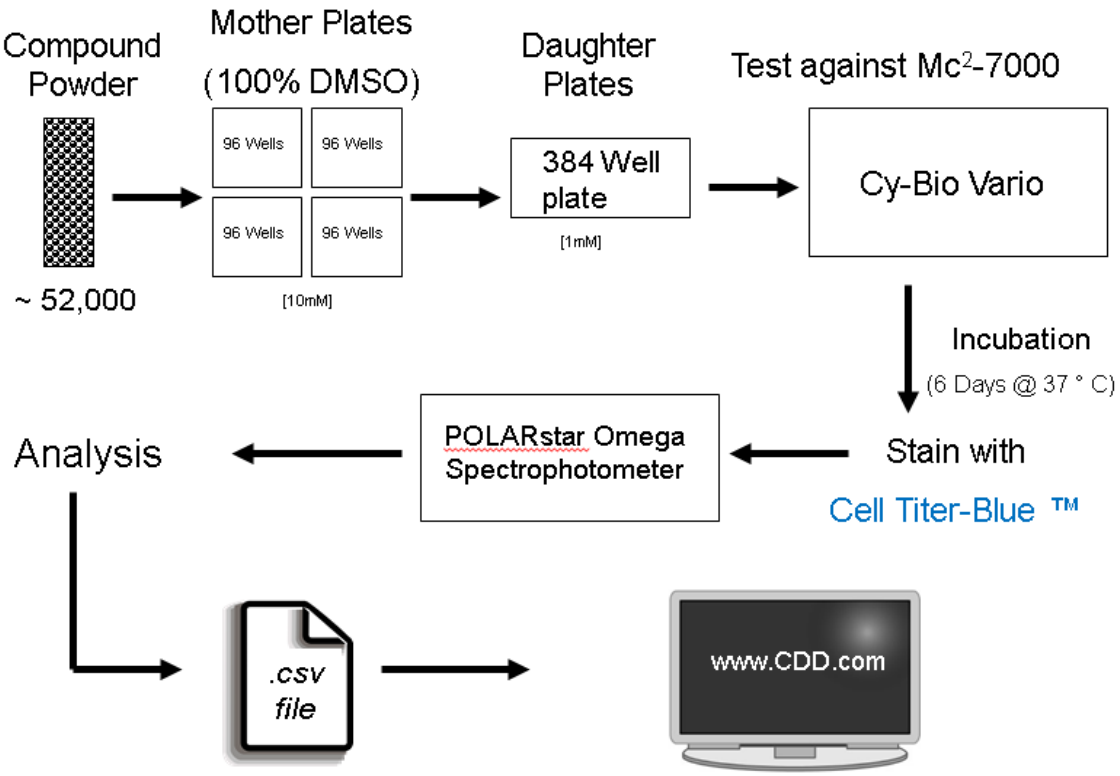


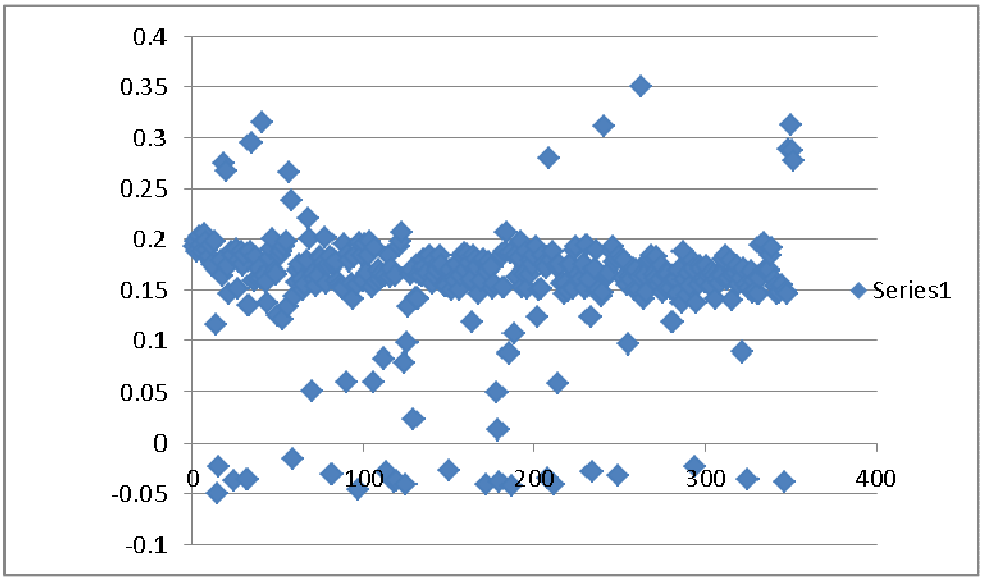
Figure 4. HTS Flow Chart.

User: USER			Path: C:\Program Files\BMG\Omega\User\Data\						File Name: 1376.dbf			
Test Name: ALAMR_BLUE-2POINTS							Date: 6/10/2009		Time: 12:07:51 PM			
ID1: Steven Diversity Library Plate 115												
Absorbance			Absorbance values are displayed as OD									
Well Row	Well Col	Content	Raw Data 573 - Raw Data 605 (-)			mean	calc by pop mean	% ctrl	variance	variance sq	sum of variance/n-1	standard dev
A	1	Sample X1	0.192			0.152767	1.256815935	125.6816	0.039233	0.001539228	0.006924276	0.083212237
A	3	Sample X3	0.197				1.289545517	128.9546	0.044233	0.001956558		
A	4	Sample X4	0.187				1.224086354	122.4086	0.034233	0.001171898		
A	5	Sample X5	0.203				1.328821015	132.8821	0.050233	0.002523354		
A	6	Sample X6	0.194				1.269907768	126.9908	0.041233	0.00170016		
A	7	Sample X7	0.187				1.224086354	122.4086	0.034233	0.001171898		
A	8	Sample X8	0.205				1.341912848	134.1913	0.052233	0.002728286		
A	9	Sample X9	0.197				1.289545517	128.9546	0.044233	0.001956558		
A	10	Sample X1	0.189				1.237178186	123.7178	0.036233	0.00131283		
A	11	Sample X1	0.182				1.191356772	119.1357	0.029233	0.000854568		
A	12	Sample X1	0.194				1.269907768	126.9908	0.041233	0.00170016		
A	13	Sample X1	0.175				1.145535358	114.5535	0.022233	0.000494306		
A	14	Sample X1	0.197				1.289545517	128.9546	0.044233	0.001956558		
A	15	Sample X1	0.116				0.759326294	75.93263	-0.036767	0.001351812		
A	16	Sample X1	-0.05				-0.327295817	-32.7296	-0.202767	0.041114456		
A	17	Sample X1	-0.024				-0.157101992	-15.7102	-0.176767	0.031246572		
A	18	Sample X1	0.178				1.165173107	116.5173	0.025233	0.000636704		
A	19	Sample X1	0.164				1.073530278	107.353	0.011233	0.00012618		
A	20	Sample X2	0.274				1.793581074	179.3581	0.121233	0.01469744		
A	21	Sample X2	0.267				1.74775966	174.776	0.114233	0.013049178		
A	22	Sample X2	0.173				1.132443525	113.2444	0.020233	0.000409374		
A	24	Sample X2	0.147				0.962249701	96.22497	-0.005767	3.32583E-05		

Figure 5. Sample Data sheet from 20 µM run.

*Sample data sheet not indicative of true data sheet size.

(Abs)



(Hit number)
Sample scatter plot from 20 μM run

Figure 5. continued.

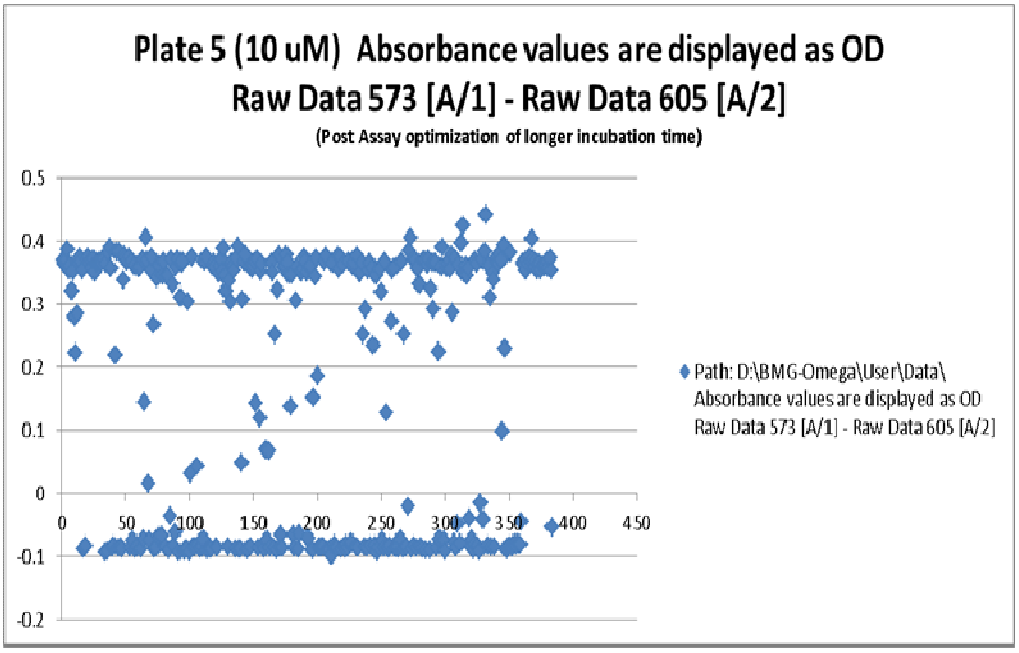


Figure 6. Example Scatter Plot (post assay optimization).

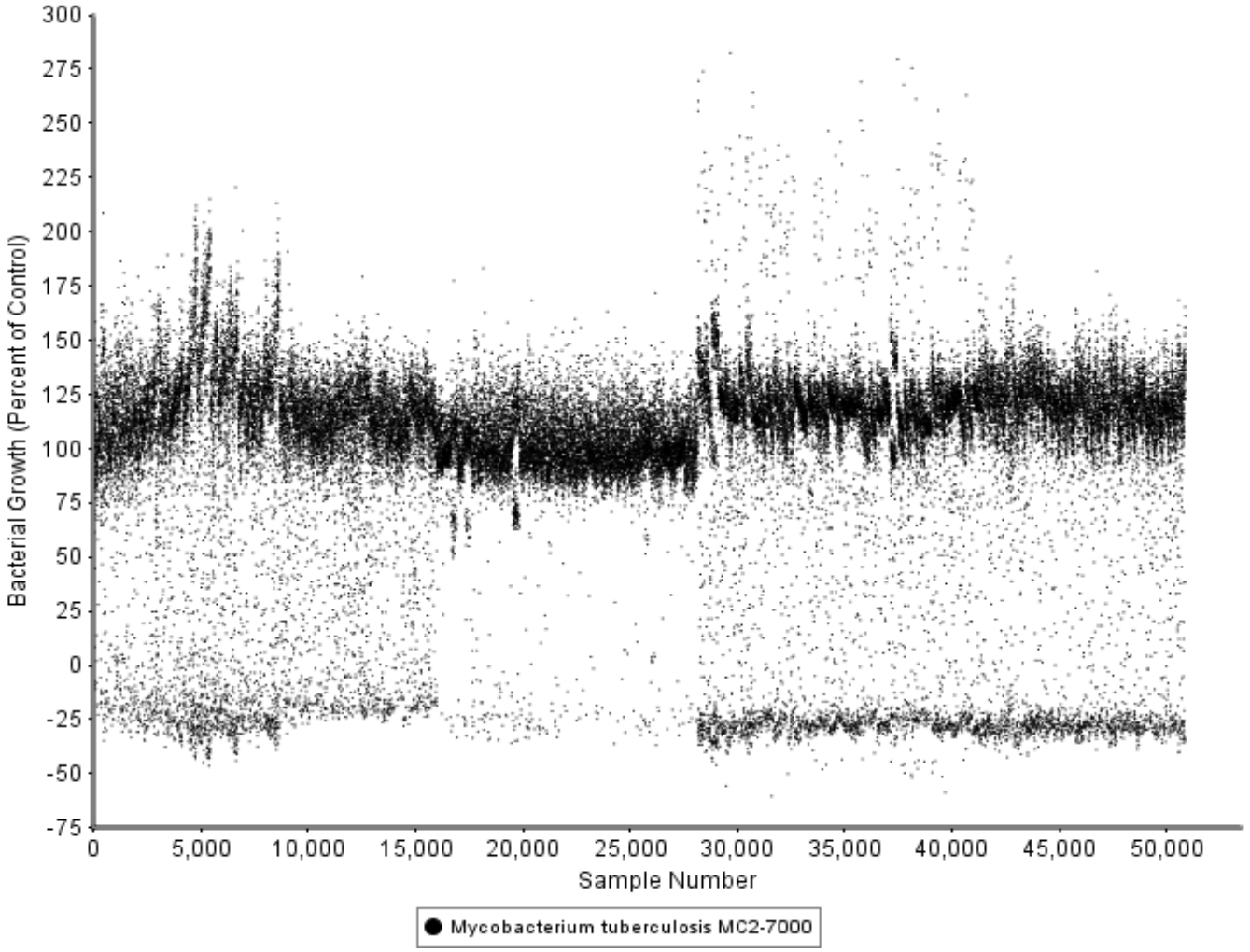


Figure 7. Whole Cell Screen of Sac-1 Diversity Library at 20 µM.

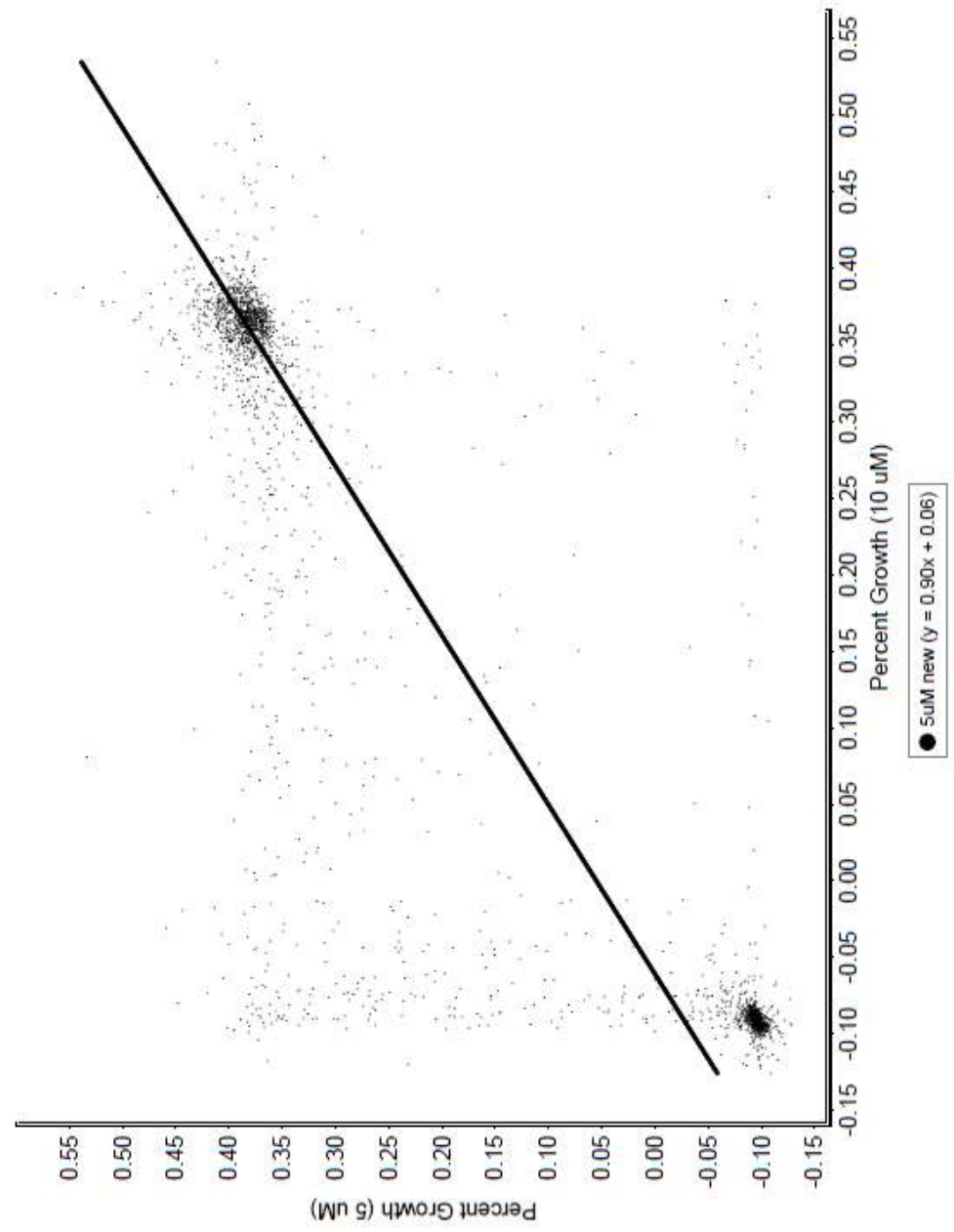


Figure 8. Assay Comparison between 10 µM run and 5 µM run.

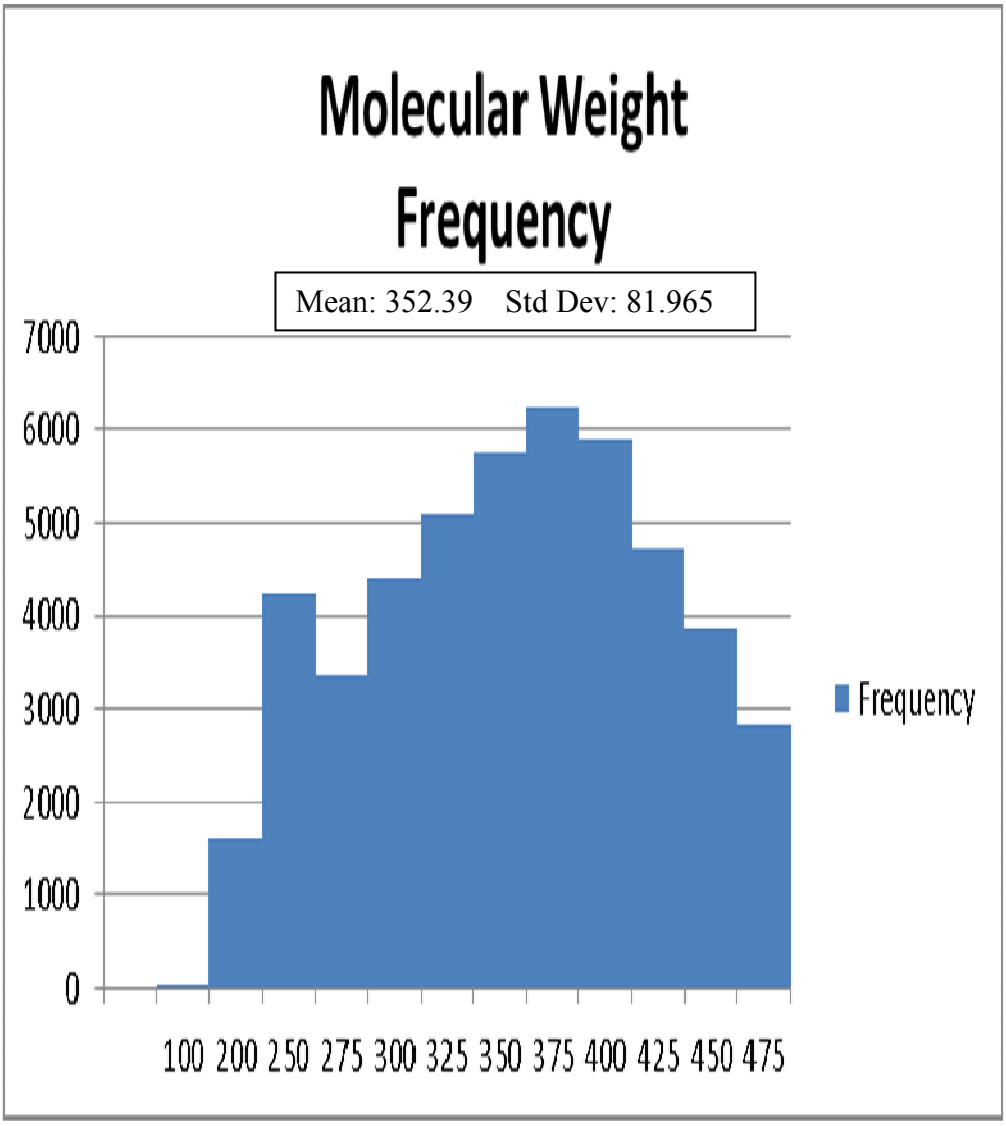


Figure 9. Library Characteristics.

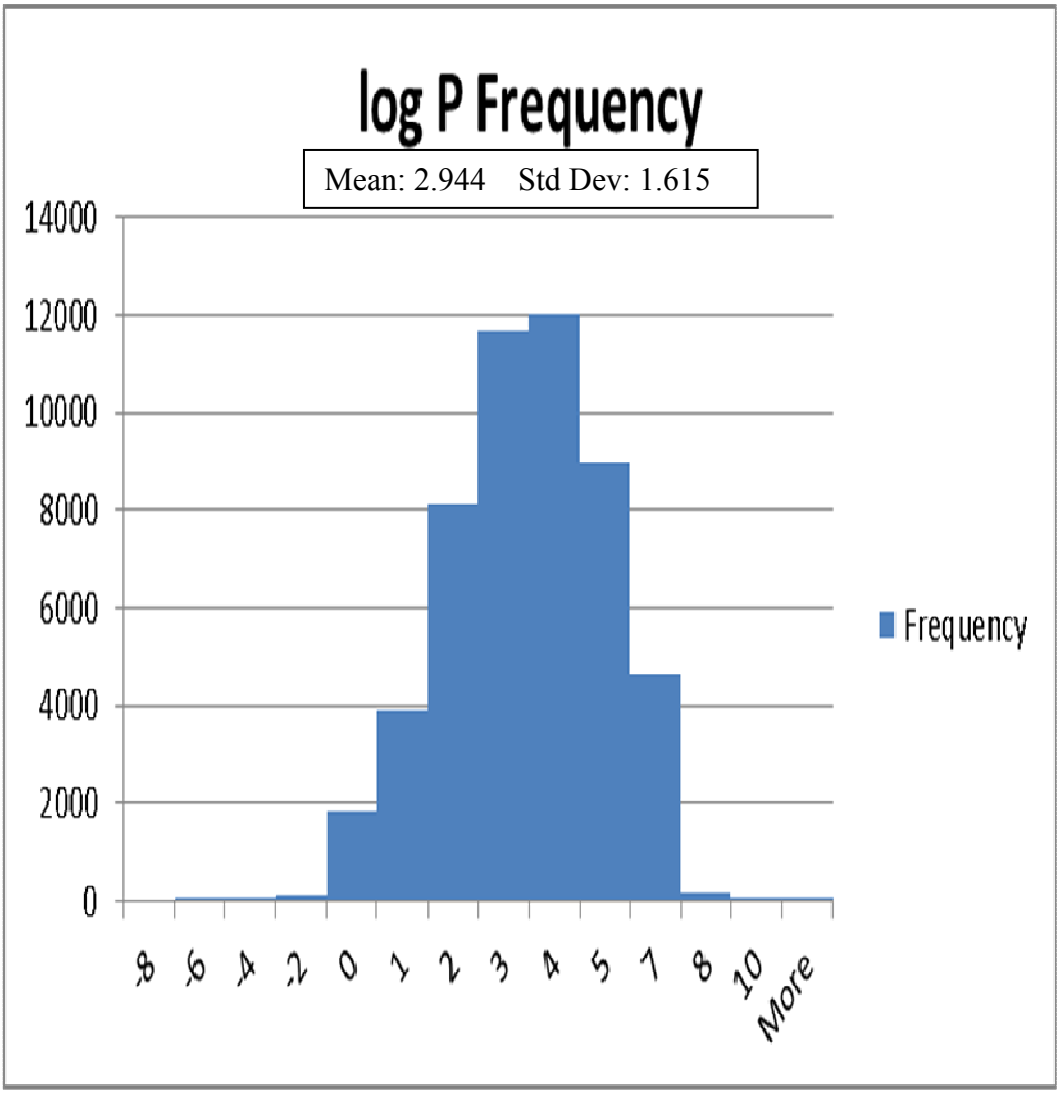


Figure 9. continued.

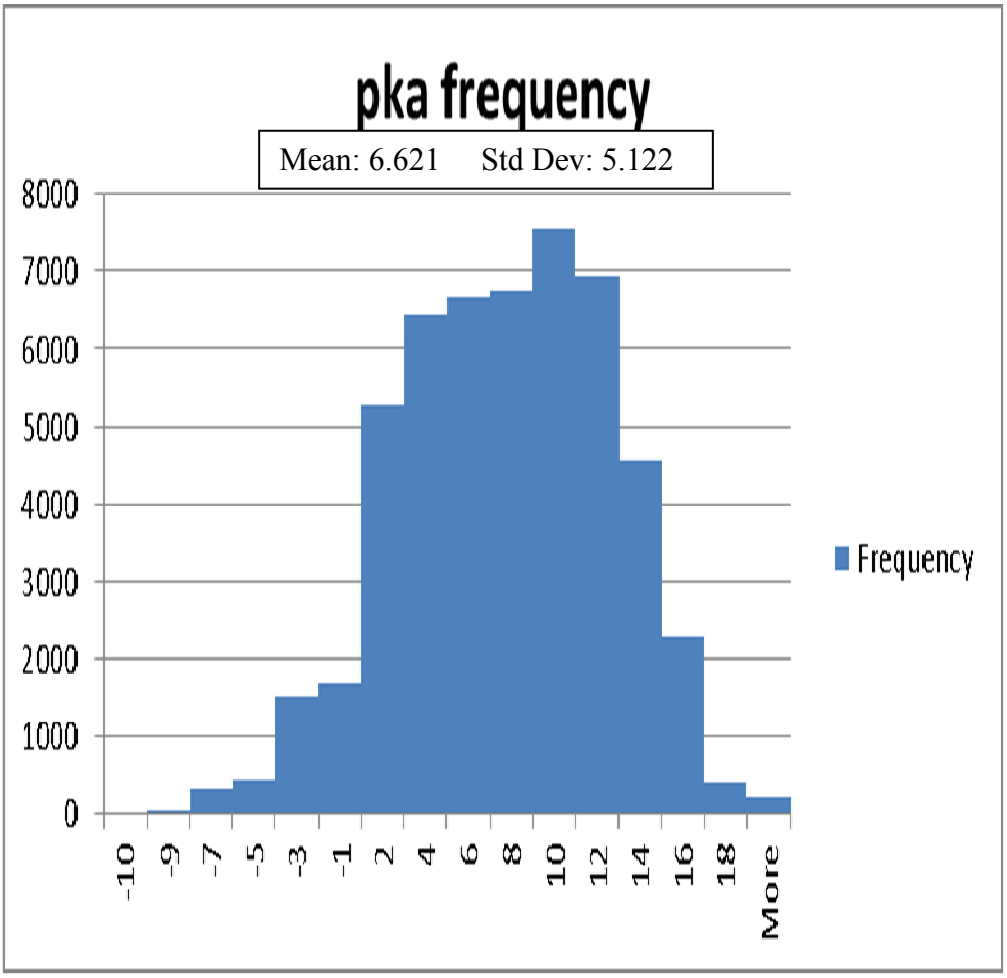


Figure 9. continued.

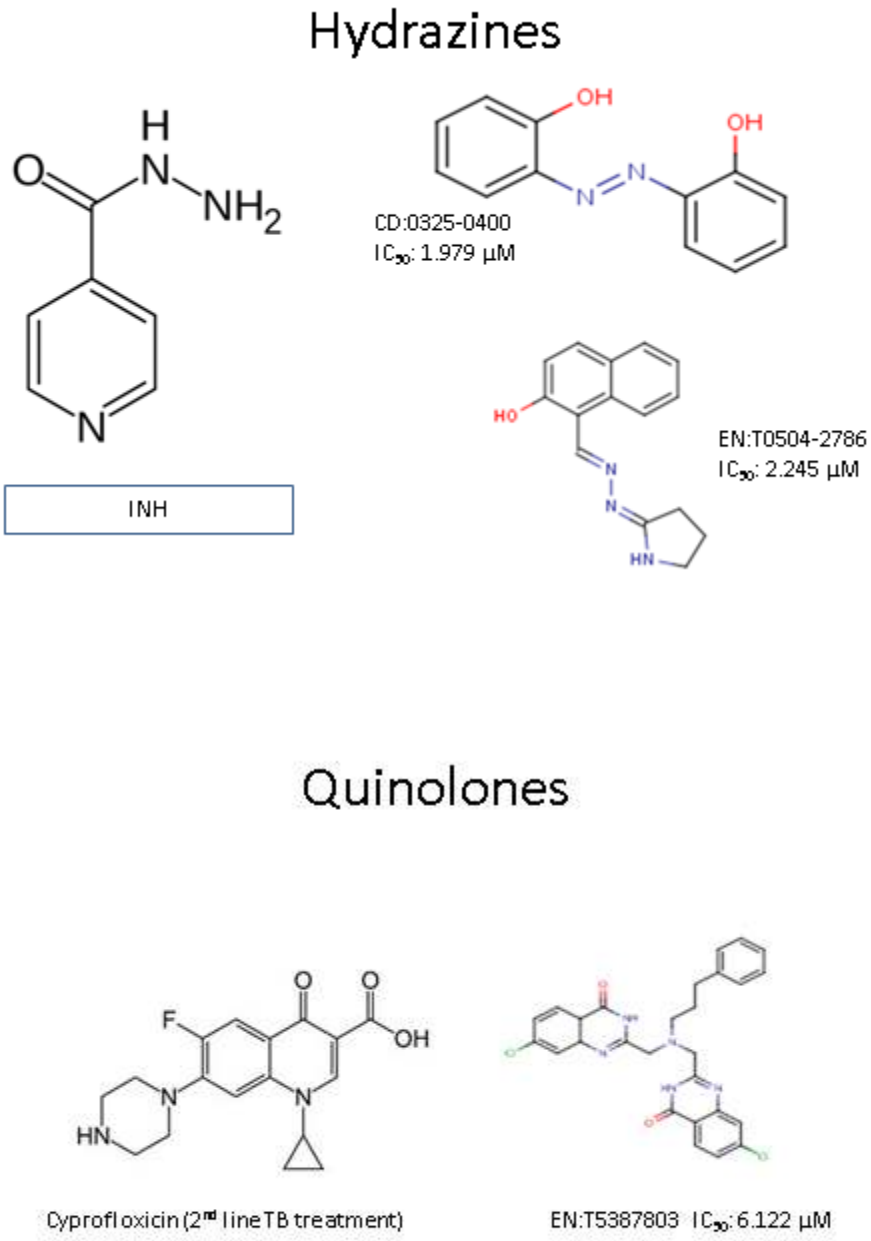


Figure 10. Notable Compounds.

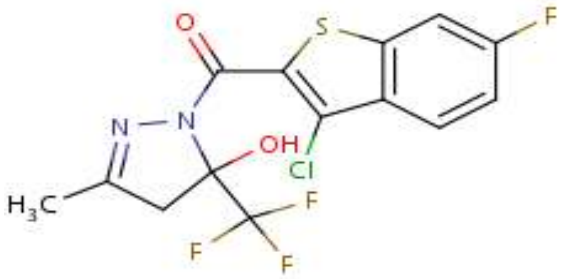
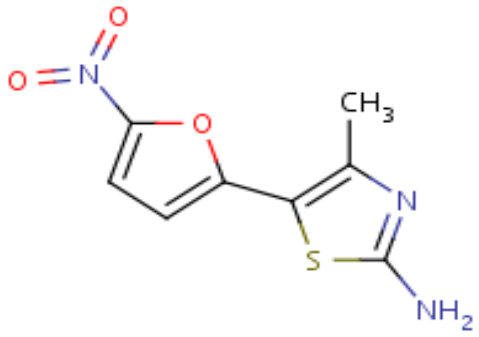
Molecule Name	Structure	IC_{50} (μM)
CD:6049-0348		11.28
EN:T0505-4171		11.99
EN:T5964414		12.59

Figure 11. Less Potent Compounds.

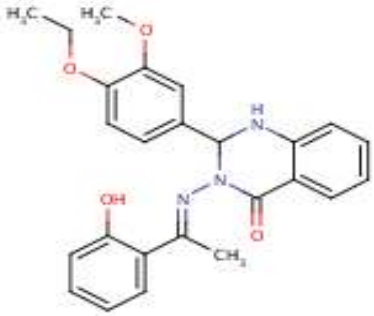
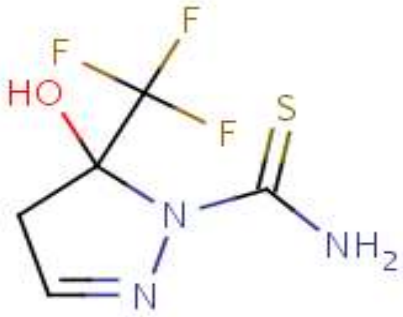
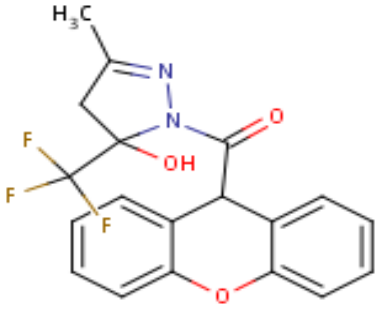
CB:7203800		13.97
EN:T0507-6380		14.89
CB:5559581		15.00
CD:6049-0117		16.47

Figure 11. continued.

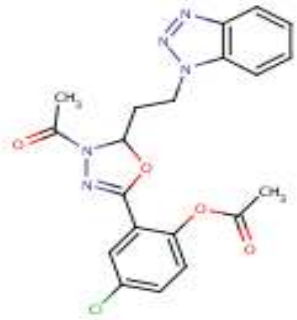
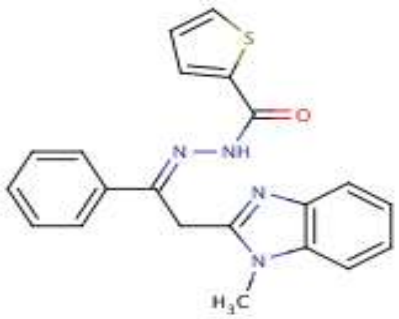
CB:5305713		17.14
CD:G642-6329		17.17
EN:T5315881		17.70
EN:T5312841		17.92

Figure 11. continued.

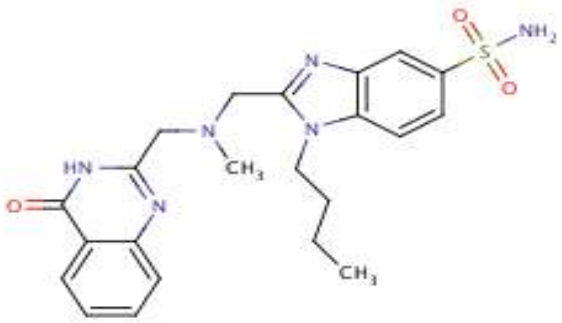
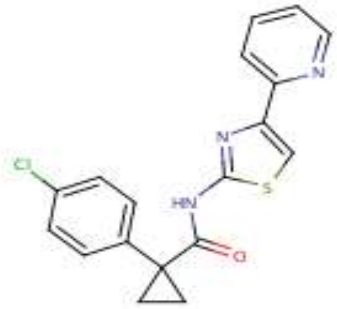
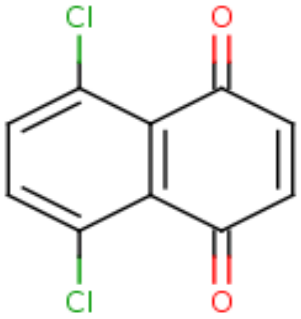
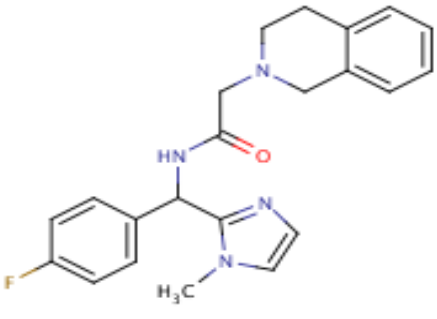
EN:T5442158		18.14
EN:T5626972		18.49
CD:R091-0019		20.78
EN:T5801947		24.07

Figure 11. continued.

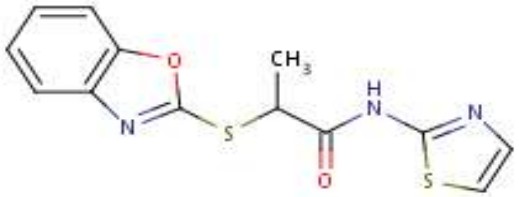
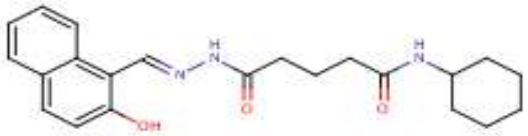
CB:9126968		26.30
CD:8008-8837		27.31

Figure 11. continued.

VITA

Name:	Steven Kenneth LaiHing
Address:	Dept. of Biochemistry and Biophysics-Texas A&M University 103 Biochemistry Bldg. 2128 TAMU College Station, TX 77843-2128
Email Address:	slaihing@gmail.com
Education:	B.S., Biochemistry, Oakwood University, 2006 M.S., Biochemistry, Texas A&M University, 2011