

## Comprehensive Mechanistic Analysis of Hits from High-Throughput and Docking Screens against $\beta$ -Lactamase

Kerim Babaoglu,<sup>†</sup> Anton Simeonov,<sup>‡</sup> John J. Irwin,<sup>†</sup> Michael E. Nelson,<sup>‡</sup> Brian Feng,<sup>†</sup> Craig J. Thomas,<sup>‡</sup> Laura Cancian,<sup>||</sup> M. Paola Costi,<sup>||</sup> David A. Maltby,<sup>§</sup> Ajit Jadhav,<sup>‡</sup> James Inglese,<sup>‡</sup> Christopher P. Austin,<sup>\*,‡</sup> and Brian K. Shoichet<sup>\*,†</sup>

Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94158-2330, NIH Chemical Genomics Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892-3370, Mass Spectroscopy Facility, Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143-0446, and Dipartimento di Scienze Farmaceutiche, Università degli Studi di Modena e Reggio Emilia, Via Campi 183, 41100, Modena, Italy

Received December 2, 2007

High-throughput screening (HTS) is widely used in drug discovery. Especially for screens of unbiased libraries, false positives can dominate “hit lists”; their origins are much debated. Here we determine the mechanism of every active hit from a screen of 70,563 unbiased molecules against  $\beta$ -lactamase using quantitative HTS (qHTS). Of the 1274 initial inhibitors, 95% were detergent-sensitive and were classified as aggregators. Among the 70 remaining were 25 potent, covalent-acting  $\beta$ -lactams. Mass spectra, counter-screens, and crystallography identified 12 as promiscuous covalent inhibitors. The remaining 33 were either aggregators or irreproducible. No specific reversible inhibitors were found. We turned to molecular docking to prioritize molecules from the same library for testing at higher concentrations. Of 16 tested, 2 were modest inhibitors. Subsequent X-ray structures corresponded to the docking prediction. Analog synthesis improved affinity to 8  $\mu$ M. These results suggest that it may be the physical behavior of organic molecules, not their reactivity, that accounts for most screening artifacts. Structure-based methods may prioritize weak-but-novel chemotypes in unbiased library screens.

### Introduction

High-throughput screening (HTS<sup>a</sup>) is widely used to discover new chemical structures for lead discovery in pharmaceutical research and for probe discovery in chemical genomics. A critical challenge in HTS is the occurrence of “false-positives”, that is, molecules that are active in the assay that operate via irrelevant mechanisms. For example, molecules may interfere with the assay signal,<sup>1</sup> act as oxidants,<sup>2</sup> chemically react with targets,<sup>1–3</sup> or form promiscuous aggregates<sup>4,5</sup> that nonspecifically inhibit the target. Several methods have been introduced to predict likely false positives based on these mechanisms,<sup>6,7</sup> leading to removal of compounds with unfavorable physical or chemical properties from screening libraries. Screening compounds in replicate, and more recently, quantitative high-throughput screening (qHTS) in titration<sup>8</sup> can decrease the incidence of false-positives. Finally, statistical analyses to more accurately model HTS results have been developed.<sup>9</sup>

Despite these developments, false positives continue to dog HTS, and their mechanisms have rarely been systematically quantified. Because the goal of most HTS campaigns is to find

new compounds for chemical optimization, false-positives may be tolerated as long as such leads are truly found. However, the lack of quantification has often left us at the mercy of intuition in prioritizing screening actives and prevented a comprehensive definition of the compounds that interact with a given target.

Another issue facing screening for both chemical biology and drug discovery is the exploration of novel chemotypes for new genomic and what are now “undruggable” targets. Though new technologies<sup>10–12</sup> allow testing of over a million compounds in a reasonable time, this represents an infinitesimal fraction of possible drug-like or lead-like compounds<sup>13,14</sup> When available, known ligands of the target can be used to bias the chemical structures in the screening library, increasing the likelihood of actives being identified, but this can be at the expense of finding novel scaffolds. Similarly, if the structure of the target has been determined, molecular docking screens can preselect compounds to be experimentally tested.<sup>15–20</sup> Theoretically, virtual methods could be used to prioritize screening libraries for testing, but the practical utility of this approach remains largely untested.

Here, we investigate these two central problems in screening through a detailed study of essentially all of the active molecules identified by quantitative high-throughput screening (qHTS), and prioritization and testing of a few chosen by docking, against a single target. Quantitative HTS and structure-based docking of over 70000 compounds were performed against the antibiotic resistance target  $\beta$ -lactamase, which has been extensively characterized enzymologically, crystallographically, and for some classes of false-positives in screening,<sup>4,5,21</sup> making it convenient to work with. We established the mechanism of every qHTS active, allowing us to determine which mechanisms were most responsible for artifactual hits and how often putatively reactive functional groups were problematic. X-ray crystal structures of enzyme–inhibitor complexes were deter-

\* To whom correspondence should addressed. Phone: 301-217-5725 (C.P.A.); 415-514-4126 (B.K.S.). Fax: 301-217-5736 (C.P.A.); 415-514-4260 (B.K.S.). E-mail: austinc@mail.nih.gov (C.P.A.); shoichet@cgl.ucsf.edu (B.K.S.).

<sup>†</sup> Department of Pharmaceutical Chemistry, University of California San Francisco.

<sup>‡</sup> NIH Chemical Genomics Center.

<sup>||</sup> Università degli Studi di Modena e Reggio Emilia.

<sup>§</sup> Mass Spectroscopy Facility, Department of Pharmaceutical Chemistry, University of California San Francisco.

<sup>a</sup> HTS, high-throughput screening; qHTS, quantitative high-throughput screening; NCGC, NIH Chemical Genomics Center; MLSMR, Molecular Libraries Screening Molecular Repository; MDH, malate dehydrogenase; NADP, nicotinamide adenine dinucleotide phosphate; MW, molecular weight; Da, daltons; rmsd, root mean squared deviation; Z-FR-AMC, Z-Phe-Arg-aminomethyl coumarin.

mined for several interesting molecules. We used the structure-based docking screen to prioritize weak-but-novel chemotypes, passed over by the experimental screen owing to lack of potency, for testing at concentrations higher than ordinarily feasible in high-throughput. The results from these studies reflect on the origins of false-positive actives in screening and offer guidance for novel scaffold discovery from screens of unbiased libraries.

## Results

We began with a quantitative HTS<sup>8</sup> campaign versus  $\beta$ -lactamase. Each of 70563 compounds in the NIH Chemical Genomics Center (NCGC) library, primarily composed of the Molecular Libraries Screening Molecular Repository (MLSMR), was assayed in concentration series for enzyme inhibition, as described.<sup>22</sup> All compounds were screened in two separate assays, once in the presence and once in the absence of the detergent 0.01% Triton X-100; in each assay, inhibition of every compound was measured no less than seven times over a concentration range from 4 nM to 30  $\mu$ M. Assaying in the presence and absence of detergent tested for aggregation-based inhibition, which detergent disrupts.<sup>23</sup> Compounds were designated as active if they produced a curve class of  $-3$  or better. These classes reflect the quality of the dose-response curves that are measured across the concentration range used by qHTS; a negative class reflects inhibition, whereas a positive curve class reflects activation of the target (the latter of which is not of interest here). The most reliable curves are in class  $-1$ , which have well-defined upper and lower baselines, and the least reliable are those in class  $-3$ , which have a single point of statistically significant inhibition at the highest concentration assayed.<sup>8</sup> For instance, a compound that begins to show inhibition at the second-lowest concentration measured (here about 20 nM) and has achieved full inhibition by the fifth concentration of the seven measured (here about 1  $\mu$ M), would have a curve class of  $-1$ , because it achieves full inhibition with full upper and lower baselines. A compound for which inhibition is only observed at the very last, highest concentration, here 30  $\mu$ M, would have a curve class of  $-3$ . Several control compounds that are known reversible inhibitors of  $\beta$ -lactamase were seeded into the screening collection and all were identified with appropriate  $IC_{50}$ s and curve classes of  $-1$  or  $-2$ . Of 1274 active molecules in the detergent-negative screen, 1204 (95%) lost activity on addition of detergent, suggesting that these molecules inhibited via promiscuous aggregation. These results were largely confirmed in secondary assays: of 17 detergent-sensitive molecules, 15 were confirmed by careful one-at-a-time assays; the two that did not were not observed to inhibit and thus do not belong in any class of inhibition.<sup>22</sup> This left 70 detergent-insensitive inhibitors from the paired qHTS assays, of which 19 were curve class  $-1$ , 36 were curve class  $-2$ , and 15 were curve class  $-3$ ; all 70 were subjected to follow-up study.

Of the 70 detergent-insensitive inhibitors, 25 were  $\beta$ -lactams. These are well-known competitive substrates or inhibitors of  $\beta$ -lactamase, acting via covalent modification of the catalytic Ser64 of the enzyme. Whereas the activity of the  $\beta$ -lactams supported the robustness of the experimental screen, as covalent inhibitors of modest novelty we considered them uninteresting as new chemical probes. They are not considered further here. Of the remaining 45 actives, 28 were repurchased from their commercial suppliers and tested in secondary assays (below and Table S1, Supporting Information). The other 17 could not be acquired from the original suppliers and so were recovered from the original mother plates. Most of these 17 shared a common

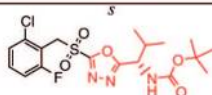
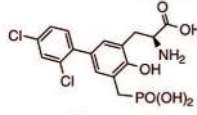
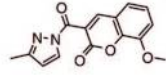
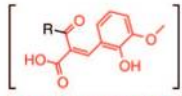
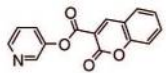
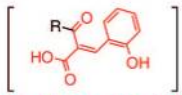
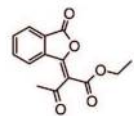
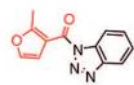
4-amino quinazoline or 4-amino pyridine core scaffold making up a related series. Three representatives of this series were resynthesized and also tested in the secondary assays.

Initially, all 45 non- $\beta$ -lactams were reassayed at 30  $\mu$ M, the highest concentration used in the qHTS, in a cuvette-based  $\beta$ -lactamase assay. A total of 25 compounds, including the three that were resynthesized, had no reproducible activity and were designated as false positives. Of note, all 25 had low-confidence dose-response curves, typically associated with single-point top-concentration activity near the three-sigma cutoff of the assay in the primary qHTS, so would not routinely be designated for follow-up.<sup>8</sup> The 20 non- $\beta$ -lactam compounds with reproducible activity in the cuvette assay were then tested for inhibition in the presence of 10-fold more detergent than was used in the qHTS, to identify colloidal aggregates that only disperse in high concentrations of detergent.<sup>23</sup> Nine such compounds were found and were classified as detergent-resistant promiscuous aggregators. Consistent with this view, these nine also inhibited the unrelated enzymes chymotrypsin, malate dehydrogenase (MDH), and cruzain in a detergent sensitive manner (Table S1, Supporting Information). We note that at 0.1% of Triton X-100 we are above the critical micelle concentration (CMC) of this detergent, and it is possible that detergent micelles are acting to sequester these nine inhibitors rather than acting to break up their aggregates, and that in fact they are working by another mechanism. Indeed, even 0.01% Triton X-100 is close enough to the CMC that one cannot be certain that micelles have not formed under the particular ionic strength of the assay, and if this is the case, other mechanisms of inhibition disruption, such as compound sequestration, may also be involved. Whereas we cannot now exclude this possibility, we do not favor it. These nine inhibitors were active against our full panel of enzymes, not simply the nucleophile active ones (as was the case, for instance, for the covalent-acting inhibitors described below), and this level of detergent does not reduce inhibition by known reversible or irreversible inhibitors of  $\beta$ -lactamase, including those we describe next.

The remaining 12 compounds were reproducible, nonaggregation-based inhibitors of  $\beta$ -lactamase. Six of these were readily identified as covalent inhibitors, partly based on their chemical structures but most definitively based on their inhibition patterns. All six inhibited not only  $\beta$ -lactamase but also at least one of a panel of counter-screen enzymes that included chymotrypsin, cruzain, and malate dehydrogenase (MDH; Table S1, Supporting Information). For all enzymes, inhibition was time-dependent. Most convincingly, five of these increased the molecular mass of  $\beta$ -lactamase by mass spectrometry as would be expected for irreversible covalent inhibitors (Table 1).

Among the most intriguing chemotypes revealed by qHTS are a collection of sulfone oxadiazoles that inhibited the enzyme potently, with  $IC_{50}$  values ranging from 0.015 to 8  $\mu$ M in the primary qHTS. The two most potent of these were retested against our panel of enzymes. Whereas the NADP-dependent enzyme MDH was uninhibited, inhibition was observed against the serine- and cysteine-active proteases chymotrypsin and cruzain (Table 2). As with  $\beta$ -lactamase, both enzymes showed time-dependent inhibition, consistent with a covalent mechanism of action. To test this hypothesis in detail, mass spectra were again taken. After preincubation with compound **1**, the MW of  $\beta$ -lactamase increased by 239 Daltons relative to the apoenzyme, suggesting displacement of the sulfone and labeling of the enzyme with the oxadiazole half of the molecule (Table 1). To investigate this at atomic resolution, an X-ray structure of the complex between  $\beta$ -lactamase and **1** was obtained by soaking

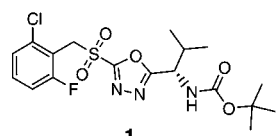
**Table 1.** Covalent Inhibitors, Change in Protein Mass and Apparent Covalent Adduct (in Red)

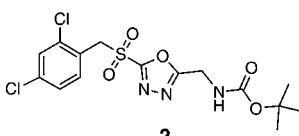
Compound #	NCGC ID #	Mol. Weight [M+H] (a.m.u.)	Change in Protein Mass (a.m.u.)	Structure	Assumed Fragment Mass (a.m.u.)
1	NCGC00067197	448	239		Exact Mass: 240.1348
5	NCGC00025080	421	421; 842; 1,261; 1,681*		Exact Mass: 419.0092
6	NCGC00054235	284	221		 Exact Mass: 221.0450
7	NCGC00062188	267	191		 Exact Mass: 191.0344
8	NCGC00070387	260	653; 781; 911; 1,042; 1171**		Exact Mass: 260.0685
9	NCGC00071986	227	109		Exact Mass: 109.0290

\* Change in mass based upon addition of 1–4 equivalency of the entire molecule. \*\* Repeating units of approx 130 atomic mass units added incrementally between 5–9 times.

**Table 2.** Inhibitory Activity of Sulfonyl-Oxadiazoles **1** and **2** versus AmpC, Chymotrypsin, MDH, and Cruzain in the Presence and Absence of Triton X-100

Compound #	IC <sub>50</sub> (μM) AmpC		IC <sub>50</sub> (μM) Chym		IC <sub>50</sub> (μM) MDH		IC <sub>50</sub> (μM) Cruzain*	
	No TX	0.1% TX	No TX	0.1% TX	No TX	0.1% TX	No TX	0.01% TX
1	< 1 μM	< 1 μM	< 1 μM	< 1 μM	> 600 μM	> 600 μM	< 1 μM	< 1 μM
2	< 1 μM	< 1 μM	< 1 μM	< 1 μM	> 600 μM	> 600 μM	95 μM	410 μM

  
**1**

  
**2**

\* Cruzain only tested at 0.01% Triton.

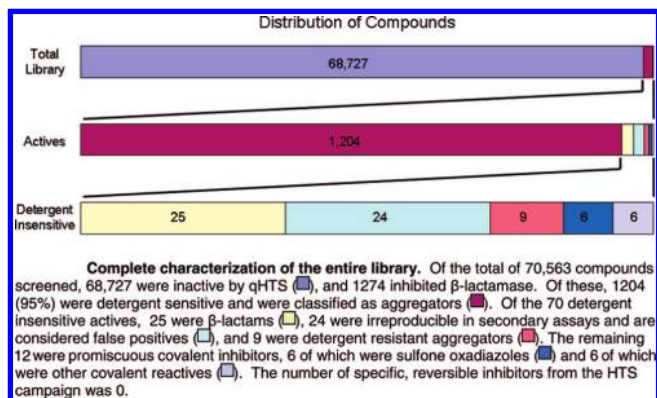
apo crystals with the inhibitor. In the 2.0 Å resolution structure, unambiguous electron density was observed showing the attachment of the oxadiazole ring to Ser64, which has displaced the sulfone moiety leaving a 239 Da fragment covalently attached to the enzyme (Figure 2).

In summary, detailed analysis allowed us to categorize all the active molecules from qHTS as either promiscuous aggregators (95%), known inhibitors (2%), irreproducible (2%), or promiscuous covalent (1%; Figure 1, Table 3, and Table S1, Supporting Information). All irreproducible actives produced poor-quality concentration–response curves in the qHTS, pre-saging the lack of robust inhibition on follow-up. No specific, reversible inhibitors were identified at the limiting concentration of 30 μM.

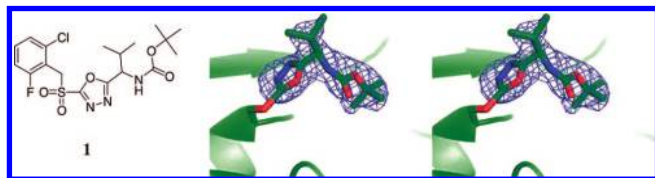
Concomitant to the screen and the detailed assessment of qHTS lead actives, we performed docking calculation on the same library, initially looking to compare the two techniques. In the absence of reversible inhibitors from the experimental screen, we turned to the docking to prioritize low affinity molecules in the library that might have been missed by the maximum concentration used in the qHTS. Each molecule was docked in several thousand orientations into a holo-structure

of β-lactamase using DOCK3.5.54; for each ligand orientation, up to 1000 ligand conformations were explored. The complementarity of every ligand configuration was evaluated for van der Waals and electrostatic interaction energies with the enzyme, corrected for ligand desolvation.<sup>24</sup> To calibrate the docking scores, we also docked known, reversible micromolar inhibitors of the enzyme, of which 20 or so are known, all analogs of a single chemotype.<sup>25,26</sup> These known inhibitors ranked among the top-scoring 500 molecules out of the 66661 docked. We thus considered these 500 as reasonable candidates for further evaluation, and each was visually examined for fit in the binding site. Based on this evaluation, 16 molecules were purchased and tested for inhibition of β-lactamase in low-throughput assays. In addition to being among the top 500 docking this, these molecules were chosen based on chemical diversity and visually compelling complementarity to the active site, weighting aspects not strictly captured by the docking calculation (for instance, formation of geometrically favorable hydrogen bonds, where DOCK3.5.54 simply calculates electrostatic interaction energies with no heuristics to weight orbital overlap). A total of 2 of the 16, compounds **3** and **4**, ranked 97th and 202nd out of 66661 docked molecules, inhibited β-lactamase with IC<sub>50</sub>





**Figure 1.** Complete characterization of the entire library. Of the total of 70563 compounds screened, 68727 were inactive by qHTS (purple) and 1274 inhibited  $\beta$ -lactamase. Of these, 1204 (95%) were detergent-sensitive and were classified as aggregators (magenta). Of the 70 detergent-insensitive actives, 25 were  $\beta$ -lactams (yellow), 24 were irreproducible in secondary assays and are considered false positives (sky blue), and 9 were detergent-resistant aggregators (orange). The remaining 12 were promiscuous covalent inhibitors, 6 of which were sulfone oxadiazoles (blue) and 6 of which were other covalent reactives (light purple). The number of specific, reversible inhibitors from the HTS campaign was 0.



**Figure 2.** Structure of the AmpC, **1** (oxadiazole), adduct to 2.0 Å resolution. Stereoview unbiased  $F_o - F_c$  density is shown around serine 64 contoured at 3 sigma with fragment of **1** modeled. Detailed crystallographic statistics are given in Supporting Information, Table S3.

values of 140 and 210  $\mu$ M, respectively, both in the presence and the absence of 0.1% Triton X-100 (Table S2, Supporting Information). These  $IC_{50}$  values were 5- to 7-fold higher than the maximum concentration screened in the qHTS and so could not have been found via the experimental approach. Nevertheless, both compounds appeared specific: neither inhibited chymotrypsin or cruzain detectably (Table S2, Supporting Information) and both were found to be competitive inhibitors of the enzyme by varying substrate concentrations in the assay; the  $K_i$  values for what were racemic compound mixtures (below) of the original docking hits were 70 and 105  $\mu$ M, respectively.

An enzyme-bound structure of one of these docking hits, **3**, was obtained by cocrystallization with  $\beta$ -lactamase (Figure 3). The ligand was clearly visible in the active site based on initial unrefined  $F_o - F_c$  difference maps. The crystallographic pose closely resembled that predicted by docking, with an rmsd of 0.9 Å. The ligand makes several polar interactions with the enzyme, all but one of which, with a bridging water, were predicted in the initial docking screen. The amino-acid carboxylate is placed in the oxyanion hole of the enzyme, hydrogen bonding with the backbone nitrogen of Ala318; such an oxyanion-amide hydrogen bond is canonical in  $\beta$ -lactam- $\beta$ -lactamase structures.<sup>25</sup> Less typically, the ligand carboxylate also hydrogen bonds with the hydroxyl of Tyr150, which is thought to be the catalytic base of the enzyme. The phthalimide ring of the inhibitor hydrogen bonds with the nucleophilic Ser64 through one amide carbonyl, while the other carbonyl oxygen makes a water mediated hydrogen-bond with Asn343. The phthalamide carboxylate occupies the “distal carboxylate” site<sup>27</sup>

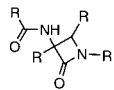
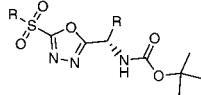
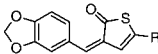
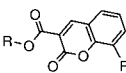
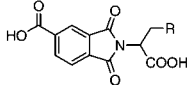
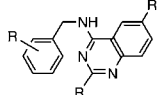
of the enzyme, hydrogen bonding with the backbone nitrogen of Gly320 and, via a water molecule, with the side chain of Ser212. The tyrosinyl side chain of the inhibitor fits against a hydrophobic wall of the active site formed by leucines 119 and 293.

These two molecules, though similar to each other, are chemically distinct from any other class of  $\beta$ -lactamase inhibitor, and represent a novel scaffold for this enzyme. Synthetic chemistry was therefore initiated to investigate their structure–activity relationships (SAR) and optimization potential. Only the *R*-isomer (the D-amino acid) is observed in the crystallographic electron-density. Consistent with this, only the *R*-enantiomer of **3** was noted to score well in the docking calculation (the *S*-enantiomer was not ranked within the top 500 hits). We note that the compound material evaluated in both the initial qHTS and in the docking-driven experiments was a commercially available racemic mixture. Consistent with best practices we had calculated both possible isomers docking assessment independently as the stereochemistry is not specified by the commercial supplier. To evaluate the issue we synthesized both the *R*- and *S*-enantiomers of **3**. Consistent with the structure and the docking, the *R*-enantiomer was 4-fold more potent than the *S*-isomer, with a  $K_i$  of 37  $\mu$ M (Table 4). To further explore the SAR, two commercially available<sup>28</sup> analogs (**11** and **15**) were purchased and tested, and seven enantiomerically pure analogs were synthesized (Table 4). Loss of the distal carboxylate (**11**) reduced affinity 4-fold, whereas removal of the amino-acid carboxylate (**12**) or replacing it with a bioisosteric tetrazole (**13**) reduced affinity (10- and 5-fold, respectively), consistent with the central role of the oxyanion hole interaction. Interestingly, extending out the amino acid carboxylate using beta amino acids (**14**) was well-tolerated, showing a less than 2-fold loss in activity, opening future areas of design. Removal of the amino-acid side chain was the most detrimental modification, reducing affinity by almost 200-fold (**15**). Large aromatic side chains, such as tryptophan (**16**) and naphthyl-alanine (**17**), were preferred over large aliphatic groups such as methionine (**18**) or leucine (**19**), and the most potent analog was the 1-naphthyl-alanine derivative (**17**) with a  $K_i$  of 8  $\mu$ M, leaving its ligand efficiency neither improved nor diminished from the initial docking hit. To investigate whether these modifications fell into a true SAR series, further crystal structures were determined of AmpC in complex with the  $\beta$ -amino acid derivative (**14**) and with the naphthyl derivative (**17**). The  $\beta$ -amino acid represents somewhat of a departure from the rest of the series, with its carboxylate further from the phthalimide core; nevertheless, both it and **17** bind essentially identically to the initial lead, demonstrating a common binding mode for this novel family of inhibitors (Figure 4).

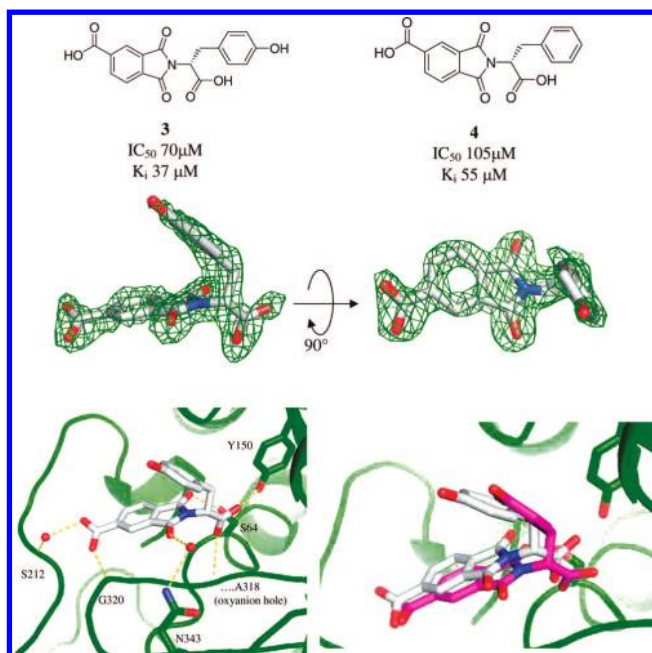
## Discussion

Three unexpected results emerge from this study. First, molecules with chemically reactive functional groups rarely inhibited the enzyme in the qHTS campaign. This is startling, given the effort devoted to identifying “reactive functionalities” in screening libraries.<sup>6,29</sup> Second, no specific, reversible inhibitors were identified in the experimental screen, likely reflecting the library’s relatively small size ( $\sim$ 70000 molecules)<sup>13,30</sup> and the lack of target bias in the chemotypes represented in it. The qHTS did unambiguously identify multiple covalent-acting  $\beta$ -lactams, the only compounds biased toward  $\beta$ -lactamase in the library (Table 3). It also rapidly identified molecules acting as promiscuous aggregates,<sup>22</sup> and this detergent-sensitivity screen may find wide use in the field; for assays that can tolerate

**Table 3.** Example of qHTS Identified Actives and Inactives and Scaffold Profiles

Structure Class	# of actives	# in collection	Potency Range[IC <sub>50</sub> ]	Lead Curve Class*	Classification of Active
	25	58	2 nM to 8 $\mu$ M	-1 (full curve)	covalent modifier
	9	9	16 nM to 8 $\mu$ M	-1 (full curve)	covalent modifier
	1	2	10 $\mu$ M	-2 (partial curve)	colloidal aggregate
	1	16	10 $\mu$ M	-2 (partial curve)	covalent modifier
	0	16	NA	NA	classified inactive (potency > 30 $\mu$ M)
	2	87	12 $\mu$ M	-2 (partial curve)	irreproducible

\* See ref 8 for an explanation of curve class designations.



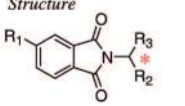













**Figure 3.** Docking discovers two novel  $\beta$ -lactamase inhibitors. (A) Structures of the two new inhibitors. (B) Initial unbiased  $F_o - F_c$  density contoured at 3 sigma from the 1.8 Å structure with compound **3** modeled. Full crystallographic statistics are given in Table S3, Supporting Information. (C) Hydrogen bonding pattern of **3** with AmpC. (D) Overlay of the crystal structure (white carbons) overlaid to the initial pose proposed by DOCK (magenta carbons). The rmsd of the heavy atoms between the two poses was 0.9 Å.

it, detergent addition will eliminate, if not all, then certainly most, colloidal aggregates before they occur. Third, molecular docking was able to prioritize specific and novel molecules whose initially very modest potencies made them opaque to the experimental screen. We consider each of these points in turn.

The few covalent hits from the qHTS, only 1% of the inhibitors and only 0.01% of the entire library, is startling and has important consequences given the effort devoted to identifying and eliminating such structures in screening libraries.<sup>6,29</sup> The one chemotype that occasionally did modify the enzyme, the sulfonyl-oxadiazoles, has received little attention, although the mass spectra and X-ray crystal structure of the oxadiazole enzyme-adduct leave little doubt as to its mechanism (Figure 2). Intriguingly, such sulfonyl-oxadiazoles are frequent hitters in the screening assay results in PubChem and are active in over 10% of the in vitro assays reported. Our determination of mechanism, combined with this overweighing among hits, suggests that this functional group, though not among those that are classically considered reactive, merits attention as a source of covalent artifact. Conversely, the dearth of such more classic covalent inhibitors might conceivably owe to a simple absence of them in the library. This does not seem to be the case; however, we identified 81 classes of classically reactive functionality in the NCGC library, including 289 molecules bearing Michael acceptors, 110 epoxides, aziridines and thioepoxides, 84 alkyl halides, 32 aldehydes, 31 dihydroxybenzenes, and close to 10000 esters, among others (Table 5 and Table S4, Supporting Information). Leaving esters and ketones aside, there were 2291 molecules bearing one of 79 “reactive” groups among the NCGC library, almost none of which inhibited  $\beta$ -lactamase.

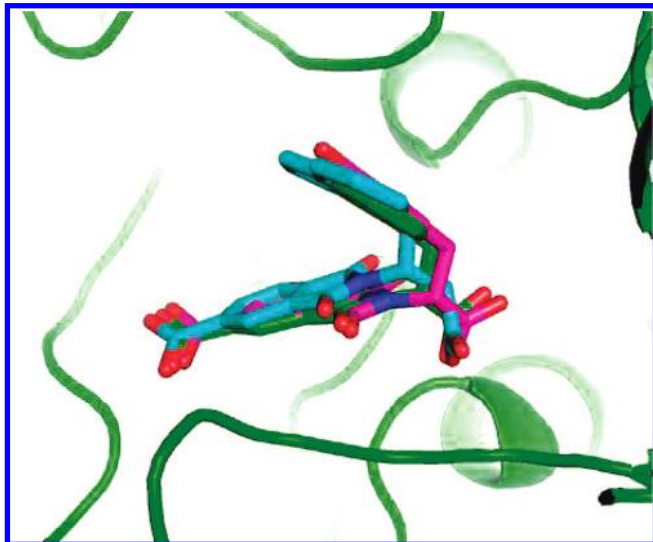
Another explanation for the few covalent inhibitors might be that  $\beta$ -lactamase is impervious to reactive chemistry. We cannot rule this out, but it is unlikely to fully explain the distribution of active molecules.  $\beta$ -Lactamase uses an activated serine nucleophile in its mechanism, which is the residue modified in the oxadiazole adduct, and whereas there may be enzymes more susceptible to covalent modification,  $\beta$ -lactamase is hardly inert. Of course, it may be that most of the over 2200 classically reactive molecules in the library are deactivated on exposure to solvent, or that at 30  $\mu$ M they have not reached a concentration sufficiently high to actually inactivate  $\beta$ -lactamase.













**Table 4.** Structure–Activity Relationships among the Phthalimide Series

Structure	Compound #	* Stereochemistry	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	K <sub>i</sub> [Apparent ( $\mu$ M)]
	3	racemic	COOH	COOH		140
	4	racemic	COOH	COOH		210
	3	R	COOH	COOH		37
	10	S	COOH	COOH		158
	11	racemic	H	COOH		830
	12	NA	COOH	H		360
	13	R	COOH			41
	14	R	COOH	—COOH		14
	15	NA	COOH	COOH	H	6900
	16	R	COOH	COOH		17
	17	R	COOH	COOH		8
	18	racemic	COOH	COOH		1300
	19	racemic	COOH	COOH		1200

mase.<sup>31</sup> These possibilities all may be illuminated by studies on more obviously sensitive proteins, such as thiol proteases, at higher compound concentrations. And of course the reluctance to either populate a library with reactive molecules or pursue them for follow-up may be sensible for downstream issues of hit-to-lead development, in vivo testing, and long-term library management. However, from a purely screening perspective, at least for a substantial portion of targets, reactive molecules may contribute little to direct artifacts from HTS, notwithstanding the attention lavished upon them.

The lack of specific, reversible actives identified from the experimental screen cannot be blamed on screen interpretation

**Figure 4.** SAR series retains the binding mode on the enzyme. Overlay of the crystal structures of compounds **3** (magenta), **14** (cyan), and **17** (green) bound to AmpC. Full crystallographic statistics are given in Table S3, Supporting Information.**Table 5.** Selection of Compounds Screened with Reactive Functional Groups<sup>a</sup>

Functional Group	Structure	Screened	Hits
2-halopyridine		50	0
4-halopyridine		13	0
sulfonyl halide		5	0
aldehyde		32	0
alkyl halide		84	0
acid halide		6	0
iso(thio)cyanate	—NCO(S)	10	0
epoxide, aziridine, thioepoxide		110	0
michael acceptor		289	3
dihydroxybenzene		31	0
anhydride		6	0
aromatic azide		5	0
trihydroxyphenyl		12	0

<sup>a</sup> Full list of functional groups searched is given in Supporting Information, Table 4.

problems, as is sometimes the case in HTS. Rather, the screen was unusually reliable, since each compound was assayed at seven concentrations, and actives were identified not on the basis



of statistical parameters but using each compound's concentration–response profile.<sup>8</sup> Consistent with this view, the qHTS unambiguously identified both  $\beta$ -lactams and the reversible inhibitors that we had seeded into the library.<sup>26,32</sup> Whereas some “actives” were irreproducible in secondary screens, these were rare, both among the promiscuous aggregators<sup>22</sup> and among the detergent-insensitive compounds investigated here. We attribute this to the high quality concentration–response curves afforded by the qHTS approach.<sup>8</sup> Consistent with the qHTS curve classification schema, compounds with the highest quality concentration–response curves (curve class 1 in the qHTS classification<sup>8</sup>) were covalent-acting  $\beta$ -lactams and sulfone oxadiazoles, which are certainly inhibitors of  $\beta$ -lactamase. Conversely, all of the compounds that were irreproducible in the secondary assays had low quality dose–response curves, typically only inhibiting at the single highest concentration assayed.

Another possible explanation for the lack of reversible inhibitors from the HTS is that AmpC  $\beta$ -lactamase is not an easily “druggable” target. Until this study, there was only one class of purely noncovalent, competitive inhibitors for AmpC,<sup>25,26</sup> though there are reversible pseudocovalent boronic acids that inhibit the enzyme in the nanomolar range, and many classes of covalent inhibitors, including several drugs (the penicillins and cephalosporins). We cannot rule out the possibility that AmpC is a difficult target to reversibly inhibit, though we do not favor this explanation.

Instead, the lack of specific inhibitors identified in the qHTS highlights the limitations of screening a relatively small and unbiased library within a large chemical possibility space. The NCGC library was designed for general use, with no single family of targets in mind, and it is not dominated by any one chemotype. Finding new chemotypes from libraries for which there is no ligand bias for the protein is an ongoing challenge for the field.<sup>33,34</sup> This explains the good track record of HTS against chemically well-explored targets such as G-protein coupled receptors and kinases and its often meager results against new genomic or even antibiotic targets<sup>35</sup> (though counter examples do exist<sup>12</sup>). Beyond covalently modifying compounds, the existing chemical space that populates most screening campaigns may simply not intersect with novel biological space. Anticipating this, the Molecular Libraries Small Molecule Repository (which is part of the NCGC screening collection) is being expanded into novel chemical space via libraries synthesized as part of the Molecular Libraries chemical diversity initiatives ([http://mli.nih.gov/technology/initiatives\\_chem\\_div.php](http://mli.nih.gov/technology/initiatives_chem_div.php)) and the Chemical Methodology and Library Development program (<http://www.nigms.nih.gov/Initiatives/CMLD>). In addition, other library technologies including natural product, combinatorial, and non-small molecule (e.g., peptoid) approaches are being incorporated into the NCGC collection, which should increase the ability to reach novel biological space as well.

Expanding the library by even an order of magnitude, however, and adding chemotypes from different sources, will only partially address this chemical space problem.<sup>13,14</sup> Another solution would be to simply screen at higher concentrations of compound, but this presents logistical difficulties, including exhaustion of source material, insolubility of compound in the assay, and an increase in the number of artifactual hits across the deck. An alternative is to prioritize a small subset of more likely chemotypes for careful testing, often at higher concentrations. One way to do this is by looking at a different part of chemical space, such as afforded by fragments.<sup>36–39</sup> Another is

by looking for complementary fits of particular library molecules to the structure of the target, as is afforded by docking.

With docking one can typically afford to test only a handful of molecules, but those that are tested can be evaluated at high concentrations under permissive substrate concentrations, allowing one to tease out interesting, derivitizable molecules for special attention in what is otherwise an undifferentiated chemical landscape. The apparent IC<sub>50</sub> of the best docking hit was initially found to be 140  $\mu$ M. However, since the docking had allowed us to consider pure molecular structure but the sold and screened “compound” was found to be a racemic mixture, the  $K_i$  of the pure *R*-isomer, once synthesized, was 37  $\mu$ M. The structure of the  $\alpha$ -phthalimido-acid inhibitor–enzyme complex suggested routes for synthetic optimization (Figure 3), and our initial efforts in this direction led to analogs with 5-fold improved affinity (Table 4). These phthalimides represent a novel series of competitive and specific inhibitors for this important antibiotic target and may merit further study.

Several technical points merit mention. Notwithstanding a substantial docking false-positive rate (14 of 16 tested), the true inhibitors were right for the right reasons; that is, the docking geometries corresponded closely to the subsequently determined crystallographic structures (Figure 3). Also, given the quality of the assay, those library molecules that did not inhibit  $\beta$ -lactamase may be considered reliable nonbinders, at least up to micromolar affinity. These molecules provide an unprecedented number of experimentally verified decoys for testing and optimizing docking methods. Finally, it bears mentioning that whereas detergent sensitivity is an appealing technique to identify aggregation-based inhibitors in screens, since it is typically reliable and lends itself to large scale application, it will for some otherwise attractive compounds be sensible to confirm this assignment by an orthogonal assay, such as dynamic light scattering,<sup>23</sup> multiple enzyme inhibition or sensitivity to target concentration.<sup>4,5</sup>

Despite its prominence in drug discovery and chemical biology, high throughput screening has rarely been subject to detailed quantification, at least in the public domain. Here, we categorize essentially every active molecule in a screen of over 70000 molecules. The results address two central problems in screening: how to prioritize hits for subsequent study, and what molecules to screen in the first place. Surprisingly, purported chemical reactivity of library molecules was uncorrelated with false-positive activity; rather, these were dominated by molecules undergoing colloidal aggregation. The lack of specific reversible actives from the HTS campaign speaks to the challenges attending on unbiased libraries. These libraries are nevertheless required when investigating new target families, or those for which drug- or tool-like compounds are unknown. This study leads to a model for exploring novel biological space with small molecules based on the complementarities between qHTS and docking. When attempting to identify small molecule modulators for a currently “undruggable” or “genomic” target for which structural information is available, it may be useful to screen and dock in parallel, using one to help interpret and guide the other. As we are confronted with an increasing number of interesting but challenging biological targets, ones with little chemical precedence in our screening libraries, and with the advent of the structural genomics programs, such combined screening strategies become at once pressing and pragmatic.

## Methods

**$\beta$ -Lactamase Assays.** The two initial HT screens were conducted in 1536 well format, with 7-point or 15-point dose–response curves

measured for every compound, looking for inhibition at a concentration of 400  $\mu$ M nitrocefin substrate in 50 mM potassium phosphate, a buffer that contributes to enzyme stability at these volumes.<sup>22</sup> The total reaction volume was 8  $\mu$ L per well. In the secondary assays undertaken here, reactions were monitored in methylacrylate cuvettes in 1 mL of reaction volume.  $\beta$ -Lactamase assays contained 1 nM AmpC  $\beta$ -lactamase, 200  $\mu$ M Nitrocefin, and varying amounts of Triton X-100 (either 0.0002 or 0.1%). Compound and enzyme were incubated for 5 min before the reaction was initiated by addition of substrate. Nitrocefin hydrolysis was monitored at 482 nm on a HP8453 UV-visible spectrophotometer at room temperature.

#### Chymotrypsin, Malate Dehydrogenase and Cruzain Assays.

Chymotrypsin and Malate Dehydrogenase were purchased from Sigma-Aldrich (St. Louis, MO). Recombinant cruzain was provided by R. Ferreira.

For chymotrypsin assays, inhibitor and 28 nM enzyme were incubated for 5 min and the reaction was initiated with 200  $\mu$ M *N*-succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide. For chymotrypsin assays without incubation, inhibitor and 200  $\mu$ M *N*-succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide were mixed, and the reaction was initiated with 28 nM enzyme. Reaction progress was monitored at 410 nm. For malate dehydrogenase assays, inhibitor and 2 nM enzyme were incubated for 5 min, and the reaction was initiated with 200  $\mu$ M oxalacetic acid and 200  $\mu$ M NADH. For malate dehydrogenase assays without incubation, inhibitor, 200  $\mu$ M oxalacetic acid, and 200  $\mu$ M b-nicotinamide adenine dinucleotide were mixed, and the reaction was initiated with 2 nM enzyme. Reaction progress was monitored at 340 nm. Cruzain activity is measured in a 96-well format fluorimetric assay, in which the cleavage of the substrate Z-Phe-Arg-aminomethyl coumarin (Z-FR-AMC) is monitored by excitation at 355 nm and emission at 460 nm. The assay is performed in 0.1 M Sodium Acetate buffer, pH 5.5, containing 10 nM enzyme, 5 mM DTT and 2.5  $\mu$ M Z-FR-AMC (0.01% TRITON X-100).

**Database Preparation.** A set of 66661 NCGC molecules were prepared for docking using the latest version of the ZINC protocol;<sup>28</sup> about 4000 library molecules were not released at the time and were not docked, thus the small discrepancy between the number of docked and screened molecules. Briefly, molecules were converted from 2D SDF to isomeric SMILES, and from these an initial 3D structure was calculated. For molecules with undefined stereochemistry, as for instance the docking hit **3**, both *R*- and *S*-isomers were calculated at up to two stereocenters, as were *E/Z*-enantiomers. A protonated form of each molecule at pH 7.0 was calculated with LigPrep (Schrodinger, LLC, New York, NY) with additional protonated and tautomeric forms calculated in the range of pH 5.75–8.25 using modified versions of LigPrep's parameter files. For each protonated form, we again used Corina to obtain a 3D model and then used AMSOL to calculate partial atomic charges and atomic desolvation energies. We used Omega (OpenEye Scientific Software, Santa Fe NM) to enumerate accessible conformations; ring conformations calculated by Corina were preserved. These ligand conformations were combined into a "flexibase" from which they were docked.

**Docking.** To prepare the site for docking, we removed all water, ligand and ion molecules from monomer B of an AmpC structure bound to a small molecule discovered previously (PDB entry 1L2S). Enzyme residues were protonated automatically<sup>25</sup> and the positions of selected protons were rotated to orientations consistent with an ultrahigh resolution  $\beta$ -lactamase structure.<sup>40</sup> The sphere set and energy potential (scoring) grids were calculated in an automated fashion, as outlined recently.<sup>25,41</sup> Briefly, docking spheres were based on atom positions from three bound inhibitors (Protein Data Bank entries 1L2S, 2HDU, and 2HDR). These were supplemented by calculated receptor-derived spheres; a total of 55 spheres were used. Spheres were labeled for chemical matching based on the hydrogen-bonding properties and charged states of nearby receptor atoms. Four scoring grids were generated: an excluded volume grid, a van der Waals potential grid, an electrostatic potential grid and a desolvation grid (unpublished). Ligand configurations were scored

using the sum of electrostatic and van der Waals interaction energies, corrected by ligand desolvation energy. Final energies reflected 100 steps of rigid-body minimization.

Sampling in DOCK3.5.54 can be varied according to user-defined parameters. To optimize these, the NCGC database was seeded with 21 known  $\beta$ -lactamase ligands and 127 experimentally determined nonbinders or "decoys". Parameters were varied to enrich the ligands and deprioritize the decoys among the rest of the database molecules. This optimization resulted in bin sizes for both receptor and ligand of 0.4 Å and an overlap size of 0.3 Å. A distance tolerance (dislim) of 1.2 Å was used to match the ligand onto the spheres. We note that this optimization was sought to optimize particular hotspots in the enzyme site; no ligand information was used to bias the search and, in fact, the ligands that were ultimately discovered were novel. A single pose with the best total energy score was saved for each docked molecule. For ligands with multiple protonation states and tautomeric forms, only the best scoring representation is retained. To improve polar complementarity between enzyme and docked ligand, the absolute magnitude of the partial atomic charges of several active site atoms were increased by 0.4 units (the net residue charges were unchanged): Ser64O $\gamma$  and HO $\gamma$ ; Ser64N and HN; Tyr150OH and HOH; Asn152O $\delta$ 1, HN $\delta$ 1, and HN $\delta$ 2; Tyr221OH and HOH; Asn289O $\delta$ 1, HN $\delta$ 1, and HN $\delta$ 2; Thr316O $\gamma$ 1 and HO $\gamma$ ; Ala318O and HN; Asn343O $\delta$ 1, HN $\delta$ 1, and HN $\delta$ 2; Asn346O $\delta$ 1, HN $\delta$ 1, and HN $\delta$ 2. For the asparagine and glutamine residues, the charge increase was divided among the protons on the amide groups. The 500 top-scoring docked molecules were visually evaluated. Molecules were chosen for experimental testing based on their docking score (i.e., only top-scoring compounds were tested), our assessment of how well they fit the site, and their chemical diversity.

**Crystallography.** The cocrystal structure of AmpC and **1** was obtained by soaking apo crystals. These were grown by vapor diffusion in hanging drops equilibrated over 30% PEG 8000, 0.1 M ammonium sulfate, 0.1 M sodium cacodylate, pH 6.7. The initial concentration of the protein in the drop was 10 mg/mL. Crystals appeared within 1 week on equilibration at 21 °C. Crystals were placed in a 6 $\mu$ L drop of reservoir solution containing 300  $\mu$ M of compound **1** and 1% DMSO and soaked for 4 h. Before data collection, crystals were flash-cooled in liquid nitrogen.

Cocrystals of AmpC in complex with compounds **3**, **14**, and **17** were grown by vapor diffusion in hanging drops equilibrated over 1.8 M potassium phosphate buffer, pH 8.7, using microseeding. The initial concentration of the protein and the compound in the drop was 3.5 mg/mL and 10 mM, respectively. Crystals appeared 1–2 weeks after equilibration at 21 °C. Before data collection, crystals were immersed in a cryoprotectant solution of 20% sucrose and 1.8 M potassium phosphate (pH 8.7) containing compound **3** for about 30 s and were flash-cooled in liquid nitrogen.

Diffraction measurements were collected on frozen crystals at beamline 8.3.1 of the Advance Light Source (ALS, Lawrence Berkeley Laboratory, CA). Initial space group determination and data collection strategy was performed with the program ELVES.<sup>42</sup> Reflections were indexed, integrated and scaled using the HKL package.<sup>43</sup> For cocrystals with **1**, the space group was *P*2<sub>1</sub>2<sub>1</sub>2 with 2 molecules in the asymmetric unit. Using the B monomer from an apo structure, an initial molecular replacement solution was obtained using PHASER.<sup>44</sup> In each of the cocrystals with **3**, **14**, and **17**, the space group was *C*<sub>2</sub> and there were two molecules in the asymmetric unit. The initial phasing model was an apo AmpC structure (PDB entry 1KE4) with water molecules and ions removed. For all structures an initial round of rigid body refinement was performed using REFMAC5<sup>45</sup> after initial phased models were positioned. Models were built and waters placed using Coot<sup>46</sup> and these were further refined with REFMAC5.

**Chemical Synthesis.** All reagents were purchased from commercial sources and used without further purification. Microwave reactions were performed on a Biotage Initiator Sixty. Analytical LC/MS were performed on a Waters Acquity UPLC, and purities determined by integration of ELSD or total UV absorbance signals. Purifications were performed on a Waters mass directed LC. NMR



spectra were obtained on a Varian 400 MHz NMR. HRMS data were supplied by The Proteomics and Mass Spectrometry Facility at NIDDK/NIH/DHHS.

The members of the library were synthesized according to Scheme S1 (Supporting Information) by a modification of a previously reported procedure.<sup>47</sup> Trimellitic anhydride (1 equiv) and the appropriate amino acid (1 equiv) were weighed into 0.5–2 mL microwave vials. Diethylene glycol dimethyl ether (diglyme, 0.5 mL) was added to each vial and the vials sealed. The reactions were heated via microwave irradiation to 200 °C for 10 min and cooled to room temperature. Reactions were analyzed by LC/MS and diluted with DMSO (2.5 mL), and the products were purified to greater than 90% purity by reverse phase preparative HPLC or LC/MS (acetonitrile/water/0.2% formic acid).

The following modifications were used as necessary: In cases where amino acids were available as HCl salts, 1 equiv of triethylamine was added to the reaction mixture. In cases where reactions were incomplete by LC/MS after 10 min due to insolubility of the amino acid, an additional 0.5 mL of diglyme was added, and the reaction heated for a further 10 min at 200 °C, followed by purification.

Characterization of molecules synthesized available in Supporting Information.

**Mass Spectroscopy.** Protein samples were analyzed by injection of 1.0  $\mu$ L onto a Phenomenex Onyx monolithic column (0.1  $\times$  150 mm) held at 50/50 water/acetonitrile with 0.1% formic acid added to each solvent. The column flow rate was 1.0  $\mu$ L/minute. The column effluent was introduced to an ABI QSTAR XL hybrid quadrupole orthogonal time-of-flight mass spectrometer equipped with a MicroIon Spray source. Spray tip voltage was 5.5 kilovolts and the nebulizing gas value was at 1. The mass spectrometer analyzed the masses from 500 to 1800 with an accumulation time of 2.0 s. Typically, several accumulations were averaged to produce the spectrum acquired. Data analysis and mass determinations were accomplished using the instrument's BioAnalyst software.

**Accession Codes.** The coordinates and structure factors for the described structures have been deposited in the Protein Data Bank with the following accession codes: 2PU4 (1), 2PU2 (3), 2R9X (14), and 2R9W (17).

**Acknowledgment.** Supported by NIH Grants GM71630 and GM59957 (to B.K.S.), and a Ruth Kirschstein NRSA fellowship GM076883 (to K.B.). We thank R. Ferreira for assistance with cruzain assays. The UCSF Mass Spectrometry Facility (A. Burlingame, Director) is supported by NIH NCRR BRTP 01614. B.Y.F. is supported by a Kozloff research fellowship and by a UCSF School of Pharmacy fellowship.

**Supporting Information Available:** Schemes and tables detailing chemical characterizations are provided, as is a full table of the reactive functionality found in the screening library. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Walters, W. P.; Ajay; Murcko, M. A. Recognizing molecules with drug-like properties. *Curr. Opin. Chem. Biol.* **1999**, *3*, 384–387.
- Huth, J. R.; Sun, C.; Sauer, D. R.; Hajduk, P. J. Utilization of NMR-derived fragment leads in drug design. *Methods Enzymol.* **2005**, *394*, 549–571.
- Rishton, G. M. Reactive compounds and in vitro false positives in HTS. *Drug Discovery Today* **1997**, *2*, 382–384.
- McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **2002**, *45*, 1712–1722.
- McGovern, S. L.; Helfand, B.; Feng, B.; Shoichet, B. K. A specific mechanism for nonspecific inhibition. *J. Med. Chem.* **2003**, *46*, 4265–4272.
- Rishton, G. M. Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discovery Today* **2003**, *8*, 86–96.
- Roche, O.; Schneider, P.; Zuegge, J.; Guba, W.; Kansy, M.; Alanine, A.; Bleicher, K.; Danel, F.; Gutknecht, E. M.; Rogers-Evans, M.; Neidhart, W.; Stalder, H.; Dillon, M.; Sjogren, E.; Fotouhi, N.; Gillespie, P.; Goodnow, R.; Harris, W.; Jones, P.; Taniguchi, M.; Tsujii, S.; von Der Saal, W.; Zimmermann, G.; Schneider, G. Development of a virtual screening method for identification of "Frequent Hitters" in compound libraries. *J. Med. Chem.* **2002**, *45*, 137–142.
- 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*, 11473–11478.
- Gribbon, P.; Lyons, R.; Laffin, P.; Bradley, J.; Chambers, C.; Williams, B. S.; Keighley, W.; Sewing, A. Evaluating real-life high-throughput screening data. *J. Biomol. Screening* **2005**, *10*, 99–107.
- Stockwell, B. R. Exploring biology with small organic molecules. *Nature* **2004**, *432*, 846–854.
- Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **2004**, *432*, 855–861.
- Inglese, J.; Johnson, R. L.; Simeonov, A.; Xia, M.; Zheng, W.; Austin, C. P.; Auld, D. S. High-throughput screening assays for the identification of chemical probes. *Nat. Chem. Biol.* **2007**, *3*, 466–479.
- Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.
- Oprea, T. I. Current trends in lead discovery: are we looking for the appropriate properties. *Mol. Diversity* **2002**, *5*, 199–208.
- Li, S.; Gao, J.; Satoh, T.; Friedman, T. M.; Edling, A. E.; Koch, U.; Choksi, S.; Han, X.; Korngold, R.; Huang, Z. A computer screening approach to immunoglobulin superfamily structure and interactions: discovery of small non-peptidic CD4 inhibitors as novel immunotherapeutics. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 73–78.
- Schnecke, V.; Swanson, C. A.; Getzoff, E. D.; Tainer, J. A.; Kuhn, L. A. Screening a peptidyl database for potential ligands to proteins with side-chain flexibility. *Proteins* **1998**, *33*, 74–87.
- Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44*, 1035–1042.
- Hartshorn, M. J.; Verdonk, M. L.; Chessari, G.; Brewerton, S. C.; Mooij, W. T.; Mortenson, P. N.; Murray, C. W. Diverse, high-quality test set for the validation of protein–ligand docking performance. *J. Med. Chem.* **2007**, *50*, 726–741.
- Shoichet, B. K. Virtual screening of chemical libraries. *Nature* **2004**, *432*, 862–865.
- Gruneberg, S.; Wendt, B.; Klebe, G. Subnanomolar inhibitors from computer screening: A model study using human carbonic anhydrase II. *Angew. Chem., Int. Ed.* **2001**, *40*, 389–393.
- Feng, B. Y.; Shelat, A.; Doman, T. N.; Guy, R. K.; Shoichet, B. K. High-throughput assays for promiscuous inhibitors. *Nat. Chem. Biol.* **2005**, *1*, 146–148.
- Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P. A high-throughput screen for aggregation-based inhibition in a large compound library. *J. Med. Chem.* **2007**, *50*, 2385–2390.
- Feng, B. Y.; Shelat, A.; Doman, T. N.; Guy, R. K.; Shoichet, B. K. High-throughput assays for promiscuous inhibitors. *Nat. Chem. Biol.* **2005**, *1*, 146–148.
- Lorber, D. M.; Shoichet, B. K. Hierarchical docking of databases of multiple ligand conformations. *Curr. Top. Med. Chem.* **2005**, *5*, 739–749.
- Powers, R. A.; Morandi, F.; Shoichet, B. K. Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. *Structure (London)* **2002**, *10*, 1013–1023.
- Tondi, D.; Morandi, F.; Bonnet, R.; Costi, M. P.; Shoichet, B. K. Structure-based optimization of a non-beta-lactam lead results in inhibitors that do not up-regulate beta-lactamase expression in cell culture. *J. Am. Chem. Soc.* **2005**, *127*, 4632–4639.
- Babaoglu, K.; Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2006**, *2*, 720–723.
- Irwin, J. J.; Shoichet, B. K. ZINC—A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
- Walters, W.; Namchuk, M. Designing screens: How to make your hits a hit. *Nat. Rev. Drug Discovery* **2003**, *2*, 259–266.
- Oprea, T. I.; Allu, T. K.; Fara, D. C.; Rad, R. F.; Ostropovici, L.; Bologa, C. G. Lead-like, drug-like or "Pub-like": how different are they. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 113–119.
- Inglese, J.; Johnson, D. L.; Shiau, A.; Smith, J. M.; Benkovic, S. J. Subcloning, characterization, and affinity labeling of Escherichia coli glycylamide ribonucleotide transformylase. *Biochemistry* **1990**, *29*, 1436–1443.
- Morandi, F.; Caselli, E.; Morandi, S.; Focia, P. J.; Blazquez, J.; Shoichet, B. K.; Prati, F. Nanomolar inhibitors of AmpC beta-lactamase. *J. Am. Chem. Soc.* **2003**, *125*, 685–695.

- (33) Spencer, R. W. High-throughput screening of historic collections: Observations on file size, biological targets, and file diversity. *Biotechnol. Bioeng.* **1998**, *61*, 61–67.
- (34) Macarron, R. Critical review of the role of HTS in drug discovery. *Drug Discovery Today* **2006**, *11*, 277–279.
- (35) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. Drugs for bad bugs: Confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discovery* **2007**, *6*, 29–40.
- (36) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **2004**, *47*, 3463–3482.
- (37) Erlanson, D. A.; Wells, J. A.; Braisted, A. C. Tethering: Fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199–223.
- (38) Verdonk, M. L.; Hartshorn, M. J. Structure-guided fragment screening for lead discovery. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 404–410.
- (39) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discovery* **2004**, *3*, 660–672.
- (40) Chen, Y.; Minasov, G.; Roth, T. A.; Prati, F.; Shoichet, B. K. The deacylation mechanism of AmpC  $\beta$ -lactamase at ultrahigh resolution. *J. Am. Chem. Soc.* **2006**, *128*, 2970–2976.
- (41) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–6801.
- (42) Holton, J.; Alber, T. Automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1537–1542.
- (43) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326.
- (44) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 458–464.
- (45) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240–255.
- (46) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- (47) Staubli, A.; Ron, E.; Langer, R. Hydrolytically degradable amino acid-containing polymers. *J. Am. Chem. Soc.* **1990**, *112*, 4419–4424.

JM701500E