ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Increasing Chemical Space Coverage by Combining Empirical and Computational Fragment Screens

Sarah Barelier,[†,#] Oliv Eidam,[†,¶,#] Inbar Fish,[†,‡] Johan Hollander,[§] Francis Figaroa,[§] Ruta Nachane,[§] John J. Irwin,[†,⊥] Brian K. Shoichet,*[,†,⊥] and Gregg Siegal*[,§,∥]

[†]Department of Pharmaceutical Chemistry, University of California, San Francisco, 1700 4th St., Byers Hall, San Francisco, California 94158, United States

[‡]Department of Biochemistry and Molecular Biology, George S. Wise Faculty of Life SciencesTel-Aviv University, Ramat Aviv, Israel

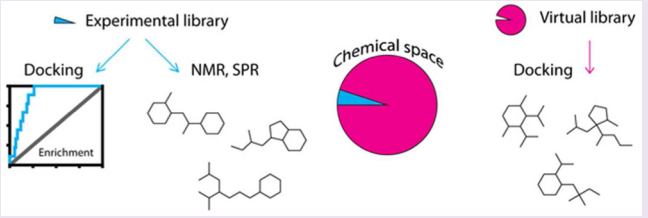
[§]ZoBio, Eisteinweg 55, 2300-RA Leiden, The Netherlands

[∥]Leiden University, Eisteinweg 55, 2300-RA Leiden, The Netherlands

[⊥]Leslie Dan Faculty of Pharmacy, University of Toronto, Donnelly Centre Suite 604, 160 College Street, Toronto, Ontario, Canada M5S 3E1

Supporting Information

ABSTRACT: Most libraries for fragment-based drug discovery are restricted to 1,000−10,000 compounds, but over 500,000 fragments are commercially available and potentially accessible by virtual screening. Whether this larger set would increase chemotype coverage, and whether a computational screen can pragmatically prioritize them, is debated. To investigate this question, a 1281-fragment library was screened by nuclear magnetic resonance (NMR) against AmpC β-lactamase, and hits were confirmed by surface plasmon resonance (SPR). Nine hits with novel chemotypes were confirmed biochemically with K_I values from 0.2 to low mM. We also computationally docked 290,000 purchasable fragments with chemotypes unrepresented in the empirical library, finding 10 that had K_I values from 0.03 to low mM. Though less novel than those discovered by NMR, the docking-derived fragments filled chemotype holes from the empirical library. Crystal structures of nine of the fragments in complex with AmpC β-lactamase revealed new binding sites and explained the relatively high affinity of the docking-derived fragments. The existence of chemotype holes is likely a general feature of fragment libraries, as calculation suggests that to represent the fragment substructures of even known biogenic molecules would demand a library of minimally over 32,000 fragments. Combining computational and empirical fragment screens enables the discovery of unexpected chemotypes, here by the NMR screen, while capturing chemotypes missing from the empirical library and tailored to the target, with little extra cost in resources.



Fragment-based screening and optimization are now widely used in drug discovery,[1] fortified by the registration of the first drug originating from a fragment-based screen.[2] In such screens, low-molecular weight compounds (150−300 Da)[3] are sought as early hits, which are then optimized for affinity, permeability, and related pharmacological properties. The low molecular weight of fragment molecules imposes practical challenges, as it typically limits their affinities to the mid-micromolar to low-millimolar range. However, judged by their ligand efficiency (LE), ΔG_b/heavy atom count (HAC), fragments have advantages over other actives from early discovery and can often be optimized for affinity without sacrificing their favorable physical properties.[4,5] Also, the combinatorial collapse of diversity at small molecular sizes allows fragment libraries to cover chemical space many orders of magnitude better than larger libraries, such as those used in high-throughput screens (HTS).[6,7]

The collapse of chemical diversity at the fragment level, combined with the need to use low-throughput biophysical assays to detect low-affinity binding,[8,9] has led to small fragment libraries (1,000−10,000 compounds).[10,11] Several of these have been optimized for diversity[10] and can recapitulate the chemotypes present in drug-like actives for several targets,[12,13] leading to active molecules in multiple screens.[14−17] Still, this is not the same as saying that fragment libraries cover most of biorelevant chemical space. As there are over 700,000 fragments that are commercially available, fragment screens may miss interesting and readily accessible chemotypes.

In principle, compounds unrepresented in any particular empirical screening library may be accessed computationally. Molecular docking can sample all available compounds and prioritize those that sterically and energetically fit target sites.[18] Concerns about reliability, however, have limited the use of docking in fragment discovery: fragments can adopt multiple

Received: March 3, 2014
Accepted: May 7, 2014
Published: May 7, 2014

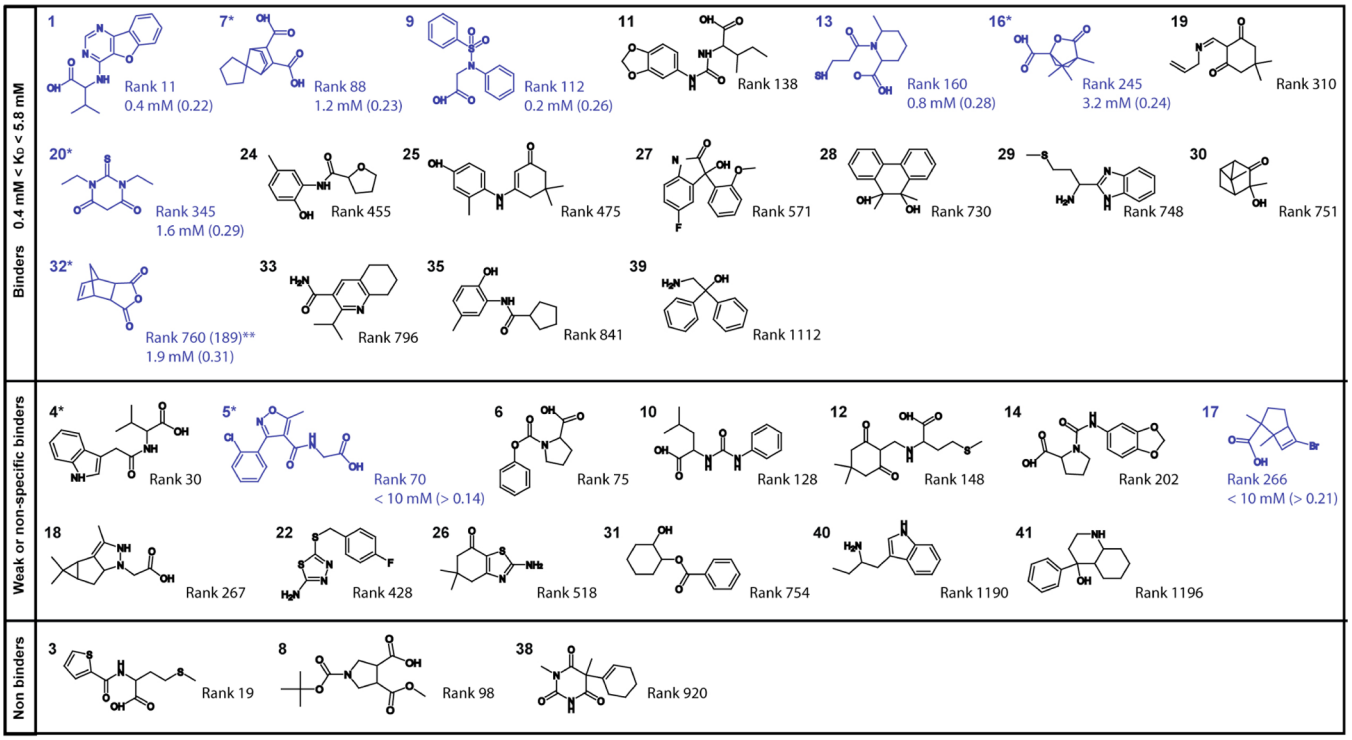


Figure 1. SPR status, docking rank, and inhibitory activity for 34 hits discovered by NMR. Inhibitors are in blue. The K_I is indicated, followed by the ligand efficiency, in parentheses. *Competitive for binding in the presence of benzo[*b*]thiophene-2-boronic acid (Supplementary Figure 1) in a secondary TINS assay. **New rank for fragment **32** docked as a diacid (Supplementary Table 3).

orientations in the binding site,[19] and scoring functions optimized for larger, drug-like molecules may be inappropriate for fragments.[20] In several fragment screens, docking has uncovered potent hits,[21] and predicted docked structures have been confirmed by subsequent crystallography.[22] Still, few studies have compared docking and empirical fragment screens directly and prospectively.[23]

We thus thought it interesting to compare an empirical screen of a fragment library with a docking screen of the same library, run in parallel against the same target. We screened an experimental fragment library of 1,281 molecules, using target-immobilized NMR screening (TINS) to detect binding.[24] We wondered whether the docking screen would prioritize the same active molecules found empirically, and whether the fragment library would illuminate chemotypes unknown for the target. More germane to this study, we wondered if, notwithstanding its diversity, the 1,281 experimental fragment library would miss chemotypes that might be prioritized by docking a much larger library of commercially available fragments. To investigate these questions at atomic resolution, we targeted the model enzyme and drug target, AmpC β-lactamase. AmpC has been extensively studied for mechanism and biophysics[25−27] and has served as a model system for different drug discovery approaches, including HTS,[28] structure-based screening,[29] and covalent inhibition.[30] The enzyme, which lends itself to facile crystallography and enzymology, is the most widespread resistance determinant to β-lactam antibiotics, such as penicillin, and several investigational drugs against AmpC have entered late-stage clinical trials. The well-behaved nature of this enzyme allowed us to investigate binding by three techniques (TINS, SPR, and enzymological K_I/K_D) and to determine the structures of nine new enzyme-fragment complexes by crystallography and compare them to the docking-predicted structures. Liabilities found in the docking by comparing it to the empirical screen and to the experimental structures, opportunities to cover more chemical space using docking, and complementarities between the computational and empirical approaches will be discussed.

RESULTS AND DISCUSSION

Target-Immobilized NMR Screening of 1,281 Fragments from the ZoBio Library. A subset of the ZoBio internal fragment library consisting of 1,281 molecules was screened against AmpC β-lactamase using TINS,[24] blind of the docking results, and 41 hits were confirmed to bind to AmpC in a replication experiment (hit rate 3.2%) (Supplementary Table 1). Six of these hits acted competitively by TINS with a known active-site inhibitor, benzo[*b*]thiophene-2-boronic acid (compound **S3** in Supplementary Figure 1).

Binding by Surface Plasmon Resonance. Of the 41 NMR hits, 35 were studied by SPR in a secondary, confirmatory assay (the other six were no longer available) (Supplementary Table 2). K_D values could be determined for 19 fragments (0.4 mM < K_D < 5.8 mM). Another 13 showed binding, but it was too weak to allow reliable determination of K_D values. Only three compounds were characterized as nonbinders, in substantial agreement with the NMR screening. All six active-site competitive NMR hits were confirmed by SPR.

AmpC Inhibition. As with most enzymes, inhibitor binding affinities for AmpC are equivalent to competitive K_I values, by linkage equilibrium. Pragmatically, inhibition is also the relevant functional read-out for the enzyme. Therefore, 34 of the 35 NMR hits tested by SPR were investigated for AmpC inhibition (one was no longer available) (Supplementary Figure 2). Of these, nine fragments had K_I values below 10 mM, with the

dx.doi.org/10.1021/cb5001636 | ACS Chem. Biol. 2014, 9, 1528−1535

most potent having a K_I of 0.2 mM (Figure 1 and Table 1). Seven of these fragments (**1**, **7**, **9**, **13**, **16**, **20**, and **32**) had well-defined SPR binding curves, with the two less potent ones (**5** and **17**) having weak SPR signals (Figure 1 and Supplementary Table 2). Ligand efficiencies ranged from 0.14 to 0.31 (Figure 1 and Table 1). Compared with known AmpC inhibitors, the highest pairwise Tanimoto coefficients (Tc) (EFCP_4 fingerprints) ranged from 0.16 to 0.28 (average 0.21), indicating high topological novelty for the NMR-derived fragments (Table 1). For the other 25 compounds, no measurable inhibition was detected up to a concentration of 10 mM or to the solubility limit of the compound. For one of these 25, binding to the protein was nevertheless observed by X-ray crystallography at the surface, about 25 Å away from the active site (fragment **41**, below).

Table 1. Inhibitory Activity and Chemotype Novelty of the Inhibitors Found by NMR and Virtual Screening

NMR hits				docking hits			
ID	K_I (mM)	LE[a]	Tc[b]	ID	K_I (mM)	LE[a]	Tc[b]
1	0.4	0.22	0.22	**44**	1.7	0.29	0.4
5	<10	>0.14	0.27	**45**	<5	>0.21	0.22
7	1.2	0.23	0.17	**46**	0.4	0.33	0.36
9	0.2	0.26	0.28	**47**	<10	>0.19	0.19
13	0.8	0.28	0.2	**48**	0.2	0.34	0.52
16	3.2	0.24	0.16	**50**	1.3	0.24	0.35
17	<10	>0.21	0.21	**53**	0.07	0.43	0.26
20	1.6	0.29	0.18	**54**	0.03	0.42	0.42
32	1.9	0.31	0.21	**55**	0.7	0.33	0.49
				57	1.0	0.24	0.32
				60	0.07	0.35	0.42

[a]Ligand efficiency. [b]Highest pairwise Tanimoto coefficient (EFCP_4 fingerprints) to a known AmpC inhibitor.

Docking the NMR Library. In parallel with the NMR screen, the fragment library was docked against the active site of AmpC. Whereas other pockets exist on the AmpC surface, they are much smaller than the active site,[29] and all structurally characterized inhibitors bind in the active site. Naturally, the noninhibitory NMR hits might bind elsewhere on the enzyme, but trying to anticipate this would be outside of the normal docking protocol, where a key binding site is typically targeted. Additionally, we were open to this sort of illuminating result from the empirical screen, which is one of its strengths. In the docking, the top NMR hit ranked 11th out of 1281, and the overall enrichment for the 41 NMR binders, using the area under the curve (AUC),[31] was 0.66 (Supplementary Figure 3). The top 10 ranking ZoBio compounds were all anions that complemented the catalytic site well but did not bind by TINS; these may be considered docking false positives. If we restrict ourselves to the nine fragments that inhibited AmpC (active site binders), four ranked in the top 10% of the docking list (> 4-fold enrichment over random at 10%). Intriguingly, the NMR hit with the lowest (worst) rank among the nine inhibitors (fragment **32**, rank 760) is in fact a diacid in solution, not an acid anhydride as it was represented in the library (Supplementary Table 3). Docking **32** in its diacid form changes its rank to 189, placing all active site binders in the top 27% of the docking list, with an AUC of 0.87 (Supplementary Figure 3).

Docking a 290,000 Chemically Dissimilar Fragment Library for New Chemotypes. To explore the chemical space that is not covered by the experimental library, a set of 290,225 commercially available fragments, dissimilar to the ZoBio compounds (Tc ≤ 0.4, using ECFP_4 fingerprints), was docked against the enzyme. As is often true with docking and even empirical screening,[32] it was impractical to follow up all hits with detailed experiments. Therefore, an 18-compound subset of the top ranking molecules was selected to assess biochemical activity. These compounds were representative of the top 500 docked molecules (top 0.17% of the library), with ranks from 7 to 490 out of 290,225 (Figure 2 and Supplementary Table 4). In addition to their physics-based docking scores and their dissimilarity to the ZoBio set, these fragments were selected for their chemical diversity, a widely used criterion,[32] and for hydrogen bonding with key active site residues (Ser64, Ala318, Asn152). Fragments that had been docked with incorrect ionization states or strained conformations were deprioritized, removing artifactual hits. The final 18 molecules well-represented the top 500 docked molecules overall: for instance, they had an average of 15.1 heavy atoms and an average net charge of −1.0, versus 15.4 and −0.9 for the top 500 hits. In the AmpC activity assay (Supplementary Figure 2), 10 of them had K_I values below 10 mM, with the most active having a K_I of 30 μM; ligand efficiencies ranged from 0.19 to 0.43 (Figure 2 and Table 1). Tc values to known AmpC ligands ranged from 0.19 to 0.52, with an average of 0.36, using

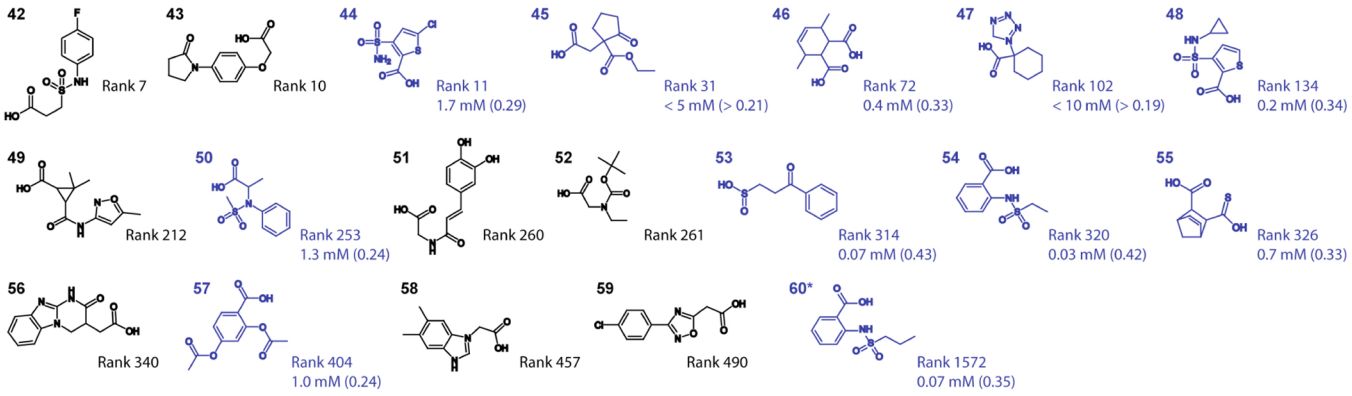


Figure 2. Docking ranks and inhibitory activity for 18 commercial fragments discovered by docking. Fragment **60**, a close analogue of fragment **54** that was used for crystallization with AmpC, is also shown. Inhibitors are in blue. The K_I is indicated, followed by the ligand efficiency, in parentheses.

dx.doi.org/10.1021/cb5001636 | ACS Chem. Biol. 2014, 9, 1528−1535

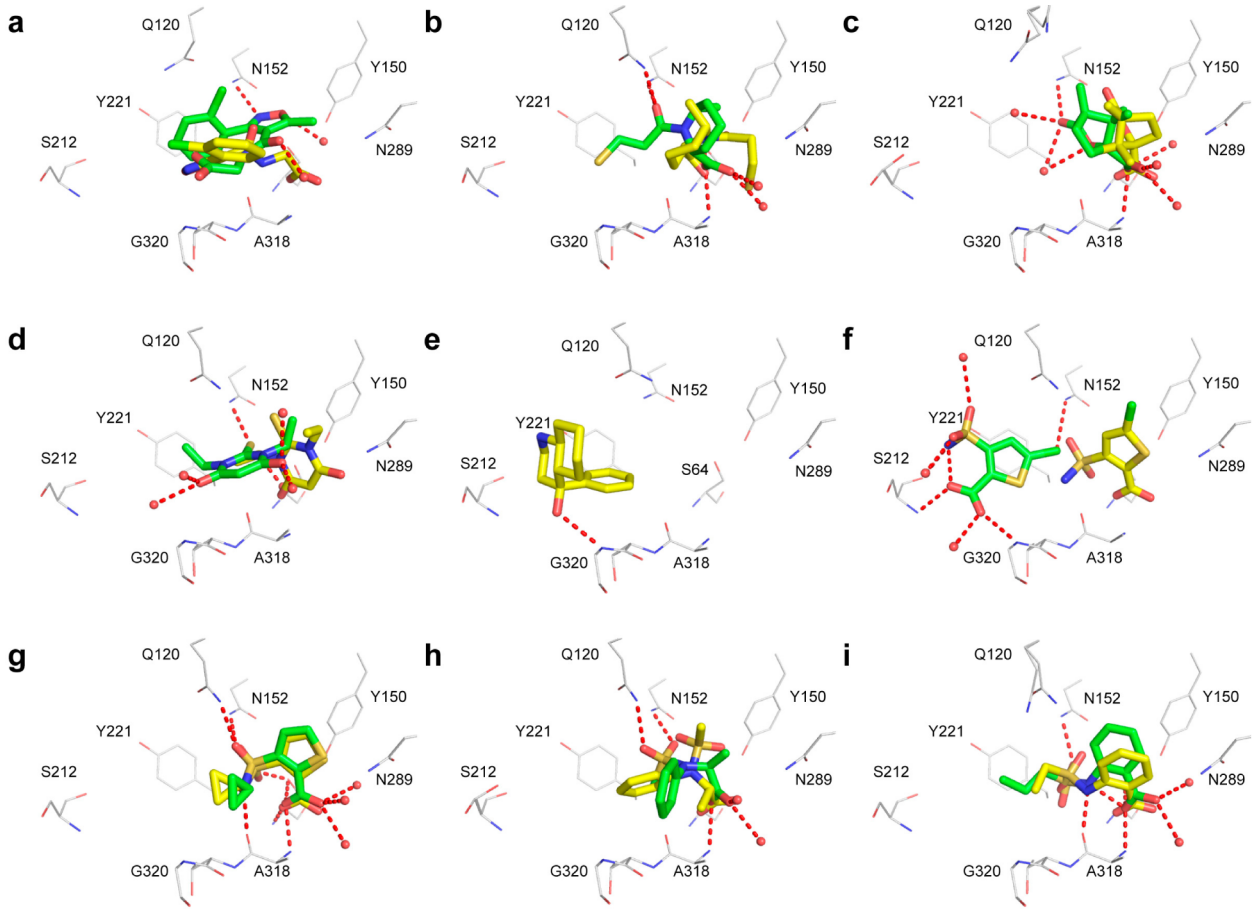


Figure 3. Comparison of docking-predicted (yellow) and crystallographic fragment geometries (green) for five NMR hits (a–e) and four docking hits (f–i): (a) **5**, (b) **13**, (c) **16**, (d) **20**, (e) **41**, (f) **44**, (g) **48**, (h) **50**, and (i) **60** superposed on **54**. Protein residues are depicted with gray carbon atoms, crystallographic water molecules as red spheres, and hydrogen bonds as red dashed lines.

ECFP_4 fingerprints (Table 1); this is substantially higher, indicating less novelty, than the 0.21 average Tanimoto coefficient observed for the NMR-derived fragment hits.

Comparison of the Docking Poses to Crystal Structures. The structures of five NMR (**5**, **13**, **16**, **20**, and **41**) and four docking fragments (**44**, **48**, **50**, and **54**) in complex with AmpC were determined by X-ray crystallography, with resolutions ranging from 1.32 to 2.28 Å. The location of all ligands was unambiguous in initial $F_o - F_c$ difference electron density maps (Supplementary Figure 4 and Supplementary Table 5), except for fragment **54**. To clarify the likely structure of **54**, we determined that of **60**, a close analogue differing only by the replacement of an ethyl by a propyl group, and determined its structure at 1.42 Å (Figure 2 and Supplementary Figure 4).

The correspondence of the ZoBio inhibitor structures with the predicted docking poses was spotty (Figure 3). The docking poses of fragments **13** and **16** recapitulated the interaction between the key inhibitor carboxylate, Ser64 and Ala318 (Figure 3b and c). Even here though, several secondary interactions were missed, such as the interaction of the ketone moieties with Asn152 and Gln120, leading to root-mean-square deviations (rmsd) values of 3.3 and 2.5 Å for **13** and **16**, respectively. For fragment **5** the differences were larger. This fragment was docked to place its carboxylate into the oxyanion hole (Figure 3a), an orientation that has been consistently observed for ligands bearing this functionality in previous AmpC structures.[33] In the crystal structure, however, the carboxylate points out toward solvent, resulting in poor correspondence between the docked pose and the crystal structure (rmsd 5 Å). Meanwhile, the thioxopyrimidine **20** was modeled to dock with one aryl oxygen as anionic (calculated pK_a is 7.4) and to interact with the enzyme's oxyanion hole (Figure 3d). In the crystal structure, however, the ligand appears to bind in its neutral form and adopts a different orientation, making only one hydrogen bond with Asn152 while the aryl oxygens hydrogen bond with structural water molecules (rmsd 3.5 Å). Finally, fragment **41** was observed to bind AmpC in both the TINS and SPR experiments but did not inhibit the enzyme. Upon determination of the crystal structure, clear density appeared at the surface of the protein, 25 Å from the catalytic site, where the fragment interacts with Gly36, a water molecule, and a phosphate ion (Supplementary Figure 4e). Fragment **41** ranked poorly, 1196/1281, in the active-site focused docking screen (Figure 3e).

There was better correspondence between the crystal structures of the docking-derived fragments and their predicted poses. The predicted pose for fragment **48** recapitulated the crystallographic geometry with an rmsd of 0.8 Å (Figure 3g). For fragment **50**, the docking pose predicted the crystal structure less accurately (rmsd 2.7 Å), with the sulfonamide making different interactions in each structure, though the key carboxylate is positioned similarly and overall the two poses overlap (Figure 3h). It is worth mentioning that both AmpC/

dx.doi.org/10.1021/cb5001636 | ACS Chem. Biol. 2014, 9, 1528–1535

48 and AmpC/**50** structures show a second ligand bound to the distal site, but with lower occupancy and weaker density (Supplementary Figure 4g and h). Fragment **54** is the most potent fragment inhibitor of AmpC at 30 μM, but the active site density was not defined well enough to allow its unambiguous placement. We turned to a close analogue, **60**, that differs from **54** by the addition of one methyl group on its distal end and inhibits AmpC with a K_I of 70 μM (docking rank 1572/290,225). Consistent with the docking prediction, the X-ray structure of **60** closely superposes with the docked pose of **54** (Figure 3i). Finally, in the crystal structure, fragment **44** binds to the distal pocket of the active site, defined by interactions with Ser212, Tyr221, and Gly320, not in the oxyanion hole as anticipated by docking. Intriguingly, **44** is a close analogue of **48**, which does bind in the oxyanion hole by crystallography (above). The discrepancy between the docked and observed poses of fragment **44** may be more apparent than real, a point to which we return.

In summary, docking predicted the key interaction between the carboxylate warhead and the catalytic Ser64 for two of the five NMR hits (**13** and **16**). Two other NMR hits adopted unexpected conformations and/or orientations in the active site (**5** and **20**), and compound **41** does not target the active site. Three of the four docking hits adopted crystallographic orientations that recapitulated the docking poses (**48**, **50**, and **54**) whereas the crystallographic and docking poses for fragment **44** clearly disagree. For all of the active site binders, both NMR- and docking-derived, at least one pose within 1.3 Å to the crystallographic geometry was sampled in the docking, even if it was not the highest scoring.

How Many Fragments Are Necessary to Represent Biologically Relevant Space? The discovery of potent fragments by docking, whose chemotypes were not covered by the empirical library, is consistent with the idea that many empirical fragment libraries miss relevant chemotypes simply because too few molecules are included in the library. Conversely, in most fragment screens, pragmatic lead matter has been discovered (certainly this was true here, against AmpC), and against some targets, fragment libraries have captured all the core chemotypes of the optimized molecules ultimately advanced to the clinic, irrespective of their origin.[12] This observation might suggest that regardless of what is missed in fragment screening, the libraries are large enough to be pragmatic. Still, it seemed interesting to quantify how many fragments would be necessary to simply cover the known biorelevant chemotypes. To investigate this question, we sought to determine the number of commercially available fragments that are substructures of three sets of biorelevant molecules: (1) FDA-approved drugs, (2) drugs and metabolites, and (3) drugs, metabolites, and natural products.[34] No two similar molecules were included in any set to ensure diversity.

A total of 458,329 ZINC fragments were compared to these three sets of bioactive molecules;[35] fragments were accepted as substructures if they matched a full substructure of any molecule, with a small tolerance of variability. To tile the bioactive molecules with fragments, allowing for some tolerance, we insisted on Tversky similarity of $\geq$ 0.6; this metric is widely used to compare fragments to larger molecules[36] (Figure 4 and Supplementary Table 6). Having found all purchasable fragments that match substructures among drugs, metabolites, and natural products (Table 2, column 3), we applied the dissimilarity criterion to arrive at a diverse set (Figure 4 and Table 2, column 4). Over 6,000

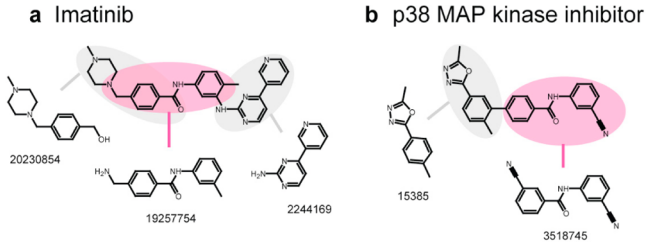


Figure 4. Fragments tiling substructures of biogenic molecules. (a) Three representative ZINC fragments that are substructures of Imatinib. (b) Three representative ZINC fragments that are substructures of DB07833 (p38 MAP Kinase inhibitor). Fragment 19257754 in Imatinib is similar to fragment 3518745 in DB07833 in chemical path fingerprints (CP Tc = 0.725; ECFP_4 Tc = 0.402), and only one of them was kept in the maximally diverse fragment set (see column 4 in Table 2).

Table 2. Number of Diverse Fragments Tiling Biologically Relevant Small Molecules

biological space	molecules	purchasable fragments[a]	dissimilar fragments[a]
(1) FDA approved drugs	1,948	15,096	6,019
(2) FDA approved drugs + metabolites	57,203	89,482	27,205
(3) FDA approved drugs + metabolites + natural products	251,353	117,526	32,373

[a]This column counts only fragments also matching bioactive molecules (see Methods). ZINC codes and SMILES strings for each fragment subset (column 4) are available as Supporting Information and at http://zinc.docking.org/full_space_fragments.

fragments are substructures of the FDA-approved drugs only. To tile all drugs, metabolites and natural products, 32,323 fragments are needed. Full-coverage fragment libraries have thus been created and made available here (http://zinc.docking.org/full_space_fragments/ and as Supporting Information).

DISCUSSION

A key observation to emerge from this study is the complementarity of empirical and structure-guided fragment screens. A strong suit of the empirical screen was its illumination of wholly new chemotypes for β-lactamase, many apparently binding in new sites (Figure 1). This has been seen often in fragment-based screens, one of whose strengths is the ability to find chemotypes that are under-represented in screens of larger molecules.[37] A strength of the computational screen was its prioritization of potent molecules (Figure 2 and Table 1), often substantially more so than those found empirically, and its ability to highlight chemotypes unexplored empirically. Both techniques also had weaknesses, to which we return, but these were often compensated by the other approach.

The nine inhibitor fragments emerging from the NMR screen had chemotypes previously unknown among AmpC inhibitors, consistent with the ability of fragment screens to explore new areas of chemical space.[7] This was even more true of the fragments detected both by NMR and by SPR, but showing no competition with a known active-site ligand and no enzyme inhibition (Figure 1). Most of these likely bind to AmpC pockets other than the active site, a supposition supported by the crystal structure of the **41**/AmpC complex (Figure 3e and Supplementary Figure 4e). This illustrates the ability of fragment screens to suggest not only novel

dx.doi.org/10.1021/cb5001636 | ACS Chem. Biol. 2014, 9, 1528−1535

chemotypes but also new binding sites, something active site-targeted docking cannot achieve. This is an advantage, for instance, when seeking allosteric modulators. Admittedly, the ZINC library was specifically designed to avoid refinding ZoBio-like molecules, so the area of chemical space that the ZoBio library explored was excluded from our search. Still, despite several detailed docking studies,[22,29,33] we have not previously found series of this type.

If empirical fragment screens can find enough novel chemotypes, why bother with docking; is not empirical fragment screening sufficient? Three arguments support combining the two techniques. First, by interrogating binding hotspots with libraries that are 2−3 orders of magnitude larger, docking can find more ligand-efficient and therefore potentially more optimizable fragments. Thus, 10 of 18 high-ranking docked fragments inhibited AmpC, and for several of their K_I and ligand efficiency (LE) values were as good as 30 μM and 0.42 kcal/mol/HAC, substantially better than that achieved by the NMR-derived fragment hits. The difference in potency is unlikely to represent simply physical properties as, for instance, the molecular weight of the two screening sets largely overlap, with the NMR set ranging from 100 to 310 Da (median 210 Da), and the docking set ranging from 46 to 250 Da (median 214 Da). Second, as efficiently as a particular fragment library covers chemical space, it can only go so far with 1,000−10,000 molecules. Just to include commercially available fragments that are represented in biogenic molecules, one would need at least 32,373 diverse compounds. Any smaller library will miss fragments that have been sampled by therapeutics or by nature (Supplementary Table 6). Third, to both maximize diversity and respect size constraints, empirical libraries must choose among related analogues, whereas different analogues may be better suited to different targets. This is illustrated by the crystal structures of fragments **44**, **48**, and **60**. These fragments are closely related, and ordinarily only one of them would be represented in a diverse library. However, the small fragment **44** binds at a distal site, while the more elaborated fragments **48** and **60** bind at the catalytic site, consistent with the docking predictions (Figure 3f, g, and i). Whereas all three inhibit AmpC with good ligand efficiencies (0.29−0.35) and fit their local pockets well, only the catalytic pocket has a geometry that supports further elaboration of this series (Supplementary Figure 5).[29,38] Substantial changes in binding mode or binding site upon fragment optimization have been previously observed,[38−40] reflecting the relatively low affinities and specificities of these molecules.[41] To ensure that the right instance of a chemotype is sampled, and this will change from target to target, one can either include multiple analogues in a library, expanding it still further, or be alert to changes in geometry as initial fragment hits are elaborated upon.

Certain caveats bear airing. First, there is a limit to the number of fragments that can be pragmatically tested, and the size of current fragment libraries has repeatedly been shown to be sufficient to find interesting chemotypes. Indeed, a lesson of this study is that even a relatively small fragment library, carefully chosen, can find novel and attractive chemotypes. Second, the failure of docking to prioritize these novel molecules partly reflects failures to predict specific orientations, as with fragments **5**, **20**, and **41**, whose X-ray structures do not correspond with the docking predictions (Figure 3a, d, and e). Even when the NMR fragments were more or less correctly posed, as with fragments **13** and **16**, their relatively low scores place them far below the top-ranking docking hits. This, in turn, reflects well-known problems with docking scoring and sampling, the improvement of which remains an area of active research.[42−44] When fragments had poor affinities, as with **5**, or mediocre scores, as with **20**, or both, as with **41**, there was poor correspondence between the docking and crystallographic poses. There was better correspondence between the docking and crystallographic poses of the docking-derived fragments, reflecting both their selection on the basis of highly favorable docking scores, and their affinities, which were typically higher than those of the NMR-derived hits.

Empirical screening, such as the TINS approach used here, will thus remain at the heart of fragment-based discovery. What this study suggests is that empirical and structure-based screening can complement each other effectively, filling gaps left by either technique used alone. Empirical screening can find chemotypes and geometries without precedent among extant inhibitors, even in a system as heavily targeted as β-lactamase. Computational screening can access a larger chemical space, prioritizing scaffolds and chemotypes absent from the empirical library. One of the great strengths of the fragment-based approach is that it can dramatically expand the sampling of ligand chemotypes.[6,7] This may be further fortified by accessing essentially all of known biogenic chemotypes, by structure-based interrogation of the over 500,000 fragments that are commercially available. A lesson of this study is that the two approaches may be pragmatically combined, increasing chemical space coverage and the potency of the fragment hits without sacrificing their novelty and with little extra cost in resources.

METHODS

Target-Immobilized NMR Screening (TINS) against AmpC β-Lactamase. AmpC was expressed and purified as described.[45] A subset of the ZoBio library (1,281 commercially available fragments) was screened for AmpC ligands using TINS,[24] applying mixes of 3−9 compounds (500 μM) to the immobilized protein (50 μM). Hits from the primary screen were tested for competition with benzo[*b*]-thiophene-2-boronic acid (compound **S3**, Supplementary Figure 1) in a secondary TINS screen (Supporting Methods).

SPR Experiments. Fragment binding to AmpC was measured on a Biacore T200 instrument (GE Healthcare). Fragments were assayed at 6 concentrations between 23 and 750 μM in steps of 2x dilutions and, if necessary, at a maximal concentration of 1500 μM (Supporting Methods).

Docking. The fragment sets were prepared using the standard protocol used for ligands in the ZINC database.[35] Fragments from the ZINC fragment set had Tanimoto coefficients (Tc) of 0.4 or lower (ECFP_4 fingerprints) compared to any ZoBio fragment screened experimentally. All docking calculations were carried out with DOCK 3.6[46,47] and solvent-excluded volume ligand desolvation,[31] using a 1.94 Å crystallographic structure of AmpC β-lactamase (PDB code 1L2S) (Supporting Methods).

AmpC β-Lactamase Biochemical Assays and Competition Experiments. Enzyme inhibition was measured by the method of initial rates. Fragments with K_I values below 10 mM were considered active (Supporting Methods).

Crystal Growth and Structure Determination. AmpC structures in complex with **16** and **50** were obtained by co-crystallization. AmpC structures in complex with **5**, **13**, **20**, **41**, **44**, **48**, **54**, and **60** were obtained by soaking the crystals into the respective ligand solution. All structures were determined by molecular replacement (Supporting Methods and Supplementary Table 5).

Chemoinformatics. The three subsets defined as biologically relevant molecules were (1) FDA-approved drugs; (2) subset 1, plus experimental drugs and natural metabolites; (3) subsets 1, 2, plus all compounds defined as biogenic molecules in ZINC.[35] A substructure

dx.doi.org/10.1021/cb5001636 | *ACS Chem. Biol.* 2014, 9, 1528−1535

search on 458,329 fragments from ZINC was performed using RDKIT (rdkit.org) (Tversky weights $\alpha = 1$, $\beta = 0$, and $\gamma = 0.6$). Only fragments with Tc ≤ 0.7 in ChemAxon path fingerprints and Tc < 0.7 in ChemAxon ECFP_4 fingerprints were kept (Supporting Methods).

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

AmpC β-lactamase in complex with **5** (4KZ5), **13** (4KZ6), **16** (4KZ7), **20** (4KZ8), **41** (4KZ9), **44** (4KZ3), **48** (4KZA), **50** (4KZB), and **60** (4KZ4).

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: bshoichet@gmail.com.

*E-mail: g.siegal@chem.leidenuniv.nl.

Present Address

[¶]Pharma Research and Early Development (pRED), F. Hoffmann-La Roche Ltd., CH-4070 Basel, Switzerland.

Author Contributions

[#]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank A. Doak for purifying AmpC and M. Fischer and A. Doak for reading this manuscript. Supported by U.S. NIH grant GM59957 (to B.K.S.).

■ ABBREVIATIONS

HAC, heavy atom count; NMR, nuclear magnetic resonance; rmsd, root-mean-square deviation; SPR, surface plasmon resonance; TINS, target-immobilized NMR screening

■ REFERENCES

(1) Erlanson, D. A. (2012) Introduction to fragment-based drug discovery. *Top. Curr. Chem. 317*, 1−32.

(2) Bollag, G., Hirth, P., Tsai, J., Zhang, J., Ibrahim, P. N., Cho, H., Spevak, W., Zhang, C., Zhang, Y., Habets, G., Burton, E. A., Wong, B., Tsang, G., West, B. L., Powell, B., Shellooe, R., Marimuthu, A., Nguyen, H., Zhang, K. Y. J., Artis, D. R., Schlessinger, J., Su, F., Higgins, B., Iyer, R., D/'Andrea, K., Koehler, A., Stumm, M., Lin, P. S., Lee, R. J., Grippo, J., Puzanov, I., Kim, K. B., Ribas, A., McArthur, G. A., Sosman, J. A., Chapman, P. B., Flaherty, K. T., Xu, X., Nathanson, K. L., and Nolop, K. (2010) Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature 467*, 596−599.

(3) Jhoti, H., Williams, G., Rees, D. C., and Murray, C. W. (2013) The 'rule of three' for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discovery 12*, 644−645.

(4) Hopkins, A. L., Groom, C. R., and Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today 9*, 430−431.

(5) Hajduk, P. J. (2006) Fragment-based drug design: how big is too big? *J. Med. Chem. 49*, 6972−6976.

(6) Fink, T., and Reymond, J.-L. (2007) Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: Assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery. *J. Chem. Inf. Model. 47*, 342−353.

(7) Leach, A. R., and Hann, M. M. (2011) Molecular complexity and fragment-based drug discovery: ten years on. *Curr. Opin. Chem. Biol. 15*, 489−496.

(8) Giannetti, A. M. (2011) From experimental design to validated hits a comprehensive walk-through of fragment lead identification using surface plasmon resonance. *Methods Enzymol. 493*, 169−218.

(9) Campos-Olivas, R. (2011) NMR screening and hit validation in fragment based drug discovery. *Curr. Top. Med. Chem. 11*, 43−67.

(10) Lau, W., Withka, J., Hepworth, D., Magee, T., Du, Y., Bakken, G., Miller, M., Hendsch, Z., Thanabal, V., Kolodziej, S., Xing, L., Hu, Q., Narasimhan, L., Love, R., Charlton, M., Hughes, S., van Hoorn, W., and Mills, J. (2011) Design of a multi-purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *J. Comput.-Aided Mol. Des. 25*, 621−636.

(11) Campobasso, N. (2012) Picking up the fragments at GlaxoSmithKline: No longer a last chance effort for lead ID, presented at the Fragment-Based Lead Discovery Conference, Sept 23−26, 2012, San Francisco, CA.

(12) Roughley, S. D., and Hubbard, R. E. (2011) How well can fragments explore accessed chemical space? A case study from heat shock protein 90. *J. Med. Chem. 54*, 3989−4005.

(13) Siegel, M. G., and Vieth, M. (2007) Drugs in other drugs: a new look at drugs as fragments. *Drug Discovery Today 12*, 71−79.

(14) Silvestre, H. L., Blundell, T. L., Abell, C., and Ciulli, A. (2013) Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery. *Proc. Natl. Acad. Sci. U.S.A. 110*, 12984−12989.

(15) Friberg, A., Vigil, D., Zhao, B., Daniels, R. N., Burke, J. P., Garcia-Barrantes, P. M., Camper, D., Chauder, B. A., Lee, T., Olejniczak, E. T., and Fesik, S. W. (2013) Discovery of potent myeloid cell leukemia 1 (Mcl-1) inhibitors using fragment-based methods and structure-based design. *J. Med. Chem. 56*, 15−30.

(16) Christopher, J. A., Brown, J., Dore, A. S., Errey, J. C., Koglin, M., Marshall, F. H., Myszka, D. G., Rich, R. L., Tate, C. G., Tehan, B., Warne, T., and Congreve, M. (2013) Biophysical fragment screening of the beta1-adrenergic receptor: identification of high affinity arylpiperazine leads using structure-based drug design. *J. Med. Chem. 56*, 3446−3455.

(17) Geitmann, M., Elinder, M., Seeger, C., Brandt, P., de Esch, I. J., and Danielson, U. H. (2011) Identification of a novel scaffold for allosteric inhibition of wild type and drug resistant HIV-1 reverse transcriptase by fragment library screening. *J. Med. Chem. 54*, 699−708.

(18) de Graaf, C., Rein, C., Piwnica, D., Giordanetto, F., and Rognan, D. (2011) Structure-based discovery of allosteric modulators of two related class B G-protein-coupled receptors. *ChemMedChem 6*, 2159−2169.

(19) Ben-David, M., Wieczorek, G., Elias, M., Silman, I., Sussman, J. L., and Tawfik, D. S. (2013) Catalytic metal ion rearrangements underline promiscuity and evolvability of a metalloenzyme. *J. Mol. Biol. 425*, 1028−1038.

(20) Marcou, G., and Rognan, D. (2007) Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model. 47*, 195−207.

(21) Schrader, F. C., Glinca, S., Sattler, J. M., Dahse, H. M., Afanador, G. A., Prigge, S. T., Lanzer, M., Mueller, A. K., Klebe, G., and Schlitzer, M. (2013) Novel type II fatty acid biosynthesis (FAS II) inhibitors as multistage antimalarial agents. *ChemMedChem. 8*, 442−461.

(22) Chen, Y., and Shoichet, B. K. (2009) Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol. 5*, 358−364.

(23) Chen, D., Ranganathan, A., Ijzerman, A. P., Siegal, G. D., and Carlsson, J. (2013) Complementarity between in silico and biophysical screening approaches in fragment-based lead discovery against the A2A adenosine receptor. *J. Chem. Inf. Model. 53*, 2701−2714.

(24) Siegal, G., and Hollander, J. G. (2009) Target immobilization and NMR screening of fragments in early drug discovery. *Curr. Top. Med. Chem. 9*, 1736−1745.

(25) Vakulenko, S. B., Golemi, D., Geryk, B., Suvorov, M., Knox, J. R., Mobashery, S., and Lerner, S. A. (2002) Mutational replacement of Leu-293 in the class C Enterobacter cloacae P99 beta-lactamase confers increased MIC of cefepime. *Antimicrob. Agents Chemother. 46*, 1966−1970.
(26) Nukaga, M., Kumar, S., Nukaga, K., Pratt, R. F., and Knox, J. R. (2004) Hydrolysis of third-generation cephalosporins by class C β-lactamases. *J. Biol. Chem. 279*, 9344−9352.
(27) Drawz, S. M., Taracila, M., Caselli, E., Prati, F., and Bonomo, R. A. (2011) Exploring sequence requirements for C(3)/C(4) carboxylate recognition in the Pseudomonas aeruginosa cephalosporinase: Insights into plasticity of the AmpC beta-lactamase. *Protein sci. 20*, 941−958.
(28) Babaoglu, K., Simeonov, A., Irwin, J. J., Nelson, M. E., Feng, B., Thomas, C. J., Cancian, L., Costi, M. P., Maltby, D. A., Jadhav, A., Inglese, J., Austin, C. P., and Shoichet, B. K. (2008) Comprehensive mechanistic analysis of hits from high-throughput and docking screens against beta-lactamase. *J. Med. Chem. 51*, 2502−2511.
(29) Powers, R. A., Morandi, F., and Shoichet, B. K. (2002) Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. *Structure 10*, 1013−1023.
(30) Eidam, O., Romagnoli, C., Dalmasso, G., Barelier, S., Caselli, E., Bonnet, R., Shoichet, B. K., and Prati, F. (2012) Fragment-guided design of subnanomolar beta-lactamase inhibitors active in vivo. *Proc. Natl. Acad. Sci. U.S.A. 109*, 17448−17453.
(31) Mysinger, M. M., and Shoichet, B. K. (2010) Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model. 50*, 1561−1573.
(32) Peng, Z. W., Gillespie, P., Weisel, M., So, S. S., So, W. V., Kondru, R., Narayanan, A., and Hermann, J. C. (2013) A Crowd-Based Process and Tool for HTS Hit Triage. *Mol. Inform. 32*, 337−345.
(33) Teotico, D. G., Babaoglu, K., Rocklin, G. J., Ferreira, R. S., Giannetti, A. M., and Shoichet, B. K. (2009) Docking for fragment inhibitors of AmpC beta-lactamase. *Proc. Natl. Acad. Sci. U.S.A. 106*, 7455−7460.
(34) Over, B., Wetzel, S., Grutter, C., Nakai, Y., Renner, S., Rauh, D., and Waldmann, H. (2013) Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem. 5*, 21−28.
(35) Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., and Coleman, R. G. (2012) ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model. 52*, 1757−1768.
(36) Hudson, S. A., McLean, K. J., Surade, S., Yang, Y. Q., Leys, D., Ciulli, A., Munro, A. W., and Abell, C. (2012) Application of fragment screening and merging to the discovery of inhibitors of the Mycobacterium tuberculosis cytochrome P450 CYP121. *Angew. Chem., Int. Ed. 51*, 9311−9316.
(37) Makara, G. M. (2007) On sampling of fragment space. *J. Med. Chem. 50*, 3214−3221.
(38) Babaoglu, K., and Shoichet, B. K. (2006) Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol. 2*, 720−723.
(39) Barelier, S., Pons, J., Marcillat, O., Lancelin, J. M., and Krimm, I. (2010) Fragment-based deconstruction of Bcl-xL inhibitors. *J. Med. Chem. 53*, 2577−2588.
(40) Good, A. C., Liu, J., Hirth, B., Asmussen, G., Xiang, Y., Biemann, H. P., Bishop, K. A., Fremgen, T., Fitzgerald, M., Gladysheva, T., Jain, A., Jancsics, K., Metz, M., Papoulis, A., Skerlj, R., Stepp, J. D., and Wei, R. R. (2012) Implications of promiscuous Pim-1 kinase fragment inhibitor hydrophobic interactions for fragment-based drug design. *J. Med. Chem. 55*, 2641−2648.
(41) Barelier, S., and Krimm, I. (2011) Ligand specificity, privileged substructures and protein druggability from fragment-based screening. *Curr. Opin. Chem. Biol. 15*, 469−474.
(42) Repasky, M. P., Murphy, R. B., Banks, J. L., Greenwood, J. R., Tubert-Brohman, I., Bhat, S., and Friesner, R. A. (2012) Docking performance of the glide program as evaluated on the Astex and DUD datasets: a complete set of glide SP results and selected results for a new scoring function integrating WaterMap and glide. *J. Comput.-Aided Mol. Des. 26*, 787−799.
(43) Neves, M. C., Totrov, M., and Abagyan, R. (2012) Docking and scoring with ICM: the benchmarking results and strategies for improvement. *J. Comput.-Aided Mol. Des. 26*, 675−686.
(44) Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem. 31*, 455−461.
(45) Usher, K. C., Blaszczak, L. C., Weston, G. S., Shoichet, B. K., and Remington, S. J. (1998) Three-dimensional structure of AmpC β-lactamase from Escherichia coli bound to a transition-state analogue: Possible implications for the oxyanion hypothesis and for inhibitor design. *Biochemistry 37*, 16082−16092.
(46) Irwin, J. J., Shoichet, B. K., Mysinger, M. M., Huang, N., Colizzi, F., Wassam, P., and Cao, Y. (2009) Automated docking screens: a feasibility study. *J. Med. Chem. 52*, 5712−5720.
(47) Lorber, D. M., and Shoichet, B. K. (2005) Hierarchical docking of databases of multiple ligand conformations. *Curr. Top. Med. Chem. 5*, 739−749.