nature chemical biology

Deconstructing fragment-based inhibitor discovery

Kerim Babaoglu & Brian K Shoichet

Fragment-based screens test multiple low–molecular weight molecules for binding to a target[1–4]. Fragments often bind with low affinities but typically have better ligand efficiencies (ΔG_{bind}/heavy atom count) than traditional screening hits[5]. This efficiency, combined with accompanying atomic-resolution structures, has made fragments popular starting points for drug discovery programs[2,6–13]. Fragment-based design adopts a constructive strategy: affinity is enhanced either by cycles of functional-group addition or by joining two independent fragments together. The final inhibitor is expected to adopt the same geometry as the original fragment hit. Here we consider whether the inverse, deconstructive logic also applies—can one always parse a higher-affinity inhibitor into fragments that recapitulate the binding geometry of the larger molecule? Cocrystal structures of fragments deconstructed from a known β-lactamase inhibitor suggest that this is not always the case.

To investigate the deconstructive hypothesis for fragment binding, we turned to a 'lead-like' inhibitor of AmpC β-lactamase, asking whether it can be parsed into component fragments that will still bind the enzyme, in geometries corresponding to those observed in the original lead. The lead AmpC inhibitor (**L1**, **Fig. 1**)[14] seemed to provide a good template for these questions. The molecule has a K_i of 1 μM, a satisfactory ligand efficiency of 0.38 kcal mol^{-1} and a molecular weight of 343 Da, placing it in the lead-like range.

We deconstructed the lead **L1** into three fragments, **F1**, **F2** and **F3**, which were components of the larger molecule (**Fig. 1**). **F1** is the 2-carboxythiophene, which as part of **L1** binds the oxyanion hole of AmpC, and **F2** and **F3** are different versions of the benzoic acid sulfonamide at the other end of **L1**. The three fragments inhibited the enzyme competitively in the millimolar range, with apparent K_i values of 40 mM, 19 mM and 10 mM for **F1**, **F2** and **F3**, respectively. We note that enzymatic K_i values may not reflect K_d values, as it is not certain, on the basis of the crystal structures, that all of the fragments would lead to full enzyme inhibition. Another indication of affinity is the 12.5-mM concentration of each fragment present in the crystallizing buffer. Each fragment was cocrystallized with AmpC, with the exception of **F2**, which was soaked into an apoprotein crystal, and their crystal structures were determined to 2.1, 2.2 and 1.2 Å, respectively (see **Supplementary Table 1** online for crystallography statistics). Electron densities for these ligands were unambiguous in initial $F_o - F_c$ maps calculated without ligand present (see **Supplementary Fig. 1** online for all electron density maps).

Contrary to our expectation, none of the three fragments bound in a position that recapitulated its placement in the larger **L1** lead. Instead, the fragments probed two entirely new sites (**Fig. 2**). In the original X-ray structure of the full **L1**, the thiophene carboxylate bound at the heart of the oxyanion hole of β-lactamase, long thought to be a hot spot for substrate and inhibitor recognition[15,16]. Indeed, one of our concerns in beginning this study was that this hot spot would act as a sink for all of the fragments, as they each bear carboxylates from different parts of **L1**. However, none of the fragments bound in this site. Rather, the thiophene carboxylate **F1** bound in an entirely new site never previously observed in AmpC, with its carboxylate group 8 Å away from the oxyanion-hole position it adopts as part of **L1** (**Fig. 2**). This new site is immediately adjacent to the core of the active site and involves a salt bridge with Arg148 and π-π stacking interactions with Tyr150. The pocket is only fully evident in the presence of the ligand; it is formed by the partial unwinding of an α-helix and the insertion of Lys290 back into the protein, where it interacts with both the ligand and the protein. This level of conformational change is rarely seen in class C β-lactamases[17], and initially we were tempted to dismiss it as a curiosity. However, a second of the fragments, **F3**, bound this site and invoked the same conformational change, implying that the site offers important recognition elements to ligand functional groups (**Figs. 2** and **3**). Also, the crystallization conditions were no different than those of either the apo structure or other holo structures, none of which undergo this conformational change, suggesting it is not an artifact of crystallization but a specific accommodation to these fragments. AmpC is typically a rigid enzyme relatively unperturbed by inhibitor binding; its accommodation to the fragments, when the lead **L1** causes no such

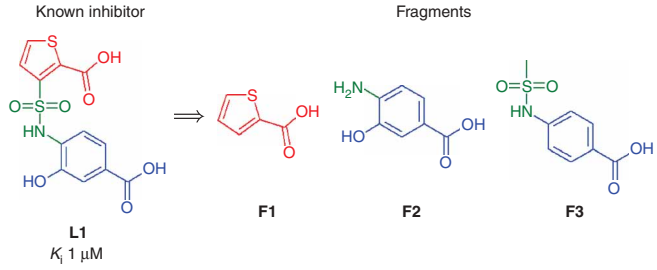

Figure 1 Deconstruction of a β-lactamase inhibitor into fragments. The known inhibitor **L1** was divided into three commercially available fragments, **F1**, **F2** and **F3**, each containing an aryl carboxylate.

Department of Pharmaceutical Chemistry, University of California San Francisco, 1700 4th St., MC 2550, San Francisco, California 94158-2330, USA. Correspondence should be addressed to B.K.S. (shoichet@cgl.ucsf.edu).

Received 19 June; accepted 30 August; published online 29 October 2006; doi:10.1038/nchembio831

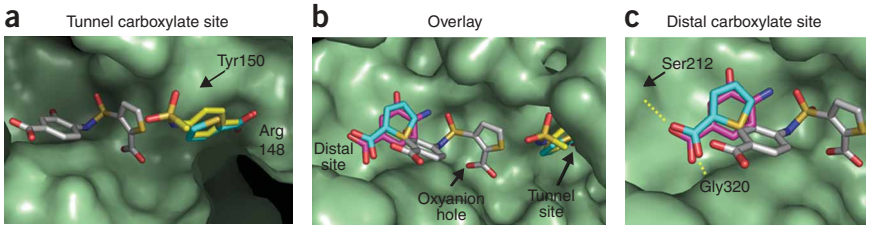

Figure 2 Fragment binding geometries. Crystal structures were obtained for each of the three fragments (colored carbons: cyan, **F1**; magenta, **F2**; yellow, **F3**; white, **L1**) bound to β-lactamase. (**a**) Binding modes of fragments **F1** and **F3** compared to the lead **L1**. (**b**) Overlay of **F1**, **F2**, **F3** and **L1**. (**c**) Binding modes of fragments **F1** and **F2** compared to **L1**.

change, is noteworthy. The origins and prediction of conformational change remain an area of active research[18]. For the purposes of this study, it is sufficient to note that the changes we observed arose as a result of a new binding mode for fragments that had the option of binding where the lead did, and that these changes produced a new and unusual pocket in the active site of the enzyme.

Fragment **F2** bound in a second previously unexplored carboxylate-binding site (**Fig. 2c**). **F1**, in a second binding mode, was also seen in this site, strengthening its characterization as a carboxylate-binding site. **F2** was in the same region as the distal carboxylate in **L1** that it was meant to imitate, but it was displaced by a full ring, so that the carboxylate moieties were 4 Å from the original position. In this new distal site, the carboxylates in the fragments were held by two hydrogen bonds with the backbone amides of Ser212 and Gly320.

How complex must a fragment of **L1** be before it retains the original binding orientation of the larger inhibitor from which it is derived? To investigate this question, we determined the structure of a fourth fragment composed of a carboxythiophene with an amide side chain resembling the sulfonamide side chain of **L1** (**F4**, **Fig. 4**); this fragment has an apparent K_i of 5 mM. If one considers the lead **L1** to have three functional groups—an oxyanion hole–binding aryl carboxylate, an amide or sulfonamide and a distal aryl carboxylate—then the fragment **F4** has two of these three groups. Unlike the simpler carboxythiophene (**F1**), **F4** bound with its carboxylate in the oxyanion hole and recapitulated the orientation of the lead **L1**. This orientation was stabilized by addition of a second hydrogen bond formed by the carbonyl of the amide group, which bound in a canonical amide-recognition region seen across β-lactamases that was also exploited by the sulfonamide in **L1**. Thus, it may be possible for fragments of **L1** to recapitulate its binding geometry, but not before a substantial proportion of the specificity determinants of the larger inhibitor have been recapitulated, in the proper context. (**F3**, for instance, also has two of the three functional groups of **L1**, but, lacking the ability to bind in the oxyanion hot spot, it does not adopt the **L1** geometry.)

This finding speaks to the balance between complexity and diversity in fragment screening libraries in general. As fragments shrink, chemical space collapses and the number of molecules that must be included in a fragment library falls correspondingly. But this is true only to the point where fragments behave differently from the larger molecules from which they have been notionally deconstructed. In the case studied here, the binding modes of the simple fragments **F1**–**F3** differed from that adopted by the larger lead molecule **L1** from which they were decomposed. However, a more complex yet similar thiophene did bind in the **L1** mode. The implication is not that fragments should be either more or less complex; indeed, the equally complex molecules **F3** and **F4** differ in their ability to recapitulate the larger molecule. Rather, what should be taken from our observations is that, individually, both complex libraries and diverse libraries miss useful binding events that can be exploited in the development of small fragments into larger, more potent, drug-like molecules.

It is sensible to consider the caveats to this study. Most prominently, we have investigated the deconstruction of only one inhibitor of one enzyme. It may be that the lead inhibitor (**L1**) deconstructed here is suboptimal; perhaps the fragments tell us that a better lead can be built. This may be especially true of the fragments binding in the 'distal' site (**Fig. 2**), which may suggest an alternative placement of the second anionic group on **L1**. That admitted, **L1** is a solid lead-like inhibitor with good physical properties (molecular weight, hydrophobicity and hydrogen bond potential), satisfactory ligand efficiency and relatively potent activity in whole cells. It is, in fact, the only noncovalent β-lactamase inhibitor described in the literature. The failure of its fragments to recapitulate its binding is not a trivial fault of the lead.

A more general caveat is that deconstructive logic will sometimes hold for other inhibitors. In fact, in the one other case we know of where this has been explored structurally, it proved possible to deconstruct a substrate of thymidylate synthase into fragments that recapitulate the binding of the larger molecule[19]. Still, there is no strong theoretical reason to expect that deconstructive logic will generally apply, and it is easy to imagine how nonadditive effects could combine in a molecule whose binding and affinity emerge only once a critical number of functional groups are present. Indeed, there are several small molecules whose affinity far exceeds the sum of their parts. These include biotin[20], several CDK inhibitors[21] and stromelysin inhibitors[22]. In substrate recognition, too, distinct functional groups can combine with pronounced nonadditivity[23].

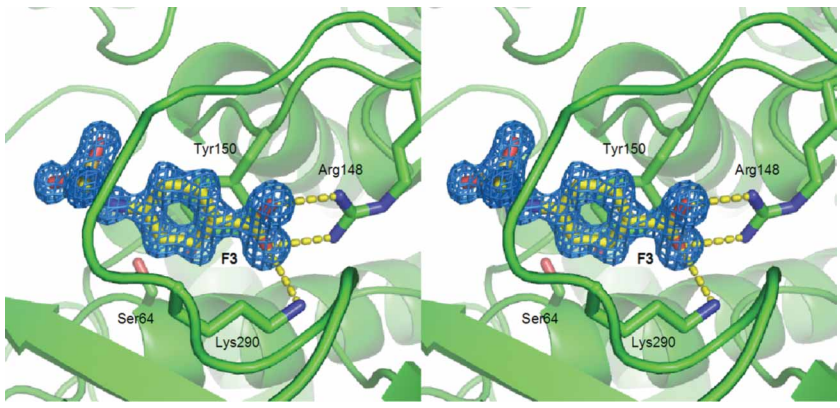

Figure 3 Detailed stereo view of tunnel carboxylate site in the structure of **F3** (yellow carbons) bound to β-lactamase (green carbons). Interacting side chains and active site Ser64 are shown as sticks to orient the viewer. The unbiased $F_o - F_c$ electron density is also shown, contoured at 3 σ.

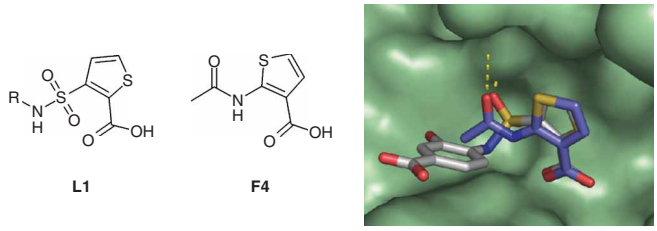

Figure 4 Increasing complexity restores binding orientation. Shown is crystal structure of fragment **F4** (purple carbons) overlaid on that of the lead **L1** (white carbons).

Fragment-based design uses a constructive logic, where additions to or combinations of fragments improve affinities while leaving the geometries of the original molecules relatively unperturbed. The lesson of this study is that the converse deconstructive logic need not hold; fragments of a larger inhibitor need not recapitulate its binding. This suggests that there will be gaps in the molecules constructed from fragments, with some good inhibitors missed by this strategy. For such molecules, enzyme recognition is an emergent, indivisible or at least not fully divisible property that depends on the presence of a critical number of covalently bonded functional groups. The existence of such inhibitors is no strike against the molecules that are constructed from fragments—far from it—but it does point to regions of chemical space that the constructive approach will not visit.

METHODS

Fragments. We purchased fragments **F1** and **F2**, **F3**, and **F4**, respectively, from Sigma-Aldrich, Epsilon Chimie and Key Organics. We used all compounds as supplied by the manufacturers without further purification.

Enzymology. We dissolved the inhibitors in DMSO at a concentration of 1 M; more dilute stocks were subsequently prepared as necessary. For kinetic measurements, we performed the assay using nitrocefin as substrate in 0.3 M sodium cacodylate buffer (pH 6.5) and monitored it in an HP8453 UV-Vis spectrophotometer at 480 nM. The K_m of nitrocefin for AmpC in this buffer was 531 μM. We made concentrated stock solutions from lyophilized powder and determined the concentration of AmpC spectrophotometrically; this enzyme was expressed and purified as described[24]. The concentration of enzyme in all reactions was 0.9 nM. We obtained K_i values by comparing progress curves in the presence and absence of inhibitor.

X-ray crystallography. We grew cocrystals of AmpC in complex with compounds **F1**, **F3** and **F4** by vapor diffusion in hanging drops equilibrated over 1.8 M potassium phosphate buffer (pH 8.7) using microseeding techniques. The initial concentration of the protein in the drop was 3.5 mg ml^{-1}, and the concentration of the fragment in each case was 12.5 mM. Crystals appeared 1–2 weeks after equilibration at 21 °C. For compound **F2**, we placed apo crystals, grown as described above, in a solution of 1.8 M potassium phosphate (pH 8.7) saturated with the fragment and soaked overnight. Before data collection, we immersed crystals in a cryo-protectant solution of 20% sucrose and 1.8 M potassium phosphate (pH 8.7) for about 30 s and flash-cooled them in liquid nitrogen.

We collected diffraction data from frozen crystals at the Advanced Light Source (Lawrence Berkeley Laboratory). We measured all data sets from single crystals, and conducted initial space group determination and created data collection strategies using ELVES[25]. We indexed, integrated and scaled the data using the HKL package[26]. For all structures, the space group was C_2, with two molecules in the asymmetric unit. We used an apo-AmpC structure (PDB 1KE4) with water molecules and ions removed for the initial phasing models. To obtain initial phases, we positioned the models using rigid-body refinement and then did an initial refinement using REFMAC5 (ref. 27). For model building and water placement, we used Coot[28], then did additional rounds of refinement with REFMAC5. All models showed good Ramachandran statistics, AmpC–**F1** having 92.9% in the most favored region and 7.1% in the additional allowed region, AmpC–**F2** having 91.6% in the most favored region and 8.4% in the additional allowed region, AmpC–**F3** having 91.7% in the most favored region and 8.3% in the additional allowed region, and AmpC–**F4** having 91.9% in the most favored region and 8.1% in the additional allowed region.

Accession codes: Protein Data Bank: coordinates and structure factors for the β-lactamase complex structures determined here have been deposited with accession codes 2HDQ (**F1**), 2HDR (**F2**), 2HDS (**F3**) and 2HDU (**F4**).

Note: Supplementary information is available on the Nature Chemical Biology website.

ACKNOWLEDGMENTS

This work was supported by US National Institutes of Health grant GM59957 (to B.K.S.) and Ruth L. Kirschstein National Research Service Award fellowship GM076883 (to K.B). We thank B. Feng, J. Irwin, A. Graves and Y. Chen for reading the manuscript.

AUTHOR CONTRIBUTIONS

K.B. and B.K.S. designed the experiments and wrote the manuscript together. K.B. did all of the actual experimental work.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturechemicalbiology
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Erlanson, D.A., McDowell, R.S. & O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **47**, 3463–3482 (2004).
2. Erlanson, D.A., Wells, J.A. & Braisted, A.C. Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
3. Verdonk, M.L. & Hartshorn, M.J. Structure-guided fragment screening for lead discovery. *Curr. Opin. Drug Discov. Devel.* **7**, 404–410 (2004).
4. Rees, D.C., Congreve, M., Murray, C.W. & Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discov.* **3**, 660–672 (2004).
5. Hopkins, A.L., Groom, C.R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **9**, 430–431 (2004).
6. Boehm, H.J. *et al.* Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *J. Med. Chem.* **43**, 2664–2674 (2000).
7. Shuker, S.B., Hajduk, P.J., Meadows, R.P. & Fesik, S.W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
8. Gill, A. New lead generation strategies for protein kinase inhibitors—fragment based screening approaches. *Mini Rev. Med. Chem.* **4**, 301–311 (2004).
9. Card, G.L. *et al.* A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. *Nat. Biotechnol.* **23**, 201–207 (2005).
10. Hartshorn, M.J. *et al.* Fragment-based lead discovery using X-ray crystallography. *J. Med. Chem.* **48**, 403–413 (2005).
11. Huth, J.R., Sun, C., Sauer, D.R. & Hajduk, P.J. Utilization of NMR-derived fragment leads in drug design. *Methods Enzymol.* **394**, 549–571 (2005).
12. Nienaber, V.L. *et al.* Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **18**, 1105–1108 (2000).
13. Blaney, J., Nienaber, V. & Burley, S.K. Fragment-based lead discovery and optimization using X-ray crystallography, computational chemistry, and high-throughput organic synthesis. in *Fragment-based Approaches in Drug Discovery* (eds. Jahnke, W. & Erlanson, D.) 215–248 (Wiley, New York, 2006).
14. 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.* **127**, 4632–4639 (2005).
15. Usher, K.C., Shoichet, B.K., Blaszczak, L., Weston, G.S. & Remington, J.R. The three dimensional structure of AmpC β-lactamase from *Escherichia coli* bound to a transition-state analog: possible implications for the oxyanion hypothesis and for inhibitor design. *Biochemistry* **37**, 16082–16092 (1998).
16. Powers, R.A. & Shoichet, B.K. Structure-based approach for binding site identification on AmpC beta-lactamase. *J. Med. Chem.* **45**, 3222–3234 (2002).
17. Crichlow, G.V., Nukaga, M., Doppalapudi, V.R., Buynak, J.D. & Knox, J.R. Inhibition of class C beta-lactamases: structure of a reaction intermediate with a cephem sulfone. *Biochemistry* **40**, 6233–6239 (2001).
18. Davis, A.M. & Teague, S.J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Edn Engl.* **38**, 736–749 (1999).
19. Stout, T.J., Sage, C.R. & Stroud, R.M. The additivity of substrate fragments in enzyme-ligand binding. *Structure* **6**, 839–848 (1998).
20. Green, N.M. Avidin. *Adv. Protein Chem.* **29**, 85–133 (1975).

21. Congreve, M.S. *et al.* Detection of ligands from a dynamic combinatorial library by X-ray crystallography. *Angew. Chem. Int. Edn Engl.* **42**, 4479–4482 (2003).
22. Hajduk, P.J. *et al.* Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. *J. Am. Chem. Soc.* **119**, 5818–5827 (1997).
23. Miller, B.G. & Wolfenden, R. Catalytic proficiency: the unusual case of OMP decarboxylase. *Annu. Rev. Biochem.* **71**, 847–885 (2002).
24. Powers, R.A. *et al.* The complexed structure and antimicrobial activity of a non-beta-lactam inhibitor of AmpC beta-lactamase. *Protein Sci.* **8**, 2330–2337 (1999).
25. Holton, J. & Alber, T. Automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci. USA* **101**, 1537–1542 (2004).
26. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
27. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
28. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).