

Why Use DOCK?

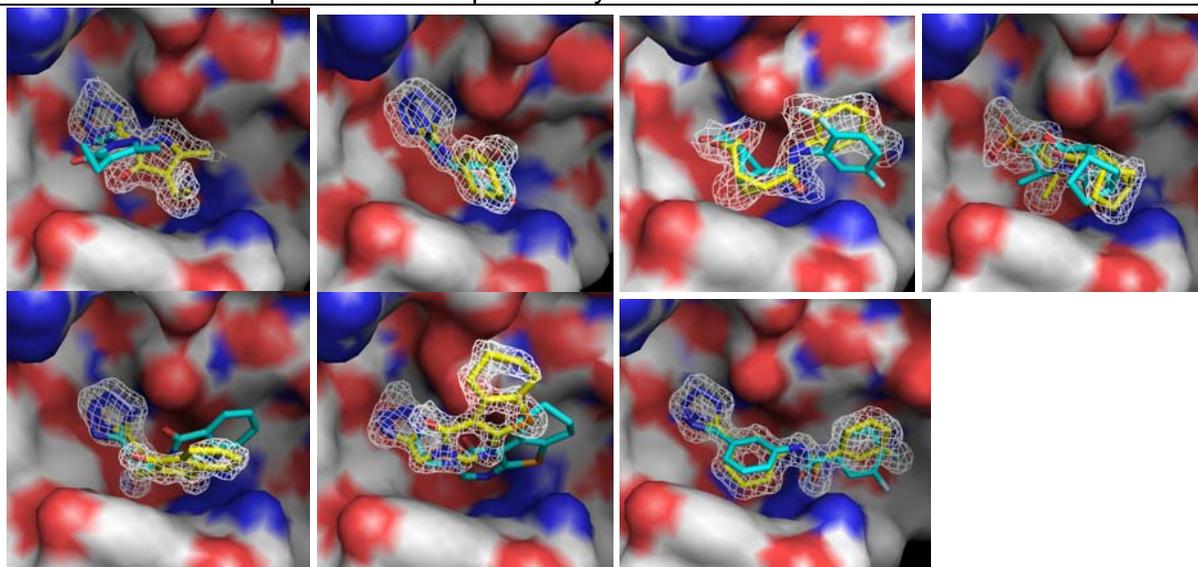
DOCK is academic software developed over many years by investigators in Tack Kuntz's lab, and for DOCK3.5 derivatives by the Shoichet group. Like any docking program, it has substantial liabilities—it under-samples degrees of freedom, treats receptor flexibility poorly, and its scoring function is crude, making it prone to both false-positives and false-negatives. It can neither predict true binding affinities nor can it even reliably rank order compounds. What it can do, in favorable circumstances, is enrich likely ligands from very large molecular databases, prioritizing those that are likely to bind for experimental testing. Its most solid support comes from the papers that have been published showing its utility. The most convincing of these are those where DOCK screens for novel ligands, these are tested experimentally, and the binding modes of the new ligands are compared to subsequently determined x-ray crystal structures. In the Shoichet lab, we use the DOCK3.5.54 version of the program, and we will focus on it here, even though DOCK6, the latest Kuntz-lab version of the code, is much more widely used, and certainly more user-friendly. Four features recommend DOCK3.5.54:

- A. It is fast, fully sampling a flexible ligand into a target in several seconds on a current core/cpu. On even a modest cluster, DOCK can screen a Million compound library in about a week.
- B. It uses a physics-based scoring function, allowing for optimization toward ever-more-reasonable physical models as computer power increases. This has made it useful for fragment screening, an area of much current interest.
- C. Whereas in its native form, DOCK3.5.54 has many variables to set, requiring expert use, it may be wrapped in a form that makes it usable by non-experts, allowing it to be directly deployed across a site (DOCK Blaster).[Irwin, 2009]
- D. **More novel, docking hits have been compared to subsequent x-ray structures from DOCK virtual screens than all other docking programs combined.**

It is perhaps the final point that is most compelling. Since 2002, 10 papers have appeared in which novel ligands have been predicted by docking, from our lab alone, and subjected to detailed mechanistic testing, typically involving crystallography. These include:

1. Chen Y, Shoichet BK. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nature Chemical Biology* **5**, 358-364 (2009).

A fragment library was screened, and new fragments were tested experimentally. The docking predicted structures were compared to subsequent x-ray results:

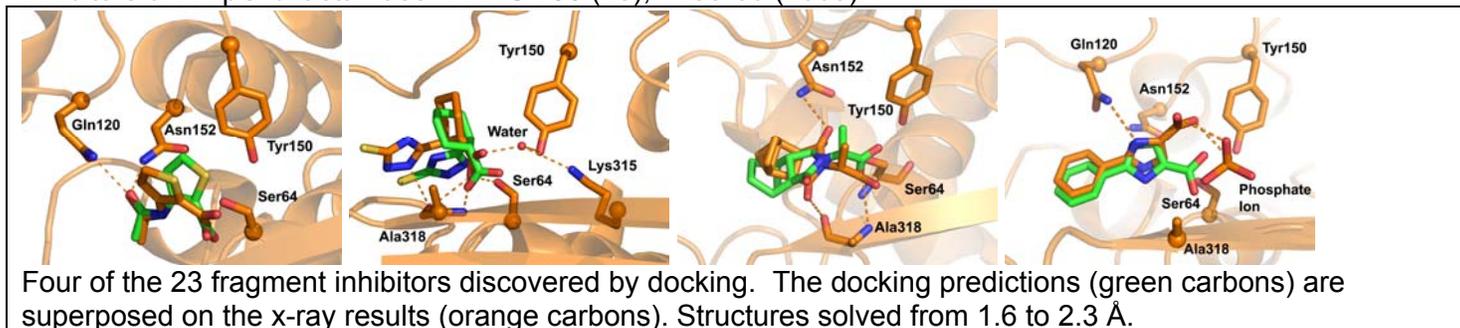


New fragment inhibitors of CTX-M β -lactamase. Docking predictions (blue) and x-ray results (yellow). Most structures to better than 1.8 Å resolution.

2. Kolb P, Rosenbaum DM, Irwin JJ, Fung JJ, Kobilka BK, Shoichet BK. Structure-based discovery of β 2-adrenergic receptor ligands. *PNAS* **106**, 6843-6848 (2009).

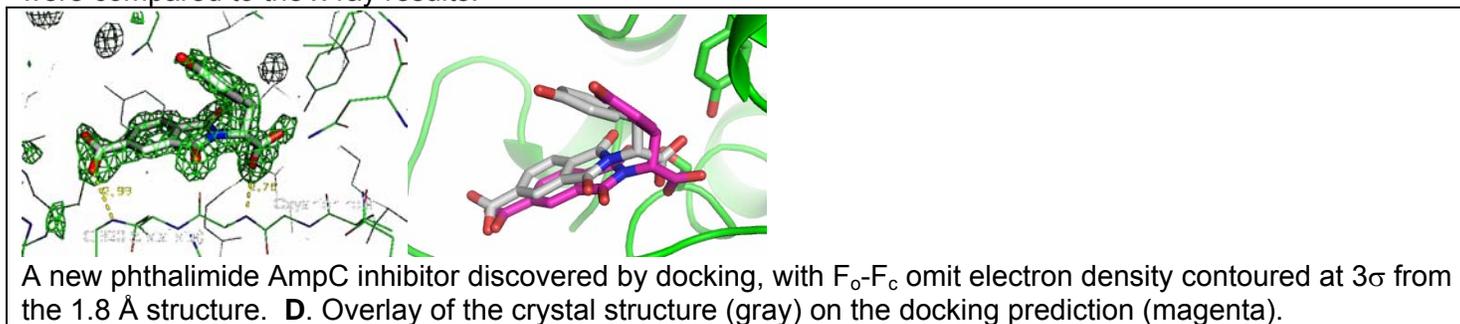
A Million lead-like inhibitors were docked against the structure of the GPCR β 2 Adrenergic receptor. Six new inverse-agonists were found, the best of which had a K_i of 9 nM.

3. Teotico DG, Babaoglu K, Rocklin GJ, Ferreira RS, Giannetti AM, Shoichet BK. Docking for fragment inhibitors of AmpC β -lactamase. *PNAS* **106** (18), 7455-60 (2009).



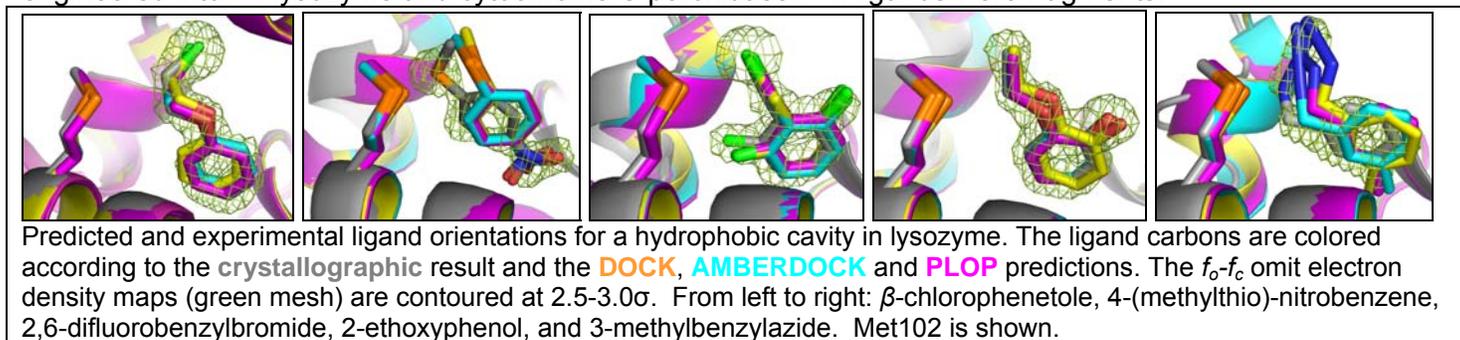
4. Babaoglu K, et al. Austin CP, Shoichet BK. Comprehensive mechanistic analysis of hits from high-throughput and docking screens against beta-lactamase. *J Med Chem* **51**, 2502-11 (2008).

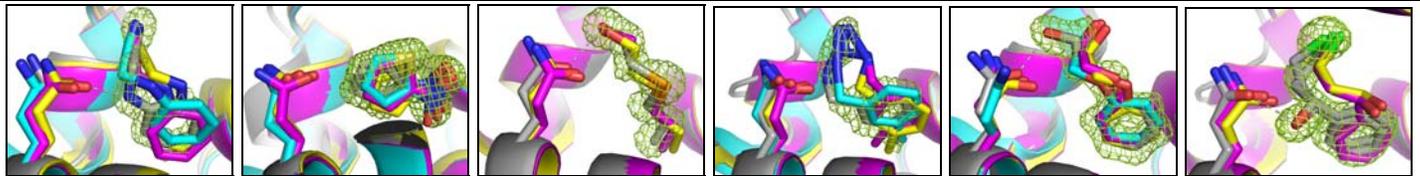
Parallel DOCK and HTS campaigns were conducted against the MLSMR library. The HTS hit rate, in the end, was 0%, whereas docking found inhibitors overlooked by the experimental screen. The docking predictions were compared to the x-ray results:



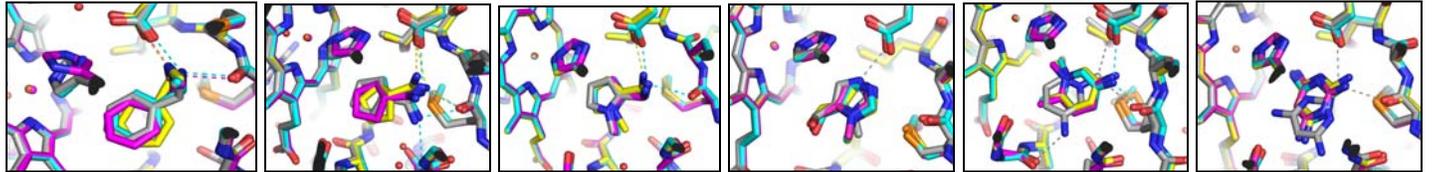
5. Graves AP, Shivakumar DM, Boyce SE, Jacobson MP, Case DA, Shoichet BK. Rescoring docking hit lists for model cavity sites: predictions and experimental testing. *J Mol Biol* **377**, 914-34 (2008).

Detailed comparison of DOCK predictions to subsequent crystallographic results in a model cavities engineered into T4 lysozyme and cytochrome C peroxidase. All ligands were fragments.



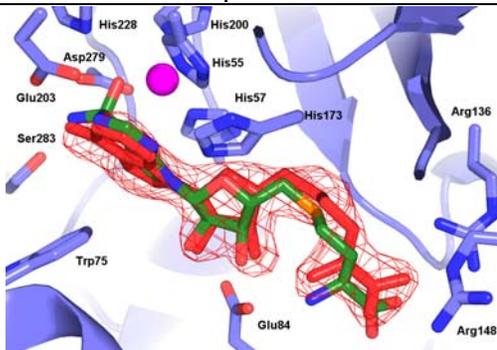


Predicted and experimental ligand orientations for a polar cavity. Left to right: n-phenylglycinonitrile, 2-nitrothiophene, 2-(n-propylthio)ethanol, 3-methylbenzylazide, 2-phenoxyethanol, and 3-chloro-1-phenyl-1-propanol. Structures colored as in C.2.6. Gln102 is shown, as are electron density maps.



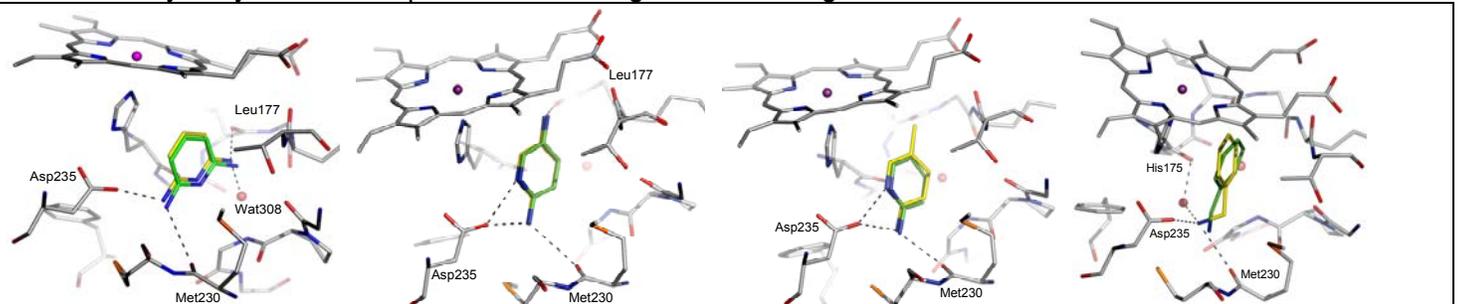
Docking poses overlaid on x-ray results for ligands of the anionic W191G cavity (6 of 11 structures shown). Left to right: n-methylbenzylamine, cyclopentane carboximidamide, 1-methyl-pyrrol)-methylamine, 1-methyl-5-imidazolecarboxaldehyde, 1,2-dimethyl-pyridin-5-amine, pyrimidine-2,4,6-triamine. Hydrogen bonds to Asp235 are shown.

6. Hermann JC, Marti-Arbona R, Fedorov AA, Fedorov E, Almo SC, Shoichet BK, Raushel FM. Structure-based activity prediction for an enzyme of unknown function. *Nature* **448**,775-9 (2007).
Use of DOCK to predict the substrate of an enzyme of previously unknown function.



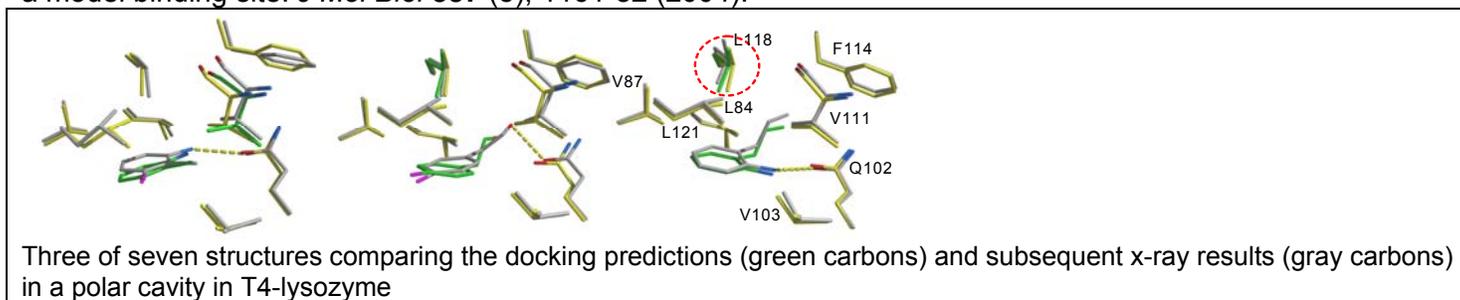
Docking prediction (green, transition-state form) superimposed on subsequent x-ray result for s-adenosyl homocysteine (green, inosyl product) as a substrate for TM0936 (blue).

7. Brenk R, Vetter SW, Boyce SE, Goodin DB, Shoichet BK. Probing molecular docking in a charged model binding site. *J Mol Biol* **357**, 1449-70 (2006).
Detailed comparison of DOCK predictions to subsequent crystallographic results in an engineered model anionic cavity in cytochrome C peroxidase. All ligands were fragments.

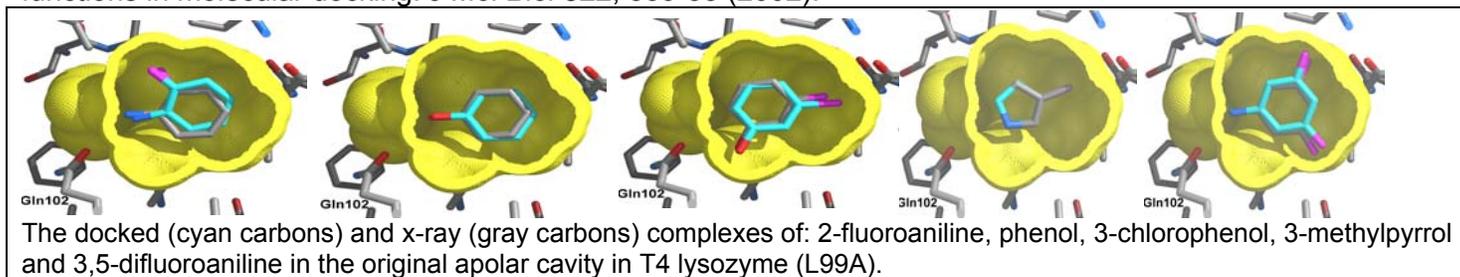


Four of 14 x-ray structures determined for docking-predicted ligands of the W191G anionic cavity. The docking predictions (carbons in green) are overlaid on the x-ray result (carbons in yellow).

8. Wei BQ, Weaver L, Ferrari AM, Matthews BM, Shoichet BK. Testing a flexible-receptor docking algorithm in a model binding site. *J Mol Biol* **337** (5), 1161-82 (2004).



9. Wei BQ, Baase WA, Weaver LH, Matthews BW, Shoichet BK. A model binding site for testing scoring functions in molecular docking. *J Mol Biol* **322**, 339-55 (2002).



10. Powers RA, Shoichet BK. Structure-based approach for binding site identification on AmpC β -lactamase. *J Med Chem* **45**, 3222-34 (2002).

