# Journal of Biomolecular Screening

**Here Be Dragons: Docking and Screening in an Uncharted Region of Chemical Space**

Ruth Brenk, John J. Irwin and Brian K. Shoichet

The online version of this article can be found at:
http://jbx.sagepub.com/cgi/content/abstract/10/7/667

Additional services and information for *Journal of Biomolecular Screening* can be found at:

**Email Alerts:** http://jbx.sagepub.com/cgi/alerts

**Subscriptions:** http://jbx.sagepub.com/subscriptions

**Reprints:** http://www.sagepub.com/journalsReprints.nav

**Permissions:** http://www.sagepub.com/journalsPermissions.nav

**Citations** (this article cites 22 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://jbx.sagepub.com/cgi/content/abstract/10/7/667#BIBL

# Here Be Dragons: Docking and Screening in an Uncharted Region of Chemical Space

**RUTH BRENK, JOHN J. IRWIN, and BRIAN K. SHOICHET**

To compare virtual and high-throughput screening in an unbiased way, 50,000 compounds were docked into the 3-dimensional structure of dihydrofolate reductase prospectively, and the results were compared to a subsequent experimental screening of the same library. Undertaking these calculations demanded careful database curation and control calculations with annotated inhibitors. These ultimately led to a ranked list of more likely and less likely inhibitors and to the prediction that relatively few inhibitors would be found in the empirical screen. The latter prediction turned out to be correct, with arguably no validated inhibitors found experimentally. Subsequent retesting of high-scoring docked molecules may have found 2 true inhibitors, although this remains uncertain due to experimental ambiguities. The implications of this study for screening campaigns are considered. (*Journal of Biomolecular Screening* 2005:667-674)

**Key words:** high-throughput screening, HTS, virtual screening, molecular docking, database preparation

## INTRODUCTION

**D**OCKING SCREENS HAVE RARELY been compared directly to empirical, high-throughput screens (HTS) in the literature (but see references 1-4). This is especially true when the virtual screen is conducted prospectively, before the HTS results are known, and where both techniques use exactly the same molecules. This was the extraordinary situation afforded by the competition described in this issue. By comparing the docking predictions to the HTS results, we hoped to probe key questions, including: how well does docking enrich ligands over random screening, what ligands does it enrich, and what false negatives does it have? These questions implicitly evaluate docking based on HTS results, but we may also consider the reverse: how many false negatives are in the empirical screen, and how likely are we to find any hits among the screened database in the first place? In this event, we feel we can evaluate only the last of these.

We were asked to describe our docking campaign versus dihydrofolate reductase (DHFR) based on two features: a 2-fold enrichment of hits among the top 5% of the docking ranked compounds and our prediction that there might be few hits overall. A 2-fold enrichment meant that 10% of the 96 inhibitors discovered by HTS were predicted among the top-ranking 5% (about 5000 compounds) of the docked molecules. The value of this very modest enrichment is further lessened by skepticism about the 96 HTS hits themselves. Of these, 53 showed a dose response, but none were competitive, which would be expected of molecules docked against the substrate binding site of DHFR.[5] Experiments in our own lab suggested that at least some of these inhibitors were promiscuous, aggregation-based inhibitors (data not shown). Indeed, on careful retesting, the screening group concluded that they had essentially no true hits. Thus, we discount enrichment of hits and focus here on the second criterion, the prediction that there would be relatively few hits from the HTS. In particular, we describe the methods and interpretations that led to our evaluation of the database as a source of DHFR inhibitors, with special attention to preparation of the databases for docking. We also consider whether any general lessons can be drawn for evaluation or prioritization of a database for its likelihood to contain ligands.

## METHODS

### Database preparation

Two databases were used in this study: a training database and a test set database. The training database consisted of 50,000 compounds previously screened for DHFR inhibition,[6] augmented with annotated DHFR inhibitors from the MDL Drug Data Report (MDDR; MDL Inc., San Leandro, CA) and supplemented by several ligands from the literature, such as methotrexate and folate.

University of California, San Francisco, Department of Pharmaceutical Chemistry, 1700 4th Street, San Francisco, CA 94143-2550.

Combined with the 12 inhibitors found in the previous study,[6] this made for 204 annotated DHFR inhibitors among the about 50,000 decoy molecules of the training database. The test set database was composed of 50,000 ChemBridge molecules that were to be screened both by HTS and by docking; the properties of these molecules as DHFR inhibitors were unknown.

To convert the databases into a format suitable for docking, we followed a prototype protocol for preparing the ZINC database.[7] Starting with SDF-format files, the databases were filtered using OpenEye's filter program (OpenEye Software, Santa Fe, NM) to remove counter ions. Energetically accessible tautomers were generated using the program Agent 1.2.[8] A properly protonated form of each molecule was generated with the filter program using default protonation and charging rules. Additional forms were produced by creating subsets of molecules containing key functional groups (e.g., pteridine and imidazole analogs) and using the filter program to apply our own charging rules. Partial atomic charges, desolvation energies, and van der Waals parameters were calculated as described.[9-11] Conformations were sampled using Omega (OpenEye) and stored in a hierarchical flexibase.[11,12] In this flexibase approach, the largest rigid fragment of any given database molecule is used as an anchor to define the possible orientations of all conformations of that molecule in the protein site. To further increase sampling, we extended the flexibase to represent multiple rigid fragments for each database molecule, each of which is sequentially used as the guide fragment for sampling orientations. This extension is particularly useful for ligands, where the largest rigid fragment does not form key interactions with the protein, as is the case with several DHFR ligands.

### Template selection

DHFR undergoes conformational changes along the reaction coordinate.[13] Because we sought competitive inhibitors, we considered only structures in which a ligand and the cofactor were present (PDB codes 1RA2, 1RA3, 1RC4, 1RX2, 1RX4). Two of these, 1RA2 and 1RX2, are essentially the same, and so we used 1RA2, which was determined to higher resolution (1.6 Å). In the remaining structures, the nicotinamide part of the cofactor is displaced from its binding site. The structure 1RA3 otherwise resembles 1RA2, whereas in 1RX2 and 1RC4, the M20 loop (the flexible active site loop) blocks the nicotinamide pocket. Because all *Escherichia coli* DHFR complexes are determined with similar ligands, we searched Relibase[14] for DHFR complexes from other species to guide our choice of representative structures. Only in the *Candida albicans* DHFR structures were ligands found that occupy part of the space made available by displacing the cofactor.[15] These ligands resemble some inhibitors found in the training database.[6] Superposing these complexes on the *E. coli* complexes with the displaced cofactor showed that only 1RA3 can accommodate these large ligands. Therefore, we decided to use this structure, in addition to 1RA2, as a template for docking. Analysis of the *E. coli* DHFR complexes revealed that all ligands form hydrogen bonds to water molecule 302 (Figs. 1a, 1b). Therefore, we kept this water molecule as a rigid part of the receptors.

### Receptor preparation

Polar hydrogens were added to the DHFR structures using MOLOC, and their positions were minimized using the MAB force field.[16] AMBER charges were subsequently assigned to the protein atoms. Inspection of DHFR complexes suggested that the carboxylate group of Asp27, water molecule 302, and the carbonyl groups of Ile5 and Ile94 formed key interactions with the ligands (Figs. 1a, 1b). To favor these interactions during docking, we redistributed the AMBER-based partial charges of appropriate atoms in these residues as described.[17] These adjustments increase the effective dipoles of these groups without altering the net charge on any residue. In particular, the partial charges of the carbonyl oxygens of the affected isoleucines and of the oxygen of Wat302 were decreased (made more negative) by 0.4 electrons and those of the carboxyl oxygens of Asp27 by 0.2 electrons each. Correspondingly, the charge on the amide hydrogen of the isoleucines was increased by 0.4 electrons, and the charge on the hydrogens of Wat302 were each increased by 0.2 electrons. All subsequent calculations followed our standard techniques.[17-20] A grid-based excluded volume map was calculated using DISTMAP.[21] CHEMGRID was used to calculate an AMBER-based van der Waals potential for the receptor.[9] DelPhi[22] was used to calculate an electrostatic potential for the receptor, using an internal dielectric of 2 and an external dielectric of 78. To approximate the effect of ligand binding, the effective dielectric of the binding site was reduced by identifying the volume occupied by ligand atoms as a low dielectric region.[23] Ligand atoms from the crystal structures, augmented with SPHGEN spheres,[23] were used as receptor matching positions to dock molecules in the site. Spheres within hydrogen-bonding distance to the carboxylate group of Asp27 or carbonyl oxygens of Ile5 and Ile95 were labeled as hydrogen-bond donors; all other spheres were labeled as neutral.[24] A solvent occlusion map was calculated to account for partial ligand desolvation (B. Shoichet, unpublished results).

### Docking

DOCK 3.5.54[10,11] was used to flexibly dock small molecules into the active site of each receptor model. To sample ligand orientations, ligand and receptor bins were set to 0.5 Å, and overlap bins were set to 0.4 Å; the distance tolerance for matching ligand atoms to receptor matching sites ranged from 1.1 to 1.2 Å. Sphere and atom labeling[24] was used to constrain atoms used for orientation. Each docking pose was evaluated for steric fit. Compounds passing this filter were scored for electrostatic and van der Waals complementarity and corrected for desolvation (B. Shoichet, unpublished results).[10] For each compound, only the best-scoring database representation (tautomer, protonation state, multiple ring alignment) was retained in the final docking hit list.
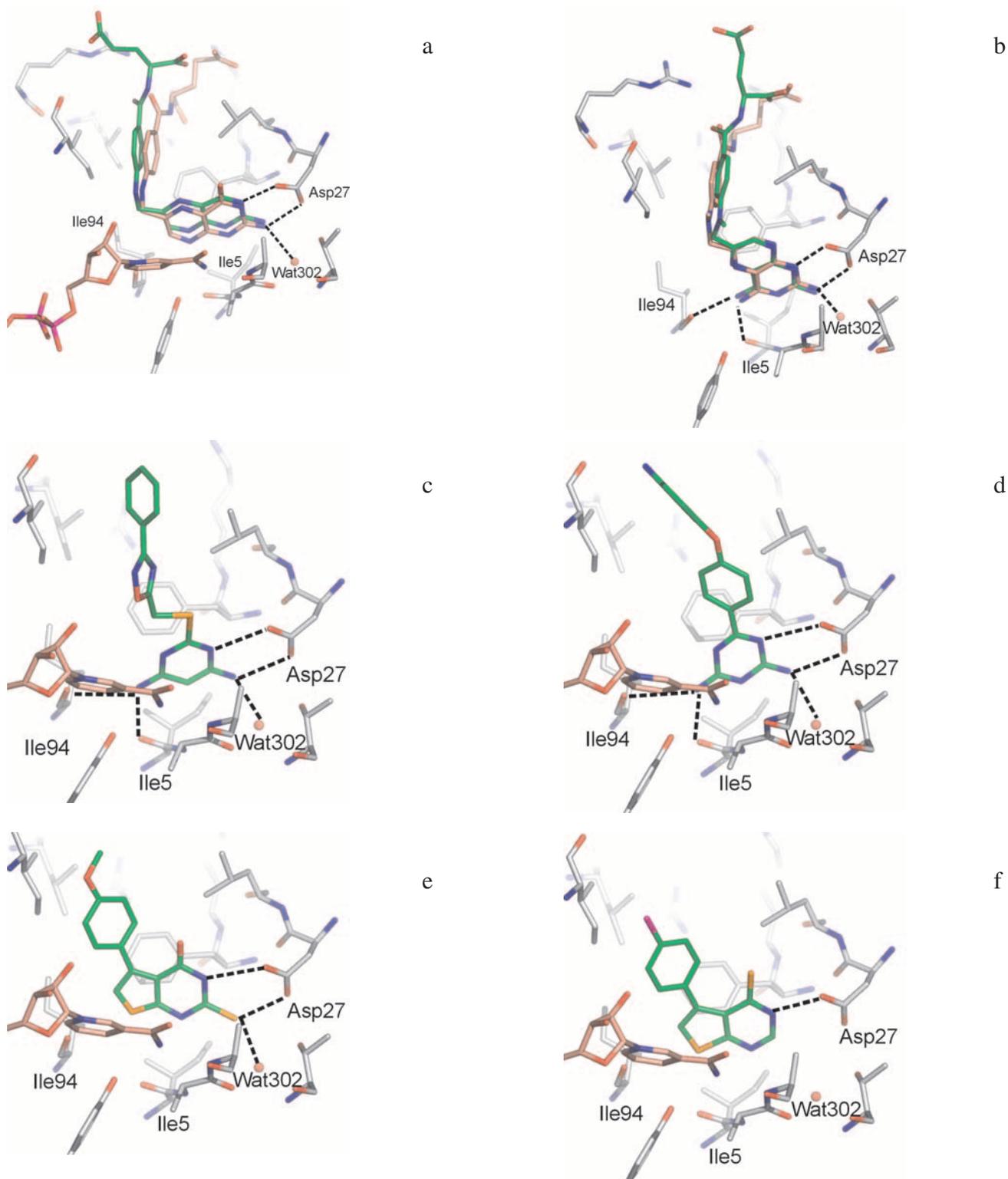
**FIG. 1.** Docked geometries of known and predicted dihyrdofolate reductase (DHFR) ligands. Unless otherwise noted, carbon atoms are gray, oxygen atoms are red, nitrogen atoms are blue, and sulfur atoms are orange. (**a, b**) Superposition of the docked pose (green carbon atoms) with the crystallographically determined binding mode (pink carbon atoms) for folic acid (**a**) and methotrexate (**b**). (**c-f**) Docked binding modes (green carbon atoms) of some representative compounds from Table 1. Compounds 1b (**c**) and 2a (**d**) form the same hydrogen-bond network with DHFR as found in the methotrexate complex if they are protonated at N1. (**e**) The thiol tautomer of 3a forms the same hydrogen-bond network as folic acid. (**f**) 4a forms only 1 hydrogen bond to Asp27. Figures were produced using PyMOL (www.pymol.org).

*Inhibition assay*

The HTS is described elsewhere in this issue. Retesting of selected compounds was done by the same lab as had performed the HTS assays.[5]

**RESULTS**

*Database preparation*

A high-quality database for docking is characterized by the presence of relevant tautomers and correctly protonated compounds. This is complicated by the fact that ligand binding can be associated with $pK_a$ shifts and changes in tautomer equilibria. Indeed, such $pK_a$ shifts are important for several well-known DHFR inhibitors, including methotrexate, on binding to DHFR.[25] Because of the absence of reliable methods to apply to thousands of molecules rapidly, we used fast but arguably crude rule-based methods for enumerating likely protonation states and tautomers.[7] These rules resulted in ionizable groups, such as pteridines and quinazolines, being represented in both their charged and neutral forms. After enumerating likely tautomers and protonation states and assigning multiple rigid fragments, the dockable database contained an average of 2.7 entries per chemical structure.

*Control calculations*

Our 1st goal was to ensure that we could produce ligand poses close to those observed in DHFR crystal structures. Therefore, we redocked the crystallographic ligands back into their own receptor (Figs. 1a, 1b). The buried parts of the ligands (the two ring systems plus the connecting linker) match well with the crystallographically determined binding mode, with a root mean square distance of 0.8 Å for the substructure of methotrexate and 0.9 Å for the substructure of folic acid.

The next goal was to ensure that we could enrich known ligands from among the decoys in the training database. All compounds were docked against both DHFR receptor conformations (1RA2 and 1RA3). Each structure performed well: 57% and 58% of known inhibitors were found in the top 1% of the docking ranked database, respectively (Fig. 2). Slightly better enrichment was obtained if both individual runs were merged (60% of known ligands in the top 1% of the database). This combined hit list contained the best score for each compound from both runs. If only the training database without the additional MDDR ligands is considered, 8 of the 12 ligands discovered in the HTS of this database[6] score in the top 2.5% of the database; 4 of these are even among the top 25 compounds.

*Docking the test database*

With these control calculations in hand, we turned to docking the test database into *E. coli* DHFR, looking to make prospective predictions. The same methods were used, resulting in a DOCK-
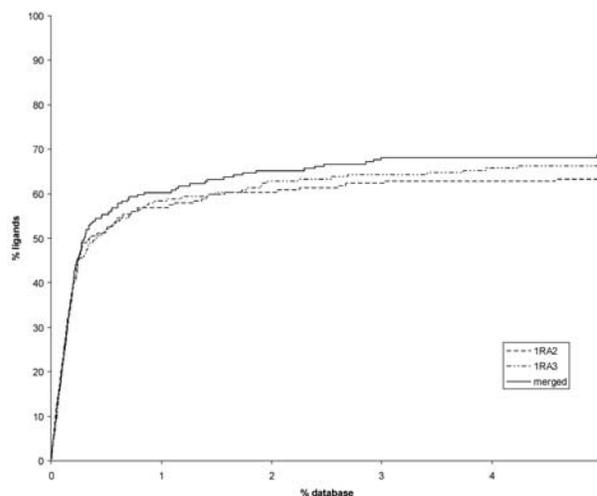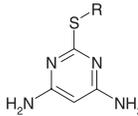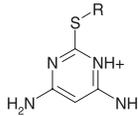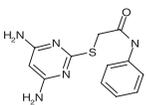


**FIG. 2.** Enrichment of 204 known binders by docking the training database against 1RA2 (dashed line), 1RA3 (dashed-dotted line), or both sites and storing only the better scoring compound in the final hit list (solid line). The cumulative percentage of known binders versus the percentage of the rank-ordered database is plotted. For instance, the 1% point on the *x*-axis means the top 1% of the docking ranked database.

ranked list of 46,720 compounds (the remaining compounds were not ranked owing to steric clashes). Subsequent visual inspection of the top 5% of the docking ranked compounds allowed us to deprioritize unlikely complexes and to rescue interesting compounds that did not score at the very top of the hit list. This is common practice in docking screens and can occasionally compensate for problems in docking scoring functions. This inspection revealed that some putative hits formed unfavorable interactions with Asp27, which we identified as a key DHFR recognition element. For instance, several top-scoring hits place benzene moieties next to this aspartate without any polar groups to interact with it. Without removing these hits, we disfavored such visually uncompelling molecules by moving about 50 of what we considered more interestin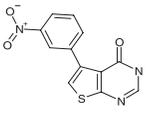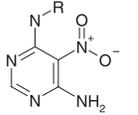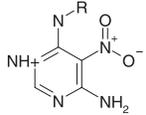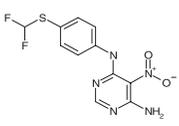g molecules ahead of them in the overall hit list. Some examples of our favored inhibitors are shown in Table 1. After these hand-selected molecules, we simply added the remaining molecules in the database, ranked exclusively based on score.

Notwithstanding considerable effort to rescue misranked compounds, in the end we were left with a list that seemed inferior to the high-scoring ligands in the training database. In particular, problems of correct assignment of tautomerization and protonation equilibria made us skeptical that there were good competitive inhibitors among our top hits and arguably in the test database, at all. For instance, the sulfanyl-pyrimidine-4,6-diamines (compounds 1a and 1b in Table 1) and the triazenes (compounds 2a and 2b in Table 1) can form the same hydrogen-bond pattern with Asp27 and Ile5 as observed in the DHFR-methotrexate complex, assuming they are protonated (Figs. 1c, 1d). However, their $pK_a$ values are about 4.9,[26,27] about half a unit lower than that of methotrexate[28]

**Table 1.** Representative Compounds of the Test Database That Had a Good Docking Pose and Score

| # | Class | Required State for Binding | Members in Class | Representative Structure | Docking Rank[a] | Concerns |
|---|---|---|---|---|---|---|
| 1a | | | |  | 10 | low pK$_a$ (4.9 modeled |
| 1b |  | | 24 |  | 4 | |
| 2a | | | |  | 1 | low pK$_a$ (4.9) |
| 2b |  | | 2 |  | 2 | |
| 3a | | | |  | 905 | Energetically less favorable tautomer required for binding |
| 3b |  | | 3 |  | 1027 | |
| 4a | | | |  | 54 | Poor interactions with Asp27 |
| 4b |  | | 7 |  | 71 | |
| 5 |  | | 1 |  | 80 | Poor interactions with Asp27, nitro group lowers the pK$_a$ |
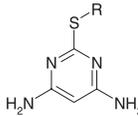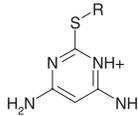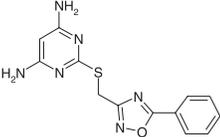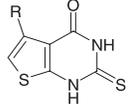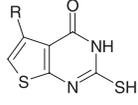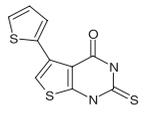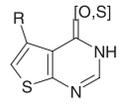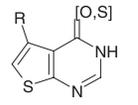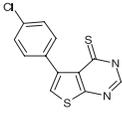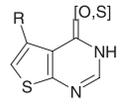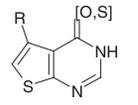
a. Raw rank from docking, before expert reranking of 50 compounds.

and far from neutral pH. For the protonated form of the molecules to bind to the enzyme, a $pK_a$ shift is required. To compensate for the energy necessary to shift this ionization equilibrium, the protonated molecules would have to have a comparatively low (favorable) $K_i$ value to inhibit at the concentration used in the HTS. Similarly, the thiouracil analogs (compounds 3a and 3b in Table 1) resemble folic acid and fit the DHFR pocket well. Depending on the tautomeric state they adopt in the binding pocket, they can form a double hydrogen bond to Asp27, as seen in the DHFR–folic acid complex (Fig. 1e). However, the required tautomeric form is, at least for thiouracil itself, not the most stable one.[29,30] Thus, whereas they were docked as this tautomer, we were dubious that this structure was really energetically accessible. Therefore, we doubted that these compounds would inhibit DHFR. Another high-scoring class was that of the pyrimidinones and pyrimidinethiones (compounds 4a and 4b in Table 1), which also resemble folic acid but lack an exocyclic hydrogen-bond donor to interact with Asp27 (Fig. 1f). Therefore, despite their relatively good shape and complementarity to the binding pocket, they seemed unlikely to inhibit DHFR. This is consistent with the observation that none of the 5 related compounds in the training database inhibited DHFR in the initial HTS.[6] Finally, the aminopyrimidine (compound 5 in Table 1) resembles some known DHFR inhibitors except that this molecule, like the pyrimidinones and -thiones (compounds 4a and 4b in Table 1), is missing an exocyclic hydrogen-bond donor to interact with Asp27. In addition, the electron withdrawing effect of the nitro group lowers the $pK_a$ of the titratable ring nitrogens, making it unlikely to be protonated in the assay. Without being protonated, the compound could not form hydrogen bonds to Asp27, suggesting that this molecule is most likely not inhibiting DHFR.

These doubts led us to predict, in our final report, that there may be few experimental hits to emerge from the HTS assays on these 50,000 compounds against this target. As the penultimate paragraph of our final report, we wrote, "We note that the scores of the ligands in the training set were often higher than those of even the top scoring ligands from the test set. *There may be relatively few binders in the test set*" (emphasis added).

### Experimental screening

The HTS of the test database is described elsewhere in this issue. In brief, 96 hits where found. Of these, 53 showed a dose response, but none were competitive as would be expected of molecules docked against the substrate binding site of DHFR.[5] Ultimately, it was concluded that none were reliable inhibitors of the enzyme.[5]

### Supplementary predictions

Whereas the penultimate paragraph of our report expressed skepticism as to the likelihood of finding good hits among the test database, the last paragraph proposed molecules for retesting that we felt might have been false negatives from the training database tested in the original, preliminary study (Table 2).[6] Subsequently,

based on the observation that none of our top-scoring molecules from the test set actually inhibited, we supplemented the 5 molecules from the training database with 5 test set compounds that we found particularly interesting. Ultimately, we suggested 10 molecules, 5 from the training database and 5 from the test database, for retesting at higher concentrations. These were compounds 1a through 2b and compound 4a (Table 1) and compounds 6 through 10 (Table 2). Two of these, compounds 2b and 9, did show modest inhibition: 36% and 31% at 46 and 45 µM, respectively. Experimental limitations made it impossible to determine at this point whether these compounds were competitive, so the interpretation of these two modest inhibitors as "successes" for docking must remain tentative.
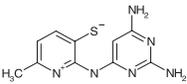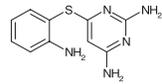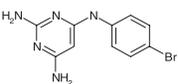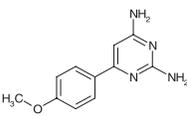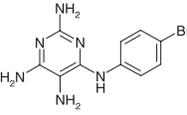
### DISCUSSION

We wish to emphasize three results from this study. First, a combination of careful control calculations and critical inspection of our hit lists suggested that there might be few inhibitors found from the experimental screen. Thus, in our final report, we noted that "there may be relatively few binders in the test set." This turned out to be the case, as no competitive inhibitors were found whatsoever, and the characteristics of the other inhibitors that were found are open to question. Second, recognition of this deficit depended on careful attention to the tedium of database preparation; several of the techniques deployed may be useful to other investigators. Third, and most tentatively, docking may have identified false negatives from the HTS results, and correspondingly, several HTS hits may be genuine, identifying docking false negatives. This may reflect the previously discussed complementarities between the two screening approaches.[2,4,31]

We performed two kinds of control experiments: We tested the ability of DOCK to predict binding modes of known ligands close to the crystallographically observed ones and its ability to enrich ligands among a database of decoys. In both tests, DOCK performed well (Figs. 1 and 2). This made us confident that our protocol was capable of retrieving new inhibitors. Docking the test database suggested what were, on first glance, interesting compounds with good complementarity to the binding pocket (Table 1). However, closer inspection of these hits revealed issues with all of them. Low $pK_a$ values, energetically unfavorable tautomeric states, or suboptimal interactions to Asp27 curbed our enthusiasm for these compounds as inhibitors of *E. coli* DHFR. Consistent with our concerns, no competitive inhibitors were found in the subsequent experimental screen.

Of perhaps underestimated importance for docking is ligand database preparation, especially the prediction of $pK_a$ values, stable tautomers, and low-energy conformations. Overlooking accessible protonation states and tautomers leads to false-negative predictions, whereas generating unrealistic protonation states or tautomers can lead to false-positive predictions. Until more reliable, fast methods applicable to many thousands of molecules become available, we must use loose, rule-based methods for enu-

**Table 2.** Predicted Ranks for Compounds in the Training
Database That Might Have Been Missed
in the Initial High-Throughput Screening

| # | Structure | Rank | Residual Activity (%) |
|---|-----------|------|------------------------|
| 6 | | 4 | 100 |
| 7 | | 7 | 100 |
| 8 | | 9 | 100 |
| 9 | | 15 | $62 \pm 4$ @ 75 µM, $69 \pm 9$ @ 45 µM |
| 10 | | 47 | 100 |

merating tautomers and protonation states. By carefully inspecting the hit list, unrealistic high-scoring molecules can be removed. Rescuing molecules that are buried in the hit list because the state required for binding was not present in the database is much more difficult. Our experience with database preparation for several docking projects, including this one, has led to rules that have been applied to generate the ZINC database, now available to the community (http://blaster.docking.org/zinc).[7]

The most tentative point to emerge from this study is the identification of HTS false negatives by docking. We found some promising compounds in the training and test databases that did not show inhibition in the initial HTS. When compounds 1a to 2b and 4a (Table 1) and 6 to 9 (Table 2) were retested at higher concentration, compounds 9 and 2b showed weak inhibition (so weak, in fact, that they would not have been detected at the original HTS concentration; therefore, *false negative* is perhaps a misnomer). Because of assay constraints, it cannot be confirmed that these are

competitive inhibitors. If in fact these molecules do bind in the active site, it would be consistent with the view that molecular docking is a complementary method to HTS and that both techniques may find molecules missed by the other.[2,4,31]

The prediction that there might be few true inhibitors from the empirical screen is the most solid of our results but also the most negative. Even such a negative prediction could be useful, for instance, to focus libraries for screening. To make this application general would demand algorithmic effort—for this project, we relied heavily on expert editing of the docking results. From the standpoint of future competitions, it does seem that some efforts along these lines would be useful, if only to ensure that the test database contains some true, validated ligands. Also, it would be helpful to continue to allow for follow-up studies on molecules that scored well in the computational screen but that did not appear to bind in the HTS assays, as this may provide a rare view into the problem of false negatives, assisting both communities.

## ACKNOWLEDGMENTS

## REFERENCES

1. Jenkins JL, Kao RY, Shapiro R: Virtual screening to enrich hit lists from high-throughput screening: a case study on small-molecule inhibitors of angiogenin. *Proteins* 2003;50:81-93.
2. Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, et al: Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J Med Chem* 2002;45:2213-2221.
3. Evensen E, Eksterowicz JE, Stanton RV, Oshiro C, Grootenhuis PD, Bradley EK: Comparing performance of computational tools for combinatorial library design. *J Med Chem* 2003;46:5125-5128.
4. Paiva AM, Vanderwall DE, Blanchard JS, Kozarich JW, Williamson JM, Kelly TM: Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis*. *Biochim Biophys Acta* 2001;1545:67-77.
5. Elowe NH, Blanchard JE, Cechetto JD, Brown ED: Experimental screening of dihydrofolate reductase yields a "test set" of 50,000 small molecules for a computational data-mining and docking competition. *J Biomol Screen* 2005;10:653-657.
6. Zolli-Juran M, Cechetto JD, Hartlen R, Daigle DM, Brown ED: High throughput screening identifies novel inhibitors of *Escherichia coli* dihydrofolate reductase that are competitive with dihydrofolate. *Bioorg Med Chem Lett* 2003;13:2493-2496.

7.  Irwin JJ, Shoichet BK: ZINC—a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 2005;45:177-182.

8.  Pospisil P, Ballmer P, Scapozza L, Folkers G: Tautomerism in computer-aided drug design. *J Recept Signal Transduct Res* 2003;23:361-371.

9.  Meng EC, Shoichet B, Kuntz ID: Automated docking with grid-based energy evaluation. *J Comp Chem* 1992;13:505-524.

10. Wei BQ, Baase WA, Weaver LH, Matthews BW, Shoichet BK: A model binding site for testing scoring functions in molecular docking. *J Mol Biol* 2002;322:339-355.

11. Lorber DM, Shoichet BK: Flexible ligand docking using conformational ensembles. *Protein Sci* 1998;7:938-950.

12. Lorber DM, Udo MK, Shoichet BK: Protein-protein docking with multiple residue conformations and residue substitutions. *Protein Sci* 2002;11:1393-1408.

13. Sawaya MR, Kraut J: Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence. *Biochemistry* 1997;36:586-603.

14. Hendlich M, Bergner A, Gunther J, Klebe G: Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions. *J Mol Biol* 2003;326:607-620.

15. Whitlow M, Howard AJ, Stewart D, Hardman KD, Chan JH, Baccanari DP, et al: X-ray crystal structures of *Candida albicans* dihydrofolate reductase: high resolution ternary complexes in which the dihydronicotinamide moiety of NADPH is displaced by an inhibitor. *J Med Chem* 2001;44:2928-2932.

16. Gerber P, Müller K: MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J Comput Aided Mol Des* 1995;9:251-268.

17. Powers RA, Morandi F, Shoichet BK: Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. *Structure (Camb)* 2002;10:1013-1023.

18. Ferrari AM, Wei BQ, Costantino L, Shoichet BK: Soft docking and multiple receptor conformations in virtual screening. *J Med Chem* 2004;47:5076-5084.

19. McGovern SL, Shoichet BK: Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J Med Chem* 2003;46:2895-2907.

20. Wei BQ, Weaver LH, Ferrari AM, Matthews BW, Shoichet BK: Testing a flexible-receptor docking algorithm in a model binding site. *J Mol Biol* 2004;337:1161-1182.

21. Shoichet B, Bodian DL, Kuntz ID: Molecular docking using shape descriptors. *J Comp Chem* 1992;13:380-397.

22. Gilson MK, Honig BH: Calculation of electrostatic potentials in an enzyme active site. *Nature* 1987;330:84-86.

23. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE: A geometric approach to macromolecule-ligand interactions. *J Mol Biol* 1982;161:269-288.

24. Shoichet BK, Kuntz ID: Matching chemistry and shape in molecular docking. *Protein Eng* 1993;6:723-732.

25. Matthews DA, Alden RA, Bolin JT, Freer ST, Hamlin R, Xuong N, et al: Dihydrofolate reductase: x-ray structure of the binary complex with methotrexate. *Science* 1977;197:452-455.

26. Gutierrez MD, Lopez R, Romero MA, Salas JM: Spectroscopic studies of some Pd(Ii), Pt(Ii), Ag(I), and Au(Iii) complexes of 4,6-diamino-2-thiopyrimidine and 4,6-diamino-2-methylthiopyrimidine—structure and binding-site determination. *Can J Chem* 1988;66:249-255.

27. Tashiro T: The ionization constants of 2-substituted 4,6-diamino-s-triazines: the applicability to the Hammett and Taft equations. *J Heterocyclic Chem* 2002;39:615-622.

28. Albert A, Brown DJ, Cheeseman G: Pteridine studies: part III. The solubility and the stability to hydrolysis of pteridines. *J Chem Soc* 1952:4219-4232.

29. Rostkowska H, Szczepaniak K, Nowak MJ, Leszczynski J, Kubulat K, Person WB: Studies of thiouracils: 2. Tautomerism and infrared-spectra of thiouracils—matrix-isolation and abinitio studies. *J Am Chem Soc* 1990;112:2147-2160.

30. Yekeler H: Ab initio study on tautomerism of 2-thiouracil in the gas phase and in solution. *J Comput Aid Mol Des* 2000;14:243-250.

31. Klebe G: Virtual screening: an alternative or complement to high throughput screening? Preface. *Perspect Drug Discov Des* 2000;20:vii-xi.

Address reprint requests to:
*Brian K. Shoichet*
*University of California, San Francisco*
*Department of Pharmaceutical Chemistry*
*1700 4th Street*
*San Francisco, CA 94143-2550*

*E-mail:* shoichet@cgl.ucsf.edu