ACS Medicinal Chemistry Letters
Viewpoint
pubs.acs.org/acsmedchemlett

Nexus Between Protein–Ligand Affinity Rank-Ordering, Biophysical Approaches, and Drug Discovery

M. Ashley Spies*

Department of Biochemistry, Carver College of Medicine, and Division of Medicinal and Natural Products Chemistry, College of Pharmacy, University of Iowa, 115 South Grand Avenue, Iowa City, Iowa 52242, United States

ABSTRACT: The confluence of computational and biophysical methods to accurately rank-order the binding affinities of small molecules and determine structures of macromolecular complexes is a potentially transformative advance in the work flow of drug discovery. This viewpoint explores the impact that advanced computational methods may have on the efficacy of small molecule drug discovery and optimization, particularly with respect to emerging fragment-based methods.

The ability to affinity rank-order small molecules against a drug target is, arguably, the ne plus ultra of structure-based drug discovery (SBDD). However, there are numerous caveats that impede the facile application of the current existing toolbox of approaches toward achieving this goal. It has been well established that docking scoring functions, while often performing well at determining binding poses of ligand to protein targets, are not adequate for determining relative binding affinities of a library of compounds.[1] This is not surprising, given the many assumptions that go into a classical in silico screening campaign. Usually the goal of such campaigns is the parsing of binders from nonbinders. However, affinity rank-ordering is the realm of atomistic, physics-based computational approaches that often employ molecular dynamics (MD) sampling methods, which are relatively intensive and thus most amenable to studies utilizing 10s or 100s of compounds. Furthermore, the efficacy of such approaches often synergizes with the quality of the biophysical data on the macromolecular target. The purpose of this viewpoint is not to review or compare the many methods that are able to reasonably achieve protein–ligand affinity rank-ordering, but to bring attention to the fact that such computational and biophysical methods are becoming a vital, yet often overlooked, component in modern methods for inhibitory scaffold discovery and optimization.

Before accepting it on faith that this would truly be a worthy achievement for medicinal chemistry, let us examine what a satisfactory solution to this problem would look like, and why it would have a profound effect on the modern drug discovery work flow. Shirts et al.[2] have astutely quantified the premium to the medicinal chemist due to accurate affinity predictions of small molecules binding to protein-targets, which was based on the largely normal distribution observed for a large (>80 000) sampling of the additions of chemical substituents to lead compounds, across 30 different protein targets at Abbott laboratories.[3] The implication of this distribution is that only 8.5% of the transformations in the various optimization campaigns lead to affinity changes greater than 1.4 kcal/mol (~1 pK_i unit) and 1% at values of 2.8 kcal/mol. Shirts et al. then demonstrated that computational affinity rank-ordering with standard error values of 1.0 and 0.5 kcal/mol would provide 36% and 50% chances, respectively, of obtaining a 1.4 kcal/mol enhancement in affinity.[2] Achieving affinity rank-ordering at the <1 kcal/mol accuracy level requires an array of computational methods that are not currently available as turnkey implementations but nevertheless are routinely accomplished. In the final analysis, the real benefit of accurate affinity rank-ordering will be manifested in the resource and time commitment that it takes to surmount the <1 kcal/mol threshold across a range of drug targets.

However, there's more to this story than simply the efficiency with which one obtains enhancements in binding free energy, this is particularly true when considering the increasing success and advantages of employing fragment-based drug discovery (FBDD) campaigns, as opposed to the classical optimizations that are at the heart of the statistical analyses outlined above. Fragment-based approaches to drug discovery have been successful in obtaining lead compounds, often against highly intractable targets, which are optimized to have excellent drug-like properties.[4,5] FBDD is a generic term for a family of strategies that seek to optimize initial hits from low molecular weight compounds, which initially have weak potencies but high ligand efficiencies (LE, binding energy/non-hydrogen atom). FBDD is a more efficient search of chemical space than high-throughput screening (HTS; which employs medium-sized compounds of greater complexity) because larger portions of potential chemical space may be readily searched.[5] The caveat here is, of course, that the very weak binding must be detectable and amenable to characterization.

Structural biology, biophysical methods, and computational chemistry often play central roles in FBDD. FBDD is quite different from classical HTS (and in silico screening of medium sized molecules) in a number of ways, perhaps most importantly in its reliance on iterative cycles of rationally incrementing the potency of the initial (weak) fragment hit while maintaining or improving a host of drug-like descriptors (e.g., high LE, low clogP, high LiPE, etc.). Although the successes of FBDD are too numerous to list, the field reached a milestone with the FDA approval of Vemurafenib (Zelboraf), which was developed using these approaches.

Published: September 19, 2013

 © 2013 American Chemical Society

Although in silico docking is a relatively standard method used in many FBDD work flows, its utility is limited to fragment placement since affinity rank-ordering is not within the realm of docking scoring functions. However, the use of physics-based methods for calculation of the affinity of fragments and their derivatives during optimization has received far less attention than for simple docking. Indeed, this is the area where in silico methods may have the most impact on drug design and discovery. A highly generalized and hypothetical work flow for rational structure-based discovery and lead optimization (which may or may not employ FBDD) is shown in Figure 1, which envisions the nexus between computational affinity rank-ordering (blue section), compound testing, and optimization (green and yellow sections, respectively), which may include a number of approaches such as linking fragments, growing derivative libraries, or bioisostere hopping. The value-added benefit of employing MD-based sampling to achieve affinity rank ordering in the work flow shown in Figure 1 is manifold. First, the same statistical analysis presented above for enriching binders, given a reasonable accuracy (~0.5−1 kcal/mol), would be especially valuable in the optimization of a fragment hit since different fragment hits will yield diverse derivative libraries with distinct distributions of affinities. A key caveat of the linking and growing strategies used in FBDD is knowing how the protein changes shape in the complex. Indeed, the use of MD-based sampling methods to determine accurate affinity rank-ordering has the capacity to not only determine what fragment to use for optimizing LE, but information on how the target protein structurally and dynamically responds to a particular growing or linking strategy. A leading FBDD group recently indicated that fragment binding events often trigger conformational changes in the target protein and even the formation of new pockets.[4]

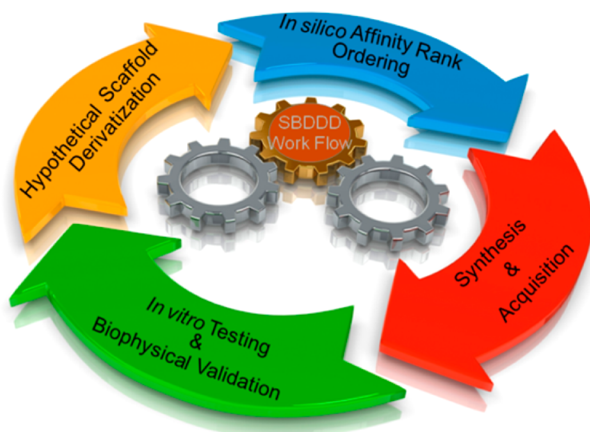

Figure 1. Flowchart for a biophysically driven drug discovery cycle.

A prime example of the need for conformational sampling to achieve relative binding affinities with a set of noncongeneric small molecule inhibitors is the case of glutamate racemase, an antibacterial target, where high-resolution structural data is necessary but not sufficient for accurate affinity prediction.[6,7] The use of advanced MD approaches for sampling large structural changes in the ensemble (i.e., beyond local thermal fluctuations) was instrumental in achieving useful affinity rank-ordering in the glutamate racemase system,[6] which was also employed in a successful optimization campaign, using a similar strategy as that outlined in Figure 1.[7] The key to effectively using the extant structural information on glutamate racemase is the recognition that major structural alterations exist and happen to be critical in the enzyme's natural catalytic cycle. Furthermore, one of the most complicated features of such a system is that subtle differences in the structures of the small molecule lead compounds either capture or induce significant changes in the protein target, in an idiosyncratic fashion. Nevertheless, these problems are soluble with either exhaustive conformational sampling, or as in the case of glutamate racemase, a solution may be found from employing the right combination of MD simulations and docking.[6]

Another major issue to emerge in surmounting the challenges presented by flexible systems such as glutamate racemase is dealing with nonobvious changes in solvation that arise from large structural alterations in the protein target ensemble. These may be reflected in the changes in the protein solvation energy itself or even more problematic, ligand-dependent changes in the local water structure in the small molecule's binding pocket. There is an increasingly voluminous body of experimental and computational research that supports a highly idiosyncratic ligand-dependent change in water structure, especially in deep enzyme active sites. Nonobvious changes in the water structure and the implications on small molecule affinity rank-ordering is an area where computational approaches offer tangible benefits. For example, the seminal MD study by Helms and Wade established that a crystallographic water in a complex of cytochrome P450cam with an inhibitor (2-phenyl-imidazole), which mediates hydrogen-bonding between protein and ligand, provides -11.6 ± 6.6 kJ/mol of stabilization energy.[8] Alternatively, the complex of the natural substrate, camphor, whose crystal structure does not show a water in this cavity, is solvated with a penalty of $+15.8 \pm 5$ kJ/mol.[8] Recent experimental studies, employing high-resolution X-ray crystallography, on a series of complexes of tRNA-guanine transglycosylase with *lin*-benzoguanines and *lin*-benzohypoxanthines revealed the presence of a water cluster in the active site, whose specific structure was strongly tied to the observed inhibition constants for this series.[9] A similar idiosyncratic importation of water clusters has now been documented for glutamate racemase, where experimental and computational work indicated a wide range of ligand-specific water-mediated contacts in a buried polar active site.[10]

MD-based sampling and experimental biophysical methods are being employed in increasingly synergistic ways to benefit SBDD, but the full potential will be only be realized when and if these methods are more routinely integrated into drug discovery work flows. Accurate and meaningful protein−ligand rank-ordering has undeniable benefits to the medicinal chemist, as outlined above, and can only be achieved across a range of target classes using methods that are significantly more intensive than standard scoring functions but well within reach of many research groups today.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: michael-spies@uiowa.edu.

Notes

Views expressed in this editorial are those of the author and not necessarily the views of the ACS.

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The author thanks Dr. Katie Whalen for assistance in the composition of this manuscript.

ABBREVIATIONS

MD, molecular dynamics; FBDD, fragment-based drug discovery; HTS, high throughput screening; LE, ligand efficiency; LipE, lipophilic efficiency; SBDD, structure-based drug design

REFERENCES

(1) Head, M. S. In *Drug Design: Structure- and Ligand-Based Approaches*; Merz, K. M., Jr., Ringe, D., Reynolds, C. H., Eds.; Cambridge University Press: New York, 2010; p 98.

(2) Shirts, M. R.; Mobley, D. L.; Brown, S. P. In *Drug Design: Structure- and Ligand-Based Approaches*; Merz, K. M., Ringe, D., Reynolds, C. H., Eds.; Cambridge University Press: New York, 2010; p 61.

(3) Hajduk, P. J.; Sauer, D. R. Statistical analysis of the effects of common chemical substituents on ligand potency. *J. Med. Chem.* **2008**, *51*, 553−564.

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

(5) Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **2012**, *51*, 4990−5003.

(6) Whalen, K. L.; Chang, K. M.; Spies, M. A. Hybrid steered molecular dynamics-docking: An efficient solution to the problem of ranking inhibitor affinities against a flexible drug target. *Mol. Inf.* **2011**, *30*, 459−471.

(7) Whalen, K. L.; Chau, A. C.; Spies, M. A. In silico optimization of a fragment-based hit yields biologically active, high-efficiency inhibitors for glutamate racemase. *ChemMedChem* **2013**, DOI: 10.1002/cmdc.201300271.

(8) Helms, V.; Wade, R. C. Thermodynamics of water mediating protein-ligand interactions in cytochrome P450cam: a molecular dynamics study. *Biophys. J.* **1995**, *69*, 810−824.

(9) Barandun, L. J.; Immekus, F.; Kohler, P. C.; Tonazzi, S.; Wagner, B.; Wendelspiess, S.; Ritschel, T.; Heine, A.; Kansy, M.; Klebe, G.; Diederich, F. From *lin*-benzoguanines to *lin*-benzohypoxanthines as ligands for *Zymomonas mobilis* tRNA−guanine transglycosylase: Replacement of protein−ligand hydrogen bonding by importing water clusters. *Chemistry* **2012**, *18*, 9246−9257.

(10) Whalen, K. L.; Spies, M. A. Flooding enzymes: quantifying the contributions of interstitial water and cavity shape to ligand binding using extended linear response free energy calculations. *J. Chem. Inf. Model.* **2013**, DOI: 10.1021/ci400244x.