ELSEVIER

Ligand efficiency and fragment-based drug discovery

Scott D. Bembenek[1], Brett A. Tounge[2] and Charles H. Reynolds[2]

[1] Johnson & Johnson Pharmaceutical Research and Development, L.L.C., 3210 Merryfield Road, San Diego, CA 92121, United States
[2] Johnson & Johnson Pharmaceutical Research and Development, L.L.C., Welsh and McKean Roads, Spring House, PA 19477, United States

The use of fragment-based drug discovery (FBDD) has increased in recent years since it is more likely to produce a better optimized compound of lower molecular weight. Ligand efficiency (LE) has become important for assessing fragments, HTS hits, and resulting optimized ligands. LE is useful for comparing ligands of equal molecular weight, but is ineffective for comparisons of ligands of differing molecular weight. LE has a strong dependence on molecular size, which has led us to develop a size-independent efficiency score termed fit quality. Evaluating FBDD examples from the literature using LE and fit quality, we find that, in general, the LEs of starting fragments are greater than those of larger, more elaborated, structures. Fit quality scores, however, tend to improve upon optimization of the fragments.

Introduction

High-throughput screening is a common strategy in drug discovery for identifying compound 'hits' against a new target. Pharmaceutical companies have invested heavily in infrastructure to test ever-larger compound collections. Screening large numbers of compounds is a reasonable strategy when there is limited knowledge of how the desired target and its potential ligand will interact. Even the largest corporate compound collection is, however, intrinsically incomplete, containing on the order of 10^5–10^6 compounds, whereas the number of synthesizable compounds has been estimated to be upward of 10^{60} [1,2]. Clearly, the variety of chemical structures that are actually sampled in a routine screen of a corporate collection is tiny by comparison. Further, it has been shown [3] that the probability of finding a compound with the right combination of features for binding, dramatically decreases as the complexity of the molecule increases. Even if a potent compound is found, it is often suboptimal in its binding to the target along with other important drug-like properties, such as bioavailability, absorption, and solubility. Finally, the hit rate from standard HTS campaigns is often low [4] and many hits fail to progress in the optimization process [5,6].

Fragment-based drug discovery offers a potential solution to these limitations. Screening molecular fragments, rather than drug-sized molecules, allows one to explore a dramatically larger portion of chemical structure space with many fewer compounds. The hit rate associated with screening fragments has been observed to be higher than standard HTS screens [4]. Nonetheless, those properties that a fragment hit must exhibit in order to constitute a good starting point for drug discovery are, as yet, unclear. Some of the same factors that are important for HTS hits are likely to apply, but fragments will undoubtedly have some unique requirements. We will focus on ligand efficiency (LE) [7–10] and its application to fragment hits, and the evolution of efficiency as fragments are optimized and elaborated to provide potent ligands.

High-throughput screening

The rise of parallel synthesis in the late 1980s and early 1990s helped to drive significant growth in the number of compounds in corporate compound collections. Most major pharmaceutical companies have in-house compound collections numbering in hundreds of thousands, or even millions, of compounds. Screening this number of compounds was made possible by the development of ever higher-throughput screening (HTS) methods. The objective of screening these large collections was to provide medicinal chemists with starting points for synthetic optimization of drug candidates. The logic is simple: if you examine more compounds, your chances of finding a compound that will bind to a target of interest improves by sheer force of numbers. Large

Corresponding author: Bembenek, S.D. (sbembene@its.jnj.com)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2008.11.007

investments and logistic efforts have been made to see this vision forward. The pharmaceutical industry has and continues to rely heavily on HTS for lead identification. It is undeniable that leads have been found from HTS.

However, the hits that are identified are often far from optimal in the way they bind the protein target and in their physiochemical properties. Poor physiochemical properties warrant concern when one considers that the major sources of attrition in early development are unfavorable physiochemical properties like poor solubility, low permeability, and metabolic instability [11,12]. It is likely that increased molecular weight plays a significant role in these problems [13]. Making compounds of increased molecular weight and complexity is most probably an artifact of the desire to map a larger portion of diversity space and the activity one gains from good hydrophobic (van der Waals) interactions with the protein. Unfortunately, these compounds become part of the screening collection and are subsequently identified by HTS as hits. Thus, many programs often start with leads that are already too large by historical standards.

Aside from these problems there are also the issues with low hit rates and the failure of many hits to move forward in the lead refinement process [5,6]. Unfortunately, sampling the complex molecular space of drug-sized molecules with a small number of molecules in the screening collection favors a low hit rate. The complexity of this molecular space also makes sampling a rich diversity of compounds impossible. One begins to realize that the brute force solution of designing larger screening libraries is ultimately intractable. Thus, a method that would allow for: identifying molecules with improved physiochemical properties; a higher hit rate, and greater sampling of diversity space while maintaining a screening library of tractable size is clearly desirable. As a result, the development and implementation of fragment-based drug discovery (FBDD) has accelerated in recent years.

Fragment-based drug discovery

In FBDD [14–18], one identifies a key fragment, or set of fragments, that binds to the desired target. Since the fragments are small by design, their binding affinities will be low and, thus, optimization is necessary. Since they are small, the medicinal chemist has much more room to operate, however, before the molecular weight becomes unreasonably high. Further, it has been hypothesized that the fragments may bind more efficiently to a small region of the protein active site and provide a better starting point for medicinal chemistry. Optimization usually involves linking multiple fragments together or adding to a fragment (growing).

Over 25 years ago, William Jencks conceptualized fragment-based drug discovery [19]. In his paper, Jencks considered the Gibbs free energy of binding of a molecule in terms of its constituent parts along with a correction term that accounted for the free energy changes resulting from rotational and translational losses in degrees of freedom. In support of this idea, a study by Nakamura and Abeles found that the inhibition of HMG-CoA reductase could be understood in terms of the linkage of two fragments where each bound to unique sites of the enzyme [20]. Jencks's paper also served to help elucidate an earlier study where biotin was deconstructed into representative fragments that were found to bind, albeit weakly, to streptavidin [21]. Nonetheless, the practical implementation of Jencks's formulism posed two major challenges: (1) identifying fragments and (2) subsequently linking them together to form a viable ligand. These obstacles prevented the incorporation of fragment-based drug discovery into modern drug discovery for over a decade.

In 1996, Abbott Laboratory researchers Shuker, Hajduk, Meadows, and Fesik published the first practical implementation of FBDD called SAR by NMR [22]. The first step of their approach involved screening fragments to identify those that bound to the protein. Binding was determined by chemical shift changes in the NMR spectra following addition of the fragment to the protein. Once a lead had been identified, another screen was conducted using analogs of the initial fragment to optimize binding to this particular site of the protein. Next, a ligand was sought that would interact with another site proximal to the initial binding site. Optimization for binding to this site was done, as before, by screening analogs. Once the optimal fragments for each site had been found, NMR or X-ray crystallography was used to determine their position in the protein. This information was then used to link the respective fragments, with the goal of producing a high affinity ligand. This sparked a revolution for fragment-based drug discovery with many papers emerging on the topic.

In many respects, FBDD is intuitive. Many drug targets contain distinct binding sites for each piece of a ligand. Much of the fundamental chemistry of synthesizing a ligand involves modular architecture, and the process of optimization often involves focusing on key regions of the ligand separately. Clearly the combination of a well-characterized target and fragment-based methods favors the identification of novel molecules with improved affinity, selectivity, and physiochemical properties.

We have already discussed the intractable problem of sampling the complex diversity space associated with drug-sized molecules. In principle, fragment-based methods can offer a solution. For example, imagine the problem of a ligand that is composed of three modular pieces where each component has N possibilities. Sampling this diversity space entails N^3 permutations, which means N^3 molecules must be made and screened. The alternative route is to screen all N substituents as unlinked fragments. This step acts as a filter, eliminating any chemical moieties that do not bind. As a result, a smaller set of linked compounds could be made. This greatly improves our combinatorial problem.

A first-principles computational analysis suggests that there are roughly 14×10^6 stable, synthetically feasible, small molecules with a molecular weight less than or equal to 160 Da. This number increases to 44×10^6 when one considers stereoisomers but excludes three-member and four-member rings and elements other than carbon, oxygen, nitrogen, and halogens [2]. While this is still a large number, it is much smaller than the diversity space of synthesizable compounds, which is upward of 10^{60} [1]. Consequently, we are now in a range where one can construct a screening library that will sample a significant part of this fragment diversity space. For example, a modest screening library of 10^3–10^4 fragments represents a significant fraction of the lead-like space at 0.001–0.01. This is in contrast to a typical HTS library containing 10^5–10^6 compounds that still encompasses a small fraction at $\sim 10^{-55}$ of the total estimated drug-like space.

Further, one expects the hit rate from a fragment-based screen to be higher than the traditional HTS hit rate. According to Hann's model of molecular complexity [3] there is an increased probability

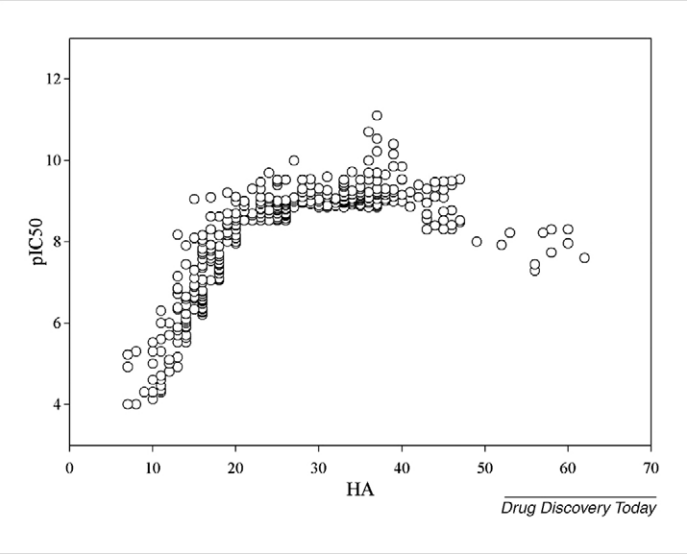


FIGURE 1

Plot of only the most potent inhibitors as a function of the number of heavy atoms (HA). The 'maximal affinities' as measured by pIC_{50} increase rapidly up to 20 heavy atoms, but plateau beyond 25.

of finding a molecular entity that binds to a target when the entity is of low molecular weight and less complex. A recent study [4] validated Hann's model empirically, finding hit rates of 0.001–0.151% for identification of ligands from HTS with an IC50 threshold in the micromolar range, compared with a 3% hit rate for a fragment screen using NMR with an affinity threshold in the millimolar range. This is also consistent with the explanation put forward for the observed fall-off in ligand efficiencies as molecular size increases [10].

Clearly, FBDD offers several attractive aspects compared to HTS. First, fewer compounds need to be screened. Second, the lower complexity results in a higher hit rate for fragments since smaller ligands (i.e. fragments) bind with greater inherent efficiencies. Third, since fragment-based design is guided by structural validation (usually by NMR or X-Ray), optimization should require the synthesis of many fewer compounds. Finally, since one is starting with fragments of low molecular weight, it is more likely that the physiochemical properties will at least start in a more desirable range [13].

Ligand efficiency (LE)

Ligand efficiency [7–10] is simply the binding free energy for a ligand divided by its molecular size. In our work we have chosen to follow the lead of the Kuntz lab [23] and define the size as the number of heavy (non-hydrogen) atoms. Potency can be defined equally well as pIC_{50}, pK_i, or ΔG. LE is a useful metric for measuring the impact on activity of the addition of more molecular bulk. Molecules that achieve a given potency with fewer heavy atoms are by definition more efficient. Using atoms efficiently in drug discovery is important, owing to the realization that larger ligands have an inherent disadvantage in terms of many physicochemical properties [13].

A large number of protein-ligand complexes with regard to their ligand efficiencies have been recently analyzed [9]. This analysis showed a distinct non-linearity in average, or maximal, potency as molecular size increased. The empirical data seem overwhelming in this regard. Focusing on only the best (i.e. most potent) ligands at each size reveals this trend very clearly (Figure 1). The same phenomenon can also be illustrated by plotting the ligand efficiencies (i.e. pK_i/HA) versus molecular size (Figure 2a). Ligand efficiency drops off very rapidly as molecular size increases until a plateau is reached at very large sizes (i.e. above 45 heavy atoms). Two physical phenomena that underlie this effect have been proposed [10]. The first is that the accessible surface area per atom of a ligand that is available to make interactions with a protein active site decreases with increasing size. This is a consequence of ever-greater buried surface area in larger, especially more branched, ligands. The second is that as ligands become larger, they must accommodate more constraints in the ligand binding site that result in inevitable structural compromises. Both factors reduce efficiency in larger ligands.

The fact that ligand efficiency is dependent on size makes direct comparison across wide size ranges problematic. An alternative metric called 'fit quality' to address this issue has been suggested

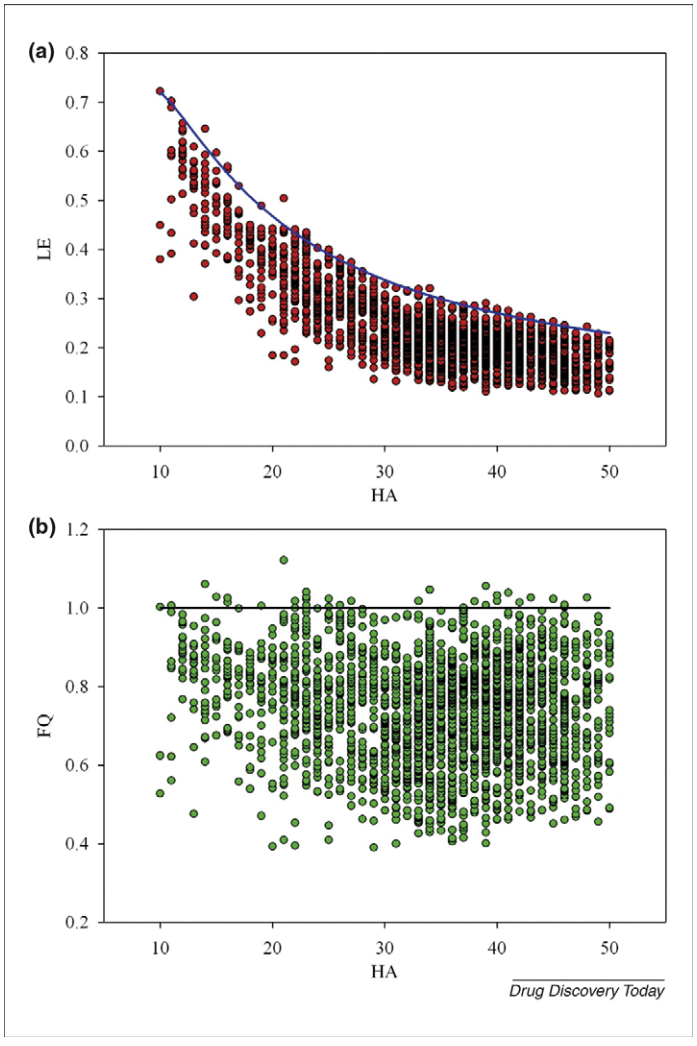


FIGURE 2

(a) Ligand efficiency as a function of heavy atoms for the K_i dataset is shown in the red circles. These values were scaled using the fit represented by the blue line (a) to produce the fit quality metric (green) shown in **(b)**. Fit quality scores around 1 (black line in (b)) indicate a near optimal ligand binding affinity for a given number of heavy atoms.

TABLE 1

Starting fragments and their optimized (from fragment linking or fragment growing) analogs for a variety of targets.

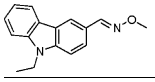
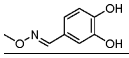
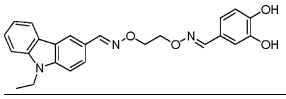
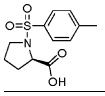
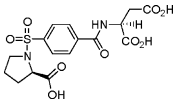
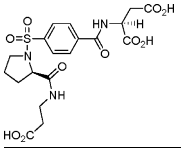
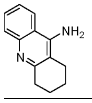
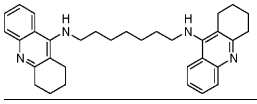
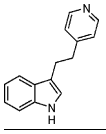
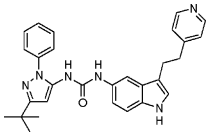
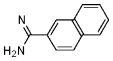
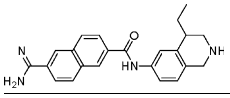
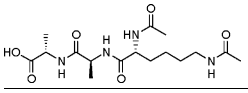
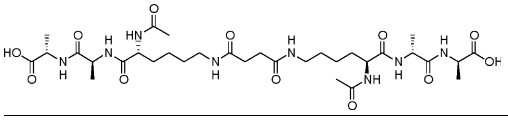
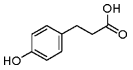
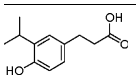
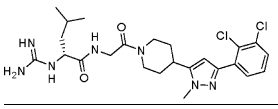
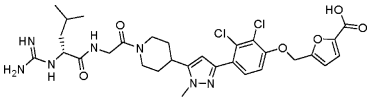
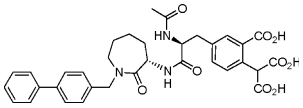
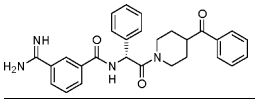
Molecule	Target	pKi	pIC50	LE	FQ
	c-SRC		4.40	0.232	0.476
	c-SRC		4.39	0.366	0.630
	c-SRC		7.19	0.232	0.705
	Thymidylate synthase		2.96	0.164	0.324
	Thymidylate synthase		4.62	0.154	0.456
	Thymidylate synthase		6.48	0.191	0.623
	AChE	6.22		0.415	0.714
	AChE	9.40		0.254	0.886
	p38		4.48	0.264	0.497
	p38		6.85	0.190	0.650
	Urokinase	5.23		0.402	0.693
	Urokinase	8.20		0.293	0.821
	Vancomycin	5.32		0.205	0.540
	Vancomycin	8.96		0.172	0.772
	FABP4		3.23	0.269	0.463
	FABP4		5.00	0.333	0.574
	IL-2		5.52	0.158	0.527

TABLE 1 (*Continued*)

Molecule	Target	pKi	pIC50	LE	FQ
	IL-2		7.22	0.160	0.648
	SH2		2.60	0.200	0.344
	SH2		8.52	0.223	0.760
	Factor Xa	3.70		0.411	0.708
	Factor Xa	7.80		0.223	0.744

[9]. Fit quality is a scaled ligand efficiency that centers optimal binders (best fit quality) near 1.0 by scaling the raw ligand efficiencies. The fit quality score (Eqs. (1) and (2)) is given by:

$$\text{LE\_Scale} = 0.0715 + \frac{7.5328}{\text{HA}} + \frac{25.7079}{\text{HA}^2} - \frac{361.4722}{\text{HA}^3} \quad (1)$$

$$\text{FQ} = \frac{\text{LE}}{\text{LE\_Scale}} \quad (2)$$

where HA is the number of heavy atoms, FQ is the fit quality, and LE is the ligand efficiency. A plot of the fit and the resulting FQ scores for the Ki dataset is shown in Figure 2. Since the fit quality is normalized with respect to size, identifying ligands with good and exceptional ligand efficiencies is easily accomplished. The best ligand efficiencies for a given atom count are those that have FQ scores near 1.0 while those with poor ligand efficiencies for their size fall much lower on the scale (e.g. FQ = 0.4). Therefore, the FQ score gives guidance throughout all phases of lead optimization by accessing whether atoms are being budgeted wisely with regard to their impact on potency.

Ligand efficiency of fragments

Calculation of the ligand efficiency and fit quality score for a dataset of 22 literature fragments and molecules resulting from optimization of these fragments over a variety of targets is presented here (Table 1) [17]. For the analysis here a fragment is considered, in more general terms, to be a starting substructure or template that a final ligand is built from, thus warranting the inclusion of some of the larger fragments in Table 1 such as those for IL-2 and Vancomycin. In calculating the fit quality scores, a value of 15 was used as a lower bound on the heavy atom count when calculating the LE_Scale. The definition of a lower bound is subjective, but truncating the scaling function at 15 makes sense because the data for compounds below 15 heavy atoms are sparse and dominated by targets, such as carbonic anhydrase, that contain special features (e.g. zinc binding site). It also makes sense to truncate the function at the other extreme, that is, 50 heavy atoms, for analogous reasons. The ligand efficiencies and fit quality versus the molecular size are given in Figure 3. Since there is a

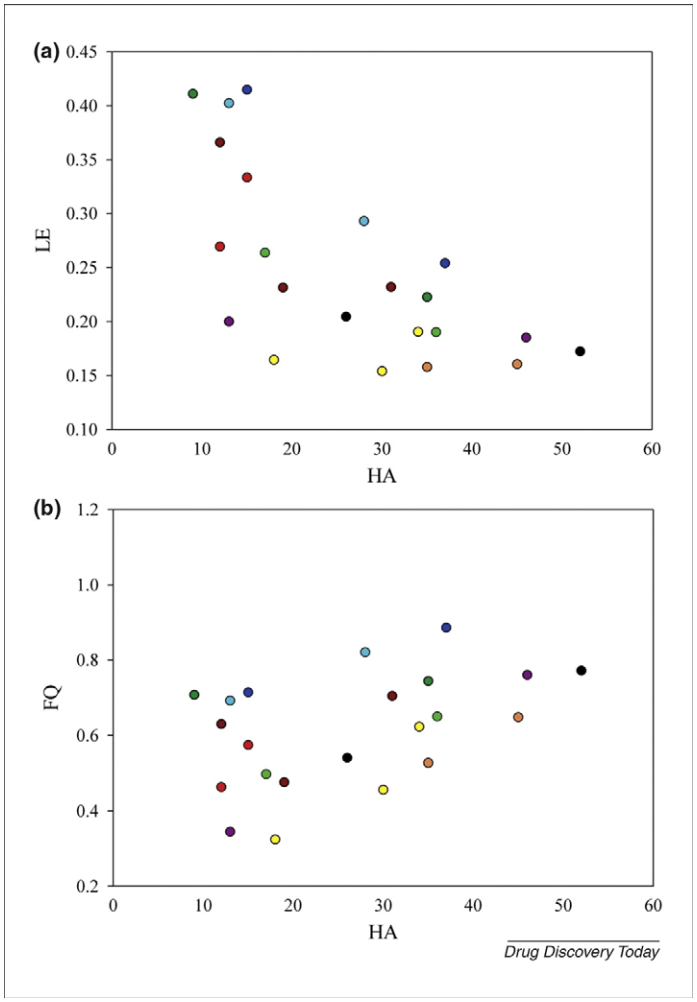


FIGURE 3

(a) Ligand efficiencies and **(b)** fit quality scores for a set of 22 literature molecules composed of fragments and the molecules resulting from optimization of these fragments for a variety of targets (● Factor Xa, ● AchE, ● Urokinase, ● c-SRC, ● FABP4, ● p38, ● SH2, ● Thymidylate synthase, ● Vancomycin, ● IL2).

limited amount of data and in order to more clearly reveal any trends, we show all the data (Ki and IC50) together in both these plots, even though the IC50 values are less transferable across datasets.

Analysis of these data provides some key insights into fragment-based design. First an overall decline in ligand efficiency as size increases is seen, just as was observed previously for ligands from traditional medicinal chemistry programs. While the number of data points is relatively small, perhaps raising concerns about the statistical validity, there does not seem to be a dramatic shift toward higher efficiencies for the fragments as compared to the general population of literature compounds found in a similar size regime. Indeed the fragment efficiencies are generally modest in comparison with some of the most potent ligands observed in the literature. In addition even within the fragment-based programs, the ligand efficiencies are not maintained in the larger ligands assembled from the fragments. By comparison the fit quality scores are maintained across the board, or in most cases even improved as we move from fragments to the more optimized structures. This provides some validation of the fit quality score and illustrates how it differs from the raw ligand efficiencies. It is likely that more optimization has occurred as the fragments are elaborated (grow in size) and their fits in the active site might be expected to be more refined. Therefore, we suggest that, unlike the raw ligand efficiency alone, the fit quality score can be used as a measure of efficiency across the entire optimization process from initial fragment hit to optimized clinical candidate.

This work raises some interesting questions with regard to fragment-based screening. Given the modest fit quality scores observed for the initial fragment hits one might question whether these starting points are really more optimal than leads from more traditional screening approaches. However, the improvement in fit quality scores in the optimized ligands may argue that optimization of these small fragments does proceed more efficiently. The data at this point are probably too limited for any firm conclusions.

Conclusions

The landscape of fragment-based drug discovery has changed dramatically. Starting from a theoretical concept, FBDD has evolved into an important new tool for drug discovery. Unlike traditional high-throughput screening, fragment screening provides a way of sampling a wide range of chemical space while screening only a small number of compounds. In theory, ligand efficiency provides a metric for assessing the quality of the hits found in fragment-based screening and helps assess the additional contribution functionality makes to the overall activity of a compound as it is optimized. We have found, however, that ligand efficiency is related intrinsically to molecular size, so it is difficult to compare ligand efficiencies for molecules of disparate sizes. Therefore, we have proposed a new scaled efficiency metric called fit quality that allows for such comparison.

The utility of this new metric was illustrated in the analysis of fragment optimization data from the literature. The efficiencies of fragments tend to be higher than the larger compounds assembled from the fragments. The size-independent fit quality scores show different behavior. In general, the more optimized larger ligands have better fit quality scores than the initial fragments from which they were constructed. Thus, fit quality may be a superior metric over the full course of a fragment-based program.

References

1 Bohacek, R.S. *et al.* (1996) The art and practice of structure-based drug design: a modeling perspective. *Med. Res. Rev.* 16, 3–50

2 Tobias Fink, *et al.* (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

3 Hann, M.M. *et al.* (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* 41, 856–864

4 Schuffenhauer, A. *et al.* (2005) Library design for fragment based screening. *Curr. Top. Med. Chem.* 5, 751–762

5 Mestres, J. and Veeneman, G.H. (2003) Identification of 'latent hits' in compound screening collections. *J. Med. Chem.* 46, 3441–3444

6 Gribbon, P. and Andreas, S. (2005) High-throughput drug discovery: what can we expect from HTS? *Drug Discov. Today* 10, 17–22

7 Hopkins, A.L. *et al.* (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* 9, 430–431

8 Abad-Zapatero, C. and Metz, J.T. (2005) Ligand efficiency indices as guideposts for drug discovery. *Drug Discov. Today* 10, 464–469

9 Reynolds, C.H. *et al.* (2007) The role of molecular size in ligand efficiency. *Bioorg. Med. Chem. Lett.* 17, 4258–4261

10 Reynolds, C.H. *et al.* (2008) Ligand binding efficiency: trends, physical basis, and implications. *J. Med. Chem.* 51, 2432–2438

11 Darvas, F. *et al.* (2002) *In silico* and *ex silico* ADME approaches for drug discovery. *Curr. Top. Med. Chem.* 2, 1287–1304

12 Di, L. and Kerns, E.H. (2003) Profiling drug-like properties in discovery research. *Curr. Opin. Chem. Biol.* 7, 402–408

13 Lipinski, C.A. (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today: Technol.* 1, 337–341

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

15 Erlanson, D.A. *et al.* (2004) Fragment-based drug discovery. *J. Med. Chem.* 47, 3463–3482

16 Hajduk, P.J. *et al.* (2007) A decade of fragment-based drug design: strategic advances and lessons learned. *Nature Rev. Drug Discov.* 6, 211–219

17 Rees, D.C. *et al.* (2004) Fragment-based lead discovery. *Nature Rev. Drug Discov.* 3, 660–672

18 Carr, R.A.E. *et al.* (2005) Fragment-based lead discovery: leads by design. *Drug Discov. Today* 10, 987–992

19 Jencks, W.P. (1981) On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci.* 78, 4046–4050

20 Nakamura, C.E. *et al.* (1985) Mode of interaction of beta-hydroxy-beta-methylglutaryl coenzyme A reductase with strong binding inhibitors: compactin and related compounds. *Biochemistry* 24, 1364–1376

21 Green, N.M. (1975) *Avidin. Adv. Protein Chem.* 29, 85–133

22 Shuker, S.B. *et al.* (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531–1534

23 Kuntz, I.D. *et al.* (1999) *Proc. Natl. Acad. Sci. U S A* 96, 9997–10002