Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 17 (2007) 4258–4261

The role of molecular size in ligand efficiency

Charles H. Reynolds,[a,*] Scott D. Bembenek[b] and Brett A. Tounge[a]

[a]*Johnson & Johnson Pharmaceutical Research and Development, L.L.C., Welsh and McKean Roads, PO Box 776, Spring House, PA 19477, USA*
[b]*Johnson & Johnson Pharmaceutical Research and Development, L.L.C., 3210 Merryfield Road, San Diego, CA 92037, USA*

Received 11 April 2007; revised 7 May 2007; accepted 10 May 2007
Available online 17 May 2007

Abstract—Ligand efficiency is a simple metric for assessing whether a ligand derives its potency from optimal fit with the protein target or simply by virtue of making many contacts. Comparison of protein–ligand binding affinities for over 8000 ligands with 28 protein targets shows conclusively that the average ligand binding affinities are not linear with molecular size. It is therefore important to scale ligand efficiencies by the size of the ligand, particularly where small ligands (e.g., fragments) are involved. We propose a simple 'fit quality' metric that removes this dependence.

Ligand efficiency is commonly defined as the binding free energy for a ligand divided by its molecular size.[1,2] Ligand efficiency has become an important concept in drug discovery partly due to the realization that large ligands have a decided disadvantage in terms of the molecular properties necessary for bioavailability.[3] More recently interest has also been driven by the emergence of fragment-based approaches to drug discovery where weakly binding fragments are used as a starting point for drug design. Ligand efficiency is a useful concept when trying to assess the relative value of fragments for follow up in fragment-based drug design.

Most medicinal chemistry programs are built on the simple idea that there is a certain amount of additivity between different functional groups in a given ligand. Thus it is common to attempt to optimize different regions in a lead structure somewhat independently with the goal of obtaining the best combination of R groups. One might reasonably expect that the potency of a ligand would, to a first approximation, be a linear combination of the affinities of the constituent pieces. Andrews et al.[4] exploited this idea to define group equivalents for common functional groups that could be combined to predict the likely average affinity of a ligand. The goal was to use this reference average affinity to assess the 'goodness of fit' of a prospective ligand. This work was followed by a key contribution from the Kuntz lab,[5] where affinities were examined for a variety of ligands against many different targets. This analysis showed several trends. First, certain types of ligands had much better intrinsic affinities by virtue of the nature of the protein–ligand interaction (e.g., covalent bonds, metals, or small ions). Second there appeared to be a flattening of the plot of free energy of binding versus molecular size for the most potent ligands. The latter conclusion was confounded to some degree by the fact that many of the small ligands involved metals, ions, and potential covalent interactions. Nevertheless, this work was largely responsible for defining ligand efficiency as the ratio of the affinity (ΔG, pK_i, pIC_{50}) divided by the number of heavy atoms. While other metrics have been suggested this is probably the most straightforward definition of ligand efficiency.

We have revisited the concept of maximal ligand efficiency using the much greater amount of protein–ligand affinity data that is now available. We have also focused on systems most relevant to drug discovery. For example we have omitted the metals and small ions that gave the enormous efficiencies seen previously. We initially analyzed protein targets separately, but it soon became apparent that the trends are essentially the same across all of the protein families we examined.

We extracted the affinity data used in this study from the BindingDB database[6] developed at the University of Maryland Biotechnology Institute. In order to obtain a

Keywords: Ligand efficiency; Fragment based design; Binding affinity.
* Corresponding author. Tel.: +1 215 628 5675; fax: +1 215 628 4985; e-mail: CReynol1@prdus.jnj.com

0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.05.038

Table 1. Protein targets for K_i and IC_{50} datasets

Target	Dataset[a]	Target	Dataset[a]
AchE	K_i/IC_{50}	MAO-B	K_i/IC_{50}
CYP19	K_i/IC_{50}	Plasmepsin-II	K_i
CA-I	K_i	Thrombin	K_i
CA-II	K_i/IC_{50}	BuChE	IC_{50}
CA-IV	K_i/IC_{50}	PKA	IC_{50}
Caspase-1	K_i/IC_{50}	FGFR-1	IC_{50}
Caspase-3	K_i/IC_{50}	GSK-3	IC_{50}
CDK-2	K_i/IC_{50}	HIV-1 RT	IC_{50}
CDK-4	K_i/IC_{50}	Neuraminidase-A	IC_{50}
Factor-Xa	K_i/IC_{50}	Neuraminidase-B	IC_{50}
HIV-1 protease	K_i	PTP1B	IC_{50}
MMP-1	K_i	VEGFR-2	IC_{50}
MMP-9	K_i	Farnesyl transferase	IC_{50}
MAO-A	K_i/IC_{50}	DPP-IV	K_i/IC_{50}

[a] Present in K_i, IC_{50} or both datasets.

good statistical sampling of protein–ligand complex affinities, we selected 28 protein targets as given in Table 1. We extracted two sets of ligand data for separate analysis: K_i and IC_{50}. The K_i dataset consisted of 2581 ligands and the IC_{50} dataset consisted of another 6072. The K_i values range from 0.01 pM to 213 μM. The reported IC_{50} values range from 0.01 nM to 5.5 mM. The number of atoms ranges from 6 to 78 for the K_i set and 7 to 62 for the IC_{50} set.

The ligand efficiencies were calculated as described above. The K_i or IC_{50} values were converted to pK_i or pIC_{50} values and then divided by the number of heavy atoms (HA) in the ligand (i.e., pK_i/HA). These ligand efficiency, heavy atom, and affinity values were used in all of the subsequent analysis.

Plots of the affinities versus size are given for the K_i and IC_{50} datasets in Figures 1 and 2, respectively. Examination of these figures shows behavior very similar to that seen in the Kuntz paper, except it is a bit less dramatic because of the absence of metals and small ions. While, of course, ligands can have a wide range of activities at any size, if one concentrates on the most active ligands for a given number of heavy atoms it is clear the change

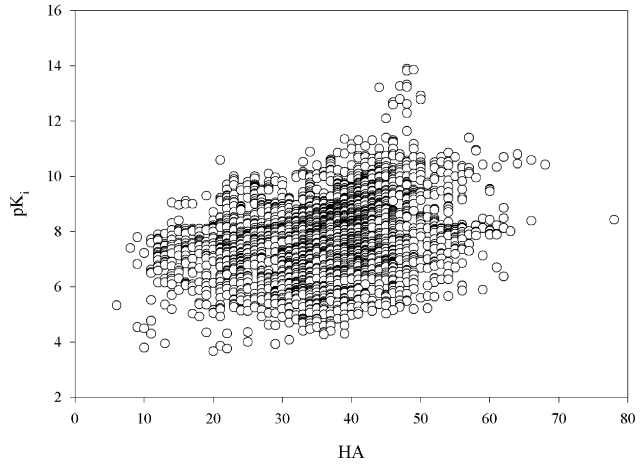

Figure 1. Plot of pK_i versus number of heavy atoms.

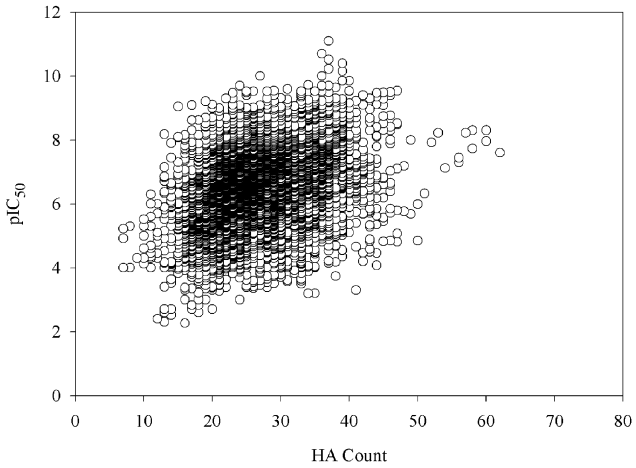

Figure 2. Plot of pIC_{50} versus number of heavy atoms.

in affinity for the best ligands ('maximal affinity') is not linear with size. In order to illustrate this point, Figure 3 shows only the few most potent ligands for any given ligand size.

Another way to show this trend is to plot the ligand efficiency values against the number of heavy atoms. If ligand binding affinity were linearly related to size then ligand efficiency would be constant as a function of size (number of heavy atoms). As can be seen from Figures 4 and 5, the best (and average) ligand efficiency drops dramatically as the size increases, especially in the early part of the curve. Thus, while we observe an average ligand efficiency for both datasets (Table 2) that is in line with the sometimes quoted value of 0.3 for an efficient binder, this value applies very poorly to small ligands. In the case of ligands under 20 heavy atoms much larger efficiencies are not only achievable, they are actually quite common. It is remarkable how consistent this result is across a wide spectrum of protein targets and ligands. Further the IC_{50} values track remarkably well with the presumably more reliable K_i data.

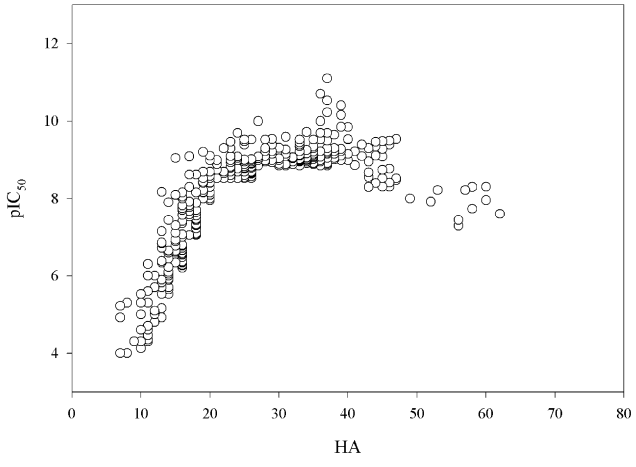

Figure 3. Plot of only the most potent ligands at each size. The 'maximal affinities' as measured by pIC_{50} increase rapidly up to 20 heavy atoms, but plateau beyond 25.

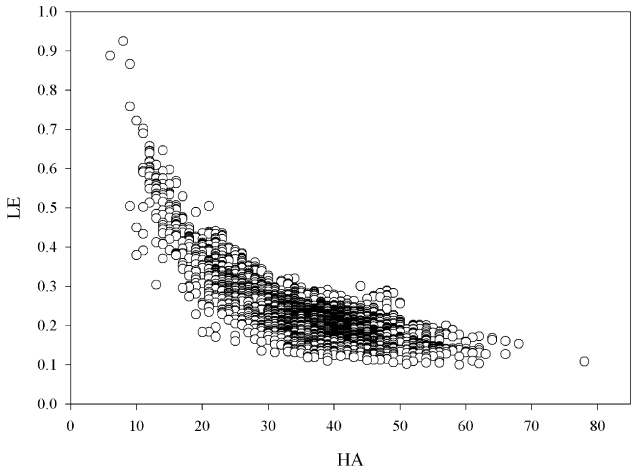

Figure 4. The ligand efficiency (pK_i/HA) as a function of number of heavy atoms for the K_i dataset. Ligand efficiency falls off dramatically between 10 and 25 heavy atoms.

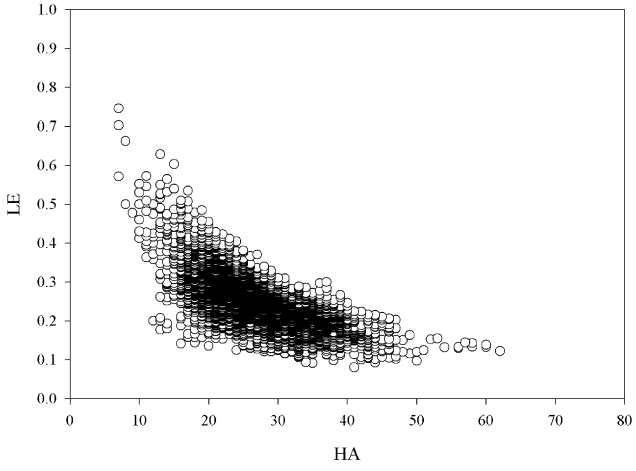

Figure 5. The ligand efficiency (pIC_{50}/HA) as a function of heavy atoms for the IC_{50} dataset. Ligand efficiency shows a similar precipitous decline between 10 and 25 heavy atoms.

Table 2. Statistics for the K_i and IC_{50} datasets

Dataset	Avg. pK_i/pIC_{50}[a]	Avg. HA[a]	Avg. LE[a]
K_i	7.8 (1.5)	36.1 (10.4)	0.24 (0.09)
IC_{50}	6.6 (1.2)	26.8 (6.8)	0.26 (0.07)

[a] HA, number of heavy atoms; LE, ligand efficiency. Standard deviation in parentheses.

Far from being an interesting curiosity, these results have important implications for the application of ligand efficiency. This is particularly true for fragment-based discovery where small ligands are examined routinely. These results show conclusively that ligand efficiency values must be normalized if one proposes to use them over a wide span of molecular sizes. Put simply, while a ligand efficiency of 0.3 might be considered very good for a ligand with 30 heavy atoms, it would be very poor for a ligand with fewer than 15 heavy atoms. There are many possible explanations for this observation. Andrews et al. explicitly included a crude entropy correction factor in their group additivity scheme. The assumption is that larger more flexible ligands might be forced to pay a larger entropic penalty for binding. It has also been suggested that as molecules become more complex the probability of fitting in an optimal way becomes smaller.[7,8] While beyond the scope of this paper, our own analysis indicates that a key factor is the lack of a linear relationship between the ligand surface area available for interaction with a protein and its atom count. Indeed, there are many factors that might contribute to this effect, but given the consistency we see across ligands and protein classes, we believe the empirical evidence for this general relationship between size and ligand efficiency is compelling.

As mentioned earlier, this effect is especially important when assessing smaller compounds such as might arise from a high-throughput screening campaign or in particular fragment-based drug design. The correct assessment of the quality (i.e., efficiency) of ligand binding for small ligands is very sensitive to proper scaling based on the number of heavy atoms. This is particularly true since fragments fall in the range where the change in efficiency with size is at its steepest, and it is certainly very different from average values derived from more typical drug-sized molecules.

We have proposed dealing with this effect in one of two ways. First, one can just use plots such as Figures 4 and 5 as a guideline directly to adjust expectations. Armed with this information medicinal chemists can better judge whether a particular efficiency is good, bad, or average based on its location on the plot. Another approach is to modify the ligand efficiency metric so that it takes the size into account.

This latter approach can be accomplished in any number of ways. We have chosen to adopt a 'fit quality' scale where each ligand efficiency value is scaled based on the maximum ligand efficiency for the heavy atom count for that ligand. To establish a scaling factor (LE_Scale), we fit the top ligand efficiency versus heavy atom count curve for the IC_{50} set to a simple exponential function (blue line in Fig. 6a). The resulting equation is

$$\text{LE\_Scale} = -0.064 + 0.873 * e^{(-0.026*\text{HA})}, \qquad (1)$$

where HA is defined as the number of heavy atoms in the molecule. So, for each ligand efficiency score (LE) we calculated the fit quality (FQ) as

$$\text{FQ} = \text{LE}/\text{LE\_Scale}. \qquad (2)$$

The exponential function was chosen because it is relatively simple and provides an empirically good fit. An alternative would be to bin the ligand sizes and use the average observed maximal ligand efficiencies within each bin as LE_Scale for that bin. In either case the objective is to scale the most efficient ligands (Fig. 6a) so that they have a normalized score of 1.0 regardless of size (Fig. 6b).

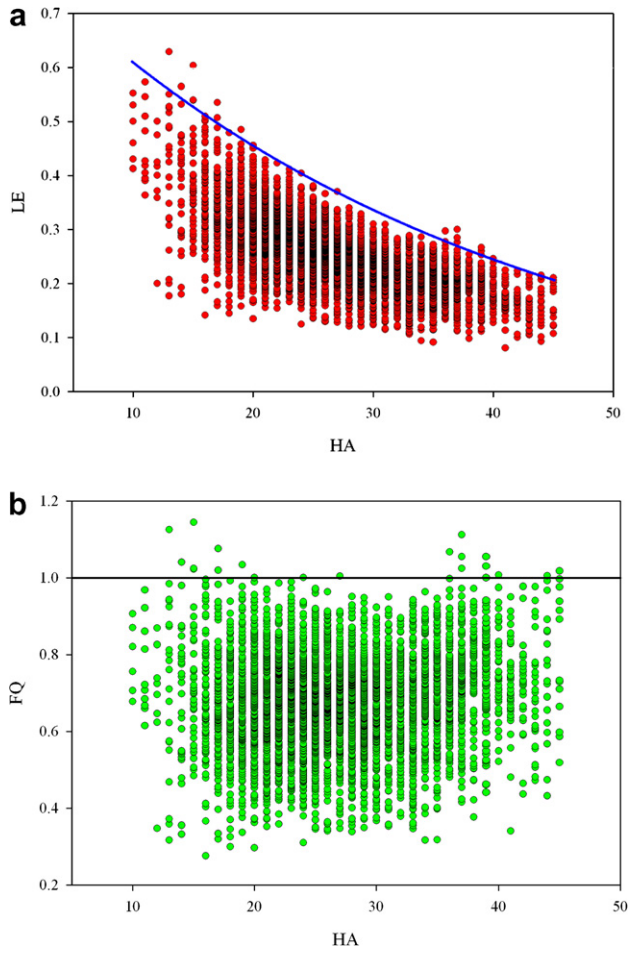

Figure 6. (a) The original ligand efficiency as a function of heavy atoms for the IC_{50} dataset is shown in the red circles. These values were scaled using the fit represented by the blue line (a) to produce the fit quality metric (green) shown in (b). Fit quality scores around 1 (black line in (b)) indicated a near optimal ligand binding affinity for a given number of heavy atoms.

A plot of the fit quality along with the original ligand efficiency values is shown in Figure 6 for all ligands with 10–45 heavy atoms. This new size normalized 'fit quality' score can be easily interpreted by medicinal chemists for ligands of any size. Fit quality scores near 1.0 (scores can exceed 1) indicate near optimal ligand binding, while low fit quality scores are indicative of sub-optimal binding. Using this criterion one can look at a chart such as Figure 6b and readily identify compounds that have near optimal affinities for their size, including a few with fit quality scores above one that must be regarded as extraordinarily efficient.

We chose to limit the range of data used to fit LE_Scale to include heavy atom counts between 10 and 45. These heavy atom boundaries were chosen to reflect limitations of our dataset. For any given molecular size (HA), affinity data are only available over a finite range. This can affect the upper and lower limits observable for ligand efficiency depending on the dataset. For example when examining Figure 4 the lowest ligand efficiency observed for a ligand with 10 heavy atoms is 0.35 for this dataset. This limit is simply due to the fact that there are no ligands in this set that are less potent than 0.2 mM. While ligands with lower potency might exist, in practice IC_{50} values are not usually obtained for such weak binders. Similarly when considering ligands larger than 45 heavy atoms the maximum ligand efficiency in our dataset is approximately 0.21. Higher ligand efficiencies for a ligand this large would require the presence of femtomolar binders in our dataset. Binding affinities in the femtomolar range are rare and in any case beyond the sensitivity of most practical assays. The boundaries of 10 and 45 encompass most of the region of interest to us and avoid using the extreme values for either very small or very large ligands.

In conclusion, an extensive examination of protein–ligand binding affinities shows a strong influence of molecular size on the average and maximal computed ligand efficiencies. This appears to be a general trend that is observed over a wide range of protein targets and ligands. Moreover the same trend is observed for both K_i and IC_{50} data. The curve for individual targets can be shifted slightly depending on the nature of the protein active site, but the overall trend is the same. It is important to take this dependence into account when assessing the efficiency of ligands, particularly small fragments as might be employed in fragment-based drug design. We have proposed a size normalized efficiency scale, fit quality (FQ), that can incorporate ligand efficiency and size into a single metric.

Acknowledgment

The authors acknowledge the bindingDB website (www.bindingdb) developed by the Gilson group at the University of Maryland Biotechnology Institute. It is an invaluable resource for this type of research.

References and notes

1. Hopkins, A. L.; Groom, C. R.; Alex, A. *Drug Discovery Today* **2004**, *9*, 430.
2. Abad-Zapatero, C.; Metz, J. T. *Drug Discovery Today* **2005**, *10*, 464.
3. Lipinski, C. A. *Drug Discovery Today: Technologies* **2004**, *1*, 337.
4. Andrews, P. R.; Craik, D. J.; Martin, J. L. *J. Med. Chem.* **1984**, *27*, 1648.
5. Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997.
6. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Res.* **2007**, *35*, D198.
7. Hann, M. M.; Leach, A. M.; Harper, G. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856.
8. Murray, C. W.; Verdonk, M. L. *J. Comput. Aided Mol. Des.* **2002**, *16*, 741.