Enthalpic Efficiency of Ligand Binding

György G. Ferenczy*,† and György M. Keserű*,‡

Sanofi-Aventis CHINOIN, 1−5. Tó u, Budapest, Hungary, H-1045, and Discovery Chemistry, Gedeon Richter Plc., 19−21. Gyömrői út, Budapest, Hungary, H-1103

Received April 1, 2010

The thermodynamics of ligand−protein binding has received much attention recently. In the present contribution we focus on the enthalpic component of binding. The dissociation constant, pK_d, was decomposed into enthalpic and entropic components ($pK_d = pK_H + pK_S$), and pK_H, defined as $pK_H = -\Delta H/(2.303 \cdot RT)$ was used to characterize the enthalpy contribution to binding. It was found that the maximal achievable pK_H decreases with increasing molecular size. This is in contrast to maximal pK_d that increases with molecular size until it achieves a plateau. Size-independent enthalpic efficiency (SIHE) was defined as SIHE = $pK_H/40 \cdot HA^{0.3}$, with HA being the number of heavy atoms. SIHE allows a size unbiased comparative binding characterization of compounds. It can find use in hit and lead selection and also in monitoring optimization in drug discovery programs. The physical background of decreasing maximal pK_H with molecular size is discussed, and its consequences to drug discovery are analyzed. It is concluded that the feasibility of simultaneous optimization of affinity and enthalpy diminishes with increasing molecular size. Consequently, binding thermodynamics considerations are to be applied primarily in hit prioritization and hit-to-lead optimization.

INTRODUCTION

Ligand efficiency measures are commonly used to characterize the affinity of ligands to their protein targets. The concept is designed to assess how well a ligand is optimized toward its target, to compare the binding ability of various ligands, and also to foresee what affinity can be achieved by medicinal chemistry optimization. It was shown that ligand efficiency defined as a per heavy atom binding affinity decreases with ligand size.[1−3] The decrease is steep for up to 20−25 heavy atoms and is modest for larger molecules. Various alternative ligand efficiency metrics were proposed as markers in the drug discovery process.[4,5] An analysis of the maximal available ligand efficiency as a function of molecular size led to the introduction of fit quality score[2,3] and size-independent ligand efficiency.[6] They were introduced to facilitate the comparison of ligands with different sizes.

Early phases of drug discovery are typically driven by potency optimization but other factors influencing developability become increasingly important at later stages.[7] It has been generally recognized that potency optimization is usually accompanied by unfavorable change in other properties, most notably the increase of molecular weight and lipophilicity.[8−10] This property shift was proposed to originate from the optimization practice in drug discovery projects,[7] where potency is primarily increased by the introduction of lipophilic groups and thereby increasing the entropic component of binding. A remedy to this situation is to increase the enthalpic component of the binding at the expense of the entropic term. This idea calls for the control of binding enthalpy that prompted Ladbury introducing enthalpic efficiency as the per heavy atom binding enthalpy.[11,12]

In the present contribution we explore the possibility of thermodynamically characterizing ligand binding in detail. Contrary to previous analyses,[2,3,6] we focused our study on binding thermodynamics data exclusively. Based on this data set a relationship between molecular size and maximal available enthalpy is set up, and this is compared to the established relationship between size versus ligand affinity and efficiency. A size-independent enthalpic efficiency (SIHE) measure is introduced that makes it possible to compare the binding thermodynamics of various size ligands. Implications of the molecular size dependence of the enthalpy are discussed, and the utility of SIHE together with an example is presented.

METHODS

Decomposition of pK_d into Enthalpic and Entropic Components. Here we show that pK_d, the negative logarithm of the dissociation constant, can be decomposed into enthalpic and entropic components, and these components are useful to characterize the thermodynamics of ligand binding.

We start with the relation between the binding free energy (ΔG_{assoc}) and K_d

$$\Delta G_{assoc} = -RT\ln K_{assoc} = RT\ln K_d \quad (1)$$

Using the enthalpy and entropy components of ΔG

$$\Delta H - T\Delta S = RT\ln K_d \quad (2)$$

This is equivalent with

$$\frac{\Delta H}{RT} - \frac{\Delta S}{R} = \ln K_d \quad (3)$$

that can be further transformed to yield

* Corresponding authors. E-mail: gyorgy.ferenczy@sanofi-aventis.com (G.G.F.) and gy.keseru@richter.hu (G.M.K.).
† Sanofi-Aventis CHINOIN.
‡ Gedeon Richter Plc.

10.1021/ci100125a
Published on Web 08/04/2010

$$\left[\frac{-\Delta H}{2.303 \cdot RT}\right] + \left[\frac{\Delta S}{2.303 \cdot R}\right] = \mathrm{p}K_\mathrm{d} \quad (4)$$

Then we define $\mathrm{p}K_\mathrm{H}$, the enthalpic component of $\mathrm{p}K_\mathrm{d}$

$$\frac{-\Delta H}{2.303 \cdot RT} = \mathrm{p}K_\mathrm{d}^{\mathrm{enthalpy}} = \mathrm{p}K_\mathrm{H} \quad (5)$$

and $\mathrm{p}K_\mathrm{S}$, the entropic component

$$\frac{\Delta S}{2.303 \cdot R} = \mathrm{p}K_\mathrm{d}^{\mathrm{entropy}} = \mathrm{p}K_\mathrm{S} \quad (6)$$

and thus we can write

$$\mathrm{p}K_\mathrm{H} + \mathrm{p}K_\mathrm{S} = \mathrm{p}K_\mathrm{d} \quad (7)$$

where $\mathrm{p}K_\mathrm{H}$ is a measure of the enthalpic component of binding. It can be derived directly from ΔH, and owing to its analogy with $\mathrm{p}K_\mathrm{d}$, it is an appropriate measure to be used in drug discovery. We will use $\mathrm{p}K_\mathrm{H}$ in the subsequent analysis. Note, that more negative enthalpy (ΔH) corresponds to more favorable binding and to a higher $\mathrm{p}K_\mathrm{H}$. In the forthcoming discussion, binding characterized with high $\mathrm{p}K_\mathrm{H}$ will be referred to as high enthalpy binding.

Data Collection and Processing. Binding thermodynamics data were collected from three databases: BindingDB,[13] PDBcal,[14] and Scorpio.[15] Altogether, data for 1232 complexes were found that refer to 648 unique ligands. Some ligands appeared in complex with more than a single protein, while some ligand−protein complexes appeared more than once. Compounds of RNA complexes and those containing other than C, N, O, P, S, H, and halogen atoms were removed. Highly polar compounds that are out of the interest of drug discovery programs were also removed; *a*log $P > -3$ criterium was applied. Biotin that has an extreme affinity toward streptavidin was also discarded. Finally, 812 ligand−protein complexes were investigated (Supporting Information). All $\mathrm{p}K_\mathrm{d}$, $\mathrm{p}K_\mathrm{H}$, and $\mathrm{p}K_\mathrm{S}$ values were calculated from ΔG, ΔH, and $T\Delta S$ data as described above.

RESULTS AND DISCUSSION

Analysis of Binding Thermodynamics Data. The analysis was started with the investigation of $\mathrm{p}K_\mathrm{d}$ as the function of the number of heavy atoms (HA) (Figure 1a). Similar plots with a larger number of data points were presented by Reynolds et al.[2] using $\mathrm{p}K_\mathrm{i}$ and IC_{50} data. Although these parameters could deviate significantly from $\mathrm{p}K_\mathrm{d}$ values measured thermodynamically, we performed a comparison to assess how our reduced number of data points represents the more complete data set. The analysis is restricted to the range of 10−50 HAs that span the size of the drug-like molecules of interest. The maximal $\mathrm{p}K_\mathrm{i}$ values as a function of HA in Figure 1 of ref 2 starts at around HA = 8 and rises up to about 40, where it has a plateau near to $\mathrm{p}K_\mathrm{i} = 12$. The maximum $\mathrm{p}K_\mathrm{d}$ values of the HA−$\mathrm{p}K_\mathrm{d}$ plot on Figure 1a also starts at about $\mathrm{p}K_\mathrm{d} = 8$ and reaches its maximum of about $\mathrm{p}K_\mathrm{d} = 12$ at HA = 40, although the increase is less smooth than in ref 2. In order to make a more quantitative comparison of the maximal affinities in our data set and those in ref 2, we fitted a function to HA−$\mathrm{p}K_\mathrm{dmax}$. The $\mathrm{p}K_\mathrm{dmax}$ values were extracted from the HA−$\mathrm{p}K_\mathrm{d}$ function by dividing HA into intervals of HA = 5 and by picking the largest $\mathrm{p}K_\mathrm{d}$ values from each interval. We chose the functional form proposed by Nissink[6] rather than that of ref 2, as the former leads to a simpler definition of a size-independent measure of ligand efficiency. Thus we sought for a relationship between HA and $\mathrm{p}K_\mathrm{dmax}$ in the form of

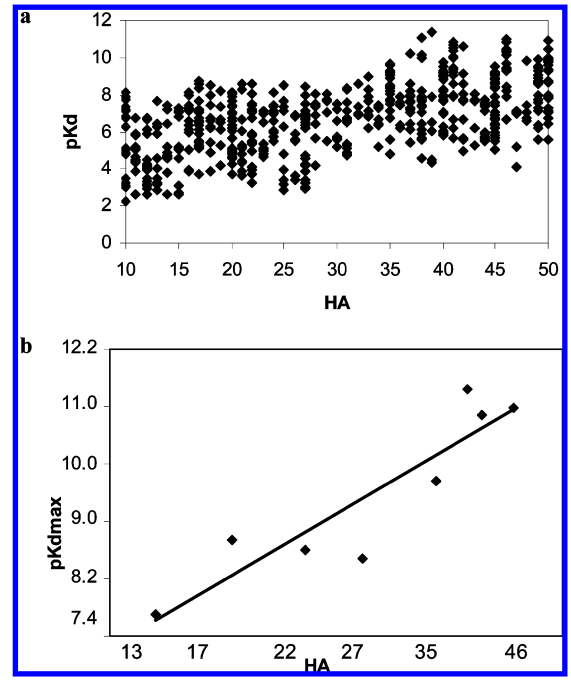

Figure 1. (a) Plot of $\mathrm{p}K_\mathrm{d}$ versus number of heavy atoms (HA). (b) Logarithmic plot of $\mathrm{p}K_\mathrm{d}$ of the most potent ligands versus the number of HAs.

$$\ln(pK_\mathrm{dmax}) = a\ln(\mathrm{HA}) + b \quad (8)$$

This can be written as

$$\ln\left(\frac{\mathrm{p}K_\mathrm{dmax}}{\mathrm{HA}}\right) = (a - 1) \cdot \ln(\mathrm{HA}) + b \quad (9)$$

that corresponds to Nissink's eq 3a in ref 6. Both eqs 8 and 9 can be used to fit the *a* parameter, however, the statistical parameters of the two fits will be different. Using our $\mathrm{p}K_\mathrm{d}$ data set, we obtain $a = 0.29$ and $b = 1.27$ (Figure 1b) that are close to the more complete $\mathrm{p}K_\mathrm{i}$ data set parameters (slope: $a - 1 = -0.73$, intercept: $b = 1.40$) of ref 6. The similarity between the HA−$\mathrm{p}K_\mathrm{imax}$ and HA−$\mathrm{p}K_\mathrm{dmax}$ functions shows that our data set is representative enough and allows the analysis of $\mathrm{p}K_\mathrm{d}$, $\mathrm{p}K_\mathrm{H}$, and $\mathrm{p}K_\mathrm{S}$ values as functions of the number of HAs.

Figure 2a shows the HA−$\mathrm{p}K_\mathrm{H}$ plot, and Figure 2b shows the ln(HA)−ln($\mathrm{p}K_\mathrm{Hmax}$) plot. (In analogy to highest affinity compounds of $\mathrm{p}K_\mathrm{dmax}$, we identify compounds of $\mathrm{p}K_\mathrm{Hmax}$ as highest enthalpy compounds.) $\mathrm{p}K_\mathrm{H}$ covers a wide range with extremes at −13 and +20. This is in contrast to $\mathrm{p}K_\mathrm{d}$ that is always positive and extends up to about 12 (cf. Figures 1a and 2a). The HA dependence of $\mathrm{p}K_\mathrm{Hmax}$ is also different from that of $\mathrm{p}K_\mathrm{dmax}$ (cf. Figures 1b and 2b; note the different scale of the vertical axes). While $\mathrm{p}K_\mathrm{dmax}$ shows an increasing trend up to about 40 HAs, $\mathrm{p}K_\mathrm{Hmax}$ exhibits a decreasing trend, i.e., the achievable maximal binding enthalpy decreases with increasing molecular size. Replacing $\mathrm{p}K_\mathrm{dmax}$ by $\mathrm{p}K_\mathrm{Hmax}$ in eq 8 and fitting the parameters to the points in Figure 2b results in $a = -0.3$ and $b = \ln(40)$ and thus

$$pK_{\mathrm{Hmax}} = 40 \cdot \mathrm{HA}^{-0.3} \quad (10)$$

Using this result we can proceed to define a size-independent enthalpic efficiency

$$\mathrm{SIHE} = \frac{\mathrm{p}K_{\mathrm{H}}}{40} \cdot \mathrm{HA}^{0.3} \quad (11)$$

The HA−SIHE plot is shown on Figure 3. Highest enthalpy compounds have SIHE values around 1 independent of their size. Thus SIHE is a good measure of the enthalpic efficiency. A higher SIHE is better, its limit is around 1, and the comparison of molecules is not biased by their differing size.

Here we note that enthalpic efficiency (EE) was previously defined as EE = ΔH/HA.[11] According to our analysis, the maximal ΔH increases (less favorable) with size, and thus the maximal EE increases even faster. This has to be taken into account when the binding enthalpy content of different size ligands is compared.

It is worth comparing the SIHE of eq 11 with a definition of size-independent ligand efficiency (SILE). Nissink[6] proposed the formula

$$\mathrm{SILE} = \mathrm{p}K_{\mathrm{i}} \cdot \mathrm{HA}^{-0.3} \quad (12)$$

based on the pK_{i} data set of Reynolds et al.[2] (Note that in contrast to our definition of SIHE, this SILE is not normalized to 1.) As it was discussed earlier, the size dependence of the affinity (pK_{d}) of our data set is similar to that of the pK_{i} of the Reynolds et al. set. Thus the different powers of HA in eqs 11 and 12 well reflect the different size dependence of enthalpy (pK_{H}) and affinity (pK_{d}). The positive sign in the power of HA in SIHE results in a factor $\mathrm{HA}^{0.3}$ that increases with increasing molecular size, as it is required to compensate for the decrease of pK_{Hmax}, the maximal available enthalpy. On the other hand, the opposite is true for SILE, where HA appears with a negative power.

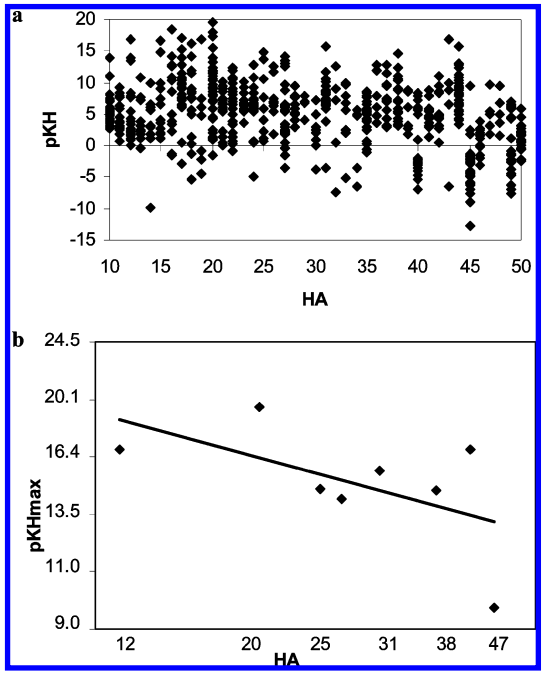

Figure 2. (a) Plot of pK_{H} versus the number of HA. (b) Logarithmic plot of maximal pK_{H} versus the number of HAs.

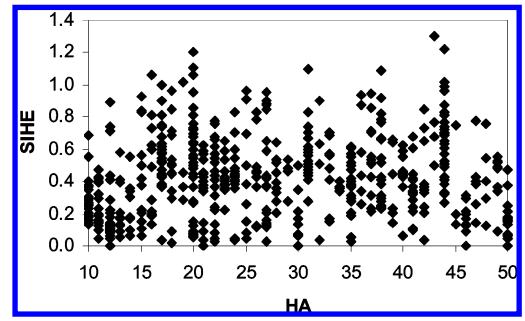

Figure 3. Plot of size-independent enthalpic efficiency (SIHE) versus number of HAs.

Interpretation of the Size Dependency of Ligand Binding Enthalpy. The binding of a ligand to its protein target is a complex process; the initially solvated ligand and protein form interactions, and water molecules previously participated in the solvation become part of the bulk solvent. The enthalpy component of the free energy change accompanied by this process is dominated by polar interactions, including hydrogen bonds (H-bonds). A gain in binding enthalpy is realized by specific interactions formed between the ligand and the protein and also by the released water molecules being more beneficial than those present before the protein−ligand binding. The reasons why such favorable interactions are more difficult to achieve as the size of the ligand increases (eq 10) are discussed below.

As polar interactions are sensitive to the relative position of the interacting partners optimal interactions can only be achieved when ligand and protein structures fit well. As the size of the ligand increases, it is less likely that such a good fit can be realized. This was demonstrated by a simple model in ref 16, demonstrating that the chance of useful interactions falls dramatically with the complexity of the interacting partners. Another factor to be taken into account is that the interaction with the small water molecules is less restricted by steric constraints, and thus solvation becomes more favorable than ligand−protein association as ligand size increases.

The existence of a single distinguished binding site, as proposed in ref 17, is also in favor of ligands with limited size. It is argued that the loss of rigid-body entropy upon binding represents a significant barrier to binding, and in most targets, it can be overcome only at a single site. These critical binding regions (hot spots) are sensitive to structural changes, and they provide an indispensable contribution to binding. We assume that enthalpy dominates such binding events. The extension of ligands to adjacent sites gives a relatively less important contribution to the binding, and this is in line with the observed decrease of binding enthalpy of larger ligands. Rejto and Verkhivker[18] also proposed the existence of a unique site of binding that is characterized by an unfrustrated energy landscape and thus a unique binding mode of small core fragments. Although computational docking did not find outstanding scores for these sites,[18] we propose that they correspond to the critical binding regions of ref 17 and are associated with important enthalpy contributions.

It is instructive to decompose pK_{dmax} into its pK_{H} and pK_{S} components. We emphasize that this enthalpy component of the highest affinity compounds does not correspond to the maximal achievable binding enthalpy (pK_{Hmax}). In order

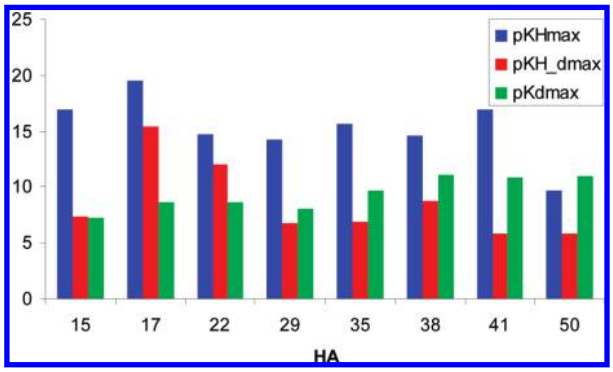

Figure 4. The pK_H of highest enthalpy compounds (pK_{Hmax}), and pK_H and pK_d of highest affinity compounds (pK_{H_dmax} and pK_{dmax}) versus number of HAs.

Table 1. Selected Property Means of Large (30 < HA < 50) High-Affinity (pK_d > 10) and High-Enthalpy (pK_H > 11) Compounds

	high-affinity compounds	high-enthalpy compounds
heavy atoms	42.1	37.6
molecular weight	602	542
*a*log *P*	3.3	0.2
H-bond acceptors	7.5	9.2
H-bond donors	3.7	3.7
charged atoms	0.2	1.4
apolar surface area (Å^2)	414	330

to remember this distinction, we will designate the enthalpy component of pK_{dmax} as pK_{H_dmax}. Note, that pK_{H_dmax} values were extracted from our data set in the following way: The five highest affinity compounds for a given HA were taken, and the one with the highest pK_H was selected as pK_{H_dmax}. In this way, pK_{H_dmax} represents the maximal enthalpy among the maximal affinity compounds. Figure 4 shows together pK_{Hmax}, pK_{H_dmax}, and pK_{dmax} as funcions of the number of HAs. The enthalpy content of maximal affinity compounds decreases with increasing compound size (cf. red and green bars), while pK_{H_dmax} exceeds pK_{dmax} for smaller ligands, it is below pK_{dmax} for larger ligands. In other words, binding of highest affinity ligands is basically enthalpy driven when ligands are small, and it becomes entropy driven when ligands are large. On the other hand, the enthalpy of highest enthalpy ligands (pK_{Hmax}) is consistently higher than pK_{dmax} (cf. blue and green bars). This shows that high binding enthalpy can be achieved even for large compounds, and the question arises, while the high enthalpy is not accompanied with high affinity. A structural comparison of large-size high-affinity and high-enthalpy compounds (Table 1) reveals that the latter are significantly more polar, as it is manifested by their lower log *P* and higher number of H-bond donors and charged atoms. This is in line with the observation that the transfer of polar ligands from water into the crystalline phase (the latter is assumed to have interactions similar to ligand-protein complexes) is enthalpically favorable and is accompanied by entropic penalty.[19]

Summarizing the analysis of Figure 4, it can be concluded that the enthalpy content of high affinity binding decreases with increasing ligand size and that high binding enthalpy can be achieved with increased polarity that tends to decrease the affinity for larger compounds. These findings have important implications for compound selection and optimization in drug discovery programs as it is discussed below.

Consequences to Optimizations in Drug Discovery. It is well documented that drug candidates tend to have increased size and lipophilicity, and this property shift is basically responsible for ADMET (absorption, distribution, metabolism, excretion, toxicity)-related problems[20] and for an increased attrition rate. More polar compounds whose binding to their target is characterized by a highly favorable enthalpy are expected to be less susceptible to ADMET issues,[22,23] partly owing to their decreased promiscuity. We found as a tendency that size increase is accompanied by a decreasing enthalpy component. On the other hand, it was observed that affinity increases with compound size up to about 40 HAs (Figure 1 and refs 2 and 3) and that highly active compounds are, on average, more complex.[21] Due to the opposite slopes of the maximal affinity (pK_{dmax}) and the corresponding enthalpy (pK_{H_dmax}) vs HA, these lines have an intersection between 20 and 30 HAs that is the typical size of lead compounds.[24] The increasing separation of pK_{dmax} and pK_{H_dmax} for larger compounds suggests that the chance of enthalpy driven affinity optimization diminishes rapidly. The situation in drug discovery programs may be more favorable as leads typically have a pK_d lower than pK_{dmax}, and this may be accompanied by a pK_H higher than that of pK_{H_dmax}. Nevertheless, the opposite slopes of the affinity and enthalpy vs HA functions represent a limit for enthalpy driven affinity optimization, and data suggest that the critical size is similar to the size of lead compounds. Consequently, affinity improvements achieved for compounds around or above this size are typically realized by entropy driven optimizations. In other words, entropic effects are making greater contributions in larger molecules that is reasonable given the trend in medicinal chemistry optimizations to add more "grease" (and size) to increase affinity.[25] This finding provides a thermodynamic rationale for unfavorable property shifts and gives further rational to keeping the molecular size under control in drug discovery programs. Another important consequence is that lead generation provides more opportunity for enthalpy optimizations than late-phase lead optimization.

We argue that it is rational focusing binding thermodynamics studies to hit discovery and hit-to-lead optimization. This strategy would provide viable leads with high binding enthalpy that might improve the success rate of lead optimization due to multiple reasons. First, leads identified in this way have a greater chance to interact with thermodynamic hot-spots in the active site. Since most proteins seem to have a single energetic focal point having a decisive contribution to the binding energy,[17,18] the identification of hot spots is crucial, which improves the odds of the subsequent extensive multidimensional optimization. Second, monitoring binding thermodynamics contributes to the localization of key binding motifs, and therefore, the resulting leads contain structural elements critical in ligand binding. Finally, enthalpy driven optimization of enthalpy prioritized hits would provide better quality leads in terms of their physicochemical profile.

In accordance with the intersection of pK_{dmax} and pK_{H_dmax} curves, the improvement in affinity typically achieved at the lead optimization phase is entropically driven resulting in unfavorable property shifts. Medicinal chemistry teams are

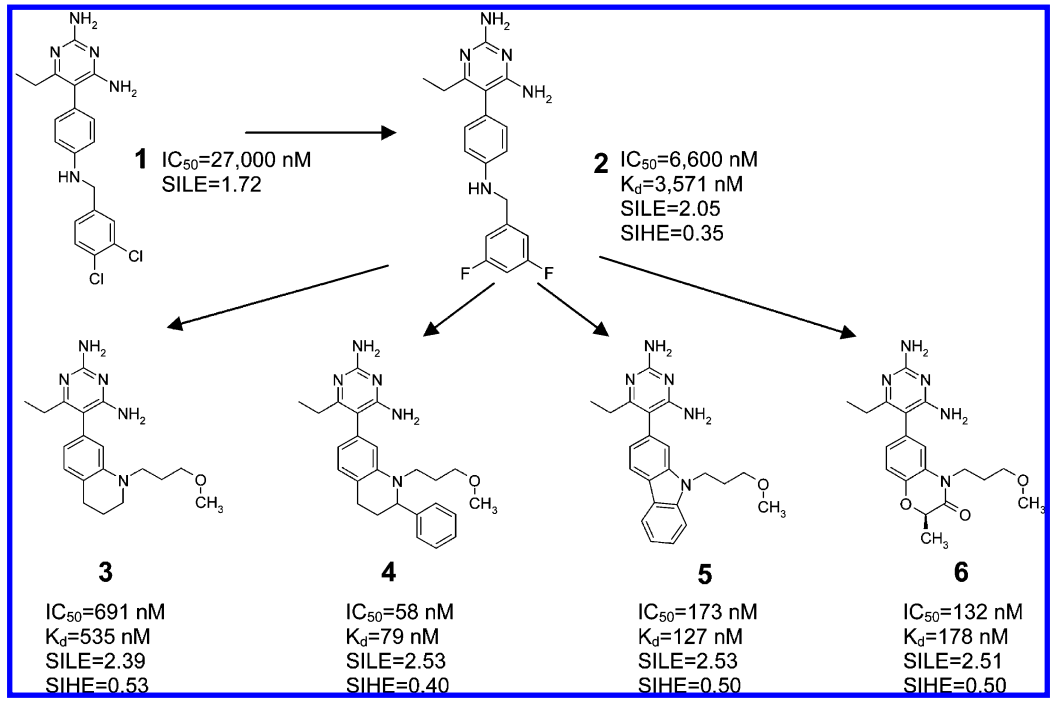

Figure 5. Prioritization of scaffolds in the early phase optimization of renin inhibitors. SILE values were calculated from K_d data except for compound **1** where IC_{50} was used.

therefore faced with the challenging trade off between affinity and ADMET-related physicochemical properties. It seems that increasing molecular weight and lipophylicity associated with the present medicinal chemistry practice is thermodynamically determined. Consequently, the selection of leads with high enthalpic efficiency is crucial for delivering good quality development candidates.

We demonstrated earlier that ligand efficiency (LE) remains constant during hit-to-lead optimization.[7] Furthermore, multidimensional optimization of leads typically decreases the ligand efficiency. These findings are rationalized by the confronting requirements of high affinity and high binding enthalpy and points to the importance of starting the optimization with reasonably sized and appropriately polar compounds (hits) that have a high binding enthalpy component. Furthermore, it is important to keep the size increase in check since it almost inevitably accompanies affinity improvements achieved in optimization campaigns. Fragment-based approaches coupled with monitoring binding thermodynamics would fit perfectly with this concept.

Recognizing that pure enthalpic optimization is an unrealistic objective for larger compounds also implies that it is not feasible to optimize against SILE and SIHE simultaneously. Consequently, both of these metrics should be used when prioritizing compounds for medicinal chemistry follow-up. This situation is exemplified by a set of renin inhibitors published recently.

Pfizer reported the early phase optimization of renin inhibitors started from a 27 μM high-throughput screening (HTS) hit (**1**) (Figure 5).[26] The surroundings of the hit were first investigated by a focused library approach that led to the identification of **2**. Since the X-ray analysis of the renin-**2** complex revealed the S3 subpocket empty,[26] several conformationally constrained scaffolds with the proposed methoxypropyl side chain were synthesized and tested (compounds **3**−**6** in Figure 5). It was demonstrated that the side chain fits to the subpocket[26] that, in addition to the restricted flexibility of the core, resulted in significant improvement in the affinity. Analyzing SILE and SIHE values, we attempt to rationalize scaffold selection for follow-up studies. SILE values of scaffolds **4**−**6** are above that of **3**, showing the formers as preferred in terms of ligand efficiency. Comparing SIHE values for these scaffolds, however, gives priority to scaffolds **5** and **6.** Although the original paper[26] identified scaffold **3** as a lead, subsequent studies revealed that the team optimized scaffold **6** further[27] and pushed its structural analogue into preclinical development.[28]

CONCLUSION

The dissociation constant of a protein−ligand complex, pK_d, can be decomposed into enthalpic and entropic components, $pK_d = pK_H + pK_S$. The term pK_H, defined as $pK_H = -\Delta H/(2.303 \cdot RT)$, measures the enthalpy contribution to binding. It was found that the maximal achievable pK_H (pK_{Hmax}) decreases with increasing molecular size. This is in contrast to the maximal pK_d (pK_{dmax}) that increases with molecular size until it achieves a plateau. The binding enthalpy of maximal affinity compounds (pK_{H_dmax}) also decreases with molecular size. We found that binding is typically enthalpy dominated for small ligands and becomes entropy dominated for larger compounds. These findings have the implications for drug discovery programs that: (i) compounds with enthalpy dominated binding thermodynamics are the preferred starting points for optimization, and (ii) the feasibility of enthalpic-driven affinity optimization diminishes with increasing compound size, and thus, it is to be performed up to the lead identification. Size-independent enthalpic efficiency (SIHE) defined as $SIHE = pK_H/40 \cdot HA^{0.3}$, with HA being the number of heavy atoms, is a suitable measure to support size unbiased compound selection in hit prioritization and to monitor the binding enthalpy in hit-to-lead optimization.

Note Added after ASAP Publication. Corrections were made to formulas in the Abstract and Conclusions, and to refs 7, 22, and 23 in the version published ASAP August 4, 2010. The corrected version was published ASAP on August 23, 2010.

Supporting Information Available: ΔG, ΔH, and $T\Delta S$ data and calculated pK_d, pK_H, and pK_S values for 812 ligand−protein complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The maximal affinity of ligands. *Proc. Natl Acad. Sci. U.S.A.* **1999**, *96*, 9997–10002.
(2) Reynolds, C. H.; Bembenek, S. D.; Tounge, B. A. The role of molecular size in ligand efficiency. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4258–4261.
(3) Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand Binding Efficiency: Trends, Physical Bias, and Implications. *J. Med. Chem.* **2008**, *51*, 2432–2438.
(4) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430–431.
(5) Abad-Zapatero, C.; Metz, J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discovery Today* **2005**, *10*, 464–469.
(6) Nissink, J. W. M. Simple Size-Independent Measure of Ligand Efficiency. *J. Chem. Inf. Model.* **2009**, *49*, 1617–1622.
(7) Keserű, G. M.; Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discovery* **2009**, *8*, 203–212.
(8) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 3743–3748.
(9) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is There a Difference between Leads and Drugs? A Historical Perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308–1315.
(10) Morphy, R. The Influence of Target Family and Functional Activity on the Physicochemical Properties of Pre-Clinical Compounds. *J. Med. Chem.* **2006**, *49*, 2969–2978.
(11) Ladbury, J. E. Enthalpic efficiency and the role of thermodynamic data in drug development: possibility or a pipeline dream. *Eur. Pharm. Rev.* **2007**, *12*, 59–62.
(12) Ladbury, J. E.; Klebe, G.; Freire, E. Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat. Rev. Drug Discovery* **2010**, *9*, 23–27.
(13) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.
(14) Li, L.; Dantzer, J. J.; Nowacki, J.; O'Callaghan, B. J.; Meroueh, S. O. PDBcal: A Comprehensive Dataset for Receptor-Ligand Interactions with Three-Dimensional Structures and Binding Thermodynamics from Isothermal Titration Calorimetry. *Chem. Biol. Drug Des.* **2008**, *71*, 529–32.
(15) Scorpio, Structure-Calorimetry of Reported Protein Interactions Online; School of Crystallography, Birkbeck, University of London: London; http://www.biochem.ucl.ac.uk/scorpio/scorpio.html. Accessed January 26, 2010
(16) Hann, M. M.; Leach, A. R.; Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.
(17) Murray, C. W.; Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput.-Aided Mol. Design* **2002**, *16*, 741–753.
(18) Rejto, P. A.; Verkhivker, G. M. Unraveling principles of lead discovery: from unfrustrated energy landscapes to novel molecular anchors. *Proc. Natl. Acad. Sci. U S A.* **1996**, *93*, 8945–50.
(19) Habermann, S. M.; Murphy, K. P. Energetics of hydrogen bonding in proteins: A model compound study. *Protein Sci.* **1996**, *5*, 1229–1239.
(20) Gleeson, M. P. Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. *J. Med. Chem.* **2008**, *51*, 817–834.
(21) Selzer, P.; Roth, H.-J.; Ertl, P.; Schuffenhauer, A. Complex molecules. Do they add value. *Curr. Opin. Chem. Biol.* **2005**, *9*, 310–316.
(22) Smith, G. F. Medicinal chemistry by the numbers: The physicochemistry, thermodynamics and kinetics of modern drug design. *Prog. Med. Chem.* **2009**, *48*, 1–29.
(23) Ferenczy, G. G.; Keserű, G. M. Thermodynamics guided lead discovery and optimization. *Drug Discovery Today*, **2010**, accepted for publication.
(24) Oprea, T. L.; Allu, T. K.; Fara, D. C.; Rad, R. F.; Ostopovici, L.; Bologa, C. G. Lead-like, drug-like or "Pub-like": how different are they. *J. Comput.-Aided Mol. Design* **2007**, *21*, 113–119.
(25) Carlson, H. A.; Smith, R. D.; Khazanov, N. A.; Kirchhoff, P. D.; Dunbar, J. B., Jr.; Benson, M. L. Differences between High- and Low-Affinity Complexes of Enzymes and Nonenzymes. *J. Med. Chem.* **2008**, *51*, 6432–6441.
(26) Sarver, R. W.; Peevers, J.; Cody, W. L.; Ciske, F. L.; Dyer, J.; Emerson, S. D.; Hagadorn, J. C.; Holsworth, D. D.; Jalaie, M.; Kaufman, M.; Mastronardi, M.; McConnell, P.; Powell, N. A.; Quin, J., 3rd; Van Huis, C. A.; Zhang, E.; Mochalkin, I. Binding thermodynamics of substituted diaminopyrimidine renin inhibitors. *Anal. Biochem.* **2007**, *360*, 30–40.
(27) Powell, N. A.; Ciske, F. L.; Cai, C.; Holsworth, D. D.; Mennen, K.; Van Huis, C. A.; Jalaie, M.; Day, J.; Mastronardi, M.; McConnell, P.; Mochalkin, I.; Zhang, E.; Ryan, M. J.; Bryant, J.; Collard, W.; Ferreira, S.; Gu, C.; Collins, R.; Edmunds, J. J. Rational design of 6-(2,4-diaminopyrimidinyl)-1,4-benzoxazin-3-ones as small molecule renin inhibitors. *Bioorg. Med. Chem.* **2007**, *15*, 5912–5949.
(28) Thomson Reuters Integrity; Thomson Reuters: Barcelona, Spain; http://integrity.prous.com/integrity/xmlxsl/. Accessed January 26, 2010.

CI100125A