ACS Medicinal Chemistry Letters

LETTER

pubs.acs.org/acsmedchemlett

# Thermodynamics of Ligand Binding and Efficiency

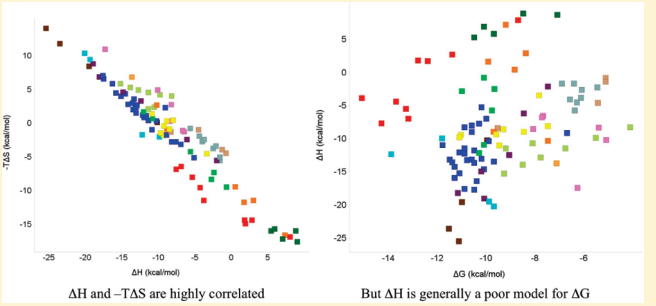
Charles H. Reynolds*,† and M. Katharine Holloway‡

†Ansaris, Four Valley Square, 512 East Township Line Road, Blue Bell, Pennsylvania 19442, United States

‡Merck Research Laboratories, P.O. Box 4, West Point, Pennsylvania 19486, United States

Supporting Information

**ABSTRACT:** Analysis of the experimental binding thermodynamics for approximately 100 protein−ligand complexes provides important insights into the factors governing ligand affinity and efficiency. The commonly accepted correlation between enthalpy and $-T\Delta S$ is clearly observed for this relatively diverse data set. It is also clear that affinity (i.e., $\Delta G$) is not generally correlated to either enthalpy or $-T\Delta S$. This is a worrisome trend since the vast majority of computational structure-based design is carried out using interaction energies for one, or at most a few, ligand poses. As such, these energies are most closely comparable to enthalpies not free energies. Closer inspection of the data shows that in a few cases the enthalpy (or $-T\Delta S$) is correlated with free energy. It is tempting to speculate that this could be an important consideration as to why some targets are readily amenable to modeling and others are not. Additionally, analysis of the enthalpy and $-T\Delta S$ efficiencies shows that the trends observed for ligand efficiencies with respect to molecular size are primarily a consequence of enthalpic, not entropic, effects.




**KEYWORDS:** Ligand efficiency, enthalpy entropy compensation, ligand binding, free energy, drug design, structure-based design

The primary goal of any drug discovery effort is to increase the potency of prospective compounds to the greatest degree possible while also optimizing the many other physical properties necessary for drug development. In most cases, little effort is invested in understanding the underlying thermodynamic quantities of enthalpy and entropy that are responsible for determining free energy. More recently, it has become common to recast binding free energies in terms of ligand efficiency, that is, the potency per heavy atom.[1−5] Ligand efficiency and the thermodynamic properties ($\Delta H$ and $\Delta S$) can provide insights into ligand binding that go beyond simple comparisons of potency. It has been suggested that more efficient ligands and ligands with more negative enthalpies of binding[6−8] provide better starting points for lead optimization. While these two concepts have generally been considered independently, an argument will be provided here that they are related.

With regard to enthalpy and entropy of binding, Friere and others[6,9] have presented an intriguing argument that drugs (ligands) that bind predominantly due to favorable enthalpies enjoy certain advantages over drugs where binding is driven predominantly by entropy. Some of the arguments in favor of enthalpy are that enthalpic interactions improve selectivity due to their geometric specificity, and they are inherently more efficient since they tend to be larger in magnitude than entropic effects. The second point already hints to an effect on overall ligand efficiency. The case for ligand efficiency has been made extensively in the literature and is now well accepted. Recent work[4,8,10,11] has shown that the standard definition of ligand efficiency is problematic for comparing ligands of disparate size because ligand efficiency is itself a function of molecular size. One question that arises in this connection is the degree to which enthalpy and entropy contribute to the observed decrease in average efficiencies as ligands become larger. Entropy is commonly regarded as a problem for larger ligands since presumably more conformational degrees of freedom should be lost upon binding. Indeed, this was the logic employed by Andrews[2] for including a correction factor in his group additivity scheme based on the number of rotatable bonds. In a previous paper, computational analysis of the contribution of conformational entropy showed no discernible trend[4] with respect to ligand size.

The experimental measurement of the thermodynamics of ligand binding can be routinely accomplished through techniques such as isothermal titration calorimetry (ITC).[12] However, because of the complexity, quantity of protein required, and low-throughput of ITC experiments, the enthalpy and entropy of ligand binding are not routinely measured. Through the work of a number of primarily academic laboratories, there has been a steady growth in the availability of this data for a variety of protein targets. Some of this data has been archived in the publicly available BindingDB database.[13] We have employed thermodynamic data for approximately 100 compounds across 14 target classes from BindingDB to analyze the effect of enthalpy and entropy on ligand binding and efficiency.

The Binding thermodynamics data for 102 ligand−protein complexes used in this study are included as Supporting

**Received:** January 11, 2011
**Accepted:** March 15, 2011
**Published:** March 23, 2011

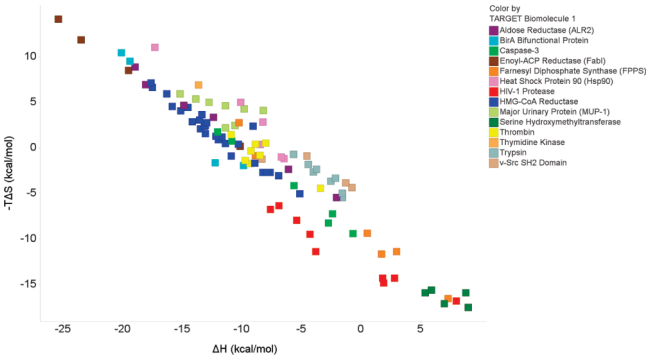



**Figure 1.** Plot of enthalpy ($\Delta H$) vs entropy ($-T\Delta S$).

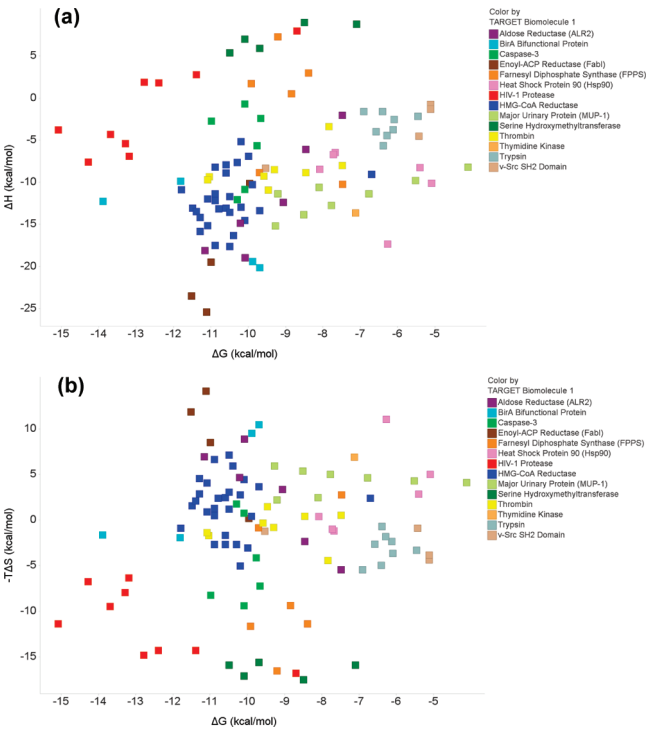



**Figure 2.** (a) Plot of enthalpy ($\Delta H$) and (b) plot of entropy ($-T\Delta S$) vs free energy ($\Delta G$).

Information. In addition to the enthalpies and entropies of binding, the ligand efficiencies have been computed based on free energy (as is typically done), enthalpy, and $-T\Delta S$. This provides considerable insight into the factors that drive the previously observed trends in ligand efficiency. In addition, the thermodynamic properties for specific protein targets such as HIV-1 protease, aldose reductase, HMG Co-A reductase, and caspase-3 were examined more closely. These targets were chosen because they have very different active sites in terms of size and polarity, and the observed trends in thermodynamic properties differ significantly. The plots and statistics were generated using Spotfire.[14]

A widely observed feature of protein−ligand binding thermodynamics is the seeming tug-of-war between enthalpy and entropy.[15−18] In general, protein−ligand complexes that exhibit more negative enthalpies of binding do so at the cost of more positive $-T\Delta S$ terms and vice versa. This enthalpy−entropy compensation effect is clearly evidenced in the protein−ligand complexes in the current study. A plot comparing $\Delta H$ to $-T\Delta S$ shows a remarkable correlation (Figure 1).

The ligands with the most favorable entropies of binding actually have positive enthalpies of binding. Conversely, the most favorable enthalpies have very unfavorable entropies of binding. Even though the $\Delta H$ and $-T\Delta S$ values are highly correlated, this does not mean that the enthalpy or entropy of binding is a good proxy for the free energy. Plots comparing $\Delta G$ to either $\Delta H$ or $-T\Delta S$ show no such correlation (Figure 2).

Intuitively, one would expect that ligands that have very specific interactions of the kind expected to produce more negative enthalpies would be more rigid and, therefore, less favorable entropically. It is also true that numerically the enthalpy and entropy must be negatively correlated if the free energy is held constant since free energy is a sum of $\Delta H$ and $-T\Delta S$. The distribution of enthalpies and entropies ($-T\Delta S$) differs, with almost all reasonable binders having negative enthalpies of binding, while only slightly more than half of the ligands in this particular data set have negative $-T\Delta S$ values. The plots of enthalpy and entropy versus free energy (Figure 2) underline the hazards of using either as a surrogate for free energy.[19−21] For example, the interaction enthalpies in Figure 2a overall show no correlation with free energy. This is troubling since most modeling studies use interaction energies computed from one low energy structure, that is, essentially enthalpy.

A closer examination of Figure 2, however, does show (Figure 3a,b) that in some specific targets enthalpy and free energy do appear to be reasonably well correlated (e.g., HIV-1 protease and aldose reductase). Least-squares fits of this data confirm (Table 1) that the free energies for these targets are highly correlated with enthalpy. This may explain why these targets have been amenable to quantitative modeling.[22,23] Other individual targets, such as HMG CoA reductase and caspase-3, are consistent with the overall trend, having no discernible correlation between enthalpy and entropy (Figure 3c,d). It should also be noted that in the cases where enthalpy and free energy are correlated, the same is true for entropy, a consequence of enthalpy−entropy compensation. Thus, any theories as to why certain targets show a better correlation between enthalpy and free energy must also consider that the correlation may also be due to some aspect of the entropy ($-T\Delta S$). Indeed, in the case of HIV-1 protease, it might be argued that this protein is "well behaved" in terms of the correlation between computed interaction energies and affinity because affinity is dominated by $-T\Delta S$, presumably primarily displacement of water.

It is clear from the statistics in Table 1 that there is a wide variation in the predictive power of enthalpies with respect to free energies of binding. This is a factor that must be considered when modeling protein−ligand interactions.

It is interesting to examine the relationship of each thermodynamic property with respect to molecular size (Figure 4). One might argue based on Figure 4 that there is a very slight trend toward the most favorable enthalpies being observed in the small molecule regime and the most favorable entropies being observed in the large molecule regime. However, these trends appear tenuous at best for the raw $\Delta H$ and $-T\Delta S$ values.

The trends become clearer when the enthalpies and entropies of binding are converted to enthalpy and entropy efficiencies, that is, $\Delta H$/(number of heavy atoms) and $-T\Delta S$/(number of heavy atoms). Just to show that this relatively small data set is not unusual, the typical ligand efficiencies ($\Delta G$/number of heavy atoms) are


dx.doi.org/10.1021/ml200010k |ACS Med. Chem. Lett. 2011, 2, 433–437

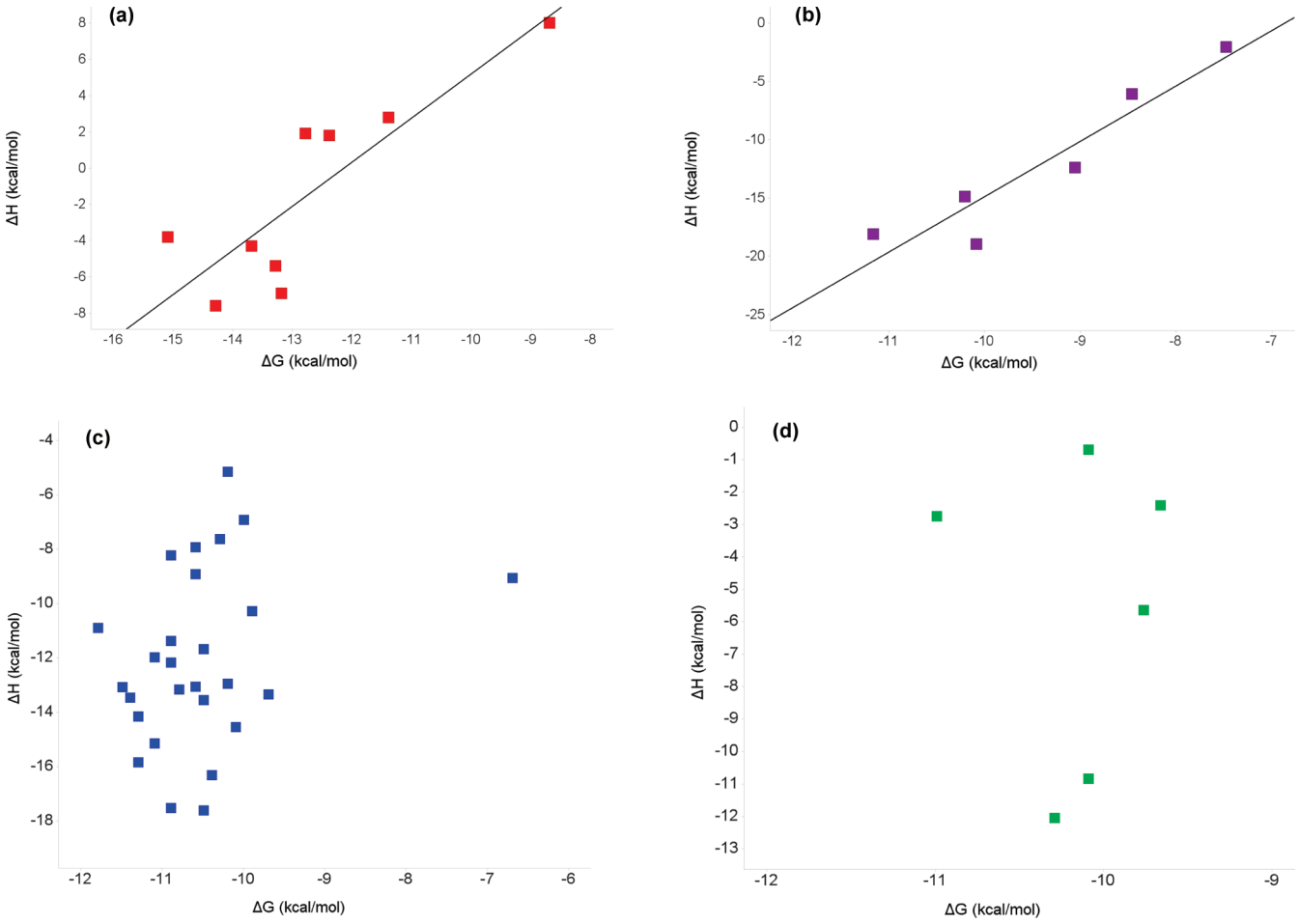



**Figure 3.** Plot of $\Delta H$ vs $\Delta G$ for (a) HIV-1 protease, (b) aldose reductase, (c) HMG CoA reductase, and (d) caspase-3.

**Table 1. Correlation of $\Delta H$ with $\Delta G$**

| | $N$ | $r^2$ |
|---|---|---|
| aldose reductase | 6 | 0.88 |
| enoyl ACP | 4 | 0.85 |
| v-SRC SH2 domain | 4 | 0.82 |
| serine hydroxymethyl transferase | 5 | 0.74 |
| HIV-1 protease | 9 | 0.73 |
| MUP 1 | 8 | 0.66 |
| BirA | 4 | 0.64 |
| thrombin | 8 | 0.34 |
| HSP 90 | 6 | 0.15 |
| farnesyl diphosphate synthase | 6 | 0.10 |
| HMG CoA reductase | 27 | 0.08 |
| caspase 3 | 6 | 0.00 |
| trypsin | 8 | 0.00 |
| thymidine kinase | 1 | - |

plotted in Figure 5. These ligand efficiencies follow the same trend observed with previous data sets of decreasing optimal efficiencies with increasing size.[4,10,11]

The enthalpy and entropy efficiencies are given in Figure 6a,b, respectively. It is immediately apparent that enthalpy and entropy behave differently. In general, the enthalpy efficiencies exhibit a curve very similar to that seen for the free energies (Figure 5). The most optimal enthalpy efficiencies deteriorate significantly as molecular size (number of heavy atoms) increases as do, to a lesser degree, the average enthalpy efficiencies. For example, the average enthalpy efficiencies for heavy atom counts of 0−19, 20−39, and 40−77 are −0.66, −0.33, and −0.13, respectively. By contrast, the entropy efficiencies exhibit very different behavior. Strictly speaking, the optimal entropy efficiencies also become worse as molecular size increases but by a much smaller margin. Furthermore, even this slight loss of efficiency as size increases is due to just four ligands targeting farnesyl diphosphate. Otherwise, the optimal entropy efficiency is little changed across the entire range of molecular sizes. Interestingly, these same four ligands have unusually unfavorable enthalpy efficiencies, another case of enthalpy−entropy compensation that results in middle of the road free energy efficiencies (Figure 5). The distribution of entropy efficiencies is very broad for small ligands and becomes significantly narrower as the ligands become larger. However, the average entropy efficiencies are relatively unchanged across the full range of molecular sizes, with an overall average value near 0. This can be shown by comparing the average entropy efficiencies for different size ranges, as was done above for the enthalpy efficiencies. The average entropy efficiencies for heavy atom counts of 0−19, 20−39, and 40−77 are +0.06, +0.01, and −0.10, respectively. This analysis of the experimental ITC data is consistent with


dx.doi.org/10.1021/ml200010k |ACS Med. Chem. Lett. 2011, 2, 433–437

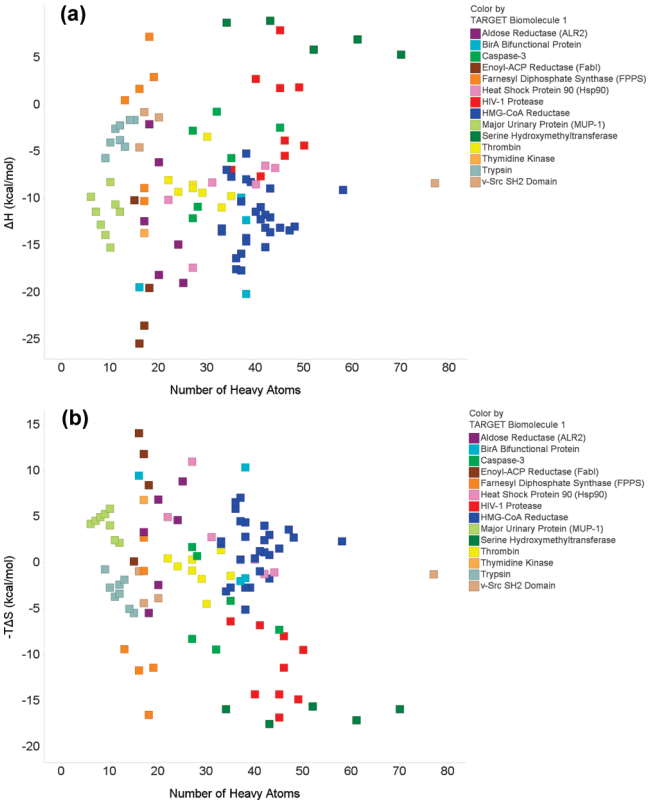



**Figure 4.** (a) $\Delta H$ and (b) $-T\Delta S$ vs molecular size (number of heavy atoms).

earlier calculations[4] in suggesting that the binding entropies ($-T\Delta S$) are not driving the difference in ligand efficiencies observed for small and large ligands. The effect appears to be primarily enthalpic—a result that might seem surprising but is also consistent with a recently reported analysis of binding enthalpies.[8]

This analysis of the enthalpy and entropy efficiencies is consistent with previous work in suggesting that the large change, in average (or maximal) ligand efficiencies, observed as the size of the ligand is varied is intrinsically an enthalpic effect not, as sometimes supposed, an entropic effect. Previous calculations found no trend with respect to size in the entropy of binding due to loss of conformational entropy in the ligands.[4] These results are also consistent, albeit perhaps not proof, of the hypothesis put forward based on simple computational models that the size effect on enthalpy is a result of increasing ligand complexity and the need to satisfy multiple geometric constraints simultaneously.[4,24] The plot of entropy efficiencies versus size (Figure 6b) indicates that average entropy effects may indeed be essentially linear with respect to size, although the distribution is extremely broad for small- to moderate-sized ligands.

In conclusion, we have analyzed the measured enthalpies and entropies of binding for approximately 100 protein−ligand complexes to gain an understanding of some of the trends in these thermodynamic properties. Examination of the enthalpies and entropies shows that there is a very strong negative correlation between these thermodynamic properties. This is consistent with conventional drug discovery wisdom and is of course partly a consequence of the relationship between $\Delta H$, $-T\Delta S$, and $\Delta G$. Plots of these properties with respect to molecular size show a

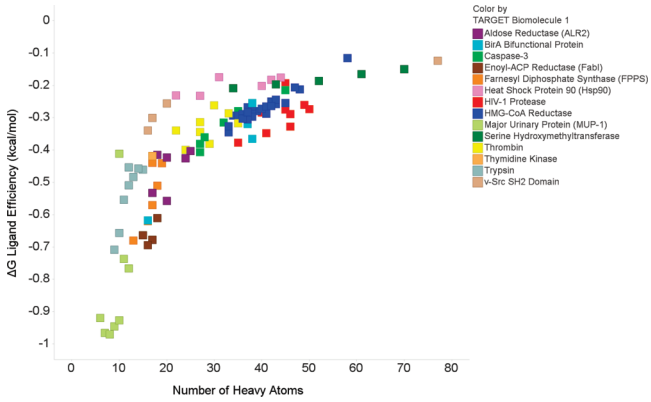



**Figure 5.** Free energy ligand efficiencies, $\Delta G$/(number of heavy atoms), vs number of heavy atoms.

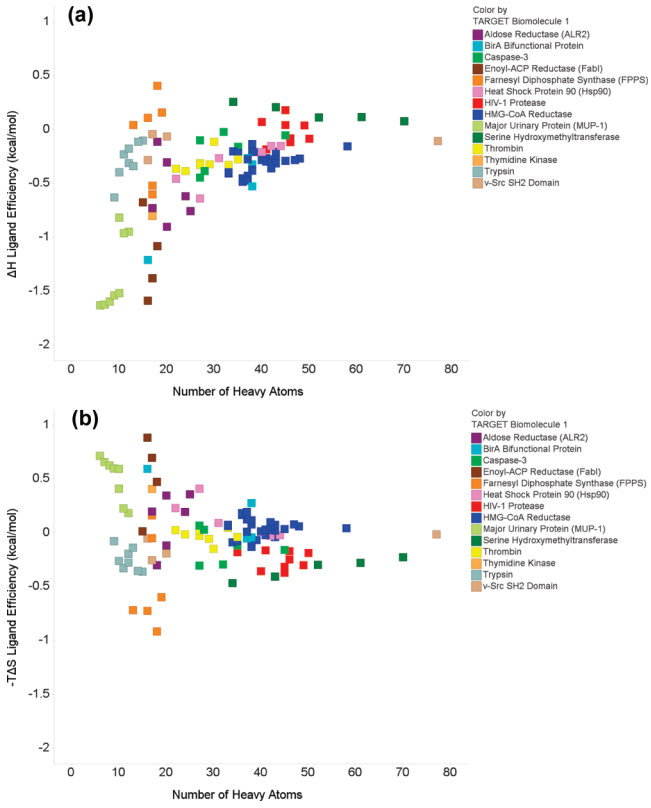



**Figure 6.** (a) Enthalpy efficiency, $\Delta H$/(number of heavy atoms), and (b) entropy efficiency, $-T\Delta S$/(number of heavy atoms), vs number of heavy atoms.

number of trends that are significant for drug discovery. First, there is a general trend, probably most easily seen by examining the enthalpy and entropy efficiencies, that on average enthalpy dominates the affinity of most small ligands. It is also apparent that the strong size dependence on ligand efficiencies (average or optimal) reported in the literature is mostly a consequence of enthalpy, not entropy. This is consistent with previous computational results that implicated enthalpy rather than conformational entropy.[4] The entropy contributions for the current set of ligands are interesting in that the average contribution per heavy atom is remarkably consistent, and very near 0, across the full range of molecular sizes.


dx.doi.org/10.1021/ml200010k |ACS Med. Chem. Lett. 2011, 2, 433–437

In the case of small ligands, the variance is very large, but it appears to converge to a small range of values for larger ligands. This is somewhat counterintuitive relative to conventional arguments that larger more floppy ligands might be expected on average to be penalized in terms of entropy and may show that conformational entropy is a relatively small contributor to binding relative to other factors, such as displacing waters from the active site. The latter effect might be expected to be more correlated with molecular size and more important for larger ligands in more hydrophobic binding sites.

Finally, while it is true that the correlation between $\Delta H$, or $-T\Delta S$, with free energy is overall very poor, in many individual cases the correlation is quite good. This is an important consideration for efforts to model protein−ligand interactions. In most cases, the energies used in modeling studies are generated from a single configuration (e.g., pose or conformation) or, at best, a few critical configurations and might be most appropriately considered an approximate enthalpy. On the basis of the analysis above, these calculations might be expected to provide good estimated affinities in cases where $\Delta H$ (and by association $-T\Delta S$) are correlated with affinity. However, in cases where this is not true, apparently a common situation (Figure 2a), the results would be expected to be very poor. This provides ample food for thought as to the importance of using computed free energies rather than simple interaction energies in structure-based drug design.

## ■ ASSOCIATED CONTENT

**Supporting Information.** Excel spreadsheet containing the structures, thermodynamic values, and BindingDB reference numbers. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*E-mail: creynolds@ansarisbio.com or chuck_reynolds@me.com.

## ■ ACKNOWLEDGMENT

C.H.R. acknowledges many stimulating discussions with Scott Bembenek, Alan Gibbs, and Brett Tounge.

## ■ REFERENCES

(1) Abad-Zapatero, C.; Metz James, T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discovery Today* **2005**, *10*, 464–469.

(2) Andrews, P. R.; Craik, D. J.; Martin, J. L. Functional group contributions to drug-receptor interactions. *J. Med. Chem.* **1984**, *27*, 1648–1657.

(3) 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.

(4) Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand binding efficiency: Trends, physical basis, and implications. *J. Med. Chem.* **2008**, *51*, 2432–2438.

(5) Abad-Zapatero, C. Ligand efficiency indices for effective drug discovery. *Expert Opin. Drug Discovery* **2007**, *2*, 469–488.

(6) Freire, E. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today* **2008**, *13*, 869–874.

(7) Freire, E. A thermodynamic approach to the affinity optimization of drug candidates. *Chem. Biol. Drug Des.* **2009**, *74*, 468–472.

(8) Ferenczy, G. G.; Keseru, G. M. Enthalpic efficiency of ligand binding. *J. Chem. Inf. Model.* **2010**, *50*, 1536–1541.

(9) 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.

(10) 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.

(11) Nissink, J. W. Simple size-independent measure of ligand efficiency. *J. Chem. Inf. Model.* **2009**, *49*, 1617–1622.

(12) Velazquez Campoy, A.; Freire, E. ITC in the post-genomic era...? Priceless. *Biophys. Chem.* **2005**, *115*, 115–124.

(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) *Spotfire*. TIBCO Software Inc., 2000−2010.

(15) DeLorbe, J. E.; Clements, J. H.; Teresk, M. G.; Benfield, A. P.; Plake, H. R.; Millspaugh, L. E.; Martin, S. F. Thermodynamic and structural effects of conformational constraints in protein-ligand interactions. Entropic paradoxy associated with ligand preorganization. *J. Am. Chem. Soc.* **2009**, *131*, 16758–16770.

(16) Dunitz, J. D. Win some, lose some: Enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol.* **1995**, *2*, 709–712.

(17) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Entropy-Enthalpy Compensation in Solvation and Ligand Binding Revisited. *J. Am. Chem. Soc.* **1998**, *120*, 4526–4527.

(18) Lafont, V.; Armstrong, A. A.; Ohtaka, H.; Kiso, Y.; Mario Amzel, L.; Freire, E. Compensating enthalpic and entropic changes hinder binding affinity optimization. *Chem. Biol. Drug Des.* **2007**, *69*, 413–422.

(19) Chen, W.; Chang, C. E.; Gilson, M. K. Calculation of cyclodextrin binding affinities: Energy, entropy, and implications for drug design. *Biophys. J.* **2004**, *87*, 3035–3049.

(20) Gilson, M. K.; Zhou, H.-X. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.

(21) Zhou, H. X.; Gilson, M. K. Theory of free energy and entropy in noncovalent binding. *Chem. Rev.* **2009**, *109*, 4092–4107.

(22) Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; deSolms, S. J. A priori prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site. *J. Med. Chem.* **1995**, *38*, 305–317.

(23) Ferrari, A. M.; Degliesposti, G.; Sgobba, M.; Rastelli, G. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Bioorg. Med. Chem.* **2007**, *15*, 7865–7877.

(24) 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.