



# Is there a link between selectivity and binding thermodynamics profiles?

Ákos Tarcsay<sup>1</sup> and György M. Keserü<sup>2</sup>



<sup>1</sup>Discovery Chemistry, Gedeon Richter Plc, 19-21 Gyömrői út, H-1103 Budapest, Hungary

<sup>2</sup>Research Centre for Natural Sciences, Hungarian Academy of Sciences, 2 Magyar tudósok krt, H-1117 Budapest, Hungary

**Thermodynamics of ligand binding is influenced by the interplay between enthalpy and entropy contributions of the binding event. The impact of these binding free energy components, however, is not limited to the primary target only. Here, we investigate the relationship between binding thermodynamics and selectivity profiles by combining publicly available data from broad off-target assay profiling and the corresponding thermodynamics measurements. Our analysis indicates that compounds binding their primary targets with higher entropy contributions tend to hit more off-targets compared with those ligands that demonstrated enthalpy-driven binding.**

## Introduction

High on-target affinity and designed selectivity against off-targets are usually the key points in the target product profile of many discovery programs and, consequently, these are among the most desired objectives of multiparameter medicinal chemistry optimizations. Potency optimizations are generally carried out by introducing apolar or polar substituents and subsequently monitoring the binding affinity (expressed in  $K_i$  or  $IC_{50}$  values). High specificity, however, does not demand high affinity [1]. Improving the binding affinity can be achieved by enthalpy- or entropy-driven optimization that covers substantially different thermodynamic profiles. Apparently, enthalpy and entropy changes are linked by the widely observed enthalpy–entropy compensation, although its impact has been recently challenged [2]. Binding affinity shows the quantity of the ligand–protein interactions via the Gibbs free energy of binding, whereas the corresponding thermodynamic profile describes the quality of the interactions.

The relationship between the knowledge encoded in Gibbs free energy of binding and its components, enthalpy and entropy, can be explained by the analogy of the projection. Constellations of stars such as the Cassiopeia are plane projections having a graph-type pattern. Stars, however, are not located at the same distance, some stars are much closer than the others and therefore the projection has a hidden dimension. Taking this distance

dimension into account makes the plane a 3D object. Constellations were used for efficient navigation for hundreds of years, and improving the binding free energy drove medicinal chemistry programs in the past decades. Space travelers, however, should use the information from the third dimension for successful navigation and, similarly, thermodynamic profiles provide beneficial information on the interactions for medicinal chemists.

Ligand–protein interactions involve attractive forces and hydration effects. Properly positioned polar groups contribute to specific interactions, such as H-bonds, salt-bridges, polar–polar interactions and nonclassical interactions such as  $\sigma$ -hole-mediated halogen bonding that result in enthalpy gain. To exploit this enthalpy reward the binding partners should be in optimal orientation, because the binding energy is highly sensitive to the distance and the angle of the interacting atoms [3,4]. Nonpolar groups typically form weaker, less-oriented and less-specific interactions such as van der Waals contacts and  $\pi$ – $\pi$  stacking [5]. Changes in desolvation entropy are favorable in both cases, but the desolvation of polar groups is associated with unfavorable desolvation enthalpy. For example, the desolvation enthalpies of OH and NH functionalities are 36.4 kJ/mol and 33.0 kJ/mol, respectively – in the range of the enthalpy gain realized with polar interactions – whereas for a methyl group the corresponding value is only 2.4 kJ/mol [6,7]. If the interactions with binding-site water molecules do not override the primary ligand–protein interactions the affinity gain achieved by the introduction of polar groups is generally enthalpy biased.

Corresponding author: Keserü, G.M. ([gy.keseru@ttk.mta.hu](mailto:gy.keseru@ttk.mta.hu))

By contrast, introduction of nonpolar substituents typically results in entropic reward that is mainly mediated by desolvation effects. Suboptimal positioned polar moieties would not be exploited in terms of enthalpy gain. The positional sensitivity of enthalpic optimization can be exemplified with a HIV-1 protease inhibitor pair. Saquinavir and TMC126 have the same number of polar groups; however saquinavir binding is associated with unfavorable ( $\sim 5$  kJ/mol) binding enthalpy but TMC126 binding is significantly more enthalpic ( $\Delta H \sim -50$  kJ/mol) owing to the better orientation of its polar groups [7]. It should also be noted that binding-site water molecules have a complex influence on thermodynamics signatures [8–10]. Orientation of polar groups largely influences their specific interactions compared with nonpolar functional groups that are introduced to fill apolar cavities. The latter types of interactions show less dependence on distance and are less sensitive to orientation. As a result, optimization of binding affinity is more straightforward by hydrophobic moieties. Noninteracting or suboptimal positioned polar atoms are charged by the unfavorable desolvation enthalpy and thus generally result in decreased affinity. Accordingly, enthalpy-driven optimization is considered to be significantly more challenging compared with the entropy-driven process. Favorable binding energy can be achieved by entropy-driven approaches such as the introduction of nonpolar groups around apolar protein surfaces.

Replacement of unstable water molecules within hydrophobic pockets is mostly driven by entropy changes, although enthalpy gain coupled with water replacement by apolar moieties has also been reported [8,9]. The effect of binding-site waters has been recently reviewed by using WaterMap for solvation energetic calculations [11]. Selectivity between dopamine D2 and D3 receptors and kinase targets was also successfully rationalized by the analysis of binding-site water molecules [12,13]. Therefore, computational approaches can significantly facilitate the design of selective compounds, if high-quality crystal structures are available. Furthermore, the combination of experimental and computational approaches can rationalize unique cases where apolar

contacts contribute to the favorable binding enthalpy in a protein-binding site occluded from solvent water [14].

The quality of interactions and the accompanying binding thermodynamics profile impact selectivity against off-targets [15]. Enthalpically optimized compounds possess carefully positioned ligand-binding-site atom pairs to achieve the desired gain in binding enthalpy. Considering a different binding pocket presented in an off-target protein, the designed interactions will not be able to yield the enthalpic contribution to binding free energy because of the improper orientation of the ligand. Since the very same desolvation penalty of the polar atoms must be paid, the off-target affinity of the ligand will be limited. By contrast, entropically optimized compounds have fewer positional constraints and desolvation of the apolar moieties can result in entropy gain as a result of the lower dependence from the binding environment. These compounds have therefore higher propensity to form attractive interactions with off-targets. In this review, we investigate this hypothesis by analyzing the thermodynamic and selectivity profiles of optimized compounds and marketed drugs.

### Binding thermodynamics and selectivity optimization HIV-1 protease

The relationship between the binding thermodynamics properties of a closely related pairs of compounds published by Kawasaki and Freire serves as an illustrative example of the impact of thermodynamics on selectivity [15]. The thermodynamics profile was measured on the primary target HIV-1 protease, and cathepsin D and pepsin were monitored as antitargets. In the first case, a subtle change such as the introduction of two methyl groups into a phenyl moiety resulted in  $-11.2$  kJ/mol gain in binding free energy owing to the more favorable enthalpy contribution of the methylated derivative (Fig. 1). This effect is a result of the optimal occupancy of a small cavity around the aryl moiety that is well oriented and the methyl groups can form desirable contacts. The selectivity toward pepsin and cathepsin D increased from 12 to 157 and 72 to 2464, respectively. In the second pair the thioether moiety was replaced by the sulfonyl-methyl group that resulted in a 1.2 kJ/mol decrease in binding free energy. However, the binding enthalpy improved from  $-34.3$  kJ/mol to  $-50.6$  kJ/mol, and the entropy contribution decreased by 11.2 kJ/mol. The introduced sulfonyl group establishes a strong H-bond with Asp30 of the protease, as is evident in the crystal structure. The selectivity against pepsin and cathepsin D increased by seven and ninefold, respectively. The authors suggested that maximal selectivity can be achieved by introducing a few very strong H-bonds toward the primary target protein. H-bonds have very rigorous distance and angular constraints. Consequently, suboptimal H-bonds formed with the off-target protein are penalized and this results in a larger decrease in the corresponding binding free energy. The overall picture of the four compounds suggests that, as the enthalpy contribution to binding free energy is increased, the compounds are more specific to the primary target. It is interesting to note that, among these four compounds, the highest affinity one does not have the highest selectivity; instead it is the one with the most favorable binding enthalpy. Although there is no theoretical background to support the linear correlation between these quantities, linear correlation coefficients ( $r^2$ ) between  $\Delta H_{\text{protease}}$  and  $\Delta\Delta G$  values obtained for pepsin and cathepsin D were significant

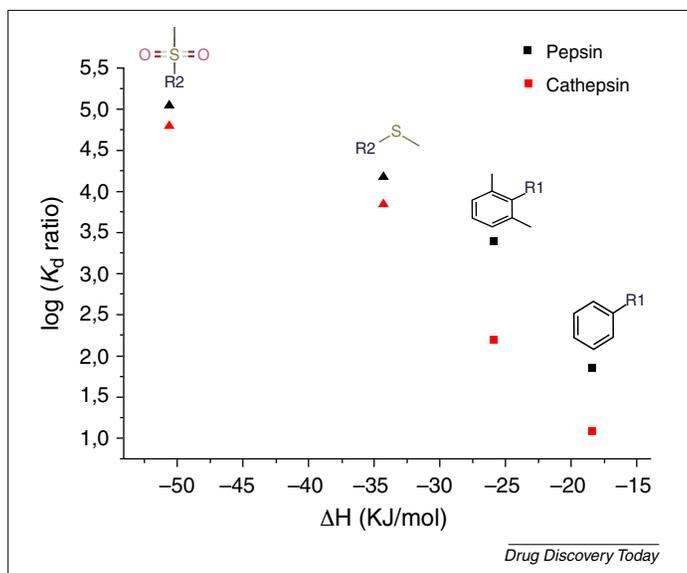


FIGURE 1

Correlation between binding-free-energy difference and binding enthalpy for HIV-1 protease inhibitors.

(0.9 and 0.93, respectively).  $\Delta G_{\text{protease}}$  and  $\Delta\Delta G$  values were somewhat lower being 0.87 and 0.77, for pepsin and cathepsin D, respectively.

#### Matrix metalloproteinase

Matrix metalloproteinase (MMP)12 inhibitors were optimized by using X-ray crystallography and thermodynamics measurements [16] while monitoring selectivity against MMP13. The highest selectivity was 5 kJ/mol in terms of  $\Delta G$  that was achieved by the most enthalpic compound ( $\Delta H = -40.4$  kJ/mol) and, again, this was not the highest-affinity compound considering the MMP12 target. The linear correlation coefficient ( $r$ ) between the  $\Delta G_{\text{MMP12}}$  and  $\Delta\Delta G_{\text{MMP12-MMP13}}$  values was found to be  $-0.15$  ( $P = 0.85$ ), whereas the linear correlation coefficient between the  $\Delta H_{\text{MMP12}}$  and  $\Delta\Delta G_{\text{MMP12-MMP13}}$  is higher:  $-0.68$  ( $P = 0.32$ ). Observations on this limited congeneric ligand set further strengthen the view that not the affinity but the binding enthalpy has a higher contribution to off-target selectivity.

#### Aldose reductase

Aldose reductase (ALR2) is a promising therapeutic target to prevent late complications of diabetes. An optimal drug candidate should possess a high level of selectivity for ALR2 over the related aldehyde reductase (ALR1) [17]. To obtain a comprehensive overview of the binding event, X-ray crystallography and thermodynamics measurements were carried out for ALR2. Based on human X-ray structures published, the six studied ligands can be grouped into those penetrating to the specificity pocket and others leaving the specificity pocket closed. According to the X-ray structures, the favorable enthalpy of IDD393 might be the result of specific polar contacts between the nitro moiety of the ligand and the Ser302 residue. This was further strengthened by the Ser302Arg mutation, where the binding enthalpy of IDD393 dropped by 19.4 kJ/mol. Similarly, comparing the structures of sorbinil and fidarestat revealed that sorbinil possesses a strong H-bond with the backbone NH of Leu300, not present in the case of fidarestat. This observation has been further supported by the decreased enthalpic contribution of sorbinil binding by 8.2 kJ/mol in the Leu300Pro mutant protein. In both cases, the increased selectivity against ALR1 might be the result of the specific enthalpic contact with the primary target that is less than optimal in the case of the antitarget ALR1.

By contrast, we did not find correlation between selectivity and binding enthalpy for the whole dataset. This might be explained by the different physicochemical profiles of the two ligand classes. Ligands occupying the selectivity pocket are carboxylic acids, and share similar pharmacophore features, whereas those bound outside the pocket (e.g. sorbinil and fidarestat) are small, weak NH acids with fewer pharmacophore elements compared with compounds in the first group.

#### Thrombin

Medicinal chemistry optimization of triazole- and tetrazole-containing sulphonamide-type thrombin inhibitors was published by Siles *et al.* [18]. The most promising compound had 828-times selectivity against trypsin. Evaluation of tryptase and chymase selectivity revealed a favorable selectivity profile, because 0.95 and 0.98 residual enzyme activities were determined in the presence of 10  $\mu\text{M}$  inhibitor. Thermodynamics profiles of thrombin

binding, the primary target were measured by isothermal titration calorimetry (ITC). The highly selective lead compound was found to bind thrombin with favorable ( $-38.1$  kJ/mol) enthalpy contribution and unfavorable (3.3 kJ/mol) entropy. Owing to the significant selectivity against other related human serine proteases and the encouraging thermodynamic profile, this lead compound serves as a high-quality starting point for further optimization. Although, in this study, only the top-ranked compound was thermodynamically characterized, the high selectivity and the corresponding enthalpy-driven binding fits into the proposed relationship of these properties.

#### Cannabinoid receptors

Binding thermodynamics of agonists and antagonists of cannabinoid (CB)1 and CB2 receptors were determined by van't Hoff analysis [19] – a methodology that is generally considered to be less reliable than ITC measurements and highly influenced by the heat capacity change of the system. In this case, similar to many G-protein-coupled receptor (GPCR) targets, binding thermodynamics separates the ligands by their functional activity. Binding of the five agonist compounds was entirely entropy driven, whereas binding of the three antagonists was mainly realized by favorable enthalpy contribution. The relationship between thermodynamics signatures and functional activities is a topic of a high number of studies [20] and is therefore beyond the scope of the present paper. Among the two investigated CB receptors, the highest affinity target was evaluated as a primary target and the lowest affinity target as a secondary one. Selectivity was defined as the difference in binding free energy between the primary and the secondary target. The enthalpy contribution calculated for the primary target was compared to the selectivity observed resulting in linear correlation coefficient ( $r$ ) of  $-0.81$  ( $P = 0.016$ ) for all of the eight ligands. Accordingly, ligands possessing more-favorable binding enthalpy contribution had higher selectivity regarding the two CB receptors. Among the eight ligands, ACEA, 2-Fl-AEA and CP55940 had different physicochemical profiles and pharmacophore sets. ACEA and 2-Fl-AEA are arachidonic acid derivatives, whereas CP55940 is an octane derivative. Therefore, their binding mode might be significantly different from that of the remaining five ligands which might explain their distinct thermodynamics profiles. Leaving these ligands out resulted in higher linear correlation coefficient ( $r = -0.95$ ,  $P = 0.012$ ) between the  $\Delta\Delta G$  and  $\Delta H$  values. According to these observations, the relationship between selectivity and enthalpy contribution seems to be valid on GPCRs with thermodynamics profiles obtained from van't Hoff analysis for compounds having a similar binding mode.

#### Nucleic acid binding

A specific DNA aptamer that recognized L-argininamide was discovered by systematic evaluation of ligands by exponential enrichment (SELEX) [1]. This construct had approximately 100-fold selectivity for L-argininamide over several other arginine analogs and amino acids. The highly selective binding of L-argininamide to the DNA was found to be relatively weak:  $-21.3$  kJ/mol only. However, the binding was accompanied by a large, favorable enthalpy, in the range of  $-36$  to  $-38$  kJ/mol, and unfavorable binding entropy, in the range of 15–17 kJ/mol. Based on this observation, it was concluded that high specificity can be achieved

without high affinity, if the binding is mediated by a large enthalpy contribution.

Thermodynamics signatures of amiloride binding to an abasic (AP site) site in RNA and DNA has been reported recently [21]. Despite the typically promiscuous binding of aminoglycoside antibiotics to various RNA targets, amiloride was found to bind strongly and selectively to an AP site of the RNA duplex. The thermodynamics measurements on AP-RNA revealed that the  $-45.2$  kJ/mol binding free energy is composed of favorable ( $-69.0$  kJ/mol) enthalpy and unfavorable ( $23.8$  kJ/mol) entropy contributions. Interestingly, the amiloride affinity to AP-DNA was 78-times lower:  $-34.3$  kJ/mol. Such remarkable preference of amiloride binding to RNA relative to DNA is characteristic compared to typical small DNA-binding ligands [21]. Analysis of the binding thermodynamics data measured for RNA and DNA revealed that the unique selectivity is associated with a large ( $10.9$  kJ/mol) enthalpy difference, whereas the entropy contribution was almost equivalent. As a consequence, selectivity between the two targets was explicitly the result of the enthalpy change.

#### Ligand-binding protein

Antidigoxigenin antibodies are administered to remove overdosed digoxin, which has a narrow therapeutic window. A very interesting approach was published recently in which protein-binding sites were computationally designed to bind digoxigenin (DIG) [22]. The binding pocket was engineered *in silico* to have specific, energetically favorable H-bonds and van der Waals interactions along with high overall shape complementarity. The best construct obtained was able to bind DIG with extremely high affinity ( $541$  pM), similar to that of antidigoxin antibodies. ITC and X-ray crystallography revealed that the best host site can form three specific H-bonds with the ligand that is associated with favorable ( $\sim -45.3$  kJ/mol) enthalpy and favorable ( $\sim -7.5$  kJ/mol) entropy. The designed protein has 29-, 372- and 3216-fold DIG preference against structurally highly similar compounds: digitoxigenin, progesterone and  $\beta$ -oestradiol, respectively. To assess the role of specific H-bonds with DIG ligand on selectivity the Tyr101 H-bond-donor interacting with DIG C-ring oxygen was mutated to Phe. This mutation resulted in lower affinity for DIG ( $K_f$ :  $39$  nM) and reduced selectivity  $\sim 0.1$ -,  $0.8$ - and  $43$ -fold for digitoxigenin, progesterone and  $\beta$ -oestradiol, respectively. Mutation of the other important H-bonding partner, Tyr43 to Phe, also decreased the affinity for DIG ( $59$  nM) and the selectivity against the three investigated compounds (DIG had  $12$ -,  $1.3$ - and  $254$ -fold selectivity against digitoxigenin, progesterone and  $\beta$ -oestradiol, respectively). These experiments confirmed that the selectivity of the artificial ligand-binding protein for DIG was conferred through the designed H-bonding interactions. Therefore, this study is a plausible example for the role of highly oriented specific interactions that results in favorable binding enthalpy and high specificity. Considering medicinal chemistry aspects, designing ligands that can form optimal, enthalpically favorable H-bonds with the target-binding site might similarly result in high specificity against different binding environments having lower complementarity and suboptimal H-bonding pattern.

#### Thermodynamic profile of marketed drugs

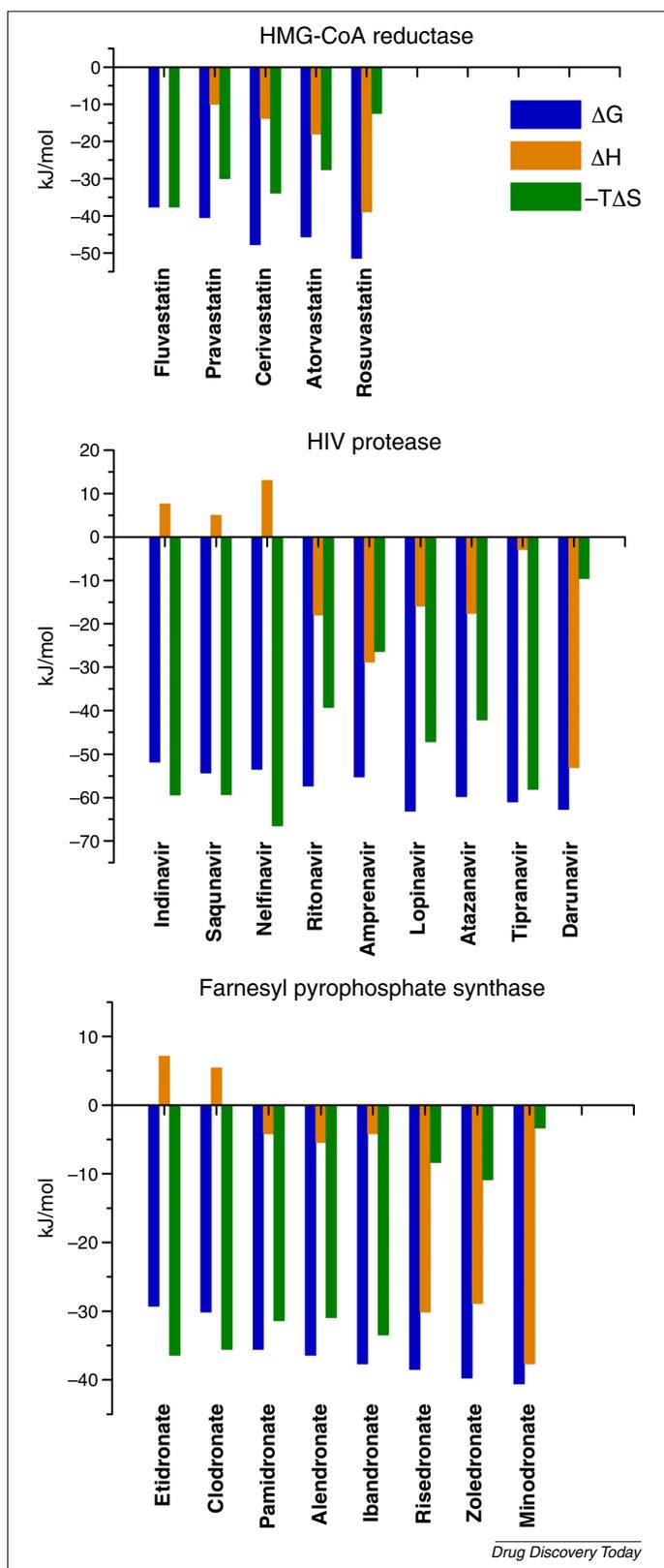
A retrospective analysis of thermodynamic signatures of marketed drugs was published for a series of statins binding to 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase, small-molecule

inhibitors of HIV protease [10,23] and bisphosphonate inhibitors for farnesyl pyrophosphate synthase (FPPS) [24]. Investigating the compounds in terms of their market entry time revealed a clear trend regarding the enthalpy contribution of the binding free energy (Fig. 2).

Binding to their primary target, first-in-class drugs were typically entropy driven, with positive or negligible enthalpy contribution in all the three cases. During the time-course of drug evolution, as novel drugs with the same mechanism of action are introduced to the market, they must show advantages over the predecessors. Those drugs that were better than the previous compounds in the class can be characterized with sequentially increasing binding enthalpy contribution. In the case of statins, for the first-in-class fluvastatin the binding enthalpy is  $0$  and the binding entropy is  $-37.6$  kJ/mol, whereas for rosuvastatin – the last entry in this analysis – the corresponding values are  $-38.9$  kJ/mol and  $-12.5$  kJ/mol, for enthalpy and entropy, respectively. Considering the protease inhibitors, indinavir binding is realized by  $-59.4$  kJ/mol favorable entropy and the enthalpy contribution is  $7.6$  kJ/mol. By contrast, darunavir binding is mainly enthalpy driven ( $-53.1$  kJ/mol), and the entropy contribution is  $-9.6$  kJ/mol only. The tendency is also obvious for bisphosphonate FPPS inhibitors. Etidronate binding is entirely entropy driven, with  $-36.4$  kJ/mol, associated with unfavorable ( $7.1$  kJ/mol) enthalpy. By contrast, minodronate binding is enthalpy driven ( $-37.7$  kJ/mol) and the entropy contribution is  $-3.3$  kJ/mol only. The underlying reasons for the success of the more-enthalpic compounds might include better physicochemical parameters, such as lipophilicity and solubility, arising from the more-polar moieties that are necessary for favorable enthalpic binding. However, we should emphasize that these drugs were optimized without controlling their binding thermodynamics and obviously there are a vast amount of other parameters influencing the success of the given drug. Although we are aware of the limitations of thermodynamics signatures, increasing enthalpy contribution to binding free energy during drug evolution in a given class is significant.

The most common adverse effect associated with statins is myalgia that is thought to be linked to myocyte HMG-CoA reductase inhibition. One possibility to avoid this side-effect is the hepatocyte-selective distribution of statins that can be achieved by decreasing passive cell permeability by increasing the hydrophilicity. In this respect, the most enthalpic (i.e. rosuvastatin) has the highest hydrophilicity and hepatoselectivity. Myocyte  $IC_{50}$  is  $926$ -times higher compared with the hepatocyte value [25]. Hepatoselectivity of less-enthalpic binders such as pravastatin, atorvastatin and cerivastatin is somewhat lower; these drugs show  $444$ -,  $144$ - and  $4.1$ -fold selectivity, respectively. Therefore monitoring binding thermodynamics becomes an integrated part of statin optimizations [26]. Optimizing the atorvastatin scaffold, a promising lead compound was derived possessing  $-74$  kJ/mol enthalpy contribution and higher than  $1000$ -fold hepatic selectivity.

Selectivity against off-targets can be another reason for the success of the best-in-class compounds. In the case of HIV protease inhibitors it was shown that more-enthalpic compounds possess higher adaptability to drug-resistant mutants along with more-enhanced selectivity to the off-target cathepsin D [27,28]. Interestingly, no correlation was observed between adaptability and binding affinity, however a reproducible correlation was found



**FIGURE 2**  
Thermodynamic profile of marketed drugs.

between the logarithm of the corresponding  $K_d$  ratio and the proportion of binding enthalpy contribution to the binding affinity in the wild-type protease [27]. The benefit of enthalpic optimization was also shown during the evaluation of indinavir,

nelfinavir, saquinavir, ritonavir, KNI764 and KNI272 in terms of the inhibition of resistant mutant V82F/I84V. In this case, compounds that form optimized H-bonds and, thus, realized more-enthalpic binding to the wild-type HIV protease show higher residual inhibition against the resistant mutant of the entropy-driven binders. Here, specific and enthalpically favorable contacts provided more-desirable adaptability [29].

### Selectivity profile analysis of marketed drugs

Our evaluation on the link between selectivity and binding thermodynamics profiles indicated that enthalpy-driven compounds show typically higher selectivity than entropy-driven analogs with a similar binding mode to primary targets. To expand the domain of the selectivity–thermodynamics relationship, our intention was to challenge this hypothesis on broad selectivity profiles. Therefore, we have collected examples possessing binding thermodynamics data on the primary target and these were subjected to broad off-target assay profiling. The thermodynamics data were collected from literature including BindingDB, PDBCal and Scorpio databases [30]. *In vitro* profiling data were collected from the coherent DrugMatrix database available through ChEMBL (<https://ebi.ac.uk/chembl/db>) and the CerepBioprint (Cerep) profile [31]. We have selected these databases to ensure that the compounds were tested on the same number of assays within the same laboratory conditions to minimize the noise originating from the inter-laboratory differences. By contrast, the thermodynamics data pooled from various sources might include inter-laboratory errors of thermodynamics measurements.

Nineteen compounds acting on six targets fulfilled our criteria (Table 1). Five compounds are HIV-1 protease inhibitors, five are dopamine D2 receptor ligands, four are HMG-CoA reductase inhibitors, two act on histamine H1 receptors, two compounds are beta-blockers and, finally, we included a DNA-gyrase inhibitor. HIV-1 protease and HMG-CoA reductase inhibitors were evaluated in terms of their thermodynamics profiles [10,23]. It was proposed that favorable interplay between enthalpy and entropy is reflected in their progress to the market. It was suggested that enthalpically more-favorable drugs tend to be the best-in-class compounds, whereas the entropy-driven binders represent pioneer, first-in-class drugs. In accordance, evaluating the enthalpic efficiency as the measure of ligand–protein complementarity has also been discussed in the literature [10,32].

First, promiscuity and thermodynamics relationships were evaluated within target groups. Table 1 shows that binding of three HIV-1 protease inhibitors: nelfinavir, indinavir and saquinavir, is entropy driven (see supplementary material Table S1 online for more detail). Ritonavir binding is also entropy driven, but the enthalpy contribution is more favorable than that for the first group. Amprenavir binding is characterized by balanced entropy–enthalpy contributions. The change from entropy-driven binding to a more-balanced thermodynamic profile is also reflected in the selectivity profile. Amprenavir hits only two targets out of ~134 involved in the DrugMatrix panel and three out of 185 on Cerep. By contrast, saquinavir hits 11 targets on the DrugMatrix assay panel and 18 on the Cerep panel.

The enthalpies of the five drugs acting on the dopamine D2 receptor possess significant  $-0.92$  ( $P = 0.026$ ) linear correlation coefficient ( $r$ ) with the number of hit targets on the Cerep profile.

TABLE 1

Compounds with thermodynamic and broad specificity assay profiles. Calculated octanol–water partition coefficient (AlogP), heavy atom count (Nh) and molecular weight (Mw) are indicated.

	Drug	Target	Drug matrix	Drug matrix $\Delta\Sigma G_{\text{off}}$	Cerep	$\Delta G^a$	$\Delta H^a$	$-T\Delta S^a$	AlogP	Nh	Mw	Refs
1	Nelfinavir	HIV-1 protease	7	-204		-53.5	13.0	-66.5	5.3	40	567.8	[10]
2	Indinavir	HIV-1 protease	3	-102	7	-51.8	7.6	-59.4	3.1	45	613.8	[10]
3	Saquinavir	HIV-1 protease	11	-338	18	-54.3	5.0	-59.3	3.7	49	669.9	[10]
4	Ritonavir	HIV-1 protease	8	-277	15	-57.3	-18.0	-39.3	5.0	50	720.9	[10]
5	Amprenavir	HIV-1 protease	2	-64	3	-55.2	-28.8	-26.4	2.4	35	506.6	[10]
6	Flupenthixol	Dopamine D2			52	-47.7	15.2	-62.9	4.82	30	434.52	[33]
7	Haloperidol	Dopamine D2	18	-682	27	-53.2	-12.8	-40.4	3.76	26	375.9	[33]
8	Alizapride	Dopamine D2			13	-42.3	-50.8	8.6	1.87	23	315.4	[33]
9	Metoclopramide	Dopamine D2	6	-214	19	-41.4	-54.8	13.4	1.78	20	299.8	[33]
10	Sulpiride	Dopamine D2	2	-76	10	-41.9	-88.6	46.7	0.7	23	341.4	[33]
11	Fluvastatin	HMG-CoA reductase	1	-36	5	-37.6	0.0	-37.6	4.2	30	411.5	[10]
12	Cerivastatin	HMG-CoA reductase	0	0	4	-47.7	-13.8	-33.9	4.2	33	459.6	[10]
13	Pravastatin	HMG-CoA reductase	0	0	1	-40.5	-10.5	-30.0	2.2	30	424.5	[10]
14	Atorvastatin	HMG-CoA reductase	0	0		-45.6	-18.0	-27.6	5.6	41	557.6	[10]
15	Clozapine	Histamine H1	26	-1100	44	-47.9	72.0	-119.9	3.42	23	326.8	[34]
16	Diphenhydramine	Histamine H1	11	-398	29	-43.6	22.6	-66.2	3.38	19	255.4	[34]
17	Pindolol	Beta-blocker	1	-43	9	-49.6	-21.3	-28.3	1.93	18	248.3	[35]
18	Isoproterenol	Beta-blocker	1	-33	5	-50.2	-143.2	92.9	1.1	15	211.6	[35]
19	Novobiocin	DNA gyrase			0	-42.7	-51.8	9.2	3.45	44	612.6	[30]
Linear correlation	Drug matrix ( $n = 16$ )	1.00	-0.99 ( $<0.001$ ) <sup>b</sup>	0.95	-0.26	0.55 (0.026) <sup>b</sup>	-0.58 (0.020) <sup>b</sup>	0.19	-0.07	-0.06		
	Drug matrix $\Sigma\Delta G$ ( $n = 16$ )	-0.99 ( $<0.001$ ) <sup>b</sup>	1.00	-0.95 ( $<0.001$ ) <sup>b</sup>	0.21	-0.55 (0.027) <sup>b</sup>	0.57 (0.022) <sup>b</sup>	-0.16	0.13	0.12		
	Cerep ( $n = 17$ )	0.95 ( $<0.001$ ) <sup>b</sup>	-0.95 ( $<0.001$ ) <sup>b</sup>	1.00	-0.12	0.54 (0.025) <sup>b</sup>	-0.54 (0.023) <sup>b</sup>	0.38	-0.20	-0.19		
Spearman rank correlation	Drug matrix ( $n = 16$ )	1.00	-0.99 ( $<0.001$ ) <sup>b</sup>	0.92 ( $<0.001$ ) <sup>b</sup>	-0.42	0.47	-0.63 (0.009) <sup>b</sup>	0.10	0.04	0.05		
	Drug matrix $\Sigma\Delta G$ ( $n = 16$ )	-0.99 ( $<0.001$ ) <sup>b</sup>	1.00	-0.92 ( $<0.001$ ) <sup>b</sup>	0.37	-0.45	0.61 (0.013) <sup>b</sup>	-0.06	0.00	0.00		
	Cerep ( $n = 17$ )	0.92 ( $<0.001$ ) <sup>b</sup>	-0.92 ( $<0.001$ ) <sup>b</sup>	1.00	-0.15	0.46	-0.57 (0.019) <sup>b</sup>	0.20	-0.26	-0.24		

<sup>a</sup>Thermodynamic data are presented in kJ/mol.

<sup>b</sup>P-values are indicated in parentheses. Statistically significant correlations are marked with red color.

<sup>a</sup> Thermodynamic data are presented in kJ/mol.

<sup>b</sup> P-values are indicated in parentheses. Statistically significant correlations are marked with red color.

The entropy-driven binding of flupenthixol is translated into high promiscuity, hitting 52 targets (Table 1). By contrast, the enthalpy-driven binding of sulpiride highlights the enhanced complementarity to the target-binding site, and results in significantly reduced promiscuity. Ligands of the dopamine D2 target show univocal tendencies on the DrugMatrix and the Cerep profile.

In the case of HMG-CoA reductase inhibitors, fluvastatin binding is entropy driven, whereas cerivastatin, pravastatin and atorvastatin binding have increased enthalpy contribution. Accordingly, fluvastatin, cerivastatin and pravastatin hit five, four or one target on the Cerep assay panel, respectively. The increasing selectivity is in-line with the entropy–promiscuity relationships, because the decreasing binding entropy results

in lower promiscuity. Binding of the histamine H1 ligands is entropy driven. Accordingly, clozapine and diphenhydramine are highly promiscuous compounds hitting 26 and 11 targets on DrugMatrix, and 44 and 29 targets on Cerep profile, respectively.

Thermodynamics signatures of beta-blockers revealed that pindolol binding is balanced in terms of enthalpy and entropy contributions, whereas isoproterenol binding is entirely enthalpy driven. Thermodynamics profiles are in-line with their medium and low promiscuity, respectively. Our last example is novobiocin, a selective compound characterized by enthalpy-driven binding and correspondingly no off-target activity on the Cerep panel. This compound is specific, with no promiscuity issue reported.

Next, we investigated the whole dataset that represents broad chemical diversity and spans six targets. To assess the relationship between thermodynamics profiles and observed hit rates, linear and rank correlation coefficients were calculated. Sum of ranking differences (SRD) values were also calculated (see supplementary material Tables S2,3 online) [36]. Although differences in binding-site characteristics and measurement conditions might impact the results of this analysis, we found that compounds hitting higher numbers of targets have more-remarkable entropy and typically less-favorable enthalpy contributions. It is worth mentioning that higher affinity achieved by entropy-driven optimization does not necessarily result in high selectivity (significant negative correlation coefficients), in contrast to lower affinity but higher enthalpy contributions. Lopinavir, atazanavir and amprenavir clearly exemplify this statement. Results acquired for this limited dataset revealed that binding enthalpy and entropy tend to correlate with the number of off-targets (Table 1; see also supplementary Figure S1 online). In the case of SRD calculations the entropy-based ranking was found to be significantly different from a random distribution, with *P* value of 0.017 and 0.016 for DrugMatrix and Cerep datasets. More-favorable binding enthalpy and less-favorable binding entropy might result in higher specificity and lower promiscuity. Enthalpy and entropy correlation was found to be  $-0.99$  ( $P < 0.001$ ), in accordance with the phenomenon of entropy–enthalpy compensation. Although correlations were found to be statistically significant, it is worth mentioning that they are generally weak and were obtained from a limited dataset preventing its over-interpretation. Leaving one compound out generally does not change the correlation coefficients ( $\pm 0.02$ ) except for clozapine. Linear correlation between enthalpy and promiscuity without clozapine drops to 0.32 and 0.39 for DrugMatrix and Cerep data, respectively.

To compare continuous variables we calculated the sum of binding energies on off-targets using the activity values presented

in DrugMatrix (Table 1). In contrast to the simple sum of targets hit, the summed affinity avoids biases owing to high hit rates with limited affinities compared with low hit rates with high affinity. Because the DrugMatrix database contains equivalent numbers of assay results for each ligand, binding energies can be summed. This analysis resulted in similar tendencies to that of the hit target type assessment.

The quality of optimized compounds is usually quantified by changes in the physicochemical profile. In terms of these parameters, enthalpy-driven optimization is generally preferred over entropy-driven optimization [10,15,30,32,37,38]. Ligand promiscuity, as well as thermodynamic signature of binding [39], is also interrelated to physicochemical properties. It was shown that three physicochemical features: basic character, molecular weight and lipophilicity ( $\log P/\log D$ ), have the highest influence on promiscuity [40]. It is interesting to note that lipophilicity ( $\text{Alog}P$ ) generally showed higher correlations with the off-target occurrences compared with molecular size descriptors, such as heavy atom count (Nh) and molecular weight (Mw) (Table 1). However, all of these correlations were statistically nonsignificant for this dataset (Table 1). The general correlation between size and potency is well established [41,42]; in our case the correlation between binding energy and Mw was found to be also remarkable ( $r = -0.45$ ,  $P = 0.052$ ). Linear correlation ( $r$ ) between entropy and lipophilicity ( $\text{Alog}P$ ) was found to be more pronounced:  $r = -0.63$  ( $P = 0.004$ ).

The correlation between binding enthalpy and lipophilic ligand efficiency ( $\text{LLE} = \text{pAct} - \log P$ ) has been recently investigated by Shultz [43]. Therefore we have collected all the cases shown in Tables 1–4 to evaluate the enthalpy–LLE correlation. The linear correlation coefficient ( $r$ ) was found to be  $-0.501$  ( $P = 0.001$ ), resulting in significant correlation for the investigated 37 cases. However correlation using the off-target assay-profiling data collected in Table 1 did not result in significant correlation with DrugMatrix ( $r = -0.003$ ) and Cerep ( $-0.257$ ) sets.

TABLE 2

## Binding thermodynamics data of matrix metalloproteinase (MMP)12 ligands.

Compound	$\Delta G_{\text{MMP12}}$ (kJ/mol)	$\Delta H_{\text{MMP12}}$ (kJ/mol)	$-\Delta S_{\text{MMP12}}$ (kJ/mol)	$\Delta\Delta G_{\text{MMP13-MMP12}}$ (kJ/mol)
1 (1) <sup>a</sup>	-41.1	-38.0	-3.1	1.0
2 (3) <sup>a</sup>	-43.9	-35.6	-8.3	2.3
3 (4) <sup>a</sup>	-49.2	-37.0	-12.2	1.0
4 (6) <sup>a</sup>	-46.2	-40.4	-5.9	5.0

<sup>a</sup>Compound identifier in the original article [16].

TABLE 3

## Binding thermodynamics and selectivity data of aldose reductase (ALR2) ligands.

Compound	$\Delta G_{\text{ALR2}}$ (kJ/mol)	$\Delta H_{\text{ALR2}}$ (kJ/mol)	$-\Delta S_{\text{ALR2}}$ (kJ/mol)	$\Delta\Delta G_{\text{ALR1-ALR2}}$ (kJ/mol)
<b>Selectivity pocket occupant</b>				
Zopolrestat	-46.0	-58.1	12.1	17.7
Compound 2 <sup>a</sup>	-42.5	-48.5	6.0	>23.8 (rat)
IDD388	-42.7	-59.0	16.3	14.9
IDD393	-42.2	-81.2	39.0	21.0 (fluoro derivative)
<b>Non-occupant</b>				
Sorbinil	-37.9	-54.7	16.8	2.4
Fidarestat	-46.7	-79.5	32.8	11.9

<sup>a</sup>Compound identifier in the original article [17].

**TABLE 4**  
**Binding thermodynamics data of cannabinoid (CB) receptor ligands.**

Compound	$\Delta G_{(\text{primary})}$ (kJ/mol)	$\Delta H_{(\text{primary})}$ (kJ/mol)	$-\Delta S_{(\text{primary})}$ (kJ/mol)	$\Delta\Delta G_{(\text{primary-secondary})}$ (kJ/mol)
WIN55212	−46.8 (CB2)	27	−73.8	2.6
JWH015	−41.4 (CB2)	48	−89.4	2.6
ACEA	−47.8 (CB1)	59	−106.8	8.2
2-FI-AEA	−46.1 (CB1)	17	−29.1	6.2
CP55940	−51.2 (CB1)	56	−107.2	3.3
AM630	−41.2 (CB2)	−19	−22.2	8.0
AM281	−45.9 (CB1)	−35	−10.9	13.7
AM251	−48.8 (CB1)	−52	3.2	13.3

Regarding the application domain of the concept of increasing selectivity by thermodynamics optimization, it is important to emphasize that the key point is the rational optimization of the interaction pattern of the protein–ligand complex with specific contacts. The experimentally determined entropy and enthalpy values measure the sum of the changes related to the complex formation, including solvation terms, protein and ligand conformational changes and protein–ligand enthalpic contacts. The flexibility of the protein target and the nature of the binding site also have crucial impact, as exemplified by the entropy–enthalpy transduction theory [44]. Decoupling the protein–ligand entropy and enthalpy contribution is therefore a difficult task that has not been solved entirely to date. Accordingly, it is highly recommended to deploy experimental and computational approaches synergistically to understand the biophysical background of the free energy changes.

### Concluding remarks

The pioneer hypothesis of Kawasaki and Freire [15] has been evaluated here regarding the interplay between selectivity and thermodynamics profiles. According to their study, ligand selectivity can be achieved by favorable binding enthalpy, because it is a straightforward measure of protein–ligand complementarity [10,15]. Highly oriented interactions accomplished by enthalpic interactions result in higher bias toward the primary target, making compounds less promiscuous. We showed that structurally diverse ligands of several validated drug targets (HIV protease, HMG-CoA reductase, D2,  $\beta$ -adrenergic and histamine H1 receptors) possessing broad selectivity profile data support this concept.

The limited number of cases collected and discussed here strengthens the link between selectivity and thermodynamics and facilitates the generation of more thermodynamics data on compounds with wide ranges of selectivity assay data. The objective of the present review is to stimulate further debate supporting or challenging this hypothesis by publishing further experimental data. From the thermodynamics point of view, drugs are acting in an open system; therefore emphasizing that the recent observations were made under equilibrium conditions is crucial. Under physiological conditions the binding kinetics might also influence the selectivity profile realized *in vivo* [45–47].

From a practical point of view, monitoring binding thermodynamics at project milestones can facilitate the selection of the higher-quality compounds. Desirable hit and lead compounds having the highest enthalpy among the chemical series might be in-line with the greater complementarity with the primary binding site, therefore impacting selectivity profile. Owing to the complex nature of the binding event and the difficulties of understanding, the thermodynamics background of the SAR contributes to the obstacles of successful optimization toward higher enthalpy. Structure-based approaches, experimental and computational, have crucial roles in rationalizing these modifications. Based on the current state of our understanding, this hypothesis can be exploited during medicinal chemistry programs optimizing affinity and selectivity in parallel.

### Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.drudis.2014.09.014>.

### References

- Chaires, J.B. (2008) Calorimetry and thermodynamics in drug design. *Annu. Rev. Biophys.* 37, 135–151
- Chodera, J.D. and Mobley, D.L. (2013) Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* 42, 121–142
- Wilcken, R. *et al.* (2013) Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J. Med. Chem.* 56, 1363–1388
- Bissantz, C. *et al.* (2010) A medicinal chemist's guide to molecular interactions. *J. Med. Chem.* 53, 5061–5084
- Zhou, P. *et al.* (2012) Specific noncovalent interactions at protein–ligand interface: implications for rational drug design. *Curr. Med. Chem.* 19, 226–238
- Sergio, C. *et al.* (1981) Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Sol. Chem.* 10, 563–595
- Ruben, A. *et al.* (2006) Overcoming roadblocks in lead optimization: a thermodynamic perspective. *Chem. Biol. Drug Des.* 67, 2–4
- Snyder, P.W. *et al.* (2011) Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17889–17894
- Biela, A. *et al.* (2012) Water makes the difference: rearrangement of water solvation layer triggers non-additivity of functional group contributions in protein–ligand binding. *ChemMedChem* 7, 1423–1434
- Ladbury, J.E. *et al.* (2010) Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat. Rev. Drug Discov.* 9, 23–27
- Yang, Y. *et al.* (2013) Approaches to efficiently estimate solvation and explicit water energetics in ligand binding: the use of WaterMap. *Expert Opin. Drug Discov.* 8, 277–287

- 12 Newman, A.H. *et al.* (2012) Molecular determinants of selectivity and efficacy at the dopamine D3 receptor. *J. Med. Chem.* 55, 6689–6699
- 13 Robinson, D.D. *et al.* (2010) Understanding kinase selectivity through energetic analysis of binding site waters. *ChemMedChem* 5, 618–627
- 14 Barratt, E. *et al.* (2005) Van der Waals interactions dominate ligand–protein association in a protein binding site occluded from solvent water. *J. Am. Chem. Soc.* 127, 11827–11834
- 15 Kawasaki, Y. and Freire, E. (2011) Finding a better path to drug selectivity. *Drug Discov. Today* 16, 985–990
- 16 Bertini, I. *et al.* (2007) Exploring the subtleties of drug–receptor interactions: the case of matrix metalloproteinases. *J. Am. Chem. Soc.* 129, 2466–2475
- 17 Steuber, H. *et al.* (2008) Merging the binding sites of aldose and aldehyde reductase for detection of inhibitor selectivity-determining features. *J. Mol. Biol.* 379, 991–1016
- 18 Siles, R. *et al.* (2011) Synthesis and biochemical evaluation of triazole/tetrazole-containing sulfonamides against thrombin and related serine proteases. *Bioorg. Med. Chem. Lett.* 21, 5305–5309
- 19 Merighi, S. *et al.* (2010) Binding thermodynamics at the human cannabinoid CB1 and CB2 receptors. *Biochem. Pharmacol.* 79, 471–477
- 20 Raffa, R.B., ed. (2001) *Drug-Receptor Thermodynamics: Introduction and Applications*, Wiley
- 21 Sato, Y. *et al.* (2012) Strong and selective binding of amiloride to an abasic site in RNA duplexes: thermodynamic characterization and microRNA detection. *Angew. Chem. Int. Ed. Engl.* 51, 6369–6372
- 22 Tinberg, C.E. *et al.* (2013) Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* 501, 212–216
- 23 Freire, E. (2008) Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* 13, 869–874
- 24 Kawasaki, Y. *et al.* (2014) Thermodynamic evaluation of the binding of bisphosphonates to human farnesyl pyrophosphate synthase. *Chem. Pharm. Bull.* 62, 77–83
- 25 Park, W.K. *et al.* (2008) Hepatoselectivity of statins: design and synthesis of 4-sulfamoyl pyrroles as HMG-CoA reductase inhibitors. *Bioorg. Med. Chem. Lett.* 18, 1151–1156
- 26 Sarver, R.W. *et al.* (2008) Thermodynamic and structure guided design of statin based inhibitors of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *J. Med. Chem.* 51, 3804–3813
- 27 Ohtaka, H. and Freire, E. (2005) Adaptive inhibitors of the HIV-1 protease. *Prog. Biophys. Mol. Biol.* 88, 193–208
- 28 Ohtaka, H. *et al.* (2004) Thermodynamic rules for the design of high affinity HIV-1 protease inhibitors with adaptability to mutations and high selectivity towards unwanted targets. *Int. J. Biochem. Cell Biol.* 36, 1787–1799
- 29 Velazquez-Campoy, A. *et al.* (2001) The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. *Arch. Biochem. Biophys.* 390, 169–175
- 30 Ferenczy, G.G. and Keserü, G.M. (2010) Enthalpic efficiency of ligand binding. *J. Chem. Inf. Model.* 50, 1536–1541
- 31 Xiangyun, W. and Nigel, G. (2012) Comparing measures of promiscuity and exploring their relationship to toxicity. *Mol. Informatics* 31, 145–159
- 32 Ferenczy, G.G. and Keserü, G.M. (2012) Thermodynamics of fragment binding. *J. Chem. Inf. Model.* 52, 1039–1045
- 33 Kilpatrick, G.J. *et al.* (1986) The thermodynamics of agonist and antagonist binding to dopamine D2 receptors. *Mol. Pharmacol.* 30, 226–234
- 34 Wittmann, H.J. *et al.* (2009) Contribution of binding enthalpy and entropy to affinity of antagonist and agonist binding at human and guinea pig histamine H(1)-receptor. *Mol. Pharmacol.* 76, 25–37
- 35 Weiland, G.A. *et al.* (1979) Fundamental difference between the molecular interactions of agonists and antagonists with the beta-adrenergic receptor. *Nature* 281, 114–117
- 36 Kollár-Hunek, K. and Héberger, K. (2013) Method and model comparison by sum of ranking differences in cases of repeated observations (ties). *Chemom. Intell. Lab. Syst.* 127, 139–146
- 37 Ferenczy, G.G. and Keserü, G.M. (2010) Thermodynamics guided lead discovery and optimization. *Drug Discov. Today* 15, 919–932
- 38 Hann, M.M. and Keserü, G.M. (2012) Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discov.* 11, 355–365
- 39 Olsson, T.S. *et al.* (2008) The thermodynamics of protein–ligand interaction and solvation: insights for ligand design. *J. Mol. Biol.* 384, 1002–1017
- 40 Tarcsay, Á. and Keserü, G.M. (2013) Contributions of molecular properties to drug promiscuity. *J. Med. Chem.* 56, 1789–1795
- 41 Gleeson, M.P. (2011) Probing the links between in vitro potency, ADMET and physicochemical parameters. *Nat. Rev. Drug Discov.* 10, 197–208
- 42 Leeson, P.D. and Springthorpe, B. (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 6, 881–890
- 43 Shultz, M.D. (2012) The thermodynamic basis for the use of lipophilic efficiency (LipE) in enthalpic optimizations. *Bioorg. Med. Chem. Lett.* 23, 5992–6000
- 44 Fenley, A.T. *et al.* (2012) Entropy–enthalpy transduction caused by conformational shifts can obscure the forces driving protein–ligand binding. *Proc. Natl. Acad. Sci. U. S. A.* 109, 20006–20011
- 45 Copeland, R.A. *et al.* (2006) Drug–target residence time and its implications for lead optimization. *Nat. Rev. Drug Discov.* 5, 730–739
- 46 Swinney, D.C. (2009) The role of binding kinetics in therapeutically useful drug action. *Curr. Opin. Drug Discov. Devel.* 12, 31–39
- 47 Núñez, S. *et al.* (2012) Target–drug interactions: first principles and their application to drug discovery. *Drug Discov. Today* 17, 10–22