View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository of the Academy's Library

The role of binding thermodynamics in medicinal chemistry optimizations

György G. Ferenczy and György M. Keserű
Research Center for Natural Sciences, Hungarian Academy of Sciences, 2 Magyar tudósok krt. H-1117 Budapest, Hungary

Abstract

Background: Ligand binding thermodynamics has been attracted considerable interest in the past decade owing to the recognized relation between binding thermodynamic profile and the physicochemical and druglike properties of compounds. **Discussion:** • Affinity improvements in drug discovery optimizations can be either enthalpically or entropically driven. In this review the relation between optimization strategies and ligand properties are presented based on the structural and thermodynamic analysis of ligand-protein complex formation. **Conclusions:** The control of the binding thermodynamic profile is beneficial for the balanced affinity and physicochemical properties of drug candidates and early phase optimization gives more opportunity to this control.

Drug discovery optimizations and molecular obesity

The objective of drug discovery projects at the preclinical stage is to find compounds with balanced properties that include high affinity towards the target, sufficient specificity and selectivity and advantageous physicochemical and pharmacokinetic profile. The evolution of the chemical starting point to a clinical candidate is a result of a multiparametric optimization process. While increasing the binding affinity is of primary importance in early optimizations the monitoring and improvement of other druglike properties are also part of the optimization process from the outset and they become even dominant with the advance of the optimization. It is well documented [1,2,3] that affinity improvement tends to increase molecular size and lipophilicity, leading to large, hydrophobic compounds, a phenomenon called molecular obesity [4,5]. This type of compound has a greater chance being promiscuous [2,6,7], having low solubility and suboptimal ADMET properties [8]. Therefore, it is of utmost importance to understand the relationship among the various properties to be optimized simultaneously and it has been increasingly recognized that binding thermodynamics (Box 1) may serve as a link among these properties. In addition to its obvious relationship to binding affinity binding thermodynamics impacts other important druglike parameters such as the physicochemical profile, binding selectivity, specificity and promiscuity. It has been demonstrated that enthalpic compunds have typically better profile of physicochemical parameters than that of the high entropy compounds [9]. Freire and coworkers showed that enthalpic compounds have higher chance being selective against off-targets [10]. Compounds binding their targets with higher entropy contributions tend to hit more off-targets compared to those ligands that had enthalpically-driven thermodynamics profile [11]. Thermodynamic investigations might therefore help both in the understanding of interconnecting relations between molecular properties and also in controlling them.

Affinity optimization is typically realized by introducing new atoms in the molecule. This process is validated by the observation that the available maximal affinity of ligands toward protein targets increases with increasing ligand size as it was first observed by Kuntz et al. [12] and later for a larger data set by Reynolds et al. [13]. The same trend was also identified using affinities from thermodynamic measurements [14]. This already indicates that care has to be taken in affinity improvements in order to avoid excessive size and lipophilicity increase. At this point, it is appropriate to analyze ligand-protein interactions and the process of complex formation to understand the detailed relationship between ligand size and affinity.

Ligand-protein binding is a complex, multistep process that includes the desolvation of the partners, their conformational change and the formation of new interactions between them. An important component of the process is the reorganization of water networks as a consequence of the release of water molecules from the solvation shells of the binding partners. All these steps contribute to the enthalpy and entropy changes that accompany binding and both contributions are affected by the structure of the ligand and the way it binds to the protein. The increasing amount of structural and thermodynamic data makes it possible to analyze how the interactions and the thermodynamic profile depend on various quantities like ligand size, lipophilicity and affinity. This, in turn, allows us better understanding of the evolution of these latter quantities in the course of optimization.

• Size-dependence of ligand-protein interactions

The interaction of fitting polar groups of the ligand and the protein is potentially highly beneficial and can significantly contribute to the binding free-energy primarily by enthalpic gain [15]. The formation of polar interactions, however, is accompanied by unfavorable contributions that come from the desolvation of the polar groups and the decreased mobility of the interacting moieties. These predominantly entropic unfavorable contributions can only be compensated by the enthalpy gain of polar interactions if the geometrical arrangement of atoms is near to optimal. H-bonds are probably the most important polar interactions and their energy is highly sensitive to the relative positions of the participating atoms. An analysis of the optimal geometry H-bonds (donor-acceptor distance not larger than 3 Å and D-H…A angle is not smaller than 160 °) in ligand-protein complexes of the Protein Data Bank [16] revealed [17] that the average number of such H-bonds is around two even for small molecules having ~15 heavy atoms and the average number does not increase significantly with increasing ligand size. The appearance of optimal geometry H-bonds in complexes of small molecules is in line with the high free-energy gain associated with these H-bonds. Their number, however, does not increase for larger ligands and this finding can be rationalized by the high sensitivity of the H-bonding energy to the geometry of the interacting atoms that prevents the formation of further optimal geometry H-bonds that could beneficially contribute to the binding free-energy. Larger molecules, however, are able to bind with larger free-energy gain as it was discussed above, thus they have to find another way to achieve high affinity. Apolar desolvation appears to be able to contribute favorably to the binding of large molecules. Indeed, correlation was found between the apolar surface area buried upon complex formation and the binding free-energy [18], and their relation suggests 1 kJ/mol free-energy gain upon the burial of 20 $Å^2$. The apolar surface area and the corresponding estimated binding free-energy was calculated for ligands in the Protein Data Bank and they are depicted as a function of heavy atom count in Figure 1. The amount of free-energy coming from apolar desolvation for a molecule of 20 heavy atoms is at most 15 kJ/mol according to Figure 1

and it corresponds to the lower limit of the estimated entropy loss upon binding [19]. This suggests that apolar desolvation alone is unable to ensure the binding of small molecules. However, it becomes increasingly important as molecular size increases. Since polar interactions are unable to increase their favorable contribution to binding (see above) it is apolar desolvation that becomes the driving force of ligand-protein binding in the case of large ligands.

- ## Size-dependence of ligand-protein binding thermodynamics

As it was discussed above, polar interactions are able to contribute to binding primarily by enthalpic gain while the contribution of apolar desolvation is dominantly entropic. Then the involvement of these effects in the formation of ligand-protein complexes has to be reflected in the thermodynamic profile of the binding. Plotting the average enthalpic and entropic contributions to binding free energy as a function of ligand size (Figure 2) shows the expected trend. Enthalpy dominates for smaller ligands [14,20] in line with the importance of polar interactions for these molecules and entropy becomes dominant for large molecules as a consequence of the increased contribution of apolar desolvation.

The desolvation contribution to binding is related to the surface area as it is supported by the correlation between buried apolar surface area and the binding free-energy [18] and also by a model that includes apolar surface area to calculate binding enthalpy [21]. Detailed thermodynamic and structural analyses of the water network and its perturbation by ligands can distinguish water molecules whose repulsion from the binding site is favorable or unfavorable in terms of their contribution to binding free energy and its components [22,23,24,25,26]. Although an assignment of free-energy and its components to water molecules is an approximation it contributes to our understanding of the role of water in ligand binding. It appears that the enthalpic component of water repulsion can be both favorable and unfavorable while the entropic component is basically favorable [26,27,28]. This finding is in agreement with the increasing role of entropy gain for larger ligands.

The trend of decreasing enthalpic contribution with increasing ligand size is also observed for the highest affinity compounds. While available affinity increases with ligand size the corresponding enthalpic contribution diminishes for large ligands as it is shown in Figure 3. This plot was generated by assigning the compounds to bins defined by heavy atom counts and the highest pK_d is shown for each bin together with the maximal pK_H of those compounds of the bin whose affinity is in the top 5% (most enthalpic binders among the top affinity compounds). It appears that small size high affinity compounds are able to bind with favorable enthalpy while the contribution of enthalpy diminishes with increasing ligand size. Large ligands are able to bind with higher affinity by higher entropic contribution.

Optimization strategies and binding thermodynamics

The basic challenge in affinity improvement from thermodynamics point of view is to overwrite enthalpy-entropy compensation [29]. Optimizations can then be classified as enthalpically and entropically driven; enthalpically driven optimizations achieve this goal by dominantly pK_H increase, while pK_S increase dominates in entropically driven optimizations. Since pK_H can be improved by the introduction of new polar interactions, enthalpically driven optimizations have the advantage to

produce improved affinity compounds with balanced physicochemical properties and advantageous pharmacokinetic profile. However, designing new beneficial polar interactions is difficult that makes enthalpically driven optimization challenging [15]. By contrast, entropically driven optimizations increase affinity by improving pK_S. This can be achieved by introducing apolar groups into the ligand so that they match the protein binding site and thus primarily increasing favorable entropy by desolvation. Another way to increase pK_S is to reduce ligand flexibility typically achieved by applying chain-ring strategies leading to larger, more complex molecules. Both lipophilicity and complexity increase is generally more straightforward than the design of new beneficial polar interactions and thus entropically driven optimizations are easier to realize. However, entropy dominated optimizations are coupled with extensive increase in molecular size and lipophilicity as it is shown by the comparison of these parameters for enthalpy and entropy dominated binders [14,9] and these features adversely affect physicochemical properties and pharmacokinetic profiles (see e.g. refs 1 and 3).

Enthalpically driven optimization is preferred over entropically driven optimization, although, as argued above, it is far more difficult to pursue. Moreover, enthalpically driven optimization appears to be more feasible for smaller compounds as it is shown in Figure 2 and Figure 3. This suggests that increasing the enthalpy content of binding is effective in the early stages of the optimization and furthermore that enthalpic starting points are beneficial. Measuring and comparing the binding enthalpy of ligands at decision points like hit and lead selection and monitoring the enthalpy in early optimizations is expected to advantageously affect the quality of optimized compounds. Late stage optimizations often improve potency at the expense of increased binding entropy and tend to produce more complex and more lipophilic compounds. This is illustrated in Figure 4 that shows the average binding enthalpy and entropy as a function of affinity. Low affinity compounds bind enthalpically and the entropic contribution on average is near to zero. The majority of these low affinity compounds are fragments containing not more than 22 heavy atoms and these small polar compounds bind enthalpically as it is discussed later. On the other hand, the thermodynamic profile above pK_d=*8* starts to be seriously biased towards entropy and this suggests that affinity increase in that range threaten the quality of physicochemical properties and pharmacokinetic profile.

• Thermodynamic rationale of fragment based approaches

The chemical starting points of fragment-based drug discovery programs are small, low complexity, polar compounds with typically low affinity. Owing to their low affinity, compounds have to be screened at high concentration that requires high solubility. In line with their small size and polarity that allows fragments to form few, good quality polar interactions without significant contribution from apolar desolvation fragment hits typically bind with favorable enthalpy [17,30,31,32,33,34,35] (Figure 5). The few compounds in Figure 5 that bind with unfavorable enthalpy are all anionic species with highly unfavourable desolvation profile. Another factor that affects binding thermodynamics is the rigid body entropy loss that is estimated to be 15-20 kJ/mol [19] and fragments are able to compensate this entropy loss at a lesser extent than larger compounds. It is worth also noting that the perturbation of the water network upon binding although associated with significant contribution to binding thermodynamics [36,37] is less important for fragments than for larger size ligands. Summarizing, fragments have a binding thermodynamic profile characterized by enthalpy domination and this, first, distinguishes them from larger ligands and, second, forms the thermodynamic rationale of fragment based approaches. Fragments are highly appropriate starting

points of drug discovery programs as they have already important polar interactions formed with the hot spot of the protein [17] and due to their enthalpic character provide large operational freedom for a balanced optimization.

Ligand efficiency indices and binding thermodynamics

• Ligand efficiency and enthalpic efficiency

Drug discovery optimizations face the challenge of simultaneously improving affinity, selectivity, physicochemical properties and pharmacokinetic profile. Both decision points like hit and lead selection and the optimization process require the complex characterization of compounds in order to compare them. Ligand efficiency indices are appropriate tools for such characterizations since they are composite measures that carry complex information, yet they can be easily obtained. Various efficiency indices have been proposed (Table 1) some of them are more widely accepted and used than others [38,39,40].The most widely used index is ligand efficiency (LE) that expresses the ratio of affinity and ligand size. It was originally defined as $LE = \frac{\Delta G}{HA}$ [41] although pK_d, pK_i and pIC_{50} are often used instead of ΔG (Table 1). It has to be recognized that although *LE* is the affinity normalized to the number of atoms, its maximal available value does depend on ligand size [13] and this fact biases the comparison of different size ligands. This recognition led to the definition of size-independent ligand efficiency, $SILE = \frac{affinity}{HA^{0.3}}$, whose maximum available value is independent of ligand size [43]. Indeed, it was demonstrated [44] that *LE* does not exhibit a clear trend in optimizations, while *SILE* shows a significant increase in successful optimizations and thus it is an effective tool in hit and lead selection and also in monitoring optimizations.

The enthalpy content of binding contains information about the type and quality of ligand-protein interaction and this prompted the definition of enthalpic efficiency $EE = \frac{\Delta H}{Q}$, where *Q* is either the number of heavy atoms or the molecular mass [30,45]. An alternative index is specific $EE = \frac{\Delta H}{Npol}$ that measures the enthalpy change relative to the number of polar atoms [30]. The maximal available value of both enthalpic efficiency and specific enthalpic efficiency heavily depend on ligand size and a size-independent version, $SIHE = \frac{pK_H}{40} HA^{0.3}$, was defined [14] for the unbiased comparison of the binding enthalpy content of different sized ligands. Although the available enthalpy may be target dependent it does not affect compound comparison based on SIHE values against the same target as it is typically required in drug discovery optimizations. However, the maximal available binding enthalpy does not change monotonically with ligand size (Supplementary Figure S1) and therefore neither *EE* nor *SIHE* is appropriate to compare very small compounds against those having more than 15 heavy atoms. It has to be also noted that a division by the number of heavy atoms have a damping effect and thus both *LE* and *EE* cover a decreasing range with increasing heavy atom count. This has the consequence that *LE* and *EE* contain less information for large compounds, but size-independent measures, like *SILE* and *SIHE*, do not suffer from this deficiency as their definition ensures ligand size independent maximal values (Supplementary Figure S2). This difference between the size dependent and size independent measures is apparent already around the fragment limit of 22 heavy atoms.

The calculation of ligand efficiency requires the knowledge of affinity, and the calculation of enthalpic efficiency requires the knowledge of enthalpy. While affinity measurements are part of the standard drug discovery setup thermodynamic measurements are less common and this currently restricts the use of enthalpic indices. It is worth also noting that the decreasing trend of maximal available ligand efficiency (Supplementary Figure S2c) resembles to the trend observed for enthalpic efficiency. The maximal available enthalpy decreases with ligand size [17] and so does the maximal enthalpic efficiency [20]. Since this decrease is not fully compensated by an entropic efficiency increase the maximal available ligand efficiency decreases.

• Lipophilic efficiency

Lipophicity is an important property of ligands as it affects affinity, selectivity and ADMET properties. Lipophilicity tends to increase in entropically driven drug discovery optimizations and its thermodynamic background was discussed above. Selectivity and specificity typically decrease with increasing lipophilicity [2], while the relationship between lipophilicity and ADMET properties is more complex. Less lipophilic compounds tend be more soluble [46], less impacted by liver metabolism and hepatic clearance [47], and less toxic [48]. On the other hand, however, they show low permeability across biological membranes [49] and highly affected by renal clearance [50]. These observations suggest that lipophilicity has an optimal range to achieve during drug discovery optimizations and this is also supported by the analysis of discovery datasets that identified optimal lipophilicity ranges for oral drugs between 0.8 and 3.2 logP units [51] and between 2 and 3.2 logP units [9].

The narrow range of optimal lipophilicity has serious consequences on both the identification of chemical starting points and on their optimization. Lipophilic indices are designed to control the balance with respect to the lipophilicity and other properties like affinity and size. Definitions of commonly used lipophilic indices are shown in Table 1.

Ligand lipophilicity efficiency (*LLE*), historically the first lipophilic index, was proposed to have target values ~5-7 or larger [2]. These values were recommended on the basis of the *clogP*~2.5 and potency range~1-10 nM taken as average values for oral drugs. Compounds at hit identification and early optimizations typically do not achieve this *LLE* range and other lipophilic indices that also include ligand size were proposed. Lipophilicity corrected ligand efficiency (*LELP*) [1] describes the price of ligand efficiency paid in *logP*. The target values for *LELP* were proposed to be between -10 and +10 based on the widely accepted lower limit of ligand efficiency (0.3), and on the lipophilicity range between –3 and +3. LLE_{AT} was proposed to characterize small screening hits, like fragments and it is defined to have a target value (0.3) and dynamic range similar to ligand efficiency (*LE*) [52].

Lipophilic indices are obtained from calculated *logP* values, measured affinities and, eventually, measures of ligand size. They do not explicitly include thermodynamic quantities but *logP* that characterizes lipophilicity is related to the entropic contribution of apolar desolvation upon ligand-protein complex formation, the latter correlating with the apolar surface area buried in the complex [18] (see above). In this respect, lipophilic indices and enthalpic indices are expected to contain overlapping information, the former having the advantage that they can be obtained without thermodynamic measurements. Establishing a relationship between efficiency indices and binding enthalpy would support enthalpic optimization by easily calculated indices. LLE and binding enthalpy

was shown to exhibit some association [53] although the relationship observed appear to be system specific and semi-quantitative in the majority of the cases studied. Our experience with multiple *LLE-ΔH* datasets suggests that the correlation is much improved for congeneric compound series. The extent and utility of the relationship between LLE and binding enthalpy merit further studies.

Table 2 presents ligand property changes of different types of medicinal chemistry optimizations. Data for successful fragment optimizations leading to clinical candidates and successful lead optimizations leading to marketed drugs are shown together with data of optimizations performed under thermodynamic control. It is instructive to see that thermodynamic optimizations are able to improve potency with a modest increase in molecular size without notable change in logP and this is reflected in an almost constant LELP and an improving LLE. In this way, thermodynamic optimizations produce average property changes that resemble to successful optimizations.

Optimization case studies

• Renin inhibitors

Diaminopyrimidine-type renin inhibitors were discovered by an HTS campaign at Pfizer identifying **1** having high micromolar affinity (Figure 6) [54].

Parallel synthesis of a 450-membered focused library around the diaminopyrimidine core resulted in **2**, a low micromolar renin inhibitor. Structural analysis of the renin-**2** complex revealed that the large S2 hydrophobic pocket and the smaller hydrophobic S3 subpocket were unoccupied. Since preliminary studies to fill the S2 pocket failed **2** was first tethered by a tetrahydroisoquinoline (**3**) and a benzoxazinone (**4**) ring system which were extended by a methoxypropyl side-chain toward the S3 subpocket [54,55]. Both compounds showed substantially higher affinity relative to **2** with a particular advantage of **3** having lower IC_{50} and K_d as compared to that of **4**. Ligand efficiency was improved similarly in both cases (Supplementary Table S1), however the improvement was more significant in SILE than LE that is the direct consequence of size dependency. Comparing lipophilic efficiency values, however, indicated a large difference between **3** and **4**. Since the benzoxazinone (**4**) is much less lipophilic than the tetrahydroisoquinoline (**3**) both LLE (larger is better) and LELP (smaller is better) showed preference for **4**. In fact, **4** is located in the most preferred region of the LLE-LELP plot (Supplementary Figure S3) characterized by oral drugs and compounds having sufficiently good human pharmacokinetic profile to enter into phase 2 trials [56].

Thermodynamic profiles of **3** and **4** were recorded and were compared to that of their predecessor **2** (Figure 7). While optimization from 2 to 3 was mainly due to entropic improvements conversion of 2 to **4** is an enthalpically driven process. Analysis of the X-ray structure of the renin-**3** complex revealed that the methoxypropyl sidechain reached the S3 subpocket. Since the displacement of ordered water molecules from the hydrophobic S3 subpocket was entropically favored it compensated for the entropy loss associated with the decreased flexibility of the side chain. These significant entropy effects identify the optimization of **2** to **3** as being basically entropically driven as indicated by the decreasing EE and the marginal improvement detected in SIHE (Supplementary Table S1). Although the X-ray structure of the renin-**4** complex is not publicly available, the complex of its 2,2-dimethyl-

benzoxazinone analog was crystallized and showed that in addition to filling the S3 subpocket the benzoxazinone group was involved in new polar contacts within the active site and its methyl group in position 2 formed favored van der Waals contacts. These new interactions yielded a significant gain in enthalpy (ΔH~4 kcal/mol) that was only partially compensated by the entropic penalty (TΔS~2 kcal/mol) due to the decreased flexibility and desolvation entropy. The significant enthalpy effects detected make this optimization enthalpy driven that is indicated by the much improved EE and SIHE (Supplementary Table S1). Although the enthalpically optimized compound (**4**) is somewhat less potent than that obtained by entropically driven optimization (**3**) its size and lipophilic efficiency are superior. This rationalizes that optimized compounds reported from the diaminopyrimidine chemotype typically contain benzoxazinone rather than tetrahydroisoquinoline rings [57]. Further optimization of **4** was focused on the central region and the S3 subpocket. Introducing the lipophilic difluorophenyl moiety into position 2 of the benzoxazinone ring resulted in **5** with improved potency (IC_{50}) but interestingly kept the affinity (K_d) virtually constant. Although LE does not change too much and SILE improved a bit a large drop in LLE and LELP was detected (Supplementary Table S1). The thermodynamic analysis confirmed this step entropically driven with significant reward in binding entropy connected with remarkable enthalpy loss (Figure 7a). Figure 7b summarizes the thermodynamics of the optimization path. Although the last step from **4** to **5** was entropically driven the large enthalpic step from **2** to **4** makes compound **5** more enthalpic (c.f. the ΔH axis) and more potent (c.f. ΔG axis) than **3** indicating **5** as a better lead. Further optimization of the S3 sidechain of **5** finally led to compound **6** having somewhat less affinity but improved ADME properties. **6** showed good bioavailability both in rat (74%) and dog (19%) triggering its selection for preclinical development as identified from the Integrity database [57].

- ## Tankyrase inhibitors

A Novartis team described a series of tankyrase TNKS1/TNKS2 inhibitors inhibiting Wnt pathway signaling. This set of compounds is exemplified by XAV939 (**7**) having low nanomolar affinity towards both TNKS isoforms (Figure 8) [58,59]. Compound 7, however, showed moderate selectivity towards other PARPs, particularly against PARP2 and had moderate metabolic stability that limited its in vivo application. In addition, thermodynamic profiling of **7** revealed that despite of the favourable enthalpic contributions the compound binds entropically to TNKS2 (Figure 9a). The Novartis team therefore started an optimization program from **7** that aimed the improvement of potency, selectivity, solubility and metabolic stability [60]. Since **7** is structurally similar to many PARP inhibitors they introduced a number of modifications to the bicyclic core of **7** from which the dihydropyran analogue (**8**) was somewhat less potent but significantly more polar (Figure 8).

Decreased lipophilicity of **8** resulted in better solubility and increased metabolic stability while maintaining LLE and improving LELP (Supplementary Table S2). In agreement with these improvements the thermodynamic profile of **8** showed that the compounds bind enthalpically (Figure 9a) as reflected in improved EE and SIHE values (Supplementary Table S2). One of the major advantages of **8**, however, was the improved selectivity over PARP1 and 2 that prompted the team to investigate its analogues further. X-ray analysis of **7** confirmed that the core occupies the nicotinamide and diphosphate pockets. In replacing the trifluoromethyl group by a number of

aromatic substituents with different polarity, it turned out that this part of the molecule has no direct interactions with the protein but instead the substituents are exposed to the bulk solvent. Based on this observation the team suggested introducing amphipathic groups that on one hand might form hydrophobic interactions with the target and on the other hand fine-tune the lipophilicity of the compounds. To achieve this goal the trifluoromethylphenyl substituent of the dihydropyran core was replaced by the phenylethyl-carboxamido sidechain of a moderately active screening hit resulting **9** (Figure 8).

X-ray crystallography of the screening hit complexed with TNKS1 revealed that the compound binds to the NAD+ donor site and its sidechain, particularly the thiophene moiety forms van der Waals interactions with the hydrophobic nook. Compound **9** unifies the advantageous contribution of the dihydropyran core with the beneficial effect of the sidechain on potency. Although ligand efficiency indices were getting worse a bit comparing the thermodynamic profile of **9** to its predecessor (**8**) shows a significant improvement in binding enthalpy (Figure 9a) as indicated by the much improved EE and SIHE values (Supplementary Table S2). The enthalpic reward realized in **9** is likely the effect of a new hydrogen bond formed with the amide carbonyl and also the van der Waals contacts of the thiophene moiety. As a result the team identified a promising enthalpic lead having good selectivity profile but moderate potency and still low metabolic stability.

Further optimization of **9** therefore involved the variation of the metabolically vulnerable thiophene and phenethyl moieties. This activity was facilitated by the identification of a screening hit that binds not only in the nicotinamide pocket but reaches the adenosine site as well. X-ray analysis of its TNKS1 complex revealed that the (fluorophenylpiperidinyl)methanone sidechain of the hit forms advantageous van der Waals interactions with the groove in the adenosine pocket that suggested replacing the sidechain of **9** with that of the hit. In fact, compound **10** designed on this basis showed high potency, good selectivity and much improved metabolic stability (Figure 8). The introduction of the more polar side chain made the compound less lipophilic that together with the increase in potency resulted in improved ligand efficiency values. Although LE did not change too much there was some improvement in SILE and more importantly both LLE and LELP were getting significantly better (Supplementary Table S2). It is important to note that all of these improvements were achieved with keeping the enthalpic character of compound virtually unchanged (Figure 9a) and **10** represents the first TNKS inhibitor forming interactions simultaneously with the nicotinamide, adenosine and nook pockets of the enzyme.

In the last step of the optimization the metabolically vulnerable thiophene moiety of **10** was changed to a cyclopropylmethyl group (Figure 8, **11**) that did not influence the potency significantly and kept ligand efficiency constant (Supplementary Table S2), however, improved the oral exposure of the compound ten-fold. The introduction of the apolar cyclopropyl substituent increased logP/logD as reflected in slightly worse lipophilic efficiency indices (Supplementary Table S2, LLE and LELP). In agreement with these changes the thermodynamic profile of **10** became more entropic (Figure 9a), but still predominantly enthalpically driven. The somewhat decreased EE and SIHE values (Supplementary Table S2) indicate this change in binding thermodynamics that can be attributed to the desolvation of the hydrophobic cylopropylmethyl moiety. Finally, improving lipophilic efficiency further, the fluorophenyl group of **11** was changed to methoxyphenyl that resulted in **12** (Figure 8, TNKS656) with improved LLE and LELP values while maintaining the ligand efficiency (LE, SILE). Comparing the thermodynamic profile of **12** to **11**, a slight improvement in the binding enthalpy was

detected that is reflected in the improved enthalpic efficiency (EE, SIHE) of the compound (Supplementary Table S2). **12** was then subjected to extensive in vivo characterization, showed good exposure and moderate bioavailability and more importantly was found to be effective on adenocarcinomas obtained from a transgenic mouse model with Wnt signaling activation.

Summarizing the optimization process we depicted the compounds **7**-**12** on the ΔG-ΔH graph (Figure 9b). The optimization process has been started from the entropic binder **7** that was converted to a compound with strong enthalpy driven binding in two steps (from **7** to **8** and **8** to **9**). Adjustment of physicochemical and ADMET properties involved only small changes in the thermodynamic profile (relative to **9**) that were both entropic and enthalpic and resulted in compound **12** with a significant enthalpic character. It is important to note that in agreement to our observation [14] significant improvement in binding enthalpy was achieved in the early phase of the optimization with less complex compounds. Furthermore, the Novartis tankyrase dataset with PARP1 and PARP2 selectivity data supports the hypothesis that enthalpic compounds are generally more selective [11]. Although the thermodynamic profiles of key compounds were recorded the authors used lipophilic efficiency, particularly LLE as the major design tool in this optimization. The LLE-LELP plot (Supplementary Figure S4) provides a straightforward validation for this approach since all the compounds except for **9** are located in the most preferred region identified by marketed drugs and Phase 2 compounds.

• Matrix metalloprotease 12 (MMP12) inhibitors

Matrix metalloproteinases are zinc containing proteases having numerous high affinity inhibitors that mostly bind to the substrate binding groove forming interactions with the Zn-site and also in the hydrophobic S1' pocket. Bertini and coworkers investigated the binding thermodynamics of a number of N-hydroxy-2-(phenylsulfonamido)acetamides (Figure 10) [61] and analyzed their thermodynamic profiles in relation to their binding mode. This type of inhibitors can be derived from acetohydroxamic acid (**13,** AHA), a fragment that chelates the Zn ion directly. Although this fragment has moderate affinity in the low millimolar range it showed excellent ligand efficiency and lipophilic efficiency (Supplementary Table S3). Isothermal titration calorimetry experiments revealed **13** an enthalpic binder that might serve as a suitable starting point for optimization. Linking the phenylsulfonamide fragment that binds in the S1' cavity [62] and **13** resulted **14** with significantly improved affinity. This key step of optimization improves both ligand efficiency (LE and SILE but to a different extent) and lipophilic efficiency values significantly (Supplementary Table S3). From a thermodynamic point of view fragment linking was associated with a large improvement in binding enthalpy (Figure 11). Enthalpic efficiency changed accordingly, however, the improvement in SIHE was more pronounced relative to changes in EE. This effect, together with the similar situation found for LE and SILE, underlines the importance of size independent metrics when comparing ligands with largely different size.

Further optimization of **14** was first realized by introducing different substituents to the para position of the benzenesulfonamide ring. The authors investigated the 4-F (**15**), the 4-methoxy (**16**) and the 4-phenyl (**17**) analogues. Introduction of the fluorine substituent did not improve the potency, on the

other hand the binding affinity of both **16** and **17** was improved. Analysis of the corresponding X-ray structures revealed that the substituted phenyl group fits into the S1' pocket. The methoxy derivatives penetrate somewhat deeper into the S1 pocket as compared to the unsubstituted derivative. This suggests that the methoxy group probably displaces water molecules from the hydrophobic S1' pocket that would result in some gain in binding entropy. Although the methoxy group forms van der Waals contacts it also disturbs the optimal H-bond interactions of the sulfonamide group and the coordination of the hydroxamic acid. The fluoro substituent introduces less perturbation, while the second phenyl group of **17** represents a new site for hydrophobic interactions but more importantly decreases the hydrophobic interactions of the benzenesulfonamide ring within the S1' pocket. Thermodynamic profiles are in accordance with these structural changes (Figure 11a). The largest entropic reward was observed for **15**, however, the binding of the compound is still enthalpically driven. The introduction of the methoxy substituent improved the binding enthalpy further and therefore this transformation is considered as enthalpically driven optimization. The second phenyl group in **17** provides entropic reward due to improved desolvation that makes this optimization step being entropically driven. In summary, from the thermodynamic point of view the most enthalpic methoxy derivative (**16**) can be considered as the preferred starting point for further optimization. Interestingly, enthalpic efficiency indices did not discriminate effectively this subset of compounds (Supplementary Table S3). Analyzing ligand efficiency indices LE and SILE prefer **15** and **16** over **17** and a similar trend was observed for both LLE and LELP. This information together with enthalpic efficiency indices suggests selecting **16** for further optimization.

In the next round the authors introduced substituents to the sulfonamide nitrogen and the carbon atom of the acetamido moiety. The sulfonamide nitrogen was substituted by isobutyl (**18**), 2-hydroxyethyl (**19**) and 2,3-dihydroxypropyl (**20**) groups while C-substituents involved hydroxymethyl (**21**) and 2-hydroxyethyl (**22**) groups in D configuration. Introduction of the hydrophobic isobutyl group increased the affinity of **18**. Ligand efficiency (LE) was slightly decreased, however, SILE improved a bit (Supplementary Table S3). The enthalpic efficiency (SIHE) remains virtually constant, but LELP was getting worse due to the increased lipophilicity. The thermodynamic profile of **18** is in line with these findings. Improvement of the binding affinity has been achieved by entropically driven optimization as indicated by the beneficial change in binding entropy (Figure 11a). In contrast, introducing the polar hydroxyethyl group (**19**) kept the ligand efficiency (SILE) constant and improved the enthalpic efficiency (SIHE) to some extent. The observed reward in binding enthalpy could be attributed to van der Waals contacts formed between the ethyl spacer and Pro238. More importantly, the terminal hydroxyl group forms H-bonds with a number of water molecules located at the entrance of the substrate binding site. The polar character of the substituent helped control the lipophilicity and improved the LLE value. Similar trends were observed for the LLE of the 2,3-dihydroxypropyl (**20**) analogue that improved further due to the more polar character of **20** relative to **16**. Binding of **20** was found to be enthalpically driven, however, contrary to **19** here the binding entropy was also improved marginally. C-substituted analogues (**21** and **22**) were found to be almost equipotent to N-substituted compounds. Contrary to latter ligands improvements in the affinity of **21** and **22** was due to entropic rewards. Considering similar changes in ligand efficiency parameters observed for both N- and C-substituted analogues the most enthalpic analogue (**19**) having also the highest SIHE is suggested for further studies.

The optimization path has been summarized on the ΔG-ΔH graph (Figure 11b). The optimization process was started from the enthalpic fragment binder **13** that was converted to a compound with strong enthalpically driven binding (**14**). Optimization of the aromatic substituent was entropically driven, here we can select the compound with maximal binding enthalpy (**16**). Finally, from the comparative thermodynamic profiling of N- and C-substituted analogues (**18**-**22**) the most enthalpic compound (**19**) was achieved by enthalpically driven optimization. Again, in agreement to our observation [14] significant improvement in binding enthalpy was achieved in the early phase of the optimization with less complex compounds that in this particular case was realized by linking two fragments.

- ## Heat shock protein 90 (HSP90) inhibitors

Recently Astex reported that combined fragment screening using both NMR and X-ray crystallography against HSP90 yielded a new aminopyrimidine starting point (Figure 12) with high micromolar affinity (**23**) [63]. The initial fragment hit showed reasonable ligand efficiency and also acceptable lipophilic efficiency (Supplementary Table S4). X-ray analysis of the HSP90-**23** complex revealed that the fragment is located in a deep pocket forming hydrogen bonds with Asp93 and four water molecules. The binding mode of the aminopyrimidine ring resembled to that of the adenosine ring of ADP having two small hydrophobic pockets in close proximity. Since **23** was twisted around the bond connecting the pyridine and the aminopyrimidine ring this conformation was maintained by introducing an ortho methoxy group when replacing the original pyridine ring to benzene. Introducing a methyl group to the 2-position of the aminopyrimidine ring yielded **24** with increased potency and improved ligand efficiency (both LE and SILE) while lipophilic efficiency was kept on a reasonable level (Table 5). Next there are two options to improve the affinity of the fragment further. In one case a chlorine atom was introduced to the ortho position of the 2-methoxyphenyl ring (**25**) that was replaced by another methoxy group resulting the dimethoxyphenyl analogue (**26**). The other option replaces the methyl group of the aminopyrimidine ring by a chlorine (**27**) and the introduction of the second methoxy group yields the chloro analogue (**28**) of **26**. Comparing ligand efficiencies compounds on the second path are more beneficial in terms of both LE and SILE and also the lipophilic efficiencies are better for **27** and **28** relative to **25** and **26** (Supplementary Table S4). On the first path the introduction of the 2-chloro substituent to the phenyl ring was associated with a significant reward in the binding enthalpy (Figure 13), however the conversion of **25** to **26** was coupled to a large entropic reward. Following the other path the methyl to chloro replacement from **24** to **27** is enthalpically driven. It is appealing to speculate that the enthalpy gain observed in replacing Me by Cl is owing to the halogen bond that can be clearly identified in the corresponding X-ray structure [63]. Similar to the first path the introduction of the second methoxy group to the phenyl ring (from **27** to **28**) is entropically driven. Considering the overall change in binding enthalpy and entropy along the path from **24** to **26** and **28** it seems that **28** is more enthalpic than **26** that makes the former more suitable for further optimization. In fact, the optimized derivatives reported from these laboratories contain chloro rather than methyl substituent on the aminopyrimidine core.

Future perspective

The recognized relation between binding thermodynamics and the physicochemical and druglike properties of ligands ensures that the measurement and analysis of binding thermodynamic profiles remain valuable tools in drug discovery projects. The predominant experimental technic for measuring thermodynamic profiles is expected to remain isothermal titration calorimetry as it is able to provide ΔG, ΔH, and ΔS in a single experiment for soluble proteins. The limitations of ITC measurements, namely throughput and the amount of protein required have been significantly alleviated by recent technological developments [65,35] and this is likely to increase the role of thermodynamic measurements both in hit identification and in optimization. This trend is expected to apply also for fragment based drug discovery since recent results have clearly demonstrated the feasibility of applying ITC in this domain [32,34,35].

Binding thermodynamics and structural analysis, most often by X-ray crystallography, are typically applied together in an attempt to rationalize binding thermodynamics by structural features. Although some trends on how ligand and protein structures affect binding thermodynamic profiles have been set up, we are currently unable to rationally design binding thermodynamics and even the interpretation of thermodynamics profile is often challenging. The interplay of diverse events among them the breaking and formation of polar and van der Waals interactions, solvent structural changes and the alteration of the available configuration space for the system results in the observed thermodynamic profile and further data are to be accumulated and analyzed in order to better understand the subtle events that contribute to binding thermodynamics. Thermodynamic and structural studies are expected to be jointly applied also in the future since they are both required to provide detailed information on the nature and quality of interactions and on the changes upon complex formation.

Our current knowledge on the relation between structural features and binding thermodynamic profiles refer to ligand-protein complexes in general. More research are needed to clarify whether various protein families behave differently, whether soluble and membrane proteins exhibit distinct binding thermodynamics and whether other drug targets like DNAs have different binding thermodynamic characteristics.

Binding kinetics is also in the forefront of the current interest in drug discovery [64] owing to its influence on drug pharmacology via the residence time of the drug-target complex. Since binding kinetics studies k_{on} and k_{off}, the association and dissociation rate constants that also determine binding thermodynamics the two fields will mutually profit from the continuing interaction.

Executive summary

- Binding thermodynamics together with structural information are able to characterize the quality of interactions in ligand-protein complexes.
- Affinity improvement achieved by the introduction of apolar groups or by increased rigidity typically increases the entropic contribution of binding free energy and tends to deteriorate physicochemical and pharmacokinetic profiles by size and lipophilicity increase.

- The introduction of polar groups into the ligands so that they form favorable interactions with the protein leads to dominantly enthalpic gain in the binding free energy and allows a better control of molecular size and lipophilicity.
- Favorable polar interactions are restricted to a narrow range of the geometrical arrangement of the interacting atoms and this makes the rational design of enthalpic optimization challenging.
- Binding thermodynamic profiles exhibit ligand size dependent trends; Small ligands, like fragments, typically bind with favorable enthalpy owing to their small size and polarity, while the entropic contribution to binding gains importance with increasing ligand size and increasing affinity.
- The control of the binding thermodynamic profile is beneficial to the physicochemical and druglike properties of drug candidates and such a control is more viable in the early phases of optimizations.
- When thermodynamic and structural data are not accessible then efficiency indices can provide similar information on interactions quality. Ligand efficiency and lipophilic efficiency indices are typically easily available and they are able to contribute to balanced optimizations particularly in the early phases.

Bibliography

1. Keserű GM. Makara GM. The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discov.* 8, 203-212 (2009).
2. Leeson PD. Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov*. 6, 881-890 (2007). * A seminal paper on the control of physicochemical properties during lead optimization programs.
3. Ferenczy GG. Keserű GM. Thermodynamics guided lead discovery and optimization. *Drug Discov. Today* 15, 919–932 (2010).
4. Keserű GM. 5th Drug Design & Lead Discovery Conference 2009: Lead Finding Strategies and Optimization Case Studies. *Drugs of the Future* 35, 143-153 (2010).
5. Hann MM. Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Comm.* 2, 349–355 (2011).
6. Hopkins AL. Mason JS. Overington JP. Can we rationally design promiscuous drugs? *Curr. Opin. Struct. Biol.* 16, 127–136 (2006).
7. Tarcsay Á. Keserű GM. Contributions of molecular properties to drug promiscuity. *J Med Chem.* 56, 1789-1795 (2013).
8. Van de Waterbeemd H. Smith DA. Jones BC. Lipophilicity in PK design: methyl, ethyl, futile. *J. Comput. Aided Mol. Des.* 15, 273–286 (2001).
9. Hann MM. Keserü GM. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discov.* 11, 355-365 (2012).
10. Kawasaki Y. Freire E. Finding a better path to drug selectivity. *Drug Discov Today* 16, 985-990 (2011).
11. Tarcsay Á. Keserű GM. Is there a link between selectivity and binding thermodynamics profiles? *Drug Discov Today*. 20, 86-94 (2015).
12. Kuntz ID. Chen K. Sharp KA. Kollman PA. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9997-10002 (1999).
13. Reynolds CH. Bembenek SD. Tounge BA. The role of molecular size in ligand efficiency. *Bioorg. Med. Chem. Lett.* 17, 4258–4261 (2007).
14. Ferenczy GG. Keserű GM. Enthalpic efficiency of ligand binding. *J. Chem. Inf. Mod.* 50, 1536–1541 (2010). * This study describes first that in contrast to binding free energy the binding enthalpy is decreasing with increasing size of the compounds.
15. Freire E. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* 13, 869–874 (2008).
16. Berman HM. Westbrook J. Feng Z. Gilliland G. Bhat TN. Weissig H. Shindyalov IN. Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* 28, 235−242 (2000).
17. Ferenczy GG. Keserű GM. Thermodynamics of Fragment Binding. *J. Chem. Inf. Mod.* 52, 1039-1045 (2012).
18. Olsson TS. Williams MA. Pitt WR. Ladbury JE. The thermodynamics of protein–ligand interaction and solvation: insights for ligand design. *J. Mol. Biol.* 384, 1002–1017 (2008).
19. Murray CW. Verdonk ML. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput.-Aided Mol. Design* 16, 741−753 (2002).
20. Reynolds CH. Holloway MK. Thermodynamics of Ligand Binding and Efficiency. *ACS Med. Chem. Lett.* 2, 433–437 (2011).

21. Luque I. Freire E. Structural Parameterization of the Binding Enthalpy of Small Ligands. *Proteins Struct Funct Genet* 49, 181–190 (2002).
22. Homans SW. Water, water everywhere — except where it matters? *Drug Discov. Today* 100, 534-539 (2007).
23. Abel R. Young T. Farid R. Berne BJ. Friesner RA. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* 130, 2817-2831 (2008).
24. Krimmer SG. Betz M. Heine A. Klebe G. Methyl, Ethyl, Propyl, Butyl: Futile But Not for Water, as the Correlation of Structure and Thermodynamic Signature Shows in a Congeneric Series of Thermolysin Inhibitors. *ChemMedChem* 9, 833–846 (2014).
25. Ichihara O. Shimada Y. Yoshidome D. The Importance of Hydration Thermodynamics in Fragment-to-Lead Optimization *ChemMedChem* 9, 2708–2717 (2014).
26. Huggins DJ. Quantifying the entropy of binding for water molecules in protein cavities by computing correlations. *Biophys. J.* 108, 928-936 (2015).
27. Mason JS. Bortolato A. Congreve M. Marshall MH. New insights from structural biology into the druggability of G proteincoupled receptors. *Trends Pharmacol. Sci.* 33, 249-260 (2012).
28. Snyder PW. Locket MR. Moustakas DT. Whitesides GM. Is it the shape of the cavity, or the shape of the water in the cavity? *Eur. Phys. J. Special Topics* 223, 853-891 (2014).
29. Ruben AJ. Kiso Y. Freire E. Overcoming roadblocks in lead optimization: a thermodynamic perspective. *Chem. Biol. Drug Des.* 67, 2–4 (2006). * This paper points out the impact of binding thermodynamics on medicinal chemistry optimizations.
30. Ladbury JE. Klebe G. Freire E. Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat. Rev. Drug Discov.* 9, 23–27 (2009).
31. Dolezal O. Doughty L. Hattarki MK. Fazio VJ. Caradoc-Davies TT. Newman J.. Peat TS. Fragment Screening for the Modelling Community: SPR, ITC, and Crystallography. *Aust. J. Chem.* 66, 1507–1517 (2013).
32. Silvestrea HL. Blundell TL. Abell C. Ciulli A. Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery *Proc. Natl. Acad. Sci. U.S.A.* 110, 12984-12989 (2013).
33. Neumann L. Ritscher A. Müller G. Hafenbradl D. Fragment-based lead generation: identification of seed fragments by a highly efficient fragment screening technology. *J Comput Aided Mol Des* 23, 501–511 (2009).
34. Scott AD. Phillips C. Alex A. Flocco M. Bent A. Randall A. O'Brien R. Damian L. Jones LH. Thermodynamic Optimisation in Drug Discovery: A Case Study using Carbonic Anhydrase Inhibitors *ChemMedChem* 4, 1985–1989 (2009).
35. Torres FE. Recht MI. Coyle JE. Bruce RH. Williams G. Higher throughput calorimetry: opportunities, approaches and challenges. *Curr. Opin. Struct. Biol.* 20, 598–605 (2010).
36. Biela A. Betz M. Heine A. Klebe G. Water makes the difference: rearrangement of water solvation layer triggers non-additivity of functional group contributions in protein-ligand binding. *ChemMedChem* 7, 1423-1434 (2012).
37. Klebe G. Applying thermodynamic profiling in lead finding and optimization. *Nat. Rev. Drug Discov.* 14, 95-110 (2015).
38. Hopkins AL. Keserű GM. Leeson PD. Rees DC. Reynolds CH. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discovery* 13, 105−121 (2014). ** This paper summarizes the use of ligand efficiency metrics in medicinal chemistry optimizations.

39. Shultz MD. Improving the Plausibility of Success with Inefficient Metrics. *ACS Med. Chem. Lett.* 5, 2–5 (2014).
40. Murray W. Erlanson DA. Hopkins AL *et al.* Validity of Ligand Efficiency Metrics. *ACS Med. Chem. Lett.* 5, 616-618 (2014).
41. Hopkins AL. Groom CR. Alex A. Ligand efficiency; a useful metric for lead selection. *Drug Discov. Today* 9, 430-431 (2004). ** This paper introduces ligand efficiency (LE) as a useful tool for controlling molecular size in affinity optimizations.
42. Abad-Zapatero C. Ligand efficiency indices for effective drug discovery. *Exp. Opin. Drug Discov.* 2**,** 469–488 (2007).
43. Nissink JWM. Simple Size-Independent Measure of Ligand Efficiency. *J. Chem. Inf. Model.* 49, 1617–1622 (2009).
44. Ferenczy GG. Keserű GM. How are fragments optimized? A retrospective analysis of 145 fragment optimizations. *J Med Chem.* **56**, 2478-2486 (2013).
45. Ladbury JE. Enthalpic efficiency and the role of thermodynamic data in drug development: possibility or a pipeline dream. *Euro. Pharmaceut. Rev*. 12, 59–62 (2007)
46. Kramer C. Heinisch T. Fligge TT. Beck B. Clark. T. A consistent dataset of kinetic solubilities for early-phase drug discovery. *ChemMedChem* 4, 1529-1536 (2009).
47. Johnson TW. Dress KR. Edwards M. Using the Golden Triangle to optimize clearance and oral absorption. *Bioorg. Med. Chem Lett*. 19, 5560-5564 (2009).
48. Hughes JD. Blagg J. Price DA. Bailey S. Decrescenzo GA. Devraj RV. Ellsworth E. Fobian YM. Gibbs ME. Gilles RW. Greene N. Huang E. Krieger-Burke T. Loesel J. Wager T. Whiteley L. Zhang Y. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.* 18, 4872-4875 (2008).
49. Waring MJ. Defining optimum lipophilicity and molecular weight ranges for drug candidates-Molecular weight dependent lower logD limits based on permeability. *Bioorg. Med. Chem. Lett*. 19, 2844-2851 (2009).
50. Varma MV. Feng B. Obach RS *et al.* Physicochemical determinants of human renal clearance. *J. Med. Chem.* 52, 4844-4852 (2009).
51. Waring MJ. Lipophilicity in drug discovery. *Expert Opin Drug Discov.* 5, 235-248 (2010).
52. Mortenson PM. Murray CW. Assessing the lipophilicity of fragments and early hits. *J. Comput. Aided Mol. Des.* 25, 663-667 (2011).
53. Shultz MD. The thermodynamic basis for the use of lipophilic efficiency (LipE) in enthalpic optimizations *Bioorg. Med. Chem. Lett.* 23, 5992–6000 (2013).
54. Holswortha D. Jalaiea M. Belliottia T *et al.* Discovery of 6-ethyl-2,4-diaminopyrimidine-based small molecule renin inhibitors. *Bioorg. Med. Chem. Lett.* 17, 3575-80 (2009).
55. Powell NA. Ciske FL. Cai C *et al.* Rational design of 6-(2,4-diaminopyrimidinyl)-1,4-benzoxazin-3-ones as small molecule renin inhibitors. *Bioorg. Med. Chem.* 15, 5912-49 (2007).
56. Tarcsay A. Nyiri K. Keserű GM. Impact of lipophilic efficiency on compound quality. *J. Med. Chem.* 55, 1252–1260 (2012).
57. Thomson Reuters Integrity®. https://integrity.thomson-pharma.com/integrity/xmlxsl/ (2014)

58. Huang SM. Mishina YM. Liu S *et al.* Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* 461, 614−620 (2009).
59. Shultz MD. Kirby CA. Stams T *et al.* W. [1,2,4]Triazol-3- ylsulfanylmethyl)-3-phenyl-[1,2,4]oxadiazoles: Antagonists of the Wnt pathway that inhibit tankyrases 1 and 2 via novel adenosine pocket binding. *J. Med. Chem*. 55, 1127−1136 (2012).
60. Shultz MD. Cheung AK. Kirby CA *et al.* Identification of NVP-TNKS656: the use of structure-efficiency relationships to generate a highly potent, selective, and orally active tankyrase inhibitor. *J Med. Chem.* 56, 6495-511 (2013).
61. Bertini I. Calderone V. Fragai M *et al.* Exploring the Subtleties of Drug−Receptor Interactions: The Case of Matrix Metalloproteinases. *J. Am. Chem. Soc.* 129, 2466-75 (2007).
62. Borsi V. Calderone V. Fragai M. Luchinat C. Sarti N. Entropic contribution to the linking coefficient in fragment based drug design: a case study. *J. Med. Chem.,* 53, 4285–4289 (2010).
63. Murray CW. Carr MG. Callaghan O *et al.* Fragment-Based Drug Discovery Applied to Hsp90. Discovery of Two Lead Series with High Ligand Efficiency. *J Med. Chem.* 53, 5942-5955 (2010).
64. Copeland RA. Conformational adaptation in drug–target interactions and residence time. *Fut. Med. Chem.* 3, 1491-1501, 2011.
65. Torres F. Kuhn P. De Bruyker D *et al.* Enthalpy arrays. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9517-9522 (2004).

key words/technical terms/concepts

ADMET PROPERTIES

Absorption, distribution, metabolism, excretion and toxicological properties

THERMODYNAMIC PROFILE

The balance between enthalpic and entropic contributions to binding free energy

ENTHALPY-ENTROPY COMPENSATION

In the context of binding thermodynamics enthalpy-entropy compensation refers to the observation that structural changes of either of the ligand or the protein typically lead to significant changes in both enthalpy and entropy but these changes are largely compensating and result in a smaller change in binding free energy.

Table 1 Definitions of efficiency indices

Definition	**Reference**
$LE=\Delta G/HA \sim 1.37pIC_{50}/HA \sim pIC_{50}/HA$	41
$BEI= pIC_{50}/MW$	42
$SILE=pIC_{50}/HA^{0.3}$ or $pK_i/HA^{0.3}$	43
$LLE= pIC_{50} - logP$	21
$LELP=logP/LE$	1
$LLE_{AT}=0.111 + 1.37(LLE/HA)$	52
$EE=\Delta H/HA$ or $\Delta H/HA_{pol}$	45
$SIHE=\Delta H/(2.303RT)/40*HA^{0.3}$	14

Table 2 Changes in various molecular properties and efficiency indices in different type of medicinal chemistry optimizations

Process	Ref	N	pPot change	MW change	logP change	LE change	SILE change	LLE change	LELP change
Fragment successful	44	14	3.05	168.1	0.7	-0.02	0.84	2.6	1.5
Lead opt. successful	**Error! Reference source not found.**	60	2.08	89.9	0.05	0.01	0.85	2.1	-1.1
Thermo opt.	[a]	30	1.27	63.8	0.01	0.01	0.37	1.22	0.09

[a] Calculated from data in Supplementary Table S5

Figure captions

Figure 1 Apolar surface area (*ASA,* left vertical axis)) and the estimated binding free-energy contribution (right vertical axis) versus the number of heavy atoms (*HA*) (see text for details). Straight line indicates maximal surfaces and free-energies. Adapted from ref 17. Copyright (2012) American Chemical Society.

Figure 2 Average enthalpic (pK_H) and entropic (pK_S) contributions to binding versus heavy atom count (*HA*). Standard errors are shown by vertical bars. Data points represent 757 ligand protein complexes [14].

Figure 3 pK_d of the highest affinity compounds and pK_H of the most enthalpic high affinity compounds versus the number of heavy atoms (*HA*) (see text for further details). Standard errors are shown by vertical bars. Data points represent 757 ligand protein complexes [14].

Figure 4 Average enthalpic (pK_H) and entropic (pK_S) contributions to binding versus affinity (pK_d). Standard errors are shown by vertical bars. Data points represent 757 ligand protein complexes [14].

Figure 5 Enthalpic and entropic components of binding for 284 fragment-protein complexes [14]. The contours represent density of points.

Figure 6 Optimization of diaminopyrimidine-type renin inhibitors

Figure 7 a) Thermodynamic profile of renin inhibitors. b) Thermodynamics of the optimization path of renin inhibitors

Figure 8 Optimization of tankyrase inhibitors

Figure 9 a) Thermodynamic profile of tankyrase inhibitors. b) Thermodynamics of the optimization path of tankyrase inhibitors

Figure 10 Optimization of MMP12 inhibitors

Figure 11 a) Thermodynamic profile of MMP12 inhibitors b) Thermodynamics of the optimization path of MMP12 inhibitors

Figure 12 Optimization of aminopyrimidine-type HSP90 inhibitors

Figure 13 Thermodynamic profile of HSP90 inhibitors

Figure 1

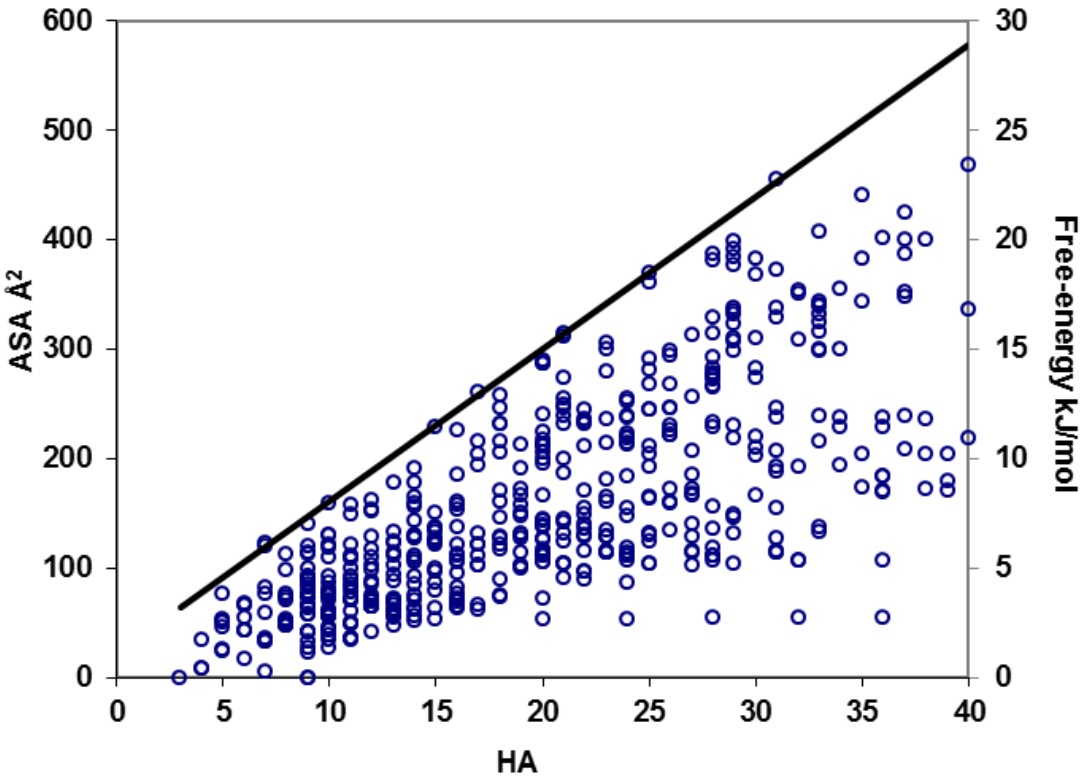

600
500
400
300
200
100
0
ASA Å²
30
25
20
15
10
5
0
Free-energy kJ/mol
0
5
10
15
20
25
30
35
40
HA

Figure 2

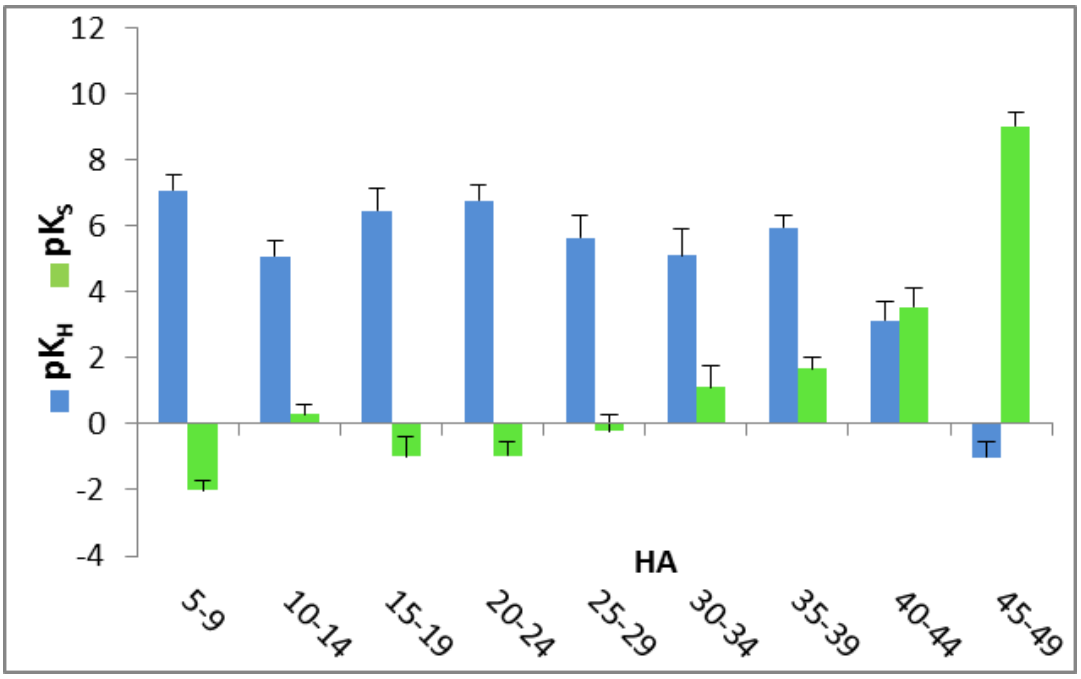

12
10
8
6
4
2
0
-2
-4
pK_H
pK_S
HA
5-9
10-14
15-19
20-24
25-29
30-34
35-39
40-44
45-49

Figure 3

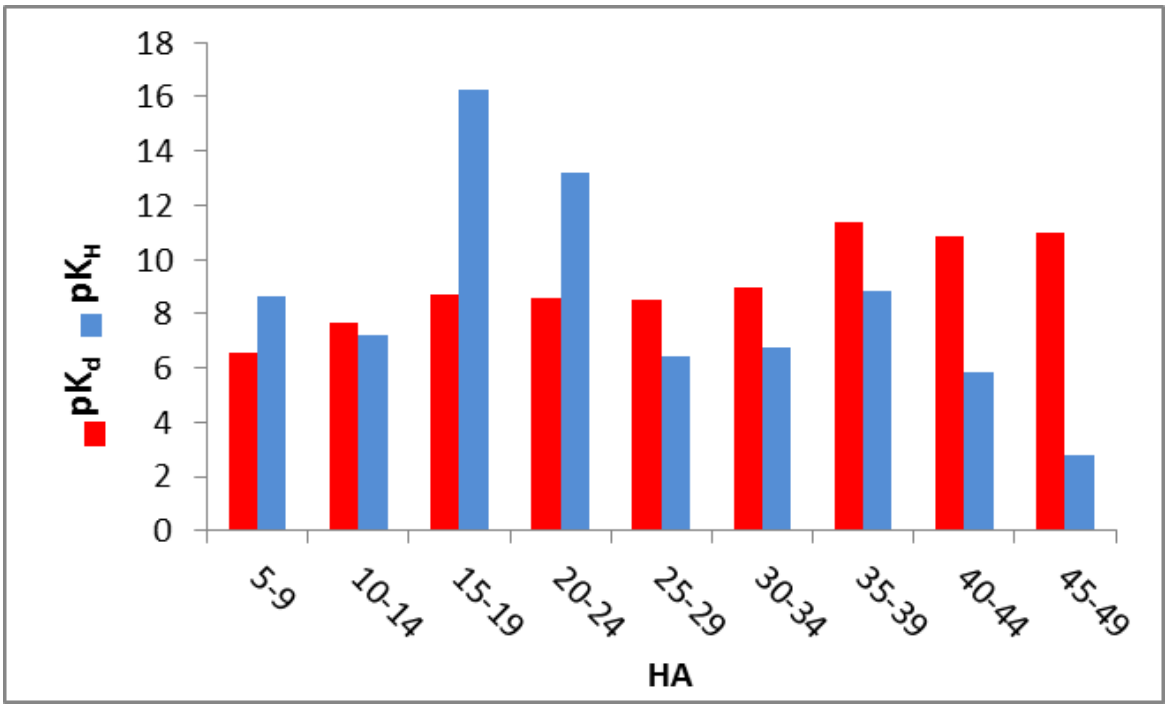

18
16
14
12
10
8
6
4
2
0
pKd
pKH
5-9
10-14
15-19
20-24
25-29
30-34
35-39
40-44
45-49
HA

Figure 4

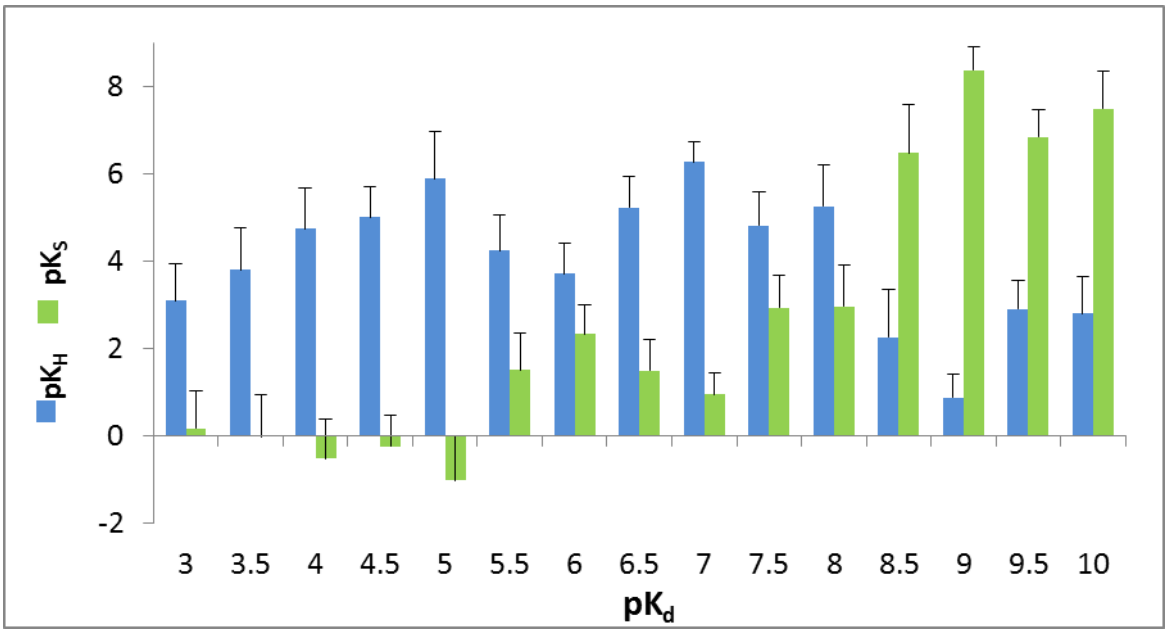


Figure 5

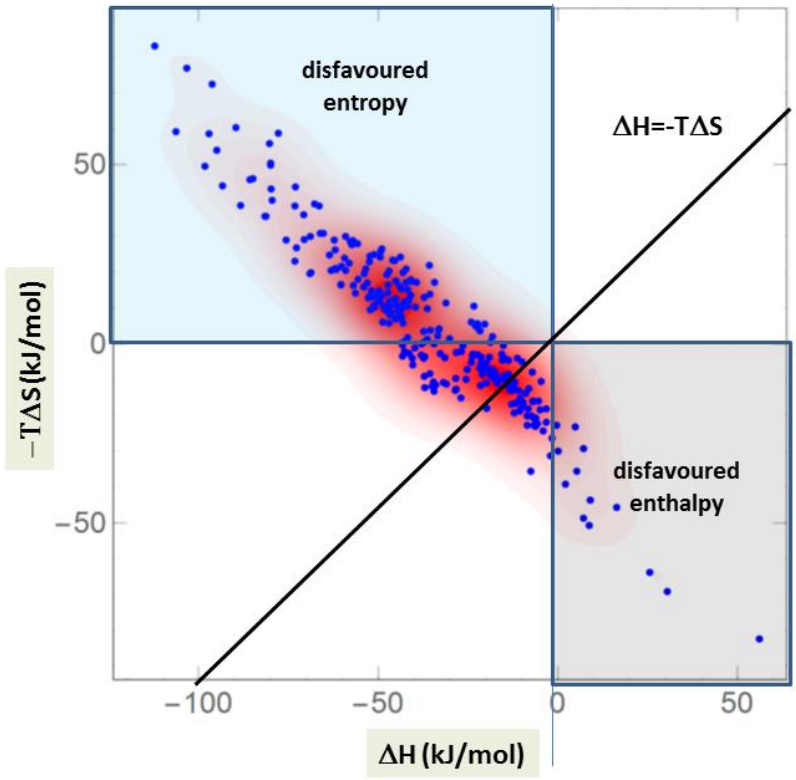
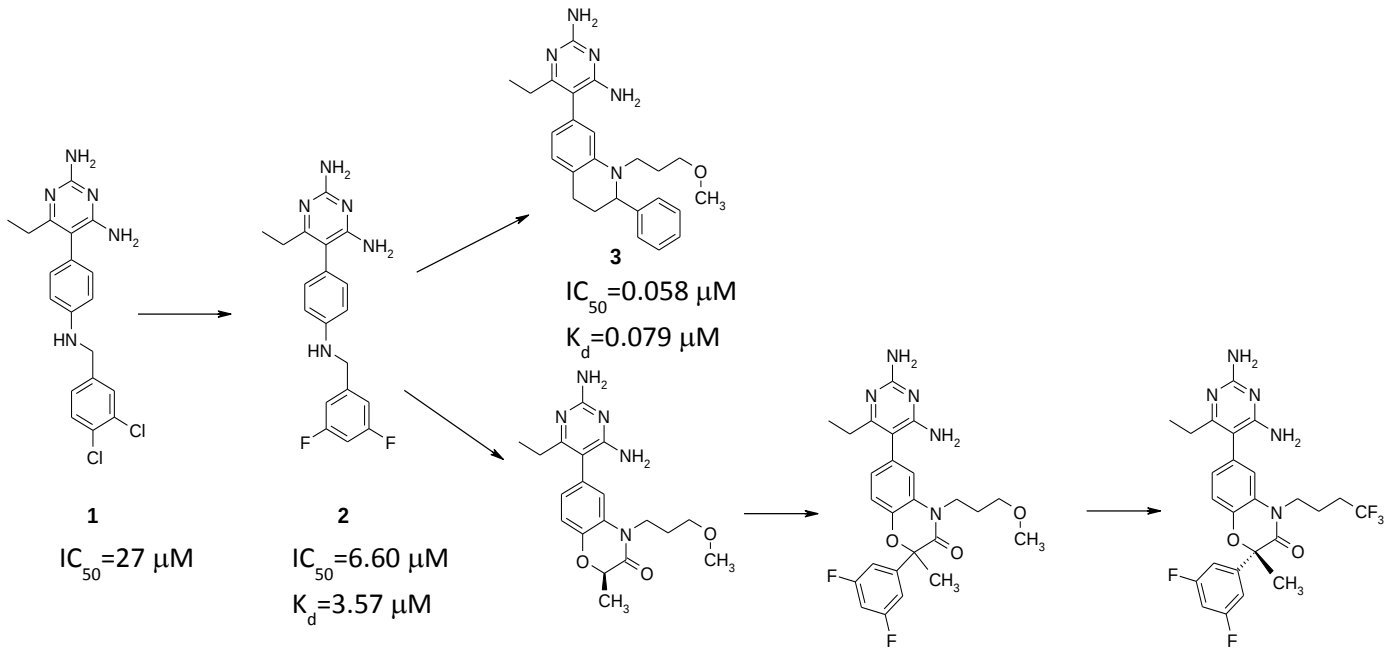

disfavoured
entropy
ΔH=-TΔS
disfavoured
enthalpy
50
0
−50
−100
−50
0
50
−TΔS(kJ/mol)
ΔH (kJ/mol)

Figure **6**



1

IC_{50}=27 μM

2

IC_{50}=6.60 μM

K_d=3.57 μM

3

IC_{50}=0.058 μM

K_d=0.079 μM

4

IC_{50}=0.132 μM

K_d=0.178 μM

5

IC_{50}=0.007 μM

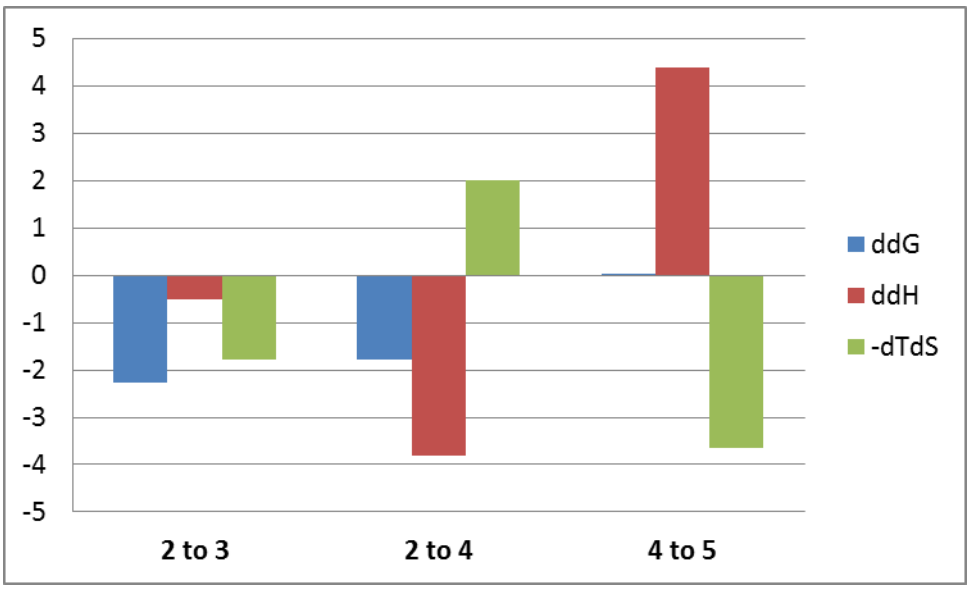
K_d=0.186 μM

6

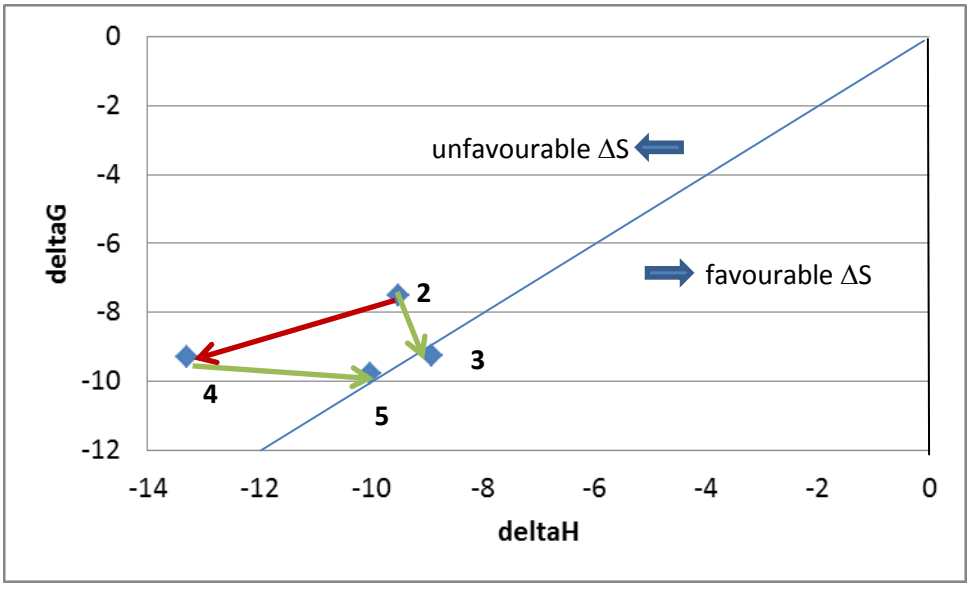
IC_{50}=0.048 μM

Figure **7**

a)

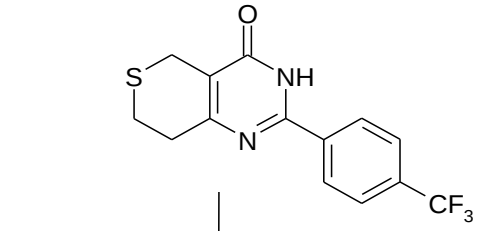


b)



ΔG=ΔH

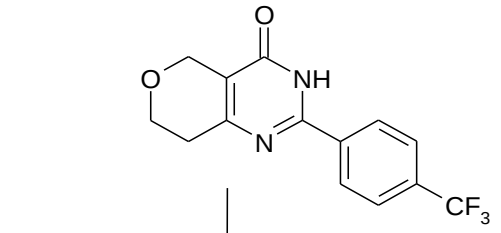
Figure **8**



7

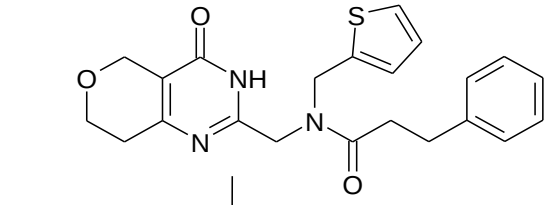
IC_{50}=0.0128 μM

K_d=0.014 μM



8

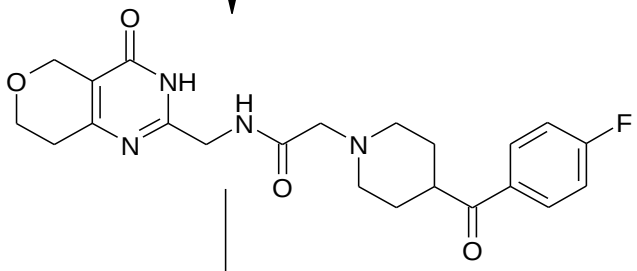
K_d=0.194 μM



9

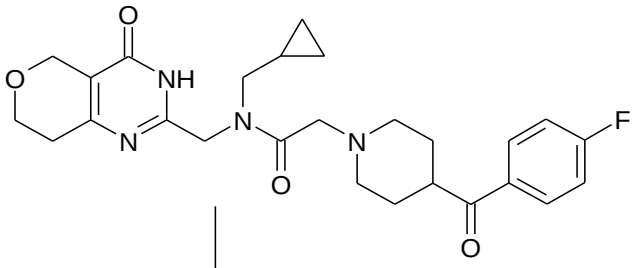
IC_{50}=1.15 μM

K_d=0.676 μM



10

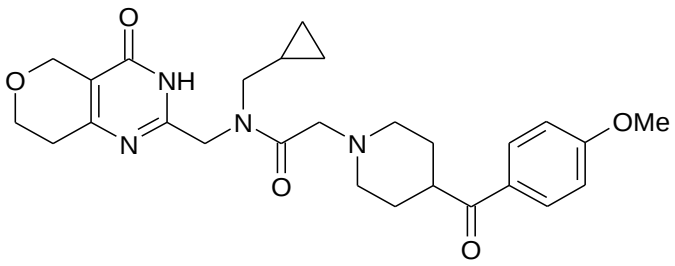
IC_{50}=0.0645 μM



11

IC_{50}=0.0347 μM

K_d=0.091 μM



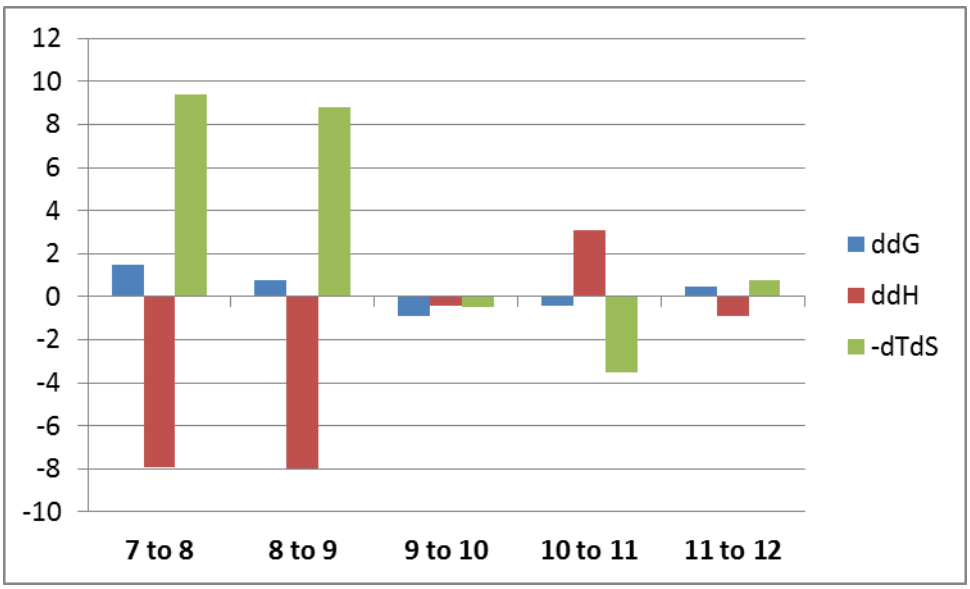
12

IC_{50}=0.0155 μM

K_d=0.083 μM

Figure **9**

a)


12
10
8
6
4
2
0
-2
-4
-6
-8
-10
7 to 8
8 to 9
9 to 10
10 to 11
11 to 12
ddG
ddH
-dTdS

b)

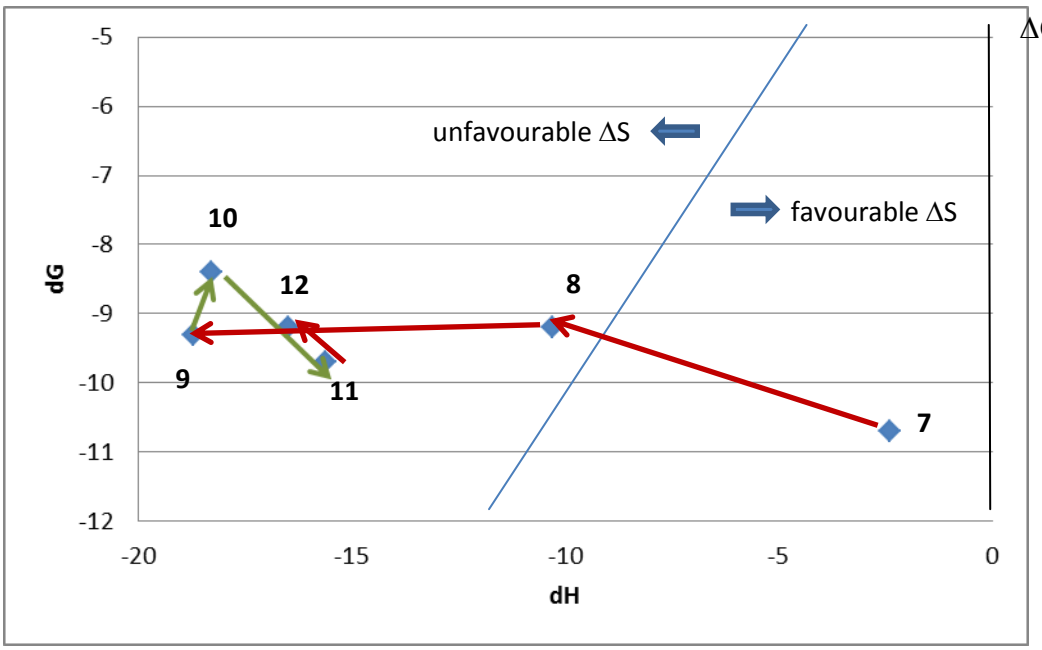

ΔG=ΔH
unfavourable ΔS
favourable ΔS
10
12
8
9
11
7
dG
dH
-5
-6
-7
-8
-9
-10
-11
-12
-20
-15
-10
-5
0

Figure **10**

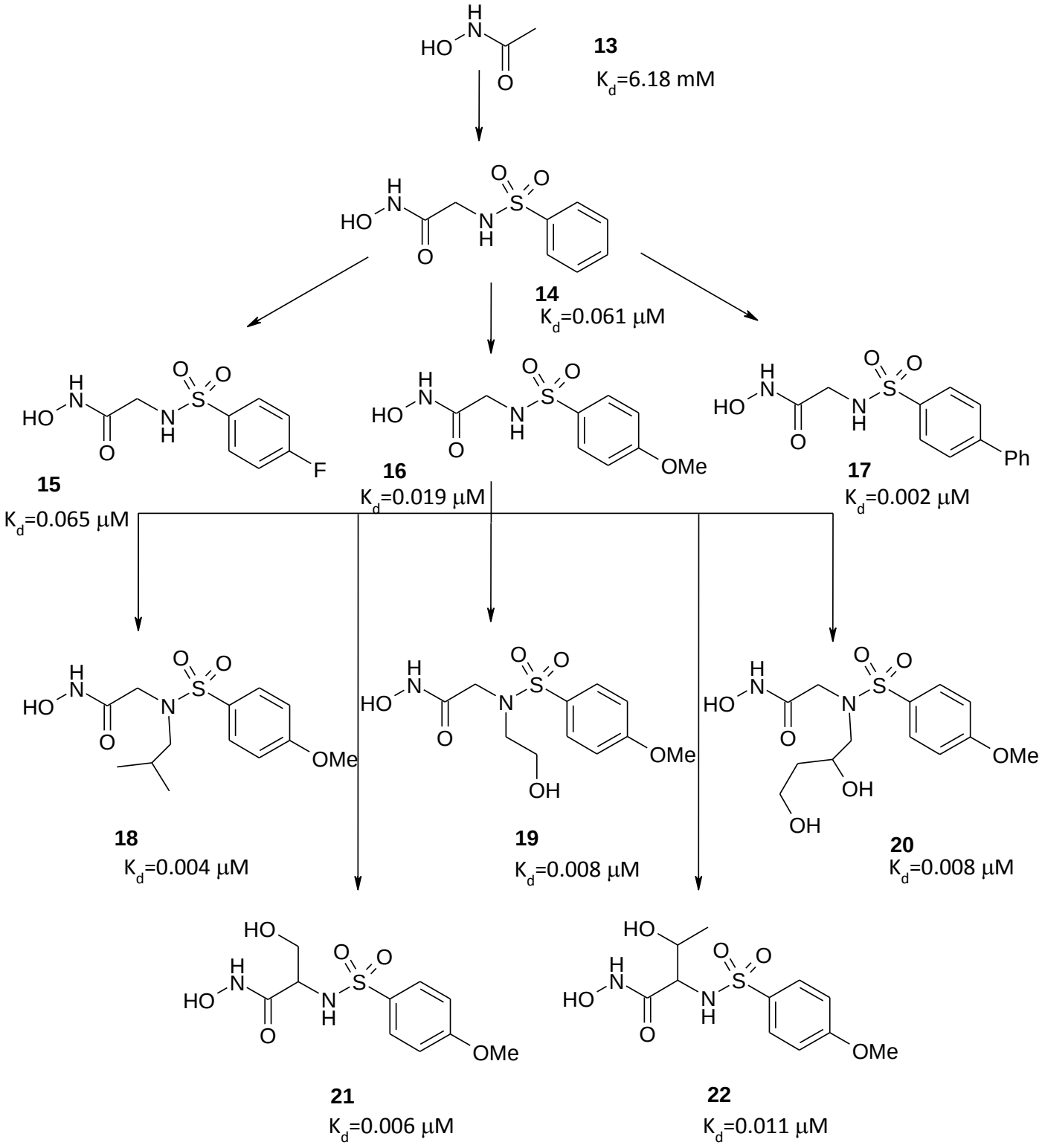
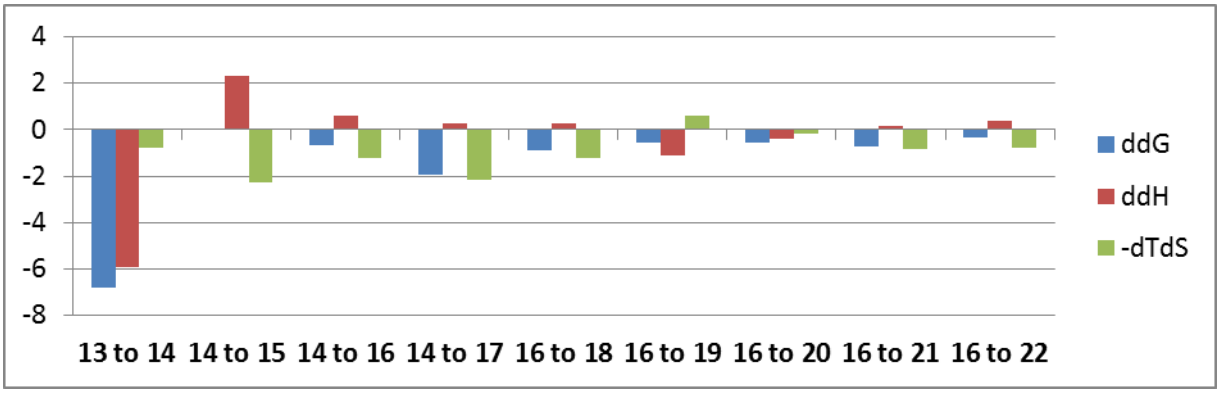

13
K_d=6.18 mM
14
K_d=0.061 μM
15
K_d=0.065 μM
16
K_d=0.019 μM
17
K_d=0.002 μM
18
K_d=0.004 μM
19
K_d=0.008 μM
20
K_d=0.008 μM
21
K_d=0.006 μM
22
K_d=0.011 μM

Figure **11**

a)


4
2
0
-2
-4
-6
-8
13 to 14
14 to 15
14 to 16
14 to 17
16 to 18
16 to 19
16 to 20
16 to 21
16 to 22
ddG
ddH
-dTdS

b)

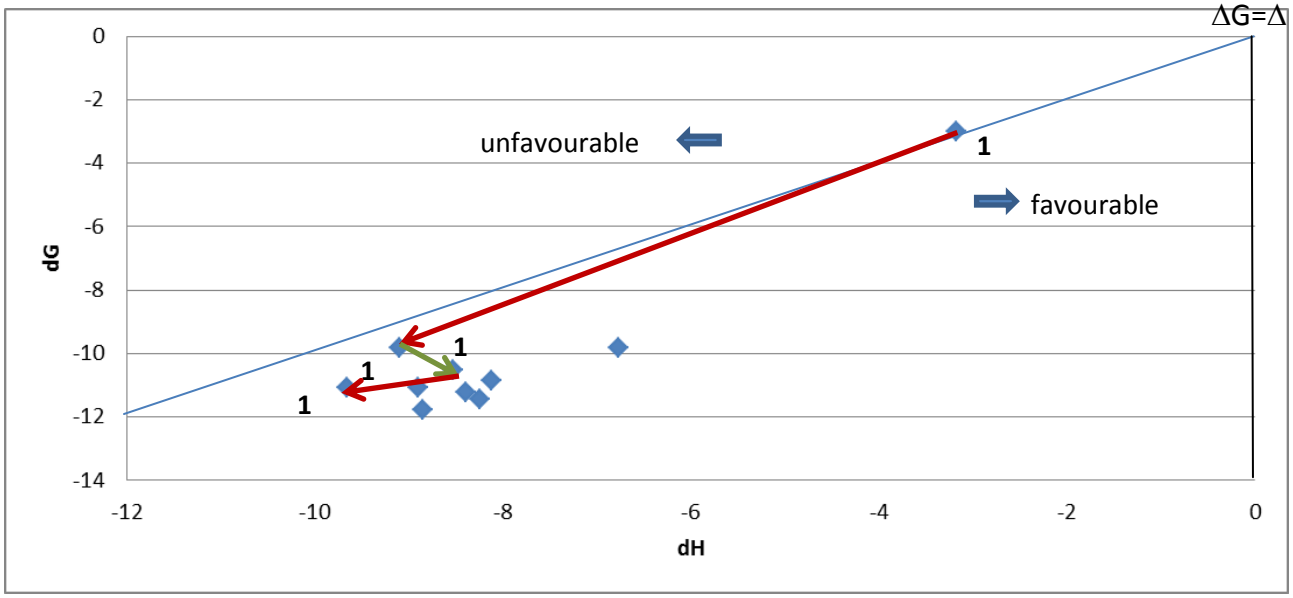
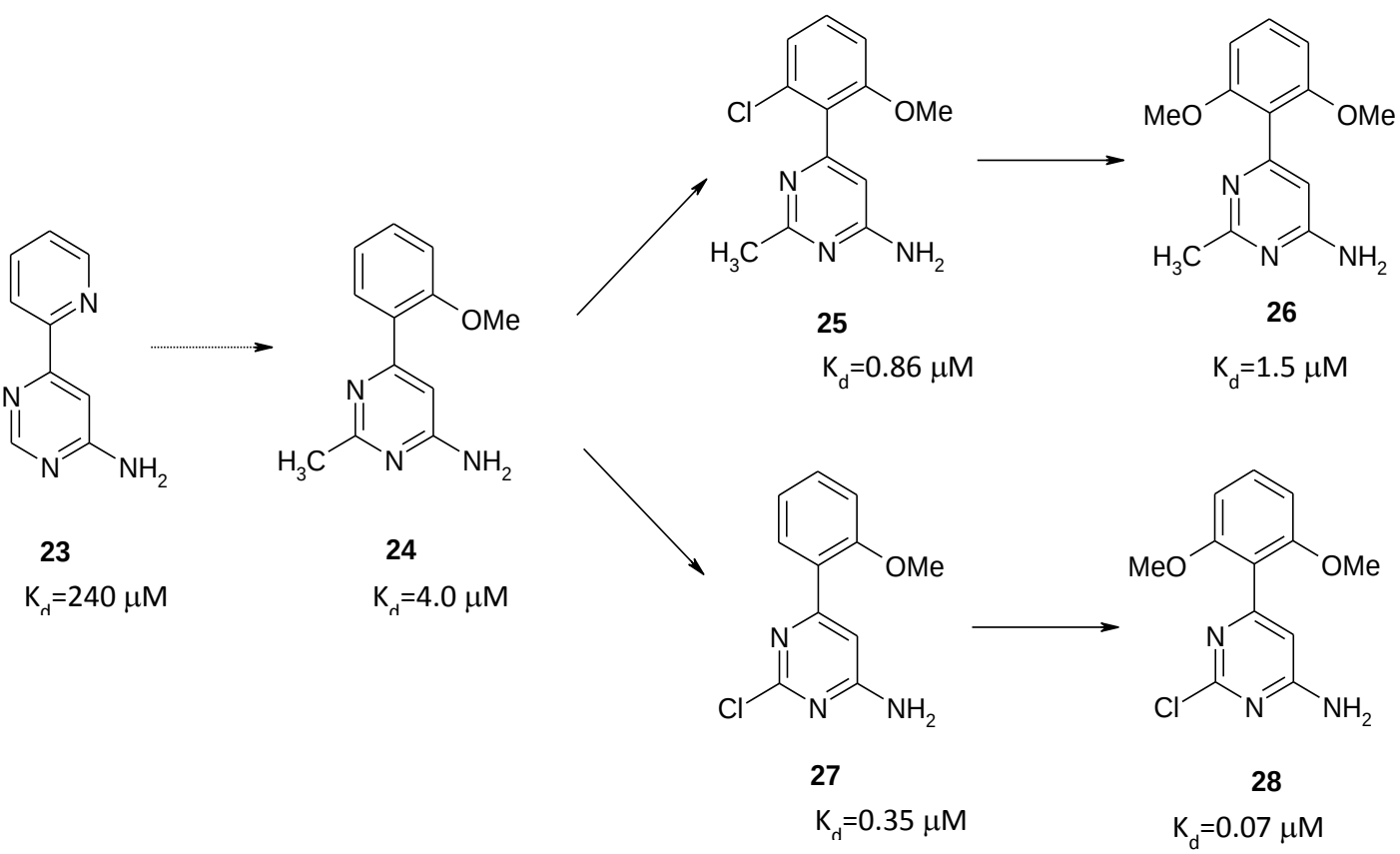

ΔG=Δ
0
-2
-4
-6
-8
-10
-12
-14
dG
-12
-10
-8
-6
-4
-2
0
dH
unfavourable
favourable
1
1
1
1

Figure **12**



23
K_d=240 μM

24
K_d=4.0 μM

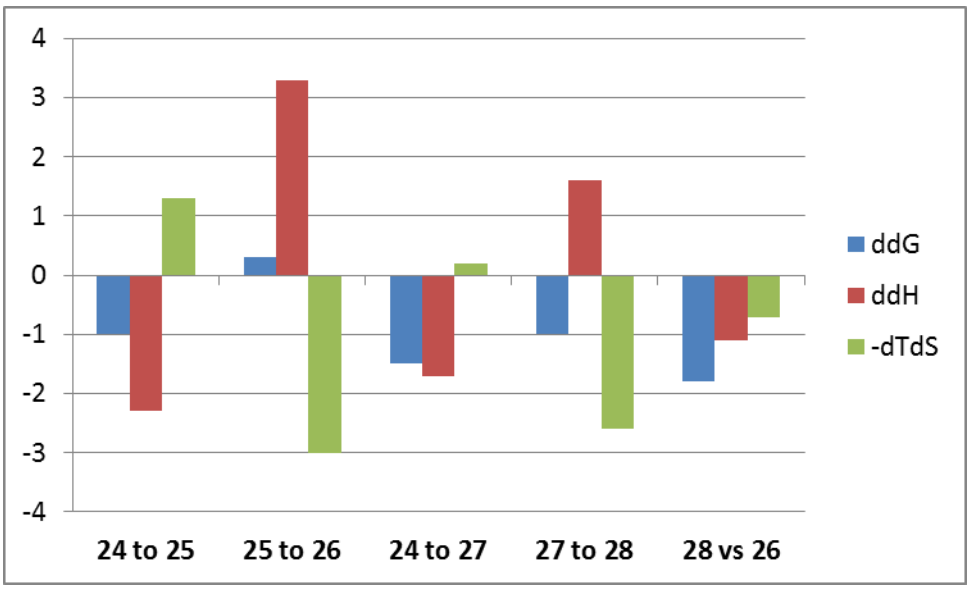
25
K_d=0.86 μM

26
K_d=1.5 μM

27
K_d=0.35 μM

28
K_d=0.07 μM

Figure **13**



Box 1

BINDING THERMODYNAMICS

The improvement of ligand affinity is equivalent with the decrease of the ligand binding free-energy (ΔG_{bind}) corresponding to the ligand-protein binding process $L + P \leftrightarrow LP$

$$\Delta G_{bind} = \Delta H - T\Delta S$$

ΔG_{bind} relates to the dissociation constant $K_d = \frac{[L][P]}{[LP]}$, often expressed as $pK_d = -\log(K_d)$ via the equation

$$\Delta G_{bind} = RTlnK_d$$

Using the above equations one can write [14]

$$pK_d = pK_H + pK_S$$

with $pK_H = \frac{-\Delta H}{2.303 \cdot RT}$ and $pK_S = \frac{\Delta S}{2.303 \cdot R}$. The use of pK_H and pK_S instead of ΔH and $T\Delta S$ has the advantage that their higher values correspond to more favorable contributions and their sum gives pK_d, a quantity generally used for characterizing ligand affinity.