ELSEVIER

A large thermodynamic dataset from Astex, AstraZeneca, Pfizer and academic labs that includes fragment–protein interactions demonstrates that, when compared with many traditional druglike compounds, fragments bind more enthalpically to their protein targets.

Binding thermodynamics discriminates fragments from druglike compounds: a thermodynamic description of fragment-based drug discovery

Glyn Williams[1], **György G. Ferenczy**[2], **Johan Ulander**[3] and **György M. Keserű**[2]

[1] Astex Pharmaceuticals, 436 Cambridge Science Park, Milton Road, Cambridge CB4 0QA, UK
[2] Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok krt. 2, H-1117 Budapest, Hungary
[3] CVMD Innovative Medicines, AstraZeneca R&D Mölndal, S-43183 Mölndal, Sweden

Small is beautiful – reducing the size and complexity of chemical starting points for drug design allows better sampling of chemical space, reveals the most energetically important interactions within protein-binding sites and can lead to improvements in the physicochemical properties of the final drug. The impact of fragment-based drug discovery (FBDD) on recent drug discovery projects and our improved knowledge of the structural and thermodynamic details of ligand binding has prompted us to explore the relationships between ligand-binding thermodynamics and FBDD. Information on binding thermodynamics can give insights into the contributions to protein–ligand interactions and could therefore be used to prioritise compounds with a high degree of specificity in forming key interactions.

Ligand size and binding thermodynamics

The maximal available binding affinity depends on ligand size and this observation appears to support the medicinal chemistry practice that adds new functional groups to improve affinity. By contrast, it has been claimed that the maximal available enthalpy gain decreases with increasing ligand size or, viewed alternatively, that medicinal chemistry optimisation has traditionally tended to enhance affinity mainly for entropic reasons [1–3]. However, this could be an over-simplification when considering the effects of small structural changes between similar ligands [4]. Although high favourable enthalpy is accompanied by high affinity in the case of small ligands, this is not necessarily true for large ligands, where higher-affinity compounds bind typically with lower enthalpy gain and further improvement in binding enthalpy is often achieved at the expense of affinity. These observations are based on the analysis of large, publicly available isothermal titration calorimetry (ITC) datasets, such as the Scorpio (http://scorpio.biophysics.ismb.lon.ac.uk/scorpio.html) and BindingDB (http://www.bindingdb.org/bind/index.jsp) databases. Because binding enthalpy broadly reflects the overall quality of protein–ligand interactions, the opposite size dependence of affinity and enthalpy could have fundamental consequences for drug discovery

Corresponding author: Keserű, G.M. (gy.keseru@ttk.mta.hu)

Glyn Williams joined Astex in 2001, to develop and apply biophysical methods to fragment-based screening and drug design. This work has involved the use of NMR, isothermal titration calorimetry and native mass apectrometry; and also contributed to the development the Astex fragment library. He is Vice President of Biophysics of Astex Pharmaceuticals. Previously, Glyn spent 11 years with Roche UK where he was responsible for biological NMR and analytical mass spectrometry. After obtaining his degree and doctorate (DPhil) in chemistry from the University of Oxford, Glyn held fellowships and lectureships in inorganic and bioinorganic chemistry at the Universities of Oxford, Sydney and London from 1983 to 1990.

György G. Ferenczy received his PhD in computational chemistry from the Eötvös University of Budapest. Following postdoctoral research at the University of Oxford, UK, and at the University of Nancy, France, he worked as a computational chemist and as a group leader first at Gedeon Richter (Budapest) and later at Sanofi (Budapest and Strasbourg). Since 2012, he is a senior research fellow at the Semmelweis University and, from 2013, at the RCNS of the Hungarian Academy of Sciences. His research interests include the development and application of computational tools for extended biochemical systems and studying molecular interactions relevant to drug discovery.

Johan Ulander currently works as Associate Principal Scientist in the computational chemistry section at Cardiovascular and Metabolic Diseases (CVMD) at AstraZeneca R&D Göteborg, Sweden. Before joining AstraZeneca he did post-doctoral research at University of California, San Diego (USCD) and University of Houston. He received his PhD in theoretical physical chemistry from Gothenburg University and has a BS in molecular biology from the University of Umeå, Sweden. He has over 10 years of experience in drug discovery from early-stage hit and target identification to late-stage drug optimisation. His interests include theoretical biophysics and statistical mechanics with applications in drug design, pharmacokinetics and pharmacodynamics.

György M. Keserű obtained his PhD at Budapest, Hungary and joined Sanofi-Aventis CHINOIN, heading a chemistry research lab. He moved to Gedeon Richter in 1999 as the Head of Computer-Aided Drug Discovery. Since 2007, he was appointed as the Head of Discovery Chemistry at Gedeon Richter and contributed to the discovery of the antipsychotic Vraylar™ (cariprazine) which has been approved and marketed in the USA from 2016. From 2013, he served as a director general of the Research Center for Natural Sciences (RCNS) at the Hungarian Academy of Sciences. Now he is heading the Medicinal Chemistry Research Group at RCNS. His research interests include medicinal chemistry and drug design. He has published over 180 papers and more than 15 books and book chapters.

http://dx.doi.org/10.1016/j.drudis.2016.11.019

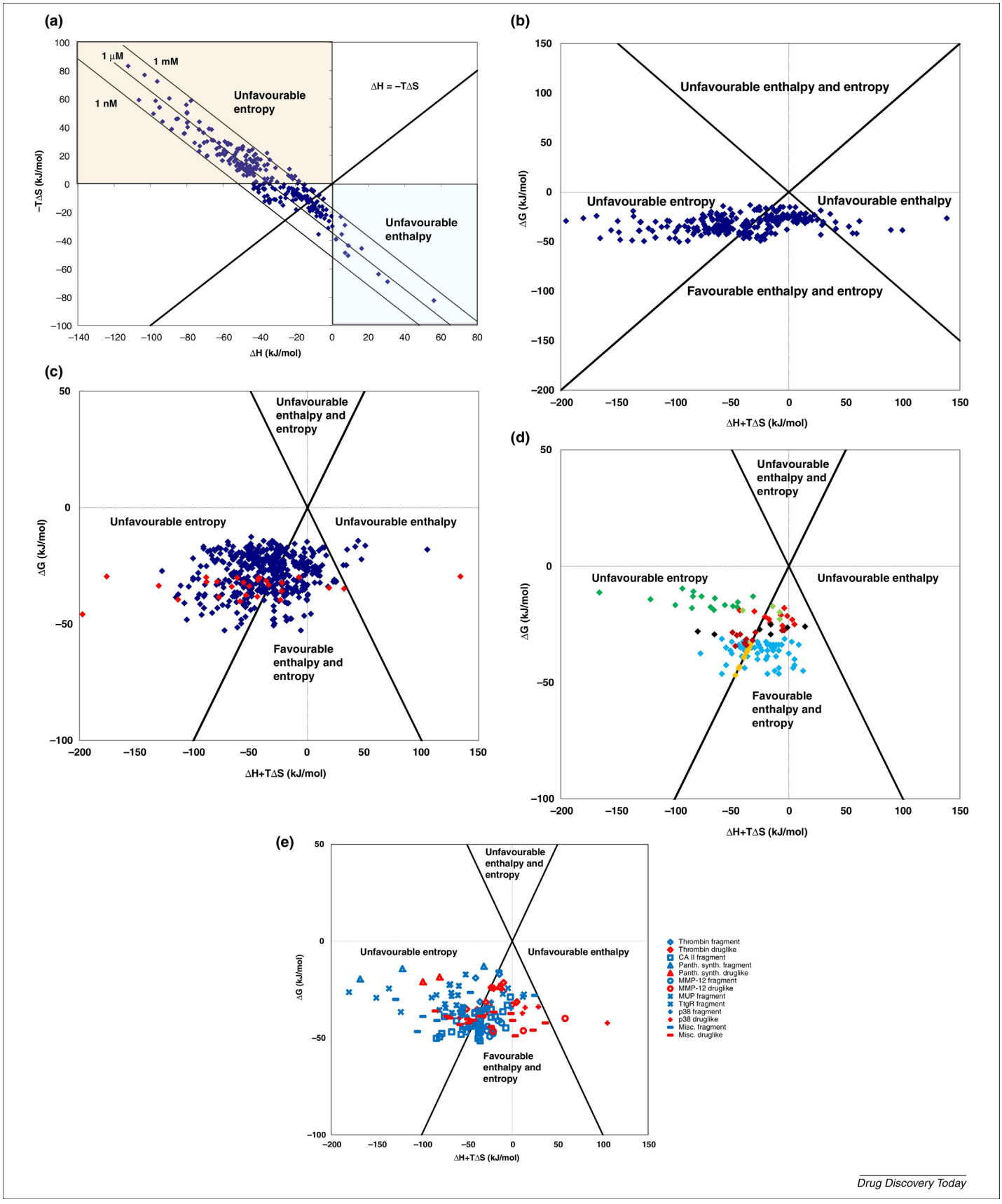


FIGURE 1

Binding thermodynamics data for fragments (also see supplementary material available online). **(a)** $\Delta H_{binding}$ versus $-T\Delta S_{binding}$ for a structurally diverse set of fragments available from the Scorpio, PDBCal and BindingDB databases. **(b)** ($\Delta H_{binding} + T\Delta S_{binding}$) versus $\Delta G_{binding}$ for the same set of structurally diverse fragments available from Scorpio, PDBCal and BindingDB databases. Note that the change of axes corresponds to a 45° rotation of the data in (a) and would allow

practices, including the identification of chemical starting points and their optimisation: (i) more enthalpic binders could prove to be more advantageous starting points for medicinal chemistry optimisations; (ii) enthalpic optimisation could be more beneficial during the early phase of optimisations; and (iii) smaller compounds could be more likely to bind enthalpically than larger compounds.

The demonstrated success of fragment-based drug discovery (FBDD) programmes and the proposed consequences prompted us to investigate the binding of fragment-size compounds by analysing their complexes in the Protein Data Bank (PDB) [1]. It was found that fragments often form a small number (two on average) of near-to-optimal geometry H-bonds. This is a consequence of their small size that allows them to form good-quality H-bonds with low steric constraints. These usually contribute decisively to the binding free energy and this contribution is predominantly enthalpic, overcoming the loss of ligand rigid-body entropy that is also associated with binding. The complexity model of Hann et al. [5] supports the view that ligands can form a limited number of optimal interactions and that their number does not increase with increasing ligand size and complexity. Indeed, it was found that the burial of 50–100 $Å^2$ polar surface area that can be associated with the formation of two H-bonds results in a significant enthalpic benefit but this benefit does not increase with increasing polar surface area burial [6].

Fragments usually bind to a confined region within a larger protein-binding site and this is identified as the hotspot. Hotspots are energetically important regions of the binding site; they can bind a diverse set of small organic compounds [7,8] and they bind fragments in a way that the extension of the fragments to larger compounds does not affect the binding pose [9]. A consequence of the small size of fragments is that their binding to the hotspot disturbs the water network to a lesser extent and in a different manner than large ligands. A particular feature of hotspots is that they are often associated with water molecules which have unfavourable excess entropy [10]. Using inhomogeneous fluid solvation theory Huggins estimated the enthalpic and entropic contributions of individual water molecules in 19 protein cavities of five proteins [11], and concluded that the contribution of entropic penalty of water molecules in protein cavities might be small to the free energy. These observations suggest that the small number of such water molecules released by fragment binding is usually unable to turn the entropy balance. This is well illustrated by the binding of a series of fragments to carbonic anhydrase where the release of water molecules plays a decisive part in determining the relative enthalpy and entropy content of binding. Nevertheless, the overall contribution of water release to the binding enthalpy is a fraction of the observed enthalpy gain and does not affect the substantially favourable enthalpy [12].

Apolar desolvation contributes to the binding of larger ligands and it was shown that the burial of ~20 $Å^2$ apolar Connolly surface upon binding typically leads to ~1 kJ/mol free energy gain [6]. Although this dominantly entropic contribution is significant for large ligands, it is less important for fragments owing to their small size and buried surface. Moreover, the rigid-body entropy loss upon ligand–protein binding amounts to ~15–20 kJ/mol [13] that must be compensated before entropically favoured binding is observed. This latter entropy loss only slightly increases with ligand size and for this reason its contribution is relatively more important for fragments than for larger ligands. The observations that fragments can achieve higher favourable binding enthalpy than larger ligands, that they can form optimal geometry H-bonds in the protein hotspot without incurring a large apolar desolvation penalty and that this occurs without significantly disturbing the water network in the binding site strongly suggest that fragments would be expected to bind enthalpically.

Experimental thermodynamics data of fragment binding

There is a large body of experimental evidence that is in line with the expectations discussed above: fragments bind to proteins with favourable enthalpy. These data come most abundantly from direct measurements of enthalpy. With recent methodological and technical developments [14], ITC experiments have become feasible for low-affinity complexes, allowing the accurate measurement of significant enthalpies ($|\Delta H_{binding}| > 5$ kJ/mol) for binding affinities in the range 100 $\mu M < K_d < \sim 1$ mM (low-c titrations) and 10 $nM < K_d < 100$ μM for direct titrations [15].

We investigated the binding thermodynamic profile of fragments reported in the biomedical literature. Binding thermodynamics data of this set of diverse 284 fragments are shown in Fig. 1a,b and indicate that the majority of the fragments bind with a favourable enthalpy change [1]. The few exceptions found in the public dataset are all charged compounds and the observed entropy dominated binding is in line with the large enthalpic penalty of de-solvation for ions (Table 1). Figure 1a gives a traditional representation of the thermodynamics binding data, where ΔH is plotted against $-T\Delta S$. Large areas of such graphs are empty, corresponding to complexes where the affinities are too weak to measure ($\Delta G > 0$ kJ/mol) or too tight to achieve with typical noncovalent interactions ($\Delta G < -60$ kJ/mol). Data in Fig. 1b and those in all subsequent figures are shown as ΔG versus $\Delta H + T\Delta S$. This change of axes leads to a 45° rotation of the data when compared with the more usual representation of ΔH versus $-T\Delta S$ in Fig. 1a. The quantity $(-1/T)$. $(\Delta H + T\Delta S)$ has a physical meaning, corresponding to the difference between the entropy created in the system (ΔS) and outside the system ($-\Delta H/T$) for a closed system undergoing a spontaneous change. Compounds with favourable enthalpy and entropy appear in the lower-middle triangle whereas those with unfavourable enthalpy are above the right diagonal and those with unfavourable entropy are above the left diagonal. Areas that contain no data points can be excluded by restricting the ΔG axis scale.

This representation also suggests the use of the ratio $(\Delta H + T\Delta S)/\Delta G$ as a measure of the enthalpic driving force. This quantity is zero when enthalpy and entropy contribute equally to the free energy

areas containing no data points to be discarded ($\Delta G < -60$ kJ/mol and $\Delta G > 0$ kJ/mol). **(c)** Fragment thermodynamics data from drug discovery programmes (blue: Astex, red AstraZeneca). **(d)** Fragments from screening efforts against different targets (green: pantothenate synthetase, light blue: p38α, orange: carbonic anhydrase, red: trypsin, black: PLP-dependent transaminase, light green: thrombin, dark red: Pqsr). **(e)** Enthalpic and entropic components of binding for complexes of neutral fragments and druglike molecules.

TABLE 1

Properties of ligands and targets from Fig. 1b,c that have unfavourable binding enthalpies

Complex	ΔH (kJ/mol)	Target (class)	Ligand MW (Da)	Ligand *c* log *P*	Calculated charge (pH 7.0)	Ionisable groups
Public 1	56.1	Osteoclast farnesyl pyrophosphate synthase (synthase)	245	−0.9	−2	Pyrophosphate
Public 2	30.6	Farnesyl diphosphate synthase (synthase)	409	−0.6	−2	Bisphosphonate
Public 3	25.6	Farnesyl diphosphate synthase (synthase)	319	0.7	−2	Bisphosphonate
Public 4	16.3	Osteoclast farnesyl pyrophosphate synthase (synthase)	283	−0.3	−2	Bisphosphonate
Public 5	9.0	Streptavidin–synthetic construct (biotin-binding domain)	270	4.6	−1	Carboxylic acid (aromatic)
Public 6	8.8	Osteoclast farnesyl pyrophosphate synthase (synthase)	272	−1.2	−2	Bisphosphonate
Public 7	7.2	Farnesyl diphosphate synthase (synthase)	272	−1.2	−2	Bisphosphonate
Public 8	7.1	Streptavidin–synthetic construct (biotin-binding domain)	242	3.6	−1	Carboxylic acid (aromatic)
Public 9	5.4	Streptavidin–synthetic construct (biotin-binding domain)	256	4.1	−1	Carboxylic acid (aromatic)
Public 10	4.8	CRP mutant (phosphocholine binding)	329	−3.0	−1	Phosphonic acid
Public 11	2.2	Farnesyl diphosphate synthase (synthase)	235	−2.1	−2	Bisphosphonate
Public 12	0.2	Concanavalin A dimer (carbohydrate binding)	299	3.6	−1	Sulfonic acid
Astex 1	43.5	Target A (protein–protein interaction)	223	3.5	−1	Carboxylic acid (aromatic)
Astex 2	12.6	HSP90 (ATPase)	249	3.4	0	Phenol
Astex 3	8.4	HCV NS3 (protease)	246	−0.9	1	Tertiary amine
Astex 4	4.6	HSP90 (ATPase)	235	3.0	0	Phenol
Astex 5	3.3	Target B (synthase)	144	1.7	0.9	Aminopyridine
AZ 1	58.5	Target C (undisclosed)	189	2.8	1	Carboxylic acid

Abbreviations: AZ, AstraZeneca; CRP, cAMP Receptor Protein; HCV, hepatitis C virus; HSP, heat shock protein; MW, molecular weight.

of binding and is positive for enthalpy-driven binding ($\Delta H < -T\Delta S$). Its value exceeds 1 when enthalpy must overcompensate an unfavourable binding entropy (compounds above the left diagonal in Fig. 1b). By contrast, a negative value of the ratio indicates an entropy-driven binding whereas values smaller than −1 correspond to a favourable entropy that overcompensates an unfavourable enthalpy (compounds above the right diagonal of Fig. 1a). Consequently, we define the dimensionless ratio $(\Delta H + T\Delta S)/\Delta G$ as the Enthalpy–Entropy Index (I_{E-E}) and use it here to indicate the enthalpy content of binding. Its advantageous feature is that it is normalised by the free energy ΔG (= $\Delta H - T\Delta S$), and so it can be used to compare compounds with millimolar to nanomolar binding affinities during the course of a hit-to-lead optimisation.

Thermodynamic binding data for a larger set of carefully selected fragments were obtained at Astex and AstraZeneca during the course of their drug discovery programmes (Fig. 1c). The Astex data shown in Fig. 1c include a total of 782 ITC measurements obtained from fragment screening hits, optimised fragment hits, fragment-derived leads and optimised leads, binding to a diverse set of 24 target proteins. This dataset complements those from public sources and shows that an overwhelming majority of compounds (98.7%) investigated at Astex bind with favourable enthalpy. The best-fit line suggests entropies of binding are, on average, 23 kJ/mol less favourable than binding enthalpies. Similar trends were found in a smaller set of fragments identified in AstraZeneca drug discovery programmes. These data also demonstrate that fragment hits bind dominantly by favourable enthalpy and this feature can be preserved while they are optimised into leads and drug candidates.

Twelve complexes from the public dataset, ten complexes of the Astex dataset and one complex from the AstraZeneca dataset display positive (unfavourable) binding enthalpies in Fig. 1b,c. More information on these complexes is provided in Table 1. It should be noted that data for five of the Astex complexes with unfavourable enthalpy have been omitted from Table 1 because the data were obtained using a displacement ITC experiment format. In these cases, the measured enthalpies also depend on an accurate knowledge of the binding enthalpy of the competitor ligand and so could contain larger errors than the complexes measured using direct ITC methods.

The ligands in the 18 complexes of Table 1 have diverse chemical structures and physicochemical properties that are reasonably representative of a set of druglike hits and leads, with $MW_{a\text{-}ve}$ = 262 Da, *c* log P_{ave} = 1.4 and a calculated charge between +1 and −2. Similarly, the protein targets are diverse and span several different target classes, containing the substrate and cofactor-binding sites of enzymes, and the protein–protein and protein–small-molecule binding domains of recognition modules. One ligand, present in Astex 5 complex, also binds to a second target in the full Astex set with $K_d \sim 200$ μM and $\Delta H \sim -20$ kJ/mol and so its unfavourable binding enthalpy to target B cannot be ascribed to the fragment alone. Conversely, all of the Astex targets in Table 1 form many complexes that have favourable enthalpies and so the observation of an unfavourable enthalpy is also not a property of the target alone.

Although the complexes of Table 1 are clearly distinct from the overwhelming majority of complexes described in Fig. 1, an understanding of their unusual thermodynamics signatures requires additional information that will probably include changes in the

protonation and hydration states of the ligand and target when the complex is formed. It should also be noted that the ligands described in Table 1 are not necessarily poor starting points for drug design. The ligands contained in the complexes Astex 2 and Astex 4 were identified by fragment screening of the ATPase domain of HSP90 and were starting points for the development of onalespib (AT13387), a molecule that is currently in Phase II clinical trials for cancer [16,17]. The ligand of Astex 4 complex corresponds to compound 19 from [14] and has $I_{E-E} = -1.5$. Addition of just two atoms to give compound 30 from [14] improved its binding affinity by almost 10^4-fold, its binding enthalpy by 39 kJ/mol and increased its enthalpic efficiency to $I_{E-E} = 0.5$. Onalespib itself binds to the HSP90 ATPase domain with an affinity of 0.7 nM, an enthalpy of −41 kJ/mol and $I_{E-E} = 0.6$ [17]. However, it should be stressed that such behaviour is unusual; as will be shown below, fragment hits generally have values of I_{E-E} close to unity ($-T\Delta S \sim 0$) and this value decreases significantly during a typical lead optimisation process.

Further thermodynamic data for several protein targets are available from fragment-screening campaigns, from fragment optimisations and from systematic analyses of congeneric series described in the literature. Figure 1d shows target-specific binding thermodynamic data that include 11 fragments tested against trypsin [18], 16 and seven fragment hits against *Mycobacterium tuberculosis* pantothenate synthetase [19] and PLP-dependent transaminase (BioA) [20], respectively, 58 fragments screened against p38α [21], 11 fragments targeting *Pseudomonas aeruginosa* PqsR protein [22], 20 congeneric fragments binding to human carbonic anhydrase II [23] and four fragments measured by direct ITC against thrombin [24]. All of these 127 fragments bind with favourable enthalpy. More recently, we compiled [25] a dataset of 138 neutral fragments (94) and druglike compounds (44) acting on 17 targets that showed the pronounced tendency of fragments to bind with more-favourable enthalpy and less favourable entropy with respect to druglike compounds binding to the same targets (Fig. 1e). It is important to note that the compounds in [25] were all evaluated by direct ITC measurement (92% of the compounds in the dataset show K_d values lower than 100 μM) and no displacement experiments were included. A statistical analysis of ΔG, ΔH and $-T\Delta S$ values on this dataset showed statistically significant differences for the enthalpic and entropic components of fragments and druglike compounds (Mann–Whitney U-test, $P < 0.005$). These data also imply that, on average, the relative contribution of the enthalpic component to fragment binding, measured using I_{E-E}, is greater than that observed for druglike compounds. There is a statistically significant difference at the $P = 0.0009$ significance level between the value of I_{E-E} for fragments and druglike compounds with medians 1.10 and 0.79, respectively (Fig. 2). Similar to publicly available databases, fragment thermodynamic data from corporate and academic drug discovery laboratories (a total of almost 1000 data points) collected for a wide variety of targets show that fragments bind with favourable enthalpy. Moreover, the binding enthalpy dominates in the large majority of cases. This clearly distinguishes fragments from larger compounds where such preference for enthalpy-dominated binding cannot be observed.

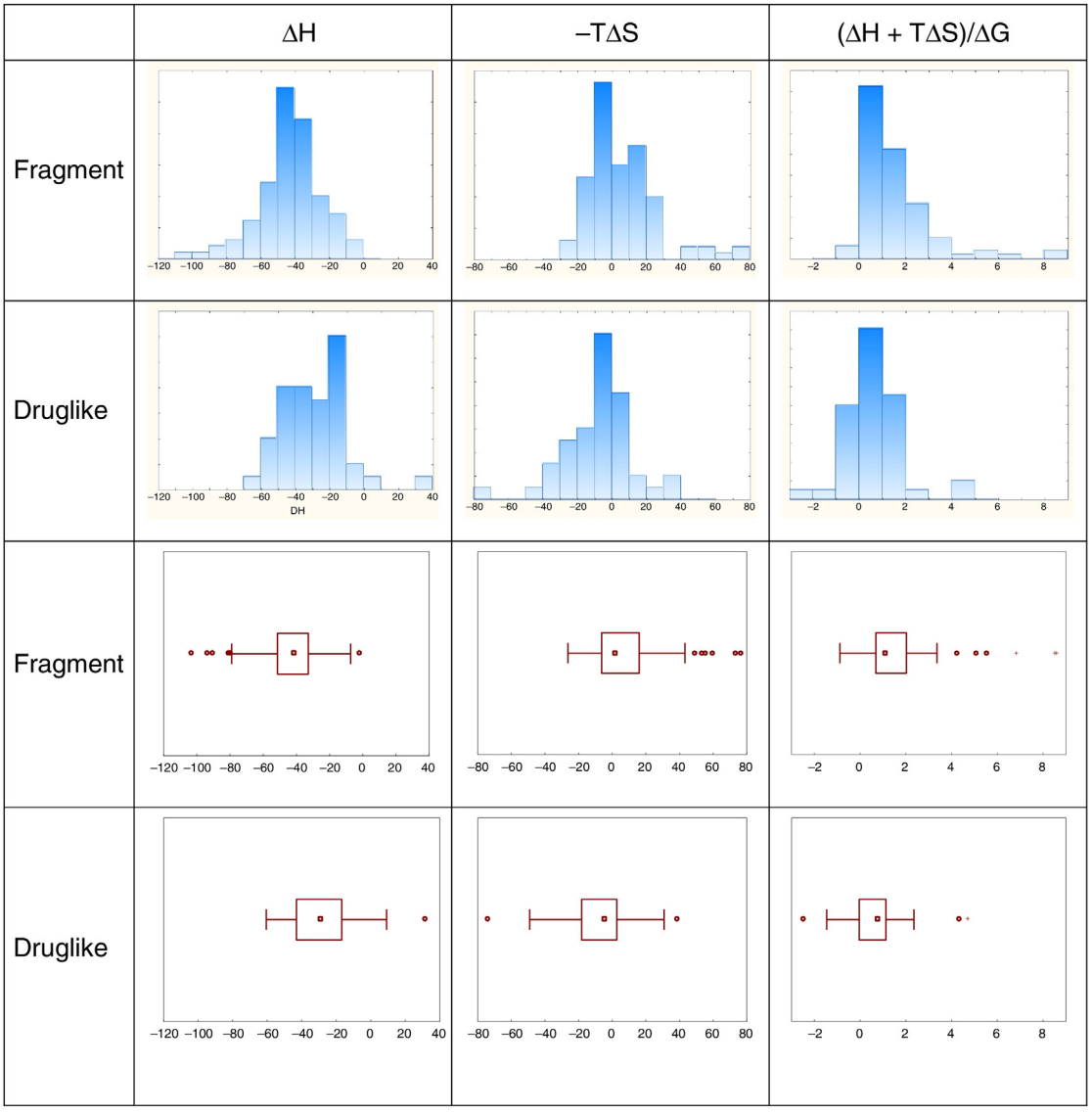
Errors in thermodynamics quantities derived from ITC

In typical ITC experiments used to generate the data described above, a sample of the protein (the titrand) is contained within a small reaction cell that is thermally insulated from the environment, at the centre of a titration calorimeter. For typical calorimeters and binding experiments, the protein concentration would be 5–10 μM and the cell volume is 0.3–1.5 ml. Small volumes of a concentrated solution of the ligand (the titrant) are then added via a syringe, which also serves to stir the solution, thus ensuring rapid mixing. If the ligand binds to the protein with a non-zero enthalpy, heat is either released or absorbed, leading to a small temperature change in the cell. An electrical heater is used to maintain a constant temperature difference between the reaction cell and a reference cell within the calorimeter, measured using a sensitive thermocouple. The change in heater power required to maintain a fixed temperature difference is then integrated over time and the result corresponds to the heat change on ligand binding in the reaction cell.

Usually, several injections are made to reach a 1:1 stoichiometry of protein and ligand and additional injections are then made to ensure that the protein-binding site is saturated. Each injection in the first phase releases a small proportion of the binding enthalpy. For a 1 ml cell containing 10 μM of protein with a typical ligand-binding enthalpy of −40 kJ/mol, each injection releases around 40 μJ of heat. To put this into perspective, this is the same amount of heat that would fall on an A4 sheet of paper in 1 s when illuminated by a 40 W bulb placed nearly 5 km away.

It is unsurprising that such calorimetric experiments require sensitive, well-maintained, properly calibrated instruments and precisely prepared solutions. Errors in the molar concentrations of titrant or titrand will result in proportionate errors in measured binding enthalpies (ΔH, kJ/mol) and dissociation constants (K_d, mol/dm^3). However, because free energies are calculated from the logarithm of K_d, the value of ΔG will contain a smaller percentage error. For example, a 25% error in the concentrations would lead to an error of ~5 kJ/mol in the calculated molar enthalpy when $\Delta H_{binding} = -20$ kJ/mol. However, a 25% error in K_d only causes an error of 0.6 kJ/mol in ΔG, which is equivalent to a 2% error in $\Delta G_{binding}$ when $K_d \sim 1$ μM, or a 4% error when $K_d \sim 1$ mM. Entropies ($-T\Delta S$) are calculated as the difference between ΔG and ΔH and so the numerical value of the entropic error will closely mirror that of ΔH, with an opposite sign. This correlation of the errors in ΔH and $T\Delta S$ measured by ITC is separate from the more familiar enthalpy–entropy compensation, in which changes in ΔG usually occur with larger and opposing changes in ΔH and $T\Delta S$ [2].

Other sources of error or variation must also be recognised and reduced. Heat can be generated simply by the dilution of the titrant into the reaction cell. This heat of dilution can be estimated from injections made after the protein is saturated and must be subtracted from all injections when the data are analysed. Finally, changes in the pH or buffer concentration during the course of the titration or between experiments can lead to changes in the protonation state of the protein or the ligand or their weak interactions with ions in solution. Both of these events can be associated with their own heat changes. The practical effects of these errors on measurements of ΔH were investigated at Astex by comparing replicate ITC data, obtained from independent ITC experiments over the course of eight drug discovery programmes. An initial search of the Astex database revealed 80 ITC datasets that were part of replicate measurements involving 30 unique ligands. The smallest number of replicate titrations was two whereas the



	Fragment	Druglike	Fragment	Druglike	Fragment	Druglike
n	94	44	94	44	94	44
10%	–66.9	–50.4	–13.3	–31.4	0.29	–0.50
25%	–51.3	–43.0	–6.3	–18.2	0.69	–0.04
Median	–41.0	–28.7	2.09	–3.65	1.10	0.79
75%	–32.8	–16.8	16.0	2.78	2.03	1.13
90%	–19.2	–10.0	29.0	14.6	3.05	1.74
Mean	–42.9	–28.2	6.58	–7.41	1.55	0.69
S.E.M.	1.87	2.84	2.06	3.07	0.17	0.19
P (MW-U)	0.0001		0.0016		0.0009	

Drug Discovery Today

FIGURE 2

Distribution and statistics of binding enthalpy (Δ*H*), binding entropy (−*T*Δ*S*) and (Δ*H* + *T*Δ*S*)/Δ*G* for neutral fragments and druglike compounds measured by direct ITC experiments. The analysis considered 94 fragments and 44 druglike compounds acting on 17 protein targets [25]. Mann–Whitney *U*-test was applied to test the difference between fragments and druglike compounds. The results show that fragments bind with more favourable enthalpy ($P = 0.0001$) and less favourable entropy ($P = 0.0016$) with respect to druglike compounds. Furthermore, the scaled difference between enthalpy and entropy of binding ((Δ*H* + *T*Δ*S*)/Δ*G*) demonstrates the increased importance of enthalpy gain for fragment binding. The box-plots show the median within the box of the 1st and 3rd quartiles together with the range of non-outlier data defined as 1.5-fold the interquartile range around the median.

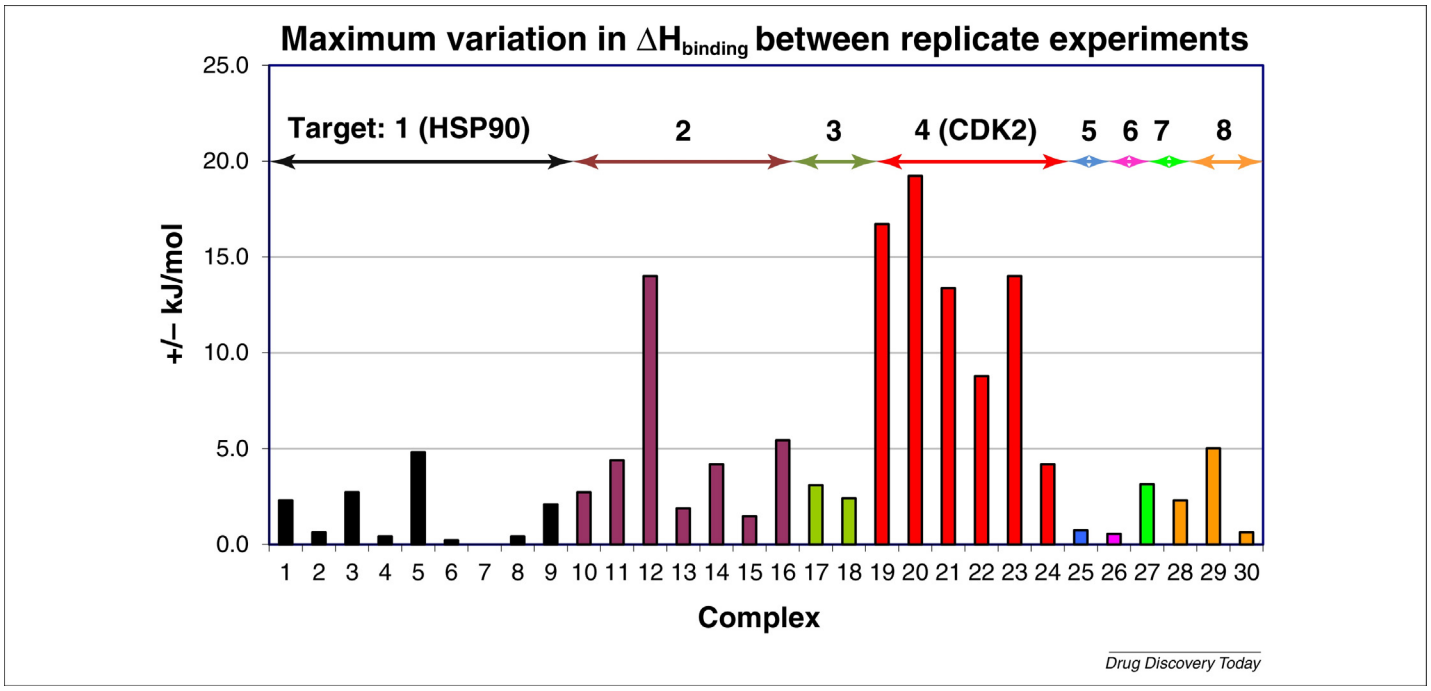


FIGURE 3

Variation in $\Delta H_{binding}$ for apparent replicate titrations within the Astex database. Eighty ITC datasets contribute to this comparison of complexes between eight protein targets and 30 ligands. The average variation between these apparent replicates is 4.7 kJ/mol.

largest was seven. The maximum variation in $\Delta H_{binding}$ measured from replicate titrations for each ligand is illustrated in Fig. 3.

For 70% of the data shown in Fig. 3, the variation in ΔH between replicate titrations is better than 5 kJ/mol. However, target 2 and target 4 (CDK2) show some variations that are greater than 12 kJ/mol. In all cases the buffer was unchanged between the replicate titrations. Closer inspection of the database showed that the largest variations between replicate measurements of

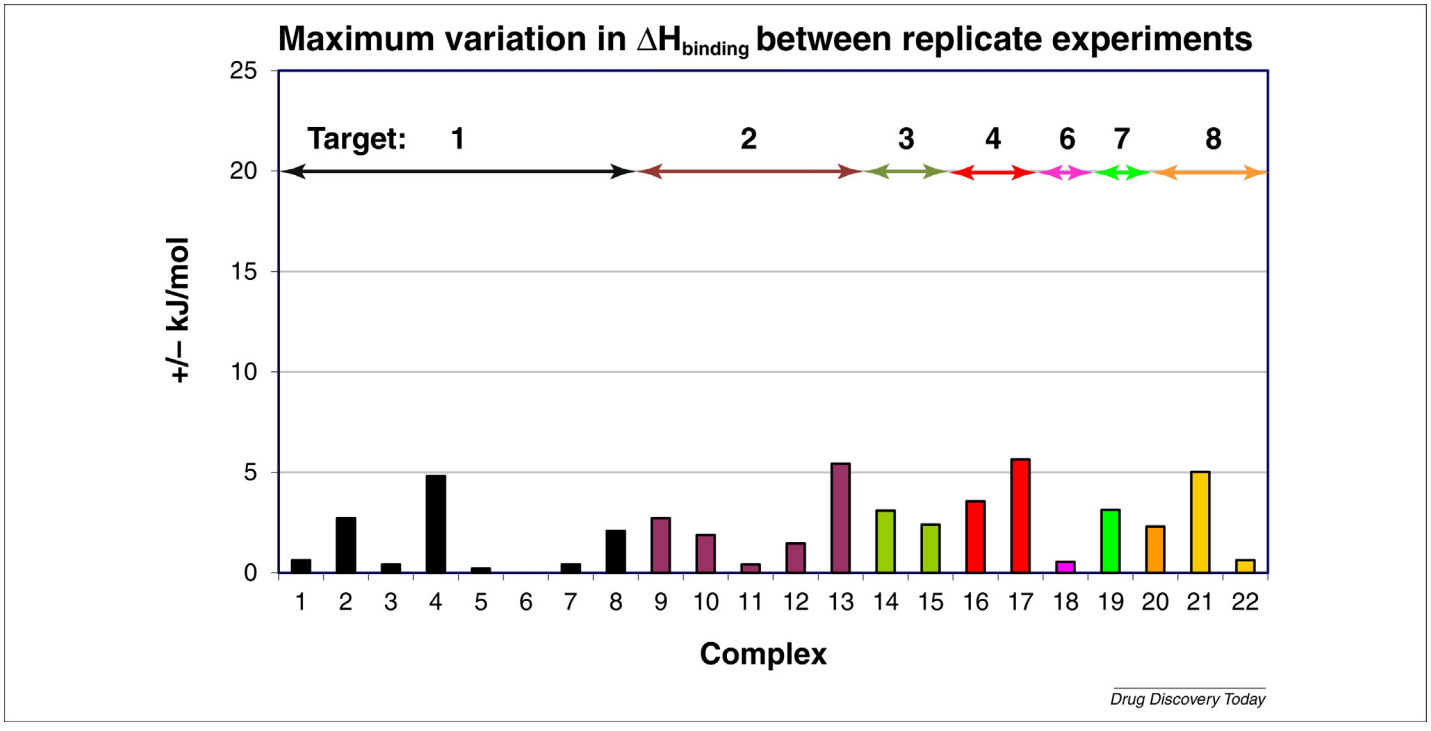


FIGURE 4

Variation in $\Delta H_{binding}$ for true replicate titrations within the Astex database. The average variation between these true replicates is 2.3 kJ/mol. Target numbering is the same as in Fig. 3. Target 5 had no true replicate data and is not present in this figure.

ΔH involved comparisons of different protein constructs (target 2: long + C-terminal tag vs. short + N-terminal tag) or different protein complexes (target 4: CDK2 vs. CDK2.cyclinA).

After removal of all data involving comparison of different forms of the target (different constructs, complexes or phosphorylation states), 56 ITC datasets remained that formed true replicate titrations for 22 ligands with seven protein targets. Target 5 has no true replicate data; the remaining targets contain 2–7 replicate titrations. The maximum variation in $\Delta H_{binding}$ measured from true replicate titrations for each ligand is illustrated in Fig. 4. This shows that the maximum variation in ΔH observed between true replicate titrations for any of the 22 complexes was 5.2 kJ/mol and the average of the maximum variations was 2.3 kJ/mol (SD = 1.8 kJ/mol). Comparison of Fig. 4 with Fig. 3 indicates that minor modifications to a protein target such as changes in construct length, post-translational modifications remote from the ligand-binding site and formation of additional protein–protein interactions can substantially change the binding enthalpy of small ligands, here by up to 10 kJ/mol. Although this analysis has focused on replicate titrations for which the expected difference in ΔH is 0, it also indicates that, within the full Astex ITC dataset, errors in ΔH measurements should be <2.3 kJ/mol on average, with 68% having errors <4.1 kJ/mol and 94% having errors <5.9 kJ/mol. Note that the majority of complexes listed in Table 1 have ΔH values >5.9 kJ/mol and so their unfavourable binding enthalpies are unlikely to be a result of experimental error.

Concluding remarks

Theoretical considerations and experimental data indicate that fragment binding is typically more enthalpically driven than the binding of fragment-derived leads and ligands derived by other drug discovery approaches. The average binding enthalpy, measured by calorimetry for a large diverse set of fragments and targets, is more favourable than the average binding entropy by an amount that agrees well with estimates of the amount of rigid-body entropy that must be surrendered when a freely rotating ligand in solution forms a geometrically constrained complex with a large molecule.

Such constraint renders fragments promising starting points for drug discovery programmes and creates a thermodynamic rationale for FBDD. It is important to remember that increasing the number and strength of high-quality interactions such as H-bonds will not necessarily result in an overall gain in enthalpy. The measured binding enthalpy is a net value and the dissection of the individual contributions might be ambiguous. Solute effects, structural flexibility and cooperativity lead to nonlinear changes in enthalpy and make enthalpy contributions of individual intermolecular contacts experimentally nonobservable. Binding enthalpy and entropy therefore should not be used as direct endpoints but, together with structural studies and free energy calculations, can deepen our understanding of ligand binding [26]. As larger molecules are designed, contributions from protein and solvent reorganisation are expected to be larger. Enthalpic gains are often partially offset by entropic losses as the complex becomes more geometrically constrained, and so changes in the enthalpy and entropy of binding tend to be negatively correlated. Starting from fragments, it is often possible to efficiently improve the affinity and binding enthalpy during early optimisation using a combination of structural (primarily X-ray) and thermodynamic (ITC) data.

The association of enthalpic interactions with high-quality binding in drug discovery has several origins. Firstly, the optimisation of geometrically constrained interactions favours the use of structure-based design which, when combined with computational methods in iterative cycles of synthesis and testing, restricts the number of chemical targets and improves the efficiency of the process. Secondly, the incorporation of more geometrical constraints into the interaction also provides a simple rationale to predict and possibly improve the selectivity of binding of a ligand to a related set of protein targets. Thirdly, the improvement of affinity using entropic gains is most closely associated with an increase in the lipophilicity of the ligand and the subsequent burial of hydrophobic surface area on binding. This has traditionally led to the generation of drug candidates of high molecular weight (≥500 Da) with poor solubility and that also bind nonspecifically to cell membranes and transport proteins. This often leads to low efficacy and the observation of metabolic liabilities in clinical trials and has been described as 'molecular obesity' [27].

The small size and low lipophilicity of a well-designed fragment library provide a large operational freedom to optimise fragment hits into development candidates with beneficial physicochemical and ADME properties, as demonstrated by recent FBDD success stories [28]. Furthermore, the requirement for libraries to be tested at high concentrations and therefore to have good aqueous solubilities favours fragments that contain several H-bond donors and acceptors. Because fragment screening also explores chemical space efficiently and usually generates multiple, chemically diverse hits, the tendency for the hits to be geometrically constrained by H-bonds allows the most common, and hence favourable, interactions with the protein to be identified and probed using new molecules that were not part of the initial screening set. This merging or growing of fragment hits is proving to be a fruitful stimulus for the synthesis of novel, small heterocycles and other chemical entities.

The tendency for fragment binding to be associated with a gain in enthalpy should not be translated into dogma. We have shown examples of two fragments that bind to their target protein, HSP90, with unfavourable enthalpies but favourable entropies, although this situation was rapidly reversed in the course of their initial optimisation using structure-based design. This is not a common occurrence among the targets studied to date; the favourable entropy of binding of these fragment hits probably has its origin in the release of several protein-bound water molecules from the active site that form part of a network of H-bonds used to recognise the purine base of its cofactor ATP. However, these fragments led to the development of a drug candidate, so such hits should not be overlooked. When modifications to the ligand improve ΔH but with no change in ΔG there is still the prospect of further increases in affinity, provided the concurrent change in binding entropy can be addressed. In these circumstances, attention might first focus on restricting the conformation of the free ligand or preserving some flexibility in the protein–ligand complex. In drug design, optimisation will require compromises between, for example, potency, safety, cellular activity and pharmacodynamic efficacy. Compounds that interact with high specificity with their target will nevertheless stand the best chance of

success. A key feature of fragment binding is the identification of near-optimal geometries for polar interactions. The better our understanding of the biophysical parameters involved the better able we will be to guide drug design towards safe and efficacious compounds.

Acknowledgements

The authors are grateful to Chris Murray and Joe Coyle (Astex) for their comments. The data presented in Fig. 1 are the work of many diligent scientists from Astex, AstraZeneca and public laboratories and whose contributions are gratefully acknowledged. This work was supported by National Brain Research Program KTIA-NAP-13-1-2013-0001.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.drudis.2016.11.019.

References

1 Ferenczy, G.G. and Keserű, G.M. (2010) Enthalpic efficiency of ligand binding. *J. Chem. Inf. Model.* 50, 1536–1541
2 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
3 Ferenczy, G.G. and Keserű, G.M. (2012) Thermodynamics of fragment binding. *J. Chem. Inf. Model.* 52, 1039–1045
4 Klebe, G. (2015) Applying thermodynamic profiling in lead finding and optimisation. *Nat. Rev. Drug Discov.* 14, 95–110
5 Hann, M.M. *et al.* (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* 41, 856–864
6 Olsson, T.S. *et al.* (2008) The thermodynamics of protein–ligand interaction and solvation: insights for ligand design. *J. Mol. Biol.* 384, 1002–1017
7 Hajduk, P.J. *et al.* (2005) Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* 48, 2518–2525
8 Vajda, S. and Guarnieri, F. (2006) Characterization of protein–ligand interaction sites using experimental and computational methods. *Curr. Opin. Drug Discov. Dev.* 9, 354–362
9 Kozakov, D. *et al.* (2015) Ligand deconstruction: why some fragment binding positions are conserved and others are not. *Proc. Natl. Acad. Sci. U. S. A.* 112, E2585
10 Ichihara, O. *et al.* (2014) The importance of hydration thermodynamics in fragment-to-lead optimization. *ChemMedChem* 9, 2708–2717
11 Huggins, D.J. (2015) Quantifying the entropy of binding for water molecules in protein cavities by computing correlations. *Biophys. J.* 108, 928–936
12 Breiten, B. *et al.* (2013) Water networks contribute to enthalpy/entropy compensation in protein–ligand binding. *J. Am. Chem. Soc.* 35, 15579–15584
13 Murray, W.C. and Verdonk, M.L. (2002) The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput. Aided Mol. Des.* 16, 741–753
14 Torres, F.E. *et al.* (2010) Higher throughput calorimetry: opportunities, approaches and challenges. *Curr. Opin. Struct. Biol.* 20, 598–605
15 Krimmer, S.G. and Klebe, G. (2015) Thermodynamics of protein–ligand interactions as a reference for computational analysis: how to assess accuracy, reliability and relevance of experimental data. *J. Comput. Aided Mol. Des.* 29, 867–883
16 Murray, C.W. *et al.* (2010) Fragment-based drug discovery applied to HSP90. Discovery of two lead series with high ligand efficiency. *J. Med. Chem.* 53, 5942–5955
17 Woodhead, A.J. *et al.* (2010) Discovery of (2,4-dihydroxy-5-isopropylphenyl)-[5-(4-methylpiperazin-1-ylmethyl)-1,3-dihydroisoindol-2-yl]methanone (AT13387), a novel inhibitor of the molecular chaperone Hsp90 by fragment based drug design. *J. Med. Chem.* 53, 5956–5969
18 Dolezal, O. *et al.* (2013) Fragment screening for the modelling community: SPR, ITC, and crystallography. *Aust. J. Chem.* 66, 1507–1517
19 Silvestre, H.L. *et al.* (2013) Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery. *Proc. Natl. Acad. Sci. U. S. A.* 110, 12984–12989
20 Dai, R. *et al.* (2015) Fragment-based exploration of binding site flexibility in *Mycobacterium tuberculosis* BioA. *J. Med. Chem.* 58, 5208–5217
21 Neumann, L. *et al.* (2011) HTS reporter displacement assay for fragment screening and fragment evolution toward leads with optimized binding kinetics, binding selectivity, and thermodynamic signature. *Methods Enzymol.* 493, 299–320
22 Zender, M. *et al.* (2013) Discovery and biophysical characterization of 2-amino-oxadiazoles as novel antagonists of PqsR, an important regulator of *Pseudomonas aeruginosa* virulence. *J. Med. Chem.* 56, 6761–6774
23 Scott, A.D. *et al.* (2009) Thermodynamic optimisation in drug discovery: a case study using carbonic anhydrase inhibitors. *ChemMedChem* 4, 1985–1989
24 Rühmann, E. *et al.* (2015) Thermodynamic signatures of fragment binding: validation of direct versus displacement ITC titrations. *Biochim. Biophys. Acta* 1850, 647–656
25 Ferenczy, G.G. and Keserű, G.M. (2016) On the enthalpic preference of fragment binding. *Med. Chem. Commun.* 7, 332–337
26 Geschwindner, S. *et al.* (2015) Ligand binding thermodynamics in drug discovery: still a hot tip? *J. Med. Chem.* 58, 6321–6335
27 Hann, M.M. (2011) Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Commun.* 2, 349–355
28 Erlanson, D.A. *et al.* (2016) Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* 15, 605–609