
Scott JS, Waring MJ. [Practical application of ligand efficiency metrics in lead optimisation](#). *Bioorganic & Medicinal Chemistry* 2018

Copyright:

© 2018. This manuscript version is made available under the [CC-BY-NC-ND 4.0 license](#)

DOI link to article:

<https://doi.org/10.1016/j.bmc.2018.04.004>

Date deposited:

11/04/2018

Embargo release date:

05 April 2019



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)

Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.

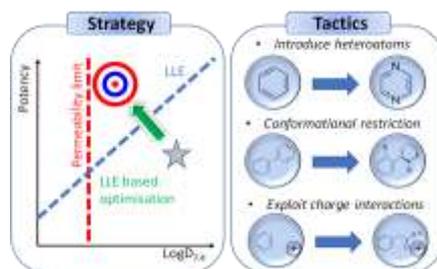
Practical application of ligand efficiency metrics in lead optimisation

Leave this area blank for abstract info.

James S. Scott^a and Michael J. Waring^b, *

^aMedicinal Chemistry, Oncology, IMED Biotech Unit, AstraZeneca, Cambridge CB4 0WG, United Kingdom

^bNorthern Institute for Cancer Research, Chemistry, School of Natural and Environmental Sciences, Bedson Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom





Practical application of ligand efficiency metrics in lead optimisation

James S. Scott^a and Michael J. Waring^{b, *}

^aMedicinal Chemistry, Oncology, IMED Biotech Unit, AstraZeneca, Cambridge CB4 0WG, United Kingdom

^bNorthern Institute for Cancer Research, Chemistry, School of Natural and Environmental Sciences, Bedson Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom

ARTICLE INFO

Article history:

Received

Received in revised form

Accepted

Available online

Keywords:

Lead optimisation

Lipophilicity

Lipophilic ligand efficiency

Ligand efficiency

ABSTRACT

The use of composite metrics that normalise biological potency values in relation to markers of physicochemical properties, such as size or lipophilicity, has gained a significant amount of traction with many medicinal chemists in recent years. However, there is no consensus on best practice in the area and their application has attracted some criticism. Here we present our approach to their application in lead optimisation projects, provide an objective discussion of the principles we consider important and illustrate how our use of lipophilic ligand efficiency enabled the progression of a number of our successful drug discovery projects. We derive, from this and some recent literature highlights, a set of heuristic guidelines for lipophilicity based optimisation that we believe are generally applicable across chemical series and protein targets.

2009 Elsevier Ltd. All rights reserved.

1. Introduction

The need to balance potency and physicochemical properties during medicinal chemistry optimisation is well established.^{1,2,3} Undesirable values of simple physicochemical descriptors such as lipophilicity (logP, logD_{7.4}) and size (molecular weight, heavy atom count) are associated with poor absorption, distribution, metabolism, elimination and toxicity (ADMET) properties such as low solubility, high metabolic clearance and increased activity at toxicological targets such as the human ether-à-go-go-related gene (hERG) channel.⁴ Nevertheless, medicinal chemists often continue to contend with large lipophilic compounds during optimisation because these typically bind well to protein targets and are therefore more likely to be found as a result of hit finding activities. This is not necessarily a catastrophic situation provided medicinal chemists are aware of the fact and can conceive of strategies to address the shortcomings of their leads and evolve them towards more desirable regions of physicochemical space.

Crude approaches to physicochemical optimisation involve the application of cut-off values for molecular weight and logP/D values such as Rule-of-5 criteria (MWt <500, logP <5).⁵ Because potency often increases with lipophilicity and molecular weight, the concept of normalising a biological potency value by descriptors of size, such as heavy atom count (HA) in the case of ligand efficiency (LE, Equation 1), or lipophilicity (lipophilic ligand efficiency, LLE or LipE, Equation 2) have been introduced.⁶ This initial concept has been expanded in an attempt

to combine multiple parameters such as size and lipophilicity (ligand efficiency dependent lipophilicity, LELP, equation 3).⁷

$$LE = -2.303 (RT/HA) \times \log K_d = (1.37/HA) \times pIC_{50} \quad (1)$$

$$LLE = pIC_{50} - \log D_{7.4} \quad (2)$$

$$LELP = clogP / LE \quad (3)$$

These concepts have attracted some criticism, chiefly due to their lack of thermodynamic basis and their underlying assumptions about the baseline relationships between their components, for example that potency should increase linearly with heavy atom count in the case of LE and that the relationship between potency and logD is linear with a slope of unity for LLE.^{8,9} These are valid considerations and make the combined assumptions that go into LELP hard to justify. The question about the validity of the metrics continues to attract debate.¹⁰

2. Selection of metrics

One important consideration is that composite metrics have been employed for two different purposes. Firstly, they have been used to facilitate comparison between different chemical series, for example to compare screening hits from different chemical series which differ from each other in their size or lipophilicity in order to determine which is most attractive. Secondly, they have been applied to the assessment of how a small structural change within

* Corresponding author. Tel.: +0-000-000-0000; fax: +0-000-000-0000; e-mail: author@university.edu

a chemical series affects potency relative to its concomitant change in size or lipophilicity. Qualitatively, these are both reasonable and helpful questions to ask. We also contend that quantitative metrics such as LE and LLE are useful in informing these decisions and their application is valid provided one is aware of their assumptions and limitations.

The focus of the majority of the research we describe here relates to lead optimisation and so we are concerned primarily with the assessment of structural changes within a series. In that regard, we have found that LLE is the most helpful optimisation parameter since, in our experience, compound optimisation has primarily been concerned with the reduction in lipophilicity of compounds which were sufficiently potent, in order to improve their ADMET properties. In that regard, we have generally not been examining changes that significantly altered the size of the compounds and hence LE has been uninformative. Moreover, we were inherently attracted to the use of LLE as a metric because of the firm belief that multi-parameter optimisation can be greatly simplified by focussing on the design of potent compounds with low lipophilicity due to the propensity of the majority of ADMET properties to be compromised when lipophilicity is high.⁴ Secondly, it is reasonable to assume that for a small molecule binding to a hydrophobic protein pocket, potency will show a positive, linear relationship with lipophilicity in the absence of any differences in polar interactions between compounds. This assumes that all compounds are within a suitable applicability domain, i.e. where the potency assay being used gives values that correlate with free energy and where lipophilicity can be measured accurately and is within an established range. It is critical to establish as far as is reasonably practical that this is the case for the compounds in question. Finally, the central assumption of the LLE parameter is that potency and lipophilicity not only correlate but do so with a slope of unity. We believe this to be a reasonable assumption but would emphasise that LLE values should not be interpreted in isolation and should be considered in the context of the absolute potency and lipophilicity changes for a given transformation within a series and an analysis of the overall trend in the data, considering compounds where lipophilicity is likely to change in the absence of other binding events.

Because the majority of ADMET properties correlate negatively with lipophilicity, the compounds with the optimal overall profile within a series would be expected to be those that were the most potent with the lowest lipophilicity i.e. highest LLE. The one, and often only, exception to this trend is permeability, which generally decreases as lipophilicity decreases. Hence, a critical component of the strategy of lipophilic optimisation is to establish the lipophilicity limit, for the series in question, at which permeability becomes too low. The position of this limit is dependent on a number of factors including size and hydrogen bonding,¹¹ hence our primary strategies for compound optimisation have been based on the idea of achieving the highest possible potency with lipophilicity values as low as the permeability limit allows.

These concepts are easy to state and many articles have highlighted the value of property based optimisation and lipophilicity control.^{2,12,13} There has been relatively little discussion, however, on how these principles might be implemented in practice within projects¹⁴ with the majority of examples restricted to individual reported studies in which the focus of the discussion is the specific outcome of the optimisation and not the implementation of the approach. Here we discuss our implementation of LLE based optimisation across a range of projects, many of which led to clinical candidates. Importantly we highlight structural changes that led to improved LLE and were critical steps in project progression. We believe that many of these

experiences are generically applicable and will be of use in future optimisations.

3. Measurement

One factor, which we consider of critical importance for LLE based analysis is to use measured lipophilicity values and not rely on calculated values. It is well established (but perhaps not widely appreciated) that calculated logP figures often vary significantly (with an average error of more than one log unit on average) from the true values.¹⁴ This variation is sufficient to render LLE's derived from calculated values meaningless. The requirement to measure logD_{7.4} values clearly requires extra experimental work, and the extra resource required may not be available to all researchers, but we would recommend that if possible, it is worth the investment for high resolution interpretation of structure activity relationships (SAR). The use of chromatographic methods for determining lipophilicity values¹⁵ may reduce the resource requirement to obtain measured values.

We encountered a prominent example of this phenomenon in our optimisation of G-protein coupled receptor 119 (GPR119) agonists for which three oxadiazole isomers were shown to have very different logD_{7.4} values, which were not predicted correctly (Figure 1). In this case, using calculated (clogP) values would have led to the conclusion that compound **1** had the highest LLE, whereas the logD_{7.4} values show that **2** has the highest. Consequently, we can conclude that **1** is gaining its superior potency through increased lipophilicity and that **2** and **3** should offer a better balance of potency and physicochemical properties. The lower logD_{7.4} of **3** overall results in the most improved solubility and hERG potency.

The above example deals with predominantly neutral compounds (no significant ionisation at pH 7.4) and so the logD_{7.4} values are not significantly different to their logP values. The difference between logD_{7.4} and clogP is due solely to inaccuracies in the clogP calculation. Ionisation may be significant for some chemical series and this needs to be considered in the application of LLE. Use of logD_{7.4} to derive LLE makes the assumption that binding of the ionised form to the target protein is negligible, which may be reasonable but needs to be considered if comparing compounds with differing pK_a values that are close to 7.4 (±1). This introduces a second problem with using calculated values because simple pK_a calculations can also carry significant errors.

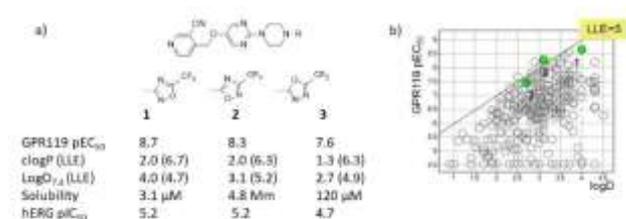


Figure 1 – a) Oxadiazole based GPR119 agonists and b) relationship between logD_{7.4} and potency across all compounds tested within this series, highlighting **1**, **2** and **3**.

4. Correlations

As stated in the introduction, a central assumption of the LLE metric, and a critical one to address if considering compounds across a range of lipophilicity values, is that the correlation between the selected measures of potency and lipophilicity is linear with a slope of unity, e.g. a unit increase in logD_{7.4} leads to a unit increase in pIC₅₀, in the absence of any additional

interactions. We believe this is a reasonable assumption for a drug binding to a lipophilic pocket, after all, logD is intended to quantify the energy associated with a compound transferring from an aqueous to a lipophilic environment, but it should always be remembered that the chosen organic phase in the logD experiment (usually 1-octanol) is only a crude surrogate of a protein pocket. It is necessary to analyse the SAR to build confidence that the assumption holds for the target in question. This is difficult, if not impossible, to do with absolute certainty because the majority of structural changes, such as additions of substituents to lead molecules, would be expected to change many other parameters to a varying degree in addition to lipophilicity and it is not possible to vary lipophilicity independently of other molecular properties – a significant problem with QSAR analyses in general. A recent analysis of >2000 pairwise changes demonstrated that addition of a single methyl group, perhaps the simplest structural change that can be made and one that would lead to an increase in logP of ~0.5 on average, could result in a potency change of ± 2 log units with an average change of zero.¹⁶ This shows the extent to which potency can be modulated, presumably by changes in lipophilicity, electronic effects, conformational restriction and steric clashes by very small structural changes.

Examination of overall trends in data where large numbers of compounds have been tested and examination of the “leading edge” of the correlation, where negative impacts of structural changes are minimised can be very informative in this regard. For example, our series of GPR119 agonists show a logD potency correlation with a leading edge which is close to LLE=5 and has a slope which approximates to unity (Figure 1b). Similarly, for a series of G-protein coupled receptor 40 (GPR40) antagonists, we observed a linear relationship between pEC₅₀ and logD_{7.4} which showed a slope of 1.1 across 2 units of logD_{7.4} (Figure 2).¹⁶ This is instructive, because we believe that the compounds bind to a protein pocket that is very hydrophobic in nature (and possesses a very hydrophobic endogenous ligand) and analysis of the SAR across the series was not suggestive of any of the compounds making additional polar interactions. In this case, the inability to break this trend meant that we were unable to optimise the series towards a drug candidate but it is also informative of what an underlying potency:lipophilicity relationship is.

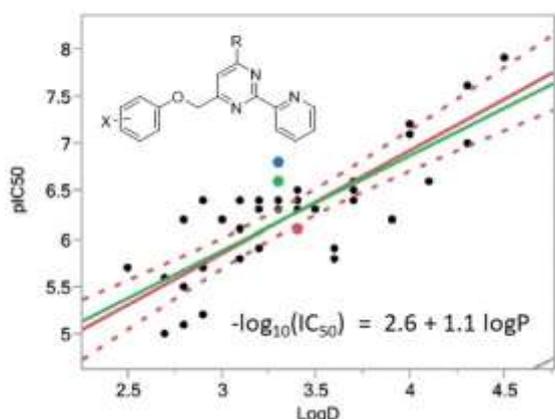


Figure 2 – Potency : LogD_{7.4} relationships in a series of GPR40 antagonists

5. Improving ADMET properties

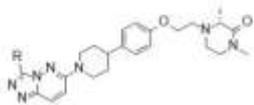
The primary rationale for optimising LLE is that this is anticipated to lead to improved overall profiles of compounds and

favour those with the best balance of potency and ADMET properties. The application of LLE vastly simplifies the optimisation problem – it is very difficult to optimise compounds against potency and all of the individual ADMET parameters that need to be considered simultaneously (solubility, metabolic stability, hERG potency etc.) because analysis of poly-dimensional SAR is difficult intellectually and subject to multiplication of experimental errors. The multiparameter optimisation problem can be simplified if one can assume that identification of potent and polar compounds will result in the majority of ADMET properties falling into place. Moreover, it also makes molecular design more rational and predictable because structural changes that manipulate lipophilicity are easy to conceive of and in prioritising synthesis targets, one might only need to speculate on which lipophilicity lowering changes might be most likely to maintain or increase potency.

As an illustrative example, we have recently been able to optimise a series of bromodomain-containing protein 4 (BRD4) inhibitors, for which the lead molecules, such as **4**, were highly lipophilic and suffered from poor solubility and metabolic stability, resulting in a lack of oral exposure (Figure 3).¹⁶ The change that addressed these problems was discovered by variation of the triazolopyrimidine substituent, which occupies a lipophilic region of BRD4 and is responsible for recognition of the methyl group of the acetylated lysine substrate. Changing the initial trifluoromethyl substituent of **4** (logD_{7.4} >4) showed that LLE could be manipulated in this region although the majority of changes resulted in reduced potency relative to **4** despite increased LLE. The highest LLE compound in this series was methoxy derivative **9**, which had reduced lipophilicity and increased potency relative to **4**. This shows that large changes in LLE are possible (>3 log units in this case) with relatively small changes to structure. It is also noteworthy that the logD_{7.4} values of these compounds do not change as their π -values would predict (CF₃ increases logD by >2.3 units, $\Delta\pi = 1.2$ and OMe by 0.4, $\Delta\pi = 0.0$). This is presumably due to the electronic interaction between the substituent and the nitrogen atoms of the heterocycle and highlights the need to obtain measured values. As a result of its high LLE, **9** had the best balance of properties, including high solubility, low microsomal turnover and desirable pharmacokinetic properties. This compound became our clinical candidate on the BRD4 project, designated AZD5153.

In our 11- β -hydroxysteroid dehydrogenase type 1 (11- β HSD1) project, we identified lead compound **11**, which was moderately lipophilic but required improvement in potency and also suffered from metabolic instability resulting in no oral exposure (Figure 4).¹⁷ Pharmacokinetic data suggested that metabolic instability was the primary reason for the lack of oral exposure although the compounds also suffered from sub-optimal solubility. Accordingly, our optimisation objectives were to increase LLE by both increasing potency and reducing lipophilicity. It was discovered that introducing an additional substituent in the 6-position of the pyridine that contained a carboxylic acid substituent addressed both of these objectives resulting in a 2.7 unit improvement in LLE and a 7 nM inhibitor, compound **12**. The reduction in lipophilicity resulted in improved solubility and metabolic stability giving 68% bioavailability in the rat. This compound ultimately was selected as the development candidate AZD4017.

Within this series, we also identified a back-up candidate that was structurally differentiated and was less prone to the formation



Compound	R	BRD4 pK _i	LogD _{7.4}	LLE	Solubility / μM	Hu heps CL ₅₀
4	-CF ₃	8.1	>4.2	<3.9	63	70
5	-H	6.0	1.9	4.1	1000	180
6	-Me	7.8	2.4	5.4	670	7.2
7	-Cl	7.9	3.1	4.8	63	24
8	-SMe	7.9	2.9	5.0	660	21
9 (AZD5153)	-OMe	8.3	2.3	6.0	>1600	<3.5
10	-NMe ₂	7.9	2.4	5.5	>1600	4.4

Figure 3 – Optimisation of a series of triazolopyridazine based BRD4 inhibitors

of acyl glucuronides - a potential risk associated with **12**.¹⁸ The cyclopropyl pyrrolidine replacement for the piperidine increased steric congestion α - to the acid and reduced the degree of glucuronidation of the carboxylate. LLE could be balanced by increasing the lipophilicity of the 3-carboxamide substituent (cyclohexyl changed to adamantyl) and reducing the lipophilicity of the pyridine 2-substituent (propylthio- replaced with methylthio-) resulting in **13**, which had a similar potency and lipophilicity to **12** (although further improved solubility) and was selected as the back-up candidate AZD6925.

The most significant improvement in LLE arose from the incorporation of the carboxylate substituent. It is interesting in this case that this leads to improved potency as well as reduced lipophilicity, with an X-ray crystal structure of **12** showing the acid positioned in close proximity to the backbone NH of Leu217 and the carboxylate of Asp259. Permeability remained acceptable for **12** (MDCK P_{app} 290 nm.s⁻¹) but other members of the series with logD<2 showed reduced flux and reduced bioavailability in rat.

11	12, AZD4017	13, AZD6925
Cell pIC ₅₀ 6.5	Cell pIC ₅₀ 8.2	Cell pIC ₅₀ 8.2
LogD _{7.4} 3.2	LogD _{7.4} 2.2	LogD _{7.4} 2.1
LLE 3.3	LLE 6.0	LLE 6.2
Rat F% 0%	Rat F% 68%	Rat F% 86%
Solubility 82 μM	Solubility 140 μM	Solubility >2000 μM

Figure 4 - Optimisation of acidic 11-βHSD1 inhibitors

Both of these examples show the improvement in solubility and metabolism resulting from reduced lipophilicity, which were achieved using structural changes that are well established to reduce logD_{7.4}. Improvements in solubility and metabolic stability with lowered lipophilicity within series are common trends. Aqueous solubility is a composite parameter derived from the activity coefficient (which is inversely correlated with lipophilicity) and lattice energy (for crystalline compounds).¹⁹ Metabolic stability is generally inversely correlated with

lipophilicity because the majority of metabolising enzymes such as CYP₄₅₀'s have evolved to recognise lipophilic substrates for detoxification and elimination, hence more lipophilic compounds bind to CYP450 enzymes with greater affinity in general.

Similar considerations apply to other ADMET parameters.⁴ We suggest that reduction in lipophilicity should be the primary strategy to address ADMET issues. Secondary strategies, such as disrupting crystal packing in the case of solubility or blocking groups for metabolism should only be considered once lipophilicity is controlled and in an appropriate range.

6. Polar substituents and heterocycles

Whilst we are usually able to rationally manipulate structures in a way that reduces lipophilicity, most commonly adding polar substituents or introducing heteroatoms, anecdotally, we believe that some strategies are superior to others in the sense that they are more likely lead to retained potency (improved LLE) or are less likely to compromise permeability as lipophilicity is reduced.

In our optimisation of activators of glucokinase (GK), we identified lipophilic lead compound **14**, which was potent but suffered from compromised ADMET properties, most notably hERG potency (Figure 5).^{19,20} As expected, these issues could be addressed by reducing lipophilicity and in this case, this was achieved in a potency neutral manner by replacing the thiazole with a pyrazole and the side chain methoxy substituent with a hydroxyl, resulting in **15**, which had significantly reduced hERG potency and became a clinical candidate AZD1092. The reduction in lipophilicity of this series led to diminished permeability as is illustrated by compromised Caco-2 P_{app} for **15**. Further characterisation of the compound revealed this reduced permeability resulted in variable fraction absorbed across species (F_{abs} 50% in rat although higher in dog) presenting a risk that human F_{abs} may be low or variable.

To address this, we embarked upon a follow up campaign focused on improving permeability whilst maintaining the other desirable properties of **15**.²¹ The primary strategy to achieve this goal was to reduce the hydrogen bonding potential of the compounds by replacing the hydroxyl group H-bond donor. One successful way to do this was to return to methyl ether derivatives but this would clearly increase lipophilicity. Therefore, concomitantly, we explored additional strategies to lower logD in other parts of the structure and found that this could be achieved without detriment to potency by replacing the pendant phenyl ring with 6-membered heterocycles. In the event, pyrazine was particularly beneficial in improving LLE (increased by 0.6 units relative to phenyl). These changes were instrumental in identifying the clinical candidate from this series **16** (AZD1656), which had a superior technical profile to **15**, including significantly improved permeability and 100% bioavailability across species and which progressed to Phase II clinical testing.

Key learning from this program was that introduction of heterocycles was a superior strategy to introducing polar, H-bonding substituents as this enabled reduction in lipophilicity without as much detriment to permeability, ultimately leading to a compound with a much better overall profile. The SAR for the introduction of heterocycles is also noteworthy, pyrazine was by far the best replacement for phenyl and the corresponding pyridines were inferior, with the 2-pyridyl isomer **17** losing potency relative to phenyl **19** and the 3-isomer **18** not showing a reduction in logD_{7.4} relative to **19** (the hydrogen bond acceptor potential of the nitrogen is significantly diminished when it is *ortho*- to a strong electron withdrawing group). Hence, the improvement offered by pyrazine could not be predicted from the

two corresponding pyridine derivatives. In our experience, diazines are often show superior LLE than the corresponding pyridines and should be explored systematically in late stage optimisation regardless of the data on the pyridine analogues; the other four possible diazines were inferior to pyrazine in this case.

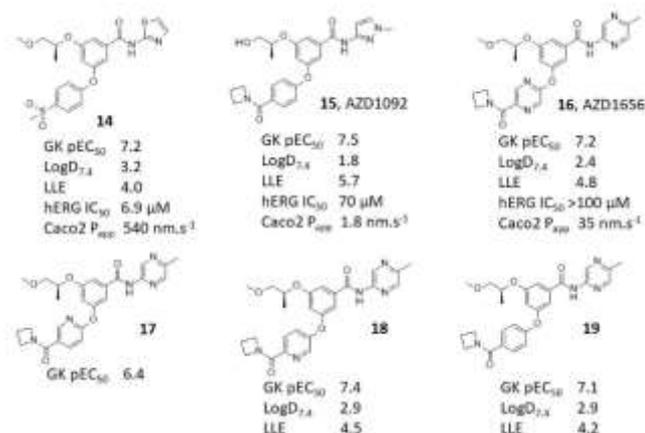


Figure 5 - LLE, hERG potency and permeability of a series of GK activators

In our search for inhibitors of doubly-mutated (DM) epidermal growth factor receptor tyrosine kinase (EGFR), we identified a series of indole-substituted pyrimidines that were potent and offered some selectivity in a cellular setting over the wild-type (WT) kinase, exemplified by compound **20** (Figure 6).²² Despite promising activity data, **20** was too lipophilic and had no oral exposure in rats. Following some of the considerations outlined above, we set out to reduce lipophilicity and found that this could be achieved without too much detriment to potency by replacing the indole with a pyrazolopyridine (introduction of heteroatoms) and by appending a basic dimethylamino group to the acrylamide (introduction of ionisable groups) resulting in **21**, which had oral exposure in the rat and served as the first in vivo probe compound for the project.

We then set out to reduce the predicted human dose of the compounds to acceptable levels by examining further improvements in LLE and we discovered that potency gains could be made by shifting the base from the acrylamide to the *para*-position of the phenyl ring (**22**), increasing potency and reducing logD_{7.4} simultaneously (3 unit improvement in LLE).²³ This not only resulted in a superior compound, but also allowed greater scope to explore structural variations in the rest of the molecule. Most critically, this allowed a reinvestigation of indole substituents on the pyrimidine, which were more potent against single mutant (SM) EGFR. A compound with equipotency against the DM and SM forms of the kinase offered a more desirable clinical profile but previously, indole substituents had been ruled out because they led to compounds that were too lipophilic. This now enabled their further investigation with a wider variety of additional substitution. The most productive variations were the combinations of substitutions of the indole nitrogen (leading to a further increase in logD_{7.4}) with variations of the pyrimidine 5-substituent (e.g. removal of the chloro), which allowed lipophilicity to be balanced and the selectivity against IGFR to be improved. This exercise led to the identification of **23**, among many other compounds, which was less potent and more lipophilic than **22** but was unique in that it showed significantly diminished insulin-like growth factor receptor (IGFR) and hERG potency – two critical selectivity targets. Whilst **23** therefore does not have the highest LLE in the series, the discovery of its unique selectivity would not have been possible without the LLE optimisation that preceded its discovery. The compound **23** was nominated as a

clinical candidate (AZD9291) and is now marketed as TAGRISSO™ (osimertinib).

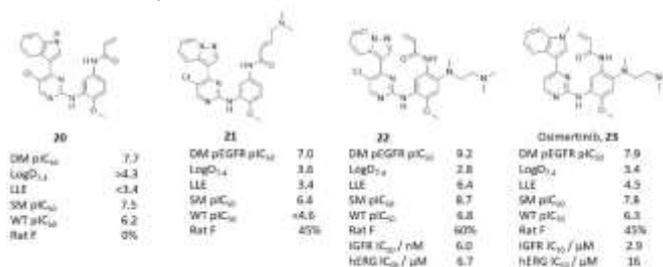


Figure 6 – Mutant EGFR inhibitors

7. Conformational restriction

A long running programme investigating diacyl glycerol acetyl transferase 1 (DGAT1) inhibitors revealed a sub-series of triaryl derivatives, which originated from the unsubstituted biphenyl compound **24** (Figure 7).¹⁵ This compound fell short of the required target potency but had very low lipophilicity and was considered a promising lead. Because **24** contains a carboxylate, which was shown to be required for potency we considered it essential for our optimisation programme that the logD_{7.4} was not increased significantly. In our experience, the critical optimisation parameter for acidic compounds is to minimise clearance of the compounds, usually to a level that they show no detectable turnover in hepatocyte incubations, because higher clearance values result in human dose predictions that are too high and carry significant uncertainty due to their low volume of distribution. The best strategy to achieve this is to make the compounds as polar as their permeability limit allows. Optimisation of carboxylic acids introduces an additional risk, specifically arising from the potential to form acyl glucuronides, which may be reactive metabolites.²⁴ On the other hand, we have recently shown that carboxylic acids carry a lower overall risk of failing in development for toxicological reasons but are more likely to fail due to poor pharmacokinetics.²⁵ Together, these factors strongly suggest that maximising LLE in acids is highly desirable as a means to minimise the propensity for metabolism and reduce the overall dose.

In our case, the permeability limit occurs at very low logD_{7.4} because the compounds are relatively small²⁶ and so we considered the best strategy for optimisation to be to increase potency whilst maintaining logD_{7.4} in a similar range to **24** but without significantly increasing molecular size. One option to achieve these ends is to explore conformational restriction to freeze out the bioactive conformation of the molecule and reduce the entropic penalty of binding. Also of consideration in this case is the array of three contiguous aryl rings and we were aware that flat, polyaromatic compounds are often associated with additional problems such as low solubility that may be worse than predicted by their lipophilicity alone.¹⁵ Taking these considerations together, one attractive avenue of SAR exploration was to introduce small substituents that would favour a non-coplanar conformations. Frequently, destabilising planar conformations of

biaryl systems with lipophilic substituents does not increase lipophilicity as much as the substituent contributions might predict.

Accordingly, we investigated the introduction of a chloro-substituent at all four possible positions on the aromatic ring. Substitution *ortho*- to the acetic acid or pyrazine substituents resulted in reduced potency relative to **24**. Substitution in the central phenyl ring *ortho*- to the other phenyl moiety (**25**) resulted in an increase in potency with only modest increase in $\log D_{7.4}$ ($\Delta LLE +0.5$), indicating that an orthogonal orientation of the two phenyl rings represents the likely bioactive conformation. The permeability of **25**, measured by MDCK P_{app} , was moderate, indicating that $\log D_{7.4}$ of 0.6 was still close to the permeability limit. Consistent with this theory, changing the chloro for a cyano substituent led to lower $\log D_{7.4}$ and poor permeability ($\log D_{7.4} = -0.4$ and MDCK $P_{app} = 12 \text{ nm}\cdot\text{s}^{-1}$). Incorporating a chloro-substituent on the distal phenyl ring *meta*- to the acetic acid (**26**) led to a greater improvement in potency ($\Delta LLE +1.8$ relative to **24**). Since the effect on conformation arising from substitution in this position is the same as that for **25**, we interpret this to mean that the chloro- of **26** is making an additional contact with the protein target leading to the greater increase in affinity. Simultaneous substitution at both of these promising positions served to reinforce the improvement and ultimately, the combination of chloro- and fluoro- substituents, combined with a methyl substituent *o*- to the carboxylate led to the optimised compound **27**. This is equipotent with **26** but has a slightly increased $\log D$, which although detrimental to LLE, results in high permeability. This compound was selected as the development candidate AZD2353.

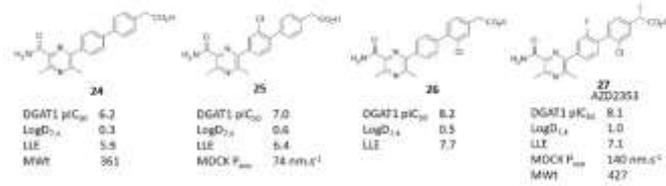


Figure 7 – Optimisation of triaryl DGAT1 inhibitors

An extreme form of conformational restriction that has been popularised recently is macrocyclisation.^{27,28} Interest in macrocycles has derived chiefly from the desire to access regions of high molecular weight for less tractable drug targets such as protein-protein interactions. We have found macrocyclisation to be a very effective strategy to increase potency without changes in lipophilicity. In our 2-hydroxy-dATP diphosphohydrolase (MTH1) project, we identified lead compound **28**, which required an increase in potency whilst maintaining its low lipophilicity (Figure 8). Analysis of the X-ray crystal structure of **28** in complex with MTH1 revealed that the methoxyethyl side chain to be in close proximity to the phenyl ring in the bound conformation. This is a high energy conformation in which repulsive interactions between these groups occur and so **28** would be expected to encounter substantial entropic penalties to binding. Covalently linking these substituents together, initially with an alkyl chain to form a macrocycle (**29**), locked the substituents in the active conformation. This resulted in a substantial increase in potency with only a small increase in lipophilicity ($\Delta LLE +2.3$). Lipophilicity could be reduced by incorporating an ether in the linking chain (**30**), which had a similar lipophilicity to **28** but was three orders of magnitude more potent, presumably due to the ligand being effectively pre-organised into the bioactive conformation.

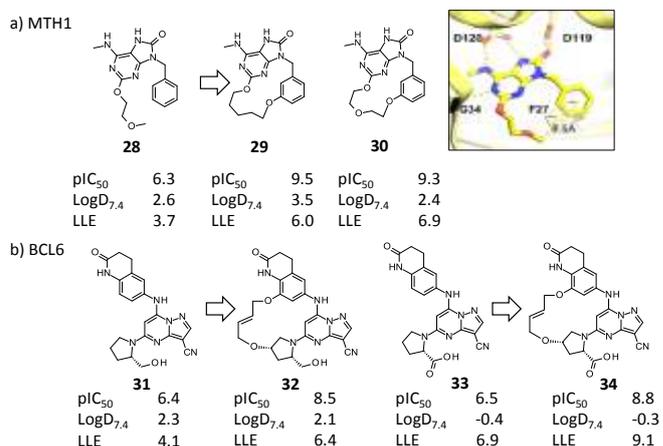


Figure 8 – Macrocyclisation as a strategy for lipophilicity neutral increases in potency in a) MTH1 inhibitors and b) BCL6 inhibitors

A second example of the effect of macrocyclisation comes from our B-cell lymphoma 6 (BCL6) inhibitor program where conformational restriction afforded significant boosts in potency for no increase in lipophilicity and consequently, increased LLE. Here, comparing acyclic precursors **31** and **33** with their corresponding macrocyclised derivatives **32** and **34**, shows that both pairs exhibit comparable increases in LLE (+2.3 and +2.2 respectively) despite having different lipophilicity values (2.1 and -0.3).²⁹

8. Recent Literature

The examples above are drawn from the authors' experiences at AstraZeneca, however, the use of LLE as an optimisation tool has become more widely accepted by medicinal chemists in recent years.^{30,31} Below, we present some noteworthy examples of this approach from the recent literature (Figure 9). Researchers at Merck describe the use of LLE in the optimisation of a series of Janus kinase 1 inhibitors from tool **35** to advanced lead **36**.³² The incorporation of a polar atom (cyclohexyl switched to tetrahydropyran) proved a key breakthrough improving both hERG and PK parameters. A second example from Merck describes the use of LLE to drive improvements in drug like properties in a series of undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase inhibitors.³³ Here an impressive 5 unit drop in lipophilicity was achieved from startpoint **37** ($\text{clogP } 7.1$) to more polar **38** ($\text{clogP } 1.9$) through the inclusion of polar functionality (aryl to heteroaryl and methyl to hydroxyl) whilst potency was maintained ($\Delta LLE +4.5$). Genentech have reported the optimisation of a pyrazole series of dual leucine zipper kinase inhibitors from **39** to in vivo tool **40** with control of LLE and polar surface area (PSA) being key drivers to improve physical properties and CNS penetration.³⁴

We also highlight examples of the use of LLE where the lead compounds have undergone significant changes in molecular size (Figure 10). An impressive example of a fragment based optimisation is described by Pfizer where a weak hit from a screen against interleukin-1 receptor associated kinase 4 (**41**) was optimised to the clinical candidate **42** (PF-06650833) with an overall ΔLLE of +4.9.³⁵ Conversely, LLE was also used in the deconstruction of a lipophilic mRNA decapping scavenger enzyme inhibitor hit **43** ($\text{clogP } 5.3$) to a smaller, more polar in vivo tool **44** (ΔLLE of +2.6) with improved physicochemical properties.³⁶ Lastly, an example of structurally enabled hybridisation of two fragment hits **45** (LLE 1.3) and **46** (LLE 1.1) for inhibition of S1 serine protease factor D with subsequent

optimisation led to compound **47** (LLE 6.9) which demonstrated *in vivo* efficacy and improved ADMET properties.³⁷

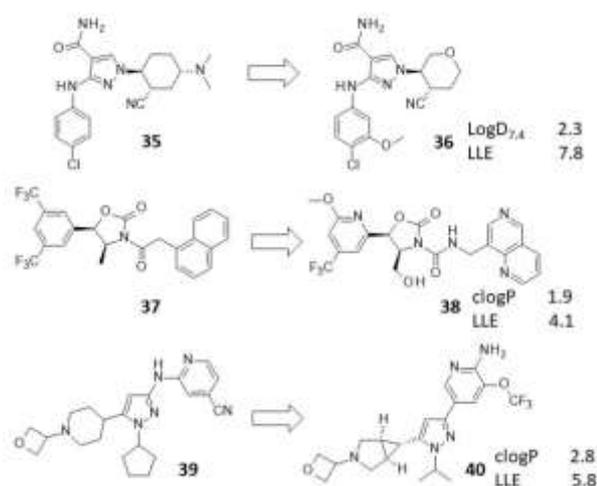


Figure 9 – Literature examples of LLE optimisation

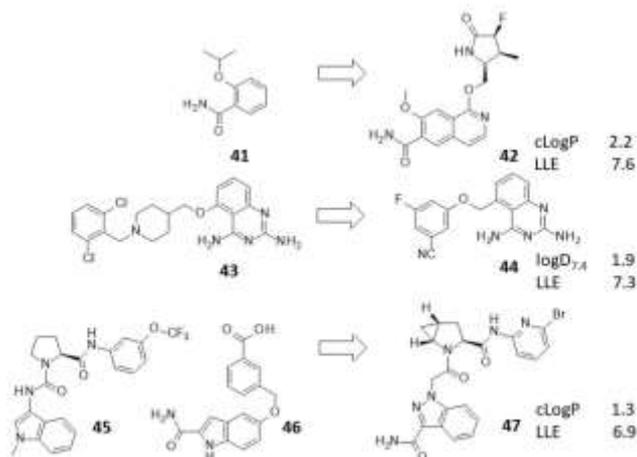


Figure 10 – LLE optimisation with changes in size

9. Conclusions

The examples we describe are not intended to be comprehensive, such an extensive review is beyond the scope of this article. However, we highlight influential aspects of our own work, along with prominent work from the literature, that illustrate considerations that we believe are generically applicable. In a general sense, the strategies we describe can be used to derive a set of heuristic guidelines, which we believe should be considered as part of any LLE based optimisation strategy.

- Address ADMET properties primarily by lipophilicity control.
- Use measured logD for analysis.
- Establish a lower limit of lipophilicity for the series based on permeability data. Target the highest possible potency with the lowest lipophilicity permitted by the permeability limit.
- Reduce lipophilicity without making significant increases in H-bonding or size to maintain / increase permeability window and target much lower regions of logD.

- Introduce heteroatoms in preference to the addition of polar substituents, these changes are more likely to be tolerated and are less detrimental to permeability.
- Avoid large increase in molecular weight. Improve potency by optimisation of existing substituents and controlling conformations.
- Exploit ionisable groups where they increase potency.
- X-ray structures may be useful in identifying polar interactions to target but experimental exploration of the SAR is needed.

Whilst we have focused our interpretation on LLE, the combination of lipophilicity with sub-structural descriptors such as aromatic ring count¹⁵ or proportions of sp³ hybridised atoms,³⁸ have also been suggested as indicators of compound “quality”. We believe that in general, the introduction of increased three-dimensional character to molecules during optimisation, which are the underlying principles of these additional metrics, is likely to be beneficial, and in line with the guidelines that we propose. We would contend, however, that their combination into single molecular descriptors compromises interpretability. Such metrics have generally been applied to analyzing large, structurally diverse datasets. Within the context of an optimisation programme, more complex analysis of SAR will be carried out, and application of combined descriptors is unnecessary.

Many of the examples we have discussed here fall within the chemical space defined by Rule of Five criteria. Optimisation of compounds that do not comply with all of these criteria may be necessary for some challenging targets and is increasingly gaining traction. In this area, it is likely that high values of LLE are harder to achieve. However, we would consider that analysis of SAR using LLE is likely to remain useful in such projects, provided it is applied with consideration of the context of the target in question.

Whilst we anticipate that healthy debate about the physical validity of LLE will continue, we have repeatedly found it to be a useful guide to simplify the problem of multiparameter optimisation and focus efforts towards identifying potent compounds with low lipophilicity. The application of these principles has been repeatedly shown to provide an effective and efficient means of identifying clinical candidates. We hope that by highlighting and classifying the key structural changes that have been beneficial in these projects, further improvements in lead optimisation strategies will be facilitated and ultimately, more successful medicines will be discovered with greater efficiency.

Acknowledgements

We gratefully acknowledge the many contributions of gifted medicinal chemists to the development of the principles and the examples we have described.

References and notes

1. Waterbeemd, H. In *Modern Methods of Drug Discovery*; Birkhäuser Basel: Basel, **2003**; pp. 243–257.
2. Leeson, P. D.; Springthorpe, B. *Nat. Rev. Drug Discov.* **2007**, *6*, 881–890.
3. Hann, M. M. *Medchemcomm* **2011**, *2*, 349.
4. Waring, M. J. *Expert Opin. Drug Discov.* **2010**, *5*, 235–248.
5. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26.
6. Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. *Nat. Rev. Drug Discov.* **2014**, *13*, 105–121.
7. Keserü, G. M.; Makara, G. M. *Nat. Rev. Drug Discov.* **2009**, *8*, 203–212.
8. Kenny, P. W.; Leitão, A.; Montanari, C. A. *J. Comput. Aided. Mol.*

- Des.* **2014**, *28*, 699–710.
9. Shultz, M. D. *ACS Med. Chem. Lett.* **2014**, *5*, 2–5.
10. Murray, C. W.; Erlanson, D. A.; Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H.; Richmond, N. J. *ACS Med. Chem. Lett.* **2014**, *5*, 616–618.
11. Waring, M. J. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2844–2851.
12. Leeson, P. D.; Young, R. J. *ACS Med. Chem. Lett.* **2015**, *6*, 722–725.
13. Mignani, S.; Rodrigues, J.; Tomas, H.; Jalal, R.; Singh, P. P.; Majoral, J.-P.; Vishwakarma, R. A. *Drug Discov. Today* **2018**, *23*, 605–615.
14. Johnstone, C. *Drug Discov. Today* **2012**, *17*, 538–543.
15. Young, R. J.; Green, D. V. S.; Luscombe, C. N.; Hill, A. P. *Drug Discov. Today* **2011**, *16*, 822–830.
16. Scott, J. S.; Birch, A. M.; Brocklehurst, K. J.; Broo, A.; Brown, H. S.; Butlin, R. J.; Clarke, D. S.; Davidsson, Ö.; Ertan, A.; Goldberg, K.; Groombridge, S. D.; Hudson, J. A.; Laber, D.; Leach, A. G.; MacFaul, P. A.; McKerrecher, D.; Pickup, A.; Schofield, P.; Svensson, P. H.; Sörme, P.; Teague, J. *J. Med. Chem.* **2012**, *55*, 5361–5379.
17. Scott, J. S.; Bowker, S. S.; DeSchoolmeester, J.; Gerhardt, S.; Hargreaves, D.; Kilgour, E.; Lloyd, A.; Mayers, R. M.; McCoull, W.; Newcombe, N. J.; Ogg, D.; Packer, M. J.; Rees, A.; Revill, J.; Schofield, P.; Selmi, N.; Swales, J. G.; Whittamore, P. R. O. *J. Med. Chem.* **2012**, *55*, 5951–5964.
18. Scott, J. S.; Barton, P.; Bennett, S. N. L.; DeSchoolmeester, J.; Godfrey, L.; Kilgour, E.; Mayers, R. M.; Packer, M. J.; Rees, A.; Schofield, P.; Selmi, N.; Swales, J. G.; Whittamore, P. R. O. *Medchemcomm* **2012**, *3*, 1264.
19. Waring, M. J.; Johnstone, C.; McKerrecher, D.; Pike, K. G.; Robb, G. *Medchemcomm* **2011**, *2*, 775.
20. Waring, M. J.; Brogan, I. J.; Coghlan, M.; Johnstone, C.; Jones, H. B.; Leighton, B.; McKerrecher, D.; Pike, K. G.; Robb, G. R. *Medchemcomm* **2011**, *2*, 771.
21. Waring, M. J.; Clarke, D. S.; Fenwick, M. D.; Godfrey, L.; Groombridge, S. D.; Johnstone, C.; McKerrecher, D.; Pike, K. G.; Rayner, J. W.; Robb, G. R.; Wilson, I. *Medchemcomm* **2012**, *3*, 1077.
22. Ward, R. A.; Anderton, M. J.; Ashton, S.; Bethel, P. A.; Box, M.; Butterworth, S.; Colclough, N.; Chorley, C. G.; Chuaqui, C.; Cross, D. A. E.; Dakin, L. A.; Debreczeni, J. É.; Eberlein, C.; Finlay, M. R. V.; Hill, G. B.; Grist, M.; Klinowska, T. C. M.; Lane, C.; Martin, S.; Orme, J. P.; Smith, P.; Wang, F.; Waring, M. J. *J. Med. Chem.* **2013**, *56*, 7025–48.
23. Finlay, M. R. V.; Anderton, M.; Ashton, S.; Ballard, P.; Bethel, P. A.; Box, M. R.; Bradbury, R. H.; Brown, S. J.; Butterworth, S.; Campbell, A.; Chorley, C.; Colclough, N.; Cross, D. A. E.; Currie, G. S.; Grist, M.; Hassall, L.; Hill, G. B.; James, D.; James, M.; Kemmitt, P.; Klinowska, T.; Lamont, G.; Lamont, S. G.; Martin, N.; McFarland, H. L.; Mellor, M. J.; Orme, J. P.; Perkins, D.; Perkins, P.; Richmond, G.; Smith, P.; Ward, R. A.; Waring, M. J.; Whittaker, D.; Wells, S.; Wrigley, G. L. *J. Med. Chem.* **2014**, *57*, 8249–8267.
24. Iwamura, A.; Nakajima, M.; Oda, S.; Yokoi, T. *Drug Metab. Pharmacokinet.* **2017**, *32*, 2–11.
25. Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; Wallace, O.; Weir, A. *Nat. Rev. Drug Discov.* **2015**, *14*, 475–486.
26. Waring, M. J. *Bioorganic Med. Chem. Lett.* **2009**, *19*, 2844–2851.
27. Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discov.* **2008**, *7*, 608–624.
28. Heinis, C. *Nat. Chem. Biol.* **2014**, *10*, 696–698.
29. McCoull, W.; Abrams, R. D.; Anderson, E.; Blades, K.; Barton, P.; Box, M.; Burgess, J.; Byth, K.; Cao, Q.; Chuaqui, C.; Carbajo, R. J.; Cheung, T.; Code, E.; Ferguson, A. D.; Fillery, S.; Fuller, N. O.; Gangl, E.; Gao, N.; Grist, M.; Hargreaves, D.; Howard, M. R.; Hu, J.; Kemmitt, P. D.; Nelson, J. E.; O’Connell, N.; Prince, D. B.; Raubo, P.; Rawlins, P. B.; Robb, G. R.; Shi, J.; Waring, M. J.; Whittaker, D.; Wylot, M.; Zhu, X. *J. Med. Chem.* **2017**, *60*, 4386–4402.
30. Meanwell, N. A. *Chem. Res. Toxicol.* **2016**, *29*, 564–616.
31. Cavalluzzi, M. M.; Mangiatordi, G. F.; Nicolotti, O.; Lentini, G. *Expert Opin. Drug Discov.* **2017**, *12*, 1087–1104.
32. Siu, T.; Brubaker, J.; Fuller, P.; Torres, L.; Zeng, H.; Close, J.; Mampreian, D. M.; Shi, F.; Liu, D.; Fradera, X.; Johnson, K.; Bays, N.; Kadic, E.; He, F.; Goldenblatt, P.; Shaffer, L.; Patel, S. B.; Lesburg, C. A.; Alpert, C.; Dorosh, L.; Deshmukh, S. V.; Yu, H.; Klappenbach, J.; Elwood, F.; Dinsmore, C. J.; Fernandez, R.; Moy, L.; Young, J. R. *J. Med. Chem.* **2017**, *60*, 9676–9690.
33. Mandal, M.; Tan, Z.; Madsen-Duggan, C.; Buevich, A. V.; Caldwell, J. P.; Dejesus, R.; Flattery, A.; Garlisi, C. G.; Gill, C.; Ha, S. N.; Ho, G.; Koseoglu, S.; Labrioli, M.; Basu, K.; Lee, S. H.; Liang, L.; Liu, J.; Mayhood, T.; McGuinness, D.; McLaren, D. G.; Wen, X.; Parmee, E.; Rindgen, D.; Roemer, T.; Sheth, P.; Tawa, P.; Tata, J.; Yang, C.; Yang, S.-W.; Xiao, L.; Wang, H.; Tan, C.; Tang, H.; Walsh, P.; Walsh, E.; Wu, J.; Su, J. *J. Med. Chem.* **2017**, *60*, 3851–3865.
34. Patel, S.; Meilandt, W. J.; Erickson, R. I.; Chen, J.; Deshmukh, G.; Estrada, A. A.; Fujii, R. N.; Gibbons, P.; Gustafson, A.; Harris, S. F.; Imperio, J.; Liu, W.; Liu, X.; Liu, Y.; Lyssikatos, J. P.; Ma, C.; Yin, J.; Lewcock, J. W.; Siu, M. *J. Med. Chem.* **2017**, *60*, 8083–8102.
35. Lee, K. L.; Ambler, C. M.; Anderson, D. R.; Boscoe, B. P.; Bree, A. G.; Brodfuehrer, J. I.; Chang, J. S.; Choi, C.; Chung, S.; Curran, K. J.; Day, J. E.; Dehnhardt, C. M.; Dower, K.; Drozda, S. E.; Frisbie, R. K.; Gavrin, L. K.; Goldberg, J. A.; Han, S.; Hegen, M.; Hepworth, D.; Hope, H. R.; Kamtekar, S.; Kilty, I. C.; Lee, A.; Lin, L.-L.; Lovering, F. E.; Lowe, M. D.; Mathias, J. P.; Morgan, H. M.; Murphy, E. A.; Papaioannou, N.; Patny, A.; Pierce, B. S.; Rao, V. R.; Saiah, E.; Samardjiev, I. J.; Samas, B. M.; Shen, M. W. H.; Shin, J. H.; Soutter, H. H.; Strohbach, J. W.; Symanowicz, P. T.; Thomason, J. R.; Trzuppek, J. D.; Vargas, R.; Vincent, F.; Yan, J.; Zapf, C. W.; Wright, S. W. *J. Med. Chem.* **2017**, *60*, 5521–5542.
36. Gopalsamy, A.; Narayanan, A.; Liu, S.; Parikh, M. D.; Kyne, R. E.; Fadeyi, O.; Tones, M. A.; Cherry, J. J.; Nabhan, J. F.; LaRosa, G.; Petersen, D. N.; Menard, C.; Foley, T. L.; Noell, S.; Ren, Y.; Loria, P. M.; Maglich-Goodwin, J.; Rong, H.; Jones, L. H. *J. Med. Chem.* **2017**, *60*, 3094–3108.
37. Lorthiois, E.; Anderson, K.; Vulpetti, A.; Rogel, O.; Cumin, F.; Ostermann, N.; Steinbacher, S.; Mac Sweeney, A.; Delgado, O.; Liao, S.-M.; Randl, S.; Rüdissler, S.; Dussauge, S.; Fettis, K.; Kieffer, L.; de Erkenez, A.; Yang, L.; Hartweg, C.; Argikar, U. A.; La Bonte, L. R.; Newton, R.; Kansara, V.; Flohr, S.; Hommel, U.; Jaffee, B.; Maibaum, J. *J. Med. Chem.* **2017**, *60*, 5717–5735.
38. Lovering, F.; Bikker, J.; Humblet, C. *J. Med. Chem.* **2009**, *52*, 6752–6756.