MedChemComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/medchemcomm

MedChemComm

ARTICLE

Insights into the impact of *N*- and *O*-methylation on aqueous solubility and lipophilicity using matched molecular pair analysis

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

T. J. Ritchie,*[a] S. J. F. Macdonald*[b] and S. D. Pickett[c]

The impact of *N*- and *O*-methylation on chromatographically measured lipophilicity and high throughput chemiluminescent nitrogen detection (CLND) aqueous solubility was studied using matched molecular pairs for data sets of amides, sulfonamides, ureas, carbamates, amines, carboxylic acids, alcohols and phenols. The extent to which solubility and lipophilicity are affected by *N*- or *O*-methylation is dependent on the nature of atoms and substituents around the nitrogen or oxygen atom. In some classes of amides, *N*-methylation unexpectedly increases solubility and lowers log $D_{7.4}$ considerably: this behaviour can be rationalised by conformational changes accompanying *N*-methylation that increase polar surface area, or by the disruption of one or more intramolecular hydrogen bonding motifs. Unlike amides, sulfonamide *N*-methylation always reduces solubility and increases lipophilicity, which again can be understood in terms of conformational effects. As expected, methylation of carboxylic acids lowers solubility and increases lipophilicity due to masking of the ionisable acidic group; however the magnitude of the reduction in solubility depends to some extent on the lipophilicity and molecular weight of the compound pairs under investigation.

Introduction

Matched molecular pair analysis (MMPA), the concept of which was introduced in 2005[1] has become a useful approach to investigate the effects of chemical transformations on various biological and physicochemical readouts, with algorithms now available to routinely fragment large data sets of compounds to generate useful MMP collections[2-3].

The results from MMP analyses have been reported by several groups and used to explore the impact of particular chemical transformations on physical properties and target assays relevant to drug discovery such as solubility,[4-9] lipophilicity,[6,8,9] protein binding,[4] membrane permeability,[5,9] P-glycoprotein efflux,[9] intrinsic clearance,[9,10] hERG inhibition,[5,6,9] metabolic stability[10,11] CYP450 inhibition,[5] and oral exposure.[4] The most appropriate statistical measures to accompany such analyses have also been discussed recently,[12] and the general topic has been reviewed.[13-15] Advantages of MMP analyses are that measured data are used, it is generally easier to predict differences in the value of a property associated with a structural change than it is to predict the value of a property directly from molecular structure[12] and the concept of chemical transformations where substituents are removed or added is well understood and frequently employed by medicinal chemists. The information generated from these studies is of considerable value to medicinal chemistry teams in lead optimisation programmes who are searching for novel analogues with an optimal balance of pharmacological potency and physicochemical properties.

In the context of the work to be discussed here, MMPs have been used to highlight the effect of methylating heteroatoms on aqueous solubility and calculated lipophilicity. We were interested in understanding more about the reported unusual behaviour of secondary amides, where *N*-methylation appeared to increase aqueous solubility despite an increase in calculated lipophilicity (clog *P*).[4] It was decided to conduct an MMP analysis on a larger set of nitrogen and oxygen heteroatom methylation substrates, whilst examining more closely the environment around each methylation site, and including chromatographically determined log $D_{7.4}$ values[16] alongside the high throughput CLND aqueous solubility measurements,[17] which are obtained by equilibrating a 5% dimethylsulfoxide (DMSO) solution (from a 10 mM stock solution) for one hour, filtering and then assaying the filtrate to determine the concentration of compound present.

Thus the following sections explore the impact of *N*-methylation on aqueous solubility and lipophilicity for a variety of functional groups (amide, sulfonamide, urea, carbamate and amine) and *O*-methylation substrates (carboxylic acids, alcohols and phenols).

[a] *TJR-Chem, Via Alberto 34C, 21020 Ranco (VA), Italy. Fax: +39 0331 975020; Tel: +39 0331 975020; E-mail: tim.j.ritchie@gmail.com.*
[b] *Fibrosis Discovery Performance Unit, GlaxoSmithKline Medicines Research Centre, Gunnels Wood Road, Stevenage SG1 2NY, UK. E-mail: simon.jf.macdonald@gsk.com.*
[c] *Computational and Structural Chemistry, GlaxoSmithKline Medicines Research Centre, Gunnels Wood Road, Stevenage SG1 2NY, UK.*

Electronic Supplementary Information (ESI) available: summary spreadsheet containing comprehensive statistical parameters for each class of methylation substrate; box plots showing influence of amine pK_a on change of log *D* and influence of log *D* of carboxylic acids on change of solubility. See DOI: 10.1039/x0xx00000x

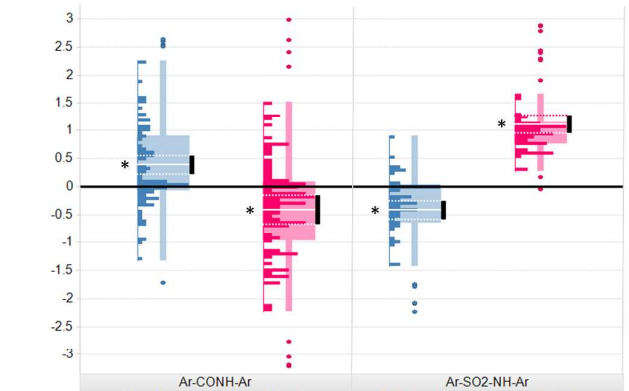
Overview of dataset

Experimental CLND solubility and chromatographic log *D* data were collated from GSK in-house data. MMPs were computed using the algorithm described by Hussain and Rea.[2] Initially all pairs where a hydrogen atom is replaced with a methyl group were retrieved without specifying the type of atom that the hydrogen was attached to; subsequently substructure searches within this set were carried out to identify specific subsets describing the *N*- or *O*-methylation of primary and secondary amides, primary and secondary sulfonamides, ureas (methylation of first or second NH), carbamates, aliphatic amines, anilines and aromatic nitrogens (e.g. pyrrole NH), where the nitrogen substituent is transformed from hydrogen to methyl, carboxylic acids, phenols and aliphatic alcohols. The various substrate classes were identified by substructure searching using SMARTS.[18] For each class of methylation substrate, the mean changes in solubility and lipophilicity were calculated together with the standard deviations, standard errors and 95% confidence intervals. A paired t test (for normally distributed values) or a signed-rank test (for non-normally distributed values) were used (JMP® software[19]) to determine whether the observed changes in solubility and lipophilicity were significantly different from zero.

To assess diversity within the MMP structures, molecules were clustered according to their fingerprint similarity as described by Gleeson et al[5]: all classes examined had at least 10 structurally diverse clusters, except where the number of pairs was less than 10.

Fig. 1 shows two representative examples of the type of results that are obtained: the change in CLND solubility and lipophilicity (log *D*) observed upon *N*-methylation of aromatic amides (i.e. amides derived from aromatic carboxylic acids and aromatic amines; left panel) and aromatic sulfonamides (derived from aromatic sulfonic acids and aromatic amines; right panel) are shown as box plots with distribution histograms. Compared to the reference line at zero (black horizontal line), *N*-methylation of aromatic amides increases solubility (mean +0.39 ± 0.86 Standard Deviations) and lowers log *D* (mean −0.42 ± 1.30; white solid lines in box plots) whilst *N*-methylation of aromatic sulfonamides lowers solubility (mean -0.42 ± 0.65) and increases log *D* (mean +1.11 ± 0.64). The plots also highlight the important point that the 'spread' or variance of the distributions from MMP analyses can vary depending on the particular methylation substrate: the experimental uncertainty in the chromatographic logD assay is estimated as being ± 0.25 log units but there is clearly a larger standard deviation and standard error associated with the change of log *D* for *N*-methylation of the amide pairs than for the sulfonamides. This may reflect more sensitivity towards the chemical environment (i.e. the substructural context) around the nitrogen atom in the former,[12] or unusual structures that result in stronger effects that appear as outliers, for example due to conformational changes upon *N*-methylation that change the way the partitioning system 'perceives' the polar and non-polar regions of the molecule. Some examples of this type of behaviour will be discussed in more detail below.

In some cases the number of pairs for particular methylations is low (less than 10) and it is important to bear in mind the statistical uncertainty that may surround such examples.[12] In addition, observed changes in solubility and lipophilicity that are small in magnitude (between -0.25 and 0.25 log units) may be inconsequential even if they are found to be statistically significant, given the inherent experimental uncertainty associated with the in vitro assays. However, we choose to include such examples in a qualitative sense to avoid loss of information and to highlight methylation substrates that could be of more interest if additional pairs were available.



	Ar-CONH-Ar		Ar-SO2-NH-Ar	
	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD
Count	105	98	60	60
Mean	0.39	-0.42	-0.42	1.11
StdDev	0.86	1.30	0.65	0.64

Figure 1. Box plots and distribution histograms showing the change in CLND solubility (blue) and lipophilicity (pink) upon *N*-methylation of aromatic amides (left panel) and aromatic sulfonamides (right panel). An asterisk indicates that the changes are statistically significantly different (p<0.05) from zero (paired t test in JMP v.10).The large shaded box represents 50% of the data either side of the median value; the shaded vertical line represents the data range for all non-outlier values; the solid white horizontal line within each box indicates the mean change observed. The dotted white lines indicate the upper and lower limits of the 95% confidence interval, which is also shown as the black vertical bar to the right of each box. If the interval is completely above or below zero, there is 95% confidence that the structural change generally causes a change in the measured property.[13]

Results and discussion

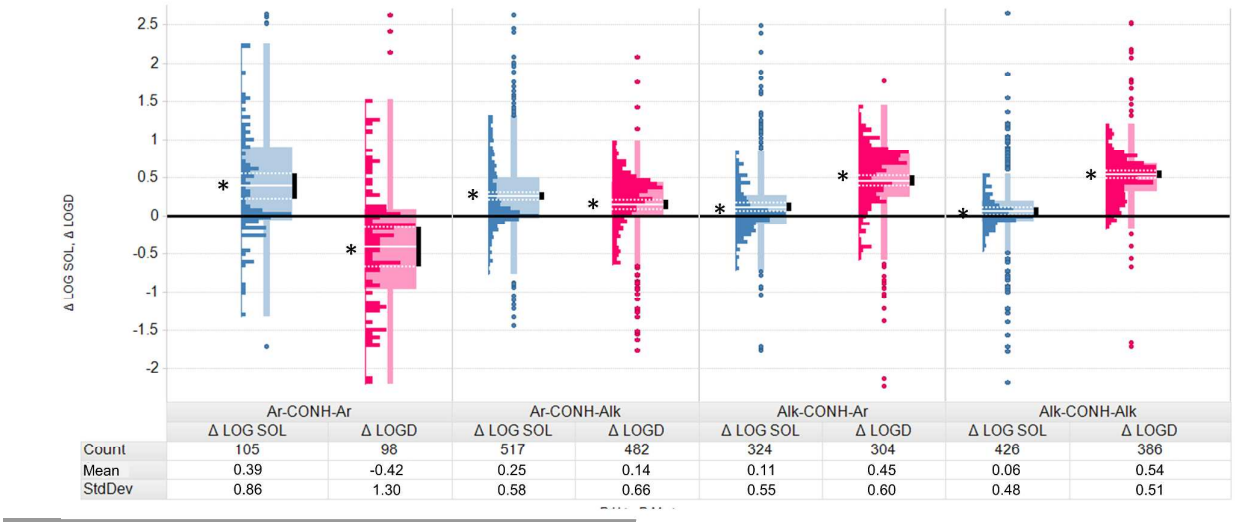
N-methylation substrates

N-methylation of amides

It has been reported previously[4] that there is an increase in log solubility upon *N*-methylation of secondary amides (+0.64 ± 0.73, n=142) despite a concomitant increase in clog *P* (mean change +0.31); this was attributed to the disruption of crystal packing that the amide NH may be involved in, decreased rigidity and planarity resulting in increased conformational mobility, which all serve to increase thermodynamic aqueous solubility. The nature of the amide substituents (i.e. aliphatic or aromatic) was not reported in this study. The current data set of secondary amides suggests the same trend (log solubility change +0.17 ± 0.58, n=1372; log *D* change +0.29 ± 0.73,

n=1270), although the magnitude of the solubility effect is lower with the increased number of data points in this study than observed previously. By comparison, if one considers a generic methyl to ethyl change to be a 'benchmark' MMP methylation transformation, the log changes in solubility and lipophilicity that result are −0.12 ± 0.45 (n=2639) and +0.57 ± 0.33 (n=2425) respectively. It should be borne in mind that the previous study measures thermodynamic solubility (agitation of compounds in 0.1 M phosphate buffer at pH 7.4 for 24 h at 25 °C) rather than kinetic solubility from a DMSO solution reported here. This may explain the difference in magnitude between the two sets with respect to the change in solubility, as the kinetic method does not take into account any solid state effects such as a reduction in lattice energy that may accompany the *N*-methylation of amides.

Using more specific SMARTS queries the secondary amides were separated into four classes based on whether the substituents on the carbonyl and amine components are aromatic or aliphatic highlight differences in the effect of *N*-methylation (see Fig. 2): amides derived from aliphatic acids show a small increase in solubility and higher log *D* upon *N*-methylation, whereas amides from aromatic acids exhibit a more pronounced solubility increase and less impact on log *D*. In the case of amides comprised of aromatic acids and anilines, log *D* decreases significantly (−0.42 ± 1.30, n=98): in this subset of amide structures, it is known that there is a strong preference for the E-(cis)-amide configuration to be populated upon *N*-methylation[20,21], observed both in molecular modelling studies with model systems and small molecule crystal structures from the Cambridge Crystallographic Database (release 2013);[22] the mean change in C-C-N-C torsion in *N*-phenylbenzamide crystal structures upon *N*-methylation is considerable (from 175.1 ± 11.7 degrees (n=622) to 29.0 ± 43.1 degrees (n=256)), reflecting the switch from Z-(trans) to E-amide geometry. Modelling studies with the model systems *N*-phenylbenzamide and *N*-methyl-*N*-phenylbenzamide (see Fig. 3) using the MOE software[23] (v.2012.10; MMFF94x forcefield[24]) suggests that the loss of planarity actually increases water-accessible polar surface area (from 39.6 to 52.9 Å^2), which leads to higher solubility and lower lipophilicity despite the loss of the polar NH and addition of a hydrophobic methyl group. This is consistent with a previous MMP study that demonstrated that the *N*-methylation of benzanilides increases aqueous solubility more significantly than other types of secondary amide due to the stability of the cis-amide conformation.[25] In a medicinal chemistry setting, this would suggest that tertiary benzamides have advantages over secondary congeners, for example in library design, being more soluble and more 3-dimensional in shape. However due to the conformational change upon *N*-methylation, biological activity is likely to be compromised if the secondary *trans* amide is part of an important pharmacophore.[26] *N*-Methylating primary aromatic amides has little effect on solubility (+0.07 ± 0.51, n=410) and increases lipophilicity (+0.45 ± 0.39, n=387); primary aliphatic amides behave similarly (solubility change +0.11 ± 0.42, n=140; logD change +0.35 ± 0.40, n=132).



	Ar-CONH-Ar		Ar-CONH-Alk		Alk-CONH-Ar		Alk-CONH-Alk	
	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD
Count	105	98	517	482	324	304	426	386
Mean	0.39	-0.42	0.25	0.14	0.11	0.45	0.06	0.54
StdDev	0.86	1.30	0.58	0.66	0.55	0.60	0.48	0.51

Figure 2. Impact on solubility (blue) and lipophilicity (pink) for various classes of amides after *N*-methylation. Ar = aromatic substituent; Alk = aliphatic substituent. An asterisk indicates that the changes are statistically significantly (p<0.05) different from zero (paired t test in JMP v.10). For an explanation of the box plot format see Fig. 1

It has been reported that the *N*-methylation of cyclic secondary amides does not increase solubility to the same extent as acyclic amides due to the cis-amide geometry enforced by the cyclic constraint.[25] This behaviour was confirmed in the current data set, particularly for amides derived from aromatic acids or anilines: *N*-methylation of cyclic amides has little effect on solubility (−0.04 ± 0.46, n=169) whereas *N*-methylation of acyclic amides increases solubility (+0.28 ± 0.62, n=778). The increase in log *D* is more pronounced in cyclic amides (+0.71 ± 0.48, n=167) relative to acyclic amides (+0.07 ± 0.79, n=717). Thus it appears that *N*-methylation of cyclic amides does not induce a conformational

change and more simply reflects the replacement of a polar hydrogen with a hydrophobic methyl group.

N-methylation of amides containing intramolecular H-bonding motifs

Molecules possessing H-bond acceptor and donor functionalities in particular orientations and proximities are often capable of forming intramolecular H-bonds (IMHBs). This structural class has received attention recently due to the observed impact of such interactions on conformational bias and physical properties such as membrane permeability, protein binding, aqueous solubility and lipophilicity.[27-34]

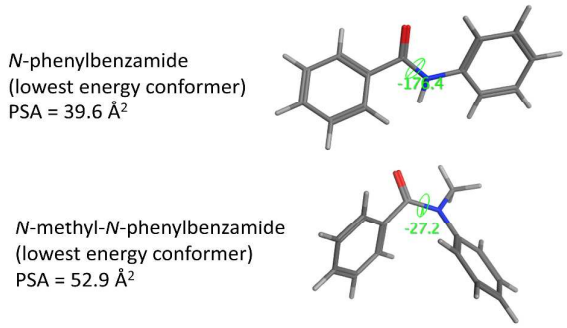


Figure 3. Change in conformational preference and water-accessible polar surface area upon *N*-methylation of *N*-phenylbenzamide. Stochastic conformer search carried out in MOE v.2012.10 using the MMFF94x forcefield. The amide C-C-N-C torsion angle is shown in green. PSA = Total water-accessible polar surface area calculated in MOE (ASA_P descriptor).

Several IMHB motifs involve the interaction between an amide NH and a suitable acceptor atom nearby, which can be a carbonyl group, an aromatic nitrogen atom, an ether oxygen or a fluorine atom (six representative examples are shown in Fig. 4). Small molecule crystal structures indicate that in the majority of cases the often planar conformations that are observed with such motifs reflect such intramolecular interactions, and one might expect that the *N*-methylation of such amides would disrupt the IMHBs and display solubility and lipophilicity changes that are different from the norm.

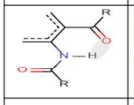
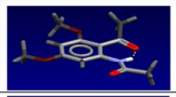
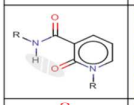
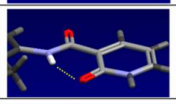
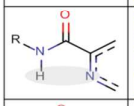
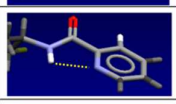
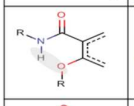
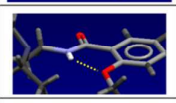
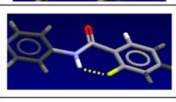
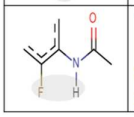
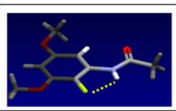
Amide IMHB motif	Crystal structure example	Crystal structure code	Corresponding SMARTS query
		CSD-YIBYUE[35]	[N;!H0]!@-[c,CX3]@[c,CX3]!@C=O (12 pairs)
		CSD-WEPCAW[36]	[N;!H0]!@-[C;X3]!@-[c,CX3]@[c,CX3]=O (5 pairs)
		CSD-GILKUJ[37]	[N;!H0]!@C!@c[n;X2] (103 pairs)
		CSD-BEZJIA[38]	OccC(=O)[NH1][c,C] (44 pairs)
		CSD-HOFVOO[39]	[NH1]!@[C;X3]!@ccF (10 pairs)
		CSD-POMVUJ[40]	[c,C;X3]!@[NH1]!@ccF (2 pairs)

Figure 4. Six amide substructures involved in intramolecular interactions with other atoms. The first column contains six amide substructures that commonly form intramolecular interactions between the amide NH and a neighbouring H-bond acceptor atom; second and third columns show a representative crystal structure illustrating the interactions and its reference code; the fourth column describes the SMARTS queries used to identify examples in the MMP data set.

The MMP acyclic amide data set was searched for these particular IMHB motifs using SMARTS strings and the effect of amide *N*-methylation on solubility and log *D* examined relative to other amides that did not possess internal H-bonds. Because the IMHB motifs above all contain aromatic substituents, amides derived from both aliphatic acids and aliphatic amines are obviously excluded in this case. The number of structures in the data set with the six types of IMHB discussed above was relatively small (19% of the total), but some interesting behaviour was seen in these derivatives. *N*-methylation of amides without IMHBs produces a moderate increase in both solubility (+0.27 ± 0.56, n=674) and log *D* (+0.22 ± 0.51, n=618) but in all the IMHB classes involving oxygen or nitrogen acceptors log *D* decreases upon *N*-methylation, although the magnitude of this effect depends on the type of IMHB (Fig. 5). Relatively few examples of the fluorine-containing IMHBs are available, so definitive conclusions cannot be made but the impact on solubility and lipophilicity appears smaller than in the other IMHB classes, perhaps reflecting the relatively weak nature of an aromatic fluorine atom as a hydrogen bond acceptor.[42]

In one particular case, albeit involving only a few pairs of molecules related to Linomide (roquinimex; lower structure in Fig. 6), log *D* decreases by -5.05 ± 3.00 (n=4) log units with a concomitant solubility change of +1.44 ± 1.32 (n=5) log units (Fig. 5, second panel from left). Log *D* changes of this magnitude prompted the retesting of the compounds involved, which confirmed that this behaviour was genuine. Such a gross change in lipophilicity can in part be explained by

inspection of the structures involved: in three of the four cases the amide is in fact involved in two distinct internal interactions, where the amide NH is hydrogen bonded to the pyridone carbonyl and the amide carbonyl interacts with the adjacent phenolic OH donor group (an interaction that is known to increase lipophilicity).[43] This results in an essentially rigid, planar structure where the polar functionality is masked (an example is the top structure shown in Fig. 6) and results in high lipophilicity (the average chromatographic log D values for these structures is 6.78 ± 2.06). Solubility is presumably also compromised by the rigid, planar structure. *N*-methylation perturbs both internal H-bonds by forcing the amide out-of-plane and exposing the remaining polar functionality resulting in the substantial decrease in lipophilicity. This phenomenon can be illustrated by the two closely related crystal structures shown in Fig. 6: the CSD-VIQFUX[44] structure exemplifies the completely flat conformation and the two internal polar interactions; the *N*-methyl derivative CSD-BEHDUO[45] crystallises as the cis-amide where both IMHBs and the planarity have been destroyed. Modelling studies with these structures confirmed that the water-accessible polar surface area (calculated in MOE) is considerably higher in the *N*-methyl analogue (134.79 Å^2) than in the N-H analogue (113.34 Å^2). Whilst this is a specific case, it serves as a reminder that under the right circumstances a small change in structure such as a methylation can result in profound changes in conformation and in physical properties.

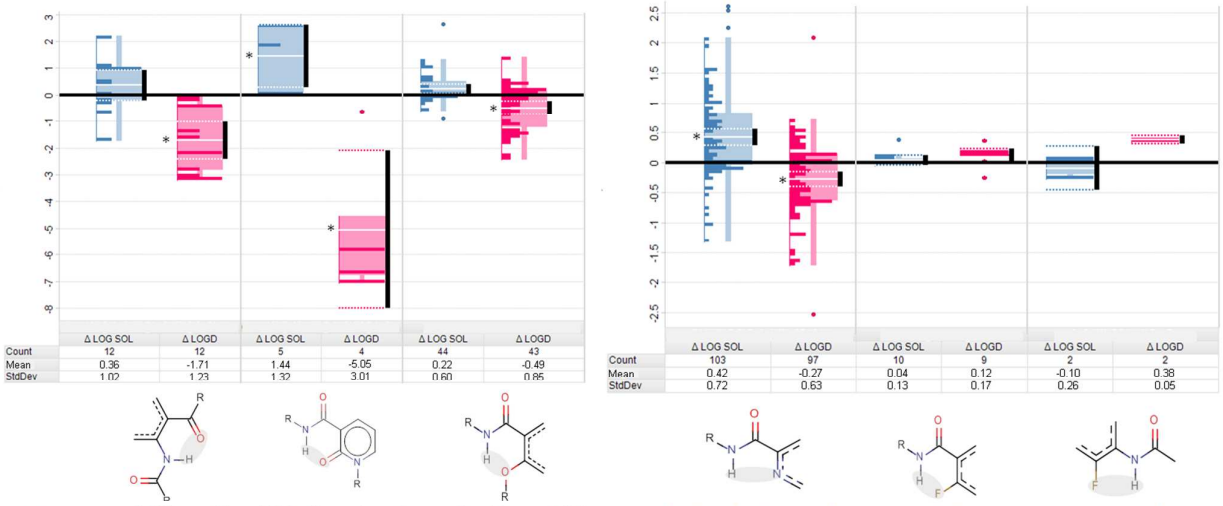


Figure 5. Impact on solubility and lipophilicity for various classes of amides containing intramolecular H-bonding motifs upon *N*-methylation. An asterisk indicates that the changes are statistically significantly ($p<0.05$) different from the mean values of amides with no IMHBs (Tukey-Kramer ANOVA, TIBCO Spotfire v.6.0.1[41]). For an explanation of the box plot format see Fig. 1.

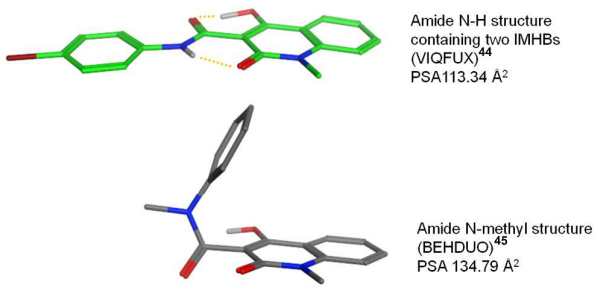


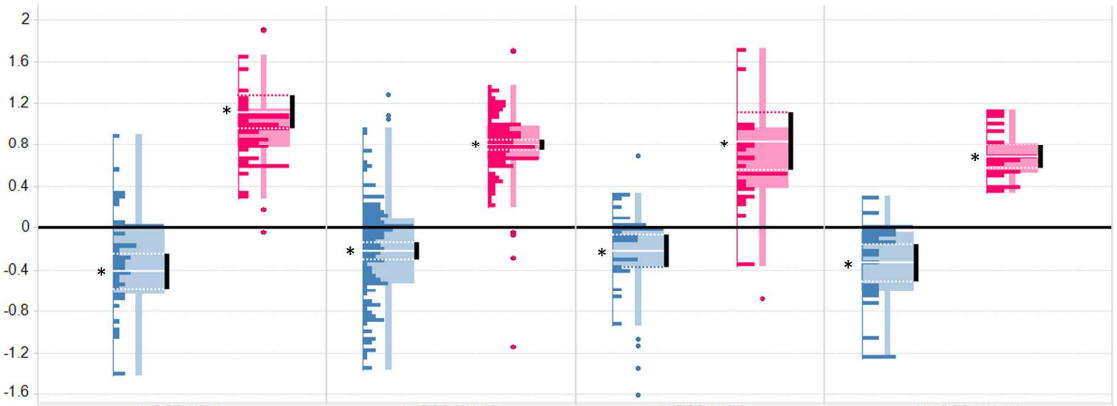
Figure 6. The disruption of intramolecular H-bonding by amide *N*-methylation. Crystal structures of two closely related pyridones illustrating the conformational change brought about by amide *N*-methylation leading to increased polar surface area.

N-methylation of sulfonamides

In contrast to amides, the *N*-methylation of sulfonamides results in more expected behaviour, with an increase in log *D* and decrease in solubility. There are some differences depending on whether they are derived from aromatic or aliphatic sulfonic acids and aliphatic or aromatic amines (Fig. 7), but these are not significantly different from one another. This behaviour can again be rationalised by conformational analysis: *N*-methylation of sulfonamides has little impact on conformation, with the C-S-N-C torsion in *N*-phenylbenzenesulfonamides changing from 64.14 ± 10.64 degrees (n=725) to 81.12 ± 18.98 (n=117) in small molecule crystal structures. *N*-methylation of aromatic sulfonamides decreases polar surface area and hence increases lipophilicity and decreases solubility. Furthermore, in some cases the sulfonamide NH group has some acidic character and is partially deprotonated at pH 7.4 so subsequent *N*-methylation will also increase lipophilicity and lower solubility due to the loss of such a solubilising moiety. This is evidenced by the larger change in chromatographic log *D* and solubility for the Ar-SO_2-NH-Ar class. *N*-methylation of primary aliphatic and aromatic sulfonamides had a lesser impact on solubility (-0.09 ± 0.35, n=9 and -0.10 ± 0.54, n=65 respectively) and lipophilicity (+0.28 ± 0.45, n=5 and +0.67 ± 0.19, n=60 respectively) than secondary sulfonamides.

MedChemComm

ARTICLE



	Ar-SO2-NH-Ar		Ar-SO2-NH-Alk		Alk-SO2-NH-Ar		Alk-SO2-NH-Alk	
	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD
Count	60	60	156	150	37	37	23	17
Mean	-0.42	1.11	-0.23	0.79	-0.22	0.83	-0.33	0.68
StdDev	0.65	0.64	0.53	0.34	0.48	0.86	0.43	0.24

Figure 7. Impact on solubility and lipophilicity for various classes of sulfonamides after *N*-methylation. Ar = aromatic substituent; Alk = aliphatic substituent. An asterisk indicates that the changes are statistically significantly ($p<0.05$) different from zero (paired t test in JMP v.10). For an explanation of the box plot format see Fig. 1.

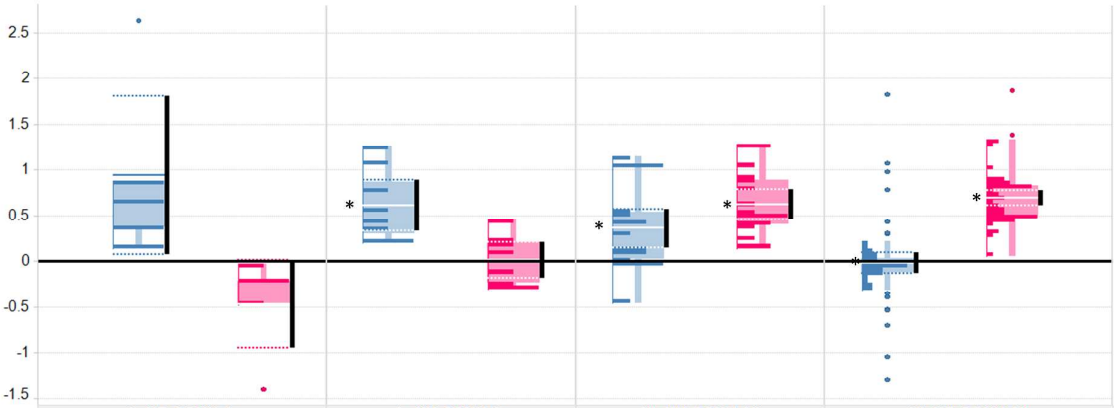
This secondary amide/sulfonamide disparity can be further illustrated by examining the behaviour of MMPs where an amide moiety is replaced with a sulfonamide. When the N-H is present, there is little impact on log solubility (+0.10 ± 0.64, n=1009) or log *D* (+0.12 ± 0.79, n=1227) when a -CONH- group is replaced by -SO_2NH-. However, the -CON(CH_3)- to -SO_2N(CH_3)- transform results in a significant decrease in log solubility (−0.52 ± 0.62, n=117) and increase in log *D* (+0.89 ± 0.56, n=153), due to the increased polarity of the *N*-methyl amide relative to the *N*-methyl sulfonamide.

_N_-methylation of ureas and carbamates

The analysis of MMPs for urea *N*-methylation (of either the first or second NH) was hampered by the limited number of example pairs in some cases. *N*-methylation of ureas derived from anilines appears to increase solubility considerably, even more so than observed for amides but due to the small number of pairs these effects do not reach significance (Fig. 8). Aromatic ureas appear to behave differently to aliphatic ureas upon *N*-methylation, but more matched pairs would be required to confirm this. Modelling studies suggest that the first *N*-methylation of aromatic ureas increases conformational freedom, and encourages a switch from trans to cis amide geometry. Water accessible polar surface area does not change significantly in this case.

MedChemComm

ARTICLE



	Ar-NR-CONH-Ar		Ar-NH-CONH-Ar		Alk-NH-CONH-Alk		Alk-NR-CONH-Alk	
	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD
Count	5	5	8	8	18	17	57	55
Mean	0.94	-0.47	0.61	0.01	0.36	0.62	-0.02	0.69
StdDev	0.99	0.55	0.40	0.28	0.46	0.35	0.45	0.31

Figure 8. Impact on solubility and lipophilicity for various classes of ureas after *N*-methylation. Ar = aromatic substituent; Alk = aliphatic substituent; R = either aromatic or aliphatic substituent. An asterisk indicates that the changes are statistically significantly ($p<0.05$) different from zero (paired t test in JMP v.10). For an explanation of the box plot format see Fig. 1.

Methylation of one nitrogen of ureas derived from aliphatic amines increases solubility (+0.36 ± 0.46, n=18) despite increasing log *D* (+0.62 ± 0.35, n=17), however methylation of the second nitrogen does not affect solubility (−0.02 ± 0.45, n=57) and increases log *D* (+0.69 ± 0.31, n=55). In this case, the distribution of pairs describing the change in log *D* is not normal (Fig. 8, far right panel, pink plot), being bimodal in nature. Upon inspection, the group of pairs with a greater than average increase in log *D* tend to be of higher lipophilicity and higher molecular weight (MW) than the pairs with a less than average change in log *D*. In fact there is a reasonable correlation between the log *D* of the methylated urea and the change in log D upon methylation ($r = 0.777$; $r^2 = 0.603$) and to a lesser extent with the MW of the urea and change in log *D* ($r = 0.428$; $r^2 = 0.183$) for this set (MW and log *D* are also correlated with each other ($r = 0.401$; $r^2 = 0.161$)). To our knowledge such an interdependence between the observed change in a property and the gross properties of the molecules has not been reported previously in MMP analyses.

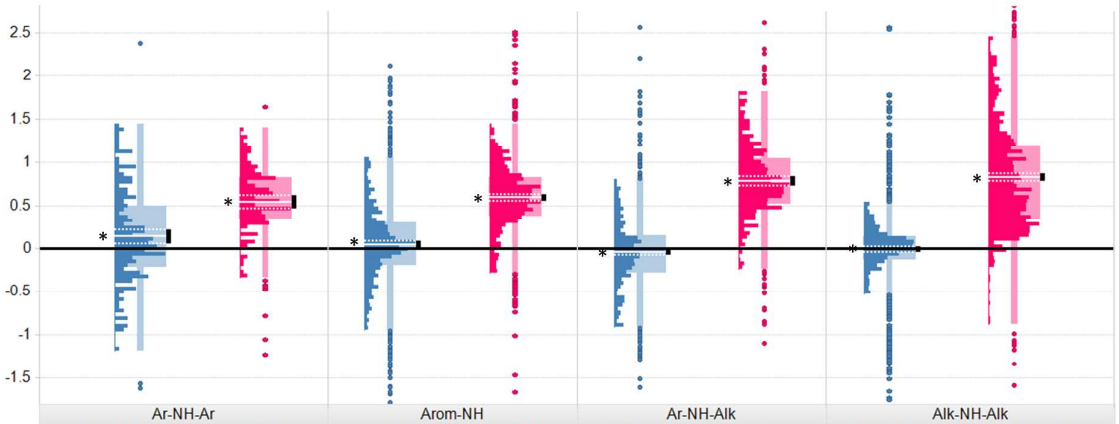
Methylation of the nitrogen in carbamates has little effect on solubility (−0.09 ± 0.40, n=50) but increases log *D* (+0.58 ± 0.52, n=40). Modelling studies with phenyl *N*-phenylcarbamate and phenyl *N*-methyl-*N*-phenylcarbamate suggest that the lowest energy structures have the same conformation.

***N*-methylation of amines**

Secondary amines were categorised as having either two aromatic substituents, two aliphatic substituents, one of each, or being part of an aromatic system (e.g. pyrrole NH). In all cases, *N*-methylation has little impact on solubility (Fig. 9). Log *D* increases in all cases, but less so in the case of amines with two aromatic substituents or part of an aromatic ring. One might expect to observe a change in the amine nitrogen pK_a upon *N*-methylation of a secondary dialkylamine to afford a tertiary amine, which was investigated using calculated and measured (where available) pK_a values for the pairs in question. There was a consistent reduction in calculated basicity of 1.10 ± 0.65 units for the amine nitrogen upon methylation;[46,47] but this does not appear to have an impact on aqueous solubility. Similar reductions in basicity are seen after *N*-methylation of the simple secondary amines piperidine (pK_a 11.22) and morpholine (pK_a 8.36) resulting in a reduction in pK_a of 1.14 (to 10.08) and 0.95 (to 7.41) respectively.[48] In the case of dialkylamine pairs the standard deviation of the distribution of log *D* change is somewhat higher than for aromatic amines and is skewed towards lower than average values (Fig. 9, right hand panel, pink plot). Inspection of the calculated pK_a values suggest that this is due to the most basic amines (pK_a >10), which tend to exhibit less than the average change in log *D* upon methylation (+0.45 ± 0.59, n=251). A box plot showing this effect in more detail is available in the Supporting Information.

MedChemComm

ARTICLE



	Ar-NH-Ar		Arom-NH		Ar-NH-Alk		Alk-NH-Alk	
	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD
Count	181	158	857	810	604	571	1121	1012
Mean	0.14	0.54	0.04	0.60	-0.04	0.78	-0.01	0.83
StdDev	0.59	0.51	0.58	0.54	0.53	0.69	0.48	0.78

Figure 9. Impact on solubility and lipophilicity for various classes of amines after *N*-methylation. Ar = aromatic substituent; Alk = aliphatic substituent. AromNH = pairs where the nitrogen atom is part of an aromatic ring. An asterisk indicates that the changes are statistically significantly ($p<0.05$) different from zero (paired t test in JMP v.10). For an explanation of the box plot format see Fig. 1.

O-methylation substrates

Solubility and log *D* data were collected for MMP pairs of aromatic and aliphatic carboxylic acids, phenols and alcohols.

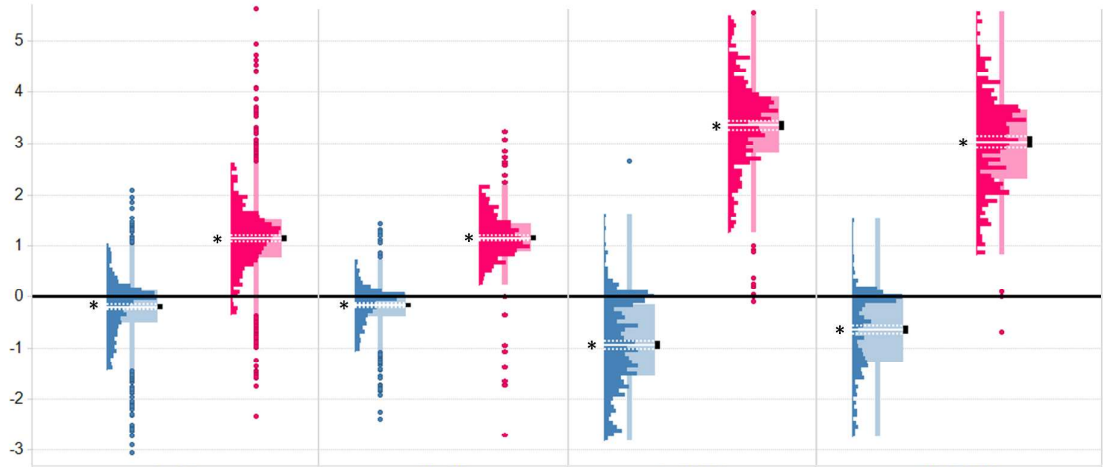
O-methylation of carboxylic acids

As would be expected, the methylation of a carboxyl group results in a significant increase in log *D* and a concomitant decrease in solubility due to the masking of the acidic ionisable centre. There is a small but significant difference observed between aromatic and aliphatic acids with solubility decreasing by −0.94 ± 0.86 (n=505) and −0.64 ± 0.78 (n=349) log units, and log *D* increasing by +3.36 ± 0.97 (n=475) and +3.02 ± 1.07 (n=314) log units respectively (see Fig. 10).

In the case of aliphatic acids and the change in solubility upon methylation it was noted that the distribution of pairs is not normal, instead forming a bimodal distribution of two groups either side of the mean value (Fig.10 far right panel, blue plot). As found with the alkyl ureas, this behaviour appears to be dependent on the physicochemical properties of the acids: the pairs with a greater than average decrease in solubility tend to be of higher molecular weight (MW) and higher lipophilicity than the pairs with a less than average change in solubility. There is some correlation between log *D* and change in solubility ($r = 0.560$; $r^2 = 0.309$) and MW and change in solubility ($r = 0.409$; $r^2 = 0.175$) for this set (note that MW and log D are also correlated with each other ($r = 0.591$; $r^2 = 0.350$). Acidic pK_a values were calculated for the carboxylic acid-containing compounds and there is no correlation between these and MW, log D or change in solubility.; thus low molecular weight (109-400) acids can accommodate *O*-methylation with less impact on solubility (mean change −0.36 ± 0.70, n=153), whereas the solubility of higher molecular weight (400-909) acids decreases more significantly (mean −0.87 ± 0.78, n=196). Similarly, the solubility change upon methylation for low log *D* (<2) acids (mean −0.21 ± 0.58, n=165) is much less than for higher log D (>2) acids (mean −0.91 ± 0.79, n=219). A box plot showing this effect in more detail is available in the Supporting Information. One explanation is that the solubility of a large lipophilic acid is driven predominantly by the ionised carboxyl group, which when masked, causes a large decrease in solubility; a small, polar acid possesses additional solubilising features, which can still retain some solubility when the carboxylic acid moiety is blocked.

MedChemComm

ARTICLE



	Ar-OH		Alk-OH		Ar-COOH		Alk-COOH	
	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD	Δ LOG SOL	Δ LOGD
Count	964	883	610	590	505	475	349	314
Mean	-0.21	1.13	-0.19	1.14	-0.94	3.36	-0.65	3.02
StdDev	0.65	0.94	0.49	0.66	0.86	0.97	0.79	1.07

Figure 10. Impact on solubility and lipophilicity for carboxylic acids, phenols and alcohols after *O*-methylation. Ar = aromatic substituent; Alk = aliphatic substituent. An asterisk indicates that the changes are statistically significantly ($p<0.05$) different from zero (paired t test in JMP v.10). For an explanation of the box plot format see Fig. 1.

O-methylation of phenols and alcohols

The *O*-methylation of phenols and alcohols results in identical behaviour with respect to changes in solubility and lipophilicity: the solubility of phenols decreases by −0.21 ± 0.65 log units (n=964); for alcohols the decrease is −0.19 ± 0.49 (n=610). Log *D* increases by +1.13 ± 0.94 (n=883) for phenols and +1.14 ± 0.66 (n=590) for alcohols. The observed behaviour of phenols is similar to the results previously reported[4] (Log solubility change −0.22 ± 0.84; n=17) but in this earlier study alcohols did not show any change in solubility (Log solubility change −0.01 ± 0.68; n=203).

As discussed above with regard to IMHB motifs, it was noted that *O*-methylation of phenols that are able to interact with an ortho carbonyl group results in a much smaller increase in log *D* (+0.12 ± 1.06, n=57), although the change in solubility was similar to other phenols (−0.14 ± 0.76, n=59).

Summary and conclusions

This analysis confirms the previous reports[4,25] that the *N*-methylation of amides tends to increase solubility despite increasing lipophilicity, but reveals that the impact of *N*-methylation of amides on aqueous solubility and lipophilicity varies considerably depending on a) the type of substituent (aromatic or aliphatic) attached to the amide, b) whether the amide is cyclic or acyclic[25], and c) whether the amide NH and carbonyl groups are involved in intramolecular H-bonds. By perturbing the rigid, planar nature of secondary aromatic amides, *N*-methylation decreases lipophilicity by increasing the overall water-accessible polar surface area. The reduction in log *D* observed in some cases may be rationalised by conformational changes upon methylation, and is supported by molecular modelling studies and crystal structure information.

Although the number of MMP examples is relatively low, the *N*-methylation of amides that result in the disruption of IMHBs (via the amide NH or carbonyl or both) appears to increase solubility and lower log *D* more than expected, and in particular cases the impact can be dramatic when more than one intramolecular interaction is interrupted. Although in this article the study of the effect of IMHBs on solubility and lipophilicity has focused primarily on amides substructures, one might expect to observe similar behaviour in molecules where there are other suitable structural features that can form intramolecular interactions.

N-methylation of ureas also appears to increase solubility considerably in some cases but more example pairs would be required to confirm this behaviour. In the case of alkyl ureas, the extent to which *N*-methylation of the second urea nitrogen

affects log *D* appears to depend on the lipophilicity and MW of the ureas.

In contrast to amides, sulfonamide *N*-methylation generally lowers solubility and increases lipophilicity. This behaviour could again be rationalised by the differing conformational preferences of sulfonamides vs. amides.

N-methylation of aliphatic and aromatic secondary amines has little impact on solubility. Log *D* increases but less so with amines with two aromatic substituents. The increase in log *D* upon methylation of strongly basic dialkylamines is less than the average.

With respect to *O*-methylation, masking a carboxylic acid significantly increases log *D* leading to a significant lowering of solubility. In the case of alkyl acids, as found with the alkyl ureas and change in log *D*, the impact on solubility depends on the gross molecular properties of the pairs involved such as molecular weight and lipophilicity. Lower molecular weight acids or those with lower chromatographic log *D* could accommodate *O*-methylation with less impact on solubility. This is an important observation with regard to the generation of esters as potential prodrugs.

The *O*-methylation of both phenols and alcohols results in a modest decrease in solubility and larger increase in lipophilicity. *O*-methylation of phenols engaged in an IMHB to an adjacent carbonyl group produced a smaller effect on lipophilicity.

As pointed out previously[5,25] the current study re-emphasises the importance of the local chemical environment around the methylation heteroatom in MMP analyses: reporting a change in solubility or lipophilicity for classes of molecules that are too generic may be misleading if a chemistry project is focused on structural frameworks that happen to behave differently from the norm, such as in the case of benzanilides and cyclic *vs*. acyclic amides.[25] This highlights the need to provide MMP data in a format that allows searching for specific chemical queries that are most relevant to the project in question. As pointed out recently, large standard deviations associated with MMP transformations suggest that further refinement may be necessary to understand more about the chemical environment around the transformation.[12] For example one could examine the individual chemotypes obtained by clustering to determine which structures exhibit unusual behaviour. However in doing so one runs the risk of having too few examples left from which to glean information with any confidence. Thus there is always a trade-off between specificity and generality that must be borne in mind, as discussed previously.[6,13,14]

Fig. 11 summarises the changes in solubility and lipophilicity upon *N*-methylation for the various compound classes discussed in this study. From this plot, it may appear logical to equate a reduction in aqueous solubility with an increase in lipophilicity and this observation has been made by several groups using MMP analysis on, for example, the impact of various substituents on solubility and lipophilicity.[4,5] However it is important to remember that whilst plots such as Fig. 11 are an appropriate way to present the relationship between the average changes in solubility and lipophilicity for the various classes of *N*-methylation substrate, plots using averaged data may suffer from correlation inflation[49] and should not necessarily be interpreted as suggesting that there is a strong correlation between change in solubility and change in lipophilicity; if one fits a straight line to the mean data in Fig. 11, the correlation between change in solubility and change in lipophilicity appears robust (r^2 = 0.722, n=21), however if all the underlying data points are included the correlation is in fact very low (r^2 = 0.062; n=5173). That is not to say that lipophilicity does not play a role in solubility: if one plots the actual solubilities of all the *N*-methylated compounds against their measured lipophilicities there is some correlation (r^2 = 0.317, n=5173). Whilst the overall correlation between change in solubility and change in lipophilicity is low across the entire data set, there are some differences between the *N*-methylation classes: the highest correlations are observed in the sets of aromatic sulfonamides (Ar-SO_2NH-Ar; r^2 = 0.284) and aromatic amides (Ar-CONH-Ar; r^2 = 0.148); the lowest were observed for alkyl ureas (Alk-NR-CONH-Alk; r^2 = 0.000) and aromatic primary amides (Ar-$CONH_2$; r^2 = 0.000).

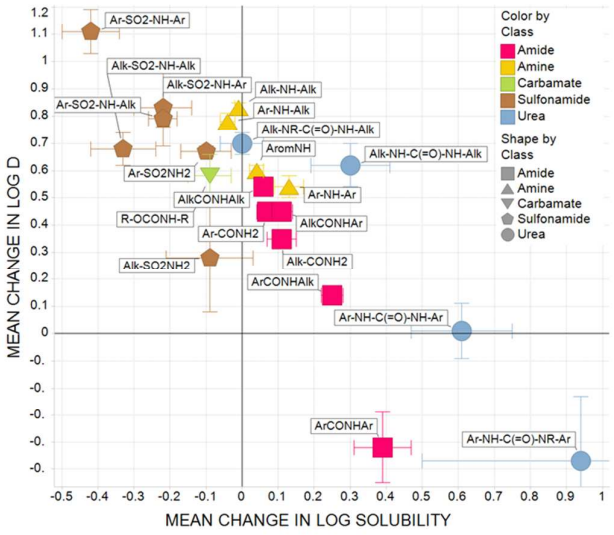


Figure 11. Scatter plot summarising the impact of *N*-methylation on aqueous solubility and lipophilicity for a variety of substrates. Points represent the mean changes observed for each class of methylation substrate, which are coloured and shaped by class. Error bars indicate the standard error of the mean. Ar = aromatic substituent; Alk = aliphatic substituent; R = either aromatic or aliphatic substituent.

Acknowledgements

The authors thank Darren Green, Stephen Swanson, John Pritchard, Lee Harrison, Ceara Rea, Martin Saunders and Klara Valko-Slegel for their help, advice and constructive comments, and the referees of the manuscript for their useful suggestions to improve the article. TJR thanks GlaxoSmithKline for funding this analysis and allowing access to the GSK MMP data collection.

References

1 P. W. Kenny and J. Sadowski. Structure Modification in Chemical Databases, in Chemoinformatics in Drug Discovery (ed T. I. Oprea), 2005, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG.

2 J. Hussain and C. Rea. Computationally Efficient Algorithm To Identify Matched Molecular Pairs (MMPs) in Large Data Sets. J. Chem. Inf. Model. 2010, **50**, 339−348.

3 D. J. Warner, E. J. Griffen and S. A. St-Gallay. WizePairZ: A Novel Algorithm To Identify, Encode, and Exploit Matched Molecular Pairs with Unspecified Cores in Medicinal Chemistry. *J. Chem. Inf. Model*. 2010, **50**, 1350−1357.

4 A. G. Leach, H. D. Jones, D. A. Cosgrove, P. W. Kenny, L. Ruston, P. MacFaul, J. M. Wood, N. Colclough and B. Law. Matched Molecular Pairs as a Guide in the Optimization of Pharmaceutical Properties; a Study of Aqueous Solubility, Plasma Protein Binding and Oral Exposure. *J. Med. Chem*. 2006, **49**, 6672−6682.

5 P. Gleeson, G. Bravi, S. Modi, D. Lowe. ADMET Rules of Thumb II: A Comparison of the Effects of Common Substituents on a Range of ADMET Parameters. *Bioorg. Med. Chem*. 2009, **17**, 5906−5919.

6 G. Papadatos, M. Alkarouri, V. J. Gillet, P. Willett, V. Kadirkamanathan, C. N. Luscombe, G. Bravi, N. J. Richmond, S. D. Pickett, J. Hussain, J. M. Pritchard, A. W. J. Cooper and S. J. F. Macdonald. Lead Optimization Using Matched Molecular Pairs: Inclusion of Contextual Information for Enhanced Prediction of hERG Inhibition, Solubility, and Lipophilicity. *J. Chem. Inf. Model*. 2010, **50**, 1872−1886.

7 L. Zhang, H. Zhu, A. Mathiowetz and H. Gao. Deep understanding of structure–solubility relationship for a diverse set of organic compounds using matched molecular pairs. *Bioorg. Med. Chem*. 2011, **19**, 5763-5770.

8 D. G. Brown, M. M. Gagnon and J. Boström. Understanding Our Love Affair with p‑Chlorophenyl: Present Day Implications from Historical Biases of Reagent Selection. *J. Med. Chem*. 2015, **58**, 2390-2405.

9 C. E. Keefer, G. Chang and G. W. Kauffman. Extraction of Tacit Knowledge from Large ADME Data Sets via Pairwise Analysis. *Bioorg. Med. Chem*. 2011, **19**, 3739−3749.

10 A. G. Dossetter. A Statistical Analysis of in Vitro Human Microsomal Metabolic Stability of Small Phenyl Group Substituents, Leading to Improved Design Sets for Parallel SAR Exploration of a Chemical Series. *Bioorg. Med. Chem*. 2010, **18**, 4405−4414.

11 A. G. Dossetter, A. Douglas and C. O'Donnell. A Matched Molecular Pair Analysis of in Vitro Human Microsomal Metabolic Stability Measurements for Heterocyclic Replacements of Di-Substituted Benzene Containing Compounds · Identification of Those Isosteres More Likely To Have Beneficial Effects. *Med. Chem. Commun*. 2012, **3**, 1164−1169.

12 C. Kramer, J. E. Fuchs, S. Whitebread, P. Gedeck and K. R. Liedl. Matched Molecular Pair Analysis: Significance and the Impact of Experimental Uncertainty. *J. Med. Chem*. 2014, **57**, 3786-3802.

13 E. Griffen, A. G. Leach, G. R. Robb and D. J. Warner. Matched Molecular Pairs as a Medicinal Chemistry Tool. *J. Med. Chem*. 2011, **54**, 7739−7750.

14 A. M. Wassermann, D. Dimova, P. Iyer and J. Bajorth. Advances in Computational Medicinal Chemistry: Matched Molecular Pair Analysis. *Drug Development Research* 2012, **73**, 518-527.

15 A. G. Dossetter, E. J. Griffen and A. G. Leach. Matched Molecular Pair Analysis in Drug Discovery. *Drug Discovery Today* 2013, **18**, 724− 731.

16 K. Valkó. Chromatographic hydrophobicity index by fast-gradient RPHPLC: a high-throughput alternative to log P/log D. *Anal. Chem*. 1997, **69**, 2022-2029.

17 S. N. Bhattachar, L. A. Deschenes and J. A. Wesley. Solubility: It's not just for Physical Chemists. *Drug Discovery Today* 2006, **11**, 1012-1018.

18 SMARTS - A Language for Describing Molecular Patterns: http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html. Daylight Chemical Information Systems, Inc., Laguna Niguel, CA.

19 JMP®, Version 10.0.2. SAS Institute Inc., Cary, NC, 1989-2014.

20 A. Itai, Y. Toriumi, N. Tomioka, H. Kagechika, I. Azumaya, and K. Shudo. Stereochemistry of N-methylbenzanilide and benzanilide. *Tet. Lett*. 1989, **30**, 6177-6180.

21 J. P. Kilburn, J. Kehler, M. Langgård, M. N. Erichsen, S. Leth-Petersen, M. Larsen, C. T. Christoffersen and J. Nielsen. N-Methylanilide and N-methylbenzamide derivatives as phosphodiesterase 10A (PDE10A) inhibitors. *Bioorg. Med. Chem*. 2013, **21**, 6053-6062.

22 F. H. Allen, *Acta Cryst*. B58, 380-388, 2002 "The Cambridge Structural Database: a quarter of a million crystal structures and rising" [DOI: 10.1107/S0108768102003890].

23 Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015.

24 T. A. Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comp. Chem*. 1996, **17**, 490-519.

25 A. M. Birch, P. W. Kenny, I. Simpson and P. R. O. Whittamore. Matched molecular pair analysis of activity and properties of glycogen phosphorylase inhibitors. *Bioorg. Med. Chem. Lett*. 2009, **19**, 850-853

26 D. Middlemiss and S. P. Watson. A medicinal chemistry case study: an account of an angiotensin II antagonist drug discovery programme. *Tetrahedron*. 1994, **50**, 13049-13080.

27 B. Kuhn, P. Mohr and M. Stahl. Intramolecular Hydrogen Bonding in Medicinal Chemistry. *J. Med. Chem*. 2010, **53**, 2601-2611.

28 P. D'Andrea, S. Mauro, M. Porcelloni, C. Rossi, M. Altamura, R. M. Catalioto, S. Giuliani, C. A. Maggi and D. Fattori. hNK2 receptor antagonists. The use of intramolecular hydrogen bonding to increase solubility and membrane permeability. *Bioorg. Med. Chem. Lett*. 2011, **21**, 1807−1809.

29 S. B. Rafi, B. R. Hearn, P. Vedantham, M. P. Jacobson and A. R. Renslo. Predicting and improving the membrane permeability of peptidic small molecules. *J. Med. Chem*. 2012, **55**, 3163−3169.

30 M. Shalaeva, G. Caron, Y. A. Abramov, T. N. O'Connell, M. S. Plummer, G. Yalamanchi, K. A. Farley, G. H. Goetz, L. Philippe and M. J. Shapiro. Integrating Intramolecular Hydrogen Bonding (IMHB) Considerations in Drug Discovery Using ΔlogP As a Tool. *J. Med. Chem*. 2013, **56**, 4870-4879.

31 D. J. P. Pinto, M. J. Orwat, S. Wang, J. M. Fevig, M. L. Quan, E. Amparo, J. Cacciola, K. A. Rossi, R. S. Alexander, A. M. Smallwood, J. M. Luettgen, L. Liang, B. J. Aungst, M. R. Wright, R. M. Knabb, P. C. Wong, R. R. Wexler and P. Y. S. Lam. Discovery of 1-[3-(Aminomethyl)phenyl]-N-[3-fluoro-2¢-(methylsulfonyl)-[1,1'-biphenyl]-4-yl]-3-(trifluoromethyl)-1H-pyrazole-5-carboxamide (DPC423), a highly potent, selective, and orally bioavailable inhibitor of blood coagulation factor Xa. *J. Med. Chem*. 2001, **44**, 566-578.

32 M. L. Quan, P. Y. S. Lam, Q. Han, D. J. P. Pinto, M. Y. He, R. Li, C. D. Ellis, C. G. Clark, C. A. Teleha, J.-H. Sun, R. S. Alexander, S. Bai, J. M. Luettgen, R. M. Knabb, P. C. Wong and R. R. Wexler. Discovery of 1-(3'-Aminobenzisoxazol-5'-yl)-3-trifluoromethyl-N-[2-fluoro-4-[(2'-dimethylaminomethyl)imidazol-1-yl]phenyl]-1H-pyrazole-5-carboxyamide hydrochloride (Razaxaban), a highly potent,

selective, and orally bioavailable factor Xa inhibitor. *J. Med. Chem*. 2005, **48**, 1729-1744.

33 D. Chopra and T. N. G. Row. Evaluation of the interchangeability of C-H and C-F groups: insights from crystal packing in a series of isomeric fluorinated benzanilides. *Cryst. Eng. Comm*. 2008, **10**, 54-67.

34 M. M. Weiss, T. Williamson, S. Babu-Khan, M. D. Bartberger, J. Brown, K. Chen, Y. Cheng, M. Citron, M. D. Croghan, T. A. Dineen, J. Esmay, R. F. Graceffa, S. S. Harried, D. Hickman, S. A. Hitchcock, D. B. Horne, H. Huang, R. Imbeah-Ampiah, T. Judd, M. R. Kaller, C. R. Kreiman, D. S. La, V. Li, P. Lopez, S. Louie, H. Monenschein, T. T. Nguyen, L. D. Pennington, C. Rattan, T. San Miguel, E. A. Sickmier, R. C. Wahl, P. H. Wen, S. Wood, Q. Xue, B. H. Yang, V. F. Patel and W. Zhong. Design and preparation of a potent series of hydroxyethylamine containing β-secretase inhibitors that demonstrate robust reduction of central β –amyloid. *J. Med. Chem*. 2012, **55**, 9009-9024.

35 G.-B. Liu, Y.-Z. Li, Y.-M. Shen, Q. Xu and J.-X. Li. *N*-(2-Acetyl-3,5-dimethoxy-phenyl)acetamide. *Acta Cryst*. 2007, E63, o2420-o2421.

36 S. Long, M. Siegler and T. Li. *N*-(3-Chloro-2-methyl-phen-yl)-2-oxo-1,2-dihydro-pyridine-3-carboxamide. *Acta Cryst*. 2006, E62, o4278-o4279.

37 J. A. Kalow and A. G. Doyle. Enantioselective fluoride ring opening of aziridines enabled by cooperative Lewis acid catalysis. *Tetrahedron*. 2013, **69**, 5702-5709.

38 R. Tanaka and N. Hirayama. Structure of bromopride. *Anal. Sci. X-ray Struct. Anal. Online*. 2004, **20**, x79-x80.

39 J. McMahon, F. P. Anderson, J. F. Gallagher and A. J. Lough. A structural systematic study of three isomers of difluoro-N-(4-pyridyl)benzamide. *Acta Cryst*. 2008, **C64**, o493-o497.

40 K. Xie, Y. Lou, J. Zheng, Q. Zhao and Y. Wei. 2'-Fluoro-3',5'-dimethoxyacetanilide. *Acta Cryst*. 2009, **E65**, o205.

41 TIBCO Spotfire®, Version 6.0.1. TIBCO Spotfire Inc., 3303 Hillview Avenue, Palo Alto, CA 94304.

42 J. A. K. Howard, V. J. Hoy, D. O'Hagan and G. T. Smith. How good is fluorine as a hydrogen bond acceptor? Tet. 1996, 52, 12613-12622.

43 Modeling Chemical Structure-Log P in Drug Design Strategies – Quantitative Approches. Edited by David J. Livingstone and Andrew M. Davis. Royal Society of Chemistry 2012.

44 I. V. Ukrainets, A. A. Tkach, L. V. Sidorenko and O. V. Gorokhova. *Khim. Get. Soedin*. 2006, 1508.

45 K. B. Dasari and T. Srikrishnan. Crystal structure and conformation of Linomide [*N*-methyl-*N*-phenyl-1,2-dihydro-4-hydroxy-1-methyl-2-oxo-quinoline-3-carboxamide]: A novel immunomodulator. *J. Chem. Cryst*. 2002, **32**, 499-504.

46 D. M. Floyd, S. D. Kimball, J. Krapcho, J. Das, C. F. Turk, R. V. Moquin, M. W. Lago, K. J. Duff, V. G. Lee, R. E. White, R. E. Ridgewell, S. Moreland, R. J. Brittain, D. E. Normandin, S. A. Hedberg and G. G. Cucinotta. Benzazepinone calcium channel blockers. 2. Structure-activity and drug metabolism studies leading to potent antihypertensive agents. Comparison with benzothiazepinones. *J. Med. Chem*. 1992, **35**, 756-772.

47 S. D. Kimball, D. M. Floyd, J. Das, J. T. Hunt, J. Krapcho, G. Rovnyak, K. J. Duff, V. G. Lee, R. V. Moquin, C. F. Turk, S. A. Hedberg, S. Moreland, R. J. Brittain, D. M. McMullen, D. E. Normandin and G. G. Cucinotta. Benzazepinone calcium channel blockers. 4. Structure-activity overview and intracellular binding site. *J. Med. Chem*. 1992, **35**, 780-793.

48 H. J. Hall. Correlation of the base strengths of amines. *J. Am. Chem. Soc*. 1957, **79**, 5441-5444.

49 P. W. Kenny and C. A. Montanari. Inflation of Correlation in the Pursuit of Drug-likeness. *J. Comput. Aided Mol. Des*. 2013, **27**, 1-13.

Graphical Abstract

The impact of N- and O-methylation on aqueous solubility and measured lipophilicity for several chemically diverse structural classes is described.