Contents lists available at ScienceDirect



Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl



Investigating the effects of the core nitrogen atom configuration on the thermodynamic solubility of 6,5-bicyclic heterocycles

Brett Cosgrove ^{a,b,*}, Kenneth Down ^a, Sophie Bertrand ^a, Nicholas C.O. Tomkinson ^b, Michael D. Barker ^a

^a Medicinal Chemistry, GlaxoSmithKline Medicines Research Centre, Stevenage, UK

^b Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow, UK

ARTICLE INFO

Keywords: ChromlogD_{7.4} Fasted State Simulated Intestinal fluid (FaSSIF) General Solubility Equation (GSE) Kinetic solubility Physicochemical properties Simulated lung fluid (SLF) Thermodynamic solubility

ABSTRACT

Physicochemical properties, such as solubility, are important when prioritising compounds for progression on a drug discovery project. There is limited literature around the systematic effects of core changes on thermodynamic solubility. This work details the synthesis of nitrogen containing 6,5-bicyclic heterocyclic cores which are common scaffolds in medicinal chemistry and the analysis of their physicochemical properties, particularly, thermodynamic solubility. Crystalline solids were obtained where possible to enable a robust comparison of the thermodynamic solubility. Other parameters such as pK_a , melting point and lipophilicity were also measured to determine the key factors affecting the observed solubility.

Introduction

Physicochemical and molecular properties such as solubility, permeability and lipophilicity are important parameters in drug discovery. They can influence the pharmacokinetic, pharmacodynamic and safety profile of a potential drug.¹ There is often a balancing act between these various parameters to find candidates with the appropriate profile for the desired target.^{2–4}

To introduce novelty and overcome physicochemical barriers, the medicinal chemist often explores alternative ring systems, with an average of six new rings systems entering the drug market each year.⁵ 6,5-Bicyclic heterocycles in particular, are ubiquitous to drug discovery and in 2019, the FDA approved multiple drugs containing this type of motif (Fig. 1).⁶ Benzimidazoles, indazoles and purine type scaffolds shown in Fig. 1 are well-established substructures, found in both fragments and drug molecules.^{6,7} A less common core found in the Bruton's tyrosine kinase inhibitor Zanubrutinib was developed through a ringmerging approach of common kinase analogues. This led to a clinical candidate with excellent selectivity and a desirable pharmacokinetic profile.⁸ Therefore, understanding the physicochemical properties of 6,5-bicyclic heterocycles is important to the medicinal chemistry community.⁹

Solubility is a critical parameter in drug discovery.¹⁰ For oral dosing, poorly soluble compounds can have reduced gut absorption and variability in drug plasma concentration which can lead to attrition during

development.¹¹ It is important to measure solubility during lead optimisation, and to be able to have strategies for improving it.¹² Solubility is reported as either thermodynamic or kinetic; these refer to two different aspects of behavior that a molecule possesses when undergoing dissolution.¹³ Kinetic solubility refers to any solubility that is associated with a metastable form of a compound.¹⁴ Thermodynamic solubility is the maximum amount of the most stable form of the compound (preferably crystalline) that can remain in solution under equilibrium conditions.¹⁵ In the literature, kinetic solubility is widely reported and accepted due to the fact that quantification can be automated effectively.¹⁶

Within the literature, methods of improving the solubility of a compound are limited and the data presented frequently does not report full information with regards to the effect of a change on each property of the molecule. A typical medicinal chemistry approach to enhancing the solubility of a compound is to introduce groups such as a basic amine or a carboxylic acid group. These simple modifications can affect other parameters.^{17,18} For instance, introducing a basic amine can lead to an increase in the volume of distribution which can lead to accumulation in certain tissues of the body,¹⁹ or even introduce other challenges such as hERG inhibition.²⁰ Adding a carboxylate group is known to be detrimental to the permeability of the molecule.²¹ Alternative approaches for improving solubility have been reported and include: increasing the sp³ character,²² distorting the crystal lattice structure²³⁻²⁴ and decreasing the aromatic ring count.²⁵ However, these approaches can also

https://doi.org/10.1016/j.bmcl.2020.127752

Received 23 September 2020; Received in revised form 11 December 2020; Accepted 13 December 2020 Available online 29 December 2020 0960-894X/© 2020 Elsevier Ltd. All rights reserved.

^{*} Corresponding author.



Fig. 1. Examples of FDA drug approvals in 2019 containing 6,5-bicyclic heterocycles.

compromise other properties of the molecule in lead optimisation. Within this work, an alternative strategy was applied to investigate the physicochemical properties of a family of bicyclic heterocycles containing a bridgehead nitrogen atom.

A range of 6,5-bicyclic heterocycles varying only in their nitrogen atom configuration were synthesised, with the substituents held constant to determine how the nitrogen atoms within the core bicyclic ring affected the properties of the molecule. The aim was to identify whether these changes would influence the solubility without introducing other liabilities into the molecules. This approach might offer alternative templates with better physicochemical properties to build a molecule around and provide a greater chance in meeting candidate quality criteria. Analysis of the physicochemical properties on crystalline batches was carried out to give a robust comparison of the thermodynamic solubility data of the derivatives prepared. Plane polarised light microscopy was initially used to detect signs of crystallinity for all final compounds (1a-5c) and X-ray powder diffraction (XRPD) was used to confirm the crystalline material. Additional properties such as pK_a , melting point and chromatographic logD were also measured to understand how subtle core changes influenced physicochemical properties.

To enable the direct comparison of the effect of altering the five cores (Fig. 2) on the physicochemical profile of the molecule, the pendant groups were kept constant. A 3,4-dimethoxybenzene substituent was incorporated in all compounds prepared. The pendant amine R' groups selected consist of typical medicinal chemistry functional groups: an aromatic ring (a), a heteroaromatic ring (b) and a fully saturated ring (c). These groups provided a range of functionality on each core to



determine how the properties were affected.

Cores 1, 4 and **5** were synthesised according to literature methods (see supplementary information for experimental details).^{26–29} **Cores 2** and **3** required synthetic development (Scheme 1 and Scheme 2).

The first step of the synthesis for compounds **2a-c** (Scheme 1) involved refluxing **2.1** with aqueous sodium hydroxide to give the tricarbonyl product (**2.2**) in good yield (62%). Compound **2.2** was heated with hydrazinecarboxamide under basic aqueous conditions to form the bicyclic heterocycle in a single step (**2.3**, 79%). The pyrimidone species (**2.3**) was subsequently chlorinated using POCl₃ then iodinated with NIS to give the key bis-halogenated intermediate (**2.5**) in good yield (62% over two steps). Compound **2.5** was further functionalised through an S_NAr reaction to give the products **2.6–2.8**. Finally, a Suzuki cross-coupling reaction with 3,4-dimethoxyphenyl boronic acid under standard reaction conditions gave compounds **2a-2c**, which were found to be crystalline after isolation.

The synthesis of compounds **3a-c** are shown in Scheme 2. Compound **3.1** and chloropropanone were heated together under forcing reaction conditions to form the cyclised product (**3.2**) in good yield (63%). It proved important to carry out the reaction under solvent free conditions to prevent the formation of multiple alkylation side products (see Supporting Information). The bicyclic heterocycle **3.2** was iodinated with NIS at 0 °C to give the key bis-halogenated compound **3.3** (79%). The penultimate step of the synthesis involved a series of S_NAr reactions to give compounds (**3.4–3.6**) in low to moderate yield (24–55%). Finally, a Suzuki cross-coupling reaction gave the compounds (**3a-3c**) as gummy solids. Multiple crystallisation techniques were performed on these products, however, only amorphous solids were obtained, as confirmed by plane polarised light microscopy and XRPD analysis.

Compounds (1a-5c) were profiled in three solubility assays (Table 1). The thermodynamic solubility was measured in two biorelevant media, Fasted State Simulated Intestinal Fluid (FaSSIF_{6.5}) and Simulated Lung Fluid (SLF_{6.9}).³⁰ Both biorelevant media mimic physiological conditions of the lungs and stomach. FaSSIF is buffered at pH 6.5 and contains 0.75 mM lecithin which comprises fatty lipids; SLF is buffered at pH 6.9 and contains both 0.75 mM lecithin and a protein, bovine serine albumin (BSA).³¹ The presence of the fatty lipids and protein influence the solubility of the compounds. For example, the lipids could potentially form micelles within the buffer and the BSA could interact with the compounds through non-covalent interactions. The kinetic solubility was measured from a 10 mM DMSO stock solution (Kinetic_{7 4}). The lipophilicity was measured using a chromatographic reverse phase method (**ChromlogD_{7 4}**) and calculated using clogP.³² The pK_a and melting points of all final compounds (1a-5c) were measured.

The individual compound solubility data shown in Table 1 was averaged for each core (1-5) and plotted in Fig. 3. The effects of the different amine monomers (a-c) are shown in Fig. 4.

All five cores had an average FaSSIF solubility that was greater than the average SLF solubility. The 0.4 log unit difference in the pH of the two media equates to a 2.5-*fold* difference. In the more acidic FaSSIF media, the equilibrium will be shifted towards the protonated state for a compound that contains a basic centre. An increased ionisation state will lead to an increase in aqueous solubility. The pK_a of each compound was measured to investigate this possibility further. However, it is important to consider the buffer composition and the effects it could have on the basicity of the compound. The presence of the lipids and BSA within the SLF buffer are likely to reduce the pK_a of the compound and in turn reduce solubility. Therefore, the measured pK_a values could be overestimated in the biorelevant buffers for measuring the thermodynamic solubility.

Core 2, a pyrazolopyrimidine motif, demonstrated the lowest average solubility of all five cores across all three solubility assays. The greatest average ChromLogD_{7.4} (5.6) was observed for this core, the high lipophilicity being associated with a low solubility. Introducing polarity into a molecule is often used to reduce the lipophilicity.³³ Interestingly,



Scheme 1. Synthetic route for the preparation of Core 2 compounds (2a-c). (a) NaOH, EtOH, 2 h, 100 °C, 62% (b) hydrazinecarboxamide, Na₂CO₃, H₂O, 1 h, 100 °C, 79% (c) PhNMe₂, POCl₃, 16 h, 110 °C, 69% (d) NIS, CH₂Cl₂, 1 h, 0 °C, 54% (e) R⁻–NH₂, NEt₃, 1,4-dioxane, 5 h, 120 °C, 35–92% (f) 3,4-dimethoxyphenyl boronic acid, XPhosPdG2 (10 mol%), Cs₂CO₃, 1,4-dioxane/H₂O (4:1), 5 h, 130 °C, 22–74%.



Scheme 2. Synthesis of Core 3 analogues (3a-3c). (a) Chloropropanone, 2.5 h, 130 °C, 63% (b) NIS, CH₂Cl₂, 4 h, 0 °C, 79% (c) R-NH₂, NEt₃, 1,4-dioxane, 5 h, 130 °C, 24–55% (d) 3,4-dimethoxyphenyl boronic acid, PdCl₂(dppf)CH₂Cl₂ adduct (20 mol%), Cs₂CO₃, 1,4-dioxane/H₂O (4:1), 5 h, 130 °C, 33–87%.

shifting the bridgehead nitrogen atom one position lowers the Chrom-LogD to 4.5 and increases the pK_aH to 5.3 (**Core 3**, an imidazopyrazine template) which significantly improves the solubility of this series (Figure 3). This illustrates the significance of the nitrogen atom configuration in the core and how this subtle transformation alters the physicochemical properties of the molecule.

Core 1, containing an imidazopyridine bicycle, was the most basic core of the five examined ($pK_aH = 6.6$). This had the greatest average FaSSIF solubility of the four crystalline cores (116 µg/mL). Interestingly, the pK_a of around 6.6 falls between the pH values of the FaSSIF and SLF media. Changing from FaSSIF to SLF the solubility drops, probably due to a decrease in the proportion of ionised compound in the SLF assay medium.

Another trend observed was the average kinetic solubility (measured from a 10 mM DMSO solution) was higher than the SLF solubility (measured from solid) for all of the crystalline compounds examined.

Medicinal chemists frequently use kinetic solubility to drive SAR as this can be established by a higher throughput assay. The data obtained within this study demonstrates a considerable drop in the thermodynamic solubility when compared to the kinetic solubility *e.g.* see Table 1 kinetic vs SLF solubility. **Core 3** was amorphous, and it did not follow the same trends in solubility as the other four cores. Therefore, it is important to develop a stable crystalline polymorph of a compound as early as possible to obtain a realistic measurement of the thermody-namic solubility.

Core 4, a pyrazolopyrimidine core, and **Core 5,** a pyrazolotriazine core, both displayed high kinetic solubilities. A significant drop in the thermodynamic solubility was observed for both cores in comparison to the kinetic solubility, emphasising the need to monitor thermodynamic, in addition to kinetic solubility.

Each core had been appended with an aromatic, a heteroaromatic and a saturated group. This provided a range of functionality to

Table 1

Physicochemical profile of compounds (1a-5c).

	Cpd	Kinetic	FaSSIF	SLF	pKa	ChromlogD _{7.4}	Мр	logS	clogP
HN' ^{R'}	1 ^a	29	6	2	6.6	4.3	203-205	-5.2	3.9
N.	1 ^b	86	109	33	6.5/4.7 ^c	4.9	85-88	-4.0	3.9
, N →	1 ^c	64	233	29	6.7	6	90–94	-4.4	4.2
R HN ^{.R'}	2^{a}	19	5	<6	3.8	4.8	194–198	-4.2	3
N AN-N	2^{b}	35	31	6	2.8/4.9 ^c	5.5	101-103	-3.3	3
жж Х	2 ^c	44	13	3	4.4	6.4	104–107	-3.6	3.3
R HN ^{.R'}	3 ^a	53	32^{b}	9 ^b	5	3.9	191-193b	-4.5	3.3
N	3 ^b	67	335^{b}	58^{b}	5.3/4.0 ^c	4.5	128-131b	-3.8	3.3
N N	3°	99	7^{b}	$10^{\rm b}$	5.6	5.2	93-95 ^b	-3.8	3.6
R HN ^{.R'}	4 ^a	79	15 ^a	1^{a}	5.4	3.7	238-240	-4.6	3
N-N	4 ^b	>107	141 ^a	39 ^a	5.4/4.1 ^c	4.3	149–151	-3.8	3
, , , , , , , , , , , , , , , , , , ,	4 ^c	>180	25 ^a	18 ^a	5.7	5.2	157–159	-4.1	3.3
R HN ^{. R'}	5 ^a	19	32	10	2.9	3.9	167–169	-3.2	2.3
N N N	5 ^b	>150	36	36	2.5/4.7 ^c	4.5	149-152	-3.1	2.3
Ä _N Ä	5 ^c	>165	29	18	3.1	5.1	143–145	-3.3	2.6

The solubility experiments were performed in at least duplicate and on crystalline material unless otherwise noted. ^an = 1, ^b = amorphous material, ^c = pK_a of *meta*-pyridine measured. Melting point (Mp, °C), Kinetic solubility_{7.4} (μ g/mL), FaSSIF_{6.5} solubility (μ g/mL), SLF_{6.9} solubility (μ g/mL), and ChromlogD_{7.4}. clogP was predicted using GSK Biobyte/Daylight tool.



Fig. 3. Compounds (**1a-5c**) measured kinetic solubility (pH = 7.4) and thermodynamic (**FaSSIF**, pH = 6.5 & **SLF**, pH = 6.9) solubility. Data is displayed as an average for each **Core 1–5**, in numerical order.** **4b**, **4c**, **5b** and **5c** contain a greater than (>) modifier in the kinetic solubility measurement. * **2a** contains a less than (<) modifier in the SLF solubility measurement.



Fig. 4. The average solubility for cores containing each monomer (**a-c** attached to the core of the molecule). **** 4b, 4c, 5b** and **5c** contain a greater than (>) modifier in the kinetic solubility measurement. *** 2a** contains a less than (<) modifier in the SLF solubility measurement.

understand how each of the groups affected the solubility. The effect of each monomer (benzamide (**a**), *meta*-pyridine (**b**) and tetrahydropyran (**c**)) on the average solubility was plotted (Fig. 4).

The cores containing the pyridine monomer (**b**) show the highest thermodynamic solubility in both FaSSIF and SLF. This was probably due to the additional basic centre resulting in a higher proportion of protonated species, thus leading to a higher observed solubility. This observation is further supported by the fact that a higher proportion of compound would be protonated in the lower pH system (FaSSIF) which was reflected in the improved solubility for compounds containing this monomer.

Compounds appended with the benzamide monomer (a) displayed the poorest solubility across all three solubility assays. This was probably due to the intermolecular hydrogen bonding interactions between molecules. Also, the additional carboaromatic ring can support π - π stacking, with the amide group planar to the aromatic ring. This led to an enhanced crystal lattice energy and as shown in Table 1, compounds appended with the benzamide monomer showed the highest melting point. Furthermore, as represented in the General Solubility Equation (GSE), a higher melting point is associated with a lower solubility, consistent with the data presented in Table 1.^{34,35}

The GSE, also known as the Yalkowsky equation, is shown in **Eq. 1**. It is used to predict the aqueous solubility (**logS**) of a given compound if the melting point (Mp) and clogP of the compound are known.³⁴ The equation is a predictive tool; variables such as crystallinity and melting point can influence the accuracy of the GSE and so it should be viewed with caution; however, it provides a particularly useful platform from which to analyse data. The equation predicts that an increase of 100 °C in melting point would reduce the intrinsic solubility by a factor of 10 (logarithmic scale) which equates to a unit lower in **logS**. It should be noted that 85% of drugs have **logS** values between -1 and -5, -1 being highly polar molecules that may have low membrane permeability and -5 being hydrophobic molecules with low aqueous solubility.³⁶

$$logS = 0.5 - clogP - 0.01(Mp - 25)$$

Eq. 1: Yalkowsky (GSE) equation.

It is important to know the physical state of the compound, as an amorphous solid does not have a definitive melting point,³⁷ whereas, a stable crystalline polymorph displays a single sharp melt. Therefore, predicting the solubility (**logS**) based on the melting point of a crystalline solid should provide a robust and reliable method for the comparison of data across a series.

The data in Table 1 shows a clear correlation between the thermodynamic solubilities of crystalline compounds and their predicted solubility values (logS). Compound 1a (FaSSIF 6 µg/mL, logS -5.2) displays the lowest solubility for Core 1 analogues and it was ranked with the lowest solubility by the GSE equation. Compound 1b (FaSSIF 109 μ g/mL, logS -4.0) is around 18-fold more soluble in FaSSIF than compound 1a, which is mirrored by a 1.2 log unit increase in logS. This trend was consistent with the majority of the crystalline compounds where the logS value was reflective of the compounds in each core. The thermodynamic solubility values for the three analogues of Core 5 were very similar and this was reflected in the narrow range of logS for this series of compounds. The exception to the trend was for Core 3, where each of the compounds were amorphous, and the predicted solubilities of these compounds did not follow the measured thermodynamic solubilities. This can be explained by the amorphous form of the compounds and the uncertainty associated with the melting points determined. This again, highlights the importance of gathering thermodynamic solubility data on crystalline compounds.

In summary, five alternative 6,5-bicyclic nitrogen containing heterocycles were synthesised and crystalline batches of compounds were obtained for 4 of the 5 cores. The physicochemical properties of compounds (1a-5c) were profiled and the solubility was found to vary significantly due to these subtle changes in nitrogen atom position. The intrinsic properties of the nitrogen atom configuration and substituents were shown to influence the solubility considerably through modifications to lipophilicity, basicity and crystallinity. For each of the crystalline cores, the analogue with the least solubility was generally the compound with the highest melting point and lowest logS. The data gathered showed that the kinetic solubility was often higher compared to the thermodynamic solubility. The measured thermodynamic solubility was dependent on the biorelevant media (FaSSIF_{6.5} vs SLF_{6.9}), as the solubility in SLF, was in all cases, lower than FaSSIF. This focus on the thermodynamic solubility of the core of crystalline compounds at an early stage in a drug discovery programme could be a useful strategy to adopt to ensure physicochemical properties are optimised in parallel with potency and selectivity for the desired target.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are thankful to Iain Reid and Shenaz Bunally for

generating the physicochemical data and advice on the assays; and Kate Dennis for XRPD analysis. We also thank the EPSRC for funding via Prosperity Partnership EP/S035990/1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi. org/10.1016/j.bmcl.2020.127752.

References

- [1] Nakashima S, Yamamoto K, Arai Y, Ikeda Y. Chem Pharm Bull. 2013;61:1228.
- [2] Morgan P, Van Der Graaf PH, Arrowsmith J, et al. Drug Discov Today. 2012;17.
- [3] Saal C, Petereit AC. Eur J Pharm Sci. 2012;47:589.
- [4] Volpe DA. Future Med Chem. 2011;3:2063. [5] Taylor RD, MacCoss M, Lawson AD. J Med Chem. 2014;57:5845.
- [6] https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-andnew-therapeutic-biological-products/novel-drug-approvals-2019 (Accessed 13.01.2020).
- [7] Giordanetto F, Jin C, Willmore L, Feher M, Shaw D. E. J Med Chem. 2019;62:3381.
- [8] Guo Y, Liu Y, Hu N, et al. J Med Chem. 2019;62:7923.
- [9] Wenlock MC, Austin RP, Barton P, Davis AM, Leeson PD. J Med Chem. 2003;46: 1250.
- [10] Valko K. Physicochemical and biomimetic properties in drug discovery. John Wiley & Sons, Inc.; 2014.
- [11] Sugano K. Int J Pharm. 2010;405:79.
- [12] Ahmad NM. Bioorg Med Chem Lett. 2016;26:2975.
- [13] Walker MA. Bioorg Med Chem Lett. 2017;27:5100.
- [14] Covarrubias CH, Vilchis-Reyes MA, Yépez-Mulia L, et al. Eur J Med Chem. 2012;52: 193.
- [15] Valko K, Butler J, Eddershaw P. Expert Opin Drug Discov. 2013;8:1225.
- [16] Hill AP, Young RJ. Drug Discov Today. 2010;15:648.
- [17] Cisneros JA, Robertson MJ, Mercado BQ, Jorgensen WL. ACS Med Chem Lett. 2017; 8:124.
- [18] Lassalas P, Gay B, Lasfargeas C, et al. J Med Chem. 2016;59:3183.
- [19] Charifson PS, Walters WP. J Med Chem. 2014;57:9701.
- Jamieson C, Moir EM, Rankovic Z, Wishart G. J Med Chem. 2006;49:5029. [20]
- [21] Thomae AV, Wunderli AH, Krämer SD. Biophys J. 1802;2005:89. [22]
- Lovering F, Bikker J, Humblet C. J Med Chem. 2009;52:6752.
- [23] Ishikawa M, Hashimoto Y. J Med Chem. 2011;54:1539.
- Wang H-L, Katon J, Balan C, et al. J Med Chem. 2007;50:3528. [24] [25] Ritchie TJ, Macdonald SJF. Drug Discov Today. 2009:14.
- [26] Catalano JG, Gaitonde V, Beesu M, et al. Tet Lett. 2015;56:6077.
- [27] Long J, Lee W-S, Chough C, Bae IH, Kim BM. J Org Chem. 2015;80:4716.
- [28] Wavhale RD, Martis EAF, Ambre PK, et al. Bioorg Med Chem. 2017;25:4835.
- [29] Mejdrova I, Chalupska D, Kogler M, et al. J Med Chem. 2015;58:3767.
- [30] Hughes JP, Rees S, Kalindjian SB, Philpott KL. Br J Pharmacol. 2011;162:1239.
- [31] Williams HD, Trevaskis LN, Charman AS, et al. Pharmacol Rev. 2013;65:315.
- [32] Liang C, Qiao JQ, Lian HZ. J Chromatogr A. 2017;1528:25.
- [33] Arnott JA, Planey SL. Exp Opin Drug Discov. 2012;10:863.
- [34] Pinal RJ. Org Biomol Chem. 2004;2:2692.
- [35] Ran Y, He Y, Yang G, Johnson JLH, Yalkowsky SH. Chemosphere. 2002;48:487.
- [36] Sanghvi T, Jain N, Yang G, Yalkowsky SH. QSAR Comb Sci. 2003;22:258.
- [37] Ran Y, Jain N, Yalkowsky SH. J Chem Inf Comput Sci. 2001;41:1208.