ELSEVIER

Getting physical in drug discovery: a contemporary perspective on solubility and hydrophobicity

Alan P. Hill[1] and Robert J. Young[2]

[1] Department of Analytical Chemistry, GlaxoSmithKline Medicines Research Centre, Stevenage, Hertfordshire SG1 2NY, UK
[2] Department of CSC Medicinal Chemistry, GlaxoSmithKline Medicines Research Centre, Stevenage, Hertfordshire SG1 2NY, UK

Suboptimal physical properties have been identified as a particular shortcoming of compounds in contemporary drug discovery, contributing to high attrition levels. An analysis of the relationship between hydrophobicity (calculated and measured) and ~100 k measured kinetic solubility values has been undertaken. In line with the General Solubility Equation, estimates of hydrophobicity, particularly ACD *c* log $D_{pH7.4}$, give a useful indication of the likely solubility classification of particular molecules. Taking ACD *c* log $D_{pH7.4}$ values together with the number of aromatic rings in a given molecule provides enhanced prediction. The 'Solubility Forecast Index' (SFI = *c* log $D_{pH7.4}$ + #Ar) is proposed as a simple, yet effective, guide to predicting solubility. Moreover, analysis of measured distribution/partition coefficient values highlighted statistically significant shortcomings in the applicability of octanol/water as a model system for hydrophobicity determination with poorly soluble compounds.

Introduction

The optimization of physical properties is fundamental to successful drug discovery [1]. Aqueous solubility is a desirable property to have in a drug molecule, facilitating delivery to and subsequent interaction with the pharmacological target [2,3]. The hydrophobicity (the commonly used etymological synonym of lipophilicity) of a compound is a measure of the preference of a compound to reside in lipid over an aqueous environment; this has an implied inverse link to aqueous solubility and is fundamental to many other interactions vital to achieving a potent and selective pharmacological action. The crux of successful drug discovery is almost invariably the achievement of a balance between hydrophobicity-driven potency and hydrophilicity-driven pharmacokinetic and/or pharmacodynamic action. An estimation of hydrophobicity is perhaps the key descriptor in the design of potential drug molecules [4]; it is fundamental to establishing structure–property relationships and to the many predictive models of pharmacokinetic parameters used in medicinal chemistry [5,6]. In spite of the availability of high-throughput methods for both the measurement and the prediction of hydrophobicity, there is growing evidence that clinical candidates are generally somewhat overly hydrophobic, which is implicated as an important factor in continuing high rates of attrition in the drug development process [7].

The established model for measuring hydrophobicity involves the partition or distribution of a compound between octanol and aqueous buffers. Thus, intrinsic hydrophobicity (log *P*, $\log_{10}$ of the partition coefficient) describes the partition of non-ionizable or unionized forms of molecules between octanol and buffer; log P is a constant for any given compound. Effective hydrophobicity (log D_{pH}, $\log_{10}$ of the distribution coefficient at a given pH) reflects the distribution of all species present between the phases for a given buffer pH (typically log $D_{pH7.4}$ is quoted, i.e. at physiological pH, 7.4). If a molecule has no ionizable centre then intrinsic and effective hydrophobicity are equivalent, regardless of pH. Of course, for ionizable compounds as pH varies, the relative proportion of species will be dependent on the pK_a of the basic and/or acid centres. There are many *in silico* packages that predict log *P* and, in conjunction with pK_a values, log *D* can additionally be estimated for any given pH. log *P* is the important factor in determining binding being implicated in the promiscuity of overly lipophilic molecules largely owing to entropy-driven non-specific binding [7], whereas log $D_{pH7.4}$ is implicated as a key descriptor in modelling of absorption and metabolism properties [4–6].

Corresponding authors:. Hill, A.P. (alan.p.hill@gsk.com), Young, R.J. (rob.j.young@gsk.com)

1359-6446/06/$ - see front matter © 2010 Elsevier Ltd. All rights reserved. doi:10.1016/j.drudis.2010.05.016

TABLE 1

Predicted log *S* (molar solubility) as a function of *c* log *P* and melting point combinations, calculated using the General Solubility Equation.

		log P						
		0	1	2	3	4	5	6
Mpt °C	50	0.25	-0.75	-1.75	-2.75	-3.75	-4.75	-5.75
	100	-0.25	-1.25	-2.25	-3.25	-4.25	-5.25	-6.25
	150	-0.75	-1.75	-2.75	-3.75	-4.75	-5.75	-6.75
	200	-1.25	-2.25	-3.25	-4.25	-5.25	-6.25	-7.25
	250	-1.75	-2.75	-3.75	-4.75	-5.75	-6.75	-7.75
	300	-2.25	-3.25	-4.25	-5.25	-6.25	-7.25	-8.25

Colouration by the aqueous solubility categories widely used within GSK (these classifications are used throughout).

Red, poorly soluble (<30 μM).

Yellow, intermediate solubility (30–200 μM).

Green, good solubility (>200 μM).

The precise interplay between solubility and other molecular properties has been at the hub of physical chemistry in drug discovery and open to debate. Indeed, recent literature [8–10] and the solubility challenge [11,12] have seen many diverse solubility models proposed, generally based on hydrophobicity in combination with other descriptors. The molecular descriptors of hydrophobicity, molecular weight and polar surface area are interrelated; thus, care needs to be taken in interpreting data sets in which each of these might be represented. The particular interrelationship between hydrophobicity and solubility is investigated in this review.

The General Solubility Equation (GSE; Eq. (1)) succinctly links solubility with hydrophobicity and melting point and has been demonstrated to be a reasonable model for predicting solubility for uncharged molecules [13].

Eq. (1) General Solubility Equation

$$\log S = -\log P - 0.01 * (M\,pt - 25) + 0.5 \quad (1)$$

The implications of this equation are illustrated in Table 1, in which log *S* values are calculated using the GSE for particular log *P* and melting point combinations pertinent to medicinal chemistry [13], coloured by solubility categories. It is clear that hydrophobicity is the dominant parameter in solubilities predicted by the GSE. For a medicinal chemist, it is easy to influence a change in log *P*, whereas the melting point is harder to predict or influence – and, indeed, in these days of high-throughput parallel chemistry, it is very rarely measured.

The colour distribution in Table 1 highlights how fine a log *P* divide there is between compounds with good and poor Solubility. It is probably not coincidental that the average drug has a *c* log *P* of 2.5 [7] – this corresponds to the upper limit of where the GSE predicts good solubility can be achieved. Likewise, poor solubility in potential drug molecules is engendered by the trend towards overly hydrophobic character recognized in contemporary medicinal chemistry [7]. Thus, it is pertinent to describe such compounds as insoluble grease balls rather than brick dust characterized by high melting lattices; this concurs with the findings of Bergström *et al.* [14] in a recent study of poorly soluble drugs. These values represent probable 'worst case' solubility because any ionizable centres would be likely to enhance solubility. The GSE can be modified to reflect the contribution of charge to solubility at any given pH by substituting log *D* for log *P*; this is exemplified for pH 7.4 in Eq. (2). Increased ionization reduces log *D* at a given pH, which consequently increases effective aqueous solubility

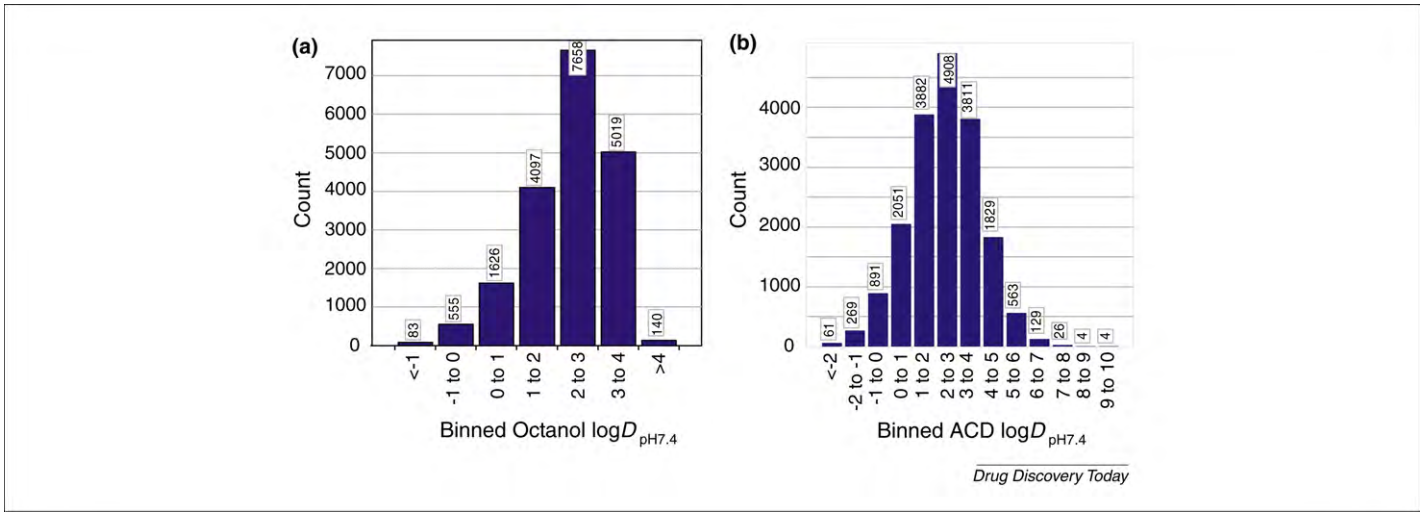

FIGURE 1

Distribution, by binning, of **(a)** measured and **(b)** calculated log $D_{pH7.4}$ values for compounds in this review.

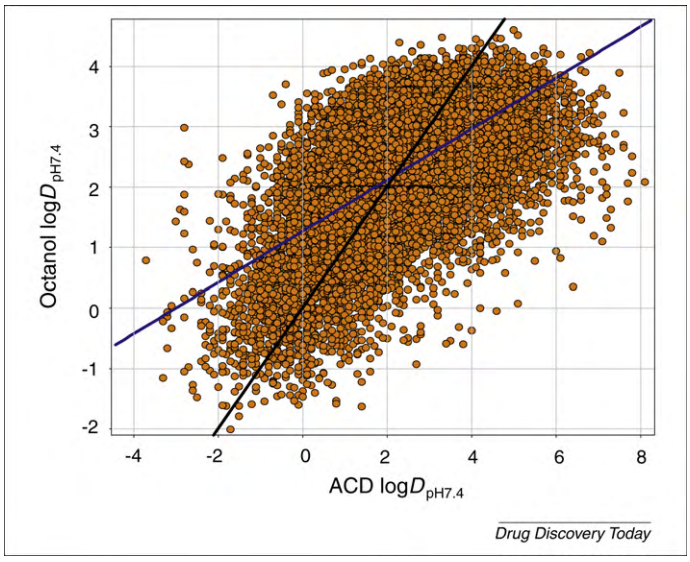

FIGURE 2

Plot of measured versus calculated log $D_{pH7.4}$ values. Line of unity, black; line of best fit, blue.

Eq. (2) GSE corrected for ionization at pH 7.4

$$\log S_{pH7.4} = -\log D_{pH7.4} - 0.01 * (M\,pt - 25) + 0.5 \quad (2)$$

To explore the relationship of hydrophobicity and solubility further, a large data set was compiled for analysis, comprising ~100 k compounds submitted for the measurement of kinetic aqueous solubility at pH 7.4. These represent a typical cross-section of experimental molecules at GSK over the past three years. The high-throughput method employed provides a good measure of the kinetic solubility of compounds and performs to a higher degree of precision than the three-box classification used in this review.

These solubilities were measured using an in-house method [15] utilizing quantification via chemi-luminescent nitrogen detection (CLND). This assay has a dynamic range between the lower detection limit of 1 μM and 500 μM, governed by the protocol's 1:20 dilution into pH 7.4 phosphate buffer solution from nominal 10 mM DMSO stock. Within this set, ~20 k compounds also had hydrophobicity values measured in the octanol–pH 7.4 buffer system using an in-house method [16]. Empirical data were supplemented with calculated values for hydrophobicity {log *P* (Daylight [17] and ACD [18] software), log $D_{pH7.4}$

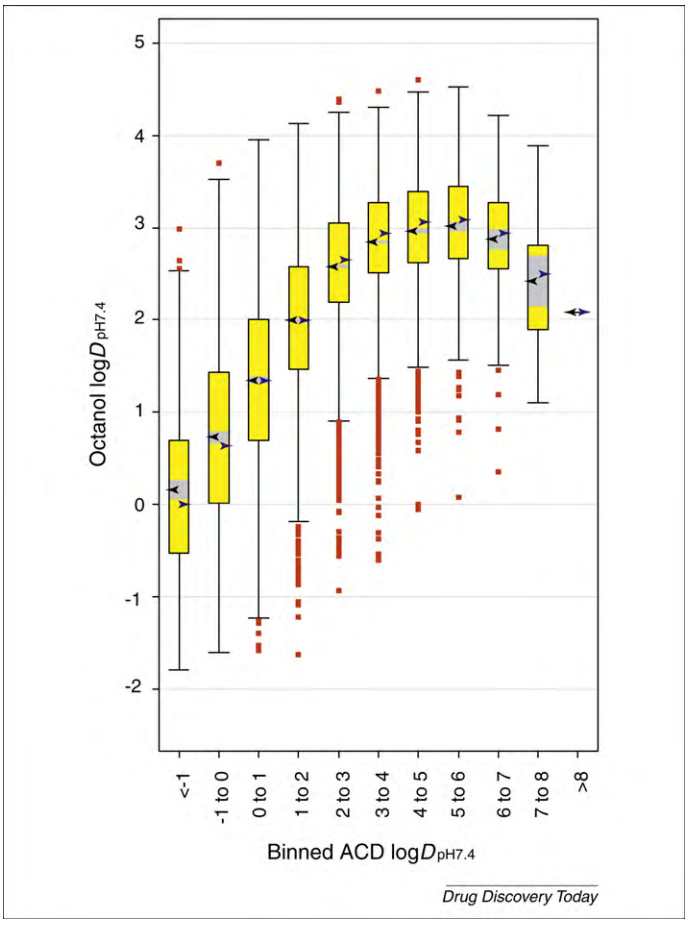

FIGURE 3

Box whisker plots comparing measured and calculated log $D_{pH7.4}$ values.

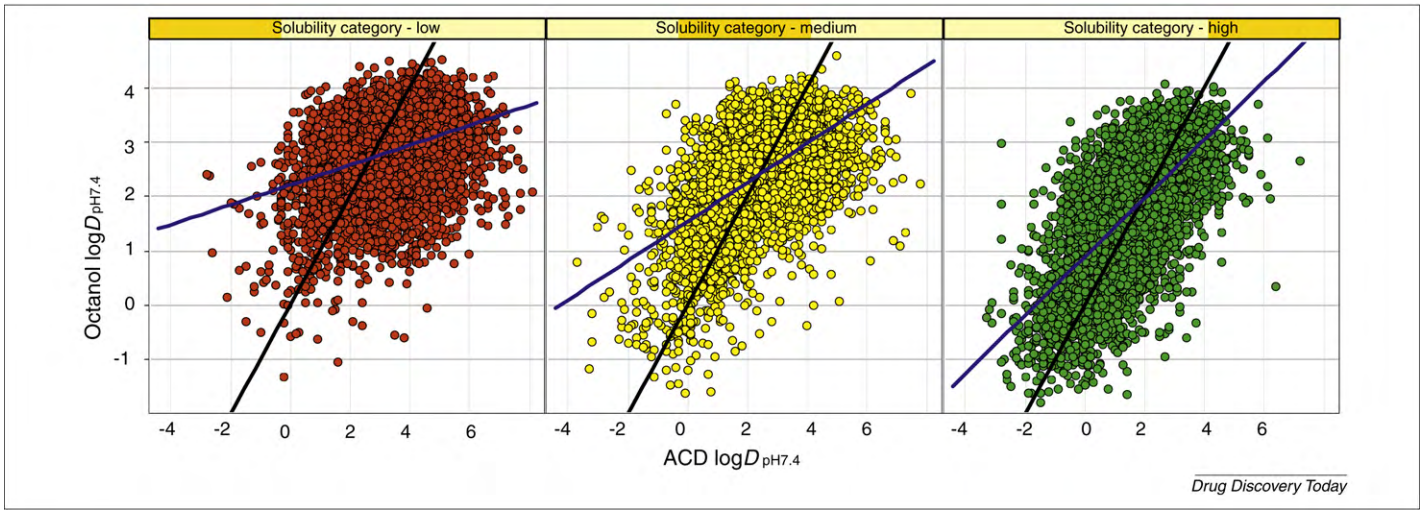

FIGURE 4

Plot of measured versus calculated log $D_{pH7.4}$ values, separated by solubility category. **(a)** Low solubility compounds (<30 μM). **(b)** Medium solubility compounds (30–200 μM). **(c)** High solubility compounds (>200 μM).

(ACD)}; molecular weight and aromatic ring count [17] were also included.

Presentation of data in the review

The outcomes of the data analyses are represented graphically in several forms. Data plots with lines of best fit and unity gave a representation of the data, albeit with a statistical analysis, which did not adequately convey the distribution of data because of the large numbers. The distribution of values was better conveyed through normalized bar graphs and box plots using binned hydrophobicity and/or solubility values, which better represent the distribution of data in a more visually amenable manner. In addition, to visualize three variables, binned multiple pie chart categorizations were employed. Solubility categories are consistently represented as described in Table 1, and numbers above bars and pies represent the number of data points in each bin.

Hydrophobicity: measured versus calculated

A disconnect between measured and calculated hydrophobicity values was immediately apparent in this review. In particular, for measured log $D_{pH7.4}$ a small dynamic range was apparent, and there was a clear plateauing of data for compounds predicted to be more hydrophobic. Whereas calculated values indicated a range of normally distributed values between −2 and +7 (Figure 1b), measured values for the same compounds were in a narrower range

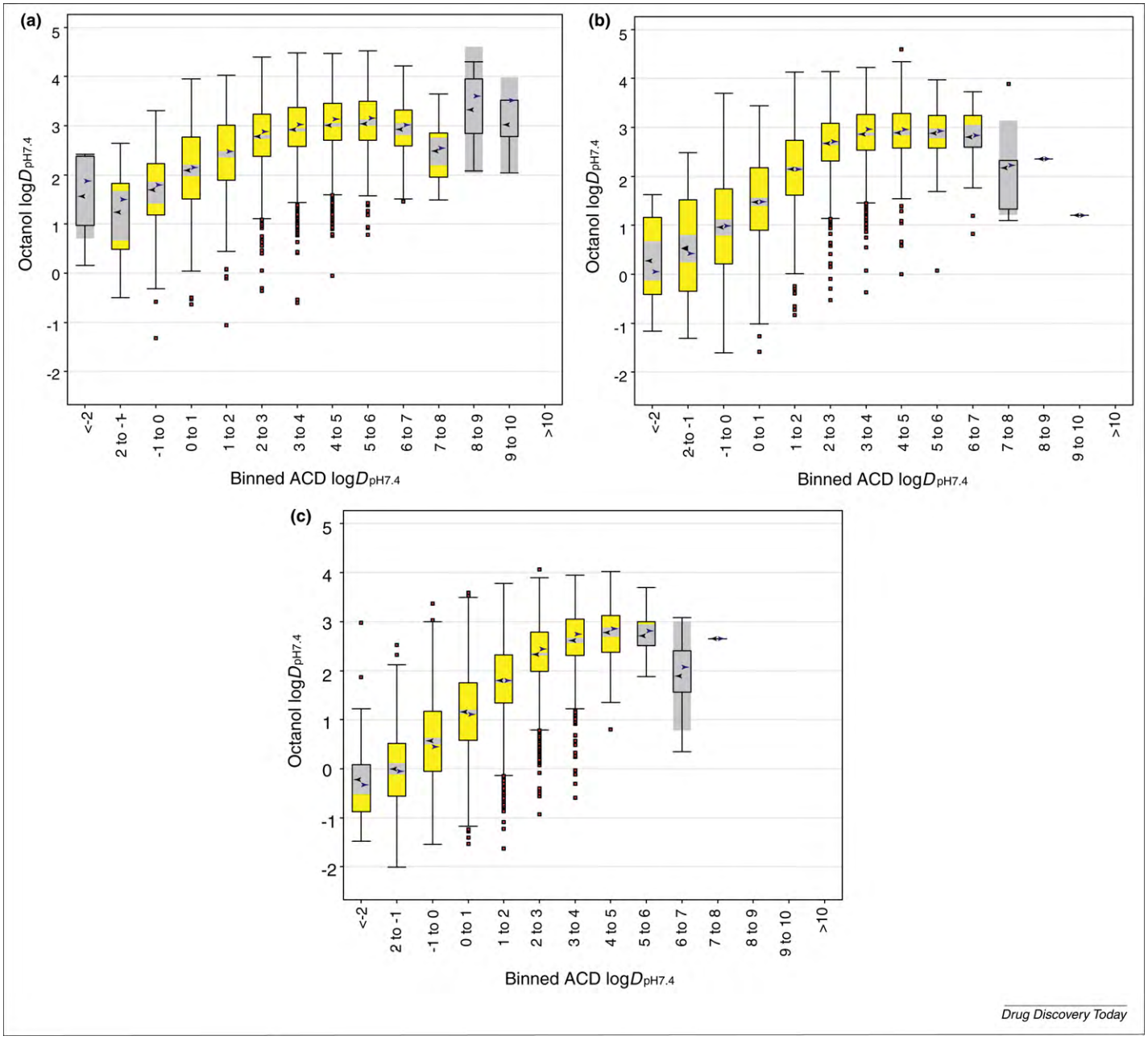

FIGURE 5

Box whisker plots comparing measured and calculated log $D_{pH7.4}$ values separated by solubility category. **(a)** Low solubility (<30 μM). **(b)** Medium solubility (30–200 μM). **(c)** High solubility (>200 μM).

between −1 and +4, with a skewed distribution (Figure 1a). This poor correlation is also evident in Figure 2, which demonstrates the disconnect between the measured and calculated hydrophobicity, particularly for compounds with high calculated hydrophobicity. The relative crossover between the lines of unity and best fit clearly support this notion. This is further reinforced in the box plot of measured versus binned predicted values (Figure 3). Starting from the most hydrophilic compounds, there is initially a statistically significant rise with increasingly hydrophobic bins; however, above *c* log $D_{pH7.4}$ values of 3–4, there is a levelling off and then a decrease.

The physical basis behind this discrepancy in measured and calculated hydrophobicity became apparent when the solubility of the compounds was considered. Poorly soluble (<30 μM) compounds showed a very poor correlation between predicted and measured log $D_{pH7.4}$, as highlighted by the line of best fit (Figure 4a; $R^2 = 0.11$, slope = 0.18) compared with the line of unity. With increasing solubility, however, the correlation clearly improves; for intermediate solubility compounds (30–200 μM), there was a marked improvement in the correlation between measured and calculated values (Figure 4b; $R^2 = 0.32$, slope = 0.36). Although still statistically poor, the best correlation (Figure 4c; $R^2 = 0.46$, slope = 0.54) was obtained for highly soluble compounds (>200 μM). These trends are further exemplified in the solubility-categorized box plots of Figure 5. Indeed, for poorly soluble compounds (Figure 5a), if log $D_{pH7.4}$ is predicted between 3 and 7, there is no statistical difference between the average measured values in each bin. These observations suggest that factors contributing to poor aqueous solubility are somehow perturbing the octanol/water system and preventing a true equilibrium from being obtained. A possible explanation for this is the tendency of poorly soluble compounds to form aggregates; this has caused issues in other systems [19].

These data strongly suggest that any measured log *D* value should be used only in the context of the solubility of the molecule in question. Paradoxically, they also indicate that for poorly soluble compounds, a calculated log *D* (or log *P*) is probably a better indicator of true hydrophobicity than a measured value. Historically, predictive models of hydrophobicity have been built on databases of measured values, so one could surmise that better

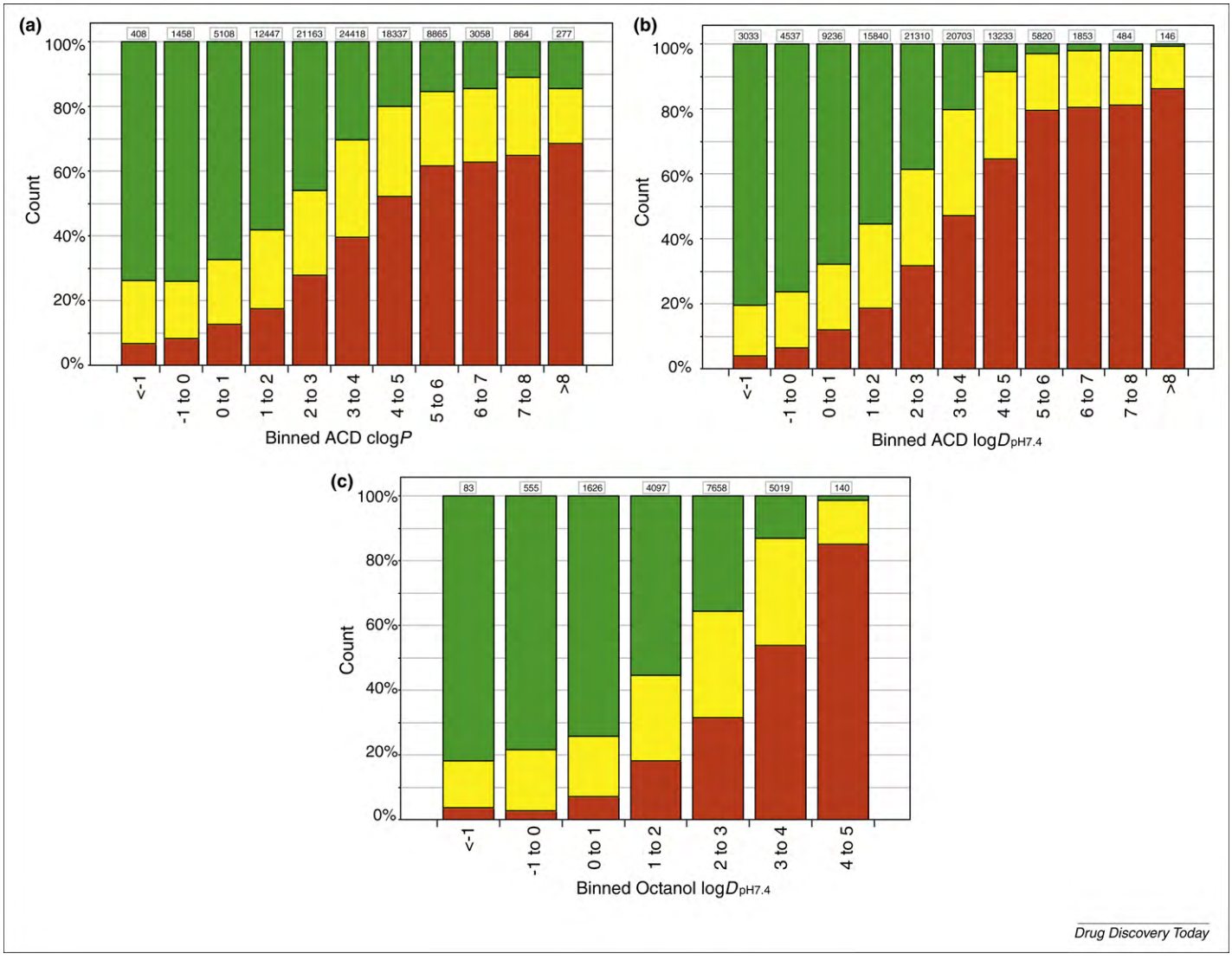

FIGURE 6

Bar charts showing distribution of solubility category as a function of binned hydrophobicity. **(a)** ACD log *P*. **(b)** ACD log $D_{pH7.4}$. **(c)** Measured log $D_{pH7.4}$.

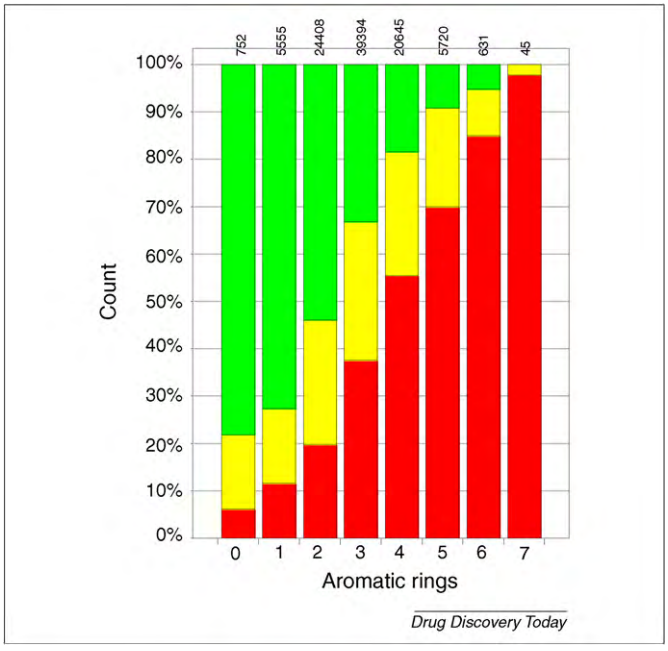

FIGURE 7

Distribution of measured kinetic solubility categories, binned by the number of aromatic rings in each structure.

models would be derived if only data from more soluble compounds were incorporated. The octanol–aqueous buffer system has been the hydrophobicity gold standard for many years, but it has clear shortcomings in assessing compounds with lesser solubility (which, as the following sections demonstrate, will be likely to have increased hydrophobicity).

Hydrophobicity and measured solubility

The distributions of measured solubility categories as a function of various hydrophobicity descriptors are displayed in Figure 6. These clearly display the expected trends of decreasing solubility with increasing hydrophobicity as predicted by the GSE, as represented in Table 1. The clearer stepped differentiation within the bands is apparent when log $D_{pH7.4}$ rather than log P is used, which reflects the conisderable contribution of ionization to solubility.

Looking at solubility classifications by means of measured log $D_{pH7.4}$ bins (Figure 6c), a similar trend is observed to that with calculated values; however, the discrimination between hydrophobicity bins is reflective of the limitations of the octanol/water system as described above. In particular, the limited population of the most hydrophobic bin should be noted.

Solubility and aromaticity: solubility Forecast Index

In recent publications, the negative impact of aromaticity on solubility has been highlighted. The effect has been quantified using three related descriptors: proportion of sp^3 atoms [20], aromatic proportion [21] and, more simply, the number of aromatic rings in a structure [22]. This impact has been rationalized by complementary explanations, invoking either an increase in lattice energy (and thus higher melting point) owing to the π-stacking of the flat rings or the reduced entropic contribution to the free energy of solvation and melting owing to the inherently more rigid molecules [23,24].

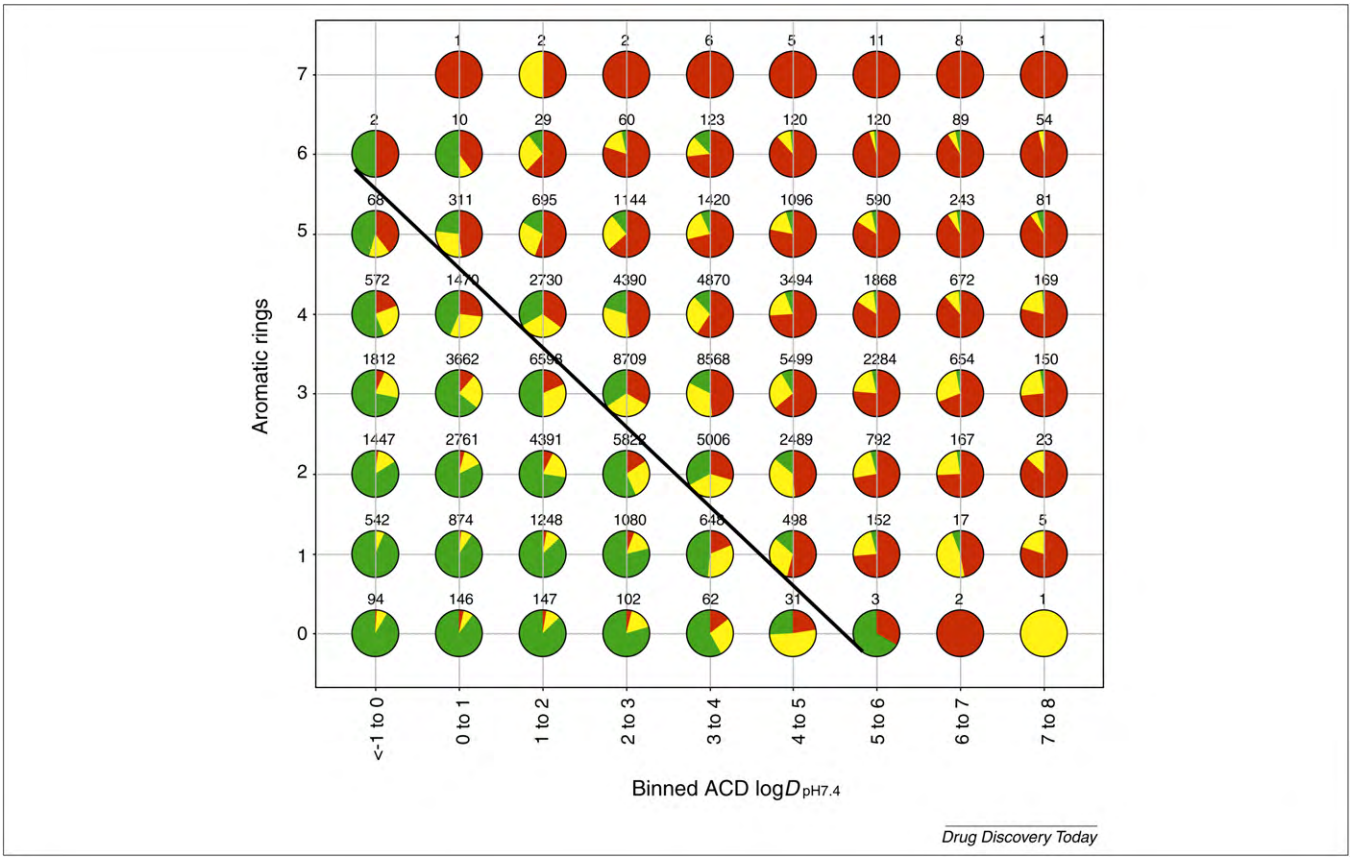

FIGURE 8

Pie chart matrix representation of solubility category as a function of ACD log $D_{pH7.4}$ and aromatic ring count.

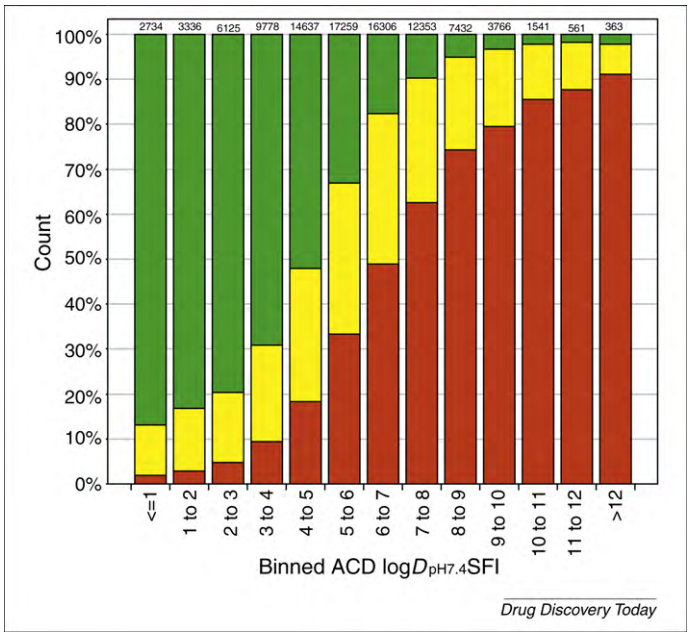

FIGURE 9

Distribution of solubility category as function of ACD log $D_{pH7.4}$ SFI.

As demonstrated in the review by Ritchie and MacDonald [22], several aromatic rings were associated with reduced solubility; this trend was very clear in this analysis, too (Figure 7). The finding that binned c log $D_{pH7.4}$ showed enhanced resolution of solubility classes (Figure 6) suggested a logical extension. Thus, the pie plot with binned *c* log $D_{pH7.4}$ and #Ar (Figure 8) clearly shows a more pronounced solubility differentiation than using c log *P* and #Ar, as employed in the earlier analysis [22]. In any given row or column of Figure 8, the proportion of poorly soluble compounds increases as either hydrophobicity or #Ar increases. This compounding effect is further emphasized by the diagonal split between regions of predominantly high and low solubility in Figure 8. In fact, the division between these regions can reasonably be described by the diagonal of *c* log $D_{pH7.4}$ + #Ar = 5. This observation led to the formulation of the simplistic Solubility Forecast Index, which supports the notion that if *c* log $D_{pH7.4}$ + #Ar < 5 then there is a reasonable chance of having good solubility.

Alternatively, the absolute SFI value can be related to a particular bin, which indicates the probability of having poor, intermediate or good solubility (Figure 9). This graded bar graph (Figure 9) can be compared with that shown in Figure 6b to show an increase in resolution when considering binned SFI versus binned *c* log $D_{pH7.4}$ alone. It is thus implicit that each aromatic ring in a molecule has a solubility penalty equivalent to an extra log unit of hydrophobicity in addition to its intrinsic hydrophobicity value [and this is the Daylight definition of #Ar, in which each aromatic ring (benzenoid or heterocyclic) is counted and fused systems (e.g. indole, naphthalene and purine) count as two rings].

In conclusion, the interrelationship of hydrophobicity and solubility has been investigated with ~100 k compounds with measured CLND solubility and ~20 k compounds with measured octanol–pH 7.4 buffer hydrophobicity values. Most notably, it has been demonstrated that the octanol–aqueous buffer model for measuring hydrophobicity does not work well for poorly soluble compounds. The octanol–buffer system works well for more hydrophilic compounds with fewer aromatic rings, but this clearly has shortcomings for more lipophilic and aromatic compounds, which, implicitly, have lesser solubility. It is noteworthy to consider that for molecules with hydrophobicity in proximity to the median values for marketed drugs (*c* log *P* = 2.5 or *c* log $D_{pH7.4}$ = 0.8) [7], these analyses strongly suggest that they are likely to be soluble and have reliable measured hydrophobicity values. Furthermore, the effect of aromatic rings on solubility was over and above their contributions to hydrophobicity, which led to simplistic concept of summing *c* log $D_{pH7.4}$ and the number of aromatic rings (Solubility Forecast Index = - *c* log $D_{pH7.4}$ + #Ar). The SFI value provides a simplistic and readily calculated indicator of the likely solubility category, implied by the distribution of measured values in that particular bin; conceptually, keeping SFI < 5 should be a powerful guide for compound design in drug discovery, which gives a high probability of securing good physical properties. Given that the average number of aromatic rings in oral drugs is 1.6 [22], their average SFI would thus be 2.4. Finally, it is clear that solubility is profoundly influenced by the hydrophobicity of a molecule, thus the limits of the GSE or any other solubility predictor must be governed by the quality of hydrophobicity estimates. As demonstrated here, the octanol–aqueous buffer model for hydrophobicity might not always be reliable.

Acknowledgements

We thank Pat McDonough for the experimental determination of octanol distribution coefficients and, in conjunction with Iain Reid, Silvia Bardoni, Patrick Day and Musundi Wabuyele, the CLND solubility measurements; Simon MacDonald, Simon Readshaw and the reviewers for their constructive critical comments on this manuscript.

References

1 Van de Waterbeemd, H. (2009) Improving compound quality through *in vitro* and *in silico* physicochemical profiling. *Chem. Biodivers.* 6, 1760–1766
2 Stegemann, S. (2007) When poor solubility becomes an issue: from early stage to proof of concept. *Eur. J. Pharm. Sci.* 31, 249–261
3 Lipinski, C.A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* 44, 235–249
4 Gleeson, M.P. (2008) Generation of a set of simple, interpretable ADMET rules of thumb. *J. Med. Chem.* 51, 817–834
5 Muresan, S. and Sadowski, J. (2008) Properties guiding drug- and lead-likeness. In *Methods and Principles in Medicinal Chemistry*. Wiley InterScience pp. 439–461
6 Lipinski, C.A. *et al.* (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 46, 3–26
7 Leeson, P.D. and Springthorpe, B. (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 6, 881–890
8 Dearden, J.C. (2006) *In silico* prediction of aqueous solubility. *Exp. Opin. Drug Discov.* 1, 31–52
9 Delaney, J.S. (2005) Predicting aqueous solubility from structure. *Drug Discov. Today* 10, 289–295

10 Kramer, C. *et al.* (2009) A consistent dataset of kinetic solubilities for early-phase drug discovery. *ChemMedChem* 4, 1529–1536
11 Llinas, A. *et al.* (2008) Solubility challenge: can you predict solubility of 32 molecules using a database of 100 reliable measurements. *J. Chem. Inf. Model.* 48, 1289–1303
12 Hopfinger, A.J. *et al.* (2009) Findings of the challenge to predict aqueous solubility. *J. Chem. Inf. Model.* 49, 1–5
13 Jain, N. and Yalkowsky, S.H. (2001) Estimation of the aqueous solubility. I. Application to organic non-electrolytes. *J. Pharm. Sci.* 90, 234–252
14 Bergström, C.A. *et al.* (2007) Poorly soluble marketed drugs display solvation limited solubility. *J. Med. Chem.* 50, 5858–5862
15 GSK in-house kinetic solubility assay: 5 μl of 10 mM DMSO stock solution diluted to 100 μl with pH 7.4 phosphate buffered saline, equilibrated for 1 h at RT, filtered through Millipore Multiscreen$_{HTS}$-PCF filter plates (MSSL BPC). The eluent is quantified by suitably calibrated flow injection CLND.
16 GSK in-house hydrophobicity assay: 10 μl of 10 mM DMSO stock solution diluted to 750 μl with octanol saturated pH 7.4 phosphate buffer and 160 μl buffer saturated octanol in a 96 well deep well block. Block sealed and inverted for 3 sets of 50 inversions, then centrifuged at 300 g for 20 min. Both phases are then quantified using generic gradient UV-HPLC.
17 Daylight software v4.9. Daylight Chemical Information Systems, Inc. Aliso Viejo, CA, USA (http://www.daylight.com/).
18 Log D suite version11. Advanced Chemistry Development, Inc. Toronto, Ontario, Canada (http://www.acdlabs.com/).
19 Jadhav, A. *et al.* (2010) Quantitative analyses of aggregation, autofluorescence, and reactivity artefacts in a screen for inhibitors of a thiol protease. *J. Med. Chem.* 53, 37–51
20 Lovering, F. *et al.* (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756
21 Lamanna, C. *et al.* (2008) Straightforward recursive partitioning model for discarding insoluble compounds in the drug discovery process. *J. Med. Chem.* 51, 2891–2897
22 Ritchie, T.J. and Macdonald, S.J.F. (2009) The impact of aromatic ring count on compound developability – are too many aromatic rings a liability in drug design? *Drug Discov. Today* 14, 1011–1020
23 Sanghvi, H. *et al.* (2003) Estimation of aqueous solubility by the General Solubility Equation (GSE) the easy way. *QSAR Comb. Sci.* 22, 258–262
24 Delaney, J.S. (2004) ESOL: estimating aqueous solubility directly from molecular structure. *J. Chem. Inf. Comput. Sci.* 44, 1000–1005

Alan Hill is currently head of the Physico-chemical properties group at Stevenage and has more than 35 years experience of physical chemistry in GlaxoSmithKline legacy companies after joining Wellcome in 1974. He has contributed notable physical insight across many programmes, in particular developing a physico-chemical model of oral absorption used in the discovery of Zolmitriptan.

Rob Young has 20 years of broad medicinal chemistry experience across several therapeutic areas at Wellcome, GlaxoWellcome and GSK. He has expertise in the application of physical methods and property-based design in medicinal chemistry, having derived and utilised structure property relationships to great effect in anticoagulant programmes in particular.