Supporting Information

Presentation of a Structurally Diverse and Commercially Available Drug Data Set for Correlation and Benchmarking Studies

Christian Sköld,[†] Susanne Winiwarter,[‡] Johan Wernevik,[¤] Fredrik Bergström,[¤] Leif Engström,[‡] Ruth Allen,[§] Karl Box,[§] John Comer,[§] Jon Mole,[§] Anders Hallberg,[†] Hans Lennernäs,[#] Torbjörn Lundstedt,[†] Anna-Lena Ungell,[‡] and Anders Karlén[†]*

[†]Division of Organic Pharmaceutical Chemistry, Department of Medicinal Chemistry, BMC, Uppsala University, Sweden. [‡]Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry, AstraZeneca R&D Mölndal, Sweden. [¤]Lead Generation; DMPK & Physical Chemistry, AstraZeneca R&D Mölndal, Sweden. [§]Sirius Analytical Instruments Ltd., Riverside, East Sussex, United Kingdom. [#] Department of Pharmacy, BMC, Uppsala University, Sweden

[*]To whom correspondence should be addressed. Phone: +46-18-471-4293. Fax: +46-18-471-4474. E-mail: anders.karlen@orgfarm.uu.se.

Table of contents

Analytically difficult compounds.

Amiloride. Although it was easy to measure the pK_a and log$P$ of amiloride, it was not possible to measure its solubility by the pH-metric CheqSol method because the compound behaved in an unexpected way. The HCl salt is initially insoluble in aqueous 0.15 M KCl at low pH. However the compound starts to dissolve at high pH before re-precipitating. However, the compound easily re-dissolves upon returning to low pH. Meanwhile, the HCl salt is soluble in de-ionized water and precipitates at high pH. The erratic pH conditions for precipitation make it difficult to obtain a reliable result. During this study a result of 216 μg/mL was obtained for the solubility of amiloride by a shake-plate method with UV detection, but this result differed significantly from another reported value (100 μg/mL).[1]

Erythromycin. Erythromycin is water-soluble with a basic pK_a, and presents no particular problems for pH-metric pK_a and log$P$. However, this compound does not behave very well in CheqSol assays, perhaps because the salt solubility is close to the intrinsic solubility. Analysis

was attempted in cosolvent and also at low ionic strength but the assay weights were too low for precipitation to occur. The shake-plate method was also unsuccessful because of the poor UV characteristics above 250 nm.

Folic acid. With three acidic and one basic group and logP_{ACD} below -2, pK_a values of folic acid look easy to measure – but it took some time to decide on the best approach. The unionized form of folic acid is insoluble in water and many cosolvents including methanol and DMSO at pH-metric concentrations. However, three of the pK_a values are UV active although the remote carboxylic acid is not. The strategy therefore was to start the assays at high pH (where the compound has -3 charge and is soluble) and titrate towards insolubility. The pK_a at 7.90 (charge transition -3 to -2) was seen both by pH-metric and pH-UV methods. The pK_a at 4.47 (charge transition -2 to -1) was seen only by the pH-metric method. The sample precipitated below this pK_a at pH-metric concentrations. However, the lower pK_a values representing charge transition -1 to 0 and charge transition 0 to +1 could be measured by the pH-UV method, which works at lower concentrations. Given the low solubility of folic acid (2.3 μg/mL) it might be expected to precipitate even at pH-UV concentrations but as shown by the supersaturation ratio of 12.8, folic acid stays in supersaturated solution for long enough for pK_a values to be measured. Hence, a combination of UV and pH-metric methods was required for successful pK_a measurement of folic acid. All attempts to measure log*P* of folic acid failed. The sample is very insoluble in octanol, even at high ratios of octanol to water, preventing reliable measurement of log*P*. Titrations were started at high pH where the compound is soluble but precipitation was observed in all octanol-containing data sets. The best estimate of log*P* based on using data prior to precipitation is a log*P* close to zero. This is only an estimate and may be unreliable. Solubility was easy to measure, once the pK_a values were known.

Glipizide. Glipizide is water-insoluble when unionized, with very slow dissolution rate even at high pH where the compound is ionized. Given the low solubility of glipizide (1.4 μg/mL) it might be expected to precipitate even at pH-UV concentrations, but as shown by the supersaturation ratio of 42.5, glipizide stays in supersaturated solution for long enough for pK_a values to be measured by the Fast D-PAS method. In this method, a 50 μL aliquot of alkaline methanolic stock solution of glipizide is titrated within two minutes over a pH range 12 to 2 in an aqueous solution of a linear buffer mixture. This fast method allowed for sufficient UV data in aqueous solution to be collected before the sample could precipitate. The change in UV absorbance versus pH was adequate for good results. Note that glipizide precipitated in the slower, standard pH-UV assays. Glipizide is moderately lipophilic, but log*P* measurement was hampered by its poor solubility in octanol at the lowest ratios of octanol/water. At least 1 mL octanol is required at pH-metric concentrations and assays must be started at high pH to make sure the compound is fully dissolved at the start of the titration. Solubility was easy to measure, once the pK_a values were known.

Levothyroxine. Although it has three pK_a values, the unionized form of levothyroxine is water-insoluble, and is also insoluble in many cosolvents including methanol and DMSO at pH-metric concentrations. It also precipitates in standard pH-UV assays. The compound is only soluble at high pH. By using Fast D-PAS, it is possible to obtain the highest two pK_a values in aqueous conditions. The lowest pK_a was inaccessible by all techniques because the compound precipitates below the second pK_a and also because this carboxylic acid pK_a is UV inactive. Although quite lipophilic, levothyroxine is fairly insoluble in octanol, and log*P* titrations must be started at high

pH in order to ensure solubilisation as well as using fairly large volumes of octanol to avoid precipitation. Nevertheless, good quality partition data is obtained for this lipophilic zwitterion showing that species XH_2 is the most lipophilic. Solubility was easy to measure, once the pK_a values were known, and levothyroxine chases equilibrium between pH 9.5–10.5.

Meclizine. pK_a, log$P$ and solubility for meclizine were difficult to measure. Meclizine shows too little change in absorbance versus pH to measure pK_a values using the pH-UV technique, and the unionized form of meclizine is insoluble in water and many cosolvents including methanol, dioxane and DMSO at pH-metric concentrations. However, it was found that meclizine was adequately soluble in MDM-water mixtures. Even though the compound precipitates at high pH in MDM-water, it is possible to obtain a result for the higher pK_a and extrapolate to aqueous conditions provided low sample concentrations are used. However, this put the concentrations below that required for reliable measurement of the lowest pK_a, which is in the signal-to-noise region of the pH scale. The lower pK_a was subsequently obtained by manual refinement of CheqSol data (see below) in MDM-water mixtures. MDM (a mixture of methanol, dioxane and acetonitrile) was recently introduced as a solvent mixture for dissolving poorly water-soluble drug-like molecules prior to analysis.[2] Meclizine is very lipophilic, and shifted to the Scherrer p_oK_a limit[3] under all ratios of octanol/water when analysed in 0.15 M KCl. The strategy for pH-metric log$P$ was to perform the measurements in conditions close to 0.01 M KCl to reduce the impact of ion-pair partitioning. This prevented the lowest octanol ratios from shifting to the Scherrer limit. The lowest pK_a shifted off-scale at the highest ratios of octanol. Nevertheless, reliable results were obtained for this very lipophilic compound. Meclizine has an insoluble HCl salt. When CheqSol is used to measure intrinsic solubility, it is a strict requirement that the salt form must be soluble during the assays. This was achieved by running assays in MDM cosolvent using the p_sK_a values determined above for the highest ionization step. The compound is a non-chaser. A good linear extrapolation to aqueous conditions was obtained. A side benefit of the CheqSol MDM assays was the ability to determine the p_sK_a values for the lower pK_a (charge transition +1 to +2) because CheqSol assays are performed at sufficiently high concentration for the p_sK_a values to be measured. Overall, meclizine is a very difficult compound to accurately measure pK_a, log$P$ and solubility.

Terfenadine. The unionized form of terfenadine is water-insoluble, and its UV absorbance is weak because the ionizable group is remote from the chromophore. Terfenadine must therefore be analysed pH-metrically, but it is insoluble at normal pH-metric concentrations in DMSO and methanol. However, it ionizes close to the mid-point of the pH scale where pH measurement is at its most sensitive. pH-metric titrations could therefore be done using very low concentrations in >35wt% methanol, obtaining the pK_a by extrapolating to aqueous conditions. Terfenadine is very lipophilic, and almost shifts to the Scherrer p_oK_a limit unless low ratios of octanol/water are used. Hence, low concentrations are required when using 50 μL of octanol. Also, the p_oK_a values occur between pH 6-7 and there is a tendency to overshoot pH above pH 7.5 as there is no buffering capacity between pH 7.5-10. Nevertheless, reliable results were obtained for this lipophilic compound. Terfenadine has an insoluble HCl salt. Since CheqSol is used to obtain an intrinsic solubility it is a strict requirement that the salt form must be soluble during the assays. This was achieved by running assays in methanol cosolvent using the p_sK_a values determined above. The compound has higher kinetic solubility values than intrinsic values and is therefore a chaser. However, the solubility is so low that there is inadequate buffer capacity and therefore a buffer was introduced to stabilize the pH. The buffer must be a basic buffer to avoid salt precipitation. A

good linear extrapolation enabled the aqueous solubility result to be obtained. Overall, terfenadine is a very difficult compound to accurately measure pK_a, log$P$ and solubility.

References

(1) Bergstroem, C. A. S.; Strafford, M.; Lazorova, L.; Avdeef, A.; Luthman, K.; Artursson, P. Absorption Classification of Oral Drugs Based on Molecular Surface Properties. *J. Med. Chem.* **2003**, *46*, 558-570.

(2) Box, K.; Völgyi, G.; Ruiz, R.; Comer, J.; Takács-Novák, K.; Bosch, E.; Ràfols, C.; Rosés, M. Physicochemical properties of MDM cosolvent - a new universal cosolvent mixture for pKa determination of poorly soluble pharmaceutical compounds. *Submitted*.

(3) Avdeef, A. Assessment of distribution-pH profiles. In *Lipophilicity in Drug Action and Toxicology*; Pliska, V., Testa, B., van de Waterbeemd, H., Eds.; Wiley-VCH: Weinheim, 1996; Vol. 4, pp 109-139.