Log D Analysis using Dynamic approach

Ganesh K[1], Prashanth A[2], Esayas A[2], Shahul H[3], Wasihun A[2], Timothy K[2], Janarthanan K. [2,]*

1) Curtiss-Wright Avionics and Electronics, Dublin 14, Ireland.
2) Department of Bio-medical Engineering, Jimma Institute of Technology, Jimma University, Ethiopia
3) Department of life-sciences, B S Abdur Rahman University, Chennai, Tamilnadu, India

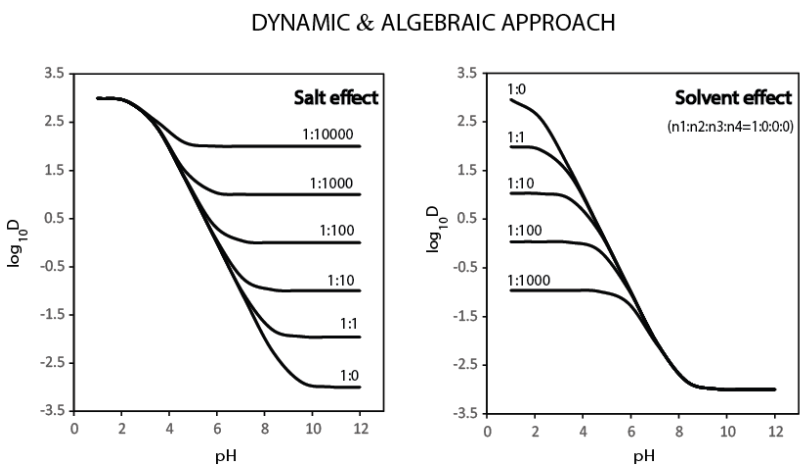
*Corresponding author: jana.jk2006@gmail.com

ABSTRACT

Log D is one of the important parameters used in Lipinski's rule to assess the druggability of a molecule in pharmaceutical formulations. It stands for the logarithm ($\log_{10}$) of the distribution coefficient (D) of a molecule partitioned between an aqueous phase (buffer solution) and a hydrophobic organic solvent phase (e.g. octanol). By definition, distribution coefficient is the ratio of the concentration of the sum of ionized and unionized species of a molecule distributed between the octanol phase and the aqueous buffer phase. Since the pH affects the ionization of a molecule, Log D value which is dependent on the concentrations of the ionized species also varies with pH. In this work, the conventional algebraic method is compared with a more generalized 'dynamic' approach to model the distribution coefficient of monoprotic, diprotic, monoalkaline, amphoteric, diamino-monoprotic and monoprotic acid in the presence of salt, co-solvent. Recently reported experimental log D data of amphoteric molecules such

as Nalidixic acid, Mebendazole, Benazepril and Telmisartan, were analyzed using both these approaches to show their equivalence.

GRAPHICAL ABSTRACT



KEYWORDS
Dynamic approach, log D analysis, monoprotic, diprotic, mono alkaline, amphoteric, salt effect, solvent effect, diamino-monoprotic molecules.

1. Introduction

Partition coefficient (P) is defined as the ratio of the concentration of a single species of a molecule, whether in ionized or unionized form, distributed between a hydrophobic phase and an aqueous phase [1,2,3,4,5]. As an example, consider, a weak monoprotic acid, HA, which can exist in two forms in an aqueous buffer phase, such as, unionized (HA_a) and ionized (A_a^-) form. If the weak monoprotic acid in an aqueous buffer phase, is equilibrated with a hydrophobic solvent (E.g. octanol), the unionized (HA_a) in the aqueous phase will be partitioned into hydrophobic phase (HA_o) with a partition

coefficient defined as, $P_{HA} = [HA_o]/[HA_a]$. Similarly, the partition coefficient for the ionized species can be written as $P_{A^-} = [A_o^-]/[A_a^-]$, where, $[A_o^-]$, $[A_a^-]$, are the concentrations of ionized species, A^-, in octanol and aqueous phase, respectively. It is less likely for a charged species like A^-, to get into an octanol phase, it forms a neutral ion pair with prevalently available cation in the aqueous solution before partitioning into hydrophobic phase. Compared to partition coefficient (P), the distribution coefficient (D) is defined as the ratio of the sum of the concentrations of both ionized and unionized species of a molecule, present in the hydrophobic organic phase to the aqueous buffer phase. The distribution coefficient of a weak monoprotic acid, is defined as, $D = ([HA_o] + [A_o^-])/([HA_a] + [A_a^-])$. Since the dissociation of the weak monoprotic acid in aqueous phase is dependent on pH value, the distribution coefficient also becomes dependent on pH. In an experiment designed to assess the lipophilicity of a molecule, the distribution coefficient (D), is measured at different pH conditions and the resultant profile of D, is fitted to a model, to obtain partition coefficients and pKa or pKb values of all species [1,2,3,4,5].

Using algebraic method the mathematical models to predict the profile of distribution coefficient of simple molecules such as monoprotic, diprotic, mono-alkaline and amphoteric can be easily derived [6]. On the other hand, when the effect of salt, solvent on the dissociation of monoprotic acid has to be taken in to account, dynamic approach could be relatively convenient than the algebraic method [3,5,7,8,9]. In this article, we explicitly, derive the dynamic models for monoprotic, diprotic, mono-alkaline, amphoteric, di-amino-monoprotic, and monoprotic in the presence of salt, co-solvent [7,8,9]. Log D profiles of recently reported amphoteric molecules such as Nalidixic acid, Mebendazole, Benazepril and Telmisartan, were analysed to show the equivalence of dynamic approach and algebraic method [10].

2. Theory

The algebraic and the dynamic expressions for amphoteric, monoprotic acid in the presence of salt (KCl), co-solvent (DMSO) are explicitly derived in this section. Additional cases such as monoprotic acid (SI.1), diprotic acid (SI.2), monoalkaline (SI.3), diamino-monoprotic amphoteric molecules (SI.7) are detailed in the supplementary information. At the outset, kinetic model that best represent the distribution of the molecule between an aqueous buffer and octanol layer will be outlined. Based on the kinetic model, the algebraic and the dynamic approaches based models will be derived.

2. 1. Amphotheric model for amino acids

2.1.1. Kinetic model for simple amino acids

Consider an amino acid ($NH_2 - R - COOH$ or HAB) containing a weak mono-protic acid ($COOH$ or HA) and a weak basic/alkaline group (NH_2 or B) distributed between an aqueous buffer and an organic hydrophobic solvent (octanol) (Figure 1A) [5,6,11]. In aqueous phase, the amino acid, $[NH_2 - R - COOH]$ or $[HAB]$, can exists in the unionized form $[NH_2 - R - COOH_a]$ or $[HAB_a]$, and ionized forms, $\left[NH_2 - R - COO^-{}_a\right]$ or $[{}^-AB_a]$, $[NH_3^+ - R - COOH_a]$ or $[HABH_a^+]$, $\left[NH_3^+ - R - COO^-{}_a\right]$ or $[{}^-ABH_a^+]$. The equilibrium among these four states can be written as (Eqns. 1-5),

$$HABH_a^+ \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} H_a^+ + HAB_a \qquad (1)$$

$$HABH_a^+ \underset{k_{-2}}{\overset{k_2}{\Leftrightarrow}} H_a^+ + {}^-ABH_a^+ \qquad (2)$$

$$HAB_a \underset{k_{-3}}{\overset{k_3}{\Leftrightarrow}} {}^{-}ABH_a^+ \tag{3}$$

$$HAB_a \underset{k_{-4}}{\overset{k_4}{\Leftrightarrow}} H_a^+ + {}^{-}AB_a \tag{4}$$

$${}^{-}ABH_a^+ \underset{k_{-5}}{\overset{k_5}{\Leftrightarrow}} H_a^+ + {}^{-}AB_a \tag{5}$$

k_1, k_{-1}; k_2, k_{-2}; k_4, k_{-4}; k_5, k_{-5}; are the forward and reverse kinetic rates for the dissociation of the species, such as $[HABH_a^+]$, $[HABH_a^+]$, $[HAB_a]$, $[{}^{-}ABH_a^+]$, into $[HAB_a]$, $[{}^{-}ABH_a^+]$, $[{}^{-}AB_a]$, $[{}^{-}AB_a]$, respectively. k_3, k_{-3} are the forward and reverse kinetic rates of the tautomerism seen between $[HAB_a]$ and $[{}^{-}ABH_a^+]$, respectively. H_a^+, is the proton dissociated from NH_3^+ or (BH^+), $COOH$ or (HA), functional groups. The four species, $[HABH_a^+]$, $[{}^{-}ABH_a^+]$, $[HAB_a]$ and $[{}^{-}AB_a]$, can be partitioned into octanol layer as given below (Eqns. 6 - 9),

$$HABH_a^+ \underset{k_{-p1}}{\overset{k_{p1}}{\Leftrightarrow}} HABH_o^+ \tag{6}$$

$$HAB_a \underset{k_{-p2}}{\overset{k_{p2}}{\Leftrightarrow}} HAB_o \tag{7}$$

$${}^{-}ABH_a^+ \underset{k_{-p3}}{\overset{k_{p3}}{\Leftrightarrow}} {}^{-}ABH_o^+ \tag{8}$$

$$^{-}AB_a \underset{k_{-p4}}{\overset{k_{p4}}{\Leftrightarrow}} {}^{-}AB_o \qquad (9)$$

k_{p1}, k_{-p1}; . k_{p2}, k_{-p2}; k_{p3}, k_{-p3}; . k_{p4}, k_{-p4}; are the forward, reverse kinetic rates, respectively, for the partitioning of $HABH_a^+$, HAB_a,$^{-}ABH_a^+$, $^{-}AB_a$, from the aqueous phase into octanol phase, as $HABH_o^+$,HAB_o,$^{-}ABH_o^+$,$^{-}AB_o$, respectively. The partition of singly charged species such as $[HABH_o^+]$, and $[^{-}AB_o]$ into the octanol layer can be largely influenced by the salt effect, which enhances ion pair formation of the charged species in the aqueous phase. The inherently neutral zwitterions $[^{-}ABH_a^+]$, gets partitioned into octanol more likely than the charged species such as $[HABH_o^+]$, and $[^{-}AB_o]$. Zwitterions exist at a particular pH called isoelectric point (pI), which is the average of its acidic group's pk_A ($COOH \rightarrow COO^- + H^+$), and the basic groups pk_B value ($NH_3^+ \rightarrow NH_2 + H^+$), i.e. $pI = (pk_A + pk_B)/2$.

2.1.2. Algebraic method for simple amino acids:

Based on the above seven equilibriums (Eqns,1,2,4,6-9), seven algebraic equations can be framed in terms of $[HABH_a^+]$, $[HAB_a]$, $[^{-}ABH_a^+]$, $[^{-}AB_a]$, $[HABH_o^+]$, $[HAB_o]$, $[^{-}ABH_o^+]$, $[^{-}AB_o]$. The distribution coefficient (D) for such a system is defined as,

$$D = \left(\frac{[HABH_o^+] + [HAB_o] + [^{-}ABH_o^+] + [^{-}AB_o]}{[HABH_a^+] + [HAB_a] + [^{-}ABH_a^+] + [^{-}AB_a]}\right)\left(\frac{1}{r}\right) \qquad (10)$$

By re-expressing $[HABH_a^+]$, $[HAB_a]$, $[^{-}ABH_a^+]$, $[HABH_o^+]$, $[HAB_o]$, $[^{-}ABH_o^+]$, $[^{-}AB_o]$, in terms of $[^{-}AB_a]$ using the seven algebraic equations and substituting those expressions into Eqn. 10, we obtain Eqn. 11,

$$D = \left(\frac{P_{HABH^+}\,[H_a^+]^2 + P_{HAB}\,K_1\,[H_a^+] + P_{^{-}ABH^+}\,K_2\,[H_a^+] + P_{^{-}AB}K_1K_2}{[H_a^+]^2 + K_1\,[H_a^+] + K_2\,[H_a^+] + K_1K_2}\right) \qquad (11)$$

In the above Eqn. 11, we made the valid assumption concerning the kinetic mechanism that, $K_1 = K_5$, and $K_2 = K_4$, for the above system. The above equation (Eqn. 11) will reduce to diprotic model (SI.4), if $K_1 = 0$, i.e. the neutral species $[HAB]$ does not exist in the kinetic mechanism.

2.1.3. Dynamic method for simple amino acids:

Based on the kinetic mechanism (Eqns. 1 - 9), the rate equation (SI.8) can be written for eight species, $[HABH_a^+]$, $[HAB_a]$, $[^-ABH_a^+]$, $[^-AB_a]$, $[HABH_o^+]$, $[HAB_o]$, $[^-ABH_o^+]$, and $[^-AB_o]$, as follows,

$$\begin{aligned}\frac{d[HABH_a^+]}{dt} = &-k_1[HABH_a^+] + k_{-1}[H_a^+][HAB_a] - k_2[HABH_a^+] \\ &+k_{-2}[^-ABH_a^+][H_a^+] - k_{p1}[HABH_a^+] + k_{-p1}[HABH_o^+]\end{aligned} \tag{12}$$

$$\begin{aligned}\frac{d[HAB_a]}{dt} = &+k_1[HABH_a^+] - k_{-1}[H_a^+][HAB_a] - k_4[HAB_a] \\ &+k_{-4}[H_a^+][^-AB_a] - k_3[H_a^+][HAB_a] + k_{-3}[H_a^+][^-ABH_a^+] \\ &-k_{p2}[HAB_a] + k_{-p2}[HAB_o]\end{aligned} \tag{13}$$

$$\begin{aligned}\frac{d[^-ABH_a^+]}{dt} = &+k_2[HABH_a^+] - k_{-2}[^-ABH_a^+][H_a^+] - k_5[^-ABH_a^+] \\ &+k_{-5}[H_a^+][^-AB_a] + k_3[H_a^+][HAB_a] - k_{-3}[H_a^+][^-ABH_a^+] \\ &-k_{p3}[^-ABH_a^+] + k_{-p3}[^-ABH_o^+]\end{aligned} \tag{14}$$

$$\frac{d[^-AB_a]}{dt} = +k_4[HAB_a] - k_{-4}[H_a^+][^-AB_a] + k_5[^-ABH_a^+] \tag{15}$$

$$-k_{-5}[H_a^+][^-AB_a] - k_{p4}[^-AB_a] + k_{-p4}[^-AB_o]$$

$$\frac{d[HABH_o^+]}{dt} = +k_{p1}[HABH_a^+] - k_{-p1}[HABH_o^+] \tag{16}$$

$$\frac{d[HAB_o]}{dt} = +k_{p2}[HAB_a] - k_{-p2}[HAB_o] \tag{17}$$

$$\frac{d[^-ABH_o^+]}{dt} = +k_{p3}[^-ABH_a^+] - k_{-p3}[^-ABH_o^+] \tag{18}$$

$$\frac{d[^-AB_o]}{dt} = +k_{p4}[^-AB_a] - k_{-p4}[^-AB_o] \tag{19}$$

By numerically integrating the above set of coupled differential equations, Eqns 12 - 19, we obtain the concentration of eight species, $[HABH_a^+]$, $[HAB_a]$, $[^-ABH_a^+]$, $[^-AB_a]$, $[HABH_o^+]$, $[HAB_o]$, $[^-ABH_o^+]$, and $[^-AB_o]$ at different time points. The resultant concentrations at equilibrium, can be substituted into the Eqn. 10, to obtain the distribution coefficient. In the above model, $[H_a^+]$ or pH, is assumed to be constant with respect to time, due to the use of appropriate buffers in $log_{10}D$ experiments to stabilize the pH of the aqueous phase. To simulate the $log_{10}D$ profile in Figure 1B, the parameters such as k_1, k_{-1}, k_2, k_{-2}, k_3, k_{-3}, k_4, k_{-4}, k_5, k_{-5}, k_{p1}, k_{-p1}, k_{p2}, k_{-p2}, k_{p3}, k_{-p3}, k_{p4}, k_{-p4}, were set to 10^{-8}, 1.0, 10^{-3}, 1.0, $10^{-(8+3)/2}$, 1.0, 10^{-3},1.0, 10^{-8},1.0, 10^{-3}, 1.0, 10^{5}, 1.0, 10^{5}, 1.0, 10^{-1}, 1.0, respectively. The initial concentrations of all the eight species $[HABH_a^+]_0$, $[^-ABH_a^+]_0$, $[HAB_a]_0$, $[^-AB_a]_0$, $[HABH_o^+]_0$, $[^-ABH_o^+]_0$, $[HAB_o]_0$ and $[^-AB_o]_0$, were set to 1.0,0,0,0,0,0, 0,0, respectively and the pH, was varied linearly between 1 and 10.

2. 2. Monoprotic acid with salt (KCl)

2.2.1. Kinetic model for monoprotic acid with salt (KCl)

Consider the distribution of a weak mono-protic acid (HA) between an aqueous solvent and an organic hydrophobic solvent (octanol) in the presence of a potassium chloride (KCl) (Figure 2A) [3,5,12,13]. In aqueous phase, the weak acid, $[HA]$, exists in the unionized form $[HA_a]$ and ionized form $[A_a^-]$. The equilibrium between these two states can be written as (Eq. 20),

$$HA_a \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} H_a^+ + A_a^- \qquad (20)$$

k_1, k_{-1} are the forward and reverse kinetic rates for the dissociation of the weak acid, $[HA]$, respectively. The salt KCl dissociates completely in aqueous solution to form $[K_a^+]$ and $[Cl_a^-]$ with a forward kinetic rate of k_2 (Eqn. 21).

$$KCl_a \overset{k_2}{\rightarrow} K_a^+ + Cl_a^- \qquad (21)$$

The unionized species, $[HA_a]$ can be partitioned into octanol layer with the forward and reverse kinetic rates as k_{p1} , k_{-p1}, as given below (Eqn. 22),

$$HA_a \underset{k_{-p1}}{\overset{k_{p1}}{\Leftrightarrow}} HA_o \qquad (22)$$

The dissociated $[K_a^+]$ and $[Cl_a^-]$ gets partitioned directly into octanol with the forward and reverse kinetic rates as k_{p2} , k_{-p2}, respectively (Eqn. 23).

$$K_a^+ + Cl_a^- \underset{k_{-p2}}{\overset{k_{p2}}{\Leftrightarrow}} {}^{+}KCl_o^- \qquad (23)$$

The ionized species $[A_a^-]$, would require its negative charge be neutralized before partitioned into octanol layer through the prevalent cation, K_a^+, coming from the dissociation of KCl. The forward and reverse kinetic rates of the partition of $[A_a^-]:[K_a^+]$ are k_{p3} , k_{-p3}, respectively, as given below (Eqn. 24),

$$A_a^- + K_a^+ \underset{k_{-p3}}{\overset{k_{p3}}{\Leftrightarrow}} {}^{+}KA_o^- \qquad (24)$$

Additionally, as seen with the monoprotic acid (SI.1), it is also possible for a fraction of the ionized species $[A_a^-]$, to get directly partitioned into the octanol layer without ion-pair formation as $[A_o^-]$, whose forward and reverse kinetic rates are given as k_{p5} , k_{-p5}, respectively, (Eqn. 25),

$$A_a^- \underset{k_{-p5}}{\overset{k_{p5}}{\Leftrightarrow}} A_o^- \qquad (25)$$

The $[H_a^+]$ ions (arising from the dissociation of HA_a, (Eq. 20)) and $[Cl_a^-]$ (from $[KCl_a]$) gets partitioned into octanol with the forward and reverse kinetic rates as k_{p4} , k_{-p4}, respectively (Eqn. 26).

$$H_a^+ + Cl_a^- \underset{k_{-p4}}{\overset{k_{p4}}{\Leftrightarrow}} {}^{+}HCl_o^- \qquad (26)$$

Finally, all the four species HA_o, ${}^{+}KCl_o^-$, ${}^{+}KA_o^-$, ${}^{+}HCl_o^-$, partitioned into the octanol layer undergo a dynamic equilibrium whose forward and reverse kinetic rates are given as k_3, k_{-3}, respectively (Eqn. 27).

$$HA_o + {}^{+}KCl_o^- \underset{k_{-3}}{\overset{k_3}{\Leftrightarrow}} {}^{+}KA_o^- + {}^{+}HCl_o^- \qquad (27)$$

2.2.2. Algebraic method for monoprotic acid with KCl:

Applying Le-Chaterlier's principle, on the above six equilibriums (Eqns. 20, 22, 23, 25, 26 & 27), and considering the mass balance for $[K_a^+]$, $[Cl_a^-]$, and $[HA_T]$, we obtain the equilibrium concentrations of $[HA_o]$, $[{}^{+}KA_o^-]$, $[A_o^-]$, $[HA_a]$ and $[A_a^-]$ (SI.5). The distribution coefficient (D) of the monoprotic system in the presence of salt is defined as,

$$D = \left(\frac{[HA_o] + [{}^{+}KA_o^-] + [A_o^-]}{[HA_a] + [A_a^-]}\right)\left(\frac{1}{r}\right) \qquad (28)$$

By substituting the analytical expressions of $[HA_o]$, $[{}^{+}KA_o^-]$, $[A_o^-]$, $[HA_a]$ and $[A_a^-]$ into Eqn. 28, we obtain,

$$D = \left(\frac{P_{HA}\, P_{HCl}\, [H_a^+] + P_A\, P_{HCl}\, K_1 + P_{HA}\, P_{KCl}\, K_3\, K_T}{P_{HCl}\, [H_a^+] + P_{HCl}\, K_1}\right) \qquad (29)$$

To simulate the $log_{10}D$ profile in Figure 2B, the parameters such as K_1, K_3, P_{HA}, P_A, P_{HCl}, P_{KCl}, were set to 10^{-3}, 10^{-8}, 10^{3}, 10^{-3}, 10^{-15}, 10^{-15}, respectively. K_T was set to, 0, 1, 10, 100, 1000 and 10000 M, to simulate 6 profiles and the pH, was varied linearly, between 1 and 10. If we set either, K_3 or $P_{KCl} = 0$, the above Eqn. 29 will reduce to that of a model of simple monoprotic acid.

2.2.3. Dynamic method for monoprotic acid with KCl:

Based on the kinetic mechanism (Eqns. 20 - 27), the rate equation can be written for the ten species, $[HA_a]$, $[A_a^-]$, $[KCl_a]$, $[K_a^+]$, $[Cl_a^-]$, $[HA_o]$, $[{}^{+}KCl_o^-]$, $[{}^{+}KA_o^-]$, $[A_o^-]$, and $[{}^{+}HCl_o^-]$, as follows,

$$\frac{d[HA_a]}{dt} = -k_1[HA_a] + k_{-1}[H_a^+][A_a^-] - k_{p1}[HA_a] + k_{-p1}[HA_o] \tag{30}$$

$$\frac{d[A_a^-]}{dt} = +k_1[HA_a] - k_{-1}[H_a^+][A_a^-] - k_{p3}[K_a^+][A_a^-] + k_{-p3}[{}^+KA_o^-] \\ -k_{p5}[A_a^-] + k_{-p5}[A_o^-] \tag{31}$$

$$\frac{d[KCl_a]}{dt} = -k_2[KCl_a] \tag{32}$$

$$\frac{d[K_a^+]}{dt} = +k_2[KCl_a] - k_{p3}[K_a^+][A_a^-] + k_{-p3}[{}^+KA_o^-] \\ -k_{p2}[K_a^+][Cl_a^-] + k_{-p2}[{}^+KCl_o^-] \tag{33}$$

$$\frac{d[Cl_a^-]}{dt} = +k_2[KCl_a] - k_{p4}[H_a^+][Cl_a^-] + k_{-p4}[{}^+HCl_o^-] \\ -k_{p2}[K_a^+][Cl_a^-] + k_{-p2}[{}^+KCl_o^-] \tag{34}$$

$$\frac{d[HA_o]}{dt} = +k_{p1}[HA_a] - k_{-p1}[HA_o] - k_3[HA_o][{}^+KCl_o^-] \\ +k_{-3}[{}^+KA_o^-][{}^+HCl_o^-] \tag{35}$$

$$\frac{d[{}^+KCl_o^-]}{dt} = +k_{p2}[K_a^+][Cl_a^-] - k_{-p2}[{}^+KCl_o^-] - k_3[HA_o][{}^+KCl_o^-] \\ +k_{-3}[{}^+KA_o^-][{}^+HCl_o^-] \tag{36}$$

$$\frac{d[^{+}KA_o^-]}{dt} = +k_{p3}[K_a^+][A_a^-] - k_{-p3}[^{+}KA_o^-] + k_3[HA_o][^{+}KCl_o^-] - k_{-3}[^{+}KA_o^-][^{+}HCl_o^-] \quad (37)$$

$$\frac{d[A_o^-]}{dt} = +k_{p5}[A_a^-] - k_{-p5}[A_o^-] \quad (38)$$

$$\frac{d[^{+}HCl_o^-]}{dt} = +k_{p4}[H_a^+][Cl_a^-] - k_{-p4}[^{+}HCl_o^-] + k_3[HA_o][^{+}KCl_o^-] - k_{-3}[^{+}KA_o^-][^{+}HCl_o^-] \quad (39)$$

By numerically integrating the above set of coupled differential equations, Eqns 30 - 39, we obtain the concentration of ten species, $[HA_a], [A_a^-]$, $[KCl_a], [K_a^+]$, $[Cl_a^-]$, $[HA_o], [^{+}KCl_o^-]$, $[^{+}KA_o^-]$, $[A_o^-]$ and $[^{+}HCl_o^-]$, at different time points. The resultant concentrations of $[HA_a], [A_a^-]$, $[A_o^-], [HA_o], [^{+}KA_o^-]$ at equilibrium, can be substituted into the Eqn. 28, to obtain the distribution coefficient. In the above model, $[H_a^+]$ or pH, is assumed to be constant with respect to time, because of the usage of appropriate buffers in $log_{10}D$ experiments to stabilize the pH of the aqueous phase. To simulate the $log_{10}D$ profile in Figure 2B, the parameters such as k_1, k_{-1},k_2, k_3, k_{-3}, k_{p1}, k_{-p1}, k_{p2}, k_{-p2}, k_{p3}, k_{-p3}, k_{p4}, k_{-p4}, k_{p5}, k_{-p5} were set to 10^{-3}, 1.0, 10^{15} (any large value, $> 10^2$), 10^{-8}, 1.0, 10^3, 1.0, 10^{-15}, 1.0, 10^{-2}, 1.0, 10^{-15}, 1.0 , 10^{-3}, 1.0 respectively. The initial concentrations of the following nine species, $[HA_a]_0, [A_a^-]_0, [KCl_a]_0$, $[K_a^+]_0$, $[Cl_a^-]_0$, $[HA_o]_0, [^{+}KCl_o^-]_0$, $[^{+}KA_o^-]_0$, $[A_o^-]_0$ and $[^{+}HCl_o^-]_0$, were initialized to 1.0, 0, K_T,0,0,0, 0,0,0,0 M respectively. Six different simulations were performed by varying the initial concentration of K_T, as, 0, 1, 10, 100, 1000, 10000 M, and the pH was varied linearly between 1 and 12.

2. 3. Monoprotic acid in the presence of co-solvent

2.3.1. Kinetic model for monoprotic acid in the presence of co-solvent

Consider the distribution of a weak mono-protic acid (HA) between an aqueous solvent and an organic hydrophobic solvent (octanol) in the presence of co-solvent (S) such as DMSO (Figure 3A) [3,4]. In aqueous phase, the weak acid, $[HA]$, exists in unionized form $[HA_a]$ and ionized form $[A_a^-]$. The equilibrium between these two states can be written as (Eqn. 40),

$$HA_a \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} H_a^+ + A_a^- \qquad (40)$$

The distribution of a co-solvent between an aqueous buffer and octanol is given by (Eq. 41),

$$S_a \underset{k_{-sp}}{\overset{k_{sp}}{\Leftrightarrow}} S_o \qquad (41)$$

k_{sp}, k_{-sp} are the forward and reverse kinetic rates for the partition of the co-solvent $[S]$ between the aqueous (S_a) and octanol phase (S_o), respectively. The co-solvation of unionized form $[HA_a]$ and $[A_a^-]$ in aqueous buffer can be written as (Eq. 42 & 43)

$$HA_a + n_1\, S_a \underset{k_{-s1}}{\overset{k_{s1}}{\Leftrightarrow}} HAS_a \qquad (42)$$

$$A_a^- + n_2\, S_a \underset{k_{-s2}}{\overset{k_{s2}}{\Leftrightarrow}} {}^-AS_a \qquad (43)$$

$k_{s1}, k_{-s1}; k_{s2}, k_{-s2}$; are the forward and reverse kinetic rates for the co-solvation of

$[HA_a]$, $[A_a^-]$ in the aqueous phase, respectively. n_1, n_2, are the number of solvent molecules required to co-solvate $[HA_a]$, $[A_a^-]$, respectively.

$[HA_a]$, and $[A_a^-]$, can be partitioned into octanol layer as shown below (Eqns. 44 & 45),

$$HA_a \underset{k_{-p1}}{\overset{k_{p1}}{\Leftrightarrow}} HA_o \tag{44}$$

$$A_a^- \underset{k_{-p2}}{\overset{k_{p2}}{\Leftrightarrow}} A_o^- \tag{45}$$

The co-solvation of unionized form $[HA_a]$ and $[A_a^-]$ in octanol phase can be written as (Eqns. 46 & 47),

$$HA_o + n_3\, S_o \underset{k_{-s3}}{\overset{k_{s3}}{\Leftrightarrow}} HAS_o \tag{46}$$

$$A_o^- + n_4\, S_o \underset{k_{-s4}}{\overset{k_{s4}}{\Leftrightarrow}} {}^-AS_o \tag{47}$$

k_{s3}, k_{-s3}; k_{s4}, k_{-s4}, are the forward and reverse kinetic rates for the co-solvation of $[HA_o]$, $[A_o^-]$ in the aqueous phase, respectively. n_3, n_4, are the number of solvent molecules required to co-solvate $[HA_o]$, $[A_o^-]$, respectively.

2.3.2. Algebraic method for monoprotic acid in the presence of co-solvent:

Applying Le-Chaterlier's principle, on the above equilibriums (Eqns. 40 - 47), and considering the law of conservation of mass for $[HA_T]$, and $[S_T]$, we solve for the analytical expressions of the ten species $[HA_a]$, $[HA_o]$, $[A_a^-]$, $[A_o^-]$, $[HAS_a]$, $[{}^-AS_a]$,

$[HAS_o]$, $[^{-}AS_o]$, $[S_a]$ and $[S_o]$ (SI.6).

The distribution coefficient (D) is defined as (Eqn. 48),

$$D = \left(\frac{[HA_o] + [A_o^-] + [HAS_o] + [^{-}AS_o]}{[HA_a] + [A_a^-] + [HAS_a] + [^{-}AS_a]}\right)\left(\frac{1}{r}\right) \qquad (48)$$

On substituting the analytical expressions for $[HA_o]$, $[A_o^-]$, $[HAS_o]$, $[^{-}AS_o]$, $[HA_a]$, $[A_a^-]$, $[HAS_a]$, $[^{-}AS_a]$ into Eqn. 48, we obtain Eqn. 49,

$$D = \left(\frac{P_{HA}[H_a^+] + P_A K_1 + P_{HA}K_{s3}[H_a^+][S_o]^{n_3} + P_A K_1 K_{s4}[S_o]^{n_4}}{[H_a^+] + K_1 + K_{s1}[H_a^+][S_a]^{n_1} + K_1 K_{s2}[S_a]^{n_2}}\right) \qquad (49)$$

2.3.3. Dynamic method for monoprotic acid in the presence of co-solvent:

The differential equations for the rate of change of different species can be framed based on the kinetic mechanism given by (Eqns. 40 - 47). The rate equation for ten species, $[HA_a]$, $[A_a^-]$, $[HAS_a]$, $[^{-}AS_a]$, $[HA_o]$, $[A_o^-]$, $[HAS_o]$, $[^{-}AS_o]$, $[S_a]$ and $[S_o]$, can be written as follows,

$$\frac{d[HA_a]}{dt} = \begin{array}{c} -k_1[HA_a] + k_{-1}[H_a^+][A_a^-] - k_{s1}[HA_a][S_a]^{n_1} + k_{-s1}[HAS_a] \\ -k_{p1}[HA_a] + k_{-p1}[HA_o] \end{array} \qquad (50)$$

$$\frac{d[A_a^-]}{dt} = \begin{array}{c} +k_1[HA_a] - k_{-1}[H_a^+][A_a^-] - k_{s2}[A_a^-][S_a]^{n_2} + k_{-s2}[^{-}AS_a] \\ -k_{p2}[A_a^-] + k_{-p2}[A_o^-] \end{array} \qquad (51)$$

$$\frac{d[HAS_a]}{dt} = +k_{s1}[HA_a][S_a]^{n_1} - k_{-s1}[HAS_a] \qquad (52)$$

$$\frac{d[^{-}AS_a]}{dt} = \quad +k_{s2}[A_a^-][S_a]^{n_2} - k_{-s2}[^{-}AS_a] \tag{53}$$

$$\frac{d[HA_o]}{dt} = \quad +k_{p1}[HA_a] - k_{-p1}[HA_o] - k_{s3}[HA_o][S_o]^{n_3} + k_{-s3}[HAS_o] \tag{54}$$

$$\frac{d[A_o^-]}{dt} = \quad +k_{p2}[A_a^-] - k_{-p2}[A_o^-] \; - k_{s4}[A_o^-][S_o]^{n_4} + k_{-s4}[^{-}AS_o] \tag{55}$$

$$\frac{d[HAS_o]}{dt} = \quad +k_{s3}[HA_o][S_o]^{n_3} - k_{-s3}[HAS_o] \tag{56}$$

$$\frac{d[^{-}AS_o]}{dt} = \quad +k_{s4}[A_o^-][S_o]^{n_4} - k_{-s4}[^{-}AS_o] \tag{57}$$

$$\frac{d[S_a]}{dt} = \quad -k_{ps}[S_a] + k_{-ps}[S_o] \tag{58}$$

$$\frac{d[S_o]}{dt} = \quad +k_{ps}[S_a] - k_{-ps}[S_o] \tag{59}$$

By numerically integrating the above set of coupled differential equations, Eqns 50 - 59, we obtain the concentration of ten species, $[HA_a]$, $[A_a^-]$, $[HAS_a]$, $[^{-}AS_a]$, $[HA_o]$, $[A_o^-]$, $[HAS_o]$, $[^{-}AS_o]$, $[S_a]$ and $[S_o]$ at different time points. The resultant concentrations at equilibrium, (i.e. concentrations at $t \to \infty$, or a long time period (>10^{10} s)), can be substituted into the Eqn. 48, to obtain the distribution coefficient. In the above model, $[H_a^+]$ or pH, is assumed to be constant with respect to time. To simulate the $log_{10}D$ profile in Figure 3B&D, the parameters such as k_1, k_{-1}, k_{s1}, k_{-s1}, k_{s2}, k_{-s2}, k_{-s3}, k_{s3}, k_{-s4}, k_{-s4}, k_{p1}, k_{-p1}, k_{p2}, k_{-p2}, k_{ps}, k_{-ps}, were set to 10^{-2}, 1.0, 10^7, 1.0, 10^4, 1.0, 10^8, 1.0, 10^5, 1.0, 10^1, 1.0, 10^{-3}, 1.0, $8 * 10^{-2}$, 1.0 respectively. n_1, n_2, n_3, n_4, were

set to 1,0,0,0, Figure 3B and n_1, n_2, n_3, n_4, were set to 1,1,1,1, Figure 3D. The initial concentrations of the eight species $[HA_a]_0$, $[A_a^-]_0$, $[HAS_a]_0$, $[^-AS_a]_0$, $[HA_o]_0$, $[A_o^-]_0$, $[HAS_o]_0$,$[^-AS_o]_0$, $[S_a]_0$ and $[S_o]_0$ were initialized to 1.0, 0,0,0,0,0,0,0,$[S_T]$,0 respectively. A total of 12 simulations were performed with the following concentration of the total solvent, $[S_T]$ as 0, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 0.5, 1, 5, 10^1, 10^2, 10^3 and the pH, was varied linearly between 1 and 12.

3. Results:

In a recent $log_{10}D$ comparison study several amphoteric molecules of pharmaceutical interest were analysed using both shake flask (octanol-buffer) and potentiometric method [10]. The experimental data on Nalidixic acid, Mebendazole, Benazepril, Telmisartan from this previous study were used in our current analysis [10]. Despite the fact that both Nalidixic acid and Mebendazole are amphoteric in nature, the experimental $log_{10}D$ resembled a typical monoprotic and monalkaline profile, respectively (Figure 4A & B). hence, they were analysed using monoprotic or monalkaline models. Benazepril is a complex amphoteric molecule with two amino and one carboxylic acid functional group, but it $log_{10}D$ profile followed a simple monoprotic-monalkaline amino acid and hence its data was analysed using a simple amphoteric model (Figure 4C). Telmisartan on the other-hand could not be explained through simpler models, hence a complex monoprotic-dialkaline model was used to explain its experimental data (Figure 4D). The optimized pK_a, pK_b, and $log_{10}P$ values were consistent with the previous studies and are summarised in Table 1 [10].

The simulation of the $log_{10}D$ profile of a monoprotic acid in the presence of salt was performed to assess the effect of salt on P_{HA}, P_A, pK_a values. The results show that only the value of P_A was linearly affected (directly proportional), whereas, the, P_{HA} and

pK_a remained constant with the addition of salt [5]. The effect of a co-solvent such as DMSO on the $log_{10}D$ profile of a monoprotic acid was assessed through simulation. The $log_{10}D$ profile was primarily influenced by the differential ability of the solvent to co-solvate different species of monoprotic acid such as $[HA_a]$, $[HA_o]$, $[A_a^-]$, $[A_o^-]$. For example, if we assume the binding affinity constant (K_{S1}, K_{S2}, K_{S3}, K_{S4}) of solvent ($[S]$) towards $[HA_a]$, $[HA_o]$, $[A_a^-]$, $[A_o^-]$, to be 1,1,1,1, respectively, and the stoichiometry n_1, n_2, n_3, n_4 to be 1,0,0,0, or 2,0,0,0 or 3,0,0,0 respectively, both P_{HA} and pK_a, were affected non-linearly, whereas, P_A, remained constant, with the addition of co-solvent ($[S_T]$). On the other-hand, if we let the stoichiometry n_1, n_2, n_3, n_4 to be 1,1,1,1, or 2,1,1,1, or 3,1,1,1, respectively, except for the stoichiometric constant, 1,1,1,1, for which pK_a, remained constant; in all other cases, the parameters, P_{HA} , P_A, and pK_a, varied non-linearly with increase in solvent concentration [3,5]. The effect of solvent on $log_{10}D$ profile became even more complex, when the degree of co-solvation of different species were assigned higher stoichiometry ratios.

4. Discussion:

Telmisartan exhibited a complex $log_{10}D$ profile and could be explained only through di-alkaline-monoprotic model. In this model eight different species were consider to exist in aqueous phase and an additional eight species to be partitioned into the octanol phase. Considering only the species present in the aqueous phase, framing a kinetic model consisting of eight species interconnected among itself through a network of equilibrium reactions is not simple. For example, if we consider each species as 'nodes' and the interconnecting equilibrium reactions as bidirectional 'edges', then the graph theory suggest a maximum of N(N-1) = 56 edges or equilibrium reactions to exist among N=8 species [14]. Under such a circumstance, developing an algebraic model based on the kinetic mechanism can be considerably simplified by considering only seven unique equilibriums that inter connect any of these eight species, and including the law of conservation of mass of the di-alkaline-monoprotic as a constraint. Additional eight equations can be derived from the equilibriums concerned with the partitioning of the

eight species between aqueous and octanol layer. On the other-hand, in the dynamic approach, the rate equation of the sixteen species will depend primarily on the network of equilibriums that is assumed to exist among the species. Even though the concentrations of the 16 species may vary during the pre-steady state condition, at equilibrium or steady state, the concentrations remains invariably the same, for equivalent kinetic mechanisms.

The model-fitting (Figure 4C & D) was carried out using algebraic method (Eqn. 11 & SI Eqn. 60) and the resultant optimized parameters were used in dynamic approach (SI Eqns. 156 -171) to simulate the $log_{10}D$ profiles. The equilibrium concentrations obtained through algebraic and dynamic methods were comparable and equivalent to a degree of two decimal points for most of the data points (Table 2). The minor discrepancies seen in the accuracy of dynamic method compared to algebraic method is caused by errors in numerical integration when solving ordinary differential equations (ODE). If ODE's are integrated over a long period of 'integral time', such as 10^8s, the 'calculation time' required to solve the ODE's numerically and the 'integration' error, will significantly increase. The significant amount of time (> 2 hrs on a 8 cpu core computer) taken to solve the system of ODE's in dynamic approach is the major limitation in implementing dynamic approach within model-fitting to routinely analyse an experimental data.

The simulation of the $log_{10}D$ profile of the monoprotic acid with salt, suggests an increase in apparent P_A', (i.e. P_A in the presence of co-solvent) value as the concentration of salt (K_T) increases. If we compare equation 29, with the simple monoprotic model, $D = \left(\frac{P_{HA}\,[H_a^+] + P_A\,K_1}{[H_a^+] + K_1}\right)$, we can obtain $P_A' = P_A + \left(P_{HA}.\frac{P_{KCl}}{P_{HCl}}.\frac{K_3}{K_1}\right) K_T$, which shows that, P_A' is linearly related to K_T, with the slope and intercept as $\left(P_{HA}.\frac{P_{KCl}}{P_{HCl}}.\frac{K_3}{K_1}\right)$ and P_A, respectively. Physically, the addition of salt, KCl_T, tend to increase the formation of ion pair, $K^+A_a^-$, in the aqueous layer through the association of the ionized species A_a^- (from monoprotic acid) and the prevalently present cation K^+ (from KCl) [3,5]. Unlike the charged species, A_a^-, the neutral ion pair $K^+{:}A_a^-$ in aqueous phase gets partitioned into

octanol as $[K^+A_o^-]$, thereby, increasing the total concentration of $[A_{o\ Total}^-]$, (where, $[A_{o\ Total}^-]= [A_o^-] + [K^+A_o^-]$), in the octanol phase. Since, $P_A = [A_{o\ Total}^-]/[A_{a\ Total}^-]$ and is directly proportional to $[A_{o\ Total}^-]$, when $[A_{o\ Total}^-]$ is increased due to partitioning of the ion pair $[K^+A_o^-]$, P_A also increases proportionally.

To understand the complex $log_{10}D$ profile of the monoprotic acid in the presence of co-solvent we compare Eqn. 49, with the model of simple monoprotic acid and obtain the expressions for apparent P'_{HA} , P'_A, and pK'_a [3]. The apparent P'_{HA} , P'_A, and pK'_a are the modified P_{HA}, P_A and pK_a, of the monoprotic acid in the presence of co-solvent, respectively. The resultant expressions, $P'_{HA} = P_{HA}\left(\frac{1+K_{S3}S_o^{n_3}}{1+K_{S1}S_o^{n_1}}\right)$, $P'_A = P_A\left(\frac{1+K_{S4}S_o^{n_4}}{1+K_{S2}S_o^{n_2}}\right)$ and $K'_1 = K_1\left(\frac{1+K_{S2}S_a^{n_2}}{1+K_{S1}S_a^{n_1}}\right)$, clearly shows that a non-linear relationship exist between P'_{HA}, P'_A, K'_1 and the co-solvent concentration in octanol (S_o) and aqueous (S_a) phase. If we consider a simpler stoichiometry of, n_1, n_2, n_3, n_4 to be 1 (or 2 or 3), 0,0,0, and solvent binding affinities, K_{S1}, K_{S2}, K_{S3}, K_{S4}, to be 1,1,1,1, the apparent P'_A remains constant, i.e. $P'_A = P_A$ and the apparent P'_{HA} and K'_1 becomes non-linearly related to solvent concentration as, $P'_{HA} = P_{HA}\left(\frac{2}{1+S_o^{n_1}}\right)$ and $K'_1 = K_1\left(\frac{2}{1+S_a^{n_1}}\right)$, respectively. In our case the apparent K'_1, decreases with increase in co-solvent concentration (i.e. $K'_1 \propto \frac{1}{S_a}$) ; and since, K'_1 is inversely related to pK'_1 through its definition, $K'_1 = 10^{-pK'_1} = \frac{1}{10^{pK'_1}}$; we can infer that K'_1 is inversely related to co-solvent concentration. Conversely, we can also state that the apparent pK'_1, increases with increase in co-solvent concentration, which is in agreement with the earlier reports [15].

Supporting Information

Supporting Information contains explicit derivation of $\log_{10}$ D for mono-protic, di-protic acid, mono-alkaline, mono-protic acid with salt, mono-protic acid with solvent, monoprotic-dialkaline molecule.

Acknowledgments

JK would like to thank Jimma University, V ClinBio Pvt Ltd for financial support; and Xavier Suburats (Universitat de Barcelona) for kindly providing the experimental data to carry out this work.

REFERENCES

[1] F.H. Clarke, Ionization constants by curve fitting: Application to the determination of partition coefficients, Journal of pharmaceutical sciences 73 (1984) 226-230.
[2] A. Avdeef, pH-Metric log P. Part 1. Difference Plots for Determining Ion-Pair Octanol-Water Partition Coefficients of Multiprotic Substances, Molecular Informatics 11 (1992) 510-517.
[3] A. Avdeef, Physicochemical profiling (solubility, permeability and charge state), Current topics in medicinal chemistry 1 (2001) 277-351.
[4] R. Mannhold, H. Kubinyi, G. Folkers, Molecular drug properties: measurement and prediction, John Wiley & Sons, 2008.
[5] P. Vladimir, B. Testa, W. Han van de, Lipophilicity in Drug Action and Toxicology, VCH, 1996.
[6] F.H. Clarke, N.M. Cahoon, Ionization constants by curve fitting: Determination of partition and distribution coefficients of acids and bases and their ions, Journal of pharmaceutical sciences 76 (1987) 611-620.
[7] J.J. Baeza-Baeza, M.C. Garcia-Alvarez-Coque, Systematic approach to calculate the concentration of chemical species in multi-equilibrium problems, Journal of Chemical Education 88 (2010) 169-173.
[8] J.J. Baeza-Baeza, M.C. García-Álvarez-Coque, Systematic approach for calculating

the concentrations of chemical species in multiequilibrium problems: inclusion of the ionic strength effects, Journal of Chemical Education 89 (2012) 900-904.
[9] R.E. Glaser, M.A. Delarosa, A.O. Salau, C. Chicone, Dynamical approach to multiequilibria problems for mixtures of acids and their conjugated bases, Journal of Chemical Education 91 (2014) 1009-1016.
[10] C. Ràfols, X. Subirats, J. Rubio, M. Rosés, E. Bosch, Lipophilicity of amphoteric and zwitterionic compounds: A comparative study of determination methods, Talanta 162 (2017) 293-299.
[11] M. Akamatsu, Y. Yoshida, H. Nakamura, M. Asao, H. Iwamura, T. Fujita, Hydrophobicity of di-and tripeptides having unionizable side chains and correlation with substituent and structural Parameters, Molecular Informatics 8 (1989) 195-203.
[12] A. Avdeef, K.J. Box, K. Takács-Novák, pH–metric log P. 6: Effects of sodium, potassium, and N-CH3-D-glucamine on the octanol-water partitioning of prostaglandins E1 and E2, Journal of pharmaceutical sciences 84 (1995) 523-529.
[13] A. Avdeef, pH-metric log P. II: Refinement of partition coefficients and ionization constants of multiprotic substances, Journal of pharmaceutical sciences 82 (1993) 183-190.
[14] O. Mason, M. Verwoerd, Graph theory and networks in biology, IET systems biology 1 (2007) 89-119.
[15] F.G. Bordwell, Equilibrium acidities in dimethyl sulfoxide solution, Accounts of Chemical Research 21 (1988) 456-463.
[16] M.U.s. Guide, The mathworks, Inc., Natick, MA 5 (1998) 333.

Tables:

Table 1 Fit parameters for Nalidixic acid, Mebendazole, Benazepril and Telmisartan based on monoprotic, monoalkaline, simple amphoteric and diamino-monoprotic acid models, respectively. The parameters such as $pK_{A,}\, pK_{B,}\, pK_{B1,}\, pK_{B2}$ are equivalent to the conventional pK_a which is related to the dissociation of the H^+ ion from the acid moieties. The $pK_{B,}\, pK_{B1,}\, pK_{B2}$ are not to be confused with the conventional pK_b which is related to the association of OH^- ion with the base moieties.

	Nalidixic acid	**Mebendazole**	**Benazepril**	**Telmisartan**
Model	monoprotic	monoalkaline	simple amphoteric	monoprotic-diamino amphoteric
Parameters				
	-1.363 ± 0.0736 (P_{HA})	-3.059 ± 0.0531 (P_{BOH})	0.458 ± 0.1075 (P_{HAB},P_{-ABH^+})	4.161 ± 0.2749 (P_{HABB}, P_{-ABBH^+},P_{-ABH^+B})
Partition coefficients (P)	1.714 ± 0.1672 (P_{A^-})	-1.511 ± 0.2575 (P_{B^+})	-1.307 ± 0.0499 (P_{HABH^+},P_{-AB})	1.458 ± 0.0935 (P_{HABBH^+},P_{HABH^+B}, $P_{-ABH^+BH^+}$,P_{-ABB})
				-71825692.165 ± 0 ($P_{HABH^+BH^+}$)
pK_A/pK_B	6.351 ± 0.1261 (pK_A)	10.248 ± 0.1716 (pK_B)	2.884 ± 0.1046 (pK_A)	5.6 ± 0.405 (pK_A)
			5.481 ± 0.1092 (pK_B)	4.439 ± 0.6262 (pK_{B_1})
				3.608 ± 0.3559 (pK_{B_2})

Table 2 Comparison of the log D values back calculated using algebraic and dynamic approach for Benazepril and Telmisartan, using simple amphoteric and diamino-monoprotic acid models, respectively. Accuracy to a degree of two decimal points was observed for most of the data points analysed through algebraic and dynamic approach.

Benazepril				Telmisartan			
		Calculated ($log_{10}D$)				Calculated ($log_{10}D$)	
pH	Experimental ($log_{10}D$)	Algebraic	Dynamic	pH	Experimental ($log_{10}D$)	Algebraic	Dynamic
1.95	0.38	0.3876	0.3864	1.95	0.24	0.2541	0.2461
2.42	0.74	0.7380	0.7373	2.42	1.07	1.0431	1.0219
3.00	1.06	1.0643	1.0662	3.00	1.99	2.0612	2.0304
3.75	1.29	1.2420	1.2469	3.75	3.32	3.1856	3.1582
4.19	1.28	1.2624	1.2675	4.19	3.65	3.6251	3.6165
4.75	1.25	1.2278	1.2310	4.75	3.63	3.9105	3.9576
5.30	0.99	1.0945	1.0926	5.75	3.96	3.7741	3.7890
6.04	0.60	0.6794	0.6700	6.04	3.84	3.6019	3.6096
6.65	0.34	0.2187	0.2092	6.89	2.85	2.8998	2.8967
7.22	-0.12	-0.1469	-0.1506	7.60	2.08	2.2678	2.2832
7.65	-0.36	-0.3177	-0.3159	8.00	1.91	1.9612	1.9963
				8.90	1.53	1.5632	1.5673
				9.72	1.57	1.4747	1.4768
				10.52	1.38	1.4597	1.4597
				10.91	1.52	1.4580	1.4421

Figure Legends:

Figure 1 (A)The kinetic model for a simple Amphoteric molecule which can exist as four species in aqueous buffer such as, $[HABH_a^+]$, $[HAB_a]$, $[^-ABH_a^+]$, and $[^-AB_a]$, and gets partitioned into octanol layer as $[HABH_o^+]$, $[HAB_o]$, $[^-ABH_o^+]$, and $[^-AB_o]$, respectively. (B) shows the simulation of the amphoteric model using algebraic method with equilibrium constants $K_1 = K_5 = 10^{-8}$ and $K_2 = K_4 = 10^{-3}$, partition coefficients, $P_{HABH^+} = 10^{-3}$, $P_{^-AB} = 10^{-1}$, and $P_{^-ABH^+} = P_{HAB} = 10^5$

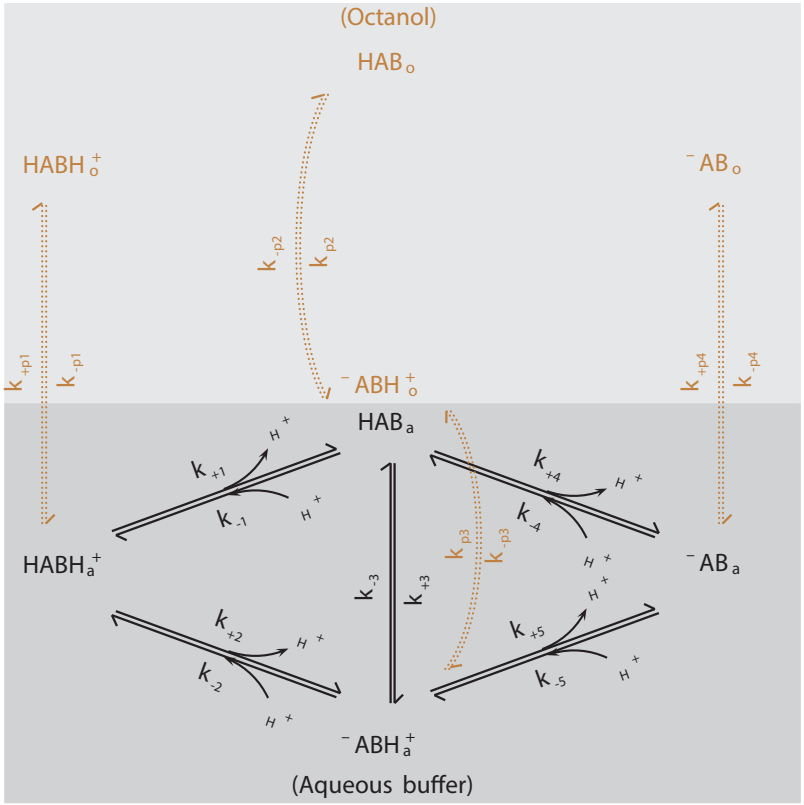
Figure 2 (A) The kinetic model for a monoprotic acid in the presence of a salt. In this mechanism, there exist a reversible equilibrium between HA_a, and A_a^- in aqueous phase and a reversible equilibrium between HA_o and $^+KA_o^-$ in the octanol phase exists. The formation of $^+KA_o^-$ in the octanol layer is mediated by the partitioning of the ion pairs $K_a^+ : Cl_a^-$, $K_a^+ : A_a^-$ and $H_a^+ : Cl_a^-$ from the aqueous to octanol layer. (B) Shows the $log_{10}D$ profiles of the monoprotic acid when the ratio of monoprotic to salt is varied between 1:0 to 1:10000. (C) shows how the P_{HA}, P_A, $pK_A$of the monoprotic acid varies as the salt concentration increases, from 0 to 10000, plotted here in log_{10} scale. It is clear that only P_A increases linearly with increase in salt concentration, whereas, P_{HA} and pK_A, remains constant. In the simulation, the P_{HA}, $P_A$and $pK_A$of the monoprotic acid were set to 10^3, 10^{-3} and 10^3 respectively, the concentration of monoprotic acid was set to 1, and the concentration of salt was set to 0,1,10,100,1000,10000 to simulate different $log_{10}D$,

profiles. The resultant profiles were fitted to the monoprotic model, $D = \left(\frac{P_{HA}\,[H_a^+] + P_A\,K_A}{[H_a^+] + K_A}\right)$, to obtain the apparent $P_{HA}{}'$, $P_A{}'$and $pK_A{}'$, which are plotted against the logarithm of salt concentration in (C)

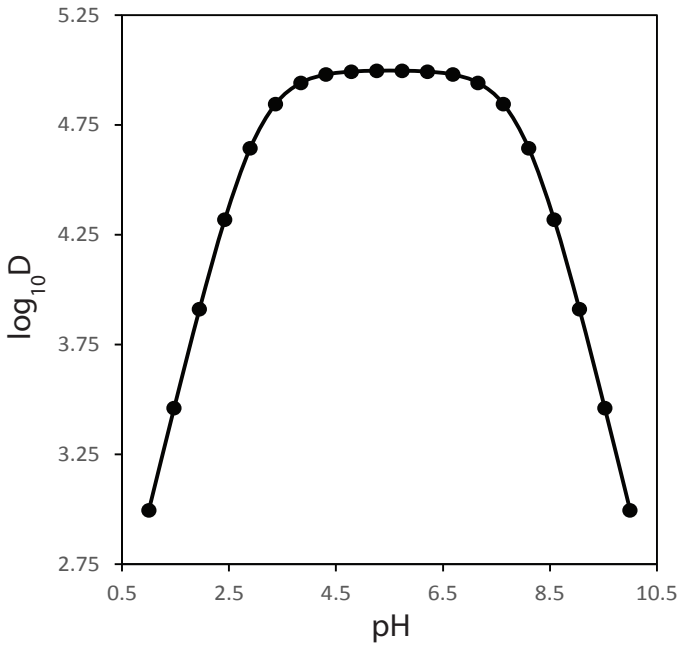
Figure 3 (A) The kinetic model for a monoprotic acid in the presence of a co-solvent ([S]). In this mechanism, there exists a reversible equilibrium between HA_a, and A_a^- in the aqueous phase, whose species also get partitioned into octanol as HA_o, and A_o^-, respectively. All four species, HA_a, A_a^-, HA_o, A_o^- are co-solvated by the solvent, present in aqueous($[S_a]$), or octanol phase ($[S_o]$) with the stoichiometry of n_1, n_2, n_3, n_4, to form HAS_a ,^-AS_a, HAS_o ,^-AS_o, respectively. (B) &(D) Shows the $log_{10}D$ profiles of the monoprotic acid when the ratio of monoprotic to solvent is varied between 1:0 and 1:1000, with the stoichiometry of n_1, n_2, n_3, n_4, to be 1:0:0:0 for (B) and 1:1:1:1 for (D), respectively. (C) & (E) shows how the P_{HA}, P_A, $pK_A$of the monoprotic acid varies as the solvent concentration increases, from 0 to 1000 (shown here in log_{10} scale). It is clear that the modulation of P_A, P_{HA} and $pK_A$is strongly dependent on the degree of the co-solvation of different species. In the simulation, the P_{HA}, $P_A$and $pK_A$of the monoprotic acid were set to 10^3,10^{-3} and 10^3 respectively, the concentration of monoprotic acid was set to 1 and the solvent concentration was set to 0, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 0.5, 1, 5, 10^1, 10^2, 10^3 to simulate 12 different $log_{10}D$, profiles. Only five of the solvent profiles corresponding to 0, 1, 10, 100, 1000 were plotted in (B) and (D). The simulated profiles were fitted to the monoprotic model, $D = \left(\frac{P_{HA}\,[H_a^+] + P_A\,K_A}{[H_a^+] + K_A}\right)$, to obtain the apparent $P_{HA}{}'$, $P_A{}'$and $pK_A{}'$, which are plotted against solvent concentration expressed in logarithm scale.

Figure 4 Experimental $log_{10}D$ analysis of (A) Nalidixic acid, (B) Mebendazole, (C) Benazepril and (D) Telmisartan using monoprotic, monoalkaline, simple amphoteric and diamino-monoprotic amphotheric models, respectively. The data were fitted using an in-house written matlab code [16].

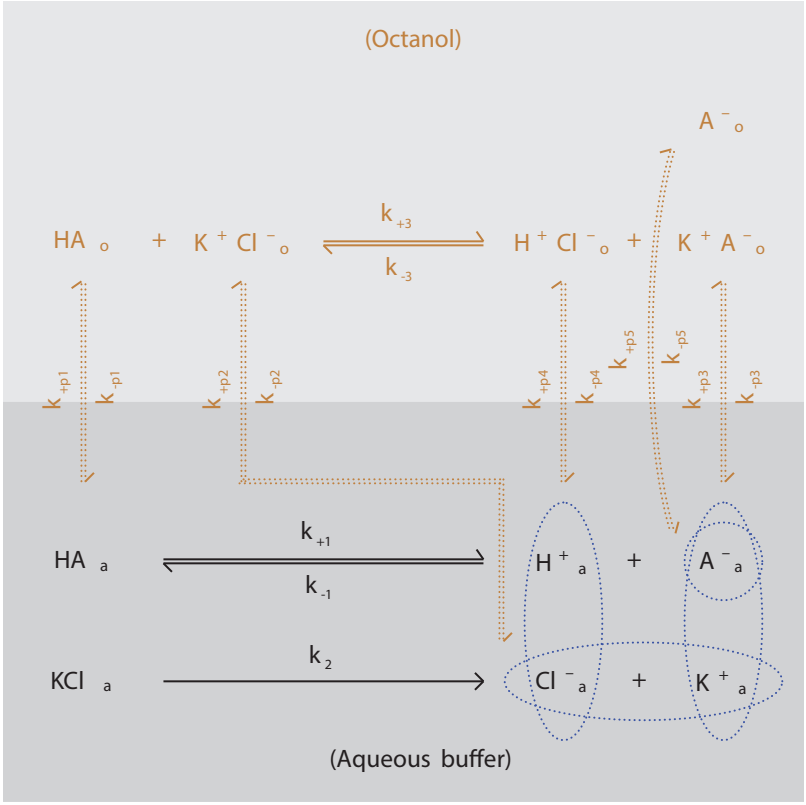
A


(Octanol)
HAB o
HABH o +
⁻AB o
k -p2
k p2
k +p1
k -p1
k +p4
k -p4
⁻ABH o +
HAB a
k +1
k -1
k +4
k -4
HABH a +
⁻AB a
k -3
k +3
k p3
k -p3
k +2
k -2
k +5
k -5
⁻ABH a +
(Aqueous buffer)

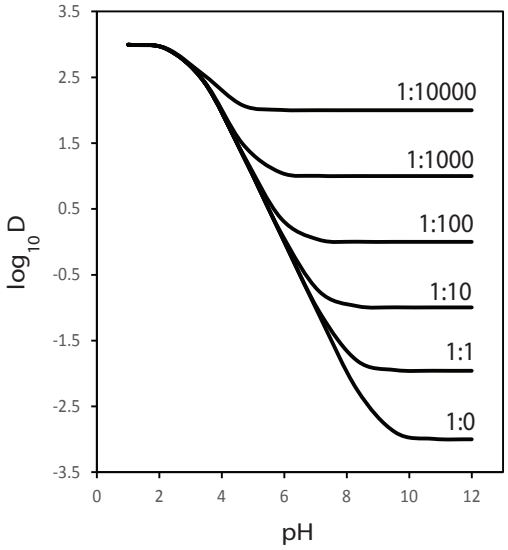
B


5.25
4.75
4.25
3.75
3.25
2.75
log10D
0.5
2.5
4.5
6.5
8.5
10.5
pH

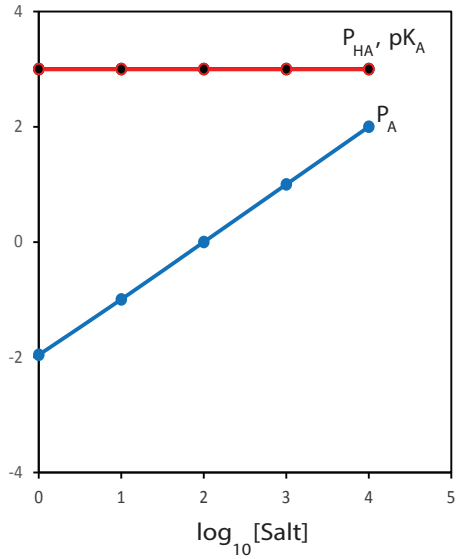
A

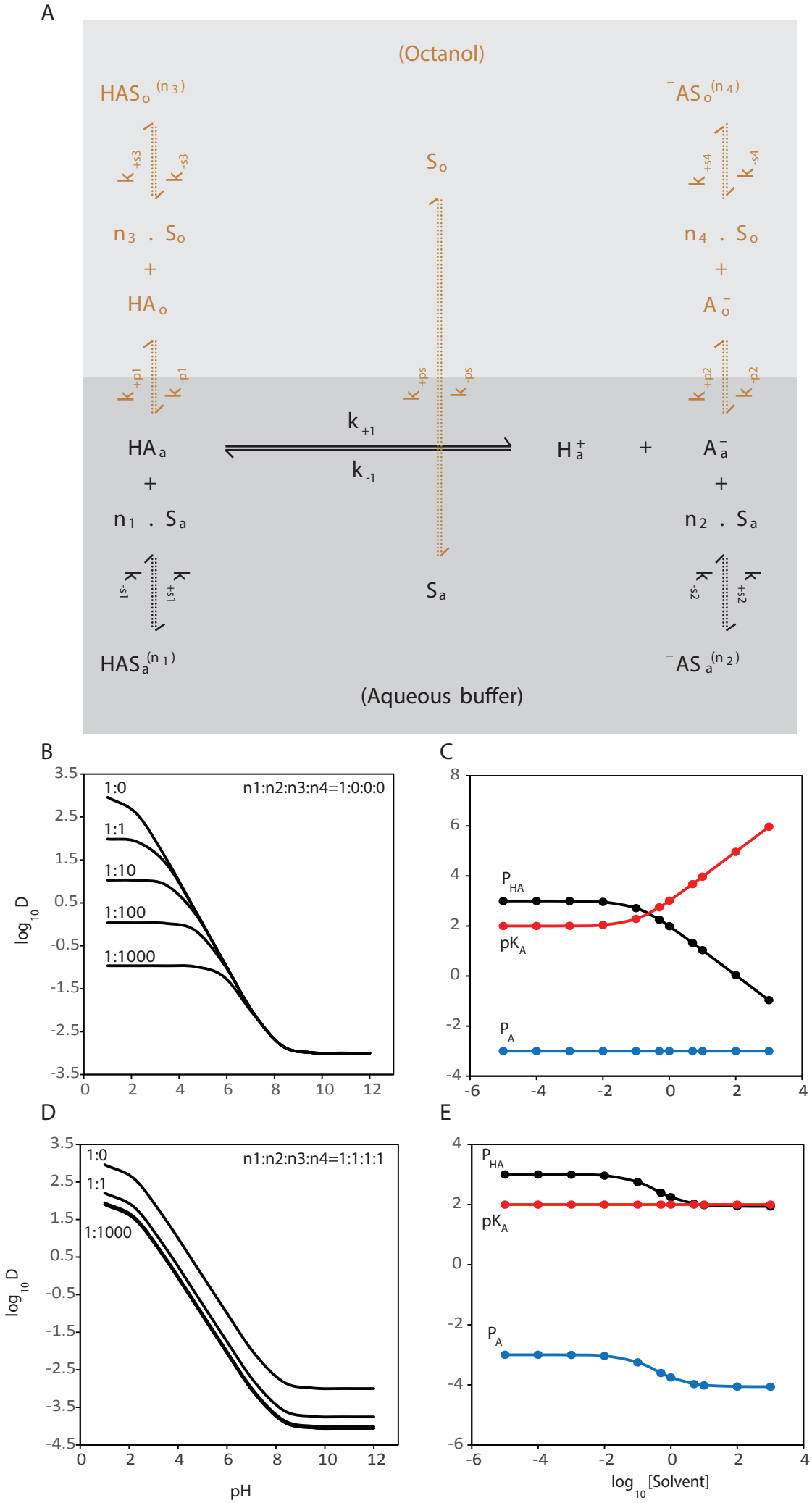

(Octanol)
A⁻o
HAo + K⁺Cl⁻o
k+3
k-3
H⁺Cl⁻o + K⁺A⁻o
k+p1
k-p1
k+p2
k-p2
k+p4
k-p4
k+p5
k-p5
k+p3
k-p3
HAa
k+1
k-1
H⁺a + A⁻a
KClaq
k2
Cl⁻a + K⁺a
(Aqueous buffer)

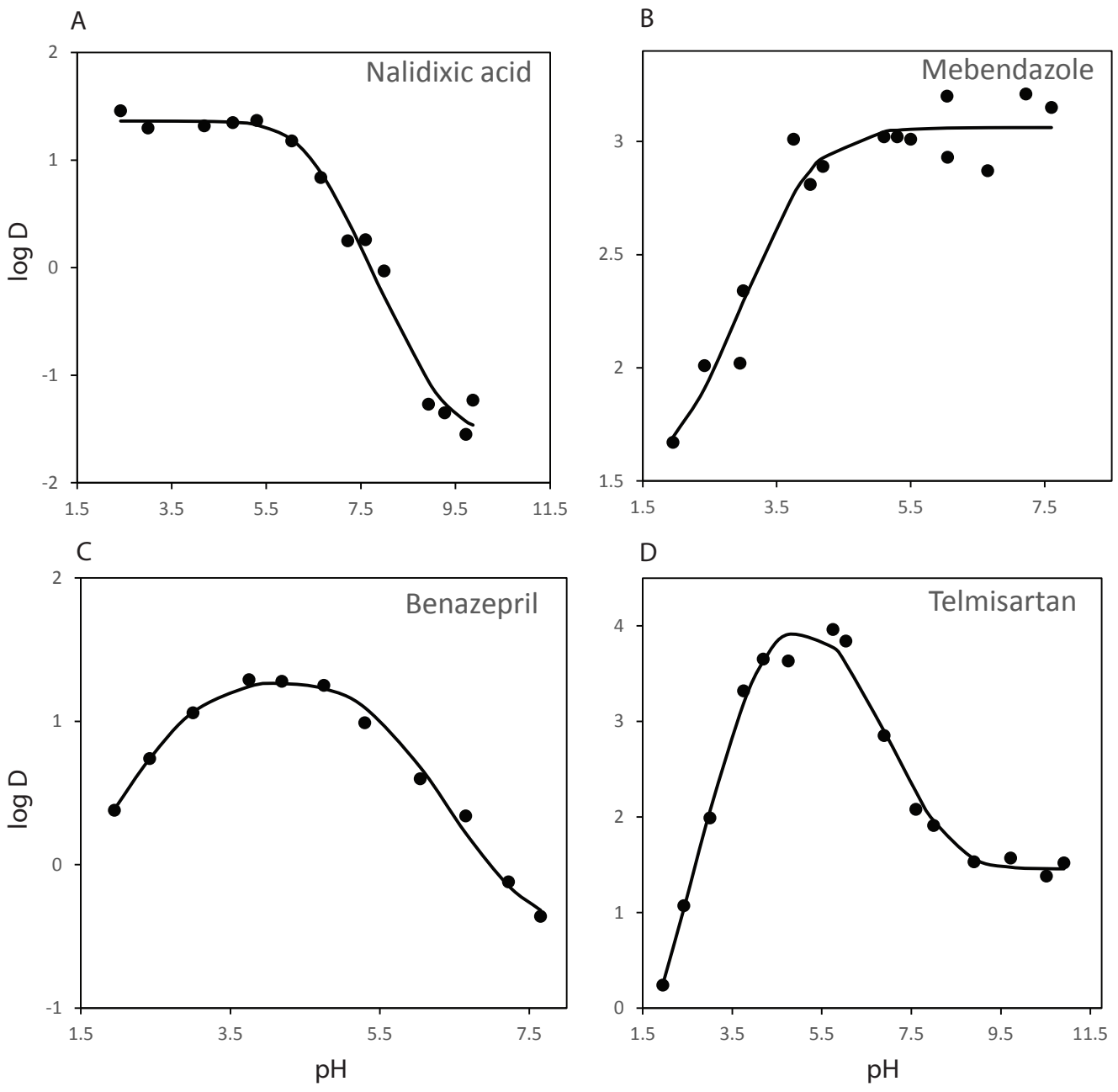
B


log10 D
pH
1:10000
1:1000
1:100
1:10
1:1
1:0

C


PHA, pKA
PA
log10[Salt]


A
(Octanol)
HASo(n3)
−ASo(n4)
So
n3 . So
+
HAo
n4 . So
+
Ao−
HAa
+
n1 . Sa
Ha+
+
Aa−
+
n2 . Sa
Sa
HASa(n1)
−ASa(n2)
(Aqueous buffer)
B
n1:n2:n3:n4=1:0:0:0
1:0
1:1
1:10
1:100
1:1000
log10 D
C
PHA
pKA
PA
D
n1:n2:n3:n4=1:1:1:1
1:0
1:1
1:1000
log10 D
pH
E
PHA
pKA
PA
log10 [Solvent]


A
Nalidixic acid
log D
2
1
0
-1
-2
1.5
3.5
5.5
7.5
9.5
11.5
B
Mebendazole
3
2.5
2
1.5
1.5
3.5
5.5
7.5
C
Benazepril
log D
2
1
0
-1
1.5
3.5
5.5
7.5
pH
D
Telmisartan
4
3
2
1
0
1.5
3.5
5.5
7.5
9.5
11.5
pH