Estimation of the octanol–water distribution coefficient of acidic compounds by microemulsion electrokinetic chromatography

Alejandro Fernández-Pumarega[1], Susana Amézqueta[1,*], Elisabet Fuguet[1,2], Martí Rosés[1]

[1]Departament d'Enginyeria Química i Química Analítica and Institut de Biomedicina (IBUB), Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, 08028, Barcelona, Spain

[2]Serra Húnter Programme. Generalitat de Catalunya. Spain

* Corresponding author: Susana Amézqueta Pérez

E-mail: samezqueta@ub.edu

Departament d'Enginyeria Química i Química Analítica and Institut de Biomedicina (IBUB), Facultat de Química, Universitat de Barcelona

c/ Martí i Franquès 1-11, 08028, Barcelona, Spain

Phone: (+34) 934021277

Fax: (+34) 934021233

Abstract

The feasibility of extending the determination of the lipophilicity of partially ionized acids (log $D_{o/w}$) by microemulsion electrokinetic chromatography (MEEKC) is tested. Theoretical considerations predict that a linear log $D_{o/w}$ vs. log k correlation can be obtained only when the neutral and ionic forms of an acid follow the same correlation equation and the slope of the correlation is unity. In practice, since the lipophilicity of the neutral acid is much higher than that of the ionic form and the correlation slope is not very different from 1, the general linear correlation for neutral compounds can be applied across most of the ionization range of the acid.

The linear correlation between log $P_{o/w}$ and log k of 20 neutral solutes has been established and extended to 6 acids used as models, tested across their full ionization range. log $D_{o/w}$-pH, and log k-pH profiles have been obtained for these 6 acids, and plotted log $D_{o/w}$ against log k for any acid at any degree of ionization. Furthermore, the log $D_{o/w}$ of the acids has been estimated from the calibration curve and log k-pH profile, and compared to values in the literature determined using reference methods such as the shake-flask one. Accurate values have been obtained using the MEEKC method when the acids are in their neutral form or partially ionized ($\alpha < 0.995$). However, this parameter is overestimated when the acids are highly or fully ionized (ionization degree, $\alpha \approx 1$). Finally, in order to test the applicability of this method, we have applied the same procedure to estimate log $D_{o/w}$ at pH=7.4 (blood physiological pH) of a set of 30 additional compounds (including partially and fully ionized acids). The results at this pH follow the same trend observed in the 6 model acids, and validate the application of the method for $D_{o/w}$ determination, except when α is very close to 1.

1. Introduction

The drug development process is a lengthy procedure that is expected to conclude with the release of a new drug onto the market. The process begins with thousands of drug candidates whose physicochemical properties are tested and evaluated, to select the most promising for preclinical and clinical testing [1]. One of the most important properties evaluated in the drug discovery process is the capacity of a compound, once in the bloodstream, to penetrate biological membranes, constituted mainly of lipid bilayers. This biological property is clearly related to lipophilicity, which can be described as the easiness of a compound to be dissolved in fats or non-polar solvents [2].

To evaluate the lipophilicity of a substance, the most widely accepted parameter is the octanol–water partition coefficient ($P_{o/w}$). The octanol–water system is widely used because of the similarity of 1-octanol to lipids (it contains a polar head and a hydrophobic chain), and its low water saturation [3].

The octanol–water partition coefficient (for neutral and fully ionized compounds) can be directly determined through the shake-flask procedure [4]. This method measures the ratio of the concentrations of the test solute in the two immiscible phases at equilibrium, according to Eq. 1, where $C_{n\text{-octanol}}$ is the concentration of the compound in the organic phase, and C_{water} is that in the aqueous phase:

$$P_{o/w} = \frac{C_{n-octanol}}{C_{water}} \qquad \text{Eq. 1}$$

For partially ionized compounds, this parameter is called the octanol–water distribution coefficient ($D_{o/w}$). In this case, the coefficient is determined using the same analytical procedure but at a buffered p*H*.

Although the shake-flask method provides a direct way to determine the log $P_{o/w}$ value, it is a tedious and time-consuming procedure. Furthermore, it is not automated. These drawbacks have led to the development of alternative methods. The most popular for acid–base compounds is the potentiometric dual phase p*H*-metric titration technique [5]. Here, the partition coefficient is calculated by considering the differences between the pK_a values obtained for two titrations performed in the presence and absence of octanol [6,7].

In the case of chromatography, the lipophilicity of a compound is correlated with the retention factor (*k*) in a chromatographic system [8]. Ishihama *et al.* [9] developed a method capable of estimating the lipophilicity of a compound that uses microemulsion electrokinetic chromatography (MEEKC). Compared to the reference shake-flask method, MEEKC is fast and simple. It is also automated and does not require large amounts or a high purity of the compounds, as it is a separation technique. These characteristics make it an ideal tool for routine analysis in the drug discovery process.

In the MEEKC technique, based on capillary electrophoresis (CE), a microemulsion (ME)-based pseudostationary phase (a charged ME with its own electrophoretic mobility) is added to the buffer solution filling the capillary. Then the compounds migrate, depending on the electrophoretic conditions and also on their partition between the ME and the aqueous phase. The ME is composed of oil droplets (the core) which are stabilized by a surfactant and a cosurfactant, whose polar heads are in contact with the aqueous phase and whose apolar tails are orientated towards the inner part of the ME [10]. The ME employed by Ishihama and coworkers [9] contained heptane as the oil (0.82%, w/w), sodium dodecyl sulfate (SDS) as the surfactant (1.44%, w/w), and 1-butanol as the cosurfactant (6.49%, w/w).

Further studies [11–15] have confirmed the correlation between the logarithm of the octanol–water partition coefficient (log $P_{o/w}$) of neutral compounds and the logarithm of the retention factor in MEEKC systems (log k). Abraham *et al.* [11] correlated log $P_{o/w}$ to log k measured in the same ME as Ishihama *et al.*, obtaining the following equation:

$$\log P_{o/w} = 1.542 + 1.276 \log k \qquad R^2 = 0.99;\ SD = 0.096;\ n = 53 \qquad \text{Eq. 2}$$

where R^2 is the determination coefficient, SD the standard deviation, and n the number of compounds.

Subirats *et al.* [12] performed the same procedure but with an ME formed of 1.30% (w/v) SDS, 8.15% (v/v) 1-butanol, 1.15% (v/v) heptane, and 5% (v/v) acetonitrile (at p*H* 7.4, in a 10 mM phosphate buffer). Acetonitrile was added to avoid co-elution of the micellar marker with highly hydrophobic compounds. The resulting correlation was:

$$\log P_{o/w} = 1.48\ (\pm 0.05) + 1.48\ (\pm 0.05) \log k \qquad R^2 = 0.96;\ n = 32 \qquad \text{Eq. 3}$$

Analyzing Eq. 2 and Eq. 3, we can state that the MEEKC method is capable of emulating the octanol–water partition system and of estimating log $P_{o/w}$ of neutral solutes through chromatographic determinations. Nevertheless, the majority of drugs released onto the market are acids or bases, which are partly or fully ionized depending on their pK_a and the medium p*H*. Some works have already estimated the log $P_{o/w}$ of acidic and basic solutes through MEEKC at a p*H* where they are not ionized [12,14]. But there are no studies regarding the estimation of log $D_{o/w}$ of totally or partially ionized compounds. Therefore, the aim of this work is to broaden the applicability of the MEEKC method to

120 estimate, in addition to log $P_{o/w}$, the log $D_{o/w}$ value of partially and totally ionized species
121 of acidic compounds.
122 Preliminary studies with a single model compound indicated that the MEEKC method
123 might be capable of estimating log $D_{o/w}$ of partially ionized acid–base compounds under
124 certain conditions [16].
125 For the present study, we selected 6 model monoprotic acids with a wide range of log $P_{o/w}$
126 values, and we studied the relationship between their retention factor and lipophilicity
127 when the acids are partially or totally ionized. Then, to validate our results, the log $P_{o/w}$
128 and log $D_{o/w}$ values of a set of 30 solutes (including neutral solutes, and partially and
129 totally ionized acids) at a p*H* equal to 7.4 (blood physiological p*H*) were estimated using
130 the proposed method. Finally, we compared the estimated log $D_{o/w}$ values (in the case of
131 partially ionized acids) and log $P_{o/w}$ values (for the neutral and fully ionized species) with
132 values reported in the literature determined using classical methods (mainly shake-flask
133 and potentiometric methods).
134
135 **2. Theory**
136 *2.1. Estimation of log $P_{O/W}$ from MEEKC retention factors. Feasibility of the extension to*
137 *partially ionized compounds.*
138 Previous work [11,12] has shown a linear relationship between log $P_{o/w}$ and log k for
139 neutral compounds (Eqs. 2 and 3), which for a neutral acid, HA, can be generalized as:
140
141 $$\log P_{\mathrm{o/w(HA)}} = q_{(HA)} + p_{(HA)} \log k_{\mathrm{(HA)}} \qquad \text{Eq. 4}$$
142

where log $P_{o/w(HA)}$ and log $k_{(HA)}$ are the logarithms of the octanol–water partition coefficient and the retention factor of fully protonated acids, respectively; and $q_{(HA)}$ and $p_{(HA)}$ are the intercept and the slope of Eq. 4, respectively.

Extension of this equation to ionic or ionizable compounds is not straightforward. To the best of our knowledge, a similar (linear) relationship has not yet been established for ionic compounds. Even if such a linear relation exists, it will probably not have the same parameters as for neutral compounds, and we should write it as:

$$\log P_{\mathrm{o/w(A-)}} = q_{(A-)} + p_{(A-)} \log k_{\mathrm{(A-)}} \qquad \text{Eq. 5}$$

where log $P_{o/w(A-)}$ and log $k_{(A-)}$ are the logarithms of the octanol–water partition coefficient and the retention factor of the fully ionized acid, respectively; and $q_{(A-)}$ and $p_{(A-)}$ are the intercept and the slope of Eq. 5, respectively.

For partially ionized compounds, the relationship is even more complex.

Both, the retention factor and the octanol–water distribution coefficient of the compound can be computed from the degree of ionization of the compound (which can be easily calculated from the pH of the medium and the pK_a' of the solute, the apparent acidity constant, and the k or $P_{o/w}$ value of the pure species, according to Eqs. 6 and 7:

$$k = (1 - \alpha)\, k_{(HA)} + \alpha\, k_{(A-)} \qquad \text{Eq. 6}$$

$$D_{o/w} = (1 - \alpha)\, P_{o/w(HA)} + \alpha\, P_{o/w(A-)} \qquad \text{Eq. 7}$$

where:

$$\alpha = \frac{10^{pH-pK'_a}}{1+10^{pH-pK'_a}} \quad \text{Eq. 8}$$

which lead to the well-known equations for the *k* vs. p*H* and $D_{o/w}$ vs. p*H* profiles:

$$k = \frac{k_{(HA)}+k_{(A-)}\cdot 10^{\mathrm{p}H-pK'_a}}{1+10^{\mathrm{p}H-pK'_a}} \quad \text{Eq. 9}$$

$$D_{o/w} = \frac{P_{o/w(HA)}+P_{o/w(A-)}\cdot 10^{\mathrm{p}H-pK'_a}}{1+10^{\mathrm{p}H-pK'_a}} \quad \text{Eq. 10}$$

or in their logarithmic forms:

$$\log k = \log \left(\frac{10^{\log\, k_{(HA)}}+10^{\log\, k_{(A-)}}\cdot 10^{\mathrm{p}H-pK'_a}}{1+10^{\mathrm{p}H-pK'_a}}\right) \quad \text{Eq. 11}$$

$$log\ D_{o/w} = \log \left(\frac{10^{\log\, P_{o/w(HA)}}+10^{\log\, P_{o/w(A-)}}\cdot 10^{\mathrm{p}H-pK'_a}}{1+10^{\mathrm{p}H-pK'_a}}\right) \quad \text{Eq. 12}$$

Combining Eqs. 4, 5, and 7, we obtain the general relationship between $D_{o/w}$ and *k* (Eq. 13).

$$D_{o/w} = (1 - \alpha)\ 10^{q_{(HA)}}\ k_{(HA)}{}^{p_{(HA)}} + \alpha\ 10^{q_{(A-)}}\ k_{(A-)}{}^{p_{(A-)}} \quad \text{Eq. 13}$$

It is evident that a linear relationship of the type:

$$\log D_{o/w} = q + p \log\ k \quad \text{Eq. 14}$$

can be obtained only when $q_{(HA)} = q_{(A-)} = q$, and $p_{(HA)} = p_{(A-)} = p = 1$. That is to say, we would expect a linear correlation between log $D_{o/w}$ and log k for acids at any pH only when the correlation for ionic and neutral compounds are the same, but also only when the slope of the correlation is close to one. Figure 1 shows an example of the variation of the lipophilicity (log $D_{o/w}$ – log $P_{o/w(HA)}$), Eq. 15, with the variation of the retention factor (log k – log $k_{(HA)}$), Eq. 16, at different degrees of ionization for a compound with a $k_{(HA)}/k_{(A-)}$ ratio of 20. The effect of several representative p values (0.5, 1.0, 1.5, and 2.0) is shown. The plots presented are easily derived from Eqs. 4, 6 and 13, assuming $q_{(HA)} = q_{(A-)}$, $p_{(HA)} = p_{(A-)} = p$, and providing values of α.

$$\log D_{o/w} - \log P_{o/w(HA)} = \log\left[(1-\alpha) + \alpha\left(\frac{k_{(A-)}}{k_{(HA)}}\right)^{p}\right] \quad \text{Eq. 15}$$

$$\log k - \log k_{(HA)} = \log\left[(1-\alpha) + \alpha\left(\frac{k_{(A-)}}{k_{(HA)}}\right)\right] \quad \text{Eq. 16}$$

The same tendency is observed for other $k_{(HA)}/k_{(A-)}$ ratios (data not shown). The relationship is completely linear only for $p = 1$, but it is close to linearity across a wide range of p values (mainly between 0.5 and 1.5, see for example Eqs. 2 and 3). The plot deviates from linearity for low values of log k (or log $D_{o/w}$), i.e., when the solute is highly ionized and the slope of the correlation is not equal to 1.

The deviation from linearity is produced when the contribution of the ionized form of the acid to log $D_{o/w}$ is significant. Since log $P_{o/w(A-)}$ is much lower than log $P_{o/w(HA)}$ [17], the relationship between log $D_{o/w}$ and log k is close to that given by Eq. 4 for most of the ionization range. Figure 2 represents the effect of the lipophilicity of the ionized form (log $P_{o/w(A-)}$) on the linearity of the plot for a typical acid with log $P_{o/w(HA)} = 4$. If log $P_{o/w(A-}$

215 $_{)}$ = 3, linearity is lost when α > 0.67; but as log $P_{o/w(A-)}$ decreases, the α value increases,
216 being approximately 0.86, 0.96, and 0.998 for log $P_{o/w(A-)}$ values of 2, 1, and 0,
217 respectively. The actual difference between log $P_{o/w(HA)}$ and log $P_{o/w(A-)}$ can be between
218 1.5 and 4.5 log units, depending on the structure of the compound and measurement
219 conditions, with a mean of about 3.15 [17]. Thus, we expect log $D_{o/w}$ vs. log k to be linear
220 for a wide range of degrees of ionization. The extent of this range is tested in the
221 experimental part of the present work.
222
223 **3. Experimental section**
224 *3.1 Equipment*
225 To perform the electrophoretic measurements, a CE 7100 system equipped with a diode
226 array from Agilent Technologies (Santa Clara, CA, USA) was used. Fused-silica
227 capillaries from Polymicro Technologies (Phoenix, AZ, USA), with an effective and total
228 length of 30 cm and 38.5 cm, respectively, were used.
229 A GLP 22 pH meter from Crison (Barcelona, Spain) was used to measure the p*H* of the
230 buffer solutions; and an ultrasonic bath from JP Selecta (Abrera, Spain) to favor the
231 dissolution of some substances.
232
233 *3.2 Reagents*
234 Sodium dihydrogen phosphate monohydrate (≥99%), dimethyl sulfoxide (≥99.9%),
235 hydrochloric acid (Tritisol^{TM} 1 N), ammonium chloride (>99.8%), and sodium hydroxide
236 (Tritisol^{TM} 0.5 N) were from Merck (Darmstadt, Germany). Methanol (HPLC grade) was
237 from Thermo Fisher Scientific (Waltham, MA, USA). Sodium dodecyl sulfate (SDS,
238 ≥99%), 1-butanol (≥99.7%), heptane (99%), sodium phosphate dodecahydrate (>98%),

239 and dodecanophenone (98%) were from Sigma-Aldrich (St. Louis, MO, USA). Disodium
240 hydrogen phosphate (99.5%) was from Baker (Phillipsburg, NJ, USA).
241 The solutes tested were of high purities and were acquired from Sigma-Aldrich, Baker,
242 Merck, Carlo Erba (Milan, Italy), Fluka (St. Louis, MO, USA), Acros Organics (Geel,
243 Belgium), and Riedel-de Haën (Seelze, Germany).
244 Water was purified using a Milli-Q plus system from Millipore (Burlington, MA, USA).
245
246 *3.3 Buffer solutions*
247 The buffer with a p*H* equal to 7.4 was prepared by mixing 0.2 M sodium dihydrogen
248 phosphate and 0.2 M disodium hydrogen phosphate solutions. The p*H* 11.5 buffer was
249 prepared by mixing 0.2 M disodium hydrogen phosphate and 0.2 M sodium phosphate
250 dodecahydrate solutions. The buffer with a p*H* of 9.5 was prepared by adding 0.5 M
251 sodium hydroxide to a 0.05 M ammonium chloride solution. The ionic strength of all the
252 buffers was 0.05 M.
253
254 *3.4 ME preparation*
255 MEs were prepared following the procedure described elsewhere in the literature [18].
256 The concentrations of each component with respect the total volume of the ME were:
257 1.30% (w/v) SDS, 8.15% (v/v) 1-butanol, and 1.15% (v/v) heptane.
258
259 *3.5 Analysis conditions*
260 We measured mobility by applying 13-14 kV. Detection was performed at λ=200, 214,
261 or 254 nm, depending on the chromophores of each compound. Injection was
262 hydrodynamic, and a pressure of 50 mbar was applied for 5s.

The compounds analyzed were dissolved in an ME:methanol solution (9:1) for MEEKC analysis, and in a water:methanol solution (9:1) for capillary zone electrophoresis (CZE) measurements, at a concentration of 200 mg L^{-1}. Dodecanophenone (200 mg L^{-1}) and dimethyl sulfoxide (0.2% v/v) were added to the test compound vials as ME and electroosmotic flow markers, respectively [19].

3.6 Calculation methods

The physicochemical properties of the compounds were obtained from the Bio-Loom database of the BioByte Corporation (Claremont, CA, USA). Retention profiles were fitted with Table Curve 2D from Systat Software Inc. (San Jose, CA, USA). Data calculations were performed using Excel from Microsoft (Redmond, WA, USA).

Mobilities (μ_i) were calculated using the following expression:

$$\mu_i = \left[\frac{1}{t_r} - \frac{1}{t_0}\right]\left[\frac{L_T L_D}{V}\right] \quad \text{Eq. 17}$$

where t_r and t_o are, respectively, the migration times of the analyte and the electroosmotic flow marker; L_T and L_D are the total and the effective capillary length; and V is the applied voltage.

Retention factors of neutral compounds were calculated from the mobilities of the compound (μ) and ME marker (μ_{ME}) by the well-known Eq. 18:

$$k = \frac{\mu}{\mu_{ME} - \mu} \quad \text{Eq. 18}$$

Application of Eq. 4 to totally or partially ionized compounds requires subtraction of the mobility of the compound in CZE, i.e., in the buffer without an ME (μ_0), according to Eq. 19:

$$k = \frac{\mu - \mu_0}{\mu_{ME} - \mu} \quad \text{Eq. 19}$$

It is not feasible to reproduce the MEEKC system without an ME and so μ_0 is usually measured in an aqueous solution with the same p*H* buffer. However, the MEEKC medium contains not only the aqueous buffer but also the surfactant, the co-surfactant and the oil, which usually have viscosities very different from water. Thus, the presence of ME components changes the viscosity of the medium, leading to inaccurate *k* values, as demonstrated in previous work [18]. Note that the viscosity of the electrophoretic medium (η) and the mobility of a compound are inversely related, according to Eq. 20 [20]:

$$\mu = \frac{q}{6\pi\eta r} \quad \text{Eq. 20}$$

where q is the charge of the ion, and r its radius.

We have proposed a viscosity correction obtained from the mobility of an ion that does not interact with the ME (the benzoate ion). The ratio of the mobilities of the benzoate ion in the ME and plain buffer ($\left(\frac{\mu}{\mu_0}\right)_{benzoate\ ion}$) is equivalent to the ratio of viscosities, and the mobility of any other ion in a plain buffer can be corrected for the viscosity changes without the need to measure the viscosities. We then determined the correct retention factor from the following equation:

$$k = \frac{\mu - \left(\frac{\mu}{\mu_0}\right)_{benzoate\ ion} \cdot \mu_0}{\mu_{ME} - \mu} \qquad \text{Eq. 21}$$

where $\left(\frac{\mu}{\mu_0}\right)_{benzoate\ ion}$ has a value of 0.76, and was measured at p*H* 11.0, when benzoic acid is fully ionized.

4. Results and discussion

4.1 log $P_{o/w}$ vs. log k correlation for neutral compounds

We established the correlation between log $P_{o/w}$ and log *k* of 20 neutral compounds, that present known and uniformly distributed log $P_{o/w}$ values [21–26] for the MEEKC system studied and the equation resulting from this correlation is:

$$\log P_{o/w} = 1.51\ (\pm 0.08) + 1.60\ (\pm 0.11) \log k \qquad \text{Eq. 22}$$

$R^2 = 0.916$; SD = 0.33; n = 20; F = 196

where F is Fisher's F parameter.

A graphical representation of this correlation can be seen in Figure 3; while the solutes, and their log $P_{o/w}$ and log *k* values are shown in Table 1. We obtained a good correlation, similar to that obtained for other SDS-MEEKC systems (Eqs. 2 and 3).

4.2 Influence of the degree of ionization on the estimation of log $D_{o/w}$

The compounds chosen to perform this study were benzoic acid, 3-bromobenzoic acid, naproxen, ketoprofen, ibuprofen, and 2,4,6-trichlorophenol. We selected them because they have known and well-defined lipophilicity-p*H* profiles (data provided in the

supplementary information). Moreover, they have pK_a values in the working pH range, and are detectable by UV-vis.

Experimental k-pH profiles of the compounds were taken from a previous study [18] and the log k-pH profiles were obtained by fitting the data to Eq. 11. Values of log $k_{(HA)}$, log $k_{(A^-)}$, and pK_a', as well as the statistics from the fits, are presented in Table 2. In addition, Figure 4 offers a graphical representation of the profiles.

A similar procedure was followed when fitting log $D_{o/w}$ values determined at different pHs from the literature (data provided in the supplementary information) to Eq. 12. Note that data from the literature are obtained in different experimental conditions (nature of the buffers, concentration of the buffers, ionic strength, temperature, etc.) so some discrepancies can be observed, especially at low pH values, where ionic-pairs between ionized acids and buffer components can be formed. Experimental conditions of literature data are also provided in Table SI-1 of the supplementary information. In the fits of the log $D_{o/w}$-pH profiles, pK_a' was fixed using the values obtained previously in the log k-pH profiles (Table 2). The estimated log $D_{o/w}$ vs. pH profiles are shown in Figure 5, while the parameters and statistics resulting from these are in Table 3. The log $D_{o/w}$-pH fits have small SD, and high R^2 and F values for all the compounds studied.

In Figure 6, the log $P_{o/w}$ values of the neutral and fully ionized species of the 6 model acids are plotted against the corresponding log k values. It can clearly be seen that all the neutral species lie within the confidence interval of the calibration curve (Eq. 22). However, this is not the case for the fully ionized species, hence their estimation using this equation is not accurate.

Next, log $D_{o/w}$ and log k for all the acids have been determined at each degree of ionization through the log k-pH and log $D_{o/w}$-pH profiles (data from Tables 2 and 3). For a given degree of ionization, log $D_{o/w}$ is graphically represented against log k (Figure 7a). The

same figure also represents the neutral calibration curve obtained in the previous section, and two extra lines corresponding to the calibration curve ±2 SD (which corresponds to the 95% confidence interval). Almost the entire set of log $D_{o/w}$ - log k values fall within this range, except for the lowest values that correspond to the highly or fully ionized species of the acids. Then we estimated log $D_{o/w}$ from the calibration curve and log k at each degree of ionization of the acid (log $D_{est.}$). The differences between the log $D_{o/w}$ values in the literature (log D_{lit}, Table 3) and the log D_{est} was calculated and plotted as a function of the ionization degree (Figure 7b).

Our results show that accurate estimates are obtained if the acid is either in its neutral form or partially ionized. In both cases, the precision of the results is similar to that reported previously for neutral compounds (Section 4.1, SD=0.33). However, the present method overestimates the $D_{o/w}$ of the highly or fully ionized species of the acids ($\alpha \geq$ 0.995). It must be noted that log $P_{o/w(A-)}$ varies depending on the capacity of the ionized compound to form ion pairs with the ions of the buffer [17,27]. Thus, literature log $P_{o/w(A-)}$ data may differ if the buffer and the conditions used in their measurements are different (such as different concentrations of buffers or nature of counter-ions, among other possibilities). Nonetheless, in all the cases our method overestimated its value. Another reason for the observed differences could be the lower retention of ionized species in MEEKC compared to that of neutral ones, which causes higher experimental error in μ measurement. Also, the larger surface between the aqueous and the lipid phase in MEEKC, compared to the classical octanol–water partition system, may lead to a higher partition into the ME than in the octanol–water system.

4.3 Estimation of lipophilicity at physiological pH (log $D_{o/w(7.4)}$)

To validate our method, we estimated the log $D_{o/w}$ value at the blood physiological pH (log $D_{o/w(7.4)}$) for ten neutral compounds, ten partially ionized acids, and ten totally ionized acids. Their log $D_{o/w(7.4)}$ values were estimated directly from their log k values using the calibration curve (Eq. 22). In the case of partially ionized acids, α was also measured via:

$$\alpha = \frac{\mu_{7.4}}{\mu_{(A-)}} \qquad \text{Eq. 23}$$

where, $\mu_{7.4}$ is the electrophoretic mobility of the compound in CZE at a pH value of 7.4, and $\mu_{(A-)}$ is the electrophoretic mobility of the fully ionized compound in CZE. The $\mu_{(A-)}$ value of these compounds was determined at a pH where the fully ionized form was present (pH = 11.5, except for phenobarbital, for which it was measured at pH = 9.5, as it has other acid–base groups that can be ionized at pH = 11.5).

Table 4 shows the log $D_{o/w(7.4)}$ values of the 30 additional substances for which it was estimated using the present method, and their comparison with values reported in the literature determined using classical methods (mostly the shake-flask procedure) (data provided in the supplementary information). The log D_{lit} values were measured under different experimental conditions and usually at room temperature. Due to their variability, reported values that differ considerably from the rest of the published data were excluded, and they were not used to obtain the average value. In the case of pentachlorophenol, log $D_{o/w(7.4)}$ is not available. However, the compound is fully ionized at this pH value, so the log D_{lit} value is determined as an average of the log $D_{o/w}$ values determined at pH values higher than 7.4. As previously for the 6 model acid, log $D_{o/w(7.4)}$ vs. log k is represented for all the compounds together with the calibration curve (Eq. 22). Furthermore, we calculated log D_{lit} – log D_{est} and plotted it against the degree of ionization (Figure 8). The values obtained via the MEEKC measurements are similar to those in the

literature when the compounds are neutral or partially ionized (presenting differences of less than two times the SD from the calibration curve). However, when the compound is highly or fully ionized ($\alpha \approx 1$) larger differences are obtained between the estimated data and those reported in the literature.

5. Concluding remarks

We obtained a linear relationship between log $P_{o/w}$ and log k for neutral compounds that is not very different from those reported in literature for similar systems. Although theory predicts that accurate log $D_{o/w}$ estimation of partially ionized acids can only be performed when the slope of the calibration curve is equal to 1.0, in practice the linear correlation can be extended to most of the ionization range of the acids tested. Therefore, it is possible to estimate $D_{o/w}$ of partially ionized acids with only the determination of k at the pH value of interest. With the MEEKC method the log $D_{o/w}$ value of an acid can be estimated with an error equivalent to that of neutral compounds for a degree of ionization up to 0.995, which corresponds to a pH of pK_a+2. However, our method overestimates the lipophilicity of highly or fully ionized acids ($\alpha \approx 1$).

ACKNOWLEDGEMENTS

Financial support from the Spanish *Ministerio de Economía y Competitividad* (CTQ2017-88179-P) and the Catalan Governments (2017SGR1074) is acknowledged. AFP wishes to thank the University of Barcelona for his APIF PhD fellowship.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

432

REFERENCES

[1] FDA, Learn about drug and device approvals. https://www.fda.gov/ForPatients/Approvals/default.htm, 2018 (accessed 25 October 2019).

[2] A. McNaught, A. Wilkinson, Compendium of chemical terminology, Second ed., Blackwell science, Oxford, 1997.

[3] C. Giaginis, A. Tsantili-Kakoulidou, Alternative measures of lipophilicity: From octanol-water partitioning to IAM retention, J. Pharm. Sci. 97 (2008) 2984–3004. doi:10.1002/jps.21244.

[4] OECD, Test No. 107: partition coefficient (n-octanol/water): shake flask method. http://www.oecd-ilibrary.org/environment/test-no-107-partition-coefficient-n-octanol-water-shake-flask-method_9789264069626-en, 1995 (accessed 25 October 2019).

[5] OECD, Partition coefficient (n-octanol/water), pH-metric method for ionisable substances. http://www.oecd.org/chemicalsafety/testing/2731134.pdf, 2000 (accessed 25 October 2019).

[6] A. Avdeef, pH-metric log P. Part 1. Difference plots for determining ion-pair octanol-water partition coefficients of multiprotic substances, Quant. Struct. Relationships. 11 (1992) 510–517. doi:10.1002/qsar.2660110408.

[7] U. Franke, A. Munk, M. Wiese, Ionization constants and distribution coefficients of phenothiazines and calcium channel antagonists determined by a pH-metric method and correlation with calculated partition coefficients, J. Pharm. Sci. 88 (1999) 89–95. doi:10.1021/js980206m.

[8] OECD, Test No. 117: partition coefficient (n-octanol/water), HPLC method. http://www.oecd-ilibrary.org/environment/test-no-117-partition-coefficient-n-

458 octanol-water-hplc-method_9789264069824-en, 2004 (accessed 25 October
459 2019).
460 [9] Y. Ishihama, Y. Oda, K. Uchikawa, N. Asakawa, Evaluation of solute
461 hydrophobicity by microemulsion electrokinetic chromatography, Anal. Chem. 67
462 (1995) 1588–1595. doi:10.1021/ac00105a018.
463 [10] R. Ryan, K. Altria, E. McEvoy, S. Donegan, J. Power, A review of developments
464 in the methodology and application of microemulsion electrokinetic
465 chromatography, Electrophoresis. 34 (2013) 159–177.
466 doi:10.1002/elps.201200375.
467 [11] M.H. Abraham, C. Treiner, M. Roses, C. Rafols, Y. Ishihama, Linear free energy
468 relationship analysis of microemulsion electrokinetic chromatographic
469 determination of lipophilicity, J. Chromatogr. A. 752 (1996) 243–249.
470 doi:10.1016/S0021-9673(96)00518-3.
471 [12] X. Subirats, H.P. Yuan, V. Chaves, N. Marzal, M. Rosés, Microemulsion
472 electrokinetic chromatography as a suitable tool for lipophilicity determination of
473 acidic, neutral, and basic compounds, Electrophoresis. 37 (2016) 2010–2016.
474 doi:10.1002/elps.201600080.
475 [13] S.K. Poole, D. Durham, C. Kibbey, Rapid method for estimating the octanol –
476 water partition coefficient (log P_{ow}) by microemulsion electrokinetic
477 chromatography, 745 (2000) 117–126.
478 [14] S.K. Poole, S. Patel, K. Dehring, H. Workman, J. Dong, Estimation of octanol-
479 water partition coefficients for neutral and weakly acidic compounds by
480 microemulsion electrokinetic chromatography using dynamically coated capillary
481 columns, J. Chromatogr. B. 793 (2003) 265–274. doi:10.1016/S1570-
482 0232(03)00321-0.

[15] J. Øtergaard, S.H. Hansen, C. Larsen, C. Schou, N.H.H. Heegaard, Determination of octanol-water partition coefficients for carbonate esters and other small organic molecules by microemulsion electrokinetic chromatography, Electrophoresis. 24 (2003) 1038–1046. doi:10.1002/elps.200390120.

[16] A. Fernández-Pumarega, S. Amézqueta, E. Fuguet, M. Rosés, Feasibility of the estimation of octanol-water distribution coefficients of acidic drugs by microemulsion electrokinetic chromatography, ADMET DMPK. 6 (2018) 55–60. doi:10.5599/admet.6.1.510.

[17] S.F. Donovan, M.C. Pescatore, Method for measuring the logarithm of the octanol-water partition coefficient by using short octadecyl-poly(vinyl alcohol) high-performance liquid chromatography columns, J. Chromatogr. A. 952 (2002) 47–61. doi:10.1016/S0021-9673(02)00064-X.

[18] A. Fernández-Pumarega, S. Amézqueta, E. Fuguet, M. Rosés, Determination of the retention factor of ionizable compounds in microemulsion electrokinetic chromatography, Anal. Chim. Acta. 1078 (2019) 221-230. doi:10.1016/j.aca.2019.06.007.

[19] E. Fuguet, C. Ràfols, E. Bosch, M. Rosés, Solute-solvent interactions in micellar electrokinetic chromatography: IV. Characterization of electroosmotic flow and micellar markers, Electrophoresis. 23 (2002) 56–66. doi:10.1002/1522-2683(200201)23:1<56::AID-ELPS56>3.0.CO;2-7.

[20] D.R. Baker, Capillary electrophoresis, Wiley, New York, 1995.

[21] A. Andrés, M. Rosés, C. Ràfols, E. Bosch, S. Espinosa, V. Segarra, J.M. Huerta, Setup and validation of shake-flask procedures for the determination of partition coefficients (log D) from low drug amounts, Eur. J. Pharm. Sci. 76 (2015) 181–191. doi:10.1016/j.ejps.2015.05.008.

[22] S. Winiwarter, N.M. Bonham, F. Ax, A. Hallberg, H. Lennernäs, A. Karlén, Correlation of human jejunal permeability (in vivo) of drugs with experimentally and theoretically derived parameters. A multivariate data analysis approach, J. Med. Chem. 41 (1998) 4939–4949. doi:10.1021/jm9810102.

[23] M. Kansy, H. Fischer, K. Kratzat, F. Senner, B. Wagner, I. Parrilla, High-throughput artificial membrane permeability studies in early drug discovery and development, in: B. Testa, H. Van de Waterbeemd, G. Folkers, R. Guy (Eds.), Pharmacokinetic optimization in drug research: biological, physicochemical, and computational strategies, Verlag Helvetica Chimica Acta, Zürich, 2001: pp. 447–464.

[24] G. Camenisch, G. Folkers, H. Van De Waterbeemd, Comparison of passive drug transport through Caco-2 cells and artificial membranes, Int. J. Pharm. 147 (1997) 61–70. doi:10.1016/S0378-5173(96)04796-5.

[25] F. Lombardo, M.Y. Shalaeva, K.A. Tupper, F. Gao, M.H. Abraham, ElogP(oct): A tool for lipophilicity determination in drug discovery, J. Med. Chem. 43 (2000) 2922–2928. doi:10.1021/jm0000822.

[26] C. Zhu, L. Jiang, T.M. Chen, K.K. Hwang, A comparative study of artificial membrane permeability assay for high throughput profiling of drug absorption potential, Eur. J. Med. Chem. 37 (2002) 399–407. doi:10.1016/S0223-5234(02)01360-0.

[27] C.S. Chen, S.T. Lin, Prediction of pH effect on the octanol-water partition coefficient of ionizable pharmaceuticals, Ind. Eng. Chem. Res. 55 (2016) 9284–9294. doi:10.1021/acs.iecr.6b02040.

Figure captions

Figure 1: Variation of the lipophilicity (log $D_{o/w}$ – log $P_{o/w(HA)}$) with the retention factor (log k – log $k_{(HA)}$) at different degrees of ionization for a hypothetical compound with a $k_{(HA)}/k_{(A-)}$ ratio of 20, and different representative p values: 0.5 (···); 1.0 (—); 1.5 (---); and 2.0 (·-·).

Figure 2: Plots of log $D_{o/w}$ against log k for compounds with a $P_{o/w(HA)}$ of 4 and $P_{o/w(A-)}$ of: 3, (—); 2 (·-·); 1 (---); and 0 (···). We calculated the log k values using Eq. 2.

Figure 3: Plot of log $P_{o/w}$ versus log k for the set of 20 neutral compounds selected for the calibration curve.

Figure 4: log k – pH profiles obtained by fitting the data from [18] to Eq. 11: a) benzoic acid, b) 3-bromobenzoic acid, c) naproxen, d) ketoprofen, e) ibuprofen, f) 2,4,6-trichlorophenol.

Figure 5: log $D_{o/w}$ – pH profiles obtained by fitting the data from the literature to Eq. 12: a) benzoic acid, b) 3-bromobenzoic acid, c) naproxen, d) ketoprofen, e) ibuprofen, f) 2,4,6-trichlorophenol.

Figure 6: Plot of log $P_{o/w}$ against log k for the neutral (□) and fully ionized (◊) species of the 6 model acids; calibration curve (Eq. 22) and ±2 SD (dotted lines) are also plotted.

Figure 7: a) Variation of log $D_{o/w}$ vs. log k at different degrees of ionization. Data for log $D_{o/w}$ and log k are taken from the profiles in Figures 4 and 5. b) Difference between log $D_{o/w}$ in the literature (log D_{lit}) and the value of log $D_{o/w}$ estimated using our present method (log D_{est}) at different degrees of ionization. The calibration curve ±2 SD (Eq. 22) is also plotted. Each line corresponds to one of the six model acids: benzoic acid (—··); 3-bromobenzoic acid (---); naproxen (—); ketoprofen (·- ·); ibuprofen (·—·); and 2,4,6-trichlorophenol (···).

Figure 8: a) Plot of log $D_{o/w(7.4)}$ against log k for a set of compounds including neutral (□), and both partially (○) and completely (◊) ionized acids. b) Difference between the log $D_{o/w(7.4)}$ in the literature (log D_{lit}) and the value of log $D_{o/w(7.4)}$ estimated using our present method (log D_{est}), according to the degree of ionization. The calibration curve ±2 SD (Eq. 22) is also plotted.

574 **Table 1:** Values of log $P_{o/w}$ and log k of the 20 solutes used for the calibration curve.

Compound	**log $P_{o/w}$** [a)]	**log k**
Acetaminophen	0.39	-0.80
Acetanilide	1.19	-0.30
Acetophenone	1.57	-0.05
Antipyrine	0.56	-0.59
Butyrophenone	2.65	0.60
Caffeine	-0.04	-0.89
Carbamazepine	2.45	0.46
Corticosterone	1.90	0.59
Coumarin	1.44	-0.09
Dexamethasone	1.74	0.44
Estradiol	4.01	1.13
Naphthalene	3.19	1.13
Hydrocortisone	1.58	0.30
Hydrocortisone-21-acetate	2.19	0.47
Lormetazepam	2.72	1.03
Prednisolone	1.83	0.32
Progesterone	3.48	1.32
Propiophenone	2.24	0.26
Testosterone	3.31	0.97
Valerophenone	3.40	0.98

575 a) From references [21–26]

576

577

578 **Table 2:** Parameters and statistics obtained from fitting log *k* vs. p*H* through Eq. 11

579 [18]. The standard deviation of each fitted parameter is shown in brackets.

Compound	pK_a'	log $k_{(A^-)}$	log $k_{(HA)}$	R^2	F	SD
Benzoic acid	3.50 (0.15)	-0.69 (0.03)	0.05 (0.04)	0.991	116	0.04
3-Bromobenzoic acid	3.50 (0.14)	-0.55 (0.05)	0.79 (0.06)	0.992	188	0.07
Naproxen	4.43 (0.07)	-0.40 (0.04)	0.92 (0.03)	0.997	532	0.04
Ketoprofen	4.19 (0.07)	-0.33 (0.03)	0.79 (0.02)	0.997	589	0.03
Ibuprofen	4.30 (0.19)	0.21 (0.06)	1.79 (0.12)	0.997	151	0.06
2,4,6-Trichlorophenol	6.28 (0.07)	-0.04 (0.05)	1.16 (0.02)	0.996	510	0.03

580

581 **Table 3:** Parameters and statistics obtained from fitting log $D_{o/w}$ vs. pH through Eq. 12
582 (data taken from the literature). The standard deviation of each fitted parameter is shown
583 in brackets. Experimental conditions of literature data are provided in Table SI-1 of the
584 supplementary information.

Compound	log $P_{o/w(HA)}$	log $P_{o/w(A-)}$	R^2	F	SD
Benzoic acid	2.01 (0.07)	-1.37 (0.09)	0.992	880	0.15
3-Bromobenzoic acid	2.91 (0.04)	-0.44 (0.05)	0.999	2696	0.06
Naproxen	3.15 (0.05)	0.05 (0.10)	0.985	1056	0.16
Ketoprofen	3.10 (0.04)	-1.95 (4.47)	0.994	1604	0.10
Ibuprofen	4.16 (0.10)	0.10 (0.31)	0.894	161	0.41
2,4,6-Trichlorophenol	3.67 (0.08)	1.03 (0.16)	0.950	246	0.27

585

586 **Table 4:** The log $D_{o/w(7.4)}$ values estimated using the present method (log D_{est}) and in the

587 literature (log D_{lit}) for compounds at different degrees of ionization.

Compound	Ionization degree (α)	log $D_{est.}$	log $D_{lit.}$ [a)]	log $D_{lit.}$-log D_{est}
3-nitroaniline	0	1.27	1.39	0.12
Aminopyrine	0	0.86	0.63	-0.23
Benzocaine	0	1.84	1.89	0.05
Bromazepam	0	2.13	1.65	-0.48
Diazepam	0	3.27	2.62 ± 0.28	-0.65
Griseofulvin	0	2.83	2.28 ± 0.13	-0.55
Hexanophenone	0	3.67	3.69	0.02
Isoniazid	0	-0.61	-0.75 ± 0.14	-0.14
Methoxsalen	0	2.02	1.97	-0.05
Thymol	0	3.04	3.34	0.30
Bumetanide	1	1.26	0.10 ± 0.30	-1.16
Diclofenac	1	1.86	1.17 ± 0.07	-0.69
Diflunisal	1	1.75	0.76	-0.99
Fenbufen	1	1.84	0.61 ± 0.03	-1.23
Flurbiprofen	1	1.70	0.89 ± 0.03	-0.81
Gemfibrozil	1	2.16	1.20	-0.96
Glyburide	1	2.21	2.19 ± 0.00	-0.02
Indomethacin	1	2.07	0.98 ± 0.13	-1.09
Mefenamic acid	1	2.20	2.03 ± 0.04	-0.17
Pentachlorophenol	1	2.59	1.83	-0.76
3-Nitrophenol	0.13	1.44	1.52	0.08
4-Nitrophenol	0.61	0.86	1.38	0.52
Butylparaben	0.20	3.23	3.32	0.09
Ethylparaben	0.13	2.03	2.44	0.41
Methylparaben	0.13	1.46	1.98	0.52
Omeprazole	0.10	1.98	2.30 ± 0.11	0.32
Phenobarbital	0.54	0.61	1.12 ± 0.03	0.51
Propylparaben	0.14	2.64	3.01	0.37
sulfamethazine	0.42	-0.56	-0.43	0.13
Theophylline	0.10	-0.33	-0.04 ± 0.01	0.29

588 a) References are provided in the supplementary information

589
590
591
592
593
594
595
596
597
598
599

Figure 1

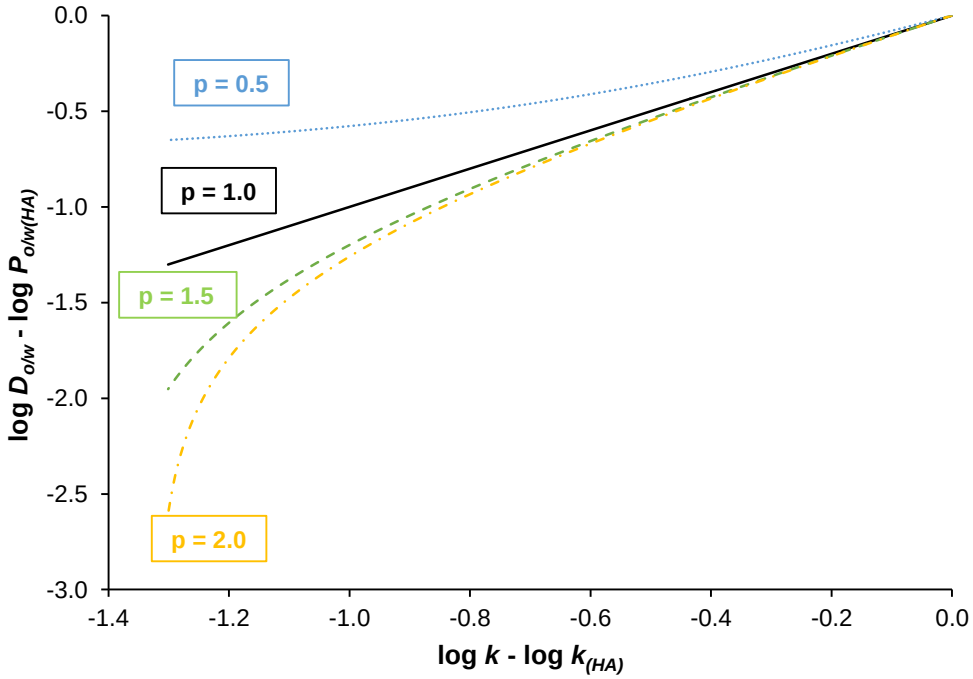

Figure 2

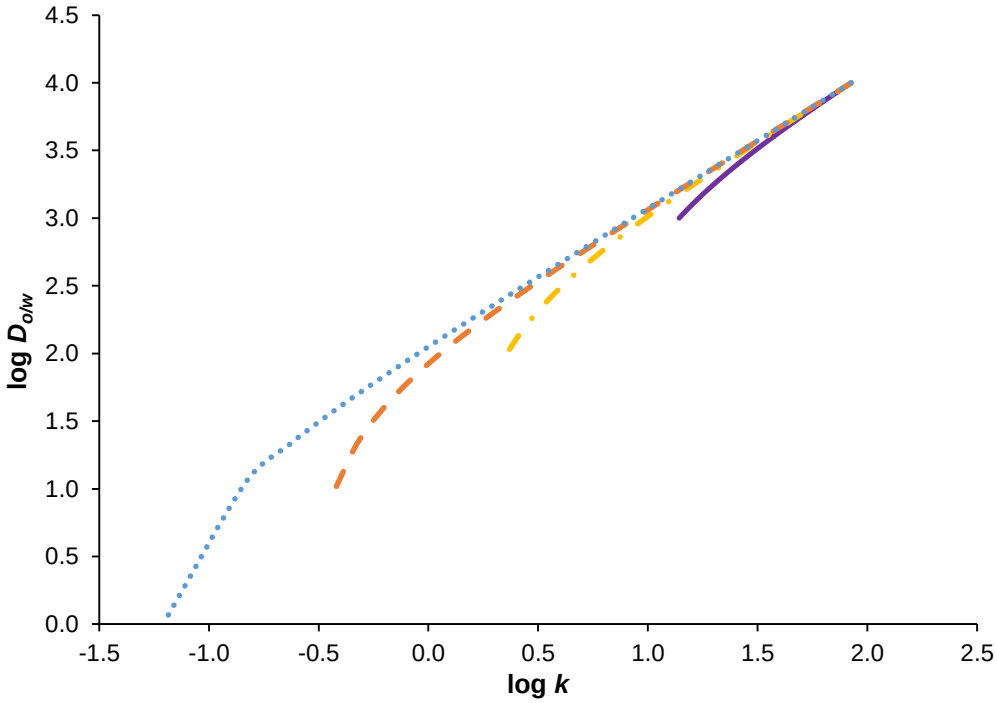

Figure 3

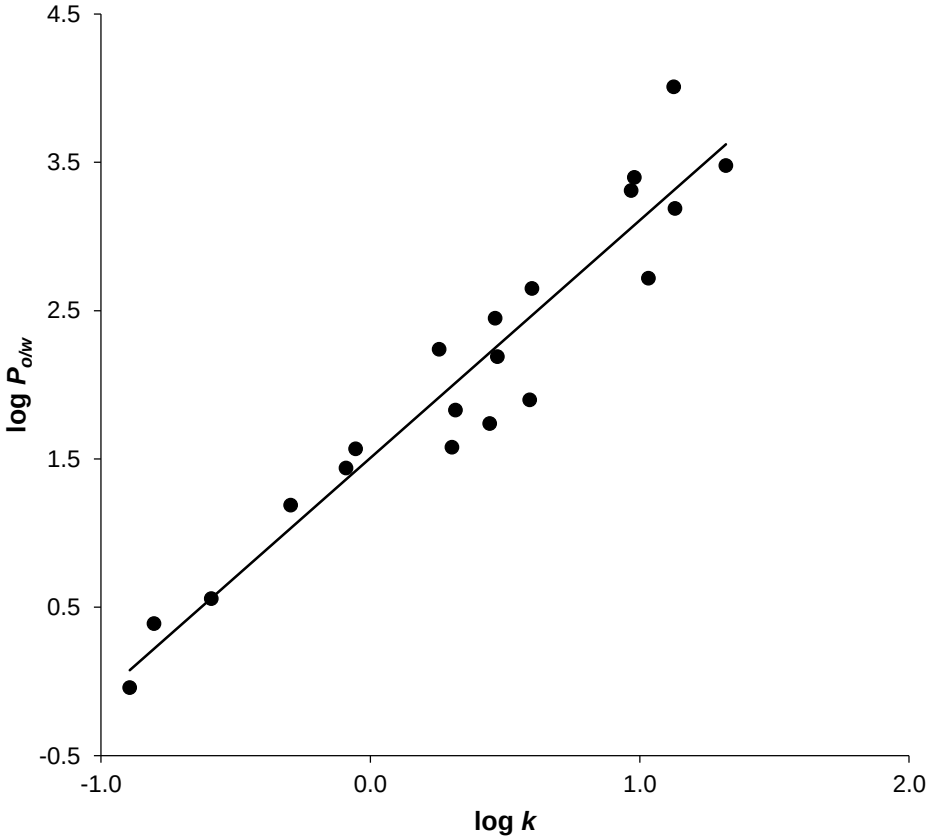

Figure 4

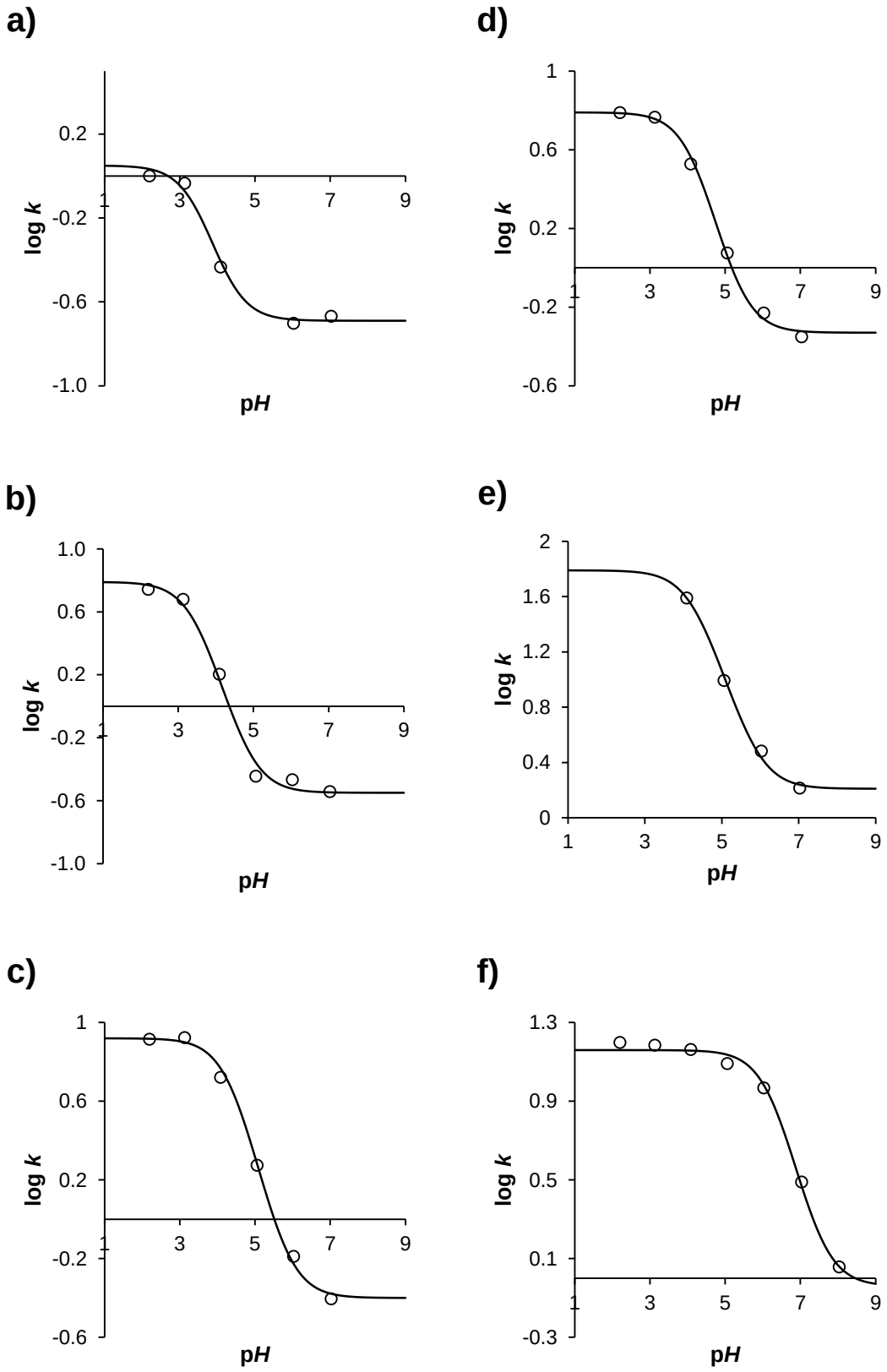
a)
0.2
-0.2
-0.6
-1.0
log k
1
3
5
7
9
pH
d)
1
0.6
0.2
-0.2
-0.6
log k
1
3
5
7
9
pH
b)
1.0
0.6
0.2
-0.2
-0.6
-1.0
log k
1
3
5
7
9
pH
e)
2
1.6
1.2
0.8
0.4
0
log k
1
3
5
7
9
pH
c)
1
0.6
0.2
-0.2
-0.6
log k
1
3
5
7
9
pH
f)
1.3
0.9
0.5
0.1
-0.3
log k
1
3
5
7
9
pH

Figure 5

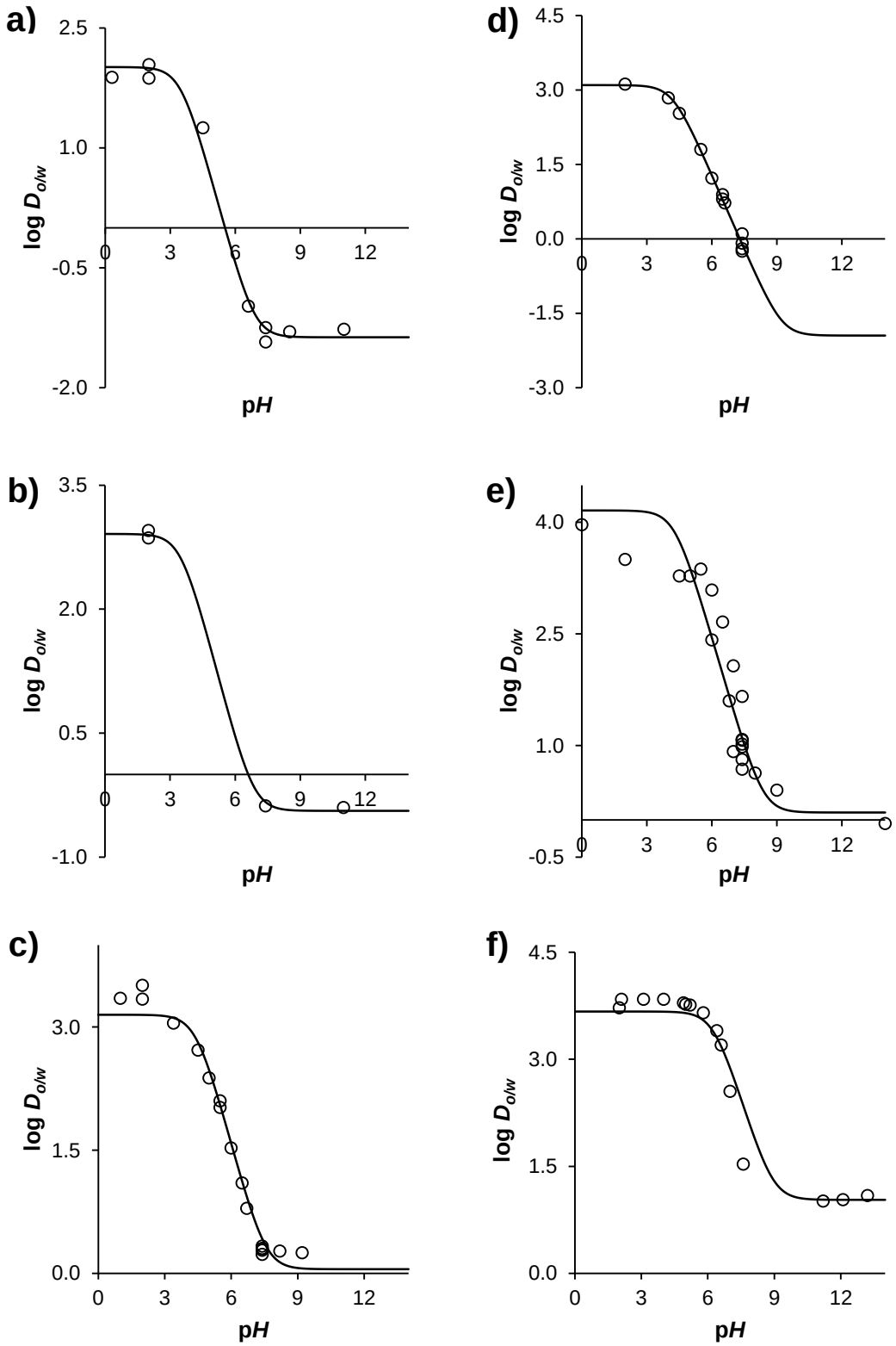

Figure 6

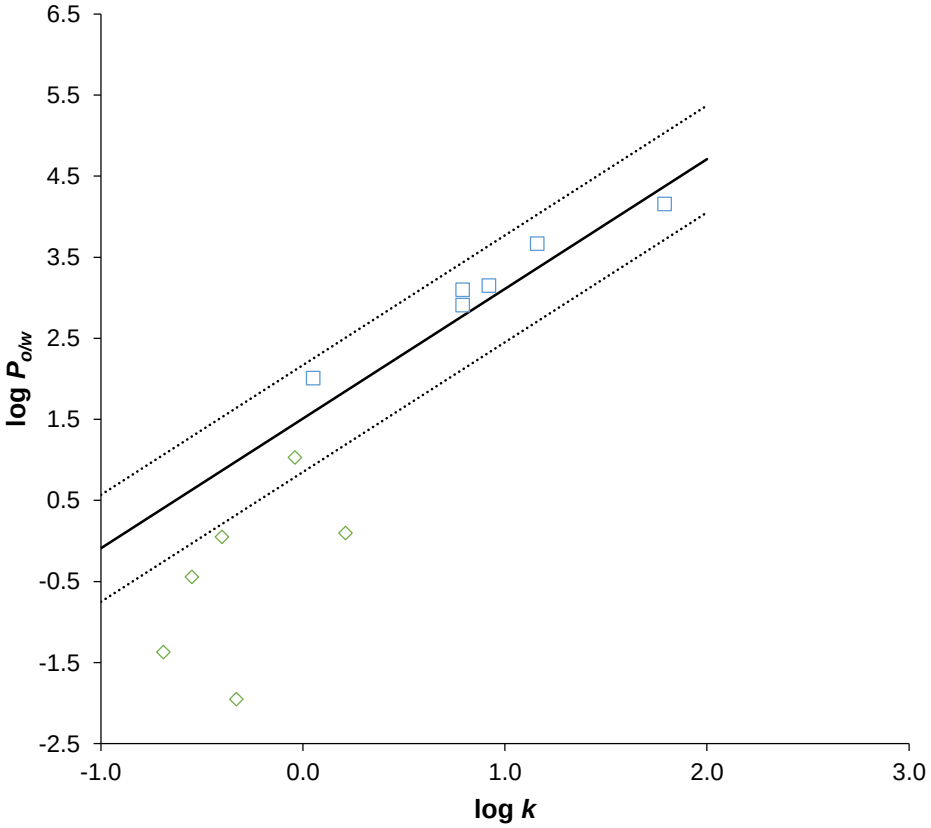

Figure 7

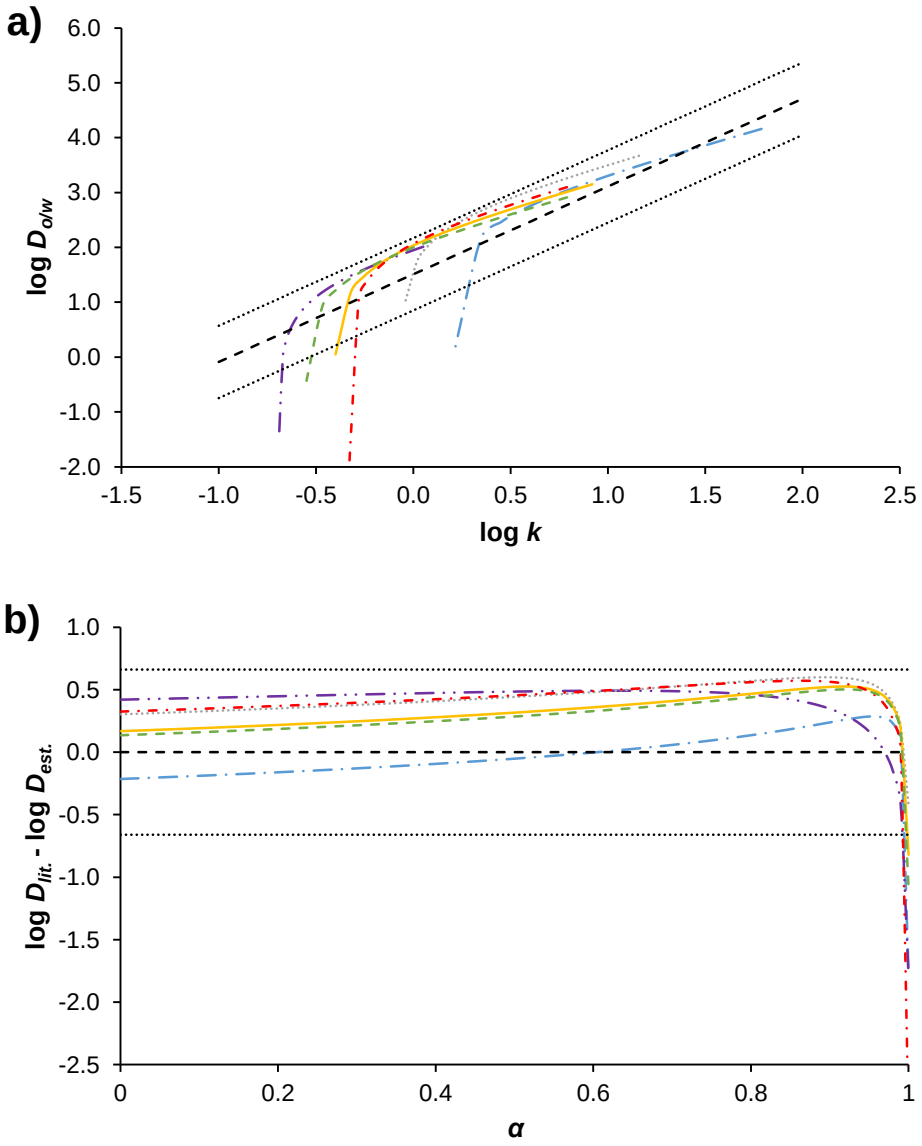

Figure 8