Talanta 97 (2012) 355–361

Contents lists available at SciVerse ScienceDirect

Talanta

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

Determination of n-octanol/water partition coefficients of weak ionizable solutes by RP-HPLC with neutral model compounds

Shu-ying Han, Jun-qin Qiao, Yun-yang Zhang, Hong-zhen Lian *, Xin Ge **

State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry & Chemical Engineering and Center of Materials Analysis, Nanjing University, 22 Hankou Road, Nanjing 210093, China

ARTICLE INFO

Article history:
Received 1 February 2012
Received in revised form
12 April 2012
Accepted 21 April 2012
Available online 30 April 2012

Keywords:
n-Octanol/water partition coefficient (K_{ow})
Apparent *n*-octanol/water partition coefficient (K''_{ow})
Reversed-phase high performance liquid chromatography (RP-HPLC)
Dual-point retention time correction (DP-RTC)
Weak acidic compounds
Aristolochic acids (AAs)

ABSTRACT

The utilization of neutral compounds as model compounds is put forward for determination of the *n*-octanol/water partition coefficient (K_{ow}) of highly hydrophobic, weak acidic compounds by reversed-phase high performance liquid chromatography (RP-HPLC). It is based on a linear relationship between the logarithm of apparent *n*-octanol/water partition coefficient ($\log K''_{ow}$), expressing hydrophobicity of acidic solutes more accurately, and the logarithm of RP-HPLC retention factor of the solutes corresponding to the neat aqueous fraction of mobile phase ($\log k_w$). The availability of neutral model compounds was theoretically tested for this novel protocol. Moreover, a high consistency of linear $\log K''_{ow}$ –$\log k_w$ correlations was demonstrated between a mixed training set of neutral and acidic model compounds, and a training set of neutral model compounds. It is proved in theory that for a certain set of compounds investigated, all derived linear relationships between $\log K''_{ow}$ and $\log k_w$ have a unit slope and the same intercept, regardless of mobile phase pH. This model was applied to measure $\log K_{ow}$ of lipophilic aristolochic acid I (AA I) and aristolochic acid II (AA II). Log K_{ow} values for AA I and AA II are 4.45 ± 0.07 and 3.99 ± 0.06, respectively. To the best of our knowledge, this is the first report on experimental $\log K_{ow}$ data for AAs. The proposed strategy solves the problem of lacking suitable acidic model compounds with reliable experimental K_{ow} in determining K_{ow} of lipophilic acidic solutes by RP-HPLC.

1. Introduction

Lipophilicity, generally expressed by logarithm of *n*-octanol/water partition coefficient, $\log K_{ow}$, constitutes a physicochemical property of paramount importance in medicinal chemistry. It plays an essential role in absorption, distribution, metabolism and excretion (ADME) characteristics of drugs while also affecting their pharmacodynamic and toxicological profiles, such as extent of plasma protein binding, accumulation in tissues and unpredictable poisonous mechanism [1,2].

Many experimental methods have been developed for K_{ow} determination, among which reversed-phase high performance liquid chromatography (RP-HPLC) is most frequently adopted and has been recommended by Organisation for Economic Co-operation and Development (OECD) [3]. It profits from the linear Collander equation between $\log K_{ow}$ and $\log k_w$, the logarithm of retention factor (k) of analytes obtained by extrapolating to neat aqueous fraction of the mobile phase [4]. Compared to traditional experimental methods, e.g., the shake-flask method (SFM) [5] and slow-stirring method (SSM) [6], RP-HPLC is faster and less expensive because it does not require any quantification of concentration, and only retention time (t_R) needs to be measured. Also, it is more accurate, especially for highly hydrophobic compounds. In addition, it has been proved that RP-HPLC shows much higher reliability in determining K_{ow} for most chemicals with complex structures in comparison to theoretical calculations based on a fragmental constant or atomic contribution approach [7]. However, the RP-HPLC method is highly dependent on accurate measurement of t_R. Silica-based stationary phase collapses unavoidably by continual flushing of aqueous mobile phase during the course of usage, causing a shift of t_R for an analyte under the same chromatographic condition in different measuring periods. Therefore, t_R should be rectified to eliminate this deviation, especially in the study of retention behavior. Several rectification protocols have been proposed to standardize

Abbreviations: K_{ow}, *n*-Octanol/water partition coefficient; RP-HPLC, Reversed-phase high performance liquid chromatography; k_w, P-HPLC retention factor of the solute corresponding to the neat aqueous fraction of mobile phase; K''_{ow} , Apparent *n*-octanol/water partition coefficient; AA I, Aristolochic acid I; AA II, Aristolochic acid II; ADME, Absorption, Distribution, metabolism and excretion; OECD, Organization for Economic Co-operation and Development; k, Retention factor; SFM, Shake-flask method; SSM, Slow-stirring method; t_R, Retention time; SP- , DP- or MP-RTC, Single, dual or multi point-retention time correction

* Corresponding author. Tel.: +86 25 83686075; fax: +86 25 83325180.
** Corresponding author. Tel.: +86 25 83592466; fax: +86 25 83592723.
E-mail addresses: hzlian@nju.edu.cn (H.-z. Lian), gexin@nju.edu.cn (X. Ge).

0039-9140/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.04.045

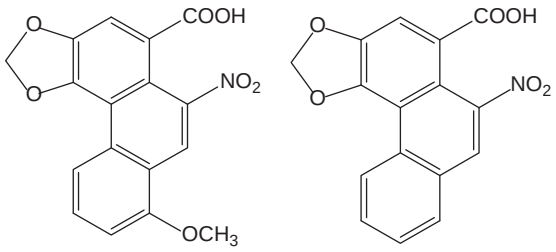

Fig. 1. Structures of aristolochic acids I and II.

t_R, such as single- (SP-), dual- (DP-) and multi-point (MP-) retention time correction (RTC) methods [8–10], making the evaluation of K_{ow} more reliable.

In general, employing the Collander equation for K_{ow} determination by RP-HPLC is limited to neutral solutes. For weak ionizable acidic compounds, the apparent *n*-octanol/water partition coefficient, K''_{ow}, was proposed to correct K_{ow} in order to describe the hydrophobicity of these acidic solutes more precisely. The improved linear relationship relating $\log K''_{ow}$ with $\log k_w$ compared to $\log K_{ow}$ has been established and successfully applied to K_{ow} measurement for weak acidic compounds [11–13]. The K''_{ow} range of model compounds is commonly considered to cover all compounds in question. However, for acidic solutes with high hydrophobicity, there is a shortage of suitable lipophilic weak acids with reliable experimental K_{ow} values as model compounds. Aristolochic acids (AAs) represent a mixture of structure-related nitrophenanthrene carboxylic acid derivatives [14], with 6-nitrophenanthro-[3,4-d]-1,3-dioxole-5-carboxylic acid (AA II) and 8-methoxy-6-nitrophenanthro-[3,4-d]-1,3-dioxole-5-carboxylic acid (AA I) as the major components (Fig. 1). The pK_a-values for AA I and AA II are 3.3 ± 0.1 and 3.2 ± 0.1, respectively, indicating relatively high acidity [15]. AA-containing herbs have many pharmacological properties, and are commonly used as traditional Chinese medicines [16]. However, in recent years many studies have indicated that AAs are responsible for aristolochic acid nephropathy, as well as Balkan endemic nephropathy and its associated urothelial cancer [17]. Currently, there are no experimental K_{ow} data for AAs available. The calculated $\log K_{ow}$ values of AA I (3.41 ± 1.46) and AA II (3.50 ± 1.04) indicate that AA II should be slightly more lipophilic than AA I, contrary to the fact that AA I contains an additional hydrophobic methoxy moiety compared to AA II.

This present work is aimed at exploring the possibility of modeling the $\log K''_{ow} - \log k_w$ relationship by a neutral training set, thereby solving the problem of shortage in suitable acidic model compounds for determining K_{ow} of highly hydrophobic acidic solutes. For this purpose, a theoretical derivation was performed to test the feasibility of a neutral training set for the $\log K''_{ow} - \log k_w$ model. Furthermore, the $\log K''_{ow} - \log k_w$ relationships obtained by a mixed training set of neutral and weak acidic compounds with reliable experimental K_{ow} values, and by a training set only containing the neutral compounds were established. Subsequently, the obtained relationships were compared to confirm the feasibility of a neutral instead of a mixed training set. To assess the potential of the model, it was validated by verification compounds and used to determine the K_{ow}-values of AAs.

2. Theoretical basis

The hydrophobicity of a weak acidic compound can be expressed by calibrating K_{ow} to K''_{ow}, with $\log K''_{ow}$ having a better linear correlation with $\log k_w$ than $\log K_{ow}$ for weak acidic compounds (*m* and *n* represent fitting parameters) [11–13,18]:

$$K''_{ow} = \frac{K_{ow}}{1+(K_{a1}/[H^+])+(K_{a1}K_{a2}/[H^+]^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]^n)} \quad (1)$$

$$\mathrm{Log}\ K''_{ow} = m \log k_w + n \quad (2)$$

The k_w-value of a weak acidic solute at a certain mobile phase pH is the weighted average of retention factors of all neutral and ionic species formed in the neat aqueous fraction of mobile phase [19]:

$$k_w = (k_0)_w x_0 + (k_1)_w x_1 + (k_2)_w x_2 + \ldots + (k_n)_w x_n \quad (3)$$

In Eq. (3), $(k_0)_w$ is the retention factor of the non-dissociated form of the acidic solute in neat aqueous phase, and $\log(k_0)_w$ has good linear correlation with the corresponding $\log K_{ow}$ according to Collander equation ($\log K_{ow} = a \log k_w + b$, where k_w refers to the retention factor of neutral solute in 100% water, and *a* and *b* are constants). The $(k_1)_w$, $(k_2)_w$ … and $(k_n)_w$ values are the retention factors of the ionic forms of the acidic solute, and x_0, x_1, x_2 … and x_n are mole fractions of the corresponding species. Eq. (3) can be transformed into Eq. (4) based on the dissociation equilibrium of weak acidic solute:

$$k_w = (k_0)_w \frac{1}{\left(1+\frac{K_{a1}}{[H^+]}+\ldots+\frac{K_{a1}K_{a2}\ldots K_{an}}{[H^+]^n}\right)} + (k_1)_w \frac{K_{a1}/[H^+]}{\left(1+\frac{K_{a1}}{[H^+]}+\ldots+\frac{K_{a1}K_{a2}\ldots K_{an}}{[H^+]^n}\right)} + \ldots + (k_n)_w \frac{K_{a1}K_{a2}\ldots K_{an}/[H^+]^n}{\left(1+\frac{K_{a1}}{[H^+]}+\ldots+\frac{K_{a1}K_{a2}\ldots K_{an}}{[H^+]^n}\right)} \quad (4)$$

For most of weak acidic compounds investigated, their K_a values of all levels are 1–2 orders of magnitude lower than $[H^+]$ of mobile phase, while $(k_1)_w$, $(k_2)_w$, … and $(k_n)_w$ are small in comparison with $(k_0)_w$. Therefore, just like Eq. (1), the approximation of k_w can be expressed as follows:

$$k_w = \frac{(k_0)_w}{1+\frac{K_{a1}}{[H^+]}+\frac{K_{a1}K_{a2}}{[H^+]^2}+\ldots+\frac{K_{a1}K_{a2}\ldots K_{an}}{[H^+]^n}} \quad (5)$$

By substitution of Eq. (1) and (5) in Eq. (2), the following equation was obtained:

$$\mathrm{Log}\ K_{ow} - m \log(k_0)_w - n = (1-m)\log\left(1+\frac{K_{a1}}{[H^+]}+\frac{K_{a1}K_{a2}}{[H^+]^2}+\ldots+\frac{K_{a1}K_{a2}\ldots K_{an}}{[H^+]^n}\right) \quad (6)$$

For non-dissociated solutes, $1+(K_{a1}/[H^+])+(K_{a1}K_{a2}/[H^+]^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]^n)=1$, thus, Eq. (6) was transferred to $\log K_{ow} = a \log k_w + b$ (for neutral compounds, $(k_0)_w = k_w$, $m=a$, $n=b$), indicating that neutral compounds can be recognized as a special state of the corresponding non-dissociated acidic compounds ($[H^+] \gg K_a$). Neutral and acidic compounds can therefore be considered as a mixed training set for the $\log K''_{ow}$ –$\log k_w$ regression. For acidic solutes, the good linear relationship between $\log K_{ow}$ and $\log(k_0)_w$ for the non-dissociated form of weak acidic solutes evinced that the polynomial on the left-hand side of Eq. (6) equaled zero at any mobile phase pH. The dissociation of weak acidic solutes cannot be ignored at the chosen pH in this study, therefore $1+(K_{a1}/[H^+])+(K_{a1}K_{a2}/[H^+]^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]^n) \neq 1$. In this instance, $m=1$ is derived. Thus, Eq. (2) can be simplified to

$$\mathrm{Log}\ K''_{ow} = \log k_w + n \quad (7)$$

In addition, by substitution of Eq. (1) and (5) in Eq. (7) at two arbitrary different pH values (e.g., $[H^+]_a$ and $[H^+]_b$), the following expressions were obtained:

$$\mathrm{Log}\left(\frac{K_{ow}}{1+(K_{a1}/[H^+]_a)+(K_{a1}K_{a2}/[H^+]_a^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]_a^n)}\right)$$

$$= \log\left(\frac{k_0}{1+(K_{a1}/[H^+]_a)+(K_{a1}K_{a2}/[H^+]_a^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]_a^n)}\right)+n_a \quad (8)$$

$$\mathrm{Log}\left(\frac{K_{ow}}{1+(K_{a1}/[H^+]_b)+(K_{a1}K_{a2}/[H^+]_b^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]_b^n)}\right)$$
$$= \log\left(\frac{k_0}{1+(K_{a1}/[H^+]_b)+(K_{a1}Ka_2/[H^+]_b^2)+\ldots+(K_{a1}K_{a2}\ldots K_{an}/[H^+]_b^n)}\right)+n_b \quad (9)$$

Eqs. (8) and (9) can be transformed into

$$\mathrm{Log}\,K_{ow}-\log\left(1+\frac{K_{a_1}}{[H^+]_a}+\frac{K_{a_1}K_{a_2}}{[H^+]_a^2}+\ldots+\frac{K_{a_1}K_{a_2}\ldots K_{a_n}}{[H^+]_a^n}\right)$$
$$= \log k_0-\log\left(1+\frac{K_{a_1}}{[H^+]_a}+\frac{K_{a_1}K_{a_2}}{[H^+]_a^2}+\ldots+\frac{K_{a_1}K_{a_2}\ldots K_{a_n}}{[H^+]_a^n}\right)+n_a \quad (10)$$

$$\mathrm{Log}\,K_{ow}-\log\left(1+\frac{K_{a_1}}{[H^+]_b}+\frac{K_{a_1}K_{a_2}}{[H^+]_b^2}+\ldots+\frac{K_{a_1}K_{a_2}\ldots K_{a_n}}{[H^+]_b^n}\right)$$
$$= \log k_0-\log\left(1+\frac{K_{a_1}}{[H^+]_b}+\frac{K_{a_1}K_{a_2}}{[H^+]_b^2}+\ldots+\frac{K_{a_1}K_{a_2}\ldots K_{a_n}}{[H^+]_b^n}\right)+n_b \quad (11)$$

The same items on both sides of Eqs. (10) and (11) can be expunged:

$$\mathrm{Log}\,K_{ow} = \log k_0 + n_a \quad (12)$$

$$\mathrm{Log}\,K_{ow} = \log k_0 + n_b \quad (13)$$

Thus, $n_a = n_b$ was deduced by comparing Eq. (12) with Eq. (13), proving that the intercept n in Eq. (7) was constant for a certain training set.

The above theoretical derivation indicated that as long as linearity of the $\log K_{ow}''$–$\log k_w$ correlation is established, the slope of this linearity always equals to one. As claimed by Minick et al. [20], and Giaginis et al. [21], for neutral compounds, the slope of an equation relating n-octanol/water partitioning ($\log K_{ow}$) with chromatographic retention ($\log k_w$) is an estimate of how closely the free energies of the processes compare, and a unit slope in such a plot implies that these two processes are homo-energetic, i.e., the free energy changes are identical. Analogously, a unit slope for the $\log K_{ow}''$ –$\log k_w$ correlation suggested that for a mixed set containing both neutral and acidic solutes, the apparent n-octanol/water partitioning and chromatographic retention are also homo-energetic processes, implying that $\log k_w$ can accurately simulate $\log K_{ow}''$. Moreover, the constant intercept n indicates that this relationship is independent of mobile phase pH used.

Since neutral compounds can be considered as the corresponding non-dissociated compounds, with its K_{ow}'' equals to K_{ow} at arbitrary pH, it is reasonable to presume that the $\log K_{ow}''$ –$\log k_w$ relationship can be obtained by a training set containing only neutral compounds. This can be applied to determine $\log K_{ow}''$ values of weak acidic solutes, and K_{ow} can then be obtained through Eq. (1). In this way, the difficulty in searching for suitable high hydrophobic acidic model compounds can be overcome.

3. Materials and methods

3.1. Materials

Water for mobile phase was Wahaha purified water (Wahaha Group, Hangzhou, China). The mobile phases were prepared from methanol (HPLC grade, Merck, Darmstadt, Germany), acetic acid (analytical-reagent grade, Sinopharm Group Chemical Reagent, Shanghai, China) and water. Reference substances of AA I and AA I/AA II mixture (with purity of 96%) were purchased from Sigma (Sigma-Aldrich, St. Louis, USA) and Acros (Acros Organics, New Jersey, USA), respectively. Table 1 lists all substances investigated in this experiment, all with the purities of 96% or higher. Purities were confirmed by RP-HPLC and all compounds were used without further purification. Stock solutions of the compounds were prepared in methanol (ca. 0.5 mg mL^{-1}) and stored in refrigerator before use.

3.2. Apparatus

A LabTech 600 LC (Lab-Tech Instru., Beijing, China) was employed consisting of a Rheodyne 7725i injector valve equipped with a 10-μL loop (Cotati, CA, USA), a HB-230A incubator (Hanbon Sci. & Tech., Huai'an, China) and a UV–vis 600 Detector (Lab-Tech) set at the respective optimum absorption wavelength for each eluted compound. The chromatographic column used was a Kromasil C_{18}, 5 μm, 150 mm × 4.6 mm i.d. (Hanbon) maintained at 30 °C. Data acquisition and processing were performed on a LC workstation 2006 (Lab-Tech). All experimental retention times were obtained by averaging the results of at least two independent injections at 1.0 mL min^{-1} mobile phase flow rate.

The pH values of mobile phases were measured with a SevenMulti electrochemical analytical meter (Metter-Toledo Instru., Schwerzenbach, Switzerland). The electrode system was calibrated using aqueous standard reference buffers of pH 2.00 and 4.01 at 25 °C (Mettler-Toledo). All pH readings were done in $^{w}_{w}$pH scale, i.e., the pH of aqueous fraction before mixing it with organic modifiers [11].

3.3. Procedure

All compounds studied were eluted by the mobile phase consisting of methanol and water at different pH adjusted by acetic acid (pH 3.20, 4.00 and 4.80). At each pH, a minimum of four isocratic methanol–water ratios were operated according to lipophilicity of the analyte. The t_R-value was recorded at each methanol–water ratio, and corrected by DP-RTC using anisole and hexamethylbenzene as "anchor compounds". The k value was calculated according to the equation $k=(t_R-t_0)/t_0$, where t_0 was determined by using sodium nitrate eluted on the "standard column" [10]. For each solute, the logarithm of k was plotted against the volume fraction of methanol (φ), and $\log k_w$ of the solute was subsequently obtained by extrapolation of retention factor to neat aqueous mobile phase *via* the Snyder–Soczewinski equation [4]. The K_{ow}'' value of each weak acidic solute was calibrated by Eq. (1) (See Table 2). For neutral compounds, the symbol K_{ow}'' was also used in order to keep uniformity with acidic ones, although it equals to K_{ow} at arbitrary pH. Fitting equations relating $\log K_{ow}''$ with $\log k_w$ derived from a mixed training set comprising 11 neutral and 10 acidic solutes, and from a neutral training set containing 11 neutral solutes, were built at different pH. The experimental procedure was performed in triplicate.

The statistical analysis for regression model was accomplished by SPSS V16.0.0 (SPSS, Chicago, Illinois, USA) and MATLAB Software V7.10.0 (R2010.a) (The MathWorks, Natick, MA, USA).

4. Results and discussion

4.1. Internal validation of $\log K_{ow}''$ –$\log k_w$ relationship derived from a neutral training set

A mixed training set of 11 neutral and 10 weak acidic compounds, and a neutral set containing 11 neutral ones, were studied for the $\log K_{ow}''$ –$\log k_w$ relationship. Table 2 lists $\log K_{ow}''$

Table 1
Log K_{ow} and pK_a of investigated compounds.

Compound	Log K_{ow}[a]	Compound	Log K_{ow}[a]	pK_a
Anisole	2.04 [22]	2-Chlorobenzoic acid	2.05 [32]	2.90 [39]
Benzyl chloride	2.30 [23]	4-Chlorobenzoic acid	2.66 [33]	3.99 [40]
Ethylbenzene	3.15 [24]	1-Naphthalenecarboxylic acid	3.10 [34]	3.69 [41]
2-Methylnaphthalene	4.00 [25]	2-Naphthalenecarboxylic acid	2.78 [31]	4.16 [41]
1,2,4-Trichlorobenzene	4.02 [26]	(+)-2-(Methoxy-2-naphthyl)propionic acid	3.18 [35]	4.15 [42]
1,2,3-Trichlorobenzene	4.05 [26]	4-(1-Methylethyl)benzoic acid	3.28 [31]	4.35 [43]
1,3,5-Trichlorobenzene	4.19 [27]	2,4,6-Trichlorophenol	3.62 [36]	5.84 [44]
Hexamethylbenzene	4.75 [25]	2,4,6-Tribromophenol	3.96 [36]	5.95 [44]
Pentachlorobenzene	5.17 [25]	3,5-Diiodo-2-hydroxybenzoic acid	4.20 [37]	2.30 [45]
Hexachlorobenzene (HCB)	5.73 [28]	Pentachlorophenol	5.00 [38]	4.69 [46]
p,p′–DDT	6.36 [29]	2-Methylbenzoic acid	2.18 [29]	3.98 [38]
Toluene	2.61 [30]	2,3,4,6-Tetrachlorophenol	4.10 [35]	5.22 [47]
p,p′–DDD	5.69 [31]	AA I	3.41 ± 1.46	3.3 ± 0.1 [15]
		AA II	3.50 ± 1.04	3.2 ± 0.1 [15]

[a] Only reliable SFM/SSM data were adopted (excluding AA I and AA II), whose log K_{ow} values were calculated using Advanced Chemistry Development (ACD/Labs) Software V8.14 for Solaris (1994–2007 ACD/Labs).

Table 2
Comparison of log K''_{ow} values calculated through Eq. (1) at different pH and literature log K_{ow} of weak acidic compounds.

Compound	Log K_{ow}	Log K''_{ow}		
		pH 3.20	pH 4.00	pH 4.80
Model compounds				
2-Chlorobenzoic acid	2.05	1.57	0.92	0.14
4-Chlorobenzoic acid	2.66	2.59	2.35	1.79
1-Naphthalenecarboxylic acid	2.78	2.98	2.62	1.96
2-Naphthalenecarboxylic acid	3.10	2.73	2.55	2.05
(+)-2-(Methoxy-2-naphthyl)-propionic acid	3.18	3.13	2.95	2.44
4-(1-Methylethyl)benzoic acid	3.28	3.25	3.12	2.70
2,4,6-Trichlorophenol	3.62	3.62	3.61	3.58
2,4,6-Tribromophenol	3.96	3.96	3.96	3.93
3,5-Diiodo-2-hydroxybenzoic acid	4.20	3.25	2.49	1.70
Pentachlorophenol	5.00	4.99	4.92	–
Verification compounds				
2-Methylbenzoic acid	2.18	2.11	1.87	1.30
2,3,4,6-Tetrachlorophenol	4.10	4.10	4.07	3.96
Sample compounds				
AA I [a]	4.45 ± 0.07	4.24	3.70	2.85
AA II [a]	3.99 ± 0.06	3.73	3.15	2.32

[a] Log K_{ow} was the average value determined by K''_{ow} through Eq. (1), while log K''_{ow} was the average value estimated by log $K''_{ow} - \log k_w$ regressions modeled by the neutral training set at different mobile phase pH listed in Table 3.

values of 10 weak acidic solutes at different pH values, as well as their literature log K_{ow} values for the non-dissociated form. This result indicated that the dissociation of weak acidic compounds could not be neglected at experimental pH, thus the Collander equation established for neutral compounds was not suitable in this work. Instead, K_{ow} was substituted by K''_{ow} for these 10 weak acidic model compounds, resulting in an improved linear correlation relating log K''_{ow} with log k_w. The 11 neutral model compounds were combined with these acidic compounds, forming a mixed training set. The internal validation results are listed in Table 3. The high statistical significance of the log $K''_{ow} - \log k_w$ fittings at various pH indicated that neutral and weak acidic substances could be considered as a mixed training set. Fig. 2 illustrates log $K''_{ow} - \log k_w$ and log $K_{ow} - \log k_w$ fitting lines for the mixed training set at pH 3.20. The outlier in Fig. 2b from line track of log $K_{ow} - \log k_w$ corresponds to 3,5-diiodo-2-hydroxybenzoic acid possessing a smaller pK_a value, i.e., stronger dissociation ability. This point falls within the 95% confident intervals for log $K''_{ow} - \log k_w$ (Fig. 2a), indicating that K''_{ow} is more suitable than K_{ow} to describe the linear correlation between lipophilicity and RP-HPLC retention of ionizable acidic compounds. Furthermore, Table 3 lists the log $K''_{ow} - \log k_w$ regressions obtained by the neutral training set obtained at different pH values. The high consistency between regressions obtained by the two training sets at each pH confirmed our hypothesis, i.e., a neutral training set can be successfully applied to log K''_{ow} determination of weak acidic solutes, implying that acidic model compounds can be omitted in the log $K''_{ow} - \log k_w$ fittings. It also should be noticed in Table 3 that although an increase of pH from 3.20 to 4.80 continuously weakened linearity of both log $K''_{ow} - \log k_w$ and log $K_{ow} - \log k_w$ of the mixed training set, log $K''_{ow} - \log k_w$ correlations at different pH still showed much better linearity, and furthermore, much higher consistency of fitting parameters including both slope and intercept than log $K_{ow} - \log k_w$ regressions did. The slightly weaker linearity of log $K''_{ow} - \log k_w$ at high pH (4.80) may be attributed to the questionable plots of acidic solutes having small pK_a (2 log units lower than pH, i.e., p$K_a < 2.8$). However, this problem could be avoided by using a neutral training set to model the log $K''_{ow} - \log k_w$ relationship, since neutral compounds are non-dissociated at arbitrary pH. It can be seen from Table 3 that the log $K''_{ow} - \log k_w$ regressions obtained by neutral training set at different pH are almost identical.

As summarized in Table 3, m and n of log $K''_{ow} - \log k_w$ regressions obtained both by mixed and neutral training sets at different pH are all approximately one ($\sim$1.10) and about 0.50, respectively. Although the presented experimental results verified the theoretical derivation decided in the Theoretical Basis section, there was still a slight deviation of slope from one, which may be attributed to other secondary interactions between solutes and stationary phase/eluent besides hydrogen-bonding, such as the stereo-chemical structure of the solutes [48]. However, for the investigated compounds, the good linearity of the regressions indicated these interactions were systemic, having no influence on the accuracy of the model in lipophilicity measurement. In addition, an invariable negative intercept of log-$K''_{ow} - \log k_w$ relationships obtained in this work reflected a systematic stronger affinity of the solutes to ODS stationary phase than to bulk octanol. Inversely, the large deviation from unit for the slopes, as well as the changing intercepts of log $K_{ow} - \log k_w$ correlations obtained by the mixed training set at different pH are also listed in Table 3, implying that the free energy changes between n-octanol/water partitioning and RP-HPLC retention are different, and this discrepancy becomes more distinct with pH increase. The reason for the lower quality of log $K_{ow} - \log k_w$ regressions is probably due to the lower hydrophobic interaction of ionized specie(s) on ODS stationary phase, as well as their

Table 3
Relationships of log K_{ow}'' (log K_{ow}) − log k_w derived from a mixed and a neutral training set, respectively, at various mobile phase pH.

pH	Mixed training set		Neutral training set
	Log K_{ow}'' (y) − log k_w (x)	Log K_{ow}'' (y) − log k_w (x)	Log K_{ow}'' (log K_{ow}) (y) − log k_w(x)
3.20	$y=(1.09\pm0.03)\,x-(0.45\pm0.14)$ $n=21$, $R^2=0.982$, $R^2_{cv}=0.978$, S.D.=0.17, $F=1010.89$	$y=(1.02\pm0.06)\,x-(0.10\pm0.25)$ $n=21$, $R^2=0.930$, $R^2_{cv}=0.918$, S.D.=0.32, $F=253.98$	$y=(1.09\pm0.04)\,x-(0.46\pm0.19)$ $n=21$, $R^2=0.986$, $R^2_{cv}=0.979$, S.D.=0.17, $F=628.05$
4.00	$y=(1.11\pm0.05)\,x-(0.50\pm0.20)$ $n=21$, $R^2=0.961$, $R^2_{cv}=0.956$, S.D.=0.27, $F=472.25$	$y=(0.94\pm0.07)\,x+(0.32\pm0.27)$ $n=21$, $R^2=0.906$, $R^2_{cv}=0.884$, S.D.=0.37, $F=183.57$	$y=(1.10\pm0.04)\,x-(0.51\pm0.16)$ $n=21$, $R^2=0.990$, $R^2_{cv}=0.982$, S.D.=0.14, $F=874.11$
4.80	$y=(1.05\pm0.06)\,x-(0.27\pm0.21)$ $n=20$, $R^2=0.949$, $R^2_{cv}=0.941$, S.D.=0.36, $F=333.92$	$y=(0.75\pm0.08)\,x+(1.20\pm0.28)$ $n=20$, $R^2=0.843$, $R^2_{cv}=0.794$, S.D.=0.47, $F=96.44$	$y=(1.11\pm0.04)\,x-(0.53\pm0.17)$ $n=20$, $R^2=0.989$, $R^2_{cv}=0.981$, S.D.=0.15, $F=833.79$

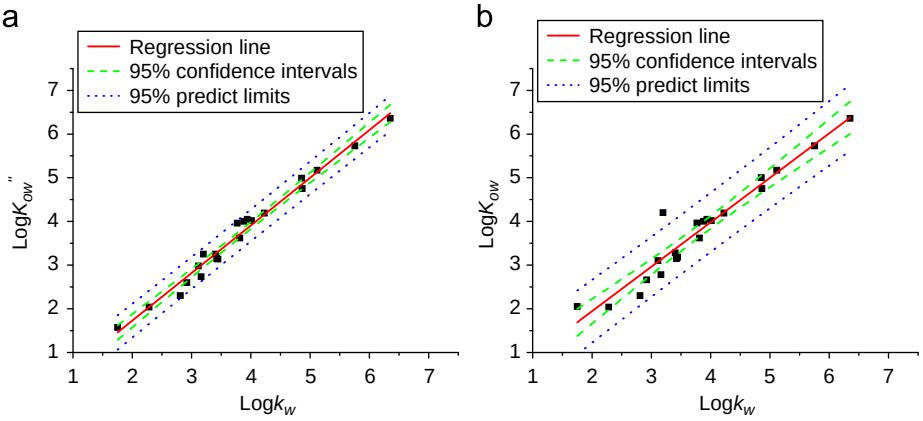

Fig. 2. Correlation of log K_{ow}'' − log k_w (a) and logK_{ow} − log k_w (b) of the mixed training set at mobile phase pH 3.20. Regression equations are listed in Table 3.

Table 4
Validation for log K_{ow}'' − log k_w regressions respectively obtained by the mixed and neutral training sets.

Compounds	pK_a	Training set	Log K_{ow}[a]							
			SFM/SSM	RP-HPLC					ACD/Lab	Error
				Mobile phase pH						
				3.20	4.00	4.80	Average	Error		
Toluene	–	Mixed	2.73	2.71	2.72	2.75	2.73	0.00	2.68 ± 0.17	−0.05
		Neutral		2.70	2.69	2.64	2.68	−0.05		
p,p′–DDD	–	Mixed	5.69	5.68	5.64	5.61	5.64	−0.05	5.39 ± 0.36	−0.30
		Neutral		5.66	5.59	5.64	5.63	−0.06		
2-Methylbenzoic acid	3.91	Mixed	2.46	2.40 (2.32)	2.29 (1.95)	2.40 (1.46)	2.36	−0.10	2.35 ± 0.21	−0.11
		Neutral		2.39 (2.31)	2.27 (1.92)	2.23 (1.28)	2.30	−0.16		
2,3,4,6-Tetrachlorophenol	5.22	Mixed	4.10	3.92 (3.91)	3.91 (3.88)	3.99 (3.85)	3.94	−0.16	4.17 ± 0.35	0.07
		Neutral		3.90 (3.90)	3.87 (3.84)	3.94 (3.80)	3.90	−0.20		

[a] Log K_{ow}'' values obtained by log K_{ow}'' − log k_w regressions at different mobile phase pH for weak acidic compounds are listed in parentheses.

silanophilic interaction in chromatographic retention process. With increasing pH, more anions arising from the corresponding acidic solutes are prone to interact with free residual silanol sites on the stationary phase through hydrogen-bonding interaction. The *n*-octanol/water partitioning of the non-dissociated form of the solute does not undergo these secondary interactions, thus these two processes are not homo-energetic, leading to inferior statistics for the correlation.

4.2. External validation

The reliability of the model was further inspected by an external validation. Toluene, p,p′-DDD, 2-methylbenzoic acid and 2,3,4,6-tetrachlorophenol with reliable SFM/SSM K_{ow} values (log K_{ow} ranged from 2 to 6) were employed as verification compounds for log K_{ow}'' –log k_w fittings obtained by mixed and neutral training sets, respectively, at various mobile phase pH. As shown in Table 4, for both models, the determined log K_{ow}'' of two acidic verification compounds, 2-methylbenzoic acid and 2,3,4,6-tetrachlorophenol, gradually decrease as pH increases, and this variation of log K_{ow}'' becomes even larger when the analyte has stronger acidity. However, mobile phase acidity nearly has no effect on log K_{ow}'' values of neutral verification compounds, toluene and p,p′-DDD. The reason for this phenomenon could be interpreted as follows: K_{ow}'' is proposed as a weighted average of the hydrophobicity of all neutral and ionic species of the solute, and the suppression of dissociation for acidic analyte becomes weaker as pH increases, especially for a fairly strong acid, which consequently increases the proportion of ionic specie(s), thus weakening the hydrophobicity of the solute, and decreasing the K_{ow}'' value. Whereas, log K_{ow} values listed in Table 4, obtained from corresponding log K_{ow}'' by Eq. (1) are almost invariant for

each compound at different pH. Moreover, these log K_{ow} data obtained by utilizing two sets of model compounds are also nearly identical, indicating that there is no discrepancy between the two sets in determining K_{ow} values for both neutral and acidic solutes. In addition, the results presented in Table 4 shows good agreement between simulated log K_{ow} and SFM/SSM values for all four verification compounds (absolute error $\leq$ 0.19 log unit), revealing a high accuracy of proposed log K''_{ow} $-$log k_w model. The log K_{ow} values calculated by ACD/Lab software are also listed in Table 4. It can be observed that the software-computed K_{ow} values have acceptable deviation for neutral and acidic compounds having simple structures. However, this calculation tends to be inaccurate for compounds showing several functional groups with interactions, e.g., p,p′-DDD. For relatively complex molecules with large or unusual fragments for which theoretical values are not available in the list provided by the method or in the software library, as well as the complex interactions among acidic solutes, organic solvents and ion-suppressors are the possible error resources for software calculation.

4.3. Determination of log K_{ow} for AAs

Since the linearity of log K''_{ow} and log k_w obtained by a neutral training set has been confirmed both theoretically and experimentally, and the log K''_{ow} $-$log k_w relationship of 11 neutral model compounds has been validated to be reliable and accurate for determining K_{ow} of structurally-related acidic compounds, using this model log K_{ow} values of AAs were determined (Table 2) and compared with C log P values (Table 1). The results suggested that ACD/Lab software, which is thought to be the most accurate theoretical calculation for acidic compounds [6], produced log K_{ow} values for AA I and AA II which had large deviations when compared to the values obtained using the models derived herein. This indicates that the theoretical calculations still have large defect in prediction of log K_{ow} for relatively complex acidic molecules due to the lack of fragmental values and inaccuracy of descriptors profiling interactions among acidic solutes, organic solvents such as n-octanol and water in the database. In addition, the questionable calculated log K_{ow} values order of AA I and AA II was reversed by the proposed RP-HPLC method, and the recommended log K_{ow} for AA I and AA II are 4.45 $\pm$ 0.07 and 3.99 $\pm$ 0.06, respectively.

5. Conclusions

In this paper, the feasibility of using neutral model compounds to model the relationship between the apparent n-octanol/water partition coefficient (log K''_{ow}) and the extrapolated RP-HPLC retention factor (log k_w) for K_{ow} measurement of weak acidic compounds was described and confirmed by experiment. It also proved in theory that for a training set, if the dissociation of acidic solute was only considered, the slope and intercept of the log K''_{ow} $-$log k_w relationship are equal to one and a constant, respectively, regardless of mobile phase pH. The unit slope indicates the apparent n-octanol/water partitioning and chromatographic retention are homo-energetic processes. Although in practice the regression possessed a slope slightly deviating from one, good linearity of the log K''_{ow} $-$log k_w model, as well as satisfactory external validation results indicated that other secondary interactions were systemic, having no influence on determination accuracy of the model for both neutral and acidic solutes. The K_{ow} values of AAs were determined by the derived log K''_{ow} $-$log k_w correlation of a neutral training set. As far as we know, these experimental K_{ow} data of AA I and AA II are reported for the first time. The proposed novel and convenient protocol by relating chromatographic retention with apparent n-octanol/water partitioning based on a neutral training set solved the problem of lacking suitable acidic model compounds for K_{ow} determination of weak acidic solutes with high hydrophobicity.

Acknowledgments

This work was supported by National Natural Science Foundation of China (90913012), National Basic Research Program of China (973 program, 2009CB421601, 2011CB911003), National Natural Science Foundation of China (20575027), National Science Funds for Creative Research Groups (21121091), and Analysis and Test Fund of Nanjing University. Authors thank Professor Dr. Huwei Liu, College of Chemistry and Molecular Engineering, Peking University for his kind help.

References

[1] C. Hansch, A.J. Leo, D. Hoekman, Fundamentals and applications in chemistry and biology in Exploring QSAR, American Chemical Society, Washington DC, 1995.
[2] H. van de Waterbeemd, D.A. Smith, K. Beaumont, D.K. Walker, J. Med. Chem. 44 (2001) 1313–1333.
[3] OECD, Guideline for Testing of Chemicals, no. 117: Partition Coefficient (n-octanol/water), High Performance Liquid Chromatography Method, 1989.
[4] L.R. Snyder, J.W. Dolan, J.R. Gant, J. Chromatogr. 165 (1979) 3–30.
[5] OECD, Guideline for Testing of Chemicals, no. 107: Partition Coefficient (n-octanol/water), Flask Shaking Method, 1981.
[6] A. Finizio, M. Vighi, D. Sandroni, Chemosphere 34 (1997) 131–161.
[7] A.J. Leo, C. Hansch, D. Helkins, Chem. Rev. 71 (1971) 525–538.
[8] K. Petritis, L.J. Kangas, P.L. Ferguson, G.A. Anderson, L. Pasa-Tolic, M.S. Lipton, K.J. Auberry, E.F. Strittmatter, Y. Shen, R. Zhao, R.D. Smith, Anal. Chem. 75 (2003) 1039–1048.
[9] O. Kohlbacher, S. Quinten, M. Sturm, B.M. Mayr, C.G. Huber, Angew. Chem. Int. Ed. 45 (2006) 7009–7012.
[10] S.Y. Han, J.Q. Qiao, Y.Y. Zhang, L.L. Yang, H.Z. Lian, X. Ge, H.Y. Chen, Chemosphere 83 (2011) 131–136.
[11] H.Z. Lian, W.H. Wang, D.N. Li, J. Sep. Sci. 28 (2005) 1179–1187.
[12] X. Ming, S.Y. Han, Z.C. Qi, D. Sheng, H.Z. Lian, Talanta 79 (2009) 752–761.
[13] S.Y. Han, X. Ming, Z.C. Qi, D. Sheng, H.Z. Lian, Anal. Bioanal. Chem. 398 (2010) 2731–2743.
[14] D.B. Mix, H. Guinaudeau, M. Shamma, J. Nat. Prod. 45 (1982) 657–666.
[15] X. Fu, Y. Liu, W. Li, Y. Liao, M. Zhai, H. Liu, Proceedings of the 22nd International Symposium on Microscale Bioseparations and Methods for Systems Biology, Dalian, 2008.
[16] M.Y. Shen, C.L. Liu, G. Hsiao, C.Y. Liu, K.H. Lin, D.S. Chou, J.R. Sheu, Planta Med. 74 (2008) 1240–1245.
[17] A.P. Grollman, S. Shibutani, M. Moriya, F. Miller, L. Wu, U. Moll, N. Suzuki, F.A. ernandes, T. Rosenquist, Z. Medverec, K. Jakovina, B. Brdar, N. Slade, R.J. Turesky, A.K. Goodenough, R. Rieger, M. Vukelić, B. Jelaković, Proc. Natl. Acad. Sci. USA 104 (2007) 12129–12134.
[18] A. Berthod, Anal. Chem. 71 (1999) 879–888.
[19] C. Horváth, W. Melander, I. Molnar, Anal. Chem. 49 (1977) 142–154.
[20] D.J. Minick, D.A. Brent, J. Frenz, J. Chromatogr. 461 (1989) 177–191.
[21] C. Giaginis, S. Theocharis, A. Tsantili-Kakoulidou, Anal. Chim. Acta 573–574 (2006) 311–318.
[22] K.S. Rogers, A. Cammarata, J. Med. Chem. 12 (1969) 692–693.
[23] N. Kurihara, M. Uchida, T. Fujita, M. Nakajima, Pest. Biochem. Physiol. 2 (1973) 383–390.
[24] J. Iwasa, T. Fujita, C. Hansch, J. Med. Chem. 8 (1965) 150–153.
[25] H. Kubinyi, in: E. Jucker (Ed.), Progress in Drug Research, Birkhauser, Basel, 1979, pp. 97–198.
[26] H. Watarai, M. Tanaka, N. Suzuki, Anal. Chem. 54 (1982) 702–705.
[27] C. Chiou, Environ. Sci. Technol. 19 (1985) 57–62.
[28] J. De Bruijn, F. Busser, W. Seinen, J. Hermens, Environ. Toxicol. Chem. 8 (1989) 499–512.
[29] S.W. Karickhoff, D. Brown, T. Scott, Water Res. 13 (1979) 241–248.
[30] N. El Tayar, H. van de Waterbeemd, M. Gryllaki, B. Testa, W.F. Trager, Int. J. Pharm. 19 (1984) 271–281.
[31] Y.Z. Da, K. Ito, H. Fujiwara, J. Med. Chem. 35 (1992) 3382–3387.
[32] H. Tomida, Chem. Pharm. Bull. 26 (1978) 2824–2831.
[33] F.H. Clarke, J. Pharm. Sci. 73 (1984) 226–230.
[34] M.J. Kamlet, R.M. Doherty, M.H. Abraham, Y. Marcus, R.W. Taft, J. Phys. Chem. 92 (1988) 5244–5255.
[35] C. Hansch, A.J. Leo, Substituent Constants for Correlation Analysis in Chemistry and Biology, Wiley, New York, 1979.
[36] M. Stockdale, M.J. Selwyn, Eur. J. Biochem. 21 (1971) 565–574.
[37] E. Freese, B.C. Levin, R. Pearce, T. Sreevalsan, J.J. Kaufman, W.S. Koski, N.M. Semo, Teratology 20 (1979) 413–439.

[38] B. Scheele, Chemosphere 9 (1980) 293–309.
[39] J.A. Dean (Ed.), Lange's Handbook of Chemistry, 15th ed.,McGraw-Hill, New York, 1999.
[40] A.V. Willi, W. Meier, Helv. Chim. Acta 39 (1956) 318–322.
[41] J.F.J. Dippy, S.R.C. Hughes, J.W. Laxton, J. Chem. Soc. (1954) 1470–1476.
[42] M.I. La Rotonda, G. Amato, F. Barbato, C. Silipo, A. Vittoria, Quant. Struct.-Act. Relat. 2 (1983) 168–173.
[43] J.W. Baker, J.F.J. Dippy, J.E. Page, J. Chem. Soc. (1937) 1774–1779.
[44] P.A. Johansson, Acta Pharm. Suec 14 (1977) 363–376.
[45] H. Gehlen, J. Rinck, Z. Phys. Chem. 237 (1968) 388–400.
[46] K. Chamberlain, A.A. Evans, R.H. Bromilow, Pestic. Sci. 47 (1996) 265–271.
[47] J. Drahonovsky, Z. Vacek, Collect. Czech. Chem. Commun. 36 (1971) 3431–3440.
[48] D. Vrakas, I. Panderi, D. Hadjipavlou-Litina, A. Tsantili-Kakoulidou, QSAR Comb. Sci 24 (2005) 254–260.