ELSEVIER

International Journal of Pharmaceutics 246 (2002) 135–142

international journal of pharmaceutics

www.elsevier.com/locate/ijpharm

Mathematical representation of apparent dissociation constants in aqueous–organic solvent mixtures

Abolghasem Jouyban [a,*], Hak-Kim Chan [b], Brian J. Clark [c], William E. Acree, Jr. [d]

[a] *School of Pharmacy, Tabriz University of Medical Sciences, Tabriz 51664, Iran*
[b] *Faculty of Pharmacy, The University of Sydney, Sydney NSW 2006, Australia*
[c] *School of Pharmacy, University of Bradford, Bradford BD7 1DP, UK*
[d] *Department of Chemistry, University of North Texas, Denton, TX 76203-5070, USA*

Received 21 January 2002; received in revised form 2 July 2002; accepted 3 July 2002

Abstract

A mathematical model for calculating apparent acid dissociation constants (pK_a) in hydroorganic mixtures with respect to the concentration of organic solvent in a binary mixture is proposed. The correlation ability of the proposed model is evaluated by employing pK_a value of 75 different weak acids in 13 water-cosolvent systems. The results show that the equation is able to correlate the pK_a values with an overall mean percentage differences (MPD) of $0.52 \pm 0.43\%$. In order to test the prediction capability of the model, four experimental pK_a values for each data set have been employed to train the model, then the pK_a values at other solvent compositions predicted and the overall MPD obtained is $1.41 \pm 1.15\%$. The applicability of the model to correlate/predict pK_a values of structurally related drugs in a given binary solvent has been shown. The obtained overall MPD for correlation and prediction capabilities are 1.60 ± 2.16 and $2.89 \pm 3.22\%$, respectively.

Keywords: Cosolvency; Acid dissociation constant; Hydroorganic mixtures

1. Introduction

Solvent blending or cosolvency affects important characteristics of a solute in a solution including its solubility, stability and dissociation constants. In analytical methods such as capillary electrophoresis (CE) or reversed phase liquid chromatography (RPLC), the analyst usually adds an organic solvent to affect the solubility and/or separation efficiency of the analytes. It is well known that the pK_a of the analytes and the pH of the running buffer/mobile phase strongly affect migration/retention time in CE or RPLC, hence provides key information for an analyst in method development stage. In the pharmaceutical industry, the knowledge of numerical values of the pK_a of the solutes with low aqueous solubility

* Corresponding author. Tel.: +98-411-3341315; fax: +98-411-3344798
E-mail address: jouyban@tbzmed.ac.ir (A. Jouyban).

0378-5173/02/$ - see front matter
PII: S0378-5173(02)00355-1

could be useful information to predict the ionic solubility of a solute in solvent mixtures.

CE although a relatively new technique, has become an important separation procedure in pharmaceutical analysis. It has been used to separate charged analytes ranging from small organic and inorganic ions to macromolecular species such as DNA fragments and proteins. This powerful technique provides high separation efficiency, selectivity, separation capacity, flexibility and relatively low operational cost. The most important factors that can affect the separation in CE are the type and concentration of the background electrolyte, pH, ionic strength, type and concentration of organic solvent, applied voltage and temperature. Some of these factors also lead to other effects in CE. For example organic solvents can affect the solubility of the analytes, the electroosmotic flow, viscosity of the running buffer, dielectric constant of the background electrolyte, dissociation constant of the analytes and also silanol groups of the capillary, the current and Joule heating effects. However, when it comes to obtaining the best set of conditions for an assay, the analyst has often resorted to a trial and error approach to method optimisation. Since CE generally operates with charged analytes, the effect of pH and pK_a variations is one of the most important factors in method development in CE where adding an organic solvent is required in order to optimise the operational conditions. Any mathematical model to calculate pK_a values in mixed solvent buffers could provide useful information for the analyst.

This paper presents a mathematical representation of dissociation constants in aqueous binary mixtures. The infinite number of different solvent compositions, which can be prepared from a particular binary system, precludes calculation of the pK_a values in all possible solvent compositions using a mathematical model. Also the possibility of prediction of the dissociation constants in mixed solvents based on a minimum number of experimental data points has been evaluated which could speed up the method development process where the knowledge of pK_a values in mixed solvent are required.

2. Theoretical treatment

A dissociation reaction of a monoprotic acid (HA) in a solvent can be represented as:

$$HA \leftrightarrow H^+ + A^- \qquad K_a = \frac{|H^+||A^-|}{|HA|}$$

The logarithm of K_a can be expressed as:

$$2.303RT \log K_a = \mu_{H^+} + \mu_{A^-} - \mu_{HA} \tag{1}$$

where μ denotes the chemical potential of the species.

The chemical potential in aqueous cosolvent mixtures can be expressed as the mole fraction of the solvents:

$$\mu^m_{H^+} = X_c\mu^c_{H^+} + X_w\mu^w_{H^+} + A_0X_cX_w + A_1X_cX_w(X_c - X_w) \tag{2}$$

$$\mu^m_{A^-} = X_c\mu^c_{A^-} + X_w\mu^w_{A^-} + B_0X_cX_w + B_1X_cX_w(X_c - X_w) \tag{3}$$

$$\mu^m_{HA} = X_c\mu^c_{HA} + X_w\mu^w_{HA} + C_0X_cX_w + C_1X_cX_w(X_c - X_w) \tag{4}$$

where subscripts m, c and w denote mixed solvent, pure cosolvent and pure water, respectively, X is the mole fraction of the solvents, and A, B and C are solute–solvent and solvent–solvent interaction terms. These terms represent the two-body and three-body interactions in the solution (Acree, 1992).

Summation of Eqs. (2)–(4) yields:

$$\mu^m_{H^+} + \mu^m_{A^-} - \mu^m_{HA} = X_c(\mu^c_{H^+} + \mu^c_{A^-} - \mu^c_{HA}) + X_w(\mu^w_{H^+} + \mu^w_{A^-} - \mu^w_{HA}) + (A_0 + B_0 - C_0)X_cX_w + (A_1 + B_1 - C_1) \times [X_cX_w(X_c - X_w)] \tag{5}$$

Replacing the corresponding equals from Eq. (1) into Eq. (5) and appropriate rearrangements give:

$$2.303RT \log K^m_a = X_c(2.303RT \log K^c_a) + X_w(2.303RT \log K^w_a) + (A_0 + B_0 - C_0)X_cX_w + (A_1 + B_1 - C_1) \times [X_cX_w(X_c - X_w)] \tag{6}$$

Since $(A_0+B_0-C_0)$, $(A_1+B_1-C_1)$ and $2.303RT$

are constant values at a given temperature and $pK_a = -\log K_a$, it is possible to simplify Eq. (6) as:

$$pK_a^m = X_c pK_a^c + X_w pK_a^w - W_0 X_c X_w - W_1[X_c X_w (X_c - X_w)] \quad (7)$$

where $W_0 = (A_0 + B_0 - C_0)/(2.303RT)$ and $W_1 = (A_1 + B_1 - C_1)/(2.303RT)$. It is obvious that one can use the volume/weight fractions of the solvents instead of the mole fractions. Although the numerical values of the curve-fitting parameters affected by employing volume or weight fractions, in practice using the volume fractions in solvent blending is more common than the weight and/or mole fractions. Therefore, it is suggested to use the volume fraction (f) based model as:

$$pK_a^m = f_c pK_a^c + f_w pK_a^w + K_0 f_c f_w + K_1 f_c f_w (f_c - f_w) \quad (8)$$

where K_0 and K_1 are the curve-fit parameters. The numerical values of K_0 and K_1 can be computed by fitting the experimental values of $(pK_a^m - f_c pK_a^c - f_w pK_a^w)$ against $f_c f_w$ and $f_c f_w (f_c - f_w)$ by using a no intercept least squares analysis. In these calculations it should be noted that one could employ more curve-fit parameters in order to provide improved accuracy of results. The general form of the proposed equation can be expressed as:

$$pK_a^m = f_c pK_a^c - f_w pK_a^w + f_c f_w \sum_{q=0}^{n} K_q (f_c - f_w)^q \quad (9)$$

In some cases, the numerical values of pK_a^c are not available. In these cases it is possible to rewrite Eq. (8) as:

$$pK_a^m = f_w pK_a^w + J f_c + K_0 f_c f_w + K_1 f_c f_w (f_c - f_w) \quad (10)$$

where J, K_0 and K_1 are the model constants. These are computed via fitting $(pK_a^m - f_w pK_a^w)$ against f_c, $f_c f_w$ and $f_c f_w (f_c - f_w)$.

In a recent paper (Barbosa et al., 1997), a linear relationship between the solvent composition and pK_a values has been presented. According to this paper, if the solute has no preference between the molecules of water and organic solvent, pK_a values in the mixture can be described by:

$$pK_a^m = f_c pK_a^c + f_w pK_a^w \quad (11)$$

where pK_a^c is the apparent pK_a value in the neat organic solvent. A possible theoretical justification for Eq. (11) can be provided by employing the chemical potential approach explained above. However in real solutions, due to preferential solvation, the pK_a values can deviate from Eq. (11).

To assess the accuracy of the proposed equation, the experimental pK_a values were fitted into the equation. The mean percentage differences (MPD) between experimental and calculated pK_a values are considered as an accuracy criterion. MPD defined as:

$$\mathrm{MPD} = \left(\frac{100}{N}\right) \sum \left| \frac{pK_a^{\text{calculated}} - pK_a^{\text{observed}}}{pK_a^{\text{observed}}} \right| \quad (12)$$

where N is the number of experimental data points in each set. The individual percentage deviation (IPD) is also calculated by:

$$\mathrm{IPD} = 100 \left| \frac{pK_a^{\text{calculated}} - pK_a^{\text{observed}}}{pK_a^{\text{observed}}} \right| \quad (13)$$

3. Results and discussion

In order to evaluate the accuracy of the proposed model, available pK_a data in different concentrations of the organic solvents including more than 5 data points were collected from the literature. The details of the collected data including the cosolvent, the solute, the number of experimental data points in each set, the maximum f_c, MPD for correlative and predictive analyses and the references for experimental data are shown in Table 1.

To evaluate the correlation ability of the model, whole data points in each set have been fitted to the model and back-calculated pK_a values have been used to calculate MPDs. This numerical method has been called correlative analysis. For Eq. (8) or Eq. (10) possessing three constant terms (i.e. pK_a^w, K_0 and pK_a^c or J), the least MPD value (0.11%) has been observed for benzoic acid in propylene glycol–water and the highest MPD value (5.89%) has been observed for butibufen in isopropanol–water mixtures. As noted in Theore-

Table 1
Details of data sets and MPD for different numerical analyses

No.	Cosolvent	Solute[a]	N[b]	Max f_c	MPD for correlative Eq.[c]		MPD for predictive Eq.[d]	Reference
					Three constant term Eq.	Four constant term Eq.	Four constant term Eq.	
1	Acetone	ACES (pK_{a_2})	6	0.50	0.37	0.31	0.94	Azab, 1993
2		HEPES (pK_{a_2})	6	0.50	0.34	0.28	0.86	Azab, 1993
3		PIPES (pK_{a_2})	6	0.50	0.37	0.31	0.95	Azab, 1993
4		Tricine (pK_{a_2})	6	0.50	0.40	0.30	0.92	Azab, 1993
5		MES (pK_{a_2})	6	0.50	0.41	0.34	1.05	Azab, 1993
6		BES (pK_{a_2})	6	0.50	0.30	0.27	0.89	Azab, 1993
7	Acetonitrile	Chlorthiazide (pK_{a_2})	6	0.70	0.49	0.41	1.09	Barbosa et al., 1998
8		Chlorthiazide (pK_{a_1})	6	0.70	1.25	1.19	3.31	Barbosa et al., 1998
9		Ciprofloxacin (pK_{a_2})	7	0.70	0.98	0.88	1.62	Barbosa et al., 2001
10		Enoxacin (pK_{a_2})	9	0.70	0.97	0.98	1.55	Barbosa et al., 2001
11		Fleroxacin (pK_{a_1})	8	0.50	0.42	0.42	0.81	Barbosa et al., 2001
12		Fleroxacin (pK_{a_2})	9	0.70	1.03	0.52	1.23	Barbosa et al., 2001
13		Flumequine (pK_{a_1})	8	0.70	0.61	0.47	1.60	Barbosa et al., 2001
14		Furosemide (pK_{a_1})	6	0.70	1.64	1.20	3.66	Barbosa et al., 1998
15		Nalidixic acid	6	0.70	0.91	0.26	1.26	Barbosa et al., 2001
16		Norfloxacin (pK_{a_1})	9	0.50	0.77	0.55	1.22	Barbosa et al., 2001
17		Norfloxacin (pK_{a_2})	11	0.70	1.37	1.02	2.41	Barbosa et al., 2001
18		Ofloxacin (pK_{a_2})	10	0.70	1.11	0.93	1.54	Barbosa et al., 2001
19		Pipemidic Acid (pK_{a_2})	7	0.70	1.70	0.66	1.28	Barbosa et al., 2001
20	Dioxane	ACES (pK_{a_2})	7	0.55	0.27	0.28	0.64	Azab, 1993
21		HEPES (pK_{a_2})	7	0.55	0.24	0.25	0.59	Azab, 1993
22		MES (pK_{a_2})	7	0.55	0.29	0.31	0.71	Azab, 1993
23		BES (pK_{a_2})	7	0.55	0.26	0.27	0.62	Azab et al., 1993
24		PIPES (pK_{a_2})	7	0.55	0.27	0.28	0.65	Azab, 1993
25		Tricine (pK_{a_2})	7	0.55	1.74	1.72	4.98	Azab, 1993
26	Dimethylacetamide	Benzoic acid	6	0.50	1.41	0.45	1.26	Rubino and Berryhill, 1986
27		*p*-Aminobenzoic acid	6	0.50	0.71	0.36	0.93	Rubino and Berryhill, 1986
28		ACES (pK_{a_2})	6	0.50	0.19	0.19	0.59	Azab, 1993
29		HEPES (pK_{a_2})	6	0.50	0.17	0.18	0.54	Azab, 1993
30	Dimethylformamide	MES (pK_{a_2})	6	0.50	0.21	0.22	0.67	Azab, 1993
31		BES (pK_{a_2})	6	0.50	0.18	0.18	0.57	Azab et al., 1993
32		PIPES (pK_{a_2})	6	0.50	0.19	0.19	0.60	Azab, 1993
33		Tricine (pK_{a_2})	6	0.50	0.28	0.16	0.47	Azab, 1993

Table 1 (*Continued*)

No.	Cosolvent	Solute[a]	N[b]	Max f_c	MPD for correlative Eq.[c]		MPD for predictive Eq.[d]	Reference
					Three constant term Eq.	Four constant term Eq.	Four constant term Eq.	
34	Dimethylsulphoxide	Benzoic acid	6	0.50	0.31	0.26	0.76	Rubino and Berryhill, 1986
35		ACES (pK_{a_2})	7	0.55	0.36	0.30	0.76	Azab, 1993
36		HEPES (pK_{a_2})	7	0.55	0.38	0.38	1.09	Azab, 1993
37		BES (pK_{a_2})	7	0.55	0.35	0.29	0.74	Azab et al., 1993
38		*p*-Aminobenzoic acid	6	0.50	0.21	0.18	0.53	Rubino and Berryhill, 1986
39		PIPES (pK_{a_2})	7	0.55	0.36	0.30	0.75	Azab, 1993
40		Tricine (pK_{a_2})	7	0.55	0.31	0.26	0.66	Azab, 1993
41		MES (pK_{a_2})	7	0.55	0.40	0.33	0.85	Azab, 1993
42	Ethanol	ACES (pK_{a_2})	6	0.50	0.76	0.25	0.76	Azab, 1993
43		HEPES (pK_{a_2})	6	0.50	0.68	0.27	0.76	Azab, 1993
44		MES (pK_{a_2})	6	0.50	0.86	0.28	0.85	Azab, 1993
45		BES (pK_{a_2})	6	0.50	0.71	0.28	0.79	Azab et al., 1993
46		PIPES (pK_{a_2})	6	0.50	0.77	0.25	0.77	Azab, 1993
47		Tricine (pK_{a_2})	6	0.50	0.77	0.16	0.50	Azab, 1993
48		Benzoic acid	6	0.50	0.54	0.16	0.47	Rubino and Berryhill, 1986
49		*p*-Aminobenzoic acid	6	0.50	0.50	0.11	0.28	Rubino and Berryhill, 1986
50		*p*-Toloic acid	6	0.50	1.01	0.07	0.17	Rubino and Berryhill, 1986
51	Glycerol	Benzoic acid	6	0.50	0.32	0.25	0.77	Rubino and Berryhill, 1986
52	Isopropanol	Butibufen	7	1.00	5.89	1.88	4.90	Rafols et al., 1997
53		Carprofen	6	1.00	4.49	0.31	1.57	Rafols et al., 1997
54		Diclofenac	6	1.00	5.15	0.50	2.67	Rafols et al., 1997
55		Fenbufen	6	1.00	4.71	0.56	2.98	Rafols et al., 1997
56		Flurbiprofen	6	1.00	4.85	0.60	3.11	Rafols et al., 1997
57		Ibuprofen	7	1.00	5.84	2.02	5.30	Rafols et al., 1997
58		Ketoprofen	7	1.00	4.46	0.53	2.05	Rafols et al., 1997
59		Naproxen	7	1.00	5.59	1.94	4.67	Rafols et al., 1997
60	Methanol	ACES (pK_{a_2})	6	0.50	0.56	0.59	1.64	Azab, 1993
61		HEPES (pK_{a_2})	6	0.50	0.45	0.44	1.20	Azab, 1993
62		MES (pK_{a_2})	6	0.50	0.63	0.66	1.85	Azab, 1993
63		BES (pK_{a_2})	6	0.50	0.47	0.47	1.26	Azab et al., 1993
64		PIPES (pK_{a_2})	6	0.50	0.64	0.67	1.84	Azab, 1993

Table 1 (*Continued*)

No.	Cosolvent	Solute[a]	N[b]	Max f_c	MPD for correlative Eq.[c]		MPD for predictive Eq.[d]	Reference
					Three constant term Eq.	Four constant term Eq.	Four constant term Eq.	
65		Tricine (pK_{a_2})	6	0.50	0.50	0.39	1.15	Azab, 1993
66	Polyethylene glycol 400	Benzoic acid	6	0.50	0.23	0.22	0.65	Rubino and Berryhill, 1986
67	Propylene glycol	Benzoic acid	6	0.50	0.11	0.09	0.23	Rubino and Berryhill, 1986
68		p-Aminobenzoic acid	6	0.50	0.34	0.35	0.93	Rubino and Berryhill, 1986
69		Acetic acid	8	0.70	1.05	0.63	1.47	Barron et al., 1999
70		Boric acid	6	0.50	0.28	0.27	0.70	Barron et al., 1999
71		Citric acid (pK_{a_1})	8	0.70	2.77	0.93	1.81	Barron et al., 1999
72	Tetrahydrofuran	Citric acid (pK_{a_2})	8	0.70	1.19	1.15	3.28	Barron et al., 1999
73		Citric acid (pK_{a3})	8	0.70	0.75	0.73	1.89	Barron et al., 1999
74		Phthalic acid (pK_{a_1})	8	0.70	1.25	1.09	2.33	Barron et al., 1999
75		Phthalic acid (pK_{a_2})	8	0.70	1.73	1.13	2.59	Barron et al., 1999
					1.14	0.52	1.41	

[a] Abbreviations are: ACES, *N*-(2-Acetamido)-2-aminoethanesulfonic acid; HEPES, *N*-(2-Hydroxyethyl) piperazine-*N*′-2-ethansulfonic acid; PIPES, Piperazine-*N*, *N*′-bis (2-ethansulfonic acid, Tricine, *N*-[Tris(hydroxymethyl) methyl] glycine; MES, 2-(*N*-Morpholino) ethansulfonic acid; BES, *N*,*N*-Bis(2-hydroxyethyl)-2-aminoethansulfonic acid; pK_{a_1}, pK_{a_2}, and pK_{a_3} are the first, second and third acid dissociation constants, respectively.

[b] N is the number of data points in each set.

[c] Mean difference between MPD for three constant term model (1.14) and four constant term model (0.52) is statistically significant ($P < 0.0005$).

[d] The number of data for predicted points in each set is ($N-4$).

tical Treatment, employing more constant terms in the model improves its accuracy. Using four constant term (i.e. pK_a^w, K_0, K_1 and pK_a^c or J) version of the model, the least MPD value (0.07%) and the highest MPD value (2.02%) have been observed for *p*-toloic acid in ethanol–water and ibuprofen in isopropanol–water mixtures, respectively. The overall MPD for three and four constant term models are 1.14 ± 1.48 and $0.52 \pm 0.43\%$ and the mean differences are statistically significant (*t*-test, $P < 0.0005$). This means that the four constant term model is capable of producing more accurate correlations.

In order to evaluate the prediction capability of the proposed model, four experimental pK_a values at $f_c = 0$, the maximum f_c (see column 6 in Table 1) and at two other f_c values (between 0 and maximum f_c with nearly constant intervals) have been used to train the model and then the trained model has been employed to predict the pK_a values at other solvent compositions. This method has been called predictive analysis. The obtained overall MPD value using four constant term model is $1.41 \pm 1.15\%$. This prediction percentage difference could be considered as acceptable error where reported maximum experimental relative standard deviation (RSD) is 3.4% (Barron et al., 1999) or 5.3% (Barbosa et al., 1998).

Fig. 1 shows the relative frequency of IPD values for correlative and predictive analyses. As expected, error percentages less than 1% show the highest frequency and error percentages greater than 4% are the lowest frequency. The IPD values

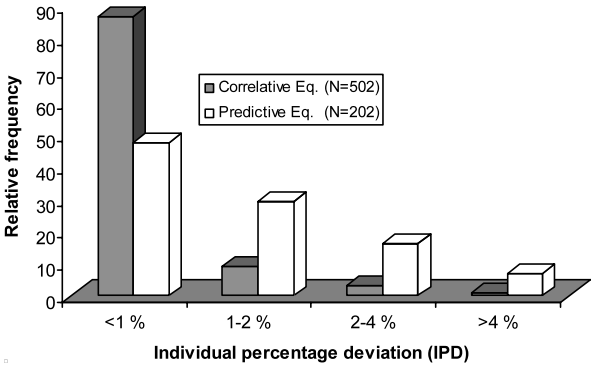

Fig. 1. Relative frequency of IPD for correlative and predictive analyses.

could be considered as a valuable criterion for predicted pK_a values when a researcher wishes to predict unmeasured pK_a in different concentrations of the cosolvent using minimum number of experimental data points. The probability of prediction error $< 1\%$ is 0.48, for 1–2% is 0.29, for 2–4% is 0.16, and for $> 4\%$ is 0.07. This means that it is possible to predict pK_a values in mixed solvents using trained model within prediction error less than experimental RSD values.

To provide a single model to correlate pK_a values for non-steroidal anti-inflammatory drugs (NSAIDs) in isopropanol–water mixtures (data set numbers 51–58 in Table 1), all 52 experimental data points have been fitted to the proposed equation, i.e. (Eq. (8)), and obtained:

$$pK_a^m = f_c pK_a^c + f_w pK_a^w - 3.145 f_c f_w - 7.633 f_c f_w (f_c - f_w) \quad (14)$$

The produced MPD value for 52 correlated data points is $1.65 \pm 2.42\%$. This low correlation error means that the model is able to predict pK_a values of structurally related drugs in a given binary solvent system using a single model. To test this hypothesis, the model is trained employing the experimental pK_a values of set numbers 51, 53, 55 and 57 and then the trained model has been used to predict the pK_a values of set numbers 52, 54, 56 and 58. The obtained MPD for predicted data points is $2.22 \pm 2.84\%$ ($N = 26$). This means that in the drug development studies for a set of structurally related drugs, the known values of pK_a in pure water and pure cosolvent (just 2 experimental data points for each drug) are required to predict pK_a values in whole composition range of the cosolvent concentration in a binary solvent and expected prediction percentage error is less than experimental RSD values.

The capability of correlating/predicting pK_a values of structurally related drugs in a given binary solvent by using the proposed model (i.e. Eq. (10)) is also shown for the first and second pK_a values of quinolones in water-acetonitrile mixtures taken from Barbosa et al. (2001). The obtained model for the first acid-dissociation constant (pK_{a_1}) using experimental data set numbers 11, 13 and 16 is:

$$pK_{a_1}^{m} = f_w pK_{a_1}^{w} + 15.101 f_c - 10.323 f_c f_w - 6.987 f_c f_w (f_c - f_w) \quad (15)$$

and for the second acid-dissociation constant (pK_{a_2}) using experimental data set numbers 10, 12 and 17–19 is:

$$pK_{a_1}^{m} = f_w pK_{a_1}^{w} + 13.500 f_c - 8.710 f_c f_w - 4.144 f_c f_w (f_c - f_w) \quad (16)$$

The calculated MPD for correlated data points using Eqs. (15) and (16) are $1.59 \pm 1.40\%$ ($N = 46$) and $2.02 \pm 2.08\%$ ($N = 25$), respectively. The pK_{a_1} values of enoxacin, enrofloxacin, danofloxacin, difloxacin, sarafloxacin and marbofloxacin in water–acetonitrile have been predicted using (Eq. (15)) and the obtained MPD is $4.17 \pm 4.09\%$ ($N = 24$). The pK_{a_2} values of enrofloxacin, danofloxacin, difloxacin, sarafloxacin and marbofloxacin in water–acetonitrile have been predicted using (Eq. (16)) and the obtained MPD is $2.19 \pm 1.83\%$ ($N = 19$).

In conclusion, the MPD value for correlative studies using 75 data sets and employing four constant term model is $0.52 \pm 0.43\%$ ($N = 502$) and for predicted data points using the trained model is $1.41 \pm 1.15\%$ ($N = 202$). These low correlation and prediction errors mean that the proposed model is able to calculate pK_a values in binary solvents within an acceptable error range. The overall MPD for correlating pK_a of structurally related drugs is $1.60 \pm 2.16\%$ ($N = 123$). The corresponding value for predicted data points using trained model is $2.89 \pm 3.22\%$ ($N = 69$). This prediction method for pK_a of structurally related drugs could provide useful information to pharmaceutical industry where a set of related drugs is synthesised and evaluated. For all these studies an analytical method such as HPLC and/or CE methods and an appropriate solvent mixture are needed. It is suggested that employing the proposed model could help a researcher to speed up these optimisation procedures.

Acknowledgements

The authors would like to thank the Australian Department of Education, Training and Youth Affairs and the University of Sydney for providing the IPRS and IPA scholarships.

References

Acree, W.E., 1992. Mathematical representation of thermodynamic properties. Part II. Derivation of the combined nearly ideal binary solvent (NIBS)/Redlich-Kister mathematical representation from a two-body and three-body interactional mixing model. Thermochim. Acta 198, 71–79.

Azab, H.A., 1993. Potentiometric determination of the second-stage dissociation constants of some hydrogen ion buffers for biological research in various water+organic solvent mixtures. J. Chem. Eng. Data 38, 453–457.

Azab, H.A., Hassan, A., Khafagy, Z.A., 1993. Potentiometric determination of the second-stage dissociation constant of *N*,*N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid in various water+organic solvent mixtures. J. Chem. Eng. Data 38, 231–233.

Barbosa, J., Berges, R., Toro, I., Sanz-Nebot, V., 1997. Dissociation constants and preferential solvation of fluoroquinolones in hydroorganic mixtures used in LC. Int. J. Pharm. 149, 213–225.

Barbosa, J., Barron, D., Beltran, J.L., Buti, S., 1998. On the role of solvent in acid-base equilibria of diuretics in acetonitrile-water mixed solvents. Talanta 45, 817–827.

Barbosa, J., Barron, D., Cano, J., Jimenez-Lozano, Sanz-Nebot, V., Toro, I., 2001. Evaluation of electrophoretic method versus chromatographic, potentiometric and absorptiometric methodologies for determining pK_a values of quinolones in hydroorganic mixtures. J. Pharm. Biomed. Anal. 24, 1087–1098.

Barron, D., Buti, S., Ruiz, M., Barbosa, J., 1999. Evaluation of acidity constants and preferential solvation in tetrahydrofuran–water mixtures. Polyhedron 18, 3281–3288.

Rafols, C., Roses, M., Bosch, E., 1997. Dissociation constants of several non-steroidal anti-inflammatory drugs in isopropyl alcohol/water mixtures. Anal. Chim. Acta 350, 249–255.

Rubino, J.T., Berryhill, W.S., 1986. Effects of solvent polarity on the acid dissociation constants of benzoic acids. J. Pharm. Sci. 75, 182–186.