ELSEVIER

Journal of Chromatography A, 680 (1994) 49–56

JOURNAL OF CHROMATOGRAPHY A

Investigation of experimental approaches to the determination of pK_a values by capillary electrophoresis

S.J. Gluck*[a], J.A. Cleveland, Jr.[b]

[a]*Analytical Sciences, 1897B Bldg, Dow Chemical Co., Midland MI 48667, USA*

[b]*Discovery Research, DowElanco, 9410 Zionsville Rd, Indianapolis IN 46268-1053, USA*

Abstract

The calculation of pK_a values from capillary electrophoresis data may be accomplished in several ways. Electrophoretic mobilities are fitted to a model which describes the dissociation versus the pH. For a linear model, a linear regression approach yields biased results. Weighted linear regression requires many replicates to determine weighting factors and is thus a time-consuming experiment. For an exponential model, non-linear regression of the electrophoretic mobilities in different, equally pH spaced buffers of the analyte gives the least biased determination. Another experimental approach to this determination is to use a permanently charged solute to correct for potential biases in the expected electrophoretic mobilities obtained between different buffers. In the buffer series chosen, there was no significant bias observed. Buffer pH may be determined by an in situ approach using an internal standard of known pK_a. However, the precision obtained is much less than using a pH meter.

1. Introduction

The measurement of electrophoretic mobility versus pH has been developed for pK_a value determination in several laboratories [1–4]. This procedure was developed for poorly soluble compounds with typical working concentrations of 10 to 1000 μM. As an example of the sensitivity of the procedure, the detection limit was 2 μM for benzoic acid [3]. Several reasonable methods are feasible for calculating the pK_a values from electrophoretic mobilities and buffer pH. In addition, we previously suggested two possible alternative calculation and experimental approaches. Corrections for potential discontinuities in electrophoretic mobilities between buffers were proposed by adding a fixed charge solute to the analysis mixture. In the other approach, an internal standard of known pK_a was added as an in situ approach to determine the pH of each buffer. These issues regarding the experimental and calculational approaches are examined in this investigation.

2. Theory

The determination is based on the principle that a solute has its maximum electrophoretic mobility when it is fully ionized, has no mobility in its neutral form, and has an intermediate, well modeled, mobility in the pH region surrounding

* Corresponding author.

SSDI 0021-9673(94)00270-O

its pK_a [1–3]. Electrophoretic mobility is calculated from the migration time of a neutral marker, t_{eof}, the migration time of the solute, t, the length of the column, L_c, the length of the column between the injection end and the detector, L_d, and the applied voltage, V, according to the relation

$$\mu = \left(\frac{L_c L_d}{V}\right)\left(\frac{1}{t} - \frac{1}{t_{oef}}\right) \quad (1)$$

There are several experimental approaches and models which apply. Derivations of these expressions were developed elsewhere [1–4].

2.1. Non-linear model

The non-linear model at 25°C is

$$pK_a = pH - \log\left(\frac{\mu}{\mu_{z^-} - \mu}\right) + \frac{0.5085 z^2 \sqrt{I}}{1 + 0.3281 a \sqrt{I}} \quad (2)$$

where pK_a is the thermodynamic pK_a, μ, the electrophoretic mobility at the pH of the buffer in the CE column, μ_{z^-} the electrophoretic mobility of the fully ionized acid, z the valency of the ion, I the ionic strength of the buffer solution, and a is the ion size parameter, generally unknown but assumed to be 5 Å [5]. The third term in the equation is equal to $-\log \gamma$, where γ is the activity coefficient of the ions in solution.

The analogous expression for a base, B is

$$pK_a = pH + \log\left(\frac{\mu}{\mu_{BH^+} - \mu}\right) - \frac{0.5085 z^2 \sqrt{I}}{1 + 0.3281 a \sqrt{I}} \quad (3)$$

The compounds used in this study were all bases and hence, only equations for bases will be shown for simplicity. Eq. 3 is rearranged for non-linear regression to

$$\mu = \frac{\mu_{BH^+} 10^{(pK_a - pH_c)}}{1 + 10^{(pK_a - pH_c)}} \quad (4)$$

The pH_c is the activity corrected pH:

$$pH_c = pH - \frac{0.5085 z^2 \sqrt{I}}{1 + 0.3281 a \sqrt{I}} \quad (5)$$

2.2. Linear model

The linear model, derived from the same equilibrium expressions as Eq. 3 is

$$\frac{1}{\mu} = \frac{1}{K \cdot \mu_{BH^+}} \cdot \frac{\gamma}{\{H^+\}} + \frac{1}{\mu_{BH^+}} \quad (6)$$

where $\{H^+\}$ is the hydrogen ion activity. The inverse of the intercept times the slope of the line gives K. The same equation is used for weighted linear regression. The standard deviations of the inverse mobilities are used for the weighting factors.

2.3. Use of ionic mobility reference

The electrophoretic mobility of a solute, modeled as a solid sphere, is usually expressed as

$$\mu = \frac{q}{6\pi\eta R} \quad (7)$$

where q is the net charge, η is the solution viscosity and R is the apparent hydrodynamic radius of the sphere. This equation is only valid in an infinitely dilute solution but we use it here to show the potential for experimental bias in pK_a determinations. In a plot of μ versus pH from experimental data, there may be individual values of μ which appear to either have random error or a bias. We refer to a bias in the mobility between separate run buffers as a discontinuous effect. Discontinuity in μ as measured from Eq. 7 would result from assumptions that q, η, and R are the same in all the buffers. For weak acids and bases, q is expected to change predictably as a function of pH. The R may be expected to change as a function of pH, a continuous effect, or buffer ion species, a discontinuous effect, although the direction of the change would be opposite for cations relative to anions. If η changes, then the effect can be quantitated and corrected by an external measurement. An approach was proposed to use an internal mobility

marker which is fully charged throughout the pH range of the experiment [3]. The anion, 4-toluenesulfonate (TSA), has a pK_a of approximately −7 and is thus fully charged throughout the experiment. The mobility of a weak base can be referenced to the mobility of the TSA and Eq. 3 becomes:

$$pK_a = pH + \log\left(\frac{\mu}{\mu_{BH^+} \cdot \frac{\mu_{TSA}'}{\mu_{TSA}} - \mu}\right) - \frac{0.5085z^2\sqrt{I}}{1 + 0.3281a\sqrt{I}} \quad (8)$$

The ratio, μ_{TSA}'/μ_{TSA} is referred to as the buffer discontinuity factor. The μ_{TSA} is the electrophoretic mobility of the TSA in the pH buffer. The μ_{TSA}' is the minimum electrophoretic mobility of the TSA in the whole buffer series. Since anions move in the opposite direction of the electroosmotic flow, their mobilities are negative and thus, the minimum electrophoretic mobility has the largest absolute value of the series. A buffer series consists of all the buffers used in a pK_a determination experiment. Thus, as an example, if the measured μ_{TSA} in a pH 6 buffer is lower than the μ_{TSA} values measured in the other pH buffers in the series, then the μ_{TSA} for the pH 6 buffer becomes μ_{TSA}'. If the only difference between all the μ_{TSA} values in the whole buffer series is the experimental error, then the inclusion of the buffer discontinuity factor in Eq. 8 will not have a significant impact on the pK_a determination.

2.4. Use of an internal electrophoretic mobility standard

The pH can be defined from the mobility of a second solute, a base, with a dissociation constant having the value pK_a'

$$pH = pK_a' - \log\left(\frac{\mu'}{\mu_{BH^+}' - \mu'}\right) + \frac{0.5085z^2\sqrt{I}}{1 + 0.3281a\sqrt{I}} \quad (9)$$

Substituting into Eq. 3 gives

$$pK = pK_a' - \log\left(\frac{\mu'}{\mu_{BH^+}' - \mu'}\right) + \log\left(\frac{\mu}{\mu_{BH^+} - \mu}\right) \quad (10)$$

In Eq. 10, it is interesting to note that the activity correction drops out.

3. Experimental

3.1. Instrument parameters

A SpectraPHORESIS 1000 (Thermo Separation Products, Fremont, CA, USA) was used for all experiments. Typically, a 2-s hydrodynamic injection was performed. Since the hydrodynamic injection rate is 6 nl/s for a 67 cm × 75 μm untreated fused-silica capillary (Polymicro Technology, Phoenix, AZ, USA), 12 nl was loaded onto the column. The separation distance, L_d, was 59.5 cm. The temperature was set at 25°C. UV absorption was monitored at 220 and 240 nm. With the instrument operating at 25 kV, typical currents were less than 20 μA.

In order to equilibrate the column and thereby minimize hysteresis effects, the following wash cycle was performed prior to each run in a sequence: (1) 2.5 min with 0.1 *M* NaOH, (2) 2.5 min with water, and (3) 3.0 min with running buffer.

Because the SpectraPHORESIS 1000 is equipped with a single reservoir for the buffer near the detector, it is not possible to match buffers at each end of the column in a sequence. Tricine (0.02 *M*, pH 7.6 to 8.1) was used as the buffer at the detector end of the column.

Buffer pH was measured using a Fisher Accuphast pH electrode with an Orion Model EA940 meter. Meter calibrations were made with Fisher NIST traceable buffer solutions.

3.2. Methods

All sample and buffer solutions were prepared using distilled, deionized, and filtered water (ASTM type I specification). The buffers used

were similar to those described in Refs. [3,4]. In the regression model evaluations, the sample consisted of 200 μM 2-aminopyridine and 1 mM mesityl oxide. In the experiments to investigate buffer discontinuity, the sample was 1 mM mesityl oxide, 100 μM α-methylbenzene and 75 μM toluenesulfonate. In the experiments with the internal standards to determine pH in situ, the sample was 1 mM mesityl oxide and 50 μM each of 2- and 3-ethylaniline.

Several routines were written in the Mathcad 4.0 (MathSoft, Cambridge, MA, USA) program to do the non-linear regressions, the linear regression and the weighted linear regression required as in Eqs. 4, 6, 8, 9 and 10. The weights for the weighted regression were the standard deviations of the inverse mobilities.

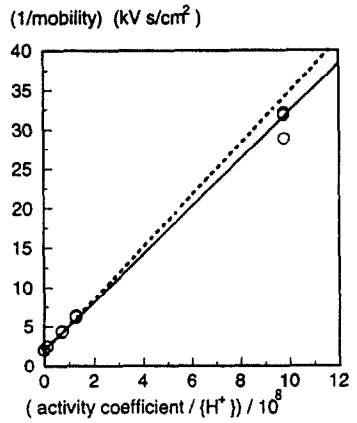

Fig. 1. Comparison of linear regression (solid line), weighted linear regression (dashed line) and the raw data (○).

4. Results

4.1. Evaluation of the different regression models

Repeatability of mobilities and migration times

At each of 6 pH values, 10 replicate determinations were performed for 2-aminopyridine. The mean electrophoretic mobilities, analyte migration times, and neutral marker migration times are listed in Table 1 along with their respective standard deviations. As the migration times of the neutral marker increased, the standard deviation also increased, indicating a potential problem with the precision of the parameters calculated from linear regression; linear regression assumes that the standard deviations of the y-values are constant.

Linear and weighted linear regression

Linear regression gave a pK_a value of 6.84 and weighted linear regression gave a pK_a value of 6.77 with $n = 60$. For comparison, the literature value is 6.71 [6]. The data are plotted in Fig. 1. The precision of the pK_a value at the 95%

Table 1
Mean and standard deviations of the electrophoretic mobilities and migration times of 2-aminopyridine and mesityl oxide

pH	Mean mobility (cm^2/Vs)	Standard deviation of mobility (cm^2/Vs)	Migration time (min)	Standard deviation of migration time (min)	Migration time of neutral marker (min)	Standard deviation of neutral marker migration time (min)
8.03	$-3.17 \cdot 10^{-5}$	$1.06 \cdot 10^{-6}$	3.46	$1.07 \cdot 10^{-2}$	3.57	$8.23 \cdot 10^{-3}$
7.15	$-1.60 \cdot 10^{-4}$	$1.93 \cdot 10^{-6}$	3.35	$1.75 \cdot 10^{-2}$	3.95	$1.66 \cdot 10^{-2}$
6.91	$-2.29 \cdot 10^{-4}$	$3.27 \cdot 10^{-6}$	3.11	$2.21 \cdot 10^{-2}$	3.89	$2.36 \cdot 10^{-2}$
6.16	$-4.10 \cdot 10^{-4}$	$3.07 \cdot 10^{-6}$	2.77	$4.22 \cdot 10^{-3}$	4.08	$1.25 \cdot 10^{-2}$
5.45	$-4.91 \cdot 10^{-4}$	$3.47 \cdot 10^{-6}$	2.80	$2.72 \cdot 10^{-2}$	4.62	$8.75 \cdot 10^{-2}$
4.88	$-5.10 \cdot 10^{-4}$	$1.38 \cdot 10^{-6}$	2.97	$3.60 \cdot 10^{-2}$	5.20	$1.12 \cdot 10^{-1}$

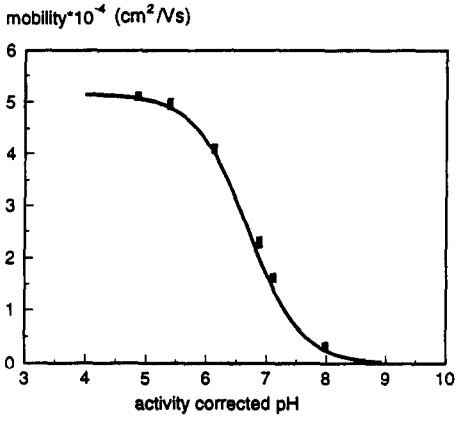

Fig. 2. Raw data and the calculated line based on the non-linear regression.

confidence level for linear regression was ±0.071, and for weighted regression the value was ±0.025. The linear regression was significantly influenced by the wider variance in mobilities in the lower pH buffers. The experiment was designed with buffer pH values equally spaced about the pK_a of the solute. In the inverse linear format of Eq. 6 required for the linear regression, this experimental design is unbalanced; hence, the lowest pH data has significant leverage over the rest of the data. Weighted regression removed this leverage by minimizing the significance of the data relative to its variance.

Non-linear regression

Non-linear regression using Eq. 4 takes advantage of the experimental design of equally spaced pH buffers about the pK_a of the solute and gives a pK_a of 6.76, in close agreement with the weighted linear regression. Fig. 2 shows the raw data and the calculated line for the non-linear regression.

4.2. Use of ionic mobility reference

Table 2 gives the pH values, the electrophoretic mobilities of α-methylbenzylamine and TSA, and the buffer discontinuity correction factors based on the TSA mobilities. With the exception of the pH 10.03 buffer, the correction factors appear to be very small relative to amounts which would make a significant difference in the final pK_a determination. Using all of the data, the pK_a value determined using the uncorrected mobilities was 9.43, and using the factored mobilities was 9.42. The discontinuity between buffers was not significant enough to impact pK_a determinations. Under the conditions of this experiment, there was not a measurable difference in electrophoretic mobility discontinuity between buffers with the exception of

Table 2
Activity corrected pH values and mobility data

Activity corrected pH	α-Methylbenzylamine mobility (cm^2/Vs)	TSA mobility (cm^2/Vs)	Discontinuity factor (μ_{TSA}'/μ_{TSA})
6.11	$-3.71 \cdot 10^{-4}$	$3.93 \cdot 10^{-4}$	1.06
6.86	$-3.66 \cdot 10^{-4}$	$3.94 \cdot 10^{-4}$	1.05
7.10	$-3.71 \cdot 10^{-4}$	$4.00 \cdot 10^{-4}$	1.04
7.49	$-3.57 \cdot 10^{-4}$	$3.86 \cdot 10^{-4}$	1.07
8.29	$-3.43 \cdot 10^{-4}$	$3.91 \cdot 10^{-4}$	1.06
8.60	$-3.16 \cdot 10^{-4}$	$3.97 \cdot 10^{-4}$	1.05
9.19	$-2.38 \cdot 10^{-4}$	$3.91 \cdot 10^{-4}$	1.06
9.55	$-1.50 \cdot 10^{-4}$	$3.95 \cdot 10^{-4}$	1.05
10.03	$-7.98 \cdot 10^{-5}$	$4.15 \cdot 10^{-4}$	1.00

The discontinuity factor is the ratio of the minumum mobility of the toluene sulfonic acid in the whole buffer series divided by the toluene sulfonic acid mobility at the pH of the run buffer

one buffer, pH 10.03 CAPS. The reason for the discontinuity for the one buffer is not understood and is likely to be an experimental error. The magnitude of the discontinuity does not impact the pK_a value determination.

4.3. Use of an internal electrophoretic mobility standard

Determination of pH via electrophoretic mobility

The electrophoretic mobility of a solute of known pK_a can, in theory, be used to determine pH. Table 3 shows the results of calculating pH from Eq. 9 using the mobilities obtained within the different buffers. The μ_{BH^+} was estimated as having a slightly greater mobility than the mobility at pH 3.05 where the solutes, 3- and 2-ethylaniline, should be fully ionized. The pK_a values of the solutes were already known. The errors in the calculated pH values along with the actual measured pH values corrected for their activities (see Eq. 5) are presented in Fig. 3. A more complete understanding of these errors can be obtained from propagation of errors analysis.

Propagation of errors in pH determination via electrophoretic mobility

Assuming the μ_{BH^+} and the pK_a to be constants in the determination of pH, the only variables are t and t_{eof}. From Eqs. 1, 3 and 5, pH may be expressed as

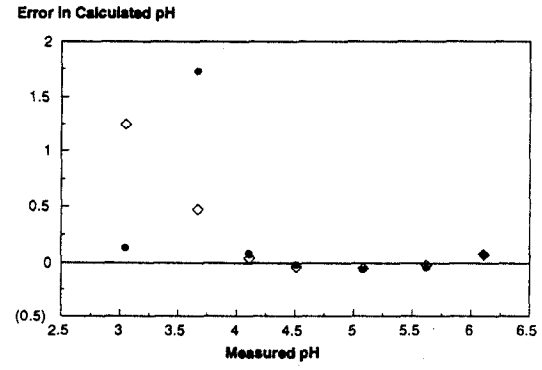

Fig. 3. Error in calculated pH from the mobilities of (●) 3- and (◇) 2-ethylaniline at different pH values.

$$\mathrm{pH} = pK_a - \log\left[\frac{\left(\frac{1}{t}-\frac{1}{t_{eof}}\right)}{\left(\frac{1}{t_{BH^+}}-\frac{1}{t}\right)-\left(\frac{1}{t}-\frac{1}{t_{eof}}\right)}\right] + \frac{0.5085z^2\sqrt{I}}{1+0.3281a\sqrt{I}} \quad (11)$$

The variance in pH determination is the sum of the variances due to each variable obtained from the experiment

$$\frac{\partial \mathrm{pH}}{\partial t} = \left(\frac{\partial \mathrm{pH}}{\partial t_{eof}}\right)^2 \sigma_{t_{eof}}^2 + \left(\frac{\partial \mathrm{pH}}{\partial t}\right)^2 \sigma_t^2 \quad (12)$$

where σ is the standard deviation. For the purposes of this error propagation, t_{eof} is assumed to be a constant.

Table 3
Calculated pH values from electrophoretic mobilities

pH Measured less	pH calculated from	
The activity correction	3-Ethylaniline mobility	2-Ethylaniline mobility
3.05	2.92	1.81
3.67	1.94	3.20
4.11	4.03	4.07
4.51	4.53	4.54
5.08	5.14	5.13
5.62	5.66	5.64
6.11	6.03	6.03

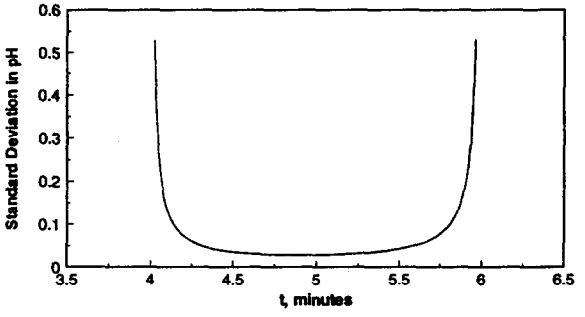

Fig. 4. Estimated standard deviation in the determination of pH by CE versus migration time where $t_{eof} = 6$ min, $\sigma_{t_{eof}} = 0.03$, $\sigma_t = 0.03$ and $t_{BH^+} = 4$ min.

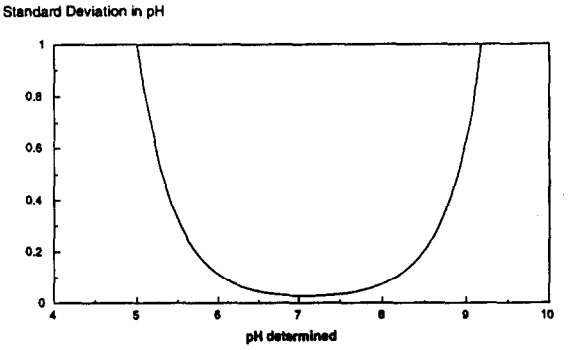

Fig. 5. Estimated standard deviation in pH versus pH measured for a solute with a pK_a equal to 7.

From Eq. 12, the partial derivatives of pH with respect to t and t_{eof} are

$$\frac{\partial pH}{\partial t} = \frac{(t_{BH^+} - t_{eof})}{2.303(t - t_{BH^+})(t - t_{eof})} \quad (13)$$

$$\frac{\partial pH}{\partial t_{eof}} = \frac{-1}{t_{eof}^2\left(\frac{1}{t} - \frac{t}{t_{eof}}\right)\ln(10)} \quad (14)$$

Substituting Eqs. 13 and 14 into Eq. 12, the variance in pH can be expressed as in Eq. 15. Using typical values ($pK = 7$, $t_{eof} = 6$ min, $\sigma_{t_{eof}} = 0.03$, $\sigma_t = 0.03$, $t_{BH^+} = 4$ min), the standard deviation of the error in pH determination is plotted in Fig. 4.

A more general way of expressing the error in the determined pH is to use actual pH values for the x-axis (Fig. 5).

As can be seen from Figs. 4 and 5, the estimated standard deviation in the calculated pH is high in the extremes where t is near t_{eof} and t_{BH^+}. This corresponds to ± 1 pH unit from the pK_a. Hence, when a solute of known pK_a is used to determine the pH, the most reliable determination will result from using mobilities taken in buffers which are within ± 1 pH unit of the pK_a values of the solutes.

pK_a determination using a second solute of known pK_a

The pK_a of 3-ethylaniline was determined with Eq. 10 using the mobilities of 2-ethylaniline at different pH values. These mobilities are listed in Table 4. The pK_a as determined by non-linear regression was 4.35, compared to a literature value of 4.37, was very sensitive to reasonable initial estimates of the μ_{BH^+}, the μ_{BH^+}' and the pK_a. The dependence on the initial estimates was so sensitive that pK_a values at least as wide as 4.1 to 4.6 could be determined with a minimized sum of squares. Using known pH values determined by a pH electrode, the procedure is insensitive to reasonable initial values of the parameters required for the non-linear regression analysis. Indeed, the variables μ and μ' are statistically highly correlated with each other, thus the regression analysis becomes mathematically unstable yielding large uncertainties in the parameter estimates.

5. Conclusions

Non-linear regression is the simplest and most precise regression procedure for determining the pK_a value of a solute by CE. Weighted regres-

$$\sigma_{pH}^2 = \frac{t^4\sigma_{t_{eof}}^2 - 2t^3\sigma_{t_{eof}}^2 t_{BH^+} + t^2\sigma_{t_{eof}}^2 t_{BH^+} + t_{eof}^4\sigma_t^2 - 2t_{eof}^3\sigma_t^2 t_{BH^+} + t_{eof}^2\sigma_t^2 t_{BH^+}}{t_{eof}^2(t - t_{eof})^2\ln(10)^2(t - t_{BH^+})^2} \quad (15)$$

Table 4
Activity correction factor for solution ionic strength, mobilities (cm^2/Vs), calculated pH based on the mobility of 2-ethylaniline and its known pK_a, and actual pH as measured from a pH meter

$-\log(\gamma_{BH^+})$	3-Ethylaniline mobility (cm^2/Vs)	2-Ethylaniline mobility (cm^2/Vs)	pH calculated from 2-ethylaniline mobility	pH measured
0.03	$3.74 \cdot 10^{-4}$	$3.50 \cdot 10^{-4}$	1.81	3.08
0.02	$3.79 \cdot 10^{-4}$	$3.30 \cdot 10^{-4}$	3.20	3.69
0.02	$3.16 \cdot 10^{-4}$	$2.38 \cdot 10^{-4}$	4.07	4.13
0.02	$2.32 \cdot 10^{-4}$	$1.45 \cdot 10^{-4}$	4.54	4.53
0.04	$1.08 \cdot 10^{-4}$	$5.64 \cdot 10^{-5}$	5.13	5.12
0.06	$4.21 \cdot 10^{-5}$	$2.04 \cdot 10^{-5}$	5.64	5.68
0.04	$1.85 \cdot 10^{-5}$	$8.14 \cdot 10^{-6}$	6.03	6.15

sion gives precise values; however, it requires many replicates to determine the standard deviations of the electrophoretic mobilities at different pH values. Linear regression is the least precise procedure because of its sensitivity to influential outliers. An experimental design appropriate for non-linear regression is to space the buffer pH values equal distances around the pK_a of the solute. Linear regression would require an experimental design using equally spaced ionic activity coefficients divided by the proton activity around the K of the solute. This design would not be linear with respect to the fraction of ionization. In the investigation of solutes of unknown pK_a, it is important to cover a wide pH range. This is most easily accomplished with a series of equally spaced pH buffers which are linear with respect to the fraction of ionization. Hence, non-linear regression is the recommended statistical procedure.

Correcting for a potential experimental bias or discontinuity in electrophoretic mobilities between different buffers in a series of buffers was investigated by using a solute with a constant mobility in all pH buffers. If a chemical or physical effect of the buffer were to cause a change in hydrodynamic radius of the solute, or if the buffers varied in viscosity, then this effect could be quantitatively corrected. In a test of this hypothesis, there was not a significant change in the mobilities of the totally ionized solute versus pH. Hence, the same pK_a, within statistical significance, was determined with both corrected and raw mobilities. For simplicity, it is recommended that future pK_a determinations do not include a constant mobility marker when using experimental conditions similar to those in this report.

The internal standard procedure of using a solute with a known pK_a for determining an unknown solute's pK_a value by CE is not recommended because it yields less precise pK_a determinations than the normal means of using buffers with their pH measured with a pH meter. The source of the imprecision is the narrow range of accurate pH prediction from the internal standard as demonstrated through experimentation and a propagation of error study. Indeed, the best pH prediction by this procedure would inherently rely on the accuracy of the literature pK_a value of the internal standard. A second reason for the technique's imprecision is the high correlation between the electrophoretic mobilities of the reference solute and the unknown solute. The recommended procedure remains to measure the pH of the running buffers with a pH meter.

References

[1] J.M. Beckers F.M. Everaerts and M.T. Ackermans, *J. Chromatogr.*, 537 (1991) 407.
[2] J. Cai, J.T. Smith and Z. El Rassi, *J. High Resolut. Chromatogr.*, 15 (1992) 30.
[3] J.A. Cleveland, M.H. Benko, S.J. Gluck and Y.M. Walbroehl, *J. Chromatogr. A*, 652 (1993) 301.
[4] S.J. Gluck and J.A. Cleveland, *J. Chromatogr. A*, 680 (1994) 43.
[5] J. Kielland, *J. Am. Chem. Soc.*, 59 (1937) 1675.
[6] D.D. Perrin, *Dissociation Constants of Organic Bases in Aqueous Solution*, Butterworths, London, 1965.