ELSEVIER

Journal of Chromatography A, 927 (2001) 19–30

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Prediction of internal standards in reversed-phase liquid chromatography
1. Initial study on predicting internal standards for use with neutral samples based on linear solvation energy relationships

Jianwei Li*

Analytical Research and Development, 3M Pharmaceuticals, 3M Center, Building 235-BE-45, St Paul, MN 55144, USA

Received 21 February 2001; received in revised form 7 May 2001; accepted 11 July 2001

Abstract

This paper describes the results of an initial study on the application of linear solvation energy relationships (LSERs) to the prediction of internal standard compounds in reversed-phase liquid chromatographic (RPLC) method development. Six neutral samples are separated on an Inertsil ODS(3) column by either acetonitrile–water or methanol–water mobile phases under either isocratic or linear gradient conditions. After the separation conditions are optimized, the desired positions for internal standard candidates are selected based on the ''open windows'' of the chromatograms. The compounds with the desired retention range are then predicted based on LSERs from a database consisting of more than 700 compounds with defined physicochemical properties. The prediction requires the use of LSER coefficients under the separation conditions for each sample. They are determined a priori by performing multivariable linear regression on the retention of 20 reference solutes against their physicochemical properties. It can be concluded from the study that LSER is an excellent approach to the selection of internal standard compounds for RPLC under either isocratic or gradient elution. The average prediction error is usually within 10%, but no more than 20%. Finally, LSER approach is fast and systematic, and will save a significant amount of time and resources during RPLC method development. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Internal standard; Linear solvation energy relationships; Method development; Retention prediction

1. Introduction

The precision and accuracy of quantitative chromatographic determinations depend on many factors such as the accuracy of the execution of the sample preparations, instrument precision and accuracy, robustness of analytical methods, and the accuracy of chromatographic peak height or area determination. Although the skill and technique of the analytical chemist are critical in ensuring the accuracy of the execution of sample preparation and data processing, it is equally important that all operating conditions of the instrument are maintained precisely during an experiment and are reproduced from experiment to experiment. Specifically, the design of the basic components of the chromatographic instrument, proper selection of column and detector, and precision control over operating conditions are factors that contribute to the reproducibility and accuracy of the

*Tel.: +1-651-737-0468; fax: +1-651-737-7918.
E-mail address: jli7@mmm.com (J. Li).

0021-9673/01/$ – see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0021-9673(01)01109-8

analytical result. Accordingly, it is always desirable to use a robust chromatographic method to reduce the effect of small variations. The robustness of a method can be achieved by including procedures for the improvement of both precision and accuracy.

A frequently used approach to improving method accuracy and robustness is the use of an internal standard [1]. The internal standard technique is based on the hypothesis that both instrumental variations (in injection volume, flow-rate, column condition, and temperature) and sample preparation errors (such as dilution and liquid phase extraction) can be compensated [2–6]. A more advanced use of an internal standard is to improve the recovery of solutes if the matrix effect is significant or multiple steps are involved in the sample preparation. Although LC instrumentation has been improved significantly in the last two decades, the internal standard methodology is still a frequent tool to control the robustness and accuracy of chromatographic methods. Finally, the internal standard method has become a very popular technique not only in chromatography, but also quantitative LC–MS [7] and NMR [8,9].

The rationale for the use of internal standard can be explained with a simple chromatographic theory. Peak height and area are usually the responses used in quantitative chromatography and given as [10]:

$$H = \left(\sqrt{\frac{N}{2\pi}}\right)\frac{C^0 V_{\mathrm{Inj}}}{t_0(1+k')F} \tag{1}$$

$$A = \tau C^0 = \left(\frac{V_{\mathrm{Inj}}}{F}\right)C^0 \tag{2}$$

where H is peak height, A is the peak area, t_0 is the column dead time, τ is the injection time determined by the ratio of injection volume (V_{Inj}) over flow-rate (F), C^0 is the solute concentration, N denotes column efficiency, and k' is the retention factor. If an internal standard is used, the ratios in peak height (PHR) and area (PAR) are given as:

$$\mathrm{PHR} = \frac{H_j}{H_i} = \left(\frac{C_j^0}{C_i^0}\right)\frac{(1+k_i')}{(1+k_j')} \tag{3}$$

$$\mathrm{PAR} = \frac{A_j}{A_i} = \left(\frac{\tau}{\tau}\right)\left(\frac{C_j^0}{C_i^0}\right) = \left(\frac{C_j^0}{C_i^0}\right) \tag{4}$$

where j and i denote the solute of interest and internal standard, respectively. It can be seen from Eq. (3) that the ratio of peak height is essentially the ratio of solute concentrations in the sample corrected by the ratio of retention factors. Injection volume, flow-rate, and column efficiency are canceled. Similarly, the ratio of peak area is just the ratio of solute concentrations in the sample. Thus, it can be concluded that the reproducibility of injection volume and the accuracy of flow-rate can be compensated by the internal standard. By similar arguments, the effect of temperature variation (change in retention and peak height, for example), column instability (change in column efficiency, N) can also be compensated by internal standards [6]. Finally, the same principle applies to the compensation of sample preparation errors.

Despite the importance of an internal standard in quantitative chromatography, the search for an internal standard compound for a specific separation is often very difficult. The selection of an internal standard for a quantitative chromatographic method requires careful planning and experimental justification. The internal standard peak must be completely resolved from sample component peaks. For application of an internal standard method to quantitative determinations of drugs, potential interference (metabolites and degradation products) with the peak of the candidate internal standard should be studied. Due to the resolution requirement for an internal standard candidate, a convenient way to insert an internal standard in a chromatogram has always been a challenge to the analytical chemist. Most of the selections have been accomplished by an empirical approach. Depending on the requirements for resolution and analysis time, this procedure can be very time-consuming and requires a lot of effort. A satisfactory internal standard may sometimes be obtained without expensive investigation, when homologues and/or analogues of the sample component can be obtained commercially or synthesized conveniently.

There are limited studies on the prediction of internal standard compounds in chromatography. The first approach focused on the use of homologous series. These series included anilides [11], alkylbenzamides [12,13], and alkyl aryl phenones [14]. In RPLC, the logarithm of retention factor of a homolo-

gous series is linearly related to a carbon chain length. Thus, the retention of a homologue can be predicted based on the retention data of other homologous members. This homologue can be subsequently evaluated for an internal standard candidate under the separation conditions for a sample.

Yamauchi et al. studied the use of phenol and nitrobenzene derivatives for internal standards in RPLC [15–17]. Many phenol and nitrobenzene derivatives were carefully selected for potential internal standard candidates. They were chromatographed under different mobile phase compositions, and subsequently arranged based on their elution order. When a separation is requested for an internal standard, the location of an internal standard in the chromatogram and retention times of these derivatives are matched to suggest possible internal standard candidates.

A more rigorous prediction of an internal standard was suggested by Skelly et al. [18]. They generated a database of 90 potential internal standards arranged in a retention time order. The retention times were obtained by gradient elution RPLC with acetonitrile (ACN)–water as the mobile phase on a reference column. Moreover, a marker solution is needed to help the selection of internal standard candidates. The marker solution has four compounds (acetanilide, α-tetralone, benzophenone, and *trans*-stilbene). When a selection of an internal standard compound is needed for a sample, the sample and marker solutions are injected under the same gradient conditions (the column may be different). The sample chromatogram is then examined for an ''open window'' (a clean region in a chromatogram). The retention times of the database are first adjusted based on the retention times of the four solutes in the marker solution, and the prediction of internal standard candidates is then made based on the ''open window''.

The homologous series may have the least application for internal standards if the resolution requirement is too tight because the change in retention may be too large for one carbon unit shift in the homologues. The phenol and nitrobenzene derivatives are usually very good internal standard candidates due to their availability and UV profiles. However, the retention of the compounds under different mobile phase conditions should be available before prediction can be made. The approach of Skelly et al. is more accurate, but it was developed for ACN/water mobile phase only. The theory used in this approach was not totally accurate, and therefore, accuracy of the prediction may suffer. Accordingly, all the studies were not based on the quantitative relationships between retention in chromatography and the properties of solutes.

This study evaluates the application of LSERs for much more accurate prediction of internal standards. Based on the desired position for an internal standard in a sample chromatogram, the possible candidates can be predicted very quickly. This technique requires the establishment of the dependence of retention on the physicochemical properties of solutes *prior to* the prediction. However, the relationship can be easily determined for both isocratic and gradient elution using a set of judiciously selected reference solutes. It has been well established both theoretically and experimentally that retention in chromatography can be linearly correlated with the solutes' properties. Moreover, this technique applies to both normal and reversed-phase liquid chromatography.

This is an initial study on the application of LSERs for the prediction of internal standards. Although other properties of compounds such as the UV profiles, purity, toxicity, and availability are also important, they are not the focus of the study. This study will mainly evaluate the accuracy of the LSER approach for the prediction of internal standards. Although in this study, neutral samples are evaluated under either isocratic or gradient condition, this approach will be applied to other types of RPLC separations. We have established a database consisting of more than 700 compounds with defined properties for the selection of candidate internal standards. We will mainly select aromatic compounds as internal standard candidates due to their UV profiles.

2. Theoretical

The prediction of internal standard compounds in this study is based on LSERs. It has been well established both theoretically and experimentally that the retention in isocratic chromatography can be

correlated with the physicochemical properties of solutes [19–43]:

$$\mathrm{Ln}(k') = \mathrm{Ln}(k'_0) + m\frac{V_2}{100} + s\pi_2 + a\Sigma\alpha_2^{\mathrm{H}} + b\Sigma\beta_2^{\mathrm{H}} \quad (5)$$

where the subscript 2 denotes a solute, $\mathrm{Ln}(k'_0)$ is the regression constant, V_2 is the solute McGowan characteristic molar volume, π_2 is the solute dipolarity/polarizability, $\Sigma\alpha_2^{\mathrm{H}}$ and $\Sigma\beta_2^{\mathrm{H}}$ are the solute ''overall'' or ''effective'' hydrogen bond acidity and basicity, respectively, and m, s, a and b are the coefficients determined by multiple linear regression analysis. V_2, π_2, $\Sigma\alpha_2^{\mathrm{H}}$ and $\Sigma\beta_2^{\mathrm{H}}$ are called the solutes' descriptors, and they represent the physicochemical properties of solutes. It is noted that the coefficients in Eq. (5) depend on the combination of the type of column and the mobile phase composition. It is noted that Eq. (5) applies to only neutral solutes and its error is usually within 0.04–0.06 log unit.

Based on Eq. (5), the chromatographic selectivity (α) can be described as follows:

$$\begin{aligned}\mathrm{Ln}(\alpha) &= \mathrm{Ln}\left(\frac{k'_j}{k'_i}\right) \\ &= m\left(\frac{V_j - V_i}{100}\right) + s(\pi_j - \pi_i) + a\left(\Sigma\alpha_j^{\mathrm{H}} - \Sigma\alpha_i^{\mathrm{H}}\right) \\ &\quad + b\left(\Sigma\beta_j^{\mathrm{H}} - \Sigma\beta_i^{\mathrm{H}}\right) \quad (6)\end{aligned}$$

where k'_j and k'_i are the retention factors of solutes j and i, respectively; V_j, π_j, $\Sigma\alpha_j^{\mathrm{H}}$ and $\Sigma\beta_j^{\mathrm{H}}$ are the descriptors for solute j; and V_i, π_i, $\Sigma\alpha_i^{\mathrm{H}}$ and $\Sigma\beta_i^{\mathrm{H}}$ are the descriptors for solute i. Eq. (6) indicates that, although the absolute retention depends on the intercept, the chromatographic selectivity is not affected. This implies that if two columns have different carbon loads, the selectivity may not be affected for neutral solutes.

For linear gradient elution, retention time (not $\mathrm{Ln}(k')$) also linearly correlates well with the properties of solutes [44]:

$$t = I + m'\frac{V_2}{100} + s'\pi_2 + a'\Sigma\alpha_2^{\mathrm{H}} + b'\Sigma\beta_2^{\mathrm{H}} \quad (7)$$

where I is the intercept; and m', s', a', and b' are also LSER coefficients, but different from those in Eqs. (5) and (6). The relative error of Eq. (7) is usually within 10%. The difference in retention time for two neutral solutes is given as:

$$\begin{aligned}t_j - t_i &= m'\left(\frac{V_j - V_i}{100}\right) + s'(\pi_j - \pi_i) \\ &\quad + a'\left(\Sigma\alpha_j^{\mathrm{H}} - \Sigma\alpha_i^{\mathrm{H}}\right) + b'\left(\Sigma\beta_j^{\mathrm{H}} - \Sigma\beta_i^{\mathrm{H}}\right) \quad (8)\end{aligned}$$

where all symbols designate the same meaning as in Eq. (6).

Obviously, before the prediction of internal standards can be made, we need to first determine the LSER coefficients. After the column and mobile phase compositions are finalized for a separation, the retention ($\mathrm{Ln}(k')$ or time) for a set of judiciously selected reference solutes is measured. By performing multivariable linear regression between retention and solutes' descriptors (Eqs. (5) and (7)), LSER coefficients (m, s, a and b or m', s', a' and b') can be obtained for the specific column and mobile phase composition. Alternatively, the LSER coefficients for isocratic elution on C_{18} columns can be calculated based on literature data [45]. The second requirement for the prediction is a database that includes the candidate compounds. This database should include as many compounds as possible with measured physicochemical properties. Aromatic compounds are preferred due to their UV profiles. A database consisting of more than 700 compounds has been established based on literature data [46].

The prediction of the internal standards can be performed in two ways based on the mathematical relationships above, depending on if the absolute or relative retention is used. If the LSER coefficients are obtained directly for the column and mobile phase composition, both prediction methods should generate the same results. However, if LSER coefficients are obtained indirectly by, for example, literature data, the prediction based on the relative retention should be more accurate.

2.1. Prediction based on absolute retention

After the LSER coefficients are obtained, the retention of any neutral solute can be predicted if the properties of the solute are known. When a sample chromatogram is optimized during method development, a decision should be made as to the position of the internal standard. Once the desired retention is

determined, the retention of several hundred compounds in a database can be easily computed by a spreadsheet program based on Eqs. (5) and (7). The retention is next sorted in either increasing or decreasing order. Then the compounds that meet the retention requirement (for example, ±10% of the target value) are selected as internal standard candidates. These compounds are then acquired and their retention times determined experimentally. The compound with the closest retention to the desire position will be selected as the internal standard if other conditions are equal.

2.2. Prediction based on relative retention (or selectivity)

Based on Eqs. (6) and (8), the criterion of the prediction is based on the selectivity of the internal standard candidate relative to a reference compound for isocratic separations (or relative retention time for linear gradient elution). After the separation conditions are optimized for a sample, a reference solute (e.g. phenol or nitrotoluene) and its retention should also be obtained. Based on the retention of the reference solute and desired position of an internal standard, their retention factors and selectivity are then computed for isocratic separations. For linear gradient elution, retention time difference is determined. The selectivity or retention time difference is then the criterion for the selection of compounds as candidates of the desired internal standard.

Once the criterion is determined, the selectivity (Eq. (6)) or relative retention time (Eq. (8)) of several hundred compounds in a database can be easily computed by a spreadsheet program. The selectivity is next sorted in either increasing or decreasing order. The compounds that meet the selectivity or relative retention requirement (for example, ±10% of the target value) are selected as the internal standard candidates. There could be many compounds to meet the requirement. However, only a small subset of these (mainly aromatic compounds) is examined for suitability. These compounds are then acquired and their retention times determined experimentally. The compound with the closest retention will be the internal standard if other conditions are equal.

It should be recognized that there is always an error associated with the prediction and a range of compounds should be selected and evaluated experimentally. The best compound is finally chosen by experiment.

It is obviously important to estimate the model error for the prediction. It can be seen in Table 3 that the relative standard error in absolute retention time for gradient elution is usually within 10%. This level of error has been observed before [44]. However, for isocratic elution, the relative error in retention factor can be as high as 20% (Table 3). The estimation of the model error in selectivity or relative retention time (Eqs. (6) or (8)) will be more difficult because it depends on the difference in retention (by error propagation analysis). The model error for the relative retention may be larger, and the absolute retention models (Eqs. (5) and (7)) should be more accurate.

3. Experimental

3.1. Chromatographic instrumentation and separation conditions

All experiments were performed on a HP 1090 chromatograph equipped with a ternary pump, autosampler, and diode-array detector (Agilent technologies, Wilmington, DE). A computer-based workstation (HP Chemstation) was used not only to control the instrumentation, but also to collect chromatographic data. An Inertsil ODS(3) column (150×4.6 mm, 5 μm) (Metachem Technologies, Torrance, CA) was used for all separations. The mobile phases were ACN/water and MeOH/water. Both isocratic and gradient separation conditions were evaluated, depending on the nature of the sample. The samples and their separation conditions are summarized in Table 1.

3.2. Prediction of internal standards

After the separation conditions were optimized for a sample, the LSER coefficients were then determined before the prediction of internal standard candidates under the same separation conditions. The LSER coefficients were determined by multivariable linear regression between Ln(k') or retention time of

Table 1
Separation conditions for the samples used in the study

Solutes in sample[a]	Sample identification[b]	Separation conditions			
		Mobile phase composition	Flow rate (ml/min)	Injection volume (μl)	Detection wavelength (nm)
Phenol Ethylparaben (~1 mg/ml in ACN)	A: Isocratic	40% ACN	2	20	280
Estrone Estriol	B: Isocratic	40% ACN	2	20	230
Predisone 4-Androstene-3,17-dione (~0.3 mg/ml in 50:50 ACN/MeOH)	B: Gradient	30–100% MeOH in 15 min	2	20	230
4-Chlorotoluene Propylbenzene (~0.5 mg/ml in ACN)	C: Isocratic	75% ACN	2	15	230
4′-Aminoacetophenone Acetophenone (~0.3 mg/ml in ACN)	D: Isocratic	40% MeOH	2	15	230
Benzophenone 2-Phenylphenol 4-Isopropyl-3-methylphenol (~0.3 mg/ml in ACN)	E: Isocratic	65% MeOH	2	15	230
3,5-Dimethylphenol 4-Chloroacetophenone 3-Methylacetophenone 4-Chlorobenzoic acid (~0.5 mg/ml in ACN)	F: Gradient	20–100% ACN in 30 min	2	10	230

This table summarizes the separation conditions for the samples. There are two gradient separations. The column was Inertsil ODS (3) and the instrumentation was an HP 1090 chromatograph.

[a] The value in the parenthesis is the approximate concentration.

[b] Sample identification by a simple letter for convenience. Also indicated in this column is the separation mode.

reference solutes against the their descriptors (Eqs. (5) or (7)). Twenty reference solutes (Aldrich, Milwaukee, WI) were used in the study, and they are indicated in Table 2. The solutions of the reference solutes were prepared in ACN (concentrations ~0.10 to 0.5 mg/ml), and they are stable for up to 2 years in refrigerated condition. Table 3 summarizes the LSER coefficients.

The prediction for isocratic separations was based on the selectivity (relative to phenol) for isocratic elution, while the absolute retention time was used for gradient elution. The prediction and experimental results are shown in Table 4.

3.3. Generation of the database

The database of 700 compounds was easily established by scanning the tables in Ref. [46] into Microsoft Word (then copy to Excel) or Excel. About 50% of the compounds are aromatic compounds, and they were mainly used in the study. This database is also available from the author. A very large database (>3000 compounds) is available through Sirius Analytical Ltd (East Sussex, UK) that has developed a software package that includes the prediction of solute descriptors [52,53].

4. Results and discussion

4.1. Isocratic separations

Fig. 1 shows the chromatogram for the separation of phenol and ethylparaben with 40% ACN as the

Table 2
Reference solutes used in the study and their physicochemical properties

Solute[a]	$(V/100)$[b]	(π^*)[c]	$(\Sigma\alpha^H)$[d]	$(\Sigma\beta^H)$[e]
Acetophenone (1)	1.014	1.00	0	0.48
Aniline (2)	0.816	0.96	0.26	0.41
Anisole (3)	0.916	0.75	0	0.29
Benzaldehyde (4)	0.873	1.00	0	0.39
Benzene (5)	0.716	0.52	0	0.14
Benzonitrile (6)	0.871	1.11	0	0.33
Benzyl acetate (7)	1.214	1.06	0	0.65
Benzyl cyanide (8)	1.012	1.15	0	0.45
Bromobenzene (9)	0.891	0.73	0	0.09
Butylbenzene or propylbenzene (10)	1.28/1.139	0.51	0	0.15
p-Chloroanisole (11)	1.038	0.86	0	0.24
4-Chlorophenol (12)	0.898	1.08	0.67	0.2
p-Cresol (13)	0.916	0.87	0.57	0.31
Fluorobenzene (14)	0.734	0.57	0	0.10
Methyl benzoate (15)	1.073	0.85	0	0.46
2-Nitrotoluene (16)	1.032	1.11	0	0.27
4-Nitrotoluene (17)	1.032	1.11	0	0.28
Phenol (18)	0.775	0.89	0.60	0.30
3-Phenyl-1-propanol (19)	1.198	0.90	0.30	0.67
Propiophenone (20)	1.154	0.95	0	0.51

This table summarizes the physicochemical properties of reference solutes used to determine the LSER coefficients. They are also called solvatochromic parameters [46–51] or simply descriptors.

[a] The number in parenthesis indicates the assigned solute number for convenience.

[b] Solute's McGowan characteristic molar volume [46].

[c] Solute's dipolarity/polarizability [46].

[d] Solute's effective acidity [46].

[e] Solute's effective basicity [46].

Table 3
LSER coefficients obtained for each sample

Sample no.	Elution type	LSER coefficients[a]					R^2[b]	RSE[c] (%)
		Int.	m/m'	s/s'	a/a'	b/b'		
A	Isocratic	−0.17	4.59	−0.73	−1.24	−4.99	0.996	8
B	Isocratic Gradient	0.28	17.81	−4.68	−1.70	−14.77	0.988	5
C	Isocratic	−0.56	2.10	−0.55	−1.02	−2.34	0.988	9
D	Isocratic	−0.98	7.83	−1.85	−0.68	−6.01	0.979	19
E	Isocratic	−0.93	4.68	−1.34	−0.76	−4.12	0.986	17
F	Gradient	0.94	19.78	−3.07	−4.72	−20.63	0.996	4

This table summarizes the LSERs coefficients obtained for the separation conditions in Table 1 by the solutes in Table 2.

[a] LSER coefficients for either isocratic or gradient elution.

[b] Correlation coefficient.

[c] Relative Standard Error in retention time or retention factor computed by the SE of the linear regression over the average of retention. Specifically, for gradient elution, RSE is computed as: $\mathrm{RSE} = 100\left(\frac{\mathrm{SE}}{\bar{t}}\right)$ where $\bar{t}$ is the average retention of all 20 solutes. However, for isocratic elution, it is computed as: $\mathrm{RSE} = 100\left(\frac{\delta k'}{\bar{k'}}\right) = 100\delta[\mathrm{Ln}(k')] = 100\mathrm{SE}$.

Table 4
Results of prediction of internal standards for each sample

Sample indication[a]	Desired selectivity or retention	IS predicted[b]	Prediction accuracy		
			Predicted	Measured	Average error[c]
A	$\alpha_{IS/Phenol}$ =	3-Methoxyphenol	1.3	1.02	20
(Fig. 1)	1.4	4-Fluorophenol	1.4	1.24	
		2-Fluorophenol	1.5	1.25	
B	$\alpha_{IS/Phenol}$ =	3,5-Dimethylphenol	2.9	2.7	5
(Figs. 2 and 6)	3.0	2,5-Dimethylphenol	3.0	3.1	
		2,3-Dimethylphenol	3.1	3.0	
	t_{IS} = 9.0 min	Benzene	8.5	8.88	8
		2,6-Dimethylaniline[d]	8.5	7.08	
		2,6-Dimethylphenol	9.0	8.11	
		1-Naphthol	9.2	9.00	
		2-Nitrotoluene	9.5	9.23	
C	$\alpha_{IS/Phenol}$ =	Benzyl bromide	3.80	3.69	8
(Fig. 3)	4.0	Toluene	3.91	4.31	
		Bromobenzene	4.20	4.79	
D	$\alpha_{IS/Phenol}$ =	4-Fluorophenol	1.48	1.41	21
(Fig. 4)	1.5	3-Ethylpyridine[e]	1.48	2.45	
		Benzaldehyde	1.53	1.71	
		3-Methoxyphenol	1.67	1.13	
E	$\alpha_{IS/Phenol}$ =	3,5-Dimethylphenol	3.20	2.79	12
(Fig. 5)	3.3	2-Iodophenol	3.36	2.95	
		1-Naphthol	3.38	3.58	
		4′-Methylacetophenone	3.39	3.04	
		2,3-Dimethylphenol	3.43	2.91	
		2,6-Dimethylphenol	3.46	2.97	
		Methyl benzoate	3.47	3.54	
F	t_{IS} = 10.0 min	2,6-Dimethylphenol	9.5	9.72	5
(Fig. 7)		1-Naphthol	9.8	10.87	
		Methyl benzoate	10.1	10.28	
		3′-Methylacetophenone	10.6	10.09	
		Benzene	10.6	11.12	

The table summarizes the prediction results and comparison to experimental values.

[a] The figure in parenthesis shows the corresponding chromatogram for the sample. Absolute retention times for predicted compounds are indicated in the figures.

[b] Predicted compounds as internal standard candidates.

[c] Average error (%) computed by: $100\left|\frac{\alpha_{Predicted}-\alpha_{Measured}}{\alpha_{Measured}}\right|$ or $100\left|\frac{t_{Predicted}-t_{Measured}}{t_{Measured}}\right|$.

[d] It may not be neutral under mobile phase conditions, so the prediction error can be large.

[e] This compound is not included in the calculation, and may undergo secondary interactions with residual silanols.

mobile phase. The position of the internal standard candidate is selected at 2.8 min based on the ''open window'' in the middle of the chromatogram. Relative to the position of phenol, the desired selectivity is 1.4. Table 4 shows both the predicted and measured selectivity for three candidate compounds. Their retention times are also indicated in Fig. 1. Also included in Table 4 are the average errors of the prediction. It can be seen in Table 4 that, although the absolute difference in selectivity is not significant, the average prediction error is about 20%. This level of error is within the range of the LSER model (Eq. (6)). However, by examining the chromatogram in Fig. 1, the retention of 2-fluorophenol is clearly very close to the desired internal standard position. This compound is very stable, and has an excellent UV profile. It can be used as the internal standard. It should be noted that the relative prediction error may

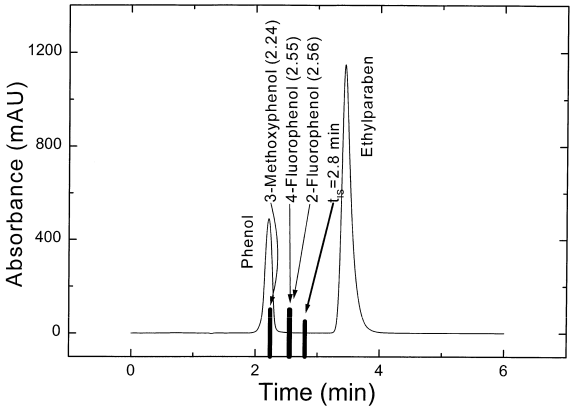

Fig. 1. Chromatogram for the separation of phenol and ethylparaben (sample A) by ACN/water mobile phase under isocratic conditions. The separation conditions are indicated in Table 1. The retention of predicted compounds is also included in the figure. Short bar indicates the position of desired internal standard. The desired position of the internal standard is at 2.8 min.

be larger for early eluting peaks due to relatively high measurement error in retention time. This example also indicates that if the predicted compounds are not matching the desired position, the predicted range can be increased.

Fig. 2 shows the chromatogram for the separation of the steroid mixture (sample B, 40% ACN). The polarity difference of the mixture is quite large, resulting in a significant discontinuity (''open window'') in the elution order. The internal standard is chosen at 5 min and the selectivity is 3.0 relative to phenol (Table 4). The ''open window'' for an internal standard in this separation is very wide, and a larger prediction error can be tolerated. It can be seen in Fig. 2 and Table 4 that the predicted selectivity and retention are in excellent agreement with the measured values. The average prediction error is as small as 5%. This example clearly demonstrates the accuracy of the LSER predictions. 2,3-Dimethylphenol eluting at 5.1 min is an excellent compound as an internal standard due to its UV profile, solubility, and availability.

The chromatogram for the separation of 4-chlorotoluene and *n*-propylbenzene (sample C, 75% ACN) is shown in Fig. 3. The selected location for internal standard candidates is 2 min (retention factor is about 1.5), and corresponding selectivity is 4.0. It can be observed from Fig. 3 and Table 4 that the predicted selectivity is very close to the measured values, and the average error is only 8%. Because the ''open window'' for the internal standard is relatively wide (from 1.5 to 2.5 min), all of them can be used as the internal standard of the separation, irrespective of their toxicity and volatility.

Fig. 4 shows the chromatogram of the separation of 4′-aminoacetophenone and acetophenone (sample D, 40% MeOH). The desired position for the internal standard is at 5 min (the ''open window'' is large),

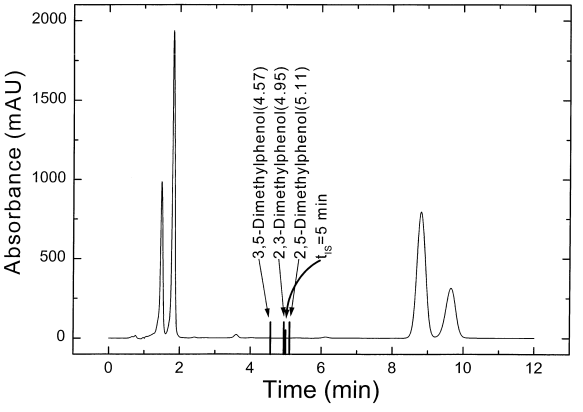

Fig. 2. Chromatogram for the separation of the steroid mixture (sample B) by ACN/water mobile phase under isocratic conditions. The peaks are not identified. The desired position of the internal standard is at 5.0 min. All other conditions same as in Fig. 1.

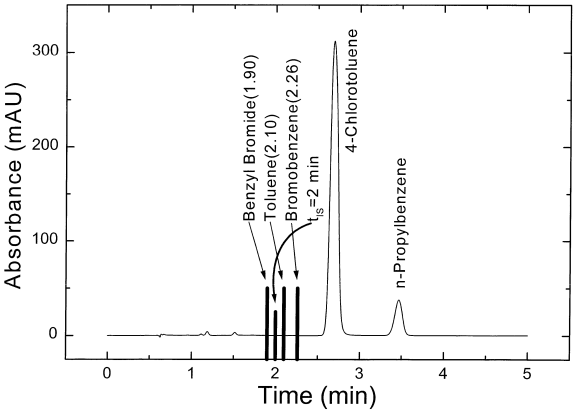

Fig. 3. Chromatogram for the separation of 4-chlorotoluene and *n*-propylbenzene (sample C) by ACN/water mobile phase under isocratic conditions. The desired position of the internal standard is at 2.0 min. All other conditions same as in Fig. 1.

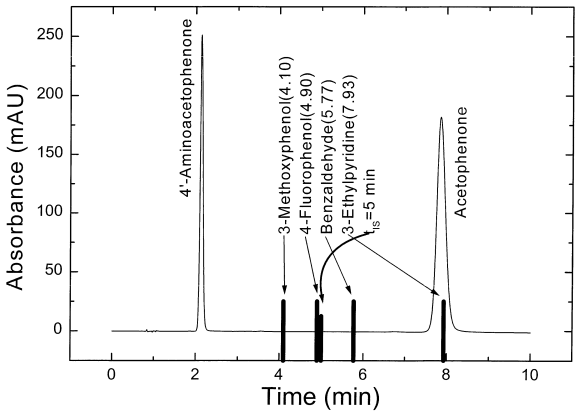

Fig. 4. Chromatogram for the separation of 4′-aminoacetophenone and acetophenone (sample D) by MeOH/water mobile phase under isocratic conditions. The desired position of the internal standard is at 5.0 min. All other conditions same as in Fig. 1.

and the relative selectivity is 1.5. It can be seen in Fig. 4 and Table 4 that the average prediction error is about 21%, that is within the limit of LSER model error. The measured retention for 3-methoxyphenol is unusually small for no obvious reason. However, 4-fluorophenol eluting at 4.9 min is very close to the position. It should be an excellent internal standard compound.

The chromatogram for the separation of sample E (65% MeOH) is shown in Fig. 5. The desired position for an internal standard is 3 min, and relative selectivity is 3.3. It can be observed in Table 4 that the average error is as small as about 12%. 4′-Methylacetophenone and 1-naphthol elute at positions that are very close to the ideal location.

It can be concluded from the results obtained above that the retention and selectivity for predicted compounds in general agree very well with the measured values within the experimental and model errors. The largest error is around 20%. This indicates that, if the width of the ''open window'' for a desired position of an internal standard is above 20%, the LSER approach can aid in selecting internal standard candidates more accurately than the empirical approaches.

4.2. Gradient elution

Fig. 6 shows the chromatogram for the separation of the steroid mixture (sample B) by gradient elution (30–100% MeOH in 15 min). It is clear that the ''open window'' in the middle of the chromatogram for the gradient elution is much narrower than the one in Fig. 2. The desired position for an internal standard is at 9 min. It can be seen in Table 4 and Fig. 6 that the average prediction error for retention is only 8%. Except that 2,6-dimethylaniline may be partially ionized due to its small pK_a, the retention times of all other compounds are very close to the

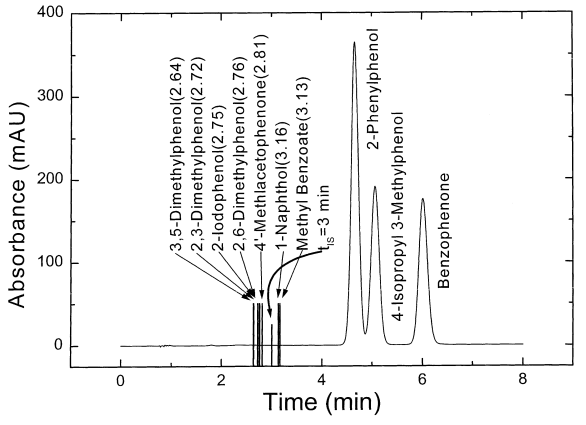

Fig. 5. Chromatogram for the separation of an aromatic mixture (sample E) by MeOH/water mobile phase under isocratic conditions. The desired position of the internal standard is at 3.0 min. All other conditions same as in Fig. 1.

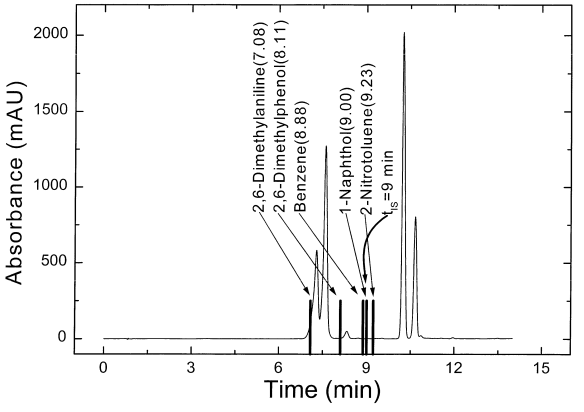

Fig. 6. Chromatogram for the separation of the steroid mixture (sample B) by MeOH/water mobile phase under linear gradient conditions. The peaks are not identified. The desired position of the internal standard is at 9.0 min. All other conditions same as in Fig. 1.

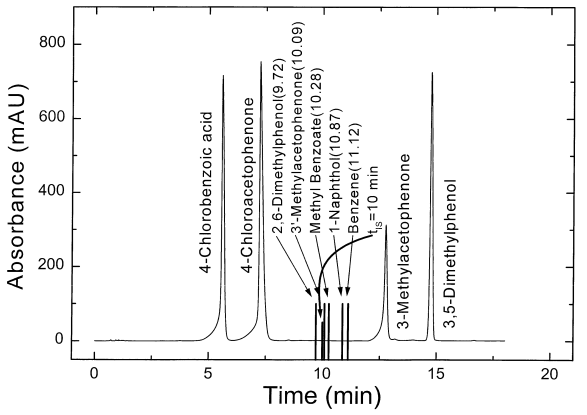

Fig. 7. Chromatogram for the separation of an aromatic mixture (sample F) by ACN/water mobile phase under linear gradient conditions. The desired position of the internal standard is at 10.0 min. All other conditions same as in Fig. 1.

desired position. 1-Naphthol elutes at the exact position of the ideal internal standard.

The chromatogram for the separation of an aromatic mixture (sample F, 20–100% ACN in 30 min) is shown in Fig. 7. There is an ''open window'' in the middle of the chromatogram, and the desired position for an internal standard is at 10 min. It can be seen in Table 4 and Fig. 7 that the average prediction error is only 5%. 2,6-Dimethylphenol and 3′-methylacetophenone are excellent compounds as the internal standard for the separation.

Accordingly, we can also conclude that the LSER approach works accurately for the prediction of internal standards for gradient elution in RPLC. We also expect that this approach will work not only for other types of mobile phases including ternary mobile phases, but also for other types of RPLC columns.

The ongoing study will further explore the use of LSERs for the prediction of internal standards for other types of separations by RPLC.

5. Conclusions

This study describes and evaluates a quick and systematic LSER approach to the prediction of internal standard candidates in RPLC method development. The current study is focused on the separation of neutral samples by RPLC under either isocratic or linear gradient conditions (mobile phases are simply ACN/water or MeOH/water). The procedures for the prediction are very simple and described as follows:

(1) After the separation conditions are optimized for each sample, the desired position for an internal standard is determined based on the ''open window'' of the chromatogram.

(2) Determine the LSER coefficients for the separation by measuring the retention of 20 judiciously selected reference solutes under each separation condition and performing multivariable linear regression on the retention against the properties of the solutes.

(3) Predict the internal standard candidates based on either relative or absolute retention from a database (consisting of more than 700 compounds with defined properties). Select the compounds with the closest retention to the desired position while taking into account of the toxicity, purity, UV profile, and availability of the selected compounds.

It has been shown by this study that LSER is an excellent approach to the fast and systematic selection of internal standard compounds for RPLC method development. The average prediction error has been shown to be no more than 20%. This approach will save a significant amount of time and resources compared with the traditional trial-and-error method. Finally, the approach should apply to other types of RPLC columns, and binary and ternary mobile phase compositions.

It should be mentioned that the use of this approach needs an initial investment in the preparation of reference solutions and establishment of a database. The reference solutions by ACN are usually very stable (up to 2 years in refrigerated condition), and no new preparation is necessary on a daily basis. The database is available from the author or can be generated by scanning the tables in Ref. [46] into a spreadsheet program.

Acknowledgements

The author would like to acknowledge the support for this work by the Department of Analytical Research and Development at 3M Pharmaceuticals.

References

[1] L.R. Snyder, J.J. Kirkland, J.L. Glajch, Practical HPLC Method Development, 2nd ed., Wiley, New York, 1997, p. 657.
[2] K.H. Dudley, in: C.E. Pippenger, J.K. Penry, H. Kutt (Eds.), Antiepileptic Drugs: Quantitative Analysis and Interpretation, Raven Press, New York, 1978, p. 19.
[3] R.E. Leitch, J. Chromatogr. Sci. 9 (1971) 531.
[4] C.D. Chandler, G.R. Gibson, W.T. Bolleter, J. Chromatogr. 100 (1975) 185.
[5] E.J. Kikta Jr, A.E. Stange, J. Chromatogr. A 138 (1977) 41.
[6] J. Wieling, P.M.J. Coenegracht, C.K. Mensink, J.H.G. Jonkman, D.A. Doornbos, J. Chromatogr. 594 (1992) 45.
[7] A.L. Burlingame, R.K. Boyd, S.J. Gaskell, Anal. Chem. 70 (1998) 647R.
[8] M. Godejohann, A. Preiss, C. Mugge, Anal. Chem. 70 (1998) 590.
[9] G. Maniara, K. Rajamoorthi, S. Rajan, G.W. Stockton, Anal. Chem. 70 (1998) 4921.
[10] J. Li, Ph.D. Thesis, Appendix A, Purdue University, West Lafayette, IN, USA, 1994.
[11] N.B. Perry, E.J. Burgess, S.D. Lorimer, J.W. Van Klink, Phytochem. Anal. 7 (1996) 263.
[12] A. Sharma, W.D. Conway, R.M. Straubinger, J. Chromatogr. B 655 (1994) 315.
[13] M. Verzele, L. Use, M. Van Kerrebroek, J. Chromatogr. 282 (1984) 333.
[14] G.M. Ferron, W.D. Conway, W.J. Jusko, J. Chromatogr. B 703 (1997) 243.
[15] S. Yamauchi, H. Mori, J. Chromatogr. 515 (1990) 305.
[16] S. Yamauchi, J. Chromatogr. 635 (1993) 61.
[17] S. Yamauchi, J. Chromatogr. A 704 (1995) 323.
[18] N.E. Skelly, S.W. Barr, A.P. Zelinko, J. Chromatogr. 535 (1990) 199.
[19] A. Wang, L.C. Tan, P.W. Carr, J. Chromatogr. A 848 (1999) 21.
[20] J. Zhao, P.W. Carr, Anal. Chem. 70 (1998) 3619.
[21] L.C. Tan, P.W. Carr, M.H. Abraham, J. Chromatogr. A 752 (1996) 1.
[22] L.C. Tan, P.W. Carr, J. Chromatogr. A 799 (1998) 1.
[23] J.H. Park, J.J. Chae, T.H. Nah, M.D. Tang, J. Chromatogr. A 664 (1994) 149.
[24] J.H. Park, P.W. Carr, M.H. Abraham, R.W. Taft, R.M. Doherty, M.J. Kamlet, Chromatographia 25 (1988) 373.
[25] M.J. Kamlet, R.M. Doherty, M.H. Abraham, P.W. Carr, R.F. Doherty, R.W. Taft, J. Phys. Chem. 91 (1987) 1996.
[26] P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, W. Melander, C. Horvath, Anal. Chem. 58 (1986) 2674.
[27] P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, M.H. Abraham, Anal. Chem. 57 (1985) 2971.
[28] R. Kaliszan, M.A. van Straten, M. Markuszewski, C.A. Cramers, H.A. Claessens, J. Chromatogr. A 855 (1999) 455.
[29] M. Reta, P.W. Carr, P.C. Sadek, S.C. Rutan, Anal. Chem. 71 (1999) 3484.
[30] J. Zhao, P.W. Carr, Anal. Chem. 72 (2000) 302.
[31] L. Li, P.W. Carr, J.F. Evans, J. Chromatogr. A 868 (2000) 153.
[32] D.S. Seibert, C.F. Poole, Chromatographia 41 (1995) 51.
[33] J. Li, P.W. Carr, Anal. Chim. Acta 334 (1996) 239.
[34] J. Li, D.A. Whitman, Anal. Chim. Acta 368 (1998) 141.
[35] J. Li, T. Robison, Anal. Chim. Acta 395 (1999) 85.
[36] J.H. Park, M.H. Yoon, Y.K. Ryu, B.E. Kim, J.W. Ryu, M.D. Jang, J. Chromatogr. A 796 (1998) 249.
[37] J.H. Park, P.W. Carr, J. Chromatogr. 465 (1989) 123.
[38] M. Cuenca-Benito, S. Sagrado, R.M. Villanueva-Camanas, M.J. Medina-Hernandez, J. Chromatogr. A 814 (1998) 121.
[39] P.G. Muijselaar, H.A. Claessens, C.A. Cramers, Anal. Chem. 69 (1997) 1184.
[40] M.H. Abraham, Anal. Chem. 69 (1997) 613.
[41] J.D. Weckwerth, P.W. Carr, Anal. Chem. 70 (1998) 1404.
[42] G.O. Cantrell, R.W. Stringham, J.A. Blackwell, J.D. Weckwerth, P.W. Carr, Anal. Chem. 68 (1996) 3645.
[43] S. Yang, M.G. Khaledi, J. Chromatogr. A 692 (1995) 301.
[44] J. Li, B. Cai, J. Chromatogr. A 905 (2000) 35.
[45] K. Kiridena, C.F. Poole, Chromatographia 48 (1998) 607.
[46] M.H. Abraham, H.S. Chadha, G.S. Whiting, R.C. Mitchell, J. Pharm. Sci. 83 (1994) 1085.
[47] M.H. Abraham, Quantitative Treatments of Solute/Solvent Interactions, in: P. Politzer, J.S. Murray (Eds.), Theoretical and Computational Chemistry, Vol. 1, Elsevier, Amsterdam, 1994, p. 83.
[48] M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chromatogr. 587 (1991) 213.
[49] M.H. Abraham, G.S. Whiting, J. Chromatogr. 594 (1992) 229.
[50] M.H. Abraham, J. Chromatogr. 644 (1993) 95.
[51] M.H. Abraham, J. Phys. Org. Chem. 6 (1993) 660.
[52] J.A. Platts, D. Butina, M.H. Abraham, A. Hersey, J. Chem. Inf. Comput. Sci. 39 (1999) 835.
[53] J.A. Platts, M.H. Abraham, A. Hersey, D. Butina, J. Chem. Inf. Comput. Sci. 40 (2000) 71.