Journal of Chromatography A, 1218 (2011) 218–228

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Interactions between minimum run time, modifier concentration, and efficiency parameters in a high performance liquid chromatography separation

T.L. Chester*, A.M. Stalcup

Department of Chemistry, University of Cincinnati, PO Box 210172, Cincinnati, OH 45221-0172 USA

ARTICLE INFO

Article history:
Received 21 May 2010
Received in revised form 22 October 2010
Accepted 8 November 2010
Available online 16 November 2010

Keywords:
HPLC
Modeling
Optimization
Parameter interactions

ABSTRACT

We modeled and studied the separation of uracil, nicotinamide, resorcinol, theobromine, theophylline, and caffeine on four C-18 columns of different lengths packed with the same stationary phase using water/methanol mobile phase at one temperature. Predictions of retention times and peak widths were compared with experimental results and were found to be sufficiently accurate for performing optimization calculations. With limits set on the required resolution and on maximum values for pressure and flow rate, calculations were performed for numerous virtual column lengths seeking the smallest possible analysis time for each length while allowing methanol concentration and flow rate to vary as required to minimize run time. Predictions were experimentally verified for the column lengths actually available. These calculations revealed the dependence of best-possible analysis time on column length, modifier concentration, flow rate, and pressure for the real system that was modeled, and provided insight into parameter interactions with respect to analysis times meeting the needs and limits specified. We show that when these parameters are considered in concert, rather than individually, conventional guidelines regarding setting their values may not always lead to the optimum.

1. Introduction

High performance liquid chromatography (HPLC) methods and particularly reversed-phase liquid chromatography (RPLC) methods that minimize analysis costs while delivering the user's separation requirements are of great economic importance. This is especially so in industrial laboratories facing high numbers of analyses and a perpetual need to reduce expense. Since costs, such as instrument acquisition, and laboratory space, are fixed, shortening the analysis time of each method and increasing the work efficiency of each instrument are useful strategies to lower costs. Hence, building work efficiency into HPLC methods when they are developed is highly desirable. Computer-based modeling is highly beneficial in developing HPLC methods [1].

1.1. Simulation, improvement, optimization, and constraints

In considering various strategies and the use of computerized tools in HPLC, we must clearly distinguish between simulation (or predictive modeling), improvement, and optimization [2]. *Simulation* can be used to aid and accelerate method development by substituting virtual (or predicted) chromatograms, computer-generated from a mathematical model, instead of experimentally testing every idea. This strategy can greatly shorten the time required to achieve method *improvement* if the underlying models are sufficiently accurate. An experienced user can employ a simulator to help find a desirable if not optimal outcome.

The word *optimization* is often used in liquid chromatography literature to indicate that the speed or work efficiency of a separation has been improved. However, Leibniz clearly referred to finding the best possible outcome, not just improvement, in his principle of the optimum [3]. Therefore, HPLC optimization occurs not necessarily when a separation has been improved, but when a separation has been improved to the fullest extent possible so that no further improvements can be made without changing the constraints existing in the system.

Constrained optimization refers specifically to determining maxima and minima of functions (or systems) with limits imposed on the values of certain variables or combinations of variables [4]. In HPLC, the usual goal of optimization is to minimize the analysis time required to achieve a desired separation. In seeking minimum analysis time it is common to allow excess resolution (*Rs*) to be diminished to save time if the peaks can still be adequately separated to satisfy the purposes of analysis. Thus, the minimum-acceptable *Rs* we specify around each peak becomes a constraint in the optimization of analysis time. Other constraints may also exist, such as pressure and flow rate limits, mobile-phase modifier concentration limits, and particle size and column dimension limits. All constraints considered collectively (that is, the constraint set) define the requirements and boundaries of the optimization.

* Corresponding author. Tel.: +1 513 310 7375.
E-mail address: Thomas.Chester@uc.edu (T.L. Chester).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.11.013

1.2. Single-factor-at-a-time vs. multifactor optimization

In general, single-factor-at-a-time (or univariate) optimization, in which the influences of individual parameters are sequentially examined one-at-a-time while the other parameter values are fixed, rarely finds the optimum in a system unless all the changing parameters are mutually independent [5]. A multifactor (or multivariate) optimization, in contrast, requires all adjustable parameters under consideration to be varied at the same time so that the effects of parameter interactions are included in the observations [5]. The unique combination of adjustable parameter values that optimizes the performance, subject to the choice of parameters allowed to change and to the constraints, is sought. In systems with a high degree of interaction among the adjustable parameters, the optimum may require values in individual parameters that are counter-intuitive and which make no sense when they are considered one at a time. For example, if the column length is initially shorter than optimal, then lengthening the column, combined with appropriate changes in flow rate and modifier concentration, will be necessary to minimize analysis time; lengthening the column to go faster is counter-intuitive but is well understood from the perspective of balancing the column length, flow rate, and pressure [6].

1.3. Regular and irregular samples

We will consider reversed-phase HPLC using a binary mobile phase in which the modifier is designated as component *B*. Samples are classified as *regular* or *irregular* depending on solute retention behavior as the modifier concentration (%*B*) is varied [1,7]. *Regular* samples maintain the same elution order and relative peak spacing with little or no change in separation factors (α values) when %*B* is changed. Such samples may be successfully separated over a large, continuous range of %*B* values. However, samples containing chemically dissimilar solutes are often *irregular* [1,7]. These samples exhibit significant changes in α values and may change elution order as %*B* is varied. Whenever the elution order changes, there may be several discontinuous %*B* ranges (or *windows*) where successful separations are possible.

1.4. Recommended RPLC method development practice

The usual recommended approach to isocratic RPLC method development is a sequence of steps [1]: (1) a column is chosen that provides a sufficient initial number of plates. (2) An initial mobile-phase modifier (*B*) is chosen, and %*B* is adjusted to give retention factors (k) in the preferred range of $1 \leq k \leq 10$ (this range can be increased to $0.5 \leq k \leq 20$ if necessary). (3) Selectivity adjustments are made to improve the α values for closely spaced or co-eluting peaks. (4) The plate number is adjusted to provide a compromise between critical *Rs*, analysis time, and the pressure required.

An initial plate number around 10,000 is usually sufficient for step 1 [1]. When setting the initial value of %*B* in step 2, the mobile phase must be made weak enough that the first peak of interest is sufficiently retained. Peaks of no interest may be allowed to elute earlier. However, if the remaining peaks are outside the range $0.5 \leq k \leq 20$ then a gradient separation is usually recommended [1,7]. In step 3, when the solutes are nonionic, selectivity adjustments entail the choice of the modifier and its concentration, temperature, and even the column type; when ionic or ionizable solutes are involved, then pH, buffer type and concentration, and ion-pair concentrations can also be adjusted [1]. In this step %*B* may have to be readjusted to maintain $0.5 \leq k \leq 20$ for all peaks except early eluting peaks of no interest. The ideal outcome of step 3 is to increase *Rs* and decrease run time [1].

Steps 2–4 are often called *optimizing retention*, *optimizing selectivity*, and *optimizing plate number* (or *optimizing efficiency*). Some degree of subjectivity may be required with this procedure, particularly for irregular samples when %*B* is changed in step 3 to find a good compromise between selectivity (which strongly contributes to *Rs*) and relative retention (which strongly affects run time).

This effective approach and its variations have been widely used, have led to thousands of usable methods, and are widely taught (for example [1,7–15]). Even though effective and efficient methods are routinely produced, from an optimization perspective we should expect that this sequential procedure will lead to the global optimum only if the steps are mutually independent. The relative retention and selectivity parameters (stationary phase, mobile phase components, %*B*, temperature, pH, etc.) are highly independent of the efficiency parameters (column dimensions, particle size, and flow rate) except for their influences on viscosity and pressure. Therefore, optimizing the efficiency parameters after optimizing the selectivity parameters, as described above, will generally lead to an acceptable if not optimal outcome for regular samples. However, for irregular samples, where selectivity and %*B* are interrelated, changing %*B* changes the α values for at least some of the peak pairs. If the critical peak pair is involved in these changes, then the number of plates required to accomplish the separation will also change with %*B*. So, for irregular samples, setting %*B* independently of the efficiency parameters may be problematic for finding the optimal combination of parameter values (even though the outcome of such a procedure may be considered acceptable).

Once the major selectivity parameters (column, mobile phase components, pH, temperature, etc.) are chosen, the optimal combination of column length, flow rate, and %*B* (and particle size, if desired) can be quickly made in a completely objective manner with consideration of resolution requirements and various practical limits (such as pressure) using a constraint-based multivariate optimizer operating within a simulator. This capability also provides a convenient means to study the complexity of parameter interactions in detail for real separations.

There is a long history of considering some HPLC parameters in concert. For example, Guiochon considered the interactions of column length, flow rate, and pressure [6]. He showed that (in the context of lengthening the column from low values when flow rate and column length are the changing parameters) the fastest analysis will occur at the highest pressure allowed using a column of the length required to provide the desired number of plates. In this consideration, %*B* was not varied, so there was no possibility of changes in relative retention, selectivity, or of interactions between mobile phase strength and selectivity. Also, extra-column broadening was ignored. Schoenmakers et al. developed software specifically for HPLC system optimization, but they did not adjust %*B* in their calculations [2,9,10].

More recently, a variety of workers have used “Poppe plots” to relate particle size, column dimensions, pressure, and the time required to achieve a given number of plates [16–22]. Sophisticated chromatography simulators are capable of considering many HPLC parameters in concert. For example, the influence on resolution of temperature, gradient time (which for gradients is analogous to %*B* for isocratic separations), ternary mobile-phase composition, and pH can be displayed in a multidimensional graphical representation, and the optimal parameter combination for maximizing *Rs* can be located visually [23]. Optimizing efficiency is performed as the next step in this approach.

Selectivity adjustment is the most important task in developing a new method. However, not all HPLC optimization efforts are necessarily aimed at producing the best method from among all possibilities; practicality often dictates less work. As one example, if a preservative is substituted in a commercial product and the new preservative interferes with an active ingredient peak in an existing assay method, then the workers may choose to adapt the existing method to accommodate the product formulation change

Table 1
Important system dimensions.

Injector loop i.d.	0.5 mm
Inlet tube i.d.	0.3 mm
Inlet tube length	24 cm
Outlet tube i.d.	0.3 mm
Outlet tube length	51 cm
Detector inlet tube i.d.	0.3 mm
Detector inlet tube length	16 cm
Detector cell path	10 mm
Detector cell volume	8 μL

rather than to develop a totally new method. This choice is driven by business needs and risk assessment. Even in new method development, practicality may steer workers toward using standardized stationary and mobile phases to minimize workplace complexity and changeover effort when particular HPLC setups can be used to execute more than one method. In any event, there is always a need to find the best combination of flow rate and %*B* (and sometimes column length) that minimizes analysis time and otherwise meets the analysis requirements whether a method is developed from scratch with an exhaustive selectivity optimization or is simply adjusted.

We examined the separation of an irregular sample in detail to reveal the interrelations between %*B*, flow rate, column length, and pressure with constraints on *Rs*, pressure, and flow rate. This was not a method development effort, but a study of the optimization and interactions of %*B* and the efficiency parameters for the sample, stationary phase, and mobile phase components we selected. For the optimization part of this work we used a multivariate optimizer embedded within a chromatogram simulator. We also studied the locus of best-possible analysis times vs. column length for this separation to reveal the complexity of parameter interactions in detail and gain further insight into HPLC parameter selection.

2. Experimental

We used a dual pump HPLC system with high pressure mixing (Shimadzu Scientific Instruments, Inc., Columbia, MD, USA). It consisted of an SCL-10Avp system controller, an SPD-10Avp dual-wavelength UV–vis detector, a SIL-10AF autosampler, two LC-10ATvp pumps, and a DGU14-A degasser. This was under control of EZStart 7.3 SP1 software operated from a generic personal computer running Windows XP Professional (Microsoft Corporation, Redmond, WA, USA). Although all the present work was isocratic, we used the gradient capability of the HPLC system to blend and deliver mobile phase at the desired composition: bottle *A* contained HPLC-grade water, and bottle *B* contained HPLC-grade methanol. We made no effort to correct the flow rate for excess volume of mixing.

The columns were a matched set of Ascentis C18, all 4.6-mm inside diameter, and with lengths of 30, 50, 100, and 150 mm (Supelco, Sigma–Aldrich, Bellefonte, PA, USA). The columns were packed with 3-μm-diameter, 100-Å-pore stationary phase from a single lot. A column oven was not available, so the columns were operated at ambient temperature in an air conditioned laboratory maintained at 24.5 ± 0.5 °C.

The important physical dimensions of the HPLC system are shown in Table 1. Partial-loop injection of 1-μL volumes was used. The detection wavelength was 274 nm.

Solutes were uracil, resorcinol, nicotinamide, and the alkaloids theobromine, theophylline, and caffeine. These were dissolved in 95/5 (by volume) water/methanol to make a test solution approximately 0.3 mg/mL for each solute except theobromine which was made at 0.1 mg/mL.

Formulas for predicting retention time, temporal peak width at the detector, and peak resolution were derived from first principles or from well established empirical estimates and are summarized in Appendix A. We estimated solute retention factors as a function of modifier concentration from screening data by regression of the common logarithm of the experimental retention factor ($\log_{10} k$) for each solute against modifier concentration using an empirical second-order polynomial function [24]. We estimated retention times from retention factors and the void time derived from the screening experiments and corrected to the column length and flow rate under consideration. Predicted *Rs* values were calculated directly from our estimates of retention times and peak widths. We used the proportionalities from the Darcy equation [25,26] and the comprehensive flow equation [27] for predicting pressure as parameter values are changed. We assumed the entire pressure drop was due to the column, and we ignored the pressure drop in extra-column tubes. We further assumed the mobile phase viscosity was constant for all water–methanol ratios, and we used the highest inlet pressure observed during screening (8.053 MPa for 40% methanol, 1 mL/min, with the 50-mm column) in calculating pressure with other parameter settings. In this way the actual pressures we encountered when verifying predictions of separations were always less than or equal to the predicted pressure (to prevent unexpectedly encountering a pressure limit).

We developed Microsoft Excel 2003 and Excel 2007 workbooks for performing calculations and simulating chromatograms. The workbooks are functionally similar to that reported earlier [28,29], but were entirely rewritten to better utilize Excel capabilities, improve computational efficiency, and add several new capabilities beyond the scope of this report. This new implementation also estimates the peak broadening caused by detector internal tubing, which was omitted from the earlier work. Both flow-related broadening and volume-related broadening in the injector are considered, but only the volume of the detector cell is used in estimating its contribution to broadening; the cell was otherwise assumed to provide enough turbulence to render Poiseuille flow and the associated flow-dependent broadening negligible [30] but not so much turbulence that the cell acts as an exponential dilution chamber.

Optimization was accomplished using the Excel Solver Add-In within the Excel workbooks. We are able to simultaneously vary column length, column diameter, flow rate, % methanol, particle size, injection volume, detector cell volume, and dimensions of inlet and outlet tubes. Constraints can be set on all of these parameters. We can also set constraints for pressure and for *Rs*. The *Rs* constraint may be set individually for every peak with respect to its nearest neighbor. The optimization target can be any single objective function to be minimized, maximized, or driven to a target value. Possible targets include analysis time, peak capacity, mobile phase volume required, critical (that is, smallest) *Rs*, total cost, or any other measure that can be specified with an objective function. Note that if the target is the maximization of peak capacity, *Rs*, or any other function that improves with time, then it is necessary to add maximum-allowed analysis time as a constraint. In this work, we usually sought the combination of % methanol and flow rate that minimized the retention time of the last peak for each column length considered while providing or exceeding a specified minimum *Rs*. In addition to calculations at specific real or virtual column lengths, we performed additional calculations in which column length was included as a changing parameter in order to find the column-length – % methanol – flow rate coordinates of optima subject to various practical constraints.

3. Results and discussion

Development of fast methods must include selectivity considerations, but in this work our goal was not method development

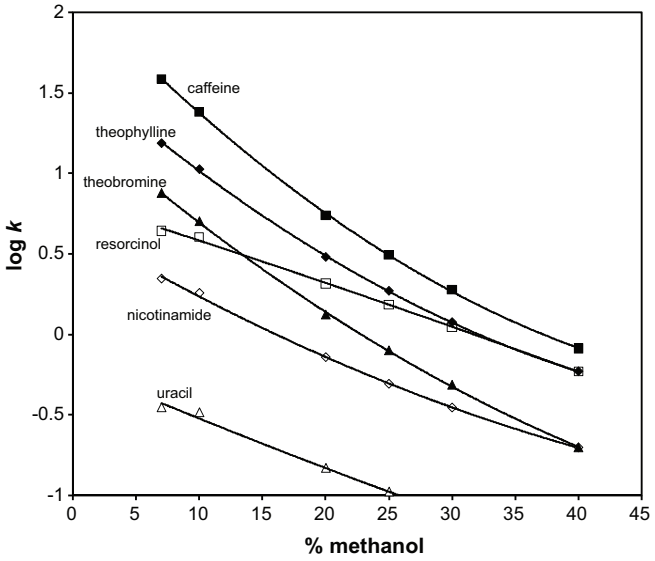

Fig. 1. Retention model developed using the 50-mm column. The points are experimental data, and the lines are regressions.

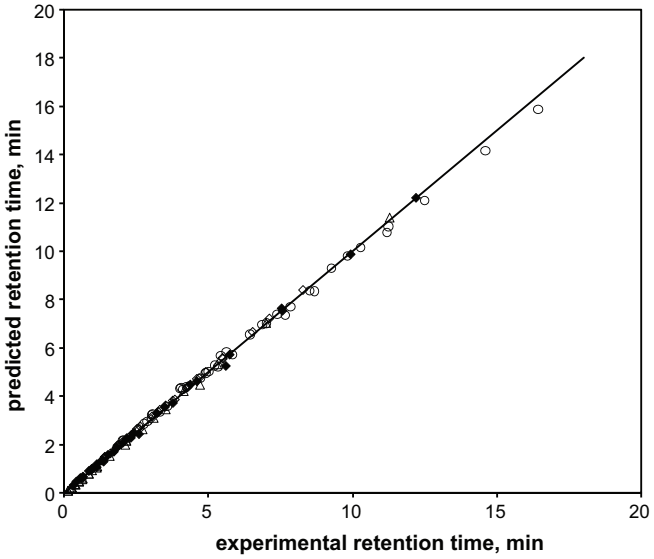

Fig. 2. Retention times predicted from the model developed using the 50-mm column compared with experimental retention times for corresponding conditions. Column lengths are 30 (△), 50 (♦), 100 (◊) and 150 mm (○) for 22 combinations of flow rate, column length, and methanol concentration not included in the model. The solid line represents perfection in making predictions, and vertical deviations of points from the line represent prediction errors.

but to optimize analysis time with respect to column length, flow rate, and methanol concentration and to explore their interactions after making the selectivity choices (specifically the stationary phase, mobile phase components, and temperature). Our stationary-phase–mobile-phase combination (Ascentis C18 plus water–methanol) was chosen for convenience and because the solutes are irregular in this system. A better method can be developed with a thorough exploration of selectivity parameters followed by efficiency optimization, but we focused on efficiency optimization and interactions for the system we used.

3.1. Modeling and accuracy testing

3.1.1. Retention times

We made duplicate injections of the test solution into mobile phases containing 10, 20, 25, 30, and 40% methanol at 1 mL/min using the 50-mm column. The root mean square error (RMSE) of retention time deviations for paired injections was 0.004 min. The largest retention time deviation observed for a pair of injections was 0.007 min. Thus, the system and these data appeared to be sufficiently stable and reproducible for modeling retention time. The results, expressed as the logarithm of retention factors, are charted in Fig. 1 along with regression lines (this figure includes an additional model point at 7% methanol that was added later to improve predictions when methanol concentrations were below 10%).

We used this model to predict retention times for 22 new sets of conditions for our four columns, and we compared these predictions with experimental results. These conditions encompassed changes in % methanol, flow rate, and column length. There were 132 peaks altogether. When the model derived from the 50-mm column was reapplied to five new sets of conditions on the same 50-mm column, the RMSE comparing predicted and experimental retention times was 0.09 min (for 30 peaks). When the 50-mm-column model was applied to make predictions over all four columns, the RMSE comparing predicted and actual retention times was 0.13 min (or about 8 s). The predicted retention times are plotted against the actual retention times in Fig. 2. The largest deviation observed occurred for the most-retained peak in the data set: the predicted retention time was 15.86 min or 0.56 min sooner than the actual retention time of 16.43 min. This chromatogram was performed using the 150-mm column, which eluted most peaks later than predicted. This suggests the possibility of a systematic deviation of either this column compared to the others, a deviation of the model, or a deviation of the temperature at the time of this work (a 1 °C temperature change resulted in a retention time shift of approximately 0.3 min for a 16-min peak). Such deviations are not particularly troublesome because they affect all the peaks similarly, thereby tending to preserve the relative peak spacing and *Rs* as long as the deviations are small. Overall, the accuracy of the retention time prediction is sufficient for our purposes even though a column oven was not available.

3.1.2. Peak widths

Fig. 3 plots predicted peak widths (four standard deviations) against experimental peak widths (measured at 13.4% of the peak height, which corresponds to four standard deviations for Gaussian peaks) for the 22 test chromatograms. We omitted uracil from this comparison because its experimental width was frequently perturbed by solvent disturbances near the void volume. The largest relative deviations occurred for the narrowest peaks (with low retention factors and high flow rates on the 30-mm column, see Section 3.5). Some bias in predicted peak widths is apparent, and a more sophisticated approach or an empirical or semi-empirical approach would produce better accuracy. However, the bias we observed does not change the general trends and, in most cases, means the resolution in the experimental chromatograms is actually slightly better than the predicted resolution. Peak width inaccuracies are caused by several sources:

(1) The Atwood–Golay estimation of broadening in connecting tubes [30] is applicable only to smooth, straight tubes. Curves and kinks in real tubes introduce additional radial mixing that reduces experimental broadening compared to our predictions. So the Atwood–Golay estimates represent the worst case, and actual broadening is expected to be smaller.
(2) The viscosity of water–methanol mixtures ranges between about 1.1 and 1.6 cP over the compositions used in this work [1,14]. Solute diffusion rates in dilute liquid solutions are proportional to molecular weight to the one-half power and are inversely proportional to the liquid viscosity [31]. However, we chose to use only one compromise value of the solute dif-

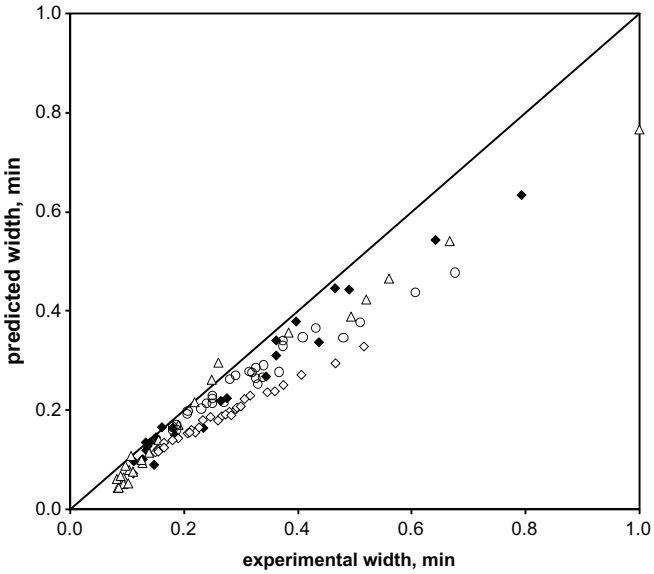

Fig. 3. Predicted peak widths (four standard deviations) vs. the corresponding experimental peak widths evaluated at 13.4% of the peak height (which corresponds to four standard deviations for Gaussian-shaped peaks). Column lengths are 30 (△), 50 (♦), 100 (◊) and 150 mm (○) for 22 combinations of flow rate, column length, and methanol concentration not included in the model. The solid line represents perfection in predictions, and vertical deviations of points from the line represent prediction errors.

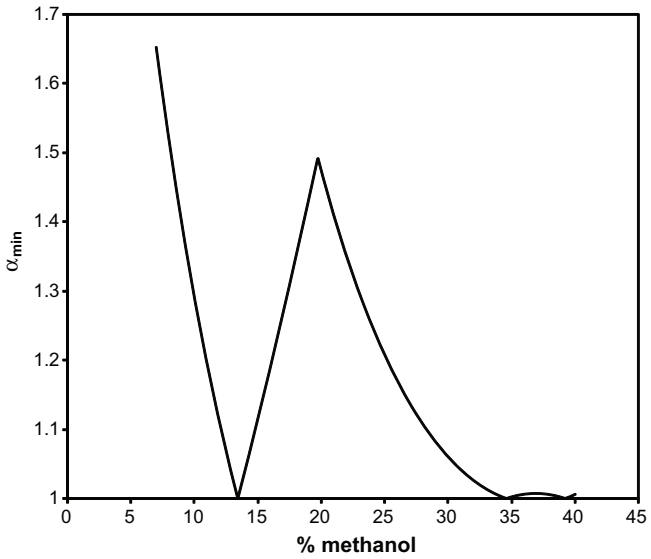

Fig. 4. Critical-selectivity window diagram for the model.

fusion coefficient for all calculations: 4.86×10^{-6} cm^2/s. Thus, we could expect the largest absolute deviations from predicted peak widths to occur for the largest solutes at relatively high methanol concentrations (where the viscosity is highest and the diffusion rates are lowest) on the longest columns (where the peaks are widest).

(3) We set the coefficients of the Kennedy–Knox equation [32] to the generally accepted values of $A = 1$, $B = 2$, and $C = 0.05$ [1,14,33]. These are typical values, and we deliberately made no effort to adjust them to improve our prediction results. We assumed in making peak width predictions that the columns would not deviate from their expected plate heights and that the peaks would be perfectly Gaussian in shape.

(4) At very high flow rates we expect eddies and exponential mixing to increase in the detector cell, and this will contribute both additional peak width and tailing.

Column-to-column differences can always be expected, and at the time of this work the 100-mm column already had several months of use and several hundred injections while the other three columns were used here for the first time. Fig. 3 suggests a systematically larger deviation from the predicted width for the 100-mm column than for the others. Peaks widths from 0.1 to about 0.4 min are usually predicted adequately for our purposes. All peak widths for the practical optima we report in Section 3.3 are in this range, so despite our inability to accurately predict peak widths far from optima, our accuracy is adequate for studying interactions in the vicinity of optima in this work and for observing general trends.

It is possible to specify an empirical asymmetry factor for each peak and alter the predicted peak shape accordingly. Setting such a tailing specification for each peak is a desirable feature for commercial software. However, because our emphasis in this work was on exploring parameter interactions, this feature was not incorporated as part of our model.

3.2. Methanol concentration windows

Fig. 1 shows that our sample is highly irregular: $\log k$ values for resorcinol and nicotinamide change at much different rates with respect to methanol concentration than do the alkaloids. This results in large, continuous changes in α values as the methanol concentration is varied. Theobromine and resorcinol coelute when the mobile phase is 13.4% methanol. Below this concentration resorcinol elutes before theobromine, and above this concentration theobromine elutes before resorcinol. Likewise, resorcinol coelutes with theophylline around 34% methanol, and nicotinamide coelutes with theobromine around 40% methanol. This behavior produces two windows with respect to methanol concentration, one window below 13% and the other between 13% and 34%, where the selectivity and retention factors may be high enough that all six peaks can be well separated. This is shown in Fig. 4 which is a critical-selectivity window diagram patterned after Laub and Purnell [34] but adapted here to examine the effects of mobile phase composition. In this figure we consider all possible pairs of solutes and plot the critical (or lowest) separation factor (α_{min}) for the entire solute set vs. methanol concentration. At least one additional selectivity window exists above 40% methanol, but retention there is too low to accomplish a practical separation.

The value of α_{min} improves as the methanol concentration is decreased from 13%. The α_{min} value becomes adequate to achieve a complete separation around 11% methanol and continues improving as the methanol concentration is lowered further. However, the k value for caffeine is 20 at 11% methanol. Because this is the usually recommended upper practical limit for isocratic separations, this window appears to be impractical at first inspection. We will return to it in Section 3.5.

3.3. Model and experimental results in the 13–34%-methanol window

To study parameter interactions, we performed a series of optimization calculations in this window for 42 column lengths distributed approximately evenly from 30 to 150 mm. The optimization goal was to minimize analysis time while producing $Rs \geq 2.5$ for all the peaks except uracil, which we allowed to elute in or very near the void volume as long as it was adequately resolved from the other solutes. We performed these calculations allowing

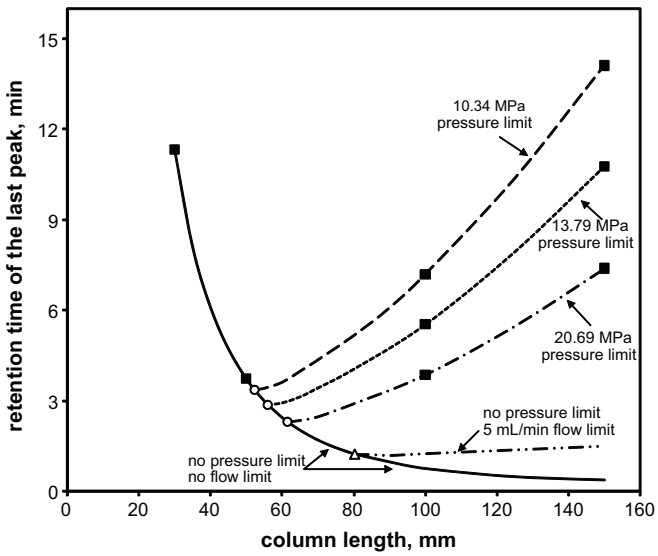

Fig. 5. Loci of best-possible analysis times vs. column length in the 13–34%-methanol window for $Rs \geq 2.5$ for all peaks (except uracil). The parameter space is five dimensional, and this figure is the projection into the analysis-time–column-length plane. The solid line is the projection of the locus without pressure or flow-rate limits. The intersection of this locus with three pressure limits is indicated by the open circles. An additional pressure-limited locus begins at each of these intersections and is plotted for each of the three pressure limits. The intersection with a flow-rate limit is indicated by the open triangle, and a flow-rate-limited locus begins at this intersection. The black squares indicate the solutions for the actual column lengths available in this work.

the methanol concentration and the flow rate to seek the combination of values necessary to minimize the analysis time for each column length while providing or exceeding the target *Rs* for every peak as measured against its nearest neighbor (the assignment of nearest neighbor is made in terms of resolution and is continuously tracked to account for shifts in relative spacing and for elution order changes with %*B* changes for irregular samples). We also applied various flow and pressure limits. The resulting loci of best-possible analysis times vs. column length are shown in Fig. 5.

The solid line in the figure is the unlimited locus (that is, without considering a pressure limit or flow limit) of best-possible analysis times vs. column length for our system. At each column length the values for methanol concentration and flow rate are those required to produce the shortest analysis time, so the partial derivatives of analysis time with respect to methanol concentration and with respect to flow rate are both zero everywhere along this locus.

For our system, the best-possible analysis time for the 30-mm column is 11.3 min. From there, the best-possible analysis time decreases monotonically as the column length is increased and the values of methanol concentration and flow rate are changed as required to minimize analysis time at each new column length. This analysis-time–column-length locus can be thought of as spanning a four-dimensional space (column length, methanol concentration, flow rate, and time), and the solid line in Fig. 5 is the projection of this locus onto the column-length–time plane. If we also keep track of pressure, then a fifth dimension is necessary to fully describe this analysis time locus in terms of all the parameters of interest. In Section 3.4 we will address the parameters not shown in this figure.

The predicted pressure required for the optimal solution using the 30-mm column is very low (0.986 MPa or 143 psi). Both flow rate and pressure smoothly increase along the solution locus as the column is lengthened. Eventually a pressure limit is reached. This may be either a pressure limit specified to protect the system, or a physical capability limit of the system. The results of applying three different pressure limits are shown in Fig. 5, and the intersections of the pressure-limited loci with the unlimited locus are indicated with the open circles.

After reaching a pressure limit, further lengthening of the column is not allowed to produce higher pressure, so a new pressure-limited locus emerges from the unlimited locus at the point where the pressure limit is encountered. In Fig. 5, pressure-limited loci of best-possible analysis times increase monotonically with column length. Therefore, in this case, the optimal column lengths occur at the intersection of the unlimited locus and each pressure-limited locus. The coordinates of these optima are, of course, also subject to the values of all other parameters not being varied for this figure.

The locations of these optima are in agreement with Guiochon's observation that the fastest analysis will occur at the highest pressure allowed using a column of the length required to provide the necessary plates [6]. This conclusion is true here even though we also allowed the methanol concentration (and the solute retention factors) to vary as required to minimize analysis time for every column length and we continuously corrected for extra-column effects on predicted peak widths.

The flow rate is quite low (0.20 mL/min) for the 11.2-min optimal solution using the 30-mm column. The flow rate increases monotonically along the unlimited locus as the column is lengthened. The flow rate will eventually reach its practical limit if the pressure limit is not encountered first. The intersection of the unlimited locus with the 5-mL/min flow-rate limit is indicated by the open triangle in Fig. 5, and a flow-rate-limited locus emerges from the intersection.

In performing optimization as part of method development, it is not necessary to study interactions in detail as we have done here. Instead, if column length is included with methanol concentration and flow rate as changing variables then the optimizer will usually find the appropriate solution for the constraints specified. In some instances the optimizer may find a local optimum (see Section 3.4), so in practice it is wise to test every solution by running the optimizer from several different starting combinations of parameters.

After finding the ideal combination of column length, methanol concentration, and flow rate, the best practical solution must be sought from among commercially available column lengths. We usually investigate the available column lengths on both sides of the optimal column length. This is simply a matter of specifying the column length of interest and then running the optimizer again to find the best corresponding values of the remaining changing variables for this fixed column length. Various practical solutions using three different pressure limits for the column lengths available to us in this work are indicated by the black squares in Fig. 5. The predicted and experimental chromatograms at these conditions are compared in Fig. 6.

The best practical solution in this work takes 3.73 min using the 50-mm column and requires 8.50 MPa (1233 psi). The predicted and experimental chromatograms are shown in Fig. 6b. The 100-mm column provides a slightly longer solution, 3.85 min at 20.69 MPa (3000 psi), Fig. 6e. If the pressure limit is made higher, faster solutions are possible with these columns.

An important observation is that the fastest analysis time may not require all the available pressure when there are only a few column lengths to choose from. Another important observation is that if method development were to be performed experimentally for this problem without knowledge of retention and selectivity behavior, and if a researcher chose to begin with a 150-mm column and with a methanol concentration producing retention factors in the range of 1–10 for all the solutes except uracil, no amount of work in this system would produce a separation faster than 7.4 min when using a 20.69 MPa (3000-psi) pressure limit, Fig. 6h. The existence of a low-pressure solution saving half the analysis time using a

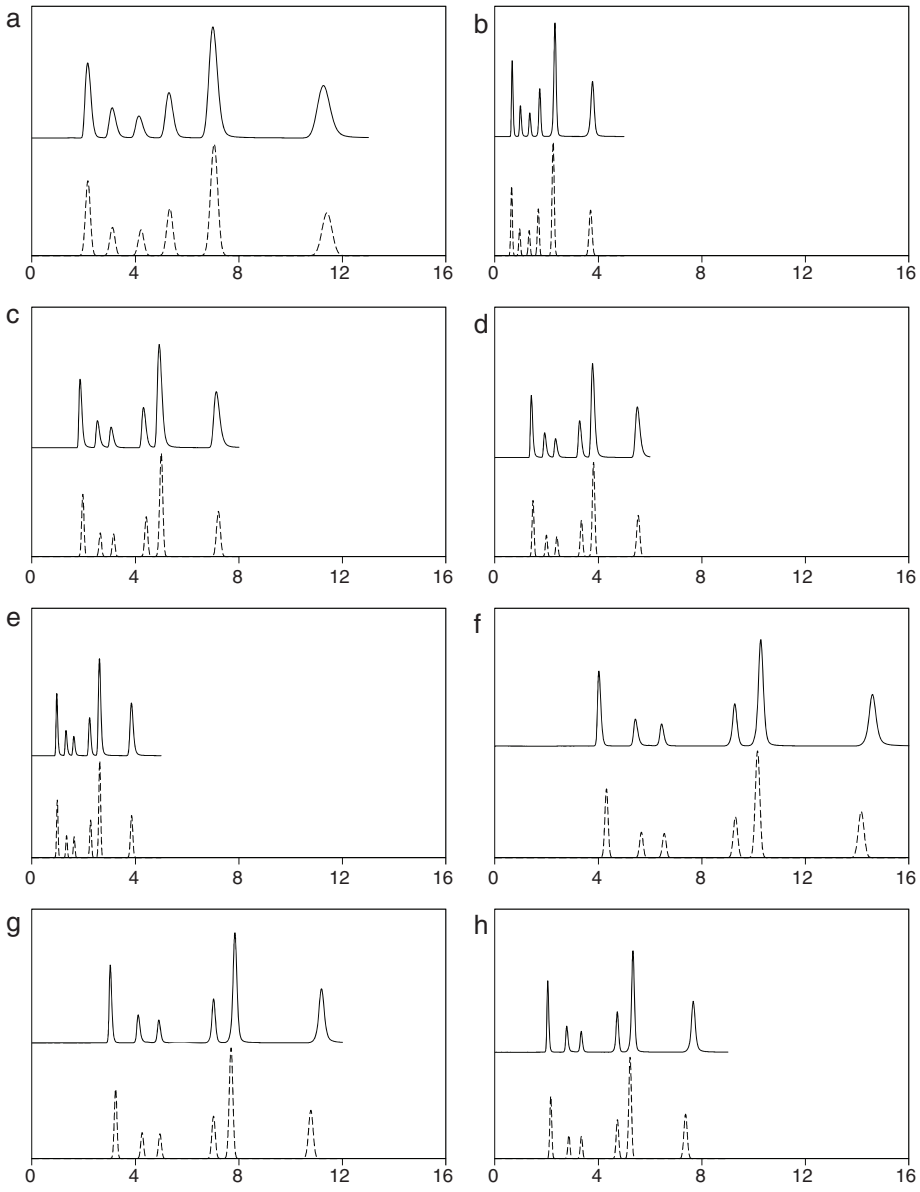

Fig. 6. Comparisons of experimental chromatograms (solid) and predicted chromatograms (dashed) for the conditions indicated with solid squares in Fig. 5. (a) 30-mm column, 19.6% methanol, 0.204 mL/min; (b) 50-mm column, 19.4% methanol, 1.056 mL/min; (c) 100-mm column, 24.7% methanol, 0.642 mL/min; (d) 100-mm column, 24.4% methanol, 0.856 mL/min; (e) 100-mm column, 23.9% methanol, 1.284 mL/min; (f) 150-mm column, 26.3% methanol, 0.428 mL/min; (g) 150-mm column, 26.1% methanol, 0.571 mL/min; and (h) 150-mm column, 25.8% methanol, 0.856 mL/min.

50-mm column, Fig. 6b, may not be obvious to a researcher who has used only the 150-mm column. However, this solution can be found quickly by exploring with a simulator or performing an optimization within a simulator as we have shown.

Fig. 7 plots the best-possible analysis time vs. column length for our HPLC system when 3.5 is specified as the required *Rs* between peaks (again omitting uracil). The overall behavior is quite similar to the situation for $Rs \geq 2.5$ except that there is no solution possible for column lengths shorter than 40 mm. The best practical analysis time is 4.75 min with the 100-mm column at 20.69 MPa (3000 psi).

3.4. Parameter interactions

More insight is available about the path of the analysis-time–column-length locus through the parameter space by adding the corresponding values of the remaining parameters to Fig. 5. This is done in Fig. 8a. Similarly, we can also plot projections of the locus into the other planes of the parameter space and include indications of the unseen variables. This is done in Fig. 8b–d; we omitted the pressure- and flow-rate-limited loci for clarity.

These figures reveal that along this locus all the parameters except methanol concentration change monotonically, and that the behavior is surprisingly complicated (the parameters plotted on the abscissas of Fig. 8b–d are not independent; column length is the only independent parameter, and all other parameters and the analysis time are dependent on the column length in these representations. The same sort of treatment can be performed by declaring any other changing variable the independent variable and calculating the minimum-time locus and the corresponding values of the remaining variables).

The most complicated behavior exists with respect to methanol concentration, Fig. 8b. There are no optimal solutions requiring less than 19.35% methanol in the 13–34%-methanol window (meaning that every separation at the corresponding column length can be performed in less time at a higher methanol concentration), and there are two analysis time minima for each methanol concentra-

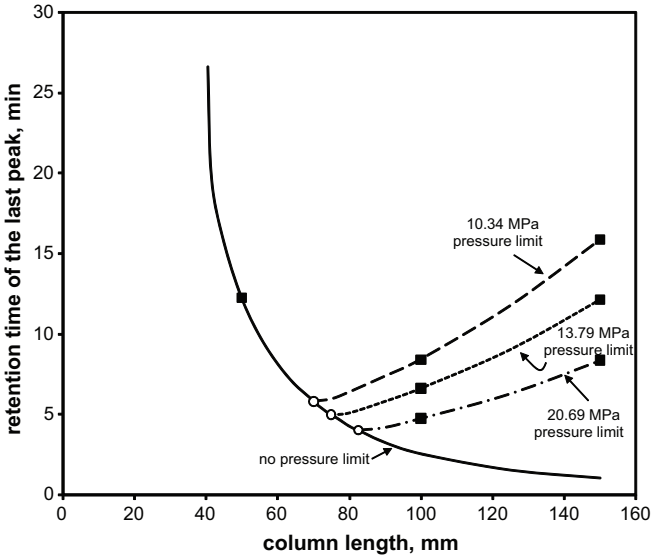

Fig. 7. Loci of best-possible analysis times vs. column length in the 13–34%-methanol window for $Rs \geq 3.5$ for all peaks (except uracil). Additional loci for three pressure limits and a flow-rate limit are also shown as in Fig. 5. The black squares indicate the solutions for the actual column lengths available in this work.

tion above 19.35%. One of these is a local optimum, not the global optimum we seek. The shortest times correspond to the longest columns. For example, at 19.65% methanol a 30-mm column would require 11.3 min to separate the solutes with $Rs \geq 2.5$, but a 100-mm column can perform the separation in less than 4 min at about the same methanol concentration. It would be rare that we would allow method optimization to be driven by a desire to control modifier concentration unless the modifier were very expensive or if we had a phase behavior limit such as buffer precipitation.

3.5. Modeling and experimental results below 13% methanol

Preliminary calculations suggested that a fast separation is possible using a methanol concentration less than 13% in spite of the high retention factor of caffeine. We performed additional screening at 7% methanol and added this point to the previous model to improve the prediction accuracy below 10% methanol. The loci of best solutions with various limits for this window are given in Fig. 9. The practical optimum at 20.69 MPa (3000 psi) for a 30-mm column predicts a 2.48-min separation using 8.27% methanol with a flow rate of 4.28 mL/min.

An experimental separation at these conditions showed peak broadening about 50% higher than predicted for uracil, nicotinamide, and resorcinol. This error may be due to turbulent mixing in the detector cell or in the fittings in the sample path. A too-high diffusion coefficient value in the model and a too-low Kennedy–Knox *C* term may also be contributing. Such high flow rates are uncommon with 4.6-mm-diameter columns, so it is not clear if additional investigations to improve predictions under these conditions are justified. Regardless, we were surprised that the retention factor for caffeine was 32 at the calculated optimum.

The actual pressure for this experimental separation was 15.45 MPa. This pressure is well below what was predicted (20.69 MPa) due to the much lower mobile phase viscosity when the methanol concentration is below 10% compared to its value at 40% methanol (which was our reference condition for calculating pressure). We empirically corrected our pressure prediction, increased our target *Rs* for nicotinamide to 3 (in hopes of adding

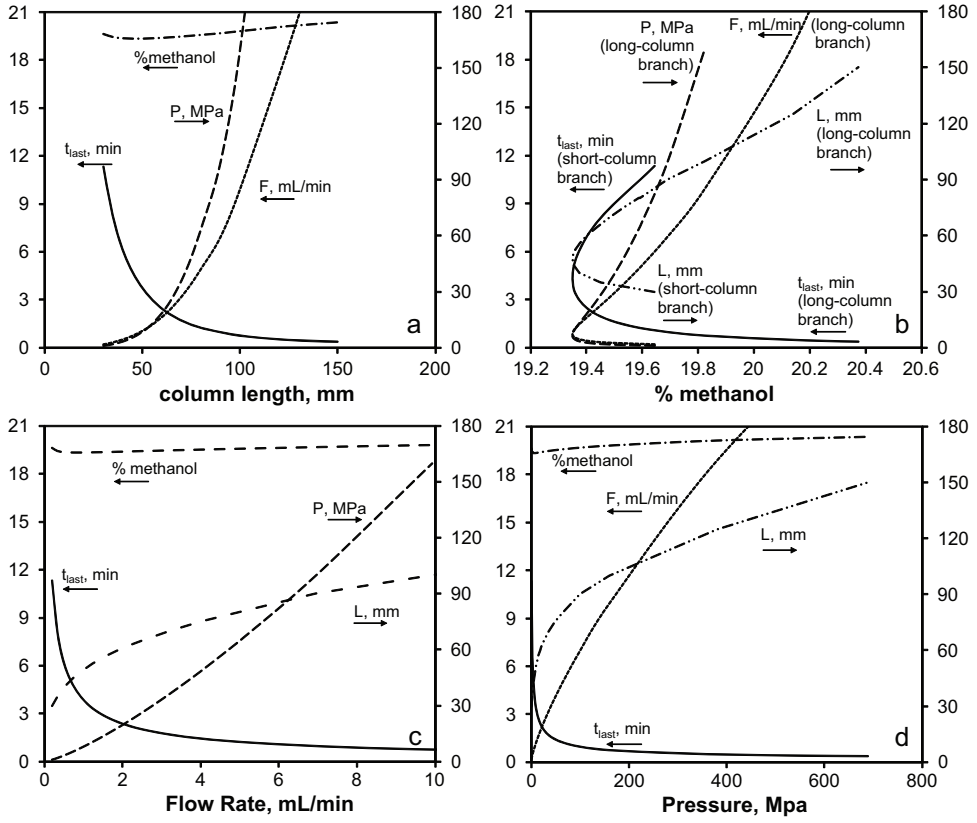

Fig. 8. The locus of best-possible analysis times vs. column length at $Rs \geq 2.5$ projected into four planes of the 5-dimensional parameter space. In each projection, the values of the remaining parameters corresponding to the solution at the analysis-time–abscissa coordinate are added. In (b) the labels for the short-column branches of the pressure and flow rate curves were omitted for clarity.

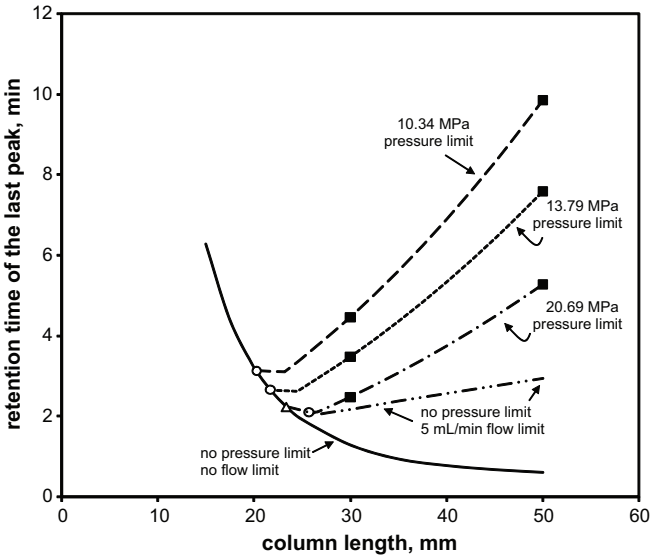

Fig. 9. Loci of best-possible analysis times in the window below 13% methanol for $Rs \geq 2.5$ for all peaks (except uracil) projected into the analysis-time–column-length plane of the 5-dimensional parameter space. Additional loci for three pressure limits and a flow-rate limit are also shown as in Fig. 5. Note that here the flow-rate limit was encountered before the 20.69-MPa pressure limit. The black squares indicate the solutions for the actual column lengths available in this work.

enough *Rs* to an actual chromatogram to achieve baseline separation), and calculated the optimum again. The resulting prediction and the corresponding experimental chromatogram are shown in Fig. 10, which achieves a baseline separation in less than 3 min. The experimental pressure remained below the limit, but the 5-mL/min flow rate limit was applied. The caffeine retention factor is 39. This is so far from the normally prescribed maximum (which is 20 [1,7]) that we have to question the validity of this guideline when selectivity increases rapidly with decreasing modifier concentrations as is the case here below 13% methanol. Note also that if selectivity remains constant with decreasing modifier strength, then optimization will lead to more ordinary retention factors, and if selectivity improves with increasing *%B* then optimization will likely lead to smaller retention factors. Regardless, conventional guidelines may not apply to finding optima for irregular samples.

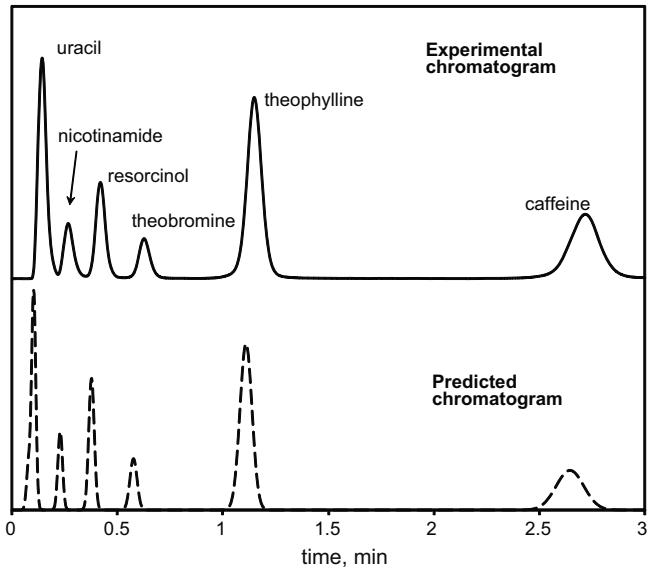

Fig. 10. Experimental (upper) and predicted (lower) chromatograms for the best practical baseline separation below 13% methanol. The conditions are: 30-mm column length, 6.9% methanol, 5 mL/min. The *k* value for caffeine is approximately 39.

In the 13–34%-methanol window, the global optimum for a given pressure limit occurred on the unlimited locus when the pressure limit was encountered. However, for the window below 13% methanol, the fastest predicted separation at a given pressure limit does not occur at this intersection. Using the 10.34-MPa (1500-psi) locus in Fig. 9 as an example, as the column length is increased from low values, the pressure limit is reached when the column length is 20.23 mm. If we continue lengthening the column while minding the pressure limit, analysis times continue to improve until the column length reaches the optimal value of 23.13 mm. This behavior results because selectivity of the *Rs*-limiting solutes is changing rapidly with methanol concentration. Uracil and nicotinamide form the *Rs*-limiting solute pair when the methanol concentration is less than 7.97% which is required on the solution locus for column lengths shorter than 23.13 mm, but for longer column lengths and higher corresponding methanol concentrations the *Rs*-limiting pair switches to resorcinol and theobromine. This switch occurs at the point of the slope discontinuity in the locus. All four of these peaks are equally resolved at this discontinuity, which is the global optimum for the varied parameters at this pressure limit. Similar behavior occurs for the 13.79-MPa (2000-psi)-limited locus. For the 20.69-MPa (3000-psi) locus we encounter the 5-mL/min flow-rate limit before reaching the pressure limit. This is the first time we have seen a separation in which the global optimum does not exist at the point where the pressure limit is encountered along the unlimited locus, and thereby proves that the optimum does not always occur at this intersection.

4. Conclusions

We modeled retention based on experimental observations, and assessed the accuracy of simulating chromatograms combining this model with straightforward calculations based on well-established equations. We made no empirical corrections to improve accuracy, but used only familiar theoretical assumptions regarding peak broadening. Accuracy of our retention time and peak width predictions, even without a column oven, is sufficient to lead us toward optimal conditions within realistic constraints. We experimentally showed that a model made from one column can be successfully applied to other columns of different lengths packed with the same stationary phase.

Systematic retention time inaccuracies may occur in the model, but they tend to cause predictions among similar peaks to be off in the same direction thereby limiting the impact on predicted *Rs*, especially when the solutes are chemically similar. When dissimilar solutes with different slopes in their retention models fail to elute at the predicted *Rs*, their dissimilarity provides an opportunity to improve the separation, often with only a minor change in *%B* (plus changes in the other parameters, as necessary, to arrive at an optimum given the resolution constraint of the problem).

We also studied the complexity of our separation and examined interactions of *%B* and several efficiency parameters. While a one-parameter-at-a-time optimization of these same parameters can obviously provide performance improvements, HPLC is clearly too complicated and its parameters too interacting for irregular samples to expect such a strategy to lead to the global optimum within a realistic amount of work. Here we examined the optimum analysis time locus over five dimensions to reveal its behavior. Without some knowledge of how the system performance behaves over all parameters, a user working from an arbitrary starting point or from an existing method may not realize that a large performance improvement is possible with a different combination of parameters. However, the level of detail shown here is unnecessary

in routine work because the desired optimum can be calculated directly from the model while bypassing interaction plots like Figs. 5 and 7–9. In practice, once the screening separations are finished, the model can be generated and the desired optimum directly calculated in a few minutes.

We continue to be surprised by the specific parameter values that, when combined, give optimal solutions to separations. This occurs because HPLC practices regarding parameter value selection were developed mostly by considering the parameter influences one at a time rather than including the interactions, and because the guidelines were developed at a time when considering several parameters simultaneously was not possible except by time-consuming experimentation. Simulation and optimization tools provide the benefits of exhaustive experimentation easily and automatically.

We showed a very fast separation that provides baseline separation and stays within practical limits using what many chromatographers would consider an absurdly high retention factor. This occurred primarily because the sample is highly irregular in its retention behavior, but irregularity is common in mixtures of chemically dissimilar solutes. Guidelines regarding acceptable parameter values may often be incorrect.

As we continue to work with unusual combinations of parameter values as in Fig. 10 to achieve practical benefits, we will no doubt encounter additional broadening effects not adequately estimated in the current model. We already see unmodeled broadening when flow rates approach 5 mL/min. Heating due to the work required to transport mobile phase through a column will not always be negligible, particularly as stationary phase particles get smaller and higher pressure drops and faster mobile phase velocities are required. Nonuniform column temperature will lead to different flow behavior and additional broadening mechanisms in the column [35]. HPLC will become even more complicated when these effects become significant and as more parameters are required and included in optimization efforts. The difficulty of empirical experimental optimization and the need for accurate models and computerized tools for simulation and optimization can only increase.

Acknowledgement

We gratefully acknowledge the support of Wayne Way (Sigma–Aldrich) for providing the matched columns used in this work.

Appendix A. Equations and methods used for predicting retention times and peak widths

Retention factors on the column are calculated from experimental retention times as

$$k = \frac{t_r - t_m}{t_m - t_{ex}} \tag{A1}$$

where k is the retention factor, t_r is the observed retention time for the solute, t_m is the experimentally observed void time for the column installed in the system (that is, the void time for the column plus the void time for the extra-column system components), and t_{ex} is the void time for the extra-column system components for the screening conditions.

A retention model for a solute is built on screening data acquired using one column at an arbitrary reference flow rate. Data are acquired using several different values of modifier concentration (%B). The retention factor for a solute at a particular %B is predicted using an empirical relation [24],

$$\log k = a + b(\%B) + c(\%B)^2 \tag{A2}$$

where a, b, and c are coefficients found by regression of screening data. Retention times can then be predicted by solving Eq. (A1) for t_r and adjusting for flow rate and column length differences from the screening conditions.

The reduced plate height (h) is

$$h = \frac{H}{d_p} \tag{A3}$$

where H is the plate height for the column and d_p is the stationary phase particle diameter; h is estimated using the equation of Kennedy and Knox [32]:

$$h = A\nu^{1/3} + \frac{B}{\nu} + C\nu \tag{A4}$$

where A, B, and C are coefficients set to 1, 2, and 0.05, respectively. The variable ν is the reduced mobile phase velocity on the column,

$$\nu = \frac{Fd_p}{\varepsilon D_M} \tag{A5}$$

where F is the flow rate, d_p is the stationary phase particle diameter, ε is the total column porosity, and D_m is the solute diffusion coefficient. N can then be estimated:

$$N = \frac{L}{hd_p} \tag{A6}$$

where L is the column length.

We used standard deviation to describe peak width. The standard deviation from the column (in time units) is

$$\sigma_C(t) = (t_r - t'_{ex})N^{-1/2} \tag{A7}$$

where N is estimated at the current flow rate and t'_{ex} is the extra-column void time evaluated at the current flow rate.

The standard deviation from transport through extra-column tubes in the system is calculated using the equation of Atwood and Golay [30]:

$$\sigma_T(t) = \frac{V_T}{F} N_T^{-1/2} \left(1 + \frac{3}{N_T}\right)^{-1/4} \tag{A8}$$

Here N_T is the plate number for the tube and is given by $N_T = 75.40 D_M L_T/F$, and L_T is the tube length. We did this separately for the injection loop tube (but only the length occupied by the sample), the column inlet tube, and the column outlet tube. Transport through the detector cell was ignored (see text).

The standard deviations from the injection volume and from the detector cell volume are

$$\sigma_{inj}(t) = \frac{1}{\sqrt{12}} \frac{V_{inj}}{F} \tag{A9}$$

and

$$\sigma_{det}(t) = \frac{1}{\sqrt{12}} \frac{V_{det}}{F} \tag{A10}$$

where V_{inj} and V_{det} are the injection volume and detector cell volume, respectively [36].

All the contributors of standard deviation are assumed to be independent, and the predicted peak width seen in the detector is estimated by summing the variances,

$$\sigma(t) = \sqrt{\sum_i (\sigma_i(t))^2} \tag{A11}$$

where i is the index representing the various sources of broadening.

References

[1] L.R. Snyder, J.J. Kirkland, J.W. Dolan, Introduction to Modern Liquid Chromatography, 3rd ed., John Wiley & Sons Inc., Hoboken, NJ, 2010, pp. 19–86 and 475–498.

[2] P.J. Schoenmakers, N. Dunand, A. Cleland, G. Musch, T. Blaffert, Chromatographia 26 (1988) 37–44.
[3] G.W. Leibniz, in: L.E. Loemker (Ed.), Philosophical Papers and Letters. A Selection Translated and Edited, second ed. (second print), Kluwer Academic Publishers, Dordrecht, 1989, p. 61.
[4] Optimization Theory, in: The Penguin Dictionary of Mathematics, 2003, Retrieved from: http://www.credoreference.com/entry/penguinmath/optimization_theory, October 28, 2009.
[5] F.H. Walters, L.R. Parker Jr., S.L. Morgan, S.N. Deming, Sequential Simplex Optimization, CRC Press, Inc., Boca Raton, 1991, pp. 31–35.
[6] G. Guiochon, Anal. Chem. 52 (1980) 2002–2008.
[7] L.R. Snyder, J.W. Dolan, High-Performance Gradient Elution, John Wiley & Sons, Hoboken, 2007, pp. 19–21.
[8] P.J. Schoenmakers, M. Mulholland, Chromatographia 25 (1988) 737.
[9] P.J. Schoenmakers, N. Dunand, J. Chromatogr. 485 (1989) 219–236.
[10] P.J. Schoenmakers, A. Peeters, R.J. Lynch, J. Chromatogr. 506 (1990) 169–184.
[11] C.H. Bryant, A. Adam, D.R. Taylor, R.C. Rowe, Anal. Chem. Acta 297 (1994) 317–347.
[12] M.A. Stone, J. Liq. Chromatogr. Relat. Technol. 30 (2007) 605–647.
[13] V.L. McGuffin, in: E. Heftmann (Ed.), Chromatography: Fundamentals and Applications of Chromatography and Related Differential Migration Methods – Part A. Fundamentals and Techniques, sixth ed., Elsevier, Amsterdam, 2004, pp. 6–9.
[14] L.R. Snyder, J.J. Kirkland, J.L. Glajch, Practical HPLC Method Development, second ed., John Wiley & Sons, Inc., New York, 1997, pp. 410–431.
[15] J.W. Dolan, LC–GC North Am. 27 (2009) 606–610.
[16] H. Poppe, J. Chromatogr. A 778 (1997) 3–21.
[17] S.-T. Popovicia, P.J. Schoenmakers, J. Chromatogr. A 1073 (2005) 87–91.
[18] A. de Villiers, F. Lestremau, R. Szucs, S. Gelebart, F. David, P. Sandra, J. Chromatogr. A 1127 (2006) 60–69.
[19] P.W. Carr, X. Wang, D.R. Stoll, Anal. Chem. 81 (2009) 5342–5353.
[20] Y. Zhang, X. Wang, P. Mukherjee, P. Petersson, J. Chromatogr. A 1216 (2009) 4597–4605.
[21] S. Feketea, K. Ganzlera, J. Fekete, J. Pharm. Biomed. Anal. 51 (2010) 56–64.
[22] E. Oláh, S. Fekete, J. Fekete, K. Ganzler, J. Chromatogr. A 1217 (2010) 3642–3653.
[23] H.-J. Molnár, K.E. Rieger, Monks, J. Chromatogr. A 1217 (2010) 3193–3200.
[24] P.J. Schoenmakers, H.A.H. Billiet, L. de Galan, J. Chromatogr. 185 (1979) 179.
[25] H. Darcy, Les Fontaines Publiques de la Ville de Dijon, Dalmont, Paris, France, 1856.
[26] B.F. Karger, L.R. Snyder, C. Horvath, An Introduction to Separation Science, John Wiley & Sons, New York, 1973, p. 90.
[27] J.C. Giddings, Unified Separation Science, John Wiley & Sons, New York, 1991, pp. 62–65.
[28] T.L. Chester, J. Chromatogr. A 1016 (2003) 181–193.
[29] T.L. Chester, S.O. Teremi, J. Chromatogr. A 1096 (2005) 16–27.
[30] J.G. Atwood, M.J.E. Golay, J. Chromatogr. 218 (1981) 97.
[31] R.C. Reid, J.M. Prausnitz, B.E. Poling, The Properties of Gases and Liquids, fourth ed., McGraw-Hill, Inc., New York, 1987, pp. 598–611.
[32] G.J. Kennedy, J.H. Knox, J. Chromatogr. Sci. 10 (1972) 549.
[33] C.F. Poole, S.K. Poole, Chromatography Today, Elsevier, Amsterdam, 1991, p. 362.
[34] R.J. Laub, J.H. Purnell, J. Chromatogr. 112 (1975) 71–79.
[35] G. Guiochon, J. Chromatogr. A 1126 (2006) 6–49.
[36] R.P.W. Scott, J. Liq. Chromatogr. Relat. Technol. 25 (2002) 2567–2587.