Computer Simulation for the Simultaneous Optimization of Any Two Variables and Any Chromatographic Procedure

P. Haber, T. Baczek, and **R. Kaliszan**
Department of Biopharmaceutics and Pharmacodynamics, Medical University of Gdansk, Gen. J. Hallera 107, 80-416 Gdansk, Poland

L.R. Snyder, J.W. Dolan, and **C.T. Wehr**
LC Resources Inc., 2930 Camino Diablo, Suite 110, Walnut Creek, CA 94596

Abstract

Computer software that allows the simulation of any chromatographic separation as a function of simultaneous changes in any one or two variables that can affect sample separation order (selectivity) is described. For one example, an application is described for the simultaneous variation of the mobile phase pH and gradient time in reversed-phase liquid chromatography. The accuracy of such predictions is examined for a sample mixture of 17 substituted benzoic acids and anilines, and requirements for an acceptable predictive accuracy are summarized. In a second example, the separation of three peptides by capillary electrophoresis is optimized.

Introduction

One major challenge in the separation of complex samples by reversed-phase liquid chromatography (RP-LC) is the selection of experimental conditions that can provide acceptable band spacing and resolution. The simultaneous optimization of two separation variables is especially effective but potentially tedious in practice. However, this tediousness can be overcome by the use of an appropriate experimental design with a computer program that allows for the prediction of separation for intermediate conditions. This approach was introduced by Kirkland et al. in 1980 (1) for the use of mobile phases containing methanol (MeOH), acetonitrile (ACN), tetrahydrofuran (THF), and water. This general procedure has since been extended to the use of other pairs of separation variables, examples of this are summarized in Table I. The commercial software described in this study enables the use of this approach for any pair of separation variables and for any chromatographic procedure, such as liquid chromatography (LC), capillary electrophoresis (CE), and supercritical fluid chromatography (SFC). The experimental design for one such application—optimization of mobile phase pH and the isocratic volume percentage of the organic solvent in mixtures of the water solvent A and the organic solvent B (%B)—in RP-LC is illustrated in Figure 1. A minimum of three different pH values (e.g., 3.0, 3.5, and 4.0) are required, but additional experiments can be added (for instance, pH 4.5) to allow for the predictions of separation for a wider pH range.

With one exception, the procedures in Table I involved isocratic separation; however, an almost identical approach can be used for corresponding gradient separations (e.g., [7] versus [1], or [8] versus [6]). In this case, gradient time

Table I. Examples of High-Performance Liquid Chromatographic Method Development That Involve the Simultaneous Variation of Two Experimental Conditions

Example	Experimental conditions
1	Quaternary-solvent mobile phase using (MeOH), (THF), (ACN), and water with solvent strength and run time held constant (reference 1)
2	Ternary-solvent mobile phase using MeOH, ACN, and water or MeOH, THF, and water with solvent strength allowed to vary (reference 2)
3	Binary-solvent mobile phase with pH and ion-pairing reagent concentration varied (reference 3)
4	Column type and mobile phase with percentage of organic allowed to vary (reference 4)
5	Mobile phase pH and percentage organic allowed to vary (reference 5)
6	Temperature and gradient steepness allowed to vary (reference 6)

(t_G) replaces %B in the design of Figure 1. The theory of gradient elution also allows for the use of gradient experiments in the prediction of corresponding isocratic separations (9) such as separations in which all conditions are the same except when values of %B replace t_G and the initial and final %B values for the gradient. For example, beginning with an experimental design for the separation as a function of pH and t_G (as in Figure 1), it is possible to predict isocratic separation as a function of pH and %B without the need for additional isocratic experiments.

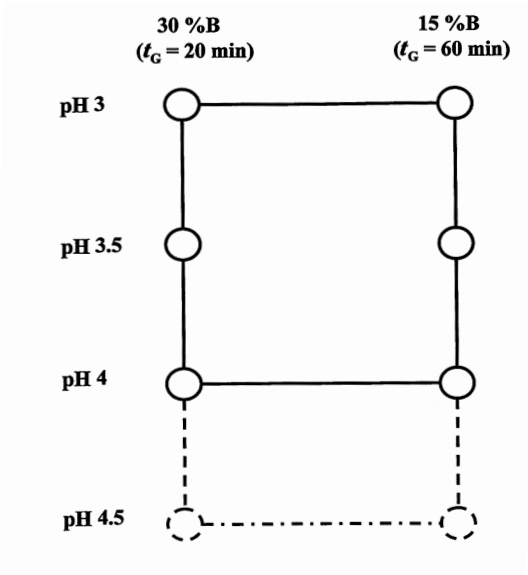

Figure 1. Illustration of an experimental design for optimizing pH and either t_G or %B. Individual circles represent conditions of pH and t_G to be used in experiments for computer simulation, and broken circles refer to optional additional experiments.

There are significant advantages in using initial gradient experiments for the prediction and optimization of isocratic separations as opposed to beginning with isocratic experiments. Thus, it is desirable that initial isocratic experiments provide retention factors (k) for all bands in the sample that fall within a certain range of values, for instance, $0.5 < k < 20$ (10). Because retention varies with both the %B and the second variable used to optimize the band spacing (e.g., pH and temperature), choosing conditions for the initial experiments of Figure 1 may not be straightforward. This is illustrated in Figure 2 with a mixture of eight substituted benzoic acids. For separation at pH 3.0, appropriate values of %B can be found by trial-and-error or other means. In this case, mobile phases of 15 and 30 %B provide a suitable range in k and partial separation ($R_s > 0.6$) of all bands. However, when these same %B values are chosen for pH 6.0, there is extensive overlap of most bands, especially for 30 %B. The latter two separations (Figures 2C and 2D) are unsuitable for use in a computer simulation because of the uncertainty in the exact retention times of individual solutes. This problem can be circumvented by decreasing %B for the two runs at pH 6.0, but additional trial-and-error experiments would then be required.

For the case of initial gradient experiments, suitable gradient retention factors (9) are predictably determined by the choice of gradient conditions; thus, there is no need for trial-and-error experiments as a means of adjusting k. As a result, it suffices to carry out two gradients with different t_G at each pH value. This is illustrated in the corresponding gradient separations in Figure 3. Although there is greater compression of the chromatogram at pH 6 versus 3 and some band overlap, all four of the separations in Figure 3 would be suitable for use as input in a computer simulation.

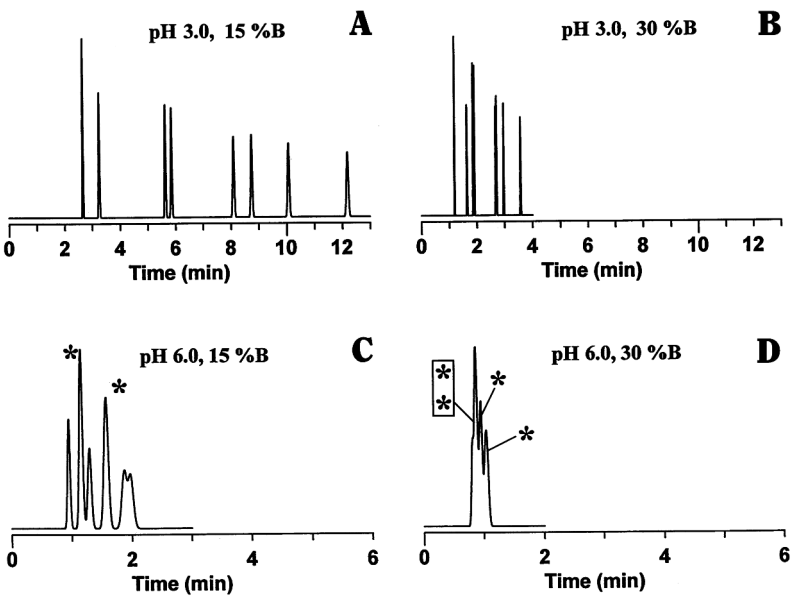

Figure 2. Illustrations of the difficulty in choosing isocratic conditions for the experimental design of Figure 1. The sample is the eight substituted benzoic acids of the Experimental section. The computer simulations are also based on data from the Experimental section. Also indicated are instances of overlapped band-pairs (*) and triplet (**).

In the present study, we have examined the simultaneous optimization of the pH level and t_G. It has been established elsewhere (11) that isocratic predictions from gradient data are generally less accurate than gradient predictions. However, one additional experiment can be used to correct errors in such isocratic predictions (11).

Several reports describe the theory of RP-LC retention as a function of mobile phase pH (12–15). Three experiments in which only pH was varied allowed for the predictions of retention for other pH values based on the applicability of the Henderson–Hasselbach equation. However, this relationship appears to be of marginal reliability, especially for the separation of basic solutes (14,15) that can be retained by ion exchange onto ionized silanols as well as by a reversed- phase process. The use of the Henderson–Hasselbach equation for predicting separation as a function of pH can be suc-

cessful for interpolation within a narrow range of pH (12,13), but the exploration of a wide pH range would then be precluded. An alternative approach (as in the present study) would be to use a more general fitting expression (e.g., cubic spline) for retention versus pH that would allow for the use of a larger number of experiments in which pH is varied. This leads to the experimental design illustrated in Figure 1 for the simultaneous optimization of pH and either t_G or isocratic %B. A minimum of three experiments with varying pH are required, but a total of as many as 16 experiments are allowed by the present software (eight different pH values, each with two different values of t_G).

Given that gradient retention can be predicted as a function of pH for each of the two t_Gs in the experiments of Figure 1, retention at any given pH can then be predicted as a function of t_G (9). Isocratic retention can likewise be predicted from a well-known empirical relationship (16):

$$\log k = \log k_w - S\phi \qquad \text{Eq. 1}$$

where k_w refers to the extrapolated value of k for water as mobile phase ($\phi = 0$), ϕ is the volume fraction of organic in the mobile phase (equal to 0.01 %B), and S is virtually constant for a given solute with only ϕ varying. Therefore, the gradient experiments of Figure 1 allow for the prediction of values of k_w and S as a function of pH, which in turn allows isocratic predictions of retention as a function of pH and %B. Values of resolution (R_s) are calculated by the computer program from predicted retention times and experimental bandwidths. In the absence of (preferred) experimental bandwidth data, the program can estimate RP-LC bandwidths from the experimental conditions.

The range in pH for which predictions are possible is limited to the values chosen initially (as in Figure 1); for example, no extrapolation is allowed for outside of the original pH range. The increments in pH (increments of 0.5 in Figure 1) must be small enough for accurate interpolations. Previous isocratic studies (12,13) in which only the pH level was varied and the sample consisted of either acidic or basic components suggest that acceptable predictive accuracy results for increments in pH are as large as 1.0. However, it can be anticipated that when both acids and bases are present as the sample components and both pH and either t_G or %B are varied, smaller increments in pH may be necessary. This was found true for the present sample.

Experimental

Equipment and procedures

High-performance liquid chromatographic separation

The high-performance liquid chromatographic (HPLC) system was an LC Model 1 (Waters Associates, Milford, MA) with a dwell volume of 4.3 mL. Conditions were as follows: a 15- × 0.46-cm Inertsil C18 column (GL Sciences, Tokyo, Japan), solvent A consisted of potassium citrate adjusted with citric

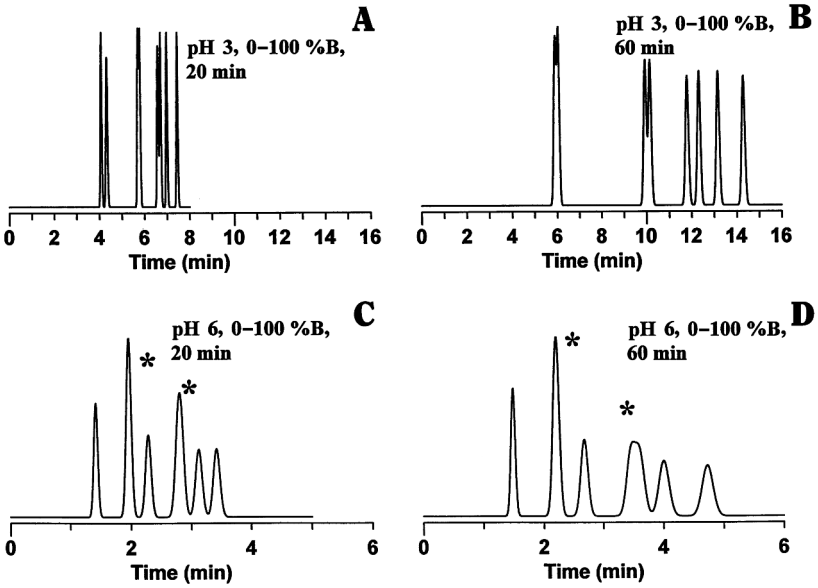

Figure 3. Illustrations of the ease in choosing gradient conditions for the experimental design in Figure 1. The sample is the eight substituted benzoic acids of the Experimental section. The computer simulations are also based on data from the Experimental section. Indicated are instances of overlapped band pairs (*).

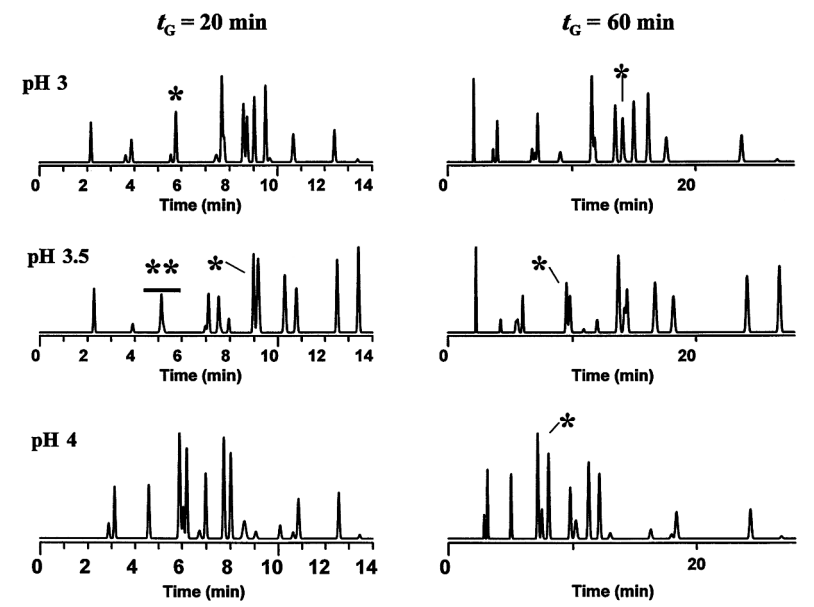

Figure 4. Initial experimental runs for the optimization of mobile phase pH and t_G or %B. Indicated are instances of overlapped band pairs (*).

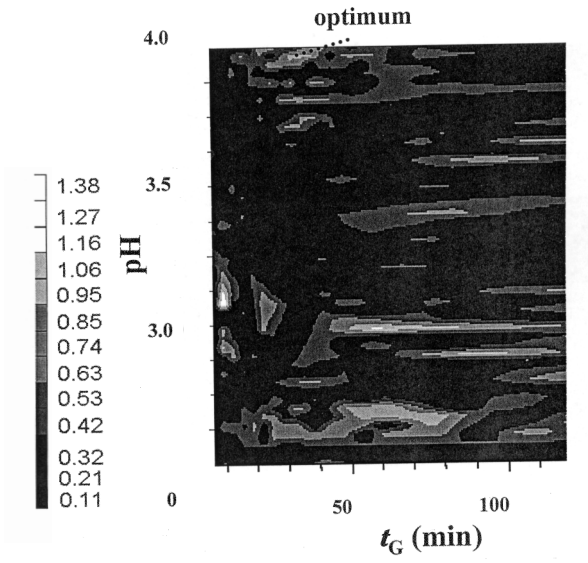

Figure 5. Resolution map for the sample and conditions of Figure 2.

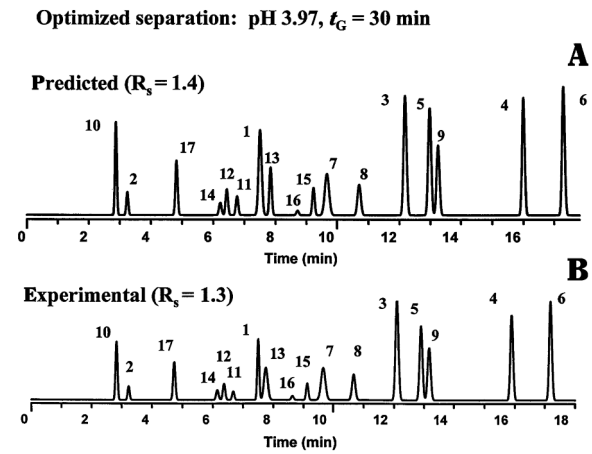

Figure 6. Predicted and actual optimum separations from the resolution map of Figure 3. Conditions are the same as in the Experimental section, except t_G = 30 min and the pH is 3.97. The actual separation is reconstructed from experimental retention data.

Table II. Summary of Errors in the Prediction of Separation as a Function of Mobile Phase pH and t_G

		Average error in R_s for indicated pH values of input runs	
pH	t_G (min)	pH levels 3.0, 3.5, and 4.0 (good*)	pH levels 3.0, 4.0, 5.0, and 6.0 (poor †)
3.25	20	0.2	1.0
	60	0.3	1.3
3.50	40	0.2	1.5
	45	0.2	1.3
3.60	40	0.4	1.6
	45	0.3	1.4
3.75	20	0.2	1.4
	60	0.4	1.5
Average		0.3	1.4

* Good input conditions signify runs with pH values that do not change by more than 0.5.
† Bad runs change by at least 1.0 pH level.

acid to required pH (25mM), and solvent B consisted of ACN (35°C, 2 mL/min). Gradients were 5–100 %B in 20 and 60 min.

CE separation

A BioFocus 3000 CE system (BioRad, Richmond, CA) was used with a 24-cm × 50-mm-i.d. capillary and ultraviolet detection at 200 nm. Sodium phosphate was used as buffer, with the pH and buffer concentration varied as described when variation occurred. The capillary temperature was controlled at 20°C with an applied voltage of 5 kV. Sample introduction was for 3 s at 5 psi.

Samples

HPLC separation

The sample contained the following 17 components: 3-methylaniline (compound 1), 4-methylaniline (compound 2), 4-chloroaniline (compound 3), 3,4-dichloroaniline (compound 4), 3-chloroaniline (compound 5), 3,5-dichloroaniline (compound 6), *N*-ethylaniline (compound 7), 3,5-dimethylaniline (compound 8), 2-chloroaniline (compound 9), 2-nitrobenzoic acid (compound 10), 3-cyanobenzoic acid (compound 11), 2-fluorobenzoic acid (compound 12), 3-nitrobenzoic acid (compound 13), 2-chlorobenzoic acid (compound 14), 3-fluorobenzoic acid (compound 15), 2,6-dimethylbenzoic acid (compound 16), and phthalic acid (compound 17).

CE separation

This sample consisted of three peptides: bradykinin, a thyrotropin-releasing hormone, and leu-enkephalin.

Computer simulation

DryLab 2000 software (LC Resources, Walnut Creek, CA) was used for the predictions of separation as a function of pH and either t_G or %B. This general-purpose program allowed for the simultaneous variation of any two separation variables for any chromatographic or electrophoretic procedure (e.g., LC, CE, and SFC). When one of the starting variables was t_G, predictions of either isocratic or gradient separation were possible. The simultaneous optimization of temperature and either t_G or %B has been reported (8).

For predictions of CE separation, band migration times and bandwidths for each band were entered manually into the DryLab 2000 software; however, it was also possible to transfer this data automatically from some data systems into DryLab 2000. For applications in which retention or bandwidth models or both were not assumed, both retention times and bandwidths were interpolated using a cubic spline fit.

Results and Discussion

HPLC separation

Gradient predictions from gradient data

Two different sets of starting experiments were investigated using pH increments of

either 0.5 or 1.0. Figure 4 shows the starting experiments for three pH values (3.0, 3.5, and 4.0) with increments of 0.5. Entry of these data (experimental conditions, retention times, and band widths) into the DryLab program allowed for the prediction of separation as a function of pH and gradient time. The choice of optimum conditions was readily made on the basis of a resolution map for these separations, as shown in Figure 5. Maximum resolution was indicated for a pH of 3.97 and t_G of 30 min (R_s = 1.4). Although Figure 5 shows a narrow range in pH and t_G for acceptable resolution, little difficulty was experienced in matching the predicted separation, as shown in Figure 6 (all peaks in Figures 6A and B have the same retention order). There was reasonable agreement between these two chromatograms; however, this was not a critical test of the accuracy of the computer simulation because the optimum pH of 3.97 was quite close to one of the experimental input values (pH 4.0).

Further predictions were made for values of pH and t_G other than those of Figure 4 for the six input runs. These results are summarized in the third column of Table II. The average error for all predictions of resolution was 0.3 for R_s, which is considered acceptable for the purposes of method development (11). Figure 7 compares actual and predicted separations for one of the examples of Table II (pH 3.25, t_G = 60 min).

In a second set of experiments, initial experiments were carried out for pH 3, 4, 5, and 6, and predictions for the same conditions as in Table II were repeated. Figure 8 compares actual

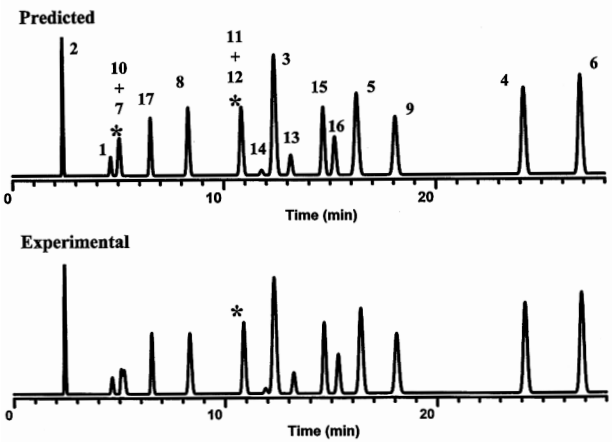

Figure 7. Predicted and actual separations based on the experiments of Figure 2 (pH levels of 3.0, 3.5, and 4.0; t_G = 20 and 60 min). Conditions are the same as in the Experimental section, except pH is 3.25 and t_G = 60 min. The actual separation is reconstructed from experimental retention data. Indicated are instances of overlapped band pairs (*).

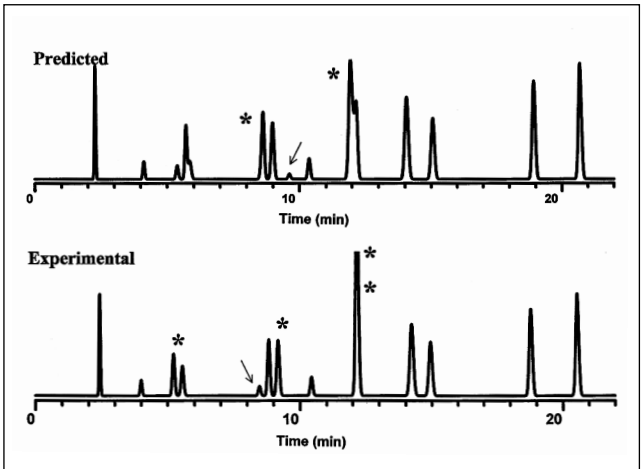

Figure 8. Predicted and actual separations based on experiments similar to those of Figure 2 (pH is 3.0, 4.0, 5.0, and 6.0; t_G= 20 and 60 min). Conditions are the same as in the Experimental section, except pH is 3.5 and t_G = 40 min. The actual separation is reconstructed from experimental retention data. Indicated are instances of overlapped band pairs (*).

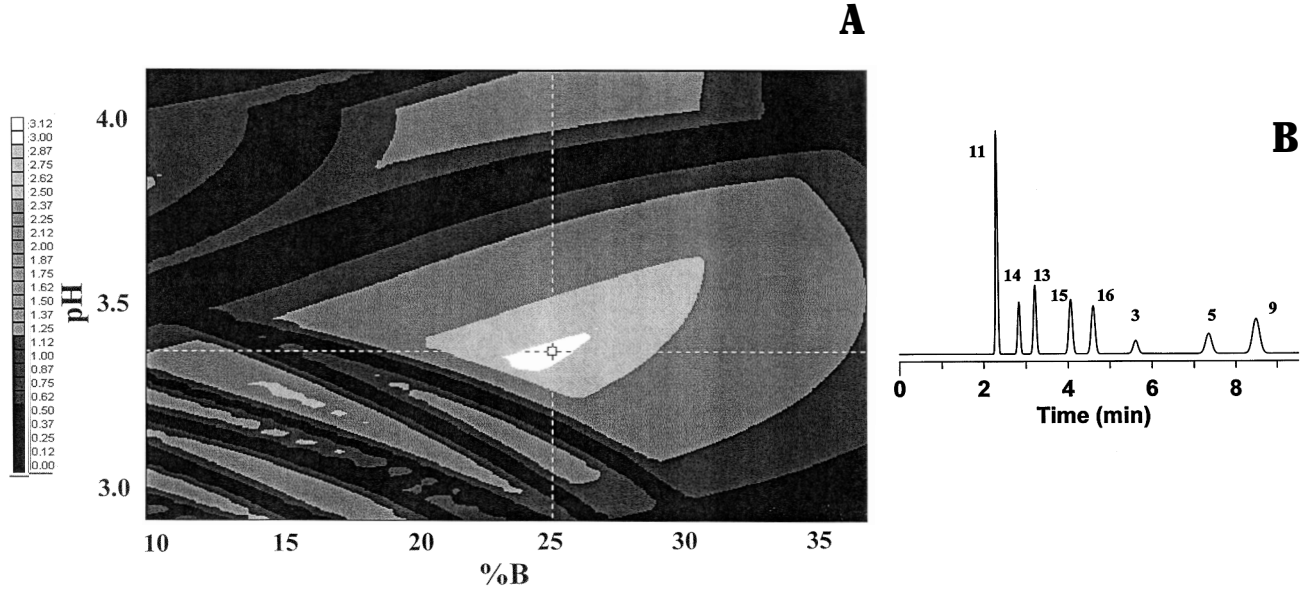

Figure 9. Isocratic predictions of separation as a function of pH and %B based on the gradient experiments of Figure 4. The samples used were compounds 3, 5, 9, 11, and 13–16. Resolution map, A, and the predicted separation for maximum resolution, B (R_s = 3.1, 25 %B, and pH 3.35).

and predicted separations for one of the examples in Table II (pH 3.50, t_G = 40 min) based on these initial four pH values. As seen in the last column in Table II, errors in R_s were now much larger (the average error in R_s was 1.4, which is unacceptably large). The "extra" peaks (marked by arrows in Figure 8) should also be noted. The error for these predictions of Table II was approximately the same, whether the value of t_G was the same as in the initial experiments (20 and 60 min) or not (40 and 45 min). This suggests that error arises mainly from the prediction of retention as a function of pH rather than t_G. This agrees with other studies (11), which show that predictive errors associated with a change in t_G are usually small.

For previous studies (12,13) in which only pH was varied, it is indicated that predictive errors of 0.4 R_s units or less can be expected, even for a pH increment as large as 1.0 unit. The present sample (composed of approximately an equal number of acidic and basic solutes) may represent a worst-case situation. Thus, as pH is varied, acidic and basic solutes move in opposite directions within the chromatogram, which leads to rapid changes in resolution for small changes in pH. The practical conclusion we draw from Figure 8 and the latter example is that accurate predictions of the effect of varying pH require narrow pH increments in the input runs (no more than 0.5 pH units if both acids and bases are present in the sample, but possibly as large as 1.0 pH units if only acids or bases are present).

Isocratic predictions from gradient data

The use of gradient data to optimize pH and %B for isocratic separation is illustrated in Figure 9. Isocratic separation of the entire sample (compounds 1–17) was not possible because of the wide range in retention of these compounds. Therefore, compounds 3, 5, 9, 11, and 13–16 (which elute within a narrower range of k values) were selected as example. Experimental data for pH 3, 3.5, and 4 and t_G = 20 and 60 min were used as inputs for the computer simulation, which resulted in the resolution map of Figure 9A. Maximum resolution occurred for 25 %B and pH 3.35 (shown by the crosshairs in Figure 9A). The predicted separation (R_s = 3.1) is shown in Figure 9B. This result was not confirmed experimentally, but the prediction can be expected to be less reliable than in the example in Figure 7; some other related examples can also be observed (11). However, such predictions of isocratic separation from the gradient data for when pH and %B varies can be improved using one additional experiment whose conditions can be predicted by computer simulation, as described previously (11).

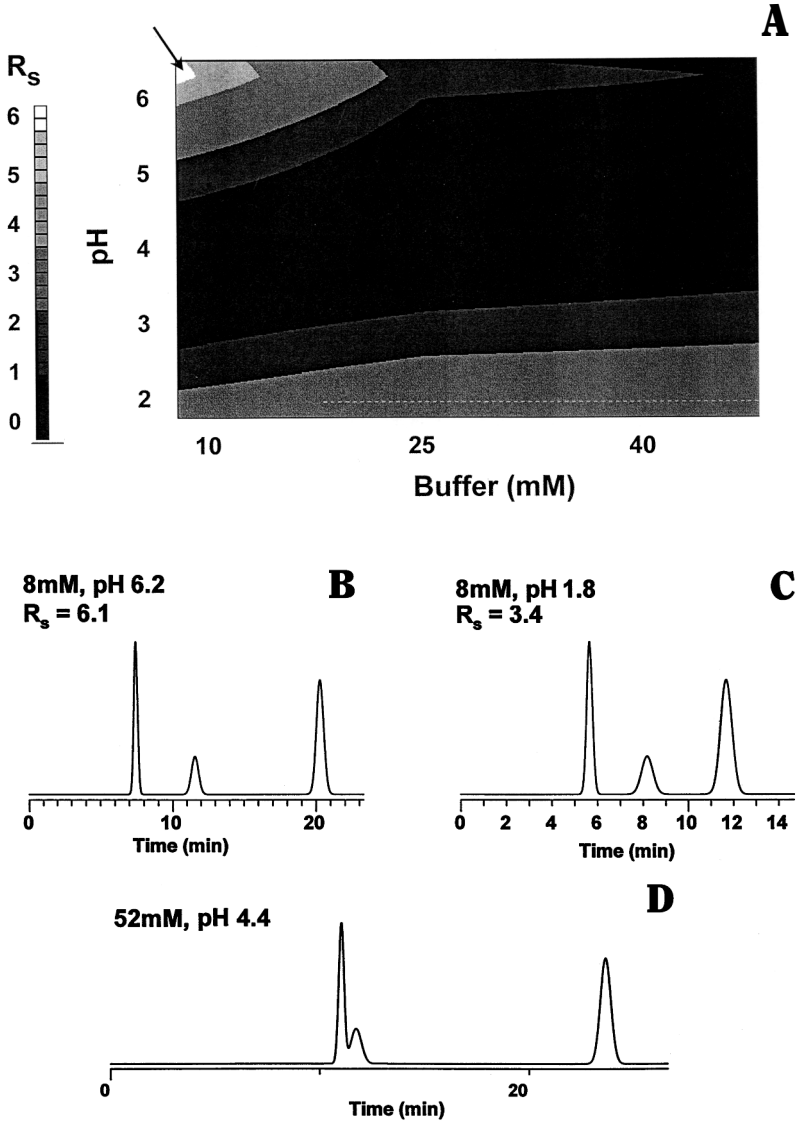

Figure 10. Separation of a mixture of three peptides (bradykinin, thyrotropin-releasing hormone, and leu-enkephalin) by CE. Resolution map as a function of pH and buffer concentration, A; separation for conditions giving maximum resolution, B; separation for conditions giving a shorter run time, C; and separation for conditions that give poor separation, D.

CE separation

The three peptides described in the Experimental section were separated at three different pH values and three different buffer concentrations. Migration times and bandwidths from these nine experiments were entered into the DryLab 2000 software for subsequent predictions of separation as a function of the conditions. A resolution map as a function of pH and buffer concentration is shown in Figure 10A. Maximum resolution (R_s = 6.1) occurred for buffer concentrations at 8mM and pH 6.2 (shown by the arrow). The predicted separation is shown in Figure 10B. Because the latter separation provided greater resolution than needed, it was possible to choose other conditions that gave acceptable resolution with a shorter time. An example of this is shown in Figure 10C for a 8mM buffer concentration and pH 1.8 (R_s = 3.4). In this case, the run time was approximately half that of

Figure 10B. A final example of the predicted separation is shown in Figure 10D for conditions that give poor separation (52mM, pH 4.4). None of these predictions were verified experimentally. This example was intended only to illustrate the application of the software for separations other than those by HPLC. However, because the latter predictions were based on simple interpolation of retention and bandwidth data, acceptable accurate predictions would necessarily result if conditions (e.g., pH and buffer concentration) for the starting experiments were spaced closely enough (this was not the case for the example in Figure 10).

Conclusion

It is generally recognized that the simultaneous optimization of two different separation conditions can be effective for the separation of many samples by means of HPLC or other chromatographic methods. The efficient application of this approach requires a good experimental design and the use of computer software that can predict separation on the basis of a relatively small number of initial experiments. It is also advantageous to have an unrestricted choice of the chromatographic procedure (e.g., HPLC, CE, or SFC) and the two variables that are to be optimized. Software was described that meets these requirements.

Two examples of this approach for chromatographic method development were shown. In the first illustration, the separation of a 17-component mixture (substituted anilines and benzoic acids) was optimized as a function of mobile phase pH and t_G. Six initial calibration experiments varying in pH and t_G were used for subsequent computer simulations. It was found for these samples that acceptable predictions of separation required a pH spacing of the calibration runs that were less than or equal to pH increments of 0.5.

In a second example, a sample containing three peptides was separated by CE as a function of buffer concentration and pH.

Acknowledgments

This study was supported in part by a Small Business Innovation Research grant from the National Institutes of Health (U.S. Department of Health and Human Services). The columns used were gifts from GL Sciences.

References

1. J.L. Glajch, J.J. Kirkland, K.M. Squire, and J.M. Minor. Optimization of solvent strength and selectivity for reversed-phase liquid chromatography, using an interactive mixture-design statistical technique. *J. Chromatogr.* **199:** 57–79 (1980).
2. J.W. Weyland, C.H.P. Bruins, and D.A. Doornbos. Use of three-dimensional minimum α-plots for optimization of mobile phase composition for RP-HPLC separation of sulfonamides. *J. Chromatogr. Sci.* **22:** 31–39 (1984).
3. R.C. Kong, B. Sachok, and S.N. Deming. Combined effects of pH and surface-active-ion concentration in reversed-phase liquid chromatography. *J. Chromatogr.* **199:** 307–316 (1980).
4. J.J. DeStefano, J.A. Lewis, and L.R. Snyder. Reversed-phase HPLC method development based on column selectivity. *LC/GC Mag.* **10:** 130–139 (1992).
5. B. Bourguignon, P.F. de Agular, M.S. Khots, and D.L. Massart. Optimization in irregularly shaped regions: pH and solvent strength in reversed-phase liquid chromatography separations. *Anal. Chem.* **66:** 893–904 (1994).
6. J.W. Dolan, L.R. Snyder, N.M. Djordjevic, D.W. Hill, D.L. Saunders, L. Van Heukelem, and T.J. Waeghe. Simultaneous variation of temperature and gradient steepness for reversed-phase HPLC method development. I. Application to 14 different samples using computer simulation. *J. Chromatogr. A* **803:** 1–31 (1998).
7. J.J. Kirkland and J.L. Glajch. Optimization of mobile phases for multisolvent gradient elution liquid chromatography. *J. Chromatogr.* **255:** 27–39 (1983).
8. R.G. Wolcott, J.W. Dolan, and L.R. Snyder. Computer simulation for the convenient optimization of isocratic reversed-phase liquid chromatography separations by varying temperature and mobile phase strength (%B). *J. Chromatogr. A* **869:** 3–25 (2000).
9. L.R. Snyder and J.W. Dolan. The linear-solvent-strength model of gradient elution. *Adv. Chromatogr.* **38:** 115–87 (1998).
10. L.R. Snyder, J.L. Glajch, and J.J. Kirkland. *Practical HPLC Method Development*, 2nd ed. Wiley-Interscience, New York, 1997, pp 31–34.
11. J.W. Dolan, L.R. Snyder, R.G. Wolcott, P. Haber, T. Baczek, and R. Kaliszan. Reversed-phase separation of complex samples by optimizing temperature and gradient time. III. Improving the accuracy of computer simulation. *J. Chromatogr. A* **857:** 41–68 (1999).
12. J.W. Dolan, D.C. Lommen, and L.R. Snyder. HPLC computer simulation based on a restricted multi-parameter approach. *J. Chromatogr.* **535:** 55–74 (1990).
13. J.A. Lewis, D.C. Lommen, W.D. Raddatz, J.W. Dolan, L.R. Snyder, and I. Molnar. Computer simulation for the prediction of separation as a function of pH for reversed-phase HPLC. I. Accuracy of a theory-based model. *J. Chromatogr.* **592:** 183–195 (1992).
14. P.J. Schoenmakers, R. Tijssen, J.W. Dolan, D.C. Lommen, and L.R. Snyder. Modelling retention of ionogenic solutes in liquid chromatography as a function of pH for optimization purposes. *J. Chromatogr. A* **656:** 577–90 (1993).
15. D. Sykora, E. Tesarova, and M. Popl. Interactions of basic compounds in reversed-phase liquid chromatography. Influence of sorbent character, mobile phase composition, and pH on retention of basic compounds. *J. Chromatogr. A* **758:** 37–51 (1997).
16. K. Valko, L.R. Snyder, and J.L. Glajch. Retention in reversed-phase liquid chromatography as a function of mobile phase composition. *J. Chromatogr.* **656:** 501–520 (1993).
17. L.R. Snyder. Changing reversed-phase HPLC selectivity: which variables should be tried first? *J. Chromatogr. B* **689:** 105–115 (1997).
18. L.R. Snyder. New approaches to HPLC method development. *Today's Chemist at Work* **5:** 29–34 (1996).
19. J.W. Dolan, L.R. Snyder, N.M. Djordjevic, D.W. Hill, L. Van Heukelem, and T.J. Waeghe. Reversed-phase separation of complex samples by optimizing temperature and gradient time. I. Peak capacity considerations. *J. Chromatogr. A* **857:** 1–20 (1999).
20. P.L. Zhu, J.W. Dolan, L.R. Snyder, D.W. Hill, L. Van Heukelem, and T.J. Waeghe. Combined use of temperature and solvent strength in reversed-phase gradient elution. III. Selectivity for ionizable samples as a function of sample type and other separation conditions. *J. Chromatogr. A* **756:** 51–62 (1996).

Manuscript accepted June 21, 2000.