

Introduction to Chromatography: Instructor's Guide

David T. Harvey

Introduction

This learning module provides an introduction to chromatography with an emphasis on understanding basic chromatographic terms and measurements, and understanding the factors that affect the quality of a separation. The treatment is independent of the form of chromatography and, although data from GC and HPLC experiments are used to illustrate topics presented in some investigations, no emphasis is placed on particular chromatographic techniques. The module consists of the following nine investigations:

- A First Look at a Chromatogram
- A Closer Look at the Separation
- Partition Coefficients and Retention Factors
- Theoretical Plates
- A First Look at Chromatographic Resolution
- A Closer Look at Chromatographic Resolution
- Improving Resolution Through Efficiency
- Improving Resolution Through Selectivity
- Improving Resolution Through Retention

The learning module is programmed in R (see www.r-project.org) using the **Shiny** package, which makes it possible to include interactive features. Each investigation consists of a brief introduction, an explanation of any controls—sliders, radio buttons, and point selection tools—that are available to the user, and an explanation of the type of output produced by the underlying code. Each investigation also includes several questions to answer.

The purpose of this document is to provide instructors with additional background on the program's features and data sets, to provide representative examples of the results students might generate, and to provide suggestions of possible ways to make use of the module's investigations.

Some Background Details on the Learning Module's Data

The chromatograms included with these investigations—both full and partial—are simulated using functions written in R. The chromatograms in Investigations 1-4 model the separation as a countercurrent extraction using the general approach outlined in the paper “Tutorial: Simulating Chromatography with Microsoft Excel Macros,” the full reference for which is Kadjo, A.; Dasgupta, P. K. *Anal. Chim. Acta* **2013**, *773*, 1-8. The reason for using this approach is that it makes it possible to visualize the separation on the column as it takes place, a central aspect of these investigations.

The two peaks in the partial chromatograms included in Investigations 5 and 6 are simulated as Gaussian functions where the means are defined by the retention times ($\bar{X} = t_r$) and the standard deviations are defined by the widths at the baseline ($\sigma = w/4$). For the calculations in Investigation 6, the value of t_m is set to 1 min. The retention factor for the later eluting solute, k_B , is used to calculate $t_{r,B}$ and then this value and α are used to calculate $t_{r,A}$. The width of each peak is calculated using the number of theoretical plates, N , and the analyte's retention time. It is important to note here that an increase in the number of theoretical plates in these investigations is not accomplished by an increase in column length as this will affect an increase in the value of k_b , which, for this investigation, must remain constant.

The data in Investigation 7 are drawn from the paper “The Evaluation of the Parameters in the van Deemter Equation,” the full reference for which is Moody, H. W. *J. Chem. Educ.* **1982**, *59*, 290-291. In this paper,

the retention times and peak widths for a single analyte, 2-butanone, were measured using different flow rates of the carrier gas and the average height of a theoretical plate calculated and reported.

The data in Investigation 8 are drawn from the paper “Optimization of HPLC and GC Separations Using Response Surfaces: Three Experiments for the Instrumental Analysis Laboratory,” the full reference for which is Harvey, D. T.; Byerly, S.; Bowman, A.; and Tomlin, J. *J. Chem. Educ.* **1991**, 68, 162-168. In this paper, retention times for each analyte are reported for three pure mobile phases—20% v/v methanol (M), 16% v/v acetonitrile (A), and 10% v/v tetrahydrofuran (T)—for three 0.50:0.50 binary mixtures (MA, MT, and TA), and for a single 0.33 : 0.33 : 0.33 ternary mixture (MAT). An analyte’s retention time for any other combination of the three pure mobile phases is calculated using the following empirical model

$$t_r = a_M X_M + a_T X_T + a_A X_A + a_{MA} X_M X_A + a_{MT} X_M X_T + a_{AT} X_A X_T + a_{MAT} X_M X_A X_T$$

where the X terms are the volume fractions of the three pure mobile phases and the seven a terms are the model’s coefficients, which are calculated from the seven standard runs.

The isocratic data in Investigation 9 are drawn from the paper “Mobile Phase Effects in Reversed-Phase Chromatography: I. Concomitant Dependence of Retention on Column Temperature and Eluent Composition,” the full reference for which is Melander, W. R.; Bor-Kuan, C.; Horv  th, C. *J. Chromatogr.* **1979**, 185, 99-109. The retention times for the gradient elutions were simulated using the general method outlined in the supplementary materials to the paper “An Advanced, Interactive, High-Performance Liquid Chromatography Simulator and Instructor Resources,” the full reference to which is Boswell, P. G.; Stoll, D. R.; Carr, P. W.; Nagel, M. L.; Vitha, M. F.; Mabbott, G. A. *J. Chem. Educ.* **2013**, 90, 198-202.

Investigation 1

This investigation presents students with a simple chromatogram and guides them in developing a common vocabulary for discussing chromatography and in developing an understanding of why the separation of a mixture is possible.

Question 1. Clicking and dragging over a peak returns the following results:

parameter	peak 1	peak 2	peak 3	peak 4
retention time (sec)	100	200	400	1000
peak height (a.u.)	1	2.82	1.15	0.42
peak width (sec)	NA	56.05	139.29	380.49
peak area (a.u. • sec)	1	100	100	100

Note: the first peak is too narrow to return a value for the peak width.

Question 2. Students should recognize that the first of the four peaks is very different from the other three peaks, particularly with respect to its peak width—which is too narrow to measure—and its peak area. They also should recognize that its retention time of 100 sec is equal to the column’s length (100 mm) divided by the mobile phase’s flow rate (1 mm/sec), which suggests that the solutes that contribute to the peak do not interact with the stationary phase; although the phrase “non-retained solutes” may not occur to them at this time, students should recognize what this term means when it appears in a later investigation and be able to connect it back to their observation here.

Question 3. Although students will not yet have command of specific vocabulary, such as partition coefficients and retention factors, they should recognize that an analyte can move through the column only when it is in the mobile phase. Furthermore, students should also recognize that when an analyte interacts with the stationary phase it is not moving with the mobile phase. Because it is reasonable to assume that analytes interact with the stationary phase in different ways, students should realize that the analytes will

move through the column at different rates.

Question 4. Given that the peaks for the three analytes have identical areas, students should recognize that all three analytes must have the same concentration.

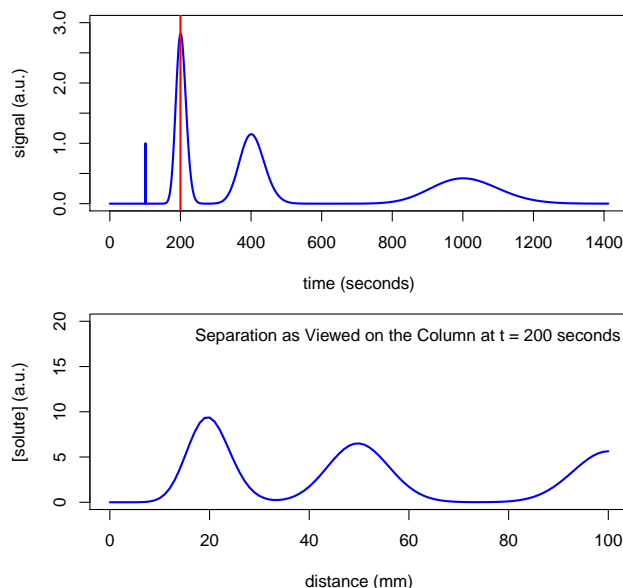
Question 5. Although they will not yet have command of specific terms, such as longitudinal diffusion, students should not be surprised that the longer it takes an analyte to pass through the column the more it will spread out in space. Given equal areas for the peaks, students should recognize that an increase in a peak's breadth requires a decrease in its height.

Investigation 2

The purpose of this investigation is to have students consider how a separation develops by examining how the separation on the column changes with time (as selected by a slider). Using the available radio buttons, students can choose to examine the analytes individually or collectively, and can choose to view the distribution of the analytes between the stationary phase and the mobile phase.

Question 1. Students should recognize that the analytes appear in reverse order in the two views because the analyte with the shortest retention time in the detector view is the analyte that first reaches the end of the column in the column view.

Question 2. Students should note that an analyte's peak becomes broader in width and shorter in height as it passes through a column and recognize that this happens for the same reasons outlined in Question 5 of Investigation 1. Students also will see, more directly, how the analyte's elution from the column generates a signal at the detector, as seen here at $t = 200$ sec.



Question 3. Students should note that the area under the curve for the mobile phase is smaller than the area under the curve for the stationary phase. Students should understand that area is proportional to concentration and predict that Y is more soluble in the stationary phase than it is in the mobile phase.

Question 4. Students should note that the individual concentration profiles for all three analytes are symmetric along the column's axis, which should suggest to them that the movement of an analyte between the mobile phase and the stationary phase is an equilibrium process. If students have a hard time recognizing this as an equilibrium process, it should become clear when they work on Question 1 of Investigation 3.

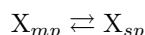
Question 5. Students should see that the difference between the analytes in terms of their respective concentrations in the stationary phase and in the mobile phase is consistent with their retention times; in

particular, analyte X, which is equally soluble in the two phases elutes first, and analyte Z, which is much more soluble in the stationary phase than in the mobile phase, elutes last.

Investigation 3

Following on Investigation 2, the purpose of this investigation is to make more concrete for students the process that controls an analyte's distribution between the mobile phase and the stationary phase, and to add the partition coefficient and the retention factor to their list of an analyte's characteristic chromatographic properties.

Question 1. Although their choice of symbols might differ from those given here, students should recognize that the equilibrium reaction is



where mp signifies the mobile phase and sp signifies the stationary phase. The equilibrium partition coefficient, K_X , is

$$K_X = \frac{[X]_{sp}}{[X]_{mp}}$$

Setting the slider to a distance of 75 mm, for example, shows that

$$[X]_{sp} = [X]_{mp} = 3.209$$

which makes $K_X = 1$; concentrations at other positions along the column give the same value for K_X . For the other two analytes, students should find that $K_Y = 3$, a value consistent with their more qualitative response to Question 3 of Investigation 2, and that $K_Z = 9$

Question 2. Beginning with the equation for f_{mp} , students should recognize that dividing the numerator and the denominator by the analyte's formula weight, FW , converts f_{mp} from a mass ratio to a mole ratio

$$f_{mp} = \frac{(\text{mol X})_{mp}}{(\text{mol X})_{mp} + (\text{mol X})_{sp}}$$

which they can then rewrite as

$$f_{mp} = \frac{[X]_{mp} V_{mp}}{[X]_{mp} V_{mp} + [X]_{sp} V_{sp}}$$

where V_{mp} and V_{sp} are, respectively, the volumes of the mobile phase and the stationary phase. Factoring out $[X]_{mp} V_{mp}$ from the numerator and denominator

$$f_{mp} = \frac{1}{1 + \frac{[X]_{sp} V_{sp}}{[X]_{mp} V_{mp}}}$$

substituting in K_X for the ratio $[X]_{sp}/[X]_{mp}$, and simplifying leaves them with the desired final equation

$$f_{mp} = \frac{1}{1 + K_X \times \frac{V_{sp}}{V_{mp}}} = \frac{1}{1 + k_X}$$

Question 3. From Question 1 of Investigation 1, students know that $t_m = 100$ sec and that $t_{r,X} = 200$ sec; thus,

$$k_X = \frac{t_{r,X} - t_m}{t_m} = \frac{200 \text{ sec} - 100 \text{ sec}}{100 \text{ sec}} = 1.00$$

The retention factors for the other two analytes are $k_Y = 3.00$ and $k_Z = 9.00$. Students should note that when $K = k$, as is the case here, that the volumes of the stationary phase and the volume of the mobile phase must be equal.

Investigation 4

This investigation introduces the concept of a theoretical plate as a means of focusing a student's attention on the processes that cause an analyte's peak to broaden as it moves through the column. It is okay if a student does not appreciate fully the sources of band broadening at the end of this investigation as they will return to this topic in Investigation 7.

Question 1. Using the first peak for analyte X as an example, the characteristic values for the two columns are

parameter	$N = x$	$N = 2x$
retention time (sec)	200	200
peak height (a.u.)	2.82	3.98
peak width (sec)	56.05	39.915
peak area (a.u. • sec)	100	100

Students should note that an increase in the number of theoretical plates leads to a decrease in the width of an analyte's chromatographic peak and an increase in its height; it also shows that the number of theoretical plates does not affect the analyte's retention time or its peak area. Analytes Y and Z show similar results, although the measured retention time for analyte Z is 1000 sec when $N = x$ and 999 sec when $N = 2x$.

Question 2. The simplest way for students to think about how the number of theoretical plates affects peak width is this: as an analyte moves down the column, the number of theoretical plates over which the analyte is distributed is independent of the thickness of the plates; thus, if an analyte is distributed over y plates, those plates will encompass a smaller linear distance on the column and, therefore, a smaller range of elution times at the detector when $N = 2x$ than when $N = x$. Because area is proportional to concentration, which is independent of N , a decrease in peak width must result in an increase in peak height if the peak area is to remain constant. Finally, an analyte's retention time is determined by the equilibrium constant—expressed as the partition coefficient, K , or as the retention factor, k —that describes its relative solubility in the mobile phase and the stationary phase; students should recognize that this equilibrium constant is not a function of N .

Question 3. Students may or may not recognize the contributions of longitudinal diffusion, multiple paths, and mass transfer, and they almost certainly will not use these specific terms at this point. The intention here simply is to have them start thinking, in preparation for Investigation 7, about what happens to individual particles of an analyte as they pass through the column. Nevertheless, students should recognize that the broadening of a chromatographic peak as the analyte moves through the column is not unexpected.

Question 4. For the first chromatogram, we find that

$$N_Z = 16 \left(\frac{1000 \text{ sec}}{380.49 \text{ sec}} \right)^2 = 110$$

and for the second chromatogram, we find that

$$N_Z = 16 \left(\frac{999 \text{ sec}}{266.68 \text{ sec}} \right)^2 = 224$$

As expected, N_Z essentially is $2\times$ greater for the second chromatogram.

Note: the simulation itself uses $N = 100$ and $N = 200$; the difference between the actual and calculated values for N likely is the consequence of using such a small number of plates and the limitations this places on using the binomial distribution to model the separation as a countercurrent extraction when the equation for N assumes a Gaussian distribution.

Investigation 5

With Investigation 5, students turn their attention to the resolution between two chromatographic peaks, with an initial focus on the factors that affect resolution and how we calculate and report resolution; Investigation 6, which follows, takes an even closer look at resolution.

Question 1. Students should recognize that for a constant Δt_r , resolution is independent of the retention times of the individual analytes; thus, for example, when $w_X = 1.0$ and $w_Y = 1.2$, Δt_r is 0.91 when the individual retention times are 8.0 min and 9.0 min, and when they are 7.0 min and 8.0 min. An increase in Δt_r , however, does increase the resolution.

Question 2. Students should recognize that smaller peak widths improve resolution. Students also should recognize that maintaining a constant total width, $w_X + w_Y$, results in a constant resolution; thus, when $\Delta t_r = 1.0$, the resolution is 0.91 when $w_X = 1.0$ and $w_Y = 1.2$, and when $w_X = 1.05$ and $w_Y = 1.15$.

Question 3. With a little experimentation, students should recognize that resolution is independent of the concentrations of the two analytes.

Question 4. From the first two questions, students should recognize that resolution is directly proportional to Δt_r and that it is inversely proportional to the sum of w_X and w_Y ; thus

$$R \propto \frac{\Delta t_r}{w_X + w_Y}$$

By collecting data for several sets of conditions, students should be able to propose one of these final equations

$$R = \frac{2 \times \Delta t_r}{w_X + w_Y} = \frac{\Delta t_r}{0.5(w_X + w_Y)}$$

Question 5. Answers here will vary, but a resolution of 1.5 generally is accepted as baseline resolution. Students should recognize that any combination of retention times and peak widths that yields $R \geq 1.5$ will result in baseline resolution and recognize that concentration is not important.

Question 6. Students should recall that the retention factor, k , is a function of retention time, and that the retention factor is proportional to the partition coefficient, K ; thus, they should recognize that they can increase Δt_r by finding a mobile phase/stationary phase combination that affects k for one analyte more than it affects k for the other analyte.

Question 7. Students should recall that increasing the number of theoretical plates results in smaller peak widths; however, beyond suggesting that they can increase N by using a longer column, students may not yet have the specific language in hand to address this in terms of the other factors that contribute to band broadening; see, for example, the earlier comments in the suggested response to Question 3 of Investigation 4. Students will return to this topic in Investigation 7.

Investigation 6

Continuing from the previous investigation, which introduced the resolution between two peaks in terms of their retention times and their peak widths, students now explore an additional way to express resolution, which defines it in terms of the number of theoretical plates, N , the relative retention factors of the two analytes, α , and the retention factor for the later eluting analyte, k_b . The investigations that follow this one provide a closer look at how these three factors are used to improve resolution.

Questions 1. Students should recognize that the equation's third term, $\frac{k_b}{k_b+1}$, represents the column's general retention of analytes because it accounts for the time needed to elute the last of the two analytes. Students also should recognize that the equation's second term, $\frac{\alpha-1}{\alpha}$, must represent the relative selectivity for the two analytes given that $\alpha = \frac{k_b}{k_a}$. Finally, students are left with the first term, $\frac{\sqrt{N}}{4}$, as being associated with the efficiency with which the analyte moves within the mobile phase and between the mobile phase and the stationary phase.

Question 2. For k_b , students should recognize that changing its value from 3.0 (the default value)

$$\frac{k_b}{k_b + 1} = \frac{3}{3 + 1} = 0.75$$

to 40.0 (the maximum value)

$$\frac{k_b}{k_b + 1} = \frac{40}{40 + 1} = 0.98$$

improves the resolution by a factor of about 30%, which likely is not sufficient to resolve the two analytes.

Using the same approach, for α students should predict that changing its value from 1.1 (the default value) to 2.0 (the maximum value) improves resolution by a factor of roughly 500%, which likely is more than sufficient to resolve the two analytes. For N , students should predict that changing its value from 1000 (the default value) to 6000 (the maximum value) improves resolution by about 250%, which should nearly resolve or fully resolve the two analytes.

By adjusting the sliders, students can confirm all three predictions.

Question 3. Answers here likely will vary, but students should recognize that to increase k_b without increasing α requires that all analytes remain on the column longer, which they can accomplish by decreasing the flow rate or by choosing a stationary phase/mobile phase combination that affects all analytes equally. Students also should recognize that increasing α while maintaining a constant k_b will require finding a stationary phase/mobile phase combination that affects the retention time of just the earlier eluting analyte. Finally, students should recognize that using a longer column will increase the value of N . Although students may recall from Investigation 4 that we can effect an increase in N if we can decrease the height of a theoretical plate, they likely will not have well-formed ideas on what factors control the height of a theoretical plate; this is the subject of Investigation 7. *Note: the calculations for the separation shown here assume that an increase in N is accomplished by increasing the length of the column.*

Question 4. Answers here likely will vary, but students should recognize from a plot of resolution as a function of N that little improvement in resolution is possible for values of $k_b > 10$; students should recognize, as well, that an increase in k_b comes at the expense of a longer analysis time.

For α , students should recognize that increasing its value for one pair of analytes comes at the expense of decreasing its value for another pair of analytes.

For N , students should recognize that improving resolution by using a longer column comes at the expense of a longer analysis time.

Investigation 7

In this investigation—the first of three that consider ways to improve the separation of two analytes—students consider the meaning of efficiency and how it affects resolution.

Question 1. Students should recognize that increasing the length of the column increases the time it takes to elute the analytes, which they can see by returning to Investigation 6 and observing how a change in N affects the retention times for the two analytes.

Question 2. Students should recognize that the existence of different pathways through the column is independent of time and, therefore, independent of the mobile phase's flow rate; thus, they should find it easy to assign this as the A term in the van Deemter equation. The remaining two terms likely will require more thought on the part of students, but with some nudging where needed, students should recognize that diffusion within the mobile phase (longitudinal diffusion) increases the longer the analyte remains in the column; thus, this factor is less important at high flow rates and is, therefore, inversely proportional to u and is assigned as the B term in the van Deemter equation. This leaves, of course, C as the term that accounts for the kinetics of the analyte's movement between the stationary phase and the mobile phase; with some hints, as needed, students should be able to explain that a slow mobile phase flow rate ensures that the partitioning of the analyte between the mobile phase and the stationary phase remains at equilibrium as it passes through the column.

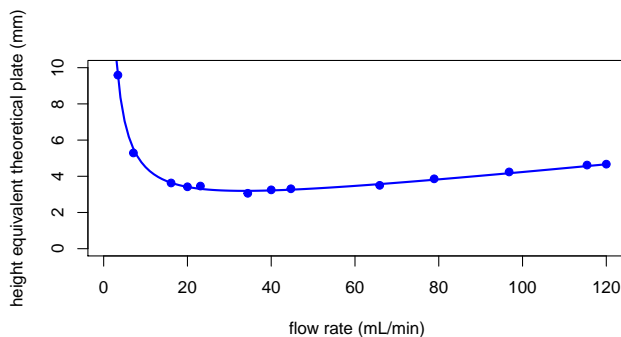
Question 3. For the multiple pathways term (A), students should have little difficulty recognizing that it will depend on particle size as larger particles will allow for a greater variety of path lengths. For the longitudinal diffusion term (B), students should recognize that it will depend on the analyte's diffusion coefficient in the mobile phase. Finally, for the mass transfer term (C), students should recognize that it will depend on the analyte's diffusion coefficients in the mobile phase and in the stationary phase; they also should recognize, perhaps with some gentle guidance, that it depends on the thickness of the stationary phase, the particle size, and the diameter of the column.

Question 4. By taking the derivative of the van Deemter equation with respect to the flow rate and setting it equal to zero

$$\frac{\delta H}{\delta u} = -\frac{B}{u^2} + C = 0$$

students can rearrange and solve for u to obtain $u = \sqrt{\frac{B}{C}}$. Substituting this back into the van Deemter equation and simplifying gives $H = A + 2\sqrt{BC}$.

Question 5. With some playing around, students should be able to find a good fit of the van Deemter equation to the data. Shown here is the result using $A = 1.61$ mm, $B = 26.6$ mm • mL/min, and $C = 0.0236$ mm • min/mL, with a total residual error of 0.12.



Under these conditions, the optimum mobile phase flow rate is

$$u = \sqrt{\frac{B}{C}} = \sqrt{\frac{26.6 \text{ mm} \cdot \text{ml/min}}{0.0236 \text{ mm} \cdot \text{min/mL}}} = 33.6 \text{ ml/min}$$

and the corresponding height of a theoretical plate is

$$H = A + 2\sqrt{BC} = 1.61 \text{ mm} + 2\sqrt{(26.6 \text{ mm} \cdot \text{ml/min}) \times (0.0236 \text{ mm} \cdot \text{min/mL})} = 3.19 \text{ mm/plate}$$

The number of theoretical plates, therefore, is

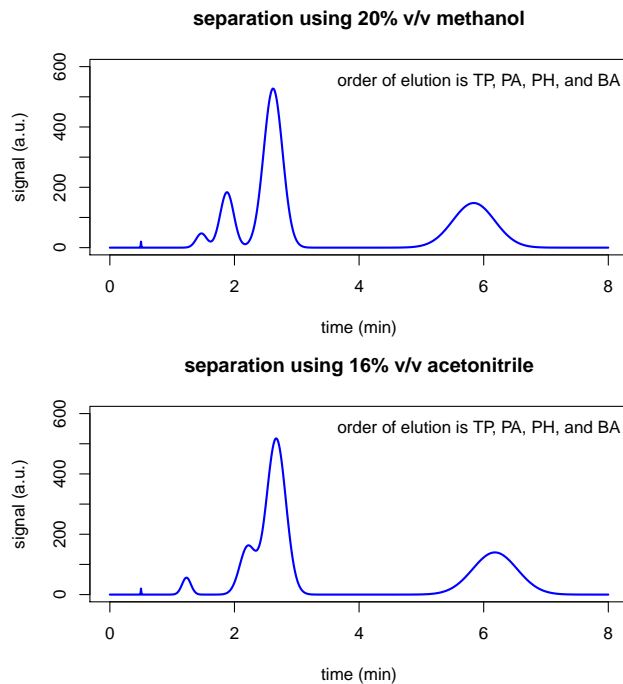
$$N = \frac{L}{H} = \frac{5000 \text{ mm}}{3.19 \text{ mm/plate}} = 1570 \text{ plates}$$

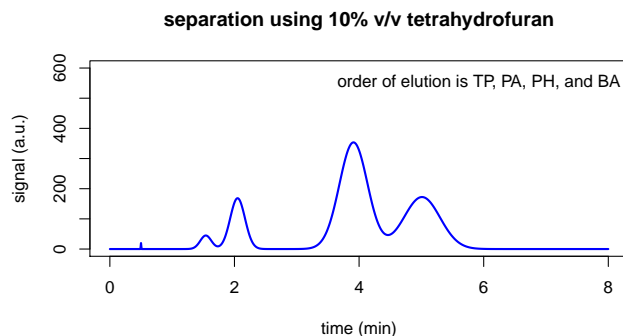
Investigation 8

Following on a consideration of how to improve resolution by increasing efficiency, in this investigation students explore how to improve resolution by controlling selectivity.

Question 1. Because the analytes differ in their relative solubilities in the three different mobile phases, students can reasonably expect that changing the composition of the mobile phase will affect each analyte's retention time.

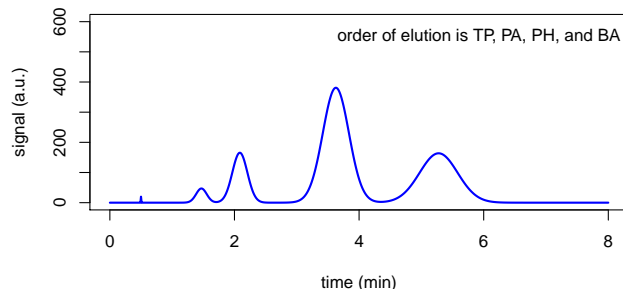
Question 2. The three chromatograms using the pure mobile phases are shown here; note that the order of elution is the same for all three chromatograms.



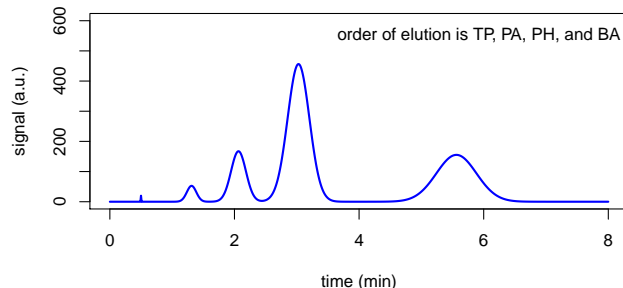


Students should recognize that the retention time for *p*-hydroxybenzoic acid shows the greatest sensitivity to the mobile phase's composition, particularly when using THF instead of MeOH or ACN. Benzoic acid shows the second greatest sensitivity to the mobile phase's composition. Terephthalic acid and *p*-aminobenzoic acid are much less sensitive to the mobile phase's composition.

Question 3. There are many possible mobile phase combinations that will work, one of which, as shown here, is 33% THF and 67% ACN.



In general, any binary mixtures of THF and ACN from about 60% ACN to 85% ACN will work; in the middle of this range, we can replace some THF with methanol; for example, a separation using a mobile phase of 10% MeOH, 75% ACN, and 15% THF is shown here

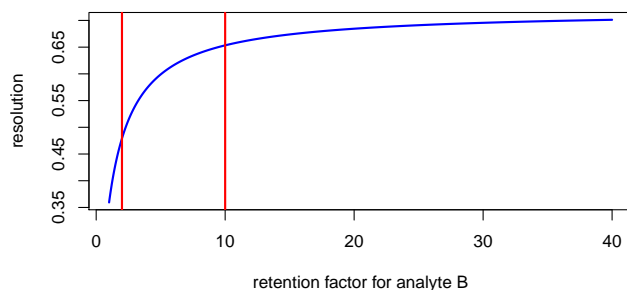


Question 4. As noted above, there are a wide range of conditions that produce a good separation of these analytes; thus the separation is relatively insensitive to changes in the composition of the mobile phase. Because we may not be able to control the composition of the mobile phase precisely, we want to use conditions where small variations in the composition of the mobile phase do not affect the quality of the separation.

Investigation 9

In this last investigation, students explore the control of overall retention as a means for improving resolution.

Question 1. The plot below shows the effect of k_b on resolution when α is 1.1 and when N is 1000 with the red lines set at $k_b = 2$ and at $k_b = 10$. Students should have no difficulty recognizing that the rate of improvement in resolution decreases quickly for $k_b \geq 10$.

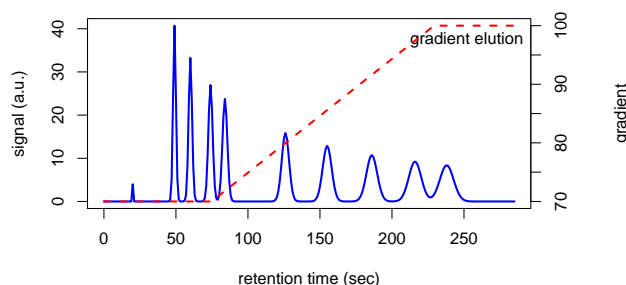


Question 2. Students should recall that $t_r = kt_m + t_m$, which gives a retention time of 60 sec for an analyte with a k_b of 2 and a retention time of 220 sec for an analyte with a k_b of 10.

When using 100% ACN, the last analyte in the mixture elutes with a retention time of approximately 45 sec and when using 70% ACN, the last analyte elutes with a retention time of approximately 430 sec.

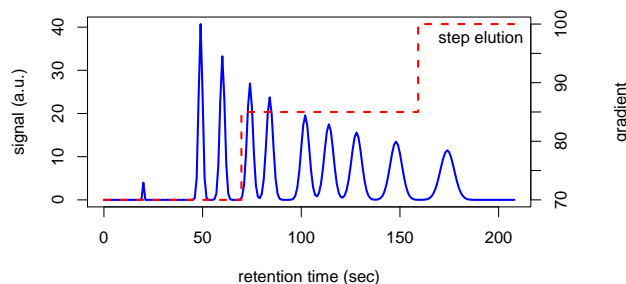
Students should recognize that the general elution problem is this: a condition that favors a short analysis time, such as using 70% ACN for the mobile phase, may resolve late eluting analytes but at the cost of failing to resolve early eluting solvents. Choosing a condition that will resolve the early eluting analytes comes at the cost of a substantial increase in analysis time.

Question 3. A baseline separation that has the last analyte eluting within 220 sec is not possible with a linear gradient, although a nearly baseline separation is possible with the last analyte eluting at approximately 240 sec with the following gradient: 70% ACN from $t = 0$ sec to $t = 75$ sec; a linear ramp to 100% ACN until $t = 230$ sec; hold at 100% ACN until the analysis is complete.



Estimating retention times from the chromatogram gives $k = \frac{t_r - t_m}{t_m} = \frac{50 - 20}{20} = 1.5$ for the first peak and of $k = \frac{240 - 20}{20} = 11$ for the last peak.

A step gradient from 70% to 85% ACN at $t = 70$ sec, followed by a second step gradient from 85% ACN to 100% ACN at $t = 160$ sec provides nearly baseline resolution of all analytes within a total run time of under 200 sec.



Estimating retention times from the chromatogram gives $k = \frac{t_r - t_m}{t_m} = \frac{50 - 20}{20} = 1.5$ for the first peak and of $k = \frac{175 - 20}{20} = 7.75$ for the last peak.