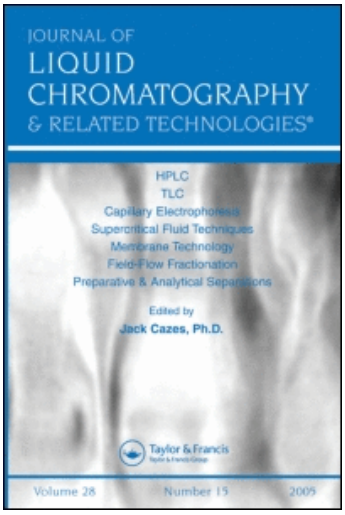

Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Mathematics of Optimization and Scaling for the Practicing Chromatographer

Mark A. Stone[a]
[a] Dow Pharmaceutical Sciences Inc., Petaluma, California, USA

To cite this Article Stone, Mark A.(2007) 'Mathematics of Optimization and Scaling for the Practicing Chromatographer', Journal of Liquid Chromatography & Related Technologies, 30: 5, 605 — 647
To link to this Article: DOI: 10.1080/15376490701190891
URL: http://dx.doi.org/10.1080/15376490701190891

Journal of Liquid Chromatography & Related Technologies®, 30: 605–647, 2007

ISSN 1082-6076 print/1520-572X online
DOI: 10.1080/15376490701190891

Mathematics of Optimization and Scaling for the Practicing Chromatographer

Mark A. Stone
Dow Pharmaceutical Sciences Inc., Petaluma, California, USA

Abstract: A substantial body of literature exists on the issues of optimization and scaling in HPLC. However, much of this literature is written in a theoretical manner such that it is difficult for the practicing chromatographer to apply the conclusions to problems in the laboratory. In addition, the analyst would have to search numerous publications to find all of the formulas that are needed when optimizing or scaling a method. In this review, an attempt is made to address these topics in a manner that is easy to follow, with all equations given in the form that is most convenient for the laboratory chemist, and with all necessary formulas available in a single source. The focus of the review is on the mathematical concepts relevant to optimization and scaling. Issues that are generally not approached mathematically, such as the selection of the mobile phase and stationary phase, are not discussed.

Keywords: Optimization, Scaling, van Deemter effects, Extra-column effects, Relative retention times, Linear gradients, Dwell volume, Isocratic segments, Method transfer

Part I: Basic Optimization and Scaling Relationships

Part II: Gradient Scaling: Keeping Relative Retention Times Constant with Gradient Methods

Part III: Applying the Basic Relationships to Gradient Methods

INTRODUCTION

There are several factors that govern the selection or modification of the linear velocity, the column length, the column diameter, and the diameter of the

Address correspondence to Mark A. Stone, Dow Pharmaceutical Sciences Inc., Petaluma, CA 94954, USA. E-mail: mstone@dowpharmsci.com

stationary phase particle. These factors include optimizing the efficiency and resolution of the method, increasing the speed of analysis, increasing sensitivity, or reducing the amount of mobile phase consumed and waste generated. It is then important to consider what scaling may be needed, in order to minimize band broadening due to extra-column effects. The purpose of this article is to present the relevant issues in a straightforward manner, with all relationships ultimately given in terms of linear velocity or flow rate, column length, column diameter, and particle diameter, such that they can be easily applied in the laboratory.

In Part II, procedures will be discussed for keeping relative retention times constant with gradient methods. This is important once a gradient separation has been developed to the point where the analyst wants to maintain the relative retention times of the peaks constant while any additional optimization steps are carried out. These considerations will also be useful when optimizing a method where peak identification based on relative retention times has already been established. In order to keep the relative retention times constant, it will be necessary to make adjustments with respect to the gradient ramps as well as the system dwell volume and any isocratic hold segments in the method.

In Part III, we discuss the application of the basic relationships (addressed in Part I) to gradient methods. Because of the more complex interdependency of variables, gradient methods can be more difficult to deal with. However, it will be shown that, when the relative retention times are kept constant (as per the procedures discussed in Part II), the average gradient capacity factor, the gradient compression factor, and the optimal linear velocity will also remain constant and, hence, the situation is considerably simplified.

The discussion is somewhat lengthy, with detailed information provided for the interested reader. However, optimization and scaling are ultimately a matter of utilizing a handful of simple formulas. An effort was made to include all critical relationships, as well as guidelines regarding their application, in the Summary section for easy reference.

PART I: BASIC OPTIMIZATION AND SCALING RELATIONSHIPS

Preliminaries: Plate Height and Pressure Drop[1]

A column plate height term (H) will appear in most of the formulas discussed in this article; therefore, a brief discussion is in order. The value of the column plate height is given by the van Deemter equation, which measures the

[1]This section assumes that the optimal linear velocity of a packed column does not change as a function of column diameter, as predicted by the van Deemter equation. This assumption is valid as long as the ratio of column diameter to particle diameter is greater than a value of about 30: which is generally true for columns with diameters > 0.2 mm.

contribution of various kinetic processes to broadening of the chromatographic peaks. Equation (1) gives a form of the van Deemter equation utilized by Scott;[1] where λ is a packing factor; d_p is the particle diameter; γ is an obstruction factor; D_S and D_M are the diffusivities of the analyte in the stationary and mobile phases, respectively; d_f is the thickness of the stationary phase film; u is the linear velocity of the mobile phase; and $f_1(k)$ and $f_2(k)$ represent expressions that are functions of the capacity factor (k).

$$H = 2\lambda d_p + 2\gamma D_M/u + [f_1(k)d_p^2/D_M]u + [f_2(k)d_f^2/D_S]u \qquad (1)$$

It may be noted that the second term of the van Deemter equation is inversely proportional to linear velocity, whereas the third and fourth terms are directly proportional to linear velocity. This leaves us with the result, well known to chromatographers, that the sharpest peaks will be obtained within a certain linear velocity range, and that decreased performance will be obtained if the linear velocity is either too far above or below this range. Therefore, in many cases, we will want to adjust the flow rate so that we are working at, or at least near, the optimal linear velocity (though this is not always the case).

Figure 1 shows van Deemter plots for four different particle diameters. The dots indicate the minimum plate height and optimal linear velocity for each curve. The optimal linear velocities are also summarized in Table 1, along with the corresponding optimal flow rates for four different column diameters.

The complexity of the van Deemter equation makes it difficult to establish simple rules for optimization and scaling. Therefore, the following two simplifications will be very useful throughout the discussion. These relationships

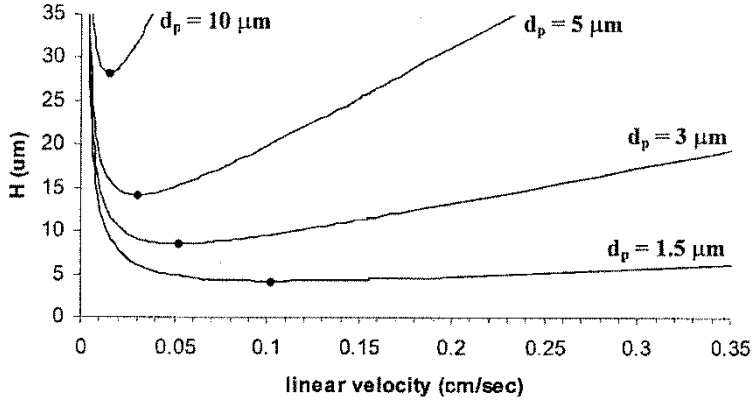

Figure 1. Van Deemter Plots. Calculated for an analyte with a molecular weight of 150 g/mole, diffusivity in the mobile phase = 7.5×10^{-6} cm^2/sec, diffusivity in the stationary phase = 3.0×10^{-6} cm^2/sec, k = 5, stationary phase film thickness = 20Å, analysis temperature = 25°C. Calculations were based on a 100% aqueous mobile phase. Linear velocities would become somewhat higher if acetonitrile were added to the mobile phase and somewhat lower if methanol were added.

Table 1. Optimal linear velocity and flow rate

Particle diameter (μm)	Optimal linear velocity (cm/sec)	Optimal flow rate (mL/min)			
		4.6 mm ID	3.2 mm ID	2 mm ID	1 mm ID
10	0.015	0.105	0.051	0.020	0.005
5	0.031	0.216	0.105	0.041	0.010
3	0.052	0.363	0.176	0.069	0.017
1.5	0.103	0.719	0.348	0.136	0.034

Based on the same conditions summarized under Figure 1. Flow rates were calculated using Equation (4) with a total porosity (ε) of 0.7, which is generally considered appropriate for porous particles. The shaded areas are not usable because the excessive pressures that result from such small particles lead to substantial viscous heat generation. In these cases, smaller diameter columns must be used.

were derived by minimization of the van Deemter equation.[2] Relationship (2) serves as a good approximation when we are operating near the optimal linear velocity, and relationship (3) is a good approximation when we are near the minimum plate height.

$$u_{opt} \propto 1/d_p \tag{2}$$

$$H_{opt} \propto d_p \tag{3}$$

The parameter over which the analyst has direct control is the mobile phase flow rate. Therefore, it is important to understand the relationship between the flow rate and the linear velocity. In this way, the flow rate can be set so as to obtain the desired linear velocity. Equation (4) gives the relationship between the flow rate (F) and linear velocity (u) of the mobile phase and the inner diameter of the column (d_c). The parameter ε is the total porosity, which is generally taken as 0.7 for porous packings and 0.4 for non-porous packings. The multiplier in the equation, simply serves to put all parameters into convenient units.

$$F(mL/min) = 0.15\pi\varepsilon u(cm/sec)\, d_c(mm)^2 \tag{4}$$

If we insert Equation (2) into Equation (4), and drop the constant terms, we obtain the following general relationship for scaling flow rate. In cases where the linear velocity of the original method is in the optimal range, and the analyst wants to stay in the optimal range when changes are made, Equation (5) indicates how the flow rate would need to be changed with respect to changes in the column diameter or the diameter of the stationary phase particles.

$$F_{opt} \propto d_c^2/d_p \tag{5}$$

As we will see in the following section, there are times when the analyst may choose to operate above the optimal linear velocity, such as when trying

to optimize the speed of analysis or when we want to increase sensitivity when using a mass-sensitive detector. This is especially true when working with small diameter particles, as they have fairly shallow van Deemter curves above the optimal region (see Figure 1), and, therefore, one can have the benefits of a higher velocity while observing only a small increase in the plate height.

From the discussion above, it is clear that there can be substantial benefits to working with smaller diameter particles. In fact, when trying to optimize an analytical method, reducing the diameter of the stationary phase particles will often be the preferred approach, as it simultaneously accomplishes several beneficial effects (increased efficiency and resolution, increased speed, increased sensitivity). In some cases, it will be beneficial to use higher linear velocities or longer columns. All of these issues will be addressed in the next section. For the present, it is important to realize that our ability to make these changes will be limited by the maximum pressure tolerance of the system. Therefore, it is necessary to consider the pressure drop across the column. The Kozeny-Carmen equation demonstrates how the pressure drop across the column is affected by the linear velocity (u) or the flow rate (F) of the mobile phase, the mobile phase viscosity (η), the column length (L), the column diameter (d_c), and the diameter of the stationary phase particles (d_p). Equations (6) and (7) present forms of the pressure drop equation in terms of the flow rate or linear velocity, respectively. These equations are based on a flow resistance parameter of 1,000 (contained within the multipliers), which is generally appropriate for liquid chromatography: for both porous and non-porous particles.[3] An effort was made to put all parameters in their most commonly used units, with pressure given in psi. Viscosity values for various pure solvents and solvent mixtures can be found in Appendix II of Reference [4].

$$\Delta P(\text{psi}) = \frac{30778F(\text{mL/min})\eta(\text{cP})L(\text{cm})}{d_c(\text{mm})^2 d_p(\mu\text{m})^2} \tag{6}$$

$$\Delta P(\text{psi}) = \frac{14504\,\varepsilon\, u(\text{cm/sec})\eta(\text{cP})L(\text{cm})}{d_p(\mu\text{m})^2} \tag{7}$$

If the constant terms are dropped, we can write the following simple proportionalities:

$$\Delta P \;\alpha\; FL/d_c^2 d_p^2 \tag{8}$$

and

$$\Delta P \;\alpha\; uL/d_p^2 \tag{9}$$

If we consider the combined effects of Equations (5) and (6), it becomes clear that, if we change the particle diameter and, at the same time, modify the

flow rate so as to stay at the optimal linear velocity, the pressure will change inversely with the particle diameter cubed: a very strong dependence!

Optimizing the Method

The topic of optimization in HPLC has been discussed in several previous publications.[2,5–15] These articles offer a very thorough coverage of the topic, and provide the foundation for our basic understanding. However, these treatments are somewhat advanced, such that it is difficult for the chromatographer to readily apply the conclusions to problems in the laboratory. The present discussion attempts to review these issues in a way that is convenient for the laboratory chemist, and with all final relationships expressed in terms of the linear velocity or flow rate, the diameter and length of the column, and the diameter of the stationary phase particles. These formulas can be used to guide the initial development process or to optimize an existing method.

Efficiency and Resolution

Efficiency (N) carries information about two of the factors that affect the quality of a separation. Specifically, these are the kinetic factors that contribute to band broadening (i.e., the plate height, H) and the contribution of column length (L) to the separation.

$$N = L/H \tag{10}$$

Resolution (R) can be defined by Equation (11).[16] The resolution equation contains all four of the factors that affect the quality of a separation, including the effect of plate height and column length, as well as the extent to which the analytes are retained (represented by the "k-term"), and the selectivity of the system (represented by the "α-term"). The parameter k_2 is the capacity factor of the second peak in a given pair, and α is the separation factor which is defined as the ratio of the two capacity factors (k_2/k_1). Because it encompasses all the relevant parameters, the resolution is a more relevant measure of the separation power that will be observed for a given application.

$$R = \left(\frac{\sqrt{L}}{4\sqrt{H}}\right)\left(\frac{k_2}{1+k_2}\right)\left(\frac{\alpha-1}{\alpha}\right) \tag{11}$$

From these relationships, it is obvious that, to maximize efficiency and resolution, we want to increase the length of the column and/or minimize the plate height. From the previous section, we know that minimizing the plate height is accomplished by reducing the diameter of the particles and working at the optimal linear velocity.

We can combine Equations (10) and (11) with Equation (3) to obtain expressions in terms of the particle diameter. And, since the present focus is not on the mobile and stationary phases, we can treat the k-term and the α-term as constants. The result of these manipulations is:

$$N \alpha L/d_p \tag{12}$$

$$R \alpha (L/d_p)^{1/2} \tag{13}$$

Reducing the plate height (by reducing the particle diameter) is the preferred method for increasing efficiency and resolution. However, the pressure drop across the column will limit the extent to which we can do this. In cases where the pressure is limiting, the remaining option is to increase both column length and particle diameter, with the former increased by an amount somewhat greater than the latter. This will simultaneously cause efficiency and resolution to increase and the pressure drop to decrease. Therefore, substantial increases in efficiency and resolution will be possible, along with costs related to the longer column (reduced speed, reduced sensitivity, increased solvent consumption) and the larger particle diameter (further reduction in sensitivity).

The relationship of efficiency and resolution to particle diameter that is implied by relationships (12) and (13) becomes less exact in cases where we do not operate at (or near) the minimum plate height, as relationship (3) is only a good approximation in this region.

Relationships can also be derived that give the maximum efficiency and resolution that are obtainable. If we rearrange Equation (7) so as to calculate the maximum column length that corresponds to a given pressure maximum and linear velocity, we obtain:

$$L_{max}(cm) = \frac{\Delta P_{max}(\,psi) d_p(\mu m)^2}{14504 \varepsilon u(cm/\,sec)\eta(cP)} \tag{14}$$

If we then plug this into Equation (10) we get:

$$N_{max} = \frac{0.689 \Delta P_{max}(\,psi) d_p(\mu m)^2}{\varepsilon u(cm/\,sec)\eta(cP) H(\mu m)} \tag{15}$$

And, plugging into Equation (11) we get:

$$R_{max} = \left(\frac{0.0431 \Delta P_{max}(\,psi) d_p(\mu m)^2}{\varepsilon u(cm/\,sec)\eta(cP) H(\mu m)}\right)^{1/2} \left(\frac{k_2}{1+k_2}\right)\left(\frac{\alpha - 1}{\alpha}\right) \tag{16}$$

Speed of Analysis

The simplest way to develop expressions for the speed of analysis is to begin with the basic relationship that

$$t_r = L(1 + k)/u$$

If we apply this relationship to a peak (real or hypothetical) that elutes at the end of the run, whose capacity factor we will call k', we can write the following:

$$\text{Analysis time} = L(1 + k')/u \tag{17}$$

We can divide Equation (10) by Equation (17) to obtain an equation for plates per unit time. This equation was first presented in the early 1980's.[17,18]

$$N/\text{time} = u/H(1 + k') \tag{18}$$

Similarly, we can divide Equation (11) by Equation (17) to obtain a relationship for resolution per unit time (recall that k_2 is the capacity factor corresponding to the second peak in a given peak pair and k' is as defined above). To the author's knowledge, this relationship has not been presented before, though it is arguably one of the most important attributes to the practicing chromatographer.

$$R/\text{time} = \left(\frac{u}{4(LH)^{1/2}(1 + k')}\right)\left(\frac{k_2}{1 + k_2}\right)\left(\frac{\alpha - 1}{\alpha}\right) \tag{19}$$

If we drop the constant terms from Equations (17), (18), and (19), and substitute d_p in place of H, we can write the following proportional relationships:

$$\text{Analysis time} = L/u \tag{20}$$

$$N/\text{time} \ \alpha \ u/d_p \tag{21}$$

$$R/\text{time} \ \alpha \ \frac{u}{(Ld_p)^{1/2}} \tag{22}$$

Notice that Equation (19) suggests that, in order to obtain the maximum resolution per unit time, we want a relatively high value of k_2, and a low value of k'. Clearly, it is not possible to accomplish both of these simultaneously. The implication is that the optimal speeds will be obtained with capacity factors in the intermediate range (perhaps values around 5). Furthermore, this suggests that increasing or decreasing the capacity factors will generally not be a particularly effective way to change the speed of a separation. Reducing the particle diameter is a beneficial approach, as it will increase the speed of analysis and total resolution simultaneously. This approach should be evaluated first, especially if the original method utilizes large diameter particles. Increasing the linear velocity will carry a tradeoff to the extent that it moves us away from the optimal part of the van Deemter equation (i.e., will cause H to increase in relationships (18) and (19)). However, for small diameter particles, this tradeoff is not significant, and increasing the linear velocity is, in fact, a very effective option for increasing speed in these cases. Here again, the pressure tolerance of the system will

limit our ability to effect these changes. Reducing column length will increase the resolution per unit time, while reducing the pressure. However, there will be a tradeoff in terms of the absolute efficiency and resolution.

In cases where we have just the necessary degree of resolution but want to increase the speed of analysis, the simplest approach is to reduce both the column length and the particle diameter by the same factor (x). In this case, the efficiency and resolution will remain constant (see relationships (12) and (13)) as has been noted by others.[3,9] Assuming we also increase the flow rate so as to stay in the optimal linear velocity range (as per relationship (5)), the resolution per unit time will increase by x^2. However, the pressure will also increase by x^2, thereby limiting the extent to which we can make these modifications.

Sensitivity

The sensitivity of an analytical method is a function of the signal-to-noise ratio, where the signal is generally defined as the height of a chromatographic peak. The following formula gives peak height (h) as a function of the area (A) and retention time (t_r) of a given peak, and the plate height (H) and length (L) of the column.[4]

$$h = \frac{A\,L^{1/2}}{t_r(2\pi H)^{1/2}} \tag{23}$$

We can substitute $L(1 + k)/u$ in place of t_r, which results in:

$$h = \frac{A\,u}{(1 + k)(2\pi L\,H)^{1/2}} \tag{24}$$

Here again, we can replace H with d_p, as per relationship (3), and drop the constant terms. For a concentration-sensitive detector, the product of the peak area and linear velocity will be constant. Thus, when we drop the constant terms we will also drop Au. In addition, since noise is either constant or will be accounted for (see Table 2 and its footnote), we can express the relationship as the signal-to-noise ratio (S/N). Hence, we will replace h with S/N. Relationship (25) will be the starting point, from which we will derive the working relationships for concentration-sensitive detectors.

$$S/N\ \alpha\ 1/(Ld_p)^{1/2} \tag{25}$$

At this point, there are a few additional issues that must be accounted for. First, we need to address the question of how the peak height would be affected by changes in the diameter of the column. This, in turn, depends on whether or not we have a limited volume of sample. There is also the effect of scaling the detection cell volume.

In the section on extra-column effects, the issue of scaling the injection volume in relation to the dimensions of the column will be addressed. It

Table 2. Information for absorbance and fluorescence detectors[a,b]

Type of detector	Transmitted energy is proportional to	Signal-to-noise ratio is proportional to
Absorbance		
Conventional cell	c^2/b^2	c
Light-guiding cell	c	$c^{1/2}$ b
Fluorescence		
Broadband light source	c	Roughly proportional to cell volume, cb. Dependent on detector design.
Laser light source	Not dependent on flow cell geometry	b

[a]Table 2 and the qualifying statements below were provided by Anthony Gilby, Waters Corporation, Milford, MA.[20] The signal-to-noise ratio is also affected by the radiant energy of the lamp. However, as the present focus is on optimization and scaling, this will be considered constant.

[b]a) "Transmitted energy" as used here is the raw detector signal generated by the photo-detector or by the individual pixels in the case of a PDA. Transmitted energy depends on flowcell parameters as shown in Table 2. It also depends on lamp output, photo-detector response and grating efficiency, all of which vary strongly with wavelength. UV absorbance detectors are generally designed to maximize transmitted energy in the region of most analytical interest, 200 to 280 nm.

b) S/N is made up of two components, both of which depend on flow cell geometry. The signal is the peak height measured in absorbance units (AU), which depends on pathlength as per Beer's Law. In the shot-noise limited case considered here, the peak-to-peak noise on a section of the detector's baseline, also measured in AU, depends on the energy transmitted through the cell. The peak-to-peak noise is reduced as the transmitted energy increases.

c) Although shot-noise limited performance is ideal, other sources of noise can often be larger than the shot noise, in which case these noise sources dictate the performance of the detector. The two following situations are commonly encountered:

Depending on the design of the detector and the application, noise from the detector electronics may be larger than the inherent shot noise. In this case, changes in c or b which reduce transmitted energy will result in a larger reduction in S/N than indicated in the Table above. This commonly occurs when transmitted energy is already low, for example through choice of wavelength or when the mobile phase absorbs.

If the detector's baseline noise is dominated by noise from the separation system fluidics (pump pulsations, imperfect gradient formation, reagent addition), changes to cell geometry which affect the transmitted energy may, surprisingly, have little or no effect on the observed baseline noise. An unstable light source can have a similar effect. When present, this type of noise sets a lower limit to the peak-to-peak absorbance noise. This behavior is quite common when transmitted energy is high (see note 1 above).

(*continued*)

will be shown that, when the volume of sample is not limited, the injection volume should be scaled in relation to $d_c^2 (Ld_p)^{1/2}$ (relationship (33)). When the injection volume is scaled in proportion to the square of the column diameter, the mass of analyte placed on column will change by the same factor as the volume of mobile phase in which it is being diluted. Therefore, the concentration of analyte and, hence, the peak height from a concentration-sensitive detector will not change. However, when we scale the injection volume in proportion to $(Ld_p)^{1/2}$, this will increase or decrease the mass of analyte placed on column without any accompanying change in dilution. To account for this, we need to add an $(Ld_p)^{1/2}$ term to the numerator of relationship (25). When dealing with a spectroscopic detector, we must also consider the issue of scaling the detection cell volume. This will also be addressed subsequently (see relationship (35)). However, the effect of this on the signal-to-noise ratio depends on how the pathlength (b) and cross section (c) of the detection cell will be modified, as well as on the type of detector being used. The right-most column of Table 2 provides multipliers that show how the signal-to-noise ratio depends specifically on the dimensions of the detection cell for various types of spectroscopic detection systems. Thus, we add an M term to the equation to represent the appropriate multiplier. After simplification, the result is the following:

$$S/N \; \alpha \; M \tag{26}$$

The values in the table are based on the assumption that the signal-to-noise ratio is shot-noise limited, which is generally the case for small chromatographic peaks in a well designed and maintained detector and fluidic system. The relationship between the signal-to noise ratio and the dimensions of the detection cell is a very complex topic (which is well beyond the scope of this paper). The values in Table 2 are offered as a first approximation that applies in idealized cases.

Notes to Table 2 (Continued)

d) The entry for S/N for a conventional absorbance detector in Table 2 will seem at odds with intuition: that reducing the pathlength leads to a reduction in S/N as well as a reduction in peak height; and vice versa. The S/N does change with pathlength when proportional noise is significant (see note c), or when a change in pathlength does not result in a change in transmitted energy to the degree expressed by the c^2/b^2 entry in Table 2. This can happen when the detector's optical design is optimized for one particular pathlength only.

e) In the case of Laser Induced Fluorescence, the noise level will typically be set by the stability of the laser source.

f) Detector and flow cell design varies from one manufacturer to another, and these rules should only be used as a rough guide to predict the effect of changing flow cell parameters in a particular detector. A more detailed discussion is beyond the scope of this paper. The reader is referred to Reference [20].

If we were dealing with the same detector, but in a case where the injection volume will not be scaled (most likely because there is only a small limited volume of sample available), we would have to account for the fact that the concentrations of the analytes would change when the diameter of the column is changed. This is because the amount of mass placed on column is not changing, but the volume in which we are diluting it changes in proportion to the square of the column diameter. To account for this, we would need to add a d_c^2 term to the denominator of relationship (25). The issue on the detection end is the same as that discussed above; hence, the multiplier (M) is again added. The resulting relationship is:

$$S/N \; \alpha \; M/d_c^2(Ld_p)^{1/2} \tag{27}$$

In cases where the detection cell will not be modified, or when dealing with a concentration-sensitive detector that does not utilize a detection cell (for example a conductivity detector), the above discussion is equally relevant, with the exception of the issues regarding the detection cell. Hence, in these cases, we would use relationships (26) and (27) without the multiplier (M).

Mass-sensitive detectors respond to the mass flux of analyte passing through the detector per unit time. Hence, the chromatographic peak area will be constant and the peak height will increase or decrease in direct proportion to the linear velocity of the mobile phase as it enters the detector.[21,22] Therefore, when we drop the constant terms from Equation (24), we will drop A but not u and, hence, the starting point for mass-sensitive detectors will be:

$$S/N \; \alpha \; u/(Ld_p)^{1/2} \tag{28}$$

In cases where the volume of sample is not limited, we would scale the injection volume in proportion to $d_c^2 \; (Ld_p)^{1/2}$ as per relationship (33). So, the mass placed on-column and, therefore, the peak height from a mass-sensitive detector would also change in proportion to this quantity. In order to account for this, we would add $d_c^2 \; (Ld_p)^{1/2}$ to the numerator of relationship (28). Simplifying, we obtain the following:

$$S/N \; \alpha \; u \, d_c^2 \tag{29}$$

If working with a mass-sensitive detector in cases where the injection volume will not be adjusted (again, most likely because there is only a small limited volume of sample available), the mass placed on-column would be constant. For this case relationship (28) would be the appropriate expression.

It should be noted that the conclusions and relationships discussed here apply regardless of whether we are using peak height or peak area quantitation. In either case, narrower peaks will give better signal-to-noise ratios and, hence, improved sensitivity.

Solvent Consumption and Waste Generation

The amount of solvent consumed or waste generated is simply given by the product of the analysis time and the flow rate. A more rigorous way to express this (which may be more useful for optimization purposes) is by the product of Equation (17), and the flow rate expressed as $u\ d_c^2\ \pi\ \varepsilon/4$ (see Equation (4); the difference is that, here, we are dividing by four instead of multiplying by 0.15. The reason for this is that we have eliminated the factor that accomplishes the unit conversions, as this is not useful in this context). When we perform the multiplication, we obtain the following relationship:

$$\text{Solvent Consumed/Waste Generated} = \pi\,\varepsilon(1 + k')\,d_c^2\,L/4 \qquad (30)$$

If the constant terms are dropped, we obtain the following relationship:

$$\text{Solvent Consumed/Waste Generated}\ \alpha\ d_c^2\,L \qquad (31)$$

Therefore, we can reduce solvent consumption and waste by reducing the diameter and/or the length of the column. Although reductions in the column diameter will have a more substantial effect, it will be seen that several of the extra-column effects are more sensitive to reductions in column diameter than to reductions in column length. Therefore, some combination of the two may be preferrable.

Scaling to Minimize Extra-Column Effects

Once decisions are made about how the method will be optimized, it is necessary to consider what scaling may be necessary in order to minimize band broadening due to extra-column effects. In what follows, formulas are discussed that dictate how various attributes should be scaled in response to changes in the length or diameter of the column, the diameter of the stationary phase particles, or the linear velocity of the mobile phase. The basis for this discussion will be a paper published in 1975 by Martin, Eon, and Guiochon, which presented equations for a variety of extra-column effects.[5] However, these issues have also been addressed by DiCesare, Dong, and Atwood,[18] Sternberg,[23] Hartwick and Dezaro,[24] Meyer,[25] Chervet, Ursem, and Salzmann,[26] and various others that will be mentioned throughout this section.

The issues discussed in this section only need to be considered when scaling down, i.e., reducing the dimensions of the column, as that is when there is the potential for the extra-column effects to worsen the separation.

Injection Volume

The contribution of the injection process to the observed width of chromatographic peaks depends on both the quality and the volume of the injection. A "perfect" injection is defined as when the injected volume is delivered to the mobile phase stream as a narrow plug, and does not mix with the mobile phase. Martin, Eon, and Guiochon,[5] building on the work of Karger, Martin, and Guichon,[27] derived a formula which determines the maximum volume that can be injected (V_i) for a given degree of band broadening, represented by θ. Specifically, θ^2 is the fraction of peak broadening that is observed. Other parameters in Equation (32) are the inner diameter (d_c), length (L), and plate height (H) of the column; the capacity factor of a given analyte (k); a parameter (K) which is a measure of the quality of the injection; and the total porosity of the column (ε) (defined previously). The K factor is equal to $\sqrt{12}$ for a perfect injection, but real world values are generally in the range of 1 to 3.[18,27,28] The equation presented below is modified from Ref. [5] in that L/H was substituted for N in order to make the dependence on column length explicit.

$$V_i = \theta K \pi \varepsilon (1 + k) d_c^2 (LH)^{1/2}/4 \tag{32}$$

If we plug relationship (3) into Equation (32), and drop all unnecessary terms, we obtain the following relationship which shows how the linear velocity should be scaled in relation to the dimensions of the chromatographic column:

$$V_i \propto d_c^2 (L d_p)^{1/2} \tag{33}$$

Given a certain fraction of peak broadening that is determined to be acceptable (θ^2), Equation (32) would tell us what the maximum injection volume would be. This equation also suggests that, if we change the dimensions of the column while keeping the injection volume constant, the degree of band broadening due to the injection process (θ^2) would change. If we were to actually calculate a result for V_i, we would need to know the values of all the parameters in Equation (32), including the capacity factor, which would be different for each analyte. However, for the purpose of scaling, it is generally not necessary to do this calculation. We are only interested in the conclusion given by relationship (33) that, to maintain the band broadening due to the injection process constant, the injection volume would need to be scaled in proportion to the square of the column diameter, the square root of the column length, and the square root of particle diameter.

Relationship (3) will be used in several derivations in the rest of this section. It should be recalled that this relationship is a good approximation only when we are operating near the minimum plate height. It follows that, in cases where we deviate significantly from this region, the relationship between injection volume and particle diameter that is indicated by

relationship (33) would become less exact, the extent of which would depend on how far removed we are. However, the relationship to column diameter and column length would still be accurate. As mentioned previously, the analyst may choose not to use the simplification provided by relationship (3) in cases like this and simply leave the equations in terms of H; and then use a kinetic plot, such as Figure 1, to obtain a good approximation of how H would change in a given situation.

Finally, it should be mentioned that additional band broadening can occur if the sample is dissolved in a solvent that is significantly stronger (chromatographically) than the mobile phase.[29,30] The reason for this is that the effective capacity factor may be reduced during the time that the analytes are being transferred onto the column.

Detector Cell Volume

Equation (34) gives the detector cell volume (V_d) as a function of several parameters, as previously defined. This equation was derived by Martin et al.[5] building on the work of Sternberg.[23]

$$V_d = \theta \pi \varepsilon (1 + k)\, d_c^2 (LH)^{1/2}/4 \tag{34}$$

Combining Equation (34) with relationship (3), and dropping unnecessary terms, we obtain the following scaling relationship:

$$V_d \;\alpha\; d_c^2 (L d_p)^{1/2} \tag{35}$$

The reason for the requirement on the volume of the detector cell is that if the volume of the cell is too large in relation to the volume of the peaks, the mixing which occurs within the cell can cause a significant deterioration of the chromatography. The critical factor is the fraction of the peak which can exist in the detector cell at any point in time. Logically, it is the smaller peaks of interest that are most susceptible to band broadening by this process.

System Dead Volume

System dead volume refers to voids in the fittings (or even the column itself). This issue has been discussed in the literature, both from a practical and theoretical standpoint.[23,31–36] However, these references do not provide a formula describing how dead volume should be scaled relative to the dimensions of the chromatographic column. This is most likely due to the fact that instrument dead volume is not something that can be easily controlled. However, it is for exactly this reason that dead volume can have a very significant effect in chromatography. Therefore, it would be useful to have some quantitative idea what the effect of dead volume would be when the dimensions of the column are reduced.

Dead volume contributes to band broadening by exactly the same mechanism described above for the detector cell. And, for dead volume regions that are downstream of the column, the relationship between the allowable dead volume and the dimensions of the column is the same as that described for the detector cell and, hence, relationship (35) can also be used in these cases. The derivation of Equation (34) (by Martin et al., Ref. [5]) was based on the assumption that the detector cell is a perfect mixer (i.e., the concentration within the cell becomes immediately uniform). This is generally a reasonable assumption for a detector cell; however, when dealing with dead volume, the assumption is somewhat shaky. Relationship (35) should, therefore, be thought of only as an approximation when applied to system dead volume downstream of the column.

With respect to dead volume regions upstream of the column, relationships (34) and (35) are not relevant, as none of these parameters have any effect on the widths of the peaks upstream of the column. The dead volume upstream of the column is actually more important than dead volume on the back end, simply because the volumes of the peaks are smallest at the head of the column. However, as there is no relationship between band broadening due to upstream dead volume and the dimensions of the column, it is not relevant in a discussion of scaling and will not be discussed further.

Detector Time Constant

The detector time constant is defined as the time a detector takes to respond to 63.2% of a sudden change of signal.[21] Simply put, the higher the time constant the slower the response time and, therefore, less readings are obtained across a peak. An insufficient number of readings across a peak may adversely affect efficiency, resolution, and sensitivity. With modern instrumentation, the time constant may be controlled within certain limits. Martin et al.,[5] building on of the work of Schmauch[21] and McWilliam and Bolton,[37] offered a definition for the detector time constant (τ) in terms of retention time and the plate height of the column.

$$\tau = \theta(1 + \mathrm{k})(\mathrm{LH})^{1/2}/\mathrm{u} \tag{36}$$

If we are operating at (or near) the minimum plate height at all times, relationship (3) will be valid, and we may substitute d_p in place of H. Dropping the unnecessary terms, we obtain the following expression:

$$\tau \,\alpha\, (\mathrm{Ld_p})^{1/2}/\mathrm{u} \tag{37}$$

It has been shown that the number of points required to sufficiently define a peak varies from 9 (for symmetrical guassian peaks) to 32 (for peaks that are significantly non-guassian).[38] However, it has been suggested that, in most cases, 15 points will be sufficient.[4] Therefore, scaling the time constant as described above may not be necessary in cases where the change in column dimensions leaves us with more than 15 points across the smallest peak of interest.

Injection Time

Martin et al. also derived an equation for the maximum injection time (t_{inj}).[5] The equation presented below is a modification of this equation; L/H has been substituted for N and $L(1 + k)/u$ has been substituted for t_r, in order to make the relationship to column length explicit.

$$t_{inj} = \theta K(1 + k)(LH)^{1/2}/u \tag{38}$$

It may be noted that this equation is identical to the equation for the time constant, except for the addition of a K term, which is related to the shape of the injection band at the column inlet. Hence, when the constant terms are dropped, we find that the injection time depends on the same variables as the time constant.

$$t_{inj} \alpha (Ld_p)^{1/2}/u \tag{39}$$

Length and Diameter of Connecting Tubing

Band broadening in the connecting tubing occurs due to longitudinal diffusion and mobile phase mass transfer effects. Formulas have been derived to quantify this effect based on the Golay equation which describes the flow of a fluid in open tubes.[5] Due to the high linear velocities of the mobile phase in the tubes, as well as the fact that the tubes have a rough surface and are generally not perfectly straight, the flow through the tubes is non-laminar. As a result, the band spreading is less than that which is predicted by the Golay equation.[40,41] Having studied the problem extensively, Neue has suggested the following alternative approach: where σ^2 is the variance of the band broadening due to the tubing, h is the plate height corresponding to the band broadening in the tubing, and l and d are the length and diameter of the connecting tubing, respectively.[3]

$$\sigma^2 = h l \pi^2 d^4/16 \tag{40}$$

We now need to express this in terms of the column dimensions, the column plate height, and the observed fraction of peak broadening. Following the approach of Martin et al., we can express the variance of the connecting tubing as follows:[5]

$$\sigma^2 = \theta^2 V_r^2 H/L \tag{41}$$

Given that $V_r = \pi \varepsilon u d_c^2 t_r/4$ and $t_r = L(1 + k)/u$, we can re-write this as:

$$\sigma^2 = \theta^2 \pi^2 \varepsilon^2 (1 + k)^2 d_c^4 L H/16 \tag{42}$$

Finally, by setting the right half of Equation (40) equal to the right half of Equation (42), and simplifying, we obtain:

$$l d^4 = \theta^2 \varepsilon^2 (1 + k)^2 d_c^4 L H/h \tag{43}$$

The value of h is generally between 2 and 10;[42,43] it is reasonable to approximate its value as being equal to 6. For the purpose of scaling, however, all that is important is the following proportional relationship, where we have again substituted d_p in place of H:

$$l\,d^4 \alpha\, d_c^4\, L\, d_p \tag{44}$$

Applying the Relationships to Monolithic Columns

Monolithic columns utilize a continuous porous solid as the stationary phase, as an alternative to the conventional packed bed. These columns are becoming increasingly popular; therefore, it will be useful to address the question of whether the relationships presented in this article apply to monolithic columns. Parameters such as column length, column diameter, and capacity factor are equally relevant when dealing with monolithic columns. Hence, insofar as these variables are concerned, the relationships presented in this article are entirely appropriate for monolithic columns. What is not obvious is the question of which parameter would replace the particle diameter, as monoliths are not composed of discrete particles. For pressure drop, and related equations, one could use the equivalent particle diameter concept that has been suggested by some.[44–46] This means that a particle diameter would be determined such that the permeability of a monolithic column is equivalent to that of a conventional packed bed column with the stated particle diameter. With respect to dispersion issues, the equivalent particle diameter approach is not as useful, as it would not allow comparison of monolithic columns and conventional columns. The simplest approach for dispersion related issues would be to leave the equations in terms of plate height and utilize kinetic calculations or plots (such as Figure 1) to determine what the plate height (H) values would be in a given situation.

PART II: GRADIENT SCALING: KEEPING RELATIVE RETENTION TIMES CONSTANT WITH GRADIENT METHODS

Linear Gradients

With gradient separations, changes in the gradient time, flow rate, or dimensions of the column can affect the relative retention times of the peaks. In some cases, it will be desirable to keep the relative retention times constant when changing these variables. This is important once a gradient separation has been developed to the point where the analyst wants to maintain the relative retention times of the peaks constant while any additional optimization

steps are carried out. This will also be useful when optimizing a method where peak identification based on relative retention times has already been established. With simple linear gradients, this can be accomplished by keeping the ratio of the gradient volume to column volume constant. We may write this as follows: where t_G is the gradient time, i.e., the time interval over which the mobile phase composition is ramped. This rule has been observed empirically[47–49] and can be derived using linear solvent strength theory.[49]

$$(t_G F)/(d_c^2 L) \tag{45}$$

The parameters can be adjusted in any way and, as long as the ratio remains constant, the relative retention times of the peaks will remain constant. However, assuming the linear velocity is in the optimal range, it may be preferrable to keep the linear velocity constant. In these cases, the recommended approach is to scale the flow rate in proportion to the square of the column diameter, and scale the gradient time in direct proportion to column length. The initial and final mobile phase compositions should always be kept constant.

It should be noted that, when dealing with separations of large molecules, it may be less important to scale column length as per relationship (45). The standard chromatographic process involves the analytes continuously partitioning back and forth between the stationary and mobile phases. However, with larger molecules, the situation approaches what is sometimes referred to as on/off behavior. This is where the analytes remain associated with the stationary phase until a certain mobile phase strength is reached, at which point they are rapidly desorbed and eluted.[50] It has been observed that, for separations of large molecules, column length often has less significance than that which is generally observed with small molecules: which is consistent with the on/off mechanism.[51–60] The phenomenon is fundamentally related to the number of binding sites between the analyte and the stationary phase and, therefore, is not strictly dependent on molecular weight. However, as a rough guide, it can be stated that conventional partitioning behavior will generally be observed for analytes with molecular weights below 2000, full on/off behavior will generally be observed for analytes with molecular weights above 10,000, and intermediate behavior will often be observed between 2000 and 10,000.[61] Clearly, this does not apply to size exclusion, and some exceptions may be observed with ion exchange separations in cases where large molecules interact with the stationary phase via a small number of binding sites. The significance of this with respect to scaling is that, in cases where on/off behavior is dominant, the analyst may observe that changing the length of the column will have a minimal affect on relative retention times. This is because the "effective column length" is not changing.

Instrument Dwell Volume and Isocratic Segments at the Beginning of the Run

Dwell volume is defined as the volume within the HPLC system from the point where the solvents are mixed to the head of the column. The dwell time is the time that it takes for the mobile phase to move through this region. The instrument dwell volume can be determined by removing the column and running a gradient from 100% methanol (A solvent) to methanol with 0.1% acetone (B solvent) over 20 minutes. If this is done at a wavelength of 265 nm, the chromatogram will essentially be a visual picture of the gradient. The time when the absorbance is halfway between the minimum and the maximum is then determined and, from this time, 10 minutes (half the gradient time) is subtracted. The result will be the dwell time. Dwell volume is then given by the product of dwell time and the flow rate ($V_D = t_D \times F$).[4] When dealing with gradient methods, the dwell volume of the system will affect the spacing of the chromatographic peaks. Therefore, adjustments will be needed if the relative retention times are to be kept constant.

First, it is important to understand the mechanism by which this occurs. Consider that, as the strength of the mobile phase increases during the course of a gradient run, the dwell volume will cause a delay between when each incremental change is effected, and when this change reaches the head of the column. As a result, the retention times of the peaks will be delayed. This is identical to the effect of a programmed isocratic hold segment at the beginning of the run, which would cause exactly the same type of gradient delay.

The key is that this phenomenon does not affect all peaks equally. Peaks which elute at the very beginning of the gradient are less delayed, because these components generally migrate through the column under the initial conditions. Peaks which elute towards the middle or end of the gradient will generally realize the full effect of the delay. Hence, in this way the relative spacing of the peaks is affected. The spacing between adjacent peaks would be most noticeable between these two regions, i.e., between the early to the middle part of the gradient ramp.

The total delay is due to the sum of the dwell time (t_D) and any programmed isocratic hold segment at the beginning of the run ($t_{iso,\ beginning}$). Therefore, it is the quantity ($t_D + t_{iso,\ beginning}$) that is important. However, it will be more convenient to express the quantity as ($V_D/F + t_{iso,\ beginning}$) in terms of the two things one would potentially adjust: the dwell volume and the initial isocratic hold time. It will be necessary to scale this quantity with respect to changes in the flow rate, as well as changes in the length of the column. These issues are discussed in the following two paragraphs, respectively.

Assume that we reduce the flow rate by a certain factor by application of relationship (45). The result is that the dwell time will increase by this same factor, simply because it takes the mobile phase this much longer to travel through the system dwell volume. If we want to keep the delay constant so

as to keep the relative retention times constant, we would need to reduce the dwell volume to bring the quantity ($V_D/F + t_{iso,\ beginning}$) back to its original value. Alternatively (and much more easily), we could reduce the initial isocratic hold time to accomplish the same goal. In other words, the above quantity must be kept constant.

$$V_D/F + t_{iso,beginning} = \text{constant} \tag{46}$$

If we scale as per relationship (45), resolution will change in proportion to the square root of column length, and the distance between the peaks will change in direct proportion to column length, just as for isocratic methods (as will be demonstrated in Part III). From the discussion above, we know that changes in the gradient delay, caused by the dwell volume and initial isocratic hold, can also cause certain peaks to become more separated, or to move closer together. Since both of these factors will affect the spacing of the peaks, it follows that they must be changed by the same amount if all relative retention times are to be maintained constant. Hence, when column length is changed, the quantity ($V_D/F + t_{iso,\ beginning}$) must be changed by the same factor.

This can be captured by adding a column length term to the right side of Equation (46), resulting in:

$$V_D/F + t_{iso,beginning} = L \times \text{constant} \tag{47}$$

This tells us how the dwell volume and/or initial isocratic hold time need to be adjusted in relation to changes in F or L, so as to keep relative retention times constant. Simple algebra allows us to state this in a more convenient way. Specifically, we can properly scale with respect to dwell volume and initial isocratic hold segments by maintaining the following ratio constant:

$$(V_D + t_{iso,beginning} \times F)/FL \tag{48}$$

If the flow rate is scaled in direct proportion to the square of the column diameter (which is often the case) we can re-write this expression as follows:

$$(V_D + t_{iso,beginning} \times F)/d_c^2 L \tag{49}$$

Relationship (49) is equivalent to the conclusion presented by others.[49,62,63] However, relationship (48) is more general, as it does not make any assumptions about how flow rate will be scaled.

In cases where the adjustment cannot be made by changing the initial isocratic hold time (because it would have to be reduced by an amount greater than its original value), there is an alternative that can be evaluated before attempting to actually modify the dwell volume of the HPLC. Some modern systems allow the analyst to delay the sample injection relative to the start of the gradient.[4,63,64] This approach can be used when the flow rate and/or the length of the column will be reduced. The "effective dwell volume" will be reduced by the product of the delay time and the flow rate (the new flow rate, if the flow rate is changed).

Isocratic Segments in the Middle of a Complex Gradient Method

Isocratic segments in the middle of a complex gradient (such as the segment labeled 'hold 2' in Figure 2) will delay the subsequent gradient ramp ('ramp 2' in Figure 2), in exactly the same way as the dwell volume and initial isocratic hold segment delay the first gradient ramp. It is fairly easy to extend the treatment above to deal with this issue.

In the previous discussion, it was only the instrument dwell volume term that had a dependence on flow rate. This is obviously not relevant here. However, the connection between isocratic segments and column length does apply. Hence, isocratic segments in the middle of a complex gradient ($t_{iso,\ middle}$) should be scaled in direct proportion to any changes in the length of the column. In other words, the following ratio should be maintained constant. It may be noted that this is the same type of expression that would result if the dwell volume term was dropped from relationship (48).

$$t_{iso,\ middle}/L \tag{50}$$

Only isocratic segments that are followed by a gradient ramp will affect relative retention times by the mechanism described. Therefore, adjusting the isocratic segment at the end of the method is not required in order to maintain relative retention times constant.

Sample Calculation

An example may help to clarify the issues discussed in the last few sections. Consider a gradient which starts with a 4 minute hold at a composition of

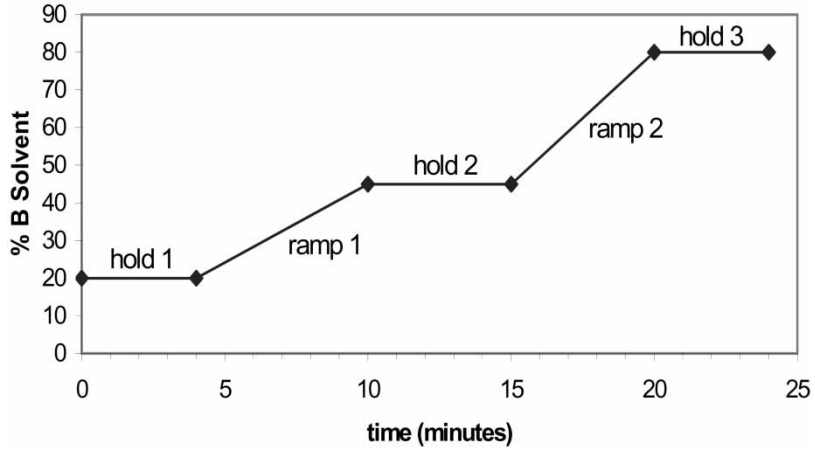

Figure 2. Sample gradient.

80/20 A/B, then ramps to 55/45 A/B between 4 minutes and 10 minutes, followed by a second isocratic hold from 10 to 15 minutes, then another ramp to 20/80 A/B between 15 and 20 minutes, with a final isocratic hold from 20 to 24 minutes. This gradient is depicted in Figure 2. Suppose that the flow rate was 1 mL/min, and the column has a diameter of 4.6 mm and a length of 10 cm. Finally, suppose that the HPLC used in this experiment has a dwell volume of 1 mL. This gradient is more complex than that which one would normally want (or need) to use. This was intentional, however, as it helps to illustrate the procedure.

In this exercise we will consider what adjustments should be made if we were to reduce the diameter of the column from 4.6 to 3.2 mm, and reduce the length of the column from 10 cm to 5 cm. It is logical to first evaluate the linear gradient ramps using relationship (45). This will determine how much the flow rate will change, which is information we need to apply relationship (48) to address the dwell volume and initial isocratic hold time.

Ramp 1

If the linear velocity is in the optimal range, and we want to keep it there, the recommended use of relationship (45) is to scale the flow rate in proportion to the square of column diameter, and to scale the gradient time in proportion to column length (in all cases, the initial and final mobile phase concentrations should be kept constant). The change in the square of the column diameter can be calculated as:

$$d^2_{c\,\text{original}}/d^2_{c\,\text{new}} = 4.6^2/3.2^2 = 2.1$$

The flow rate in the original method was 1 mL/min. Since the square of the column diameter will be reduced by a factor of 2.1, the new flow rate will be:

$$1/2.1 = 0.48\,\text{mL/min}$$

The gradient time will be scaled in direct proportion to column length. Therefore, the gradient time will be reduced by a factor of 2. Since the gradient time for Ramp 1 in the original method was 6 minutes, the new gradient time for Ramp 1 will be 3 minutes.

Ramp 2

The scaling of ramp 2 would be accomplished in the same way as for ramp 1. The change in flow rate has already been established. Thus, all that remains is to reduce the gradient time for ramp 2 in proportion to the change in column length. In the original method, Ramp 2 takes place over an interval of 5 minutes, therefore, in the new method Ramp 2 should take place over 2.5 minutes.

Dwell Volume and Hold 1

The dwell volume must be evaluated together with the initial isocratic hold segment, using relationship (48). The first step is to determine how much the product of flow rate and column length will change on going to the new column. With flow given in mL/min and length given in cm, we can write:

$$(FL)_{original} = 1 \times 10 = 10$$

$$(FL)_{new} = 0.48 \times 5 = 2.4$$

Therefore, the quantity (FL) will be reduced by a factor of:

$$10/2.4 = 4.2$$

Relationship (48) suggests that the quantity $(V_D + t_{iso,\ beginning} \times F)$ must change by this same factor. In the original method this quanity is equal to:

$$(1 + 4 \times 1) = 5$$

Therefore, with the revised method, the quantity must equal:

$$5/4.2 = 1.2$$

By simple algebra we can determine that this can be accomplished by reducing $t_{iso,\ beginning}$ (i.e., Hold 1) from 4 minutes to 0.42 minutes, such that the quantity becomes:

$$(1 + 0.42 \times 0.48) = 1.2$$

In this way, we have made the necessary adjustment without having to take on the difficult task of actually changing the dwell volume of the system.

Hold 2

Relationship (50) suggests that the time corresponding to isocratic segments in the middle of a complex gradient should be scaled in direct proportion to column length. We are reducing column length by a factor of 2, therefore, we must reduce $t_{iso,\ middle}$ from 5 to 2.5 minutes.

Hold 3

No scaling is required for the hold segment at the end of the run, because hold segments that are not followed by a gradient ramp do not affect the relative retention times of the peaks, and most peaks will elute prior to this point anyway. For the purpose of this exercise, we will leave Hold 3 unchanged.

The total run time for the new method would be 12.42 minutes. A plot of this revised method is shown in Figure 3. The same scale as Figure 2 is used to allow for easy comparison.

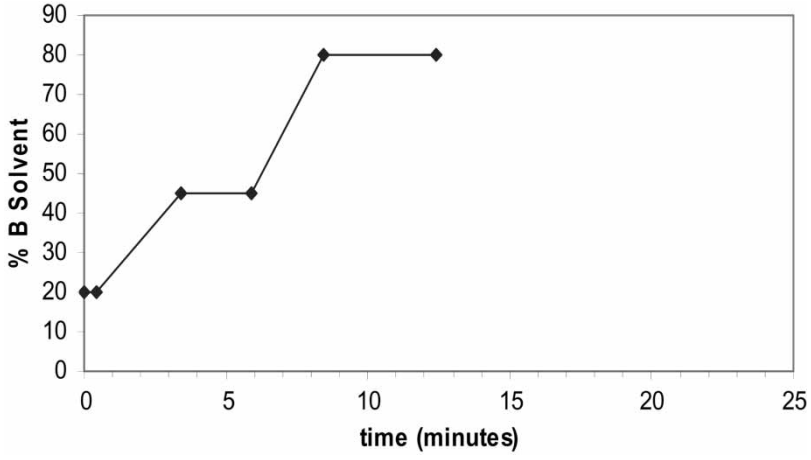

Figure 3. Scaled gradient method.

Gradient Scaling During Method Transfer

Relationship (48) can also be used in the context of a method transfer, where the column dimensions would be constant, but significant differences may exist in the instrument dwell volumes. By use of this relationship, it is possible to cancel the effect of any difference in the dwell volumes of the two HPLC systems. This may be important in cases where a method shows unacceptable chromatography when run on a different HPLC system, or when a method which identifies all components on the basis of their relative retention times (such as a related substances assay) must be run on a different system. Here again, this correction can often be made simply by modifying the isocratic hold time at the beginning of the run. The goal is that the quantity $(V_D + t_{iso,\ beginning} \times F)$ is the same with the new HPLC system as on the previous system.

The alternative of delaying the injection is also possible here if we are transferring to a system with a larger dwell volume. The injection should be delayed by an amount equal to the difference in the dwell volumes of the two systems divided by the flow rate $[(V_{D,\ new} - V_{D,\ original})/F]$; but, it may be easier to simply run a few experiments and determine what works.

PART III: APPLYING THE BASIC RELATIONSHIPS TO GRADIENT METHODS

Interdependency of Variables in a Gradient Method

At this point it will be useful to revisit the conclusions of Part I with respect to their applicability to gradient methods. Fundamentally, all of the relationships

discussed in Part I are equally valid for gradient methods. The complication is that, with gradient methods, the mobile phase composition and, hence, the capacity factor, is changing during the course of the run. Linear solvent strength theory discusses gradient separations in terms of the average capacity factor (k^*), which may be defined as follows, as derived from Equations (2.11) and (2.13) in Ref. [64]. If the initial mobile phase composition is kept constant (which will generally be the case during optimization and scaling) this relationship applies to all peaks, including early eluters.[64]

$$k^* = 3.48\, t_G\, F/\Delta\phi\, S\pi\varepsilon\, d_c^2\, L \tag{51}$$

The gradient version of the plate height (H^*) would have the same form as Equation (1), except that we would replace k with k^* and an average value would have to be used for the mobile phase diffusivity (D_M). Similarly, the resolution of a gradient method can be written in a way that is analogous to the isocratic expression, except for the addition of the gradient compression factor (G), as follows:

$$R^* = \left(\frac{\sqrt{L}}{4G\sqrt{H^*}}\right)\left(\frac{k_2^*}{1+k_2^*}\right)\left(\frac{\alpha^*-1}{\alpha^*}\right) \tag{52}$$

The gradient compression factor (G) addresses the fact that the mobile phase composition is increasing during a gradient analysis and, therefore, the trailing part of a peak experiences a slightly stronger mobile phase, at any point in time, than the leading part of the peak. As a result, the analytes in the trailing portion of the peak will migrate somewhat faster than the analytes in the leading part of the peak. It has been predicted that this phenomenon would cause some compression (or sharpening) of the peaks.[65] This topic has been somewhat controversial. Some have suggested that the gradient compression which is actually observed in practice is small or nonexistent (or $G \approx 1$).[57,66–71] Recently, a study was conducted which definitively proved that gradient compression does occur (or $G < 1$).[72] The extent to which gradient compression occurs is a function of the steepness of the gradient. Therefore, when adjustments are made such that k^* decreases (i.e., we go to a steeper gradient), there will be a "hidden benefit" as there will be an increase in gradient compression (or a reduction in G) which results in higher resolution; and the opposite will be true when k^* increases. A further discussion of gradient compression is beyond the scope of this paper. However, to give some idea of the magnitude of the effect, the most recent (and well controlled) study showed that, with linear gradients, improvements in resolution were generally between 9 and 32%.[72] Given that gradient compression causes the chromatographic peaks to be narrower than what would be predicted from the van Deemter equation, the gradient compression factor can be thought of as a correction

factor for the plate height, such that, in the gradient version of any optimization or scaling relationship, the H term should be multiplied by G^2.

By looking at Equations (51) and (52), it is clear that there is a more complex interdependency of variables in gradient HPLC, as compared to isocratic separations. With gradient methods, the capacity factor is a function of the gradient time, the flow rate, and the dimensions of the column. As a result, the effect of certain changes will not be as predictable with gradient methods. Consider, for example, that if we decrease column length, Equation (52) suggests that this will have a negative effect on resolution. However, Equation (51) suggests that we will simultaneously increase the gradient capacity factors, which will have a positive effect on resolution. This, in turn, will cause a decrease in gradient compression (or an increase in G) which would have a negative effect. Hence, the net effect will be a result of the competition between these processes.

Suppose, now, that we increase the linear velocity such that we are beyond the optimal region of the van Deemter curve. The plate height will increase, which will have a negative effect on resolution. However, k^* will also increase, which will have a positive effect on resolution. And this, in turn, will cause gradient compression to decrease (increase in G), which has a negative effect. Here again, the net effect will be determined by these competing processes. This demonstrates why the optimal linear velocity of a gradient method will generally be somewhat higher than that which is predicted by the van Deemter equation, because the optimal linear velocity of a gradient method is dependent on this "capacity factor effect" in addition to the van Deemter effects that apply to all chromatographic methods. This has been demonstrated in several publications.[19,62,71,73,74] The "capacity factor effect" is more pronounced at lower values of t_G, simply because these generally correspond to lower values of k^* where the dependence of resolution on the capacity factor is stronger. Hence, the optimal linear velocity will increase as the gradient time is decreased.

Although the focus of this article is on the mathematical tools needed to optimize and scale chromatographic separations, the complex interdependency of the variables makes it difficult to use these tools when first developing a gradient method from scratch. Computer programs are now available that can assist the analyst in this process. Alternatively, a step-wise experimental approach such as the following can be used (this is a combination of conventional approaches and the author's own experience).

1. For gradient times that are predicted to be between 20 and 90 minutes, start with a 15 cm 3 μm column, or a 10 cm 1.7 μm column. If shorter gradient times are anticipated, start with a shorter column, and vice versa.
2. Run a gradient from 5% to 95% organic over a period of 30 minutes. Adjust the initial and final mobile phase composition to eliminate any wasted time at the beginning or end of the chromatogram.

3. Evaluate the gradient time. Increase the gradient time to improve resolution, or decrease gradient time to eliminate wasted space in the chromatogram, and hence, increase the speed of analysis. If desired, add isocratic hold segments in the method so as to further optimize the separation.
4. The optimal linear velocity will change as the gradient time changes, due to the "capacity factor affect" (discussed above). Therefore, it would be helpful to have an easy way to "ballpark" what the linear velocity should be for each gradient time that's evaluated. Although it is possible to calculate the optimal linear velocity for a given t_G (Ref. [71]) this is cumbersome. The author has found that the following rule of thumb can offer an easy (though admittedly very approximate) way to estimate what linear velocity should be used for a given gradient time.
   - Assume that, for gradient times ≥ 2 hours, the optimal linear velocity is equivalent to that predicted by the van Deemter equation. Further, assume that below 2 hours each reduction of the gradient time by a factor of 4 requires a doubling of the linear velocity.

   Once the gradient time is established, some quick experiments should then be conducted to find the true optimal linear velocity.
5. Optional: evaluate different column lengths and/or particle diameters, if necessary, to optimize resolution, or speed, or sensitivity, etc.

The remainder of Part III will focus on optimizing an already existing method and then performing any requisite scaling. By inspection of relationships (45) and (51), it is obvious that, when relationship (45) is used to keep the relative retention times constant, k^* will also remain constant. It follows from this, in turn, that the gradient compression factor and the optimal linear velocity will also remain constant. Finally, relationship (3), which was used in many of the derivations in Part I, will be valid (relationship (3) was derived by minimization of the van Deemter equation;[2] this process treats the capacity factor as a constant). In these cases, the situation is simplified and the relationships between the variables in the optimization and scaling relationships are analogous to those for isocratic methods. Conversely, the analyst may decide that they want to change some of the variables independently. In these cases, it must be realized that the relative retention times of the peaks, k^*, G, and the optimal linear velocity will change, and that relationship (3) is no longer a safe assumption, making the effect of any manipulations more difficult to predict. In these cases, it may be best to resort back to the stepwise approach summarized above.

Effect of Instrument Dwell Volume and Isocratic Segments

In the previous section, it was argued that, when relationship (45) is followed, the optimization and scaling relationships may be directly

applied to gradient methods. This discussion centered on the issue of linear gradient segments. The issues of isocratic segments and dwell volume must now be addressed. The primary effect of instrument dwell volume or programmed isocratic segments at the beginning of the run is to increase the amount of time that the analytes spend in the stationary phase (with early eluters being an exception). Although there is, theoretically, some longitudinal band broadening that occurs in the stationary phase, it is generally insignificant and can be ignored.[75] Therefore, these factors do not significantly affect the widths of chromatographic peaks. Isocratic segments in the middle of a gradient are uncommon (and, hence, not much of a concern). However, in cases where they exist, the same argument can be made, in that they delay the subsequent gradient ramp, thereby increasing the time that the more retained analytes spend in the stationary phase. Finally, an isocratic segment at the end of the run is generally not significant, as almost all analytes will be eluted by this point. It follows from this that—to a rough first approximation—the presence of dwell volume and isocratic hold segments does not affect the validity of the optimization and scaling relationships that will be discussed in this section (with early eluting analytes being an exception). Furthermore, with the exception of the relationships for resolution and resolution per unit time, it does not matter whether or not we utilize relationships (48) and (50).

Optimizing the Method

Efficiency and Resolution

The expression for efficiency in gradient methods (N^*) has a form analogous to that for isocratic methods, except that H is replaced with H^*, which is then multiplied by the gradient correction factor squared:

$$N^* = L/H^*G^2 \tag{53}$$

Efficiency is not a terribly useful concept with gradient separations because it does not reflect the effect that various alterations can have on k^*. Therefore, the resolution equation is recommended when optimizing gradient separations. The gradient version of the resolution equation was presented earlier (Equation (52)).

As mentioned above, if relationship (45) is applied, k^*, G, and the optimal linear velocity will also remain constant. If the analyst decides to alter certain parameters independently (for example increasing t_G in order to increase k^*, and hence, resolution), the situation becomes more complicated. In these cases, it may be difficult to use the mathematical tools and it may be preferrable to utilize the stepwise approach described above. For example, relationships (12) and (13) will not be valid in these cases, as they

assume constant k^*. In fact, if we reduce the length and particle diameter by the same factor (with all other parameters held constant), the resolution will improve in most cases, simply because the reduction of column length carries less of a penalty in gradient methods. This has been demonstrated in the literature.[19,62,64]

Lastly, it should be understood that, when the relative retention times of all peaks are not kept constant (as per the procedures discussed in Part II), this can affect the spacing between the peaks, which effectively changes resolution. Therefore, when Equation (52) is used, it is important to also use relationships (48) and (50), as necessary.

Speed of Analysis

To obtain the gradient version of Equations (17), (18), and (19) we would replace k' with k'^*, k_2 with $k_2{}^*$, H with H^*, and α with α^*. Then G^2 would be added to the denominator of relationship (18) and G would be added to the denominator of relationship (19). Again, resolution is a more meaningful parameter with gradient methods; hence, it is recommended to evaluate speed in terms of resolution per unit time. As discussed in Part I, making changes that affect the capacity factors will generally not be the best way to improve resolution per unit time. Therefore, the general approach to increase the speed of a gradient method would be to first reduce t_G while increasing the linear velocity and/or reducing column length, such that k^* remains constant. With small diameter particles, increasing the linear velocity would be preferrable, as this will have less of an effect on the total resolution than reducing the length of the column. This approach will increase the speed of analysis and, simultaneously, maintain relative retention times, k^*, gradient compression, and the optimal linear velocity constant. Recall that, when we are evaluating the resolution per unit time of a gradient method, it is important to also use relationships (48) and (50) as necessary.

As mentioned in Part I, the simplest approach, in cases where we have just the necessary degree of resolution, but want to increase the speed of analysis, is to reduce both column length and particle diameter by the same factor (x). In this way, the efficiency and resolution will remain constant (see relationships (12) and (13)). Of course, when utilizing this approach for gradient methods, it is also necessary to apply relationship (45). Assuming that we increase the flow rate so as to stay in the optimal linear velocity range (as per relationship (5)), this approach will increase the resolution per unit time by x^2. However, the pressure will also increase by x^2. This approach can be somewhat confusing with gradient methods; therefore, it is useful to break it down into steps:

1. Make a decision concerning how much L and d_p will be reduced.
2. Make a decision concerning how the flow rate will be adjusted. If we are in the optimal linear velocity range, and want to stay there, adjust the flow

rate inversely with the change in d_p (as per relationship (5)). However, with small diameter columns, the best speed will be obtained at linear velocities above the optimal.
3. Once we know how much the flow rate and length of the column will change, make the necessary adjustment in gradient time so as to keep ratio (45) constant.
4. Apply relationships (48) and (50) as necessary.

Sensitivity

In the beginning of the discussion on sensitivity, Equation (24) was derived. To make this equation valid for gradient methods, we would need to replace H with H^* and k with 0.5 k^*, which is the capacity factor of a peak just prior to its elution from the column. Lastly, we would need to add the gradient compression factor (G) to the denominator. We use 0.5 k^* because, in this context, what is significant is the linear velocity with which the peak is moving at the end of the column. And this, in turn, is a function of the capacity factor that applies at the end of the column (this issue will be discussed in more detail below). From the previous discussion, we know that H^*, k^*, and G will not be affected by column dimensions, when scaling is done as per relationship (45); therefore, these may be treated as constants just as was done for isocratic methods. The conclusion is that all the relationships derived for sensitivity may be applied to gradient methods as well, when scaling as per relationship (45). Here again, when we make adjustments such that k^* changes, the situation becomes more complicated.

It is logical that, in most cases, 0.5 k^* will be smaller than the capacity factors of peaks that elute in the middle and end of a typical isocratic method. This is the reason for the inherently better sensitivity of gradient methods for these peaks.

Solvent Consumption and Waste Generation

The conclusion reached in Part I was that solvent consumption and waste generation are a function of the length and diameter of the column. This conclusion is, in no way, dependent on whether or not the composition of the mobile phase is changing during the run. Thus, the conclusion may be extended to gradient methods with no stipulations, except that we would replace k' with k'^* in Equation (30).

Scaling to Minimize Extra-Column Effects

In order to "convert" to the gradient version of the extra-column effects equations, we would need to replace H with H^*G^2 (as was described above). However, the k terms that appear in these equations will be handled

in a somewhat different manner. Instead of replacing k with k^*, we will replace it with k_o or $0.5\,k^*$, i.e., the capacity factors at the beginning of the column or the end of the column, respectively. The reason for this is easy to understand. Chromatographic peaks move at a lower velocity while on the column which has a stationary phase, than when in the injection valve, connecting tubing, or detection cell, which do not have a stationary phase. As a result, there is a focusing effect which occurs as peaks are transferred onto the column, and a band broadening that occurs as peaks exit the column. The more retained a peak is (higher k), the lower is the on-column velocity and, hence, these focusing and band broadening affects are more pronounced. The significance of the capacity factor terms in the extra-column effects equations is to account for these phenomena. Therefore, what is important in these equations is the capacity factor at the head of the column (k_o) or the capacity factor at the end of the column ($0.5\,k^*$). Specifically, k_o would be the relevant parameter for the equations for maximum injection volume or speed of injection, as well as when evaluating the maximum length and diameter of connecting tubing upstream of the column; whereas, $0.5\,k^*$ would be the relevant parameter in the expressions for detector cell volume, detector time constant, when evaluating the maximum length and diameter of connecting tubing downstream of the column. As discussed above, $0.5\,k^*$ will also be appropriate in the sensitivity equations, as this is what determines the band broadening as the peaks are transferred into the detector.

For example, the gradient version of the detection cell volume equation would be:

$$V_d^* = \theta\,\pi\,\varepsilon(1 + 0.5\,k^*)\,d_c^2\,(LH^*)^{1/2}G/4 \tag{54}$$

Given that relationship (3) is valid for gradient methods, when we utilize relationship (45), we can combine relationship (3) and Equation (54). The result is the same scaling relationship that was generated for isocratic methods (relationship (35)).

In reality, the focusing effect discussed above would render the effect of certain attributes largely insignificant. Specifically, the effect of the injection volume, the speed of injection, and the connecting tubing upstream of the column (all relationships for which k_o would have been the appropriate parameter) will usually not be critical for gradient methods (with the exception of very early eluting peaks). In fact, it is generally possible to significantly increase the injection volume when the sample is dissolved in a solvent that is chromatographically weaker than the intial mobile phase.[64,76]

Finally, the presence of dwell volume and isocratic hold segments in the method do not significantly affect the validity of the scaling relationships. In fact, they are even less important than they were in the case for the optimization relationships. This is because, for the scaling relationships, all that matters is the capacity factor at the beginning or the end of the column. The extent to which there is instrument dwell volume or isocratic hold segments within the method does not affect k_o or $0.5\,k^*$.

SUMMARY

In this review, an attempt has been made to discuss all of the issues relevant to optimization and scaling in HPLC. Optimization decisions are generally governed by concerns such as maximizing efficiency and resolution, increasing the speed of analysis, improving sensitivity, or reducing the amount of mobile phase consumed and waste generated. It is, then, important to consider what scaling may be needed in order to minimize band broadening due to extra-column effects. The focus of the article has been on the mathematical concepts relevant to optimization and scaling. Hence, issues such as the selection of mobile phase and stationary phase (which are not generally approached mathematically) were not discussed.

Method optimization relationships can be used to make decisions about how the linear velocity and column dimensions should be selected or changed. It was demonstrated that the required flow rate, in order to achieve a certain linear velocity, can be calculated as follows (relationship (4) in the text):

$$F(\text{mL/ min}) = 0.15\pi\varepsilon\, u(\text{cm/sec})\, d_c(\text{mm})^2$$

It was shown that, when the linear velocity is in the optimal range, and the analyst wants to keep it there while additional optimization steps are carried out, the flow should be adjusted in relation to the column diameter and particle diameter as follows (relationship (5)). Exceptions to this are when we are trying to optimize the speed of analysis with smaller diameter particles or when trying to increase sensitivity with a mass-sensitive detector. In these cases, we would choose to work above the optimal velocity and, hence, deviate from relationship (5).

$$F_{opt} \;\alpha\; d_c^2/d_p$$

In many cases, the preferred approach for optimization is to reduce the particle diameter, as this will improve the total resolution, the speed of analysis, and the sensitivity, simultaneously. However, reducing the particle diameter has a significant penalty in terms of pressure. It should be noted that, when the flow is adjusted so as to stay at the optimal linear velocity, the pressure will change inversely with the cube of the particle diameter. The relationships for pressure drop are reiterated below (relationships (6) and (7) in the text).

$$\Delta P\,(\text{psi}) = \frac{30778F(\text{mL/min})\eta\,(\text{cP})\,L(\text{cm})}{d_c(\text{mm})^2 d_p(\mu\text{m})^2}$$

$$\Delta P\,(\text{psi}) = \frac{14504\,\varepsilon\, u(\text{cm/sec})\eta\,(\text{cP})\,L(\text{cm})}{d_p(\mu\text{m})^2}$$

Optimization relationships are summarized in Table 3, along with comments regarding their application. The dependence of sensitivity on column diameter is a particular source of confusion and, hence, warrants a brief comment. It is often stated that reducing the diameter of the column will result in improved sensitivity. In fact, this is only true when dealing with a concentration-sensitive detector, in cases where there is only a small volume of sample available for injection. Even then, there are things which may partially offset this effect. If dealing with spectroscopic detectors (which contain detection cells), it will generally be required to reduce the dimensions of the cell when the dimensions of the column are reduced (as per relationship (35)); and, there will often be some signal-to-noise penalty associated with this reduction. The right-most column of Table 2 provides multipliers which describe the dependence of the signal-to-noise ratio on the pathlength (b) and cross section (c) of the detection cell (these multipliers are based on idealized conditions). Finally, when the diameter of the column is reduced, most of the extra-column effects issues become more problematic. This can result in band broadening which will also carry a signal-to-noise penalty due to the reduction in peak height. It is important that the analyst is aware of these issues when making decisions about whether to reduce the diameter of the column.

The optimization relationships presented in the text are valid for gradient methods. However, with a gradient separation, there is a more complex interdependency of variables that makes optimization somewhat more difficult. It was shown that, if we apply relationship (45), the situation is considerably simplified as:

- The relative retention times of the peaks will remain constant (if relationships (48) and (50) are also used, as necessary).
- The average gradient capacity factor (k^*) will remain constant.
- The degree of gradient compression (given by G) will remain constant.
- The optimal linear velocity will remain constant.
- Relationship (3) will still be valid, allowing substitution of d_p in place of H when we are operating near the minimum plate height.

If the analyst decides to adjust various parameters independently (i.e., does not follow relationship (45)), the effect of these adjustments will be more difficult to predict and it will be difficult to use the mathematical formulas. In these cases, it is probably best to utilize the stepwise procedure described in the text to optimize a gradient method (pages 631–632).

The optimal linear velocity of a gradient method will generally be somewhat higher than that which is predicted by the van Deemter equation, due to the fact that increasing the linear velocity also results in an increase in the average gradient capacity factor (see relationship (51)). This "capacity factor effect" becomes more significant as we go to shorter gradient times (t_G), because these generally correspond to smaller values of k^* where the dependence of resolution on the capacity factor is stronger. Hence, the optimal linear velocity of a gradient method will increase as the

Table 3. Basic optimization relationships for packed columns[a]

Attribute	Relationship	Comments
Efficiency and resolution. Relationships (12) and (13)	$N \alpha L/d_p$ $R \alpha (L/d_p)^{1/2}$	Efficiency and resolution can be increased by increasing the column length and/or reducing the particle diameter (while staying at the optimal linear velocity). Reducing the particle diameter is the preferred option as increasing column length has a cost in terms of sensitivity, analysis time, and solvent consumption. When pressure is limiting, the remaining option is to increase both column length and particle diameter, with the former increased by an amount somewhat greater than the latter. This will, simultaneously, cause efficiency and resolution to increase while the pressure drop will decrease. Therefore, substantial increases in efficiency and resolution will be possible: with obvious costs in terms of sensitivity, analysis time, and solvent consumption.
Speed of analysis. Relationships (20), (21) and (22)	Analysis time α L/u N/time α u/d_p R/time α u/$(Ld_p)^{1/2}$	To increase the speed of analysis, the analyst should first evaluate reducing the particle diameter, especially if the original method utilizes a large particle. Increasing the linear velocity will carry a tradeoff to the extent that it moves us away from the optimal part of the van Deemter curve. However, for small diameter particles this is not significant, and increasing the linear velocity is, in fact, the most effective way to increase the speed of analysis in these cases. Reducing column length will increase the resolution per unit time, but will carry some tradeoff in terms of total efficiency and resolution. In a situation where we have just the necessary amount of resolution but want to increase the speed of analysis, the simplest approach is to reduce the column length and particle diameter by the same factor (x). In this case efficiency and resolution will remain constant (see relationships (12) and (13)) while the time of analysis would also decrease by x^2 (assuming we adjust the linear velocity to stay in the optimal range). However, the pressure drop will also increase by x^2, limiting the extent to which we can make these modifications. A step-wise procedure for applying this technique to gradient methods is provided in the text (page 634–635).

(*continued*)

Table 3. Continued

Attribute	Relationship	Comments
Sensitivity. Relationships (26), (27), (28), and (29)	$S/N \alpha M$	Should be used for spectroscopic detectors, when the injection volume will be scaled as per relationship (33). For appropriate values of M see Table 2. If the detection cell volume will not be modified M will be constant, and can be dropped. Should also be used for non-spectroscopic concentration-sensitive detectors (e.g. conductivity), if M is dropped.
	$S/N \alpha M/d_c^2 (Ld_p)^{1/2}$	Should be used for spectroscopic detectors, when the injection volume will not be scaled. If the detection cell volume will not be modified M will be constant, and can be dropped. Should also be used for non-spectroscopic concentration-sensitive detectors (e.g. conductivity), if M is dropped.
	$S/N \alpha u d_c^2$	Should be used for mass-sensitive detectors when the injection volume will be scaled as per relationship (33).
	$S/N \alpha u/(Ld_p)^{1/2}$	Should be used for mass-sensitive detectors when the injection volume will not be scaled.
Consumption and waste. Relationship (31)	Solvent consumed/waste generated $\alpha d_c^2 L$	To reduce solvent consumption and waste we can reduce column diameter or column length. Although column diameter will have a more substantial affect, several of the extra-column effects are more sensitive to reductions in column diameter than reductions in column length. Therefore, some combination of the two may be advisable.

[a]These relationships apply to gradient methods as well, when scaling is done as per relationship (45). If relationship (45) is not applied the situation becomes more complicated as k^*, G, and the optimal linear velocity will all be variable, and relationship (3), which was used in many derivations, is no longer valid. When evaluating the resolution or resolution per unit time of a gradient method, it may also be necessary to utilize relationships (48) and (50): see comments in Table 5 as to when relationships (48) and (50) are necessary.

gradient time is decreased. A rule of thumb for estimating this was suggested in the text (see item 4 on page 632).

It was suggested in the text that, to a first approximation, the presence of instrument dwell volume and isocratic segments in a gradient method will

Table 4. Basic scaling relationships for packed columns

Attribute	Scaling relationship	Comments
Injection volume. Relationship (33)	$V_i \alpha d_c^2 (Ld_p)^{1/2}$	These issues only need to be considered when reducing the dimensions of the column (L, d_c, and d_p). These relationships apply to gradient methods as well, as long as relationship (45) is applied. It should be recalled that the following issues are generally not critical with gradient methods (with the exception of early eluting peaks) due to the focusing affect which results from the weak mobile phase at the beginning of the run: • the quality and volume of the injection • the injection time • the connecting tubing between the injection system and the column • the affect of any dead volume regions upstream of the column.
Detector cell volume and down-stream dead volume.[a] Relationship (35)	$V_d \alpha d_c^2 (Ld_p)^{1/2}$	
Detector time constant.[b] Relationship (37)	$\tau \alpha (Ld_p)^{1/2}/u$	
Injection time. Relationship (39)	$t_{inj} \alpha (Ld_p)^{1/2}/u$	
Diameter and length of connecting tubing. Relationship (44)	$l\, d^4 \alpha d_c^4\, L\, d_p$	

[a]This relationship should only be considered a first approximation for down-stream dead volume. Although dead volume regions upstream of the column are generally more problematic than down-stream dead volume, there is no relationship between allowable upstream dead volume and the dimensions of the column. Hence, there is no formula for upstream dead volume and it is not relevant in a discussion of scaling.

[b]Scaling the detector time constant will only be necessary when a change in column dimensions causes the number of readings across the smallest peak of interest to drop below a value of about 15.

have an insignificant effect on peak width and, therefore, do not affect the validity of the optimization relationships in Table 3.

Many of the formulas discussed in this article were derived using relationship (3), which assumes that we are operating at (or near) the minimum plate height at all times. In some cases, this assumption will not be valid, most likely because the analyst may decide to work above the optimal velocity so as to increase the speed of analysis or increase sensitivity when using a mass-sensitive detector. With the smaller diameter particles (which are generally used for modern analysis), this is not generally a significant concern because, with these particles, one can work above the optimal velocity and still be close to the minimum plate height (see Figure (1)). However, if

Table 5. Keeping relative retention times constant with gradient methods

Attribute to be scaled	Scaling relationship	Comments
Linear gradients. Relationship (45)	$(t_G\ F)/(d_c^2\ L)$	The ratio should be kept constant and the initial and final mobile phase compositions should be kept constant. Use of this relationship allows us to keep relative retention times, k^*, the degree of gradient compression (G), and the optimal linear velocity constant. It will also ensure that relationship (3), which was used in many of the derivations, is valid. In cases where the linear velocity is in the optimal range, and the analyst wants to keep it there, the preferred approach is to scale the flow rate in proportion to the square of the column diameter and to scale the gradient time in direct proportion to column length.
Dwell-volume and isocratic segments at the beginning of the run. Relationship (48)	$\frac{(V_D + t_{iso,\ beginning} \times F)}{F\ L}$	The ratio should be kept constant. Application of this relationship is required in order to maintain relative retention times constant, by making adjustments to account for instrument dwell volume and any isocratic hold segments at the beginning of the run. It is also required when evaluating the resolution or resolution per unit time of a gradient method that contain early eluting components. Relationship (45) should first be applied so we know how much the flow rate will change.
Isocratic segments in the middle of a complex gradient method. Relationship (50)	$t_{iso,\ middle}/L$	The ratio should be kept constant. Application of this relationship is required in order to maintain relative retention times constant when the gradient method contains an isocratic hold segment in the middle of the run. It is also required when evaluating the resolution or resolution per unit time of gradient methods containing an isocratic segment in the middle of the run.

working above the optimal velocity when using larger diameter particles, relationship (3) will become a less valid approximation. In these cases, it may be preferrable to leave the expressions in terms of H. A kinetic plot such as in Figure 1 could then be used to obtain a good approximation of what the value of H would be (or one could actually try to calculate it from the van Deemter equation, though this is generally not convenient).

Once the decision about changing column dimensions is made, the second step is to use the scaling relationships, summarized in Table 4, to make decisions about what scaling may be necessary, in response to the changes in column dimensions. Once again, when dealing with gradient methods, the situation becomes more complicated if relationship (45) is not used. The presence of dwell volume and isocratic hold segments in the method are even less important for scaling relationships than was the case for the optimization relationships. This is because, for the scaling relationships, all that matters is the capacity factor at the beginning or the end of the column. The presence of instrument dwell volume or isocratic hold segments within the method does not affect this.

In Part II, relationships were discussed that allow one to maintain constant relative retention times when working with gradient methods. These relationships are summarized in Table 5. Relationship (45) gives the proper way to scale with respect to linear gradient ramps. Relationships (48) and (50) give the proper way to scale with respect to the dwell volume of the HPLC and isocratic hold segments within a complex gradient method. As discussed above, application of relationship (45) also allows the basic optimization and scaling relationships (summarized in Tables 3 and 4) to be extended to gradient methods: with relationships (48) and (50) only being needed when evaluating the resolution or resolution per unit time of a gradient method.

Relationship (48) allows us to avoid having to modify the dwell volume of the HPLC (which is very difficult), by adjusting the initial isocratic hold time instead. If this option is not workable (e.g., due to an insufficient initial hold step in the original method), an alternative is to program the system to delay the injection of the sample. This approach can be used when the flow rate and/or the length of the column will be reduced. The "effective dwell volume" will be reduced by the product of the delay time and the flow rate (the new flow rate, if the flow rate is changed).

Relationship (48) can also be useful in the context of a method transfer when transferring to a system with a larger dwell volume. Delaying the injection can be used when transferring to a system with a smaller dwell volume.

REFERENCES

1. Scott, R.P.W. *Introduction to Analytical Gas Chromatography*; Marcel Dekker Inc.: New York, 1998.

2. Scott, R.P.W. *Liquid Chromatography Column Theory*; John Wiley and Sons: New York, 1992.
3. Neue, U.D. *HPLC Columns: Theory, Technology, and Practice*; Wiley-VCH Inc.: New York, 1997.
4. Snyder, L.R.; Glajch, J.L.; Kirkland, J.J. *Practical HPLC Method Development*, 2nd Ed.; Wiley-Interscience: New York, 1997.
5. Martin, M.; Eon, C.; Guiochon, G. Study of the pertinency of pressure in liquid chromatography, II. problems in equipment design. J. Chromatogr. **1975**, *108*, 229–241.
6. Giddings, J.C. Anal. Chem. **1964**, *36*, 1890.
7. Giddings, J.C. Anal. Chem. **1965**, *37*, 60.
8. Martin, M.; Blu, G.; Eon, C.; Guiochon, G. J. Chromatogr. Sci. **1974**, *12*, 438.
9. Martin, M.; Eon, C.; Guiochon, G. Study of the pertinency of pressure in liquid chromatography. I. Theoretical analysis. J. Chromatogr. **1974**, *99*, 357–376.
10. Martin, M.; Eon, C.; Guiochon, G. Study of the pertinency of pressure in liquid chromatography. III. A practical method for choosing the experimental conditions in liquid chromatography. J. Chromatogr. **1975**, *110*, 213–232.
11. Tsuda, T.; Novotny, M. Band-broadening phenomenon in micro capillary 44tubes under the conditions of liquid chromatography. Anal. Chem. **1987**, *50*, 632–634.
12. Scott, R.P.W.; Kucera, P. Mode of operation and performance characteristics of microbore columns for use in liquid chromatography. J. Chromatogr. **1979**, *169*, 51–72.
13. Halasz, I. Mass transfer in ideal and geometrically deformed open tubes. II. Potential application of ideal and coiled open tubes in liquid chromatography. J. Chromatogr. **1979**, *173*, 229–247.
14. Knox, J.H.; Gilbert, M.T. Kinetic optimization of straight open-tubular liquid chromatography columns. J. Chromatogr. **1979**, *186*, 405–418.
15. Guiochon, G. Optimization in liquid chromatography. In *High Performance Liquid Chromatography, Advances and Perspectives*; Horvath, C., (ed.); New York, 1980; Vol. 2, 1.
16. Purnell, J.H. J. Chem. Soc. **1960**, 1268.
17. DiCesare, J.L.; Dong, M.W.; Ettre, L.S. Very high speed liquid column chromatography. The system and selected applications. Chromatographia **1981**, *14*, 257–268.
18. DiCesare, J.L.; Dong, M.W.; Atwood, J.G. Very high speed liquid chromatography. II. Some instrumental factors influencing performance. J. Chromatogr. **1981**, *217*, 369–386.
19. Neue, U.D.; Alden, B.A.; Iraneta, P.C.; Mendez, A.; Grumbach, E.S.; Tran, K.; Diehl, D.M. In. *Handbook of Pharmaceutical Analysis by HPLC*; Ahuja, S., Dong, M., (eds.); Elsevier: Amsterdam, 2005; Chapt. 4.
20. Gilby, A. *Encyclopedia of Separation Science, First On-Line Update. Chromatography: Liquid: Detectors: Ultraviolet and Visible Detection*; Elsevier: Amsterdam; In press.
21. Schmauch, L.J. Response time and flow sensitivity of detectors for gas chromatography. Anal. Chem. **1959**, *31*, 225–230.
22. Halasz, I. Concentration and mass flow rate sensitive detectors in gas chromatography. Anal. Chem. **1964**, *36*, 1428–1430.
23. Sternberg, J.C. Extracolumn contributions to chromatographic band broadening. In *Advances in Chromatography*; Giddings, J.C., Keller, R.A.(eds.); Marcel Dekker: New York, 1966; Vol. 2.

24. Hartwick, R.A.; Dezaro, D.D. *Microcolumn High-Performance Liquid Chromatography*; Kucera, P., (ed.); Elsevier: Amsterdam, New York, 1984; Chapt. 3.
25. Meyer, V.R. High performance liquid chromatographic theory for the practitioner. J. Chromatogr. **1985**, *334*, 197–209.
26. Chervet, J.P.; Ursem, M.; Salzmann, J.P. Instrumental requirements for nanoscale liquid chromatography. Anal. Chem. **1996**, *68*, 1507–1512.
27. Karger, B.L.; Martin, M.; Guiochon, G. Role of column parameters and injection volume on detection limits in liquid chromatography. Anal. Chem. **1974**, *46*, 1640–1647.
28. Huber, J.F.K.; Hulsman, J.A.R.J.; Meizer, C.A.M. Quantitative analysis of trace amounts of estrogenic steroids in pregnancy urine by column liquid-liquid chromatography with ultraviolet detection. J. Chromatogr. **1971**, *62*, 79–91.
29. Hawkins, C.; Dolan, J.W. Understanding split peaks. LCGC **2003**, *21*, 1134–1138.
30. Dolan, J.W. But I can't change the method. LCGC **2004**, *22*, 524–528.
31. Alhedai, A.; Martire, D.E.; Scott, R.P.W. Column "dead volume" in liquid chromatography. Analyst **1989**, *114*, 869–875.
32. Colin, H.; Krstulovic, A.; Bounine, J.P.; Guiochon, G. The importance of dead volume in the optimization of separations in reversed phase liquid chromatography using ternary solvents. Chromatographia **1983**, *17*, 209–214.
33. Bakalyar, S.R.; Olsen, K.; Spruce, B.; Bragg, B.G. LCGC **1988**, *6*, 876–882.
34. Bakalyar, S.R.; Olsen, K.; Spruce, B.; Bragg, B.G. How to make tube connections between injectors, columns, and detectors in liquid chromatography. Rheodyne Technical Note *#9*, 1988.
35. Miller, J.M. HPLC troubleshooting experiment: extra-column dead volume. Am. Lab News Ed. **2000**, *32*, 16–17.
36. Dolan, J.W.; Berry, V.V. Liq. Chromatogr. HPLC Mag. **1984**, *2*, 20–21.
37. McWilliam, I.G.; Bolton, H.C. Instrumental peak distortion. I. Relaxation time effects. Anal. Chem. **1969**, *41*, 1755–1762.
38. Cram, S.; Chesler, S.N. Effect of peak sensing and random noise on the precision and accuracy of statistical moment analyses from digital chromatography data. Anal. Chem. **1971**, *43*, 1922–1933.
39. Scott, R.P.W.; Kucera, P. The design of column connection tubes for liquid chromatography. J. Chromatogr. Sci. **1971**, *9*, 641–644.
40. Golay, M.J.E.; Atwood, J.G. Early phases of the dispersion of a sample injected in poiseuille flow. J. Chromatogr. **1979**, *186*, 353–370.
41. Atwood, J.G.; Golay, M.J.E. Dispersion of peaks by short straight open tubes in liquid chromatography systems. J. Chromatogr. **1981**, *218*, 97–122.
42. Thesis, Uwe Neue, University of Saarlandes, 1976.
43. In house studies conducted at Waters Corporation.
44. Tallarek, U.; Leinweber, F.C.; Seidel-Morgenstern, A. Fluid dynamics in monolithic adsorbents: Phenomenological approach to equivalent particle dimensions. Chem. Eng. Technol. **2002**, *25*, 1177–1181.
45. Leinweber, F.C.; Lubda, D.; Cabrera, K.; Tallarek, U. Characterization of silica based monoliths with bimodal pore size distribution. Anal. Chem. **2002**, *74*, 2470–2477.
46. Leinweber, F.C.; Tallarek, U. Chromatographic performance of monolithic and particulate stationary phases. hydrodynamics and adsorption capacity. J. Chromatogr. A **2003**, *1006*, 207–228.
47. Engelhardt, H.; Elgass, H. Optimization of gradient elution, separation of fatty acid phenacyl esters. J. Chromatogr. **1978**, *158*, 249–259.

48. Jandera, P.; Churacek, J. *Gradient Elution in Column Liquid Chromatography*; Elsevier: Amsterdam, 1985; 95–97.
49. Dolan, J.W.; Snyder, L.R. Maintaining fixed band spacing when changing column dimensions in gradient elution. J. Chromatogr. A **1998**, *799*, 21–34.
50. Geng, X.; Regnier, F.E. Retention model for proteins in reversed phase liquid chromatography. J. Chromatogr. **1984**, *296*, 15–30.
51. Vanecek, G.; Regnier, F.E. Variables in the high-performance anion-exchange chromatography of proteins. Anal. Biochem. **1980**, *109*, 345–353.
52. Regnier, F.E. High performance liquid chromatography of bioplolymers. Science **1983**, *222*, 245–252.
53. Lindgren, G.; Lundstrom, B.; Kallman, I.; Hansson, K.A. Physical characteristics and properties of new chromatographic packing materials for the separation of peptides and proteins. J. Chromatogr. **1984**, *296*, 83–95.
54. Moore, R.M.; Walters, R.R. Protein separations on reversed phase high performance liquid chromatography mini columns. J. Chromatogr. **1984**, *317*, 119–128.
55. Englehardt, H.; Mueller, H. Optimal conditions for the reversed phase chromatography of proteins. Chromatographia **1984**, *19*, 77–84.
56. Yang, Y.B.; Harrison, K.; Carr, D.; Guiochon, G. Factors effecting the separation and loading capacity of proteins in preparative gradient elution high performance liquid chromatography. J. Chromatogr. **1992**, *590*, 35–47.
57. Snyder, L.R.; Stadalius, M.A. In *High Performance Liquid Chromatography–Advances and Perspectives*; Horvath, Cs., (ed.); Academic Press: Orlando, FL, 1986; Vol. 4.
58. Snyder, L.R.; Stadalius, M.A.; Quarry, M.A. Gradient elution in reversed phase HPLC separation of large molecules. Anal. Chem. **1983**, *55*, 1412A–1430A.
59. Stadalius, M.A.; Quarry, M.A.; Snyder, L.R. Optimization model for the gradient elution separation of peptide mixtures by reversed phase high performance liquid chromatography. application to method development and the choice of column configuration. J. Chromatogr. **1985**, *327*, 93–113.
60. Dong, M.W.; Grant, J.R.; Larsen, B.R. BioChromatography **1989**, *4*, 19.
61. Alpert, A. (PolyLC Inc.) and Nguyen, R. (Grace Vydac), Private communications.
62. Neue, U.D. *Advances in Chromatography*; Brown, P.R., Grushka, E. (eds.); Marcel Dekker: New York, 2001; Vol. 41, 115.
63. Schellinger, A.P.; Carr, P.W. A practical approach to transferring linear gradient methods. J. Chromatogr. A **2005**, *1077*, 110–119.
64. Snyder, L.R.; Dolan, J.W. *High-Performance Gradient Elution*; Wiley: New York, 2006.
65. Snyder, L.R.; Saunders, D.L. J. Chromatogr. Sci. **1969**, *7*, 195.
66. Stuart, J.D.; Lisi, D.D.; Snyder, L.R. Separation of mixtures of ophthalaldyhyde-derivatized amino acids by reversed phase gradient elution. Accuracy of computer simulation for predicting retention and bandwidth. J. Chromatogr. **1989**, *485*, 657–672.
67. Snyder, L.R.; Dolan, J.W. *Advances in Chromatography*; Brown, P.R., Grushka, E., (eds.); Marcel Dekker: New York, 1998; Vol. 38, 115.
68. Snyder, L.R.; Dolan, J.W.; Grant, J.R. Gradient elution in high performance liquid chromatography. I. theoretical basis for reversed phase systems. J. Chromatogr. **1979**, *165*, 3–30.
69. Snyder, L.R. In. *High-Performance Liquid Chromatography, Advances and Perspectives*; Horvath, C., (eds.); Academic Press: New York, 1980; 207.
70. Stadalius, M.A.; Gold, H.S.; Snyder, L.R. Optimization model for the gradient elution separation of peptide mixtures by reversed-phase high performance

liquid chromatography. Verification of band width relationships for acetonitrile-water mobile phases. J. Chromatogr. **1985**, *327*, 27–45.
71. Neue, U.D. Theory of peak capacity in gradient elution. J. Chromatogr. A **2005**, *1079*, 153–161.
72. Neue, U.D.; Marchand, D.H.; Snyder, L.R. Peak compression in reversed phase gradient elution. J. Chromatogr. A **2006**, *1111*, 32–39.
73. Cheng, Y.; Lu, Z.; Neue, U. Ultrafast liquid chromatography/ultraviolet and liquid chromatography/tandem mass spectrometric analysis, Rapid. Comm. Mass Spectrom. **2001**, *15*, 141–151.
74. Neue, U.D.; Mazzeo, J.R. A theoretical study of the optimization of gradients at elevated temperatures. J. Sep. Sci. **2001**, *24*, 921–929.
75. Cazes, J.; Scott, R.P.W. *Chromatography Theory*; Marcel Decker: New York, 2002.
76. Mills, M.J.; Maltas, J.; Lough, W.J. Assessment of injection volume limits when using on-column focusing with microbore liquid chromatography. J. Chromatogr. A **1997**, *759*, 1–11.

Received September 28, 2006
Accepted November 21, 2006
Manuscript 6980A