Journal of Chromatography, 459 (1988) 43–63
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1453

DESIGN OF OPTIMIZED HIGH-PERFORMANCE LIQUID CHROMATOGRAPHIC GRADIENTS FOR THE SEPARATION OF EITHER SMALL OR LARGE MOLECULES

III*. AN OVERALL STRATEGY AND ITS APPLICATION TO SEVERAL EXAMPLES

B. F. D. GHRIST**
Medical Products Department, E. I. Du Pont de Nemours & Co., Concord Plaza, Wilmington, DE 19898 (U.S.A.)
and
L. R. SNYDER*
LC Resources Inc., 26 Silverwood Court, Orinda, CA 94563 (U.S.A.)

SUMMARY

Recommendations are presented for an efficient approach to the design of optimized gradients for complex samples using computer simulation. Examples based on the separation of polyaromatic hydrocarbon and ribosomal protein mixtures are shown.

INTRODUCTION

The preceding paper[1] described how different sample types (cases I–III) respond to various changes in gradient conditions. Based on that discussion, this paper outlines a general approach to the design of gradients for maximum resolution and/or minimum run time for each sample type, especially case III samples (where the band spacing varies with the gradient conditions).

DESIGN OF OPTIMAL GRADIENTS

Case I samples

This is illustrated in Figs. 1 and 2 in ref. 1, assuming a sample with a molecular weight of about 200. Increasing the gradient time (with other conditions fixed) leads to a progressive increase in resolution, until the average k' value for sample bands ($\bar{k} = 1/1.15b$) equals about 10. A further increase in t_G beyond this point leads to minimal additional improvement in resolution. At the same time, bands broaden in approxi-

* For Parts I and II, see refs. 4 and 1.
** Present address: Eli Lilly Co., Indianapolis, IN 46285, U.S.A.

0021-9673/88/$03.50

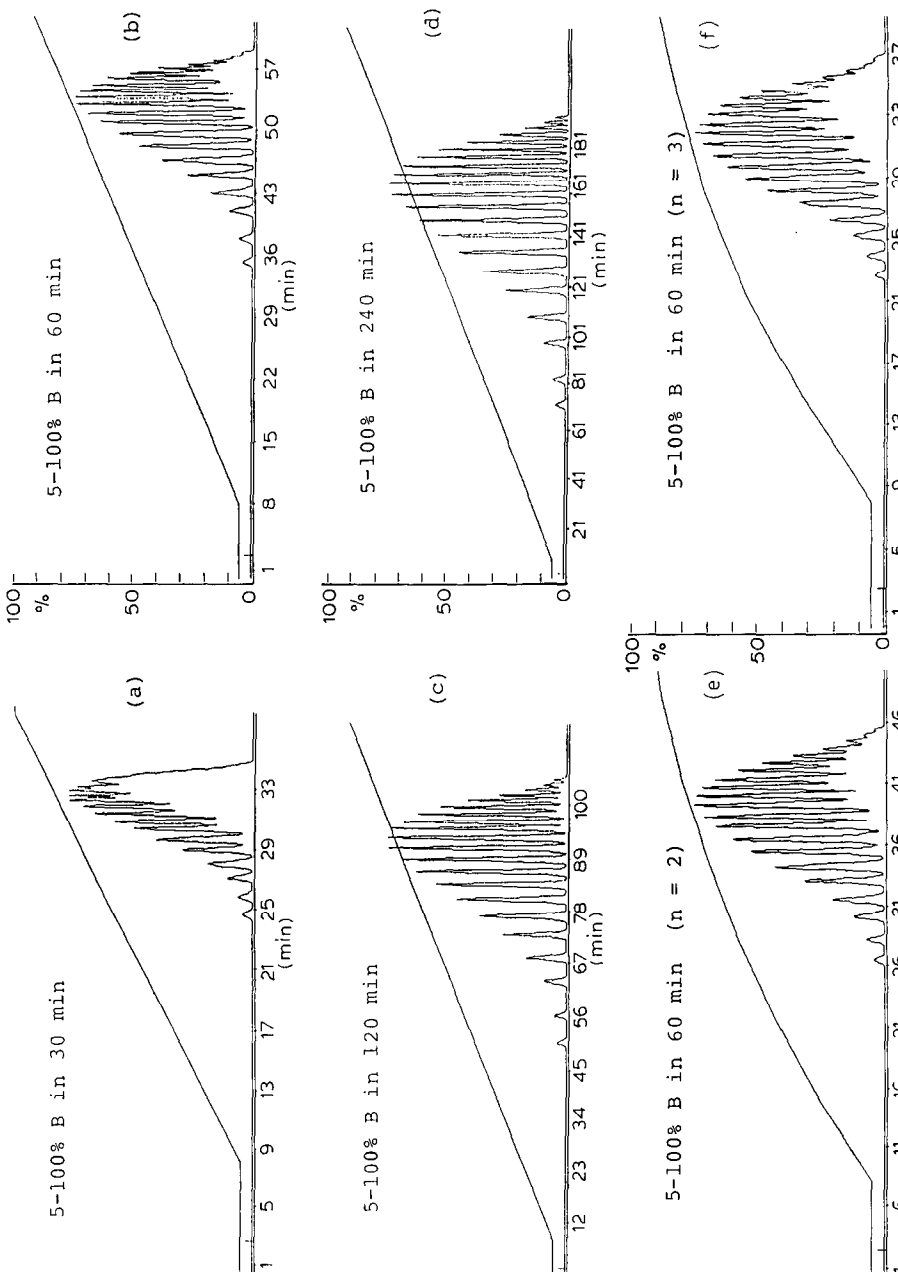


Fig. 1. Effect of gradient time and shape on the separation of a sample described by case II; computer simulations using DryLab G. Conditions: V_D = 5.5 ml; 25 × 0.46 cm I.D. column; flow-rate, 1 ml/min; solute characteristics described by eqn. 3.

mate proportions to t_G, meaning that there will be a significant loss in detection sensitivity for longer gradient times (not shown in Fig. 1 in ref. 1; all peak heights there are normalized to a fixed maximum value). This means that there is an optimal intermediate gradient time, equal to 50–100 min for this example ($\bar{k}$ = 7–14).

For samples that can be classified as case I, there are no reversals of peak position as the gradient time is changed. The effect of changing the starting value of φ (φ_0) while holding the gradient steepness (b) constant is shown in Fig. 2 in ref. 1. If the flow-rate and column dimensions are unchanged, the gradient time t_G must be reduced in proportion to $\Delta\varphi$ for b to remain constant, because

$$b = (V_m S/F)(\Delta\varphi/t_G) \tag{1}$$

Fig. 2 in ref. 1 can be better understood in terms of eqn. 8 in ref. 1:

$$\varphi_e - \varphi_0 = 1/S \tag{2}$$

The value of φ_0 (equal to φ^* for a linear, single-segmented gradient) must be sufficiently different from φ at elution (φ_e) to prevent initial elution of the sample by the starting mobile phase if the separation is to be unaffected by the choice of φ_0. In this instance S = 5, so that $\varphi_e - \varphi_0$ should be greater than 0.20 (20% B).

As an example, consider the resolution of the first two bands of the chromatograms in Fig. 2 in ref. 1. The average value of φ_e for these two bands is about 35% B (see gradient overlay, which represents φ at the column outlet). This means that when φ_0 exceeds 35 − 20 = 15% B, the resolution of the first two bands will be decreased.For values of $\varphi_0 < 0.15$, the relative retention and resolution of the two bands should be unaffected. The resolution R_s of these two bands as a function of φ_0 is as follows: when φ_0 = 0.05, 0.20, 0.30, 0.40 and 0.50, R_s = 5.8, 5.7, 5.5, 4.8 and 2.7, respectively. A significant decrease in resolution is seen to occur for $\varphi_0 > 0.30$. That is, eqn. 8 appears to be conservative for estimating the maximum value of φ_0 for samples of this type (case I). The final value of φ in the gradient has no effect on the appearance of the chromatogram, as long as all bands are eluted before the end of the gradient. This is illustrated by Fig. 2f and d in ref. 1.

Segmented gradients are not usually advantageous for improving the resolution of case I samples. However, segmented gradients can be used to reduce the run time whenever later portions of the chromatogram have much higher resolution (lower density of bands), by using steeper segments for the elution of these bands.

Case II samples

Homologous or oligomeric samples often correspond to case II and show plots of log k' *vs.* φ that extrapolate to about −1.0 for φ = 1. We have therefore carried out computer simulations with DryLab G for a series of hypothetical compounds where

$$S = \log k_w + 1 \tag{3}$$

and values of log k_w change by equal increments (as in Fig. 3b in ref. 1). The effect of gradient time on the separation of the latter sample is shown in Fig. 1a–d. Sample resolution increases regularly with increase in t_G, as in the simulations in Fig. 1 in ref. 1 for a case I sample. However, the chromatograms in Fig. 1a–d show initial bands that

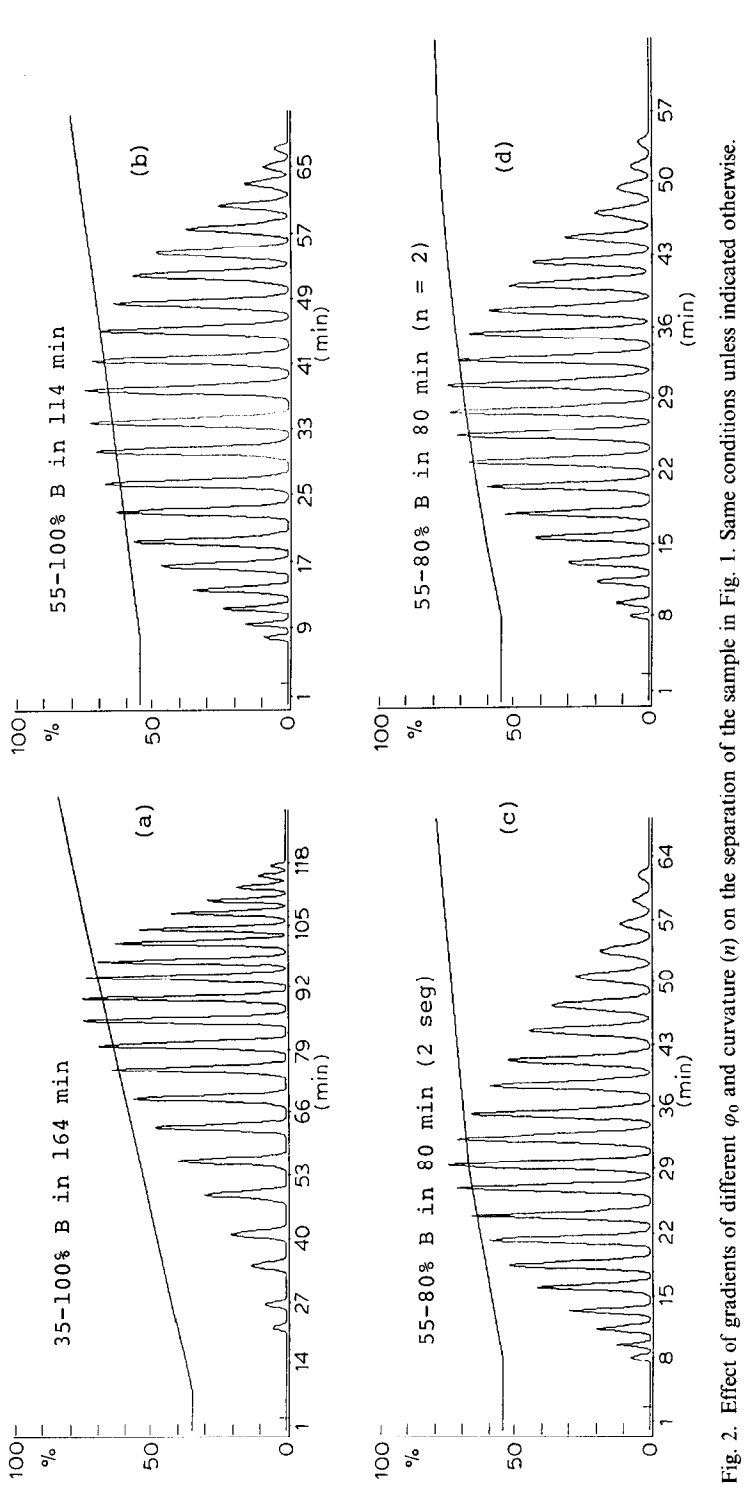


Fig. 2. Effect of gradients of different φ_0 and curvature (n) on the separation of the sample in Fig. 1. Same conditions unless indicated otherwise.

are overresolved, even for fairly short gradient times, whereas later bands are only marginally resolved even for very long gradients (240 min, Fig. 1d).

For examples of this type it is generally recommended to try a convex gradient. Such gradients are usually described by equations of the form

$$\varphi = 1 - [1 - (t/t_G)]^n \tag{4}$$

However, it is seen in Fig. 1e and f that the use of convex gradients (of varying convexity; $n = 2$ and 3) with a gradient time of 60 min does *not* improve the separation obtained by a linear gradient in the same time (Fig. 1b). The reason is that the initial steep portion of these convex gradients rises to a high value of φ before the sample can be eluted from the column. The resolution of the sample is in turn determined by the values of φ_e (or k_f) for each band pair.

Fig. 2a and b show further separations of the sample in Fig. 1, where the value of φ_0 is varied while the gradient steepness (value of b) is maintained the same as in the separation in Fig. 1d ($t_G = 240$ min). The use of $\varphi_0 = 55\%$ B (Fig. 2b) yields similar resolution at the two ends of the chromatogram, so as to equalize the resolution considerably over the entire chromatogram.

Now we can test the effects of a convex gradient for a similar gradient range and steepness. In Fig. 2c a two-step convex gradient is used, with a modest improvement in resolution for the end of the chromatogram and further equalization of overall resolution. The separation based on this two-step gradient is little different from that obtained with a corresponding continuous convex gradient as defined by eqn. 1 (Fig. 2d).

We can draw the following conclusions from the simulations in Figs. 1 and 2:

(1) Samples corresponding to case II show a continuous improvement in resolution as the gradient time is increased, especially for later eluted bands; however, the gradient time required for adequate resolution of the final bands in the sample may be prohibitive, if a linear gradient from 5 to 100% B is used.

(2) The most effective strategy is to increase φ_0 to the point where the resolution of initially eluted bands is no greater than that of bands eluted near the end of the chromatogram; this can be combined with an increase in t_G so as to increase the resolution of both early and later eluted bands to an acceptable level.

If the gradient time is excessive at this point, or resolution is still inadequate, either a two-segment gradient or a convex gradient (eqn. 1 with $n = 2$) can be used to advantage; however, the resulting improvement in separation may be minor.

Although it is also of interest to consider gradients of increasing curvature (increased n values in eqn. 1), it appears that increasing n is really equivalent to decreasing t_G, with little other effect on the separation. Thus a gradient for $n = 3$ has virtually the same shape as a gradient for $n = 2$, if t_G for the $n = 3$ gradient is increased 1.3-fold. Therefore, the separation resulting from a gradient with $n = 3$ should be similar to that obtained with a gradient where $n = 2$ and the gradient time is 1.3-fold shorter. This is shown in Fig. 3 for the same sample as in Figs. 1 and 2. The resulting retention times and resolution are essentially the same in the two instances (each gradient could be terminated at 46 min, when all bands have left the column).

It appears from the preceding examples and discussion that (a) the use of convex gradients is generally of little advantage compared with two-segment gradients of

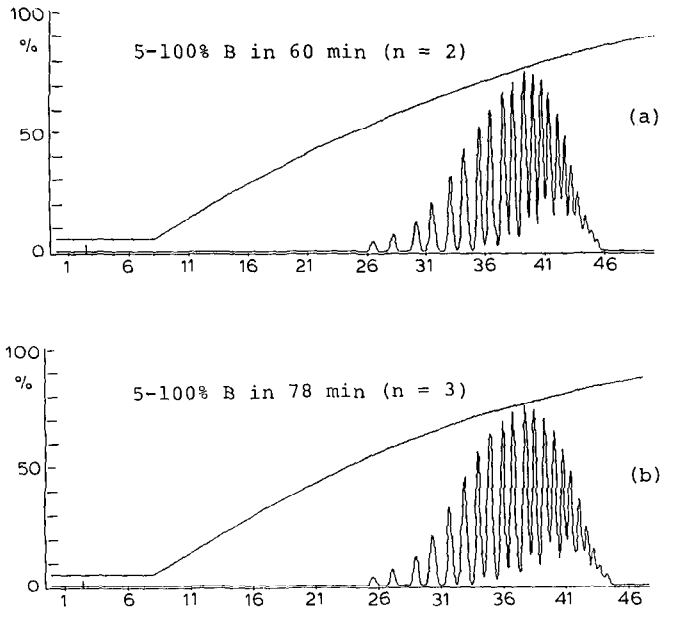


Fig. 3. Effect of starting %B and gradient shape on the separation of the sample in Fig. 1. Same conditions unless indicated otherwise.

similar gradient range and time* and (b) there is no advantage in changing the curvature of convex gradients (*i.e.*, using different values of n in eqn. 1). This considerably simplifies our study of gradient shape as it relates to separation, particularly for case II samples. Our further discussions can therefore focus on segmented gradients, as opposed to the continuously curved gradients described by eqn. 1.

Case III samples: example of sixteen polyaromatic hydrocarbons

We shall now apply some of the principles discussed above to the design of optimal gradients for two different samples that have been described previously: (a) a sixteen-component sample of polyaromatic hydrocarbons (PAHs) and (b) a twenty-component sample of 30S ribosomal proteins. Because of the practical differences that result for samples of differing molecular weight, it will prove useful to compare and contrast these two examples.

The PAH sample was described in ref. 2, but with little discussion of how an optimal gradient was obtained. The two initial runs used as experimental inputs for computer simulation are shown in Fig. 4a and b (computer simulations; original chromatograms shown in ref. 2). Preliminary examination of these two chromato-

* The latter conclusion concerning curved gradients (*i.e.*, they are not very useful) should be qualified for oligomeric mixtures having a wide range of molecular weight. The sample used in Figs. 1–3 corresponds to a mixture with a maximum molecular weight of less than 1000. As the sample molecular weight (and corresponding values of S) become larger, the advantage of segmented or curved gradients may increase.

grams shows that there are three critical band pairs: 3–4, 9–10 and 14–15. Values of R_s (for a 10 000-plate column) are superimposed on the critical band pairs in each chromatogram. Usually the best first step in computer simulation is to examine a relative resolution map (RRM) as a function of gradient time t_G. This is shown in Fig. 4c, and it appears that a gradient time of *ca.* 20 min is optimal.

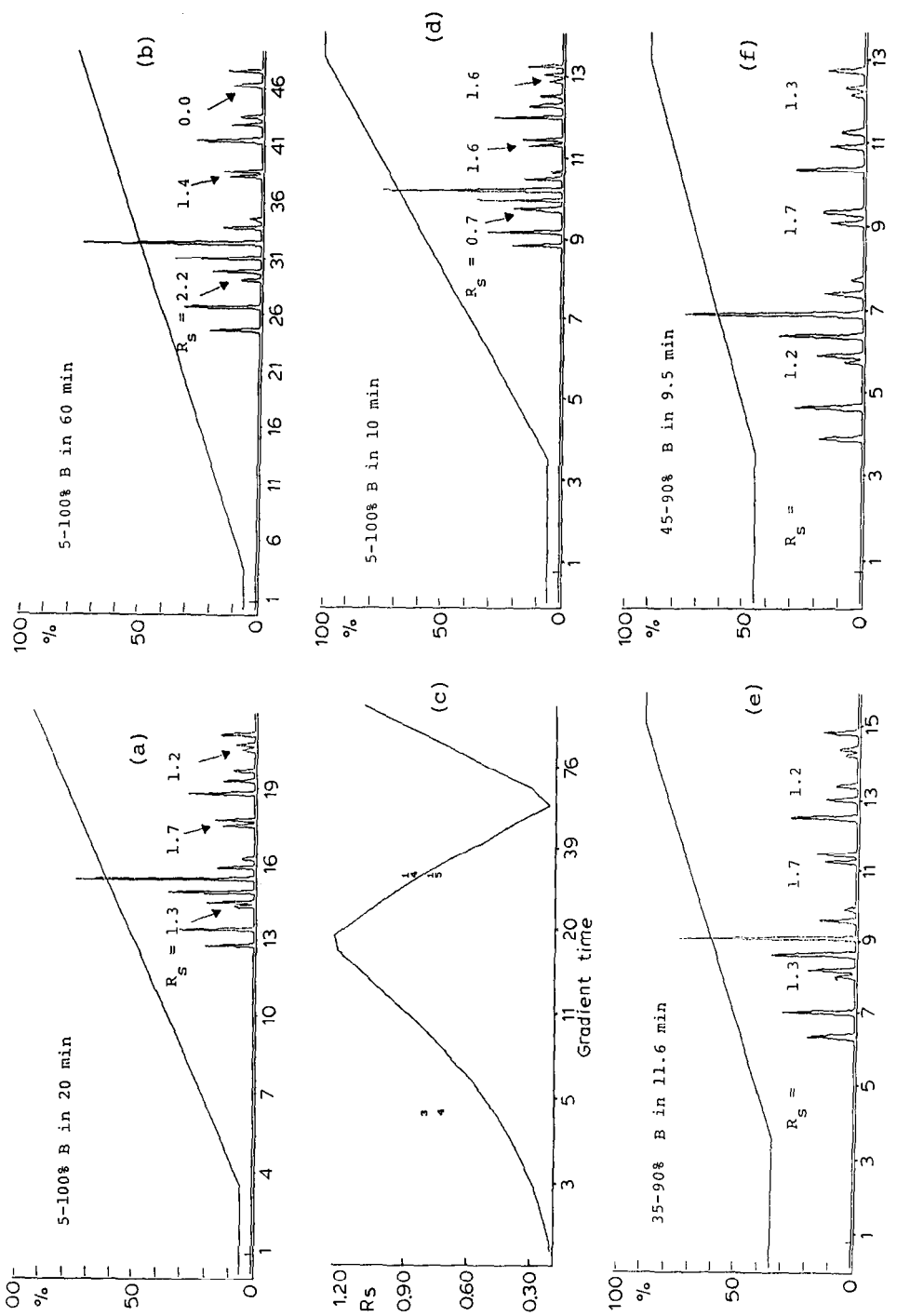
The next step is to adjust the gradient range for maximal resolution in the minimal time. This can be done by trial-and-error, using computer simulation. Fig. 4e–g illustrate this process, with the conclusion that a 45–90% gradient in 9.5 min is probably near optimal. Alternatively, DryLab G will recommend an appropriate gradient range: 42–94% B in this instance, with the same gradient steepness (t_G = 11 min). At this point (Fig. 4f) it is seen that a minimum resolution of R_s = 1.2 can be achieved, in a time of 9.5 min per sample (the total analysis time is closer to 19 min if column re-equilibration is taken into account[4]).

Returning to Fig. 4a and b, it is seen that bands 3 and 4 are better resolved at higher values of t_G (60 min), whereas bands 14 and 15 are better resolved at low values of t_G. This suggests that a shallow gradient to elute bands 1–4, followed by a steeper gradient to elute the remainder of the sample, might prove better than a simple linear (unsegmented) gradient in this instance. A further example of these trends in resolution with gradient time is shown in Fig. 4d for t_G = 10 min.

It is usually instructive at this point to examine the resolution of critical band pairs as a function of gradient time, using computer simulations. Table I summarizes data for the PAH sample obtained in this way. The resolution of band pair 3–4 continues to increase with increasing t_G, the resolution of band pair 9–10 is more or less independent of t_G and the resolution of band pair 14–15 is a maximum for a gradient time of about 10 min (9.5%/min).

A good strategy for designing an optimal gradient is to first optimize the gradient steepness for maximal resolution of the front end of the chromatogram. This would suggest an initial gradient steepness that is as low as possible (Table I). However, this turns out to be wrong for two reasons. First, a very shallow gradient for the elution of bands 3 and 4 will adversely affect the resolution of bands 9 and 10 and 14 and 15, even though a steeper (second) segment is used to elute these compounds. This can be seen as follows. Fig. 8 in ref. 1 suggests that the second segment must begin about 25% before the elution of bands 9 and 10 or 14 and 15 if their separation is not to be affected by the first (shallow) segment. However, bands 3 and 4 elute only 15% earlier than bands 9 and 10.

A second reason not to make the initial gradient so shallow is that there is usually little point in increasing the resolution of band pair 3 and 4 beyond the maximal resolution that can be obtained for bands 9 and 10 and 14 and 15 (R_s = 1.6). This suggests a starting gradient (first segment) with a steepness of about 3%/min and a second segment with a steepness of about 10%/min. As bands 3 and 4 are eluted in the initial gradient with φ_e = 50%, this suggests a gradient of 5–52% B in 15 min and 52–100% B in 5 min. Fine tuning of this gradient (by trial and error) then resulted in the chromatogram shown in Fig. 4h. Finally, this gradient was shortened by starting at a higher value of φ_0 (trial-and-error simulations), giving the final separation of Fig. 4i. Further trial-and-error modification of the gradient was attempted without a significant further increase in resolution. A separation similar to that in Fig. 4i was confirmed


5-100% B in 20 min
(a)
R_s = 1.3
1.7
1.2
5-100% B in 60 min
(b)
R_s = 2.2
1.4
0.0
5-100% B in 10 min
(d)
R_s = 0.7
1.6
1.6
45-90% B in 9.5 min
(f)
R_s =
1.2
1.7
1.3
35-90% B in 11.6 min
(e)
R_s =
1.3
1.7
1.2
(c)
Rs
Gradient time
%

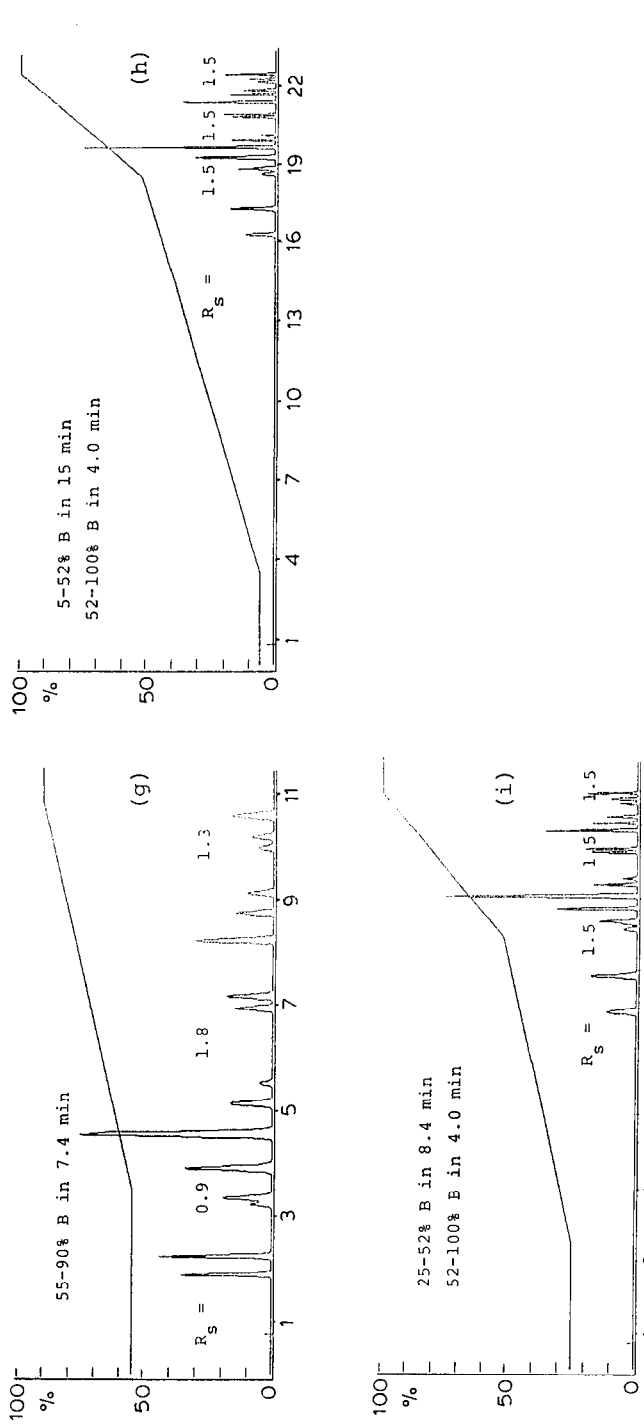


Fig. 4. Separation of a sixteen-component PAH sample as a function of gradient conditions. Computer simulations (DryLab G) based on experimental data in ref. 2.

TABLE I

RESOLUTION OF CRITICAL BAND PAIRS FOR A PAH SAMPLE AS A FUNCTION OF GRADIENT TIME

Other conditions given in ref. 2.

t_G (min)	R_s (N = 10 000) for indicated band pair (%/min)*			
	3–4	9–10	14–15	% B/min
5	0.3	1.4	1.4	19
10	0.7	1.6	1.6	9.5
20	1.3	1.7	1.2	4.7
40	1.8	1.6	0.5	2.4
60	2.2	1.4	0.0	1.6

* Change in %B per minute (gradient steepness).

experimentally in ref. 2. The separation in Fig. 4i exhibits a 25% increase in minimal resolution (R_s = 1.5) compared with the run of Fig. 4f (R_s = 1.2)*.

Case III samples: example of twenty 30S ribosomal proteins

Editing initial experimental runs. Two experimental runs were carried out initially, as shown in Fig. 11a and b in ref. 1. Individual bands were matched between the two runs as described in refs. 2–4. Comparison of a simulated run as in Fig. 14d in ref. 1 with the corresponding experimental run (Fig. 14b in ref. 1) can be used to confirm that bands in run 2 have been matched to those in run 1, prior to inputting final retention times into the computer**. For this example, there is a close similarity between the experimental and simulated chromatograms (Fig. 5b and d), confirming the accuracy of peak matching.

A similar comparison for run 1 may also be useful for picking up any errors in retention time (it is assumed that band areas from run 1 were entered into DryLab G). This is illustrated in Fig. 5 for the same chromatograms as in Fig. 11a and c in ref. 1. The initial computer simulation (Fig. 5b) differs in minor respects from the experimental chromatogram (Fig. 5a). Minor adjustments in the retention times of three bands (by 0.1–0.2 min) results in a better match (Fig. 5c) with Fig. 5a.

The further correction of the input data for run 1 (as in Fig. 5c *vs.* 5b) is not essential to accurate computer simulation, but it can have a favorable impact. The errors in Fig. 5a arise from band overlap (see discussion in ref. 4), and in such instances it is advisable to correct the value of t_g for the overlapped band (*i.e.*, so as to improve the accuracy of the predicted resolution of run 1).

The correction of minor errors in retention time (as described above) becomes more important as the sample complexity increases. As will be seen, samples such as the 30S ribosomal proteins can be difficult to resolve completely, and this places additional emphasis on the accuracy of computer simulation.

* The run time is greater in Fig. 4i (12.4 min) than in Fig. 4f (9.5 min), but the separation in Fig. 4i is still superior; adjustment of the column plate number (flow-rate, column length) can be used to normalize the run time, which then provides greater resolution for the conditions in Fig. 4i than those in Fig. 4f.

** A useful first step in computer simulation is to adjust the resolution of predicted chromatograms so as to match one of the starting experimental runs. This was done in this instance by using the 240-min run (Fig. 5b and d). The resulting value of N for the experimental column corresponded to 2500 plates.

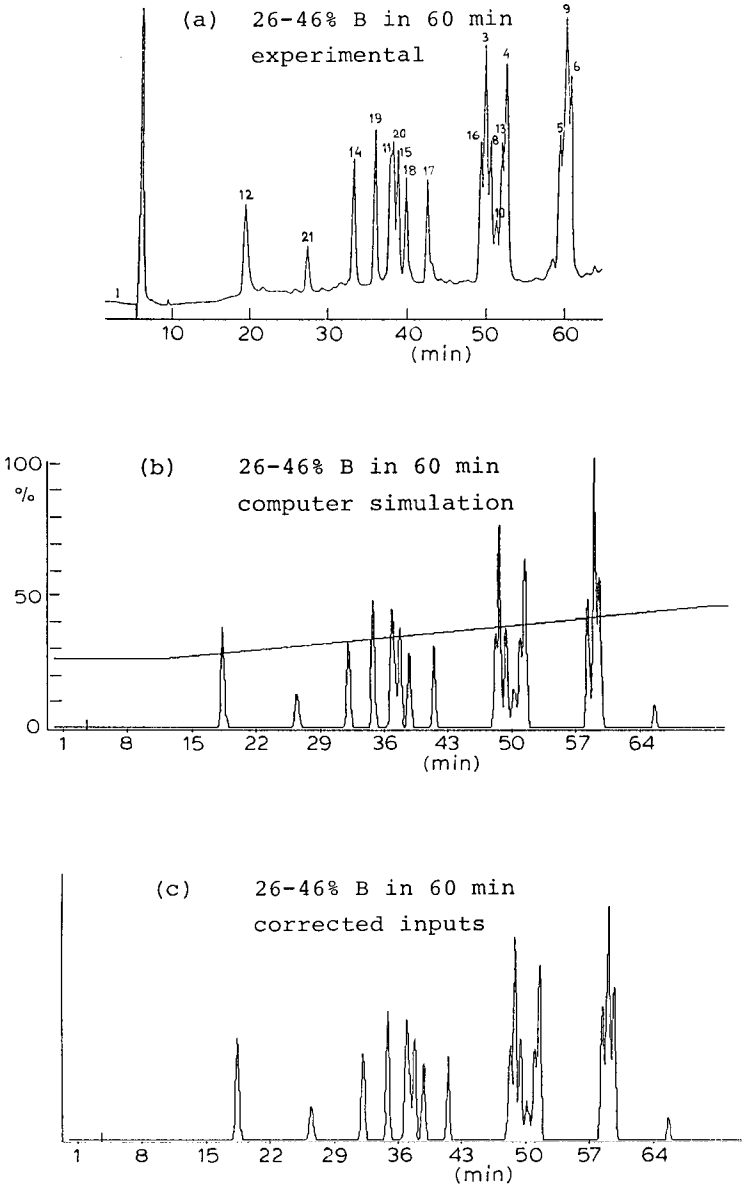


Fig. 5. Use of corrected ribosomal protein retention times for a 60-min run to improve computer simulations. (a) Experimental run; (b) simulation based on retention time data from (a); (c) simulation based on empirically corrected retention times. Conditions as in Fig. 14 in ref. 1 except where indicated otherwise.

Subdividing the chromatogram. A good initial step for any sample is to examine a relative resolution map as in Fig. 4c for the PAH sample. Fig. 6a shows such a map (RRM) for the ribosomal protein sample (all 20 bands, $N = 2500$). A maximum resolution of $R_s = 0.7$ is possible, but only for a long gradient time ($t_G \approx 8$ h). This suggests that a segmented gradient may be preferable.

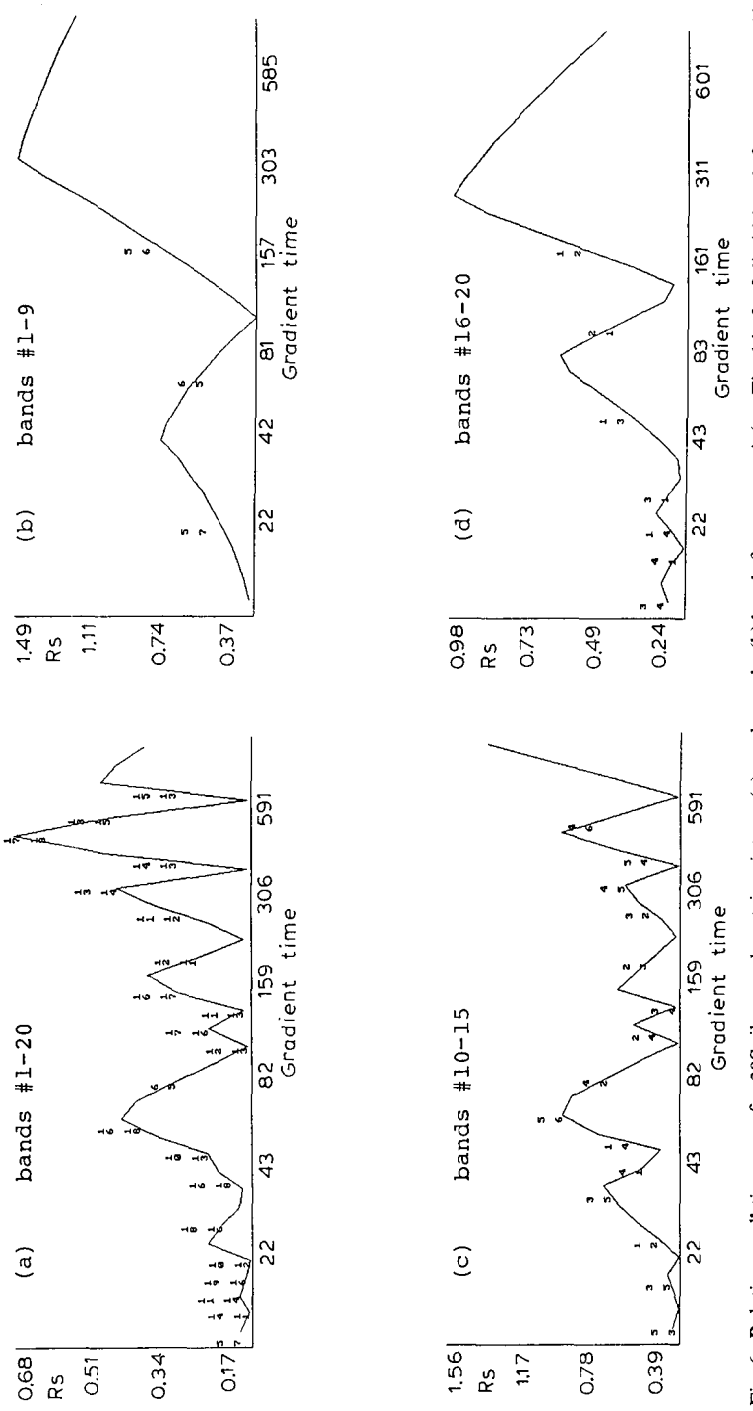


Fig. 6. Relative resolution maps for 30S ribosomal protein mixture. (a) total sample; (b) bands from group A (see Fig. 14 of ref. 1); (c) bands from group B; (d) bands from group C. Conditions as in Fig. 5 except where indicated otherwise.

For complex samples such as this, we should attempt (if possible) to divide the chromatogram into groups of adjacent bands. As seen in Fig. 11b and d in ref. 1 (t_G = 240 min), there are three such groups of bands, A, B and C, consisting of bands 1–9, 10–15 and 16–20, respectively. There are no crossovers of bands between these three groups; as a first step, we can therefore treat each group separately. Having grouped the sample bands in this way, the next step is to look at RRMs for each band set.

Fig. 6b–d show the RRMs for groups A–C. A minimum resolution is predicted of R_s = 1.5 for group A, 1.5 for group B and 1.0 for group C. This suggests that we should be able to achieve a minimum resolution of R_s = 1.0 for the entire sample by optimizing the gradient segments for each group. We shall see that this is possible. The (predicted) optimal separation of each group is shown in Fig. 7a, c and d. Fig. 7b shows a sub-optimum for group B (shorter gradient) that we shall examine later. The very long gradient time suggested for group B (1200 min, Fig. 7c) need not be a problem, inasmuch as the actual gradient segment can be very much shorter than for the full 26–46% B range (about 150 min, as seen in Fig. 7c).

Designing the gradient. The best strategy is usually to optimize each gradient segment sequentially, beginning with the first group of compounds (bands 1–9 in this instance). Reference to Fig. 7a suggests a gradient of 20% per 350 min or 0.057%/min for this segment, with the segment ending at about φ^* = 33% (the φ_e value of band 9). Values of φ_e can be determined from the retention time t_g of the band. For a linear gradient

$$\varphi_e = \varphi_0 + \Delta\varphi[(t_g - t_D - t_0)/t_G] \tag{5}$$

For band 9, t_g = 141 min and $t_D + t_0$ = 11.5, so that φ_e = 26% + [(129.5/350) · 20] = 33%.

Similar calculations can be used to determine φ_e for segmented gradients, or values of φ_e can be read (approximately) from chromatograms such as that in Fig. 7, which have a gradient overlay that is corrected for t_D and t_0. The first segment is therefore calculated as 26–32.6% B in 115 min (φ^* = 32.6% was fine-tuned during the addition of an additional segment to the gradient; see below). The resulting separation will be the same as in Fig. 7a, as neither φ_0 nor the gradient steepness (b) have been changed.

The next step is to optimize the second gradient segment for group B compounds. The optimal gradient in Fig. 7c for this group of bands suggests a gradient steepness of 20% per 1200 min = 0.0167% B/min, *e.g.*, 32.6–46% B in 800 min. Use of this second gradient segment (in combination with the first segment above) gave a separation that is inferior to that observed in Fig. 7c, because of the influence of the first segment on the separation of bands eluted by the second segment. As discussed earlier, this can generally be corrected for by varying the slope of the second segment.

Trial-and-error adjustments yielded an optimal slope of 0.0136% B/min for the second segment, and this two-segment gradient gave a resolution of bands 1–15 that was equivalent to that provided in Fig. 7a and c. Band 13 (the last eluted band of group B under these conditions) is eluted with a retention time of 310 min and a φ_e value of 36% B. Correcting for the effects of t_D and t_0 on the gradient at the outlet of the column, this suggests a second gradient segment of 32.6–35% B in 177 min. The

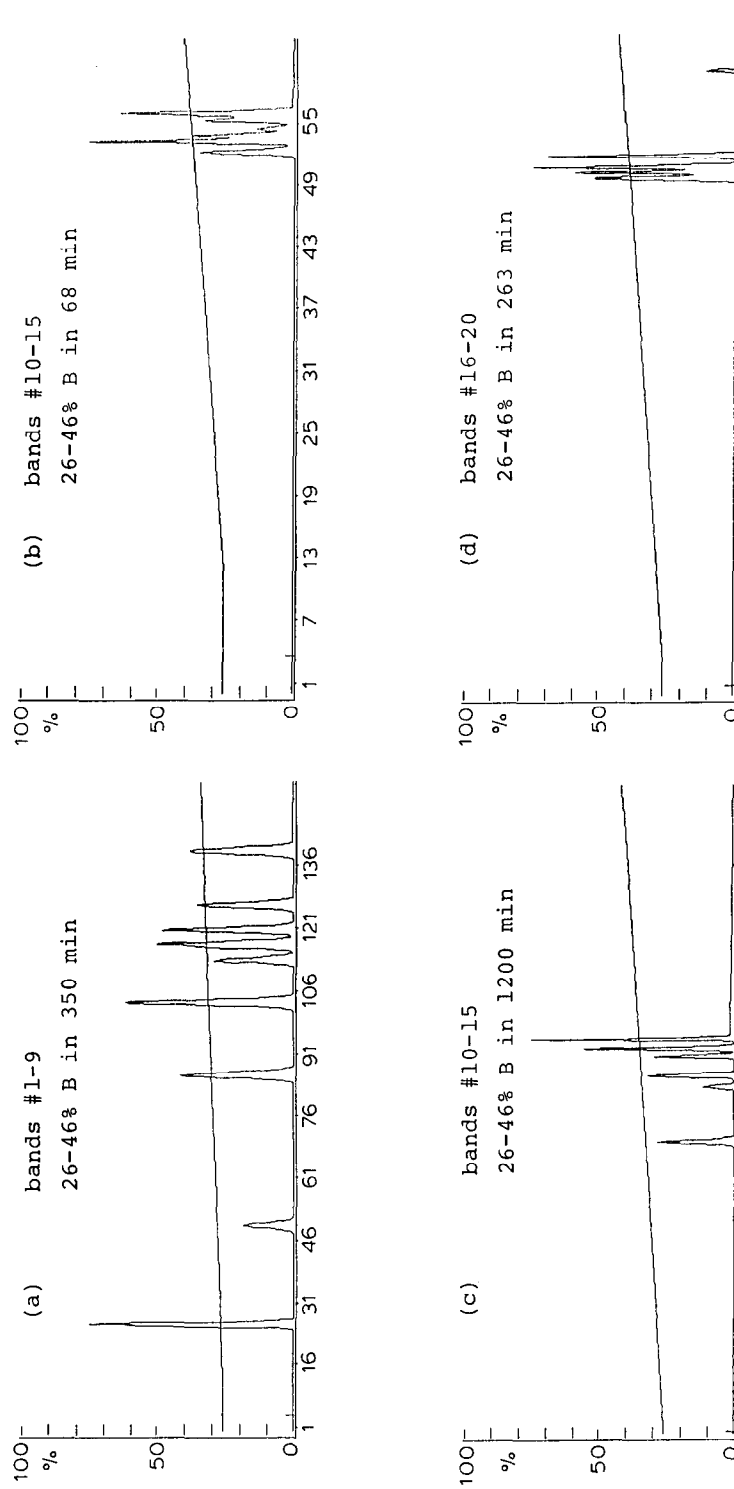


Fig. 7. Separation of various ribosomal protein band groups using optimized gradient times for 26–46% B gradients. Conditions as in Fig. 5 except where indicated otherwise.

resolution of group B (R_s = 1.2) obtained with this segment (in combination with the first segment) was close to that obtained in Fig. 7c (R_s = 1.5). Fine tuning of the gradient at this point did not result in any improvement of this two segment separation.

The optimal gradient slope for the final segment (to elute group C) is predicted from Fig. 7d to be about 0.076%. The addition of this segment to the first two again

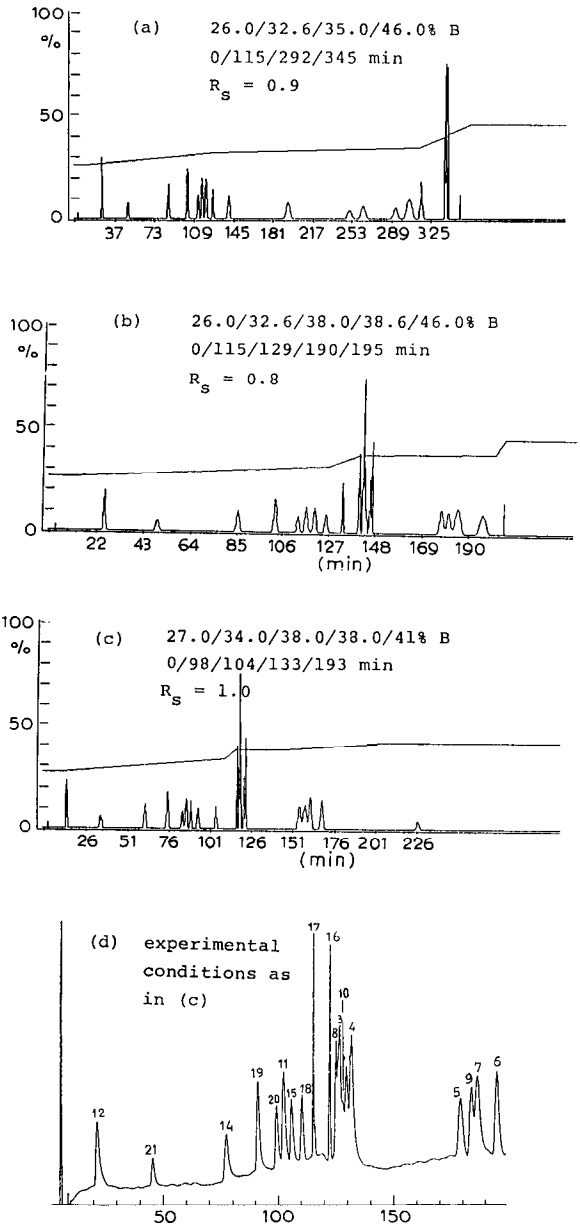


Fig. 8. Optimized separation of 30S ribosomal proteins using different gradients. See text for details.

gave a separation of group C that was inferior to that in Fig. 7d. However, fine tuning yielded an optimal gradient steepness of 0.21%/min, and this three segment gradient did provide a resolution of group C (R_s = 0.9) which is close to that predicted in Fig. 7d (R_s = 1.0).

The final three-segment gradient gave an overall minimum resolution of R_s = 0.9 (Fig. 8a) which is close to that inferred from Fig. 7 (R_s = 1.0). The separation in Fig. 8a is seen to be considerably better than is possible with any single-segment gradient (Fig. 6a, minimal resolution R_s = 0.7). The separation time (345 min total) is also less than that required for the best single-segment gradient (t_G = 500 min). The effort required to design this particular gradient (apart from the initial two experimental runs) amounted to a few hours of computer time. A similar procedure based on trial-and-error experimental runs would be impractical, probably requiring several weeks.

Elsewhere, we have reported the separation of all twenty components of this sample, using a complex four-segment gradient[4]. This gradient was developed by trial and error (computer simulations), before the recommendations arrived at in this study were implemented. It is interesting to see that this latter gradient (Fig. 8c) is actually superior (minimal R_s = 1.0) to that in Fig. 8a, and requires a much shorter time. This suggests that an experienced chromatographer can use (many) trial-and-error computer simulations to arrive at good final gradients. However, the systematic approach outlined above requires less effort (and experience) than the more empirical approach that resulted in the separation in Fig. 8c.

Comparison of the predicted separation in Fig. 8c with the actual chromatogram in Fig. 8d for these conditions shows only fair overall agreement. The observed minimal resolution was about R_s = 0.8 (*vs.* 1.0 predicted), and the retention times of earlier bands differ significantly, between experimental and predicted chromatograms. These errors in prediction are believed to be due to (a) rounding of the gradient by the equipment and (b) changes in column retention between the initial and final experimental runs. Attention to the recommendations in ref. 4 should result in better agreement in other instances.

Example of 32 50S ribosomal proteins

A similar approach to the above was next tried for the separation of the 32 50S ribosomal proteins, an even more complex sample. Details of this work will be described elsewhere, but Fig. 9 compares the predicted (optimal) separation using a four-segment gradient with the final experimental chromatogram*. As discussed in ref. 5, the agreement between these two chromatograms is good ($\pm$ 0.5% in t_g, $\pm$ 8% in R_s). The conditions for Fig. 9 allow the separation of 31 of the 32 proteins in this sample with a minimal resolution of about 0.8 (one band pair cannot be resolved with this column, regardless of the gradient conditions).

Other observations

Evaluating the chromatogram. The evaluation of a chromatogram during the design of a gradient by computer simulation is important to decisions made in the inevitable trial-and-error fine tuning of the gradient at each step. Some workers will

* Fig. 9b shows only the confirmed 50S ribosomal protein bands, which are numbered in Fig. 9c. Additional bands in Fig. 9c correspond to other compounds that were ignored in this study.

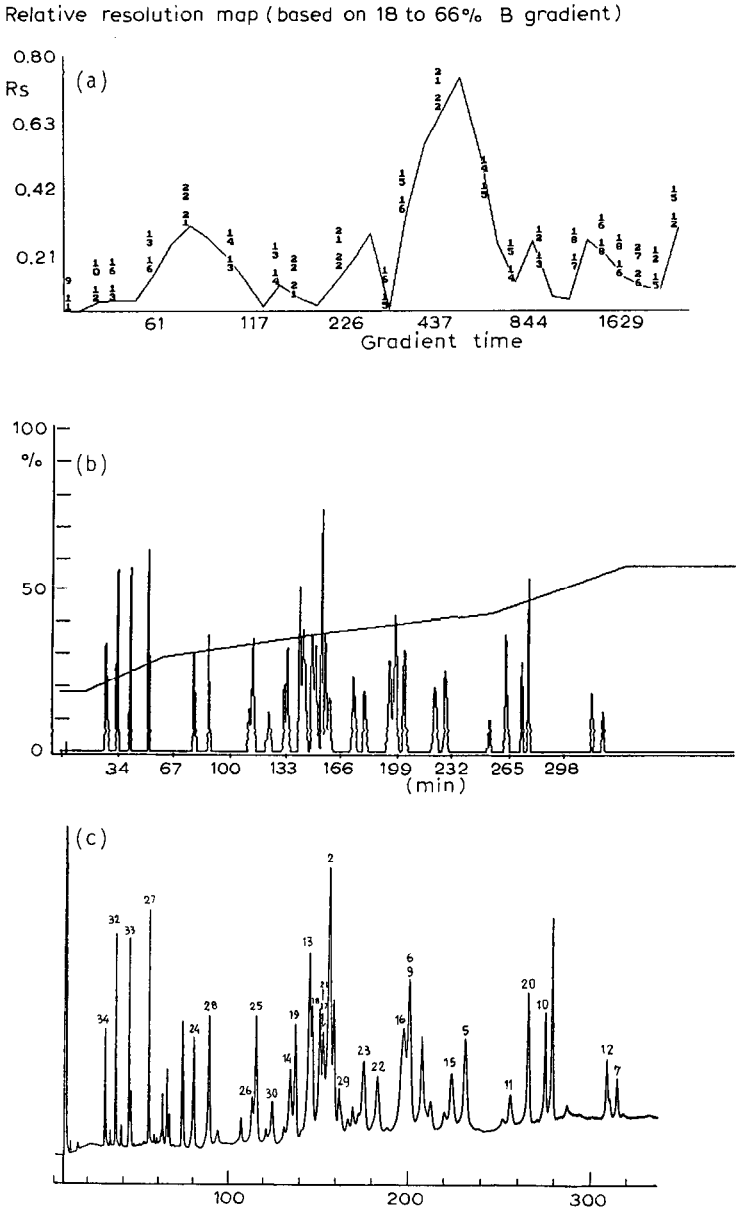


Fig. 9. Optimized separation of 50S ribosomal protein sample. (a) Resolution map; (b) simulated chromatography; (c) experimental chromatogram. Conditions: 18, 29, 37, 43, 58% B in 0, 46, 142, 241, 320 min; other conditions as in the Experimental section in ref. 4. $N = 900$.

prefer to use tables of R_s values for each band in the chromatogram (these are provided by DryLab G for each run, *e.g.*, as in Table I). Changes in the gradient usually lead to an increase in R_s for one or more critical bands and a decrease in resolution of other critical bands. Often the best choice of gradient will equalize the R_s values of two

critical band pairs, and tables of R_s values therefore inform us (a) in which direction a particular gradient should be changed for best results and (b) how close we are to an optimal result.

Other workers may prefer to examine the actual chromatograms that are predicted for each gradient to be evaluated. However, complex samples such as this often make visual interpretation of the chromatogram difficult. Overlapping bands combined with the low resolution of the computer screen can lead to generally confusing data. One means of improving this situation, which is not available experimentally, is to increase the predicted resolution of each separation by some large and arbitrary amount. This is illustrated in Fig. 10 for the group B bands (Fig. 11 of ref. 1) of the 30S ribosomal protein sample above. Here, the effect of gradient time (or steepness) on the separation (26–46% B gradients) is examined for a plate number of 160 000; these separations have a resolution that is eight times that observed experimentally. In the examples in Fig. 10 it is possible to see clearly changes in relative band position as the gradient steepness is changed; band 4 (marked with an asterisk) migrates from position 1 to position 4 as the gradient time is changed from 30 to 240 min. Bands 2 and 3 are initially well resolved (t_G = 30 min), but gradually approach each other and merge at a gradient time of 240 min.

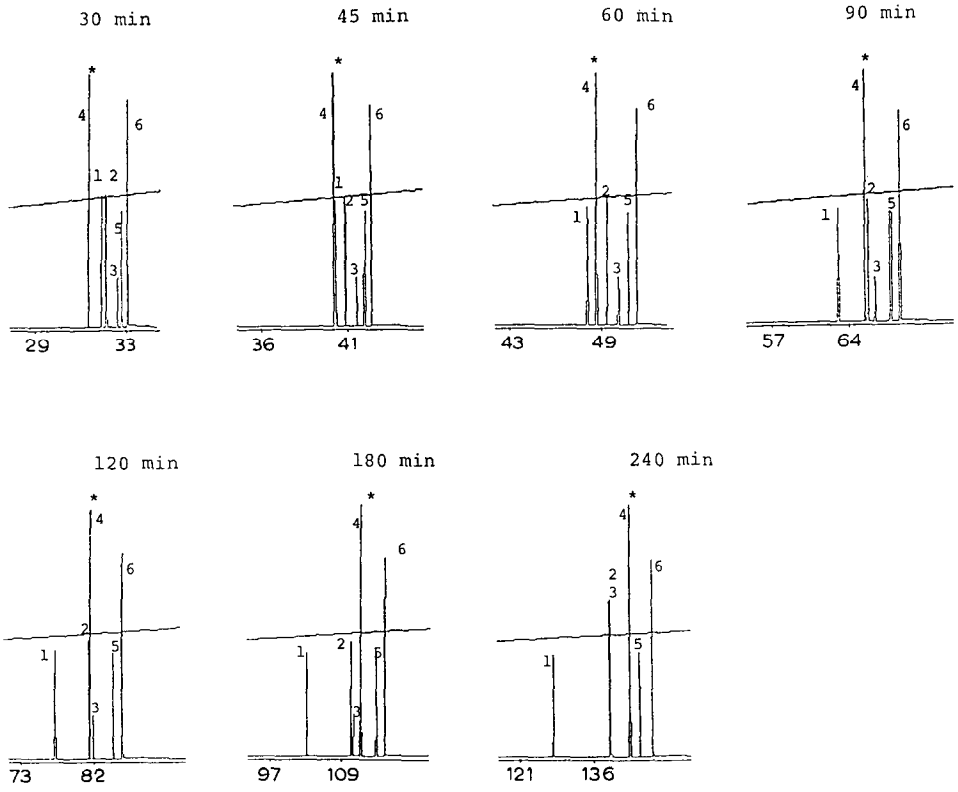


Fig. 10. Use of high-resolution computer simulations to track bands as the gradient time is varied. Ribosomal protein sample; conditions as in Fig. 5 except where indicated otherwise.

Alternative gradients. The examples in Fig. 8 suggest that more than one approach to gradient optimization may often be successful. In fact, the separation of a given sample such as this can usually be achieved (with similar resolution and in a similar time) by an almost endless number of gradients. For example, Fig. 6c suggests that group B can be separated with nearly adequate resolution (minimal $R_s = 0.8$) in a much shorter time ($t_G = 68$ min), as shown in Fig. 7b. This suggests adjusting the steepness of the second segment to a value of about 0.3% B/min, instead of the value of 0.017% B/min suggested by Fig. 7c. A similar approach to that for the development of the gradient in Fig. 8a yielded the predicted separation in Fig. 8b. An overall minimum resolution of $R_s = 0.8$ (*vs.* $R_s = 0.9$ in Fig. 8a) was obtained, in about half the total time (195 *vs.* 345 min). For some applications, this separation might prove preferable to that in Fig. 8a.

The considerable effort spent in designing the separation in Fig. 8a may not be necessary in all instances. The data in Fig. 6 lead immediately to the predicted separations of each group in Fig. 7. Depending on what compounds in the sample are of interest, any of these group separations might have been suitable for the desired application. However, with a little additional effort (*ca.* 1 h), a single separation (Fig. 8a) can provide equivalent results for the total sample.

Recommendations for designing an optimal gradient by computer simulation

Our experience in applying computer simulation to a number of different examples (described here and elsewhere) is summarized below. It is assumed that (a) the inital experimental data have been collected in such a way as to maximize the accuracy of computer predictions of separation, (b) the data have been entered into the computer correctly (bands matched between two initial runs) and (c) any discrepancies between the experimental and corresponding simulated chromatograms have been addressed (as in the example in Fig. 5). That is, we should first do everything possible to insure the success of computer simulation for the sample at hand. The individual computer simulation steps required to optimize the gradient conditions are as follows (assumes the use of DryLab G or equivalent software):

(1) First, examine a resolution map for the entire sample, *e.g.*, as in Fig. 4c for the PAH sample. Many samples (especially those with fewer than fifteen components) can be adequately resolved by using a linear (unsegmented) gradient with an optimal steepness or value of t_G. Other samples (case II) where the resolution decreases continuously from the beginning to the end of the chromatogram can be handled similarly, except that a two-segment gradient may provide marginally better resolution.

(2) Once an optimal gradient steepness has been chosen, trim the gradient range (reduce $\Delta\varphi$) to save time. This can be done by trial and error (as in Fig. 2 in ref. 1) or DryLab G will provide specific recommendations. Often it is necessary to re-optimize gradient time after the final gradient range is selected (repeat step 1).

(3) For more demanding samples, such as the case III examples described here, a segmented gradient will often be preferred. There are two possible approaches, illustrated by the PAH and ribosomal protein samples. Which approach should be followed depends on an initial examination of the experimental chromatograms. One approach (as in the 30S ribosomal protein example) is to divide the sample into distinct groups, *e.g.*, A–C in Fig. 11d in ref. 1, if this is possible. Alternatively (as with the PAH

sample), one can see whether there are a small number of critical band pairs in the sample. These can be identified initially on the basis of the RRM (Fig. 4c) and by looking for other marginally resolved band pairs (when t_G is varied). In the PAH example, band pairs 3–4, 9–10 and 14–15 were selected in this way.

(4) For samples such as the PAHs, it is recommended to first carry out computer simulations for those bands (only) prior to and including the first critical band pair. The gradient range and steepness for these compounds can then be optimized. In this connection, the use of an RRM (for the bands in question) can save time. This step is similar to step 1.

(5) A second segment can be added next, just following elution of the first critical band pair. The steepness of this segment can be varied in trial-and-error fashion to optimize the resolution of the second critical band pair. This procedure can be repeated as necessary for a third, fourth, etc., critical band pairs and associated gradient segments. During this process, one should be aware that an optimized preceding segment can compromise the resolution in the following segment. This will be less of a problem with protein samples, because of small values of $\varphi_e - \varphi^*$, and more of a problem with small molecules in the 200–500 Da molecular weight range. In these instances, it may be advantageous to de-optimize a preceding segment (*i.e.*, reduce the value of φ^* so as to allow a maximum value of minimum resolution for the two segments together).

(6) For samples that can be broken into distinct groups (especially samples of higher molecular weight), each group can be examined separately to determine an approximately optimal gradient steepness (%B/min). The use of RRMs is useful for this purpose, as illustrated in Fig. 6 for the ribosomal proteins. Now proceed in similar fashion as for steps 4 and 5 so as to build up a suitable multi-segmented gradient. Again it will sometimes be profitable to de-optimize a preceding segment in order to improve the resolution of a subsequent segment, so as to achieve a maximal value of minimal resolution (for both segments). Similarly, it is sometimes best to end a preceding segment prior to elution of the last band in that group, especially when the last band is well resolved from the critical band pair in this segment. That is, the value of φ^* should sometimes be smaller than the value required to elute the last band of a group.

These recommendations plus the discussion of the preceding two examples should facilitate the computer-assisted design of optimized gradients. Ideally, these rules could form the basis of a computer-assisted gradient-optimization scheme. However, it is not yet possible to reduce this procedure to sufficiently simple steps that can be programmed into a computer, as the overall process is still too complex and not yet well enough understood.

Procedures to avoid. In our experience, there are also some approaches or strategies which seem less useful. Attempting to understood the details surrounding the separation of critical band pairs in the chromatogram generally leads to "paralysis by analysis". Fig. 10 for one part of an actual chromatogram illustrates this point; one can easily imagine the incredible complexity of separation for this group of compounds as a function of (a) preceding gradient segments, (b) the starting point (φ^*) of the present segment and (c) the duration (t_G) of this segment (in which the sample group is eluted).

Another problem to avoid is attempting an overly precise adjustment of gradient conditions so as to maximize the minimal resolution. Improvements of less than 0.1

unit in R_s are generally not worthwhile, except as they accumulate to give greater improvements in separation. The actual experimental run will often deviate from the predicted separation by ± 0.1 R_s units. Also, a major advantage of computer simulation is its ability to suggest that further improvements in gradient conditions will not be useful, when several successive attempts at improving sample resolution prove unsuccessful.

CONCLUSIONS

Recommendations for the best approach to designing an optimal gradient for a given sample are provided, based on the use of computer simulation and DryLab software. Computer simulation greatly simplifies the task of obtaining an adequate separation of complex samples containing a large number of components (*e.g.*, fifteen or more compounds). The application of computer simulation appears especially worthwhile for biological samples containing peptides or proteins.

The design of acceptable gradients for some samples (case I) is straightforward; only the gradient steepness and range need to be optimized, and there are no changes in band spacing to confuse the chromatographer. Samples composed of homologous or oligomeric series (case II) are slightly more challenging, but again the proper choice of gradient steepness and range is usually adequate. In some instances a two-segment gradient will provide better overall resolution and/or a shorter run time; curved gradients are seldom required or worthwhile.

Samples that exhibit changes in band spacing when gradient conditions are varied (case III) offer the greatest opportunity for maximizing resolution by selecting the best gradient. Often such samples benefit from multi-segmented gradients, so that the gradient steepness (and band spacing) can be optimized at different parts of the chromatogram. Because separation within a given gradient segment is affected by previous segments in the gradient, the design of the overall gradient can be challenging. Various techniques for simplifying this procedure are described.

REFERENCES

1 B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 25.
2 J. W. Dolan, L. R. Snyder and M. A. Quarry, *Chromatographia*, 24 (1987) 261.
3 B. F. D. Ghrist, B. S. Cooperman and L. R. Snyder, in F. E. Regnier and K. M. Gooding (Editors), *HPLC of Biological Macromolecules: Methods and Applications*, Marcel Dekker, New York, 1989.
4 B. F. D. Ghrist, B. S. Cooperman and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 1.
5 B. F. D. Ghrist, *Ph.D. Thesis*, University of Pennsylvania, Department of Biochemistry, 1988.