Journal of Chromatography, 485 (1989) 569–579
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 22 059

PEAK TRACKING IN HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY BASED ON NORMALIZED BAND AREAS

A RIBOSOMAL PROTEIN SAMPLE AS AN EXAMPLE[a]

I. MOLNAR*, R. BOYSEN and P. JEKOW
Institute of Applied Chromatography, Blücherstrasse 22, D-1000 Berlin 61 (F.R.G.)

SUMMARY

Complex chromatograms that result from reversed-phase gradient elution often exhibit changes in band order when the gradient steepness is changed. This complicates the interpretation of the resulting separation, and prevents the application of computer simulation for method development. A simple procedure based on normalized band areas was used to match bands between runs where the gradient steepness has been changed. In one example involving a *Thermus aquaticus* ribosomal protein sample, it was possible to find an additional band that was not apparent in two initial experimental runs with different gradient slopes.

INTRODUCTION

The separation of mixtures of peptides and/or proteins by reversed-phase gradient elution has proved to be a powerful technique, the advantages including the ability to control band spacing by changing the gradient steepness[1,2]. In this way complex mixtures (where the number of components approaches the peak capacity of the separation) can be manipulated to take maximum advantage of the available space in the chromatogram[3–5]. However, this tendency of the chromatogram to change with gradient steepness can also be a disadvantage, in that it may be difficult for the chromatographer to follow the relative movement of various bands in the chromatogram as the gradient conditions are varied. If computer simulation (*e.g.*, DryLab G based on two initial experimental runs[6]) is used as an aid in method development for such separations, peak tracking between two or more experimental runs is also required.

Several approaches to peak tracking have been described for such situations. Issaq and McNitt[7] described the use of band areas as an aid in matching peaks between runs. Glajch *et al.*[2] reported the use of an algorithm that was based on relative retention and band area as a means of tracking bands in peptide samples. Lankmayr and co-workers[8–10] used a similar procedure based on fuzzy set theory. Other workers[11–13] have made use of the discriminating power of a diode-array

[a] Dedicated to Professor Dr. Csaba Horváth for his 60th birthday.

detector as a means of matching bands between two or more chromatograms. Each of these peak-tracking strategies can be useful in some instances, but can fail in others.

EXPERIMENTAL

Equipment and software

The gradient high-performance liquid chromatographic (HPLC) system used was designed for protein analysis, based on two pumps, a gradient programmer and mixing chamber (Knauer, Berlin, F.R.G.), an autosampler (Promis, Spark Holland, Emmen, The Netherlands) and a spectrophotometric detector (Knauer). Chromatograms and band area measurements were obtained using a detector–PC interface (Nelson Analytical, Zug, Switzerland) in connection with a V386 Personal Computer (Victor, Frankfurt, F.R.G.) and an NEC P6 printer (NEC, West Berlin, F.R.G.). The DryLab G-PLUS software used is from LC Resources (European supplies Molnar, Berlin, F.R.G.).

Materials

NaH_2PO_4 and H_3PO_4 were purchased from Laborat (Berlin, F.R.G.). Solvents were of HPLC grade (Merck, Darmstadt, F.R.G.) and degassed with helium before use. The benzoic acid esters (methyl, ethyl, *n*-propyl, *n*-butyl and *n*-pentyl) were from Alldrich (Munich, F.R.G.).

Samples

The preparation and separation of the 50S ribosomal proteins from *Thermus aquaticus* (wildtype, YT1; CAMR, Porton, U.K.) was carried out according to the procedure described previously[14].

Chromatographic conditions

Benzoic acid esters. The sample was a mixture of the five C_1–C_5 *n*-alkyl benzoates. Solvent A was 50 m*M* NaH_2PO_4–50 m*M* H_3PO_4 in water (pH 2.1), solvent B was acetonitrile and gradients were carried out from 5 to 100% B in times of 6, 13, 19 and 57 min (runs 1–4) at a flow-rate of 2.0 ml/min (pressure = 17.4 MPa). The column was Supersphere RP–18 (25 × 0.46 cm I.D.), 4-μm particles (Molnar).

Ribosomal proteins. The Solvent A was 500 m*M* NaH_3PO_4–500 m*M* H_3PO_4 in water (pH 2.1), solvent B was a 50:50 (v/v) mixture of 250 m*M* NaH_2PO_4–250 m*M* H_3PO_4 in water (pH 2.1) with acetonitrile and gradients were carried out from 40 to 100% B in times of 90 and 270 min (runs 5 and 6) at a flow-rate of 1.5 ml/min (pressure = 18 MPa) at room temperature. A third gradient in 180 min (the other conditions being the same) was subsequently run to check the prediction of a "hidden" band in the first two runs.

RESULTS AND DISCUSSION

Peptide and protein samples are often separated by low-pH gradients from water to acetonitrile on reversed-phase columns. For these separations it appears that band areas augmented by relative retention should provide sufficient information to allow accurate peak tracking in most instances. Where there is doubt, computer-

simulation software can provide additional assistance[3], as will be illustrated here. We assume that it is desired to match bands between two gradient runs differing only in gradient steepness. Under these conditions, it is likely that absorptivity changes between runs, which is a possible problem in the use of band areas, will be minimal. The data given here confirmed this assumption.

It is first necessary to record the peak areas of all distinct bands in the chromatogram, including partly resolved peaks. Where the baseline is not well defined, these areas can be improved by trial-and-error reintegration using the Nelson software. We begin by noting that the total peak area for the two chromatograms being compared should be identical (within experimental error). However, this is not essential for the use of band areas in peak tracking.

The total peak areas, $(A_T)_1$ and $(A_T)_2$ for runs 1 and 2, respectively, are determined, as $A_T = \sum_{i=1}^{n} A_i$, and their ratio is calculated:

$$R_T = (A_T)_2/(A_T)_1 \tag{1}$$

The areas of corresponding (presumably the same compound i) bands in run 1 (A_1) and run 2 (A_2) are also ratioed:

$$R_i = (A_i)_2/(A_i)_1 \tag{2}$$

If we have matched two bands correctly, R_i should equal R_T. Alternatively, the ratio $C_i = R_i/R_T$ corrects for differences in total area between the two runs and is therefore

TABLE I

PEAK AREAS AND RATIOS FOR BENZOATE ESTERS

Benzoate Solute	*Peak areas, A_i ($\times 10^{-3}$)*			
	Run 1 (6 min)[a]	*Run 2 (13 min)*[a]	*Run 3 (19 min)*[a]	*Run 4 (57 min)*[a]
Methyl	1026	1152	1080	1434
Ethyl	622	692	652	866
n-Propyl	537	602	551	760
n-Butyl	466	502	465	647
n-Pentyl	442	475	439	613
Total area, A_T	3093	3423	3187	4320
	Peak-area ratios, $(A_i)_2/(A_i)_1$			
	Run 1/2	*Run 2/3*	*Run 3/4*	
Methyl	0.89	1.07	0.75	
Ethyl	0.89	1.06	0.75	
Propyl	0.89	1.09	0.73	
n-Butyl	0.93	1.08	0.72	
n-Pentyl	0.93	1.08	0.72	
R_T	0.90	1.07	0.74	

[a] Gradient.

a better indicator of band sameness (when $C_i = R_i/R_T \approx 1.00$) ($C_i$ = identification coefficient).

Benzoate esters

The C_1–C_5 *n*-alkyl benzoates are easily identified in their respective chromatograms, *i.e.*, there are no possibilities of changes in relative retention with change in the gradient conditions. Table I summarizes band-area measurements for these compounds in four runs of varying gradient time (6, 13, 19 and 57 min), together with calculated values of the ratios R_i for each band and each adjacent pair of gradient runs. In Table II we correct for differences in total band area among the various runs by calculating the ratios R_i/R_T. These ratios are seen to equal 0.99 ± 0.03 (1 S.D.), which is indicative of the repeatability of band-area measurements in this study. Hence the use of band-area ratios shows considerable promise for peak tracking in the comparison of chromatographic runs differing only in gradient steepness.

TABLE II

CORRECTED PEAK-AREA RATIOS FOR BENZOATE ESTERS, $R_i/R_T = C_i$

Benzoate Solute	*Run 1/2*[a]	*Run 2/3*[b]	*Run 3/4*[c]
Methyl	0.99	1.00	1.01
Ethyl	0.99	0.99	1.01
n-Propyl	0.99	1.02	0.99
n-Butyl	1.03	1.00	0.98
n-Pentyl	0.92	1.00	0.98

[a] 13-min gradient (run 2) *vs.* 6-min gradient (run 1).
[b] 19-min gradient *vs.* 13-min gradient.
[c] 57-min gradient *vs.* 19-min gradient.

Ribosomal proteins

Chromatograms for the separation of this sample with only gradient steepness varying are shown in Figs. 1 (90-min gradient) and 2 (270-min gradient). It is apparent that peak tracking is considerably more demanding than with the alkyl benzoate sample. There are many bands in the chromatograms, several band pairs are poorly resolved and the baseline is often poorly defined. Chromatograms such as these are typical of such samples, *e.g.*, see the examples in refs. 3 and 5.

Several bands in Figs. 1 and 2 can be matched on the basis of relative retention and (especially) band areas. However, other regions of the chromatogram present definite problems, *e.g.*, the 28–37-min region of Fig. 1 and the corresponding 75–96-min region of Fig. 2 (indicated as bands C_1–C_{12}). Consider the six fairly obvious bands B_1–B_6 first. Table III summarizes the band areas and band-area ratios as defined previously. The corrected ratios $C_i = R_i/R_T$ for these bands are shown at the end of Table III. For all six bands, the average ratio is 1.00 ± 0.04 (1 S.D.), *i.e.*, the match is comparable to that for the alkyl benzoates in Table III. However, poorer agreement for correctly assigned bands can be expected for the remaining bands in Figs. 1 and 2 owing to poorer resolution, smaller bands in some instances and other factors.

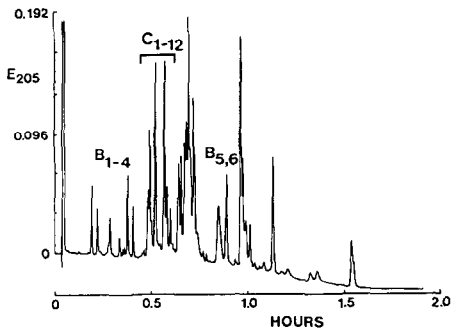


Fig. 1. Separation of ribosomal protein sample by reversed-phase gradient elution. 90-min gradient; see Experimental for other conditions.

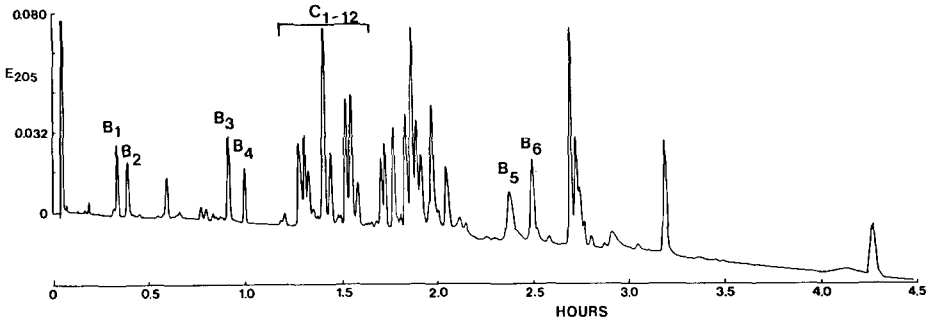


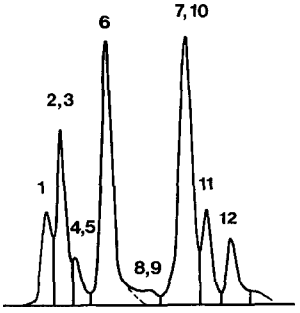
Fig. 2. Separation of ribosomal protein sample by reversed-phase gradient elution. 270-min gradient; see Experimental for other conditions.

TABLE III

RETENTION AND PEAK AREA DATA FOR EASILY MATCHED BANDS IN RIBOSOMAL PROTEIN RUNS (FIGS. 1 AND 2)

Peak No.	*Retention times of ribosomal proteins (min)*		*Individual peak areas of ribosomal proteins, A_i ($\times 10^{-3}$) and R_i for runs 5 and 6*			*Individual peak-area ratios/ total peak-area ratios of well behaving ribosomal proteins, C_i, for run 5/6*
	Run 5	*Run 6*	*Run 5*[a]	*Run 6*[b]	*R_i (5/6)*	
B1	11.53	20.05	1093	976	$R_{B1} = 1.12$	1.04
B2	13.38	23.32	972	897	$R_{B2} = 1.08$	1.00
B3	22.78	55.22	1800	1607	$R_{B3} = 1.12$	1.04
B4	24.57	60.15	795	784	$R_{B4} = 1.01$	0.94
B5	51.15	142.67	2999	2703	$R_{B5} = 1.11$	1.03
B6	53.50	149.82	2854	2803	$R_{B6} = 1.02$	0.94
A_T			10 513	9770		

[a] 90 min gradient.
[b] 270 min gradient.



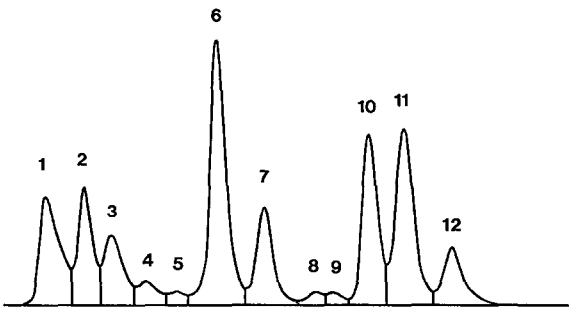


Fig. 3. Expanded chromatogram for bands C_1–C_{12} in Fig. 1 (90-min gradient).

Fig. 4. Expanded chromatogram for bands C_1–C_{12} in Fig. 2 (270-min gradient).

Bands C_1–C_{12} in Figs. 1 and 2. This region of the chromatogram will provide a real test of our ability to carry out peak tracking on the basis of band areas. In order to see better the separation of these bands, the chromatogram is expanded in Figs. 3 (90-min gradient; *cf.*, Fig. 1) and 4 (270-min gradient; *cf.*, Fig. 2). As separations are generally better with longer gradient times, it is usually best to look first at the longer gradient (Fig. 4). Here we can see twelve distinct bands, which are numbered for comparison with their counterparts in Fig. 3. The band assignments (numbers) in Fig. 3 are not immediately obvious, but have been made on the basis of the following discussion.

Table IV summarizes the band area comparison, carried out in the same way as earlier (Tables II and III). Since there are only eight distinct bands in Fig. 3, com-

TABLE IV

PEAK-AREA DATA FOR PEAKS C_1–C_{12} IN 90-MIN (RUN 5) AND 270 MIN (RUN-6) CHROMATOGRAMS IN FIGS. 3 AND 4

Peak No.	*Peak area (run 5)*	*Peak No.*	*Peak area (run 6)*	R_i *(5/6)*
C_1	1672	C_1	2026	0.83
		C_2	1588	
		C_3	1135	
(C_2 + C_3)	2695	C_2 + C_3	2723	0.99
		C_4	387	
		C_5	175	
(C_4 + C_5)	757	C_4 + C_5	562	1.25
C_6	4875	C_6	4068	1.20
		C_8	185	
		C_9	175	
(C_8 + C_9)	339	C_8 + C_9	360	0.94
		C_7	1449	
		C_{10}	2395	
(C_7 + C_{10})	5892	C_7 + C_{10}	3844	1.53[a]
C_{11}	1615	C_{11}	2985	0.54[a]
C_{12}	1386	C_{12}	943	1.47[a]
A_T	19 231		17 511	

[a] C_i values deviate significantly from 1.0.

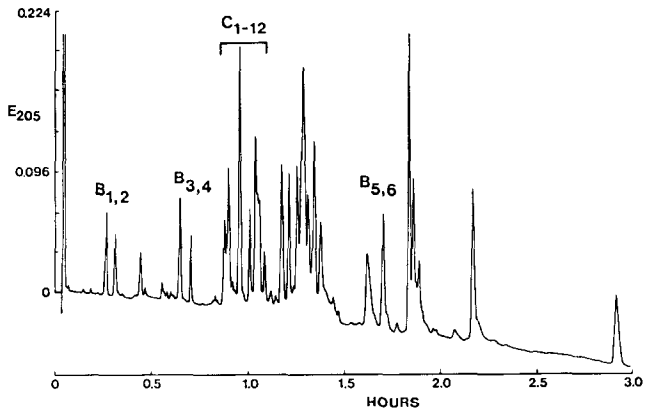


Fig. 5. Separation of ribosomal protein sample by reversed-phase gradient elution. 180-min gradient; see Experimental for other conditions.

pared with twelve in Fig. 4, it is apparent that there are several overlaps in the chromatogram in Fig. 3. Various obvious combinations were tried, resulting in the pairings summarized in Table IV. The ratios R_i/R_T shown in the last column in Table IV are reasonably close to the expected value of 1.00 in most instances. However, the last three values are clearly out of line, suggesting that something is wrong with these assignments.

For the last band (C_{12}), inspection of the chromatogram suggested that the band area may not have been integrated properly owing to an initially poor baseline assignment by the data system. Unfortunately, separations such as this often exhibit baselines of questionable quality owing to the need for high-sensitivity detector settings, contamination of the sample and reagents, etc. In these instances it is useful to explore whether any "reasonable" baseline can result in a good match of the bands areas between the two runs, before assuming that the bands in question do not match. Using the reintegration feature of the Nelson software, it was possible to bring band C_{12} into agreement in this instance. However, bands C_7, C_{10} and C_{11} require different handling, as any reasonable readjustment of the baseline did not bring the band areas into agreement between the two runs.

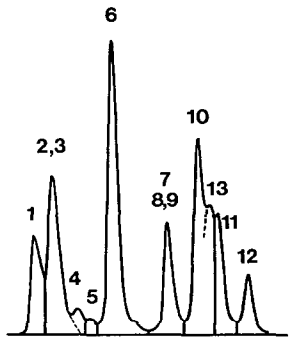


Fig. 6. Expanded chromatogram for bands C_1–C_{12} in Fig. 5 (180-min gradient).

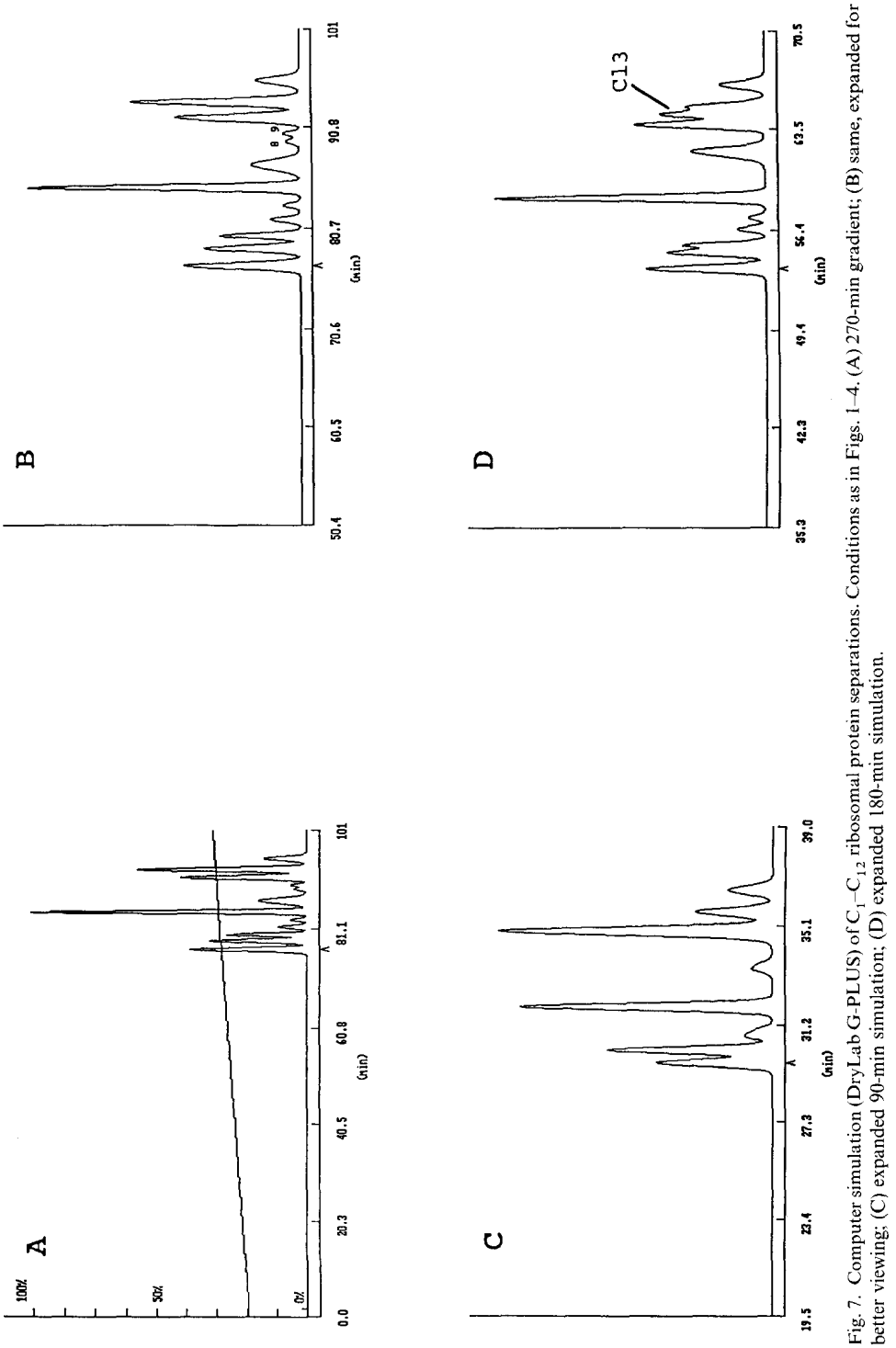


Fig. 7. Computer simulation (DryLab G-PLUS) of C_1–C_{12} ribosomal protein separations. Conditions as in Figs. 1–4. (A) 270-min gradient; (B) same, expanded for better viewing; (C) expanded 90-min simulation; (D) expanded 180-min simulation.

The relative areas of bands C_7, C_{10} and C_{11} suggest (on further study) that an additional sample component is involved in this group. It appears that this "hidden" protein overlaps bands $C_7 + C_{10}$ in the 90-min run (Fig. 3) and band C_{11} in the 270-min run (Fig. 4). If this is so, then an intermediate run (*e.g.*, 180-min gradient) would be expected to show the presence of the hidden band (see discussion in refs. 3–5). Figs. 5 (whole chromatogram) and 6 (region including bands C_1–C_{12}) show the resulting separation with the 180-min gradient. It is apparent from Fig. 6 that an additional band is now resolved between bands C_{10} and C_{11}, which we have designated C_{13}. In addition, bands C_8 and C_9 in Fig. 4 have merged with C_7 in Fig. 6. Bands C_2 and C_3 are also overlapped in Fig. 6.

Computer simulations. The above process of peak tracking, including the discovery of "hidden' bands (*e.g.*, C_{13}) can be facilitated by the use of computer simulation with the DryLab software. It can also be shown that additional "hidden" bands are unlikely in this region of the chromatogram. Finally, some other applications of computer simulation for this sample will be illustrated.

DryLab G-PLUS is a recently released program that can be used for method development based on reversed-phase gradient elution. This new software package has some useful features that will be illustrated here in the separation of this ribosomal protein sample. The use of DryLab G-PLUS requires experimental data for two runs with different gradient times; the chromatograms in Figs. 3 and 4 were used for this purpose. Retention times, band areas and experimental run conditions for these two runs were entered into DryLab G-PLUS.

The next step, particularly where peak tracking is of interest, is to adjust the column plate number (by computer simulation) so as to match the resolution of an initial experimental run. Figs. 7A and B show the simulated chromatogram for the 270-min run after this adjustment. Comparison with Fig. 4 shows generally good agreement. Similarly, the agreement of the chromatogram of Fig. 3 (90-min run) with the simulation in Fig. 7C is also reasonable.

Now we are ready to compare the 180-min run of Fig. 6 with a simulation by DryLab G-PLUS (Fig. 7D). The 63–65-min regions of these two chromatograms are comparable, confirming the presence of band C_{13}. There is no indication of any additional "hidden" bands in the further comparison of Figs. 6 and 7D. Hence the use of computer simulation indicates that there are thirteen (and only thirteen) bands in this region.

Suppose it were of interest to isolate the "hidden" band C_{13} for further study. The separation in Fig. 6 is not very good, so the question arises of whether we can select another gradient time that would provide a better resolution of C_{13} from adjacent bands. Computer simulation is ideal for this purpose. With DryLab G-PLUS, we can "mark" band C_{13} and then plot the resolution for this band alone (overlapped with other possible bands in the sample) as a function of gradient steepness or time. Fig. 8A shows the resulting "resolution map" provided by DryLab G-PLUS. We see first that C_{13} is totally overlapped by band C_{10} or C_{11} (resolution equal to zero) for gradient times of 90 and 270 min, which we already knew. With a further increase in gradient time, band C_{13} passes through band C_{12} and eventually leaves the column last. With a gradient time of about 1200 min, C_{13} is well separated from the other bands of the sample ($R_8 \approx 1.7$)[a].

[a] If there is any question of band C_{13} overlapping later bands in the separation in Fig. 5, retention data for those bands should also be entered prior to computer simulation.

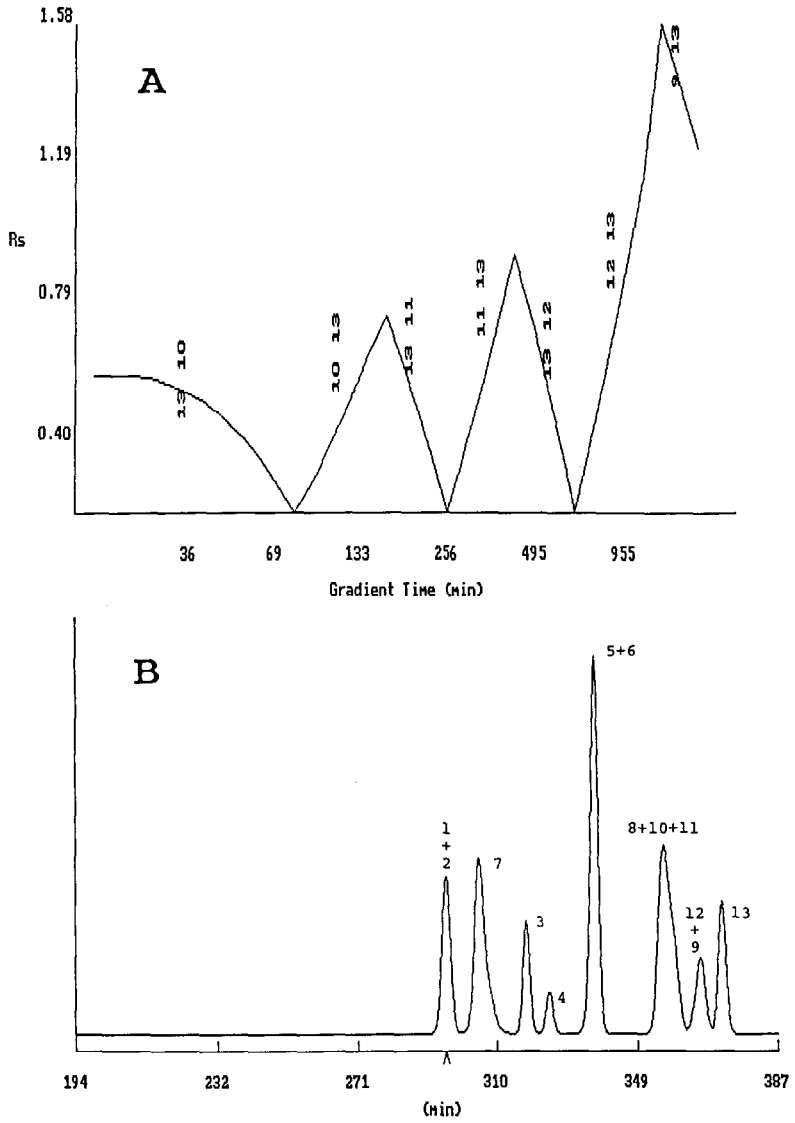


Fig. 8. Use of computer simulation to isolate the "hidden" band C_{13}. (A) resolution map (20–50% B gradient) for C_{13} and adjacent bands; numbers (*e.g.*, 10 13; 13 11) indicate critical band pairs of lowest resolution; (B) simulated chromatogram for "best" gradient time of 1200 min. Other conditions as in Figs. 1–4.

The simulated chromatogram for a gradient time of 1200 min is shown in Fig. 8B. It can be expected that C_{13} will be easily separated from bands C_1–C_{12} in a time of about 6 h (the gradient can be stopped short of 1200 min). Alternatively, bands C_{11} + C_{13} from the 270-min run (Fig. 4) could have first been recovered as a mixture, then resolved with a very short gradient (*e.g.*, 90 min).

CONCLUSIONS

It has been shown that relative band areas can be used effectively to track bands between gradient runs that differ only in steepness (or gradient time). This process can be further aided by the use of computer simulation, where a third experimental run is compared with the chromatogram predicted on the basis of the first two runs. Computer simulation can also be used to aid in the development of a final gradient procedure, or to answer specific questions such as how best to isolate a given band from the chromatogram.

SYMBOLS

A_i	peak area of component i
$(A_T)_1$, $(A_T)_2$	total peak area of runs 1 and 2; equal to $(\Sigma A_i)_1$, and $(\Sigma A_i)_2$ respectively
C_i	identification coefficient for a band in run 1 *vs.* a band in run 2, equal to R_i/R_T; ideally, $C_i = 1.00$
R_T	total peak-area ratio of run 2 *vs.* run 1; equal to $(A_T)_2/(A_T)_1$
R_i	individual peak-area ratio; equal to $(A_i)_2/A_i)_1$

ACKNOWLEDGEMENTS

We thank V. A. Erdmann for the ribosomal sample and L. R. Snyder for helpful discussions. This work was supported by the Bundesministerium für Forschung und Technologie, Project Nr. 0317107B, the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

REFERENCES

1 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 296 (1984) 31.
2 J. L. Glajch, M. A. Quarry, J. F. Vasta and L. R. Snyder, *Anal. Chem.*, 58 (1986) 280.
3 B. F. D. Ghrist, B. S. Cooperman and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 1.
4 B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 25.
5 B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 43.
6 J. W. Dolan, D. C. Lommen and L. R. Snyder, *J. Chromatogr.*, 485 (1989) 91.
7 H. J. Issaq and K. L. McNitt, *J. Liq. Chromatogr.*, 5 (1982) 1771.
8 M. Otto, W. Wegscheider and E. P. Lankmayr, *Anal. Chem.*, 60 (1988) 517.
9 E. P. Lankmayr, W. Wegschneider and K. W. Budna, *J. Liq. Chromatogr.*, 12 (1989) 35.
10 E. P. Lankmayr, W. Wegscheider, J. Daniel-Ivad, I. Kollosvary and G. Csonka, *J. Chromatogr.*, 485 (1989) 557.
11 J. K. Strasters, H. A. H. Billiet, L. de Galan, B. G. M. van de Ginste and G. Kateman, *J. Liq. Chromatogr.*, 12 (1989) 3.
12 J. C. Berridge, *Techniques for the Automated Optimization of HPLC Separations*, Wiley, Chichester, 1986.
13 A. F. Fell, H. P. Scott, R. Gill and A. C. Moffat, *Chromatographia*, 16 (1982) 69.
14 I. Molnar, R. I. Boysen and V. A. Erdmann, *Chromatographia*, 28 (1989) 39.