Journal of Chromatography, 327 (1985) 93–113
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 586

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

M. A. STADALIUS* and M. A. QUARRY
Biomedical Products Department, E. I. du Pont de Nemours & Co., Concord Plaza, Wilmington, DE 19898 (U.S.A.)
and
L. R. SNYDER
Lloyd R. Snyder, Inc., 2281 William Court, Yorktown Heights, NY 10598 (U.S.A.)

INTRODUCTION

Previous papers in this series[1,2] have presented a general model for the separation of peptide and protein samples by reversed-phase high-performance liquid chromatography (HPLC) using gradient elution. This model begins with the general theory of band broadening in isocratic liquid chromatography as developed by previous workers (*e.g.*, refs. 3–8), adds the corresponding theory which relates gradient and isocratic separation[9–14], and then generalizes the overall treatment so as to apply to both large and small molecules. Given information concerning the isocratic separation of a given peptide or protein in a defined HPLC system, our approach allows prediction of corresponding separations by gradient elution. This has been confirmed for both retention-time measurements[1] and for bandwidth data[2]. It is also possible to extract isocratic separation parameters from gradient runs[12,13], which in turn allows predictions of gradient separation (for any experimental conditions) without recourse to generally less convenient isocratic separations. Finally, as is further developed in this paper, previous experimental data can be generalized so as to allow predictions of gradient separation in the absence of experimental isocratic data. Only the molecular weight of the sample is required, and rough estimates of this quantity suffice for accurate calculations of separation as a function of experimental conditions.

In this paper we will illustrate the utility of our model for practical application and for gaining further insight into the fundamental nature of these gradient separations of peptides and proteins. Specifically we will examine the following areas: (i) a general approach to method development for the separation of complex mixtures of peptides and/or proteins; (ii) the dependence of column performance on the configuration of the column (dimensions, diameter and pore size of the column-packing particles) and sample molecular weight; (iii) adaptation of the model for use with a

0021-9673/85/$03.30

specific column configuration that is optimally suited for separations of peptide samples with molecular weights in the 300–20 000 dalton range.

A glossary of all symbols used is given in the preceding paper[2].

THEORY*

The model used by us is described in detail in refs. 1 and 2. It should be emphasized that this model does not assume any particular "mechanism" or retention process for the partitioning of peptide molecules between the stationary and mobile phase (*e.g.*, ref. 15), nor does the model consider differences in solute retention that results from changes in mobile phase composition (*e.g.*, refs. 16 and 17) apart from variation in the volume fraction φ of organic solvent. Our model does assume that the separation system is "well behaved" as discussed in the preceding paper[2]; *i.e.*, secondary retention processes that lead to band asymmetry and/or band broadening (*e.g.*, slow sorption/desorption kinetics) are absent, slow configurational changes of sample molecules can be ignored, the solute exists in a single equilibrium state during elution (usually a denatured state), etc. These assumptions will often be invalid in a given separation, particularly of higher molecular weight samples (*e.g.*, see refs. 18–21). In these cases the practical goal is to confirm that such problems exist, then select chromatographic conditions that will minimize "nonideal" effects and provide the best possible separation of the sample.

Optimizing the separation

The following discussion can be better understood if we compare this approach to analogous procedures for method development in the isocratic separation of small molecules (*e.g.*, ref. 22). In the latter case we first vary solvent strength so as to optimize overall sample retention; *i.e.*, to achieve $1 \leqslant k' \leqslant 10$ for all sample components. In gradient elution we can achieve the same effect by varying gradient conditions, since the effective or average capacity factor $\bar{k}$ is given by

$$\bar{k} = t_G F / 1.15\, V_m\, \Delta\varphi\, S \tag{1}$$

Here t_G is the gradient time, F is the flow-rate, V_m is the column dead volume, $\Delta\varphi$ is the change in volume fraction φ of the strong solvent (usually acetonitrile) during the gradient, and S is equal to $-\mathrm{d}(\log k')/\mathrm{d}\varphi$ for the solute under isocratic conditions. The value of S depends mainly on solute molecular weight M; to a first approximation we can write[1]:

$$S = 0.48\, M^{0.44} \tag{2}$$

For protein digests, the value of M might be about 1000, and S is then equal to 10. In any case, given an estimate of M for the sample, S can be predicted and experimental conditions selected (eqn. 1) to yield an optimum value of $\bar{k}$ (usually $2 \leqslant \bar{k} \leqslant$

* For explanation of symbols see also ref. 2, pp. 43 and 44.

10). This procedure contrasts with isocratic method development, where the optimum solvent strength for a given sample must be determined by trial and error.

Following selection of the right solvent strength in isocratic separation, the separation time is then determined by the retention time for the last-eluted band. In gradient elution the separation time t_G is chosen (*e.g.*, 15–60 min), and for the initial sample separation a broad gradient range should be selected; *e.g.*, φ varying from $0.05 \leqslant \varphi \leqslant 0.60$. Typically the sample will not totally fill the chromatogram, with either the beginning and/or the end of the chromatogram being devoid of major peaks or bands of interest. At this point the gradient range can be narrowed in order to save time; *e.g.*, $0.15 \leqslant \varphi \leqslant 0.40$ for a given sample. The gradient time should also be reduced proportionately so as to keep $\bar{k}$ constant; *i.e.*, $(t_G/\Delta\varphi) = \text{constant}$. In this way the original separation will be maintained, since changes in the initial and final values of φ for the gradient only affect bands that elute near the beginning or end of the gradient.

The next option in isocratic method development is to increase the plate number N for better overall resolution of the sample. Typically this will be done while holding mobile phase composition constant, which means no change in band spacing. In the isocratic separation of small molecules, N can be increased by increasing column length L, and to a lesser extent by decreasing the mobile phase flow-rate F. Essentially the same approach is recommended in the gradient elution separation of peptide samples. However certain differences *vs.* isocratic separation should be noted. In gradient elution a change in L or F by itself will also change $\bar{k}$. This is seen in eqn. 1 where $\bar{k}$ depends on F and upon V_m — which is usually proportional to L. A further difference relates to the larger values of M that are typically found for peptide/protein samples *vs.* other samples. Thus the Knox equation (see refs. 1 and 2).

$$h = Av^{1/3} + B/v + Cv \tag{3}$$

predicts that changes in flow-rate (and reduced velocity v) will have a greater effect on h and upon N when v is large. Larger molecules have smaller diffusion coefficients D_m, which means that corresponding HPLC separations of these molecules will exhibit larger values of v. Therefore a reductions in F will generally result in a larger increase in N for separations of peptides and proteins (*vs.* that of small molecules). Since changes in F are more convenient than increase in L, the first step in increasing N for gradient separations of peptide/protein samples should be a decrease in F while holding $\bar{k}$ constant. This can be achieved by increasing t_G in proportion to any decrease in F (see eqn. 1); *i.e.*, by holding the gradient volume $V_G = t_G F$ constant.

When flow-rate is reduced to a sufficiently small value, so that the reduced velocity approaches 3–10, further reduction in F is no longer advantageous (N approaches a maximum). At this point the best strategy is to increase column length for further increase in N and resolution. However the product $(t_G F/L)$ must be maintained constant if $\bar{k}$ and band spacing are to remain unchanged (eqn. 1, noting that V_m is proportional to L for column lengths of the same type).

The steps outlined above for optimizing the separation lead to increase in sample resolution without changes in band spacing. That is, we increase peak capacity PC which is given as

$$\text{PC} = (2.3/4)\,(S\Delta\varphi)\,N^{1/2}[\bar{k}/(1 + \bar{k})] \tag{4}$$

Here S refers to the average value for various compounds in the sample, equal to the value given by eqn. 2 using the average value of M for the sample. Once there appears to be adequate room within the chromatogram for all sample peaks to fit (50–100% more space than is required for the peaks so far separated), band spacing can be changed so as to resolve band-pairs that overlap at this point. This is achieved by holding other conditions constant while varying flow-rate over a range of 5 to 10-fold[14]. Flow-rate changes can also be used to increase detection sensitivity for selected peaks, since peakwidth decreases and peakheight increases in proportion to decrease in $(1 + k')$, while k' decreases with decrease in F according to eqn. 1*. Band spacing can be further adjusted by changing the mobile phase solvent (*e.g.*, isopropanol replacing acetonitrile), by varying pH, by adding an ion-pairing reagent (*e.g.*,

TABLE I

OPTIMIZING THE SEPARATION OF PEPTIDE/PROTEIN SAMPLES BY MEANS OF REVERSED-PHASE GRADIENT ELUTION

Column: DuPont Bioseries PEP-PR1; mobile phase: acetonitrile–water with 0.1% trifluoroacetic acid (amine additive optional, see ref. 2).

Step 1	Run standard gradient: 5–60% (v/v) acetonitrile–water, t_G = 45 min, $F = 0.10S$ ($\bar{k}$ = 5).
Step 2	Check for non-chromatographic contributions to bandwidth by comparing bandwidths from Step No. 1 with values calculated from the model summarized in ref. 2; minimize these effects if possible before proceeding.
Step 3	Reduce the gradient range $\Delta\varphi$ to avoid empty regions at beginning or end of chromatogram (*e.g.*, 15–50% acetonitrile); reduce t_G proportionately (keep $t_G/\Delta\varphi$) constant (step No. 3 can be combined with No. 4 to minimize the number of separate runs).
Step 4	Increase peak capacity to provide enough space within chromatogram for all bands to fit with adequate resolution (a) increase t_G and decrease F proportionately (hold $t_G F$ constant); (b) increase column length L while increasing F [hold ($t_G F/L$) constant].
Step 5	Vary band spacing to resolve all band-pairs by varying F with other variables held constant.
Step 6	Improve detection sensitivity by minimizing band-volume (see text).
Step 7	Further improve bandspacing by changes in mobile phase composition (optional); start over with Step No. 1 for each new mobile phase composition.

hexane sulfonate), etc.[16,17]. However this requires that the entire optimization procedure described above be repeated from the beginning. Table I summarizes the present approach to method development for the gradient separation of peptide/protein samples.

Column plate number and configuration

The model developed in refs. 1 and 2 for the gradient separation of large

* Here the value of k' should be k_r, the capacity factor of the band as it elutes from the column; see later discussion.

biomolecules describes column plate number as a function of certain column parameters: $a' \times 1.1$, the value of the Knox parameter B (eqn. 3) when $k' = 0$; b', the surface diffusion parameter, equal to the ratio of solute diffusion coefficients in the stationary phase *vs.* the mobile phase; ρ, the restricted diffusion parameter, equal to the factor by which D_m is decreased due to the combination of large-diameter solute molecules and small-diameter pores within the column packing. These three parameters (a', b' and ρ) are actually interrelated as can be seen from a preliminary analysis to be presented later, based upon the general analysis of ref. 8. To summarize that treatment, both b' and ρ will decrease together as the solute molecule within the pores of the packing becomes more crowded. This is observed for data summarized in the preceding paper[2]. For a given column, values of the column parameters a', b' and ρ should be strictly a function of sample molecular weight M. For the present 15-nm pore C_8 column studied by us this appears to be the case. With a' empirically equal to 1.1 as in ref. 8, values of the other column parameters are given within experimental error by

$$b' = 0.45 - 0.089 \log M \tag{5}$$

and

$$\log \rho = 0.25 - 0.19 \log M \tag{6}$$

With these latter relationships the model of refs. 1 and 2 allows the prediction of relative resolution (peak capacity and peak volume) for any sample and set of separation conditions, providing we can estimate the sample molecular weight M. This can be useful in various ways as illustrated in later sections. A computer program for application of this model with the Bioseries PEP-RP1 column is described in the Appendix.

Detector sensitivity

Detection sensitivity or peak height is determined by the volume which contains the band as it elutes from the column. The 1-sigma peak volume σ_v is given[1] as

$$\sigma_v = (JG)\, V_m(k_f + 1)\, N^{-1/2} \tag{7}$$

where J is the factor by which bands are widened anomalously in gradient elution[2], G is the compression factor that describes the narrowing of gradient bands (ref. 9, pp. 288–289), and k_f is the value of k' for a band in gradient elution as it leaves the column. The latter is given (assumes "gradient" conditions, see ref. 1) as

$$k_f = \bar{k}/2 \tag{8}$$

For $0 \leqslant \bar{k} \leqslant 5$ the quantity (JG) is approximately constant (equal to 1.15 ± 0.05) by calculation from values of J and G as a function of $\bar{k}$). If F is decreased, k_f decreases (eqns. 1 and 8) and N usually increases. This means that a decrease in F

will normally decrease σ_v and increase detection sensitivity. If t_G is increased while F is decreased proportionately ($t_G F$ constant), $\bar{k}$ will remain constant (no change in band spacing, see ref. 14) so that both resolution and detection sensitivity will increase for the whole sample. This therefore constitutes a useful strategy in improving the separation.

EXPERIMENTAL

All runs were made with a Hewlett-Packard Model 1090 liquid chromatograph. Bioseries PEP-RP1 columns (Du Pont) were used: 8 × 0.62 cm, 5-μm-diameter particles bonded with C_8 (15-nm pores). The column temperature was 40°C, and gradients were run with mixtures of 0.1% (v/v) morpholine and 0.125% (v/v) trifluoroacetic acid (TFA) in water (solvent A), and 0.1% (v/v) morpholine plus 0.1% (v/v) TFA in acetonitrile (solvent B).

RESULTS AND DISCUSSION

The present model (Appendix) adapted for use with the Du Pont Bioseries PEP-RP1 column (eqns. 5 and 6 and experimental values of $x = 0.67$, $a' = 1.1$, $L = 8$ cm, $d_c = 0.62$ cm and $d_p = 4.8$ μm) allows us to do several different things:

(1) evaluate a given separation to see if bandwidths are close to expected values; if experimental σ_v values are significantly larger, non-chromatographic effects are involved and should be corrected first (non-well-behaved system);

(2) predict what change in resolution (peak capacity or band width) will occur as a result of change in any separation variable; if the resulting separation improvement is not worthwhile, some other change in conditions should be explored instead;

(3) further study "surprising" predictions of the model to better see what is actually happening; the model can be interrogated for intermediate calculations such as reduced velocity, diffusion coefficient, plate number, etc. In this way a good physical understanding of these effects can result;

(4) develop an optimum approach to method development for the reversed-phase gradient elution separation of mixtures of peptides and/or proteins; specific possibilities can be evaluated using the model in order to eliminate unprofitable options;

(5) anticipate problems with equipment limitations: extra-column effects, inability to provide low flow-rates, presence of mixing volumes that distort the gradient, etc.

(6) explore the performance of different column configurations so as to design optimum columns for specific applications.

We will examine each of these possibilities.

Method development with the Bioseries PEP-RP1 column

The approach discussed in the Theory section will be illustrated for a typical peptide sample: a tryptic digest of lysozyme. We will also provide calculated results based on our model for comparison and interpretation in each step.

Initial separation. Fig. 1 shows the initial separation of the lysozyme digest. A broad gradient (5–60% (v/v) acetonitrile) was selected, and a $\bar{k}$ value of 6 was chosen initially. The average molecular weight of the sample was estimated equal to 1100 from the known structure of lysozyme and the assumption that trypsin cuts the molecule at every position containing a lysine or arginine residue. This suggests that S equals 10.4. For an initial t_G value of 30 min, and a V_m value for this column of 1.44 ml, F should equal (6 × 1.15 × 1.44 × 0.55 × 10.4)/30 or 1.89 ml/min (eqn. 1). This flow-rate was chosen in the run of Fig. 1. After obtaining the separation of Fig. 1, the average band-volume σ_v was obtained from this chromatogram and compared with the value predicted by the model: 0.077 ml (expt) *vs.* 0.090 ml (calc). This agreement was deemed adequately close, and method development was continued (step 2 of Table I).

Narrowing the gradient range. It is seen in Fig. 1 that the early and late regions of the chromatogram are devoid of major peaks. Therefore the gradient range can be shortened so as to eliminate these portions of the chromatogram and reduce the separation time at this stage. Beginning at t_o (arrow No. 1 in Fig. 1) we see that the next 4 min (t_o arrow No. 2) are devoid of major peaks. Likewise there are no major peaks after arrow No. 3 (16 min after t_o). The interval 4–16 min corresponds to 12–34% (v/v) acetonitrile, which is the gradient interval selected for the second run. The gradient time t_G is shortened in proportion to $\Delta\varphi$ to maintain the separation constant at this point: $t_G = 30\ (34 - 12)/(60 - 5) = 12$ min. Fig. 2 shows the resulting separation. At this point we have saved time but have not improved the separation.

Increasing peak capacity. Next we increase ovcerall resolution and peak capacity (also sensitivity) by increasing t_G while decreasing F proportionately. Fig. 3 shows corresponding separations under approximately the same conditions as for Fig. 2 (10–40% acetonitrile gradient, t_G equal 15 min, $F = 2$ ml/min, $\bar{k} = 5.8$ in Fig. 3a). Fig. 3b has $t_{5G} = 30$ min and $F = 1$ ml/min, and Fig. 3c has $t_G = 60$ min and

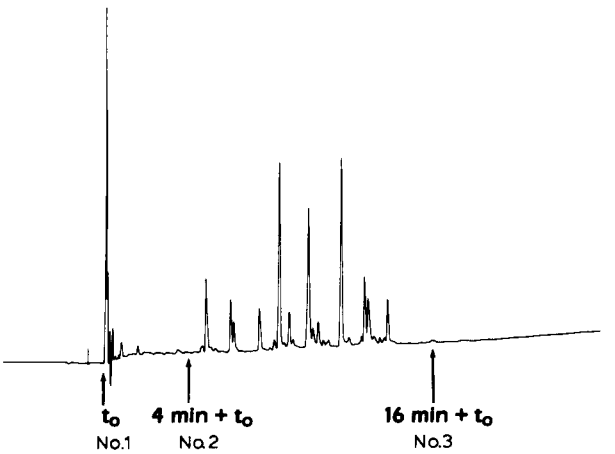

Fig. 1. Initial separation of lysozyme digest. Single Bioseries PEP-RP1 column, gradient from 5% solvent A to 60% solvent B (water–acetonitrile plus TFA–morpholine). Gradient time is 30 min, flow-rate is 1.89 ml/min, $\bar{k}$ equal 6.0.

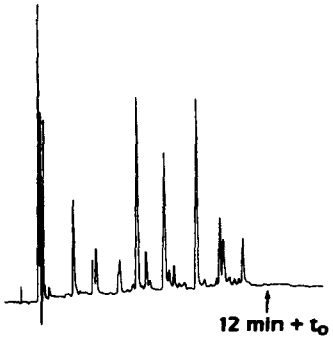

Fig. 2. Separation of lysozyme digest: reducing the gradient range. Conditions as in Fig. 1, except gradient from 12% solvent A to 34% solvent B, gradient time equal 12 min.

$F = 0.5$ ml/min (other conditions the same). Some improvement is evident for initial increase in t_G, but resolution approaches a maximum for gradient times of 60 min or larger. The reason for this diminishing effect of gradient time on resolution is that the column plate number N is approaching its maximum value at the lower flow-rates associated with t_G values of 60–120 min. This is made clearer in Fig. 4a, where we use our model to calculate resolution (expressed as peak capacity PC) as a function

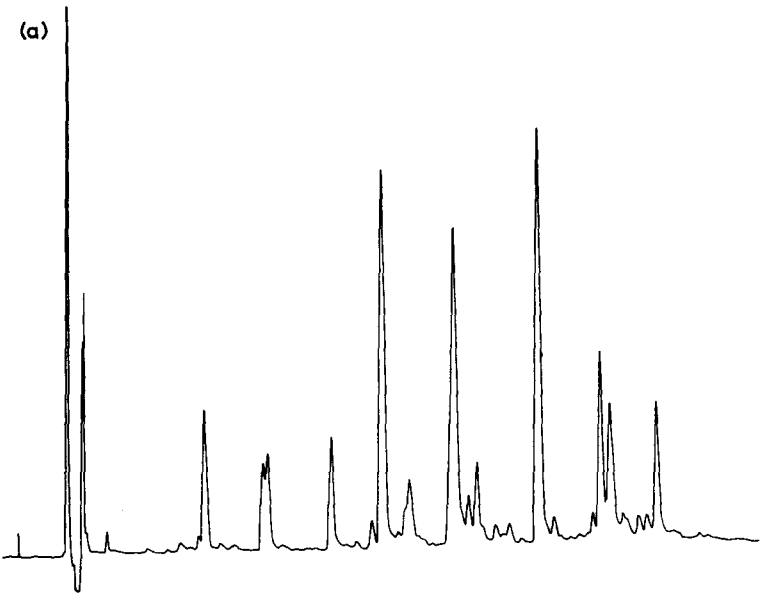

Fig. 3.

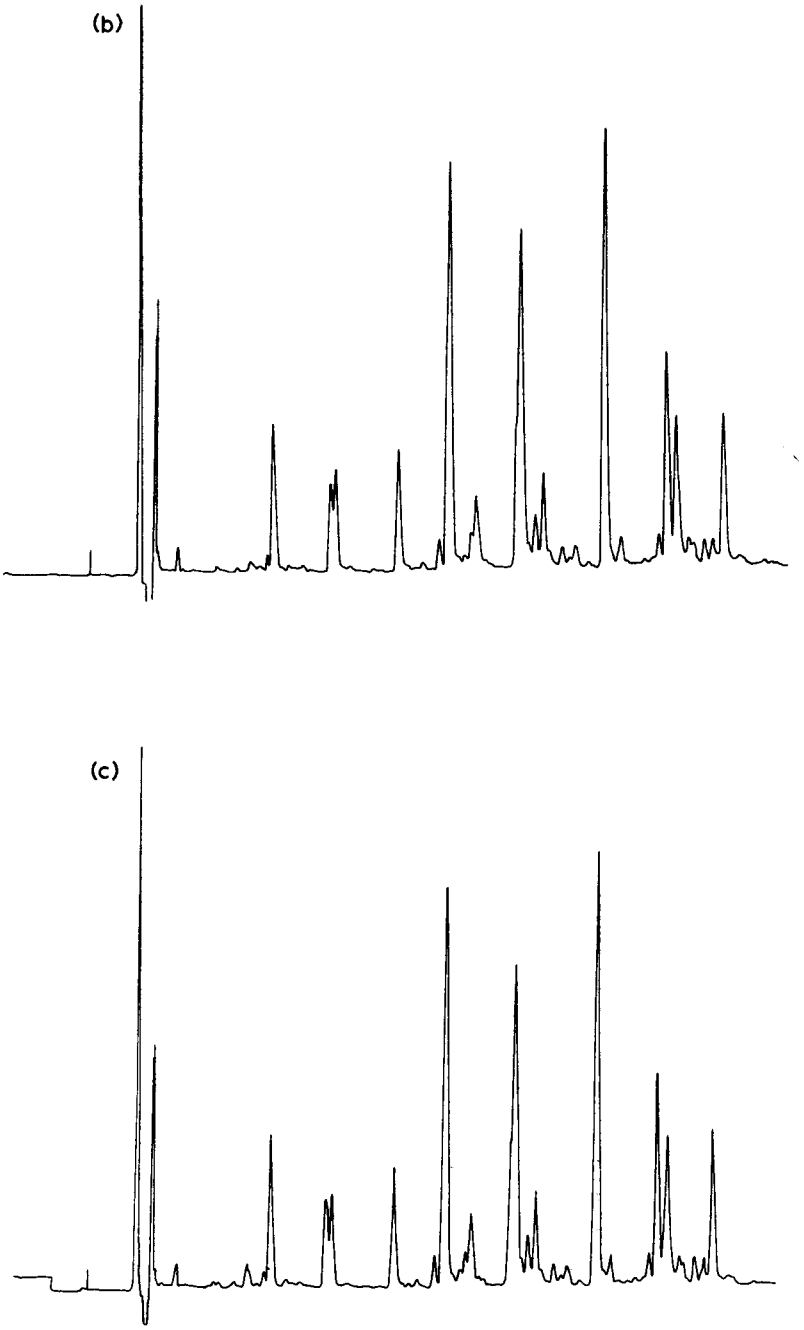

Fig. 3. Separation of lysozyme digest: increasing peak capacity at constant $\bar{k}$ (no change in band spacing). Single Bioseries PEP-RP1 column, gradient from 10% solvent A to 40% solvent B. (a) t_G = 15 min, F = 2 ml/min; (b) t_G = 30 min, F = 1 ml/min; (c) t_G = 60 min, F = 0.5 ml/min; $\bar{k}$ equal 5.8.

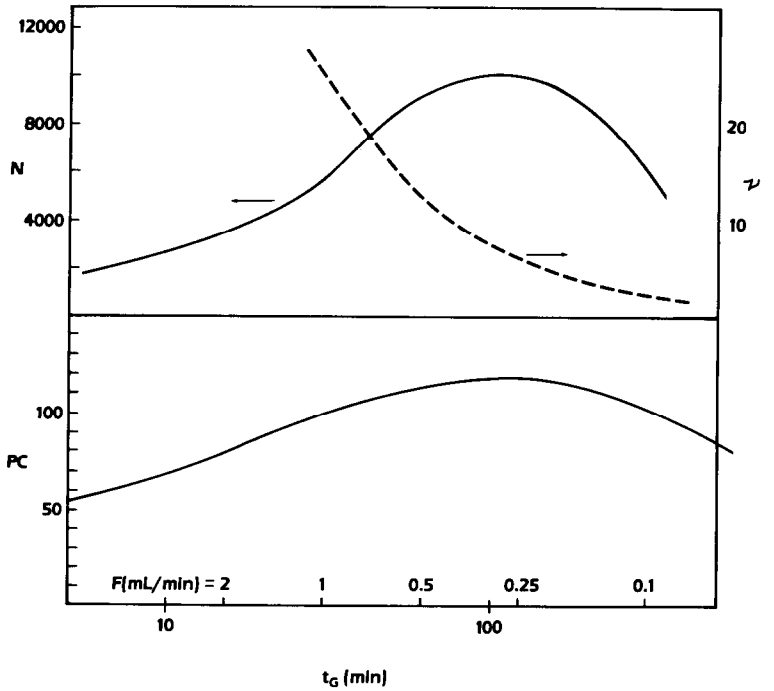

Fig. 4. Calculated values from model of peak capacity PC, plate number N and reduced velocity v *vs.* gradient time t_G. Assumes flow-rate varied to hold $\bar{k}$ constant ($\bar{k} = 6$), $\Delta\varphi = 0.3$, $\eta_{25} = 0.91$, temperature = 25°C, single Bioseries PEP-RP1 column ($L = 8$ cm), and $M = 1000$.

of t_G and F for the runs of Figs. 2 and 3*. We see here that PC has a maximum value for t_G equal 120 min, and then decreases for longer gradient times. We also see that peak height (proportional to $1/\sigma_v$) likewise is a maximum at 120 min. The plate number N and the reduced velocity v are plotted in Fig. 4b. N achieves a maximum value of 9800 plates for a reduced velocity of about 6.

Further increase in peak capacity by increase in column length. At this stage in the separation ($t_G = 60$ min, Fig. 3c), a significant further increase in N requires an increase in L. We can use the model to assess the relative improvement possible with different column lengths. Fig. 5 shows peak capacity as a function of both column length and gradient time, for columns of 8, 16 and 24 cm lengths (corresponding to 1, 2 or 3 columns in series). The maximum peak capacity of 118 for a single column is seen to increase to 165 for 2 columns in series and 205 for 3 columns. These represent significant improvements in resolution, but note also that that the gradient time for these maximum peak capacity values is increased in proportion to column length (4 h for 2 columns, 6 h for 3 columns). The advantage of longer columns for the lysozyme digest separation is illustrated in Fig. 6b, for two columns in series ($t_G = 60$ min, $F = 1$ ml/min, $\bar{k} = 5.8$, other conditions as in Fig. 2). Considerably

* The model calculations of Figs. 4, 5 and 8 are based on conditions which are slightly different from those chosen for experimental runs in Figs. 1–3, 6, 7 and 9, as can be seen in the captions for these figures. The model calculations are based on rounded values of $\Delta\varphi$ (0.30) and of M (1000).

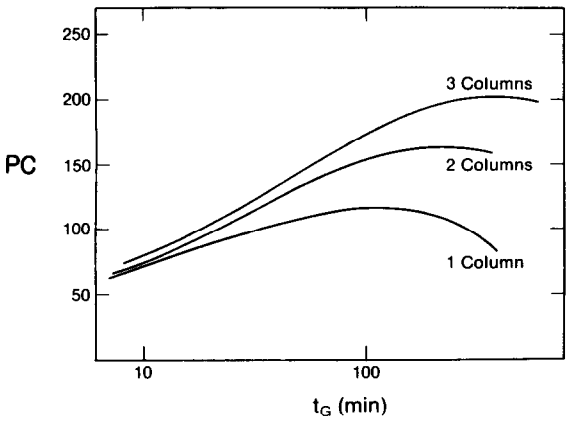

Fig. 5. Effect of column length on peak capacity as calculated from model. Conditions as in Fig. 4, except varying number of columns in series (1 column = 8 cm, 2 columns = 16 cm, 3 columns = 24 cm); $\bar{k}$ held constant at 6.0 by varying F (eqn. 1).

improved resolution *vs.* the corresponding 1-column separation of Fig. 6a (equal Fig. 3c) is observed. Thus a doublet in Fig. 6a (No. 1) is split into a triplet in 6b, a single peak in 6a (No. 2) becomes a doublet in 6b, a shoulder in 6a (No. 3) resolves into two peaks in 6b and all peaks at the end of the chromatogram of 6a (No. 4) are noticeably better resolved in 6b.

Note also in Fig. 5 that column length has little effect on PC or resolution when t_G is small (< 30 min); see also discussion of ref. 24). On the other hand, Fig. 5 suggests that the separation of Fig. 4b (2 columns, t_G = 60 min) can be further improved by increasing t_G to 2.5 h (PC = 162 *vs.* 140).

Variation of band spacing by changing flow-rate. At this point in the separation (Fig. 6b) we can change the relative positions of various bands in the chromatogram by changing flow-rate. This is illustrated for the two-column separation by changing F from 1.0 ml/min (Fig. 7a) to 2.5 ml/min in Fig. 7b. Several major changes in peak position are visible, with better separation of some bands in the run of 7a, and better separation of others in run 7b. Complete separation of the major peptides is possible, by selecting an optimum value of F as described in ref. 14.

Increasing detection sensitivity. In the above example the major peptides may not all be of equal interest. In this case an increase in detection sensitivity for an individual peptide may be more important than the complete resolution of other bands in the sample. Or for another sample which presents less difficulty in separation, the primary emphasis in method development may be on maximizing peak-heights. Eqn. 7 is the key to achieving this objective. According to this equation we desire to maximize $(1/\sigma_v)$ or the term $N^{1/2}/V_m(k_f + 1)$. The smallest value of V_m is achieved with a single column, which is the preferred choice for maximizing sensitivity (note that peak height is proportional to $N^{1/2}/V_m$ or to $L^{-1/2}$). A small value of k_f corresponds to small values of F, while N is a maximum for some intermediate (but small) value of F. These relationships are illustrated in Fig. 8a using calculations based on our model. Values of $(1/\sigma_v)$ are plotted *vs.* t_G for three different $\bar{k}$ values: 0.5, 1.5 and 6. Also shown (Fig. 8b) are corresponding peak capacity values PC. In

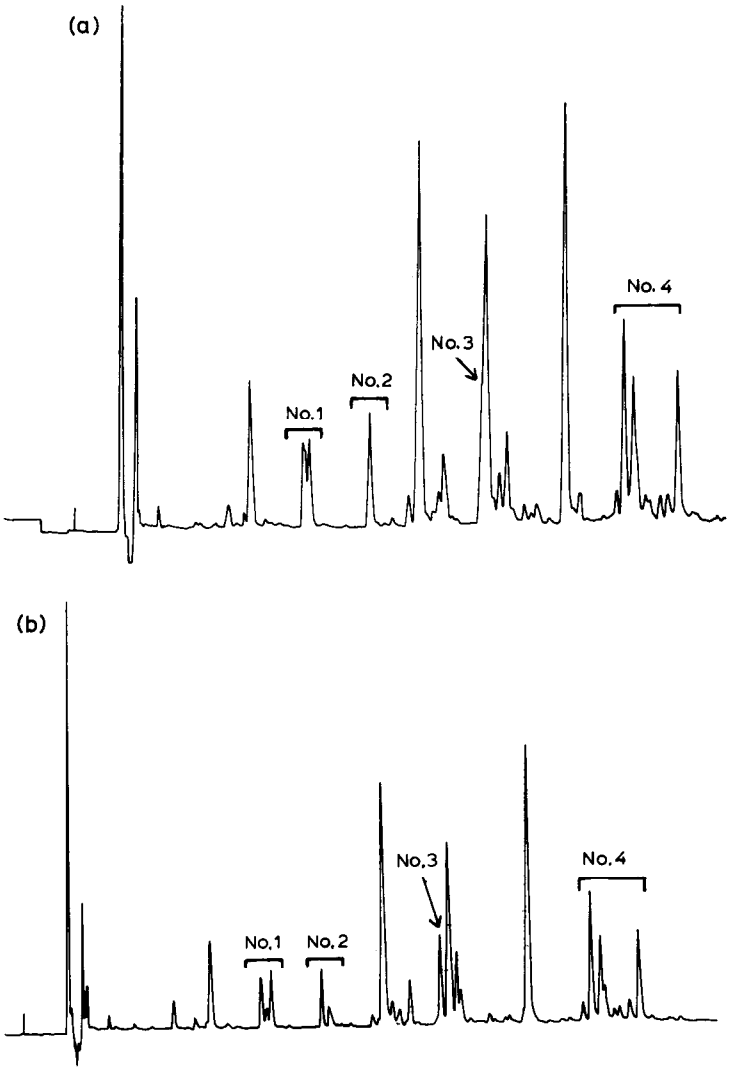

Fig. 6. Separation of lysozyme digest: further increase in peak capacity by increasing column length. Conditions as in Fig. 3 except as noted. (a) Same separation as Fig. 3c; (b) two columns in series (L = 16 cm), F increased from 0.5 ml/min in a to 1 ml/min in b ($\bar{k}$ equal 5.8 in each run).

general we see that peak heights increase for smaller $\bar{k}$ (and k_f) values as expected; maximum peakheights are observed at lower t_G values as $\bar{k}$ decreases (other factors the same). This is exactly analogous to the separation of small molecules by isocratic elution, where early peaks (short separation time) are narrower and taller *vs.* later peaks in the chromatogram. This is further illustrated in Fig. 9 for the separation of the lysozyme digest under conditions (t_G and F) that yield roughly maximum sensitivity for $\bar{k}$ values of 5.8 (Fig. 9a), 2.9 (Fig. 9b) and 1.5 (Fig. 9c). The injected sample is 40 μl in Fig. 9a, 20 μl in 9b and 10 μl in 9c — yet average peakheights decrease

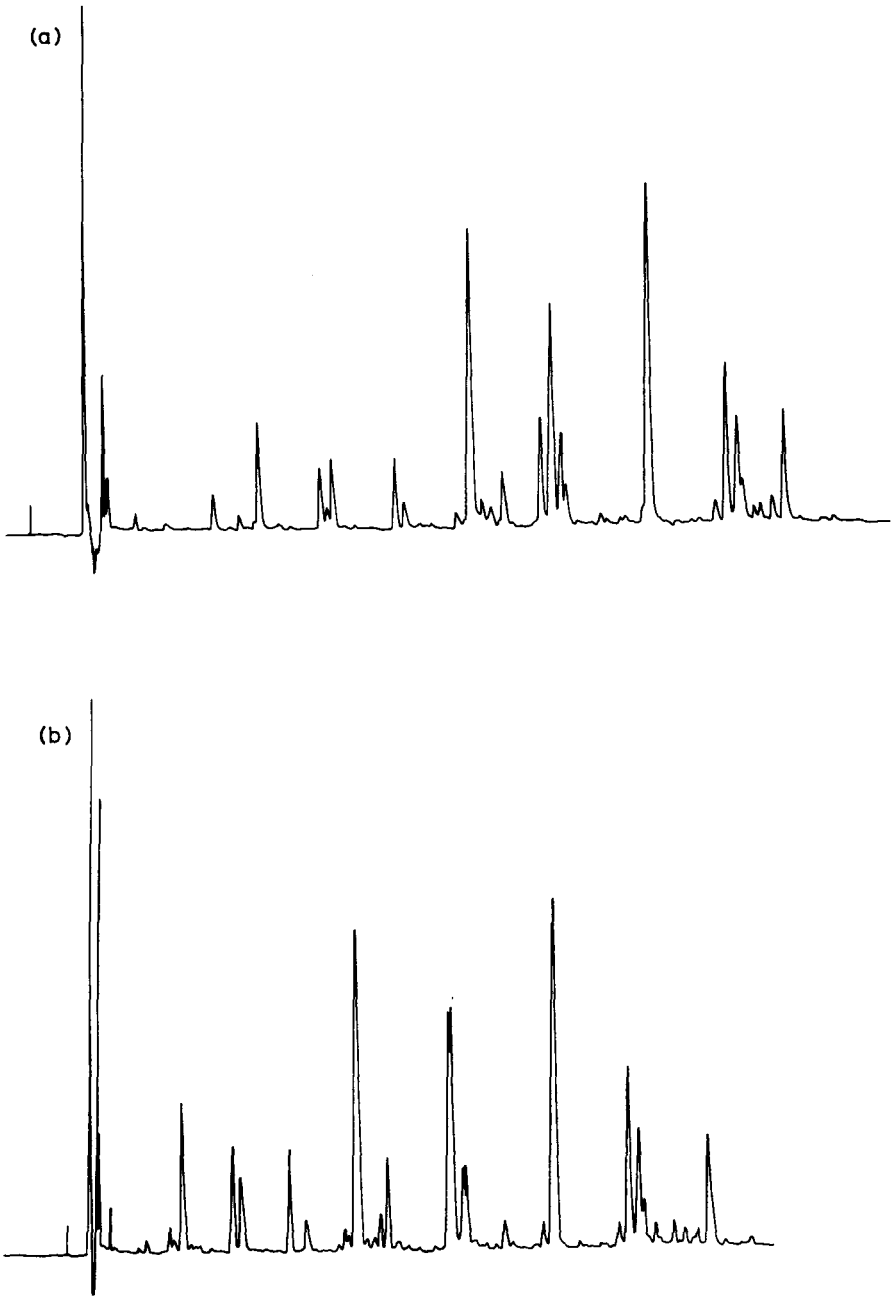

Fig. 7. Separation of lysozyme digest: variation in band spacing (separation selectivity) by changing flow-rate. (a) Separation of Fig. 3c; (b) same, except flow-rate changed from 0.5 ml/min (a) to 2.5 ml/min (b).

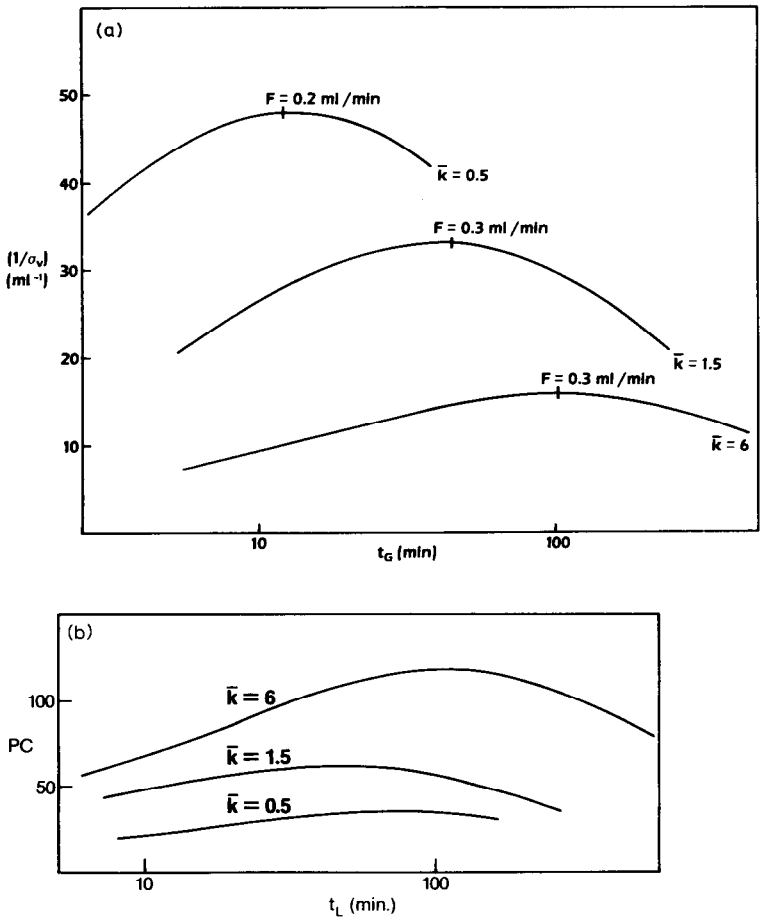

Fig. 8. Maximizing detection sensitivity by varying gradient time t_G and flow-rate ($\bar{k}$ varies). Calculated from model, using conditions of Fig. 3 except as noted. (a) Reciprocal band width ($1/\sigma_v$ — proportional to peak height) *vs.* t_G for different values of $\bar{k}$ (determined by choice of F and t_G); (b) corresponding values of peak capacity for conditions of a.

only slightly from run 9a to 9c. That is, we achieve a major increase in peakheight and sensitivity by operating at lower values of t_G and $\bar{k}$.

Experimental peak capacity values PC were obtained from several of the preceding lysozyme separations using the relationship

$$PC = t_G/4\sigma_t \tag{9}$$

Here σ_t refers to the band width (1 sigma) in time-units (min). These experimental PC values were compared with values calculated from the present model:

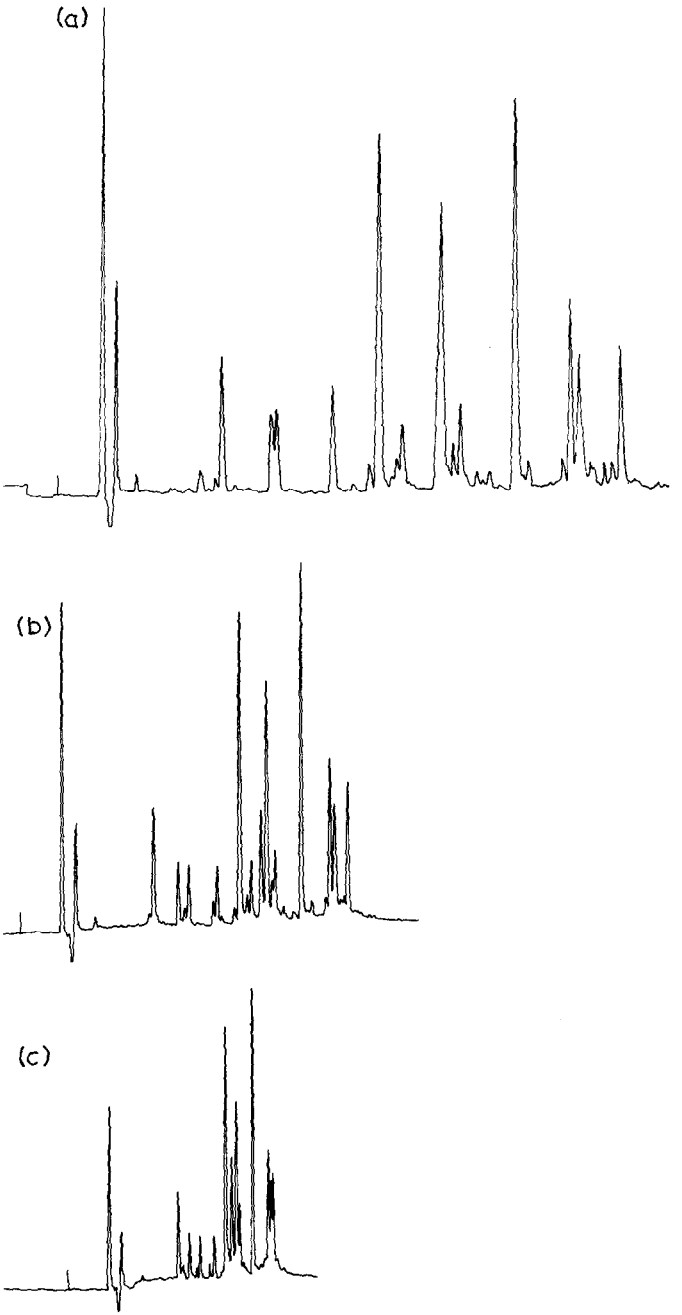

Fig. 9. Separation of lysozyme digest: increase in detection sensitivity by decreasing gradient time. Conditions as in Fig. 3 except as noted. (a) Same separation as Fig. 3c: t_G = 60 min, F = 0.5 ml/min, 40 μl sample injected; (b) same as in a, except t_G = 30 min, with 20 μl sample; (c) same as in a, except t_G = 15 min, with 10 μl sample.

Run	t_G (min)	F (ml/min)	L (cm)	PC	
				expt.	calc.
Fig. 3a	15	2	8	96	82
3b	30	1	8	112	103
3c	60	0.5	8	122	118
6b	60	1	16	146	145

Effect of sample molecular weight

In isocratic separations it is generally observed that an increase in solute molecular weight leads to a decrease in plate number (*e.g.*, refs. 21 and 23). This suggests that separations of larger peptide or protein molecules will be poorer and/or slower than corresponding separations of small molecules (see also discussion of ref. 24). There are several reasons for the less efficient separation of higher molecular weight samples by reversed-phase gradient elution, and it is important to distinguish among these effects if practical separations are to be achieved with minimum effort:

(i) comparison of N values at similar flow-rates, corresponding to larger v values for larger molecules;

(ii) increasing likelihood that the sample is not "well-behaved"[18–21], due to physico-chemical processes not recognized by the model;

(iii) slower diffusion of large molecules within small pores of the column packing.

N vs. v for large molecules (i). We have noted that larger molecules have smaller values of D_m and resulting larger values of v *vs.* the case of smaller molecules (and similar flow-rates). This dependence of v on M results generally in larger values

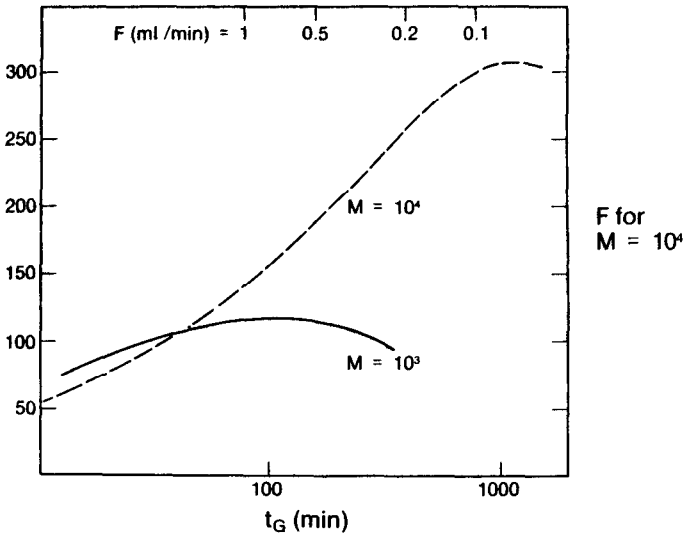

Fig. 10. Effect of sample molecular weight M on peptide separations, as calculated by model. Conditions assumed as in Fig. 4, except dashed curve is for M = 10 000 (*vs.* 1000 for solid curve).

of h and smaller values of N (eqn. 1) for large molecules. It also leads to maximum values of N and peak capacity at lower flow-rates and larger gradient times t_G, as illustrated in Fig. 10 for the present Bioseries PEP-RP1 column and samples having molecular weights of 1000 and 10 000, respectively. Whereas the maximum peak capacity for the 1000-dalton sample is achieved at a gradient time of 2 h, the corresponding optimum gradient time for the 10 000-dalton sample is about 20 h ($\bar{k}$ = 6 in both cases). On the other hand, for gradient times in excess of an hour, the peak capacity of the latter separation (M = 10 000, $\bar{k}$ = 6) is generally larger than for the lower-molecular-weight samples; this surprising result arises as follows. For a sufficiently low (*i.e.*, optimum) flow-rate, large molecules give approximately the same values of N as for small molecules at higher flow-rates (eqn. 3 and discussion of ref. 24). However peak capacity is also proportional to the value of S for the sample (eqn. 4) — which will increase with M (eqn. 2).

Non-well-behaved systems (ii). Various processes noted previously (secondary retention, slow configurational changes, multiple molecular states, etc., as in refs. 18–21) can cause experimental N values to exceed expected values by large factors. These effects can be recognized by comparing experimental σ_v values with those predicted by the model — experimental values that are greater by 50% or more suggest "non-well-behaved" bands, which can often be corrected by appropriate changes in the separation procedure. In some cases secondary retention of basic amino acid residues on silanol groups of the packing seems to be involved, which can be markedly reduced by addition of triethyl amine or morpholine to the mobile phase[2,25].

Combination of large molecules and small pores (iii). The diffusion of large molecules within narrow pores is generally slower than the diffusion of the same molecules in the bulk mobile phase. This effect is approximately described by a model based on spherical solute molecules diffusing within cylindrical pores (ref. 26; see also refs. 21 and 27). The latter model predicts that diffusion in small pores will be reduced (*vs.* bulk mobile phase) by a factor equal to

$$\rho^* = 1 - 2.10r + 2.09r^3 - 0.95r^5 \qquad (10)$$

Here r is the ratio: (solute Stokes diameter)/(pore diameter). The restriction factor ρ^* of eqn. 10 can be equated approximately to the restricted-diffusion parameter ρ of our present model — given by eqn. 6 for the Bioseries column as a function of solute molecular weight. At high flow-rates where the C term of eqn. 3 is dominant, N will be proportional to ρ, other factors equal.

Optimized column configuration for peptide separations

Consider next the various factors that enter into the design of optimum column configurations for the separation of peptide mixtures ($500 \leqslant M \leqslant 20\,000$). By column configuration is meant (i) the dimensions of the column, (ii) the diameter of the particles, and (iii) the pore-size/surface-area of the packing. The following factors are relevant to this discussion:

(1) the maximum peak capacity and sample resolution of which the column is capable under practical conditions;

(2) maximum detection sensitivity;

(3) compatability of the column with typical HPLC equipment, which varies

in extra-column band-broadening and flow-rate range;

(4) maximum recovery of separated sample;

(5) maximum column stability;

(6) maximum column loadability or capacity, allowing separation of larger amounts of sample per run.

Some compromise among the above goals will be necessary. For example, a large pore diameter favors the resolution and recovery of high-molecular-weight samples, but the lower surface area of such a packing will mean decreased column loadability for preparative separations and (frequently) less stable columns. Similarly the use of heavily-bonded polymeric phases (see ref. 22) may provide increased stability of the packing, but often at the expense of decreased column efficiency and lower peak capacities. In examining the above goals with respect to optimum column configuration, we will use the present model and the Bioseries column packing (eqns. 5 and 6) as a reference point for comparison.

Column pore size vs. performance. The effect of pore size on the separation is given mainly by the column parameter ρ — which is a function of M (eqn. 6). The Bioseries column has the following values of ρ for various sample molecular weights:

M	ρ
1000	0.48
10 000	0.31
20 000	0.27

At high flow-rates (large reduced velocities) N will be proportional to ρ. This means that the Bioseries columns shows to less advantage for molecules bigger than 20 000 daltons. However at lower flow-rates used for maximum resolution (Figs. 4, 5, 8 and 10) this effect will be less important. As Fig. 10 shows, the present column can do a good job in separating rather large molecules. In any case, the present 15-nm pore packing appears to be a good compromise for separating samples in the peptide molecular weight range (500–20 000 daltons).

A more important reason for preferring large-pore packings for the separation of larger molecules is that higher recoveries of sample are generally obtained (*e.g.*, refs. 17, 28 and 29).

Column particle size vs. performance. For the separation of large molecules, particle size should generally be as small as possible — if separation time is to be minimized (*e.g.*, ref. 24). Meek and Rossetti[23] have shown this for the case of peptide separations on 5- *vs.* 10-μm particles, and the present model confirms this*. However the use of particles smaller than 5 μm leads to problems with column plugging, unless stringent measures are taken. This difficulty is compounded for the case of "dirty" samples of biological origin, and for this reason the use of 5-μm particles in the Bioseries column seems to represent a good compromise.

Column dimensions vs. performance. Column length affects mainly the maximum peak capacity and related gradient time. This is illustrated for the case of the Bioseries column and $M = 1000$ as follows (see Fig. 5):

---

* The model suggests that an optimum particle diameter for samples of higher molecular weight ($500 \leqslant M \leqslant 20\,000$) is less than 2 μm. Apart from the difficulty in working with such columns, there is a problem in combining small particle diameters with larger pore diameters.

Column length (cm)	*Maximum peak capacity*	*Gradient time*
2	59	15–30 min
8	118	1–2 h
32	236	4–8 h

Very short columns have been used in separations of proteins, where the separation is often not improved by substituting a longer column (*e.g.*, 2-cm *vs.* 25-cm columns in ref. 30). However this is an artifact of the use of shorter (non-optimum) gradient times and reduced peak capacities — as illustrated in Fig. 5. When maximum sample resolution is required, particularly for lower molecular weight samples, longer columns are usually advantageous. The choice of an 8-cm length for the Bioseries column reflects the difficulty of average peptide sample separations; these typically require separation times of a few hours. Longer columns, where necessary, can be assembled by connecting 8-cm lengths in series. It should also be noted that shorter columns yield lower pressure drops, which are generally advantageous. The Bioseries column will normally be operated at pressures well below 1000 p.s.i.

Column diameter is chosen to maximize detection sensitivity and for compatibility with presently used HPLC equipment. The diameter of the column has no effect on peak capacity or resolution, if mobile phase flow-rate is reduced in proportion to decrease in column cross-sectional area. For maximum sensitivity we desire minimum values of σ_v, which according to eqn. 7 means minimum values of the column volume V_m (which is proportional to cross-sectional area). However the peak volume σ_v should not be smaller than the extra-column band broadening σ_{ec} of the HPLC system. Thus the final peak volume resulting from separation plus extra-column effects is

$$\sigma_{vt}^2 = \sigma_v^2 + \sigma_{ec}^2 \tag{11}$$

As a result, peak heights will vary with σ_v and σ_{ec} as follows:

σ_v/σ_{ec}	*Peak height* $(1/\sigma_v)$	*Loss in PC due to* σ_{ec} *(%)*
10	0.10	0
3	0.32	5
1	0.71	29
0.3	0.96	71
0	(1.00)	100

The achievement of σ_v values smaller than σ_{ec} ($\sigma_v/\sigma_{ec} < 1$) has only a minor effect on sensitivity, while resolution (or PC) rapidly decreases. Therefore there is little point in trying to reduce peak volumes below $\sigma_v = \sigma_{ec}$. A typical value of σ_{ec} for well-plumbed HPLC systems (with standard 8-μl flow cells) is about 0.02 ml[30], suggesting a maximum value of $(1/\sigma_v)$ equal to 50 ml^{-1}. Examination of Fig. 8a shows that the Bioseries column can produce maximum peak heights for $\bar{k}$ equal to about 0.5*. The Bioseries column configuration will generally be run with flow-rates

* The use of equipment with microbore capability (as for the system used in this study) allows further reduction in σ_v and increase in detection sensitivity. For such systems the use of narrower-diameter columns would prove beneficial, as in the case of other HPLC separations[31].

of 0.5–5 ml/min (see Figs. 4, 8 and 10) and with gradient volumes V_G greater than 15 ml. Present HPLC equipment is for the most part capable of providing acceptable gradients under these conditions (see discussion of ref. 10). The use of narrower, shorter columns would in most cases require special equipment for optimum performance.

CONCLUSIONS

A general model has been developed and experimentally validated for predicting the separation of peptide and protein samples as a function of experimental conditions in reversed-phase gradient elution[1,2]. This model has been applied here to a specific column (Du Pont Bioseries PEP-RP1), in order to provide a general approach to method development for the routine separation of peptide mixtures (500–20 000 daltons). The model can be used in conjunction with this approach to calculate values of average bandwidth for any combination of experimental conditions, and this capability can greatly facilitate the process of method development. Thus it is possible to optimize both overall sample resolution and detection sensitivity as a function of gradient conditions. A similar approach to method development is suggested for the separation of higher-molecular-weight samples.

The same model can also be used to explore different column configurations for the separation of samples of different molecular weight. The pore diameter of the column packing is of major importance, with a 15-nm-pore packing appearing optimum for separations of samples in the 500–20 000 dalton range. Larger solute molecules are better separated on columns having pores in the 30 to 50-nm range. The particle diameter of the packing is likewise critical, with 5-μm or smaller particles being optimum. However very small particles are less practical for the separation of samples of biological origin, because of risk of plugging the column. At the moment 5-μm particles appear to represent a good practical compromise. Column length is an important parameter, in contrast to earlier reports. While column length has little effect on separation for short gradient times, maximum sample resolution (or peak capacity) can only be achieved by longer columns and corresponding long gradient times. Here we recommend a column length of about 8 cm for typical separations of peptide samples, although higher resolution is possible by connecting several such columns in series. Column diameter directly affects assay sensitivity, and indirectly affects sample resolution for presently used HPLC systems (which have significant extra-column band-broadening). For commonly used equipment, a column diameter of about 0.6 cm appears generally best. This value allows near maximum detection sensitivity, without requiring microbore equipment to preserve separation.

REFERENCES

1 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 296 (1984) 31.
2 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 327 (1985) 27–46.
3 J. C. Giddings, *Dynamics of Chromatography*, Marcel Dekker, New York, 1965.
4 J. H. Knox and M. Saleem, *J. Chromatogr. Sci.*, 7 (1965) 614.
5 G. J. Kennedy and J. H. Knox, *J. Chromatogr. Sci.*, 10 (1972) 549.
6 J. H. Knox, *J. Chromatogr. Sci.*, 15 (1977) 352.
7 J. H. Knox and H. P. Scott, *J. Chromatogr.*, 282 (1983) 297.

8 R. W. Stout, J. J. DeStefano and L. R. Snyder, *J. Chromatogr.*, 282 (1983) 263.
9 L. R. Snyder, in Cs. Horváth (Editor), *High-Performance Liquid Chromatography*, Vol. 1, Academic Press, New York, 1980, p. 208.
10 M. A. Quarry, R. L. Grob and L. R. Snyder, *J. Chromatogr.*, 285 (1984) 1.
11 M. A. Quarry, R. L. Grob and L. R. Snyder, *J. Chromatogr.*, 285 (1984) 19.
12 M. A. Quarry, R. L. Grob and L. R. Snyder, *Anal. Chem.*, submitted for publication.
13 L. R. Snyder, *Anal. Chem.*, submitted for publication.
14 J. L. Glajch, M. A. Quarry and L. R. Snyder, *Anal. Chem.*, submitted for publication.
15 X. Geng and F. E. Regnier, *J. Chromatogr.*, 296 (1984) 15.
16 W. S. Hancock and J. T. Sparrow, in Cs. Horváth (Editor), *High-Performance Liquid Chromatography*, Vol. 3, Academic Press, New York, 1983, p. 49.
17 M. T. W. Hearn, in Cs. Horváth (Editor), *High-Performance Liquid Chromatography*, Vol. 3, Academic Press, New York. 1983, p. 87.
18 S. A. Cohen, S. Dong, K. P. Benedek and B. L. Karger, in I. Chaiken, M. Wilchek and I. Pasikh (Editors), *Symposium Proceedings, 5th International Symposium on Affinity Chromatography and Biological Recognition*, Academic Press, New York, 1984, p. 479.
19 S. A. Cohen, K. P. Benedek, S. Dong, Y. Tapuhi and B. L. Karger, *Anal. Chem.*, 56 (1984) 217.
20 K. A. Cohen, K. Schellenberg, K. Benedek, B. L. Karger, B. Grego and M. T. W. Hearn, *Anal. Biochem.*, 140 (1984) 223.
21 M. T. W. Hearn, *J. Chromatogr.*, 296 (1984) 61.
22 L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, Wiley-Interscience, New York, 1979, 2nd edn.
23 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.
24 L. R. Snyder, M. A. Stadalius and M. A. Quarry, *Anal. Chem.*, 55 (1983) 1412A.
25 K. A. Cohen, J. Chazaud and G. Calley, *J. Chromatogr.*, 282 (1983) 423.
26 R. R. Walters, *J. Chromatogr.*, 249 (1982) 19.
27 C. N. Satterfield, C. K. Colton and W. H. Pitcher, *Am, Inst. Chem. Eng. J.*, 19 (1973) 628.
28 J. M. Di Bussolo, *Amer. Biotechnology Lab.*, June 1984, p. 20.
29 M. A. Stadalius, *PhD Thesis*, University of Delaware, June 1984.
30 R. W. Stout, J. J. DeStefano and L. R. Snyder, *J. Chromatogr.*, 261 (1983) 189.
31 N. H. C. Cooke, K. Olsen and B. G. Archev, *LC Magazine*, 2 (1984) 514.