Journal of Chromatography, 296 (1984) 31–59
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 358

OPTIMIZATION MODEL FOR THE GRADIENT ELUTION SEPARATION OF PEPTIDE MIXTURES BY REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

VERIFICATION OF RETENTION RELATIONSHIPS

M. A. STADALIUS*
Biomedical Products Dept., E. I. Du Pont de Nemours and Co., Concord Plaza, Wilmington, DE 19898 (U.S.A.) and Department of Chemistry, University of Delaware, Newark, DE 19716 (U.S.A.)
H. S. GOLD
Department of Chemistry, University of Delaware, Newark, DE 19716 (U.S.A.)
and
L. R. SNYDER
Lloyd R. Snyder, Inc., 2281 William Court, Yorktown Heights, NY 10598 (U.S.A.)

SUMMARY

The presently accepted theory for gradient separations of small molecules has been used to develop a predictive model for peptides and proteins as samples, using reversed-phase high-performance liquid chromatography. Given the experimental conditions (gradient time, flow-rate, temperature, etc.), the molecular weight of the sample, and certain column characteristics (Knox parameters, column dimensions, particle diameter, etc.), it is possible to calculate the overall results of a given separation by gradient elution: peak capacity or average resolution, peak volume or relative peak height, etc. This information can in turn facilitate the optimized separation of any sample. The present model assumes that isocratic and gradient retention are interrelated for peptide molecules, in the same fashion as for small molecules. This assumption has been verified for various peptides and proteins and further used to gain new insight into the control of retention and band-spacing in gradient elution.

INTRODUCTION

The separation of bio-macromolecules such as nucleic acids, proteins and peptides is of major interest at the present time. Recently there have been dramatic improvements in our ability to resolve complex mixtures of these compounds by means of high-performance liquid chromatography (HPLC), particularly with columns based on ion-exchange or reversed-phase packings (*e.g.*, refs. 1–6). Several groups (*e.g.*, refs. 7–11) are engaged in further improving these separations by exploring the effect of changing the particles used to pack these HPLC columns (particle size, pore diameter, bonded-phase composition, etc.). Other workers have examined

0021-9673/84/$03.00

changes in mobile phase composition for the improvement of band shape, solute recovery and resolution (*e.g.*, refs. 12–17).

These macromolecule separations are typically carried out by means of gradient elution. The fundamental basis of the gradient separation of macromolecules is at present controversial[18], and it appears that many workers are not using gradient elution to best advantage —because they lack a good understanding of the principles of separation. A major problem is that both gradient elution and the chromatographic properties of macromolecules appear not to follow the simple rules observed for the HPLC separation of small molecules by means of isocratic elution. However a closer look at these apparent anomalies suggests that they are in fact predictable from simple "small-molecule" theory[18].

The fundamental aspects of gradient elution separation are now well understood and can be described quantitatively in simple terms[19]. Starting with this general model of gradient elution, several workers have successfully applied theory to the semi-quantitative interpretation of data from the gradient separation of peptides and proteins (*e.g.*, refs. 20–22). Quantitative agreement between theory and experiment has likewise been demonstrated for the gradient elution reversed-phase separation of polystyrenes in the molecular weight range 800 to 233,000 daltons[23,24], where isocratic retention data can be used to predict retention in gradient elution systems and *vice versa*. The corresponding prediction of bandwidth and resolution requires a quantitative understanding of column plate number as a function of separation conditions and column configuration. Recent data[25] suggest that such an understanding now exists and can be applied to the modeling of separations by gradient elution, either for large or small solute molecules.

In the present paper we will describe a quantitative model for the gradient separation of peptides and proteins by reversed-phase HPLC, starting with existing theory for the corresponding separation of small molecules[18,23]. We will also compare the predictions of this model with experimental retention data from our laboratory, as well as with data from a number of literature studies. Elsewhere we will provide a corresponding comparison of the model with data on bandwidths and resolution.

THEORY

A theory or model for the reversed-phase gradient elution of peptides should relate to the goals of separation and to the experimental conditions under our control. Table I summarizes some of the more important goals and separation variables. Under goals we require some minimum resolution of the sample or —for complex samples such as proteolytic digests— room in the chromatogram for a maximum number of separated peaks (IA). There is often interest in maximum detection sensitivity, or a minimum volume for each band (IB). We generally want short separation times (IC), but it will be seen that maximum resolution and peak capacity require longer times. Near-100% recovery of each compound is desirable, and the separated bands should have Gaussian, non-tailing shapes (ID). We prefer to use conventional equipment, characterized by standard flow-cell volumes (*e.g.*, 8 μl) and the usual pumping range (*e.g.*, 0.1–20 ml/min) (IE).

Concerning the separation variables of Table I that are controllable, these are

TABLE I

SUMMARY OF GOALS AND SEPARATION CONDITIONS FOR REVERSED-PHASE SEPARATION OF PEPTIDE SAMPLES

I.	Goals	
	A.	separation (resolution R_s or peak capacity PC)
	B.	sensitivity (peak volume, equal to $4\sigma_v$)
	C.	speed (equal to gradient time t_G)
	D.	peptide recovery, band shape
	E.	use of conventional equipment
II.	Separation variables	
	A.	Sample (molecular weight, chemical structure)
	B.	Equipment (flow-rate range F, extra-column volume σ_{ec}, detector time-constant τ, etc.)
	C.	Column configuration (dimensions L and d_c, particle diameter d_p and pore-size, composition of bonded phase, plate number N and permeability, etc.)
	D.	Experimental conditions (apart from mobile phase) (gradient time t_G, flow-rate F, column length L, gradient range $\Delta\varphi$)
	E.	Mobile phase composition (buffer, pH, choice of organic solvent, ion-pairing, etc.)
	F.	Use of denaturing conditions (temperature, addition of urea or guanidine, etc.)

individually related to our separation goals. We will not deal here with the question of peptide recovery and band shape (ID), which in turn is related to mobile phase composition (IIE) and the use of denaturing conditions (IIF); *e.g.*, ref. 26. Likewise the choice of sample (IIA) is not normally an option, but we need to know the effect of sample molecular weight on sample separation. For a given HPLC chromatograph its operating characteristics (extra-column band-broadening σ_{ec} and time-constant τ) should be known, as well as their effect on the resulting separation. The present model focuses on the effect of column configuration (IIC) and experimental conditions (IID) on separation (IA) and detection sensitivity (IIB). Experimental conditions are of primary concern to the chromatographer trying to optimize a given separation. Column configuration is the major interest of groups with responsibility for developing and evaluating new column designs for use in peptide separation. The following discussion is based on the theory for gradient elution as presented in ref. 19. Commonly used terms are defined in ref. 27 and in the glossary of this paper.

Interrelationship of gradient and isocratic retention

Values of k' in reversed-phase isocratic elution are usually given as a function of organic volume-fraction φ[19]:

$$\log k' = \log k_w - S\varphi \tag{1}$$

In some systems and especially for peptide solutes (*e.g.*, ref. 4), plots of $\log k'$ *vs.* φ are curved rather than linear as predicted by eqn. 1. In these cases eqn. 1 gives the tangent to this curve at some particular value of φ or k' (see ref. 24).

Solute retention time t_g in gradient elution is given as[19,23]

$$t_g = (t_o/b)\,[\log 2.3\, k_o b\, (t_{sec}/t_o) + 1] + t_{sec} + t_D \tag{2}$$

for macromolecular solutes and column-packings of small pore diameter. Here t_o is

the column dead-time for a small solute molecule, t_{sec} is the value of t_o for the solute in question (smaller than t_o for large, partially excluded molecules), t_D is the dwell-time of the gradient system (volumes of mixer plus all lines and other elements between mixer and column inlet, divided by flow-rate F; see ref. 28), k_o is the value of k' for the solute at the start of the gradient (in the initial mobile phase), and b is a gradient parameter defined by

$$b = S\Delta\varphi\, t_o/t_G \tag{3}$$

The quantity t_G is the gradient time and $\Delta\varphi$ is the change in φ during the gradient ($\Delta\varphi = 1$ for a 0–100% gradient). For smaller solutes and larger-pore particles, eqn. 2 can be approximated by

$$t_g = (t_o/b)\,[\log (2.3\, k_o b + 1)] + t_o + t_D \tag{4}$$

which we will use here unless noted otherwise.

Gradient retention data (t_g values) can be used to calculate corresponding isocratic retention data, as reviewed in refs. 23 and 24. Thus if a sample is run under the same gradient conditions, except that different gradient times t_{G1} and t_{G2} are used, we have the following relationships for retention times t_{g1} and t_{g2} for a given solute in the two runs (eqns. 3 and 4):

$$t_{g1} = (t_o/b_1) \log (2.3\, k_o b_1 + 1) + t_o + t_D \tag{5}$$

$$t_{g2} = (t_o/b_2) \log (2.3\, k_o b_2 + 1) + t_o + t_D \tag{6}$$

$$b_1/b_2 = t_{G2}/t_{G1} = \beta \tag{7}$$

Here t_{g1} and t_{g2} are retention times t_g for a given solute in experiments where t_G is equal to t_{G1} and t_{G2}, respectively. The quantities b_1 and b_2 are values of b in each experiment, and their ratio β is known from eqn. 3. Eqns. 5–7 involve three unknowns (k_o, b_1, b_2) which can be solved by an iterative trial-and-error approach[23]. Alternatively, for the case of peptides and proteins (where k_o is usually large), explicit solutions for k_o, b_1 and b_2 are possible (eqns. A2 and A3 of Appendix I). Given a value of b_1, a value of S can be calculated from eqn. 3. A value of k_w can be calculated from

$$\log k_w = \log k_0 + S\,\varphi_0 \tag{8}$$

where φ_0 refers to the value of φ at the beginning of the gradient. Finally, values of k_w and S thus derived from gradient data define the dependence of k' on φ in isocratic systems (eqn. 1). Where plots of log k' *vs.* φ (isocratic) are not known to be linear, eqn. 1 derived in this fashion is the tangent to the log k' *vs.* φ curve, at a value of $\varphi \equiv \bar{\varphi}$ corresponding to the mobile phase composition at the column midpoint, at the time the solute band is eluted halfway along the column. The corresponding value of k' at this point will be defined as $\bar{k}$. Both $\bar{k}$ and $\bar{\varphi}$ are of fundamental interest in

interpreting gradient separation (see below). Eqn. A9 (Appendix I) gives $\bar{\varphi}$ as a function of separation conditions and the value of t_g.

Peakwidth relationships in gradient elution

Resolution R_s in gradient elution is given as[19]

$$R_s = (\tfrac{1}{4})\,(\alpha - 1)\,N^{\frac{1}{2}}\,[\bar{k}/(1+\bar{k})] \tag{9}$$

where α is the separation factor (ratio of $\bar{k}$ values) for a band-pair when the bands are at the column midpoint, N is the column plate number at the same point in time, and $\bar{k}$ is the value of k' at the same point during separation. Thus α, N and $\bar{k}$ are each defined in terms of an equivalent isocratic separation. Our present discussion will center on maximizing R_s by maximizing the quantity $N^{\frac{1}{2}}\,\bar{k}/(1+\bar{k}) \equiv N^{\frac{1}{2}}Q$. Changes in α for purposes of increased resolution are discussed later in this paper.

Average resolution can also be defined in terms of peak capacity (PC), the number of bands that can fit into a given chromatogram (time equal to t_G) with $R_s = 1$ for all band-pairs. Peak capacity is given as

$$\mathrm{PC} = t_G/4\sigma_t = t_G F/4\sigma_v \tag{10}$$

Here σ_t is the bandwidth measured in time units (1 std. dev.), and σ_v is the bandwidth measured in volume units.

The effective k' value for a gradient separation is equal to $\bar{k}$, so far as resolution is concerned. It is given[19] as

$$\begin{aligned}\bar{k} &= 1/1.15\,b \\ &= F\,t_G/1.15\Delta\varphi\,S\,V_m\end{aligned} \tag{11}$$

cf. eqn. 3 and note that $t_o = V_m/F$. V_m is the column dead-volume, given by

$$V_m = (0.4/x)(\pi/4)\,d_c^{\,2}\,L \tag{12}$$

The quantity x is the fraction of mobile phase within the column not contained within the particle pores, d_c is the column inside diameter, and L is the column length. Eqn. 12 assumes that the volume of space between particles is equal to 40% of the total column volume (a good approximation for well-packed columns of rigid particles).

From ref. 19, σ_v can be derived as

$$\sigma_v = t_G F/2.3\,S\,\Delta\varphi\,N^{\frac{1}{2}}\,Q \tag{13}$$

Combination of eqns. 10 and 13 yields

$$\mathrm{PC} = (2.3/4)\,(S\Delta\varphi)\,N^{\frac{1}{2}}Q \tag{14}$$

For a given sample (value of S) and gradient range (value of $\Delta\varphi$), eqn. 14 predicts that maximizing peak capacity also means maximizing R_s (eqn. 9) *when α is maintained constant.* In gradient elution we can therefore use average resolution and peak capacity interchangeably as measures of "separation goodness".

It is seen in the above equations that the plate number N of the column plays a major role in both separation (R_s or PC) and detection sensitivity (or σ_v). In both cases, N should be as large as possible (other factors equal). Therefore we need to know how N varies with separation conditions, including the molecular weight of the sample. The plate number of a column can be expressed in terms of plate height H and column length L,

$$N = L/H \tag{15}$$

and H can be expressed by the Knox equation[25]:

$$h = B/v + Av^{\frac{1}{3}} + Cv \tag{16}$$

The reduced plate height h is given by

$$h = H/d_p \tag{17}$$

and the reduced velocity v is

$$v = ud_p/D_m \tag{18}$$

Here u is the linear velocity of the mobile phase moving through the column, d_p is the particle diameter, and D_m is the solute diffusion coefficient in the mobile phase. Eqns. 15–18 allow determination of N for various samples and separation conditions (see below) if we know the values of A, B and C in eqn. 16. Because values of v are typically large in peptide separations, the B/v term of eqn. 16 can be ignored; similarly, final values of N are not very sensitive to change in values of A. For well-packed columns of reversed-phase particles, A usually falls within the range $0.6 \leqslant A \leqslant 1.0$; A can be measured for a given column by fitting eqn. 16 to data for the elution of small solute molecules (see ref. 25). The major question concerns the value of C for a given peptide separation. The general theory for C as a function of experimental conditions has been expanded recently[25], allowing the prediction of C and N for peptide separations:

$$C = [(1-x+\bar{k})/(1+\bar{k})]^2 \, (D_m/D_p)/(1-x) \, 30 \, \gamma \, \rho \tag{19}$$

where

$$(D_m/D_p) = (1-x)/[(B/2\gamma) - x] \tag{20}$$

and

$$B = a' + b' \, \bar{k} \tag{21}$$

The quantity γ (tortuosity factor) can be set equal to 0.64, a' is typically *ca.* 1.1, b' is the surface-diffusion parameter (varying between 0 and 0.5), and ρ is the restricted-diffusion parameter (equal to one or less). Thus, given experimental values of x,

b' and ρ, values of C (and N) can be calculated as a function of experimental conditions. In practice x will be known (or can be set equal to 0.6–0.7), and values of b' and ρ will be determined by fitting experimental data to eqns. 19–21.

Development of the model

The quantity $\bar{k}$ is initially calculated via eqn. 11. Input values of F, t_G, column dimensions (L, d_c) and x will be available. These allow calculation of V_m (eqn. 12), and the quantity S can be estimated (see Fig. 6 and related discussion) from the approximate relationship

$$S = 0.48 \text{ (mol. wt.)}^{0.44} \tag{22}$$

for reversed-phase HPLC systems and acetonitrile as organic solvent in the mobile phase. The molecular-weight value in eqn. 22 is the estimate for the sample (average) and is not critical to the following calculations. If a value of S is known for the compound(s) in question, this value can be substituted for S from eqn. 22.

A value of N for the system is estimated next. The first step is to calculate the coefficient C from eqns. 19–21 as described above. The sample molecular weight can be used to estimate the solute diffusion coefficient in water at 25°C:

$$D_{w,25} = 10^{-5} [2.2 \text{ (mol. wt.)}^{-1/3} + 62/\text{(mol. wt.)}] \tag{23}$$

Eqn. 23 is empirical, but it accurately predicts the D_m values of various proteins in water at 25°C (40–900 kilodaltons[29]), and it mimics the Wilke–Chang equation[27] for molecular weights less than 1000 daltons. Eqn. 23 is empirically corrected for viscosity η and temperature T (°K) via the Wilke–Chang equation:

$$D_m = D_{w,25} (0.9/\eta) (T/298) \tag{24}$$

For acetonitrile–water mobile phases, the value of η is empirically related to the η value at 25°C (see ref. 30 for η values) as

$$\eta = \eta_{25} (298/T)^6 \tag{25}$$

TABLE II

APPROXIMATE AVERAGE VALUES OF η_{25} FOR ACETONITRILE GRADIENTS WITH INITIAL AND FINAL ORGANIC VOLUME FRACTION IN MOBILE PHASE AS φ_o AND φ_f, RESPECTIVELY

φ_o	η_{25} *for values of* φ_f				
	φ_f = *0.2*	*0.4*	*0.6*	*0.8*	*1.0*
0.00	0.90	0.92	0.91	0.85	0.77
0.20		0.95	0.91	0.83	0.75
0.40			0.87	0.77	0.68
0.60				0.67	0.59

Eqns. 23–25 thus yield a value of D_m for the HPLC system in question. Because separation temperatures for reversed-phase peptide separations are generally carried out near 25°C, and because optimum peptide separations begin with the sample in a denatured state, this approximate temperature correction of D_m should normally be adequate. The value of η_{25} assumed will be the average η_{25} value for the range in φ encountered during the gradient (Table II).

Here φ_o and φ_f refer to values of φ (acetonitrile) at the beginning and end of the gradient. For most acetonitrile–water gradients used to separate peptides and proteins, we can assume $\eta_{25} = 0.9$.

The linear velocity u is calculated from

$$u = LF/V_m \tag{26}$$

which with the value of D_m just obtained allows calculation of reduced velocity ν from eqn. 18. Values of A, C and ν can be substituted into eqn. 16 (ignoring B/ν) to give h, which with eqns. 15 and 17 gives the plate number N. Finally the peak capacity PC (or average resolution) and the band volume σ_v (inversely proportional to detection sensitivity or peak height) can be obtained from eqns. 13 and 14.

Additionally, the column pressure drop can be calculated, given the column flow-resistance factor (usually equal to 500–1000), as outlined in ref. 31. All calculations in the present model can be accommodated by a small programmable calculator such as the TI Model 59 (480 steps).

For well-designed HPLC equipment and columns of conventional particle size and dimensions, extra-column band-broadening will not normally be significant in separations of peptides and proteins. However, the extra-column contribution to σ_v (σ_{ec}) and the detector time-constant τ for the system can be corrected for. After values of H are determined using eqns. 15–18, corrected H values are calculated via eqns. 5–5b of ref. 32.

EXPERIMENTAL

Apparatus

The liquid chromatograph used in this study was a DuPont series 8800 with UV spectrophotometer (DuPont, Wilmington, DE, U.S.A.). Retention times for all isocratic and gradient data were obtained in real time using the DuPont PDP-10 DART Analysis System.

Reagents

HPLC grade acetonitrile (J. T. Baker, Phillipsburg, NJ, U.S.A.) and purified water from a Milli-Q system are the mobile phase solvents. Reagent grade trifluoroacetic acid (TFA) (Eastman Kodak, Rochester, NY, U.S.A.) was purified weekly by distillation over chromium trioxide. Gold-label morpholine (Aldrich, Milwaukee, WI, U.S.A.) and HPLC-grade phosphoric acid (Fisher Scientific, Fair Lawn, NJ, U.S.A.) were used as obtained. The peptides and proteins (leucine enkephalin, bradykinin, angiotensin, glucagon, insulin, ribonuclease A and lysozyme) were obtained from Sigma (St. Louis, MO, U.S.A.).

Over time, insulin is irreversibly converted into its desamido form. This con-

version is catalyzed by an acidic environment. Since the desamido form interfered with the peakwidth determinations at very low k' and $\bar{k}$ values, the desamido form of insulin was isolated preparatively and used in the isocratic and gradient study. The procedure is described below.

Column

The experimental packing used in this study is a 5 μm, 15 nm pore silica-based particle bonded with octyldimethylchlorosilane (C_8) and capped with trimethylchlorosilane. The 8.0 × 0.62 cm I.D. column configuration was packed using standard slurry-packing procedures.

Procedure

Protein and peptide retention data were obtained with the column described above in the isocratic and gradient mode. The aqueous modifier (A) contained 0.1% morpholine and 0.125% trifluoroacetic acid (TFA) so that the final pH was 2.2. Both 0.1% morpholine and 0.1% TFA were added to the organic modifier (acetonitrile or B). Samples were dissolved in the aqueous mobile phase with a resulting concentration of *ca.* 0.05 mg/ml for each component. The sample size, injected in to a Rheodyne valve with a 50-μl loop, varied between 5 and 15 μl. The column was thermostated at 35°C unless specified.

Conversion of insulin into its desamido form was accomplished by allowing 0.4 mg/ml of bovine insulin to stand at room temperature in an aqueous solution of 0.25% phosphoric acid. After 50% of the sample had been converted into the desamido form, 2-μl aliquots, which had been adjusted to pH 2, were injected onto the standard column described above and separated using a standard gradient (30–35% B/30 min; 1.50 ml/min). The collected fraction was then chromatographed isocratically at 68%A–32%B for a purity check. If the skew for that peak exceeded 0.30 the sample was re-purified using the same gradient conditions.

The column dead time t_o, was taken as the retention time for the unretained peak, uracil. All gradient elution experiments were carried out using linear solvent gradients. Gradient conditions are listed in Table III.

Interpretation of retention data from gradient elution studies required that the dwell time, t_D, of the HPLC system be known (t_D is the time required for the mobile phase to reach the column inlet). The dwell time was obtained by determining the extrapolated onset of a gradient whose mobile phase absorbance increases with volume fraction of the B reservoir. Note that the column was replaced with an equivalent length of 0.01-in. I.D. stainless-steel tubing. Details of this procedure are available in ref. 28.

EVALUATION OF THE PRESENT MODEL

We will next examine the ability of the present model to predict separation in reversed-phase gradient elution systems for peptides and proteins.

Experimental data from present study

Isocratic and gradient retention data were collected for seven peptides and a wide range of experimental conditions to confirm the reliability of the present model.

TABLE III

GRADIENT RETENTION DATA FOR SEVEN PEPTIDES AND PROTEINS

Conditions as in Table IV except where noted otherwise.

Solute	*Mol. wt.*	φ_o	φ_f	*F (ml/min)*	t_G *(min)*	t_g *(min)*	$t_g^{o\star}$		*log* $k_w^\star$	$S^{\star\star}$	*S.D. of fit****
							Expt	*Calc*			
Leucine enkephalin	600	0.05	0.45	1.0	10	13.26	9.0	8.9	3.18	11.3	0.006
		0.05	0.40	1.0	12	15.07	11.3	11.2			
		0.05	0.35	1.0	20	20.57	18.4	18.7			
		0.15	0.43	1.0	35	16.93	30.2	30.8			
		0.15	0.45	4.0	20	5.83	14.4	14.1			
		0.15	0.45	2.0	40	11.17	28.4	28.2			
		0.15	0.40	3.0	60	9.89	44.1	40.7			
Bradykinin	630	0.10	0.40	1.0	20	16.76	17.9	17.9	3.36	12.8	0.007
		0.15	0.43	1.0	35	16.22	29.5	29.7			
		0.15	0.45	4.0	20	5.77	14.4	13.7			
		0.15	0.40	3.0	60	11.03	39.2	40.4			
		0.15	0.45	2.0	28	9.91	21.2	20.5			
Angiotensin	1300	0.05	0.45	1.0	10	14.10	9.8	9.7	4.74	15.2	0.003
		0.05	0.40	1.0	12	16.33	12.5	12.5			
		0.05	0.35	1.0	20	23.63	21.5	21.7			
		0.15	0.43	1.0	35	23.84	37.1	37.2			
		0.15	0.45	4.0	20	9.72	18.3	18.0			
		0.15	0.45	2.0	40	18.66	35.9	35.9			
		0.15	0.40	3.0	60	23.73	57.9	57.1			
		0.15	0.45	2.0	28	15.31	26.6	26.3			

Glucagon	3500	0.10	0.40	1.0	20	24.73	25.9	25.8	8.10	22.3	0.004
		0.15	0.43	1.0	35	31.91	45.2	45.6			
		0.15	0.45	4.0	20	14.47	23.1	22.9			
		0.15	0.40	3.0	60	42.34	76.5	77.0			
		0.15	0.45	2.0	28	22.04	33.3	32.9			
Insulin	6000	0.05	0.45	1.0	10	14.86	10.6	10.6	7.87	22.7	0.005
		0.05	0.40	1.0	12	17.58	13.8	13.8			
		0.05	0.35	1.0	20	26.86	24.7	24.7			
		0.15	0.43	1.0	35	30.68	43.9	43.6			
		0.15	0.45	4.0	20	13.82	22.4	21.9			
		0.15	0.45	2.0	40	27.70	44.9	43.7			
		0.15	0.40	3.0	60	42.36	76.5	73.3			
		0.15	0.45	2.0	28	21.09	32.3	31.4			
Ribonuclease A	12,500	0.10	0.40	1.0	20	20.73	21.9	21.8	12.3	41.3	0.003
		0.15	0.43	1.0	35	25.38	38.6	38.8			
		0.15	0.45	4.0	20	11.49	19.9	19.7			
		0.15	0.40	3.0	60	34.03	68.2	67.8			
		0.15	0.45	2.0	28	17.62	28.9	28.2			
Lysozyme	14,000	0.05	0.45	1.0	10	15.81	11.6	11.7	15.4	39.2	0.006
		0.05	0.40	1.0	12	19.62	15.8	15.4			
		0.15	0.43	1.0	35	37.59	50.8	50.4			
		0.15	0.45	4.0	20	17.59	26.2	25.9			
		0.15	0.45	2.0	40	34.54	51.8	51.9			
		0.15	0.40	3.0	60	55.56	89.7	89.9			
		0.15	0.45	2.0	28	25.84	37.1	36.9			

* Calculated from values of t_g : $t_g^o = t_g + [\varphi_o/(\varphi_f - \varphi_o)]\, t_G - 5.5/F$ (dwell time = 5.5 min).
** Best fit to experimental t_g values.
*** Standard deviation of fit of calculated and experimental t_g values, expressed in units of φ.

TABLE IV

ISOCRATIC RETENTION DATA FOR INSULIN AS A FUNCTION OF MOBILE PHASE COMPOSITION φ AND FLOW-RATE F (IN ml/min)

Acetonitrile–water mixtures with 0.1% morpholine and 0.125% trifluoroacetic acid added (pH 2.2); 8 × 0.62 cm I.D. column packed with 15-nm-pore C_8-silica (d_p = 5 μm); temperature, 30°C.

φ	k'				
	F = 0.5	F = 1.0	F = 2.0	F = 4.0	F = 6.0
0.29		12.4	20.5	29.2	35.0
0.30		6.4	9.1	12.0	13.9
0.31	3.4	3.8	7.0		
0.32		2.8	3.5	4.2	4.8
0.34	1.2	1.2	1.6	2.0	2.5

We first studied the retention of insulin in the isocratic system of Table IV. It is seen that apparent k' values in this system vary by as much as three-fold with change in flow-rate, an effect not normally observed in the separation of small molecules by HPLC. Other workers have made similar observations (unreported data) and it is clear that these peptide–HPLC systems are more complex than simple theory suggests. However, the practical consequences of this complexity are minimal. We have further studied this dependence of k' on flow-rate for other systems, finding generally that the effect is dependent on the mobile phase selected, is generally more pronounced for larger peptides, and seems to be more significant for small-pore packings. As discussed elsewhere[28], variations in k' for macromolecules and other HPLC systems are conveniently measured in terms of the equivalent change in φ to maintain k' constant. On this basis, a two-fold change in k' for insulin in Table IV is equivalent to only a 0.013 change in φ (eqns. 1 and 22), which is relatively minor. In further comparisons of the isocratic data of Table IV with corresponding gradient data, k' values for a flow-rate of 2 ml/min will be assumed as representative.

Table III summarizes gradient retention data for the same system as in Table IV and seven different peptides varying in molecular weight from 600 to 14,000 daltons. Note that experimental conditions are varied over a wide range of values: flow-rates of 1–4 ml/min, gradient times t_G° ($t_G/\Delta\varphi$; *i.e.*, corrected to a 0–100% basis) of 25–240 min, gradient volumes $V_G^\circ = Ft_G^\circ$ of 25–720 ml. These data allow a stringent test of the ability of the present model to describe retention in systems such as that of Tables III and IV.

The various experiments of Table III involve different values of the starting (φ_o) and final (φ_f) values of φ for the gradient. It is convenient to convert the resulting experimental values of t_g to a 0–100%-gradient basis, with correction for t_D:

$$t_g^\circ = t_g + [\varphi_o/(\varphi_f - \varphi_o)]\, t_G - t_D \tag{27}$$

Here t_g° is the value of t_g that would have resulted if φ_o = 0.00 and φ_f = 1.00*. Table III lists resulting t_g° values calculated from experimental t_g values via eqn. 27. These

* For the present system, t_D (min) equals 5.5 divided by F in ml/min.

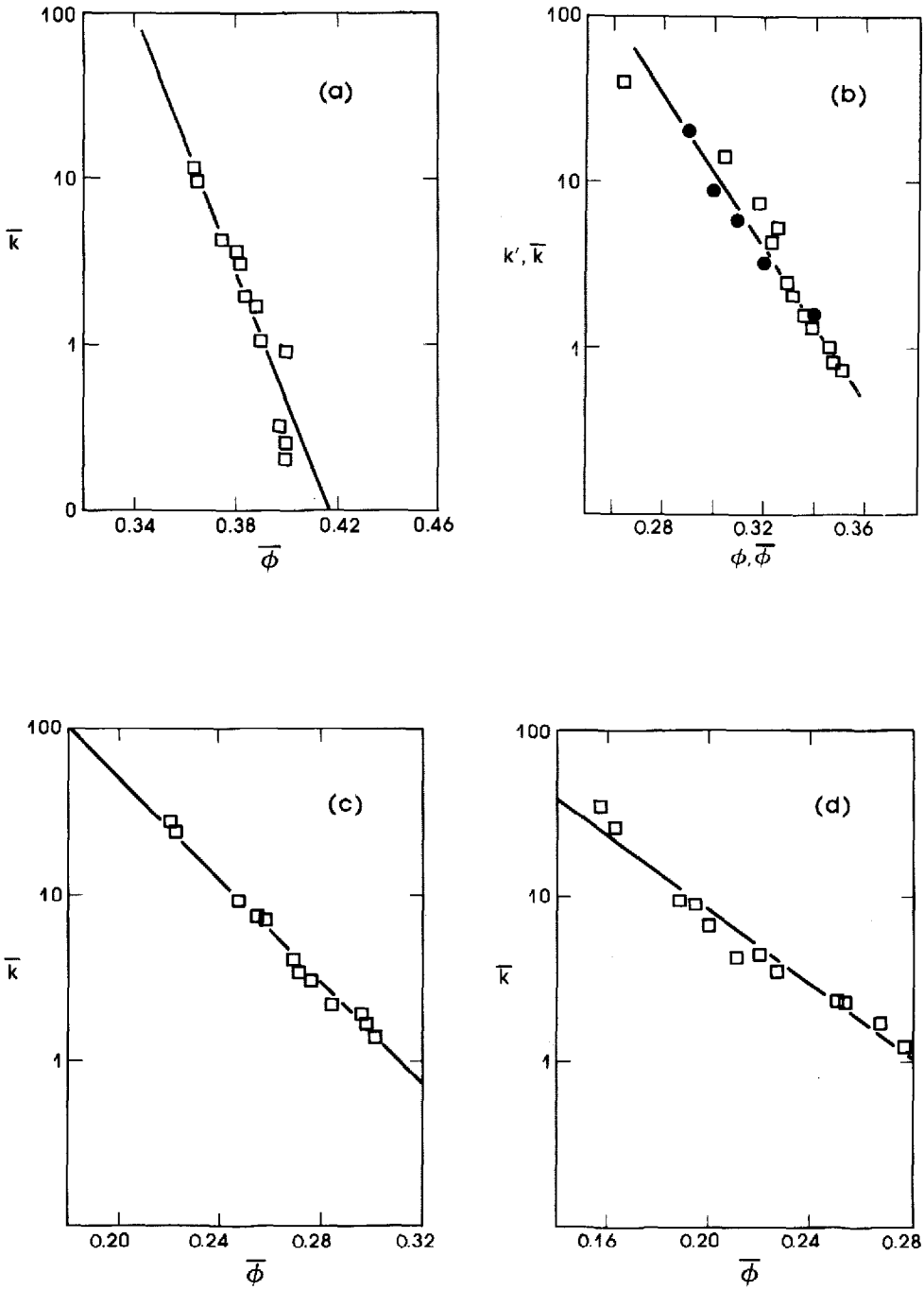

Fig. 1. Derived plots of k' *vs.* φ based on gradient data of Table III, as described in text and Appendix I. (a) Lysozyme, (b) insulin, (c) angiotensin, and (d) leucine enkephalin; ——, best fit to data points (eqn. 1); □, data derived from gradient data; ●, isocratic data for $F = 2$ (Table IV).

t_g^o values can be used (see eqns. 5–7 and further discussion of Appendix I) to calculate isocratic values of $\bar{k}$ *vs.* $\bar{\varphi}$, as well as values of k_w and S for each solute. Table III summarizes these latter parameters. Fig. 1 shows resulting plots of $\bar{k}$ *vs.* $\bar{\varphi}$ (or k' *vs.* φ) for four of the seven peptides of Table III. Solute molecular weight decreases in going from Fig. 1a (lysozyme) to Fig. 1d (leucine-enkephalin), and the slopes S of these plots are likewise seen to decrease as predicted by eqn. 22. In each case, linear plots of log $\bar{k}'$ *vs.* $\bar{\varphi}$ are observed, as predicted by eqn. 1.

For insulin as solute (Fig. 1b), isocratic data from Table IV (F = 2 ml/min) are plotted as dark circles on top of the data derived from the gradient experiments of Table III. It is seen that these data fall close to the solid curve defined by corresponding data from the gradient experiments. The present model predicts that values of k' *vs.* φ determined by either gradient or isocratic means should give the same relationship (k' *vs.* φ) and this is seen to be the case.

The derived (best-fit) values of log k_o and S from Table III can be used to calculate values of t_g^o from eqn. 4. These t_g^o values are listed in Table III alongside the experimental values; good agreement between the two sets of data is observed. The standard deviation of the fit (units of φ) is given in the last column of Table III. The average variation between experimental and calculated t_g^o values is only ± 0.005 unit in φ (1 S.D.), which is similar to the agreement found between experimental and calculated t_g values for small molecules[24,28]. Thus the data of Tables III and IV suggest that the present model is as accurate as is generally the case for small-molecule separations.

Data of Meek and Rossetti[33]

This study is one of the more complete comparisons of gradient *vs.* isocratic elution of peptides. It is well suited for testing the present model because:

(1) Gradient conditions were varied over wide limits: $20 \leqslant t_G^o \leqslant 200$ min, $0.5 \leqslant F \leqslant 2.5$ ml/min;

(2) A wide range of peptides (3–51 residues) was studied;

(3) Isocratic plate numbers were measured for several peptides as a function of F ($1 \leqslant k' \leqslant 2$);

(4) Peak capacity and peak height values were reported for the gradient runs as a function of t_G and F;

(5) Mobile phase (buffer) conditions were initially optimized for minimum band-tailing and maximum plate numbers; therefore secondary retention effects are likely to be minimal.

In addition, this study allows derivation of S values for use in developing a general S *vs.* solute-mol.-wt. relationship.

Retention time data. Values of t_g as a function of t_G and F were obtained from Dr. Meek (unreported in ref. 33), and are summarized for insulin in Table V. The value of V_m for the column (25 × 0.40 cm I.D. of 8-nm-pore C_{18} packing) was estimated at 2.2 ml. The value of V_d (equal $t_D F$; see ref. 28) can be estimated at *ca.* 1.0 ml, from the 1-ml gradient mixer used in a high-pressure mode.

We first examined these data in terms of derived values of S and k_o, using the approach covered by eqns. 5–7 above. A previous analysis[24] has shown that the precision of derived values of S and k_o will be greater for larger ratios of t_G for the two experiments (t_{G1}, t_{G2}), so we used data from t_G^o values of 20 and 100 min, with

TABLE V

GRADIENT RETENTION DATA FOR INSULIN IN SYSTEM OF MEEK AND ROSSETTI[33]

Comparison with values calculated from eqn. 4 (S = 12.7, k_o = 5.90 · 10^5). Experimental t_g values are unreported data provided by Dr. Meek; conditions: 0–60% gradient of 0.1 M $NaClO_4$–0.1% H_3PO_4 (A) and acetonitrile (B), 25 × 0.4 cm I.D. column of 10 μm C_{18} silica (BioRad), room temperature.

Flow-rate (ml/min)	*t_g (min) [expt/(calc)]**			
	t_G^o = 20 min	*40 min*	*100 min*	*200 min*
0.5		26.9	53.9	96.8
		(26.2)	(52.7)	(94.3)
1.0	13.3	22.2	47.4	86.9
	(13.1)	(22.0)	(47.1)	(86.3)
1.5	11.8	20.5	44.8	
	(11.7)	(20.4)	(44.7)	
2.0	11.0	19.5	43.4	81.9
	(11.0)	(19.5)	(43.2)	(80.0)
2.5	10.5	18.8	42.4	
	(10.5)	(18.9)	(42.1)	

* Experimental t_g values given with calculated value (eqns. 2 and 3 in parentheses) below; the latter assumes V_D = 1.0 ml, V_m = 2.2 ml.

F varying from 0.5 to 2.5 ml/min. In the case of insulin as solute, the results obtained are listed in Table VI. The agreement of S values for these different pairs of experiments is reasonable; the slight upward trend in S for smaller F values (and lower φ values) suggests curvature of the log k'–φ plot for isocratic data, as has been found in other peptide systems[20].

The average values in Table VI of S (12.7) and log k_o (5.77) can be inserted into eqns. 3 and 4 to test the ability of these equations to predict values of t_g over a wide range in values of F and t_G, as used in this study. These calculated values of t_g are compared with experimental values from ref. 33 in Table V. The agreement observed is generally quite close, and can best be expressed in terms of differences in φ at elution: $\delta\varphi$ = ±0.007 (1 S.D.). This again compares favorably with the agreement found in similar gradient/isocratic comparisons for small molecules[28].

TABLE VI

CALCULATED VALUES OF S, log k_o AND $\bar{\varphi}$ FOR INSULIN (MEEK DATA) AS A FUNCTION OF FLOW-RATE, USING GRADIENT RETENTION DATA (t_{G1}^o = 20 min AND t_{G2}^o = 200 min) OF TABLE V

Flow-rate (ml/min)	*S*	*log k_o*	*Average φ*
1.0	14.1	6.34	0.415
1.5	13.5	6.09	0.424
2.0	12.3	5.63	0.430
2.5	10.7	5.00	0.441
Average	12.7 ± 1.5	5.77	

TABLE VII

SUMMARY OF CALCULATED S AND LOG k_o VALUES FOR PEPTIDES IN SYSTEM OF MEEK AND ROSSETTI[33]

See text for details. Symbols: TRH, thyrotropin-releasing hormone; GW, glycine tryptophan; Y_3, Tyr–Tyr–Tyr; ME, [Met]enkephalin; BK, bradykinin; P, papain; NT, neurotensin; SS, somatastatin.

Peptide	*Mol. wt. (kdalton)*	*S*	*log k_o*	*φ (range)*
TRH	0.4	8.4 ± 1.1	0.91	0.03–0.13
GW	0.3	5.2 ± 0.5	1.55	0.09–0.21
Y_3	0.5	6.8 ± 0.9	1.97	0.14–0.29
ME	0.6	6.4 ± 0.6	2.24	0.18–0.34
BK	1.1	7.4 ± 0.9	2.88	0.25–0.34
P	1.3	8.0 ± 0.7	3.44	0.29–0.43
NT	1.6	8.1 ± 0.7	3.37	0.29–0.42
SS	1.8	8.7 ± 0.7	3.86	0.33–0.41
Insulin	6.0	12.7 ± 1.5	5.77	0.39–0.48

Data for several other peptides studied in ref. 33 were treated similarly. Table VII summarizes derived values of S and log k_o for these solutes. Again, reasonable agreement of S values between experiments with the same solute is observed. In most cases the same tendency toward larger S values at lower F values and higher $\bar{\varphi}$ values was found (as for insulin above).

The experimental data of Table V (t_G^o eqn. 20 and 200 min) and corresponding data for other peptides (unreported data supplied by Dr. Meek) could be used to calculate values of $\bar{k}$ *vs.* $\bar{\varphi}$ as described in Appendix I and eqns. 5–7. Fig. 2 compares these log k' *vs.* φ plots for both isocratic and gradient elution (isocratic data provided

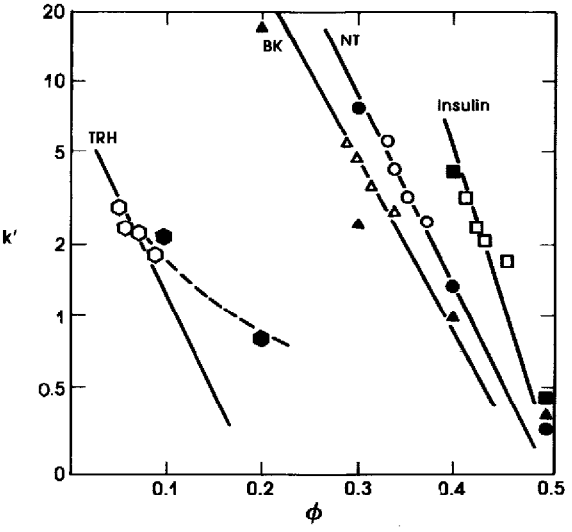

Fig. 2. Variation of capacity factor k' with organic volume-fraction φ for peptides in system of Meek and Rossetti[33]. Experimental conditions as in Table V, except φ fixed (isocratic). Solid points are isocratic data (courtesy of Meek), open points are calculated from gradient experiments as described in text. Solid curves are from S and log k' values of Table VII.

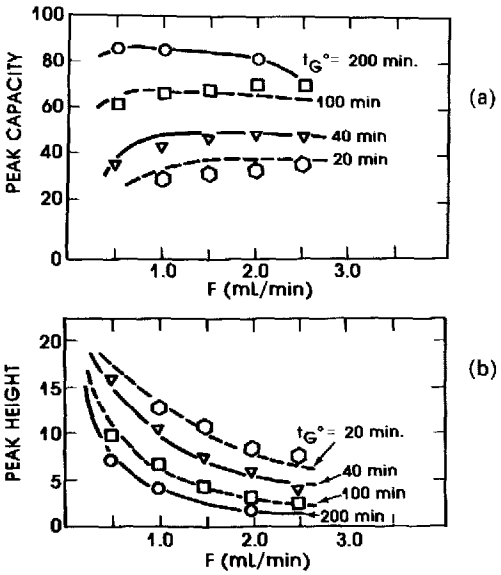

Fig. 3. Variation of (a) peak capacity and (b) peak heights with gradient time t_G° and flow-rate F (system of Meek and Rossetti[33]). Conditions as in Table V. Data points are experimental values, curves are calculated values as discussed in text (S = 10.4, b' = 0.1, ρ = 0.3, A = 0.6, x = 0.6, σ_{ec} = 0.05 ml, mol. wt. assumed to be 1100).

by Dr. Meek). Data derived from isocratic experiments are shown as dark circles, squares, etc., while corresponding data from gradient experiments are given as open circles, squares, etc. Both sets of data fall reasonably close to calculated curves from eqn. 1, using the parameters of Table VII. Thus the present model is further confirmed by these data. There is some tendency toward curvature of these plots as noted earlier, especially for TRH in Fig. 2.

Data on peak capacity and peak height. Meek and Rossetti[33] have reported data (10-μm column) on peak capacity and peak height as a function of gradient time t_G and flow-rate F. A mixture of peptides with an average molecular weight of *ca.* 1100 daltons was used. Peak capacity and relative peak height values are predictable for this system from eqns. 13 and 14, assuming that values of the column parameters (x, A, b' and ρ) are known. Elsewhere we will report a detailed analysis of these data. Here we note only that best-fit values of b' = 0.1 and ρ = 0.3 are suggested by the present model, with other assumed parameters given in Fig. 3 (S = 10.4 is obtained from eqn. 22). The resulting fit of the experimental data by the present model is shown in Fig. 3, where peak capacity (Fig. 3a) and peak height (Fig. 3b) values are plotted *vs.* mobile phase flow-rate F and gradient time t_G°. The data points are experimental values and the solid curves are those predicted by the model. The overall deviation of points from the curves is *ca.* ± 10% (1 S.D.). The study of Meek and Rossetti provides a rather thorough test of the present model; agreement between experimental data and theory is generally good.

Data of Cooke et al.[22]

These workers reported gradient separations of proteins with acetonitrile–

TABLE VIII

SUMMARY OF ISOCRATIC PARAMETERS, log k_o AND S, FOR LYSOZYME AND CARBONIC ANHYDRASE

Derived from isocratic retention data supplied by Cooke *et al.*[22] (see text for details).

Protein	*Organic solvent*	*log* k_w	S
Lysozyme	Acetonitrile	14.91	45
	1-Propanol	11.26	56
Carbonic anhydrase	Acetonitrile	19.43	48
	1-Propanol	19.18	74

water gradients from a reversed-phase column. They have also supplied us with unreported additional data from that study. The data available to us included isocratic and gradient retention data for a small number of corresponding HPLC systems, as well as peak capacity values for several proteins as a function of varying gradient time and flow-rate.

Retention time data. Isocratic retention data were provided for two proteins in both acetonitrile–water and 1-propanol–water systems. These yielded the isocratic parameters corresponding to eqn. 1 (Table VIII). Gradients were made with 0.01 M TFA–water (A) and either 0.01 M TFA–acetonitrile (B) or 0.01 M TFA–45% 1-propanol (B) with t_G^o = 100 min and a flow-rate of 1 ml/min. The column dimensions were 7.5 × 0.46 cm I.D., with d_p = 10 μm. A value of x = 0.7 was assumed. Eqns. 4 and 11 were used to calculate values of t_g for comparison with experimental values (Table IX; corrected for dwell time t_D): the above agreement between experimental and calculated t_g values is seen to be satisfactory ($\delta\varphi$ = ±0.007, 1 S.D.).

Data on peak height and peak capacity. A standard mixture of seven proteins was chromatographed with acetonitrile–water gradients and widely varying values of t_G and F. These data provide a rather stringent test of the present model, similar to

TABLE IX

SUMMARY OF EXPERIMENTAL AND CALCULATED GRADIENT RETENTION TIMES FOR LYSOZYME AND CARBONIC ANHYDRASE

Data supplied by Cooke *et al.*[22]. Gradients: linear 1%/ml, beginning at 0% organic. Value of (t_o + t_D) at 1.00 ml/min is 4.4 min.

Solute	*Organic solvent*	t_g *(min)*	
		Expt.	*Calc.*
Lysozyme	Acetonitrile	34.4	33.4
	1-Propanol	42.1	40.8
Carbonic anhydrase	Acetonitrile	20.8	20.8
	1-Propanol	26.2	26.8

that of Fig. 3 for the study of Meek and Rossetti. A preliminary application of the model to these data shows comparable agreement as in Fig. 3, thus confirming the model for peptides and proteins whose molecular weights span the range 600–70,000 daltons. A detailed analysis of these and other data will be given in a further paper. It should be stressed at this point that these comparisons of model and experimental results for peak capacity and peak height show the ability of the model to fit experimental data. No previously reported study contains sufficient experimental data to allow a rigorous check of the present model as regards band width, peak capacity, etc. Such a study based on our own data is now in progress.

Data of Van der Zee and Welling[34]

Van der Zee and Welling have reported some interesting data on the variation of t_g values of proteins in reversed-phase gradient elution as only column size is varied. These values are summarized in Table X, where the solutes studied are arranged in order of decreasing molecular weight. The authors noted that the change in t_g in going from a small column (3 × 0.3 cm I.D., V_m = 0.15 ml) to a larger column (30 × 0.46 cm I.D., V_m = 3.5 ml) increased as solute molecular weight decreased. They ascribe this to a size-exclusion (gel-filtration) phenomenon. However, such an effect is also a natural consequence of increasing S values with increase in solute molecular weight (eqn. 22), and it is useful to establish which effect actually predominates.

Even more interesting is the possibility of varying band-spacing within the chromatogram by changes in column dimensions or other experimental parameters. Thus in Table X it is seen that β-galactosidase and lysozyme have similar t_g values with the larger column (V_m = 3.5 ml): 19.9 and 19.6 min, respectively. With the smaller column (0.15 ml), however, these two compounds are well separated: t_g = 17.0 and 15.0 min, respectively.

We first examined the ability of eqns. 3 and 4 to predict the t_g values of Table X. Because of the unusual solvents used in the gradient (*n*-butanol–ethanol–2-methoxyethanol mixtures), estimation of values of S (see below) was somewhat uncertain. However, it appears that this particular solvent system fortuitously yields values of S that are similar to those found for various proteins with acetonitrile. Therefore, S values were estimated from eqn. 22, and $\Delta\varphi$ = 0.87. Values of V_m were calculated for the three columns used, assuming x = 0.7 (eqn. 12). The ratio of V_{sec} to V_m for the individual proteins (see Table X) was assumed the same as for a similar column from the literature[35]. The solute molecular weights correspond to the denatured monomeric subunit as suggested by the data of ref. 34. Finally, values of log k_o in Table X are best-fit values of eqns. 3 and 4 to the experimental t_g values.

The resulting calculated t_g values are in close agreement with the experimental values of ref. 34; the standard deviation for the entire data set of Table X is only ±0.2 min, close to the experimental precision of the data. More important is the insight that this correlation provides into changes in retention and band-spacing as a result of change in b and $\bar{k}$. Whenever experimental conditions are changed so as to cause variation in b (and $\bar{k}$), changes in relative retention can be expected for solutes of varying S value (see discussion in ref. 19). This will be true regardless of the t_{sec} values and size-exclusion effects on separation. The reason is that a change in b results in a change in $\bar{k}$, the effective k' value throughout the gradient separation.

TABLE X

CHANGE IN PROTEIN RETENTION WITH CHANGE IN COLUMN DIMENSIONS. EXPERIMENTAL CONTROL OF BAND SPACING

Experimental conditions: Nucleosil 10 C_{18} columns of various sizes (3 × 0.3, 18 × 0.3 and 30 × 0.46 cm I.D.; 1 ml/min; gradient time (linear) 40 min; 87:10:3:0.1 water–2-methoxyethanol–*n*-butanol–TFA (solvent A), 70:20:10:0.025 ethanol–*n*-butanol–2-methoxyethanol–TFA (solvent B); Van der Zee and Welling[34].

Solute	*Mol. wt.** *(kdaltons)*	V_{sec}/V_m**	S***	$\log k_o$§	t_g *(min) [expt./(calc.)]*§§			$t_{3.5}-t_{0.15}$§§§
					V_m = *0.15 ml*	*0.9 ml*	*3.5 ml*	
β-Galactosidase	120	0.57	82.4	31.2	17.0 (17.2)	18.3 (18.1)	19.9 (19.9)	2.9
Phosphorylase *b*	95	0.57	74.4	36.4	22.3 (22.2)	23.4 (23.2)	25.0 (25.1)	2.7
Bovine serum albumin	69	0.57	64.6	24.5	16.9 (17.0)	18.3 (18.1)	19.9 (20.0)	3.0
Ovalbumin	44	0.58	53.0	27.4	23.0 (23.2)	24.6 (24.4)	26.4 (26.5)	3.4
Aldolase	40	0.59	50.8	21.3	18.7 (18.7)	20.1 (19.9)	21.8 (22.0)	3.1

Carbonic anhydrase	30	0.62	44.8	19.2	18.8 (19.0)	20.6 (20.4)	22.6 (22.7)	3.8
Ferritin	20	0.67	37.5	18.7	21.9 (22.0)	23.8 (23.6)	25.8 (26.1)	3.9
Myoglobin	18	0.67	35.8	14.9	18.2 (18.1)	20.0 (19.7)	22.0 (22.3)	3.8
Hemoglobin	16	0.69	34.0	15.3	19.6 (19.7)	21.6 (21.3)	23.8 (24.0)	4.2
Lysozyme	14	0.71	32.0	11.3	15.0 (15.1)	17.0 (16.8)	19.6 (19.6)	4.6
Ribonuclease	14	0.71	32.0	8.9	11.3 (11.7)	13.3 (13.4)	16.5 (16.2)	5.2
Cytochrome *c*	11	0.73	28.8	9.9	14.8 (14.5)	16.5 (16.4)	18.9 (19.3)	4.1

* Approximate values from ref. 34 for monomer subunits of protein.
** Estimated from protein SEC data for 12-nm-pore column of ref. 35 and denaturing conditions.
*** Calculated from eqn. 22.
§ Best fit to expt. t_g values.
§§ Calculated from eqns. 3 and 4 ignoring value of t_D (assumed zero).
§§§ Difference in t_g values for V_m = 3.5 and 0.15 ml.

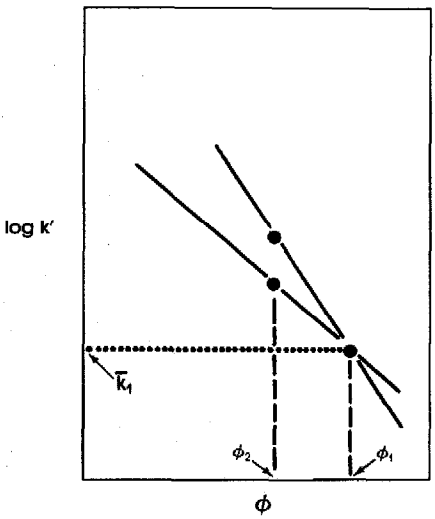

Fig. 4. Effect of a change in b (or $\bar{k}$) on band-spacing when two compounds have different S values. Here, both compounds are eluted together, for $\bar{k}_1 = 1/1.15b$ and an average φ value equal to φ_1. For a smaller value of b such that $\varphi = \varphi_2$, compounds have different $\bar{k}$ values and are separated.

And a change in $\bar{k}$ is then equivalent to a change in the average φ value during elution. The resulting effect of change in $\bar{k}$ on band-spacing is shown in Fig. 4, where log k' is plotted *vs.* φ for two solutes of differing S value. It is seen that both compounds are eluted at φ_1 (and k_1), but have different k' values at φ_2 (corresponding to different average k' values).

The change in t_g between the small (0.15 ml) and large (3.5 ml) columns of

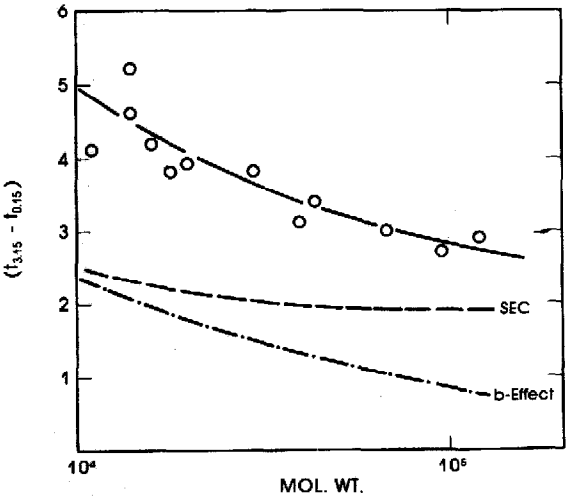

Fig. 5. Change in t_g for a change in V_m from 0.15 to 3.5 ml as a function of solute molecular weight[34]. ○, Experimental data; ——, calculated values; — — —, size-exclusion contribution to t_g change, equal to difference in t_{sec} values; - - - -, remaining contribution from difference in b values or S values for two solutes.

Table X (last column) is plotted against solute molecular weight in Fig. 5. The points are experimental values, and these are seen to agree closely with the calculated curve from eqn. 2. The dashed curve labeled "SEC" is the contribution from size-exclusion effects (difference in t_{sec} values), while the dashed curve labeled "*b*-effect" shows the effect due to change in b and $\bar{k}$ (as illustrated in Fig. 4). From this figure it is clear that the size-exclusion contribution to change in band-spacing with change in V_m is much smaller than the *b*-effect.

General equations for band-spacing as a function of column volume V_m and mobile-phase flow-rate F.

It is of value to derive general equations for the change in relative retention for two compounds that results from a change in column length or diameter (change in V_m), or from a change in flow-rate F. Eqns. 3 and 4 can be written as

$$t_g = (t_G/\Delta\varphi S) \log (2.3\ k_o\ \Delta\varphi S t_{sec}/t_G) + t_{sec} \tag{28}$$

where k_o is assumed to be large (so that $2.3\ k_o b \gg 1$), and t_D is corrected for (assumed zero in eqn. 28). For the case of a single solute separated with two different columns, so that retention times of t_{g1} and t_{g2} are observed, and corresponding values of t_{sec} equal to t_1 and t_2 are found, we have

$$t_{g2} - t_{g1} = \Delta t_g = (t_G/\Delta\varphi S) \log (t_2/t_1) + t_2 - t_1 \tag{29}$$

Now t_{sec} is related to V_{sec} as

$$t_{sec} = V_{sec}/F \tag{30}$$

and we can define the size-exclusion distribution constant (for a given solute) C_{sec} as

$$t_{sec} = C_{sec} t_o \tag{31}$$

where C_{sec} is seen to be equal to V_{sec}/V_m (Table X). Substituting these relationships into eqn. 29 then yields

$$\Delta t_g = (t_G/\Delta\varphi S) \log (V_2/V_1) + (C_{sec}/F)(V_2 - V_1) \tag{32}$$

Here, V_1 and V_2 refer to V_m values for columns 1 and 2.

Now we can determine the change in band-spacing as a result of change in V_m for two compounds A and B. Let the retention times for compound A and V_m values V_1 and V_2 be given as t_{g1a} and t_{g2a}. Likewise, for compound B and V_m values of V_1 and V_2, retention is given as t_{g1b} and t_{g2b}. The change in band-spacing or difference in retention times for A and B as a function of V_m is then given as

$$\begin{aligned} \Delta\Delta t_g &= (t_{g2a} - t_{g2b}) - (t_{g1a} - t_{g1b}) \\ &= (t_{g2a} - t_{g1a}) - (t_{g2b} - t_{g1b}) \end{aligned}$$

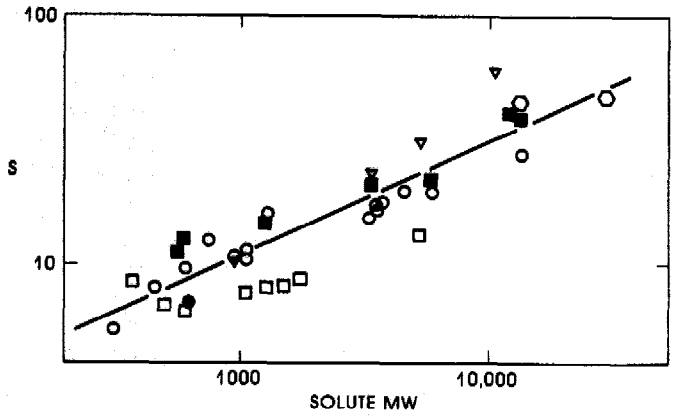

Fig. 6. Summary of peptide S values *vs.* molecular weight; acetonitrile–water as mobile phase. ■, Present study; ○, ref. 22; ▽, ref. 37; □, ref. 33; ○, ref. 20; ●, average value for steroids (unpublished data from this laboratory).

$$= (t_G/\Delta\varphi S_a) \log (V_2/V_1) + (C_a/F)(V_2 - V_1)$$
$$- (t_G/\Delta\varphi S_b) \log (V_2/V_1) - (C_b/F)(V_2 - V_1)$$
$$= \underbrace{\frac{(t_G/\Delta\varphi)(\log [V_2/V_1])(1/S_a - 1/S_b)}{\text{(RPC)}}}_{\text{(RPC)}} + \underbrace{\frac{(C_a - C_b)(V_2 - V_1)/F}{\text{(SEC)}}}_{\text{(SEC)}} \quad (33)$$

Here, S_a and S_b refer to S values for compounds A and B, and C_a and C_b are C_{sec} values for A and B. The first term on the right of eqn. 33 is the reversed-phase chromatography (RPC) effect, and the second term is the size-exclusion (SEC) effect. We have seen in Fig. 5 that the RPC term ("*b*-effect") will generally predominate.

Eqn. 33 states that the change in retention differences for two solutes A and B is related to their S values and to the difference in column V_m values, V_1 and V_2. The more the two compounds differ in molecular weight, the more their S values (eqn. 22) will differ and the greater will be the effect of change in V_m on band-spacing. Because values of C_{sec} decrease with increasing molecular weight, the RPC and SEC effects work in concert, as seen in Fig. 5.

We can proceed in similar fashion to derive the effect of change in F on band-spacing, obtaining finally

$$\Delta\Delta t_g = (t_G/\Delta\varphi)(\log[F_1/F_2])(1/S_a - 1/S_b) + (C_a - C_b) V_m (1/F_2 - 1/F_1) \quad (34)$$

Here, F_1 and F_2 refer to different flow-rates in the two experiments. Eqn. 34 is seen to be of the same form as eqn. 33, meaning that a two-fold change in either V_m or in F will have an equivalent effect on band-spacing.

Hartwick *et al.*[36] have also noted the possibility of adjusting band-spacing by change in the gradient parameter b. They optimized the separation of nucleosides and related bases by gradient elution, using gradient time to vary b and band-spacing. This is a less efficient approach than the use of V_m or F as described above, because then peak capacity and band-spacing are varied simultaneously.

Dependence of S on experimental conditions

Several workers (*e.g.*, refs. 20, 33 and 37) have noted that values of S increase with the size of the peptide or protein molecule. Fig. 6 summarizes several studies from the literature as well as data reported here, in the form of a log–log plot of S values *vs.* solute molecular weight. The resulting correlation of Fig. 6 shows some scatter (±25% in S, 1 S.D.) but generally confirms the trend in S *vs.* molecular weight as summarized by eqn. 22 (solid line of Fig. 6). A similar increase in S with increasing solute molecular weight has been noted for polystyrenes eluted by tetrahydrofuran–water[23,24]. This effect appears to be general for all solute types in reversed-phase chromatography, as will be discussed further elsewhere.

There is a trend in Fig. 6 for higher S values in the study of ref. 37 (▽ in Fig. 6) and lower S values in the study of ref. 33 (□). This suggests that S may depend to some extent on column type, although minor differences in mobile phase are also present in these various studies (but all involve acetonitrile–water). Values of S are known to vary with the organic solvent used, being larger for less polar organics (*e.g.*, see ref. 19). It is likely also that S varies with the peptide molecular structure apart from molecular weight, but this appears to be a minor effect. Eqn. 22 should be adequate for estimations of S from sample molecular weight, at least for purposes of predicting peak capacity, peak heights, etc.

CONCLUSIONS

A preliminary model is presented for describing the reversed-phase separation of peptide and protein samples by gradient elution. The model is essentially the same as that already verified for the similar separation of small organic molecules (*e.g.*, refs. 24 and 28). Application of the present model to retention data from several different laboratories (including our own) shows good agreement between experimental and calculated values — comparable to that reported earlier for the similar separation of small molecules by gradient elution. It therefore appears that these reversed-phase gradient elution separations of peptides and proteins behave in similar fashion to corresponding separations of small molecules. That is, there are no unique features to these separations of macromolecules (see discussion in ref. 18).

The present model is both detailed and comprehensive, allowing the prediction of most aspects of these separations. A later paper will further verify the peak-volume and resolution relationships for the seven peptides studied by us. It will also be shown that this model lends itself to the facile design of optimum separations and to the evaluation of columns for peptide separations.

APPENDIX I

Derivation of values of S and k_o from experimental gradient data

The derivation of values of S and k_o from two gradient retention times t_{g1} and t_{g2} obtained with differing values of t_G for the two experiments is discussed in the text (see eqns. 5–7 and related literature). For the special case where k_o is large and therefore $2.3\,bk_o \gg 1$, eqns. 21 and 21a can be approximated by

$$t_g = (t_o/b) \log (2.3\,k_o b) + t_o \qquad \text{(A1)}$$

which then allows explicit solution for k_o and b_1:

$$b_1 = (t_o \log \beta)/[t_{g1} - (t_{g2}/\beta) + t_o (1 - \beta/\beta)] \tag{A2}$$

and

$$\log k_o = (b_1/t_o)(t_{g1} - t_o) - \log (2.3 b_1) \tag{A3}$$

This approach (limited to variation in t_G only) can be extended to data such as those of Table X where both t_G and F are varied simultaneously. We start by expressing retention in terms of values of retention volume V_g, which is equal to $t_g F$. This yields equations for V_{g1} and V_{g2} as in eqn. A1:

$$V_{g1} = F_1 t_{g1} \text{ and } V_{g2} = F_2 t_{g2} \tag{A4}$$

As in the derivation of eqn. A2, we can obtain from eqn. A4

$$b_1 = -\beta (\log \beta)/(X_1 - \beta X_2)$$

where

$$X_1 = [(V_{g1}/F_1) - t_1]/t_1 \tag{A5}$$

and

$$X_2 = [(V_{g2}/F_2) - t_2]/t_2 \tag{A6}$$

Here t_1 and t_2 refer to t_o values for experiments 1 (t_{G1}) and 2 (t_{G2}). The quantity k_o can be obtained from eqn. A3, setting t_o equal to t_1.

Values of $\bar{\varphi}$ and $\bar{k}$ are also desired, as in plots of Fig. 1. As discussed in ref. 19, the retention time t_g ($t_{\frac{1}{2}}$) for elution of a band halfway across the column is given as

$$t_{\frac{1}{2}}/t_o = [(1/b) \log (1.15 k_o b + 1)] + 0.5 \tag{A7}$$

(see eqn. 23b of ref. 19). Combining this with eqn. 4 and

$$\bar{\varphi} = (t_g^0 - 0.5 t_o)/t_G^0 \tag{A8}$$

then yields

$$\bar{\varphi} = [t_g - t_o - (t_o/b) \log 2]/t_G^0 \tag{A9}$$

Values of $\bar{k}$ can be obtained from eqn. 11.

SYMBOLS

A	Knox equation constant which reflects how well a column is packed; eqn. 8.
a'	Intercept of B *vs.* k' plot: equal to *ca.* 1.1. See ref. 32.
B	Knox equation constant which reflects extent of band broadening due to longitudinal diffusion; eqns. 16 and 21.
b	Gradient steepness parameter. Steep gradients translate into large b values and corresponding small $\bar{k}$ values; eqn. 3.
b'	Surface diffusion parameter; experimentally determined by evaluating the slope of the B *vs.* k' plot. See eqn. 21 and ref. 32.
b_1, b_2	Gradient steepness parameters for a single solute and two gradients differing only in their gradient times (t_{G1} and t_{G2}); eqn. 7.
C	Knox equation constant which reflects mass transfer contributions to band broadening; eqns. 16 and 19.
C_a, C_b	Size-exclusion distribution constant (C_{sec}) for solute a and b; eqn. 33.
C_{sec}	Size-exclusion distribution constant; equal to t_{sec}/t_o.
D_m	Solute diffusion coefficient (cm^2/sec) in the mobile phase outside the particle; eqns. 23–25.
D_p	Effective solute diffusion coefficient (cm^2/sec) within the pores of the packing; eqn. 20.
$D_{w,25}$	Solute diffusion coefficient (cm^2/sec) in water at 25°C; eqn. 23.
d_c	Column inner diameter (cm).
d_p	Average particle diameter (cm).
F	Flow-rate of mobile phase (ml/sec, unless specified).
F_1, F_2	Two different flow-rate values (ml/sec).
H	Plate height (cm); equal to L/N.
h	Reduced plate height (dimensionless); equal to H/d_p.
k_o, k_w	Solute k' (capacity factor) at the initial mobile phase composition, φ_0, or water respectively; eqns. 1, 2, 4 and 8.
$\bar{k}$	Value of k' (capacity factor) for a solute when it reaches the column midpoint during gradient elution; eqn. 11.
L	Column length (cm).
N	Column plate number; equal to L/H.
PC	Maximum peak count possible if unit resolution between each peak is assumed for a specific set of gradient conditions; eqns. 10 and 14.
Q	Capacity factor contribution to R_s; defined as $\bar{k}/1+\bar{k}$.
R_s	Resolution function for gradient elution; eqn. 9.
S	Slope of the log k' *vs.* volume fraction organic (φ), eqn. 1; can be estimated using the molecular weight of the solute by eqn. 22.
S_a, S_b	Values of S for solutes a and b.
t_D	Dwell time (sec, unless specified) for the gradient system; time required for the gradient mobile phase to travel from the mixer to column inlet; eqns. 2 and 4.
t_G	Gradient time (sec, unless specified); time elapsed during the gradient run.
t_{G1}, t_{G2}	Values of t_G for two different gradient runs, resulting in different values of b (b_1, b_2) and t_g (t_{g1}, t_{g2}) for a single solute.

t_g^o Gradient time (sec, unless specified) expected for a solute if gradient steepness is maintained while $\Delta\varphi$ is normalized to 1.00; eqn. 27.

t_g Gradient retention time (sec, unless specified) for a solute; eqns. 2 and 4.

t_{g1}, t_{g2} Gradient retention times (sec, unless specified) for a single solute in two different gradient time (t_G) values (eqns. 5 and 6); or for a solute in two columns differing in their void volumes; eqn. 33.

t_{g1a}, t_{g1b}, t_{g2a}, t_{g2b} Gradient retention times (sec, unless specified) for a solute a and b in columns having void volume V_1 and V_2, other gradient conditions are identical; eqn. 33.

Δt_g Difference between two solute gradient retention times ($t_{g2} - t_{g1}$) obtained by using b values which differ only in V_m. (columns have different dimensions)

$\Delta\Delta t_g$ Difference between Δt_g value obtained by using solute A and Δt_g value obtained by using solute B; value is augmented by large differences between S_a and S_b and V_1 and V_2; eqn. 33.

t_o Column dead-time (sec, unless specified); time required for mobile phase molecules to traverse the column.

t_{sec} Retention time (sec, unless specified) for a solute under size-exclusion conditions; eqn. 2.

t_1, t_2 Single solute t_{sec} values on columns differing in their void volumes.

u Mobile phase linear velocity (cm/sec); equal to L/t_o.

V_1, V_2 V_{sec} values of two columns differing in physical dimensions; eqn. 33.

V_D Delay volume; equal to t_D/F.

V_m Column void volume; equal to the sum of the volume inside and outside the packing pores; eqn. 12.

V_{sec} Retention volume (ml) for solute under size-exclusion conditions; equal to $t_{sec}\,F$.

x Fraction of the column void volume outside the pores of the packing material.

α Separation factor, $\bar{k}$ ratio of two adjacent solute bands at column midpoint in gradient elution; eqn. 9.

β Ratio of t_{G2} and t_{G1} which is equivalent to ratio of b_1 and b_2; eqn. 7.

η Mobile phase (acetonitrile–water) viscosity at column temperature.

η_{25} Mobile phase (acetonitrile–water) viscosity at 25°C.

σ_{ec} Extra column volume contribution to band broadening (ml).

σ_t Peak bandwidth expressed in time units (sec, unless specified); eqn. 10.

σ_v Peak bandwidth expressed in volume units (ml, unless specified); eqn. 13.

ρ Restricted diffusion parameter; see ref. 32.

φ Volume fraction of organic component in the mobile phase.

φ_f Volume fraction of organic component in the mobile phase at the end of the gradient.

φ_o Volume fraction of organic component in the mobile phase at the start of the gradient.

$\Delta\varphi$ Change in organic volume fraction during the gradient run; equal to $\varphi_f - \varphi_o$.

$\delta\varphi$ Deviation in φ at elution of a solute between experimentally obtained and theoretically predicted values.

γ Tortuosity factor; reflects perturbation of molecular diffusion due to packing geometry.
τ Time constant of liquid chromatographic system.
ν Mobile phase reduced velocity (dimensionless), equal to ud_p/D_m.

REFERENCES

1 F. E. Regnier and K. M. Gooding, *Anal. Biochem.*, 103 (1980) 1.
2 J. Richey, *Amer. Lab.*, (Oct. 1982) p. 104.
3 W. S. Hancock and J. T. Sparrow, in Cs. Horvath (Editor), *High-Performance Liquid Chromatography*, Vol. 3, Academic Press, New York, 1983, p. 49.
4 M. T. W. Hearn, in Cs. Horvath (Editor), *High-Performance Liquid Chromatography*, Vol. 3, Academic Press, New York, 1983, p. 87.
5 M. T. W. Hearn, F. E. Regnier and C. T. Hehr, *Proceedings of the First International Symposium on HPLC of Proteins and Peptides*, Academic Press, New York, 1983.
6 *Proceedings of the Second International Symposium of Proteins, Peptides and Polynucleotides, J. Chromatogr.*, 266 (1983) (entire issue).
7 J. D. Pearson, W. C. Mahoney, M. A. Hermodson and F. E. Regnier, *J. Chromatogr.*, 207 (1981) 325.
8 K. J. Wilson, E. van Wieringen, S. Klauser, M. W. Berchtold and G. J. Hughes, *J. Chromatogr.*, 237 (1982) 407.
9 E. C. Nice, M. W. Capp, N. Cooke and M. J. O'Hare, *J. Chromatogr.*, 218 (1981) 569.
10 Y. Kato, K. Nahamara and T. Hashimoto, *J. Chromatogr.*, 245 (1982) 193.
11 A. J. Alpert, *J. Chromatogr.*, 266 (1983) 23.
12 M. J. O'Hare, M. W. Capp, E. C. Nice, N. H. C. Cooke and B. G. Archer, *Anal. Biochem.*, 126 (1982) 17.
13 E. Minasian, R. S. Sherma, S. J. Leach, B. Grego and M. T. W. Hearn, *J. Liquid Chromatogr.*, 6 (1983) 199.
14 W. S. Hancock, C. A. Bishop, R. L. Prestidge, D. R. K. Harding and M. T. W. Hearn, *J. Chromatogr.*, 153 (1978) 391.
15 H. P. J. Bennett, C. A. Browne and S. Solomon, *J. Liquid Chromatogr.*, 3 (1980) 1353.
16 G. E. Tarr and J. W. Crabb, *Biochemistry*, 131 (1983) 99.
17 H. Jaffe and D. K. Hayes, *J. Liquid Chromatogr.*, 6 (1983) 993.
18 L. R. Snyder, M. A. Stadalius and M. A. Quarry, *Anal. Chem.*, 55 (1983) 1412A.
19 L. R. Snyder, in Cs. Horvath (Editor), *High-Performance Liquid Chromatography*, Vol. 1, Academic Press, New York, 1980, p. 208.
20 M. T. W. Hearn and B. Grego, *J. Chromatogr.*, 255 (1983) 125.
21 M. A. Phelan and K. A. Cohen, *J. Chromatogr.*, 266 (1983) 55.
22 N. H. C. Cooke, B. G. Archer, M. J. O'Hare, E. C. Nice and M. Capp, *J. Chromatogr.*, 255 (1983) 115.
23 J. P. Larmann, J. J. DeStefano, A. P. Goldberg, R. W. Stout, L. R. Snyder and M. A. Stadalius, *J. Chromatogr.*, 255 (1983) 163.
24 M. A. Quarry, R. L. Grob and L. R. Snyder, *Anal. Chem.*, submitted for publication.
25 R. W. Stout, J. J. DeStefano and L. R. Snyder, *J. Chromatogr.*, 282 (1983) 263.
26 S. A. Cohen, Y. Tapuhi, J. C. Ford and B. L. Karger, in preparation.
27 L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, Wiley-Interscience, New York, 2nd ed., 1979.
28 M. A. Quarry, R. L. Grob and L. R. Snyder, *J. Chromatogr.*, 285 (1984) 1, 19.
29 *Scientific Tables*, 7th ed., Ciba Geigy Ltd., Basle, Switzerland, 1970, p. 580.
30 H. Colin, J. C. Diez-Masa, G. Guiochon, T. Czajkowska and I. Miedziak, *J. Chromatogr.*, 167 (1978) 41.
31 J. R. Gant, J. W. Dolan and L. R. Snyder, *J. Chromatogr.*, 185 (1979) 153.
32 R. W. Stout, J. J. DeStefano and L. R. Snyder, *J. Chromatogr.*, 261 (1983) 189.
33 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.
34 R. Van der Zee and G. W. Welling, *J. Chromatogr.*, 244 (1982) 134.
35 W. W. Yau, J. J. Kirkland and D. D. Bly, *Modern Size-exclusion Chromatography*, Wiley-Interscience, New York, 1979, p. 441 (Fig. 13.13a).
36 R. A. Hartwick, C. M. Grill and P. R. Brown, *Anal. Chem.*, 51 (1979) 34.
37 S. Terabe, H. Nishi and T. Ando, *J. Chromatogr.*, 212 (1981) 295.