Journal of Chromatography, 392 (1987) 33–49
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 19 351

HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY OF AMINO ACIDS, PEPTIDES AND PROTEINS

LXIX*. EVALUATION OF RETENTION AND BANDWIDTH RELATIONSHIPS OF MYOSIN-RELATED PEPTIDES SEPARATED BY GRADIENT ELUTION REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

MILTON T. W. HEARN* and M. I. AGUILAR
Department of Biochemistry, Monash University, Clayton, Victoria 3168 (Australia)
(First received September 15th, 1986; revised manuscript received December 12th, 1986)

SUMMARY

The gradient elution behaviour of ten peptide analogues encompassing the primary amino acid sequences adjacent to the serine-19 phosphorylation site of myosin light chain has been investigated using a 30-nm pore diameter *n*-butylsilica stationary phase and 0.1% trifluoroacetic acid–water–acetonitrile mobile phases. Quantitative expressions derived from linear solvent strength theory and general plate height theory have been used to assess the influence of gradient time and flow-rate on the relative retentions and bandwidths of these peptides. The retention behaviour of these peptide analogues was found to closely mirror the predictions of linear solvent strength gradient elution theory. Although the experimentally observed bandwidth changes over a conventional range of gradient conditions (*i.e.* t_G = 30–60 min) also correlated with bandwidth changes predicted on the basis of plate theory derived for low molecular weight organic molecules, for very steep and very shallow gradients divergencies from theory were observed. The significance of these discrepancies is discussed in relation to the influence of amino acid sequence changes rather than composition of these peptide analogues. These data allow the importance of sequential effects on the chromatographic behaviour and in particular, the resolution optimisation of these closely related peptides to be evaluated.

INTRODUCTION

Myosin light chain kinase is an enzyme which is intimately involved in the regulation of smooth muscle contraction, through its ability to phosphorylate a serine residue at position 19 in the myosin light chain[1]. The enzyme is highly specific since it does not readily phosphorylate exogenous substrates and exhibits a preference for

* For Part LXVIII see ref. 11.

0021-9673/87/$03.50

myosin light chains from the same tissue type[2,3]. Several investigations have been recently carried out on the substrate specificity requirements of the smooth muscle myosin light chain kinase using synthetic peptides corresponding to residues 11–23 in the gizzard myosin regulatory light chains[4,5]. These studies have shown that the amino acid sequence associated with the amino-terminal tripeptide (Lys^{11}-Lys^{12}-Arg^{13} ....) and the arginine residue at position 16 (Arg^{16}) have significant influences on the kinetics of peptide phosphorylation at Ser-19.

Clearly the spatial organisation, predicated by these specific amino acid residues in myosin light chain and its derived peptide analogues, determines a preferred orientation for Ser-19 in the active site of the enzyme such that the optimal K_m and V_{max} values are obtained. Whether the Ser-19 is fully exposed in these polypeptides to the surrounding solvent or only becomes accessible in the active site of the enzyme by induced fit has not been determined.

Physical methods of analysis such as circular dichroism or difference spectroscopy, have been used for many years to provide insight into the prevailing conformation(s) of peptides under defined conditions. Similarly, the *in vivo* and *in vitro* use of structural analogues in the determination of quantitative structure–activity relationships (QSARs) continues to provide valuable information about biochemical and pharmacological processes. The intimate relationship which must exist between the chromatographic behaviour of peptides and proteins and their primary, secondary and tertiary structures has however only recently been addressed. Several models pertinent to these considerations on the retention behaviour of peptides and proteins in reversed-phase (RP) and ion-exchange high-performance liquid chromatography (HPLC) have been described[6–8] for isocratic and gradient elution systems and have provided a basis for interpreting complex peak zone behaviour evident with some peptides and proteins. Under "near-equilibrium" conditions with rapidly interconverting, time-averaged structures (such as random coil forms or in some cases stable globular structures), the use of retention and band-broadening models based on linear solvent strength gradient elution concepts[9] has provided a basis for optimisation of separation methods for peptide solutes[10–15]. For example, the application of linear solvent strength gradient theory to the RP-HPLC of groups of closely related peptides derived from human β-endorphin[10], luteinising hormone releasing hormone (LHRH) and growth hormone releasing factor (GHRF)[11], growth hormones (GHs)[12] and interleukin S (ILs-II)[16] have demonstrated that the effect of various chromatographic parameters on peptide retention can be accurately rationalised in terms of specific tertiary structural characteristics of these different polypeptides. Based on these approaches, optimal choice of flow-rate or gradient time conditions can be selected leading to enhanced resolution of a complex peptide mixture through improvement in bandspacing[10–12,15]. Furthermore, quantitative expressions derived from the general plate height theory for small organic molecules have been found with some polypeptides to allow the calculation of "ideal" bandwidths which agree reasonably well with the experimentally observed bandwidth values[10–13,17]. However, in cases where chromatographic residence time effects result in changes in peptide secondary structure or where other secondary equilibrium effects operate, the experimental bandwidths have been found[6,8,11,17] to significantly exceed the values predicted on the basis that the solute behaves as a conformationally rigid species which interacts with a single class of binding sites during its passage along the column.

Evidence has also been presented[12,17–19] which has indicated that differences in the bandbroadening phenomena exhibited by some peptide solutes could be related to the induction of a significant degree of secondary conformational structure at the stationary-phase surface. The present paper explores these phenomena further with the analysis of the gradient elution data of a set of polypeptides, related to the amino-terminal sequence region around Ser-19 of the chicken gizzard myosin regulatory light chain, separated on a wide pore *n*-butylsilica as the stationary phase and a water–acetonitrile–trifluoroacetic acid solvent system.

EXPERIMENTAL

Apparatus

All chromatographic measurements were obtained using a Waters Assoc. (Milford, MA, U.S.A.) liquid chromatograph system consisting of two M6000A pumps, a WISP M710B automatic injector and an M660 gradient programmer. The detector used was an M450 variable-wavelength UV monitor operating at 215 nm and was coupled to an M730 data module.

Chemicals and reagents

Acetonitrile was HPLC grade, obtained from Millipore (Lane Cove, Australia). Trifluoroacetic acid (TFA) was obtained from Pierce (Rockford, IL, U.S.A.). Water was quartz-distilled and deionised in a Milli-Q system (Millipore, Bedford, MA, U.S.A.). Synthetic peptides corresponding to the amino-terminal residues 11–23 of chicken gizzard myosin regulatory light chain were generously provided by Dr. Bruce Kemp, University of Melbourne, Repatriation General Hospital, Heidelberg, Victoria, Austrialia. The amino acid compositions of the peptides were determined using a Durrum D500 analyser following hydrolysis in 6 *M* hydrochloric acid containing 0.1% phenol for 16 h at 110°C.

Chromatographic procedures

Bulk solvents and mobile phases were filtered and degassed under vacuum. Chromatographic measurements were made at 20°C using a Bakerbond Wide Pore C_4 Column (25 cm × 4.6 mm I.D.) with an average particle diameter of 5 μm. Linear gradient elution was carried out with 0.1% TFA in water (buffer A) and 0.1% TFA in water–acetonitrile (50:50) (buffer B) over gradient times varying between 20 and 120 min and flow-rates between 1 and 4 ml/min. Sample sizes for each polypeptide varied between 1 and 10 μg with injection volumes of 10–100 μl. No change in solute retention time for a defined gradient time was noted over this mass loading range. The column dead time, t_0, was taken as the retention time for sodium nitrate. The various chromatographic parameters b, S, $\bar{k}$, k'_0, $\bar{\psi}$, $t_{g,calc}$, PC_{exp}, C, D_m, b', ρ, $\sigma_{v,calc}$, were calculated using the Pek'N-ese program, written in BASIC language for a Hewlett-Packard HP86B computer using input values of t_0, F, t_g, t_G, $\Delta\psi$, x, η, MW, d_p, a', $\sigma_{v,exp}$, T and L as previously described[11,17].

RESULTS AND DISCUSSION

Retention relationships

A detailed description of the application to polypeptide separation of the linear solvent strength (LSS) gradient elution theory and the derivation of equations used in the present paper has been outlined previously[10,18]. Briefly, the retention time, t_g, for a polypeptide chromatographed under ideal LSS gradient elution conditions can be related to the gradient steepness parameter, b, through the expression

$$t_g = (t_0/b)\,[\log 2.3\, k_0' b] + t_0 + t_e \tag{1}$$

where t_0 is the column dead time, k_0' is the capacity factor for the solute in the initial conditions of the gradient (solvent A) and is assumed to be large for peptides and proteins; and t_e is the gradient elapse time required for the change in solvent B to reach the column inlet. In circumstances where conformational reorientation of the solutes occur a further term, t_d is required to accomodate the dwell time effect[8]. When solutes are chromatographed using several different gradient times on a particular column with the same mobile phase system at a constant flow-rate the gradient steepness parameter, b, can be derived using the relationship

$$b = t_0 \log \beta/[t_{g1} - (t_{g2}/\beta) + t_0\,(t_{G1} - t_{G2})/t_{G2}] \tag{2}$$

where t_{g1} and t_{g2} are the solute gradient retention times at gradient times t_{G1} and t_{G2}, respectively, and β is the ratio of the gradient times ($= t_{G2}/t_{G1}$). Alternatively, if chromatograms are obtained at flow-rates F_1 and F_2 whilst the gradient time is maintained constant, the b values can be determined as follows,

$$b_1 = \frac{\log (F_2/F_1)}{[X_1 - X_2\,(F_1/F_2)]} \tag{3}$$

where

$$X_1 = \frac{t_{g1} - t_{0,1}}{t_{0,1}}$$

and

$$X_2 = \frac{t_{g2} - t_{0,2}}{t_{0,2}}$$

Evaluation of b values from retention data obtained for various gradient times and solvent flow-rate conditions allows the calculation of the median capacity factors, $\bar{k}$, and the corresponding organic mole fraction, $\bar{\psi}$. Provided the retention of the polypeptide in question follows ideal LSS behaviour, k and ψ can be derived, to a first approximation from eqns. 4 and 5, namely

$$\bar{k} = 1/1.15b \tag{4}$$

$$\bar{\psi} = \left[t_{g1} - t_0 - \left(\frac{t_0}{b}\right) \log 2 \right] \Big/ (t_G/\Delta\psi) \tag{5}$$

The parameters $\bar{k}$ and $\bar{\psi}$ are formerly equivalent to the isocratic parameters $k'$and ψ which, under regular reversed-phase elution conditions are empirically related through the expression.

$$\log k' = \log k'_0 - S\psi \tag{6}$$

The tangent (or slope) S to the curve obtained in a plot of log $\bar{k}$ *versus* $\bar{\psi}$ can thus be determined at any particular value of $\bar{k}$ and the value of log $\bar{k}$ at $\bar{\psi} = 0$, (*i.e.* log k'_0) can be obtained by extrapolation using regression analysis.

The relationships summarised in eqns. 1–6 give rise to a number of predictions related to the retention behaviour of peptide solutes under linear solvent strength gradient conditions. Firstly the corresponding gradient form of eqn. 6 predicts linear dependencies of log $\bar{k}$ on $\bar{\psi}$, provided the solutes are eluted under regular reversed-phase conditions and their retention behaviour does not involve any significant participation of slow secondary equilibria associated with ionisation, surface adsorption or conformational effects. Secondly, as the b value is changed through alteration of t_G or F, selectivity changes should become evident for those peptides with different S values in which the processes of their absorption–desorption result in different association constants or kinetic rate constants, either through the participation of different classes of binding sites on the heterogeneous stationary-phase surface (*i.e.* through different binding mechanisms), through the involvement of different topographic regions on the solute surface (*i.e.* via the same mechanism but different effective hydrophobic contact areas) or through hybrid combinations of both processes. With solute analogues in which identical molecular regions are involved in the interaction with the stationary-phase surface, coincidental retention behaviour with parallelism in the log $\bar{k}$ *versus* $\bar{\psi}$ plots should be observed. In the case of closely eluting peptide solutes, with fundamentally different contact area dependencies, manipulation of secondary solution equilibria through, for example, pH, ion-pairing or solvation effects, should result in selectivity changes due to alteration of the topographical dimensions of the interactive region of the solute.

In order to further investigate the relationship between peptide structure and retention behaviour in reversed-phase HPLC and to further test the utility of the LSS retention model, summarised by eqns. 1–6, for the optimisation of the resolution of closely related peptide analogues separated by reversed-phase HPLC, gradient elution data were collected for a series of ten synthetic peptides related to the amino-terminal region (Lys^{11} .... Ser^{23}) of the chicken gizzard smooth muscle myosin light chain. The amino acid sequences of the peptides used in this study are listed in Table I and can be divided into three main sub sets. The first two groups (peptides 1–4 and 5–7) represent positional isomers in which an arginine or an alanine residue, respectively has been interchanged along the peptide chain. The third group (peptides 8–10), of which peptide 2 can also be considered a member, consists of insertion analogues with an increasing number of alanine residues. The molecular weights of the peptides range from *ca.* 1400–1600. All peptides were chromatographed singly and as mixtures on a wide pore *n*-butylsilica column using a linear gradient from 0.1%

TABLE I

STRUCTURES AND RETENTION PARAMETERS OF PEPTIDE ANALOGUES CORRESPONDING TO MYOSIN LIGHT CHAIN, RESIDUES 11–23

The sequence of myosin light chain (11–23) is .... K^{11} KRPQRATSNVFS^{23} ....

Peptide	*Sequence*	*MW*	$D_m \cdot 10^7$ *(cm²/s)*	*S*	*log* k'_0	*Summated hydrophobicities**				
						χ_1	χ_2	χ_3	χ_4	K_m**
1	KKRAAARTSNVFA	1418	2.52	15.32	2.17	2.79	−1.03	34.9	4.6	270
2	KKRAARATSNVFA	1418	2.52	15.37	2.10	2.79	−1.03	34.9	4.6	7.5
3	KKRARAATSNVFA	1418	2.52	15.41	2.21	2.79	−1.03	34.9	4.6	20
4	KKRRAAATSNVFA	1418	2.52	15.46	2.26	2.79	−1.03	34.9	4.6	114
5	AKRPQRATSNVFS	1460	2.50	15.83	2.25	0.57	−0.76	24.7	3.5	460
6	KARPQRATSNVFS	1460	2.50	15.73	2.23	0.57	−0.76	24.7	3.5	445
7	KKAPQRATSNVFA	1444	2.50	15.47	2.23	1.66	−1.54	36.1	7.7	951
8	KKRARATSNVFA	1347	2.60	15.35	2.18	2.26	−0.91	27.6	3.6	266
9	KKRAAARATSNVFA	1489	2.50	16.00	2.36	3.32	−1.15	42.2	5.6	10.1
10	KKRAAAARATSNVFA	1560	2.48	16.82	2.53	3.85	−1.27	49.5	6.6	70.5

* Summation of the hydrophobicity χ_1–χ_4 for the amino acid residues in the sequences of these peptides was carried out using coefficients discussed in refs. 22–25, respectively.

** The active constants (K_m, μmole) were determined by myosin kinase phosphorylation as described by Kemp *et al.*[4,5].

TFA in water to 0.1% TFA in water–acetonitrile (50:50, v/v). Chromatographic data was accumulated for t_G = 20, 30, 40, 60 and 120 min at a fixed flow-rate of 1 ml/min. A second set of data was obtained for separations carried out at flow-rates of 1, 2, 3 and 4 ml/min while t_G was held constant at 40 min.

Fig. 1 shows the plots of log $\bar{k}$ *versus* $\bar{\psi}$ for the myosin kinase substrate analogues in which the data from both the constant flow and the varied flow experiments are included. It is evident from Fig. 1 that essentially linear dependencies of log $\bar{k}$ *versus* $\bar{\psi}$ were observed for these peptide analogues. The correlation coefficients for the linear regression analysis of the log $\bar{k}$ *versus* $\bar{\psi}$ data for these different peptides were between 0.95 and 0.99. Even with this set of closely related peptides in which the amino acid sequences only vary slightly, it can be seen that selectivity changes occur over the range of experimental conditions used. For example, the resolution between peptides 9 and 10, which only differ by one alanine residue, was clearly

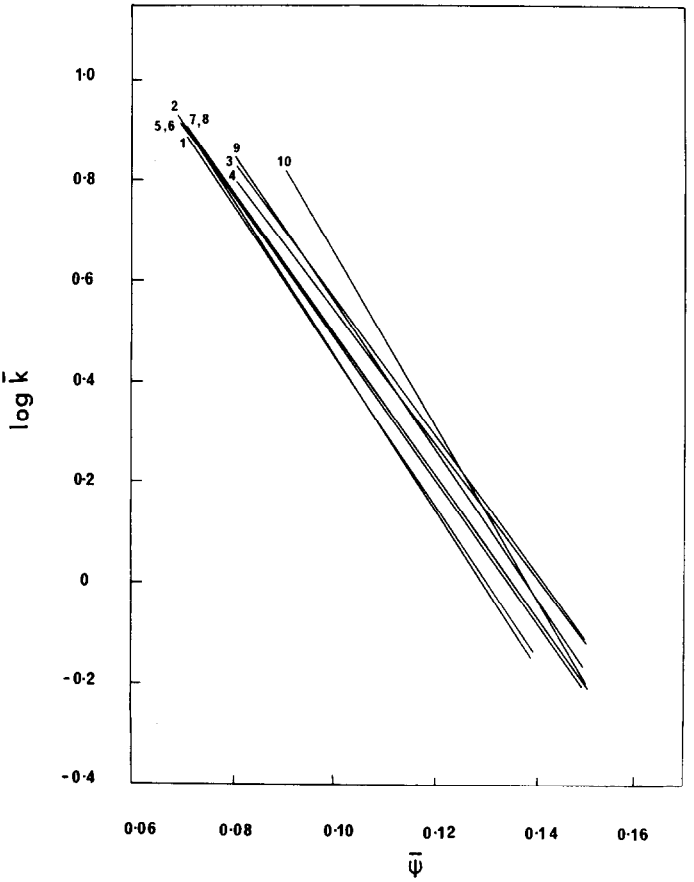

Fig. 1. Plots of log $\bar{k}$ *versus* $\bar{\psi}$ based on gradient experiments for myosin-related peptides 1–10. The plots were derived from best fit analysis to the data points (which have been excluded for clarity) (r^2 = 0.95–0.99), where t_G = 20, 30, 40, 60 and 120 min and F = 1 ml/min and F = 1, 2, 3 and 4 ml/min and t_G = 40 min. Other chromatographic methods are given in the Experimental section. See Table I for the code to peptide structure and for the derived S and log k'_0 values.

enhanced when longer gradient times or lower b values were used. These changes in selectivity are reflected in the differences between the S values for these two peptides (S = 16.00 and 16.82, respectively). Various data on the log k'_0 values and relative hydrophobicities (χ) for these peptides are listed in Table I. As is evident from these data there are small, but nevertheless significant, differences in the experimentally derived values of S and log k'_0. For example, peptide 10 (KKRAAAARATSNVFA) has the steepest slope value and largest log k'_0 value of the three peptide sets.

In previous studies[10,11,13,14] empirical relationships which link S values and the molecular weight of a polypeptide have been derived in the form $S = a\,(\mathrm{MW})^b$ where the coefficients a and b have varied according to the set of peptides and chromatographic conditions used in the analysis. While the present study provides an extremely limited range of molecular weights the derivation of an S *versus* MW relationship is still possible. When the experimental data for these peptides was used the relationship took the form $S = 0.11\,(\mathrm{MW})^{0.68}$ (r = 0.90). The average experimentally derived S value ($\bar{S}$) was 15.68 while the S value calculated from the average MW of the peptide set (MW = 1443) according to two relationships previously derived by Stadalius *et al.*[13] ($S = 0.48\,(\mathrm{MW})^{0.44}$) and Aguilar *et al.*[10] ($S = 2.99\,(\mathrm{MW})^{0.21}$) was 11.79 and 13.78, respectively. The average standard deviation between the experimental S values and the predicted S values according to these two empirical equations was 3.87 and 1.88, respectively. As we have concluded previously[10,11,18] the value of the parameter S in reversed-phase separations of polypeptides and proteins is related to the magnitude of the hydrophobic contact area and the number of interaction sites established between the solute and the stationary-phase ligands during the adsorption process. Although these sites could increase in number with the increasing molecular weight and size of the solute, it will not be the molecular weight *per se* but rather the polarity and disposition of the surface amino acid residues involved in the interaction with the stationary phase which will ultimately control the mechanistic pathway of the binding process. Therefore, if the peptide assumes any degree of preferred folding, no simple relationship will exist between S and either molecular weight or S and the summated hydrophobicity coefficients (χ), calculated according to the linear amino acid sequence. Inspection of the S value and χ values shown in Table I demonstrates that relatively complex higher order relationships must exist between S and χ even for this selection of closely related myosin light chain peptides.

Since the magnitude of log k'_0 is a measure of the free energy changes associated with the binding of the solute to the stationary phase under the initial gradient buffer A conditions, it could be anticipated on the basis of the Martin additivity equation that log k'_0 values should progressively increase with incremental increases in the solute hydrophobicity. With small peptide oligomers such as polyalanine or polyphenylalanine (n = 6), incremental increases in log k'_0 with chain length have been observed[20,21]. A similar linearised connection between log k and χ forms the basis of most chromatographic methods to determine hydrophobicity coefficients[22–25]. However, as is not surprising with polypeptides that exhibit well developed, and stabilised secondary and tertiary structures, previous studies (*cf.* ref. 10) have not supported the expected linear relationship between log k' and $\Sigma\chi$. Even in the present study, peptides 1 4, which have identical amino acid composition and therefore have the same relative hydrophobicity values (based on summated linear coefficients), had

sufficiently different log k'_0 and S values to allow resolution of these four peptides with shallow water rich gradients, despite the close similarities of their S values. The sequence of these peptides only differs in the location of an arginine residue which varies from position 4 in peptide 4 to position 7 in peptide 1. The progressive displacement of this arginine residue by alanine residues along the peptide chain clearly had sufficient influence on the retention characteristics of the peptides to allow separation. The earliest eluting peptide of the set with the smallest S and log k_0 values was peptide 1 whilst peptide 4 with the highest log k'_0 value in this group was eluted last.

In view of these subtle selectivity differences it is important to ascertain what insight into the orientation of these peptides at the non-polar stationary-phase surface does this chromatographic information provide. It is highly likely that the C-terminal moiety of peptides 1–4, incorporating the sequence TSNVFA, is involved in the binding of the solute to the hydrocarbonaceous ligands for at least two reasons. Firstly, the polarised and fully ionised KKR^{++} sequence is not expected to show any significant hydrophobic interaction with the *n*-butyl chains under the low pH elution conditions employed. Secondly, the overall stability and hydrophobicity of the favoured β-sheet structure of the C-terminal region will be enhanced by the progressive addition of the three alanine residues. This endo-insertion will increase the molecular contact area and result in stronger binding to the stationary phase. An enhancement in the binding affinity is clearly reflected in the corresponding S and log k'_0 values which increase in magnitude with the number of alanine residues adjacent to the threonine residue at position 8. A similar phenomenon also exists with the peptides 2, 8, 9 and 10, where the two arginine residues are separated by one, two, three or four alanine residues, respectively. The concommittant increase in the S and log k'_0 values for these peptides once again illustrates that the affinity of the peptide solute for the stationary-phase surface is enhanced through an extension of the hydrophobic contact area.

Previously, we have argued[6,18,20] that the nature of the secondary and tertiary structure of peptide and protein solutes at the stationary-phase surface must be taken into account when interpretation of their chromatographic behaviour in adsorption chromatography is attempted in terms of either empirical or deterministic models aimed at the optimisation of a particular separation. As is evident from the above data, even with very closely related structural analogues, the thermodynamic and kinetic processes involved in the chromatographic separation of peptides and proteins are subject to much more specific structural and conformational effects than the linear amino acid sequence would imply. These results illustrate that much more refined theoretical treatments, which can accomodate the structural hierachy of peptides and proteins during their passage along the chromatographic bed, must be developed to allow adequate prediction of the experimental behaviour of these biosolutes in reversed-phase, and other adsorption, chromatographic modes. While previous studies with β-endorphin[10,17] LHRH-[11] and GHRF-related peptides[11], growth hormones[12] and interleukin related polypeptides[16] have indicated that the linear solvent strength concepts used to derive log $\bar{k}$ *versus* $\bar{\psi}$ relationships can provide reasonable approximations to the experimental gradient elution retention data, anomalous bandbroadening effects can frequently be observed. These peak broadening effects could not be predicted on the basis that zone dispersion for the particular

peptide can be described by the plate height theory as derived for low molecular weight solutes. In order to clarify whether similar effects operated with the myosin light chain peptides, a similar comparison of the experimental and calculated bandwidths was also carried out.

Bandwidth relationships

The peak capacity for a chromatographic separation of gradient time t_G and average resolution, R_s, equal to unity for all peptide pairs can be given by

$$PC = t_G F/4\sigma_v \tag{7}$$

where σ_v is the bandwidth measured in volume units (1 S.D.). Furthermore, the relationship between σ_v and $\bar{k}$ for linear solvent systems can be expressed as

$$\sigma_{v,calc} = [(\bar{k}/2) + 1]GV_m N^{-0.5} \tag{8}$$

where G is the band compression factor which arises from the increase in solvent strength across the solute zone as the gradient develops along the column, and is given by the expression

$$G^2 = [1 + 2.3b + 1/3\,(2.3b)^2]/(1 + 2.3b)^2 \tag{9}$$

and V_m is the column void volume (equal to t_0F). Under the typical flow-rate conditions used for the separation of peptides in RP-HPLC, the plate number, N, can be approximated by

$$N = \frac{D_m t_0}{C d_p^2} \tag{10}$$

where d_p is the mean particle diameter, and C is the Knox equation parameter which accounts for resistance to mass transfer at the stationary-phase surface and can be estimated by

$$C = \frac{[(1 - x + \bar{k})/(1 + \bar{k})]^2}{15\rho a' + 15\rho b'\bar{k} - 19.2\,\rho x} \tag{11}$$

where x is the interstitial column volume fraction, here taken to be 0.67, a' is equal to 1.1, b' is the surface diffusion parameter and ρ is the restricted diffusion parameter. The diffusion coefficient of a solute in the mobile phase (D_m) can be expressed in terms of the solute molecular weight[17] and evaluated from eqn. (12):

$$D_m = \frac{8.34 \cdot 10^{-10} T}{\eta\,(\mathrm{MW})^{0.33}} \tag{12}$$

where T is the temperature (K) and η is the eluent viscosity.

The use of the general plate height theory to derive the above bandwidth relationships assumes that the solute migrates as a unique, conformationally rigid molecule. It is generally accepted that peptides and proteins can explore a variety of conformations in solution. If it is assumed that these processes or any additional

secondary equilibria are extremely rapid compared to the chromatographic separation time, then the ratio between the experimentally observed bandwidth $\sigma_{v,exp}$, and the bandwidth determined by eqn. 8, $\sigma_{v,calc}$, should approach unity over the normal operational range of retention values commonly used in optimisation studies, *i.e.* $1 < \bar{k} < 10$.

If relatively slow, time-dependent, solvent or stationary phase-induced changes in the secondary or tertiary structure of the peptide solute lead to significant alterations in molecular dimensions or surface topography as the peptide binds to, or is subsequently re-orientated at, the stationary-phase surface, resulting changes in the diffusional and interactive properties of the solute will lead to a pattern of differential zone migration not anticipated by conventional plate height theory. These increases or decreases in mass transfer properties will ultimately be revealed as experimental bandwidths which deviate from the values predicted by eqn. 8.

Theoretical bandwidths for the myosin related peptides were therefore calculated over the range of experimental conditions employed using eqns. 8–12 and compared to experimental peak widths as a function of the gradient steepness parameter, *b*. Fig. 2a shows the conventional plot of $\sigma_{v,exp}/\sigma_{v,calc}$ as a function of $1/b$ for the experiments carried out at a constant flow-rate of 1 ml/min. It can be seen that for peptides 1–4 and 8–10 the σ_v ratio reaches a minimum at relatively small $1/b$ values and then increases again at shallower gradients while the calculated bandwidths for peptides 5–7 were found to more closely approximate the experimental σ_v value over the *b* value range studied. It is evident from these results that even with a set of

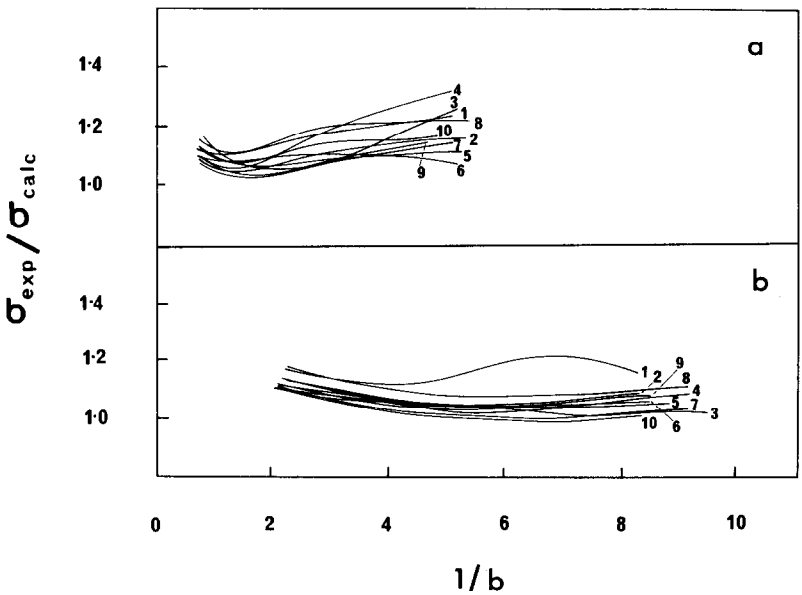

Fig. 2. Plots of $\sigma_{v,exp}/\sigma_{v,calc}$ *versus* $1/b$ for myosin related peptides 1–10. In Fig. 2a, the dependencies shown correspond to data acquired under conditions of different gradient times as in Fig. 1 at a flow-rate of 1 ml/min, while in Fig. 2b, data from the experiments at constant gradient time of 40 min with varying flow-rates were used. In both cases $\sigma_{v,calc}$ was evaluated using the eqn. $\sigma_{v,calc} = [(\bar{k}/2) + 1] G V_m N^{-0.5}$ (eqn. 8). See legend to Fig. 1 for other details.

closely-related peptides such as these analogues, deviations from "ideal" chromatographic behaviour can occur. Although the deviation between $\sigma_{v,exp}$ and $\sigma_{v,calc}$ with steep gradients, *i.e.* large b values, can be ascribed to the so-called J effect[8,11,14], detailed explanations of the underlying reasons for solute-dependent divergencies of $\sigma_{v,exp}$ and $\sigma_{v,calc}$ with very shallow gradients, *i.e.* small b values, have yet to be offered. At this stage, the conformational re-orientation of peptides at hydrocarbonaceous surfaces is one possibility which is consistent with the experimental observations and warrants further investigation. The observed divergencies in bandwidth seen with steep gradients may thus be of general chromatographic origin, associated with the phenomena of gradient generation with current instrumentation, whilst the divergencies in bandwidth seen with shallower gradients may arise as a consequence of solute-specific physico-chemical phenomena associated with solute solvation or structural stabilisation–destabilisation at liquid–solid interfaces.

Fig. 2b shows the plot of $\sigma_{v,exp}/\sigma_{v,calc}$ *versus* $1/b$ for the same set of myosin peptides where t_G was maintained constant at 40 min and the flow-rate varied between 1 and 4 ml/min. It can be seen here that the plots for each peptide are virtually superimposible. Since the total chromatographic residence times in these experiments were similar, it is thus not surprising that no significance differences in zone dispersion phenomena were observed for a specific peptide as F was varied.

Bandbroadening relationships can be further analysed through the comparison of experimental and calculated peak capacity data. Substitution of eqn. 8 into eqn. 7 results in the expression

$$PC = \frac{(\ln 10)}{4} \cdot \frac{S\Delta\psi\ (k)^{0.25}}{Fd_p} \cdot \left[\frac{D_m t_0}{C}\right]^{0.5} \tag{13}$$

When values of the other variables are known experimentally, or can be calculated, eqn. 13 allows a theoretical peak capacity value to be determined for any particular $\bar{k}$. Hence, for separations carried out under regular reversed-phase conditions, and where conformational reorientation is extremely rapid compared to the chromatographic residence time of the solute, *i.e.* when the parameters S and C are time independent, linear dependencies are anticipated between the peak capacity and $\bar{k}^{0.25}/C^{0.5}$.

Fig. 3 shows the plot of PC *versus* $\bar{k}^{0.25}/C^{0.5}$ for the myosin-related peptides with varied t_G and a constant flow-rate of 1 ml/min. In Fig. 3a, plots of PC_{calc}, calculated according to eqn. 13, *versus* $\bar{k}^{0.25}/C^{0.5}$ are presented. As is evident, essentially linear, superimposible plots for the different peptides were obtained. However, when PC was determined according to eqn. 7, *i.e.* directly as PC_{exp}, the magnitude of PC_{exp} was in all cases significantly lower than PC_{calc}, whilst at higher values of $\bar{k}^{0.25}/C^{0.5}$, all PC_{exp} *versus* $\bar{k}^{0.25}/C^{0.5}$ curves deviated from linearity (Fig. 3b). The fact that reasonable correlation is found between PC_{exp} and PC_{calc} for the shorter gradient elution experiments whilst the peak capacity deviation only becomes significant for gradient times above *ca.* 60 min duration would suggest that underestimation of the D_m or C values occurs when t_G is large. These observations are particularly relevant to peptides 1–4 which have coincident amino acid compositions, identical molecular weights and, according to eqn. 10–12, the same bandwidths.

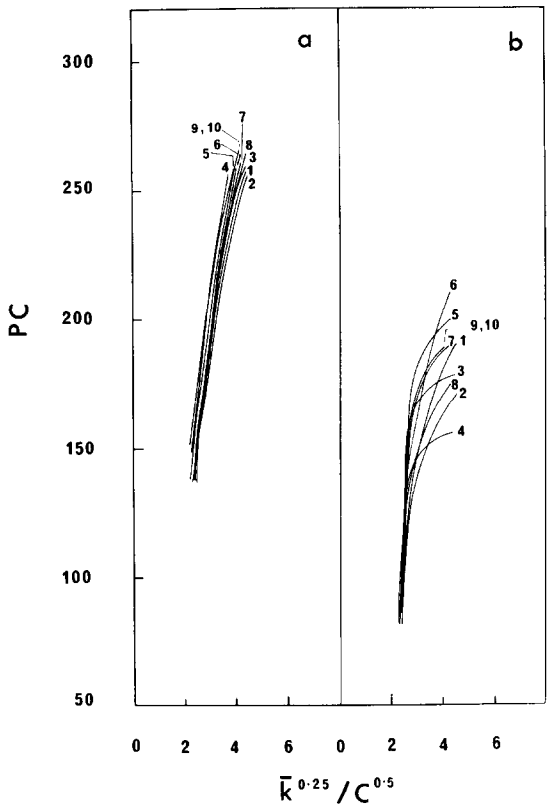

Fig. 3. Plots of peak capacity (*PC*) *versus* $\bar{k}^{0.25}/C^{0.5}$ for the myosin peptides 1–10. The data correspond to values obtained for varied t_G experiments (t_G = 20, 30, 40, 60 and 120 min) and F = 1 ml/min. The *PC* value in panel a was determined from the expression $PC = [1.15\, S\Delta\psi\, (t_0 D_m)^{0.5}\, \bar{k}^{0.25}]/2d_p\, C^{0.5}$ (eqn. 13) while the *PC* value in panel b was calculated from the eqn. $PC = t_G/4\sigma_{t,exp}$, (eqn. 7). See legend to Fig. 1 for other details.

However, the small differences in the N-terminal sequence of these peptides appear to have resulted in different kinetic and/or diffusional behaviour so that the precise values of D_m, C, etc. due to the existence of four distinct migratory species could not be predicted on the basis of molecular weight *per se*. These results further illustrate that more complex structural features must be accommodated into bandspreading model development in order to more accurately describe the RP-HPLC behaviour of peptide solutes.

The influence of flow-rate on the peak capacity for the myosin fragments is shown in Fig. 4. Whilst the shapes of the *PC versus* flow plots generally follow the anticipated profile the plots are not superimposible for all peptides. The variation in the magnitudes of *PC* values and in the position of the curve maxima in the *PC versus F* plots for the different peptides provides an indication that the kinetic processes involved in the RP-HPLC of these peptides are subject to intricate secondary retention effects. Evaluation of such differences in the retention behaviour and band-broadening phenomena when manifested by very closely related sets of peptides has

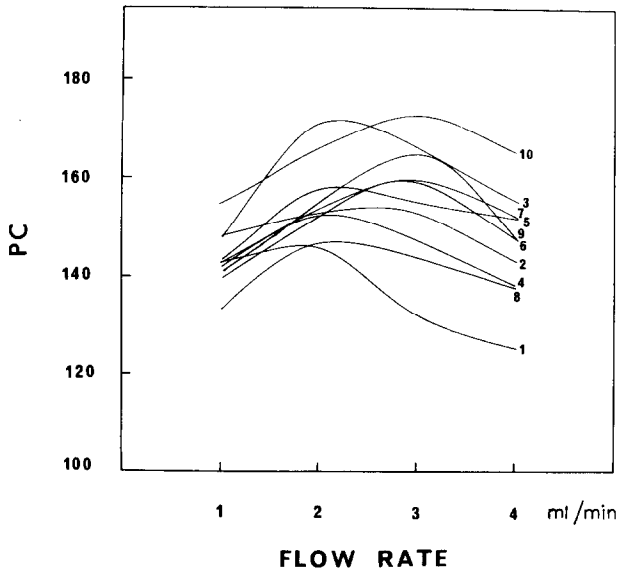

Fig. 4. Variation of peak capacity (*PC*) with flow-rate (ml/min) for the myosin related peptides 1–10. The data correspond to calculated *PC* values obtained for the varied flow experiments where t_G = 40 min and F = 1, 2, 3 and 4 ml/min, respectively. See legend to Fig. 1 for other details.

important implications for RP-HPLC optimisation in the resolution of microheterogenous proteins, and complex mixtures of enzymatic digests. For example, we have shown that quantitative evaluation of bandspacing dependencies as functions of flow-rate and gradient time enable microheterogeneous forms of the glycoprotein hormones, recombinant growth hormones and various truncated forms of the heparin binding growth factors to be readily resolved.

Moreover, a conceptual analogy exists between the interaction of a peptide or protein with its natural biological "receptor" and with any chromatographic surface in terms of the three dimensional spatial requirements for the solute to bind to its complementary ligand(s). Where such binding parallels exist, then biological binding properties such as the apparent association constant which represents quantitative estimates of function, can be correlated via linear free energy relationships to chromatographic parameters. We have previously shown[10,11,26] that the extrathermodynamic functional group effect (τ) can be approximately represented in terms of log k'_0, S and K_m through the linearised expressions

$$\tau = \Delta \log k_0 - \Delta S\psi \tag{14}$$

and

$$[\log K_m]^{-1} = e\tau + d \tag{15}$$

Consequently, plots of $[\log K_m]^{-1}$ *vs.* τ provide a basis to evaluate quantitative structure-activity relationships of polypeptides in terms of constitutional and environ-

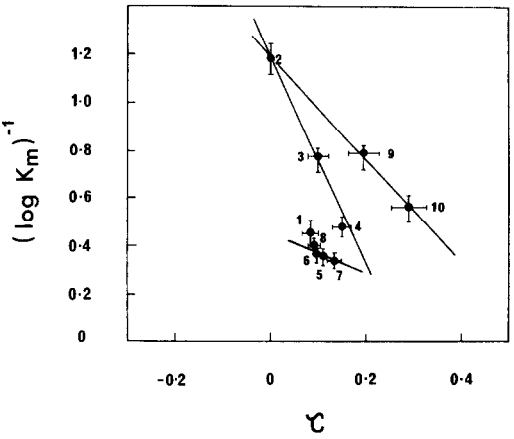

Fig. 5. Plots of the logarithms of the reciprocal of the activity coefficient, K_m, *versus* the functional group τ values for the various myosin light chain peptide analogues 1–10. Also shown are the regression lines for these peptides, grouped according to their sequential arrangement, as derived from the relationship $[\log k_m]^{-1} = e\tau + d$. The point $\tau = 0$ corresponds to peptide 2. See Table I and legend to Fig. 1 for other details.

mental factors. The corresponding plot for the ten myosin-related peptides used in the present study is shown in Fig. 5. No significant differences in the secondary structure of these peptides are evident in their circular dichroism spectra. However, peptides (1–10) can be phosphorylated to different degrees by myosin light chain kinase, with the K_m activities encompassing the range 7.5–950 μmole (Table I). The K_m values for the phosphorylation reaction are a measure of the ability of these peptides to assume a conformation with the appropriate charge distribution which will complement the active site of the kinase. The chromatographic properties of the ten myosin-related peptides will also reflect their three-dimensional structure for the duration of the separation. Although in a reversed-phase system this chromatographic structure may not be identical to the biologically relevant "receptor" structure, the potential nevertheless exists to employ chromatography as an experimental probe of the three dimensional structure of a peptide or protein. Ultimately, it should be possible to take proper advantage of the subtle differences in peptide conformation to maximise resolution whilst at the same time obtaining useful information on the functional status of the solute. When this is achieved the intimate relationships between peptide structure, peptide function and chromatographic performance will be fully characterised, and the quantitative utility of such HPLC approaches to probe structure-function dependencies fully realised.

ACKNOWLEDGEMENT

The support of the National Health and Medical Research Council of Australia to M.T.W.H. is gratefully acknowledged.

LIST OF SYMBOLS

a'	Intercept of the B *versus* $\bar{k}$ plot, taken to be equal to 1.1.
B	Knox equation constant, which arises from zone dispersion due to longitudinal diffusion.
b	Gradient steepness parameter, as defined by eqn. 2.
b'	Solute diffusion parameter, equal to the ratio at solute diffusion in the stationary and mobile phases (D_s/D_m).
C	Knox equation constant, which accounts for mass transfer contributions throughout the pore structure of the stationary phase.
D_m	Diffusion coefficient (cm^2/s) of the solute in the mobile phase.
D_s	Diffusion coefficient (cm^2/s) of the solute at the stationary-phase surface.
d	Intercept of the $[\log K_m]^{-1}$ *versus* τ plot.
d_p	Mean particle diameter (cm).
e	Slope of the $[\log K_m]^{-1}$ *versus* τ plot.
$F_1, F_2 \ldots$	Flow-rate of mobile phase at different values.
G	Band compression factor, as defined by eqn. 9.
K_m	Michaelis constant.
k'	Capacity factor for the solute determined from isocratic retention data.
$\bar{k}$	Median capacity factor for the solute determined from gradient retention data as defined by eqn. 4.
k'_0	Capacity factor determined or extrapolated at $\psi = 0$.
L	Column length (cm).
MW	Molecular weight.
N	Column plate number.
PC_{exp}	Peak capacity as defined by eqn. 7 from experimental values of t_G and $\sigma_{v,exp}$.
PC_{calc}	Peak capacity as defined by eqn. 13.
R_s	Average resolution for gradient separation.
S	Slope of the plot $\log k'$ *versus* ψ as defined by eqn. 6.
T	Absolute temperature (K).
t_d	Chromatographic dwell time (s).
t_e	Gradient elapse time (s).
$t_{G1}, t_{G2} \ldots$	Gradient times (s).
$t_{g1}, t_{g2} \ldots$	Gradient elution times for the solute under two different gradient times.
$t_{0,1}, t_{0,2} \ldots$	Column dead times for different conditions of flow.
V_m	Column void volume.
V_{max}	Maximum velocity of enzyme–substrate reaction.
x	Fraction of the column void volume in the interstitial spaces of the column bed.
α	Gradient separation factor, as defined by the ratio of $\bar{k}$ values for two adjacent solute zones.
β	Ratio of the gradient times as defined by $\beta = t_{G2}/t_{G1}$.
$\Delta\psi$	Change in organic solvent modifier mole fraction.

η	Mobile phase viscosity at a defined temperature (poise).
ρ	Restricted diffusion parameter.
$\sigma_{v1,exp}$	Peak bandwidth in volume units (ml), determined experimentally.
$\sigma_{v1,calc}$	Peak bandwidth in volume units (ml), calculated according to eqn. 8.
τ	Solute selectivity parameter ($= \ln \alpha$).
ψ	Organic solvent modifier mole fraction in isocratic system.
$\bar{\psi}$	Median organic solvent modifier mole fraction as defined by eqn. 5.
χ	Hydrophobic retention coefficient.

REFERENCES

1 R. S. Adelstein and E. Eisenberg, *Ann. Rev. Biochem.*, 49 (1980) 921.
2 R. S. Adelstein and C. B. Klee, *J. Biol. Chem.*, 256 (1981) 7509.
3 J. T. Stull, D. K. Blumenthal, P. de Lanerolle, C. W. High and D. R. Manning, in J. C. Stocklet (Editor), *Advances in Pharmacology and Therapeutics*, Vol. 3, Pergamon Press, Oxford, 1978, p. 171.
4 B. E. Kemp, R. B. Pearson and C. House, *Proc. Natl. Acad. Sci. U.S.A.*, 80 (1983) 7471.
5 B. E. Kemp and R. B. Pearson, *J. Biol. Chem.*, 260 (1985) 3355.
6 M. T. W. Hearn, in J. D. Navratil (Editor), *Chemical Separations: Separation Science and Technology*, Litarvan Press, Colorado, 1986, p. 77.
7 J. Jacobson, W. Melander, G. Vaisnys and C. Horvath, *J. Phys. Chem.*, 88 (1984) 4356.
8 M. T. W. Hearn, A. N. Hodder and M. I. Aguilar, *J. Chromatogr.*, 327 (1985) 47.
9 L. R. Snyder, in Cs. Horvath (Editor), *High-Performance Liquid Chromatography*, Vol. 1, Academic Press, NY, 1980, p. 208.
10 M. I. Aguilar, A. N. Hodder and M. T. W. Hearn, *J. Chromatogr.*, 327 (1985) 115.
11 M. T. W. Hearn and M. I. Aguilar, *J. Chromatogr.*, 359 (1986) 31.
12 M. T. W. Hearn and M. I. Aguilar, *J. Chromatogr.*, in press.
13 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 296 (1984) 31.
14 M. A. Stadalius, M. A. Quarry and L. R. Snyder, *J. Chromatogr.*, 327 (1985) 93.
15 J. L. Glajch, M. A. Quarry, J. F. Vasta and L. R. Snyder, *Anal. Chem.*, 58 (1986) 280.
16 M. Kunitani, P. Hirtzer, D. Johnston, R. Halenbeck, A. Boosman and K. Koths, *J. Chromatogr.*, 359 (1986) 391.
17 M. T. W. Hearn and M. I. Aguilar, *J. Chromatogr.*, 352 (1986) 35.
18 M. T. W. Hearn, M. I. Aguilar and A. N. Hodder, *J. Chromatogr.*, in press.
19 X. M. Lu, K. Benedek and B. L. Karger, *J. Chromatogr.*, 359 (1986) 19.
20 M. T. W. Hearn, *Adv. Chromatogr.*, 20 (1982) 1.
21 M. T. W. Hearn and B. Grego, *J. Chromatogr.*, 255 (1983) 125.
22 R. F. Rekker and H. M. de Kart, *Eur. J. Med. Chem.*, 14 (1979) 479.
23 S. J. Su, B. Grego, B. Niven and M. T. W. Hearn, *J. Liq. Chromatogr.*, 4 (1981) 1945.
24 C. A. Browne, H. P. J. Bennett and S. Solomon, *Anal. Biochem.*, 124 (1982) 201.
25 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.
26 M. T. W. Hearn, *J. Chromatogr.*, submitted for publication.