Journal of Chromatography, 458 (1988) 193–205
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1470

EFFECT OF PEPTIDE CHAIN LENGTH ON PEPTIDE RETENTION BEHAVIOUR IN REVERSED-PHASE CHROMATOGRPHY

COLIN T. MANT, T. W. LORNE BURKE, JAMES A. BLACK and ROBERT S. HODGES*
MRC Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7 (Canada)

SUMMARY

The use of amino acid retention or hydrophobicity coefficients for the prediction of peptide retention time and/or the elution order on hydrophobic stationary phases is based on the premise that amino acid composition is the major factor affecting peptide retention in reversed-phase chromatography. Although this assumption generally agrees well for small peptides (up to *ca.* 15 residues), the retention times of increasingly larger peptides are less than expected from a simple summation of retention coefficients. In the present study, we report the synthesis of four series of peptide polymers which vary significantly in overall hydrophobicity and polypeptide chain length (5–50 amino acid residues, Ac = acetyl): Ac-(G-L-G-A-K-G-A-G-V-G)$_n$-amide (n = 1–5), Ac-(G-K-G-L-G)$_n$-amide (n = 1, 2, 4, 6, 8, 10), Ac-(L-G-L-K-A)$_n$-amide (n = 1, 2, 4, 6, 8, 10) and Ac-(L-G-L-K-L)$_n$-amide (n = 1, 2, 4). From the retention behaviour of these peptide polymers on C_4, C_8 and C_{18} stationary phases under gradient elution conditions, we have clearly established the effect of polypeptide chain length and hydrophobicity on peptide retention. This, in turn, has enabled us to extend the utility of retention time prediction for peptides containing up to 50 residues by introducing a peptide chain-length correction.

INTRODUCTION

The versatility of reversed-phase chromatography (RPC) is reflected in its successful application to the isolation of peptides from a wide variety of sources. Knowledge of the contribution of individual amino acids to peptide retention behaviour on hydrophobic stationary phases, enabling prediction of elution profiles of peptides of known composition, greatly enhances the value of RPC. A major advantage of peptide retention predictions, for instance, is that the position of a peptide(s) of interest in the elution profile of a peptide mixture is narrowed down to a small section of the chromatogram, saving much time and effort in subsequent purification. In addition, information about the relative order of peptide elution from a complex mixture may be obtained. In conjunction with detection by UV absorbance of aromatic residue-containing peptides, fluorescence detection and/or amino acid-specific colour reactions, the identification of specified peptides in a complex mixture may be greatly simplified.

0021-9673/88/$03.50

A major factor governing the retention behaviour of peptides during RPC is the relative hydrophilic/hydrophobic contribution that the side-chains of individual amino acid residues make to the overall hydrophobicity of the peptide. Indeed, several research groups[1–10] have determined sets of coefficients for predicting peptide retention times during RPC, on the assumption that the chromatographic behaviour of a peptide is mainly or solely dependent on amino acid composition. Although this assumption holds up well enough for small peptides (up to *ca.* 15 residues), it could be expected that amino acid sequence may also have an effect on peptide retention. In fact, deviations from predicted retention times and/or elution order for small peptides are generally explained in terms of sequence-specific conformational differences, leading to preferential interaction sites, or anomalous stationary phase interactions[2,3,11–15]. In addition, a non-polar environment, such as a hydrophobic stationary phase, may induce helical structures in potentially helical molecules[16].

Sequence-dependent effects can be divided into two categories: conformational and nearest-neighbour effects. Nearest-neighbour effects can be defined as a reduction in the contribution that an individual side-chain makes to the overall hydrophobicity of the peptide by the close proximity of neighbouring side-chains in the amino acid sequence. For example, when comparing the peptides Gly-Leu-Gly and Leu-Leu-Leu, is the contribution of each leucine side-chain identical in both peptides, or is the average contribution of each leucine in the latter peptide reduced due to nearest-neighbour effects of adjacent leucines? Our definition of nearest-neighbour effects is that they are amino acid sequence-dependent, but independent of conformation. In other words, to prove the existence of amino acid sequence-dependent nearest-neighbour effects, the lack of any defined peptide conformation on interaction of a peptide with the reversed-phase sorbent must be demonstrated. By comparison, amino acid sequence-dependent conformational effects would be a reduction in the overall hydrophobicity of a peptide as a result of the peptide adopting a unique conformation on interacting with the hydrophobic stationary phase, compared to the hydrophobicity of the peptide if it existed as a random coil, *i.e.*, lacking a unique conformation.

Several researchers have noted that peptides larger than 15–20 residues tended to be eluted more rapidly than predicted from hydrophobic considerations alone[3,4,6,13,16–19]. This non-ideal behaviour is generally assumed to be due to stabilized secondary and tertiary structures in the polypeptide which remove certain amino acid residues from contact with the hydrophobic stationary phase. However, it is also possible that there is a peptide chain length effect on retention behaviour of polypeptides, independent of any conformational considerations. Lau *et al.*[20] reported a linear relationship between $\log_{10}$MW and peptide retention time during RPC for a series of five peptide polymers of 8–36 residues. Mant and Hodges[16] demonstrated a similar exponential relationship for a series of five peptide polymers of 10–50 residues. The effect on peptide retention of increasing peptide length decreased progressively with each ten-residue addition.

To understand peptide retention behaviour during RPC completely, it is not sufficient merely to demonstrate that various factors (nearest-neighbour, conformational, peptide chain length) have an effect on peptide retention; it is also necessary to quantitate the relative contribution each factor makes to retention behaviour. Most reported amino acid retention coefficients derived from observed peptide retention

values during RPC have been obtained by computer-calculated regression analysis of the retention times of a wide range of peptides of varied composition and length[1–7,10]. If nearest-neighbour, conformational and peptide chain length effects were present in individual peptides in these peptide mixtures, the coefficients derived by using this approach would be in error and vary among different research groups, depending on the particular peptides used. In addition, because of the low occurrence of certain amino acid residues in any small group of peptides examined, substantial errors may be created for these residues when computer-calculated regression analysis is used to generate the coefficients. A more precise approach to determining retention coefficients was developed by Guo *et al.*[8], who examined the contribution of individual amino acid residues to peptide retention on reversed-phase columns by measuring their effect on retention of a model synthetic peptide (Ac = acetyl): Ac-Gly-X-X-$(Leu)_3$-$(Lys)_2$-amide, where X was substituted by the 20 amino acids found in proteins. This approach overcame many of the problems associated with computer-calculated regression analysis and eliminated any effect of peptide chain length.

In the present study, we wished to examine the effect of polypeptide chain length on peptide retention times during RPC. We have, therefore, synthesized four series of peptide polymers (5–50 residues) of varying hydrophobicity and subjected them to RPC on C_4, C_8 and C_{18} silica-based columns. From the observed retention times of the polymer sets, we have gained a clearer understanding of the effect of both peptide chain length and overall peptide hydrophobicity on peptide retention behaviour during RPC, enabling the accurate prediction of retention times for peptides up to 50 residues in length.

EXPERIMENTAL

Materials

HPLC-grade water and acetonitrile were obtained from J. T. Baker (Phillipsburg, NJ, U.S.A.). HPLC-grade trifluoroacetic acid (TFA) was obtained from Pierce (Rockford, IL, U.S.A.). A synthetic decapeptide reversed-phase standard, S4, and a mixture of five synthetic size-exclusion standards (10–50 residues) were obtained from Synthetic Peptides Inc. (Department of Biochemistry, University of Alberta, Edmonton, Canada).

Peptide synthesis

The peptide polymers described were synthesized on a peptide synthesizer Model 430A (Applied Biosystems, Foster City, CA, U.S.A.) using the general procedure for solid-phase synthesis described by Parker and Hodges[21] and Hodges *et al.*[22].

Apparatus

The HPLC instrument consisted of a Varian Vista Series 5000 liquid chromatograph (Varian, Walnut Creek, CA, U.S.A.) coupled to an HP 1040A detection system, HP 9000 Series 300 computer, HP 9133 disc drive, HP 2225A Thinkjet printer and HP 7440A plotter (Hewlett-Packard, Avondale, PA, U.S.A.). Samples were injected with a 200-μl injection loop (Model 7125; Rheodyne, Cotati, CA, U.S.A.).

Columns

Peptide mixtures were separated on three columns: (1) SynChropak RP-4 (C_4), 250 mm × 4.1 mm I.D., particle size 6.5 μm, pore size 300 Å, carbon loading *ca.* 7.5% (SynChrom, Linden, IN, U.S.A.); (2) Aquapore RP-300 C_8, 220 mm × 4.6 mm I.D., 7 μm, 300 Å (Brownlee Labs., Santa Clara, CA, U.S.A.); (3) SynChropak RP-P (C_{18}), 250 × 4.6 mm I.D., 6.5 μm, 300 Å, carbon loading *ca.* 10% (SynChrom).

RESULTS AND DISCUSSION

Design of peptide polymers

The effect on peptide retention of increasing peptide length is clearly illustrated in the RPC profile (Fig.1) of a mixture of five synthetic peptide size-exclusion standards[16,23] on a C_{18} column at pH 2.0 [linear AB gradient at 1% B/min and 1 ml/min, where eluent A is 0.1% aq. trifluoroacetic acid (TFA) and eluent B is 0.1% TFA in acetonitrile]. The amino acid sequence of the standards is Ac-(G-L-G-A-K-G-A-G-V-G)$_n$-amide, where n = 1–5, *i.e.*, 10–50 residues in length. As reported previously by Mant and Hodges[16], the effect of increasing peptide length on the retention times of these peptides decreased progressively with each addition of a ten-residue repeating unit.

In order to examine further the effect of peptide chain length, as well as peptide hydrophobicity, on peptide retention behaviour, it was necessary to design series of peptide polymers covering a similar range of chain length, but differing in overall hydrophobicity. The design of these polymers, in turn, required the application of an accurate set of amino acid side-chain hydrophobicity parameters. The model peptide approach of Guo *et al.*[8] would be expected to have produced the most accurate set of retention coefficients currently available. Subsequent application of these coefficients resulted in the design and synthesis of three sets of peptide polymers: (a) Ac(G-K-G-L-G)$_n$-amide, where n = 1, 2, 4, 6, 8, 10 (5–50 residues); (b) Ac-(L-G-L-K-A)$_n$-amide,

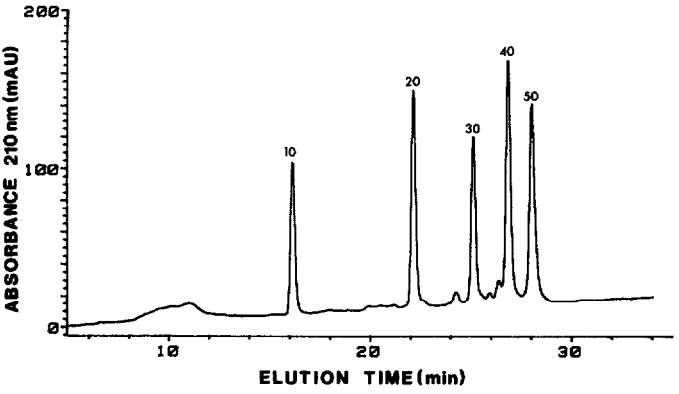

Fig. 1. RPC of a mixture of synthetic peptide polymers. Column: SynChropak RP-P C_{18} (250 mm × 4.6 mm I.D.). Mobile phase: linear AB gradient (1% B/min), where eluent A is 0.1% aq. TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; 26°C. The numbers labelling the peptide peaks denote 10, 20, 30, 40 and 50 residues, respectively. The sequence of the polymer series is Ac-(G-L-G-A-K-G-A-G-V-G)$_n$-amide, where n = 1–5.

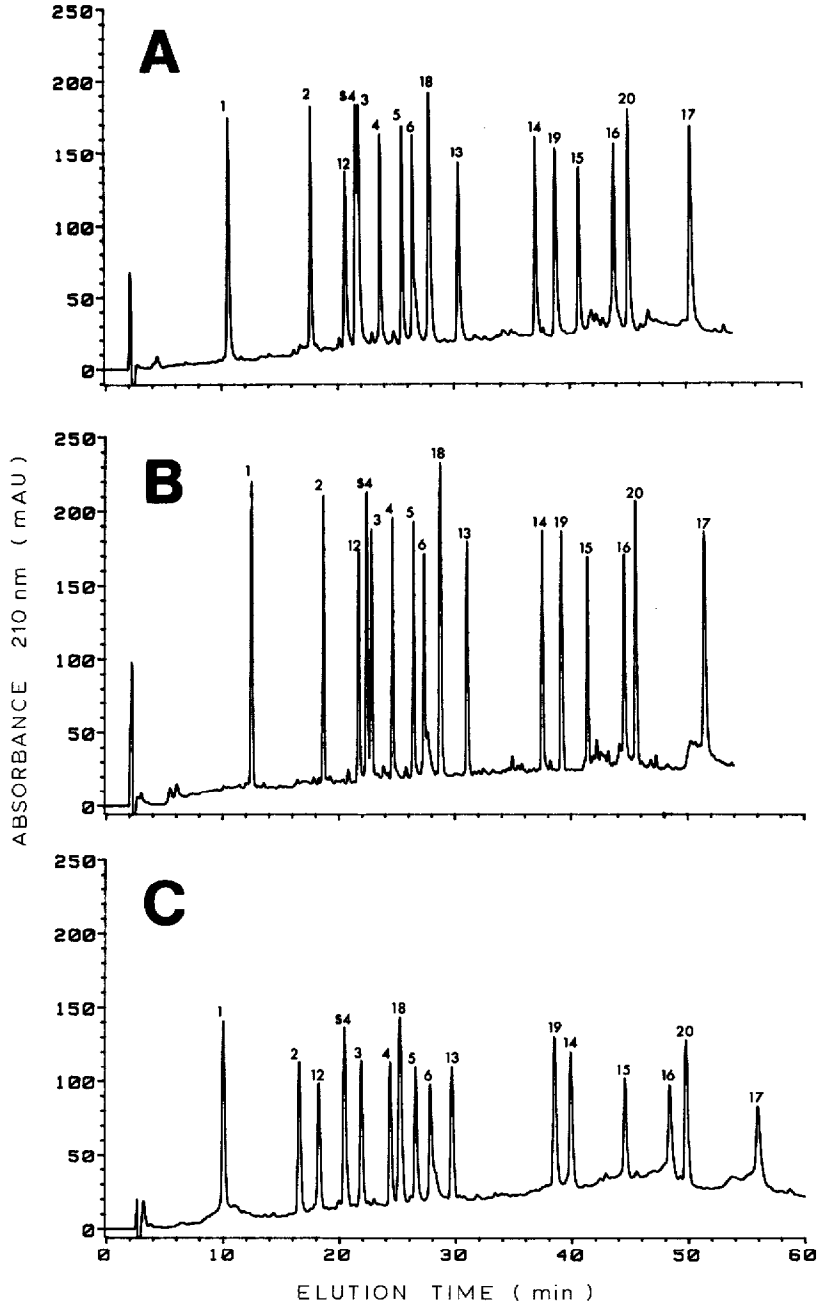

Fig. 2. RPC of a mixture of synthetic peptide polymers. Column: (A) SynChropak RP-4 C_4 (250 mm × 4.1 mm I.D.); (B) Aquapore RP-300 C_8 (220 mm × 4.6 mm I.D.); (C) SynChropak RP-P C_{18} (250 mm × 4.6 mm I.D.). Mobile phase: linear AB gradient (1% B/min), where eluent A is 0.1% aq. TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; 26°C. Numbers denote peptides listed in Table I.

where $n = 1, 2, 4, 6, 8, 10$ (5–50 residues); (c) Ac-(L-G-L-K-L)$_n$-amide, where $n = 1$–4 (5–20 residues). The retention coefficients reported by Guo *et al.*[8] (obtained with aq. TFA to TFA–acetonitrile gradients, pH 2.0) for Lys, Gly, Ala and Leu were −2.1, −0.2, 2.0 and 8.1 min, respectively, *i.e.*, in order of increasing hydrophobicity, K < G < A ≪ L. Thus, the hydrophobicity of the polymer series increased in the order, Ac-(G-K-G-L-G)$_n$-amide ("G" series) < Ac-(L-G-L-K-A)$_n$-amide ("A" series) < Ac-(L-G-L-K-L)$_n$-amide ("L" series). For the purposes of this study, each peptide is referred to by a number and letter which denote, respectively, the number of residues it contains and to which polymer series it belongs. Thus, 5G refers to the five-residue "G" series peptide, 30A refers to the 30-residue "A" series peptide, etc. The presence of a Leu residue in the five-residue repeating unit in the "G" series ensured that this series was sufficiently hydrophobic to be retained by the reversed-phase columns; the presence of a Lys residue in the five-residue repeating units of the polymer sets ensured that the peptides were soluble in 0.1% aq. TFA (pH 2.0); finally, the replacement of a Leu residue by an Ala residue in the five-residue repeating unit of the "A" series ensured that this polymer set was intermediate in hydrophobicity between the hydrophilic "G" series and the very hydrophobic "L" series.

Effect of polypeptide chain length on peptide retention time

Fig. 2 shows elution profiles of a mixture of the "G", "A" and "L" series of peptide polymers on C_4 (A), C_8 (B) and C_{18} (C) reversed-phase columns. The peptides were chromatographed under conditions identical with those employed by Guo *et al.*[8] to obtain their coefficients (linear AB gradient of 1% B/min at a flow-rate of 1 ml/min, where solvent A is 0.1% aq. TFA and solvent B is 0.1% TFA in acetonitrile). Since the coefficients of Guo *et al.*[8] were obtained on a particular column on a particular high-performance liquid chromatographic (HPLC) instrument, an internal synthetic decapeptide standard, S4, was included in each run as an internal peptide standard to correct for different columns and instrumentation[8,9]. The five size-exclusion standards (plus S4) from Fig. 1 (denoted "X" series for the present study) were also chromatographed on all three columns under the same conditions. Observed retention times for all four peptide polymer series are shown in Table I. The peptide elution profiles are very similar on all three columns. The few selectivity differences that are observed between the three columns are the result of a larger change in retention of the five-residue peptides compared to the longer 10- to 50-residue peptides.

Predicted peptide retention times were determined by use of the rules for prediction of peptide retention times, developed by Guo *et al.*[8]

$$\tau = \Sigma R_c + t_s$$

where the predicted retention time, τ, equals the sum of the retention coefficients, ΣR_c, for the amino acid residues, plus the time correction for the internal peptide standard, t_s. The value t_s is obtained by subtracting the sum of the retention coefficients for the peptide standard S4, ΣR_c^{std}, from the observed retention time of the same peptide, t_R^{std}

$$t_s = t_R^{std} - \Sigma R_c^{std}$$

thus, combining these equations:

$$\tau = \Sigma R_c + t_R^{std} - \Sigma R_c^{std}$$

TABLE I

COMPARISON OF PREDICTED AND OBSERVED RETENTION TIMES*

Peptide number	*Peptide designation***	C_4**			C_8			C_{18}		
		t_R^{obs}	τ_c§	Δt§§	t_R^{obs}	τ_c	Δt	t_R^{obs}	τ_c	Δt
1	5G	10.6	9.4	1.2	12.5	10.3	2.2	10.0	8.4	1.6
2	10G	17.7	14.8	2.9	18.7	15.7	3.0	16.6	13.8	2.8
3	20G	21.8	20.6	1.2	22.8	21.3	1.5	21.9	20.5	1.4
4	30G	23.6	22.4	1.2	24.6	23.0	1.6	24.4	22.9	1.5
5	40G	25.5	23.4	2.1	26.5	24.1	2.4	26.6	24.5	2.1
6	50G	26.4	23.9	2.5	27.4	24.6	2.8	27.8	25.6	2.2
7	10X	17.1	18.2	1.1	17.9	19.1	1.2	16.1	16.9	0.8
8	20X	22.2	23.4	1.2	22.6	24.0	1.4	22.1	23.3	1.2
9	30X	24.6	25.7	1.1	25.0	26.4	1.4	25.1	26.3	1.2
10	40X	25.9	27.0	1.1	26.3	27.8	1.5	26.8	28.4	1.6
11	50X	26.8	27.7	0.9	27.1	28.4	1.3	28.0	29.9	1.9
12	5A	20.6	19.9	0.7	21.7	20.8	0.9	18.3	18.9	0.6
13	10A	30.4	29.1	1.3	31.0	29.8	1.2	29.7	29.2	0.5
14	20A	37.0	37.6	0.6	37.5	38.3	0.8	39.9	39.1	0.8
15	30A	40.8	42.9	2.1	41.4	43.6	2.2	44.6	45.9	1.3
16	40A	43.8	46.0	2.2	44.6	46.7	2.1	48.4	50.7	2.3
17	50A	50.4	47.5	2.9	51.5	48.3	3.2	56.0	54.0	2.0
18	5L	27.9	26.0	1.9	28.7	26.9	1.8	25.2	25.0	0.2
19	10L	38.7	35.7	3.0	39.2	36.4	2.8	38.5	36.1	2.4
20	20L	45.0	47.5	2.5	45.6	48.3	2.7	49.8	49.8	0
Average error				1.7			1.9			1.4

* The observed retention times (min) were obtained under conditions of linear AB gradient elution (1% B/min), where eluent A is 0.1% aq. TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; 26°C; absorbance at 210 nm.

** Separations were carried out on a SynChropak RP-4 C_4 column (250 mm × 4.1 mm I.D.; particle size 6.5 μm; pore size 300 Å), an Aquapore RP-300 C_8 column (220 mm × 4.6 mm I.D., 7 μm, 300 Å) and a SynChropak RP-P C_{18} column (250 mm × 4.6 mm I.D.; 6.5 μm; 300 Å).

*** G, X, A and L denote the "G" series of peptide polymers [Ac-(G-K-G-L-G)$_n$-amide], the "X" series of polymers [Ac-(G-L-G-A-K-G-A-G-V-G)$_n$-amide], the "A" series of polymers [Ac-(L-G-L-K-A)$_n$-amide] and the "L" series of polymers [Ac-(L-G-L-K-L)$_n$-amide], respectively. 5G refers to the five-residue "G" series peptide; 30A refers to the 30-residue "A" series peptide, etc.

§ τ_c is the predicted retention time of a peptide, taking peptide chain length into account: $\tau_c = \Sigma R_c + t_s + (m\Sigma R_c \ln N + b)$, where ΣR_c is the sum of the retention coefficients of Guo *et al.*[8] for the amino acid residues in a peptide; N is the number of residues in a peptide; t_s is the time correction for the internal peptide standard (S4; see text for details); m and b are the slope and intercept, respectively, obtained from the semilogarithmic plots [$\tau - t_R^{obs}$ *versus* $\Sigma R_c \ln N$ (Fig. 5)] for each column.

§§ Δt is the error, in min, between the predicted (τ_c) and observed (t_R^{obs}) peptide retention times ($\tau_c - t_R^{obs}$).

The approach to determining the t_s value in the present study was a slight modification of that reported in the study by Guo *et al.*[8]. These researchers expressed the predicted peptide retention time τ as

$$\tau = \Sigma R_c + t_s + t_0$$

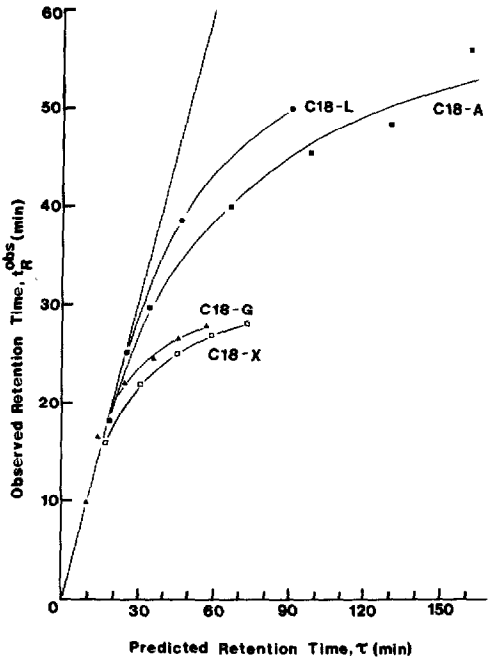

Fig. 3. Deviation of observed from predicted peptide retention times due to polypeptide chain length effect. Column: SynChropak RP-P C_{18} (250 mm × 4.6 mm I.D.). Mobile phase: linear AB gradient (1% B/min), where eluent A is 0.1% aq. TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; 26°C. Absorbance: 210 nm. C18-X, C18-G, C18-A and C18-L denote results for the "X", "G", "A" and "L" series of peptide polymers, respectively, on the C_{18} column. Sequences of the peptide series are described in the text. The straight line represents a perfect correlation between predicted and observed retention times.

where t_0 was the time correction for unretained compounds only. In the present study, the elution time for unretained compounds was not determined independently, being included instead in the overall time correction, t_s, for the internal peptide standard.

If each addition of a repeating unit in the peptide polymers increased peptide hydrophobicity in a linear manner, then a plot of observed *versus* predicted peptide retention time would also show a linear relationship. Fig. 3 illustrates that this is clearly not the case. The results shown were obtained on the C_{18} column (thus, C18-L denotes "L" series peptide polymers on the C_{18} column, etc.), but are also representative of the C_4 and C_8 columns. Although there is generally a good correlation of observed and predicted retention times for the five-residue peptides (5G, etc.) and also the ten-residue peptides of the least hydrophobic "G" and "X" series (10G, 10X), this correlation falls off rapidly as the peptide chain length increases up to 50 residues. This good correlation for small peptides reflects the work of Guo *et al.*[9], who showed excellent correlation of observed and predicted retention times of 58 peptides of 2–16 residues, 43 of these peptides lying in the 5–12 residue range. It is interesting to note that peptides 10A and 10L, though fairly small, showed significant differences between observed and predicted retention times (5 and 8 min, respectively). These peptides are quite hydrophobic, particularly as compared to peptides 10G and 10X. It should also

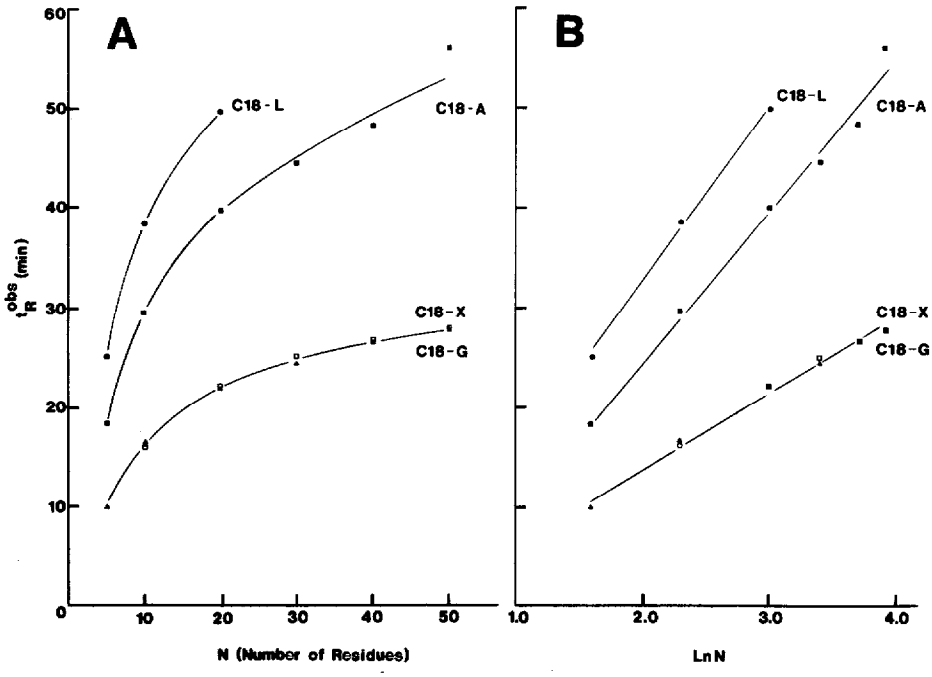

Fig. 4. Effect of polypeptide chain length on observed peptide retention times in RPC. Observed peptide retention time, t_R^{obs}, *versus* N (number of residues) (A) or ln N (B). Conditions as in Fig. 3.

be noted that the greater the hydrophobicity of a peptide at a particular chain length, the greater the deviation of observed from predicted retention time. Tchapla *et al.*[24] reported a similar increasing deviation of expected solute retention time with increasing length of solute molecule during RPC of various homologous series of non-peptide molecules. In addition, the shorter the alkyl ligand of the RPC sorbent (C_6–C_{18}), the earlier the observed deviations. The authors postulated a change in the retention process as the length of the molecules increased. However, in the present study, the deviation from linearity of the predicted *versus* observed peptide retention time plot for the C_{18} (Fig. 3), C_8 and C_4 columns all occurred at the same peptide chain length. This suggests that the above authors' explanation[24] for the retention behavior of non-peptide molecules is probably not applicable to the present study involving RPC of peptide polymers.

The relationship between observed peptide retention time on the C_{18} column and peptide chain length is illustrated in Fig. 4. These results are again representative of all three reversed-phase columns used in this study. Fig. 4A shows the non-linear relationship between observed peptide retention time and the number of residues, N, the peptides contain. Not surprisingly, given the linear relationship between τ and N, the curved profiles resulting from these plots are similar to those illustrated in Fig. 3. The exponential nature of the relationship between peptide retention time and peptide chain length is illustrated in Fig. 4B. Plotting observed peptide retention time *versus* the logarithm of the number of residues (ln N) resulted in straight-line plots with different slopes, depending on the hydrophobicity of a particular peptide polymer

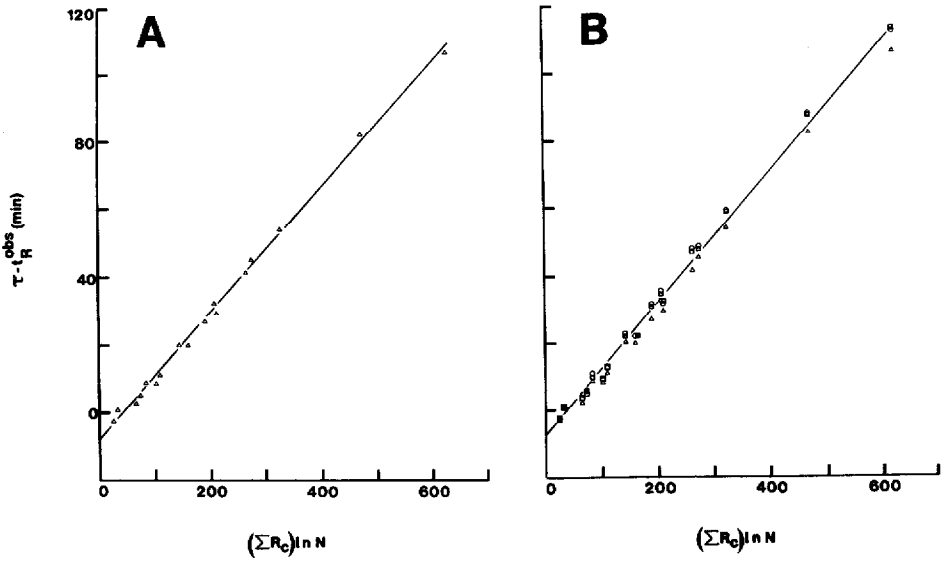

Fig. 5. Correlation of peptide retention time with peptide chain length and hydrophobicity. Predicted minus observed peptide retention time ($\tau - t_R^{obs}$) *versus* $\Sigma R_c \ln N$, where ΣR_c is the sum of the retention coefficients of Guo *et al.*[8] for the amino acid residues in a peptide, and N is the number of residues in a peptide. (A) Results for four series of peptide polymers ("X", "G", "A" and "L" series; see text for details) on a SynChropak RP-P C_{18} column (250 mm × 4.6 mm I.D.). (B) results for four series of peptide polymers on a SynChropak RP-4 C_4 column (250 mm × 4.1 mm I.D.), Aquapore RP-300 C_8 column (220 mm × 4.6 mm I.D.) and SynChropak RP-P C_{18} column (250 mm × 4.6 mm I.D.). The five-residue peptides were not included in the plots. Mobile phase conditions as described in Fig. 3. Absorbance at 210 nm.

series. The "G" and "X" series of polymers are very similar in hydrophobicity, resulting in overlapping profiles in Fig. 4. The slopes of the plots shown in Fig. 4B increased with increasing hydrophobicity of the peptide polymers, *i.e.*, "G" ≈ "X" < "A" < "L" series.

Correlation of peptide retention time with peptide chain length and hydrophobicity

From Fig. 3 it was apparent that a clearer understanding of peptide retention behaviour during RPC required clarification of the effects of both peptide chain length and peptide hydrophobicity on observed retention times. Although the observed peptide retention time data could be linearized with respect to peptide chain length (t_R^{obs} *versus* ln N) (Fig. 4B), the resulting straight-line plots diverged, with the slopes dependent on the hydrophobicity of a particular peptide polymer series.

The intimate relationship between peptide hydrophobicity and chain length and their combined effect on peptide retention behaviour is clearly illustrated in Fig. 5. Plotting predicted (τ) minus observed (t_R^{obs}) peptide retention time *versus* the product of peptide hydrophobicity (expressed as ΣR_c, the sum of the coefficients of Guo *et al.*[8]) and the logarithm of the number of residues (ln N) resulted in a single, straight-line plot. Thus, the discrepancy between predicted and observed peptide retention times is linearly related to $\Sigma R_c \ln N$. Fig. 5A demonstrates the plot for the C_{18} column. All four sets of peptide polymers fall on the line, with an overall correlation of 1.00 (determined

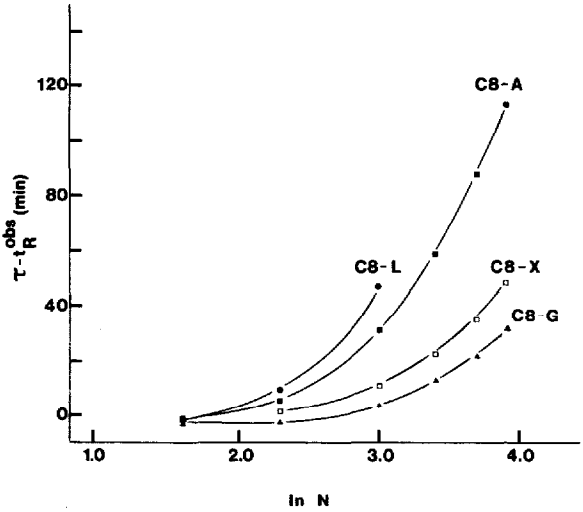

Fig. 6. Plot of predicted minus observed peptide retention time ($\tau - t_R^{obs}$) *versus* the logarithm of the number of residues (ln N). Column: Aquapore RP-300 C_8 (220 mm × 4.6 mm I.D.). Mobile phase conditions as described in Fig. 3; absorbance at 210 nm. C8-X, C8-G, C8-A and C8-L denote "X", "G", "A" and "L" series of peptide polymers, respectively, on the C_8 column. Sequences of the peptide series are described in the text.

by linear least-squares fitting). The profile shown in Fig. 5B was obtained by plotting the data for all three columns. The high correlation (r = 0.99) highlights the consistency of the ($\tau - t_R^{obs}$) *versus* ΣR_c ln N relationship on different reversed-phase columns, varying in dimensions, hydrophobic functionalities (n-alkyl chain length) and ligand density. If ln MW (logarithm of peptide molecular weight) replaces ln N in the above relationship, the correlation of the resulting plot is not as high. The five-residue peptides were not included in these plots, since there was essentially no discrepancy between their observed and predicted retention values (Fig. 3).

When the expression denoting peptide hydrophobicity (ΣR_c) is removed from the relationship producing the straight-line plots demonstrated in Fig. 5, *i.e.*, plotting ($\tau - t_R^{obs}$) *versus* ln N, the profiles for the four peptide series become non-linear and diverge (Fig. 6). Results are shown for the C_8 column only, although, once again, the plots shown are also representative of the C_4 and C_{18} columns. Fig. 6 again stresses the importance of taking the hydrophobicity of a polypeptide into account when attempting to correlate its retention time with the number of residues it contains.

Prediction of polypeptide retention time in RPC

From Fig. 5:

$$\tau - t_R^{obs} \propto \Sigma R_c \ln N$$
$$\tau - t_R^{obs} = m\Sigma R_c \ln N + b$$
$$t_R^{obs} = \tau - (m\Sigma R_c \ln N + b) \quad (1)$$

As described above

$$\tau = \Sigma R_c + t_s \quad (2)$$

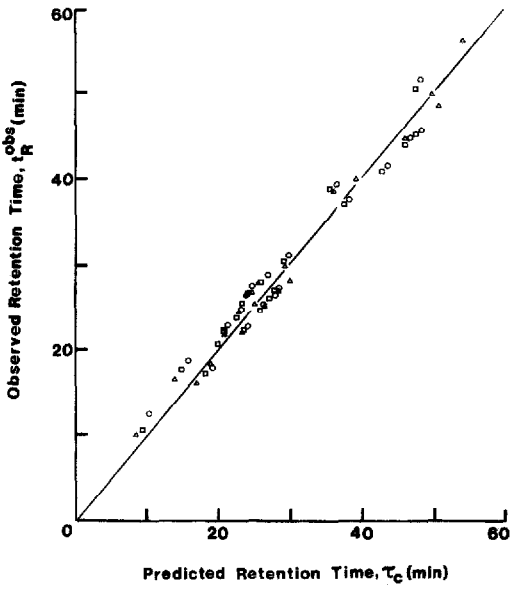

Fig. 7. Correlation of predicted and observed peptide retention times in RPC. Results shown are for four series of peptide polymers ("X", "G", "A" and "L" series; see text for details) on a SynChropak RP-4 C_4 column (250 mm × 4.1 mm I.D.), Aquapore RP-300 C_8 column (220 mm × 4.6 mm I.D.) and SynChropak RP-P C_{18} column (250 mm × 4.6 mm I.D.). The predicted retention times, taking peptide length into account (τ_c), were calculated as described in the text. Mobile phase conditions as described in Fig. 3. Absorbance at 210 nm.

where t_s is the time correction for the peptide standard, S4. Substituting eqn. 2 into eqn. 1 produces the expression:

$$t_R^{obs} = \Sigma R_c + t_s - (m\Sigma R_c \ln N + b) \quad (3)$$

When predicting the retention time of peptides, taking into account peptide chain length, t_R^{obs} in eqn. 3 becomes τ_c (predicted polypeptide retention time)

$$\tau_c = \Sigma R_c + t_s - (m\Sigma R_c \ln N + b) \quad (4)$$

where $(m\Sigma R_c \ln N + b)$ is the correction factor for peptide chain length.

Eqn. 4 was applied to retention time prediction of all four series of peptide polymers on all three reversed-phase columns used in this study, and the results are shown in Table I. The average deviations, Δt, of predicted values, τ_c, from observed peptide retention times, t_R^{obs}, were only 1.7, 1.9 and 1.4 min for the C_4, C_8 and C_{18} columns, respectively. These small deviations are indicative of the high predictive accuracy of this method for peptides up to 50 residues in length. This accuracy is again highlighted by the high correlation ($r = 0.99$) of predicted *versus* observed peptide retention times for the peptide polymers on all three columns (Fig. 7). Comparison of Fig. 7 with Fig. 3 demonstrates the impressive improvement in prediction of

polypeptide retention times when peptide chain length is taken into account. This improvement is especially gratifying, considering the stringent test conditions for this method, covering as it does an extremely wide range of peptide hydrophobicities to values far exceeding those of most peptides encountered.

This study has clearly demonstrated that, if a peptide is not subject to sequence-dependent conformation or nearest-neighbour effects, its reversed-phase chromatographic behaviour can be correlated with its amino acid composition and the number of residues in the polypeptide chain.

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council of Canada and equipment grants from the Alberta Heritage Foundation for Medical Research.

REFERENCES

1 J. L. Meek, *Proc. Natl. Acad. Sci. U.S.A.*, 77 (1980) 1632.
2 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.
3 S. J. Su, B. Grego, B. Niven and M. T. W. Hearn, *J. Liq. Chromatogr.*, 4 (1981) 1745.
4 K. J. Wilson, A. Honegger, R. P. Stötzel and G. J. Hughes, *Biochem. J.*, 199 (1981) 31.
5 C. A. Browne, H. P. J. Bennett and S. Solomon, *Anal. Biochem.*, 124 (1982) 201.
6 T. Sasagawa, T. Okuyama and D. C. Teller, *J. Chromatogr.*, 240 (1982) 329.
7 T. Sasagawa, L. H. Ericsson, D. C. Teller, K. Titani and K. A. Walsh, *J. Chromatogr.*, 307 (1984) 29.
8 D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 499.
9 D. Guo, C. T. Mant, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 519.
10 C. T. Mant and R. S. Hodges, in M. T. W. Hearn (Editor), *HPLC of Proteins, Peptides and Polynucleotides*, VCH, Weinheim, in press.
11 S. Terabe, R. Konaka and K. Inouye, *J. Chromatogr.*, 172 (1979) 163.
12 M. T. W. Hearn and B. Grego, *J. Chromatogr.*, 203 (1981) 349.
13 E. C. Nice, M. W. Capp, N. Cooke and M. J. O'Hare, *J. Chromatogr.*, 218 (1981) 569.
14 M. T. W. Hearn and M. I. Aguilar, *J. Chromatogr.*, 359 (1986) 31.
15 M. T. W. Hearn and M. I. Aguilar, *J. Chromatogr.*, 392 (1987) 33.
16 C. T. Mant and R. S. Hodges, in K. Gooding and F. Regnier (Editors), *High-Performance Liquid Chromatography of Biological Macromolecules: Methods and Applications*, Marcel Dekker, New York, in press.
17 M. J. O'Hare and E. C. Nice, *J. Chromatogr.*, 171 (1979) 209.
18 K. J. Wilson, A. Honegger and G. J. Hughes, *Biochem. J.*, 199 (1981) 43.
19 C. T. Wehr, L. Correia and S. R. Abbott, *J. Chromatogr. Sci.*, 20 (1982) 114.
20 S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 317 (1984) 129.
21 J. M. R. Parker and R. S. Hodges, *J. Protein Chem.*, 3 (1985) 465.
22 R. S. Hodges, R. J. Heaton, J. M. R. Parker, L. Molday and R. S. Molday, *J. Biol. Chem.*, 263 (1988) 11768.
23 C. T. Mant, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 397 (1987) 99.
24 A. Tchapla, H. Colin and G. Guiochon, *Anal. Chem.*, 56 (1984) 621.