Journal of Chromatography, 476 (1989) 363–375
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1630

CORRELATION OF PROTEIN RETENTION TIMES IN REVERSED-PHASE CHROMATOGRAPHY WITH POLYPEPTIDE CHAIN LENGTH AND HYDROPHOBICITY

COLIN T. MANT*, NIAN E. ZHOU and ROBERT S. HODGES
Department of Biochemistry and the Medical Research Council of Canada Group in Protein Structure and Function, 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 behaviour 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 holds up well enough for small peptides (up to *ca.* 15 residues), it is now recognized that polypeptide chain length must be taken into account when attempting to equate retention time behaviour of larger peptides and proteins with their overall hydrophobicity. In the present study, we have examined the reversed-phase retention behaviour of 19 proteins of known sequence on stationary phases of varying hydrophobicity and ligand density. From the observed protein retention behaviour on C_4, C_8 and C_{18} stationary phases under gradient elution conditions, we have been able to correlate the observed retention times of proteins ranging in molecular weight from 3500 to 32 000 dalton and in chain length from 30 to 300 residues with their overall hydrophobicity (based on retention parameters derived from small peptides) and the number of residues in the polypeptide chain. The retention behaviour of the proteins on the C_4, C_8 and C_{18} columns was also compared to that obtained on supports containing lower ligand densities (phenyl ligands). The maintenance of native or partially folded protein conformation on the phenyl columns, resulting in lower retention times than would be expected for fully denatured proteins, underlined the importance of efficient protein denaturation for satisfactory correlation of protein retention times with protein hydrophobicity. In addition, the effectiveness of increasing temperature and/or ligand density of the stationary phase in denaturing proteins was also demonstrated.

INTRODUCTION

The most widely-used mode of high-performance liquid chromatography (HPLC) for peptide separations is, by far, reversed-phase chromatography (RPC)[1]. Although RPC has, for the most part, been used for the separation of relatively small molecules (< 50 residues), its application to the separation of larger polypeptides and proteins has seen a significant increase in recent years.

0021-9673/89/$03.50

A major factor governing the retention behaviour of peptides and proteins during RPC is the relative hydrophilic/hydrophobic contribution that the side-chains of individual amino acid residues make to the overall hydrophobicity of the molecule. In fact, several research groups[2–13] have determined sets of coefficients for predicting peptide retention times during RPC on the assumption that the contribution of each residue to peptide retention is additive and that retention time is linearly related to the sum of the contribution of each residue. This assumption, that the chromatographic behaviour of a peptide is mainly or solely dependent on amino acid composition, holds up well enough for small peptides (up to *ca.* 15 residues), but several researchers have noted that peptides larger than 15–20 residues tended to be eluted more rapidly than predicted from hydrophobic considerations alone[1,4,5,7,14–17]. Attempts by various researchers[14,15,18] to correlate protein retention times with protein hydrophobicity, as expressed by the sum of Recker fragmental constants[14,15] or the sum of HPLC-derived retention parameters[18], have been largely unsuccessful.

In the present study, we have subjected 23 proteins of known sequence to RPC on five stationary phases of varying hydrophobicity and ligand density. From the observed protein retention behaviour, we have been able to correlate the observed retention times of proteins ranging in molecular weight from 3500 to 32 000 dalton and in chain length from 30 to 300 residues with their overall hydrophobicity (expressed as the sum of amino acid side-chain hydrophobicity coefficients) and the number of residues in the polypeptide chain.

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 peptide standard, S4, and a mixture of five synthetic peptide polymers [Ac-(G-L-G-A-K-G-A-G-V-G)$_n$-amide, where n = 1, 2, 3, 4 and 5] (referred to as the "X" series of peptide polymers in the text) were obtained from Synthetic Peptides Inc. (Department of Biochemistry, University of Alberta, Edmonton, Canada). Bovine insulin, bovine insulin (chain B), equine cytochrome *c*, bovine α-lactalbumin, bovine ribonuclease A, chicken avidin, chicken lysozyme, sperm whale myoglobin, sheep prolactin, papain, jack bean concanavalin A, porcine elastase, bovine α-chymotrypsinogen A, equine alcohol dehydrogenase, chicken ovalbumin, bakers yeast enolase and bovine serum albumin (BSA) were obtained from Sigma (St. Louis, MO, U.S.A.). Rabbit skeletal troponin C (RsTnC), turkey skeletal TnC, rabbit skeletal troponin I (RsTnI), rabbit skeletal troponin T (RsTnT), rabbit cardiac troponin T (RcTnT), and rabbit cardiac tropomyosin (RcTM) were prepared from tissue extracts in this laboratory.

Peptide synthesis

Three series of peptide polymers were synthesized on an Applied Biosystems (Foster City, CA, U.S.A.) peptide synthesizer Model 430A, using the general procedure for solid phase synthesis described by Parker and Hodges[19] and Hodges *et al.*[20]. The sequences of the peptides were ("G" series) Ac-(G-K-G-L-G)$_n$-amide, where n = 1, 2, 4, 6, 8, 10 (5–50 residues); ("A" series) Ac-(L-G-L-K-A)$_n$-amide, were n = 1,

2, 4, 6, 8, 10 (5–50 residues); ("L" series) Ac-(L-G-L-K-L)$_n$-amide, where n = 1, 2, 4 (5–20 residues).

Apparatus

The HPLC instrument consisted of a Hewlett-Packard (Avondale, PA, U.S.A.) HP 1090 liquid chromatograph coupled to an HP 1040A detection system, HP 9000 Series 300 computer, HP 9133 disc drive, HP 2225A Thinkjet printer and HP 7440A plotter.

Columns

Proteins were separated on five columns: (1) SynChropak RP-4 (C_4), 250 × 4.1 mm I.D., 6.5 μm particle size, 300 Å pore size, *ca.* 7.5% carbon loading (SynChrom, Linden, IN, U.S.A.); (2) Aquapore RP-300 (C_8), 220 × 4.6 mm I.D., 7 μm, 300 Å (Brownlee Labs., CA, U.S.A.); (3) SynChropak RP-P (C_{18}), 250 × 4.6 mm I.D., 6.5 μm, 300 Å, *ca.* 10% carbon loading (SynChrom); (4) Bio-Gel TSK Phenyl-5PW, 75 × 7.5 mm I.D., 10 μm, 1000 Å (Bio-Rad Labs., Richmond, CA, U.S.A.); (5) Bio-Gel TSK Phenyl RP+, 75 × 4.6 mm I.D., 10 μm, 1000 Å (Bio-Rad Labs.).

RESULTS AND DISCUSSION

In order to correlate overall hydrophobicity of a polypeptide or protein with its retention behaviour in RPC, it is important to have an accurate means of expressing this hydrophobicity. Most attempts to determine the effective contribution of each amino acid side chain and end group to the retention process in RPC from HPLC-derived data have involved computer-calculated regression analyses of the retention times of a wide range of peptides of varied composition[2–8,11,12]. Comparison of the various sets of retention indices or coefficients derived in this manner showed numerous discrepancies both in the relative order of hydrophobicities of the amino acid side chains, and in the magnitude of the contributions of specific residues[9]. A possible explanation for these discrepancies is that certain residues did not appear often enough in the various peptide mixtures used to enable an accurate determination of their contributions. In addition, possible polypeptide chain-length-dependence effects are not being taken into account since the peptides used to determine the coefficients were of a wide range of size, composition and sequence[13]. The most precise set of retention coefficients yet determined was reported by Guo and co-workers[9,10] who examined the contribution of individual amino acid residues to retention of a model synthetic peptide: Ac-G-X-X(L)$_3$-(K)$_2$-amide, where X was substituted by the 20 naturally occurring amino acids found in proteins. This method overcame the problems associated with the computer-calculated regression analysis approach, and the retention parameters derived from this work were applied to the present study.

Guo *et al.*[9] obtained their retention coefficients by subjecting their peptide analogues to linear 0.1% aqueous TFA to 0.1% TFA in acetonitrile gradients at pH 2.0, and these conditions were also used in the present study. The acidic nature (pH 2.0) of the TFA-containing mobile phase suppresses the ionization of surface silanols on silica-based columns, thereby overcoming undesirable ionic interactions between basic solutes and the column packing. In addition, TFA not only provides an acidic medium, but is an excellent protein solubilizing agent[21,22], which is used routinely in solid phase peptide synthesis to extract peptides and proteins from the resin support after cleavage.

Complete denaturation of a protein is required for full expression of its overall hydrophobicity. The hydrophobicity of a protein in its native conformation is dramatically different from its unfolded state, since the hydrophobic side-chains are buried during the folding process. The hydrophobic interactions stabilizing the three-dimensional structure of a protein must be disrupted to maximize interaction of the polypeptide chain with the reversed-phase sorbent. Lau *et al.*[22] demonstrated that the primary cause of protein denaturation during RPC is the hydrophobicity of the stationary phase which disrupts the hydrophobic interactions stabilizing the native conformation. These workers showed that even an ultra-short (C_3) sorbent with low carbon loading was able to denature very stable synthetic model proteins. These model proteins consisted of two-stranded α-helical coiled coils in which the quaternary structure was stabilized by hydrophobic interactions between the two α-helics. These proteins are probably the most stable proteins yet reported as indicated by temperature and denaturation studies in 0.1% aqueous. TFA which is a starting solvent for RPC. For example, the coiled coil consisting of two 35-residue chains was only 30% denatured at temperatures greater than 70°C[23,24]. Yet, interaction with the hydrophobic matrix during RPC caused disruption of the hydrophobic interactions that

TABLE I
PROTEINS USED IN THIS STUDY

Protein	*Molecular weight*	*N*[a]	*Relative hydrophobicity*[b]
1 Insulin (chain B)[c]	3500	30	0.49
2 Insulin	6030	51	0.79
3 Cytochrome *c*	11 700	104	1.00
4 α-Lactalbumin	14 180	123	1.76
5 Ribonuclease A	13 690	124	1.10
6 Avidin	14 330	128	1.55
7 Lysozyme	14 310	129	1.62
8 Myoglobin	17 200	153	1.71
9 RsTnC	17 960	159	2.25
10 Turkey TnC	18 000	162	2.17
11 RsTnI	20 700	178	1.89
12 Prolactin	22 550	198	2.89
13 Papain	23 430	212	2.83
14 Concanavalin A	25 570	237	3.25
15 Elastase	25 900	240	3.33
16 α-Chymotrypsinogen A	25 670	245	3.35
17 RsTnT	30 520	259	2.39
18 RcTnT	32 880	276	2.57
19 RcTM	32 000	284	3.20
20 Alcohol dehydrogenase	39 800	374	5.46
21 Ovalbumin	45 000	386	6.50
22 Enolase	46 700	436	5.97
23 BSA	66 300	582	7.82

[a] Number of amino acid residues.

[b] Expressed as ΣR_c of the protein/ΣR_c of cytochrome *c*, where ΣR_c is the sum of the retention coefficients of the amino acid residues as reported by Guo *et al.*[9].

[c] Although only polypeptides of more than 50 residues are classed as proteins, insulin (chain B) was included as a protein for the purposes of the present study.

stabilized these proteins, suggesting that this denaturation during RPC would be representative of the situation for most proteins. Thus, the 23 proteins listed in Table I were subjected to linear gradient elution (1% B/min), at a flow-rate of 1 ml/min, on analytical C_4, C_8 and C_{18} columns, where eluent A was 0.1% aqueous TFA and eluent B was 0.1% TFA in acetonitrile. Considering the work of Lau *et al.*[22], it was assumed that the extremely hydrophobic nature of these sorbents would denature these proteins. In addition, the 300-Å pore size of the three columns is very suitable for separating both peptides and proteins[1,22,25]. To standardize retention behaviour, each protein was run separately with an internal peptide standard, S4, and cytochrome *c* as a protein standard. A representative separation of a mixture of eight proteins, including cytochrome *c* (protein 3), and peptide S4 on the C_8 column is shown in Fig. 1. Elution profiles obtained on the C_4 and C_{18} columns were very similar.

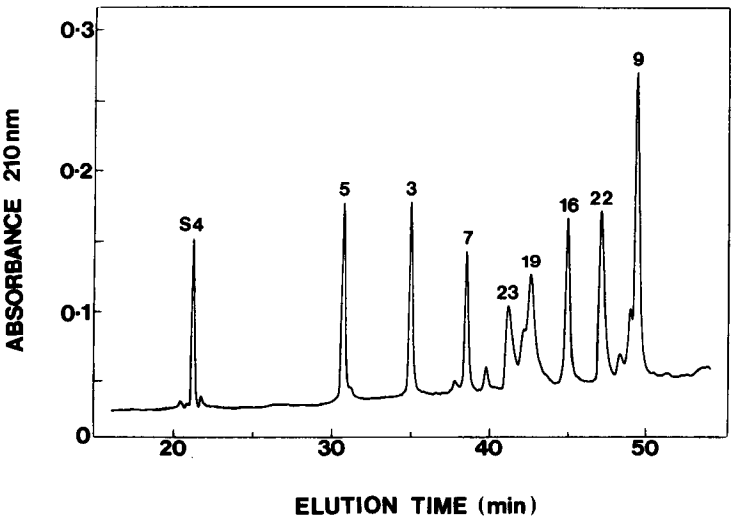

Fig. 1. RPC of a mixture of proteins. Column, Aquapore RP-300 C_8 (220 × 4.6 mm I.D.); mobile phase, linear A–B gradient (1% B/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0), flow-rate, 1 ml/min; temperature, 26°C. Numbers denote proteins listed in Table I. S4 is a synthetic reversed-phase decapeptide standard.

Effect of polypeptide chain length on retention time

There is a peptide chain length effect on retention behaviour of polypeptides, independent of any conformational considerations[26]. Thus, Lau *et al.*[22] reported a linear relationship between log molecular weight and peptide retention time during RPC for a series of five peptide polymers of 8–36 residues. Mant and Hodges[1] 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.

Fig. 2 compares the results of plotting observed retention times on the C_{18} column *versus* the logarithm of the number of residues (ln *N*) for four series of peptide polymers (Fig. 2A) and the 23 proteins listed in Table I (Fig. 2B). The peptide polymers

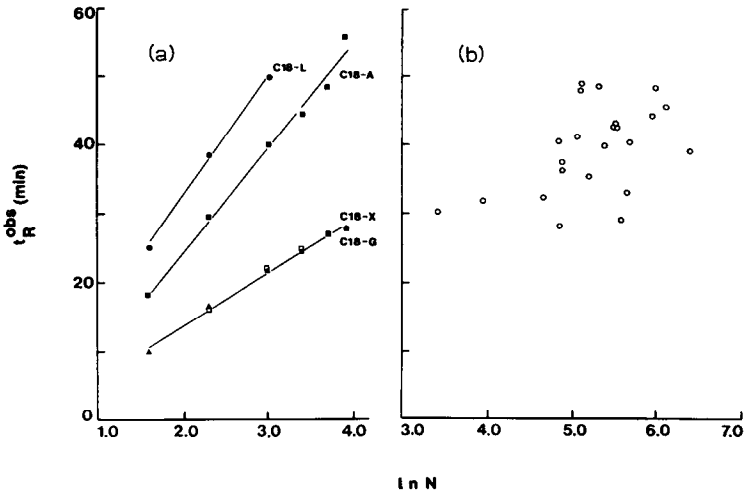

Fig. 2. Effect of polypeptide chain length on observed peptide and protein retention times in RPC. Observed retention time (t_R^{obs}) *versus* ln N (where N = number of residues) for four series of peptide polymers (a) or 23 proteins (b). Column, SynChropak RP-P C_{18} (250 × 4.6mm I.D.); mobile phase, linear A–B gradient (1% B/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; temperature, 26°C; absorbance, 210 nm. C18-X, C18-G, C18-A, C18-L denote results for the "X", "G", "A" and "L" series of peptide polymers, respectively. Sequences of the polymers are shown in the Experimental section. The proteins are listed in Table I.

(described in the Experimental section) were run under the same conditions as described above for the proteins. The plots for the polymers resulted in straight-line plots with different slopes, depending on the hydrophobicity of a particular peptide polymer series. The "G" and "X" series of polymers are very similar in hydrophobicity, resulting in overlapping profiles (Fig. 2A). The slopes of the plots in Fig. 2A increased with increasing hydrophobicity of the peptide polymers, *i.e.*, "G" ≈ "X" < "A" < "L" series. By comparison, the plot for the proteins (Fig. 2B) did not show any linear relationship. This was not surprising since, unlike the peptide polymers which increase in length and hydrophobicity in a well-defined manner, the proteins differ widely in size, composition and hydrophobicity. The results shown in Fig. 2 suggested that a clearer understanding of protein retention behaviour during RPC required clarification of the effects of polypeptide hydrophobicity, as well as chain length, on observed retention times.

Correlation of protein retention behaviour with polypeptide chain length and hydrophobicity

Predicted protein retention times were determined by use of the rules for prediction of peptide retention times developed by Guo *et al.*[9],

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

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

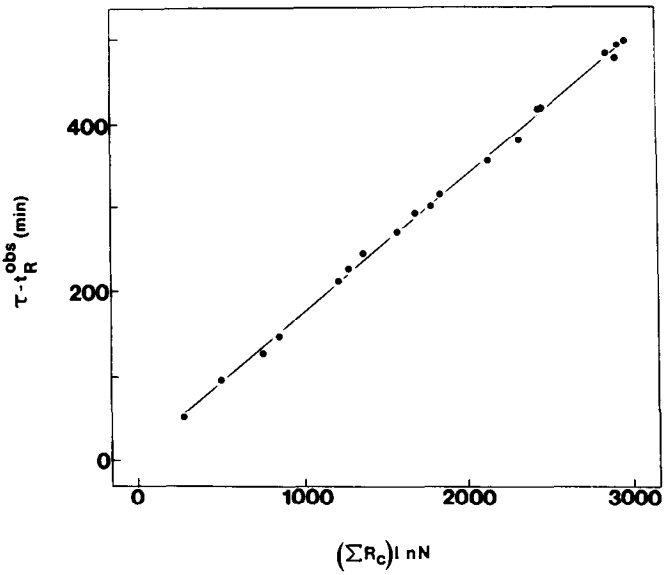

Fig. 3. Correlation of protein retention time with polypeptide chain length and hydrophobicity. Predicted minus observed protein retention time ($\tau - t_R^{obs}$) *versus* ΣR_c ln N, where ΣR_c is the sum of the retention coefficients of Guo *et al.*[9] for the amino acid residues in a protein, and N is the number of residues in the protein. Results are for an Aquapore RP-300 C_8 column (220 × 4.6 mm I.D.). Mobile phase conditions as described in Fig. 2. Absorbance at 210 nm. Results shown are for proteins 1–19 (Table I).

(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})

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

This intimate relationship between protein hydrophobicity and chain length and their combined effect on protein retention behaviour in RPC is clearly shown in Fig. 3. Plotting predicted (τ) minus observed (t_R^{obs}) protein retention time *versus* the product of protein hydrophobicity (expressed as ΣR_c, the sum of the coefficients of Guo *et al.*[9]) and the logarithm of the number of residues (ln N)[26], resulted in a single, straight-line plot (correlation, $r = 1.00$) for 19 proteins (proteins 1–19 in Table I), *i.e.*, a range of 30–284 residues in polypeptide chain length and 3500–32 000 dalton in molecular weight. Thus, the discrepancy between predicted and observed protein retention times is linearly related to ΣR_c ln N. Fig. 3 shows the plot for the C_8 column. Interestingly, the results on the C_4 and C_{18} columns (different *n*-alkyl chain lengths and ligand densities) gave very similar results. When ln molecular weight replaced ln N in the above relationship, the correlation of the resulting plot was not as high.

When the expression denoting polypeptide hydrophobicity (ΣR_c) was removed from the relationship producing the straight-line plot shown in Fig. 3, *i.e.*, plotting ($\tau - t_R^{obs}$) *versus* ln N, the profiles shown in Fig. 4 were obtained. Results are shown for both the four series of peptide polymers (Fig. 4A) and for proteins 1–19 (Table I) (Fig. 4B) on the C_8 column. When the peptide polymers were plotted in identical fashion to

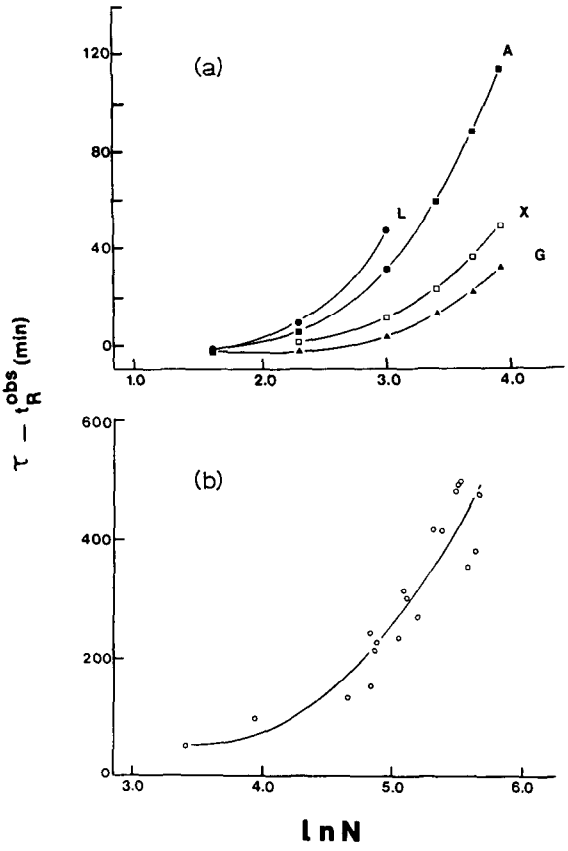

Fig. 4. Plot of predicted minus observed peptide (a) and protein (b) retention time ($\tau - t_R^{obs}$) *versus* ln N. Column, Aquapore RP-300 C_8 (220 × 4.6 mm I.D.). Mobile phase conditions as described in Fig. 2. Absorbance at 210 nm. Sequences of the peptide polymer series "G", "X", "A" and "L" are shown in the Experimental section. Results shown are for proteins 1–19 (Table I).

that shown for the proteins in Fig. 3, the result was a single, straight-line plot[26]. Removing the expression denoting hydrophobicity, ΣR_c, from the abscissa resulted in the profiles for the four peptide series becoming non-linear and diverging (Fig. 4A). The proteins showed a similar relationship (Fig. 4B), with a scattering of data points around a single, non-linear line. Fig. 4 again stresses the importance of taking the hydrophobicity of a peptide or protein into account when attempting to correlate retention time with polypeptide chain length.

Prediction of protein retention time in RPC

From Fig. 3

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

As described above

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

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 proteins, taking into account polypeptide chain length, t_R^{obs} in eqn. 3 becomes τ_c (predicted protein 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 polypeptide chain length[22].

Eqn. 4 was applied to retention time prediction of proteins 1–19 (Table I) and the results are shown in Fig. 5. The closed symbols denote observed *versus* predicted (τ) protein retention times when polypeptide chain length is not taken into account, *i.e.*, following application of the equation, $\tau = \Sigma R_c + t_s$ (see above); open symbols denote observed *versus* predicted (τ_c) retention times when chain length has been taken into account, *i.e.*, following application of eqn. 4. The solid line represents a perfect correlation between predicted and observed protein retention times. The contrast between the two sets of results is striking. It is clear that there is no correlation between predicted and observed protein retention times unless a polypeptide chain length

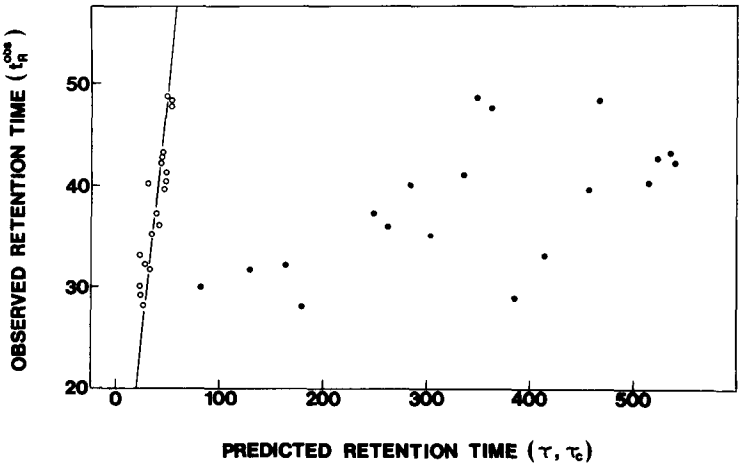

Fig. 5. Correlation of predicted and observed protein retention times in RPC. Results shown are for proteins 1–19 (Table I) on an Aquapore RP-300 C_8 column (220 × 4.6 mm I.D.). The predicted retention times either with (τ_c) or without (τ) taking polypeptide chain length into account were calculated as described in the text. Closed symbols denote observed *versus* predicted protein retention times without taking polypeptide chain length into account (τ); open symbols denote correlation when chain length has been taken into account (τ_c). The solid line represents perfect correlation between predicted and observed protein retention times. Mobile phase conditions as described in Fig. 2. Absorbance at 210 nm.

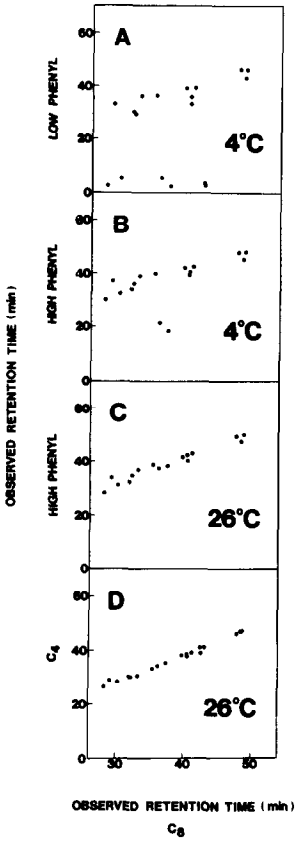

Fig. 6. Effect of hydrophobicity and ligand density of the stationary phase on the retention times of proteins in RPC. Columns, Bio-Gel TSK Phenyl-5PW (75 × 7.5 mm I.D.) (A); Bio-Gel TSK Phenyl RP+ (75 × 4.6 mm I.D.) (B and C), SynChropak RP-4 C_4 (250 × 4.1 mm I.D.) (D), and Aquapore RP-300 C_8 (220 × 4.6 mm I.D.); mobile phase, linear A–B gradient (1% B/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile; flow-rate, 1 ml/min; temperatures as shown; absorbance at 210 nm. Proteins 1–19 (Table I) were chromatographed on all four columns. Where the occasional protein produced a badly skewed peak, thereby preventing accurate retention time measurement, this data point was not used.

correction is applied. Once this correction was applied, the improvement in prediction of protein retention time was quite dramatic (an error of only ±5 min for the majority of the proteins). This improvement was especially gratifying considering the potential causes for deviations, including (1) incomplete denaturation of the protein on interaction with the hydrophobic sorbent, although this is unlikely for the proteins used in this study for the reasons given above[22], (2) formation of secondary structure on interaction with the reversed-phase sorbent[1,22], (3) preferential interaction of hydrophobic domains in the protein with the reversed-phase sorbent although this is unlikely for the proteins used in the study for the reasons given above[22] and (4) maintenance of tertiary structure in the protein by disulfide bonds, preventing accessibility of all regions of the amino acid sequence to interact with the reversed-phase sorbent.

Effect of hydrophobicity and ligand density of the bonded phase on the retention time behavour of proteins

The results shown in Figs. 3 and 5 justified the original assumption that the hydrophobic C_4, C_8 and C_{18} sorbents would effectively denature proteins. If this had not been the case, the correlation of the plot shown in these figures would not have been as good. To underline this point, the effect of less hydrophobic reversed-phase sorbents on the retention time behaviour of proteins was examined.

Proteins 1–19 in Table I were subjected to RPC on two polymer-based columns containing phenyl groups as hydrophobic ligands. One column (Bio-Gel TSK Phenyl-5PW) was developed for the related HPLC technique of hydrophobic interaction chromatography (HIC), with the phenyl groups only sparsely distributed across the support ("low phenyl"). The second column (Bio-Gel TSK Phenyl RP+) was based on the above HIC packing, but contained a much greater density of phenyl groups ("high phenyl") and was intended to function as a reversed-phase column. The proteins were chromatographed on both columns using a linear gradient (1% B/min) at a flow-rate of 1 ml/min, where eluent A was 0.1% aqueous TFA and eluent B was 0.1% TFA in acetonitrile. Fig. 6 compares the results of plotting observed protein retention time on the C_8 column at 26°C against the observed retention times on the low phenyl column at 4°C (Fig. 6A), the high phenyl column at 4°C (Fig. 6B) and 26°C (Fig. 6C), and the C_4 column at 26°C.

It would be expected that, if two hydrophobic sorbents denatured proteins to the same extent, and allowing for slight selectivity differences between the packings, then a plot of observed protein retention times on one column against observed retention times on the other would produce a linear plot with little scatter of data points. This linear relationship is clearly lacking for the low phenyl column at 4°C (Fig. 6A), where there is a wide scatter of data points. In fact, several proteins [insulin (chain B), ribonuclease A, avidin, lysozyme, concanavalin A, α-chymotrypsinogen A] were barely, if at all, retained by the column suggesting that these proteins are being separated in their native or only partially unfolded states. This result for lysozyme reflected similar results reported by Ingraham *et al.*[27]. In addition, these results showed that the acidic nature of the starting solvent, 0.1% aqueous TFA, does not necessarily denature all proteins. The low temperature (4°C) used for these runs also favoured maintenance of native protein conformation.

The effect of the more hydrophobic nature of the high phenyl column on the retention times of the proteins is clearly apparent in Fig. 6B, where the scatter of data points has lessened considerably. All proteins were now retained by the column, with just lysozyme and avidin still showing only partial unfolding. The temperature for the low phenyl and high phenyl runs shown in Fig. 6A and B, respectively, remained constant (4°C); thus, the increase in hydrophobicity of the column packing was solely responsible for the improvement in correlation of the data points shown in Fig. 6B. The results shown in Fig. 6A and B suggested that changes in reversed-phase sorbent hydrophobicity could be a useful aid to improving the resolution of protein mixtures in RPC.

The effect on protein retention time behaviour on the high phenyl column as the temperature was raised from 4°C (Fig. 6B) to 26°C is shown in Fig. 6C. Although in standard RPC the primary cause of denaturation is the hydrophobicity of the sorbent[27], an increase in temperature also has a denaturing effect; hence, the further

improvement in correlation between observed protein retention times on the high phenyl and C_8 columns seen in Fig. 6C. As noted by several researchers[27 29], such differences in protein retention times at different temperatures could also be used to improve the resolution of proteins with similar retention times.

The effect of increased sorbent hydrophobicity on the retention time behaviour of proteins at constant temperature (26°C) is again shown in Fig. 6D. The correlation of the observed protein retention times on the C_4 and C_8 columns has again improved, this time compared to the high phenyl column (Fig. 6C). In fact, the linearity of the plot shown in Fig. 6D indicated that the C_4 and C_8 sorbents denatured the proteins to the same extent. It should be noted that attempts to generate linear plots for the low and high phenyl columns similar to that shown in Fig. 3 for the C_8 column (and also representative of the C_4 and C_8 columns) met with little success, reflecting again the importance of efficient protein denaturation for satisfactory correlation of protein retention times with their overall hydrophobicity.

The polypeptide chain length effect demonstrated by the retention behaviour of peptide polymers of repeating units[26] or protein retention behaviour (this study) in RPC suggests that only a relatively small proportion of the amino acid residues are interacting with the stationary phase at any one time even when the polypeptide chain is fully unfolded. The excellent correlations shown in Figs. 3 and 5 (open circles) suggest that when the protein is completely unfolded, there is full expression of its overall hydrophobicity. Thus, there is not one specific portion of the unfolded polypeptide chain interacting with the stationary phase. Instead, all of the amino acid residues in an unfolded polypeptide chain can interact, in small portions, at any one time. Because of this polypeptide chain length effect, retention coefficients must be determined at a fixed peptide chain length and will vary depending on the length of the peptide used. If one is predicting retention times of polypeptides using coefficients derived at the same chain length as these polypeptides, then the predicted retention is simply equal to the sum of the coefficients. On the other hand, if you use coefficients derived from one chain length to predict retention times of peptides with different chain lengths, then a polypeptide chain length correction must be introduced.

CONCLUSIONS

The present study has shown that retention parameters derived from small peptides can be applied to the correlation of the overall hydrophobicity of proteins with their retention time behaviour during RPC. A clearer understanding of the three-dimensional structures of proteins on interacting with a hydrophobic sorbent will likely be required to improve further the accuracy of protein retention time prediction.

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 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 Decker, New York 1989, p. 1101.
2 J. L. Meek, *Proc. Natl. Acad. Sci. U.S.A.*, 77 (1980) 1632.
3 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.
4 S. J. Su, B. Grego, B. Niven and M. T. W. Hearn, *J. Liq. Chromatogr.*, 4 (1981) 1745.
5 K. J. Wilson, A. Honegger, R. P. Stötzel and G. J. Hughes, *Biochem. J.*, 199 (1981) 31.
6 C. A. Browne, H. P. J. Bennett and S. Solomon, *Anal. Biochem.*, 124 (1982) 201.
7 T. Sasagawa, T. Okuyama and D. C. Teller, *J. Chromatogr.*, 240 (1982) 329.
8 T. Sasagawa, L. H. Ericsson, D. C. Teller, K. Titani and K. A. Walsh, *J. Chromatogr.*, 307 (1984) 29.
9 D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 499.
10 D. Guo, C. T. Mant, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 519.
11 Y. Sakamoto, N. Kawakami and T. Sasagawa, *J. Chromatogr.*, 442 (1988) 69.
12 K. Jinno and E. Tanigawa, *Chromatographia*, 25 (1988) 613.
13 C. T. Mant and R. S. Hodges, in M. T. W. Hearn (Editor), *HPLC of Proteins, Peptides and Polynucleotides*, VCH Publishers, Weinheim, in press.
14 M. J. O'Hare and E. C. Nice, *J. Chromatogr.*, 171 (1979) 209.
15 E. C. Nice, M. W. Capp, N. Cooke and M. J. O'Hare, *J. Chromatogr.*, 218 (1981) 569.
16 K. J. Wilson, A. Honegger and G. J. Hughes, *Biochem. J.*, 199 (1981) 43.
17 C. T. Wehr, L. Correia and S. R. Abbott, *J. Chromatogr. Sci.*, 20 (1982) 114.
18 J. Heukeshoven and R. Dernick, *J. Chromatogr.*, 252 (1982) 241.
19 J. M. R. Parker and R. S. Hodges, *J. Protein Chem.*, 3 (1985) 465.
20 R. S. Hodges, R. J. Heaton, J. M. R. Parker, L. Molday and R. S. Molday, *J. Biol. Chem.*, 263 (1988) 11 768.
21 F. E. Regnier, *Methods Enzymol.*, 91 (1983) 137.
22 S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 317 (1984) 129.
23 S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Biol. Chem.*, 259 (1984) 13 253.
24 R. S. Hodges, P. D. Semchuk, A. K. Taneja, C. M. Kay, J. M. R. Parker and C. T. Mant, *Peptide Res.*, 1 (1988) 19.
25 M. Hermodson and W. C. Mahoney, *Methods Enzymol.*, 91 (1983) 352.
26 C. T. Mant, T. W. L. Burke, J. A. Black and R. S. Hodges, *J. Chromatogr.*, 458 (1988) 193.
27 R. I. Ingraham, S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 327 (1985) 77.
28 L. R. Snyder, *J. Chromatogr.*, 179 (1979) 167.
29 R. A. Barford, B. J. Sliwinski, A. C. Breyer and H. L. Rothbart, *J. Chromatogr.*, 235 (1982) 281.