Journal of Chromatography, 240 (1982) 329–340
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

CHROM. 14,754

PREDICTION OF PEPTIDE RETENTION TIMES IN REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY DURING LINEAR GRADIENT ELUTION

TATSURU SASAGAWA*
Department of Biochemistry, University of Washington, Seattle, WA 98195 (U.S.A.)
TSUNEO OKUYAMA
Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, 2-1-1 Fukazawa, Setagaya-ku, Tokyo (Japan)
and
DAVID C. TELLER
Department of Biochemistry, University of Washington, Seattle, WA 98195 (U.S.A.)
(First received December 1st, 1981; revised manuscript received January 20th, 1982)

SUMMARY

The retention behavior of 100 peptides was studied during high-performance liquid chromatography on a C_{18} column using aqueous trifluoroacetic acid as the mobile phase and acetonitrile as the mobile phase modifier in a linear gradient elution system. Retention times of the peptides were linearly related to the logarithm of the sum of Rekker's constants (R. F. Rekker, *The Hydrophobic Fragmental Constant*, Elsevier, Amsterdam, 1977, p. 301) for the constituent amino acid. Assuming this relationship, the best fit constants for this system were computed by non-linear multiple regression analysis. Using the new constants, it is possible to predict retention times for a wide variety of peptides at any slope of linear gradient, if the amino acid composition is known. It also enables accurate prediction of the retention time of peptides, whose amino acid composition is not known, after an analytical run with an alternate gradient.

INTRODUCTION

The introduction of a volatile eluent system[1,2] in reversed-phase high-performance liquid chromatography has greatly simplified the isolation of pure peptides from complex mixtures. However, selection of chromatographic conditions for purification of a given peptide is still found by time-consuming trial and error methods. *A priori* knowledge of the retention time of a given peptide would simplify the selection of chromatographic conditions. Calculations of peptide R_F values which are based on the amino acid composition were already reported by Knight[3] and Pardee[4]. Similar works using high-performance liquid chromatography were also reported[5–7]. O'Hare and Nice[6] noticed that the retention time of small peptides (< 15 residues) is linearly

0021-9673/82/0000–0000/$02.75

TABLE I

COMPARISON OF PREDICTED AND OBSERVED RETENTION TIMES

The retention times were predicted from eqn. 2, where A', C' were reported in Fig. 2. D'_j was reported in Table II (column 5, 6). The predicted retention times in parentheses were calculated from weighted fit parameters.

No.	*Peptide**	*Retention time (min)*		
		Observed	*Predicted*	
1	GGG	1.8	1.3	(1.9)
2	PG	2.5	1.3	(2.3)
3	ARKM*	2.8	1.0	(2.0)
4	TEEQ	5.0	4.7	(5.2)
5	GL-NH_2	5.3	5.5	(5.2)
6	Ac-AAA	6.1	6.5	(6.8)
7	MTAK	6.5	8.0	(6.4)
8	NLC*	6.6	7.0	(6.9)
9	MARKM*	7.5	10.4	(9.3)
10	MAR	7.8	5.3	(7.2)
11	YK	8.0	6.0	(7.5)
12	TPGSR	8.1	7.6	(7.3)
13	KYE	8.2	7.7	(9.4)
14	GY	8.5	9.9	(8.8)
15	TEAEMK	9.2	10.7	(10.3)
16	EY	9.6	9.3	(9.1)
17	HLK	9.8	12.9	(10.6)
18	FK	9.9	11.4	(10.8)
19	IRE	10.3	12.2	(11.7)
20	PL	10.3	13.8	(11.2)
21	IAE	10.9	13.1	(10.9)
22	GF	11.5	14.1	(11.9)
23	KMKDTDSEEE	11.5	9.9	(13.7)
24	AFR	12.0	13.5	(12.9)
25	DIAAK	12.0	11.7	(13.1)
26	QIAE	12.0	14.2	(11.0)
27	NIPC*	12.4	11.5	(10.6)
28	ASEDLK	13.0	11.9	(12.7)
29	EAFR	13.5	14.5	(14.3)

30	FDR	13.8	12.6	(12.8)
31	VFDKDGNGY	14.8	19.0	(19.7)
32	FKE	15.0	12.5	(12.5)
33	KVFGR	15.6	15.5	(15.6)
34	SLGQNPTEAE	15.8	17.4	(16.9)
35	GW	16.3	15.5	(15.3)
36	MIRE	16.5	16.9	(16.3)
37	SHPETLEK	16.7	19.2	(17.5)
38	HGLDNYR	17.0	18.0	(17.6)
39	WY-NH_2	17.1	16.6	(17.4)
40	Ac-ADQL	18.2	16.6	(17.8)
41	LFK	18.2	19.4	(16.2)
42	IAEFK	19.5	20.7	(19.6)
43	ADIDGDGQVNYEE	19.8	21.1	(22.4)
44	VFDKDGNGYI	20.2	19.0	(19.7)
45	ISAAELR	20.3	21.0	(20.1)
46	FESNFNTQATNR	20.3	19.0	(19.4)
47	ELGTVMR	21.2	21.0	(19.2)
48	GHHEAELK	21.3	18.9	(17.2)
49	WWC*NDGR	21.4	22.1	(24.9)
50	LQDMINE	22.0	21.0	(20.7)
51	FVQMMTAK	22.5	23.6	(22.8)
52	WWC*	23.5	23.4	(24.5)
53	QIAEFK	23.8	21.3	(20.9)
54	Ac-ADQLTEEQIAE	24.0	25.7	(26.0)
55	RSLGQNPTEAELQDM*	24.0	25.2	(25.1)
56	MIREADIDGDGQVNYEE	24.8	27.6	(29.7)
57	FLTMMAR	25.1	25.3	(24.9)
58	VDADGNGTIDFPE	25.3	24.5	(23.8)
59	HVMTNLGEK*LTDEEVDEM*	25.7	28.1	(27.7)
60	LGTVMRSLGQNPTEAE	25.8	27.5	(27.4)
61	SALLSSDITASVNC*	26.0	26.4	(25.9)
62	NTDGSTDYGILQINSR	26.9	27.6	(27.9)
63	VEADVAGHGQDILIR	26.9	29.0	(28.7)
64	FLTMMARKMKDTDSEEE	27.0	29.1	(30.7)
65	LRHVMTNLGEK*LTDE	27.0	24.7	(24.8)
66	VTVPLVSDAEC**R	27.3	27.0	(26.1)
67	VFDKDGNGYISAAELR	27.5	28.5	(30.1)
68	AFRVFDKDGNGYISAAE	28.6	29.3	(31.2)

(Continued on p. 332)

TABLE I *(continued)*

No.	*Peptide**	*Retention time (min)*		
		Observed	*Predicted*	
69	LRHVMTNLGEK*LTDEEVDE	28.6	29.6	(30.8)
70	VFDKDGNGYISAAEL	29.0	28.5	(29.5)
71	GYSLGNWVC**	29.1	29.1	(29.6)
72	IREADIDGDGQVNYEEFVQM*	29.2	28.8	(30.5)
73	Ac-ADQLTEEQIAEFK	29.2	27.7	(28.4)
74	EAFSLFDKDGDGTITTK	30.0	31.4	(31.5)
75	ALELFR	30.2	25.4	(24.6)
76	AFSLFDKDGDGTITTKE	30.4	31.4	(31.5)
77	HVMTNLGEK*LTDEEVDEMIR	33.5	31.0	(32.5)
78	NKALELFRKDIAAKYKELGYQG	34.2	34.0	(37.3)
79	PGYPGVYTEVSYHVDWIK	34.8	34.5	(36.7)
80	DDYGADEIFDSMIC**AGVPEGGK	35.9	34.8	(37.1)
81	EADIDGDGQVNYEEFVQMMTAK	37.2	31.2	(33.5)
82	INEVDADGNGTIDFPEFLTM*	37.5	33.9	(34.1)
83	KDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEK*LTDEEVDEM*	37.8	40.6	(45.8)
84	HLHENFDYDLLDNDISLLK	38.5	38.1	(40.8)
85	ASSTNLKDILADLIPKEQARIKTFRQQHGNTVVGQITVDM*	39.0	40.3	(43.7)
86	HGVTVLTALGAILK	40.5	32.1	(30.6)

87	Ac-ADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMR	42.1	41.3	(44.6)
88	SQLSAAITALNSESNFARAYAEGIHRTKYWELIYEDC**M*	42.3	42.9	(47.5)
89	Ac-ADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVM*	42.8	41.4	(45.4)
90	SLGQNPTEAELQDMINEVDADGNGTIDFPEFLTM	44.0	39.0	(42,3)
91	YLEFISEAIIHVLHSR	45.0	36.5	(38.2)
92	MARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEK*LTDEEVDEMIREADIDGDGQVN YEEFVQMMTAK	45 5	46.0	(53.8)
93	NGLAGPLHGLANQEVLVWLTQLQKEVGKDVSDEKLRDYIWNTLNSGRVVPGYGHAVLRKTDPRYTC** QREFALKHLPHDPM*	45.5	50.3	(53.8)
94	VLSEGEWQLVLHVWAKVEADVAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAEM*	45.6	47.2	(54.1)
95	SLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMAR	45.8	39.8	(43.6)
96	KMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEK*LTDEEVDEMIREADIDGDGQVNYEE FVQMMTAK	46.2	45.5	(53.1)
97	VDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEK*LTDEE VDEMIREADIDGDGQVNYEEFVQMMTAK	48.3	50.1	(58.8)
98	FKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKD TDSEEEIREAFRVFDKDGNGYISAAE	48.9	50.7	(59.0)
99	KASEDLKKHGVTVLTALGAILKKKGHHEAELKPLAQSHATKHKIPIKYLEFISEAIIHVLHSRHPGN FGADAQGAM*	50.0	49.6	(56.2)
100	Ac-ADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTID FPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEK*LTDEEVDEMIREADI DGDGQVNYEEFVQMMTAK	51.0	56.0	(66.7)

* M* = Homoserine or its lactone; C* = aminoethylcysteine; C** = carboxymethylcysteine; K* = trimethyllysine.

related to the sum of Rekker's constants[8] (which are based on the partition coefficients of free amino acids between water and octanol) for the constituent amino acids in a linear gradient elution system. Meek[7] reported a similar relation for small peptides, and extended the method to numerical analysis of the retention constants of amino acids. However, using their methods, we observed many discrepancies in the predicted retention order in a volatile eluent system and have reinvestigated procedures to predict those retention times.

EXPERIMENTAL

Materials

Almost all peptides were obtained by either enzymatic or cyanogen bromide degradation of calmodulin[9,10], lysozyme, citrate synthase[11], myoglobin, or crayfish trypsin[12]. Other oligopeptides were purchased from the indicated sources: Pro–Gly, Pro–Leu, Gly–Leu–NH_2, Gly–Tyr, Glu–Tyr, Lys–Tyr–Glu(Cyclo Chemical); Gly–Gly–Gly (Hoffmann-La Roche); Gly–Phe (California Biochemical Research); Gly–Trp (Mann Research Labs.); Trp–Tyr–NH_2 (Vega Biochemicals). The μBondapak C_{18} column (30 × 0.4 cm) was a product of Waters Assoc. Acetonitrile was obtained from Burdick & Jackson. Trifluoroacetic acid (Pierce) was used after distillation.

Methods

Retention times were measured on a μBondapak C_{18} column using a Varian Model 5000 liquid chromatograph. The mobile phase was 0.1% trifluoroacetic acid (pH 2) and the mobile phase modifier was acetonitrile containing 0.07% trifluoroacetic acid. The flow-rate was 2.0 ml/min. The concentration of the mobile phase modifier was increased linearly from 0 to 60% over 60 min (1%/min). The elution was monitored by absorption at 216 nm, and retention times were measured from the time at injection to that at the center of the eluting peak. Regression analyses were done either by a VAX computer or by a PDP-12 computer with a floating point processor. Matrix inversion was performed using double precision arithmetic.

RESULTS

Retention behavior of peptides

Retention times of 100 peptides tested are listed in Table I. The observed retention times (t_{Ri}) were plotted against $\sum_j D_j n_{ij}$, where D_j is the retention constant of amino acid j (Table II), and n_{ij} is number of residues of amino acid j in peptide i. Using Meek's constants as D_j the plot (Fig. 1a) gave a poor correlation (correlation coefficient 0.78). However, using Rekker's constants (where values for hydrophilic amino acids were slightly modified, Table II) as D_j the plotted data indicated an exponential relationship (Fig. 1b). Therefore, the data were fitted to a function of the form:

$$t_{Ri} = A \ln (1 + B \sum_j D_j n_{ij}) + C \qquad (1)$$

TABLE II

RETENTION CONSTANTS (D_j) OF AMINO ACIDS

The numbers in parentheses represent the number of amino acids used for each calculation

Amino acid	*Rekker's*	*Rekker's (modified)*	*Meek's*	*Present study (D'_j)*	
				Non-weighted	*Weighted*
Tryptophan	2.31	2.31	18.1(7)	35.8(12)	2.34(12)
Phenylalanine	2.24	2.24	13.9(18)	31.4(86)	1.71(86)
Isoleucine	1.99	1.99	11.8(4)	27.4(95)	1.38(95)
Leucine	1.99	1.99	10.0(13)	26.4(129)	1.34(129)
Tyrosine	1.70	1.70	8.2(16)	21.0(43)	1.23(43)
Methionine	1.08	1.08	7.1(11)	14.5(64)	0.85(64)
Proline	1.01	1.01	8.0(13)	7.9(33)	0.48(33)
Valine	1.46	1.46	3.3(6)	7.4(89)	0.38(89)
Threonine	−0.26	0.10	1.5(9)	7.4(111)	0.12(111)
Histidine	−0.23	−0.10	0.8(6)	8.8(38)	0.34(38)
Alanine	0.53	0.53	−0.1(8)	2.4(139)	0.13(139)
Glutamine	−1.09	0.20	−2.5(5)	3.2(59)	0.36(59)
Glutamic acid	−0.07	0.20	−7.5(4)	2.7(198)	0.27(198)
Glycine	0.00	0.10	−0.5(20)	4.0(134)	0.22(134)
Serine	−0.56	0.10	−3.7(11)	1.1(62)	0.18(62)
Arginine	–	−0.10	−4.5(10)	0.0(73)	0.26(73)
Aspartic acid	−0.02	0.10	−2.8(7)	− 0.1(165)	0.10(165)
Asparagine	−1.05	0.10	−1.6(6)	−11.3(71)	−0.45(71)
Lysine	0.52	−0.10	−3.2(9)	− 3.1(98)	0.05(98)
Carboxymethylcysteine	–	0.10	–	32.5(5)	1.57(5)
Homoserine	–	0.10	–	12.3(13)	0.23(13)
Aminoethylcysteine	–	−0.10	–	4.3(5)	0.31(5)
Trimethyllysine	–	−0.10	–	−38.1(9)	−1.38(9)
Acetyl-	–	0.00	3.9(1)	12.4(6)	0.81(6)
Amide-	–	0.00	5.0(8)	−13.2(2)	−0.56(2)

by using the method of least squares, where the modified Rekker's constants were used for D_j. The best fit for B was evaluated by plotting the correlation coefficient (of t_{Ri} and $\ln [1 + B \sum_j D_j n_{ij}]$) against B. The maximum correlation coefficient (0.97) was estimated when B was 1.1. The intercept (C) and the slope (A) were −5.6 and 13.4 respectively.

Calculation of retention constants

Assuming an equation of the form:

$$t_{Ri} = A' \ln (1 + \sum_j D'_j n_{ij}) + C' \qquad (2)$$

new retention constants (D'_j) and constants A', C' were computed from the observed retention times by non-linear multiple regression analysis[13]. As initial values for D'_j, modified Rekker's constants were multiplied by 1.1; for A' and C', 13.4 and −5.6 were used respectively. Several variations of curve fitting were applied to the data. By

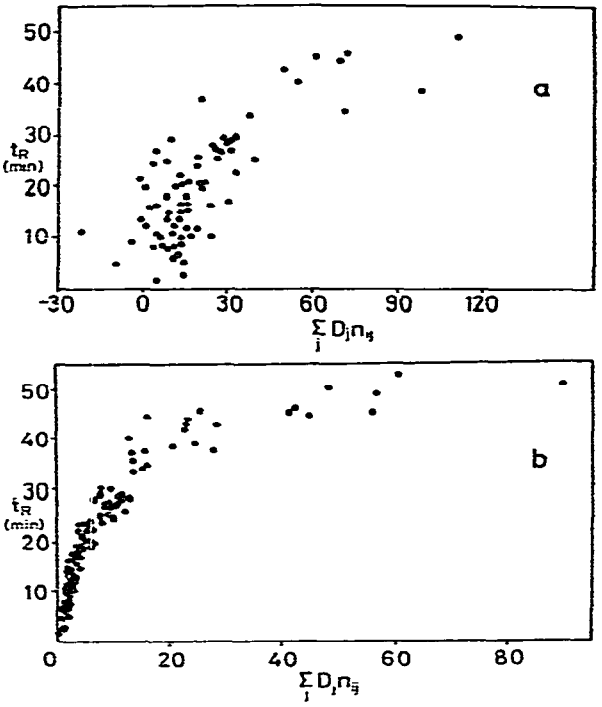

Fig. 1. Dependence of retention time (t_R) on amino acid composition. The observed retention times (Table I) were plotted against $\sum_j D_j n_{ij}$, were D_j is the retention constant of amino acid j (Table II), and n_{ij} is number of amino acid residues j in peptide i. (a) Meek's constants were used for values of D_j. Peptides which contain amino acids whose retention constants were not reported were excluded from the plot. The linear correlation coefficient was 0.78. (b) Modified Rekker's constants were used for values of D_j.

straightforward least squares analysis without weights, the parameters D'_j (Table II, column 5) obtained for the amino acids were rather different from those of Meek[7] or Rekker[8]. It was also found that the correlation coefficient of observed and predicted retention times was rather insensitive to the retention constants D'_j. For this reason, and because we wished to have a more uniform percentage deviation of observed and predicted retention times, weighted least squares were performed. The weights used were $1/N_i^2$ where N_i is the number of amino acids in the peptide. The results of this calculation are quite satisfactory. The retention constants (Table II, column 6) were of similar magnitude to those reported by other workers (Table II, columns 2, 3). The unweighted correlation of observed and predicted retention times is 0.984 while the weighted correlation is 0.982 (Fig. 2). The mean percent deviation of retention times is only 9 percent. The predicted retention times for each peptide in this study from both unweighted and weighted curve fitting methods are reported in Table I for comparison with observed values. Fig. 2 shows the plot of t_{Ri} as a function of ln $(1 + \sum_j D'_j n_{ij})$ for both unweighted (a) and weighted (b) least squares parameters. The graphs are both linear with correlation coefficients of 0.98.

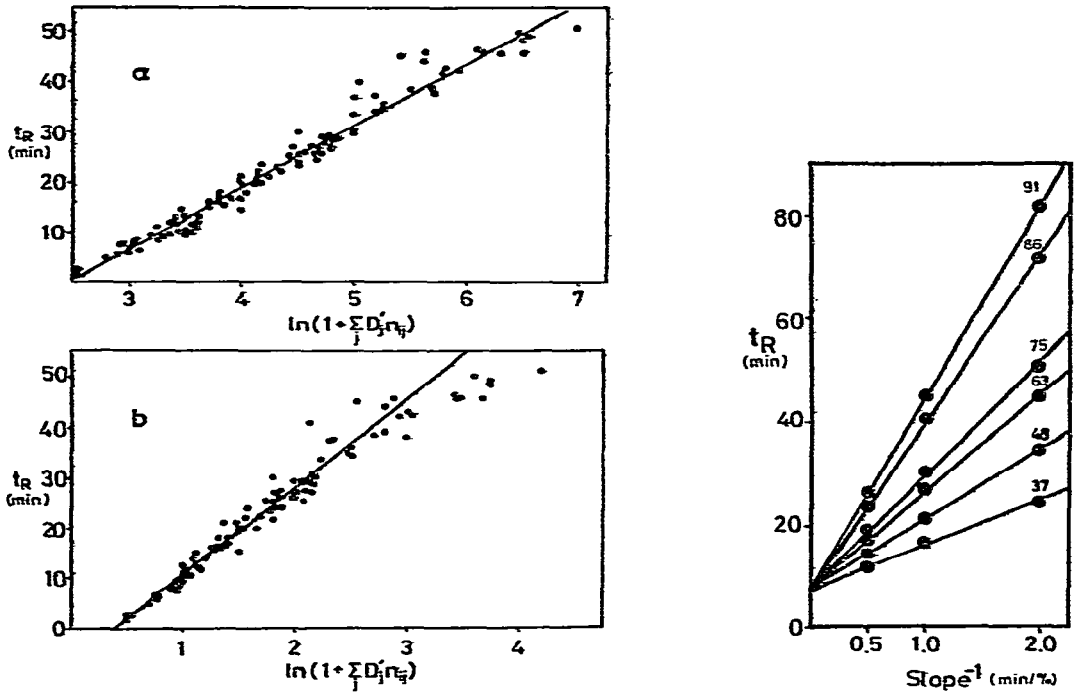

Fig. 2. Relationship between retention time (t_R) and ln $(1 + \sum_j D'_j n_{ij})$. The observed retention times (t_{Ri}) were plotted against ln $(1 + \sum_j D'_j n_{ij})$, where D'_j is the calculated retention constants either by unweighted or weighted curve fitting. (a) Unweighted fit retention constants were used for D'_j. The intercept and the slope of the straight line was −30.3 and 12.4, respectively. The correlation coefficient was 0.984. The mean percent deviation of retention time was 9.9%. (b) Weighted fit retention constants were used for D'_j. The intercept and slope of the straight line were −7.04 and 13.6, respectively. The weighted correlation coefficient was 0.981. The mean percent deviation of retention time was 9.2%.

Fig. 3. Plot of retention time (t_R) against the inverse of the slope (m) of the gradient. Retention times of six peptides (the number of the peptide corresponds to that in Table I) in different gradients were plotted against the inverse of the slope (m) of the gradient.

Effect of slope of gradient on retention time

Retention times of six different peptides were measured in the same elution system but using three different gradients, between 0.5 and 2.0%/min. The retention times (t_{Ri}) were plotted against the inverse of the slope (m) of the gradient (Fig. 3). Late eluting peptides gave straight lines which have a common intercept (0,7). Therefore, it should be possible to estimate the variation of t_{Ri} with m by equation (3):

$$t_{Ri} = F_i(1/m) + 7 \tag{3}$$

where F_i is the slope of the straight line and is characteristic for a given peptide. By combination of equations (2) and (3) one may express the retention time of a peptide in any linear gradient (0.5 to 2%/min) as

$$t_{Ri} = (1/m) \times [A' \ln (1 + \sum_j D'_j n_{ij}) + C' - 7] + 7 \tag{4}$$

TABLE III

COMPARISON OF PREDICTED AND OBSERVED RETENTION TIMES

Numbers correspond to those in Table I. The retention times were predicted by using non-weighted parameters

No.	*Peptide*	*Retention time (min)*			
		Slope (2%/min)		*Slope (1%/min)*	
		Predicted	*Observed*	*Predicted*	*Observed*
94	Myoglobin residues 1–55	26.4	26.0	45.8	45.6
99	Myoglobin residues 56–131	27.5	28.3	47.9	50.0
78	Myoglobin residues 132–153	20.1	20.6	33.2	34.2

The predictive utility of equation (4) was tested on cyanogen bromide peptides of myoglobin. The results are presented in Table III.

DISCUSSION

The object of this study has been the development of a method capable of predicting retention times for a wide range of peptides in a linear gradient elution system. Peptides of known sequence have been examined using aqueous trifluoroacetic acid as mobile phase. Both Meek[7] and O'Hare and Nice[6] reported a linear relation between t_R values of small peptides and the sum of their amino acid retention constants. However, the present study clearly shows an exponential relationship (Fig. 1b). Apparent discrepancy between the previous methods and ours may simply be due to a difference in the range of peptides investigated. In his system, Meek only tested small peptides (<29 residues long), and thus obtained data which approximated a linear relation. Our data also look linear during the first 20 min of elution. Non-linear eqn. 2 accurately describes the dependence of retention time on the amino acid composition for a wider range of peptide size. An example is given in Table IV in order to demonstrate the applicability of the above method. The reason underlying the exponential relation remains to be explained. Assuming the empirical eqn. 2, we have computed a set of retention constants for amino acids both with and without attaching weights on the observations. As was expected, both aromatic and aliphatic amino acids make large positive contribution to retention, and the relative degree of the contribution is almost the same as Meek's constants. However, some differences were found between our constants and Meek's constants (correlation coefficient between our constants and Meek's constants[7] was 0.816) which probably explain the poor correlation (Fig. 1a) of his prediction and our data. Neutral and acidic amino acids in our system showed a small positive contribution except for aspartic acid and asparagine. Meek, however, assigned negative values for almost all of these amino acids, and an especially large negative value for glutamic acid. The positive contribution of glutamic acid or glutamine in our system was clearly illustrated by the retention order in the following sets of peptides: YK < KYE, FK < FKE, AFR < EAFR, IAE < QIAE, IAEFK < QIAEFK. The apparent discrepancies between our constants and Meeks' may simply be due to a difference in the eluent system. We used trifluoroacetic acid (pH 2) in the mobile

TABLE IV

COMPARISON OF DATA FROM LITERATURE WITH PREDICTED RETENTION TIMES

Data were from O'Hare and Nice[6]. Retention times were predicted using equation 2 and the non-weighted retention constants for amino acids of Table II. Due to the different solvent system and the presence of pyroglutamic acid and cystine, it was necessary to fit four new constants. The slope and intercept of the straight line were 8.30 and −14.73, respectively. The retention constants for pyroglutamic acid and cystine were found to be −12.2 and −27.3, respectively. The correlation coefficient was 0.81 and the mean percent deviation of retention time was 12.0%.

Peptides	*Number of residues*	*Retention time (min)*	
		Predicted	*Observed*
Arginine vasotosin	9	12.8	12.0
Lysine vasopressin	9	13.1	13.0
Arginine vasopressin	9	13.9	14.0
ACTH 5-10	6	22.5	17.0
Diphenylalanine	2	20.2	18.0
ACTH 1-18	18	25.8	18.5
Met-enkephalin	5	21.7	19.0
Oxytosin	9	18.6	19.5
ACTH 4-10	7	23.8	20.5
ACTH 1-24	24	28.8	21.5
α-Endorphin	16	27.0	22.0
Leu-enkephalin	5	22.9	22.0
Insulin A(bovine)	21	22.2	22.0
Angiotensin II	8	24.3	23.0
Neurotensin	13	25.1	24.5
α-Melanotropin	13	26.4	26.0
Bombesin	14	25.0	26.0
RNAase	124	36.7	27.5
Triphenylalanine	3	23.5	28.0
Gastrin I	17	29.1	28.5
Substance P	11	26.0	29.0
Substance P 4-11	8	25.7	30.0
ACTH 1-39 (human)	39	32.8	30.5
ACTH 18-39	22	27.8	30.5
ACTH 34-39	6	24.2	31.0
Somatostatin	14	24.6	32.0
Insulin (bovine)	51	34.7	32.0
ACTH 1-39 (porcine)	39	33.5	33.0
Insulin B (bovine)	30	32.6	33.5
β-Endorphin (ovine)	31	31.3	34.0
β-Lipotropin (human)	91	38.4	34.5
Calcitonin (human)	32	31.7	34.5
Cytochrom C	104	40.1	35.0
Glucagon	29	31.7	36.0
Tetraphenylalanine	4	25.9	36.5
Calcitonin (salmon)	32	29.4	37.0
Lysozyme	129	40.4	37.5
Myoglobin	153	45.7	45.0
Melittin	25	32.3	46.0

phase, and Meek used perchloric acid (pH 2.1). However, it also should be noted that he calculated his 26 constants using only 25 peptides, which may be too small a sample for accurate estimation of the constants. During course of present study, Meek[14] revised his amino acid retention constants. Correlation coefficients between our constants and Meek's new constants are 0.844 (phosphate system), 0.821 (perchlorate system). The contribution of basic amino acids was small except for trimethyllysine. It is not so simple to explain the contribution of basic amino acids, because they will also interact with residual silanol.

For large peptides, their conformations, besides their amino acid composition, may also contribute to their retention times. This information was incorporated in the analysis by weighting each observation. The weighted fit constants predict retention times more accurately for small peptides. Probably the retention constants obtained by using weighted analysis reflect more realistic constants.

The effect of the slope of the gradient on the retention time is clearly illustrated in Fig. 3, where, for each peptide tested, a steeper gradient eluted the peptide earlier. Thus retention times were linearly related to the inverse of the slope. The intercept (7 min) corresponds closely to the delay time for the gradient front to reach the sample injection point.

These studies were carried out with a single C_{18} column to establish the empirical relationship between amino acid composition and elution time. Extension of these studies to other C_{18} columns and different elution system (using phosphate at both neutral and acidic pH values) have been recently completed, and will be reported elsewhere.

ACKNOWLEDGEMENTS

We are grateful to Dr. K. A. Walsh for his continued interest and support, and to Mr. P. A. Karplus for his discussion throughout the work. The authors also wish to thank Dr. K. Titani for donating peptides and Mr. J. Yagi for his help in data analysis. This work was in part supported by grants from the National Institutes of Health (GM 15731 and GM 13401).

REFERENCES

1 C. E. Dunlap III, S. Gentleman and L. I. Lowney, *J. Chromatogr.*, 160 (1978) 191.
2 W. C. Mahoney and M. A. Hermodson, *J. Biol. Chem.*, 255 (1980) 11199.
3 C. A. Knight, *J. Biol. Chem.*, 190 (1951) 753.
4 A. B. Pardee, *J. Biol. Chem.*, 190 (1951) 757.
5 T. Sasagawa and T. Okuyama, *Bunseki Kagaku (Jap. Anal.)*, 27 (1978) 621.
6 M. J. O'Hare and E. C. Nice, *J. Chromatogr.*, 171 (1979) 209.
7 J. L. Meek, *Proc. Nat. Acad. Sci. U.S.*, 77 (1980) 1632.
8 R. F. Rekker, *The Hydrophobic Fragmental Constant*, Elsevier, Amsterdam, 1977, p. 301.
9 W. E. Schreiber, T. Sasagawa, K. Titani, R. D. Wade, D. Malencik and E. H. Fischer, *Biochemistry*, 20 (1981) 5235.
10 T. Sasagawa, L. H. Ericsson, K. A. Walsh, W. E. Schreiber, E. H. Fischer and K. Titani, *Biochemistry*, in press.
11 D. P. Bloxham, D. C. Parmelee, S. Kumar, R. D. Wade, L. H. Ericsson, H. Neurath, K. A. Walsh and K. Titani, *Proc. Nat. Acad. Sci. U.S.*, 78 (1981) 5381.
12 K. Titani, T. Sasagawa, R. G. Woodbury, L. H. Ericsson, H. Dörsam and R. Zwilling, in preparation.
13 R. P. Bevington, *Data Reduction and Error Analysis for the Physical Science*, McGraw-Hill, New York, 1969, p. 232.
14 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.