ELSEVIER

Journal of Chromatography A, 808 (1998) 105–112

JOURNAL OF CHROMATOGRAPHY A

Calculation of peptide retention coefficients in normal-phase liquid chromatography

Tatsunari Yoshida

Scientific Instrument Division, Tosoh Corporation, Tokyo Research Center, 2743-1, Hayakawa, Ayase-shi, Kanagawa-ken 252-1123, Japan

Received 9 October 1997; received in revised form 27 January 1998; accepted 29 January 1998

Abstract

The retention of 121 peptides was studied on a TSK Amide-80 column using solutions containing 0.1% trifluoroacetic acid and an increasing linear gradient of water in acetonitrile. The contribution of each residue to retention was calculated by linear multiple regression analysis. This paper described the contribution values 'hydrophilicity retention coefficients'. The result is an index of hydrophilicity retention coefficients for normal-phase liquid chromatography, analogous to the hydrophobicity indices calculated for the reversed-phase liquid chromatography. The order of residues in the index of one mode was substantially the inverse of the others'. Using the new hydrophilicity retention coefficients, retention times could be predicted for peptides of known amino acid content and sequence.

Keywords: Retention coefficients; Structure–retention relationships; Peptides

1. Introduction

In a previous paper [1], peptide separation in normal-phase liquid chromatography (NPLC) was reported using a TSK gel Amide-80, with carbamoyl groups bonded to a silica-gel matrix. An increasing gradient of water (in 0.1% trifluoroacetic acid) was used [2–7], the inverse of reversed-phase liquid chromatography (RPLC). Under these conditions, peptides are retained through a normal-phase mechanism. Hydrophilic peptides are often not retained on an octadecyl silica (ODS) column in the RPLC mode, but separation was possible on the TSK gel Amide-80 column in the present NPLC. Separation selectivities in normal-phase and reversed-phase methods differed significantly.

RPLC on an ODS column is commonly used for the separation and purification of peptide mixtures. Chromatographic conditions for separation are usually determined by trial and error methods and thus attention has been directed to the prediction of peptide retention times [8–22]. Hydrophobicity retention coefficients [8,10,12–18] for predicting peptide retention times during RPLC have been determined based on the fact that the contribution of each residue to peptide retention is additive and retention time is linearly related to the sum of the contributions. The hydrophobicity retention coefficients [11] were found closely correlated to Rekker's constants [22] (based on partition coefficients of free amino acids for water and octanol) and useful for clarifying the retention mechanism of RPLC [8,10,11,13].

The objective of this study is to calculate an analogous set of hydrophilicity retention coefficients for the present NPLC.

0021-9673/98/$19.00
PII: S0021-9673(98)00092-2

2. Experimental

2.1. Materials

HPLC-grade acetonitrile (ACN) was obtained from Nacalai Tesque, (Kyoto, Japan), and trifluoroacetic acid (TFA) and formic acid from Wako Pure Chemical Industries, (Osaka, Japan). Milli-Q (Japan Millipore Ltd, Tokyo, Japan) was used for water purification. Most of the peptides were purchased from Sigma (St. Louis, MO) and Peptide Institute, (Osaka, Japan) and the others were obtained by enzymatic or cyanogen bromide degradation of protein (myoglobin and concanavalin A) or synthesis by a peptide synthesizer (Perkin Elmer Applied Bio systems Division). The TSK gel Amide-80 column (25×0.46 cm I.D.) was from Tosoh (Tokyo, Japan).

2.2. Apparatus

The HPLC system was a Tosoh liquid chromatograph equipped with a SC-8020 micro-computer, CCPM-II pump, UV-8020 detector, AS-8020 auto sample injector and CO-8020 column oven. Regression analysis was conducted by NEC a PC-9801 computer with a floating point processor.

2.3. Methods (chromatographic measurements)

Eluent A (the initial eluent) was 0.1% TFA in ACN–water (90:10) and eluent B, 0.1% TFA in ACN–water (55:45). The peptides were dissolved in 10 μl ACN–water–formic acid (5:45:50), followed by the addition of 40 μl initial eluent and separation by linear gradient from eluent A to eluent B over 58.33 min (0.6% water/min). The flow-rate was 1.0 ml/min. Elution was monitored by UV absorption at 215 nm. The temperature in the column oven was 40°C.

3. Results

The retention times of 121 peptides are listed in Table 1. Matrix inversion was performed using double precision arithmetic. The data were fitted to the linear relationship,

$$T_{\text{obs}} = \sum_j nij \times Dj + b$$

where T_{obs} = observed retention times; nij = number of amino acid residues j in peptide i; Dj = hydrophilicity retention coefficient j; b = retention coefficients of terminal amino and carboxyl groups.

The hydrophilicity retention coefficients (Dj) are listed in Table 2, column 1. The predicted retention time of peptides are listed in Table 1 for comparison with observed data. In Fig. 1, the observed retention time is plotted against that predicted. The correlation of observed and predicted retention times was 0.94.

4. Discussion

Chromatograms for separating peptides by the present NPLC, together with those of the RPLC, are shown in Fig. 2. Separation selectivities for the present NPLC and RPLC differed significantly. But, the elution order of peptides in the present NPLC was not only a simple reversal of that in RPLC.

This study was conducted to determine quantitative hydrophilicity retention coefficients of peptide residues in the NPLC and permit the prediction of retention times in a linear gradient system. Peptides of known sequence were examined by the present NPLC using acetonitrile–water mixed solution containing 0.1% trifluoroacetic acid. In RPLC, Meek [8] reported that retention times of phenylalanine oligomers are linearly related to the number of phenylalanine and the slope equals the retention added per residue and peptide bond, and intercept represents the contribution of the terminal amino and carboxyl groups. As with RPLC, assuming peptide retention to depend mainly on the sum of contributions to retention of the peptide's amino acid residues, the hydrophilicity retention coefficients (Table 2) were determined based on the retention times of 121 peptides by computer regression analysis.

As shown from Table 2, column 1, the basic residues, Arg (3.90), His (3.44) and Lys (2.77) made significant contributions to retention in the present NPLC. For the amide residues, Asn (3.25) and Gln (2.35), acidic residues, Asp (2.24) and Glu (1.58), and hydroxyl residues, Thr (1.73) and Ser (2.53),

Table 1
Comparison of predicted and observed retention times

No.	Observed	Predicted	Number of sequence residue
1	5.01	9.85	2 FY
2	8.06	13.02	2 GA
3	12.90	14.99	2 GD
4	11.64	14.32	2 GE
5	5.79	9.80	2 GF
6	9.55	12.58	2 GG
7	16.81	16.19	2 GH
8	5.61	11.34	2 GI
9	16.28	15.51	2 GK
10	5.53	10.43	2 GL
11	5.69	12.60	2 GM
12	16.00	15.99	2 GN
13	9.54	13.51	2 GP
14	15.18	15.09	2 GQ
15	13.17	15.27	2 GS
16	11.33	14.47	2 GT
17	6.22	10.55	2 GV
18	6.70	10.94	2 GW
19	8.35	12.63	2 GY
20	3.93	7.77	2 VF
21	14.47	18.24	3 EHP
22	6.70	9.35	3 EVF
23	7.76	10.27	3 GGL
24	8.63	10.38	3 GGV
25	15.15	21.32	3 GPR
26	6.12	10.74	3 PLG
27	6.52	8.41	3 VYV
28	18.76	17.13	4 AGSE
29	23.12	19.60	4 EAEN
30	6.68	6.69	4 FGGF
31	9.01	10.32	4 FMRF
32	18.05	16.37	4 GGYR
33	21.33	20.86	4 GHRP
34	23.33	20.97	4 GRGD
35	22.83	22.08	4 TKPR
36	14.85	14.37	4 VGDE
37	15.18	14.66	4 VGSE
38	7.79	9.80	4 WMDF
39	26.06	24.60	5 DSDPR
40	8.48	9.41	5 FLEEI
41	8.56	8.50	5 FLEEL
42	25.42	20.69	5 VTYHS
43	18.48	15.73	5 KEEAE
44	19.40	18.31	5 VEEAE
45	6.94	7.22	5 YGGFL
46	7.99	9.39	5 YGGFM
47	18.30	17.20	5 YIGSR
48	7.41	11.23	5 YPFPG
49	14.73	17.90	6 HIAPAW
50	10.65	14.48	6 HIKWPA
51	15.98	16.97	6 HDWPTI
52	9.15	9.24	6 VGVAPG

(Cont.)

Table 1. Continued

No.	Observed	Predicted	Number of sequence residue
53	19.10	18.67	7 AVPYPQR
54	25.70	23.66	7 DAEFHDR
55	10.93	10.27	7 GVYVHPV
56	19.60	19.30	7 LRRASVA
57	13.63	11.56	7 NFTYGGF
58	14.31	14.38	7 RVYIHPF
59	15.23	13.58	7 RVYVHPF
60	19.26	18.10	7 SQNYPIV
61	10.13	6.73	7 VFFAVDE
62	6.96	10.60	7 YPFPGPI
63	27.38	25.79	8 ASTTTNYT
64	17.74	16.62	8 DRVYIHPF
65	12.58	11.64	8 DYMGWNDF
66	14.44	11.85	8 EGVYVHPV
67	18.83	16.83	8 NRVYVHPF
68	19.92	16.51	8 VHLTPVEK
69	14.95	12.22	8 YGGFLRRI
70	37.02	27.86	9 EAKSQGGSN
71	18.04	19.50	9 RPPGFSPFR
72	23.35	17.53	9 WAGGDASGE
73	21.48	20.13	10 DRVYIHPFHL
74	29.44	28.78	10 EHWSYGLRPG
75	20.47	17.76	10 GNHWAVGHLM
76	20.79	16.10	10 GNLWATGHFM
77	15.47	13.63	10 HKTDSFVGLM
78	22.18	17.42	10 PHPFHFFVYK
79	16.86	12.97	10 RFKDNOSQQR
80	37.07	35.60	10 TAQYPPTFGR
81	22.83	21.22	10 CDTDPFQDSR
82	18.14	17.67	11 EADPNKFYGLM
83	19.87	20.63	11 ISRPPGFSPFR
84	14.58	16.82	11 IVMYSPTSILR
85	27.42	23.88	11 LRKKLODVHNF
86	22.25	22.13	11 M1CRPPGFSPFR
87	20.68	15.31	11 RPKPEEFFGLM
88	35.62	33.95	12 DTEDOEDOVDPR
89	22.92	23.19	12 NEKLSQLOTYIM
90	27.60	22.49	12 NSDKFPVYYPGK
91	23.42	21.09	12 YGGFMRRVGRPE
92	21.55	25.20	13 DTEDQCEDQVDPR
93	22.39	17.61	13 DRVYIHPFHLVIH
94	35.44	34.75	13 ELYENKPRRPYIL
95	26.96	25.23	13 MSSIKLIEEQITR
96	23.53	23.55	13 NLSSWIGLDDDCK
97	27.25	26.88	13 RRLIEDAEYAARG
98	20.09	16.72	13 WHWLQLKPGQPMY
99	24.09	21.53	13 YGGFLRRIRPKLK
100	18.80	13.09	14 INLKALAALAKKIL
101	29.74	28.37	16 AVSEHQLLHDGKSIYK
102	22.42	18.45	16 LKKISQYQKFALPQYR
103	29.62	26.55	16 PSOQPNOHPSQPNOOH

Table 1. Continued

No.	Observed	Predicted	Number of sequence residue
104	43.97	48.05	16 KSILKVLEALDLINEK
105	25.18	22.60	16 YGGFMTSEKSOTPLVT
106	37.08	38.07	17 KDDEKLKEFHDGGSKYR
107	23.11	20.30	17 YGGFMTSEKSQTPLVTL
108	20.30	17.12	19 IDLGIHSEWITOATGVWFR
109	21.41	19.15	19 IWGCSGLICTTAVPWNASK
110	29.12	30.95	19 NLAKGKEESLDSDLYAELR
111	20.70	16.67	20 FWLLNVLFPPHTTPAELSNK
112	20.26	20.83	20 STAPLPWPWSPAALRLLOYR
113	21.10	19.15	21 RILAVEYLKDQQLLGIWGCSR
114	32.13	35.05	24 DAEFRHDSGYQNHHOLVFFAEDVK
115	32.77	36.08	24 SYSMEHFRWGKPVGKKRRPVKVYP
116	30.83	28.73	25 VYQHQKAKPWIQPKTKVIPYVRYLM
117	20.88	21.93	26 RPAIWIDLGIHSREWITQATGVWFAK
118	31.23	41.67	28 HSDAVFTDNYTRLRKOMAVKKYLNSILN
119	21.07	23.01	28 SLARTPWAVTCFELEAVCIYACCIHSKM
120	29.77	33.99	31 YGGFTSEKSQTPLVTLFKNAIIKNAYKKGEM
121	30.14	38.40	54 VLSEGEWQLVLHVWAKVEADVAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAE

Predicted retention times were calculated from Table 2 parameters.

considerable contribution was also noted. The hydrophobic residues, Phe (−2.94), Leu (−2.31), Val (−2.19), Trp (−1.80) and Ile (−1.40) made negative contributions. Tyr (−0.11), containing both hydroxyl and hydrophobic residues made, only little negative contribution. In Fig. 3, hydrophilicity retention co-

Table 2
Retention coefficients of amino acids

Amino acid	Present	Parker's	Sasagawa's	Rekker's
Ala	0.28	2.10	0.13	0.53
Cys	0.80[a]	1.40	1.57[a]	–
Asp	2.45	10.00	0.10	−0.02
Glu	1.58	7.80	0.27	−0.07
Phe	−2.94	−9.20	1.71	2.24
Gly	−0.16	5.70	0.22	0.00
His	3.44	2.10	0.34	−0.23
Ile	−1.34	−8.00	1.38	1.99
Lys	2.77	5.70	0.05	0.52
Leu	−2.31	−9.20	1.34	1.99
Met	−0.14	−4.20	0.85	1.08
Asn	3.25	7.00	−0.45	−1.05
Pro	0.77	2.10	0.48	1.O1
Gln	2.35	6.00	0.36	−1.09
Arg	3.90	4.20	0.26	–
Ser	2.53	6.50	0.18	−0.56
Thr	1.73	5.20	0.12	−0.26
Val	−2.19	−3.70	0.38	1.46
Trp	−1.80	−10.00	2.34	2.31
Tyr	−0.11	−51.90	1.23	1.70
b value	12.90			

[a] Carboxymethylcysteine.

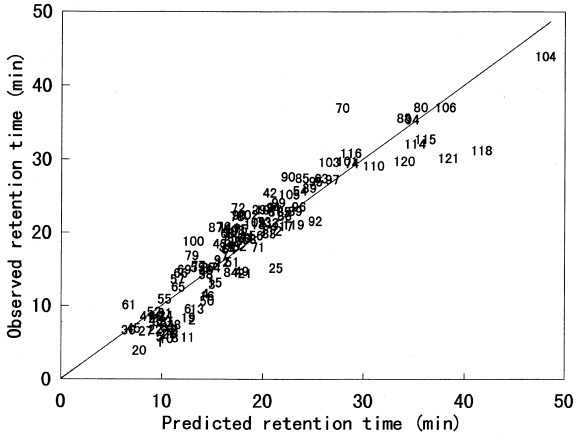

Fig. 1. Relationship between observed retention times of peptides and predicted retention times based on the present hydrophilic retention coefficients. The intercept and slope of the straight line were 0.02 and 0.99, respectively. The correlation coefficient is 0.94.

efficients are plotted against hydrophobicity retention coefficients (Table 2, column 3) reported by Sasagawa et al. [11,12]. It is evident that the

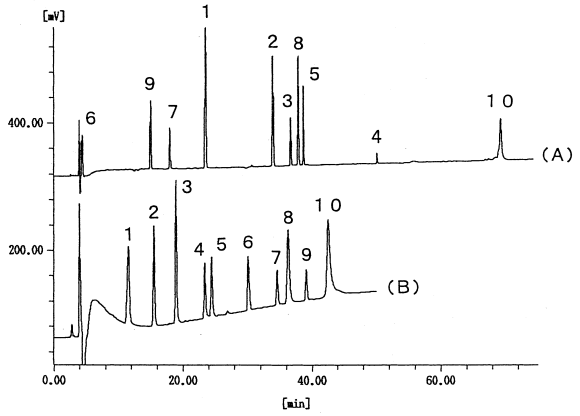

Fig. 2. Chromatograms of peptides separated on (A) TSKgel ODS-80Ts (RPLC) and (B) TSKgel Amide-80 (NPLC). The peptide mixture was separated with (A) 83.3-min linear gradients of ACN from 5 to 55% in 0.1% TFA (0.6% ACN per min.) and (B) 70-min linear gradients of water from 3 to 45% in 0.1% TFA (0.6% water per min.). Peak identifcation: 1=FY; 2=FGGF; 3=FLEEI; 4=DYMGWMDP-NH2; 5=NFTYGGF; 6=AGSE; 7=WAGGDASGE; 8=YGGFMTSQKSQTPLVT; 9=ASTTTNYT; 10=VLSEGEWQLVLHVWAKVEADVAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAE. This figure is taken from Ref. [1].

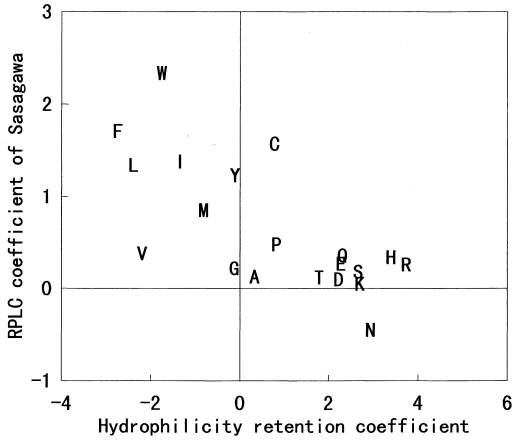

Fig. 3. Correlation of the present hydrophilic retention coefficients with those of Sasagawa. The correlation coefficient is 0.71.

hydrophilicity retention coefficients for the present NPLC are roughly the inverse of the corresponding values for RPLC but correlation is scattered. Such differences in Fig. 3 produce the differences of separation selectivities in Fig. 2, which are not simple reversal relationships.

The results of retention time prediction based on hydrophilicity retention coefficients are shown in Fig. 1 and Table 1. The correlation coefficient was 0.94 and mean percent deviation of retention time, 7.71%. The data in Fig. 1 suggests a slightly curved relationship. Although composition primarily determines retention, conformation and sequence may also have some effect [2,8,11–13].

Parameters [16] have been reported by Parker et al., whose results were derived from the retention times of 20 model synthetic peptides in RPLC. Their parameters are listed in Table 2, column 2. Hydrophilicity retention coefficients are plotted against Parker's in Fig. 4. Retention time prediction based on the work of Parker is shown in Fig. 5. The correlation coefficient was 0.72 and mean percent deviation of observed retention time, 71.99%. Parker's parameters were not useful for predicting peptide retention times during the present NPLC. The relative degree of the contribution in hydrophilicity retention coefficients is almost the same as Parker's parameters. However, some differences were found between hydrophilicity retention coefficients and Parker's. Especially, it was noted that the

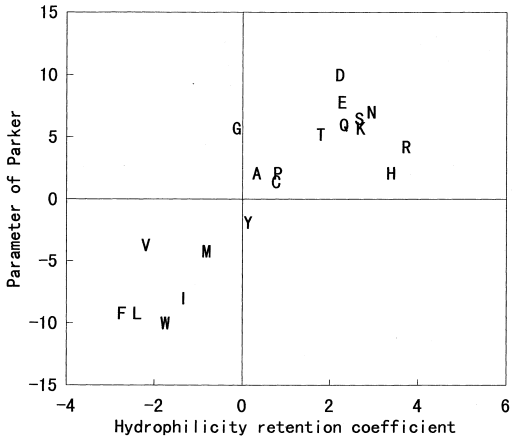

Fig. 4. Correlation of the present hydrophilic retention coefficients with those of Parker. The correlation coefficient is 0.84.

values of Arg and His in this study were higher than Parker's. Such differences explained the poor correlation of the present data and the prediction data using his parameters. In RPLC, furthermore, Sasagawa [11] reported that discrepancies of retention parameters in each case may simple be due to differences in the eluent system (for example, samples, mobile phase and column, etc.). In a similar manner, this discrepancy as shown in Figs. 4 and 5 may be assumed to be due to the difference in separation mode.

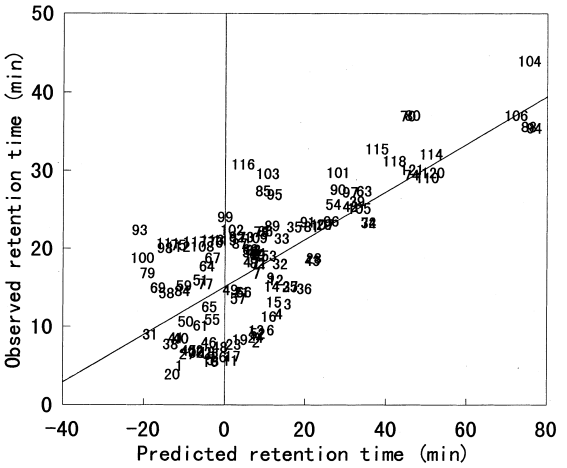

Fig. 5. Relationship between observed retention times of peptides and predicted retention times based on the hydrophilic retention coefficients of Parker. The intercept and slope of the straight line were 0.31 and 15.20, respectively. The correlation coefficient is 0.72.

5. Conclusions

Hydrophilicity retention coefficients derived from the retention times of 121 peptides of known sequence, and means were established for predicting retention times in peptide separation in the present NPLC. A plot of observed vs. predicted retention times appeared to show a slightly curved relationship. A suitable equation model in the present NPLC should perhaps be established which takes into account various interactions such as effects of terminal groups.

References

[1] T. Yoshida, Anal. Chem. 69 (1997) 3038–3043.
[2] A.J. Alpert, J. Chromatogr. 449 (1990) 177–196.
[3] B. Zhu, C.T. Mant, R.S. Hodges, J. Chromatogr. 548 (1991) 13–24.
[4] B. Zhu, C.T. Mant, R.S. Hodges, J. Chromatogr. 594 (1992) 75–86.
[5] J.A. Boutin, A. Ernould, G. Ferry, A. Genton, A.J. Alpert, J. Chromatogr. 583 (1992) 137–143.
[6] U. Bickel, T. Yoshikawa, E.M. Landaw, K.F. Faull, W.M. Pardridge, Proc. Natl. Acad. Sci. USA 90 (1993) 2618–2622.
[7] K. Furuya, K.M. Schegg, H. Wang, D.S. King, D.A. Schooley, Proc. Natl. Acad. Sci. USA 92 (1995) 12323–12327.
[8] J.L. Meek, Proc. Natl. Acad. Sci. USA 77 (1980) 1632–1636.
[9] J.L. Meek, Z.L. Rossetti, J. Chromatogr. 211 (1981) 15–28.
[10] M.J. O'Hare, E.C. Nice, J. Chromatogr. 171 (1979) 209–226.
[11] T. Sasagawa, T. Okuyama, D.C. Teller, J. Chromatogr. 240 (1982) 329–340.
[12] T. Sasagawa, L.H. Ericksson, D.C. Teller, K. Titani, K.A. Walsh, J. Chromatogr. 307 (1984) 29–38.
[13] C. Chabanet, M. Yvon, J. Chromatogr. 599 (1992) 211–225.
[14] C.A. Browne, H.P.J. Bennett, S. Solomon, Anal. Biochem. 171 (1988) 73–90.
[15] S.J. Su, B. Grego, B. Niven, M.T.W. Heam, J. Liq. Chromatogr. 4 (1981) 1745–1764.
[16] J.M.R. Parker, D. Guo, R.S. Hodges, Biochemistry 25 (1986) 5425–5432.
[17] C.T. Mant, N.E. Zhou, R.S. Hodges, J. Chromatogr. 476 (1989) 363–375.

[18] D. Guo, C.T. Mant, A.K. Taneja, J.M.R. Parker, R.S. Hodges, J. Chromatogr. 359 (1986) 499–518.
[19] D. Guo, C.T. Mant, A.K. Taneja, R.S. Hodges, J. Chromatogr. 359 (1986) 519–532.
[20] C.T. Mant, T.W.L. Burke, J.A. Black, R.S. Hodges, J. Chromatogr. 458 (1988) 193–205.
[21] C.A. Browne, H.P.J. Bennett, S. Solomon, Anal. Biochem. 124 (1982) 201–208.
[22] R.F. Rekker, The Hydrophobic Fragment Constant, Elsevier, Amsterdam, 1977, pp. 301.