The effect of column and eluent fluorination on the retention and separation of non-fluorinated amino acids and proteins by HPLC

Katherine Joyner [a], Weizhen Wang [*,b], Yihua Bruce Yu [a,c,**]

[a] Department of Pharmaceutical Sciences, University of Maryland, Baltimore, MD 21201, United States
[b] Department of Mathematics and Statistics, Wright State University, Dayton, OH 45435, United States
[c] Fischell Department of Bioengineering, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

ABSTRACT

The effect of column and eluent fluorination on the retention and separation of non-fluorinated amino acids and proteins in HPLC is investigated. A side-by-side comparison of fluorocarbon column and eluents (F-column and F-eluents) with their hydrocarbon counterparts (H-column and H-eluents) in the separation of a group of 33 analytes, including 30 amino acids and 3 proteins, is conducted. The H-column and the F-column contain the n-C$_8$H$_{17}$ group and n-C$_8$F$_{17}$ group, respectively, in their stationary phases. The H-eluents include ethanol (EtOH) and isopropanol (ISP) while the F-eluents include trifluoroethanol (TFE) and hexafluorosopropanol (HFIP). The 2 columns and 4 eluents generated 8 (column, eluent) pairs that produce 264 retention time data points for the 33 analytes. A statistical analysis of the retention time data reveals that although the H-column is better than the F-column in analyte separation and H-eluents are better than F-eluents in analyte retention, the more critical factor is the proper pairing of column with eluent. Among the conditions explored in this project, optimal retention and separation is achieved when the fluorocarbon column is paired with ethanol, even though TFE is the most polar one among the 4 eluents. This result shows fluorocarbon columns have much potential in chromatographic analysis and separation of non-fluorinated amino acids and proteins.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Conventional high-performance liquid chromatography (HPLC) uses various hydrocarbon columns (e.g., C$_4$, C$_8$ and C$_{18}$) and hydrocarbon eluents (e.g., acetonitrile, methanol, ethanol, etc.) to achieve separation of analytes [1]. As an alternative for hydrocarbon columns, fluorocarbon columns have been developed for the separation of both fluorinated- and non-fluorinated compounds [2–9]. For example, fluorocarbon columns have been used successfully in fluorous mixture synthesis [10–14]. In comparison, the use of fluorocarbon eluents is much less common [15,16]. To better understand the effect of column and eluent fluorination on analyte retention and separation, it is necessary to make side-by-side comparison of fluorocarbon columns and eluents with their hydrocarbon counterparts. Such side-by-side comparisons make it possible to separate the effect of fluorination from other factors, such polarity, size, functional groups, etc. For example,

trifluoroethanol (CF$_3$CH$_2$OH) should be compared with ethanol (CH$_3$CH$_2$OH) rather than methanol (CH$_3$OH) or acetonitrile (CH$_3$CN).

In this work, fluorocarbon column and eluents (F-column and F-eluents) are compared with their hydrocarbon counterparts (H-column and H-eluents) in a systematic fashion. Such side-by-side comparison of fluorinated column and eluents vs. non-fluorinated column and eluents allows us to reveal the effect of column and eluent fluorination on analyte retention and separation. A total of 33 analytes were used in this study, including 30 amino acids and 3 proteins. Statistical analysis is conducted on the retention time data. Through this analysis, we hope to assess the applicability of F-column and F-eluents for the separation of non-fluorinated amino acids and proteins.

Previously, statistical analyses of HPLC data have been conducted to establish the relationship between analyte structure and retention time [17–19]. Instead of focusing on the analytes, this work focuses on columns and eluents; specifically the effect of column and eluent fluorination on analyte retention and separation. The same set of analytes is used as probes to assess different (column, eluent) combinations in terms of analyte retention and separation. Statistical analysis is conducted to compare the various combinations in a pair-wise fashion. The statistical analysis involves three parameters: correlation coefficient, mean and variance.

* Corresponding author. Tel.: +1 937 775 2141.
** Corresponding author. Department of Pharmaceutical Sciences, University of Maryland, 20 Penn St., Room 635, Baltimore, MD 21201, United States.
Tel.: +1 410 706 7514; fax: +1 410 706 5017.
E-mail addresses: weizhen.wang@wright.edu (W. Wang),
byu@rx.umaryland.edu (Y.B. Yu).

The strength of dependency of two HPLC methods is quantified by their correlation coefficient, r. When two HPLC methods produce identical retention behavior among a group of analytes, $r = 1$. We call such methods parallel to each other. On the other hand, if two HPLC methods produce entirely different retention behavior among a group of analytes, $r = 0$. We call such methods orthogonal to each other. In reality, the most likely relationship between two HPLC methods is somewhere between parallel and orthogonal with $0 < r < 1$. As long as $r \neq 0$, two HPLC methods are not independent of each other.

The ability of a HPLC method to retain analytes is quantified by the retention time mean, μ. If no analyte is retained under a HPLC method, μ is zero. If all analytes are well-retained under a HPLC method, μ is large.

The ability of a HPLC method to separate analytes is quantified by the retention time variance, σ^2. If in a HPLC method all the analytes co-elute, σ^2 is zero. If in a HPLC method the analytes are well separated, σ^2 is large.

By comparing fluorocarbon column and eluents with their hydrocarbon counterparts in terms of correlation coefficient, mean and variance, the effect of column and eluent fluorination on analyte retention and separation can be revealed.

2. Experiment design

2.1. Selection of analytes

Analytes are listed in Table 1. 30 amino acids, including both natural and unnatural ones, are selected as analytes. All the amino acids are N-protected by either the Boc group (analytes **2–27**) or the Fmoc-group (analytes **3′**, **7′**, **9′** and **21′**). The reason for using N-protected amino acids is because some free amino acids are not retentive. In addition to amino acids, 3 proteins, lysozyme (**32**), myoglobin (**33**) and bovine serum albumin (**34**), are also included as analytes. Boc-aminoisobutyric acid (**1**) is used as the internal reference in all chromatographic runs.

2.2. Selection of HPLC conditions

All chromatographic runs use the two-eluent, linear gradient and constant temperature (25 °C) mode. This is the most commonly used HPLC method in the separation of amino acids, peptides and proteins [1].

2.3. Selection of columns

The H-column is a Zorbax 300 SB-C_8 column (2.1 mm × 150 mm, 5 μm pore size). The F-column is a Fluoro-Flash® column (2.1 mm × 150 mm, 5 μm pore size) from Fluorous Technologies. The H-column contains the n-C_8H_{17} group in its stationary phase while the F-column contains the n-C_8F_{17} group in its stationary phase.

2.4. Selection of eluents

As the H- and F-columns are both reversed-phase columns, eluent A is H_2O. Eluent B is either a hydrocarbon solvent (H-eluent) or a fluorocarbon solvent (F-eluent). The H-eluents include ethanol (CH_3CH_2OH, EtOH) and isopropanol (($CH_3)_2CHOH$, ISP). The fluorinated counterparts of the H-eluents, trifluoro-ethanol (CF_3CH_2OH, TFE) and hexafluoro-isopropanol (($CF_3)_2CHOH$, HFIP), are used as F-eluents for comparison. Judged by their dielectric constants ε [20], TFE ($\varepsilon = 27.68$) is more polar than its hydrocarbon counterpart EtOH ($\varepsilon = 25.30$) while HFIP ($\varepsilon = 16.70$) is less polar than its hydrocarbon counterpart ISP ($\varepsilon = 20.20$). The average dielectric constant for the two H-eluents, EtOH and ISP, is 22.7

Table 1
List of analytes.

Compound no.	Name	Symbol
1	Boc-aminoisobutyric acid	Boc-Aib
2	Boc-glycine	Boc-Gly
3	Boc-L-alanine	Boc-Ala
3′	Fmoc-L-alanine	Fmoc-Ala
4	Boc-L-valine	Boc-Val
5	Boc-L-leucine	Boc-Leu
6	Boc-L-isoleucine	Boc-Ile
7	Boc-L-norleucine	Boc-Nle
7′	Fmoc-L-norleucine	Fmoc-Nle
8	Boc-L-methionine	Boc-Met
9	Boc-L-proline	Boc-Pro
9′	Fmoc-L-proline	Fmoc-Pro
10	Boc-L-serine	Boc-Ser
11	Boc-L-threonine	Boc-Thr
12	Boc-L-cysteine	Boc-Cys
13	Boc-L-asparagine	Boc-Asn
14	Boc-L-gluamine	Boc-Gln
15	Boc-L-aspartic acid	Boc-Asp
16	Boc-L-glutamic acid	Boc-Glu
17	Boc-L-histidine	Boc-His
18	Boc-L-lysine	Boc-Lys
19	Boc-L-arginine	Boc-Arg
20	Boc-L-tryptophan	Boc-Trp
21	Boc-L-phenylalanine	Boc-Phe
21′	Fmoc-L-phenyalanine	Fmoc-Phe
22	Boc-L-tyrosine	Boc-Tyr
23	Boc-L-phenylalanine(4-F)	Boc-Phe(4-F)
24	Boc-L-phenylalanine(4-Cl)	Boc-Phe(4-Cl)
25	Boc-L-phenylalanine(4-Br)	Boc-Phe(4-Br)
26	Boc-L-phenylalanine(4-I)	Boc-Phe(4-I)
27	Boc-L-phenylalanine(4-NO2)	Boc-Phe(4-NO2)
32	Lysozyme	Lysozyme
33	Myoglobin	Myoglobin
34	Bovine serum albumin	BSA

while the average dielectric constant for the two F-eluents, TFE and HFIP, is 22.2. Therefore, by comparing TFE and HFIP together with EtOH and ISP, contribution to the observed retention time differences by polarity can be eliminated.

2.5. Selection of statistical analysis method

2.5.1. Matched-pair analysis

Statistical analysis methods depend on the type of data. In our analysis, 8 sets of data, as a result of pairing 2 columns with 4 eluents, are generated from the same set of 33 analytes. Therefore, any two of the 8 data sets form a matched pair. Because data in a matched pair experiment are from the same set of subjects, they are likely to be dependent. The strength of the dependency between two data sets is measured by the correlation coefficient. The matched-pair t-test is used to compare the means of the two data sets [21]. The Morgan-Pitman test is used to compare the variances of the two data sets in a matched-pair [22,23].

The major advantage of matched-pair samples over two-independent samples is that the former eliminates subject effects so that the numerical difference in the two samples is due to true differences between the two sampled populations rather than random error. In consequence, the resultant statistical data analysis is more efficient at identifying differences between the two populations. In other words, small differences between two matched-pair samples may be statistically significant.

Definitions of statistics of a sample: For a data set $\{x_1, x_2, \ldots, x_n\}$, the *sample* mean is given by:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{1}$$

The *sample* variance is given by:

$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x - \bar{x})^2 \qquad (2)$$

For a paired data set $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the *sample* correlation coefficient is given by:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \qquad (3)$$

Of course, the self-correlation coefficient, r_{xx}, is 1 by definition. The purpose of statistical analysis is to use sample statistics ($\bar{x}$, s_x^2 and r_{xy}) to draw conclusion regarding population statistics (μ, σ^2, and r).

2.5.2. Compare population means using sample means

To compare two population means using a paired data set $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, a paired t-test is typically employed as follows. Consider a pair of null hypothesis H_0 and alternative hypothesis H_A:

$$H_0 : \mu_x = \mu_y \quad \text{vs.} \quad H_A : \mu_x > \mu_y \qquad (4)$$

where μ_x is the mean of the population from which the sample of $\{x_1, x_2, \ldots, x_n\}$ is selected. μ_y is the mean of the population from which the sample of $\{y_1, y_2, \ldots, y_n\}$ is selected. To proceed with the paired t-test, let $d_i = x_i - y_i$ for $i = 1, 2, \ldots, n$ and obtain the data set $\{d_1, d_2, \ldots, d_n\}$. Then, reject the null hypothesis H_0 at the 0.05 test level if:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} > t_{0.05, n-1} \qquad (5)$$

Otherwise, the null hypothesis H_0 is accepted. $t_{0.05, n-1}$ is the 95th percentile of a t-distribution with $n - 1$ degrees of freedom [21].

2.5.3. Compare population variances using sample variances

To compare two population variances using a paired data set $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the Morgan-Pitman test [22,23] is applied. Consider a pair of null hypothesis H_0 and alternative hypothesis H_A:

$$H_0 : \sigma_x^2 = \sigma_y^2 \quad \text{vs.} \quad H_A : \sigma_x^2 > \sigma_y^2 \qquad (6)$$

where σ_x^2 is the variance of the population from which the sample of $\{x_1, x_2, \ldots, x_n\}$ is selected. σ_y^2 is the variance of the population from which the sample of $\{y_1, y_2, \ldots, y_n\}$ is selected. To proceed with the Morgan-Pitman test, let $u_i = x_i + y_i$ and $v_i = x_i - y_i$ for $i = 1, 2, \ldots, n$ and obtain a new paired data set $\{(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)\}$. Reject the null hypothesis H_0 at the 0.05 test level if

$$r_{uv} > t_{0.05, n-1} \qquad (7)$$

otherwise, the null hypothesis H_0 is accepted. r_{uv} is the correlation coefficient for the paired data set $\{(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)\}$ and can be calculated using Eq. (3).

3. Results and discussion

There are 8 retention time data sets with each set having 33 data points (Table 2). Therefore, there are a total of 264 data points. Statistical analysis is conducted on these data.

Table 2
Retention time values (in min).

Compound no.	H-column EtOH	H-column TFE	H-column ISP	H-column HFIP	F-column EtOH	F-column TFE	F-column ISP	F-column HFIP
2	10.6	8.4	8.3	8.5	11.3	10.9	9.7	10.6
3	13.4	9.9	10.7	11.4	14.7	17.6	15.0	13.9
3'	26.5	23.2	21.9	21.1	23.4	23.1	20.3	17.1
4	20.9	17.8	16.9	17.2	25.4	23.6	22.5	17.2
5	24.2	20.5	20.2	18.8	28.8	25.7	24.6	18.4
6	24.0	20.5	20.0	18.9	26.4	25.7	24.6	18.5
7	24.9	19.3	20.9	17.5	26.6	22.9	23.2	16.0
7'	30.9	31.1	25.4	27.6	30.4	26.5	23.4	20.5
8	19.9	17.3	16.7	17.0	20.1	21.3	19.8	15.8
9	18.2	16.7	14.4	16.9	19.8	22.8	18.7	17.3
9'	27.5	27.5	22.6	24.5	24.4	28.1	21.0	19.5
10	10.6	6.1	8.7	7.6	10.7	12.1	9.2	8.8
11	13.3	9.1	10.5	10.1	13.6	15.1	11.0	11.6
12	25.8	21.1	21.1	19.0	27.5	23.7	23.8	16.7
13	9.8	5.9	8.2	7.9	9.1	11.1	8.2	9.0
14	11.0	6.8	8.8	8.8	10.2	12.1	8.8	10.3
15	12.4	7.5	10.1	8.1	11.9	12.0	9.8	8.6
16	12.8	7.6	10.2	8.1	12.8	12.8	10.6	8.8
17	9.9	6.1	7.9	11.6	11.4	10.0	9.6	10.3
18	9.8	5.1	7.8	7.2	9.1	7.7	9.5	6.5
19	10.6	6.0	8.4	9.3	11.5	12.8	9.6	8.6
20	24.2	18.5	20.2	17.4	24.8	21.3	22.5	15.4
21	24.8	21.7	21.2	20.0	27.4	24.4	23.9	17.9
21'	27.5	30.5	25.0	26.9	26.1	26.5	22.4	19.4
22	19.6	13.8	16.1	12.7	17.7	16.2	17.8	11.6
23	26.5	20.5	21.7	18.5	28.5	24.3	24.7	17.2
24	25.1	21.3	21.8	18.7	28.5	24.3	24.7	17.2
25	29.6	25.2	24.6	22.7	29.2	25.3	25.2	18.1
26	29.6	26.5	24.8	24.0	27.1	25.3	25.1	18.0
27	25.2	21.7	21.2	19.9	28.8	26.0	25.0	18.4
32	27.3	27.2	20.9	26.8	36.4	20.9	28.0	14.2
33	27.3	31.4	19.8	34.1	33.6	21.3	23.9	12.8
34	31.7	40.2	23.7	40.9	36.6	24.6	22.9	24.5

$t_R = t_R(\text{analyte}) + t_R(\text{Boc-Aib})^* - t_R(\text{Boc-Aib})$. $t_R(\text{Boc-Aib})^*$ is the retention time of Boc-Aib (**1**) paired with Boc-Gly (**2**) in each data set.

Table 3
List of correlation coefficients between HPLC methods.

	H-column EtOH	H-column TFE	H-column ISP	H-column HFIP	F-column EtOH	F-column TFE	F-column ISP	F-column HFIP
H-column EtOH	1							
H-column TFE	0.955	1						
H-column ISP	0.990	0.924	1					
H-column HFIP	0.890	0.978	0.839	1				
F-column EtOH	0.951	0.932	0.920	0.900	1			
F-column TFE	0.923	0.860	0.938	0.770	0.874	1		
F-column ISP	0.940	0.857	0.937	0.786	0.956	0.925	1	
F-column HFIP	0.881	0.872	0.886	0.808	0.834	0.939	0.838	1

3.1. Correlation analysis

A correlation analysis is conducted for every 2 sets of retention times listed in Table 2. The pair-wise correlation coefficients are listed in Table 3.

To separate a group of diverse analytes, it is desirable to have weakly correlated HPLC methods. If all methods have high correlation coefficients, then they will produce very similar separation profiles, thereby defeating the purpose of having multiple HPLC methods. The correlation coefficient between (H-column, EtOH) and (H-column, ISP) is 0.99, i.e., r[(H-column, EtOH), (H-column, ISP)] = 0.99. The question is: with the participation of F-column and F-eluents, will r become smaller than 0.99? The answer is yes: the participation of F-column and F-eluents makes r between any two tested methods smaller than 0.99 (Table 3). The effect of column and eluent fluorination on correlation coefficients can be divided into the following three scenarios.

3.1.1. Eluent switching without column switching

On the H-column, r[(H-column, EtOH), (H-column, ISP)] = 0.99; r[(H-column, TFE), (H-column, HFIP)] = 0.98. On the F-column, r[(F-column, EtOH), (F-column, ISP)] = 0.96; r[(F-column, TFE), (F-column, HFIP)] = 0.94. These results indicate that when the same column is used, little variation in retention behavior is introduced by switching the eluents within the H- or F-family (i.e., from EtOH to ISP or from TFE to HFIP).

On the other hand, when the eluents are switched from H- to F-, much more significant variation in retention time is introduced without column switching, as can be seen in Table 3. For example, r[(H-column, ISP), (H-column, HFIP)] = 0.84 and r[(F-column, ISP), (F-column, HFIP)] = 0.84. Such weakened correlation translates into separation differences between H- and F-eluents. For example, in the H-column, Boc-Met (**8**) and Boc-Pro (**9**) are separated by 0.1 min when HFIP is used as eluent B; however, when ISP is used as eluent B, the separation is 2.3 min. In the F-column, myoglobin (**33**) and BSA (**34**) are separated by 1.0 min when ISP is used as eluent B; however, when HFIP is used as eluent B, the separation is 11.7 min.

3.1.2. Column switching without eluent switching

With H-eluents, r[(H-column, EtOH), (F-column, EtOH)] = 0.95; r[(H-column, ISP), (F-column, ISP)] = 0.94. These results indicate that when H-eluents are used, little variation in retention behavior is introduced by switching the column from H- to F-.

However, with F-eluents, r[(H-column, TFE), (F-column, TFE)] = 0.86 and r[(H-column, HFIP), (F-column, HFIP)] = 0.81. These results indicate that when F-eluents are used, significant variation in retention behavior is introduced by switching the column from H- to F-. Such weakened correlation translates into separation differences between H- and F-columns when F-eluents are used. For example, with TFE as eluent B, the separation between Boc-Lys (**18**) and Boc-Arg (**19**) increases from 1.1 to 5.1 min upon column switching from H- to F-. As another example, with HFIP as eluent B, the separation between Boc-Phe(4-I) (**26**) and Boc-Phe(4-NO$_2$) (**27**) increases from 0.4 to 4.1 min upon column switching from F- to H-.

3.1.3. Column + eluent switching

When column switching is accompanied by eluent switching, more variations in retention behavior are introduced. Such variations make it possible to separate one pair of analytes using one method and separate another pair of analytes using another method. One example is Boc-Leu (**5**) and Boc-Ile (**6**) vs. lysozyme (**32**) and BSA (**34**). **5** and **6** (Δt_R = 2.4 min) are better separated than **32** and **34** (Δt_R = 0.2 min) on the F-column with EtOH as eluent B. However, by switching the column from F- to H- and also switching eluent B from EtOH to TFE, **32** and **34** (Δt_R = 13.0 min) become better separated than **5** and **6** (Δt_R = 0.0 min). Here, r[(H-column, TFE), (F-column, EtOH)] = 0.93.

Another example is Boc-Met (**8**) and Boc-Pro (**9**) vs. Boc-Phe (**21**) and Boc-Phe(4-F) (**23**). **21** and **23** (Δt_R = 1.5 min) are better separated than **8** and **9** (Δt_R = 0.1 min) on the H-column with HFIP as eluent B. However, by switching the column from H- to F- and switching eluent B from HFIP to TFE, **8** and **9** (Δt_R = 1.5 min) become better separated than **21** and **23** (Δt_R = 0.1 min). Here, r[(H-column, HFIP), (F-column, TFE)] = 0.77. Close examination of retention time data in Table 2 reveals many examples like the ones presented here.

3.1.4. Summary of correlation analysis

When using the same column, switching the eluent from H- to F- creates much larger variation in retention behavior than switching the eluent within the H- or F-family. When switching the column from H- to F-, using F-eluents creates much larger variation in retention behavior than using H-eluents. When column switching is accompanied by eluent switching, more variations are introduced in retention behavior, as indicated by reduced correlation coefficients between HPLC methods.

It is worth pointing out that although the participation of F-column and F-eluents introduces significant variations into the HPLC retention behavior of the analytes, it falls far short from making any two HPLC methods orthogonal to each other. This is a reflection of the fact that the F-column is still a reversed-phase column.
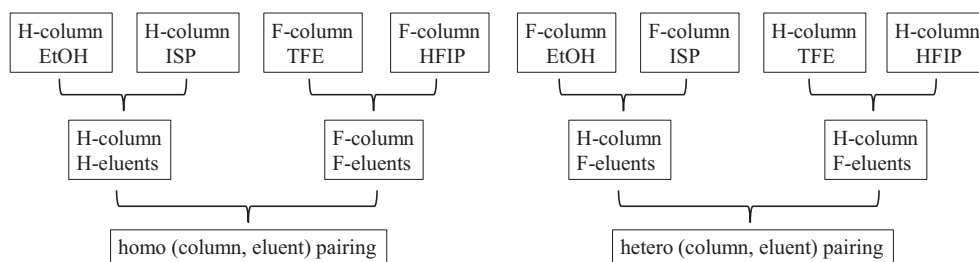
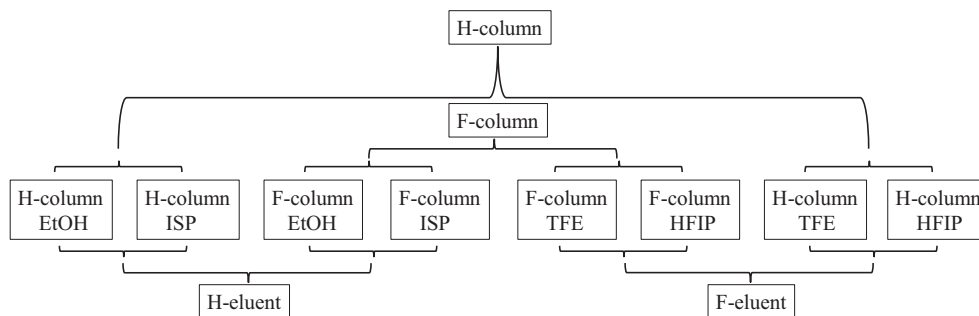**Fig. 1.** Logic flowchart for analyzing the effect of (column, eluent) pairing.

Fig. 2. Logic flowchart for analyzing the effect of column and eluent fluorination.
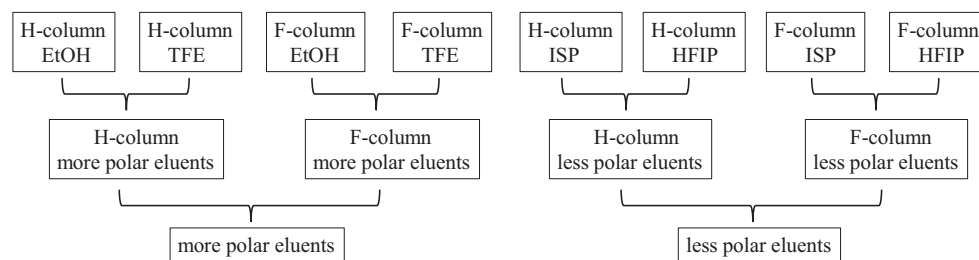

Fig. 3. Logic flowchart for analyzing the effect of eluent polarity.

3.2. Analysis of analyte retention

Having confirmed that F-column and F-eluents can lead to significant variation in the retention behavior of analytes, we now analyze the impact of column and eluent fluorination on analyte retention, which is quantified by the retention time mean. The analysis of retention time mean is conducted from three different angles: the effect of (column, eluent) pairing; the effect of column and eluent fluorination; and the effect of eluent polarity. The logic flowcharts of these three types of analyses are presented in Figs. 1–3.

3.2.1. Effect of (column, eluent) pairing on analyte retention

3.2.1.1. Retention ability of each (column, eluent) pair. Retention time means and pair-wise comparison of retention time means for the 8 (column, eluent) pairs are listed in Table 4. The size of each data set is 33. From Table 4, it is clear that, at the 0.05 test level, (F-column, EtOH) and (F-column, HFIP) are respectively the best and worst pairs for analyte retention. The implication of this result is that eluent selection is more critical for the F-column than for the H-column because the F-column has the best and the worst retention ability, depending on the eluent.

3.2.1.2. Retention ability of (H-column, H-eluents), (H-column, F-eluents), (F-column, H-eluent) and (F-column, F-column). To elimi-

nate the influence of polarity, 4 composite (column, eluent) pairs, (H-column, H-eluents), (F-column, H-eluents), (H-column, F-eluents) and (F-column, F-eluents), as shown in Fig. 1, were analyzed. Each composite pair contains 66 data points. The statistical analysis results of retention times means for these four composite pairs are listed in Table 5. At the 0.05 test level:

$$\mu(\text{F-column, H-eluents}) > \mu(\text{H-column, H-eluents})$$
$$> \mu(\text{H-column, F-eluents})$$
$$\approx \mu(\text{F-column, F-eluents}) \tag{8}$$

Therefore, when eluent polarity is eliminated as a factor, (F-column, H-eluents) has the best retention capacity and (F-column, F-eluents) has the worst retention capacity.

3.2.1.3. Retention ability of homo vs. hetero (column, eluent) pairing. At the next level, we compare homo vs. hetero (column, eluent) pairing on analyte retention. Homo pairing refers to the composite data set [(H-column, H-eluents) + (F-column, F-eluents)] as it pairs the H-column with H-eluents and the F-column with F-eluents. Hetero pairing refers to the composite data set [(H-column, F-eluents) + (F-column, H-eluents)] as it pairs the H-column with F-eluents and the F-column with H-eluents. The size of each composite data set is 132. The retention time means for the homo and hetero pairings are 18.13 min and 19.05 min,

Table 4
Pair-wise comparison of retention time means.

	H-column EtOH $\bar{x} = 20.76$	H-column TFE $\bar{x} = 17.94$	H-column ISP $\bar{x} = 16.98$	H-column FIP $\bar{x} = 17.57$	F-column EtOH $\bar{x} = 21.93$	F-column TFE $\bar{x} = 19.95$	F-column ISP $\bar{x} = 18.75$	F-column HFIP $\bar{x} = 14.81$
H-column EtOH $\bar{x} = 20.76$	=							
H-column TFE $\bar{x} = 17.94$	<	=						
H-column ISP $\bar{x} = 16.98$	<	≈	=					
H-column FIP $\bar{x} = 17.57$	<	≈	≈	=				
F-column EtOH $\bar{x} = 21.93$	>	>	>	>	=			
F-column TFE $\bar{x} = 19.95$	≈	>	>	>	<	=		
F-column ISP $\bar{x} = 18.75$	<	≈	>	≈	<	<	=	
F-column HFIP $\bar{x} = 14.81$	<	<	<	<	<	<	<	=

Comparison is made between an entry in the first column and an entry in the first row; $\bar{x}$ is in min; ">" means larger; "<" means smaller; "=" means equal;"≈" means no statistically significant difference.

Table 5
Pair-wise comparison of retention time means.

	H-column H-eluents $\bar{x} = 18.87$	H-column F-eluents $\bar{x} = 17.76$	F-column H-eluents $\bar{x} = 20.34$	F-column F-eluents $\bar{x} = 17.38$
H-column H-eluents $\bar{x} = 18.87$	=			
H-column F-eluents $\bar{x} = 17.76$	<	=		
F-column H-eluents $\bar{x} = 20.34$	>	>	=	
F-column F-eluents $\bar{x} = 17.38$	<	≈	<	=

Comparison is made between an entry in the first column and an entry in the first row; $\bar{x}$ is in min; ">" means larger; "<" means smaller; "=" means equal; "≈" means no statistically significant difference.

respectively. At the 0.05 test level,

$$\mu[(\text{H-column, F-eluents})$$
$$+ (\text{F-column, H-eluents})] > \mu[(\text{H-column, H-eluents})$$
$$+ (\text{F-column, F-eluents})] \tag{9}$$

Therefore, in terms of analyte retention, hetero (column, eluent) pairing is better than homo (column, eluent) pairing. Clearly, this conclusion is the result of fluorination as eluent polarity is balanced out on both sides of Eq. (9).

3.2.2. Effect of fluorination on analyte retention

3.2.2.1. Effect of eluent fluorination on analyte retention. To assess the effect of eluent fluorination on analyte retention, retention time data for H-eluents in both H- and F-columns are combined together to be compared with retention time data for F-eluents in both H- and F-columns. The size of each composite data set is 132. The retention time means for H- and F-eluents are 19.61 min and 17.57 min, respectively. At the 0.05 test level,

$$\mu(\text{H-eluents}) > \mu(\text{F-eluents}) \tag{10}$$

Therefore, H-eluents are more retentive of amino acids and proteins than F-eluents.

3.2.2.2. Effect of column fluorination on analyte retention. To assess the effect of column fluorination on analyte retention, retention time data for the H-column with both H- and F-eluents are combined together to be compared with retention time data for the F-column with both H- and F-eluents. The size of each composite data set is 132. The retention time means for H- and F-columns are 18.32 min and 18.86 min, respectively. At the 0.05 test level,

$$\mu(\text{H-column}) \approx \mu(\text{F-column}) \tag{11}$$

i.e., there is no statistically significant difference between H- and F-columns in analyte retention.

3.2.3. Effect of eluent polarity on analyte retention

To reveal the effect of eluent polarity on analyte retention, data from the two more polar eluents, TFE (ε = 27.68) and EtOH (ε = 25.30), are combined together to be compared with data from the two less polar eluents, ISP (ε = 20.20) and HFIP (ε = 16.70). Data from the H- and F- columns are analyzed first separately and then together.

3.2.3.1. More polar eluents vs. less polar eluents in the H-column. In the H-column, $\bar{x}$(H-column, more polar eluents) = 20.15 min; $\bar{x}$(H-column, less polar eluents) = 17.03 min. The size of each composite data set is 66. At the 0.05 test level,

$$\mu(\text{H-column, more polar eluents}) > \tag{12}$$
$$\mu(\text{H-column, less polar eluents})$$

3.2.3.2. More polar eluents vs. less polar eluents in the F-column. In the F-column, $\bar{x}$(F-column, more polar eluents) = 20.94 min; $\bar{x}$(F-column, less polar eluents) = 16.78 min. The size of each

Table 6
Pair-wise comparison of retention time variances.

	H-column EtOH $s^2 = 55.20$	H-column TFE $s^2 = 85.38$	H-column ISP $s^2 = 38.44$	H-column HFIP $s^2 = 65.61$	F-column EtOH $s^2 = 71.91$	F-column TFE $s^2 = 37.33$	F-column ISP $s^2 = 43.56$	F-column HFIP $s^2 = 19.15$
H-column EtOH $s^2 = 55.20$	=							
H-column TFE $s^2 = 85.38$	>	=						
H-column ISP $s^2 = 38.44$	<	<	=					
H-column HFIP $s^2 = 65.61$	≈	<	>	=				
F-column EtOH $s^2 = 71.91$	>	≈	>	≈	=			
F-column TFE $s^2 = 37.33$	<	<	≈	<	<	=		
F-column ISP $s^2 = 43.56$	<	<	≈	<	<	≈	=	
F-column HFIP $s^2 = 19.15$	<	<	<	<	<	<	<	=

Comparison is made between an entry in the first column and an entry in the first row; s^2 is in min^2; ">" means larger; "<" means smaller; "=" means equal; "≈" means no statistically significant difference.

composite data set is 66. At the 0.05 test level,

$$\mu(\text{F-column, more polar eluents}) > \mu(\text{F-column,} \tag{13}$$
$$\text{less polar eluents})$$

3.2.3.3. More polar eluents vs. less polar eluents. When data for the more polar eluents from H- and F-columns are combined, $\bar{x}$(more polar eluents) = 20.15 min. When data for the less polar eluents from H- and F-columns are combined, $\bar{x}$(less polar eluents) = 17.03 min. The size of each composite data set is 132. At the 0.05 test level,

$$\mu(\text{more polar eluents}) > \mu(\text{less polar eluents}) \tag{14}$$

3.2.4. Summary of retention ability

There is no statistically significant difference between H- and F-columns in retaining amino acids and proteins. H-eluents result in stronger analyte retention than F-eluents. To achieve higher retention, it is preferable to pair H-column with F-eluents and F-column with H-eluents. Eluent selection is more critical for the F-column than for the H-column. (F-column, EtOH) gives the best retention while (F-column, HFIP) gives the worst most retention. As for eluent polarity, more polar eluents lead to better retention in both H- and F-columns.

3.3. Analysis of analyte separation

Having analyzed the impact of column and eluent fluorination on analyte retention, we now analyze the impact of column and eluent fluorination on analyte separation, which is quantified by the retention time variance. Similar to the analysis of analyte retention, the analysis of analyte separation is conducted from three different angles: the effect of (column, eluent) pairing; the effect of column and eluent fluorination; and the effect of eluent polarity. The logic flowcharts of these three types of analyses are presented in Figs. 1–3.

3.3.1. Effect of (column, eluent) pairing on analyte separation

3.3.1.1. Separation ability of each (column, eluent) pair. Retention time variances and pair-wise comparison of retention time variances of the 8 (column, eluent) pairs are listed in Table 6. The size of each data set is 33. From Table 6, it is clear that, of the 8 pairs, (H-column, TFE) ($s^2 = 85.38$) and (F-column, EtOH) ($s^2 = 71.91$) are the best pairs for analyte separation; there is no statistically significant difference between them. On the other hand, (F-column, HFIP) ($s^2 = 19.15$) is by far the worst pair for analyte separation. Similar to analyte retention, eluent selection is more critical to the F-column than for the H-column for analyte separation as the F-column has the best and the worse separation ability, depending on eluents.

3.3.1.2. Separation ability of (H-column, H-eluents), (H-column, F-eluents), (F-column, H-eluent) and (F-column, F-column). To eliminate the influence of polarity, 4 composite (column, eluent) pairs, (H-column, H-eluents), (F-column, H-eluents), (H-column, F-eluents) and (F-column, F-eluents), as shown in Fig. 1, were analyzed. Each composite pair contains 66 data points. The statistical analysis results of retention times means for these four composite pairs are listed in Table 7. At the 0.05 test level

$$\sigma^2(\text{H-column, F-eluents}) > \sigma^2(\text{F-column, H-eluents}) > \tag{15}$$
$$\sigma^2(\text{H-column, H-eluents}) >$$
$$\sigma^2(\text{F-column, F-eluents})$$

Therefore, when eluent polarity is eliminated as a factor, (H-column, F-eluents) has the best separation capacity and (F-column, F-eluents) has the worst separation capacity.

3.3.1.3. Separation ability of homo vs. hetero pairing. At the next level, we compare homo vs. hetero (column, eluent) pairing on analyte separation. Homo pairing refers to the composite data set [(H-column, H-eluents) + (F-column, F-eluents)] as it pairs the H-column with H-eluents and the F-column with F-eluents.

Table 7
Pair-wise comparison of retention time variances.

	H-column H-eluents $s^2 = 49.77$	H-column F-eluents $s^2 = 74.30$	F-column H-eluents $s^2 = 59.41$	F-column F-eluents $s^2 = 34.53$
H-column H-eluents $s^2 = 47.99$	=			
H-column F-eluents $s^2 = 74.30$	>	=		
F-column H-eluents $s^2 = 59.41$	>	<	=	
F-column F-eluents $s^2 = 34.53$	<	<	<	=

Comparison is made between an entry in the first column and an entry in the first row; s^2 is in min^2; ">" means larger; "<" means smaller; "=" means equal; "≈" means no statistically significant difference.

Hetero pairing refers to the composite data set [(H-column, F-eluents) + (F-column, H-eluents)] as it pairs the H-column with F-eluents and the F-column with H-eluents. The size of each composite data set is 132. The retention time variances for the homo and hetero pairings are 42.38 min^2 and 68.06 min^2, respectively. At the 0.05 test level,

$$\sigma^2[(\text{H-column, F-eluents}) + (\text{F-column, H-eluents})] > \sigma^2[(\text{H-column, H-eluents}) + (\text{F-column, F-eluents})] \qquad (16)$$

Just like analyte retention, hetero (column, eluent) pairing is better than homo (column, eluent) pairing for analyte separation. Clearly, this conclusion is the result of fluorination as eluent polarity is balanced out on both sides of Eq. (16).

3.3.2. Effect of fluorination on analyte separation

3.3.2.1. Effect of eluent fluorination on analyte separation. To assess the effect of eluent fluorination on analyte separation, retention time data for H-eluents in both H- and F-columns are combined together to be compared with retention time data for F-eluents in both H- and F-columns. The size of each composite data set is 132. The retention time variances for H- and F-eluents are 54.76 min^2 and 54.02 min^2, respectively. At the 0.05 test level,

$$\sigma^2(\text{H-eluents}) \approx \sigma^2(\text{F-eluents}) \qquad (17)$$

Therefore, there is no statistically significant difference between H-eluents and F-eluents in analyte separation.

3.3.2.2. Effect of column fluorination on analyte separation. To assess the effect of column fluorination on analyte separation, retention time data for the H-column with both H- and F-eluents are combined together to be compared with retention time data for the F-column with both H- and F-eluents. The size of each composite data set is 132. The retention time variances for the H- and F-columns are 61.94 min^2 and 48.86 min^2, respectively. At the 0.05 test level,

$$\sigma^2(\text{H-column}) > \sigma^2(\text{F-column}) \qquad (18)$$

Therefore, H-column is better than F-column at separating amino acids and proteins.

3.3.3. Effect of eluent polarity on analyte separation

To reveal the effect of eluent polarity on analyte separation, data from the two more polar eluents, TFE ($\varepsilon = 27.68$) and EtOH ($\varepsilon = 25.30$), are combined together to be compared with data from the two less polar eluents, ISP ($\varepsilon = 20.20$) and HFIP ($\varepsilon = 16.70$). Data from the H- and F- columns are analyzed first separately and then together.

3.3.3.1. More polar eluents vs. less polar eluents in the H-column. In the H-column, s^2(H-column, more polar eluents) = 71.23 min^2; s^2(H-column, less polar eluents) = 51.27 min^2. The size of each composite data set is 66. At the 0.05 test level,

$$\sigma^2(\text{H-column, more polar eluents}) > \sigma^2(\text{H-column, less polar eluents}) \qquad (19)$$

3.3.3.2. More polar eluents vs. less polar eluents in the F-column. In the F-column, s^2(F-column, more polar eluents) = 54.76 min^2; s^2(F-column, less polar eluents) = 34.81 min^2. The size of each composite data set is 66. At the 0.05 test level,

$$\sigma^2(\text{F-column, more polar eluents}) > \sigma^2(\text{F-column, less polar eluents}) \qquad (20)$$

3.3.3.3. More polar eluents vs. less polar eluents. When data for the more polar eluents from H- and F-columns are combined, s^2(more polar eluents) = 63.20 min^2. When data for the less polar eluents from H- and F-columns are combined, s^2(less polar eluents) = 42.77 min^2. The size of each composite data set is 132. At the 0.05 test level,

$$\sigma^2(\text{more polar eluents}) > \sigma^2(\text{less polar eluents}) \qquad (21)$$

3.3.4. Summary of separation ability

H-column is better than F-column in separating amino acids and proteins. There is no statistically significant difference between H-eluents and F-eluents in analyte separation. To achieve better separation, it is preferable to pair H-column with F-eluents and F-column with H-eluents. Eluent selection is more critical for the F-column than for the H-column. (H-column, TFE) and (F-column, EtOH) give best separation while (F-column, HFIP) gives the worst separation. As for eluent polarity, more polar eluents lead to better separation in both H- and F-columns.

4. Conclusion

F-column and F-eluents introduces significant variation in the retention behavior of non-fluorinated amino acids and proteins. H-column is better than F-column in analyte separation but there is no statistically significant difference between H-column and F-column in analyte retention. H-eluents are better than F-eluents in analyte retention but there is no statistically significant different between H-eluents and F-eluents in analyte separation. More critical than column and eluents is the proper pairing of column and eluents. To achieve the best retention and separation outcome, H-column should be paired with F-eluents and F-column should be paired with H-eluents. Choosing the right eluent is more critical for the F-column than for the H-column as the F-column can achieve the best and worst retention and separation, depending on the eluents. Additionally, more polar eluents produce better analyte retention and separation in both H- and F-columns. When taking both retention and separation into account, the optimal pairing is F-column with EtOH as eluent.

5. Experimental

5.1. Materials and instruments

5.1.1. Amino acids and proteins

Amino acids were purchased from Aapptec or Novabiochem and proteins from Sigma–Aldrich. All purchased amino acids and proteins were used without further purification. All chiral amino acids have the L-configuration.

5.1.2. Eluents

EtOH was from Sigma–Aldrich (spectrophotometeric grade); ISP was from Fisher (HPLC grade); TFE and HFIP were from Oakwood Products (reagent grade). EtOH and ISP were used as purchased. TFE and HFIP were distilled before usage. Water was purified by a PURELAB Ultra Mk2 water purification system. Trifluoroacetic acid (TFA) was purchased from Oakwood Products

5.1.3. Instrumentation

Agilent Technologies 1200 Series liquid chromatography system housed in a temperature-controlled room.

5.2. Chromatographic conditions

The chromatographic conditions were: eluent A: 0.1% TFA in water; eluent B: 0.1% TFA in EtOH, or TFE, or ISP, or HFIP; gradient: 2% B/min, 0% B–100% B in 50 min; flow rate: 0.25 mL/min; column chamber temperature: 25 °C; room temperature: 25 °C. Each analyte was co-injected with the internal standard Boc-Aib (**1**).

Acknowledgements

References

[1] C.T. Mant, R.S. Hodges (Eds.), High Performance Liquid Chromatography of peptides and Proteins: Separation, Analysis, and Coformation, CRC Press, Boca Raton, 1991.
[2] M. Przybyciel, LCGC (North America) 23 (2005) 554–565.
[3] M.R. Euerby, A.P. McKeown, P. Petersson, J. Sep. Sci. 26 (2003) 295–306.
[4] C.F. Poole, H. Ahmed, W. Kiridena, C. DeKay, W.W. Koziol, Chromatographia 65 (2007) 127–139.
[5] W. Kiridena, C. DeKay, W.W. Koziol, Z. Ali, H. Ahmed, C.F. Poole, Chromatographia 63 (2006) 407–417.
[6] W. Zhang, J. Fluorine Chem. 129 (2008) 910–919.
[7] D.P. Curran, Synth. Lett. 9 (2001) 1488–1496.
[8] D.P. Curran, Z. Lou, Green Chem. 3 (2001) G3–G7.
[9] N. Xiao, Y.B. Yu, J. Fluorine Chem. 131 (2010) 439–445.
[10] Z. Luo, Q. Zheng, Y. Oderaotoshi, D.P. Curran, Science 291 (2001) 1766–1769.
[11] W. Zhang, Z. Luo, C.H.-T. Chen, D.P. Curran, J. Am. Chem. Soc. 124 (2002) 10443–10450.
[12] Z.-X. Jiang, Y.B. Yu, J. Org. Chem. 75 (2010) 2044–2049.
[13] P.C. de Visser, M. van Helden, D.V. Filippov, G.A. van der marel, J.W. Drijfhout, J.H. van Boom, D. Noort, H.S. Overkleeft, Tetrahedron Lett. 44 (2003) 9013–9016.
[14] S.M. Brittain, S.B. Ficarro, A. Brock, E.C. Peters, Nat. Biotechnol. 23 (2005) 463–468.
[15] K. Valkó, S. Espinosa, C.M. Du, E. Bosch, M. Rosés, M. Bevan, M.H. Abraham, J. Chromatogr. A 933 (2001) 73–81.
[16] T. Monde, T. kamiusuki, T. Kuroda, K. Mikumo, T. Ohkawa, H. Fukube, J. Chromatogr. A 722 (1996) 273–280.
[17] K. Bodzioch, A. Durand, R. Kaliszan, T. Bączek, Y. Vander Heyden, Talanta 81 (2010) 1711–1718.
[18] P. Zhou, F. Tian, F. Lu, Z. Shang, J. Chromatogr. A 1216 (2009) 3107–3116.
[19] N. Pfeifer, A. leinenbach, C.G. Huber, O. Kohlbacher, J. Proteome Res. 8 (2009) 4109–4115.
[20] D.R. Lide (Ed.), CRC Handbook of Chemistry and Physics, 84th ed., CRC Press, Boca Raton, 2004.
[21] R.E. Walpole, R.H. Myers, S.L. Myers, K. Ye, Probability and Statistics for Engineers and Scientists, 8th ed., Pearson Prentice Hall, Upper Saddle River, New Jersey, 2007.
[22] W.A. Morgan, Biometrika 31 (1939) 13–19.
[23] E.J.G. Pitman, Biometrika 31 (1939) 9–12.