Journal of Fluorine Chemistry 132 (2011) 114–122

Contents lists available at ScienceDirect

Journal of Fluorine Chemistry

journal homepage: www.elsevier.com/locate/fluor

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

Article history:
Received 11 October 2010
Received in revised form 3 December 2010
Accepted 6 December 2010
Available online 13 December 2010

Keywords:
HPLC
Fluorocarbon column
Fluorocarbon eluents
Paired *t*-test
Morgan–Pitman test

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_8H_{17} group and *n*-C_8F_{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.

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_3CH_2OH) should be compared with ethanol (CH_3CH_2OH) rather than methanol (CH_3OH) or acetonitrile (CH_3CN).

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).

0022-1139/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfluchem.2010.12.005

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 FluoroFlash® 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 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.

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

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:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (1)$$