Journal of Chromatography A, 1368 (2014) 125–131

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Use of individual retention modeling for gradient optimization in hydrophilic interaction chromatography: Separation of nucleobases and nucleosides

Eva Tyteca[a,*], Davy Guillarme[b], Gert Desmet[a]

[a] Vrije Universiteit Brussel, Department of Chemical Engineering, Pleinlaan 2, B-1050 Brussels, Belgium
[b] School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, 20, Boulevard d'Yvoy, 1211 Geneva 4, Switzerland

ARTICLE INFO

Article history:
Received 7 August 2014
Received in revised form
24 September 2014
Accepted 25 September 2014
Available online 13 October 2014

Keywords:
Computer-assisted method development
Individual retention modeling
Nucleobases
Nucleosides
HILIC

ABSTRACT

In this study, the separation of twelve nucleobases and nucleosides was optimized via chromatogram simulation (i.e., prediction of individual retention times and estimation of the peak widths) with the use of an empirical (reversed-phase) non-linear model proposed by Neue and Kuss. Retention time prediction errors of less than 2% were observed for all compounds on different stationary phases. As a single HILIC column could not resolve all peaks, the modeling was extended to coupled-column systems (with different stationary phase chemistries) to increase the separation efficiency and selectivity. The analytical expressions for the gradient retention factor on a coupled column system were derived and accurate retention time predictions were obtained (<2% prediction errors in general). The optimized gradient (predicted by the optimization software) included coupling of an amide and an pentahydroxy functionalized silica stationary phases with a gradient profile from 95 to 85%ACN in 6 min and resulted in almost baseline separation of the twelve nucleobases and nucleosides in less than 7 min. The final separation was obtained in less than 4 h of instrument time (including equilibration times) and was fully obtained via computer-based optimization. As such, this study provides an example of a case where individual retention modeling can be used as a way to optimize the gradient conditions in the HILIC mode using a non-linear model such as the Neue and Kuss model.

1. Introduction

To speed up the method development (MD) process in chromatography, fully or semi-automated MD software programs have been developed for reversed phase liquid chromatography (RPLC) [1–6]. These automated MD strategies for RPLC described in literature are either search-based (e.g. using the Simplex method) [7], model-based (e.g. Drylab [8], Chromsword) [9] or based on a combination of both (design of experiments (DoE), multiple linear regression (MLR) [10], predictive elution window stretching and shifting method (PEWS^2) [11]). Recently, the PEWS^2 method was successfully applied for the gradient optimization in hydrophilic interaction chromatography (HILIC) [12].

HILIC is becoming more and more popular, for the determination of hydrophilic compounds, poorly retained in RPLC conditions, and for the analysis of ionizable compounds [13]. HILIC retention can be considered as a mixed-mode mechanism, combining hydrophilic partitioning of the analytes between the organic-rich mobile phase and the water enriched layer partially immobilized on the stationary phase, compounds adsorption through hydrogen bonds, and electrostatic and ionic interactions [14]. Although both the linear reversed phase (semi-log) (Eq. (1)) and the normal phase (log-log) retention relationships have been used in literature to model HILIC retention [15], the dependency of $\ln(k)$ on both φ and $\log(\varphi)$ in the HILIC separation mode does not follow a strict linear relationship [15,16]. Based on these observations, Jin et al. proposed a mixed model combining partitioning and adsorption terms (Eq. (2)) to describe the retention behavior in HILIC [15]. Jandera and Hájek [17] introduced an extended model including one extra parameter to better describe the retention at low percentages of water. Other (reversed-phase) non-linear models such as the quadratic model and the empirical model from Neue and Kuss (Eq. (3)) have also been proposed in literature to describe the retention relationship in the HILIC separation mode [12].

$$\ln(k) = \ln(k_w) - S\varphi \quad (1)$$

* Corresponding author. Tel.: +32 2 2693617.
E-mail address: eva.tyteca@vub.ac.be (E. Tyteca).

http://dx.doi.org/10.1016/j.chroma.2014.09.065
0021-9673/

$$\ln(k) = \ln(k_w) + S_1\varphi + S_2 \ \ln(\varphi) \quad (2)$$

$$\ln(k) = \ln(k_w) + 2\ln(1 + S_2\varphi) - \frac{S_1\varphi}{1 + S_2\varphi} \quad (3)$$

where φ is the fraction of water, k_w is the extrapolated values of k for $\varphi = 0$ (i.e., pure ACN) and S is the solvent strength parameter which is a constant for a given compound and organic solvent [18,19], S_1 the slope for the non-linear models, S_2 the curvature coefficient [21].

Greco et al. reported determination coefficients R^2 above 0.99 for 14 benzoic acids using Eq. (2) [14]. In a recent study, including several analytes possessing diverse physico-chemical properties, we reported $R^2_{adjusted}$ and Q^2-values for the empirical Neue and Kuss (Eq. (3)) model, close to those of the mixed model (Eq. (2)) [12]. However, the gradient retention predictions were much less accurate in HILIC than RPLC, restricting the use of commercial software packages requiring the simulation of the retention of every peak in the chromatogram [12].

The expression for the gradient retention factor can be found by solving the fundamental gradient equation:

$$t_0 = \int_0^{t_R - t_0} \frac{dt_s}{k(\varphi)} \quad (4)$$

where t_R and t_0 are the total retention time and the column dead time, respectively [20] and t_s is the time in the stationary phase. Whereas the mixed model (Eq. (2)) no longer has an analytical solution to the fundamental gradient equation, it is one of the virtues of the Neue and Kuss-model (Eq. (3)) that it is easily amenable to an analytical solution, leading to the following expression for the effective retention factor $k_{eff} = (t_R - t_0)/t_0$ [21]:

$$k_{eff} = \frac{t_D}{t_0} + \frac{[\phi_0 + (1 + S_2\phi_0)/S_1 \ln(1 + \beta k'_w S_1(t_0 - t_D/k_0)\exp(-S_1\phi_0/1 + S_2\phi_0))]/[1 - (S_2(1 + S_2\phi_0))/S_1 \ln(1 + \beta k'_w S_1(t_0 - t_D/k_0)\exp(-S_1\phi_0/1 + S_2\phi_0))] - \phi_0}{\beta t_0} \quad (5)$$

where t_D is the dwell time, β is the gradient slope, defined as $(\varphi_e - \varphi_0)/t_G$ and k_0 is the isocratic retention factor at the start of the gradient, i.e. for $\varphi = \varphi_0$.

In a previous study, involving samples with pteridines, sugars and a drug mixture of venlafaxine, tramadol and their metabolites in HILIC we observed strongly non-linear retention relationships, that were so complex that they could not be modeled using any of the existing non-linear retention models. As a consequence, the traditional retention modeling approach [22] that is so successful and widespread used in RPLC, could not be applied and we had to switch to a model-guided search approach (so-called PEWS^2-technique [11,12]). For a sample containing compounds behaving more "nicely", individual retention modeling using a non-linear 3-parameter model (mixed model and the empirical model from Neue and Kuss) could be used to optimize the gradient conditions.

In the present contribution, we report on the optimization of a HILIC separation of twelve nucleobases and nucleosides using this individual retention modeling and the coupling of columns (with different stationary phases). In the adopted MD work flow, three different column chemistries and two mobile phase pH were tested.

The separation of nucleic acids and analogs are of great interest in pharmaceutical sciences, genomics, genetics and others [23]. Massolini and coworkers used narrow bore columns with gradient elution (ZIC-HILIC, 150 mm × 2.1 mm, 5 μm) to separate the 12 nucleobases and nucleosides included in this study in 55 min. Nikitas et al. used reversed-phase liquid chromatography (without ion paring reagent) to separate 16 nucleobases and nucleosides using two-segment gradient profiles [24]. The linear log-log retention relationship showed better prediction accuracy compared to the semi-log retention relationship (Eq. (1)), both for simple linear and multi-linear gradient profiles. Jin et al. [15] reported R^2-values around 0.999 and isocratic prediction errors smaller than 5% at 95%ACN for eight nucleosides on six different columns, using Eq. (2). However, no gradient predictions were ever made using this mixed model. Pappa-Louisi and coworkers reported on the gradient retention time predictions of amino acids in HILIC using the LSS-model (Eq. (1)) for linear, multi-linear and curved elution profiles with the same starting concentration.

2. Experimental

Water was obtained from a Milli-Q water purification system from Millipore (Bedford, MA, USA). Acetonitrile (ACN), methanol (MeOH) and formic acid were of ULC-MS grade and purchased from Biosolve (Valkenswaald, Netherlands). Ammonium acetate (>98%, Sigma–Aldrich, Belgium) buffer 10 mM was prepared by weighing adequate mass of ammonium acetate. The pH was adjusted to 3.0 or 6.0 with formic acid (FA). All compounds (thymine, uridine, inosine, adenine, uracil, hypoxanthine, cytidine, thymidine, adenosine, cytosine, guanine and guanosine) were purchased from Sigma–Aldrich. Stock solutions of 1000 ppm were prepared in MeOH except for guanosine and guanine (0.1%FA and KOH solutions, respectively). The samples were diluted in ACN + 0.1% FA and injected at 100 ppm. The injection volume was 1 μL. All experiments were performed on an Agilent Infinity 1290 system with a dwell volume of 112 μL. The columns were Supelco HILIC and Supelco OH5, i.e., bare silica and pentahydroxy functionalized silica respectively (100 mm × 2.1 mm, 2.7 μm), and Waters Acquity Amide, i.e., amide functionalized silica (100 mm × 2.1 mm, 1.7 μm). The dead times were measured at 70%ACN: 0.419, 0.381 and 0.422 min on the bare silica, the pentahydroxy and the amide stationary phase, respectively. The flow rate was 0.5 mL/min. Temperature was set 25 °C. UV-detection was set at 254 nm. The isocratic scouting runs were performed using different subgroups (see different colors in Fig. 3), to avoid overlapping peaks and ensure accurate retention time determination.

2.1. Data analysis

Matlab software (2009) was used for the estimation of the model parameters and the calculation of the optimal gradient conditions via an in-house written Matlab code. The retention factors of the isocratic scouting runs were used to obtain the model parameters (Eqs. (2) and (3)) via least squares fitting using the Matlab® routine *lsqcurvefit*. Gradient retention times were either predicted by implementing the analytical expression (Eq. (5), for the Neue–Kuss model) in Matlab-software or, for the other non-linear model (Eq. (2)), via numerical integration of the fundamental gradient equation via an in-house written Matlab routine based on the trapezoid rule.

$$t_0 = \int_0^{t_R - t_0} \frac{dt_s}{k(\varphi)} = \frac{t_D}{k_0} + \frac{1}{\beta}\int_{\varphi_0}^{\varphi_{elution}} \frac{d\varphi}{k(\varphi)} \Leftrightarrow \beta \cdot \left(t_0 - \frac{t_D}{k_0}\right) = \int_{\varphi_0}^{\varphi_{elution}} \frac{d\varphi}{k(\varphi)} = I \quad (6)$$

The percentage of ACN at elution $\varphi_{elution}$ is obtained by minimizing $I - \beta^*(t_0 - t_D/k_0)$. From $\varphi_{elution}$, the effective retention factor k_{eff} can be calculated via:

$$k_{eff} = \frac{t_R - t_0}{t_0} = \frac{t_D}{t_0} + \frac{\phi_{elution} - \phi_0}{\beta t_0} \quad (7)$$