Full length article

Applicability of retention modelling in hydrophilic-interaction liquid chromatography for algorithmic optimization programs with gradient-scanning techniques

Bob W.J. Pirok [a,b,*], Stef R.A. Molenaar [a], Rianne E. van Outersterp [a], Peter J. Schoenmakers [a]

[a] *University of Amsterdam, van 't Hoff Institute for Molecular Sciences, Analytical-Chemistry Group, Science Park 904, 1098 XH, Amsterdam, The Netherlands*
[b] *TI-COAST, Science Park 904, 1098 XH, Amsterdam, The Netherlands*

ARTICLE INFO

ABSTRACT

Computer-aided method-development programs require accurate models to describe retention and to make predictions based on a limited number of scouting gradients. The performance of five different retention models for hydrophilic-interaction chromatography (HILIC) is assessed for a wide range of analytes. Gradient-elution equations are presented for each model, using Simpson's Rule to approximate the integral in case no exact solution exists. For most compound classes the adsorption model, i.e. a linear relation between the logarithm of the retention factor and the logarithm of the composition, is found to provide the most robust performance. Prediction accuracies depended on analyte class, with peptide retention being predicted least accurately, and on the stationary phase, with better results for a diol column than for an amide column. The two-parameter adsorption model is also attractive, because it can be used with good results using only two scanning gradients. This model is recommended as the first-choice model for describing and predicting HILIC retention data, because of its accuracy and linearity. Other models (linear solvent-strength model, mixed-mode model) should only be considered after validating their applicability in specific cases.

1. Introduction

Hydrophilic interaction chromatography (HILIC) has become increasingly important for the analysis of highly polar analytes, such as antioxidants [1], sugars (e.g. glycomics [2–4]), (plant) metabolites [5–7], foodstuffs [8], and environmental pollutants [9]. The exact mechanism of retention in HILIC has been intensively investigated and it is thought to be rather complex. The currently accepted mechanism is a combination of (*i*) partitioning processes of the analytes between a water-poor organic mobile phase and a water-rich layer absorbed on a polar stationary-phase material [10], and (*ii*) electrostatic interactions between the analytes and the stationary-phase surface [11]. Therefore, HILIC can best be described as a mixed-mode retention mechanism.

To describe retention in HILIC, several retention models have been investigated. The model most commonly used in reversed-phase LC (RPLC) involves a linear relationship between the logarithm of the retention factor (k) and the volume fraction of strong solvent (φ). When a linear gradient is used in RPLC this results in so-called linear-solvent-strength conditions [12]. Already in 1979, LSS conditions have been studied and described in detail by Snyder et al. [12] and equations were also derived for situations in which analytes elute before the gradient commences or after it has been completed [12,13]. However, due to the mixed-mode retention mechanism, this linear model may be less suitable to accurately model retention in HILIC.

To describe retention more accurately across a wider φ-range, Schoenmakers et al. introduced a quadratic model [13], including relations for the retention factor for analytes eluting within and after a gradient. However, an error function was required to allow partial integration of the gradient equation. This is an impractical aspect of the relationship. Moreover, the model may show deviations from the real values when predicting outside the scanning range. An empirical model proposed by Neue and Kuss circumvented the integration problems, allowing analytical expressions to

be obtained for retention under gradient conditions [14]. A retention model based on surface adsorption has also been proposed for HILIC, predicting retention across narrow ranges of water concentrations in the eluent. To account for the observed mixed-mode behaviour found in HILIC, Jin et al. introduced a three-parameter model [15], which was found to precisely describe retention factors in isocratic mode [16]. However, similar to the quadratic model, integration of the corresponding gradient equation was significantly complicated, involving a gamma function and potentially yielding complex numbers.

For efficient method development, an underlying accurate description of the retention mechanism is crucial. Because gradient elution is an essential tool for analysing or scanning samples, accurate description of gradient-elution patterns is also essential. Computer-aided-optimization tools, such as Drylab [17] for 1D-LC, or PIOTR [18] for 1D and 2D-LC, utilize the concept of so-called "scouting" or "scanning" runs to establish retention parameters [19], from which the optimal conditions and the optimal chromatogram can be predicted. Method development in HILIC following these principles has extensively been studied by Tyteca et al. [20,21]. However, the currently employed retention models for HILIC do not allow accurate prediction of retention times of analytes eluting during or after the completion of gradients based on a very limited number of scouting measurements. This hampers the application of such optimization tools for HILIC.

In this work, we present the results of an evaluation study of each of the five above-listed models for predicting retention times in gradient-elution HILIC based on a limited number of scouting runs for a wide range of applications. First, the equation for each retention model is addressed in the context of gradient-elution chromatography. We use Simpson's Rule [22] to approximate the integration of resulting gradient equations when an exact solution does not exist, i.e. for the quadratic and mixed-mode models. The performance of each of these models in computer-aided method-development programs is assessed.

2. Theory

In the case that a solute elutes before the start of the gradient, the retention time ($t_{R,before}$) can be calculated from

$$t_{R,before} = t_0(1 + k_{init}) \tag{1}$$

where k_{init} depicts the retention factor at the start of the gradient and t_0 the column dead time. The general equation of linear gradients allows calculation of the retention time if a compound elutes during the gradient

$$\frac{1}{B} \int_{\varphi_{init}}^{\varphi_{init}+B(t_R-\tau)} \frac{d\varphi}{k(\varphi)} = t_0 - \frac{t_{init} + t_D}{k_{init}} \tag{2}$$

In this equation $k(\varphi)$ is the retention model, denoting the variation of the retention factor k with the composition parameter φ. The change in φ as a function of time (i.e. the slope of the gradient) is depicted $B\varphi = \varphi_{init} + Bt$ and τ is the sum of the system dwell time t_D, the waiting time before the gradient is programmed to start t_{init}, and t_0 ($\tau \approx t_D + t_{init} + t_0$). For useful application of gradient-elution retention prediction models in real cases, it is essential that the retention time cannot only be established if the analyte elutes during the gradient, but also if it elutes after the gradient is completed. In this case, the retention is obtained by integrating the retention model in the following equation

$$\frac{1}{B} \int_{\varphi_{init}}^{\varphi_{final}} \frac{d\varphi}{k(\varphi)} + \frac{t_R - \tau - t_G}{k_{final}} = t_0 - \frac{t_{init} + t_D}{k_{init}} \tag{3}$$

Here, t_G represents the gradient time. The application of some of the proposed HILIC retention models is complicated, because the integrals in Eqs. (2) and (3) cannot be analytically solved. The application of each of the HILIC models for gradient-elution separations is the main topic of this paper and this will be described in detail in the following sections.

2.1. Exponential model

In the exponential model (Eq. (4)), k_0 accounts for the extrapolated retention factor for φ = 0 and S denotes the change in the retention factor with increasing mobile phase strength.

$$\ln k = \ln k_0 - S\varphi \tag{4}$$

This equation is often referred to as the linear-solvent-strength (LSS) equations, because it corresponds to LSS conditions in combination with linear gradients ($\varphi = \varphi_{init} + Bt$). Schoenmakers et al. derived equations for a compound eluting during ($t_{R,gradient}$) and after ($t_{R,after}$) the gradient [13].

$$t_{R,gradient} = \frac{1}{SB} \ln \left\{ 1 + SB \cdot k_{init} \left[t_0 - \frac{t_{init} + t_D}{k_{init}} \right] \right\} + \tau \tag{5}$$

$$t_{R,after} = k_{final} \left(t_0 - \frac{t_D + t_{init}}{k_{init}} \right) - \frac{1}{BS} \left(1 - \frac{k_{final}}{k_{init}} \right) + t_G + \tau \tag{6}$$

Here, k_{final} represents the retention factor at the end of the gradient and t_G the duration of the gradient.

2.2. Neue-Kuss empirical model

The empirical model introduced by Neue and Kuss [14] is given by

$$\ln k = \ln k_0 + 2 \ln(1 + S_2\varphi) - \frac{S_1\varphi}{1 + S_2\varphi} \tag{7}$$

where the coefficients S_1 and S_2 represent the slope and curvature of the equation, respectively. Integration of the gradient equation yields

$$t_{R,gradient} = \frac{\ln F}{B(S_2 - S_1 \ln F)} - \frac{\varphi_{init}}{B} + \tau \tag{8}$$

with F defined as

$$F = S_2 B k_0 \left(t_0 - \frac{t_{init} - t_D}{k_{init}} \right) + e^{\frac{S_2\varphi_{init}}{1-S_1\varphi_{init}}} \tag{9}$$

Similarly, introducing Eq. (7) into Eq. (3) and rewriting yields

$$t_{R,after} = k_{final} \left(t_0 - \frac{t_{init} + t_D}{k_{init}} \right. $$
$$\left. + S_2 B k_0 (e^{\frac{S_2\varphi_{init}}{1+S_1\varphi_{init}}} - e^{\frac{S_2\varphi_{final}}{1+S_1\varphi_{final}}}) \right) + t_G + \tau \tag{10}$$

2.3. Adsorption model

The adsorption model is based on confined surface adsorption as used in normal-phase chromatography and is given by

$$\ln k = \ln k_0 - n \ln \varphi \tag{11}$$

where n depicts the ratio of surface areas occupied by a water and a solute molecule [16]. In the events that the compound elutes during or after the gradient retention can be calculated from [23]

$$t_{R,gradient} = \frac{\left[k_0 \left(t_0 - \frac{t_{init}+t_D}{k_{init}} \right) B(n+1) + \varphi_{init}^{n+1} \right]^{\frac{1}{n+1}}}{B}$$
$$- \frac{\varphi_{init}}{B} + \tau \tag{12}$$