Journal of Chromatography A, 1409 (2015) 79–88

Contents lists available at ScienceDirect

Journal of Chromatography A

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

A chromatographic objective function to characterise chromatograms with unknown compounds or without standards available

T. Alvarez-Segura, A. Gómez-Díaz, C. Ortiz-Bolsico, J.R. Torres-Lapasió*, M.C. García-Alvarez-Coque

Departament de Química Analítica, Universitat de València, c/ Dr. Moliner 50, 46100 Burjassot, Valencia, Spain

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form 3 July 2015
Accepted 6 July 2015
Available online 9 July 2015

Keywords:
Liquid chromatography
Complex samples
Unknown compounds
Chromatographic objective function
Peak count
Resolution

ABSTRACT

Getting useful chemical information from samples containing many compounds is still a challenge to analysts in liquid chromatography. The highest complexity corresponds to samples for which there is no prior knowledge about their chemical composition. Computer-based methodologies are currently considered as the most efficient tools to optimise the chromatographic resolution, and further finding the optimal separation conditions. However, most chromatographic objective functions (COFs) described in the literature to measure the resolution are based on mathematical models fitted with the information obtained from standards, and cannot be applied to samples with unknown compounds. In this work, a new COF based on the automatic measurement of the protruding part of the chromatographic peaks (or peak prominences) that indicates the number of perceptible peaks and global resolution, without the need of standards, is developed. The proposed COF was found satisfactory with regard to the peak purity criterion when applied to artificial peaks and simulated chromatograms of mixtures built using the information of standards. The approach was applied to mixtures of drugs containing unknown impurities and degradation products and to extracts of medicinal herbs, eluted with acetonitrile–water mixtures using isocratic and gradient elution.

1. Introduction

A main goal in the development of an analytical procedure in liquid chromatography (LC) is getting the maximal resolution through the fine-tuning of the experimental conditions, with the assistance of an optimisation protocol. When an analytical sample contains a large number of compounds, the application of trial and error assays to get complete resolution can be time-consuming, expensive, or even unfeasible and, in addition, without any guarantee of success. Computer-based methodologies are currently the most efficient tools to search the optimal separation conditions in LC [1–4]. A practical way for determining such conditions is the use of global measurements to appraise the separation quality, using chromatographic objective functions (COFs) [5–8].

Several approaches have been proposed to find out the conditions offering the best separation in situations of extremely low chromatographic resolution, where classical COFs are not able to detect even baseline resolved peaks. These approaches are based on counting the peaks in the chromatogram [9–12], and are focused on the well resolved peaks, in contrast to conventional COFs that attend mainly to those least resolved. This means that they are oriented to quantify the degree of success in the separation, and not the failure (as classical COFs do). The most popular of such approaches was proposed by Berridge [9], who defined a COF combining the number of detected peaks, a term that accounts for the Snyder's resolution between adjacent peak pairs (R_S) (see Section 2.1), and a term penalising long analysis times, in a weighted summation.

More recently, we proposed an approach that allows the optimisation of the resolution level of chromatograms in cases where conventional global criteria, such as the worst resolved peak pair or the product of elementary resolutions are not able to detect any separation [10,12]. The approach makes use of a function that counts the number of peaks exceeding a given threshold of peak purity (i.e. the well resolved peaks). The experimental conditions able to resolve the same amount of peaks were discriminated by either quantifying the partial resolution of those peaks that exceeded the established threshold, or by improving the separation of peaks below the threshold.

A higher level of complexity in the optimisation of the chromatographic resolution is represented by samples containing unknown

* Corresponding author. Tel.: +34 963544014; fax: +34 963544436.
E-mail address: jrtorres@uv.es (J.R. Torres-Lapasió).

http://dx.doi.org/10.1016/j.chroma.2015.07.022
0021-9673/

compounds (some or even all), or without standards available. This is the case of chromatographic fingerprints [11,13–15]. Most COFs reported in the literature are based on mathematical models that should be fitted with the chromatographic information obtained from standards through design of experiments. Therefore, they cannot be applied to this type of sample, or are hardly adapted in situations when only a few standards are available. Trying to find a solution to optimise the resolution of fingerprints of environmental samples, Duarte and Duarte suggested a measurement based on a valley-to-peak ratio to assess the resolution [11]. The approach presents, however, certain limitations when the chromatograms contain peaks showing far different heights and/or significant baseline noise. In this work, we propose a different COF to face these situations, which is based on the automatic measurement of the peaks, which is applicable to experimental highly complex chromatograms. The approach, which is related to the peak purity, attends to the protruding part of each visible peak with regard to the valleys that delimit it. The performance of the new COF is evaluated with regard to the peak purity for cases of study of increasing complexity.

2. Theory

2.1. Classical approaches to measure the elementary resolution

Elementary resolution criteria describe the separation between two solutes exhibiting consecutive peaks, or each particular solute from all other solutes in the sample [1–8,16–20]. For this purpose, several COFs of diverse complexity have been proposed. The simplest one, which only considers the peak position, is the selectivity factor:

$$\alpha_{i,i+1} = \frac{k_{i+1}}{k_i} \tag{1}$$

where k_i and k_{i+1} (with $k_{i+1} > k_i$) are the retention factors of two consecutive peaks. Other COFs consider the width of the individual peaks, or the full profile including the size. The Snyder's resolution R_S is the most popular [1,7,8]:

$$R_S = \frac{t_{R,i+1} - t_{R,i}}{B_i + A_{i+1}} \tag{2}$$

where $t_{R,i+1}$ and $t_{R,i}$ are the retention times of two consecutive peaks, and B_i and A_{i+1} are the right and left half-widths for peaks i and $i+1$, respectively.

Other interesting COFs measure different types of valley-to-peak ratios, which require locating a point in the chromatogram related to the valley between two consecutive peaks (which has been called the "valley point") (Fig. S1a) [20]. This point is not necessarily the minimum in the valley. The following equation is applied:

$$v_{i,i+1} = 1 - \frac{h_v}{h_T} \tag{3}$$

where h_v is the signal height (measured from the baseline) at the time of the valley point, and h_T the height measured at this time from the baseline to the straight-line obtained by joining the maxima of the two neighbouring peaks. Fig. S1a illustrates three types of valley-to-peak ratio measurements. In the first measurement (using h_{v1} and h_{T1}), the valley point is straightforwardly found by locating the minimal height between the two consecutive peaks [21]. This definition is similar to the ratio proposed by Duarte and Duarte for the quantification of the resolution of soil fingerprints [11].

2.2. Peak purity

Different COFs may score the same chromatogram differently, which brings as a consequence the selection of different chromatograms as optimal. However, the analyst may easily discriminate the resolution performance and find some chromatograms more suitable than others. The reason of these differences is that most COFs are too simple to account for all the features that the analyst considers in the assessments of the separation quality. However, some COFs may satisfactorily correlate to the analyst's appraisal of good resolution. An example of such a COF is the peak purity (Fig. S1b) [1–3,19]:

$$p_i = 1 - \frac{o_i'}{o_i} \tag{4}$$

where o_i' is the area under the peak overlapped by a hypothetical chromatogram built with the peaks of the accompanying compounds in the sample, and o_i the total peak area. A similar COF was proposed early by Schoenmakers [1], although it was scarcely applied since it required the knowledge of both the position and profile for each peak in the chromatogram. Therefore, its calculation was laborious and needed numerical computation. Nowadays, this is not a problem.

The peak purity offers unique features:

(i) It correlates with the information the analyst wishes: the interference level.
(ii) It accounts the overlapping in the whole peak providing a realistic evaluation of the separation, since the full signal including profile and size is monitored.
(iii) It is a normalised function, which facilitates the combination of the peak measurements in the whole chromatogram, and the inclusion of other quality criteria.
(iv) It qualifies individual peaks, instead of peak pairs, which makes peak weighting and exclusion simpler, and avoids identity problems related to peak reversals. This feature has enabled the proposal of new optimisation strategies, based on information not accessible using other COFs.

2.3. Proposal of a new COF

The objective of this work is the proposal of a new COF able to score the performance of chromatograms of samples where some or all compounds are unknown, or where standards are not available, making the application of the peak purity criterion unfeasible. However, the new COF is also valid for situations where all compounds are known. We have called the new COF "peak prominence", since it quantifies the protruding part of each peak in a chromatogram with regard to the valleys that delimit it, or incidentally the baseline (see Fig. 1a and Fig. S2 in the Supplementary material). The resolution level can be calculated as the prominence ratio:

$$\mathrm{pr}_i = \frac{a_{\mathrm{pr},i}}{a_i} \tag{5}$$

where $a_{\mathrm{pr},i}$ is the area of the protruding part of the peak (Fig. 1a), and a_i represents its total area (Fig. 1b). Alternatively, a peak height ratio can be used.

There is a significant difference between the new approach and the conventional valley-to-peak ratio criteria described in Section 2.1. The latter depend on the measurement of two neighbouring peaks and the valley in between, whereas the proposed COF is based on the measurement of one peak between two valleys (or the baseline). The peak prominence shares the good features of the peak purity: it is sensitive to signal profile and size, it is an easily interpretable normalised measurement, and it scores individual peaks, which facilitates the combination of the elementary