Journal of Chromatography A, 1491 (2017) 182–194

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Combination of partial least squares regression and design of experiments to model the retention of pharmaceutical compounds in supercritical fluid chromatography☆

Bertyl Andri[a,*], Amandine Dispas[a], Roland Djang'Eing'a Marini[a], Philippe Hubert[a], Patrick Sassiat[b], Ramia Al Bakain[c], Didier Thiébaut[b], Jérôme Vial[b]

[a] University of Liege (ULg), CIRM, Laboratory of Analytical Chemistry, 15 Avenue Hippocrate, B36, B-4000 Liege, Belgium
[b] Department of Analytical, Bioanalytical Sciences and Miniaturization (LSABM), Institute of Chemistry, Biology and Innovation (CBI), ESPCI ParisTech, CNRS UMR 8231, PSL Research University, 10 rue Vauquelin, Paris Cedex 05, France
[c] Department of Chemistry, Faculty of Science, The University of Jordan, P.O. Box 11942, Amman, Jordan

ARTICLE INFO

Article history:
Received 30 November 2016
Received in revised form 10 February 2017
Accepted 13 February 2017
Available online 16 February 2017

Keywords:
Supercritical Fluid Chromatography
Chemometric approach
Chromatographic behaviour
LSER descriptors
Retention prediction

ABSTRACT

This work presents a first attempt to establish a model of the retention behaviour for pharmaceutical compounds in gradient mode SFC. For this purpose, multivariate statistics were applied on the basis of data gathered with the Design of Experiment (DoE) methodology. It permitted to build optimally the experiments needed, and served as a basis for providing relevant physicochemical interpretation of the effects observed. Data gathered over a broad experimental domain enabled the establishment of well-fit linear models of the retention of the individual compounds in presence of methanol as co-solvent. These models also allowed the appreciation of the impact of each experimental parameter and their factorial combinations. This approach was carried out with two organic modifiers (i.e. methanol and ethanol) and provided comparable results. Therefore, it demonstrates the feasibility to model retention in gradient mode SFC for individual compounds as a function of the experimental conditions. This approach also permitted to highlight the predominant effect of some parameters (e.g. gradient slope and pressure) on the retention of compounds.

Because building of individual models of retention was possible, the next step considered the establishment of a global model of the retention to predict the behaviour of given compounds on the basis of, on the one side, the physicochemical descriptors of the compounds (e.g. Linear Solvation Energy Relationship (LSER) descriptors) and, on the other side, of the experimental conditions. This global model was established by means of partial least squares regression for the selected compounds, in an experimental domain defined by the Design of Experiment (DoE) methodology. Assessment of the model's predictive capabilities revealed satisfactory agreement between predicted and actual retention (i.e. $R^2 = 0.942$, slope = 1.004) of the assessed compounds, which is unprecedented in the field.

1. Introduction

Supercritical Fluid Chromatography (SFC) is bit-by-bit leaving its "niche application only" status [1] to become a more universal and truly complementary technique to LC and GC [2]. Since the robustness of the instruments also drastically improved over the recent times, routine application of the technique in highly regulated frameworks (e.g. GxP guidelines applied in the pharmaceutical field) is therefore worth considerations [3–8].

So, increased interests are nowadays devoted to the technique, which uses packed columns and carbon dioxide blended with organic co-solvent as mobile phase [9]. Thus, SFC is being recognized as fast, green, efficient, and orthogonal to liquid chromatography in terms of selectivity. It is also compatible with numerous detectors [10]. This explains why SFC was gradually applied to the analysis of different samples issued from various domains [11–20].

However, in spite of the recent improvements and fundamental SFC works, it appears that no "universal model" of the retention already exists to help the SFC method development [2]. Different

☆ Selected paper from 31st International Symposium on Chromatography (ISC2016), 28 August-1 September 2016, Cork, Ireland.

* Corresponding author.
E-mail address: bertyl.andri@ulg.ac.be (B. Andri).

http://dx.doi.org/10.1016/j.chroma.2017.02.030
0021-9673/

recent works demonstrated the interest devoted to this particular topic [21–24]. Given the numerous possible applications of SFC, it is obvious that the establishment of such model should take into account the most influential experimental parameters (typically pressure, temperature and solvent gradient slope) over a broad range of conditions. To address such study and properly define the experiments required (number of experiments and levels of studied factors), chemometric tools and especially Design of Experiments (DoE) are interesting [25,26]. Indeed, it allows building knowledge on case study in the most efficient way (i.e. taking the lesser amount of experiments, screening relevant experiments, taking account of experimental variation/uncertainty...). Application of DoE is not recent in the field of SFC and various examples have already issued [27–35]. However, most of the DoE published aim at the optimization of method parameters to e.g. address the separation of given analytes, which is not the purpose of the present study. Indeed, we decided to take advantage of this methodology to provide relevant modifications of the experimental parameters. So, each of the experiments defined could be considered as a different chromatographic system.

For the present work, a set of 15 compounds (Fig. 1) was intentionally selected to cover a broad range of known physicochemical properties (polarity, acido-basic properties...) and used as probes of the retention during the experiments. All these compounds are originating from the pharmaceutical field, where a large variety of molecules can be encountered and great interests are devoted towards SFC [17].

Due to the multivariate nature of the data (i.e. simultaneous modification of the factor's levels) and the potential interactions among the factors (i.e. pressure and temperature may impact the fluid's density, which in turn may have an effect on the elution of analytes), multivariate data analyses are required to study the responses recorded (i.e. a retention time) [36]. Multiple Least Square Linear Regression (MLR) was first applied individually to evaluate the possibility to establish retention models for each probe. Indeed, these models had to describe the link between the experimental conditions and the retention observed. So, such models should permit to appreciate the impact of the different experimental parameters on the retention of the probes. At first, this individual modelling was performed in presence of methanol and was then successfully repeated (i.e. repetition of the whole DoE) in presence of ethanol as organic modifier.

As it was possible to build individual models, the establishment of a global model of the retention for the whole set of compounds was therefore considered. This model took benefit of the observations made for the individual models (e.g. most important parameters, chromatographic behaviour...). It also had to take account of the physicochemical properties of the compounds. Indeed, given the global scope of the approach, it was mandatory to take into account the properties of each compound to explain the individual retentions observed. So, neutral LSER descriptors were used as indicators of the physicochemical properties of the compounds [37,38] to establish the models. These descriptors are determined according to multi-linear equations based on partition or retention data of compounds in different "systems" (i.e. partition system or chromatographic system). Each term of these equations is associated to a coefficient related to the system and represents given properties of the compound i.e. its polarizability (S), the effects of π and n-electrons on its polarizability (E), its capacity to create acidic or basic H bonds (A & B), its molecular volume (V) and its partition between gas and hexadecane (L). To account for the potential ionization of the pharmaceutical compounds employed, ionic conversions of the LSER descriptors were also employed to build the models [39–41].

Thanks to the extensive work of Lesellier and West, many valuable applications of LSER descriptors were already released in the field of SFC. To cite only a few, it therefore served as a basis to understand effects of mobile phase composition [42], classify analytical columns [43–45], predict separation of herbicides in SFC [22] and recently, to assess orthogonality of stationary phases with the aim of impurity profiling [41]. Other authors also successfully applied Abraham's solvation parameters to the field of SFC. Indeed, Bui et al. determined neutral LSER descriptors for a large number of pharmaceutical compounds [46].

Partial Least Square regression (PLS) was employed for the global data treatment. Indeed, PLS analysis can find the hidden structure of the dataset and therefore extract its meaningful information [47,48]. So, PLS allows the modelling of multiple responses and is able to deal with the multicolinearity (non independence of predictor variables) inside a dataset. This technique is therefore suitable for modelling based on dataset, which simultaneously contains variable experimental conditions and constant physicochemical data.

In this work, we intended to study the retention of selected probes on a 2-Ethyl Pyridine (2-EP) column. This choice was made on the basis of previous work [49]. This stationary phase was found to elute the whole set of probe compounds with comparable performances (e.g. peak capacities) without any additive required.

Given the pharmaceutical interests devoted to the use of modern SFC, working under gradient mode elution was mandatory to reflect nowadays use of modern SFC.

So, this work mainly aims at the establishment of individual and global retention models in gradient mode SFC. In order to undertake this task, chemometric tools (e.g. DoE, MLR, LSER descriptors, PLS regression...) were employed and allowed the establishment of a successful global approach to the modelling of retention in SFC.

2. Material and methods

2.1. SFC apparatus and analytical column

An Agilent 1100 series system (Agilent Technologies, Waldbronn, Germany) equipped with an auxiliary Aurora A5 fusion SFC module (Aurora SFC systems Inc., Redwood City, CA, USA) was employed for the experiments. This system was composed of a G1312A Binary pump (high pressure mixing), a G1314A Variable Wavelength detector with 14 μL high-pressure flow cell (10 mm path), a G1329A autosampler with a 5.0 μL fixed-loop and thermoregulated with a G1330B ALS Therm cooling unit. Depending of the experimental requirement, the columns was heated with a Croco-Cil RS 232 (Cluzeau Info Labo, Sainte-Foy, France) column oven or cooled in bath controlled via a Minichiller (Huber Kältemaschinenbau GmbH, Offenburg, Germany). Temperature of the coolant bath was also externally monitored with KM 12 lab thermometers (Comark Instruments Ltd, Norwich, UK).

The chromatographic column was purchased from Princeton SFC (Cranbury, NJ, USA) and consisted of a 2-ethylpyridine bonded phase: 2-EP (3.0 × 100 mm; 3.0 μm). The chromatographic system was operated with Agilent Chemstation software (Rev B.0402) with special add-on to control the auxiliary SFC module. A custom-made injector macro was written for needle wash, partial loop injection and offline loop loading in SFC mode. Data were collected and processed according to the Performance Report mode proposed by the Chemstation©.

2.2. Standards and reagents

4-Aminophenazone (99.5%), Aceclofenac (99.9%), Atenolol (99.1%), Bifonazole (99.3%) Carvedilol (99.9%), Diclofenac (99.5%), Ethylbenzene (99.7%), Etodolac (99.6%), Haloperidol (99.7%), Hydrocortisone (99.8%), Ipriflavone (99.8%), Naproxen (99.6%),