Analytica Chimica Acta 983 (2017) 67–75

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Targeted analyte deconvolution and identification by four-way parallel factor analysis using three-dimensional gas chromatography with mass spectrometry data

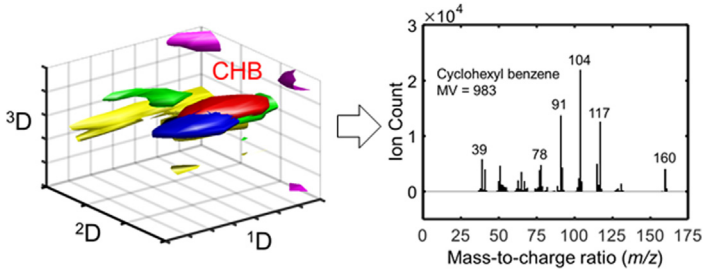
Nathanial E. Watson, Sarah E. Prebihalo, Robert E. Synovec*

Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98195, USA

HIGHLIGHTS

- A four-way PARAFAC method is presented to analyze GC^3-TOFMS data.
- The method overcomes challenges of reducing rank prior to PARAFAC analysis.
- All 42 target analyte spectra were successfully resolved and identified.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 17 February 2017
Received in revised form
1 June 2017
Accepted 6 June 2017
Available online 19 June 2017

Keywords:
Gas chromatography
Time-of-flight mass spectrometry
Parallel factor analysis
PARAFAC
Four-way data analysis
Multidimensional chromatography

ABSTRACT

Comprehensive three-dimensional gas chromatography with time-of-flight mass spectrometry (GC^3-TOFMS) creates an opportunity to explore a new paradigm in chemometric analysis. Using this newly described instrument and the well understood Parallel Factor Analysis (PARAFAC) model we present one option for utilization of the novel GC^3-TOFMS data structure. We present a method which builds upon previous work in both GC^3 and targeted analysis using PARAFAC to simplify some of the implementation challenges previously discovered. Conceptualizing the GC^3-TOFMS instead as a one-dimensional gas chromatograph with GC × GC-TOFMS detection we allow the instrument to create the PARAFAC target window natively. Each first dimension modulation thus creates a full GC × GC-TOFMS chromatogram fully amenable to PARAFAC. A simple mixture of 115 compounds and a diesel sample are interrogated through this methodology. All test analyte targets are successfully identified in both mixtures. In addition, mass spectral matching of the PARAFAC loadings to library spectra yielded results greater than 900 in 40 of 42 test analyte cases. Twenty-nine of these cases produced match values greater than 950.

1. Introduction

The body of knowledge surrounding comprehensive two-dimensional (2D) gas chromatography (GC × GC) is seemingly expanding at an ever-growing rate [1]. Moving from a theoretical possibility pioneered during the latter half of the twentieth century [2] to a novelty in the early 1990s [3], separation of complex mixtures is now almost unimaginable without two serially coupled GC columns [1]. Creative options for column selection and coupling are still being developed. In addition, GC × GC has been joined to most GC detectors, be they single- or multi-channel. Indeed, the combination of GC × GC with time-of-flight mass spectrometry

* Corresponding author.
E-mail address: synovec@chem.washington.edu (R.E. Synovec).

http://dx.doi.org/10.1016/j.aca.2017.06.017
0003-2670/

(TOFMS) has proven particularly interesting (GC × GC-TOFMS) [4–7].

When GC × GC-TOFMS is implemented, the resultant three-dimensional (3D) data enable application of various chemometric deconvolution techniques providing mathematical resolution [8,9]. Among these methods, the Parallel Factor Analysis (PARAFAC) algorithm has proven particularly useful [10–16]. PARAFAC models the data array as the linear combination sum of signals for each eluting analyte described by the outer product of chromatographic and mass spectrometric vectors in each of the acquired dimensions for an analyst determined number of components. PARAFAC requires, at a minimum, a 3D array which implicates GC × GC-TOFMS as one of the few instruments which create data of this structure naturally.

Previously, comprehensive 3D gas chromatography (GC × GC × GC or GC^3) was described as an alternate gas chromatographic configuration which enables implementation of PARAFAC on the native data [17,18]. GC^3 with PARAFAC deconvolution was shown to provide a benefit to both the detection limit and provide the ability to resolve convoluted elution signals. Recently the GC^3 instrumental design was improved to include a TOFMS as the detector creating a 3D gas chromatograph with mass spectrometric detection (GC^3-TOFMS) [19]. The principal benefits being both the added mass spectrometric selectively and the ability to identify analyte peaks based on the acquired mass spectrum. This advanced instrumental platform natively creates four-dimensional data (4D) which permits new options for visualizing and chemometrically interrogating the data array provided by complex samples.

Four-way PARAFAC, in itself, is not novel [9,20]. Four-way data can be simply created by stacking replicates of 3-way data with varying concentration profiles. The results produce a reconstruction of the three instrumental dimensions and the fourth dimension contains the concentration information in each of the replicates. This procedure of replicate stacking is even more common in 2D methods (e.g. one chromatographic dimension with multichannel detection) where the replicate stacking enables access to PARAFAC which has a minimum dimensionality of three. Instruments, like the GC^3-TOFMS, which produce native four-way data are significantly less typical, but the development of such has the potential to address emerging challenges in chemical analysis.

Herein we present a targeted method for analysis of four-way data, specifically data generated by GC^3-TOFMS. Embracing the native dimensionality of the data produced by this instrument, we explore a technique which overcomes one of the inherent challenges of applying chemometrics to more traditional GC × GC-TOFMS data: assigning a target window for application of PARAFAC [14–16]. For instance, if one were inclined to apply PARAFAC to a whole dataset, the computational intensity of the technique would quickly overcome the available computing capacity requiring that subordinate windows be selected within the dataset. In GC × GC-TOFMS this window determination step is one of the more time consuming and subjective steps of the analysis. GC^3-TOFMS enables a new paradigm by which each modulation of the first "primary" chromatographic dimension (^{1}D) produces a complete GC × GC-TOFMS chromatogram which is suitable, as a whole, for PARAFAC analysis. More simply, we suggest this method could be envisioned as a primary GC separation coupled with GC × GC-TOFMS detection. By implementing a non-polar stationary phase for the ^{1}D separation, a complex sample is subjected to an initial separation roughly equivalent to a fractional distillation whereby each fraction is analyzed by GC × GC-TOFMS. The "fractions" are the modulations from the ^{1}D separation onto the ^{2}D separation, and so on. This paradigm is demonstrated with PARAFAC, whereby data portions amenable to third order chemometrics are produced in a natural fashion and should be applicable to other chemometric methods [21].

In this study, a stack of seven consecutive GC × GC-TOFMS ^{1}D modulation "slabs" produce a window across ^{1}D without any need to window the second GC separation (^{2}D), third GC separation (^{3}D) and mass spectrometric dimensions. The GC × GC separations in each slab, i.e., ^{2}D by ^{3}D separations, are designed to provide a two-dimensional chromatographic peak capacity $n_{c,2D}$ of ~15–20 that is ideally suited to both provide ample chemical selectivity concurrent with keeping the rank suitably low for successful implementation of PARAFAC. The temporal width of the seven ^{1}D modulation window is defined to fully encompass the elution event for a target analyte, allowing for some ambiguity in ^{1}D retention time selection. Standards or retention indices [22] provide an excellent method for approximate ^{1}D target analyte window location assignment prior to PARAFAC analysis. Herein, two samples are submitted to study this approach for targeted analyte discovery and deconvolution: a standard mixture of 115 components and a diesel fuel spiked with several non-native compounds [19]. Analytical success is determined by both visual verification of the chromatographic loadings and mass spectral matching to a library mass spectrum [14,23].

2. Experimental

2.1. Data collection

The GC^3-TOFMS was operated as previously described [19] and briefly recalled here. A Pegasus 4D GC × GC-TOFMS (LECO Corporation, St. Joseph, MI) with an integrated Agilent 6890N Gas Chromatograph (Agilent Technologies, Santa Clara, CA, USA) was modified to produce a GC^3-TOFMS instrument as shown in Fig. 1. One high-speed, six port diaphragm valve (Valco Instruments Company Inc, Houston, TX, USA), upgraded by the manufacturer to operate at a maximum temperature of 325 °C, and fitted with a 5 μL sampling loop was installed in the GC oven. The high-temperature valve was utilized as the modulator between the ^{1}D and ^{2}D separations, and the stock thermal modulator was implemented between the ^{2}D and ^{3}D separations. The ^{3}D column was contained within the secondary oven in the commercial instrumental

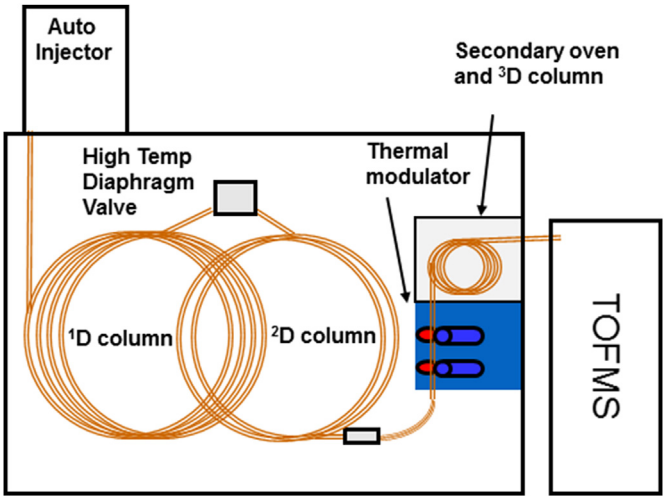


Fig. 1. Schematic of the major components of the GC^3-TOFMS instrument. A high-temperature diaphragm valve was utilized as the modulator to interface the ^{1}D column separation to the ^{2}D column separation. The thermal modulator was used interface the ^{2}D column separation to the ^{3}D column separation.