A modified multiscale peak alignment method combined with trilinear decomposition to study the volatile/heat-labile components in *Ligusticum chuanxiong Hort - Cyperus rotundus rhizomes* by HS-SPME-GC/MS

Min He[a,*], Pan Yan[a], Zhi-Yu Yang[a], Zhi-Min Zhang[b], Tian-Biao Yang[a], Liang Hong[a]

[a] *Department of Pharmaceutical Engineering, School of Chemical Engineering, Xiangtan University, Xiangtan 411105, People's Republic of China*
[b] *College of Chemistry and Chemical Engineering, Central South University, Changsha 410083, People's Republic of China*

ARTICLE INFO

ABSTRACT

Head Space/Solid Phase Micro-Extraction (HS-SPME) coupled with Gas Chromatography/Mass Spectrometer (GC/MS) was used to determine the volatile/heat-labile components in *Ligusticum chuanxiong Hort - Cyperus rotundus rhizomes*. Facing co-eluting peaks in **k** samples, a trilinear structure was reconstructed to obtain the second-order advantage. The retention time (RT) shift with multi-channel detection signals for different samples has been vital in maintaining the trilinear structure, thus a modified multiscale peak alignment (mMSPA) method was proposed in this paper. The peak position and peak width of representative ion profile were firstly detected by mMSPA using Continuous Wavelet Transform with Haar wavelet as the mother wavelet (Haar CWT). Then, the raw shift was confirmed by Fast Fourier Transform (FFT) cross correlation calculation. To obtain the optimal shift, Haar CWT was again used to detect the subtle deviations and be amalgamated in calculation. Here, to ensure there is no peaks shape alternation, the alignment was performed in local domains of data matrices, and all data points in the peak zone were moved via linear interpolation in non-peak parts. Finally, chemical components of interest in *Ligusticum chuanxiong Hort - Cyperus rotundus rhizomes* were analyzed by HS-SPME-GCMS and mMSPA-alternating trilinear decomposition (ATLD) resolution. As a result, the concentration variation between herbs and their pharmaceutical products can provide a scientific basic for the quality standard establishment of traditional Chinese medicines.

1. Introduction

The volatile components in food/herb are usually analyzed by Head Space-Solid Phase Micro Extraction coupled with Gas Chromatography/ Mass Spectrometer detection (HS-SPME-GC/MS) [1,2], a rapid, high-efficient and solvent-less method. *Ligusticum chuanxiong Hort - Cyperus rotundus rhizomes* [3,4], the common couplet medicines in herb prescriptions, is such an example, because the main active ingredient in them is volatile or semi-volatile. In addition, Z-Ligustilide, Cyperone and other bioactive components with thermo-optical instability, should be rapidly prepared through HS-SPME technology before a qualitative and quantitative determination, which would make the test results more accurate.

HS-SPME-GC/MS can bring abundant chemical information, as well as co-eluting compounds and unexpected interferences in a series of complex samples. Many chemometric methods are often used to assist in qualitative or quantitative analysis when the experimental conditions are not sufficient for complete separation of complex components in

traditional Chinese medicine. In recent years, the second-order calibration methods can play a unique "second-order advantage" in dealing with matrix generated from hyphenated GC–MS apparatus or other instruments [5,6], and obtain the components of interest in a mixture containing unknown interferences by smart "mathematical separation". The common second-order calibration methods represented by Generalized Rank Annihilation Method (GRAM) [7], Direct Trilinear Decomposition method (DTLD) [8], Parallel Factor analysis (PARAFAC) [9,10], PARAFAC2 [11,12], Multivariate Curve Resolution-Alternating Least-Squares (MCR-ALS) [13], Alternating Trilinear Decomposition (ATLD) [14], Bilinear Least Squares/Residual Bilinearization (BLLS/RBL) [15] etc. Given that in a test, several samples are often tested and compared at the same time, and thus the datasets were compiled into a three-dimensional (3D) array resolved by trilinear decomposition methodologies. ATLD is such an example, which based on Moore–Penrose generalized inverse with singular value decomposition accompanied with many advantages [14,16].

Theoretically, each component from different samples should have

* Corresponding author.
  E-mail address: dahai8214813@gmail.com (M. He).

the same retention time (RT) and peak shape, and only differs in the amounts or intensities. In reality, peak shifts usually be observed among different data matrices, although the experiment minimizes the deviation as much as possible between the results and hypothesis. The RT shift may be attribute to drift in the instrument, temperature/pressure fluctuations, aging of separation columns, especially injection mode. Even in the usage of some automated instruments, the subtle RT shifts may be observed due to inevitable experimental deviation. These shifts in peak positions may exert a detrimental effect on the construction of trilinear array. Therefore, the peak alignment of the data matrices should be performed so that the further trilinear decomposition is ensured. Numerous methods have been proposed to align RT shifts in analytical instruments, including Dynamic Time Warping (DTW) [17], Correlation Optimized Warping (COW) [18], Parametric Time Warping (PTW) [19,20], Peak Alignment using Reduced Set mapping (PARS) [21], Automated Peak Alignment by Beam Search (Auto-PABS) [22], Fuzzy Warping (FW) [23], et al. Among them, DTW is often applied in the industrial batch processes, rather than the chromatographic data. COW and PTW are time consuming due to the parameter optimization, which may change the peak shape. Zheng et al. [24] proposed a method of Chromatogram Alignment via Mass Spectrometry (CAMS), which can make good use of mass spectral information to verify the RT shift alignments in total ion chromatograms (TICs). Zhang et al. [25] brought a practical and handy peak alignment method - alignDE, which is applied to reconstruct the aligned TICs using differential evolution. Zhang et al. [26] also proposed MultiScale Peak Alignment (MSPA) method in 2012, which is a good way to align the change of nonlinear retention time shifts in TICs within an acceptable RT-deviation. In addition, a modified COW algorithm has been presented for successful PARAFAC modelling of LC–MS data [27]. The result of the alignment is much dependent on the choice of benefit function, and no general guidance was given for this choice. Bortolato et al. proposed a parallel factor analysis alignment [28], which can properly align the different data matrix. Yu et al. put forward a new time shift alignment method derived from the well-known Rank Minimization method for aligning chromatographic matrices [29]. The results indicate that the correct RT alignment by this method will be helpful for trilinear decomposition to achieve the second-order advantage. Nevertheless, the existed methods have their own characteristics, new methods need to be further proposed when dealing with multi-channel data matrices with different requirements.

In this study, we proposed a modified MSPA (mMSPA) method, with aims to rapidly align the data matrices in local chromatographic domains. Then ATLD algorithm was employed in resolution of HS-SPME-GC/MS data for complex analytic system, using *Ligusticum chuanxiong Hort - Cyperus rotundus rhizomes* as an example. Through these cases, the importance of these methods in pharmaceutical analysis is illustrated.

2. Experimental and methods

2.1. Experimental materials

Twelve samples of *Ligusticum chuanxiong Hort* were collected in the city of Du-Jiang-Yan, Peng-Zhou and Gan-Zi, the providence of Si-Chuan in 2016, respectively. Seven samples of *Cyperus rotundus rhizomes* were collected in the providence of Hu-Nan, Hu-Bei, Guang-Dong, Guang-Xi, Hai-Nan in 2016, respectively. The crude drugs were air-dried indoors and smashed before SPME enrichment. Alkane standard solutions of C_8–C_{20} (mixture no. 04070) and C_{21}–C_{40} (mixture no. 04071) were purchased from Fluka Chemika (Buchs, Switzerland).

2.2. Extraction procedure

Unprocessed sample: after drying at 40 °C and grinding, *Ligusticum chuanxiong Hort* and *Cyperus rotundus rhizomes* are mixed in proportion to 1:1.

Ethanol extraction: after soaking for 6 h, 30 g of precisely weighted unprocessed herbs with 300 ml 90% Ethanol, were reflux extracted for 1 h. The solution was centrifuged (1500 ×g for 10 min), and finally concentrated under vacuum.

Decoction: 30 g of unprocessed sample was soaked in 300 ml water for 6 h, then kept water boiling for 3 h. The solution was finally concentrated under vacuum.

Supercritical fluid extraction (SFE): an equipment (HA121-50-01 type, China) was used for the extraction of volatile/heat-labile components, designed for sample weight (150 g), working pressures (350 bar) and temperatures (60 °C), respectively. After 2 h, the extraction vessel was then depressurized to collect the product.

2.3. Enrichment of volatile/heat-labile components by SPME

1.5 g of *Ligusticum chuanxiong Hort* or *Cyperus rotundus rhizomes* or the mixed samples was added into a 20-mL glass vial and incubated for 30 min at 70 °C. Then, enrichment of volatile/heat-labile components was performed by 40/60 μm activated carbon/polydimethylsiloxane/divinylbenzene fiber (ACAR/PDMS/DVB, Zhenzheng, China) combined with a manual SPME device, which was equilibrated at 70 °C for 60 min on a heating platform. After extraction, fibers were immediately conditioned in the GC injection port for 5 min before being used for analyses.

2.4. GC–MS analysis

The procedure was performed on a GC-2010 gas chromatograph (Shimadzu, Kyoto, Japan) equipped with a QP2010 mass spectrometer (Shimadzu); A HP-5 capillary column (30 m × 0.25 mm i.d., film thickness of 0.25 μm, Agilent) was used for separation. The following oven temperature program was applied, initiated at 40 °C for 6 min, increased to 90 °C at the rate of 4 °C min^{-1}, then increased to 140 °C at the rate of 1 °C min^{-1} and finally increased to 280 °C at the rate of 8 °C min^{-1} for 3 min. The spectrometer was operated in electron impact mode; the scan range was 33–330 amu; the ionization energy was 70 eV and the scan rate was 0.50 s per scan. Injector, interface and ion source were kept at 280, 260 and 200 °C, respectively. For SPME fibers desorption, a splitless liner tube is selected. The split mode of 1:50 was used for herbs, while the split modes of 1:10 and 1:50 were used for the pharmaceutical products, respectively.

2.5. mMSPA

When facing the chromatographic signals with multi-channel detection in local domain, the mMSPA algorithm can obtain more accurate alignments of sub-matrices of interest. In this procedure, Continuous Wavelet Transform with Haar wavelet as the mother wavelet (Haar CWT) converts the chromatogram into a wavelet space to detect the peak position and peak width. Then Fast Fourier Transform (FFT) cross correlation is adopted to calculate the optimum shifts. At last, the cyclic calculation can be used for simultaneous RT correction of multiple ion chromatograms. The flow chart for the mMSPA algorithm is shown in Fig. 1, which will be discussed below.

2.5.1. Peak detection and peak width calculation

Firstly, one selective ion was observed in mass spectra corresponding to one component, and the corresponding ion profile was picked out from data matrix. Secondly, the peak location and peak width are detected based on a derivative calculation method via Haar CWT [30], and false positive peaks were eliminated by signal-to-noise ratio (SNR) setting.

As seen from Eq. (1), wavelet is a series of wavelet functions $\psi_{a, b}(t)$ deriving from wavelet generating function $\psi(t)$ regulated by scale factor a and shift factor b. The projection process of signal $s(t)$ onto the wavelet function ψ can be described as Eq. (2), and C is a matrix of wavelet