
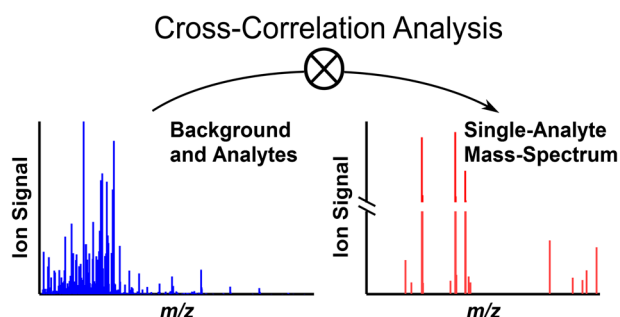


Automatic Analyte-Ion Recognition and Background Removal for Ambient Mass-Spectrometric Data Based on Cross-Correlation

Yi You,¹ Sunil P. Badal,^{1,2} Jacob T. Shelley^{1,2} 

¹Department of Chemistry and Biochemistry, Kent State University, Kent, OH 44242, USA

²Department of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute, Troy, NY 12180, USA



Abstract. Ambient mass spectrometry is a powerful approach for rapid, high-throughput, and direct sample analysis. Due to the open-air desorption and ionization processes, random fluctuations of ambient conditions can lead to large variances in mass-spectral signals over time. The mass-spectral data also can be further complicated due to multiple analytes present in the sample, background-ion signals stemming from the desorption/ionization source itself, and other

laboratory-specific conditions (e.g., ambient laboratory air, nearby hardware). Thus, background removal and analyte-ion recognition can be quite difficult, particularly in non-targeted analyses. Here, we demonstrate the use of a cross-correlation-based approach to exploit chemical information encoded in the time domain to group/categorize mass-spectral peaks from a single analysis dataset. Ions that originate from ambient (or other) background species were readily flagged and removed from spectra; the result was a decrease in mass-spectral complexity by over 70% due to the removal of these background ions. Meanwhile, analyte ions were differentiated and categorized based on their time-domain profiles. With sufficient mass resolving-power and mass-spectral acquisition rate, isolated mass spectra containing ions from the same species in a sample could be extracted, leading to a reduction in mass-spectral complexity by more than 98% in some cases. The cross-correlation approach was tested with different ionization sources as well as reproducible and irreproducible sample introduction. Software built in-house enabled fully automated data processing, which can be performed within a few seconds. Ultimately, this approach provides an additional dimension of analyte separation in ambient mass-spectrometric analyses with information that is already recorded throughout the analysis.

Keywords: Ambient desorption/ionization, Data processing, DART, FAPA, MS/MS, Cross-correlation

Received: 7 February 2019/Revised: 17 April 2019/Accepted: 3 May 2019/Published Online: 3 June 2019

Introduction

The emergence of ambient desorption/ionization mass spectrometry (ADI-MS) has enabled direct and rapid sample analysis in the open air [1, 2]. In contrast to ionization methods that require vacuum (e.g., electron ionization

[3], chemical ionization [4], synchronized discharge ionization [5], etc.), ADI approaches allow samples to be directly introduced to the ionization source or ionization beam under ambient conditions, obviating the need for many sample preparation and pretreatment steps [2]. After the first introduction of ADI-MS with desorption electrospray ionization (DESI) by Takats et al. [6] in 2004, the concept of direct sample analysis from solid/liquid surfaces has received significant attention. Among the numerous types of ambient ionization sources that have been described [7], those based on electrical discharges, such as direct analysis in real time (DART) [8, 9], have garnered particular

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13361-019-02246-y>) contains supplementary material, which is available to authorized users.

Correspondence to: Jacob Shelley; e-mail: shellj@rpi.edu

attention due to their simple design and operation [10], high flux of reagent species [11], and broad utility [12, 13].

While ADI-MS has improved throughput for qualitative, screening analyses [14–16], the mass spectra frequently contain an abundance of background ions that stem from the open-air nature of analysis, where chemical species in the laboratory environment are exposed to the ionization beam along with the analytes of interest. The significant chemical background due to these species that is encountered in ADI-MS can be detrimental to analyses through competitive ionization matrix effects [17] or by masking analyte ions of lower abundance in a mass spectrum. In terms of non-targeted analysis with ADI-MS, ions of low abundances are not commonly used due to difficulties in recognizing and identifying them [18, 19]. This latter point is particularly a problem as sample introduction in ADI-MS is near simultaneous without any separation. Thus, mass spectrometers with limited resolving power (e.g., unit-resolution quadrupoles or quadrupolar ion traps) are not capable of differentiating isobaric ions. As such, background subtraction of a blank mass spectrum would seem like an essential step to recognition and identification of unknown analytes and low-abundance species.

Unfortunately, conventional background-removal approaches assume the noise (e.g., white noise) in the time domain to be reproducible or static and include noise collection, noise filtration, and noise removal, which can require labor-intensive processes [20]. For instance, the background signals (i.e., chemical noise) encountered with chromatography coupled to MS can be removed by simple subtraction [21]. The restricted environment of chromatography-coupled MS methods assures reproducible background signals with a blank analysis. However, due to the open-air nature of ADI-MS, the time variances of ambient conditions (e.g., air current, nearby volatile species, variations of atmospheric trace gases, etc.) are imparted to the mass-spectral data [11, 22]. As such, subtraction-based background removal approaches usually are not appropriate or are difficult to implement in the processing of ADI-MS data.

In practice, processing of ADI-MS data often relies upon information acquired in only the m/z domain [6, 9, 11, 23–25]. Previously reported ADI-MS studies that used discharge-based or spray-based (e.g., transmission mode desorption electrospray ionization) [25, 26] sources all suggested similar approaches. Meanwhile, the time-domain information (i.e., ion chronograms) is usually treated as useless noise and discarded due to its complexity and large variance [11, 27, 28]. However, unique physiochemical properties of each chemical species in a sample result in different time-dependent desorption, ionization, ion-ion or ion-molecule interactions, and ion-transport behaviors. Consequently, this physiochemical information is encoded in the time domain (i.e., the ion chronograms). Thus, ions present in an ADI mass spectrum that originate from different chemical species will exhibit unique features or slight variations in their time-dependent behavior. Therefore, the development of a computational method to assess the similarities (or differences) between ion chronograms of all species

detected in an ADI-MS experiment would provide a means to automatically differentiate or categorize ions in mass spectra based on their originating chemical species.

In the present study, we explore the use of cross-correlation to gauge the similarities or differences between ion chronograms from ADI-MS analyses. Cross-correlation is a mathematical approach commonly used for noise reduction in signal processing [29–31]. In the analytical sciences, cross-correlation has been extensively used as the basis for library searching of optical spectra (e.g., infrared absorption spectroscopy) [32] as well as mass spectra [33, 34].

In addition to spectral library searching, cross-correlation is also the basis for SEQUEST to determine peptide sequences from tandem mass spectra [35–38]. In that case, experimentally measured tandem mass spectra of proteins/peptides are compared with computationally synthesized fragmentation spectra via a cross-correlation method. Essentially, the cross-correlation function suggests the similarities between two inputs. Thus, the match between measured and computed tandem mass spectra can be used to identify the most likely primary sequence.

However, it is important to note that in the case of spectral-library searching, the spectra themselves (i.e., the wavelength, frequency, m/z , etc.) are cross-correlated with previously obtained or computationally generated spectra, while the time information is not utilized. In ADI-MS, the transient ion signals corresponding to various analytes exhibit unique features in time domain. Particularly, the analyte-ion signals respond to a measurement quite differently. For instance, during the heating process in thermal desorption, the analytes with high vapor pressures are converted into the gas phase at a lower temperature, whereas the ones with low vapor pressures remain in the condensed phase. The features of the ion-dependent time-domain profiles can provide valuable insight through revealing the physiochemical properties of the analytes. In this study, the time-domain profiles of each detectable ion peak are compared to each other in order to differentiate the species. Specifically, the cross-correlation function between two ion chronograms (i.e., the time-domain profiles) is used to evaluate the similarities/differences in physio-chemical properties of their corresponding species.

Experimental

Chemicals

Solutions containing a mixture of 11 carbamate pesticides (531.2 Restek Corporation, U.S., Bellefonte, PA) were used as model sample in this study. The solution was made by diluting the stock pesticide mixture to 1 ppm in HPLC-grade acetonitrile (A996-4, Fisher Chemical, Waltham, MA). The ionization sources tested here used ultra-high purity helium (99.999%, Airgas, Radnor, PA) as a discharge gas in all cases. Commercial headache relief pills, which contained acetaminophen (250 mg/pill), aspirin (250 mg/pill), and caffeine (65 mg/pill), were obtained from a local pharmacy (Walgreen Co.,

Kent, OH). A US dollar bill was also used as a test sample. Standard γ -aminobutyric acid (GABA, Sigma-Aldrich, St. Louis, MO) was kindly donated by Dr. Leah Shriver (University of Akron, Akron, OH).

Ion Sources

A Direct Analysis in Real Time (DART) ID-CUBE (Ion Sense, Inc., Saugus, MA) was used as the ADI source in the initial evaluation of the cross-correlation approach. The ID-CUBE utilizes an OpenSpot Sample Card (Ion Sense, Inc., Saugus, MA) for sample introduction, which consists of a metal mesh held within a cardboard frame. Through user control, the sample-containing mesh is resistively heated with a current-controlled power supply for a hardware-fixed time period of 30 s. The sample introduction is strictly controlled by the heating process, and desorption/ionization does not occur without initiation of the heater current [39].

The DART-ID-CUBE was connected to the atmospheric-pressure interface of the mass spectrometers with a Vapur interface. The Vapur interface helps maintain appropriate vacuum levels of the mass spectrometer by removing much of the neutral helium from the DART, in a jet-separator-type configuration [26, 40]. This interface also improves the ion-transmission efficiency from the OpenSpot card into the mass spectrometer [39, 41]. The supplemental pumping for the Vapur interface was provided by a diaphragm pump (ME 2C NT, Vacuubrand Inc., Essex, CT) with a needle valve that restricted the supplemental pumping flow rate to approximately 8 L min^{-1} .

The use of the DART-ID-CUBE for desorption/ionization ensured predictable and reproducible sample introduction and, as a result, ion chromatograms, for testing this cross-correlation approach. In this study, sample was applied to the mesh of the OpenSpot card and air-dried prior to the thermal desorption and ionization by the source. The helium flow rate used throughout this study was 0.41 L min^{-1} . The DART discharge was powered by the mass spectrometer's built-in ESI/APCI power supply. As such, the DART currents and voltages differed based on the mass spectrometers used. The specific voltage-current combinations are discussed further below.

A Flowing Atmospheric-Pressure Afterglow (FAPA) ADI source was also used to demonstrate the applicability of this cross-correlation approach with other ionization sources and sampling approaches. The FAPA source used here was of a pin-to-capillary geometry, which has been described previously [11, 12]. The FAPA source was operated with a helium flow rate of 1.0 L min^{-1} and a discharge current of 20 mA, which required approximately 500 V, that were controlled with a mass flow controller (GR116-08, Fathom Technologies, Georgetown, TX) and high-voltage power supply (BHK-1000, Kepco, Flushing, NY), respectively. For the FAPA analysis, samples were either used as-is or deposited on glass probes; no other sample pretreatment was performed. Samples were introduced by being manually held between the exit of the

FAPA source and the inlet of mass spectrometer [10, 11]. In this study, the FAPA source was also coupled to the same Vapur interface used with the DART-ID-CUBE.

Mass Spectrometers

Initial studies were performed on high-resolution Orbitrap mass spectrometers (Exactive Plus and Q-Exactive, Thermo Scientific, Bremen, Germany) for unequivocal determination of elemental composition. In most of the presented cases, the maximum resolving power ($m/\Delta m$ at m/z 200) of 140,000 was used. To examine the importance of spectral acquisition rate (i.e., the time resolution) on this cross-correlation approach, lower mass resolving powers were used to yield the maximum possible spectral acquisition rate (ca. 7 Hz). The S-lens RF level, inlet capillary temperature, and number of microscans (i.e., hardware averages) were fixed to 50%, 320 °C, and one, respectively. The mass range used was 50 to 750 Th. The maximum injection time and automatic gain control (AGC) target were set to 5 ms and 1×10^7 counts s^{-1} , respectively. Under such conditions, the AGC target was never reached and resulted in relatively stable spectral acquisition rates throughout an analysis. With the Exactive Plus, the DART operating current was 46 μA , which required a direct-current potential of ca. 300 V.

To gauge the importance of mass resolving power on this data processing approach, samples were also analyzed with a unit-resolution, linear ion trap mass spectrometer (LTQ XL, Thermo Scientific, San Jose, CA). The maximum injection time and number of microscans for these studies were set to 5 ms and one, respectively, to achieve a maximum spectral acquisition rate of 8.6 spectra s^{-1} . The capillary temperature was set to 330 °C. Mass spectra were acquired in the range of 50 to 500 Th. With the LTQ, the DART operated at a discharge current of 99 μA , which required a direct-current potential of ca. 450 V.

Software Environment

All data from this study were processed with a program written in-house with Visual C# .Net framework v4.5 (Microsoft Cooperation, Redmond, WA). Extraction of the raw mass-spectral data from the Thermo RAW files was performed with MSFileReader Library 3.0.1 (Thermo Scientific, San Jose, CA) [42, 43]. The threshold for peak detection in counts per second was manually selected. In this study, the detection threshold for the Orbitrap mass spectrometers (Exactive Plus and Q Exactive) and LTQ XL data were set to 1000 and 10 counts s^{-1} , respectively. Notably, the ion peaks were detected from time-averaged mass spectra throughout each measurement. Cross-correlation of the individual chromatograms was performed with a fast Fourier transform (FT). The algorithms for fast Fourier transform and ion-chromatogram extraction are further described in Supplemental Information (SI).

Result and Discussion

Types of Ion Chronograms

The hardware-fixed heating configuration of the DART ID-CUBE enabled sample introduction to only occur when electrical current passed through the mesh. As such, the ion chronograms obtained with the DART-ID-CUBE are more reproducible than many ADI-MS approaches [26]. To demonstrate the time-domain features of the ion chronograms, ion chronograms for three analyte ions and one background ion were extracted from the mass-spectral data from the analysis of the mixture of 11 pesticides. The use of a 1 ppm solution of analytes resulted in explicit ion signals. As shown in the green trace of Figure 1a, the protonated molecular ion signal (MH^+) for carbofuran only appeared when the heating current was applied through the sample-containing mesh. Meanwhile, background-ion signals existed for much of the duration of the analysis, except when the sample card was initially inserted in the ion source at ~ 0.25 min (cf. Figure 1a, red trace); this

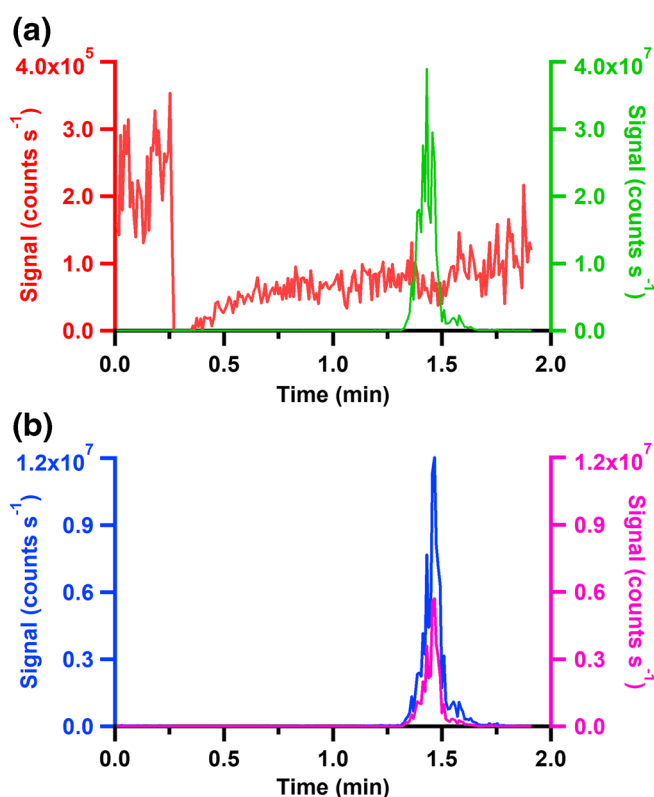


Figure 1. Chronograms of different ion types in the analysis of a 1-ppm pesticide mixture with the DART-ID-CUBE. The red trace in (a) corresponds to the chronogram of a common background ion, protonated phthalic acid anhydride at m/z 149.0235 [44]. This chronogram was smoothed with a boxcar filter of a window size of 9 points due to the large fluctuations. The green trace in (a) is the chronogram of protonated carbofuran pseudo molecular ion at m/z 222.1125. The blue and magenta traces in (b) are chronograms of protonated carbaryl (MH^+) and ammoniated carbaryl ion ($M + NH_4^+$) at m/z 202.0864 and 219.1127, respectively. Analyte ion chronograms were not smoothed

decrement is due to momentary blocking of the ion beam from entering the MS. The time-domain profiles of background ions and those of analytes, in this case, were visually distinguishable.

In contrast, ion chronograms for two protonated analytes (carbofuran and carbaryl), shown in Figure 1a, b, were visually similar. However, careful inspection reveals that the shape of the chronogram, such as the time at which maximum signal occurred (i.e., t_{max}), is slightly different between the two species. In particular, the ion chronogram of protonated carbofuran reached the signal maximum prior to that of the protonated and ammoniated carbaryl. Additionally, other morphologies of the chronograms can vary depending on their originating species (cf. Figure 2). For instance, both carbofuran and carbaryl became detectable at a time of 1.3 min. Through comparing the chronograms between their protonated ions, carbofuran exhibited a steeper rising edge (cf. Figure 2, green and blue traces). Meanwhile, other morphological parameters, such as falling edge, symmetry, local maxima, etc., are also analyte-specific. However, the deviations in the t_{max} and morphologies are not simply functions of one single physiochemical property (e.g., melting point, boiling point, phase transition enthalpy, etc.), but a more sophisticated response of a chemical species to desorption, ionization, and subsequent transport processes. In contrast to the differences between the chronograms of ions that originate from different chemical species, those ions that come from the same chemical species (e.g., protonated and ammoniated carbaryl) exhibit high similarity in the time-domain profile as well as the same t_{max} (cf. Figure 1b). Although t_{max} seemed to be quite convenient to differentiate the chronograms of some chemicals, it is not sufficient in analyses where different species share the same t_{max} . As such, utilizing the morphological information carried by the ion chronograms, i.e., the shape (including the t_{max}), provides greater discrimination.

Calculating the Cross-Correlation Function

In order to quantitatively gauge the differences and similarities between ion chronograms, the cross-correlation function was chosen. The traditional cross-correlation function is defined in the continuous regime (cf. Eq. S1). However, due to the discrete nature of mass-spectral recording, as opposed to continuous (analog) acquisition, ion signals within an ion chronogram are discretely distributed in the time domain, which necessitates the use of a discrete cross-correlation function. The cross-correlation function, C_{ab} in the discrete regime, can be written as

$$C_{ab} = \sum_{n=0}^{T/\Delta t} f(t) \times g(t \pm n \cdot \Delta t), n = 0, 1, 2, \dots, T/\Delta t \quad (1)$$

where $f(t)$ and $g(t)$ are ion chronograms (i.e., transient ion signals). The T and Δt are the total measurement time and the interval between mass-spectral acquisitions, respectively. The integer n is the sequential index of the data points. Specifically,

the term $n \cdot \Delta t$ refers to the displacement, which is commonly denoted with τ . However, this discrete cross-correlation is computationally expensive and not appropriate for this application. Instead, the Fourier relationship of cross-correlation [30, 32] provides a rapid, high-performance calculation of the cross-correlation function, which can be written as:

$$\begin{aligned} C_{ab}(\tau) &= f(t) \otimes g(t) \\ \begin{matrix} iFT \uparrow & FT \downarrow & FT \downarrow \end{matrix} & \\ C_{ab}(t) &\leftrightarrow F^*(v) \times G(v) \end{aligned} \quad (2)$$

where C_{ab} is the cross-correlation function, $f(t)$ and $g(t)$ are the input functions, the $F^*(v)$ and $G(v)$ are the Fourier transforms of the input functions, the $*$ represents the complex conjugate, v is the frequency-domain independent variable, FT is a Fourier transform, and the \otimes is the symbol to denote a cross-correlation operation. A plot of the cross-correlation function (i.e., $C_{ab}(\tau)$) against the displacement (i.e., τ) is a cross correlogram between $f(t)$ and $g(t)$. Notably, the Fourier transform converts the chemical information from the time domain into the frequency domain. Similar time-domain features will result in similar frequency components. Through cross-correlation, common features of two ion chronograms are enhanced and vice versa. Thus, cross correlograms can be used to quantitatively assess the similarities and differences in time-domain features between ion chronograms.

Types of Cross Correlograms

The cross correlograms of ADI-MS chronograms can be categorized into four types: (i) that of an ion against itself (also called autocorrelation, cf. Figure 3a), (ii) that of two ions that originated from the same chemical species (cf. Figure 3b), (iii) that of two ions that originated from different chemical species from the sample (cf. Figure 3c), and (iv) that of an analyte ion against a background ion (cf. Figure 3d). It is readily apparent in these correlograms that the τ values where the chronograms overlap, i.e., τ_{\max} , can be used as an indicator of the similarities between two detected ions.

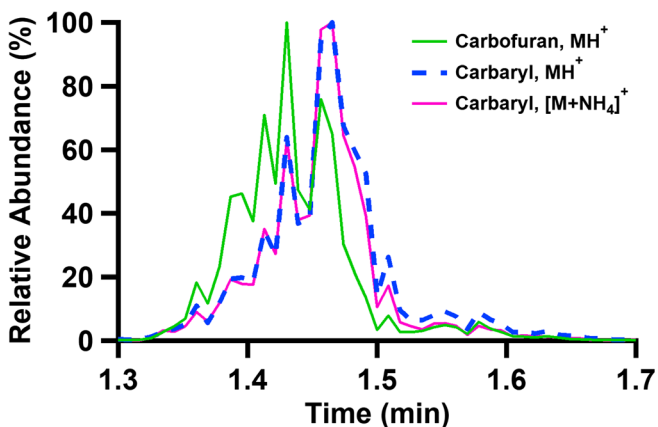


Figure 2. Expanded view of the ion chronograms for protonated carbofuran (solid green trace), protonated carbaryl (dashed blue trace), and ammoniated carbaryl (solid magenta trace)

For ions that stem from the same chemical species, the τ_{\max} is located at $\tau=0$ s because they share similar ion-production processes corresponding to desorption, ionization, and subsequent transportation of a particular analyte. In contrast, a non-zero τ_{\max} suggests that the chronograms that are being cross-correlated are not from the same chemical species. However, a $\tau_{\max}=0$ s can arise due to noise introduced during data acquisition. Some of the causes for signal fluctuations are common to all ion chronograms. Thus, a local maximum at $\tau=0$ s can always be found in a cross correlogram without smoothing, which may unnecessarily complicate the results with a false global maximum. This issue is especially true for cross correlograms of different analyte species where the true τ_{\max} is usually non-zero. Presence of the zero artifact can result in false ion categorization by recognizing analyte ions from different species as the same (cf. Figure S3). As such, the cross correlograms were smoothed with a low-pass filter in Fourier space for improved computational efficiency [45]. In the present work, a low-pass filter of ca. 0.5 Hz was found to be adequate to remove the zero artifacts and was used to smooth all cross correlograms in this work (cf. Figure 3).

Cross-Correlation-Based Approach for Background Removal

In the case of analyses performed with the DART ID-CUBE, analyte ions are expected only within the 30-s duration that heat is applied to the OpenSpot sample card. As such, if the τ_{\max} from the cross correlogram exceeded the range of ± 30 s, the ion originated from an extraneous source and was not introduced with the sample. These sorts of ions can be considered background and removed from mass spectra. In practice, an ion chronogram of any ion can be compared through cross correlation with that of an arbitrarily chosen reference ion, such as any species known to originate from the sample. It should be noted that the reference ion can be the first or last ion peak in time domain. Thus, the τ_{\max} difference can be either $+30$ s or -30 s.

To demonstrate the ability of background removal through the cross-correlation approach, a single-analyte sample was used instead of the pesticide mixture to avoid overly complicated mass spectra. A trace amount of solid γ -aminobutyric acid (GABA) was transferred onto the OpenSpot sample card with a glass probe. The mass-spectrometric data consisted of primarily two types of ions: those corresponding to the analyte, GABA, and those from background. To perform cross-correlation analysis, the chronogram for the ion of protonated GABA at m/z 104.0709 was arbitrarily chosen as the reference. From the cross correlograms between all ion chronograms and the reference ion, the mass-spectral peaks with τ_{\max} within ± 30 s were isolated and extracted from the raw mass spectrum. Namely, an ion is considered as background if its τ_{\max} exceeds ± 30 s. The time-averaged mass spectrum consisting of 1468 ion peaks, including background and analyte ions, is shown in Figure 4a. The cross-correlation-

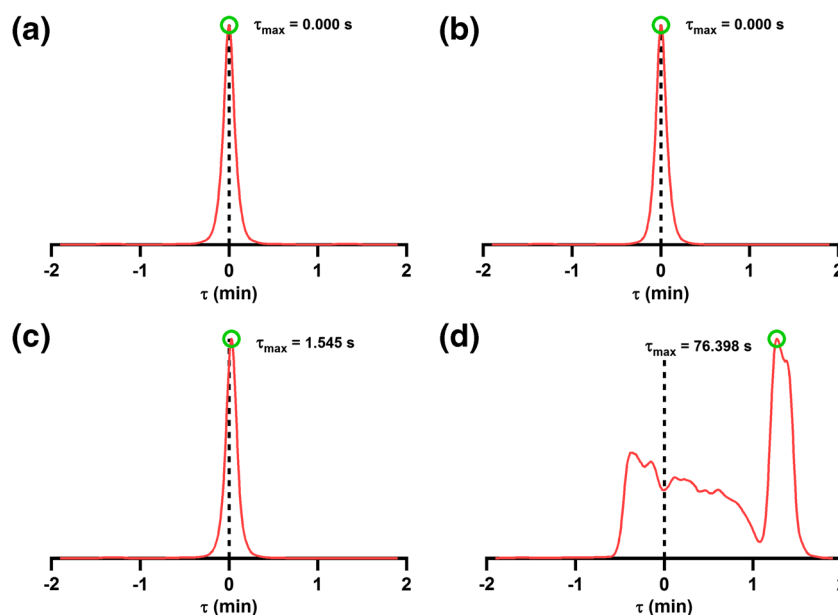


Figure 3. Cross correlograms relating different types of ions from the analysis of the pesticide mixture. (a) Auto-correlation function for protonated carbaryl, m/z 202.0864. (b) Cross-correlation between ion chronograms of protonated carbaryl (MH^+) at m/z 202.0864 and ammoniated carbaryl ($M + NH_4^+$) at m/z 219.1127. (c) Cross-correlation between ion chronograms for protonated carbaryl at m/z 202.0864, and protonated carbofuran at m/z 222.1124. (d) Cross-correlation of ion chronograms for protonated carbaryl, m/z 202.0864, and protonated phthalic acid (background) at m/z 149.0235

based background removal process flagged nearly 90% of ion peaks as background. They were removed from the raw mass spectrum, resulting in one containing only 168 ion peaks (cf. Figure 4b). Although the simplified mass spectrum is still rather complicated, analyte-ion peaks of low abundance ($< 10\%$ of the base-peak signal), such as the fragments denoted in Figure 4b, become significantly easier to recognize. In addition, other than the labeled peaks in Figure 4b, the ion peaks of high abundance (e.g., m/z 172.1330 and m/z 195.1745) were later found to be not related to GABA by further constraining τ allowance to 0 s. This method will be discussed in detail in a later section of this work.

Similarly, the cross-correlation approach was also applied to the mass-spectral data recorded with a mixture of pesticides. In this example, the chronogram for the protonated carbaryl ion (m/z 202.0863) was arbitrarily chosen as the reference. Through flagging and removing the ions with a τ_{\max} that exceeds the sample introduction window (± 30 s), the time-averaged mass spectrum of an analysis of a pesticide mixture that contained 1319 ion peaks (cf. Figure 5a) was converted to one with 321 ion peaks (cf. Figure 5c). These ion peaks were largely due to species from the sample applied to the card. However, full interpretation of a spectrum with this number of peaks would still be rather cumbersome; after all, the interpretation process is still largely manual.

In addition to high-resolution mass spectrometers, the cross-correlation approach was also applied to mass spectra recorded with a unit-resolution LTQ mass spectrometer. The reduction of mass-spectral complexity was similar

when cross-correlation approached was applied to the data acquired with a unit-resolution mass spectrometer. Note that for computing the cross-correlation function with the LTQ data, the reference ion was changed from protonated carbaryl to carbofuran MH^+ (cf. Figure S5b). Specifically, the time-averaged mass spectrum of the 1-ppm pesticide mixture originally contained 424 ion peaks (cf. Figure 5b). After cross-correlation background removal, the mass spectrum contained only 68 ion peaks (cf. Figure 5d). The removal of 86% of ion peaks drastically decreased the complexity of a mass spectrum. Unfortunately, a large portion of analyte information was lost along with the background removal. This information loss can be attributed to isobaric interferences.

The ion chronograms from a unit-resolution mass spectrometer are more strongly influenced by isobaric overlap. For instance, the chronogram of the ion at m/z 145.0 showed both background and analyte features (cf. Figure S5a). From the mass spectra acquired with the Exactive Plus mass spectrometer, at least two different species make up the feature at m/z 145.0: An ion at m/z 145.0650 was due to a fragment of carbaryl, and a species at m/z 145.0860 was present in the background. The cross correlogram (cf. Figure S5c) between the ions at m/z 145.00 (cf. Figure S5a) and the reference (cf. Figure S5b) showed a τ_{\max} that exceeded the sample introduction window (± 30 s). Although a local maximum was found at $\tau = 0.35$ s, the global maximum dominated the cross correlogram due to the low abundance of the analyte fragment ion compared to that of the background. Notably, a global or local maximum in the correlogram due to background species

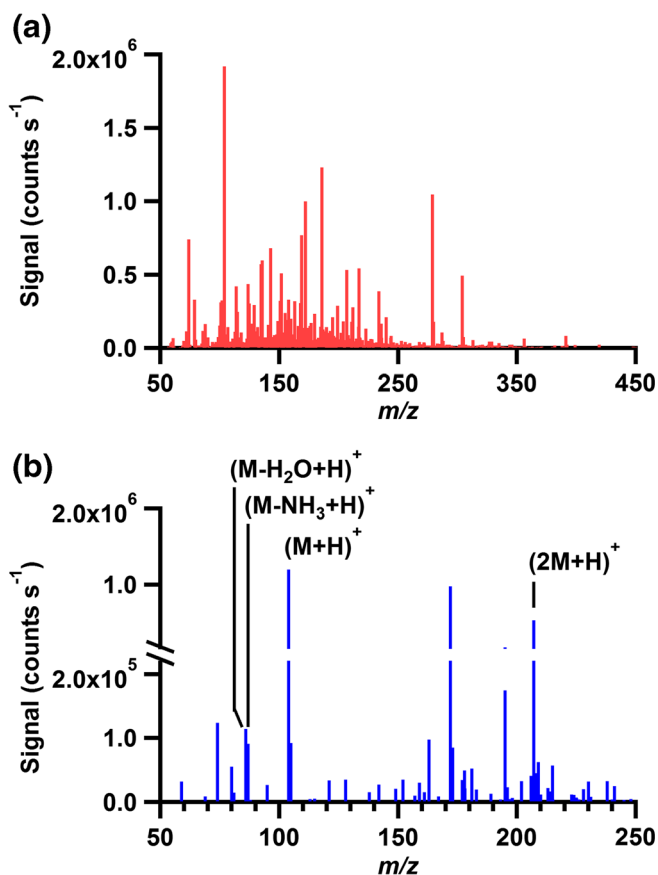


Figure 4. Cross-correlation-based background removal for the analysis of a single-component sample. The protonated GABA ion was used as the reference. **(a)** Time-averaged raw mass spectrum with 1468 ion peaks. **(b)** Time-averaged mass spectrum after background removal, which contains 168 ions peaks. Note the axis break in **(b)** to show low-abundance species retained after background removal

will exhibit a peak outside the ± 30 s range due to the strictly controlled sample introduction period used in this study. It is theoretically possible to obtain a cross correlogram that exhibits multiple peaks within the sample introduction period. Such a finding would indicate that the sample contains multiple isobaric analytes, which cannot be separated by the instrument. In this scenario, the use of a high-resolution mass spectrometer would significantly improve the quality of cross-correlation-based data analysis.

Choosing the Reference Ion

The selection of one reference ion for cross-correlation analysis is the essential step. In the present work, the reference ion was manually chosen to be one of the analytes known to be within the sample. Even in the absence of any knowledge of possible analytes within a sample, such as in non-targeted analyses, an arbitrary reference ion can be easily found after analysis. With the distinctively different time-domain features between an analyte and a background species, an analyte reference can be chosen by inspecting the shape of the ion chronograms. This

reference ion can be any species known to originate from the sample, even if the exact identity of that species is unknown. A more involved option would be to add an internal standard to the sample before analysis. As an example, caffeine could be added to this pesticide mixture. As such, the protonated caffeine at m/z 195.0877 would be expected in the time-averaged mass spectrum; its ion chronogram could, then, be used as the reference. But, it is important to stress that this data analysis approach does not require the use of standard.

Generally, a reference ion should possess a well-defined ion chronogram. In this regard, the signal of this reference ion should not be close to the noise floor of a mass spectrometer. In the present work, for either the GABA or the pesticide sample, the reference ion was chosen to be a protonated analyte. This selection criterion was based on the soft-ionization nature of ADI methods where pseudo-molecular ions are usually of highest abundance. Meanwhile, a reference ion should not possess significant background features in its chronogram. In the case of high-resolution mass spectrometers (i.e., the Orbitraps), the ion peaks in mass spectra were well-defined and isobaric overlap was minimal. However, due to the limited resolving power of the linear ion trap mass spectrometer, the ion chronograms of the analytes commonly exhibited background features (cf. Figure S5b). That was the reason why the reference ion was changed from protonated carbaryl to protonated carbofuran for the analysis of LTQ data. Although the ion chronogram of protonated carbofuran still exhibited a background feature, the analyte behavior in the chronogram was still sufficient for cross-correlation calculation. In contrast, the ion chronogram of protonated carbaryl was no longer suitable due to the large contribution of a background feature due to isobaric overlap.

Ideally, analyte ions that match the criteria stated above can be used as a reference. For instance, the use of carbaryl or carbofuran showed highly similar results while flagging the background ions. The total number of ion peaks after background removal remained the same (i.e., 321). Except for two ions at m/z 335.2579 and m/z 349.2375 were mismatched. The ion signals of these two ions were very close to the noise floor of the instrument in this case. The small differences were presumably due to the noise contained in the ion chronograms. Primarily, the low-abundance ions that were affected contained many points where zero ion signal was recorded. But, the peaks of interests were still correctly recognized as analytes. A more detailed investigation of reference-ion choice on this cross-correlation method will be discussed in a future publication.

Categorization and Extraction of Single-Component Mass Spectra

By further constraining the τ_{\max} allowance to 0 s rather than the entire sample-introduction window (i.e., ± 30 s), ions that were highly correlated with the reference ion were recognized and isolated from the time-averaged mass spectrum (cf. Figure 6a). Among the 17 ions that were found to be highly correlated with the protonated carbaryl, 12 peaks have been identified and

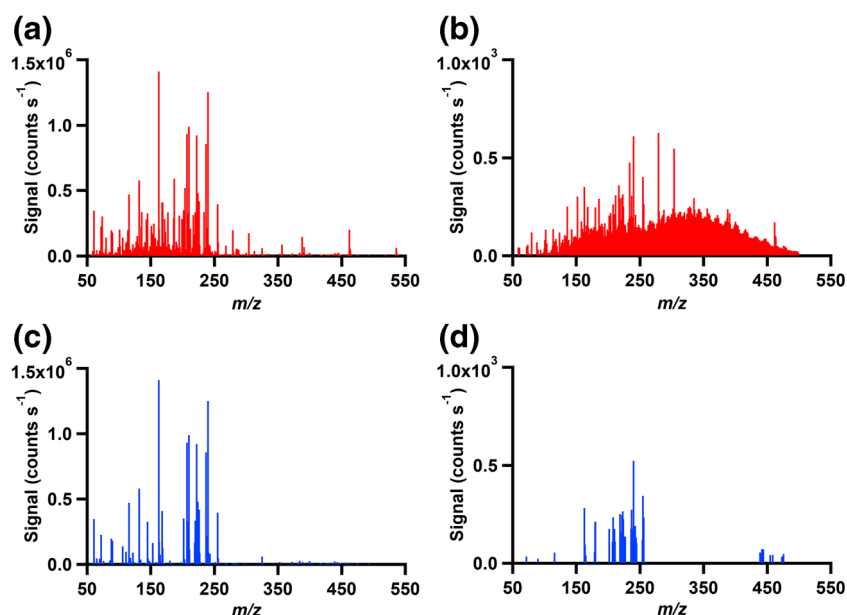


Figure 5. Cross-correlation background removal for the mass spectrum of the pesticide mixture with protonated carbaryl as the reference. (a), (b) Time-averaged raw mass spectra obtained from Orbitrap and LTQ mass spectrometers that contain 1319 and 424 ion peaks, respectively. (c), (d) The time-averaged mass spectra of the Orbitrap and LTQ mass spectrometers after background removal, which exhibit only 321 and 68 ion peaks, respectively

relate to carbaryl (cf. Table 1). Moreover, low-abundance ions near the noise floor of the instrument were recognized and extracted from the complicated mass spectrum. For instance, the ammoniated carbaryl-dimer cluster at m/z 420.1920 was preserved, which was two orders of magnitude less abundant than that of protonated carbaryl. Such low-abundance ions would be difficult or impossible to recognize manually. Other unknown species listed in Table 1 were found to be highly correlated with protonated carbaryl. Namely, the chronograms of these species shared similar time-domain patterns (cf. Figure S10). Unfortunately, the low abundance of these “unknown” ions limited the possible means for their identification beyond elemental composition from accurate mass. Based solely on elemental composition, we are unable to rationalize identities of these fragments. Additionally, it is important to note that although fragment ions of a particular analyte were found and identified, no fragmentation energy was applied to the ions within the mass spectrometer. These fragment ions were likely generated during the desorption/ionization process as well as during transport through the first-stage of the mass spectrometer [11]. All fragment ions reported here were previously reported for carbaryl [46].

Like the use of $\tau_{\max} = 0$ s allowing the recognition of reference-related ions, those stem from the same chemical species, but are different from the reference, and also show analogous features on the cross correlogram; namely, ions that originate from the same chemical species share a τ_{\max} . As an example, the ions with τ_{\max} of -2.09 s were extracted from the all-ion mass spectrum (cf. Figure 6b), in which 9 out of 17 peaks were identified and related to carbofuran (cf. Table S1). Eventually, the τ_{\max} is analyte-specific. The τ_{\max} can, thus, be used as labels for each species within a sample mixture.

Importance of Time Resolution

The cross-correlation approach relies upon variations of signals in time. As such, one critical requirement to perform cross-correlation for ion recognition is the time resolution or, in this case, the mass-spectral acquisition rate. To demonstrate the importance of spectral acquisition rate, mass spectra of the pesticide sample were acquired with different numbers of hardware averages (i.e., “microscans”) with the Orbitrap mass spectrometer. The resolving power setting of the Orbitrap mass spectrometer was set to 35,000 to obtain a higher mass-spectral acquisition rate while maintaining adequate resolving power. Meanwhile, the maximum injection time and ACG target were set to 20 ms and 10^7 counts s^{-1} , respectively. As an additional comparison, an off-line resampling with varying numbers of averaged data points was performed at the highest spectral acquisition rate of ~ 8 spectra s^{-1} .

With the model sample of a pesticide mixture used above, the number of ions that was highly correlated with carbaryl (cf. Figure S6, blue trace) and carbofuran (cf. Figure S6, red trace) was measured. Because the Exactive Plus only allows up to 10 hardware averages (i.e., microscans) to be made, the raw chronograms were resampled through averaging in order to mimic different spectral acquisition rates. The number of ions that correlated with protonated carbaryl and carbofuran did not significantly change when the spectral acquisition rate was greater than 1.4 spectra s^{-1} , which corresponds to 5 microscans of hardware averaging at a resolving power of 35,000. However, the number of correlated ions more than doubled when the acquisition rate was decreased to 0.7 spectra s^{-1} (i.e., 10 microscans). Through software averaging, it was observed

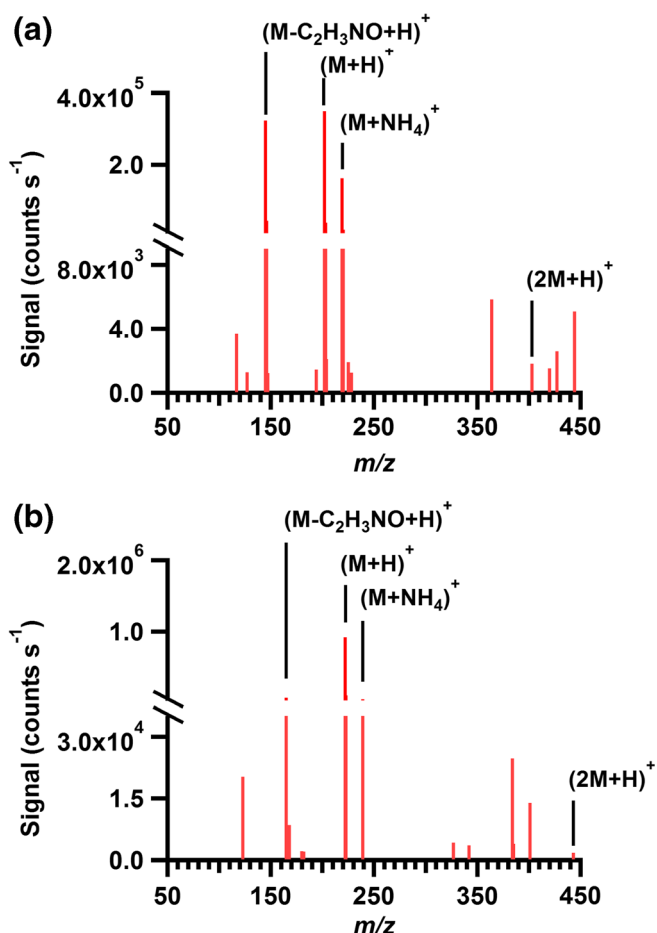


Figure 6. Isolated analyte-specific mass spectra. (a) Isolated mass spectrum of carbaryl-related ions using the cross-correlation approach. Isolated, single-analyte mass spectrum of carbaryl. (b) Isolated mass spectrum of carbofuran ($\tau_{\max} = -2.09$ s) with protonated carbaryl (m/z 202.0864) as reference. Note the break in the vertical axis to more easily view the low-abundance ions

that time-domain features of the ion chronograms were well resolved when the mass-spectral acquisition rate was faster

than 1.1 spectra s^{-1} , which showed a similar number of recognized ions. Such an observation implies that the data acquisition frequency for ion categorization was within a certain frequency band. At low spectral acquisition rates, the time-domain variation in the ion chronograms was no longer distinguishable. Thus, an adequately fast spectral acquisition rate is needed to capture higher frequency variances for correlation-based ion recognition. However, it could be that high-frequency noise, unrelated to the physicochemical nature of the analyte, might dominate the time-domain features at higher spectral acquisition rates than could be obtained here. In the previous section of this work, 12 ion peaks corresponding to carbaryl were identified with high certainty (cf. Table 1). However, less than half of the peaks were correctly isolated at a mass-spectral acquisition rate higher than 2 spectra s^{-1} . Thus, further investigation of the effect of higher spectral acquisition rates will be explored in the future with more suitable mass analyzers (e.g., time-of-flight).

Symmetry of Cross Correlograms

The fundamental idea behind the cross-correlation approach is to compare the similarities between ion chronograms. The determination of τ_{\max} depends on the features of the ion chronogram, such as the time at signal maximum and shape of chronogram. In some cases, where ions may not be categorized correctly with only τ_{\max} , the symmetry of the cross correlogram can be used as another dimension of ion recognition. To demonstrate symmetry-based ion recognition, a low-pass filter in Fourier space of ca. 1.7 Hz was used, after which a portion of ions were not separated. In this case, ion chronograms of different species with very similar t_{\max} values induced the same τ_{\max} . Under this condition, a τ_{\max} of 0 s was found on the cross correlogram (cf. Figure S7b) between protonated propoxur (m/z 210.1125) and an aldicarb fragment ion (m/z 116.0532), which are two very different chemical species. In this case, the zero artifact

Table 1. List of Ions that Were Highly Correlated with the Reference, Protonated Carbaryl with m/z 202.0864

m/z	Signal ($\times 10^3$ counts s^{-1})	Formula	Type of ion
117.0701	4.79	$^{12}C_9H_9$	Fragment ion
127.054	1.92	$^{12}C_{10}H_7$	Fragment ion
145.0649	501	$^{12}C_{10}H_9O$	Fragment ion
146.0681	50.5	$^{13}C^{12}C_9H_9O$	Fragment-ion isotope
147.0716	1.44	$^{13}C_2^{12}C_8H_9O$	Fragment-ion isotope
194.1135	1.72	–	Unknown
202.0864	419	$^{12}C_{12}H_{12}O_2N$	Molecular ion
203.0896	48.9	$^{13}C^{12}C_{11}H_{12}O_2N$	Molecular-ion isotope
204.0931	2.33	$^{13}C_2^{12}C_{10}H_{12}O_2N$	Molecular-ion isotope
219.1127	179	$^{12}C_{13}H_{15}O_2N_2$	$(NH_4)^+$, Adduct ion
220.1161	20.2	$^{13}C^{12}C_{11}H_{15}O_2N_2$	$(NH_4)^+$, Adduct-ion isotope
228.0963	1.49	–	Unknown
403.1657	2.13	$^{12}C_{24}H_{23}O_4N_2$	Dimer
420.1920	1.69	$^{12}C_{24}H_{26}O_4N_3$	$(NH_4)^+$ Adduct, Dimer
427.1690	2.90	–	Unknown
444.1949	4.16	–	Unknown

was not the cause of the overlap. Rather, these two ions reached the maximum signal at the same time and had highly correlated fluctuations otherwise (cf. Figure S7a). However, comparison of the cross correlogram between these ions (cf. Figure S7b, green trace) with the autocorrelogram (cf. Figure S7b, magenta dashed-trace) exhibited some notable differences; namely, the shape of the correlogram was not symmetric around $\tau = 0$ s.

While the symmetry, or lack thereof, can be assessed visually in this case, an automated and quantitative approach would be ideal for processing of larger datasets. The symmetry of the cross correlogram can be gauged in Fourier space. In Fourier space, the imaginary part represents the contributions from sinusoidal functions, which is an odd function. Therefore, the symmetry can be simply assessed by taking the absolute summation of the imaginary part of the cross correlogram in Fourier space, termed the symmetry index. A greater symmetry index indicates worse symmetry in the cross correlogram. The calculation of symmetry index was performed prior to the removal of the zero-artifact (i.e., applying the Fourier space low-pass filter) to prevent the loss of symmetry information. Notably, the determination and evaluation of the symmetry indices required normalization of the ion chronograms prior to computation of the cross-correlation function.

The symmetry index for propoxur at m/z 210.1125 in this case was 0 because it was the reference ion (cf. Figure 7). In contrast, the symmetry index for an aldicarb fragment at m/z 116.0532 was 336.4, which is the smallest symmetry index from all ions known to stem from aldicarb. The cross-correlation functions of ions that were more symmetric than aldicarb were considered to be related to propoxur, and vice versa. Thus, a symmetry-index threshold of 336.4 was used. Ions with a symmetry index less than 336.4 were more similar to protonated propoxur (cf. Figure 8a), while the rest of the peaks were assigned to aldicarb (cf. Figure 8b).

Notably, use of symmetry index to isolate a mass spectrum corresponding to propoxur was very similar to the one obtained through the τ -based approach when the Fourier low-pass filter was optimized (cf. Figure S11). The use of a threshold of 336.4 resulted in the most comparable isolated mass spectrum of propoxur compared to what was obtained from the τ -based method (cf. Figure S11). A threshold greater than that 336.4 may falsely categorize aldicarb-related ions into the isolated mass spectrum of propoxur. Similarly, smaller value of the threshold may exclude propoxur-relating species. Compared to τ_{\max} -based peak categorization, the symmetry-index-based method is less reliable due to its complexity. Symmetry index may not be suitable as a standalone method as the τ -based one; it can, however, provide additional information when τ_{\max} alone is not sufficiently discriminative.

It is important to note that the calculation of symmetry indices was performed prior to applying the low-pass filter in the Fourier domain. That means that the symmetry indices and the τ_{\max} -based methods are two orthogonal means in describing the similarities/differences between ion chronograms with respect to the reference. Meanwhile, ions

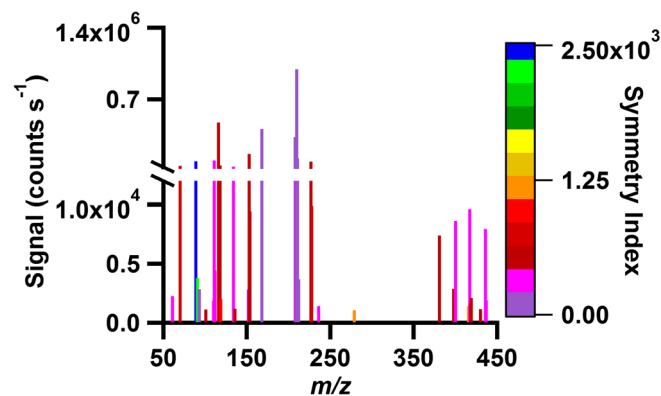


Figure 7. Isolated peak category of propoxur with symmetry index. By using the m/z -210.1125 (propoxur) chronogram as a reference, a category of peaks was determined along with symmetry indices of each peak

stemming from the same chemical origin will result in similar symmetry indices due to the inheritance of the same physiochemical behavior during desorption/ionization. As a result, the symmetry indices of propoxur-related ions exhibited symmetry indices close to 0 because protonated propoxur was used as reference. Vice versa, the ions stemming from aldicarb possess non-zero symmetry indices but were also grouped together. Unfortunately, the number of ion

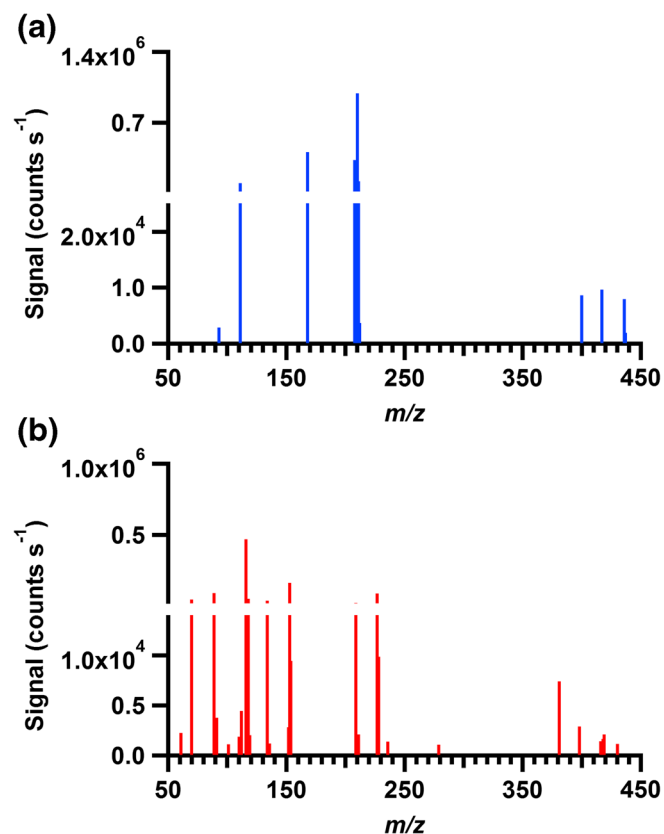


Figure 8. Separation of peak category based on the symmetry indices, where (a) is the peak category of m/z 210.1125 with symmetry indices < 336.4 and (b) is the ion peaks with symmetry indices ≥ 336.4 . The base peak in (b) is at m/z 116.0532

peaks was not sufficient to describe this distribution pattern reliably. But, a rough fitting of Gaussian curves to the symmetry-index distributions for each analyte yielded a crossing point of these curves very close to 336 (figure not shown). Thus far, the selection of threshold was still largely empirical. Yet, the fitting of symmetry-index distribution allows the determination of the threshold *in silico* and, as such, can be automated, as well.

Demonstration with Irreproducible Sample Introduction

To demonstrate the capability of the cross-correlation approach for irreproducible sample introduction that are common in ADI-MS MS analyses, a headache relief tablet and a US dollar bill were used as model samples and analyzed with a FAPA source coupled to a Q-Exactive Orbitrap mass spectrometer. The resolving power of the Orbitrap was set to 17,500 to achieve a spectral acquisition rate of 11.8 spectra s^{-1} . The mass spectra were recorded for approximately 70 s. To smooth the cross correlograms, a low-pass filter with ~ 0.7 Hz was used.

The tablet sample produced a total of 1484 ion peaks in the time-averaged mass spectrum. To perform the cross-correlation analysis, protonated acetaminophen (m/z 152.0706) was used as the reference ion. Additionally, the sample introduction window was set to ± 8 s which was estimated through observation of the total ion chromatogram. After background removal with the cross-correlation method, 342 ion peaks remained. Peaks that were highly correlated with protonated acetaminophen ($\tau_{\max} = 0$) were isolated to a mass spectrum that contained 19 ion peaks that include protonated acetaminophen and its $X+1$ and $X+2$ ^{13}C isotopes (cf. Figure 9c). From the group of ions at $\tau_{\max} = 0.81$ s, the pseudo-molecular ion of caffeine (cf. Figure 9b) was found. Notably, $X+1$ and $X+2$ ^{13}C isotopes were correctly isolated. A common caffeine fragment at m/z 138.0660 was detected and manually found in the raw mass spectrum, where its signal was 0.01% of the protonated caffeine. A mass spectrum corresponding to aspirin was isolated at $\tau_{\max} = 0.54$ s. Interestingly, the two most abundant peaks in this spectrum (cf. Figure 9a) correspond to a common aspirin fragment (m/z 121.0284) and ammoniated acetylsalicylic acid (m/z 198.0763). The protonated form of aspirin (m/z 181.0495) was also detected. However, due to the distinct time-domain profiles between the aspirin fragment and protonated aspirin, the τ_{\max} was 0.27 s (one data point) different from these two ions. The reason for the difference in chronograms between the ions originating from aspirin is unclear at this point.

In the analysis of the US bill, the paper currency was held between the ion source and MS inlet for approximately 20 s. Cocaine is commonly detected on currency bills [47]; thus, the chronogram of protonated cocaine at m/z 304.1544 was used as the reference and the sample introduction window for background removal was set to ± 20 s. Beyond the use of known sample species as reference ions, it is also possible to manually find a trace with a strong analyte feature in a few seconds to

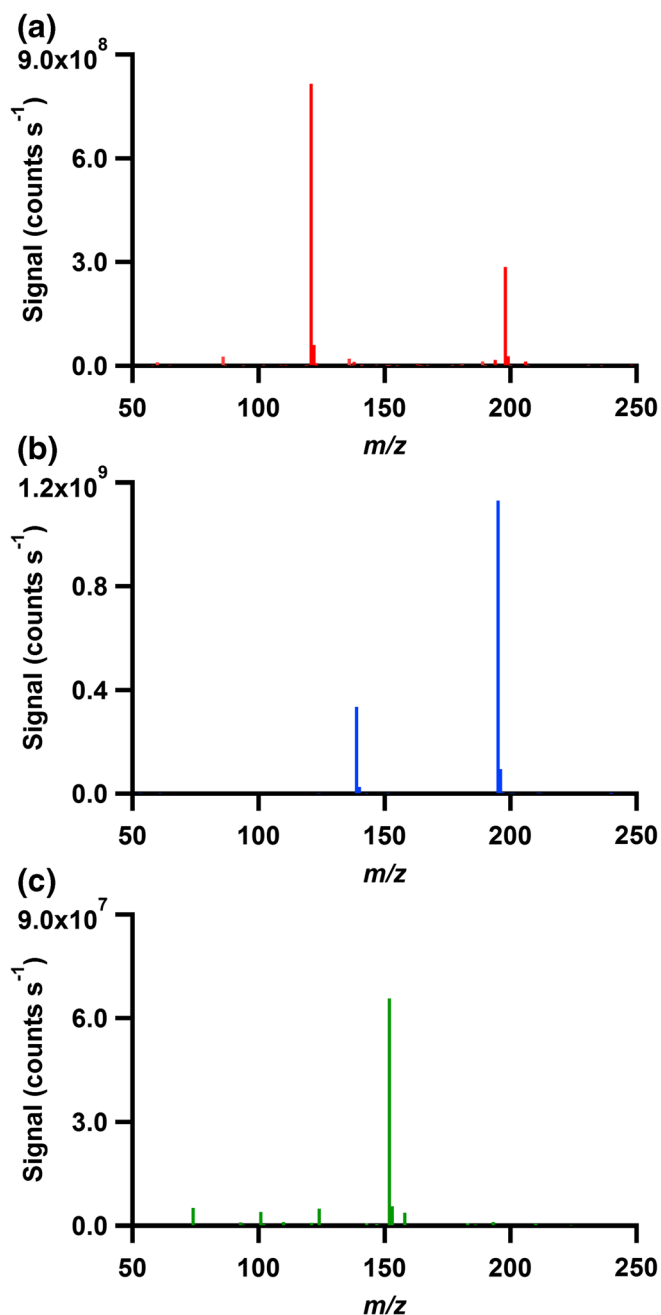


Figure 9. Isolated mass spectra of different species of the headache relief tablet. (a) Isolated mass spectrum of aspirin (m/z 121.0284). (b) Isolated mass spectrum of caffeine (m/z 195.0874). (c) Isolated mass spectrum of acetaminophen (m/z 152.0706). In this figure, maximal signals of each ion were used instead of time-averaged ones

perform cross-correlation analysis. In the case of the US dollar bill, the peak-detection threshold was set to 100 counts/s for the species with relatively low abundance (e.g., isotopes). Through background removal, 1701 out of 2081 ion peaks were found with analyte features in time domain. At this point, the peak identification remained difficult due to the large number of analyte-like ions.

The use of single-analyte mass spectra in this case drastically decreased the complexity and the amount of work during peak identification. A group of 24 ions were found to have a $\tau_{\max} = 0$ s and includes those that possess highly similar time-domain features compared to protonated cocaine, which was used as the reference ion. Within this group, protonated cocaine with two of its isotopes was found (cf. Figure S8a). In this example, one of the strong peaks (m/z 139.1109) was not due to cocaine, but the efficiency of ion identification was still significantly improved.

To demonstrate the capability of the peak-isolation method based on cross-correlation, we arbitrarily picked the ion group with $\tau_{\max} = 0.53$ s (cf. Figure S8b). With the elemental composition determined through accurate-mass measurement and the isotopic distribution, as well as sample context, the base peak within this group at m/z 192.1384 was identified as protonated diethyltoluamide (DEET) with high certainty. Similarly, other groups of ions corresponding to different τ_{\max} can be easily identified as well. For instance, the group at $\tau_{\max} = 0.46$ s likely includes methamphetamine. Meanwhile, diphenylamine and 3,4-methylenedioxyamphetamine (MDMA) were found and identified at $\tau_{\max} = -0.92$ s and 2.37 s, respectively. Notably, these identified species exhibited very low ion signal comparing to background. For instance, the ion signal of protonated DEET at m/z 192.1384 was $\sim 4 \times 10^6$ counts s^{-1} , which was one order of magnitude lower than that of a background ion at m/z 279.1586.

Moreover, based on the τ_{\max} , ions can be categorized into different groups for identification. By including the symmetry indices, the peak identification can be further simplified, as described previously. For instance, the symmetry index of the ion peak at m/z 139.1109 within the cocaine group was 5×10^3 . Compared to the symmetry index of ^{13}C -isotope, which was 264, this peak can be excluded (cf. Figure S9). Thus, the use of symmetry indices can be very useful for peak identifications within ion groups that contain large numbers of ions after the isolation process.

Conclusion

In this study, the capability of a cross-correlation approach for automated ADI-MS data processing was demonstrated. The data analysis was performed with software made in-house to achieve rapid and almost fully automatic background removal as well as analyte-ion recognition and categorization. Through the incorporation of an additional dimension of information (i.e., time), ions in a mass spectrum can be categorized based on unique physiochemical properties of the analyte molecules that are exhibited during desorption/ionization processes. Compared to extensive chromatographic separation followed by mass-spectrometric detection, this cross-correlation approach requires minimal sample preparation. In addition, the analytes within samples were introduced simultaneously.

This cross-correlation approach utilizes the features of ion signals in the time domain to separate ions based on their time-

domain profiles (on second time-scale). Use of both reproducible and irreproducible sample introduction methods suggested that cross-correlation can be a generic approach for rapid MS-data analysis as long as there are some degrees of separation of ions. Essentially, the use of the time-domain information is the solution of the instrumental response function with the physical-chemical properties of the analytes as its variable. In specific, the instrumental response function in this study is the combination of ionization/desorption, ion transportation, mass separation, and ion detection. It is, thus, possible to modify this function by using, for instance, other ionization source. In particular, conventional LC-ESI-MS requires good separation in the LC stage. However, with cross-correlation-based data analysis approach, the LC gradient would not need to be fully optimized. Details of the applications of cross-correlation-based method for LC-MS will be discussed in a future publication.

Moreover, most of the datasets shown in this work were dependent on mass spectrometers of high-resolving power. In specific, the unit-resolution mass spectrometer was not perfectly capable of separating isobaric ions. This issue can strongly affect the cross-correlation analysis when background- and analyte-ions are of very similar m/z . However, the resolving-power issue can be potentially overcome mathematically (e.g., deconvolution).

The use of time-domain information also reveals a unique perspective regarding the ionization processes, especially with the recognition of low-abundant ions. Though the co-existence of analyte fragments and cluster ions was not discussed in detail here, the recognition of such ions may provide key information to understand ionization and fragmentation processes at atmospheric pressure.

Acknowledgements

We sincerely appreciate Dr. Leah Shriver and coworkers at the University of Akron for the kind donation of the γ -aminobutyric acid (GABA) standard used in this work.

References

1. Cooks, R.G., Ouyang, Z., Takats, Z., Wiseman, J.M.: Ambient mass spectrometry. *Science*. **311**(5767), 1566–1570 (2006)
2. Monge, M.E., Harris, G.A., Dwivedi, P., Fernández, F.M.: Mass spectrometry: recent advances in direct open air surface sampling/ionization. *Chem. Rev.* **113**(4), 2269–2308 (2013)
3. Nounou, P.: Mass spectrometry of aromatic compounds. 3. Ionization and dissociation of PHENANTHRENE and METHYLPHENANTHRENE molecules by electron impact. *J. Chim. Phys.* **65**(4), 700 (1968)
4. Munson, M.S., Field, F.-H.: Chemical ionization mass spectrometry. I. General introduction. *J. Am. Chem. Soc.* **88**(12), 2621–2630 (1966)
5. Chen, T.-C., Ouyang, Z.: Synchronized discharge ionization for analysis of volatile organic compounds using a hand-held ion trap mass spectrometer. *Anal. Chem.* **85**(3), 1767–1772 (2013)
6. Takats, Z., Wiseman, J.M., Gologan, B., Cooks, R.G.: Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science*. **306**(5695), 471–473 (2004)
7. Venter, A., Nefliu, M., Cooks, R.G.: Ambient desorption ionization mass spectrometry. *TrAC Trends Anal. Chem.* **27**(4), 284–290 (2008)

8. Cody, R.B., Laramée, J.A., Durst, H.D.: Versatile new ion source for the analysis of materials in open air under ambient conditions. *Anal. Chem.* **77**(8), 2297–2302 (2005)
9. Cody, R.B., Laramée, J.A., Nilles, J.M., Durst, H.D.: Direct analysis in real time (DARTtm) mass spectrometry. *JEOL news.* **40**(1), 8–12 (2005)
10. Andrade, F.J., Shelley, J.T., Wetzel, W.C., Webb, M.R., Gamez, G., Ray, S.J., Hieftje, G.M.: Atmospheric pressure chemical ionization source. 1. Ionization of compounds in the gas phase. *Anal. Chem.* **80**(8), 2646–2653 (2008)
11. Shelley, J.T., Wiley, J.S., Hieftje, G.M.: Ultrasensitive ambient mass spectrometric analysis with a pin-to-capillary flowing atmospheric-pressure afterglow source. *Anal. Chem.* **83**(14), 5741–5748 (2011)
12. Badal, S.P., Michalak, S.D., Chan, G.C.-Y., You, Y., Shelley, J.T.: Tunable ionization modes of a flowing atmospheric-pressure afterglow (FAPA) ambient ionization source. *Anal. Chem.* **88**(7), 3494–3503 (2016)
13. Schwartz, A.J., Shelley, J.T., Walton, C.L., Williams, K.L., Hieftje, G.M.: Atmospheric-pressure ionization and fragmentation of peptides by solution-cathode glow discharge. *Chem. Sci.* **7**(10), 6440–6449 (2016)
14. Nilles, J.M., Connell, T.R., Stokes, S.T., Dupont Durst, H.: Explosives detection using direct analysis in real time (DART) mass spectrometry. *Propellants Explos. Pyrotech.* **35**(5), 446–451 (2010)
15. Vaclavik, L., Rosmus, J., Popping, B., Hajslova, J.: Rapid determination of melamine and cyanuric acid in milk powder using direct analysis in real time-time-of-flight mass spectrometry. *J. Chromatogr. A.* **1217**(25), 4204–4211 (2010)
16. Crawford, E., Musselman, B.: Evaluating a direct swabbing method for screening pesticides on fruit and vegetable surfaces using direct analysis in real time (DART) coupled to an Exactive benchtop orbitrap mass spectrometer. *Anal. Bioanal. Chem.* **403**(10), 2807–2812 (2012)
17. Schilling, G.D., Shelley, J.T., Bames, J.H., Sperline, R.P., Denton, M.B., Barinaga, C.J., Koppelaar, D.W., Hieftje, G.M.: Detection of positive and negative ions from a flowing atmospheric pressure afterglow using a Mattauch-Herzog mass spectrograph equipped with a faraday-strip array detector. *J. Am. Soc. Mass Spectrom.* **21**(1), 97–103 (2010)
18. Cajka, T., Riddellova, K., Tomaniova, M., Hajslova, J.: Ambient mass spectrometry employing a DART ion source for metabolomic fingerprinting/profiling: a powerful tool for beer origin recognition. *Metabolomics.* **7**(4), 500–508 (2011)
19. Pan, Z., Li, Y., Deng, X., Xiao, S.: Non-targeted metabolomic analysis of orange (*Citrus sinensis* [L.] Osbeck) wild type and bud mutant fruits by direct analysis in real-time and HPLC-electrospray mass spectrometry. *Metabolomics.* **10**(3), 508–523 (2014)
20. Windig, W., Phalip, J.M., Payne, A.W.: A noise and background reduction method for component detection in liquid chromatography/mass spectrometry. *Anal. Chem.* **68**(20), 3602–3606 (1996)
21. Zhu, P.J., Ding, W., Tong, W., Ghosal, A., Alton, K., Chowdhury, S.: A retention-time-shift-tolerant background subtraction and noise reduction algorithm (BgS-NoRA) for extraction of drug metabolites in liquid chromatography/mass spectrometry data from biological matrices. *Rapid Commun. Mass Spectrom.* **23**(11), 1563–1572 (2009)
22. Shelley, J.T., Wiley, J.S., Chan, G.C., Schilling, G.D., Ray, S.J., Hieftje, G.M.: Characterization of direct-current atmospheric-pressure discharges useful for ambient desorption/ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **20**(5), 837–844 (2009)
23. Zhang, Y., Ma, X., Zhang, S., Yang, C., Ouyang, Z., Zhang, X.: Direct detection of explosives on solid surfaces by low temperature plasma desorption mass spectrometry. *Analyst.* **134**(1), 176–181 (2009)
24. Chipuk, J.E., Brodbelt, J.S.: Transmission mode desorption electrospray ionization. *J. Am. Soc. Mass Spectrom.* **19**(11), 1612–1620 (2008)
25. Evans-Nguyen, K.M., Quinto, A., Hargraves, T., Brown, H., Speer, J., Glatter, D.: Transmission mode desorption electrospray ionization (TM-DESI) for simultaneous analysis of potential inorganic and organic components of radiological dispersion devices (RDDs). *Anal. Chem.* **85**(24), 11826–11834 (2013)
26. Jones, C.M., Fernández, F.M.: Transmission mode direct analysis in real time mass spectrometry for fast untargeted metabolic fingerprinting. *Rapid Commun. Mass Spectrom.* **27**(12), 1311–1318 (2013)
27. Chipuk, J.E., Gelb, M.H., Brodbelt, J.S.: Rapid and selective screening for sulfhydryl analytes in plasma and urine using surface-enhanced transmission mode desorption electrospray ionization mass spectrometry. *Anal. Chem.* **82**(10), 4130–4139 (2010)
28. Schilling, G.D., Shelley, J.T., Broekaert, J.A.C., Sperline, R.P., Denton, M.B., Barinaga, C.J., Koppelaar, D.W., Hieftje, G.M.: Use of an ambient ionization flowing atmospheric-pressure afterglow source for elemental analysis through hydride generation. *J. Anal. At. Spectrom.* **24**(1), 34–40 (2009)
29. Allen, J., Berkley, D., Blauert, J.: Multimicrophone signal-processing technique to remove room reverberation from speech signals. *J. Acoust. Soc. Am.* **62**(4), 912–915 (1977)
30. Owens, K.G.: Application of correlation analysis techniques to mass spectral data. *Appl. Spectrosc. Rev.* **27**(1), 1–49 (1992)
31. Meehan, T. K.: Multipath noise reduction for spread spectrum signals. Google Patents. Retrieved from <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19950007215.pdf>. (1994)
32. Powell, L.A., Hieftje, G.: Computer identification of infrared spectra by correlation-based file searching. *Anal. Chim. Acta.* **100**, 313–327 (1978)
33. Stein, S.E., Scott, D.R.: Optimization and testing of mass spectral library search algorithms for compound identification. *J. Am. Soc. Mass Spectrom.* **5**(9), 859–866 (1994)
34. Mueller, C., Weinmann, W., Dresen, S., Schreiber, A., Gergov, M.: Development of a multi-target screening analysis for 301 drugs using a QTrap liquid chromatography/tandem mass spectrometry system and automated library searching. *Rapid Commun. Mass Spectrom.* **19**(10), 1332–1338 (2005)
35. Eng, J.K., McCormack, A.L., Yates, J.R.: An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**(11), 976–989 (1994)
36. Yates III, J.R., Eng, J.K., McCormack, A.L., Schieltz, D.: Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem.* **67**(8), 1426–1436 (1995)
37. Tabb, D. L.; Eng, J. K.; Yates III, J. R.: Protein identification by SEQUEST. In *Proteome Research: Mass Spectrometry*, pp. 125–142. Springer (2001)
38. Eng, J.K., Fischer, B., Grossmann, J., MacCoss, M.J.: A fast SEQUEST cross correlation algorithm. *J. Proteome Res.* **7**(10), 4598–4602 (2008)
39. Chemetsova, E.S., Crawford, E.A., Shikov, A.N., Pozharitskaya, O.N., Makarov, V.G., Morlock, G.E.: ID-CUBE direct analysis in real time high-resolution mass spectrometry and its capabilities in the identification of phenolic components from the green leaves of *Bergenia crassifolia* L. *Rapid Commun. Mass Spectrom.* **26**(11), 1329–1337 (2012)
40. Yu, S., Crawford, E., Tice, J., Musselman, B., Wu, J.-T.: Bioanalysis without sample cleanup or chromatography: the evaluation and initial implementation of direct analysis in real time ionization mass spectrometry for the quantification of drugs in biological matrixes. *Anal. Chem.* **81**(1), 193–202 (2008)
41. Gross, J.H.: Direct analysis in real time—a critical review on DART-MS. *Anal. Bioanal. Chem.* **406**(1), 63–80 (2014)
42. Husen, P., Tarasov, K., Katafiasz, M., Sokol, E., Vogt, J., Baumgart, J., Nitsch, R., Ekroos, K., Ejsing, C.S.: Analysis of lipid experiments (ALEX): a software framework for analysis of high-resolution shotgun lipidomics data. *PLoS One.* **8**(11), e79736 (2013)
43. Tyanova, S., Mann, M., Cox, J.: MaxQuant for in-depth analysis of large SILAC datasets. In: Warscheid B. (ed) *Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). Methods in Molecular Biology (Methods and Protocols)*, vol 1188. Humana Press, New York, NY 351–364 (2014)
44. Guo, X., Bruins, A.P., Covey, T.R.: Characterization of typical chemical background interferences in atmospheric pressure ionization liquid chromatography-mass spectrometry. *Rapid Commun. Mass Spectrom.* **20**(20), 3145–3150 (2006)
45. Kaiser, J., Reed, W.: Data smoothing using low-pass digital filters. *Rev. Sci. Instrum.* **48**(11), 1447–1457 (1977)
46. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G. METLIN: a metabolite mass spectral database. *Ther Drug Monit [Internet].* **27**, 747–751 (2005)
47. McEwen, C.N., McKay, R.G., Larsen, B.S.: Analysis of solids, liquids, and biological tissues using solids probe introduction at atmospheric pressure on commercial LC/MS instruments. *Anal. Chem.* **77**(23), 7826–7831 (2005)