Background ion detection algorithm in ultra-high-throughput QC with mass spectrometry

Gordana Ivosev[1]; Chang Liu[1]; Alandra Quinn[2]; Jeff Chin[2]; Stephen Tate[1]; Thomas R. Covey[1]; Hui Zhang[2]
[1]SCIEX, 71 Four Valley Drive, Concord, ON, L4K 4V8 Canada; [2]Pfizer Inc., Eastern Point Road, Groton, CT, 06340

ABSTRACT

Ultra-high-throughput (1-sec-per-sample or faster) screening with MS readouts became reality with new platforms such as acoustic ejection mass spectrometry (AEMS). The full-scan mode of MS (TOF scan, Q1 scan, etc.) is an attractive approach because of the simplicity of method development and it can support both targeted and untargeted assays. It is critical to differentiate the “real signal” from the “background signal”, especially for the untargeted assays (e.g. untargeted profiling screen, purity assessment, and unknown metabolite/by-product identifications). Here, we introduce a new algorithm and workflow to automatically identify the background ions for the MS spectra data sets collected in the high-throughput QC application. These ions are correctly scaled for each sample and subtracted from the raw data in an automatic fashion.

INTRODUCTION

All pharmalogical screening depends on high-quality compound libraries, and it is highly desired to run fast and reliable QC with the compound library to validate the screening results. LC-MS is widely used for small molecule library QC however the throughput is a bottleneck for big libraries. AEMS is a new platform that provides ultra-high throughput analysis capability[1-3]. In the AEMS system, nanoliter volume samples are introduced from acoustic droplet dispensers into a continuous liquid stream drawn into the ion source by the Venturi effect created with a high velocity pneumatic nebulizer common to ESI. Since its introduction, AEMS has been applied to various high-throughput drug discovery works, including high-throughput pharmacology screening, ADME screening, bioanalysis, parallel medicinal chemistry and compound QC [2,4]. In some of these applications, the full-scan mode of MS is used for both target species analysis and the identification of non-target ions.

Since the information extraction relies on the full scan mass spectra, it would be critical to differentiate the “signal ions” from the “background”, especially for the workflows requiring the analysis of non-targeted m/z (e.g. full spectra profiling for pattern matching, purity assessment for the compound QC workflow, and by-product identifications for reaction optimizations).

There are two types of “background” mass spectra, which could cause either false positives or false negatives of the analysis results. The first type was sample independent, whose appearance is constant, not synchronized with the sample introduction. This type of the background could be from the mobile phase (carrier solvent), ion source, and/or system contaminations. During the time period of the sample plug appears, this type of the “background” mass spectra could be as the consistent level as the non-sample period, or either being enhanced (e.g., due to the pH change of the sample plug) or suppressed due to the ionization suppression from the sample matrix.

Another type of the “background” was from the matrix of the sample solution (e.g. sample solvent peaks). This type of the background would have a different intensity between the sample plug timing period vs the non-sample timing period. It will be relatively constant for the samples from the same resource/assays, but may be different from batch to batch (e.g. samples dissolved in different lots of the solvents).

There are several reported approaches for identifying the mass spectra background and getting it subtracted from the sample spectra. One way is to get the mass spectra during the non-sample period as an estimate of the “type 1 background”. The limitations of this approach include it cannot get the type 2 background estimated, and it could not compensate the potential enhancement or suppression of these ions during the sample plug timing period. In addition, in some high-throughput analysis systems (e.g. AEMS), sample signals in the time domain are close to each other, without providing a stable non-sample period for background MS spectra extraction.

Another way is to get a separate “blank” ejection to estimate both type 1 and type 2 background. However, the additional sample prep/sampling would take effort. In addition, it would be challenging to get the ideal blank samples for all workflow/assays. This method would not be able to compensate the potential background scale variation among different samples either. In this study, we introduce a new method to automatically extracting both type 1 and type 2 background and getting the correctly scaled background subtracted for each sample.

MATERIALS AND METHODS

The reference standard plates (for the generation of the reference MS spectra) and the compound QC test plates were prepared in DMSO in the 384-well format. These plates were analyzed on a research prototype of Acoustic Ejection Mass Spectrometry (AEMS) system coupling with the SCIEX TripleTOF 6600 system. The MS was run in the TOF-scan mode. The MS data files each containing 384 ejections were processed with a research-version data splitting algorithm to assign the well-position to each MS signal peak. The split MS data were analyzed against the sample information (imported as the compound ID and/or chemical formular of each well) using the research-version data processing platform. Identified background signals are subtracted from the raw data after scaling.

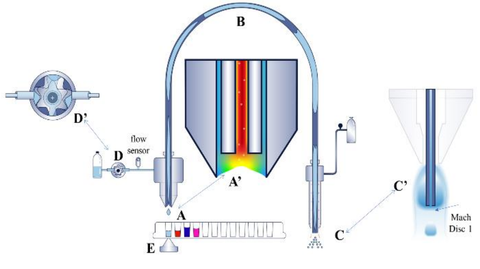

Figure 1. AEMS system. A. OPI capture port oriented downward. A’. Drawing of critical condition surface in “A”. B. OPI with a 50 cm transport tube. C. OPI venturi pump/ESI nebulizer. C’. Sonic expansion creating pressure drop. D. Fluid delivery pump. E. Acoustic dispensing upward against gravity.

RESULTS

In AEMS, the MS signal from different ejections are collected as a single data file, as shown in Figure 2. The TOF mass spectra of each ejection could be recorded for the follow up processing. In this work, the first step is to correlate the sample well position with each MS signal peak, with the synchronization of the MS signal with the sample ejection log .

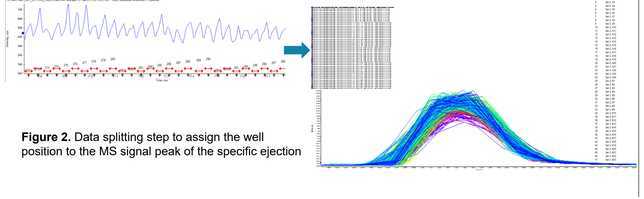

Figure 2. Data splitting step to assign the well position to the MS signal peak of the specific ejection

The mass spectra of each ejections from the same plate (with the similar background ions) are cross-compared for the automatic identification of background ion m/z, as shown in Figure 3.

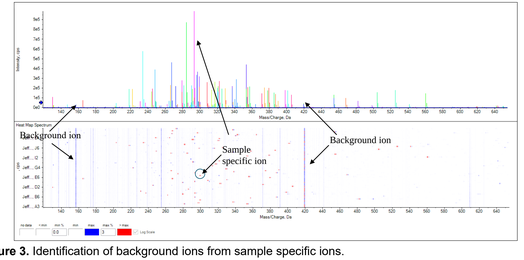

Figure 3. Identification of background ions from sample specific ions.

This method could accurately identify the background ions of both types (from sample matrix or carrier solvent). As shown in Figure 4, no matter the signal of these ions got enhanced or suppressed when the sample plug appears. It could be noted that the same ion may show different identities from various ejections. The ejection specific background scaling factor could be estimated for the subtraction process.

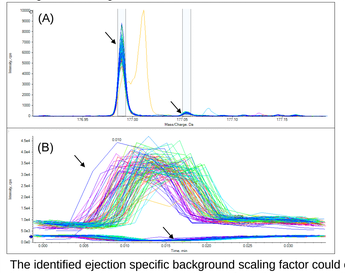

Figure 4. A. The overlapped mass spectra from multiple ejections, with two highlighted background ions in this mass range. B. Overlapped XIC chronograms of the two background ions highlighted in (A).

The identified ejection specific background scaling factor could ensure only the portion contributed from the background is subtracted, but not the part from the sample itself. For example, for the ejection containing the blue and orange color chronogram, the signal ions having the same or similar m/z with the background ions.

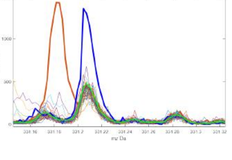

Figure 5. A. The overlapped mass spectra from multiple ejections

The identification and subtraction of the background ions from mass spectra could greatly simplify the MS spectra and provide the direct information for the assessment of compound purity. For example, Figure 6 shows two example compounds from two resources. The first compounds also showed different mass spectra from the two runs, but these distinguished ions are from the background. On the other hand, a major impurity ion was discovered on another compound.

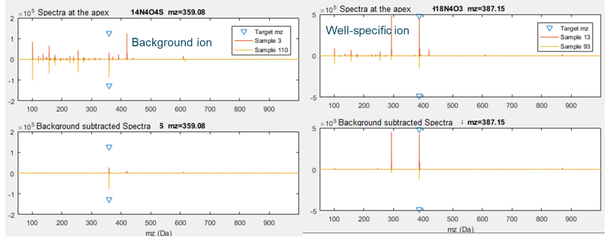

Figure 6. Original and background subtracted mass spectra for two compounds from different resources.

CONCLUSIONS

The automatic mass spectra background identification and subtraction is described in this study for high-throughput MS analysis. The background mass spectra and the sample-specific scaling factor was identified by automatic cross-comparing the samples from the same plates. This method has been used to process the data generated for compound QC workflow. After the subtraction of the identified background, the MS spectra from the well-specific target analytes is significantly simplified. The ions from the sample specific impurity or degradations could be successfully differentiated from the background for high-confident compound QC results. The platform introduced here could be used to automatically process the large-quantity high-throughput analysis data.

REFERENCES

1 Liu, et al. Anal. Chem. (2020) 92, 15818-15826.
2 Zhang, et al. Anal. Chem. (2021) 93, 10850-10861.
3 Häbe, et al. Anal. Chem. (2020) 92, 12242-12249.
4 Zhang, et al. SLAS Technol. (2021), 26, 178-188.

TRADEMARKS/LICENSING