Journal of Chromatography A, 1463 (2016) 153–161

Contents lists available at ScienceDirect

Journal of Chromatography A

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

A two stage algorithm for target and suspect analysis of produced water via gas chromatography coupled with high resolution time of flight mass spectrometry

Saer Samanipour*, Katherine Langford, Malcolm J. Reid, Kevin V. Thomas

Norwegian Institute for Water Research (NIVA), 0349 Oslo, Norway

ARTICLE INFO

Article history:
Received 30 March 2016
Received in revised form 14 July 2016
Accepted 27 July 2016
Available online 29 July 2016

Keywords:
Produced water
GC–HR-TOFMS
Dot product
Matching algorithm
Unique ion extractor
Reverse match
Suspect analysis
Target analysis

ABSTRACT

Gas chromatography coupled with high resolution time of flight mass spectrometry (GC–HR-TOFMS) has gained popularity for the target and suspect analysis of complex samples. However, confident detection of target/suspect analytes in complex samples, such as produced water, remains a challenging task. Here we report on the development and validation of a two stage algorithm for the confident target and suspect analysis of produced water extracts. We performed both target and suspect analysis for 48 standards, which were a mixture of 28 aliphatic hydrocarbons and 20 alkylated phenols, in 3 produced water extracts. The two stage algorithm produces a chemical standard database of spectra, in the first stage, which is used for target and suspect analysis during the second stage. The first stage is carried out through five steps via an algorithm here referred to as unique ion extractor (UIE). During the first step the *m*/*z* values in the spectrum of a standard that do not belong to that standard are removed in order to produce a clean spectrum and then during the last step the cleaned spectrum is calibrated. The Dot-product algorithm, during the second stage, uses the cleaned and calibrated spectra of the standards for both target and suspect analysis. We performed the target analysis of 48 standards in all 3 samples via conventional methods, in order to validate the two stage algorithm. The two stage algorithm was demonstrated to be more robust, reliable, and less sensitive to the signal-to-noise ratio (S/N), when compared to the conventional method. The Dot-product algorithm showed lower potential in producing false positives compared to the conventional methods, when dealing with complex samples. We also evaluated the effect of the mass accuracy on the performances of Dot-product algorithm. Our results indicated the crucial importance of HR-MS data and the mass accuracy for confident suspect analysis in complex samples.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Gas chromatography coupled with mass spectrometry (GC–MS) is one of the common analytical techniques for analysis of complex samples for volatile and semi-volatile compounds [1–5]. The three main approaches to perform this type of analysis are: target analysis, where the analytical standard of the analyte is available; suspect analysis, where the analytical standard is not available however information, such as exact mass and the fragmentation pattern is available for that analyte; and finally non-target analysis, where no prior information is available about the potential analytes of interest [6]. Confident detection of an analyte in a complex sample is a challenging task, particularly during suspect and non-target analysis [6,7]. The introduction of high resolution and/or high accuracy mass spectrometers improved drastically the levels of confidence in the suspect analysis, however difficulties still persist [6,8,9].

For target analysis, depending on the target analyte and the data processing tools used for analysis, few *m*/*z* values and the absolute retention time are used for identity confirmation of a target analyte in the sample [10–13]. Regarding suspect analysis, the identity confirmation is carried out employing either the direct analysis or reverse analysis [9,14,15]. Direct analysis consists of first performing mass spectral deconvolution of the suspect peak in the sample, and then comparing the deconvoluted spectra to a standard database [16–18] (e.g. Mass spectral library of National Institute of Standards and Technology, NIST [19]). As a result of the spectral comparison the chemical structures with the highest similarity score are reported as a hit list. Lu et al. demonstrated that the conventional deconvolution algorithms may cause introduction of

* Corresponding author at: NIVA, Gaustadalléen 21, 0349 Oslo, Norway.
E-mail address: saer.samanipour@niva.no (S. Samanipour).

http://dx.doi.org/10.1016/j.chroma.2016.07.076
0021-9673/© 2016 Elsevier B.V. All rights reserved.

artifacts into the final deconvoluted spectrum, depending on the complexity of the sample [20], which translates into errata library matching and scoring. In case of reverse analysis, the spectra of a chemical standard is compared to the whole chromatogram of the sample and where the analyte is present in the sample a higher level of similarity score is observed [21]. A large number of scoring systems have been developed and tested on different datasets (as reviewed by Scheubert et al. [9]). Amongst the tested scoring algorithms the dot product has been recognized as one of the most reliable matching methods, for both direct and reverse analysis [22,21,16]. The direct matching algorithms appear to be highly sensitive to the quality of deconvolution, spectral weighting function, binning step, and signal-to-noise ratio (S/N) [9,20,23]. Also the mentioned scoring systems often do not produce high enough levels of confidence in the detection [23]. The reverse matching method shown to be less sensitive to levels of S/N [9,14,24]. For example, in the study by Sinha et al. the authors were able to detect trimethylsilyl in urine samples by employing a unit mass spectra of trimethylsilyl and reverse dot product methodology [21]. The confidence in the detection for the reverse matching algorithms, is highly dependent to the quality and the levels of mass accuracy of the standard spectra [23,16]. Limited studies have focused on the matching algorithms for the GC–HR-MS data [22,24], particularly the reverse matching methodology, due to the lack of GC–HR-MS spectral database of standards.

Herein we report on a two stage algorithm for target and suspect analysis in complex samples using GC–HR-MS data. In the first stage the unique ions of a standard spectra are extracted from the raw data (via unique ion extractor algorithm, UIE) in order to produce a chemical standard database of HR spectra. In the second step the clean spectra of a target/suspect analyte is compared to the whole GC–HR-MS chromatogram of the sample employing reverse dot product methodology (via Dot-product algorithm). The comparison between the standard spectra and the sample spectra results in a similarity matrix with higher levels of similarity for the analytes which are present in the sample compared to the background signal. This approach was validated by comparing the results of the two stage algorithm to the conventional target and suspect analysis method. Higher levels of reliability and robustness were observed for the two stage algorithm when compared to the conventional methods. The validation was carried out through the analysis of 48 analytes in 3 produced water extracts. The produced water samples consisted of a total extract of produced water, the non-polar fraction of produced water, and the polar fraction of produced water. The produced water extracts provided a high level of complexity for the validation study, due to the commonalities in the fragmentation pattern of the target/suspect analytes and the background signal. The two stage algorithm proved to be able to distinguish the signal of target/suspect analytes from the background signal successfully. The two stage algorithm produced 0 cases of false positive compared to 1 via the conventional method. Moreover, this algorithm showed to be less sensitive to the levels of S/N.

2. Experimental

2.1. Chemicals

A mixture of 28 aliphatic hydrocarbons and 20 alkylated phenols were purchased from Sigma–Aldrich, Norway. A complete list of the standards is provided in the Supporting Information, Table S1. ACS grade ethanol, dichloromethane, methanol, hydrochloric acid, sodium hydroxide, and sodium sulphate were also obtained from Sigma–Aldrich. We obtained technical grade glass fiber filter (GF/C) from VWR, Norway.

For our analysis we used an extract of produced water. Produced water is a petrogenic by-product of offshore petroleum extraction. Produced water is a complex mixture containing thousands of compounds including heavy metals, hydrocarbons, phenols, organic acids, and oil production chemicals [11]. An extract of produced water at pH 2, using dichloromethane was provided by Stiftelsen for Industriell og Teknisk Forskning, Trondheim, Norway (SINTEF). Herein we refer to this sample as total extract. The extraction was performed according to the guidelines of Norwegian Environmental Protection Agency for the sampling and analysis of oil and gas [2]. In short 2.5 L of produced water was extracted employing 60 mL of dichloromethane, via liquid–liquid extraction, for three constitutive times. The final extract was dried using sodium sulphate.

An aliquot of the total extract was fractioned into polar and non-polar portions. For this fractionation, we dissolved 1 mL of the total extract into 1 L of water at pH 11, which was carried out by shaking the solution for 24 h at 150 rpm. This solution was extracted using liquid–liquid extraction with 60 mL of dichloromethane for three consecutive times. The final extract was dried on a bed of sodium sulphate. The volume of the final extract was reduced to 1 mL of dichloromethane employing a turbovap system under a gentile flow of N_2. For the non-polar fraction, the pH of the water was reduced to 1 from 11. The same liquid–liquid extraction procedure was carried out for the acidified sample. The final extract of the acidified sample was considered the non-polar fraction of the total extract.

All the extracts were stored immediately at $-20\,^{\circ}$C until analysis.

2.2. GC–HR-TOFMS analysis

We analyzed mixtures of standards at three concentration levels (2, 10, and 20 ng/mL), the total extract (i.e. the total extract of produced water received from SINTEF), and the polar and non-polar fractions of the total extract with a GC–HR-TOFMS (GCT Premier, Waters, USA) equipped with electron impact ion source (EI). The separations were carried out on a BD-5 column (30 m $\times$ 0.25 m $\times$ 0.25 mm, Agilent). All the injections were performed in splitless mode having an injection volume of 1 μL. Helium was used as the carrier gas. The TOFMS collected 2 spectra every second between 50 Da and 600 Da. The detector exhibited a resolution of $\sim$8000 at half width full range (i.e. 50–600 Da). The detector was operated at 2850 V and a filament current of $\sim$1 mA. More information about the instrumental setup is provided in Section S2 of Supporting Information.

2.3. Data analysis

The raw chromatograms were exported as netCDF files employing MassLynx (Waters, Manchester, UK). The raw chromatograms then were imported into matlab (R2015b) [25] for further processing. All the scripts for both the UIE and Dot-product algorithms were developed in matlab. As a validation tool for UIE algorithm as well as the target analysis, we used the software package TargetLynx (Waters, USA) within the Masslynx. A target analyte was considered detected in TargetLynx if we observed a positive match between the retention times $\pm$5 s and the exact mass $\pm$10 mDa of the standard and the target peak in the sample. Both the retention window and the exact mass window were selected based on the observed variabilities in our dataset for these parameters. The minimum S/N required for a positive detection was set to 10.

The S/N calculations were performed via MassLynx. The signal was defined as the 50% of the peak height whereas the noise was defined as the root mean square error of the 10 scans in one side of the peak. The ratio of these two values resulted in the S/N.

All the calculations were performed on a personal computer with an Intel i7, 2.8 GHz processor, and 16 GB of memory. The operating system was Windows 7 enterprise version.

Download English Version:

https://daneshyari.com/en/article/1198341

Download Persian Version:

https://daneshyari.com/article/1198341

Daneshyari.com