Talanta 103 (2013) 384–391

Contents lists available at SciVerse ScienceDirect

Talanta

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

Qualitative nontarget analysis of landfill leachate using gas chromatography time-of-flight mass spectrometry

Joonas Jernberg *, Jukka Pellinen [1], Anna-Lea Rantalainen [1]

Department of Environmental Sciences, University of Helsinki, Niemenkatu 73, Lahti FI-15140, Finland

ARTICLE INFO

Article history:
Received 1 August 2012
Received in revised form
24 October 2012
Accepted 29 October 2012
Available online 6 November 2012

Keywords:
Gas chromatography
Time-of-flight mass spectrometry
Nontarget analysis
Accurate mass
Landfill leachate

ABSTRACT

Nontarget analysis means that a sample is analysed without preselection of the studied analytes. While target analysis attempts to determine whether certain selected compounds are present in the sample, nontarget analysis is performed to explore what unknown compounds can be found. We developed a nontarget method using a landfill leachate sample as a complex test sample. The method was based on the use of a gas chromatograph–time-of-flight mass spectrometer (GC–TOF–MS) for final analysis and a deconvolution computer application for data processing. This nontarget analysis method was tested and validated by applying it to a landfill leachate sample spiked with 11 organic pollutants that were treated as unknowns. Sensitivity was found to be the most critical parameter affecting the success of nontarget analysis. The limit of identification (LOI) was 2500 ng L^{-1} for four of the 11 compounds, 500 ng L^{-1} for three compounds and 100 ng L^{-1} for one compound. Three compounds were not detected in any of the spiked samples. A six-stage identification process was developed based on the spiking experiments. The process was based on the forward fit value of the library hit, the number of deconvoluted ions and the accurate mass scoring of the measured ions. The process was applied to an unspiked leachate water sample. Altogether, 44 compounds were tentatively identified in the sample. Elemental compositions of 36 components were additionally determined for which an unequivocal compound identification could not be given. Nontarget analysis with GC–TOF–MS is a promising method for the qualitative analysis of complex water samples. However, we conclude that the computer application for nontarget analysis needs improvement to decrease the amount of manual work needed in the identification process.

1. Introduction

The increased global use of chemicals has caused an emerging concern among scientists and policymakers concerning resultant environmental pollution [1]. The number of chemicals used worldwide is steadily increasing, and novel compounds are continuously being synthesized. As a result, trace levels of organic compounds are found in all spheres of the environment. The emerging contaminants are structurally diverse and heterogeneous group of chemical compounds which are currently not covered by existing regulations or legislation, have not been widely studied and are believed to pose a threat to ecosystems [2]. Richardson [3] and Wille et al. [4] recently published high-quality reviews about the current status of the analysis of emerging contaminants in the environment. Some of the emerging contaminants are used in domestic households as the ingredients of common consumer products. Municipal solid waste and wastewater therefore provide a possible route for emerging contaminants to enter into the environment [5–10]. Stormwater in urban areas additionally forms a little-studied route for chemicals into the water system [11–13].

The most common GC–MS approach in use is still electron ionization (EI) and quadrupole analyser in selected ion monitoring (SIM) mode. A higher sensitivity may alternatively be gained with chemical ionization (CI) for some analytes, producing intensive adduct of molecular ion [14,15]. Environmental scientists have recently shown also an increased interest in GC tandem MS [16–19]. With these techniques, the target ions are determined before the analysis and during the data acquisition all other ions are excluded. Thus other instrumentation in which full spectrum data are collected and used to identify sample components are needed for the nontarget analysis instead of quadrupole mass spectrometers.

All ions with a mass-to-charge ratio (m/z) within the defined mass range eluting from the column and ionizing in the ionization chamber are measured in full spectrum techniques. These techniques allow the application of post-target and nontarget analysis [20]. One additional advantage of the comprehensive datasets produced in these analyses is the enablement of retrospective sample reanalysis, even a long time after data acquisition [21,22]. Modern gas chromatography-time-of-flight (GC-TOF) MS instruments are well suited for these purposes, as they provide high mass

* Corresponding author. Tel.: +358 919120378.
E-mail address: joonas.nurmi@helsinki.fi (J. Jernberg).
[1] These authors contributed equally to this work.

0039-9140/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.10.084

accuracy (typically below 5 ppm) and mass resolution (> 7000 full width at half maximum height (FWHM)) combined with high full-spectrum sensitivity and speed. High mass resolution enables the use of narrow-window extracted ion chromatograms (nw-XIC) of 20–50 mDa, which efficiently reduces background noise and enhances the signal-to-noise ratio of the analyte. In nontarget analysis using GC–TOF–MS, the peaks of previously unknown components in the sample chromatogram are extracted from the full spectrum data, using special deconvolution software that detects and combines ions arising from the same component without any previous information about the compound. The formed spectrum is then searched against the spectral library, and component identification is confirmed by comparing the measured accurate masses of molecular and fragment ions with the corresponding exact masses of the library hit compound. Applicability of GC–TOF–MS for the nontarget analysis of complex sample matrices like human breast adipose tissue [23] and honeybee samples [24] has recently been reported. More information on the analytical strategies using GC–TOF–MS can be found in an extensive review of Hernández et al. [25].

Emphasis of the previously published research using GC–TOF–MS has so far focused on the post-target analysis of different contaminants in environmental samples. Identification has been based on the utilisation of the nw-XICs, and ion ratios of the selected masses corresponding to the analytes of interest are extracted from the complete dataset [26–29]. Recently, Portolés et al. [26] reported a qualitative wide-scope post-target screening method in which approximately 150 contaminants were examined from water samples with GC–TOF–MS. The potential of nontarget screening has briefly been demonstrated in some previous papers [26,28,29], but so far only a few applications concentrating purely on nontarget analysis of aquatic samples using GC–TOF–MS have been reported [30,31].

The need for nontarget methods in environmental analysis is indisputable. However, their use is still very limited as they are considered very laborious and ineffective. Additionally, the high mass resolution instrumentation required for nontarget screening is uncommon. Spectra produced by GC-(EI)–MS contain more information for nontarget screening because of the natural fragmentation in EI compared to the LC-MS data. Availability of extensive commercial spectral libraries additionally facilitates identification. The main objective of our study was to develop a systematic procedure for the qualitative nontarget analysis of emerging contaminants in a complex liquid environmental sample, using a GC–TOF–MS instrument and deconvolution software. Furthermore, our other goals were to estimate identification reliability and explore the possible limitations of nontarget analysis. A landfill leachate sample was selected as the complex matrix for method development. Generic liquid–liquid extraction (LLE) was used without optimization for any particular group of compounds. Identification of the sample components through nontarget analyses was tested and validated, using a leachate sample spiked with a mixture of 11 semi-volatile organic compounds. Finally, the nontarget method was applied for the analysis of a real unspiked sample. In addition to the analysis itself, the performance and features of the deconvolution application were evaluated and some improvement proposals are presented.

2. Materials and methods

2.1. Reagents and reference standards

The method 526 calibration mixture of the Environmental Protection Agency (EPA), containing acetochlor (CAS 34256-82-1), cyanazine (CAS 21725-46-2), diazinon (CAS 333-41-5), 2,4-dichlorophenol (CAS 120-83-2), 1,2-diphenylhydrazine (CAS 122-66-7), disulfoton (CAS 298-04-4), fonofos (CAS 944-22-9), nitrobenzene (CAS 98-95-3), prometon (CAS 1610-18-0), terbufos (CAS 13071-79-9) and 2,4,6-trichlorophenol (CAS 88-06-2) in ethyl acetate at a concentration of 200 $\mu g\,mL^{-1}$ was purchased from Ultra Scientific (North Kingstown, RI, USA). Working and spiking solutions of the EPA-mixture were prepared in ethyl acetate and methanol using volumetric dilution. Triethyl phosphate and tris(1-chloro-2-propyl)phosphate were purchased from Dr. Ehrenstorfer GmbH (Augsburg, Germany). Dichloromethane (DCM) and *n*-hexane for organic residue analysis were purchased from Mallinckrodt Baker B.V. (Deventer, Holland). LC-MS-grade methanol (MeOH) and anhydrous granular sodium sulphate $\geq 99\%$ was purchased from Sigma-Aldrich (Munich, Germany). SupraSolv®-grade ethyl acetate was purchased from Merck KGaA (Darmstadt, Germany). Ultrapure water was produced by PURELAB Ultra from ELGA Process Water (Marlow, UK).

2.2. Sampling and sample preparation

The landfill leachate samples were collected as grab samples from a reservoir at an active municipal landfill in the City of Lahti, Southern Finland. In physical appearance the leachate was a strong-smelling and black turbid liquid. Samples were stored in the dark in 2.5-L amber glass bottles at 4 °C before analysis. Prior to extraction, the water samples were filtered with 1.6-μm GF-A and 0.7-μm GF-F fibreglass filters from Whatman (Maidstone, Kent, UK). The samples were concentrated using LLE. The laboratory blanks of ultrapure water were processed in parallel. An aliquot of 100 mL leachate sample was first extracted with 50 mL of *n*-hexane and then with 50 mL of DCM. After LLE the extracts were concentrated to about 5 mL in a rotary evaporator and transferred into glass tubes. Approx. 1 g of anhydrous sodium sulphate (dried in a muffle furnace at 400 °C for 3 h) was added to the solution, which was left to stand overnight. The sample was decanted and concentrated under a gentle nitrogen flow at 35 °C to a final sample volume of 500 μL.

2.3. Instrumentation and analytical conditions

Leachate sample analyses were performed using a gas chromatograph orthogonal-acceleration time-of-flight mass spectrometer GCT Premier (Micromass® MS Technologies, Manchester, UK) equipped with a GC Pal injection system (CTC Analytics, Zwingen, Switzerland). Separation of the sample compounds was carried out using a ZB-5MSi column (30 m × 0.25 mm × 0.25 μm) from Phenomenex (Torrance, CA, USA) with a deactivated guard column (2 m × 0.25 mm) from Phenomenex. Helium was used as the carrier gas in a constant flow mode of 1.0 $mL\,min^{-1}$. The samples were injected, using the splitless injection technique with an injection volume of 1.0 μL. The purge valve was opened after an injection time of 1 minute. The inlet temperature was 280 °C. For the DCM extract, the oven temperature program was set as follows: 30 °C (hold 1 min), to 320 °C at an increase of 10 $°C\,min^{-1}$ and hold for 8 min. For the *n*-hexane extract, the oven temperature program was as follows: 50 °C (hold 1 min), to 320 °C at an increase of 10 $°C\,min^{-1}$ and hold for 8 min. The temperature of the transfer line into the mass spectrometer was 300 °C. The TOF–MS was operated in EI mode at 70 eV. The source temperature was 200 °C and the detector voltage 2600 V. The resolution of the instrument was > 7000 FWHM. The acquisition rate was 0.09 s per scan with an interscan delay of 0.01 s between scans. The measured mass range (m/z) was 50–550 in centroid mode. A solvent delay period of 5 min was used at the beginning of the analytical run. The accuracy of mass measurements is improved in GCT Premier by using a fixed lock mass ion, which is automatically applied to correct for any possible

drift of the mass axis. The lock mass ion is produced by continuously injecting a lock mass compound into the ion source from an internal reservoir. In our study, we used the lock mass ion of m/z 218.9856 from heptacosafluorotributylamine (CAS 311-89-7).

2.4. Nontarget screening method and identification criteria

The GC–TOF–MS instrument was controlled using MassLynx V4.1 software from Waters (Milford, MA, USA). The acquired data were processed using ChromaLynx XS, one of the application managers of MassLynx. ChromaLynx XS software is a peak detection and spectral deconvolution tool with automatic library searching, screening and comparison features. In the nontarget screening method, a nw-XIC of $\pm$20 mDa was used. The maximum number of ions extracted per component was set to 8. A spectra rejection tool was used to discard the adjacent components if their spectra were very similar, thus leading the software to not report the same component several times. The rejection tool parameters were as follows: the matching factor (forward fit) was 800, the scan width was 3 and the number of significant ions was 10. The formed component spectrum was searched against the commercial NIST08 library, producing the eight most relevant library matches according to the forward fit parameter. The accurate mass scoring was done for up to six of the most intense ions. The software calculates whether the accurate mass of some substructure of the proposed compound corresponds to the accurate mass measured in user-defined limits. The result was presented as the difference between the calculated exact and the measured mass (Δm). The layout of nontarget screening results in ChromaLynx is illustrated in Fig. 1. The software firstly integrates nominal mass ion chromatograms over the defined mass range to find the location of the chromatographic peaks (Fig. 1A). Each ion of every peak is then examined, to determine whether it maximizes at that retention time. Ions that have not maximised are excluded from the spectrum, and a user-defined number of maximizing ions in decreasing order of intensity are included in the final nominal mass spectrum of the component (Fig. 1B and D). Up to eight ions can be included in one spectrum. The formed spectrum is then compared with library spectra to find a possible identification for the component (Fig. 1E and F). The software reports a user-defined number of matching library hits according to their forward fit values. Finally, the software calculates the Δm values for ions detected in the component spectrum (Fig. 1C).

To test the ability of the nontarget method to detect and identify compounds, the leachate sample was spiked with the EPA standard mixture at concentrations of 100, 500 and 2500 ng L^{-1}, analysed with GC–TOF–MS and processed using the nontarget screening method. Finally data from the unspiked leachate samples were processed using the nontarget method, without any assumptions about sample composition. For identifying unknowns, a six-stage process was developed to diminish the number of suggested components and to remove false positive identifications (Fig. 2). All components with a forward fit value of less than 700 for the best library match were rejected in the first stage. In the second stage, components containing less than three ions deconvoluted in a spectrum were rejected. In the third stage, components with an ion abundance value of less than 10 were left

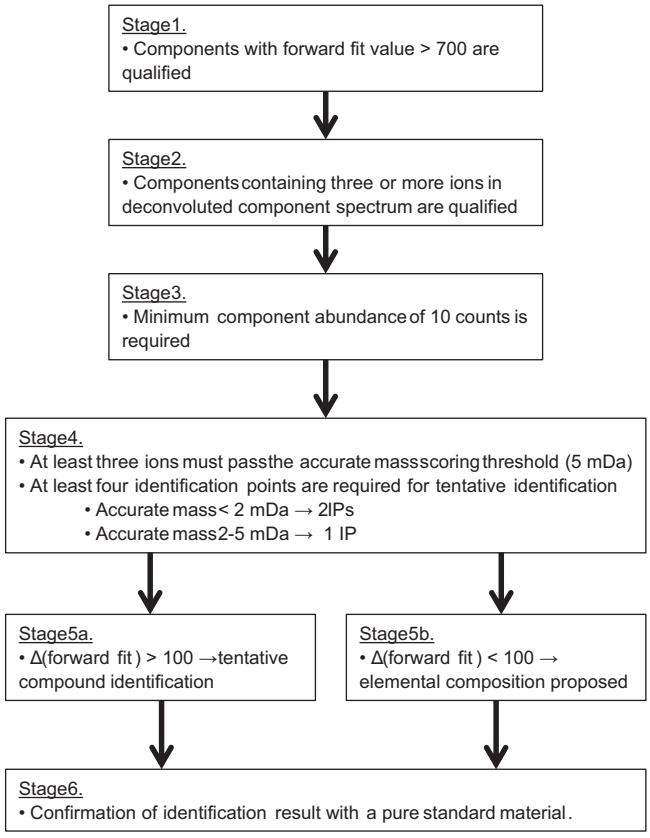

Fig. 2. Workflow of the nontarget identification process.

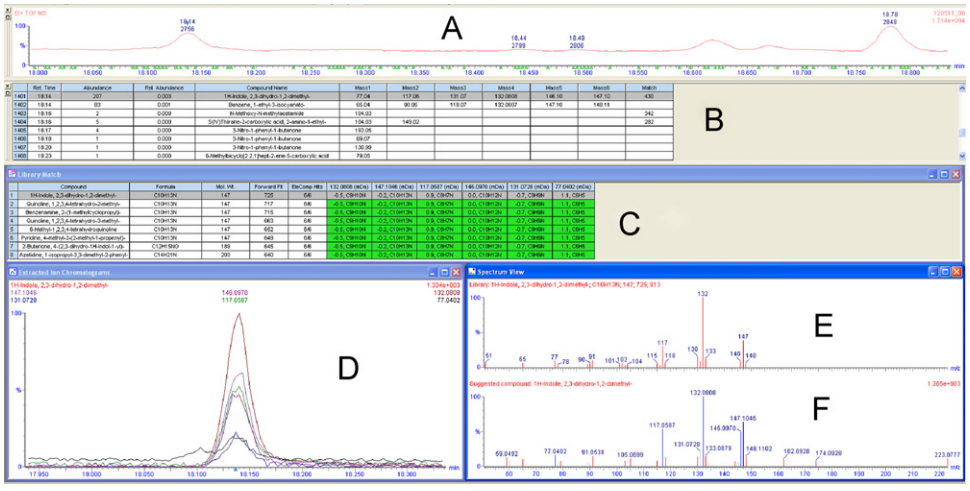

Fig. 1. The layout of nontarget screening results in ChromaLynx software. (A) Sample chromatogram, (B) components/ions deconvoluted, (C) accurate mass scoring table, (D) extracted ion chromatograms of the selected component, (E) library spectrum of the selected library hit and (F) deconvoluted sample spectrum.

out. In the fourth stage, at least three of the deconvoluted ions had to pass the Δm limit of 5 mDa and gain at least four identification points (IP) to proceed to the next stage. In the fifth stage, two alternative cases emerged. If the difference in forward fit values between the best and second best library match (Δ(forward fit)) was more than 100, the component was considered tentatively identified (stage 5a). If the Δ(forward fit) was less than 100 (i.e., components having similar EI spectra) and all components qualified had the same elemental formula, the identity of the best library match was reported to illustrate one possible component structure (stage 5b). The final (sixth) stage of the process was confirmation of the tentative identification with a pure standard material, if available.

3. Results and discussion

3.1. Method performance and validation

A relevant question in nontarget analysis is whether the analytes present in the sample can be detected without any preselection. This ability was assessed by spiking the raw leachate sample with EPA standard mixture at concentrations of 100, 500 and 2500 ng L^{-1} and processing the produced data with ChromaLynx. A diluted sample of EPA mixture (0.5 ng μL^{-1}) was additionally analysed and processed in parallel. Results of the experiment are summarized in Table 1. We define the lowest concentration at which a compound was detected and identified

Table 1
Method performance test results for the landfill leachate sample spiked at 2500, 500 and 100 ng/L.

Compound	Forward fit	m/z 1[a]	m/z 2	m/z 3	m/z 4	m/z 5
Nitrobenzene (8.02 min)		123.0320	77.0391	65.0391	51.0235	107.0371
		C6H5NO2 (M)[b]	C6H5	C5H5	C4H3	C6H5NO
Standard	830	−0.4	−1.0	0.1	0.2	0.1
2500 ng/L	689	−0.2	−0.1	−0.6	−0.3	n.d.[c]
2,4-Dichlorophenol (9.24 min)		161.9639	163.9610	97.9923	63.0235	125.9872
		C6H4O[35Cl2] (M)	C6H4O[35Cl][37Cl]	C5H3[35Cl]	C5H3	C6H3O[35Cl]
Standard	926	−0.2	−0.6	−0.6	0.4	0.5
2500 ng/L	865	−0.3	−0.3	−0.7	−0.6	0.3
500 ng/L	660	0.2	1.3	1.4	n.d.	n.d.
2,4,6-Trichlorophenol (11.88 min)		195.9249	197.9220	131.9534	96.9845	199.9190
		C6H3O[35Cl3] (M)	C6H3O[35Cl2][37Cl]	C5H2[35Cl2]	C5H2[35Cl]	C6H3O[35Cl][37Cl2]
Standard	919	−0.8	0.6	−0.5	0.9	0.6
1,2-Diphenylhydrazine[d] (15.29 min)		182.0844	77.0391	105.0453	152.0626	–
		C12H10N2 (M)	C6H5	C6H5N2	C12H8	–
Standard	895	−0.7	−2.8	−0.4	−0.2	–
2500 ng/L	901	3.8	−0.4	−1.3	n.d.	–
500 ng/L	718	0.3	−0.3	0.0	n.d.	–
100 ng/L	593	0.6	0.1	n.d.	n.d.	–
Prometon (16.54 min)		225.1590	210.1355	168.0885	183.1120	141.0651
		C10H19N5O (M)	C9H16N5O	C6H10N5O	C7H13N5O	C4H7N5O
Standard	661	−0.3	−0.8	2.8	−0.6	0.8
Terbufos (16.98 min)		230.9739	57.0704	103.0581	96.9513	153.0139
		C5H12O2PS3	C4H9	C5H11S	H2O2PS	C4H10O2PS
Standard	842	0.2	−0.3	−0.2	0.2	1.1
2500 ng/L	418	0.8	n.d.	0.3	0.9	1.6
Fonofos (17.06 min)		246.0302	108.9877	137.0190	110.0190	80.9564
		C10H15OPS2 (M)	C2H6OPS	C4H10OPS	C6H6S	H2OPS
Standard	859	0.2	0.1	1.6	−0.6	0.9
2500 ng/L	834	0.3	−2.0	0.0	0.9	0.8
500 ng/L	676	−0.4	−0.6	0.1	−0.8	0.1
Diazinone (17.20 min)		304.1011	179.1184	137.0715	152.0950	199.0636
		C12H21N2O3PS	C10H15N2O	C7H9N2O	C8H12N2O	C8H12N2O2P
Standard	744	−0.5	−0.4	−0.3	−0.8	0.4
2500 ng/L	740	−0.7	0.0	−0.1	−0.6	0.4
Disulfoton (17.32 min)		88.0347	89.0425	96.9513	141.9676	185.9938
		C4H8S	C4H9S	H2O2PS	C2H7OPS2	C4H11O2PS2
Standard	653	−0.1	−0.5	0.0	1.5	0.1
2500 ng/L	794	−0.8	−0.4	−0.2	0.8	n.d.
500 ng/L	596	0.1	−0.4	1.6	1.9	n.d.
Acetochlor (18.13 min)		132.0813	146.0970	147.1048	117.0578	162.0919
		C9H10N	C10H12N	C10H13N	C8H7N	C10H12NO
Standard	663	0.1	0.3	−0.4	2.2	0.9
2500 ng/L	467	n.d.	1.1	n.d.	1.9	n.d.
Cyanazine (19.11 min)		212.0703	213.0781	225.0655	214.0673	198.0546
		C8H11N5[35Cl]	C8H12N5[35Cl]	C8H10N6[35Cl]	C8H11N5[37Cl]	C7H9N5[35Cl]
Standard	508	0.9	0.2	−0.4	4.3	1.5

[a] Exact mass (Da), the elemental composition of the ion and experimental Δm (mDa) are reported one below the other for each ion.
[b] Molecular ion.
[c] Not detected.
[d] Compound detected as azobenzene (CAS 103-33-3).

following the six-stage process as the limit of identification (LOI) of the compound. LOI was 2500 ng L^{-1} for 4 out of 11 compounds, 500 ng L^{-1} for 3 out of 11 compounds and 100 ng L^{-1} for 1 out of 11 compounds (see Table 1). Three compounds (cyanazine, 2,4,6-trichlorophenol (2,4,6-TCP) and prometon) were not detected in any of the spiked samples. Cyanazine and prometon show high fragmentation levels in EI, which generally leads to a lower intensity of singular ions. This hinders their detection in nontarget screening, since the component spectrum is formed by combining single ions. In fact, these compounds even produced signals of low intensity in the analysis of pure EPA standard solution mixture without any sample matrix. On the contrary, compounds with the lowest LOIs only showed a few relatively intensive fragments. It can be concluded that the sensitivity of an analyte in EI is a critical parameter, affecting the success of nontarget detection. Results also show the concentration effect on the forward fit parameter value. A clear decreasing trend can be seen in forward fit values when concentrations approach LOI. At low concentrations, more characteristic ions are buried in background noise and are not detected by the software. This decreases the number of ions in a deconvoluted spectrum, weakening the probability of identification. The LOI values determined here represent a very complex sample matrix, and thus it is likely that lower LOIs are gained when analysing cleaner samples.

In a complex sample, like the leachate, co-eluting compounds may hamper identification of the sample components. This particular problem was evident with 2,4,6-TCP during the method performance test, as it was not detected in any of the spiked leachate samples despite it showing a relatively intensive molecular ion peak in its EI spectrum. The detection of 2,4,6-TCP failed due to a co-elution with some unknown component that had better sensitivity or which was present in much higher concentrations. The software was unable to separate these two compounds as individual components. Thus, a summation spectrum of two compounds was created, and obviously no relevant library hits were reported. With unknown samples, the mixing of EI spectra as a consequence of the co-eluting components is an undeniable weakness of nontarget analysis. As a result, false negative identifications may occur as in the case of 2,4,6-TCP. When noticed, the co-elution issue may sometimes be solved by changing the oven temperature program or the type of analytical column.

The Δm values of up to five characteristic ions of spiked compounds are reported in Table 1. Results show that the measured Δm values were generally below 1.0 mDa. In addition, no decline in Δm values was seen towards lower concentrations. On the contrary, the high end of the concentration range is often more critical in TOF analyses as the mass accuracy of the TOF instrument decreases with high analyte concentrations. This is due to the saturation of time-to-digital converter (TDC) widely used in orthogonal-acceleration TOF–MS. When the TDC is triggered by an ion, resulting in a time period of a few nanoseconds (i.e., dead time), during which the TDC is unable to register additional counts. When the number of ions increases (higher analyte concentration), more and more of them fall in the dead time period and so remain undetected. This results in a shift of the mass peak towards lower mass. Peak intensity is additionally suppressed because the ions have coinciding arrival times resulting in a single response. Dead time correction algorithms in the software can be used to correct this inherent feature of TDC, which is an ion counting device, to some extent. However, it still affects data when high sample concentrations are analysed and should be taken into account. In the data processing stage, ChromaLynx automatically reports whether the signal of a peak has been saturated.

3.2. Nontarget identification of leachate sample

Deconvolution software used in nontarget screening produce a substantial amount of data. According to our experiences, the majority of the data originates from sample background noise and the portion representing real analytes is quite low. Thus, it is necessary to create criteria for systematically reducing the number of proposed components and discarding false identifications. We developed a six-stage process (Fig. 2) in which the components found by the software were assessed according to certain criteria, and only qualifying components were accepted to the next stage. In the first stage, a forward fit limit of 700 was set. This selection was based on the results gained from the method performance test (Table 1). Some of the spiked analytes also showed lower forward fit values than 700; however we wanted to keep the limit somewhat higher because at lower fit values identification reliability is greatly reduced. The forward fit criterion was selected as the first stage, as it was the only criterion which could be set in the software. All other stages had to be manually processed.

The criterion in the second stage was based on the number of ions in the deconvoluted spectrum. EI usually produces some fragment ions and at least three ions were required for a spectrum to qualify. We noticed that the software reports a number of components consisting of only one ion (Fig. 1B). Generally all these originated from noise or lock mass ions and were thus disregarded at this stage.

In the third stage, all components with very low abundance (≤ 10) were discarded. The selection of this value was purely experimental, and should be optimised according to the sample matrix. The use of abundance as one criterion was based on the observation that at low concentrations the quality of a spectrum was rarely sufficient for reliable identification.

The fourth stage of the process utilizes the accurate mass feature of TOF–MS. The software proposes elemental compositions for fragment ions found and calculates their Δm values. At least three ions were demanded to pass the accurate mass scoring within a limit of 5 mDa. In addition, the calculation of IPs was added to our process to enhance the plausibility of the identification. The implementation of IPs is based on the European Commission Guideline for identification and quantification of organic residues [32]. The accumulation of IPs is based on the number of measured ions. In addition to IPs, the qualification of identification requires the fulfilment of certain criteria of ion ratios. However, the use of accurate mass measurements in terms of IPs is not included in this statute. Instead, the use of a high-resolution MS is defined but the current TOF–MS instruments do not fulfil the prescribed criterion (resolution $>$ 10,000 for the entire mass range at 10% valley). Complementary criteria for accurate mass measurements have been proposed in the literature [33,34]. These criteria have mainly been applied in target and post-target analysis, where pure standard compounds have been available and identification is usually supported by measuring ion ratios. In nontarget analysis, this is only possible after tentative identification. The calculation principles of IPs used in our study were slightly modified from a previous paper [28]. Mass error below 2.0 mDa gained two IPs per ion and mass error from 2.0 to 5.0 mDa gained one IP. At least four IPs were required for a component to be considered tentatively identified and to be reported.

Two alternative cases were possible in the fifth stage. If the difference in forward fit parameter values between the best and second best library match (Δ(forward fit)) was more than 100, the component was considered tentatively identified (stage 5a). The compounds identified in this category are summarized in Table 2. If the Δ(forward fit) was less than 100 and all components

Table 2
Compounds tentatively identified in the landfill leachate sample according to the identification process.

Retention time (min)	Tentative compound identification	CAS-number	Elemental formula	Forward fit	Number of ions	IPs
n-Hexane extract						
7.82	Methane, tert-butoxymethoxy	24209-75-4	C6H14O2	820	3	6
7.91	Butanenitrile, 4-(methylthio)-	59121-24-3	C5H9NS	711	4	7
7.96	Dimethylphenylmethanol	617-94-7	C9H12O	795	6	12
8.39	Phenylethyl alcohol	60-12-8	C8H10O	906	4	6
8.52	Phosphoric acid, diethyl pentyl ester	20195-08-8	C9H21O4P	862	4	8
8.85	Diethyl trisulphide	3600-24-6	C4H10S3	744	6	12
9.32	Methyl *n*-hexyl disulphide	64580-52-5	C7H16S2	738	5	9
9.48	1-Methyl-4-(1-hydroxy-1-methylethyl)benzene	1197-01-9	C10H14O	758	6	12
9.49	Methyl *n*-hexyl disulphide	64580-52-5	C7H16S2	713	5	10
10.59	Ethyl hexyl disulphide	64580-53-6	C8H18S2	789	3	6
11.22	3-Hexanethiol	1633-90-5	C6H14S	798	6	12
11.54	Dipropyl trisulphide	6028-61-1	C6H14S3	768	6	11
11.79	*n*-Propyl *n*-hexyl disulphide	64580-54-7	C9H20S2	739	4	8
14.36	1,6-Dioxacyclododecane-7,12-dione	777-95-7	C10H16O4	769	6	12
14.41	Dihexyl disulphide	10496-15-8	C12H26S2	708	6	12
15.44	Dihexyl disulphide	10496-15-8	C12H26S2	731	5	10
15.48	Lenthionine	292-46-6	C2H4S5	876	5	10
15.62	Benzenesulfonamide, *N*-ethyl-2-methyl	1077-56-1	C9H13NO2S	789	6	11
16.18	Benzoic acid, 2-ethylhexyl ester	5444-75-7	C15H22O2	748	4	8
16.28	Benzenesulfonamide, 4-methyl-*N*-propyl	1133-12-6	C10H15NO2S	932	4	7
16.49	Hexathiepane	17233-71-5	CH2S6	814	5	10
16.89	2-Hexanethiol	1679-06-7	C6H14S	800	5	10
17.16	Tris(1-chloro-2-propyl)phosphate[a]	13674-84-5	C9H18Cl3O4P	783	6	12
19.82	Cyclic octaatomic sulphur	10544-50-0	S8	882	5	9
Dichloromethane extract						
9.93	Butanenitrile, 4-(methylthio)-	59121-24-3	C5H9NS	871	5	11
10.42	Phenylethyl alcohol	60-12-8	C8H10O	957	4	8
10.54	Triethyl phosphate[a]	78-40-0	C6H15O4P	954	6	12
11.42	2-Piperidinone	675-20-7	C5H9NO	892	6	11
12.02	Ethanol, 2-phenoxy	122-99-6	C8H10O2	814	6	12
12.44	Ethosuximide	77-67-8	C7H11NO2	800	6	12
12.49	Caprolactam	105-60-2	C6H11NO	800	5	9
15.93	Acetamide, *N*-(2-phenylethyl)-	877-95-2	C10H13NO	796	5	8
15.98	Sulphur	13798-23-7	S6	846	3	5
16.09	Benzoic acid, 4-ethoxy-, ethyl ester	23676-09-7	C11H14O3	806	6	11
17.69	Benzenesulfonamide, *N*-ethyl-2-methyl	1077-56-1	C9H13NO2S	795	6	9
17.80	2(3H)-Benzothiazolone	934-34-9	C7H5NOS	987	4	7
18.34	Benzenesulfonamide, 4-methyl-*N*-propyl	1133-12-6	C10H15NO2S	921	5	9
18.55	Hexathiepane	17233-71-5	CH2S6	766	5	7
19.08	Benzenesulfonamide, *N*-butyl	3622-84-2	C10H15NO2S	916	6	11
19.80	Caffeine	58-08-2	C8H10N4O2	871	6	10
20.56	Pyrrolo[1,2–a]pyrazine-1,4-dione, hexahydro-3-(2-methylpropyl)-	5654-86-4	C11H18N2O2	735	6	10
20.64	Pyrrolo[1,2–a]pyrazine-1,4-dione, hexahydro-3-(2-methylpropyl)-	5654-86-4	C11H18N2O2	778	6	11
21.91	Cyclic octaatomic sulphur	10544-50-0	S8	833	5	8
24.66	Pyrrolo[1,2–a]pyrazine-1,4-dione, hexahydro-3-(phenylmethyl)-	14705-60-3	C14H16N2O2	900	6	12

[a] Identification confirmed with a standard compound.

qualified had the same elemental formula, the elemental composition was considered tentatively identified and the identity of the best library match was reported to illustrate one possible structure of the component (stage 5b). The elemental compositions identified in this category are summarized in Table 3. Since all identifications from nontarget screening should be considered tentative, the last stage of the process is the confirmation of identification with a pure standard material, if available. For this stage, the measurement of ion ratios of fragments can also be included.

In stage 5a, 24 and 20 compounds were tentatively identified in the *n*-hexane and dichloromethane extracts, respectively. The presence of triethyl phosphate and tris(1-chloro-2-propyl)-phosphate was confirmed with standard compounds. Some of the compounds were detected in both extracts. The largest group tentatively identified in the sample was disulphides. This finding is in accordance with a previous study [35] in which various disulphides (C3–C10) and some trisulphides were detected in landfill biogas, originating from anaerobic degradation of organic material. In addition, e.g., alkyl-benzenesulphonamides, organophosphorus triesters and cyclic sulphur compounds were detected in our study. From the individual compounds identified e.g., 2(3H)-benzothiazolone (CAS 934-34-9), benzenesulphoamide, *N*-butyl (CAS 3622-84-2) and tris(1-chloro-2-propyl)phosphate (CAS 13674-84-5) have also previously been detected in landfill leachate [7]. The number of ions used in accurate mass scoring and the number of IPs calculated are also presented in Table 2. The median values of the number of ions and IPs were 5 and 10, respectively. The number of IPs can be used as an indicator of identification reliability. The elemental formulas, accurate masses and Δm values of the fragments of three different compounds identified are presented as case examples in Table 4. For all tentatively identified compounds, the Δm values were generally less than 2.0 mDa which expounds the high count of IPs. Compared to conventional GC–MS measurement with low mass resolution, data acquisition with GC–TOF–MS offers a very useful additional tool to support the identification with accurate mass. However, it should be pointed out that the same final library identification was gained for three pairs of components with different retention times. As an example, compounds at retention times 14.41 and 15.44 min in *n*-hexane extract (Table 2) were both tentatively identified as dihexyl disulphide. This indicates

Table 3
Elemental compositions suggested in the landfill leachate sample according to the identification process.

Retention time (min)	Forward fit	Elemental formula	Best library match	Number of ions	IPs
n-Hexane extract					
6.56	935	C9H12	1,3,5-Trimethylbenzene	4	8
7.02	826	C9H12	1,3,5-Trimethylbenzene	3	6
7.64	713	C6H12S	2,5-Dimethyltetrahydrothiophene	4	8
7.72	952	C7H8O	4-Methylphenol	4	5
8.06	760	C10H16O	Fenchone	4	7
8.28	712	C8H10O	3,4-Dimethylphenol	4	8
8.32	907	C6H14S2	Bis(1-methylethyl)disulphide	4	8
8.43	752	C6H14S2	Ethyl *n*-butyl disulphide	4	8
8.73	776	C6H14S2	Metyl pentyl disulphide	3	6
8.82	732	C9H14O	β-Pinone	6	12
8.86	855	C8H10O	3,4-Dimethylphenol	4	8
8.89	720	C10H20O	Cyclohexanemethanol, a,a,4-trimethyl-	4	7
8.94	896	C10H16O	(R)-Camphor	6	6
9.29	744	C4H10S	1-Butanethiol	3	6
9.31	811	C10H20O	Neoisomenthol	6	11
9.54	856	C10H8	Naphtalene	3	6
9.61	763	C10H16O	Isopinocamphone	5	10
9.79	756	C7H16S2	Propyl *n*-butyl disulphide	6	12
9.86	782	C10H16O	2-Norpinanone	6	11
10.40	802	C9H10O	2,5-Dimethylbenzaldehyde	4	8
10.47	902	C9H12O	4-Propylphenol	3	6
10.91	761	C7H7ClO	2-Chloro-4-methylphenol	3	6
10.98	701	C10H14O	m-tert-butylphenol	5	10
11.05	931	C8H7N	Indole	3	6
11.18	732	C8H18S2	Dibutyl disulphide	6	12
11.58	730	C9H20S2	*n*-Propyl *n*-hexyl disulphide	3	6
11.59	710	C6H5Cl2N	2,3-Dichlorobenzenamine	4	7
12.13	845	C12H24O3	Propanoic acid, 2-methyl, 3-hydroxy-2,4,4-trimethylpentyl ester	6	11
13.46	878	C6H14S	1-Pentanethiol, 2-methyl	3	6
14.72	946	C12H17NO	Diethyltoluamide	5	10
18.14	725	C10H13N	1H-Indole, 2,3-dihydro-1,2-dimethyl	6	12
Dichloromethane extract					
9.58	856	C8H10O	1-Phenylethanol	6	11
10.23	820	C9H18O	Nonanal	6	12
10.91	940	C8H10O	2,4-Dimethylphenol	3	5
15.65	880	C9H9NO	2,6-Dimethylphenyl isocyanate	4	7
16.78	801	C12H17NO	Benzamide, *N*,*N*-diethyl-4-methyl	6	10

Table 4
Elemental formulas, accurate masses and Δm values of the fragments of three different compounds tentatively identified.

Compound	2-Propanol, 1-chloro-, phosphate	Caffeine	Hexathiepane
m/*z* 1	C2H6O4P	C8H10N4O2	CH2S4
	125.0004 (−0.6)	194.0804 (3.3)	141.9039 (−2.3)
m/*z* 2	H4O4P	C5H7N3	CH2S2
	98.9847 (0.2)	109.0640 (0.7)	77.9598 (0.0)
m/*z* 3	C8H16O4P[35Cl2]	C3H5N	S2
	277.0163 (-0.5)	55.0422 (−1.1)	63.9441 (−0.1)
m/*z* 4	C5H11O4P[35Cl]	C3H3N2	CH2S6
	201.0084 (-0.1)	67.0296 (-0.5)	63.9441 (−2.4)
m/*z* 5	C3H7O3P[35Cl]	C4H6N2	S3
	156.9821 (1.7)	82.0531 (1.4)	95.9162 (2.7)
m/*z* 6	H3O3P[35Cl]	C6H7N3O	–
	116.9508 (0.1)	137.0589 (−3.4)	–
IPs	12	10	7

that these two compounds closely resemble each other, and spectral library does not contain a spectrum for both compounds. Alternatively, the same best library hit may also be explained by isomerism in some cases. This demonstrates that the final confirmation of the identification must always rely on the use of a pure standard.

In addition to tentatively identified compounds, 36 components for which a likely chemical structure could not be assigned passed the six-stage process. Tentative elemental compositions of these compounds are reported in Table 3. These compounds showed more than one library hit with good forward fit, and thus tentative identifications could not be performed. The best library match is given in Table 3 to illustrate one possible structure of the compound. For example, the first two compounds in Table 3 can be isomers of e.g., trimethyl benzene or ethylmethylbenzene, but their exact identity cannot be confirmed without standard compounds. Different terpenoids and phenolic compounds were mainly detected in this category. In addition to the compounds in Table 2, some fairly certain compound group identifications were noticed but could not be reported, because they did not fulfil the criteria imposed. One such group was the phthalates. Their rejection was based on the inadequate number of ions in the deconvoluted spectrum (stage 2). Phthalates only show one common, very characteristic ion (*m*/*z* 149) in their EI spectra; otherwise the level of fragmentation and the intensity of molecular ion is low. Thus the determination of their structure is usually not possible on the basis of their mass spectra. Another problem was found with diverse esters. A sufficient number of ions was detected for them, but the similarity of spectra prevented identification as several library hits with different elemental formulas were gained. This is due to common fragmentation of esters, especially in the low mass range leading to similar EI spectra. These examples show that applicability of

the nontarget screening process is limited among isomers and compounds with minor fragmentation.

The general aim of nontarget analysis is to detect as many compounds in the sample as possible. However, from sampling to final data treatment, the analytical process itself is a bunch of selections, always resulting in the loss of some analytes. For example the effect of extraction solvent selection is obvious in stage 5a, as caffeine was only detected in the DCM extract and hexathiepane only in the *n*-hexane extract. Thus, it is clear that an ideal nontarget analysis method does not exist. Functioning of the software used in nontarget analysis has, however, the most significant role. Because of the immense amount of full spectrum data, it is indispensable that the software is capable of extracting the essential information. Different thresholds and settings are additionally required to efficiently filter the deconvoluted components. According to our experiences, a lot of false components from noise are deconvoluted. In addition, all identification stages subsequent to the first one had to be performed manually, as the software did not offer such filter options. When more than a few samples are processed this current approach is very time-consuming and tedious. Consequently, in our opinion limitations of the software applied form the bottleneck of this nontarget screening method. Additional tools for the nontarget screening software are thus required to obtain a more automated method, making extensive screening studies feasible.

4. Conclusions

In our research, we demonstrated the use of GC–TOF–MS data of a landfill leachate sample for nontarget analysis with deconvolution software. Results show that nontarget screening can produce valuable information about the sample composition, and reveal the presence of compounds which would not have been found using some traditional target analysis technique. Nontarget screening may be used as a parallel technique with target analysis in order to gain more information about the sample. In the next step of our research, the developed method will also be applied for other aquatic matrices, such as surface water and stormwater. Although the usefulness of nontarget analysis is unquestionable, some limitations are related to the method. The low sensitivity of an analyte, co-eluting components, complex sample background noise or decisions made by the analyst may cause an analyte to remain undetected. Unequivocal identification is not possible in some cases, e.g., with isomers. Consequently, the identification obtained from nontarget analysis should be considered as tentative until it has been confirmed with a standard compound. The main hindrance for the use of the nontarget screening method is the deconvolution software. Routine applications would require more advanced and automated features to treat large datasets efficiently. However, we believe that with future developments in instrumentation and software, increasing interest towards nontarget screening techniques will be seen.

Acknowledgements

Our research was funded by the Helsinki University Centre for Environment (HENVI), Ramboll Finland Oy, Finnish Doctoral Programme in Environmental Science and Technology, Onni and Hilja Tuovinen Foundation, Finnish Water Services Development Fund and Lake Vesijärvi Foundation. Ph.D. P. Manninen is acknowledged for his ideas and guidance during our research.

References

[1] United Nations Environment Programme. UNEP Six Priority Areas Factsheets—Harmful Substances and Hazardous Waste (2010).
[2] M. la Farré, S. Pérez, L. Kantiani, D. Barceló, TrAC, Trends Anal. Chem. 27 (2008) 991–1007.
[3] S.D. Richardson, Anal. Chem. 84 (2012) 747–778.
[4] K. Wille, H.F. De Brabander, L. Vanhaecke, E. De Wulf, P. Van Caeter, C.R. Janssen, TrAC, Trends Anal. Chem. 35 (2012) 87–108.
[5] G. Teijon, L. Candela, K. Tamoh, A. Molina-Díaz, A.R. Fernández-Alba, Sci. Total Environ. 408 (2010) 3584–3595.
[6] T. Eggen, M. Moeder, A. Arukwe, Sci. Total Environ. 408 (2010) 5147–5157.
[7] L. Badoil, D. Benanou, Anal. Bioanal. Chem. 393 (2009) 1043–1054.
[8] J. Schwarzbauer, S. Heim, S. Brinker, R. Littke, Water Res. 36 (2002) 2275–2287.
[9] R.J. Slack, J.R. Gronow, N. Voulvoulis, Sci. Total Environ. 337 (2005) 119–137.
[10] N.A. Al-Odaini, M.P. Zakaria, M.I. Yaziz, S. Surif, J. Chromatogr. A 1217 (2010) 6791–6806.
[11] E. Eriksson, A. Baun, L. Scholes, A. Ledin, S. Ahlman, M. Revitt, C. Noutsopoulos, P.S. Mikkelsen, Sci. Total Environ. 383 (2007) 41–51.
[12] M. Jartun, A. Pettersen, J. Soils Sediments 10 (2010) 155–161.
[13] S. Zgheib, R. Moilleron, G. Chebbo, Water Sci. Technol. 64 (2011) 1450–1458.
[14] R. Húšková, E. Matisová, L. Švorc, J. Mocák, M. Kirchner, J. Chromatogr. A 1216 (2009) 4927–4932.
[15] X. Xia, B.S. Crimmins, P.K. Hopke, J.J. Pagano, M.S. Milligan, T.M. Holsen, Anal. Bioanal. Chem. 395 (2009) 457–463.
[16] E. Pitarch, C. Medina, T. Portolés, F.J. López, F. Hernández, Anal. Chim. Acta 583 (2007) 246–258.
[17] F.J. Camino-Sánchez, A. Zafra-Gómez, J.P. Pérez-Trujillo, J.E. Conde-González, J.C. Marques, J.L. Vílchez, Chemosphere 84 (2011) 869–881.
[18] A. Llop, F. Borrull, E. Pocurull, J. Sep. Sci. 33 (2010) 3692–3700.
[19] S. Hashimoto, Y. Takazawa, A. Fushimi, K. Tanabe, Y. Shibata, T. Ieda, N. Ochiai, H. Kanda, T. Ohura, Q. Tao, S.E. Reichenbach, J. Chromatogr. A 1218 (2011) 3799–3810.
[20] M. Ibáñez, J.V. Sancho, F. Hernández, D. McMillan, R. Rao, TrAC, Trends Anal. Chem. 27 (2008) 481–489.
[21] J. Sancho, Ó. Pozo, M. Ibáñez, F. Hernández, Anal. Bioanal. Chem. 386 (2006) 987–997.
[22] F. Hernández, Ó.J. Pozo, J.V. Sancho, F.J. López, J.M. Marín, M. Ibáñez, TrAC, Trends Anal. Chem. 24 (2005) 596–612.
[23] F. Hernández, T. Portolés, E. Pitarch, F.J. López, J. Mass Spectrom. 44 (2009) 1–11.
[24] T. Portolés, M. Ibáñez, J.V. Sancho, F.J. Ló pez, F. Hernández, J. Agric. Food Chem. 57 (2009) 4079–4090.
[25] F. Hernández, T. Portolés, E. Pitarch, F.J. López, TrAC, Trends Anal. Chem. 30 (2011) 388–400.
[26] T. Portolés, E. Pitarch, F.J. López, F. Hernández, J. Chromatogr. A 1218 (2011) 303–315.
[27] T. Portolés, E. Pitarch, F.J. López, J.V. Sancho, F. Hernández, J. Mass Spectrom. 42 (2007) 1175–1185.
[28] F. Hernández, T. Portolés, E. Pitarch, F.J. López, Anal. Chem. 79 (2007) 9494–9504.
[29] M.J. Gómez, M.M. Gómez-Ramos, A. Agüera, M. Mezcua, S. Herrera, A.R. Fernández-Alba, J. Chromatogr. A 1216 (2009) 4071–4082.
[30] R. Serrano, J. Nácher-Mestre, T. Portolés, F. Amat, F. Hernández, Talanta 85 (2011) 877–884.
[31] T. Portolés, E. Pitarch, F.J. López, F. Hernández, W.M.A. Niessen, Rapid Commun. Mass Spectrom. 25 (2011) 1589–1599.
[32] European Commission, Off. J. Eur. Commun. L221 (2002) 8.
[33] F. Hernández, M. Ibáñez, J.V. Sancho, Ó.J. Pozo, Anal. Chem. 76 (2004) 4349–4357.
[34] M.W.F. Nielen, M.C. van Engelen, R. Zuiderent, R. Ramaker, Anal. Chim. Acta 586 (2007) 122–129.
[35] S. Rasi, A. Veijanen, J. Rintala, Energy 32 (2007) 1375–1380.