Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1019 (2003) 233–249
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Search criteria and rules for comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry analysis of airborne particulate matter

W. Welthagen [a,b], J. Schnelle-Kreis [c], R. Zimmermann [a,b,c,*]

[a] *Institute of Ecological Chemistry, GSF-Research Centre, Ingolstädter Landstrasse 1, D-85764 Oberschleißheim, Germany*
[b] *Analytical Chemistry, University of Augsburg, D-86159 Augsburg, Germany*
[c] *BIfA—Bavarian Institute of Applied Environmental Research and Technology, D-86167 Augsburg, Germany*

Abstract

Direct thermal desorption–gas chromatography–time-of-flight mass spectrometry (DTD–GC–TOFMS) and comprehensive two-dimensional (2D) gas chromatography–time-of-flight mass spectrometry (GC × GC–TOFMS) was applied for characterisation of semi-volatile organic compounds (SVOC) in fine particulate matter (PM), with a diameter of up to 2.5 μm (PM2.5), from ambient air in Augsburg, Germany. DTD–GC–TOFMS measurements on the SVOC in PM2.5 are done on a daily basis (time series over several years). The data will be used in an epidemiological study questioning the influence of SVOC in PM2.5 on ambient aerosol related health effects. The outcome of the first measurements periods is that the organic inventory in the ambient aerosol can undergo drastic fluctuations, e.g. due to meteorological influences or specific emission sources. This includes also the large fraction of chromatographically not resolved peaks (unresolved carbonaceous matter (UCM)). The UCM fraction contains about 70% of the SVOC mass in PM2.5. GC × GC–TOFMS is a suited technique to study the nature of the yet unidentified compounds forming the UCM. The considerably increased chromatographic resolution in GC × GC allows separation of many UCM compounds while the TOFMS supplies mass spectral data of all separated compounds. However, the data sets are getting enormously complex. In a typical PM2.5 sample from Augsburg more than 15,000 peaks can be detected. Thus, it is important to classify the observed GC × GC peaks by rational means. A classification procedure based on GC × GC retention times and the fragmentation patterns is suggested. With a preliminary classification procedure it is already possible to group compounds with some certainty into substance classes. After some further development, this approach can be used for classifying GC × GC data, e.g. for environmental and epidemiological studies.

Keywords: Gas chromatography, comprehensive two-dimensional; Aerosols; Environmental particles; Time-of-flight mass spectrometry; Semi-volatile organic compounds

1. Introduction

Several epidemiological studies have demonstrated the relevance of ambient particles in health effects. The results of these studies together with animal toxicology

* Corresponding author.
E-mail address: ralf.zimmermann@gsf.de (R. Zimmermann).

0021-9673/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2003.08.053

and in vitro experimental studies support the hypothesis that both physical (particle size, shape, surface) and chemical (dissolved and adsorbed chemicals, surface catalytic reactions) properties of the particles are involved in toxic, genotoxic and carcinogenic mechanisms of inhaled particulates.

It was recently recognised that inhaled ultra fine particles ($D_p < 100$ nm) are more toxic than PM10 particles [1]. Their relative large surface area and the ability to be absorbed into tissues might be important factors in cardiopulmonary toxicity. However, the number of ultra fine particles in the air is often poorly correlated with PM2.5 and even less with PM10. Thus, ultra fine particles are unlikely to explain much of the association between particulate mass and health conditions. Also the impact of surface absorbed compounds on health outcomes is not well understood yet. Specific compounds related to traffic exhaust like from diesel vehicles [2] or even diesel locomotives [3] are possibly important.

Epidemiological investigations of the influence of individual, particle bound chemical pollutants were done only for few inorganic species [4]. The role of transition metals (Fe, V, Zn) for acute reactions is under discussion [5]. However, little is known on the influence of the organic chemicals present in ambient particulate matter (PM) on the health outcomes. The influence of organic substances was evaluated by measurement of the concentration of elemental and organic carbon (EC/OC) [6]. But so far, the association of individual specific organic pollutants or groups of pollutants with health effects, occurring in the fine dust, was not examined in epidemiological studies.

For a time-series study, on the influence of organic aerosol compounds, it is necessary to have data of several compounds or groups of compounds at least with a daily resolution. Because most of the organic compounds occur in low concentrations in ambient aerosol, time-consuming analytical methods are required for their analysis. Several studies address the organic composition of ambient PM, using gas chromatography–mass spectrometry (GC–MS) for separation and identification of semi-volatile organic compounds (SVOC). GC–MS is a well-established technique for the separation and analysis of complex mixtures [7,8].

As a new analytical approach, direct thermal desorption (DTD) was employed for extracting semi-volatile species from ambient particulate matter [9,10]. Recently, we started a project for sampling and analysis of semi-volatile organic compounds associated with ambient PM2.5 on a daily basis by direct thermal desorption–gas chromatography–time-of-flight mass spectrometry (DTD–GC–TOFMS). The data of this project together with data of particle mass (PM2.5) and number-concentrations will be analysed in an epidemiological study.[1] In the epidemiological study, relative abundance data of compound classes (summarised concentrations of, e.g. *n*-alkanes, fatty acids, PAH, PAH-quinones, etc.) will be statistically correlated with health data of the population of Augsburg. The health data (myocardial infarction and sudden cardiac deaths, and survival of myocardial infarction survivors) are acquired in the framework of the KORAN project (Co-operative Health Research in the Region of Augsburg [11]). Furthermore, the data of individual molecular tracers for different sources (like hopanes as tracer of fossil fuel burning, retene as tracer for wood combustion, cholesterol as tracer for charbroiling, iso- and anteiso-hentriacontan as tracers for environmental cigarette smoke [7] or nitriles as tracers for biomass burning [8]) will be considered in the epidemiological study. In addition, certain deduced indices like carbon preference index ((CPI) the ratio of odd to even numbered alkanes as indicator for anthropogenic emissions) and the total concentration of SVOC would also be determined and used in the data analysis procedures. One drawback of the approach is, however, that only a small fraction of the SVOC compounds present in ambient particulate matter samples can be identified and assigned to compound classes. This is due to the fact that one-dimensional (1D) gas chromatography cannot provide sufficient chromatographic resolution for the separation of the large numbers of organic compounds present in ambient PM. Typical gas chromatograms of SVOC from PM exhibit broad unresolved bands (unresolved carbonaceous matter (UCM)), including the majority of the SVOC compounds. For better understanding of the nature of the organic chemical inventory hidden in the UCM-bands, it is required to apply novel analytical techniques with considerably increased resolving power.

[1] Co-operation with Institute of Epidemiology, GSF-Research Centre, A. Peters and H.E. Wichman.

Comprehensive two-dimensional (2D) gas chromatography (GC × GC) represents a new approach to enhance the GC resolution [12,13]. GC × GC is a novel technique, whereby a sample is separated (in two dimensions) with two comprehensively coupled gas chromatographic columns. Two different chromatographic mechanisms (i.e. volatility and polarity) are used to separate the compounds in the two columns. Several applications of GC × GC with fast flame ionisation detector for separation of complex mixtures, such as petrochemical mixtures [14] and some volatile organics in the atmosphere [15] have been reported.

A promising technique for analysis of organics in PM is GC × GC coupled to fast time-of-flight mass spectrometry [16–18]. Due to the increased separation of GC × GC with respect to one-dimensional GC, the mass spectra are of considerably increased quality (lower background level). The problem of skewing of mass spectra in GC–MS experiments with scanning mass analysers is also not present in time-of-flight mass spectrometry. Thus, TOFMS provide identical mass spectral patterns over a complete chromatographic peak for the same component. This property promotes deconvolution algorithms.

It is obvious that for statistical epidemiological studies, which need to be done on a routine basis, it is necessary to classify chemically related compounds into groups. Otherwise, the large amount of chemical data cannot be handled. Classifying compounds into chemical related groups also could be advantageous for analysis of the spatial, temporal or particle size dependent variability of the chemical fingerprints as well as for source characterisation.

The GC × GC-technique itself is rather well suited for group separations. The use of the two-dimensional (2D) retention data for this purpose is already described by a number of publications [19–21], and is therefore a useful parameter to include in this study as well. It relates to Giddings theory published on the concepts of two-dimensional systems [22]. Giddings showed theoretically that the key property of a separation method, which determines whether it can show the inherent structure of a mixture being separated or not is the method's dimensionality. The dimensionality of a mixture is thus the number of independent variables of every member of the mixture. When a mixture is then separated based on these independent variables, each compound will separate to a unique location on a separation plane (chromatogram). However, as the compounds are composed of molecules with discrete structures that are related, the compounds must distribute over the dimensional separation space (chromatogram) to discrete locations, which are also related to each other. Explaining Giddings' theory in GC × GC separation [23]: if a mixture is separated into one dimension, such as the boiling point fractions in petroleum samples, the alkanes with similar boiling points to the aromatics overlap and thus insignificant ordering of compounds occurs. If the variable "boiling point" is changed to "polarity", the same overlap does not occur but new overlaps are created by different "boiling point" fractions, thus the separation or ordering is still insignificant. The mixture is simply not sufficiently well ordered in any one dimension, it requires at least two independent variables to uniquely speciate the compounds of the mixture. This theory applied to two-dimensional gas chromatography thus provides the basis on which compounds are separated in a two-dimensional plane. It should however be noted that all GC separations is dependant on the "boiling point" variable, and thus the two separation systems in GC × GC could not strictly be considered as independent of one another.

Compounds belonging to the same chemical group in a chemical mixture are related to one another in some chemical or physical way, thus separating them according to these physical and chemical properties will result in compounds being separated into chemical groups, providing structured distribution patterns of the chemical groups.

Petrochemical samples contain thousands of compounds but the relative few number of chemical classes, make it ideal for a two-dimensional system like GC × GC. Aerosol samples however also contain thousands of compounds but they belong to many different compound classes. The two dimensions of GC × GC are simply not enough, therefore at least another dimension (MS) needs to be added.

The use of fragmentation patterns in mass spectral data to identify compounds and to classify compounds is a well-established technique [24] and was long used before mass spectral libraries became available. Several chemical compounds classes show unique distribution of ions (fragmentation patterns)

in the mass spectrum. These fragmentation patterns depend on the structures of the molecules ionised and the stability of the fragment ions. It is clear that the complexity of these fragmentation patterns often increases with molecule size and number of functional groups. But it still gives valuable interpretable results.

The purpose of the here presented study therefore is to improve on an already well-established GC × GC-based concept for group type separation by additionally using the mass spectral information. The de-convoluted mass spectra and the retention time of both dimensions would be the basis of sorting these peaks into chemically related groups.

In this paper, some typical results on the daily GC–TOFMS monitoring of organic compounds in the ambient aerosol (PM2.5) are given and discussed in terms of the complexity and variability of the chemical composition (SVOC) of the urban aerosol. Furthermore, results obtained with GC × GC–TOFMS on ambient aerosol (PM2.5) are presented and a first concept for classifying the huge number of chromatographically resolved peaks with respect to the two-dimension retention data and the mass spectral information is demonstrated.

2. Experimental

2.1. Sampling

Within the project particles (PM2.5) are sampled on a daily time basis at a central monitoring station in Augsburg, Germany, on quartz fibre filters (Munktell T293, Grycksbo, SE). Daily sampling started in spring 2002 and will be continued till 2005 or longer. Ambient urban air particulate matter is sampled with a PM2.5 sequential sampler (Partisol-Plus 2025, Rupprecht & Pataschnik Albany, NY, USA) at a flow rate of 1 m^3/h. Separators are used to isolate the cassettes with filters (eight in each cassette) in the filter magazine. Each filter magazine contains seven filters for daily sampling and one filter as field blank. The filter cassettes are exchanged once a week. In the laboratory, the filters are cut into 13 pieces. Each piece represents a PM2.5 aliquot of 1 m^3 sampled air. These pieces are stored deep-frozen at −18 °C until analysis.

2.2. Direct thermal desorption–gas chromatography–time-of-flight mass spectrometry (DTD–GC–TOFMS)

The samples are analysed for SVOC by direct thermal desorption–gas chromatography–time-of-flight mass spectrometry using a novel DTD interface. In detail, the liner of the GC injection interface (DTD-liner) is used as sample container for one to three filter pieces. Isotope labelled reference compounds are added for quantification.

After automated insertion of the DTD-liner and closing of the cold injector port the liner is purged with helium for 2 min. After this, the injector port is heated to 350 °C and the analytes are thermally desorbed directly from the DTD-liner onto the capillary GC column. During desorption time, the SVOC are focused on a retention gap at 50 °C oven temperature.

The analytical DTD–GC–TOFMS parameters (instrumentation: Injection port Optic III, ATAS-GL, Veldhoven, NL; GC: Agilent 6890, USA; TOFMS: Pegasus III, LECO Ltd., St. Joseph, MI, USA) were as follows. Retention gap: deactivated fused silica, 2 m, 0.22 mm i.d. (SGE, Ringwood, AUS) Column: BPX5, 50 m, 0.22 mm i.d., 0.25 mm d_f (SGE), flow: 1.0 ml/min helium, Thermal desorption: filter purge at 50 °C, split 50 ml/min (2 min hold), Temperature ramp to 350 °C (15 min hold) at 60 °C/min—split 5 ml/min, oven parameters: 50 °C (22 min hold) to 150 °C at 100 °C/min; to 330 °C (30 min hold) at 20 °C/min. The TOFMS was operated with a data acquisition frequency of 20 Hz.

For each analysis, three filter pieces are placed into a DTD-liner. Isotope labelled internal standards are added for quantification. For statistical evaluation, the concentrations of single compounds (e.g. source specific tracer substances) are summarised concentrations of compound groups (e.g. *n*-alkanes, fatty acids, PAH, PAH-quinones) and concentration of the unresolved carbonaceous matter (integral over the UCM-band) are adopted.

2.3. Two-dimensional comprehensive gas chromatography–time-of-flight mass spectrometry (GC × GC–TOFMS)

A comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometer was used

in this study (instrumentation: Injection port Optic III, ATAS-GL, Veldhoven, NL; GC: Agilent 6890, USA; TOFMS: Pegasus III, LECO Ltd., St. Joseph, MI, USA, GC × GC system: Pegasus 4D option, LECO Ltd., MI, USA). The two-dimensional comprehensive gas chromatographic system consists of two serially connected separation columns, a long (30 m) first-dimension column (with a non-polar stationary phase) and the much shorter (1.5 m), second-dimension column (with a polar stationary phase). A four-jet nitrogen modulator, with two heated and two cooled jets, was used for modulation of the eluent from the first column into the second. The modulator is situated between the two columns. On the first column, a conventional high-resolution gas chromatographic separation takes place. The modulator repeatedly focuses the eluent from this first column and re-injects it as narrow injection bands into the second column, which is run under "almost isothermal" conditions at a temperature of typically 5 °C higher than the first column's temperature. Due to the fast speed of separation of the second column, the differences of start and end temperatures of each second dimension separation cycle are negligibly small, therefore the term "almost isothermal" is used.

To maintain the separation achieved by the first column, the re-injection intervals of the modulator need to be fast (e.g. 4 s) and therefore the speed of separation in the second column also needs to be fast. The faster separation in the second column is obtained by using a column with smaller inner diameter than that of the first column. Column parameters of the sets used in this study was a 30 m × 0.25 mm i.d. × 0.25 μm d_f BPX5 from SGE as first dimension column and a 1.5 m × 0.10 mm i.d. × 0.10 μm d_f BPX50 from SGE as second dimension column.

Additional instrumentation parameters were as follows: 3 μl splitless injection at 300 °C, 1.5 ml/min column flow rate, helium carrier gas, and data acquisition rate of 100 Hz. The first dimension oven was held at 60 °C for 10 min then heated to 300 °C at 1.5 °C/min while the second oven was kept at 5 °C above first oven temperature. The modulator containing the heating jets was heated to 100 °C above oven temperatures and the pulse duration of the heated pulses was set to 300 ms. The cold pulses, cooled with a liquid nitrogen exchanger, were only switched off during the heating periods (hot pulses). A modulation period of 4 s was used.

3. Results and discussion

3.1. Direct thermal desorption–gas chromatography–time-of-flight mass spectrometry

In the PM2.5 samples, among other compounds, *n*-alkanes, long chain carboxylic acids, long chain alkyl nitriles, esters, terpenes, steranes, hopanes, substituted aromatics, PAH, oxygenated-PAH (oxy-PAH), heterocyclic aromatic compounds have been identified and partly quantified. The question, to which degree the detected long chain alkyl nitriles are created from the analog acids in presence of an ammonium source (e.g. NH_4NO_3) or from the amides during the thermal desorption [25] or the use of it as tracers for biomass burning [8] has yet to be investigated in detail.

As shown in Fig. 1A and B, the differences in the total ion current (TIC) chromatograms are substantial and mainly visible in the unresolved carbonaceous matter region. The relative changes over a 2-week period in May 2002 are depicted. A steady but relative small increase in UCM concentration from 14th to 17th is followed by a steep decrease to the 19th (Fig. 1A). In Fig. 1B, the strong changes in concentrations of the organic compounds on a daily basis within a 5-day period are highlighted. Whereas the maximum in UCM concentration in the whole period is at about 46–47 min in the chromatograms, on the middle day of this period, the 22nd, this maximum is shifted to 50–51 min. In addition, the concentration of several individual compounds are strongly increased on this particular day.

In Fig. 1C and D, examples for day-to-day fluctuations of the concentrations of some *n*-alkanes (Fig. 1C) and PAH (Fig. 1D) in the same time period are shown. The carbon preference index indicates a slight increased contribution of anthropogenic emissions on the 22nd May only. Although the concentrations of all alkanes were increased on this particular day the increase of the concentrations of the odd-numbered alkanes was higher than of the even-numbered alkanes. Some PAH's (e.g. chrysene {CRY} and benzo[*b*+*k*]fluoranthene {B*b*+*k*F}) show increased concentrations on the same day, whereas

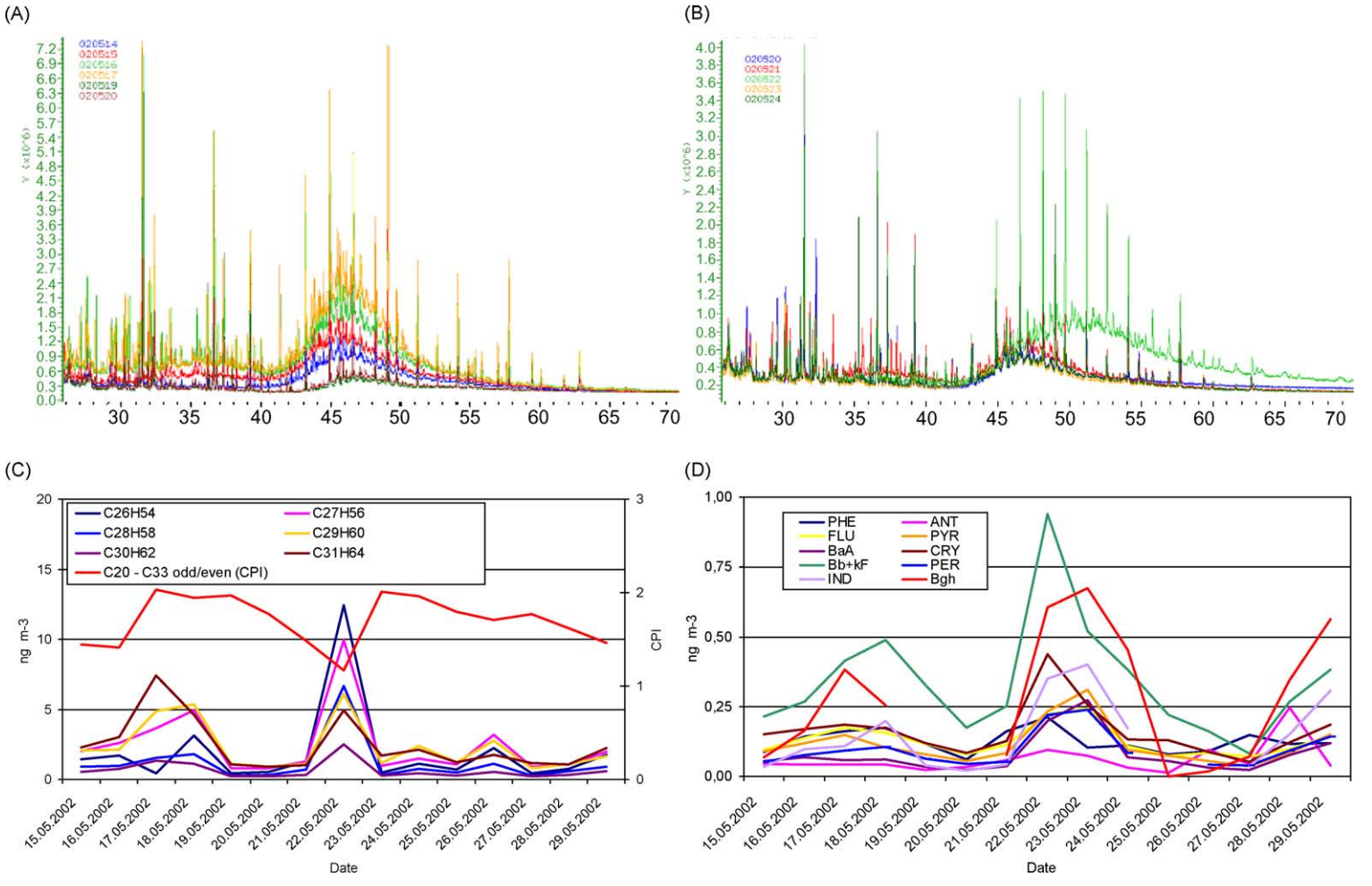

Fig. 1. Day-to-day changes in the organic inventory of urban PM: (A) concentration increase and decrease in the period 14th to 20th May; (B) strong concentration increase on a single day within a 5-day period 20th to 24th May; (C) concentrations of some *n*-alkanes and carbon preference index (CPI) in the period 15th–29th May; (D) concentrations of some PAH (PHE: phenanthrene; ANT: anthracene; FLU: fluoranthene; PYR: pyrene; BaA: benzo[*a*]anthracene; CRY: chrysene; B$b+k$F: sum of benzo[*b*]- and [*k*]fluoranthene; PER: perylene; IND: indeno[1,2,3-cd]pyrene; Bgh: benzo[ghi]perylene) in the period 15th–29th May.

most of the other PAH's (e.g. benzo[*a*]anthracene {BaA}, pyrene {PYR}, indeno[1,2,3-cd]pyrene {IND}, benzo[ghi]perylene) show a maximum in concentration on two consecutive days. Long-term monitoring and data analysis will give information about the variability of the organic chemical inventory of urban ambient fine particles (PM2.5). In this study about 250 individual compounds are identified. Most of them are quantified for the statistical evaluation. This is however only a very small fraction of the total number of SVOC associated with ambient PM. The majority of the organic compounds present in ambient PM, primarily characterised by the UCM-band in the one-dimensional chromatograms are not identified yet.

In summary, Fig. 1 clearly shows that the SVOC fraction in ambient particulate matter exhibit a rather drastic day-to-day variability. Note that the method was evaluated for reproducibility, which pointed out to be for most of the target compounds better than 10%. The large fluctuations of the SVOC inventory in the ambient aerosol implicate that a large day-by-day variability of the exposure of the Augsburg population to particle-bound SVOC is present.

Valid data on the temporal variation of the relative abundances of the compound classes, the concentration of the unresolved carbonaceous matter and the concentrations of the molecular tracers are prerequisites for epidemiological studies on short-term effects of particle associated SVOC. Unfortunately, in this study quantified compounds represent only about 15% of the total SVOC. Therefore, it is likely, that compounds that are involved in the health effects are hidden in the UCM-bands. The determined large variability in the UCM-band therefore is only a poor surrogate for these compounds. The GC × GC–TOFMS can amplify our knowledge on the compounds and compound classes present in aerosol samples. But for use in epidemiological studies the enormous amount of single compound data generated by this technique has to be reduced to reasonable sets of combined data.

3.2. Two-dimensional comprehensive gas chromatography–time-of-flight mass spectrometry (GC × GC–TOFMS)

In the investigated ambient aerosol samples (SRAM 1649a, "urban dust" and PM2.5 Augsburg samples), several thousand peaks were found in the two-dimensional chromatogram, using the automated peak identification routine based on deconvolution methods. Among various other compound classes, alkanes, alkenes, cycloalkanes, long chain carboxylic acids and esters, aldehydes, ketones, substituted aromatics, PAH, oxygenated-PAH and heterocyclic aromatic compounds were identified by library search routines (NIST). However, large numbers of unknown and probably mismatched compounds still occur. An important feature is the ordered appearance of chemically related compounds in the 2D GC × GC chromatogram. This is due to the physico-chemical similarities within compounds classes and the gradual changes of these properties with increasing molecular sizes/chain length. The ordered appearance helps to classify unknown substances. In Fig. 2, two 2D-chromatograms of an Augsburg PM2.5 sample are shown as three-dimensional (3D) landscape-plots. The 2D-chromatogram based on the total ion current (not shown here) is totally dominated by non-polar hydrocarbons, which are weakly separated by the chosen column parameters. However, many compounds that are suspect to be health relevant are occurring in rather low concentrations and are more polar than the alkanes. In order to visualise these compounds, characteristic selected ions are chosen (SI) for generating the 3D-plots that are depicted in Fig. 2. The upper 3D-plot (Fig. 2A) represents a SI-plot based on the typical molecular ions (154, 166, 178, 202, 228, 252, 276 *m*/*z*) for polycyclic aromatic hydrocarbons (PAH, e.g. phenantrene, 178 *m*/*z*; pyrene, 202 *m*/*z*; chrysene, 228 *m*/*z*; benzpyrenes, 252 *m*/*z*). The lower 3D-plot (Fig. 2B) represents a SI-plot based on the typical ions (180, 194, 204, 230 *m*/*z*) for oxygenated-PAH (e.g. 9H-fluoren-9-one, 180 *m*/*z*, 9,10-anthracenedione 180, 204 *m*/*z*). Circles in the observed chromatograms depict the PAH- and oxy-PAH peaks, respectively.

This simple "selected ion" approach to view the PAHs and oxy-PAHs, however, rely on the high stability of the aromatic, polycyclic species. Under electron impact (EI) ionisation, the aromatic compounds are relatively stable against reactions that would destroy the highly stable aromatic moieties. The high stability of aromatic compounds could be explained by electron delocalisation due to resonance effects [24]. This leads to few dominant peaks in the mass spectrum such as the molecular ions for the PAH. However, many

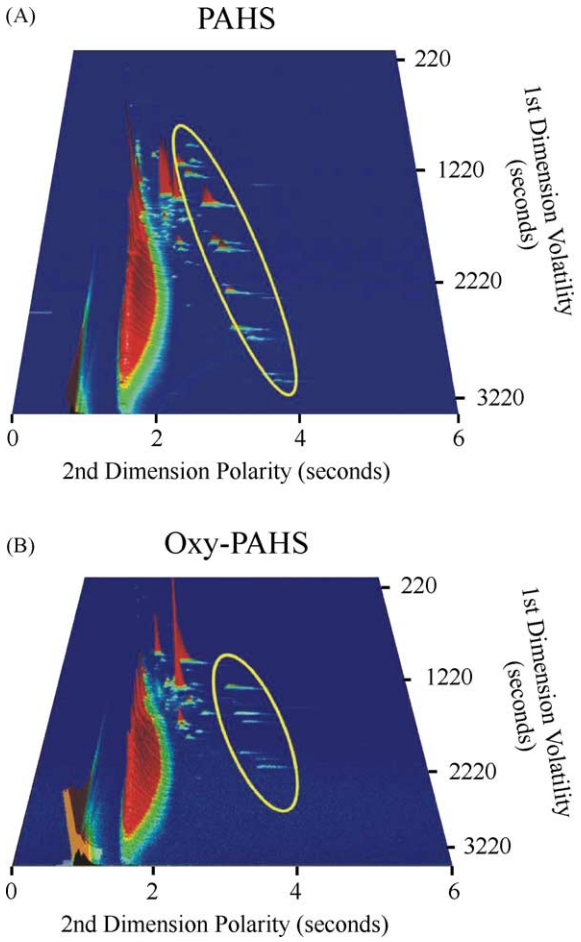

Fig. 2. Two-dimensional GC × GC–TOFMS chromatogram of an Augsburg PM2.5 sample in 3D landscape-plot: (A) selected ion (SI) representation for PAH masses (178, 202, 228, 252, 276 *m/z*); (B) selected ion representation for oxy-PAH masses (180, 194, 204, 230 *m/z*).

other chemical compounds are much less stable than PAH or oxy-PAH. For these less stable compounds, identification routines based just on a "selected ion" representation are not sufficient.

The use of the fragmentation patterns obtained from these compounds could, however provide useful identification. Pattern recognition techniques may be applied to find chemical differences in PM2.5 samples. Furthermore, the ordered appearance of the chemically related peaks in GC × GC is clearly visible. Unknown compounds and mismatched compounds could be classified (i.e. assigned to belong to a substance class or sub-class) utilising the mass spectrometric fragmentation pattern information as well as 2D-chromatographic information.

Fig. 3A shows a typical two-dimensional contour-plot of an Augsburg PM2.5 aerosol sample. The peak finding routine (based on deconvolution techniques, Chroma-TOF software, version 2.01, Leco Inc., USA) detected about 15,000 peaks. A library search (NIST) resulted that a majority of the peaks was not identifiable or exhibited library matches below 700. In order to efficiently develop and test concepts for automated peak classification, we decided to use a smaller part (Fig. 3B) of the two-dimensional plot located in a more simplistic region of the chromatogram. In this region only 1060 peaks were found. The chromatographic peak data consists of:

(i) first dimension retention times;
(ii) second dimension retention times;
(iii) peak area (analytical ion current (AIC)); and
(iv) mass spectrometric peak list.

The chromatographic peak data was exported and further analysed with external software programs (e.g. EXEL 2000 software, Microsoft Inc., USA). The peak data (i) to (iii) could be represented as a so-called bubble-plot. Each bubble represents a chromatographic peak. The bubble position is given by the retention time co-ordinates of the peak (i.e. by: (i) first retention time; and (ii) second dimension retention time). The area of the bubble corresponds to the peak area; (iii) calculated from the de-conveloted chromatographic peak. Fig. 3C shows the contour-plot of the extracted section with overlaid bubble-plot repression.

In order to evaluate the concept of substance class assignment based on fragmentation patterns, the exported mass spectrometric fragmentation pattern information was screened for characteristic ion pattern information, using the knowledge of classical mass spectroscopic analysis [24]. The rules applied in this preliminary investigation were based on ion abundances and first and second dimension retention times. The mass spectral peak selection criteria (ion abundance rules) and the retention time criteria are given in Table 1. All of the components identified by this rules were then compared to the results of the NIST library search (similarity search) Table 2. The unknowns listed in Table 2 are the components with a

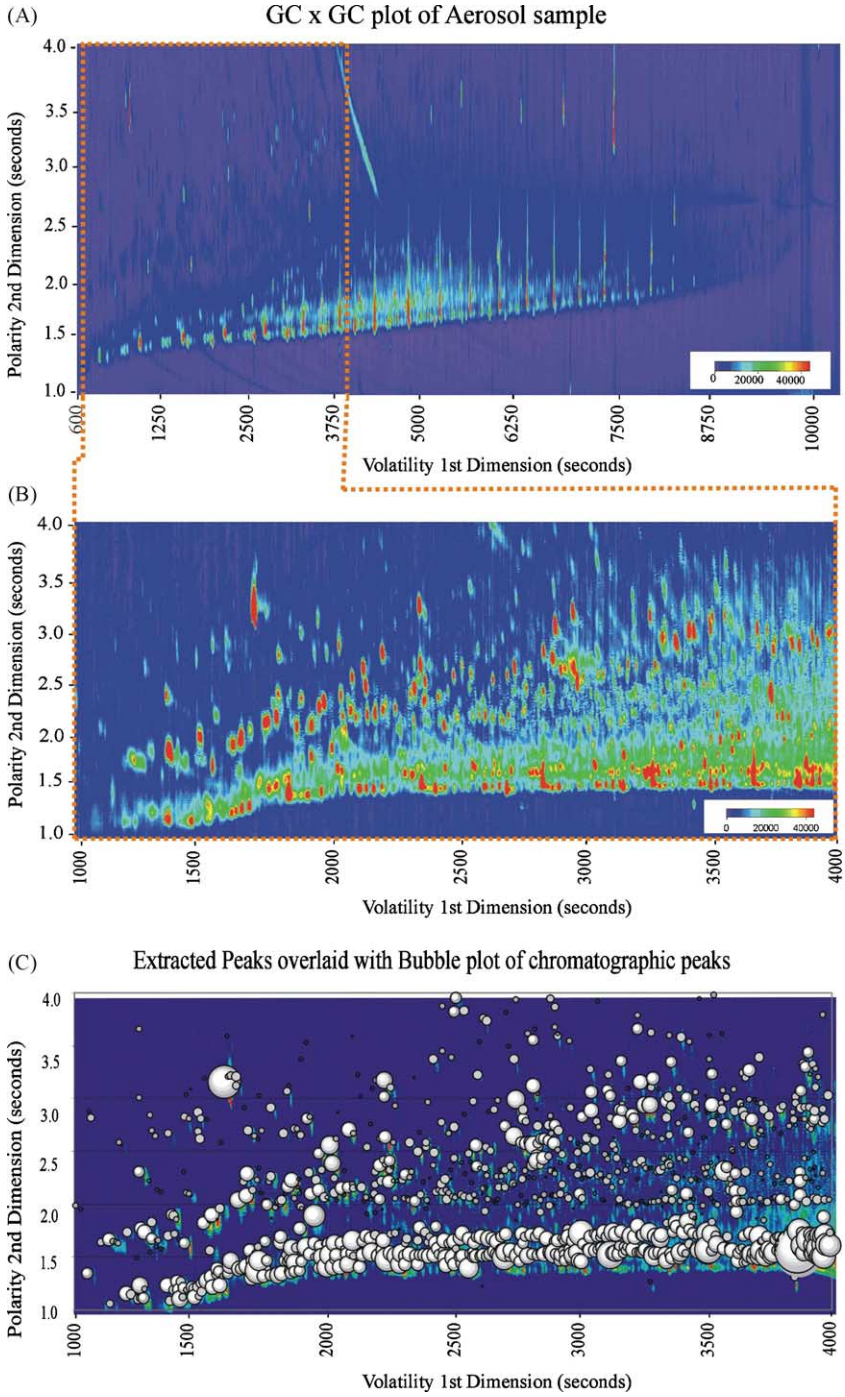

Fig. 3. Two-dimensional GC × GC–TOFMS total ion current (tic)-plot of an aerosol sample in 2D contour-plot: (A) showing the full chromatogram of the analysed aerosol with (B) the extraction of the selected section for the data analysis; (C) represents the extracted section of the above overlaid with a bubble-plot generated from the peak apexes of the same selection.

Table 1
Selection rules (ion abundance rules) for compound class identification

Compound class/group	Compound class/group identification rules	
	Mass pattern selection rules	Retention time selection rules
Alkanes	Base peak 57 or 71 with second largest peak 71 or 57	*No time rule needed* (must be within 1 and 1.5 s of second dimension)
Alkenes and cycloalkanes	Base peak 55 or 69 with both present and with three of following peaks in more than 15% relative intensity 97, 83, 70, 57, 56	Must be within 1 and 2 s of second dimension
n-Alkane acids	Peaks with base mass 60 and second highest mass 73	*No time rule*
Alkyl-substituted benzenes	(1) Peaks with mass 91 above 15% relative intensity and greater than mass 77 present above 5% relative intensity (2) Compounds with mass 77 above 25% relative intensity	*No time rule needed* (1) Generally above 2 s second dimension time with exceptions at first dimension time below 1700 s (2) For mass selection point 2 only those compounds below 2 s second dimension and below 1700 s first dimension
Polar benzenes with or without alkyl groups	Peaks with mass 77 above 25% relative intensity	Above 2 s second dimension time
Partly hydrated naphthalenes and alkanyl-substituted benzenes	Peaks with mass 91 above 15% relative intensity mass 77 above 5% relative intensity and mass 128 above 10% relative intensity	*No time rule needed* (generally above 2 s second dimension time)
Naphthalene and alkyl-substituted naphthalenes	(1) Peaks with mass 128 above 15% relative intensity and mass 77 above 5% relative intensity (2) Also peaks with base peak or in higher relative intensity than 50% masses 141, 155, 169	*No time rule needed* (generally above 2 s second dimension time)

Table 2
Confirmation by NIST library of compounds find in each group

Compound class/group	Peaks identified by mass and retention time rules	NIST			
		Confirmed	Unknown library match below 600	Mismatch	Not identified
Alkanes	104	100	4	0	5
Alkenes and cycloalkanes	184	44 alkenes 90 cycloalkanes	31	19	10 14
n-Alkane acids	4	4	0	0	0
Alkyl-substituted benzenes	229	98	114	17	2
Polar benzenes with or without alkyl groups	41	11	28	2	1
Partly hydrated naphthalenes and alkanyl-substituted benzenes	145	59	61	25	2
Naphthalene and alkyl-substituted naphthalenes	70	36	20	14	5

similarity match below 600. Mismatched components listed in Table 2 are those components that are typically misidentified by the NIST library or in some cases compounds that also fit the same rules applied to identify the group. An example of one of these mismatched components is E-2-octadecadecen-1-ol which is identified within the alkene and cycloalkane group, the mass spectra of this compound fits the group perfectly as well as the retention time specified. Future work will hopefully eliminate these exceptions.

The following different groups/rules (see Table 1) are described in more detail:

3.2.1. Aliphatic compounds (Fig. 4)

3.2.1.1. Alkanes (Fig. 4A). The fragmentation patterns in alkanes or saturated aliphatic hydrocarbons

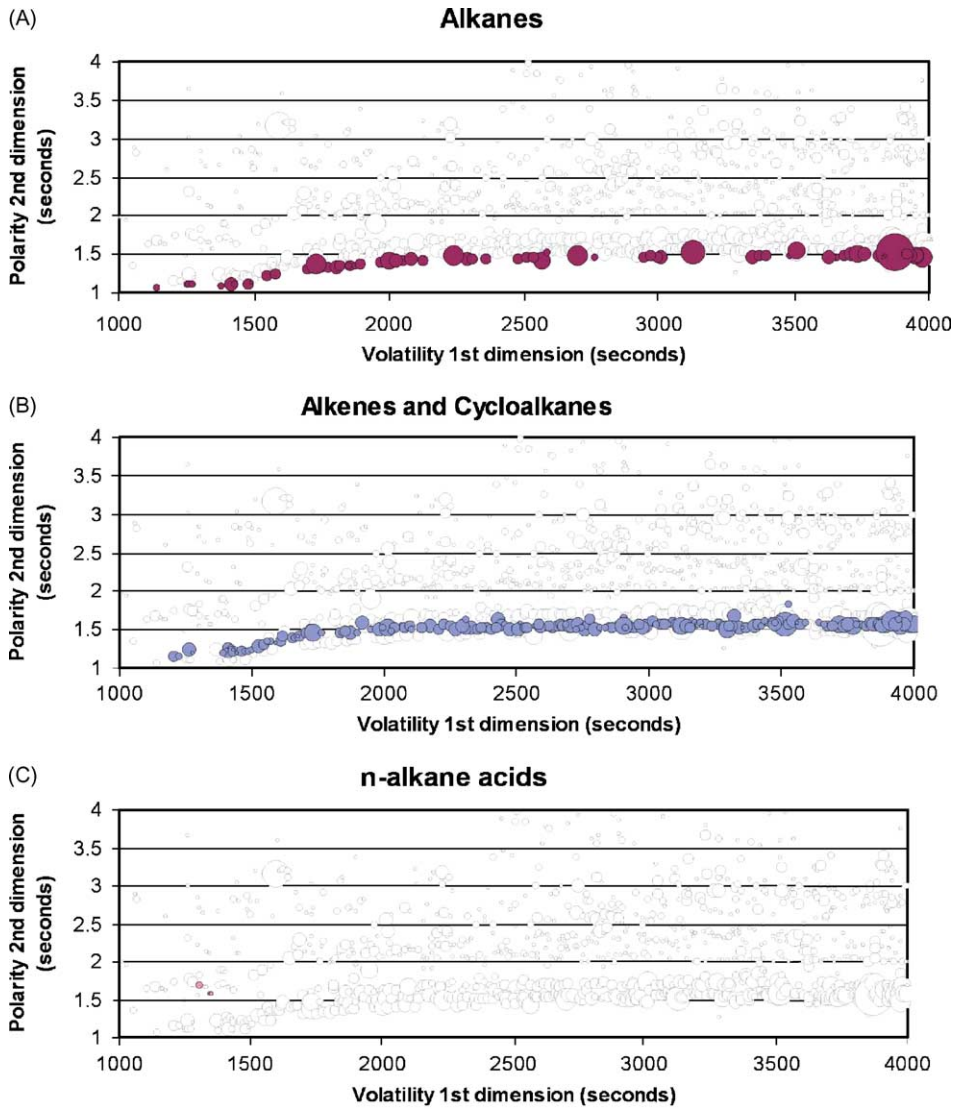

Fig. 4. The three bubble-plots of the aliphatic groups assigned by the criteria in Table 1 with: (A) the alkanes; (B) the alkenes and cycloalkanes; and (C) the *n*-alkane acids.

indicate a typical ion series of $C_nH_{2n+1}^+$ for straight chain alkanes and branched alkanes [24]. For the section of the two-dimensional plot examined, it was found that when searching for alkanes at higher than 50 *m*/*z* most alkanes have characteristic peaks at 57 and 71 *m*/*z*. The 98% of all alkanes identified by the NIST library fitted this fragmentation criterion. From the two-dimensional chromatogram, it was also easy to confirm the MS assignation of alkanes, due to the fact that they are almost unretained in the polar second column separation.

3.2.1.2. Alkenes and cycloalkanes (Fig. 4B). For unsaturated aliphatic hydrocarbons and saturated alicyclic hydrocarbons, the identification process is getting more complex. The fragmentation in alkenes due to the double bond shows increasing patterns of $C_nH_{2n-1}^+$ and $C_nH_{2n}^+$ ion series [24]. The cyclic alkanes also show a strong fragmentation pattern of $C_nH_{2n}^+$ [24]. Thus, differentiating between the two is a difficult task for search algorithms. In the approach used in this study, we focused again on *m*/*z* higher than 50 and were able to classify a large proportion of the combined two compound groups with the fragmentation criteria in Table 1. Using a retention time window became increasingly more useful, the two components have similar retention indexes in both dimensions and it is thus still not possible to separate the two, but it can be used in excluding mismatched components. Mismatched components included compounds like cyclohexanone, which fitted the same mass rules but not the retention time rules. The retention window used also contained the unretained components at 1 s up to the more polar aromatics at 2 s second dimension retention time. A separation of cycloalkanes and alkenes might be possible by inclusion of masses below 50 *m*/*z* in the future.

3.2.1.3. n-Alkane acids (Fig. 4C). The third group investigated in this preliminary study were the acids of the unbranched alkanes. The most prominent mass in the aliphatic acids is 60 *m*/*z*, this is due to γ-hydrogen rearrangement peak [24], $CRR'C(OH)_2^+$. From the spectra obtained, it was further recognised that to identify this group more accurately the second most abundant mass 73 *m*/*z* was also needed. Four acids are listed in Table 2, where only two acids were present. This is due to an error in the analysis software where second dimension peaks are combined to produce the total area of a compound. The reconstruction is complicated when there is tailing in the first dimension, which often occurs with acids on non-polar stationary phases.

3.2.2. Aromatic compounds (Fig. 5)

The aromatic groups identified in this study can be divided into four sections. When containing electro-negative substituents the dominating series are [24] $C_nH_{0.5n}^+$–$C_nH_{0.8n}^+$ ($n = 3$–6), and in the presence of electron-donating substituents or heterocyclic compounds the series $C_nH_{0.9n}^+$–$C_nH_{1.2n}^+$ is more predominant [24]. Using these characteristic fragmentation patterns, four aromatic groups were assigned.

3.2.2.1. Alkyl-substituted benzenes (Fig. 5A). The first of this group were the alkyl-substituted benzenes, this group is characterised by abundance of mass 91 *m*/*z* above 15% and also present in above 5% relative abundance mass 77 *m*/*z*. The retention window of these compounds was also limited to those eluting after 2 s in the second dimension. The group was expanded by adding some aromatic compounds with different retention and mass criteria, these were the alkylated benzenes with mass 77 *m*/*z* above 30% and retention time below 2 s second dimension and below 1700 s first dimension.

3.2.2.2. Polar benzene derivatives with or without alkyl-substituents (Fig. 5B). This group containing the aromatic benzenes with polar-substituents were identified with containing mass 77 *m*/*z* in higher than 25% relative abundance. The retention window in this extraction becomes more critical, and only compounds in the more polar region (above 2 s) of the second dimension qualify.

3.2.2.3. Partly hydrated naphthalenes and alkenyl-substituted benzenes (Fig. 5C). This class contains benzene derivatives with alkenyl-substituents as well as partly hydrated naphthalene derivatives (one benzene ring remaining). The class is characterised by the masses 77 (above 5%), 91 (above 15%) and 128 *m*/*z* (above 10%) in relative abundances. The compounds of this group elute like the other aromatic groups after 2 s on the second dimension.

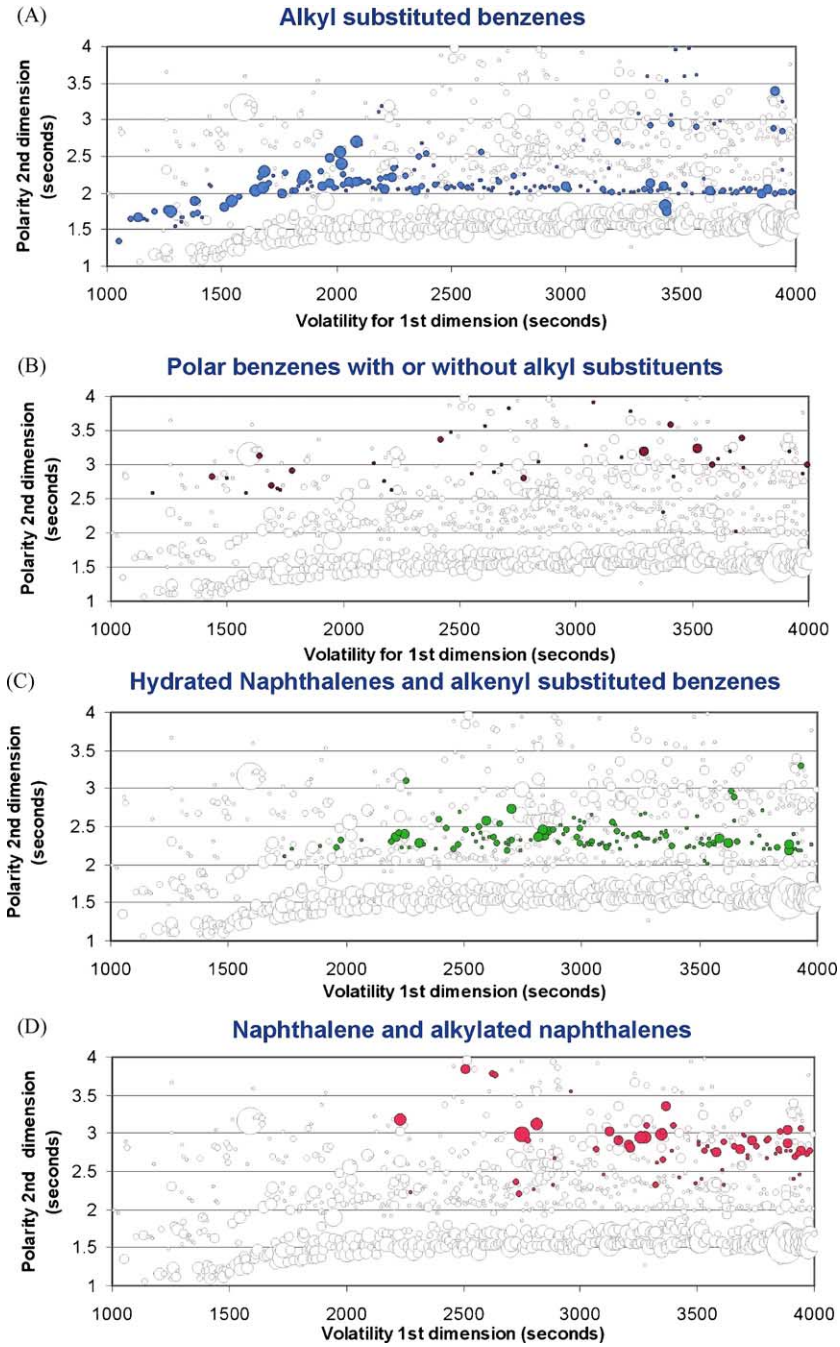

Fig. 5. The four bubble-plots created from the criteria in Table 1 for the aromatic groups: (A) the benzenes with alkyl-substitutions: (B) the polar benzene derivatives with or without alkyl-substitutions: (C) the partly hydrated naphthalenes with one benzene ring still intact and the alkenyl-substituted benzenes; and with (D) the alkylated naphthalenes.

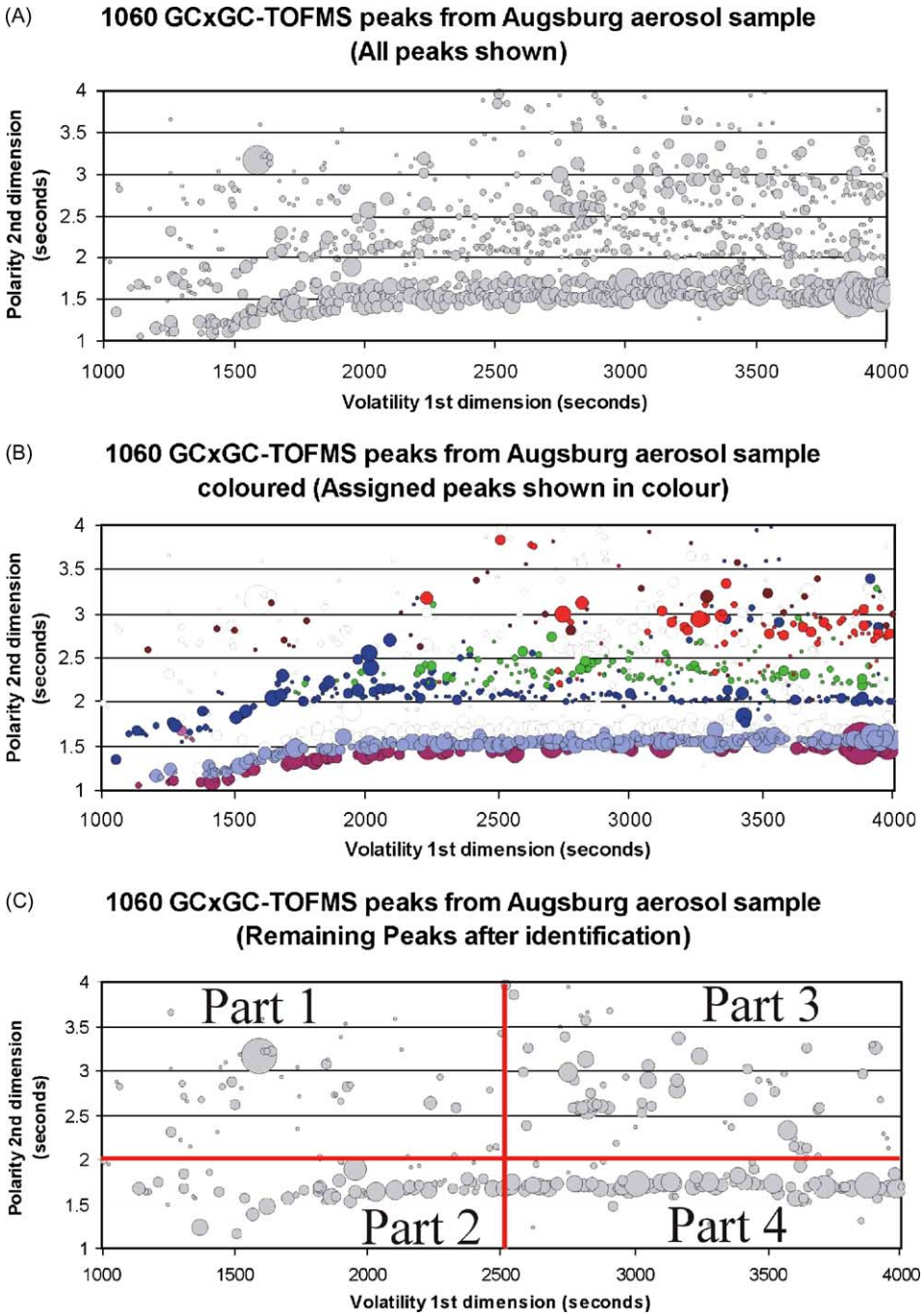

Fig. 6. (A) shows the bubble-plot of the entire peak apices used in this preliminary study for grouping; (B) indicates the different groups identified on (A); the final plot (C) shows the remaining peak apices not identified in the search criteria, and indicates the four possible groups in which these apices could be divided for identification.

3.2.2.4. Naphthalene and alkyl-substituted naphthalenes (Fig. 5D). Relative intense peaks at mass 128 *m*/*z* (above 15%) characterise the naphthalene derivatives. The alkylated naphthalenes often contain base peaks or relative high abundances of above 50% at masses 141, 155 and 169 *m*/*z*, etc. The mass 91 *m*/*z*, which is characteristic for alkylated benzene derivatives, is missing. The alkylated naphthalenes elute, like the other aromatic groups, after 2 s on the second dimension.

From these group assignments, the obtained bubble-plot from the chromatographic peak information (Figs. 3 and 6A) could now be more easily interpreted by colouring the identified peak groups (Fig. 6B). This coloured bubble plot can be used in a visual recognition of pattern changes at a glance.

For statistical evaluations, the remaining ungrouped compounds (about 35% of the total) could then be even further divided into different groups, depending on their retention time in the two dimensions as measures for their volatility and polarity. In Fig. 6C, a potential differentiation scheme for the ungrouped compounds is given. Part 1 of these groups would then be the polar more volatile components, part 2 the non-polar more volatile components, part 3 the polar less volatile components, and part 4 the non-polar less volatile components. In future applications where more groups is needed, this approach can be used even with more groups, multivariate analysis techniques could also be applied as an alternative.

4. Conclusion

With the developed search parameters, it is possible to group peaks in the 2D chromatogram into distinct chemical classes using TOF mass spectrometric information. These search routines, although being in a preliminary development stage, already show significant improvements on previously used methods in the classification of compounds.

More than 60% (by peak area) and 65% (by amount compounds “697 compounds out of the 1060 present”) of the organic inventory of the aerosol sample in the investigated region of the chromatogram could be effectively assigned to groups allowing a straight forward characterisation of the constituents of ambient particulate matter (Fig. 7) or other complex samples.

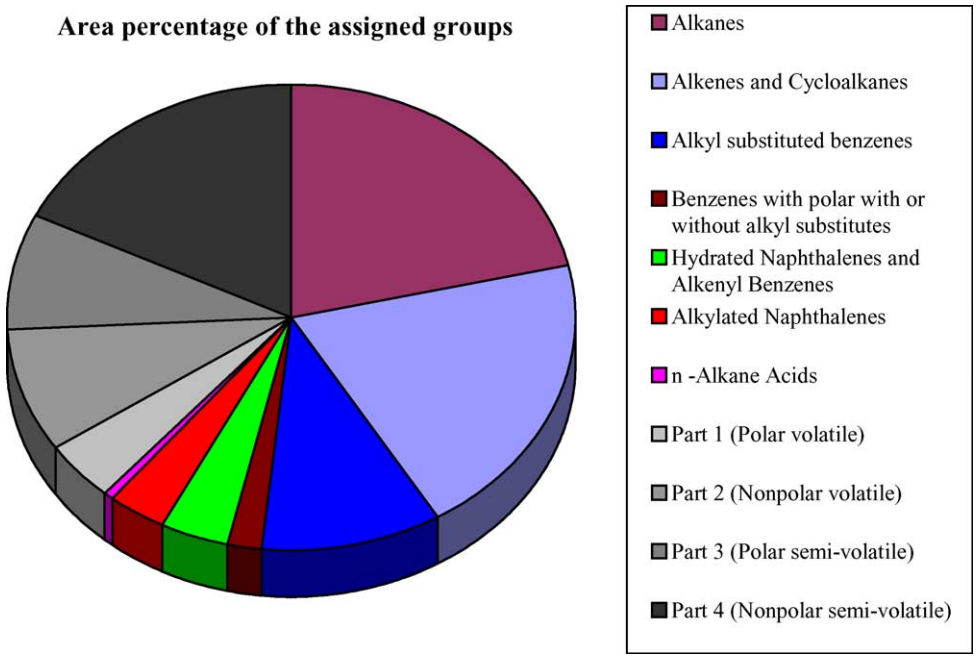

Fig. 7. Pie-graph representation of the identified groups. The different groups are evaluated according to the area of the combined peak apices in each group.

Fig. 7 shows a pie-chart diagram of the chemical class groups identified in the aerosol sample by applying the rules given in Table 1. It is of note that this initial analysis was not focused on quantitative analysis, but rather on finding possible solutions for future applications in grouping different components. The area representation in Fig. 7 should therefore be viewed in this perspective. With respect to concept of epidemiological evaluation of the health effects of the ambient particulate matter, the assigned chemical-class groups are particularly helpful. Furthermore, the groups probably aid the development of statistical models for the chemical variability in aerosols. Note, that it is likely that within the ensemble of the unidentified compounds of a chemical-class group, the majority will also belong to the same chemical class. Future work on this topic would include refinement of these searches and the inclusion of more groups.

Development of GC × GC-based methods and applications nowadays are limited by the development of suited data analysis techniques. This is in particular true for GC × GC–TOFMS with the enormity of data generated. Therefore, a universal data analysis tool for the GC × GC–TOFMS instrument, allowing a straightforward group separation, needs to be developed. This data analysis tool should allow formulating rules and constraints, which are applied on the deconvoluted GC × GC–TOFMS data set (retention times and mass spectra). Applying one or more of the different rules on the deconvoluted data set will select a sub-population of the total number of compounds. The selected compounds form a group. Ideally the groups should represent compound classes. Potential selection rules are:

(1) ion abundance rules: search for typical fragment ion patterns for compound group definition (such as given e.g. in Table 1);
(2) retention time rules: application of first and second dimension retention time constraints for compound group definition;
(3) fragmentation rules: identification of typical fragmentation patterns (e.g. typical neutral losses such as 28 *m*/*z* for loss of CO, which is a typical fragmentation process for aromatic oxygen compounds [16]);
(4) homologous series rules: identification of homologous series of compounds (e.g. increase of characteristic peaks by $n \times 14$ *m*/*z* with $n = 0, 1, 2,$ etc. for increasing number of CH_2 groups in the molecule);
(5) isotopic pattern rules: search for typical isotopic ratios (e.g. chlorine) *MS library-based rules*.

Using commercial and lab-generated MS library search routines with match factor constraints to identify compounds. Identified compounds should be assignable to groups defined by application of the rules 1–5.

This analysis tool could be applied to many research fields in analytical chemistry. User defined criteria or search routines would assist in finding unique groups that will enable GC × GC–TOFMS operators to simplify their own data. Creating new groups is a challenge as it has unlimited possibilities.

Acknowledgements

This work was carried out in co-operation with the GSF-Focus-Network “Aerosols and Health” which co-ordinates aerosol-related research within the GSF Research Centre. The Bavarian State Ministry for Regional Development and Environmental Affairs supported this work.

References

[1] D.M. Brown, M.R. Wilson, W. MacNee, V. Stone, K. Donaldson, Toxicol. Appl. Pharm. 175 (2001) 191.
[2] P. van Vliet, M. Knape, J. de Hartong, Environ. Res. 74 (1997) 122.
[3] N.N. Janssen, J. Schwartz, A. Zanobetti, H.H. Suh, Environ. Health Perspect. 110 (2002) 43.
[4] H.E. Wichmann, C. Spix, T. Tuch, G. Wölke, A. Peters, J. Heinrich, W.G. Kreyling, J. Heyder, Health Effects Institute, Cambridge, MA, Report 98, 2000.
[5] A.J. Ghio, R.B. Devlin, Am. J. Respir. Crit. Care Med. 164 (2001) 704.
[6] P.E. Tolbert, M. Klein, K.B. Metzger, J. Peel, W.D. Flanders, K. Todd, J.A. Mulholland, P.B. Ryan, H. Frumkin, J. Expo. Anal. Environ. Epidemiol. 10 (5) (2000) 446.
[7] G.R. Cass, Trends Anal. Chem. 17 (6) (1998) 356.
[8] B.R.T. Simoneit, A.I. Rushdi, M.R. Bin Abas, B.M. Didyk, Environ. Sci. Technol. 37 (1) (2003) 16.
[9] A.H. Falkowich, Y. Dudich, Environ. Sci. Technol. 35 (2001) 2326.
[10] D. Waterman, B. Horsfield, F. Leistner, K. Hall, S. Smith, Anal. Chem. 72 (15) (2000) 3563.

[11] http://www.gsf.de/KORAN/en/.
[12] J.B. Phillips, J. Beens, J. Chromatogr. A 856 (1999) 331.
[13] P.J. Marriott, R. Shellie, Trends Anal. Chem. 21 (9–10) (2002) 573.
[14] J. Blomberg, P.J. Schoenmakers, U.A.Th. Brinkman, J. Chromatogr. A 972 (2002) 137.
[15] X. Xu, J. Williams, C. Plass-Dülmer, H. Berresheim, G. Salisbury, L. Lange, J. Leliveld, Atmos. Chem. Phys. Discuss. 3 (2003) 1477.
[16] J. Dallüge, M. van Rijn, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, J. Chromatogr. A 965 (2002) 207.
[17] J. Dallüge, L.L.P. van Stee, X. Xu, J. Williams, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, J. Chromatogr. A 974 (2002) 169.
[18] T. Hyölyläinen, M. Kallio, M. Shimmo, K. Saamio, K. Hartonen, M.-L. Riekkola, in: Presentation at the First International Symposium on Two-Dimensional Gas Chromatography, Volendam, The Netherlands, 2003.
[19] J. Beens, R. Tijssen, J. Blomberg, J. Chromatogr. A 822 (1998) 233.
[20] M. van Deursen, J. Beens, J. Reijenga, P. Lipman, C. Cramers, J. High Resol. Chromatogr. 23 (2000) 507.
[21] K.J. Johnson, R.E. Synovec, Chemom. Intell. Lab. Syst. 60 (2002) 225.
[22] J.C. Giddings, Anal. Chem. 56 (12) (1984) 1258–1270.
[23] H.J. de Geus, U.A.Th. Brinkman, Trends Anal. Chem. 15 (5) (1996) 168.
[24] F.W. Mclafferty, F. Turecek, Interpretation of Mass Spectra, fourth ed., University Science Books, Sausalito, CA, 1993.
[25] D. Waterman, B. Horsfield, K. Hall, S. Smith, J. Chromatogr. A 912 (2001) 143.