Direct Inlet Probe – high-resolution time-of-flight mass spectrometry as fast technique for the chemical description of complex high-boiling samples

Uwe Käfer†‡, Thomas Gröger\*†, Christopher P. Rüger‡, Hendryk Czech†‡, Mohammad Saraji-Bozorgzad§, Thomas Wilharm∥, and Ralf Zimmermann†‡

† Joint Mass Spectrometry Centre, Comprehensive Molecular Analytics, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg †‡§

‡ Joint Mass Spectrometry Centre, Chair of Analytical Chemistry, University of Rostock, Dr. Lorenz Weg 2, 18059 Rostock, Germany;

§ Photonion GmbH, Hagenower Str. 73, 19061 Schwerin, Germany

∥ ASG Analytik-Service Gesellschaft mbH, Trentiner Ring 30, 86356 Neusäss, Germany

\*E-mail: thomas.groeger@helmholtz-muenchen.de

Abstract

Comprehensive chemical investigation of non-volatile complex mixtures, without extensive sample pretreatment, remains challenging due to the high number of constituents with different chemical properties. In past years, direct high-resolution mass spectrometry established itself as powerful technique for the detailed molecular description of ultra-complex mixtures, but was mainly used with atmospheric pressure ionization. In this study, we present a direct inlet approach with vacuum ionization and high-resolution time-of-flight mass spectrometry. Exemplary, the non-volatile fractions of crude oil were directly inserted into the ion source and volatilized under reduced pressure conditions. An applied temperature gradient enabled thermal pre-separation, according to volatility, prior to electron ionization and mass spectrometric detection. With exact mass information, peaks were assigned to elemental compositions and grouped into component classes. Moreover, the application of supervised and unsupervised statistical tools allowed differentiation of the samples on a molecular level and the identification and attribution of significant chemical features.

1. **Introduction**

Nowadays, mass spectrometry (MS) is commonly used as detector for chromatographs, in gas chromatography (GC-MS) or liquid chromatography (LC-MS). Typically, complex mixtures are separated and single components are identified with their specific retention times and mass spectrometric information. However, GC is limited by the volatility of the analytes, whereas LC requires good solubility of the sample in the liquid phase and is prone to on-column precipitation. Moreover, chromatographic separation prolongs analysis times, which may substantially limit the sample throughput.

With high-resolution mass analyzers like Orbitrap-analyzers, Fourier-transformation ion-cyclotron-resonance mass spectrometers (FT-ICR-MS) or multi-reflectron time-of-flight instruments, numerous analytical questions can be answered by direct mass spectrometry without prior chromatographic separation. In the field of Petroleomics, often a direct infusion approach is used in combination with high resolution mass spectrometry [1–5]. Here, the high mass resolving power and mass accuracy can partially compensate for the lack of chromatographic separation and different chemical classes can be discriminated by using distinct mass defects of the most common elements in petroleum samples. Amongst other things, direct infusion approaches are of particular interest because heavy petroleum fractions are not volatile enough for a GC inlet and revealed to be problematic for LC due to poor solubility in the mobile phase and irreversible adsorption to the stationary phase [6–8]. As a drawback, direct infusion is mostly limited to atmospheric pressure ionization techniques like electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI) or atmospheric pressure photo ionization (APPI) which could produce adducts and only reflect a certain fraction of the chemical space [9–11]. Other atmospheric pressure ionization techniques, which do not rely on liquid direct infusion, such as, direct analysis in real time (DART) [12] and atmospheric solids analysis probe (ASAP) [13], became popular in the recent years for direct analysis of complex samples without sample preparation. DART and ASAP were also recently used for characterization of petroleum samples [14,15] but, like the other atmospheric ionization techniques, also suffer from matrix effects. Further techniques that are regularly used for the ionization of solid samples are field desorption [16–18] and laser desorption MS [19–21]. Both techniques produce mainly fragmentation-free mass spectra but require a more sophisticated instrumentation.

A different approach is represented in the field of evolved gas analysis, where samples are thermally treated in a vaporizing unit and evolving gases, either from volatilization or pyrolysis at high temperatures, are transferred to a mass spectrometric detection via a transfer capillary [22,23]. Here, volatile components can be separated according to their boiling point prior to MS analysis. Non-volatile organic components typically undergo thermal decomposition under atmospheric conditions at temperatures higher than 400°C. Subsequently, the identification of the pyrolysis gases can be used for the investigation of structural features of the sample.

Direct inlet probe (DIP) – MS can also be classified to the field thermal analysis, because the inserted sample is heated either actively by an integrated thermoelement [24] or passively by the filament [25] directly in the ion source. In contrast to classical thermogravimetric analysis-MS (TGA-MS), the sample is volatilized under considerably reduced pressure, whereby the volatility of substances is increased. Moreover, losses due to sample recondensation in the transferline, which are reported for TGA-MS [26] cannot occur.

Being one of most established introduction techniques for MS, DIP-systems evolved from a manual insertion of the push rod to fully automated devices which are able to apply precisely controlled temperature gradients and are compatible with autosamplers [27–29]. Recently, Flego et al. used DIP coupled to nominal-mass-resolution MS with electron ionization (EI) for the investigation of different crude oils, bitumen and asphaltenes [30,31]. They differentiated several samples based on their complex mass spectrometric fingerprint and used the time-resolved MS information for correlation with physical properties like API gravity. However, the unit-mass-resolution of the single-quadrupole instrument did not allow an assignment of chemical features to the complex mass spectrometric pattern and, thus, the study was limited to a visual evaluation of defined areas in two-dimensional contour plots. In this regard, high-resolution mass spectrometry (HRMS) can substantially improve the quality of information obtained from DIP-MS analysis because of its ability to resolve isobaric components and subsequently assign elemental compositions to observed mass traces.

In this study, we present a novel approach for the mass spectrometric investigation of complex, high boiling samples by using DIP with electron ionization (EI) coupled to a multi-reflecting time-of-flight mass analyzer, which normally is used as detector for GC×GC. Exemplary, we analyzed six different non-volatile and highly complex bitumen and vacuum residues from crude oil refining, which are known to be challenging for any analytical technique. The application of a temperature gradient directly in the ion source was used to volatilize constituents with atmospheric pressure boiling points up to approximately 750°C. Furthermore, a rough separation according to boiling points was achieved prior to MS analysis. HRMS allowed discrimination of isobaric features and assignment of elemental compositions for derived mass traces. However, the applied time-of-flight mass spectrometer is not able to resolve some common mass splits < 10 mDa found in petroleum samples. Nevertheless, we will show that an inspection and further processing of some contentious mass splits will lead to a plausible assignment of elemental compositions. Collected mass spectra were used to differentiate samples based on their molecular fingerprint. Moreover, with multivariate statistical analysis, allocation of significant chemical features to samples of different origins was possible.

**2. Experimental section**

*2.1 Instrumentation*

DIP-HRMS measurements were carried out on a Pegasus® GC-HRT 4D (LECO, St. Joe, USA) equipped with direct inlet probe (SIM GmbH, Oberhausen, Germany). TGA-HRMS measurements were performed at the same instrument by coupling a STA F449 thermo-balance (Netzsch, Selb, Germany) (**Figure 1A**). For the coupling, a heated deactivated fused silica capillary (250 °C) and heated interfaces on both sides of the capillary were used to minimize re-condensation of evolved gases [32].

Mass spectra were collected in positive mode with an acquisition rate of 0.5 Hz in a mass window of 250 to 850 *m/z* in ultra-high-resolution-mode (64 reflections). In this mode an average (18 measurements) resolution of 52k for *m/z* 413.9775 was achieved. For the comparison of TGA- and DIP-HRMS, high-resolution-mode (32 reflections) was used with a mass range from 100 to 1000 *m/z* and a mass resolution of more than 30k for m/z 413.9775. The ion source temperature was kept at 250°C for all measurements. Perfluorotributylamine (PFTBA) was continuously added to the gas phase as internal standard for mass calibration. A complete overview of key instrumental parameters is listed in the Supporting Information (**Table S1**).

*2.2 Sample material*

In total, six petroleum residues from four different refineries were analyzed. More specific, bitumen SRB\_5070 and SRB\_70100 derived from refinery 1 and MB\_OXID and MB\_PMB derived from refinery 2, refer to industrial products, according to normative regulations. SRB\_5070 and SRB\_70100 represent straight run bitumen (EN 12591), while MB\_OXID is classified as oxidized (EN 13304) and MB\_PMB as polymer modified bitumen product (EN 14023). Additionally, two self-made vacuum residues, VR\_CPC and VR\_MP were produced by vacuum distillation of the crude oils “CPC blend” (Kazakhstan) and “Mittelplate” (Germany), with atmospheric equivalent initial boiling points of 570°C (VR\_CPC) and 525 °C (VR\_MP). For the distillation, an automated distillation system was used (PD 100cc & PD 200cc, Plodist, Meckenheim, Germany). In contrast to the four industrial products, VR\_CPC and VR\_MP represent samples with extreme physical properties. Supporting information contains the values for needle penetration according to DIN EN 1426 for all six samples (**Table S2**). For DIP-HRMS measurements, the samples were dissolved in dichloromethane (HPLC grade, Roth, Karlsruhe, Germany).

*2.3 DIP-measurements*

For DIP measurements, one µL of sample solutions (5 µg/µL) were manually transferred to single-use sample containers in the vessel-holder of the push rod. Before introduction into the ion source the solvent was evaporated for 2 minutes at 50°C. Subsequently, the push rod transferred the sample vessel into the ion source where a temperature gradient was applied from 50°C to 400°C (1 °C/s) with 3 minute hold at 400°C, resulting in a total analysis time of 520 s.

*2.4 Data processing & data handling*

Post-calibration of acquired mass spectra was performed with the ChromaTOF HRT software (v5.10*,* LECO, St. Joe, USA) based on internal standard (PFTBA). For further processing, a home-build MATLAB script (R2017a, the Mathworks Inc., Massachusetts, USA) was used. Briefly, each spectrum within each measurement was calibrated based on the internal calibrant (PFTBA). Minimal abundance filter for each spectrum was 50 counts, which reduced spectral noise and processing time. Found features in each spectrum were assigned to a sum formula containing C, H, N, O, and S with #C ≤ 100, #H ≤ 200, N ≤ 1; #O ≤ 2, and #S ≤ 2 within a 5 ppm error window according to the smallest mass difference. Additionally, possible elemental compositions were limited to a minimal number of 0 DBEs and a maximum number of three heteroatoms. Peaks for which no elemental composition was found according to the rules above were removed. **Figure S1** and **Figure S2** visualize achieved mass accuracy and distribution of relative errors.

*2.5 Statistical analysis*

For statistical analysis, the abundance data were first normalized to total intensity, subsequently power transformed to reduce heteroscedasticity and finally auto-scaled. Hierarchical cluster analysis (HCA), with Ward’s minimum variance algorithm and Euclidian distance as metric was used for unsupervised exploration of the relations between the samples. For principal component analysis (PCA), TIC normalized abundances were power transformed and in contrast to the first data exploration, vast-scaled to increase sensitivity for marker features. Prior to PCA, the known group structure was included in the analysis by applying analysis of variance (ANOVA) with subsequent Bonferroni adjustment to correct the alpha-error accumulation along the variables. Features with adjusted p-values <0.05 were regarded as statistically significant and selected as input variables for a subsequent PCA.

**3. Results & Discussion**

*3.1 Temperature gradient and thermal separation*

In order to vaporize substances, the temperature of the applied sample is slowly increased. Therefore, a heater and a temperature sensor are installed directly under the sample containers. This setup allows controlled heating and enables applying defined temperature gradients. Temperature limits of the DIP range from room temperature to the maximum temperature of 400 °C, although it is possible to decrease the lower limit beyond 0°C with active CO2 cooling for low boiling substances. For the study of high boiling petroleum fractions, we used a linear gradient from 50 to 400°C with a solvent evaporation step before the analysis as previously described. After the maximum temperature was reached, the temperature was held for three minutes to ensure that most of the substances were evaporated and to prevent re-adsorption of high boilers to the push rod.

Comparable to thermo-analytical techniques, such as TGA, components can be separated according to their volatility. Heating the sample in high-vacuum, the effective boiling point of components is reduced according to the Clausius-Clapeyron-relationship. Hence, high boiling components are volatilized at lower temperatures and thermal decomposition processes could be prevented to a large extent. A direct comparison between the different pressure regimes applied for DIP and TGA is shown in **Figure 1** to demonstrate the capability of the DIP-setup to analyze high boiling matrices. TGA-MS analysis showed almost no evolving gases in the range from 35°C to roughly 350°C, which gives evidence that almost no volatile constituents are present in the bitumen sample (**Figure 1B**). This result met the expectation, since bitumen is derived from the vacuum residue from crude oil distillation. At 350°C to 500°C, pyrolysis was observed, indicated by thermal fragments smaller than 400 *m/z* (**Figure 1D**) similar to previous studies [23]. In contrast, the boiling curve obtained from the DIP-MS measurement of the bitumen sample begins at 100°C DIP temperature (**Figure 1C**), and with increasing temperatures, heavier molecules were volatilized and ionized eventually (**Figure 1E**). Due to the direct introduction into the heated source at reduced pressure, recondensation of constituents is also avoided to a large extent.

Although the temperature was kept constant at 400°C at the end of a measurement, the *m/z* range of detected peaks increased, which may be explained by a delayed heat transfer from the push rod to the sample material. Ions higher than 850 *m/z* could be observed, despite the most abundant signals are from EI-fragments in the low *m/z* region. In order to estimate the boiling point reduction in the source-vacuum with DIP compared to TGA atmospheric pressure setup, a mixture of four PAHs standards was measured and the temperatures of their appearance was compared to their literature boiling points (**Figure S4**).For this estimation, a mean boiling point reduction of about 330 K was obtained. Taking this reduction into account, an initial boiling point of roughly 450 to 500 °C under atmospheric conditions for the SRB\_5070 bitumen sample can be concluded, which fits very well to the final atmospheric equivalent temperatures of crude oil refining [33].

*3.2 High resolution mass spectrometry*

While the capabilities of thermal separation under reduced pressure with direct inlet probe is already established [25], the combination with high resolution mass spectrometry enables a more detailed description of the evolved gases. However, for complex petroleum samples, such as bitumen, the most critical mass splits that have to be resolved are C3/SH4 (3.4 mDa), 13CH/N (8.1 mDa), CH2/N (12.6 mDa), 32S/O2 (17.8 mDa) and CH4/O (36.4 mDa). Especially for the discrimination of C3/SH4 ions a mass resolution higher than 100k for the investigated mass range is needed [3]. Nevertheless, even if the mass peak is not completely resolved, a shift of the peak apex towards the exact mass of the more abundant ion can be observed.

Figure **S2A** exemplary shows found signals with nominal mass 324 in three replicated measurements of VR\_CPC and their proposed assignment to an elemental composition (**Figure** **S2C**). In each of the three independent measurements, nearly all of the mass bins in the range between 324.185 and 324.200 were assigned to C25H24+ (*m/z* 324.1872). In contrast, for MB\_OXID, a mass shift of the unprocessed signals could be observed and therefore a considerable part was assigned to the C22H28S+-ion (*m/z* 324.1906). Considering the good reproducibility of the assignments within the replicates, this finding can be regarded as reasonable and both features were used for the discrimination of the samples. The proposed procedure will not completely overcome the physical limitations for the discrimination of C3/SH4 and 13CH/N- mass splits, and especially compositions with a high number of heteroatoms will remain challenging and only elemental compositions with a maximum of three heteroatoms was considered. With this approach, it was possible to assign molecular formulas for approximately 80% of detected signals with a mean mass accuracy of 1.56 ppm. From the assigned elemental composition, the amount of DBE can be derived, which represent the number of rings and double bonds in a molecule, a widely used concept in the field of Petroleomics to estimate the aromaticity of a given elemental composition. Moreover, features with the same number of DBE can be considered as homologous series because they have the same degree of unsaturation and only differ in the amount of CH2 groups [34]. Different isomeric structures are not considered for this approach.

*3.3 Electron ionization*

EI at 70 eV is a universal ionization technique with high ion yields and good reproducibility but also leads to strong fragmentation and the molecular ions often have low abundance or are not observed at all. Especially molecular structures with low bonding energies, such as alkanes and aliphatic sidechains, are fragmented to smaller ions. For aliphatic structures the most abundant ions are typically smaller fragments originating from sigma-cleavages. These characteristic radical cations for aliphatics (e.g. *m/z* 43, 57, 71) have calculated DBEs of 0.5 because of a missing hydrogen anion. Here, strong fragmentation leads to an absence of the molecular ion for large aliphatics and almost identical mass spectra are derived, which prevents efficient discrimination of the homologues with EI.

In the contour plot for DIP-HRMS measurements, all features for the CH-class with 0.5 DBE were found to have their maximum at the same time (**Figure 2A,E**), which was identical to the maximum of the total ion count, with decreasing abundance for higher homologues. These ions are derived from alkanes and alkyl-sidechains and are continuously observed over the whole run. They can be used as a sum parameter for the estimation of the global portion of saturated structural features in the sample.

In contrast, aromatic structures are most widely retained after ionization because of the stable conjugated double bond system and resonance stabilization by delocalization of the charge. Therefore, alkylated three-ring-aromatics showed Gaussian- shaped release characteristics in the contour plot by visualizing features of the CH-class with 10 DBE (**Figure 2 B,F**)**.** Moreover, the members of this homologues series were separated in respect to the increasing boiling point for the addition of CH2 units. The same concept could also be used for other classes, such as alkylated benzothiophenes (**Figure 2 C,G**) or -dibenzothiophenes (**Figure 2 D,H**), which are known to be prominent compound classes in heavy petroleum fractions.

Although fragments from EI do carry considerable structural information, they also lead to complex spectra, which can hinder the interpretation. However, if we consider EI fragmentation, two main groups of reactions are prevalent. On the one hand, the dominating simple bond cleavages form odd nominal mass fragment ions (i.e. fragment ions depicting non-integer DBE values). On the other hand, the much less common rearrangement-type of fragmentation reactions predominately form even nominal mass fragment ions (i.e. fragment ions depicting integer DBE values).  Rearrangements are typical only for ions with rather low internal energy. This can be used to filter-out the majority of fragment ions from the data set and thus enhance the visibility of molecular ions drastically by removing all features with odd nominal mass (except nitrogen compounds, Nitrogen rule [35]) or non-integer DBE values, respectively.

**Figure 2,J** visualizes all detected features from the CH, CHN, CHO and CHS-class in the SRB\_5070 sample with integer DBE values. In general, larger *m/z* were found to be volatilized at higher temperatures, which reflects a boiling-distribution of the hydrocarbon analytes. Besides the molecular weight, also the degree of aromaticity, represented by the number of DBE, influences the volatility [36,37]. Therefore, features with DBE values higher than 20 were first detected at the end of a measurement at high temperatures. Furthermore, more polar species, containing heteroatoms, also released at higher temperatures than pure hydrocarbons, which fits the Boduzynski model [36].

The capability of DIP-MS, to separate the analytes before mass spectrometric detection is limited to a boiling distribution comparable to distillation with one theoretical plate. Nevertheless, it is a valuable approach to simplify the mixture before the mass spectrometric detection, which improved the automated assignment of elemental composition due to reduced complexity and time profiles. Moreover, assignments were checked for their reliability by comparing the boiling distribution of homologues features.

*3.4 Investigation of several different vacuum residues and bitumen*

For the characterization of six different high boiling fractions from four refineries, each sample was measured three times, which allowed to estimate the combined variance caused by the injection, introduction, ionization, spectral acquisition as well as post-processing. For each measurement, the time-resolved intensities of the found features were added up to a sum spectrum and normalized to the TIC in order to minimize the variance from injected sample amount and other sources of variance. **Figure 3** shows the compound class distribution of all assigned mass traces in the six different samples as mean value derived from triplicates. The most abundant compound classes, in all samples were CH, CHN, CHO, CHS and CHS2. Regarding their relative intensities, this is in good accordance to elemental compositions of high boiling petroleum fractions found in literature[33]. Here, the universal EI reflects the elemental composition of vacuum residues and bitumen better than more selective ionization techniques, such as ESI, which is often used in Petroleomics field [38,39] and selectively ionizes polar species, thus promoting the abundance of heteroatomic constituents. However, a limited mass range from 250 to 850 *m/z* was used to focus on heavy ions, hence, substance groups which show no strong molecular ion formation or base peaks below 250 *m/z*, are underrepresented. This finding is especially true for aliphatics, which show a strong fragmentation as previously discussed. Nevertheless, significant differences in the compound class distribution could be found for the six crude oil residues or bitumen. SRB\_5070 and SRB\_70100, two straight-run bitumen from the same refinery revealed a high compositional similarity compared to samples from other refineries. For example, both showed significantly higher abundance in the CHN-class than the other samples while showing below-average abundance for pure hydrocarbons (CH-class). This similarity could indicate that both samples are derived from the same feedstocks and similar processing. A similar observation was found for the MB\_PMB and MB\_OXID samples, two bitumen samples from a different refinery, which both revealed significantly high signal intensities for pure hydrocarbon species and features from the CHO-class.

Although compound classes allow for a rough discrimination of the samples, only superficial information about the chemical composition is gained from sum information. Exemplary, VR\_CPC and MB\_OXID exhibit comparable responses for the dominant compound classes, though having vastly different physical properties. Structural features within the four most abundant compound classes CH, CHN, CHO and CHS are compared between these two samples in **Figure 4** utilizing DBE versus carbon number diagrams, which reveals significant differences in the specific molecular pattern. A complete overview of all samples can be found in the supporting information **(Figure S5-S7)**.

For all samples strong signals near the planar polycyclic aromatic hydrocarbon limit [40] were observed in the four compound classes, originating from the aromatic core structures of heavier molecules due to EI fragmentation. Despite the hard EI, also higher homologues of alkylated aromatic species were observed up to the mass range limit of 850 *m/z* for CH- and ions from CHN- or CHO-class were almost exclusively found in the aromatic region (more than four DBE), representing heteroaromatic compounds or other resonance stabilized structures such as e.g. polycyclic aromatic -anilines or -phenols. Aliphatic heteroatom containing ions revealed only low abundance or were not observed at all, which is partially caused by heteroatom loss due to fragmentation, a well-known fragmentation path for amines, alcohols, aldehydes, ketones or thiols [35]. Visual comparison of the molecular profile of the MB\_OXID and VR\_CPC bitumen samples reveals significant differences in CH-, CHN-, CHO- and CHS-class. For example, MB\_OXID-sample showed a strong abundance of alkylated -benzothiophenes (CHS-class, DBE 6) and -dibenzothiophenes (CHS-class, DBE 9) with carbon numbers from 20 to 50, while sulfur-containing compounds in VR\_CPC showed to be shifted to higher carbon numbers. The same observation could be made for CH-, CHN- and CHO-class-components, which also revealed higher average carbon numbers for the VR\_CPC sample. This observation can be explained by the high initial boiling point of VR\_CPC. While industrial bitumen, such as MB\_OXID, is typically distilled to roughly 500°C, VR\_CPC was a self-made vacuum residue with an initial boiling point of 570°C.

*3.5 Statistical analysis*

In order to demonstrate the information contained in the DIP-EI mass spectra and to comprehensively extract features specific for the different bitumen and vacuum residues, supervised an unsupervised statistical approaches were performed. First, a hierarchical cluster analysis (HCA) was conducted, as described in the experimental section, to classify all samples according to their mass spectrometric fingerprint, without previous knowledge about group structures **(Figure 5A)**. A cophenet coefficient [41] of 0.79 confirmed the adequate representation of the original data with the applied model. HCA proved the spectral similarity of replicates and distinct clustering of all six samples was found. Moreover, as suggested earlier, samples from the same refineries such as MB\_5070 and SRB\_70100 or MB\_OXID and MB\_PMB formed higher level clusters. The two self-made vacuum residues, originating from two different crude oils formed separated groups, indicating their structural dissimilarity to the four industrial bitumen products SRB\_5070, SRB\_70100, MB\_OXID and MB\_PMB. The results of the HCA suggest great potential of DIP-HRMS as fingerprinting technique for source identification, process monitoring or identification of production outliers. This is in particular true for complex matrices that are problematic to analyze with other methods, like petroleum derived high boilers.

While HCA demonstrated sufficiently high dissimilarities between the fingerprints of the bitumen samples, the subsequent data analysis includes the knowledge about sample types and aims to extract mass spectrometric features for those classes. First, ANOVA was performed with Bonferroni adjustment to remove redundant *m/z*. At a significance level of 0.05, the number of variables (*m/z*) was reduced from initially 34661 to the 1160 significant features (**Figure S8**). As previously discussed, *m/z* signals that differ in less than 10 mDa are hardly resolved with a resolution of 50k, and therefore can lead to arbitrary assignment of respective elemental compositions. In cases where no reliable assignment is possible, the variance within the triplicates increases and questionable features are filtered out of the dataset, which reduces the influence of possible processing errors. **Figure 5B** illustrates the score plot for the first two principal components, which explains 70% of the total variance, and led to grouping of the six samples in respect to their origin, similar to the HCA.

We regard 70% of explained variance as sufficient and keep two principal components for simplicity of the data visualization. Furthermore, it is known that PCA results may suffer from a higher number of variables than samples. However, we assume appropriate modelling of the data by PCA as the influence plot [42], consisting of residuals (Q) and Hotelling’s T2 does not reveal any sample with simultaneous high influence on the PCA result and insufficient representation (**Figure S9**). In the loadings plot (**Figure 5C**), the contribution of each feature to the principal components is depicted, which can therefore be used to identify relevant *m/z* for the discrimination between the samples. For example, features from the CHN class showed to be significant for SRB\_5070, SRB\_70100 and VR\_MP samples (**Figure 5D**), similar to earlier findings from the comparison of all assigned features to compound classes (**Figure 3**). Moreover, we highlighted features **(Figure 5E)**, discussed earlier for significant spectral differences of VR\_CPC. The negative loadings of PC1 are mostly hydrocarbon features with more than 40 carbon atoms, reflecting the higher average molecular mass of VR\_CPC, due to the high initial boiling point, as discussed before. Summarized, the results from multivariate analysis showed that DIP-HRMS can potentially be used as a fast and straightforward technique for revealing chemical fingerprints and identify chemical markers in complex samples.

1. **Conclusions**

In this study we presented a new approach for the analysis of complex and high boiling matrices applying a direct inlet - high resolution mass spectrometry setup (DIP-HRMS). The possibility to vaporize high boiling components directly in the ion source and to separate complex mixtures according to their volatility reduces complexity before mass spectrometric detection, which improves elemental composition assignment based on repetitive building blocks. Exact mass information and sophisticated data processing enables the identification of chemical groups and detailed description of the complex mixture. Although the C3/SH4 (3.4 mDa) and the 13CH/N (8.1 mDa) mass splits could not be resolved completely, assignment of critical sum formulas showed to be robust enough for a comparative investigation of the sampleset. Moreover, by filtering the ions according to the Nitrogen rule for integer DBE values the visibility of molecular ions was increased, for improved interpretation of spectral data. Furthermore, an application for discrimination of different vacuum residues of crude oils on the basis of their mass spectrometric fingerprint was shown. With statistical approaches these fingerprints could also be translated into identification of characteristic chemical structures. Due to the short analysis times below 10 minutes and the possibility for automation, DIP-HRMS will be a powerful technique for screening approaches on complex low volatile sample materials. Moreover, because of the versatility of time-of flight-mass spectrometry, DIP-HRMS could be implemented combined with TGA- and (GC×)GC-inlet on the same mass analyzer platform for the investigation of complex samples over the complete boiling range. The potential of the described method as fast and robust method for the investigation of complex matrices is even increasing, considering, that the next generation of high-resolution-time-of-flight instruments could achieve mass resolutions higher than 100,000 [43].

**Associated content**

Additional information is provided in the supporting information as stated in the text.

Acknowledgments

The authors thank LECO Europe and SIM GmbH for providing the Direct Inlet Probe and technical support. Financial support by ZIM (Zentrales Innovationsprogramm Mittelstand) is gratefully acknowledged.

References

[1] R. Gavard, D. Rossell, S.E.F. Spencer, M.P. Barrow, Themis: Batch Preprocessing for Ultrahigh-Resolution Mass Spectra of Complex Mixtures, Anal. Chem. 89 (2017) 11383–11390.

[2] T. Kekäläinen, J.M.H. Pakarinen, K. Wickström, V.V. Lobodin, A.M. McKenna, J. Jänis, Compositional Analysis of Oil Residues by Ultrahigh-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Energy Fuels 27 (2013) 2002–2009.

[3] C.F. Klitzke, Y.E. Corilo, K. Siek, J. Binkley, J. Patrick, M.N. Eberlin, Petroleomics by Ultrahigh-Resolution Time-of-Flight Mass Spectrometry, Energy Fuels 26 (2012) 5787–5794.

[4] A.M. McKenna, G.T. Blakney, F. Xian, P.B. Glaser, R.P. Rodgers, A.G. Marshall, Heavy Petroleum Composition. 2. Progression of the Boduszynski Model to the Limit of Distillation by Ultrahigh-Resolution FT-ICR Mass Spectrometry, Energy Fuels 24 (2010) 2939–2946.

[5] A.M. McKenna, J.M. Purcell, R.P. Rodgers, A.G. Marshall, Heavy Petroleum Composition. 1. Exhaustive Compositional Analysis of Athabasca Bitumen HVGO Distillates by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Definitive Test of the Boduszynski Model, Energy Fuels 24 (2010) 2929–2938.

[6] R.B. Boysen, J.F. Schabron, The Automated Asphaltene Determinator Coupled with Saturates, Aromatics, and Resins Separation for Petroleum Residua Characterization, Energy Fuels 27 (2013) 4654–4661.

[7] S. Gutiérrez Sama, M. Farenc, C. Barrère-Mangote, R. Lobinski, C. Afonso, B. Bouyssière, P. Giusti, Molecular Fingerprints and Speciation of Crude Oils and Heavy Fractions Revealed by Molecular and Elemental Mass Spectrometry: Keystone between Petroleomics, Metallopetroleomics, and Petrointeractomics, Energy Fuels 32 (2018) 4593–4605.

[8] C. Jarne, V.L. Cebolla, L. Membrado, K. Le Mapihan, P. Giusti, High-Performance Thin-Layer Chromatography Using Automated Multiple Development for the Separation of Heavy Petroleum Products According to Their Number of Aromatic Rings, Energy Fuels 25 (2011) 4586–4594.

[9] Y. Cho, A. Ahmed, A. Islam, S. Kim, Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics, Mass Spec Rev 34 (2015) 248–263.

[10] J.V. Headley, K.M. Peru, M.P. Barrow, Advances in mass spectrometric characterization of naphthenic acids fraction compounds in oil sands environmental samples and crude oil-A review, Mass Spec Rev 35 (2016) 311–328.

[11] D. Zhan, J.B. Fenn, Electrospray mass spectrometry of fossil fuels, International Journal of Mass Spectrometry 194 (2000) 197–208.

[12] R.B. Cody, J.A. Laramée, H.D. Durst, Versatile New Ion Source for the Analysis of Materials in Open Air under Ambient Conditions, Anal. Chem. 77 (2005) 2297–2302.

[13] C.N. McEwen, R.G. McKay, B.S. Larsen, Analysis of Solids, Liquids, and Biological Tissues Using Solids Probe Introduction at Atmospheric Pressure on Commercial LC/MS Instruments, Anal. Chem. 77 (2005) 7826–7831.

[14] W. Romão, L.V. Tose, B.G. Vaz, S.G. Sama, R. Lobinski, P. Giusti, H. Carrier, B. Bouyssiere, Petroleomics by Direct Analysis in Real Time-Mass Spectrometry, J. Am. Soc. Mass Spectrom. 27 (2016) 182–185.

[15] M. Farenc, Y.E. Corilo, P.M. Lalli, E. Riches, R.P. Rodgers, C. Afonso, P. Giusti, Comparison of Atmospheric Pressure Ionization for the Analysis of Heavy Petroleum Fractions with Ion Mobility-Mass Spectrometry, Energy Fuels 30 (2016) 8896–8903.

[16] W.L. Mead, Field ionization mass spectrometry of heavy petroleum fractions. Waxes.

[17] T.M. Schaub, R.P. Rodgers, A.G. Marshall, K. Qian, L.A. Green, W.N. Olmstead, Speciation of Aromatic Compounds in Petroleum Refinery Streams by Continuous Flow Field Desorption Ionization FT-ICR Mass Spectrometry, Energy Fuels 19 (2005) 1566–1573.

[18] P. Leinweber, H.-R. Schulten, Differential thermal analysis, thermogravimetry and in-source pyrolysis-mass spectrometry studies on the formation of soil organic matter, Thermochimica Acta 200 (1992) 151–167.

[19] C. Becker, K. Qian, D.H. Russell, Molecular Weight Distributions of Asphaltenes and Deasphaltened Oils Studied by Laser Desorption Ionization and Ion Mobility Mass Spectrometry, Anal. Chem. 80 (2008) 8592–8597.

[20] P. Rathsack, B. Wolf, M.M. Kroll, M. Otto, Comparative Study of Graphite-Supported LDI- and ESI-FT-ICR-MS of a Pyrolysis Liquid from a German Brown Coal, Anal. Chem. 87 (2015) 7618–7627.

[21] F. Trejo, J. Ancheyta, T.J. Morgan, A.A. Herod, R. Kandiyoti, Characterization of Asphaltenes from Hydrotreated Products by SEC, LDMS, MALDI, NMR, and XRD, Energy Fuels 21 (2007) 2121–2128.

[22] C.P. Rüger, T. Miersch, T. Schwemer, M. Sklorz, R. Zimmermann, Hyphenation of Thermal Analysis to Ultrahigh-Resolution Mass Spectrometry (Fourier Transform Ion Cyclotron Resonance Mass Spectrometry) Using Atmospheric Pressure Chemical Ionization For Studying Composition and Thermal Degradation of Complex Materials, Anal. Chem. 87 (2015) 6493–6499.

[23] T. Streibel, R. Geißler, M. Saraji-Bozorgzad, M. Sklorz, E. Kaisersberger, T. Denner, R. Zimmermann, Evolved gas analysis (EGA) in TG and DSC with single photon ionisation mass spectrometry (SPI-MS): molecular organic signatures from pyrolysis of soft and hard wood, coal, crude oil and ABS polymer, J Therm Anal Calorim 96 (2009) 795–804.

[24] S. Krieger, O.J. Schmitz, Non-destructive plasticizer screening using a direct inlet probe-atmospheric pressure chemical ionization source and ion trap mass spectrometry, Rapid Commun. Mass Spectrom. 28 (2014) 1862–1870.

[25] C. Flego, C. Zannoni, Direct Insertion Probe−Mass Spectrometry: A Useful Tool for Characterization of Asphaltenes, Energy Fuels 24 (2010) 6041–6053.

[26] E. Kaisersberger, E. Post, Practical aspects for the coupling of gas analytical methods with thermal-analysis instruments, Thermochimica Acta 295 (1997) 73–93.

[27] M. Regert, C. Rolando, Identification of Archaeological Adhesives Using Direct Inlet Electron Ionization Mass Spectrometry, Anal. Chem. 74 (2002) 965–975.

[28] S. Horst, O.J. Schmitz, Quantitative Analysis of Bisphenol A in Recycled Paper with a Novel Direct Inlet Probe-Atmospheric Pressure Photoionization–IonTrap-MS, J. Anal. Test. 1 (2017) 255–263.

[29] A. Ballesteros-Gómez, S.H. Brandsma, J. de Boer, P.E.G. Leonards, Direct probe atmospheric pressure photoionization/atmospheric pressure chemical ionization high-resolution mass spectrometry for fast screening of flame retardants and plasticizers in products and waste, Anal Bioanal Chem 406 (2014) 2503–2512.

[30] C. Flego, C. Carati, L.D. Gaudio, C. Zannoni, Direct Mass Spectrometry of tar sands: A new approach to bitumen identification, Fuel 111 (2013) 357–366.

[31] C. Flego, C. Zannoni, Direct Insertion Probe–Mass Spectrometry (DIP–MS) Maps and Multivariate Analysis in the Characterization of Crude Oils, Energy Fuels 27 (2013) 46–55.

[32] M. Saraji-Bozorgzad, R. Geissler, T. Streibel, F. Mühlberger, M. Sklorz, E. Kaisersberger, T. Denner, R. Zimmermann, Thermogravimetry coupled to single photon ionization quadrupole mass spectrometry: a tool to investigate the chemical signature of thermal decomposition of polymeric materials, Analytical chemistry 80 (2008) 3393–3403.

[33] R.N. Hunter, A. Self, J. Read, The Shell Bitumen handbook, 6th ed., ICE Publishing/Telford, London, 2015.

[34] A.G. Marshall, R.P. Rodgers, Petroleomics: Chemistry of the underworld, Proceedings of the National Academy of Sciences 105 (2008) 18090–18095.

[35] F.W. McLafferty, F. Turec̆ek, Interpretation von Massenspektren, Springer, Berlin, 2013.

[36] M.M. Boduszynski, Composition of heavy petroleums. 2. Molecular characterization.

[37] Klaus H. Altgelt and Mieczyslaw M. Boduszynski, Composition of heavy petroleums. 3. An improved boiling point-molecular weight relation.

[38] F. Handle, M. Harir, J. Füssl, A.N. Koyun, D. Grossegger, N. Hertkorn, L. Eberhardsteiner, B. Hofko, M. Hospodka, R. Blab, P. Schmitt-Kopplin, H. Grothe, Tracking Aging of Bitumen and Its Saturate, Aromatic, Resin, and Asphaltene Fractions Using High-Field Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Energy Fuels 31 (2017) 4771–4779.

[39] Q. Shi, D. Hou, K.H. Chung, C. Xu, S. Zhao, Y. Zhang, Characterization of Heteroatom Compounds in a Crude Oil and Its Saturates, Aromatics, Resins, and Asphaltenes (SARA) and Non-basic Nitrogen Fractions Analyzed by Negative-Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Energy Fuels 24 (2010) 2545–2553.

[40] C.S. Hsu, V.V. Lobodin, R.P. Rodgers, A.M. McKenna, A.G. Marshall, Compositional Boundaries for Fossil Hydrocarbons, Energy Fuels 25 (2011) 2174–2178.

[41] R. R. Sokal, F. Rohlf, Sokal RR, Rohlf FJ. The comparison of dendrograms by objective methods. Taxon 11: 33-40, Taxon 11 (1962).

[42] R.G. Brereton, Hotelling's T squared distribution, its relationship to the F distribution and its use in multivariate space, J. Chemometrics 30 (2016) 18–21.

[43] M.I. Yavor, T.V. Pomozov, S.N. Kirillov, Y.I. Khasin, A.N. Verenchikov, High performance gridless ion mirrors for multi-reflection time-of-flight and electrostatic trap mass analyzers, International Journal of Mass Spectrometry 426 (2018) 1–11.

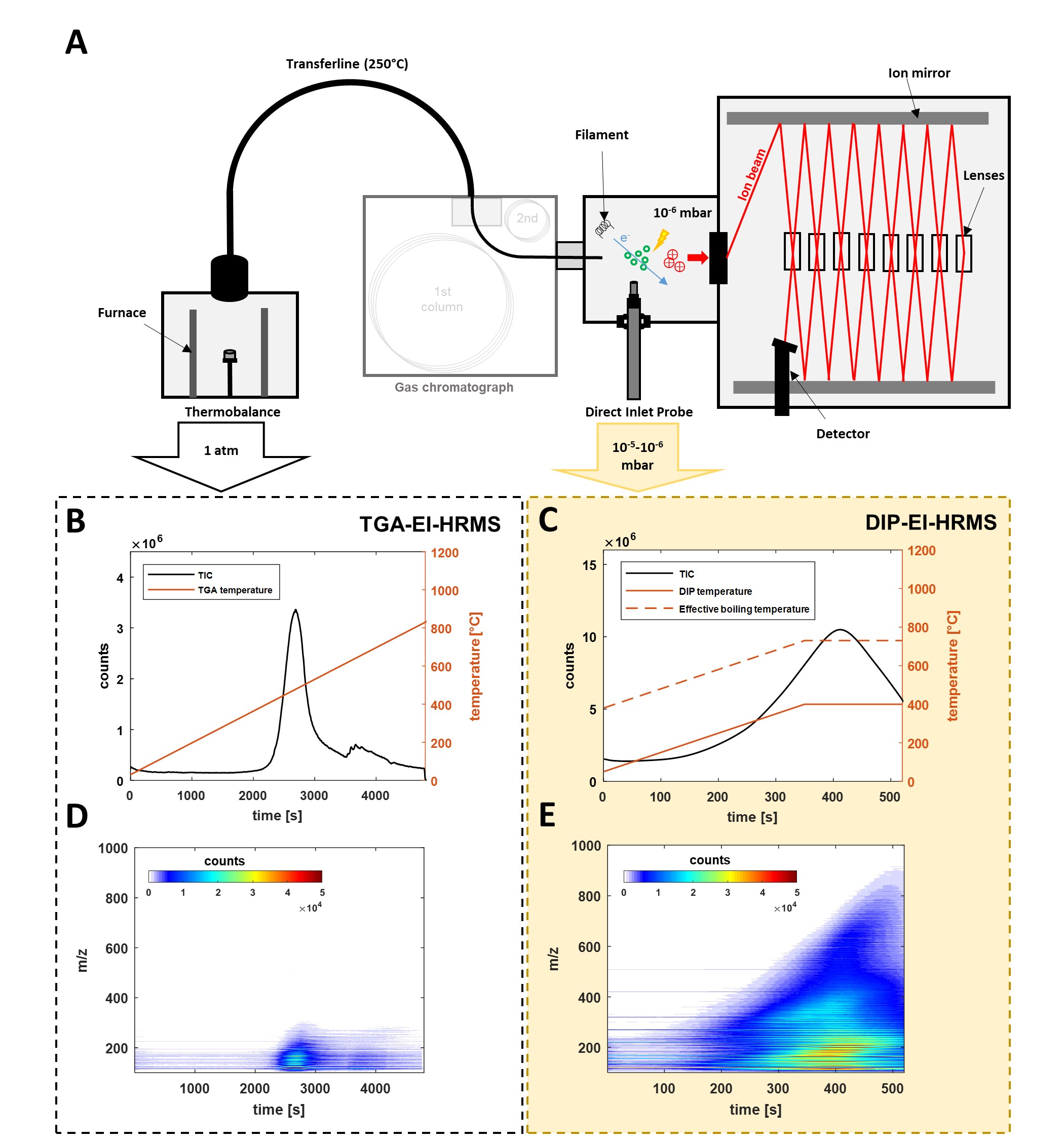


Figure 1. Instrumental setup (A) for TGA-HRMS and DIP-HRMS measurements. Time resolved total ion counts of TGA-EI-HRMS (B) and DIP-EI-HRMS (C) measurements of SRB\_5070 bitumen and the corresponding applied temperature (orange line). For DIP-also the estimated boiling point reduction is shown (dotted line). Contour plots show the time resolved mass spectra of TGA-HRMS (D) and DIP-HRMS (E) measurements.

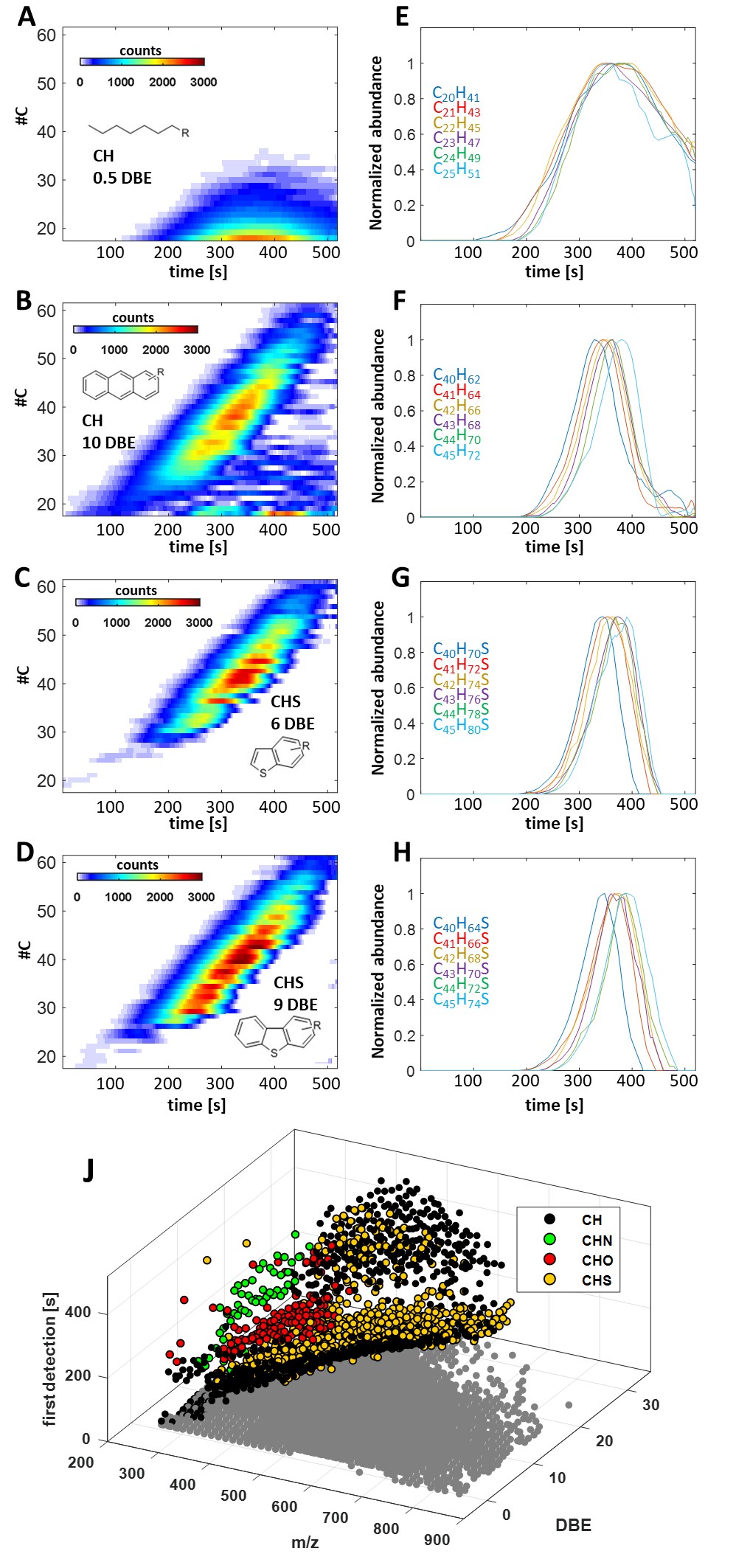


Figure 2. Contour plots for SRB\_5070, for different homologous series: A) CH-class 0.5 DBE, B) CH-class 10 DBE, C) CHS-class 6 DBE, D) CHS-class 9 DBE and a selection of corresponding extracted ion chromatograms (E,F,G,H). The three dimensional plot shows m/z vs first detection time vs DBE of features with integer DBE of the four most abundant compound classes.

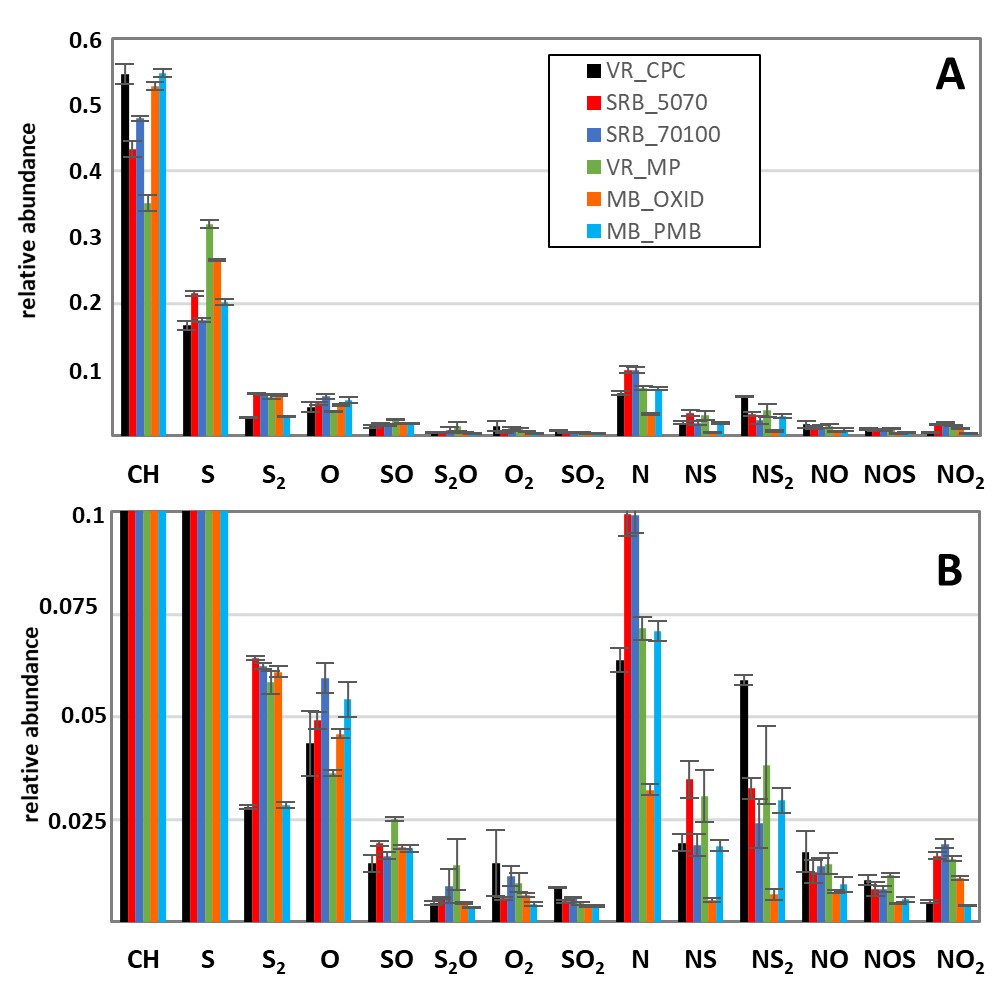


Figure 3. A) Normalized abundances of all assigned compound classes for six investigated samples. Errorbars represent the standard deviation of triplicates. B) Zoom-in for a more detailed view of compound classes with low abundance.

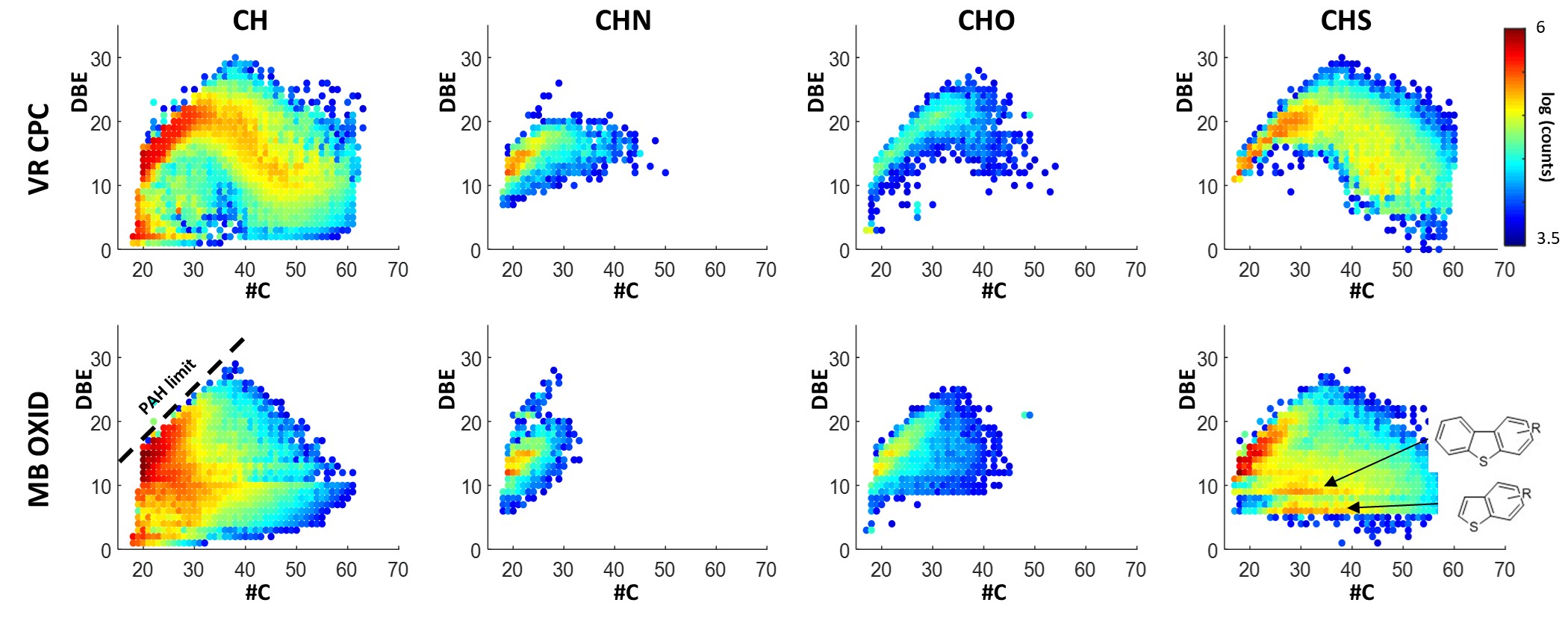


Figure 4. DBE versus carbon number (#C) diagram for the CH-, CHN-, CHO- and CHS-class for the VR\_CPC and MB\_OXID bitumen samples (only features with integer DBE values). Intensity is color-coded, applying a logarithmic scale.

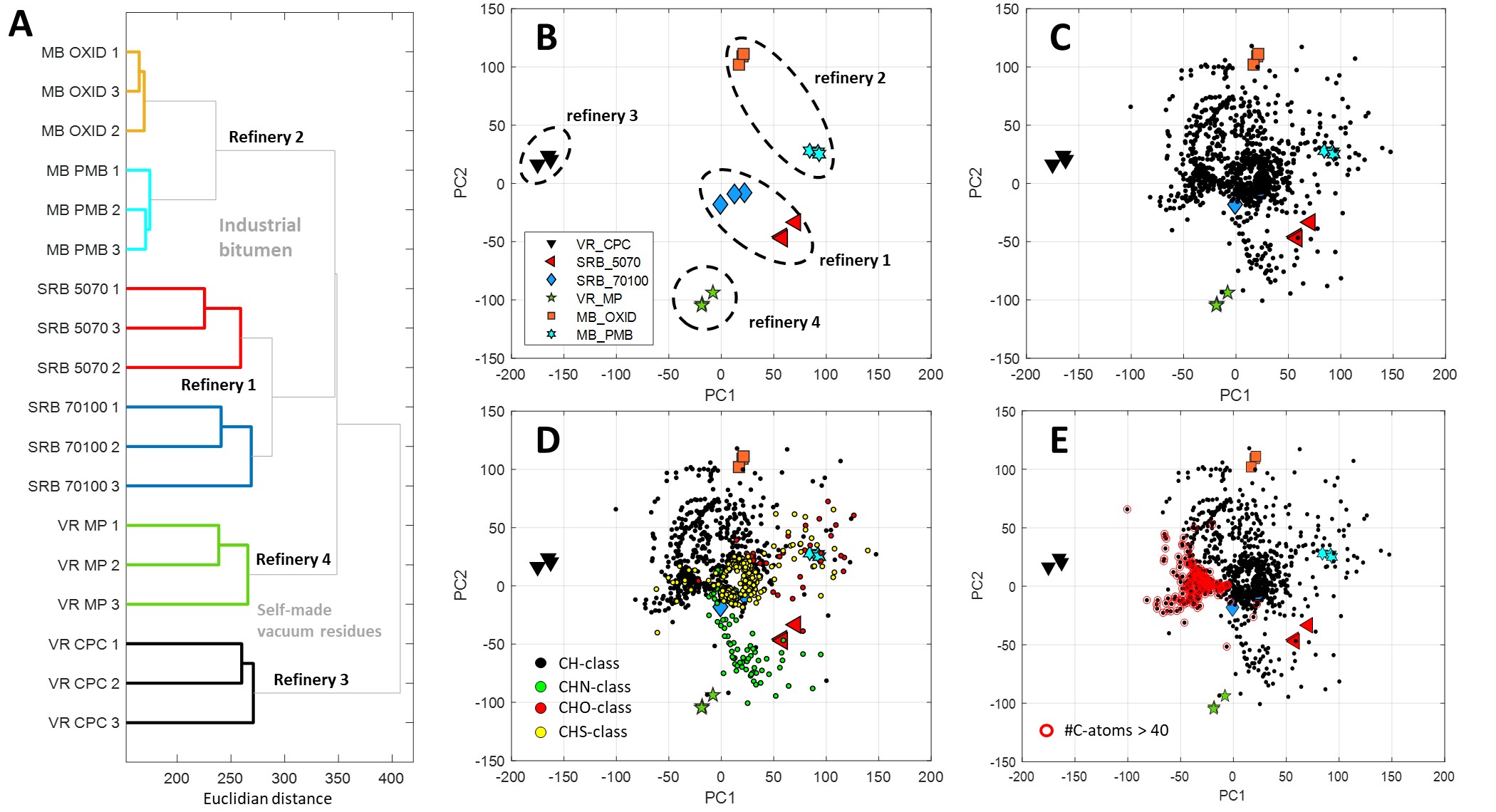


Figure 5. Dendrogramm from HCA (A) showing grouping of samples with spectral similarity. Scores plot (B) and loadings plot (C) from PCA. Colorcoded loadings plot (D) visualizes the elemental composition of features (CH-class: black dots, CHN-class: green dots, CHO-class: red dots, CHS-class: yellow dots). Loadings plot (E) with highlighted features for CH-class, #C>40 (red circles).