Screening for persistent constituents using whole petroleum substance biodegradation testing with deconvolution and non-target analysis of GCxGC-FID/TOFMS data

J. Samuel Arey,[a] Mari Creese,[b] Lisbet Sørensen,[b] Andy M. Booth,[b] Robert K. Nelson,[c] and Delina Lyon[d]

[a] Oleolytics LLC, State College, PA; [b] SINTEF Ocean, Trondheim, Norway; [c] Woods Hole Oceanographic Institution, Woods Hole, MA; [d] Concawe, Brussels, Belgium

UVCB Workshop of The Health and Environmental Sciences Institute (HESI) : Reykjavik, Iceland, Sept 18-19, 2023

Summary

Introduction. Petroleum substances pose a challenge to Persistence assessment, because many of these substances contain thousands of constituents that biodegrade at widely varying rates. Recently, we developed a workflow to screen for persistent constituents using whole petroleum substance biodegradation testing with constituent tracking methods, based on analysis by comprehensive two-dimensional gas chromatography coupled to flame ionization detector (GCxGC-FID) and to high-resolution time-of-flight mass spectrometry (GCxGC-HR-TOFMS).[1,2,3] However, chromatographic coelutions challenged our efforts to quantify, track, and identify constituents in these highly complex UVCBs.

Approach. In this study, we re-design the workflow to include deconvolution and non-target analysis approaches that improve the detection, quantification, and identification of slowly-biodegrading constituents. The screening method detects, quantifies, and identifies constituents that exceed designated criteria for: (i) the constituent concentration in the original product; and (ii) the non-degraded mass fraction of the constituent after 64 days of biodegradation. We test the screening method on a diesel fuel previously analyzed by GCxGC-FID/TOFMS.[3]

Results. For this diesel fuel, we found 50 constituents that met the screening criteria. Among these, 43 constituents (86%) are interpreted as C_{15}-C_{23} two-ring naphthenes, of which 30 likely contain a quaternary carbon. Five constituents (10%) are C_{21}-C_{23} acyclic isoprenoids, and two constituents (4%) were not successfully interpreted. Example results are shown for peak #4.

Conclusions. The re-designed workflow improves on the detection, quantification, and identification of peaks, compared to the previous method. This is illustrated by the finding that most of the screened constituents are identified as naphthenes containing quaternary carbons, a structural feature associated with recalcitrance to biodegradation. This information can be used to prioritize further testing of constituents expected to biodegrade slowly. The data analysis workflow can be extended to other endpoints requiring time-lapsed information, such as bioaccumulation potential.

Overview of Work-Flow

Analyze a diesel fuel and its biodegraded residue by GCxGC-FID and GCxGC-HR-TOFMS

Aerobically biodegrade a diesel fuel in seawater during 64 d (laboratory)[1,2] ⇨ Analyze the unweathered diesel fuel and biodegraded residue by GCxGC-FID[1,2] and by GCxGC-HR-TOFMS[3]

⇩

Quantify concentrations and extent of biodegradation of separable C_{10}–C_{30} constituents using GCxGC-FID data

Detect and quantify each GCxGC-FID peak using a recent deconvolution algorithm[4] (~8000 constituents) ⇨ Align[5] and match[6] peaks in the biodegraded residue with the fresh diesel sample (~3870 constituents) ⇨ Quantify the fraction of each peak mass that biodegraded (~3870 constituents)

⇩

Select peaks meeting designated criteria for: (i) minimum concentration of constituent in the original product; and (ii) minimum extent of biodegradation of constituent after for 64 days

50 constituents meet screening criteria for persistence potential

⇩

Identify potentially-persistent peaks by non-target analysis of GCxGC-HR-TOFMS data

Find the GCxGC-HR-TOFMS spectrum of each potentially-persistent peak using algorithms for alignment[5] and matching[6] ⇨ Deconvolute[4] GCxGC-HR-TOFMS spectra of potentially-persistent peaks and format for data analysis ⇩

Interpret chemical family and carbon number of each potentially-persistent peak by non-target analysis techniques:

- Narrow plausible chemical family and carbon number by diagnostic ions,[7,8,9] neutral losses,[9] and GCxGC elution model[10]
- Determine elemental compositions of major ions by calculating expected monoisotopic masses and isotope ratios
- Search for similar mass spectra in major libraries[11]
- Identify chemical structure motifs using Substructure Analysis[12]

Detection and Quantification of Peak #4 in the Diesel Fuel and its Biodegraded Residue

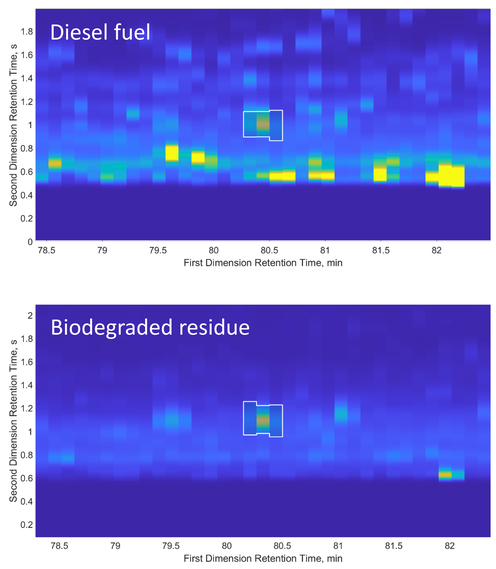

Figure 2. Heat map of the GCxGC-FID chromatogram region containing peak #4

Color: FID signal intensity, proportional to hydrocarbon constituent concentration in the diesel fuel.

White outline: approximate boundary of peak #4.

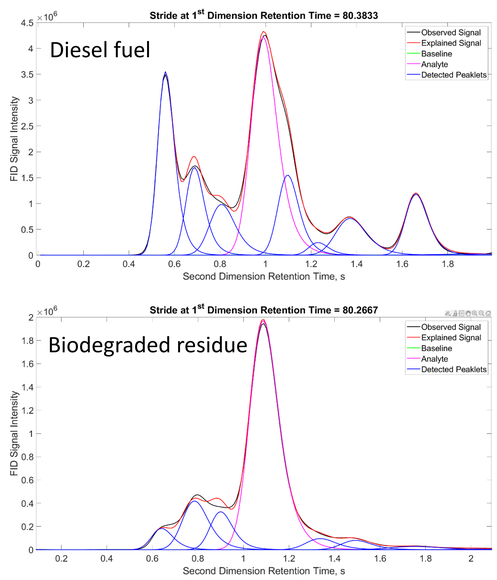

Figure 3. Deconvolution of the 2nd dimension slice of the GCxGC-FID chromatogram containing the dominant signal component of peak #4

Black line: Observed FID signal.

Pink line: FID signal attributed to peak #4.

Blue lines: FID signal attributed to other constituents.

Non-Target Analysis Results of Peak #4

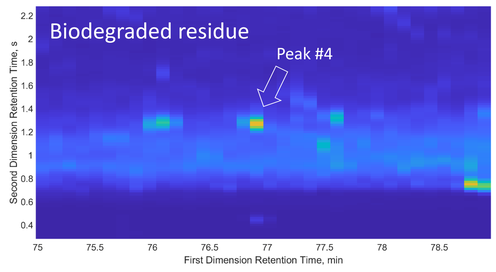

Figure 4. Heat map of the GCxGC-HR-TOFMS chromatogram region containing peak #4

Color: TOFMS signal intensity.

Interpretation: peak #4 is a $C_{15}H_{28}$ naphthene with two rings which likely contains quaternary carbon(s). This is likely a drimane, consistent with the characteristic ions 123, 137, and 193, and the presence of quaternary carbon (also see Substructure Analysis).

Molecular ion mass: 208.215

Theor. monoisotopic mass: 208.219 ($C_{15}H_{28}^{+}$)

Substructure Analysis indicates a high likelihood that peak #4 is a naphthenic and contains quaternary carbon. Substructure Analysis evaluates the chemical structure features that appear in the top 100 best-match spectra within the combined three libraries. The analysis accounts for both the frequency of substructure appearance and similarity of those spectra with respect to peak #4, in the top 100 list.

GCxGC elution model: C_{14}-C_{15}, either naphthenic or monoaromatic

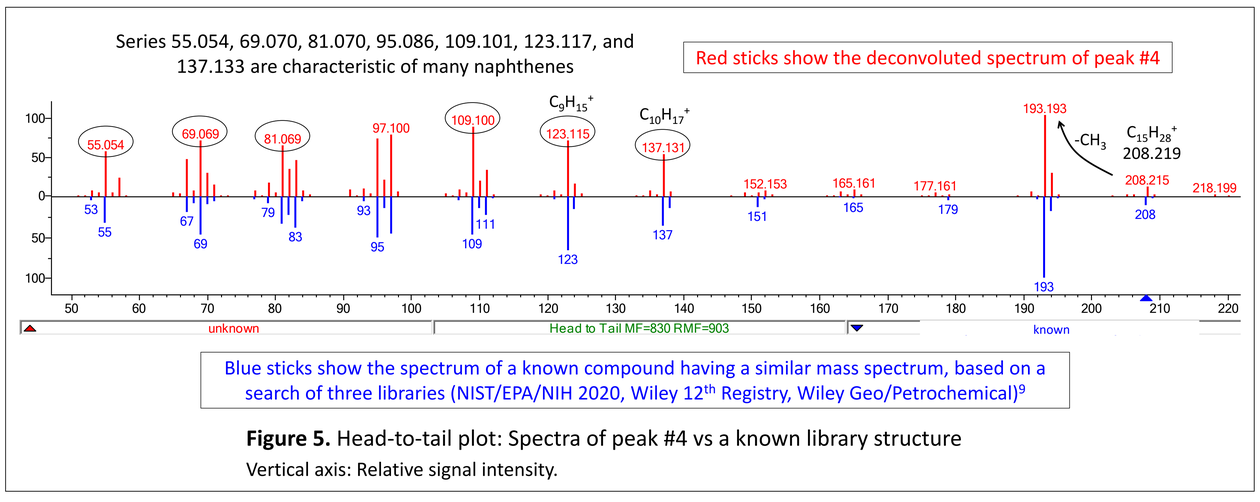

Figure 5. Head-to-tail plot: Spectra of peak #4 vs a known library structure

Vertical axis: Relative signal intensity.

References

[1] Arey, Gros, Redman, Bleich, Letinski, Connelly, and Nelson, "Quantifying biodegradation rates of thousands of petroleum constituents in seawater by GCxGC-FID" *SETAC North America 41st annual meeting (virtual)*, Nov 15-19, 2020.

[2] Booth, Sørenson, Brakstad, Ribicic, Creese, Arey, Lyon, Redman, Martin-Aparicio, Camenzuli, Wang, and Gros, "Comprehensive two-dimensional gas chromatography with peak tracking for screening of constituent biodegradation in petroleum UVCB substances" *ES&T* ASAP Aug 17, 2023.

[3] Lyon, Creese, Booth, Gros, Arey, Saunders, Saunders, Sørenson, Sourrisseau, et al., "Screening for persistence using whole petroleum substance UVCB biodegradation testing with constituent tracking", *SETAC Europe 33rd annual meeting,* Dublin, Apr 30 – May 4, 2023.

[4] Arey, GCxGC Peak Measurement, *patent pending*.

[5] Gros, Nabi, Dimitriou-Christidis, Rutler, and Arey, "Robust algorithm for aligning GCxGC chromatograms", *A. Chem.* 84, 9033-9040 (2012).

[6] Wardlaw, Arey, Reddy, Nelson, Ventura, and Valentine, "Disentangling oil weathering at a marine seep using GCxGC : Broad metabolic specificity accompanies subsurface petroleum biodegradation" *ES&T* 42, 7166-7173 (2008).

[7] Peters, Walters, and Moldowan, "The Biomarker Guide, Vol. 1", 2nd ed., Cambridge University Press, 2005.

[8] Nelson, Forsythe, Eiserbeck, Scarlett, Grice, Mullins, and Reddy, "GCxGC analysis of novel 2α-methyl biomarker compounds from a large Middle East oilfield" *Energy & Fuels* 36, 8853-8865 (2022).

[9] McLafferty and Turecek, "Interpretation of Mass Spectra, 4th ed.", University Science Books, 1993.

[10] Arey, Aparicio, Vaiopoulou, Forbes, and Lyon, "Modeling the GCxGC elution patterns of a hydrocarbon structure library to innovate environmental risk assessments of petroleum substances", *ES&T* 56, 17913-17923 (2022).

[11] NIST/EPA/NIH 2020 Library; Wiley Registry of MS Data, 12th Edition; and Wiley MS Library of Geochemicals, Petrochemicals, and Biomarkers

[12] NIST MS Search Program v. 2.4, National Institute of Standards and Technology, Standard Data Reference Program, Gaithersburg, MD