Supporting Information for:

Cracked and shucked: GC-APCI-IMS-HRMS facilitates identification of unknown halogenated organic chemicals in French marine bivalves

Authors: Randolph R. Singh[†], Yann Aminot[†], Karine Moisan[†], Hugues Preud'homme[‡], Catherine Munschy[†]

[†]Chemical Contamination of Marine Ecosystems (CCEM) Unit, Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Rue de l'Ile d'Yeu, BP 21105, 44311, Cedex 3, Nantes, France

[‡]IPREM-UMR5254, E2S UPPA, CNRS, Technopôle Helioparc, 2 Avenue P. Angot 64053 Pau Cedex 9, France

Corresponding author: Randolph Singh (randolph.singh@ifremer.fr)

Table of Contents

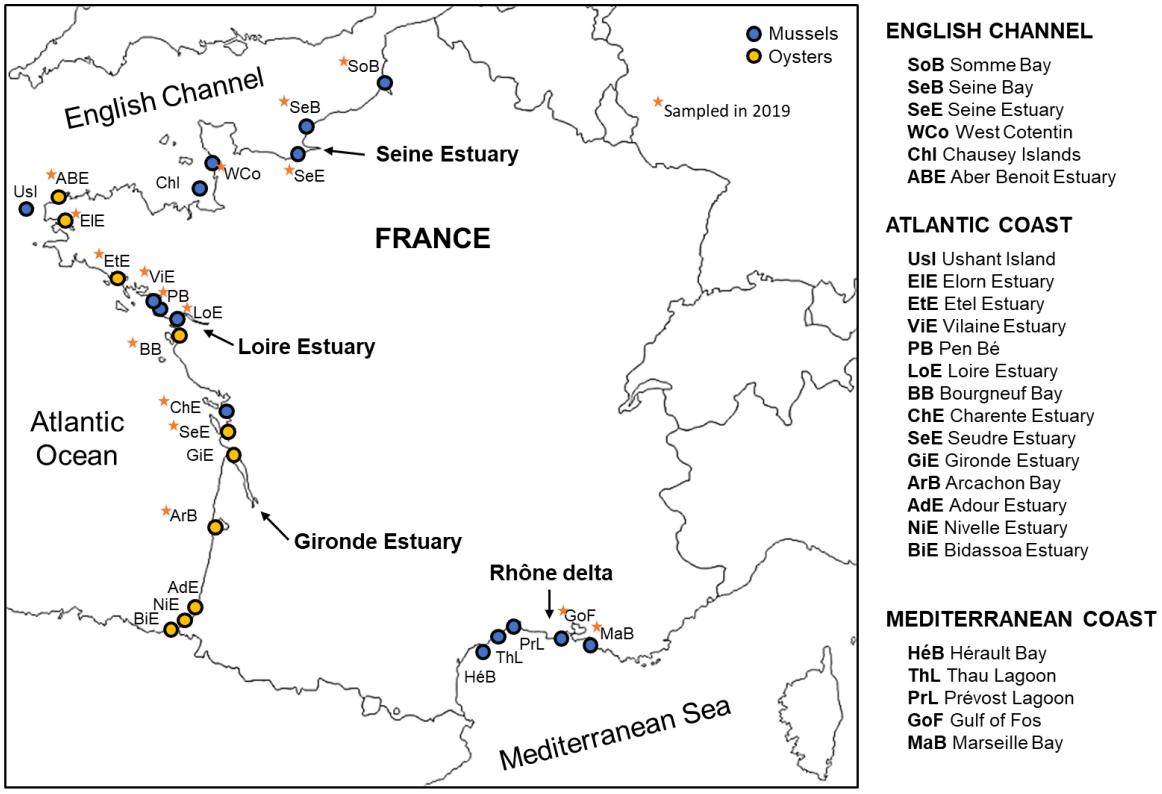
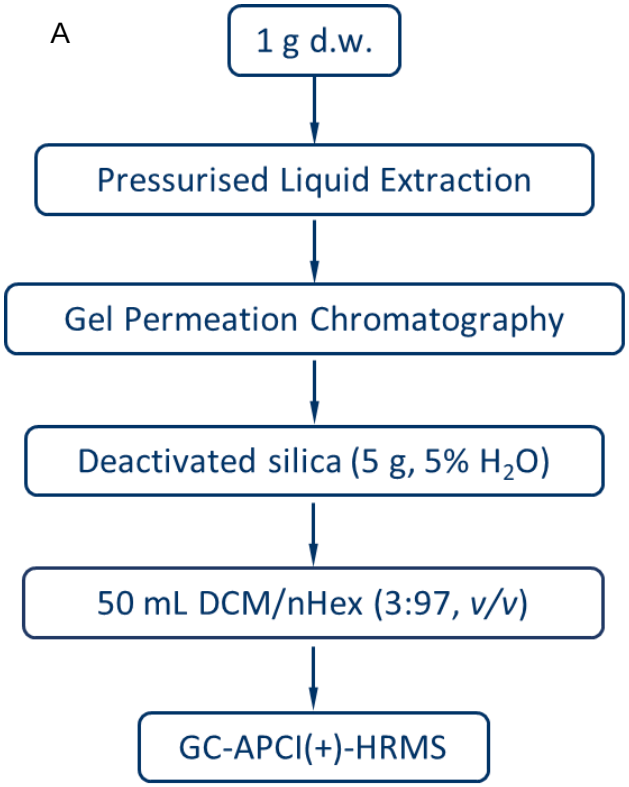


Figure S1. Sampling sites for the different mussel (blue spheres) and oyster (yellow spheres) samples collected for this study.



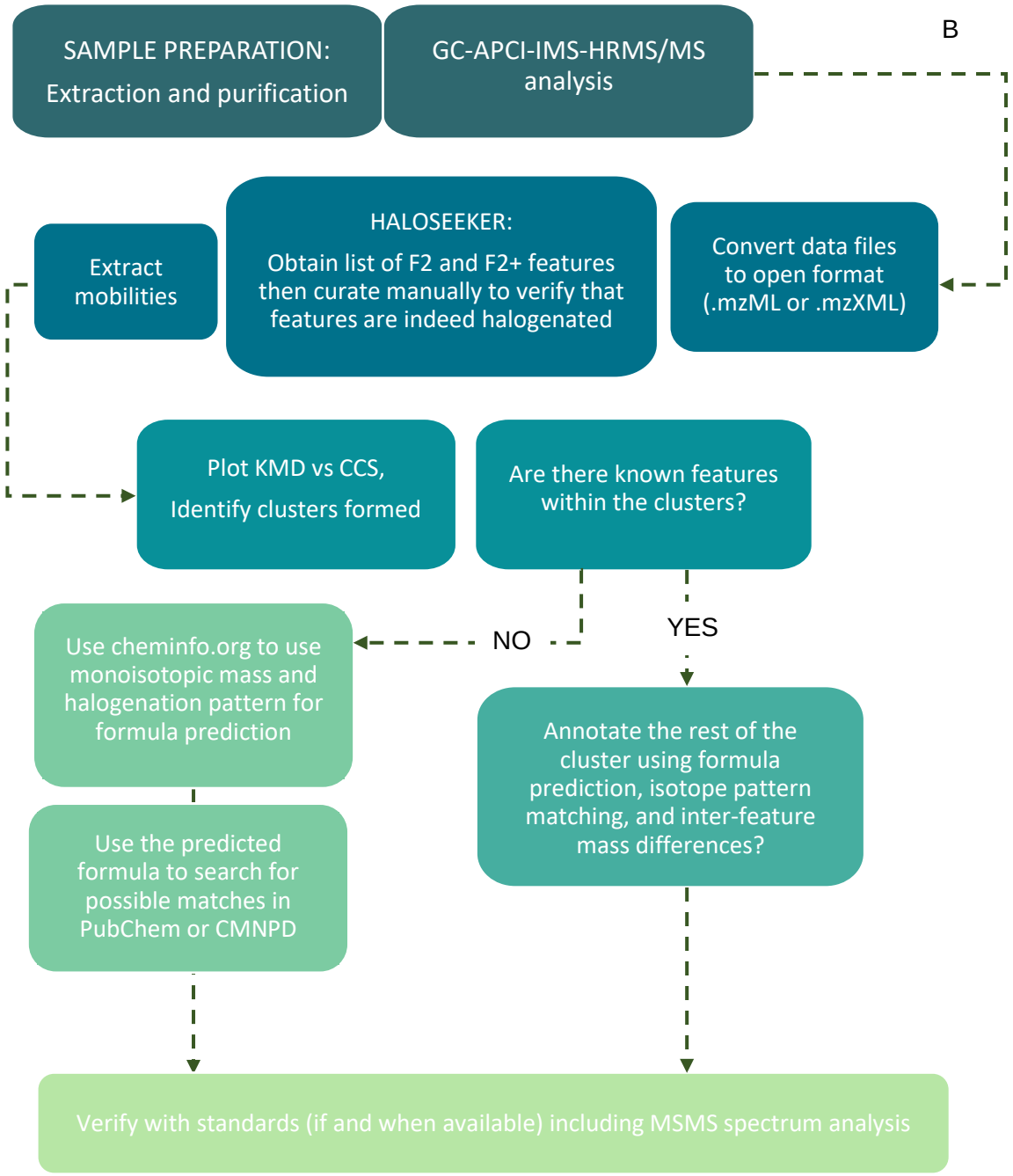


Figure S2. A. Sample preparation protocol adapted from Sun et al (2012).
B. Data processing workflow.

List of isotopically labelled standards used and quality assurance.

Several isotopically labelled standards were spiked as surrogates including: 17 ^{13}C-labelled organochlorine pesticides (^{13}C-Hexachlorocyclohexane alpha, ^{13}C-Hexachlorobenzene, ^{13}C-Hexachlorocyclohexane gamma, ^{13}C-Hexachlorocyclohexane beta, ^{13}C-Hexachlorocyclohexane delta, ^{13}C-Aldrine, ^{13}C-Isodrine, ^{13}C-Endosulfan alpha, ^{13}C-pp'DDE, ^{13}C-Dieldrine, ^{13}C-op'DDD, ^{13}C-Endrine, ^{13}C-op'DDT, ^{13}C-Endosulfan beta, ^{13}C-pp'DDD, ^{13}C-pp'DDT, ^{13}C-Endosulfan sulfate) , 18 fully ^{13}C-labelled PCBs (^{13}C CB28, ^{13}C CB52, ^{13}C CB101, ^{13}C CB138, ^{13}C CB153, ^{13}C CB180, ^{13}C CB77, ^{13}C CB81, ^{13}C CB105, ^{13}C CB114, ^{13}C CB118, ^{13}C CB123, ^{13}C CB126, ^{13}C CB156, ^{13}C CB157, ^{13}C CB167, ^{13}C CB169, ^{13}C CB189, 14 ^{13}C-labelled PBDEs (^{13}C BDE3, ^{13}C BDE15, ^{13}C BDE28, ^{13}C BDE47, ^{13}C BDE99, ^{13}C BDE153, ^{13}C BDE154, ^{13}C BDE183, ^{13}C BDE197, ^{13}C BDE207, ^{13}C BDE209, ^{13}C BDE47, ^{13}C BDE99, ^{13}C BDE153), one ^{13}C-labelled MeO-PBDE (^{13}C-MeOBDE-47), one ^{13}C-labelled dichlorocarbazole, ^{13}C-labelled triclosan and ^{13}C-labelled methyltriclosan. ^{13}C-labelled tetrachlorocarbazole was used internal standard, whose peak area was used to normalize the peak areas of the analytes to correct for inter-injection variability. The median variation in the internal standard signal was 20% compared to the average signal of the internal standard in 6 injections of a clean standard mixture. Median recoveries for the different isotopically labelled internal standards are as follows: organochlorine pesticides (71.5 %), PCBs (95.4 %), PBDEs and PBDE-related chemicals (94.6%). Mass accuracy, assessed using internal standards, was consistently below 5 ppm throughout the whole analysis.

Semi-Quantification

Identified peaks were manually integrated using El-Maven to obtain individual peak areas. Peak areas were normalized using the area of the internal standard 1,3,6,8-tetrachloro-9H-[$^{13}C_{12}$]carbazole ($^{13}C_{12}H_5Cl_4N$, monoisotopic mass: 314.9573, exact mass:316.9544). PCBs and PBDEs were quantified using their corresponding isotopically labelled standard or

of their congener. HNPs, especially newly reported HNP were semi-quantified by comparing the normalized peak area against the peak area of known members of their compound class.

Data Processing

Raw Bruker files (.d format) were converted (Figure S3) to the standard open format (.mzML) using MSConvert (Proteowizard (version: 3.0.21045-7732b6429). Peak picking, pairing, and alignment was accomplished using HaloSeeker (Figure S4) using the following deconvolution parameters: m/z tolerance (ppm) 5, s/n 10, Peakwidth (s) 6, 15, Prefilter 3, 10000, m/z center function wMean, Baseline check no, Integration by CWT no, Noise 0, m/z difference 0.001, tR tolerance (s) 1, m/z tolerance (mDa) 0.1, Referent sample pos 4082_19-4043_13F1, tR tolerance 1, m/z tolerance (mDa) between samples 1, Distance function cor_opt, Response 1, Gap init 0.3, Gap extend 2.4, Factor diagonal 2, Initiating penalty 0, Local alignment yes, Fragment/adduct, Percentage of the width at FWHM 60, Threshold for EIC correlation (R2) 0.75, Method for grouping peaks hcs, Mutliplier of the standard deviation 6, P-value threshold 0.05. (Describe picking F2+ features).

Unknown feature annotation was performed using a suite of open tools described below; to facilitate description of the data analysis method employed the m/z = 467.7767 will be used as an example (Figure S5). The experimental monoisotopic mass of an unknown feature obtained from HaloSeeker, was used as an input within cheminfo.org for formula prediction considering both electron loss and proton adduct formation as the governing ionization mechanisms in APCI+. Annotation of potentially chlorinated/ brominated features was accomplished using a combination of formula prediction and isotopologue ratio matching. Formula were predicted within 10 ppm of the experimental monoisotopic mass (primary input data, Figure S5) using the following elements and range: C0-30H0-50N0-6O0-5F0-50Cl0-14Br0-10I0-4. The predicted formulas that match the monoisotopic mass, within an error margin of 10 ppm, were prioritized using the degree of halogenation (eg number of chlorines or bromines) that can be inferred from the relative ratio of the isotopologue pattern and relative

ion intensities. Theoretical centroid isotopologue distribution and ratios were calculated using enviPat (https://www.envipat.eawag.ch/index.php) at a mass resolving power of 50,000. Analysis of the isotopologue pattern for m/z = 467.7767 suggests that the unknown feature has 3 bromines and 1 chlorine. A comparison of the experimentally measured isotopologue distribution and their respective percent abundance to the exact mass to the theoretical isotopologue distribution is shown in Figure S6.

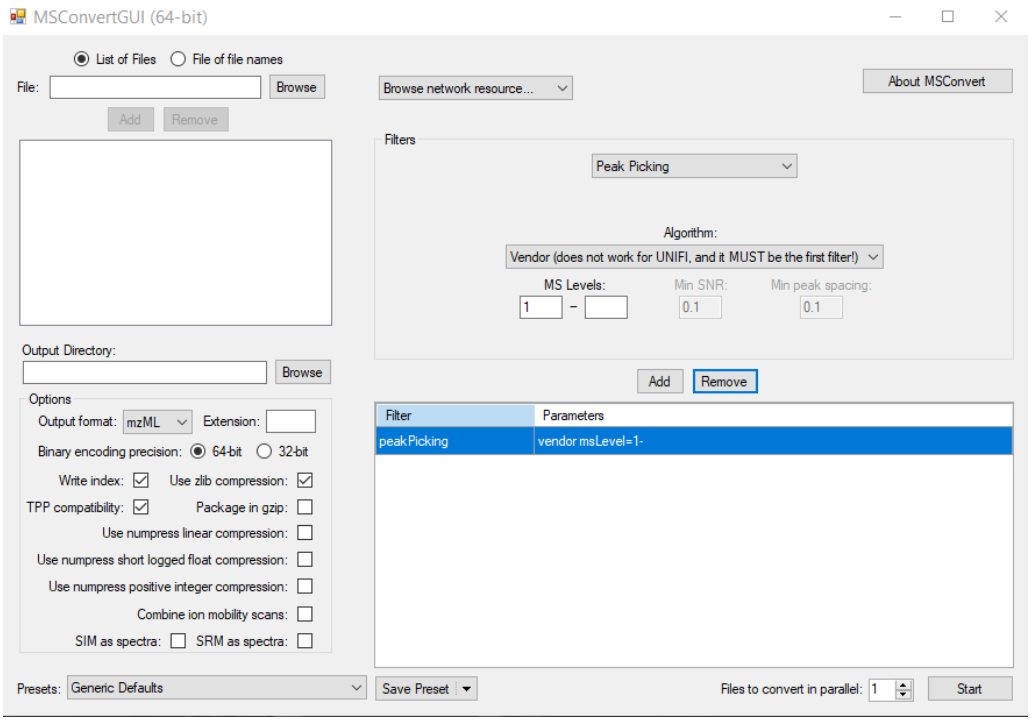


Figure S3. MS Convert settings for converting ddMS2 Bruker files to .mzml.

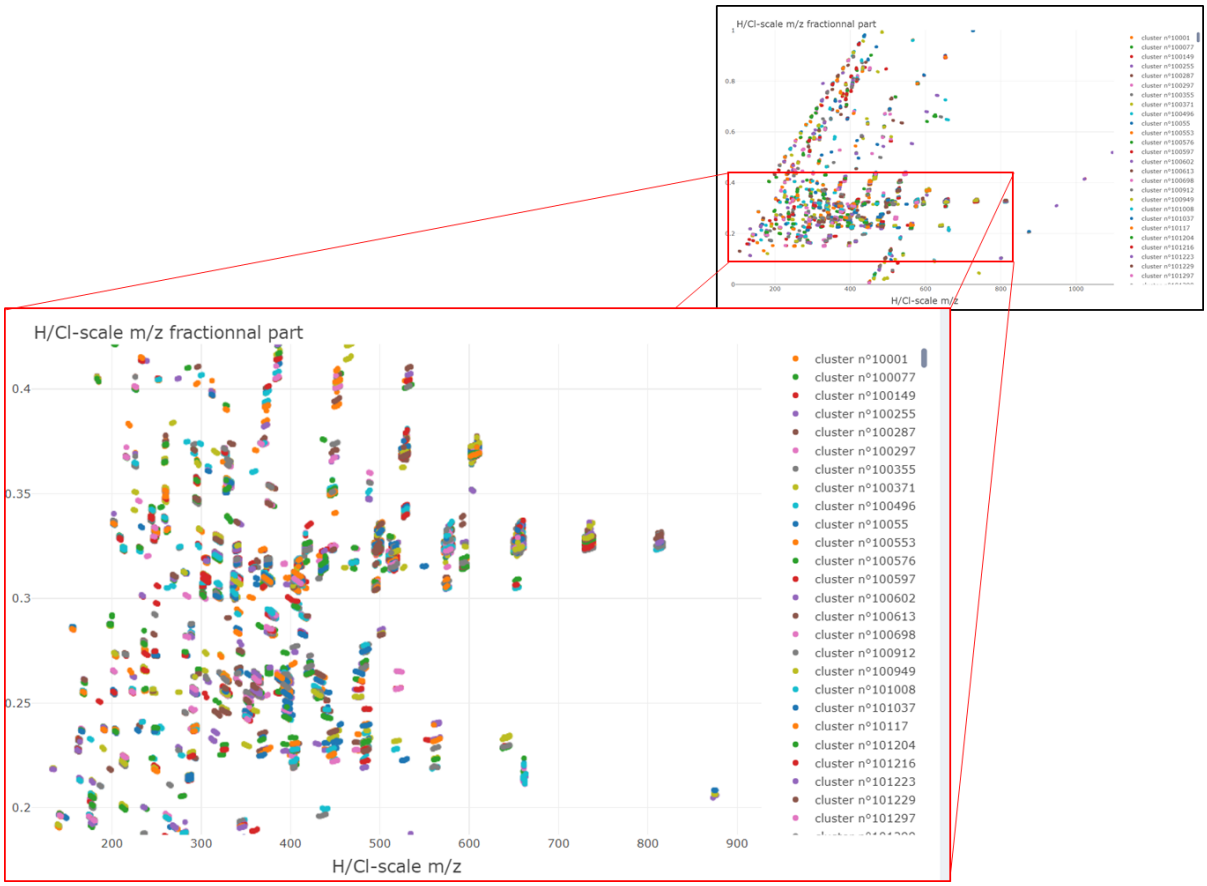


Figure S4. Screenshot of polyhalogenated features picked by HaloSeeker (upper right) from the bivalve data including a zoomed in version (bottom left) highlighting region of interest.

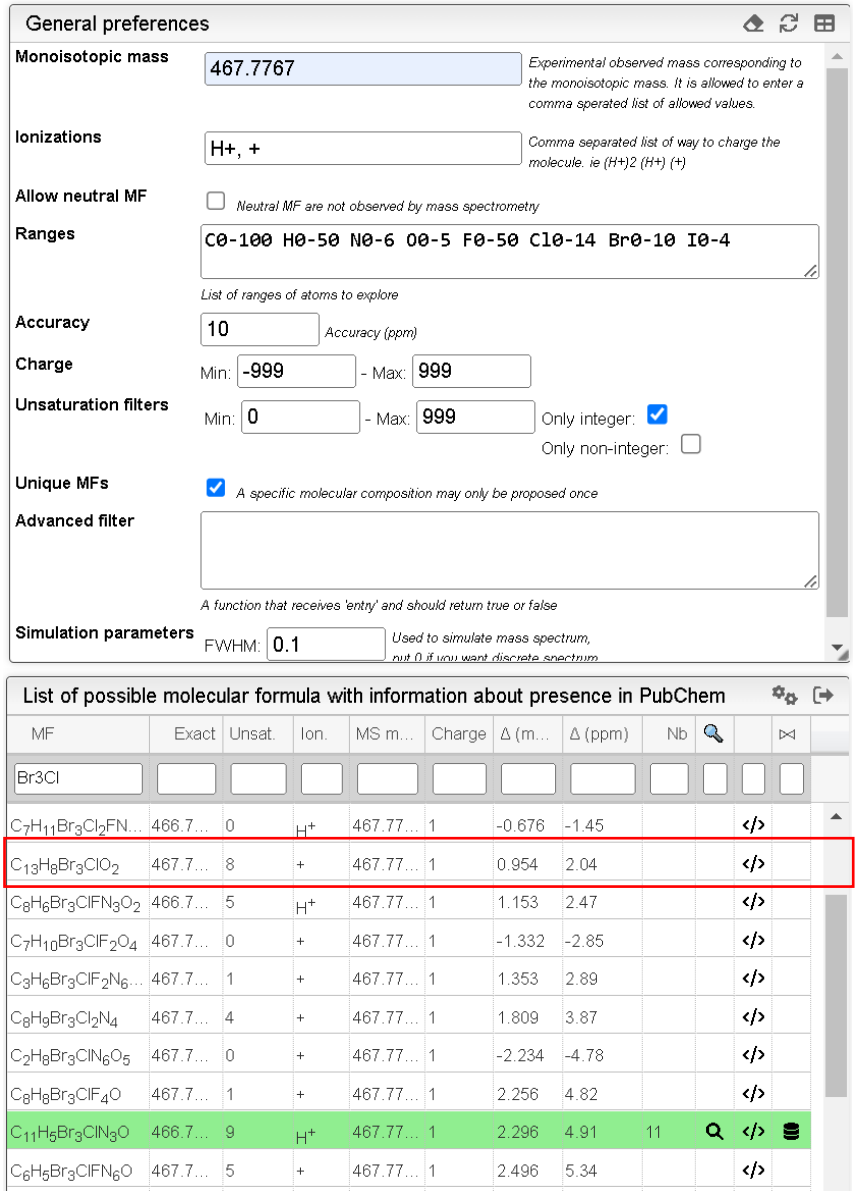


Figure S5. From the list of potential formulas that satisfy the specified criteria for prediction, the degree of halogenation (in this example the unknown was identified to follow a tribrominated monochlorinated pattern) was used to narrow down the options (orange box). Among the potential formulas, one had a database match within PubChem (within red box), structures of which are shown in the lower part of the figure (green box).

Among the formulas that match the input mass and the Br_3Cl isotopologue profile, one entry has a database match (Figure S5, highlighted in green). The online tool provides a link showing a list of chemicals and their respective structures (Figure S7) within PubChem (largest chemical database, can be openly accessed!) that match the unknown Br_3Cl compound. Unfortunately, manual fragmentation of the putative structures was not able to match the experimental MSMS. While not a success story, this example highlights the need for caution when it comes to accepting database matches based solely on MS1 information. Ultimately, the annotation of this unknown feature is facilitated through the use of its Kendrick Mass Defect and its CCS.

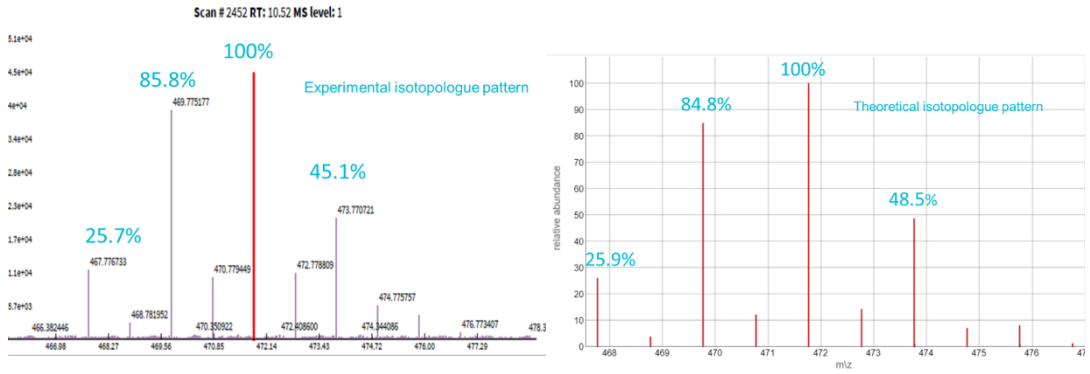


Figure S6. Comparison of the experimental (top) vs theoretical (bottom) isotopologue ratios of the predicted formula showing close agreement and thus potential accuracy of the predicted formula. The relative abundance of the isotopologues were calculated relative to the intensoid mass.

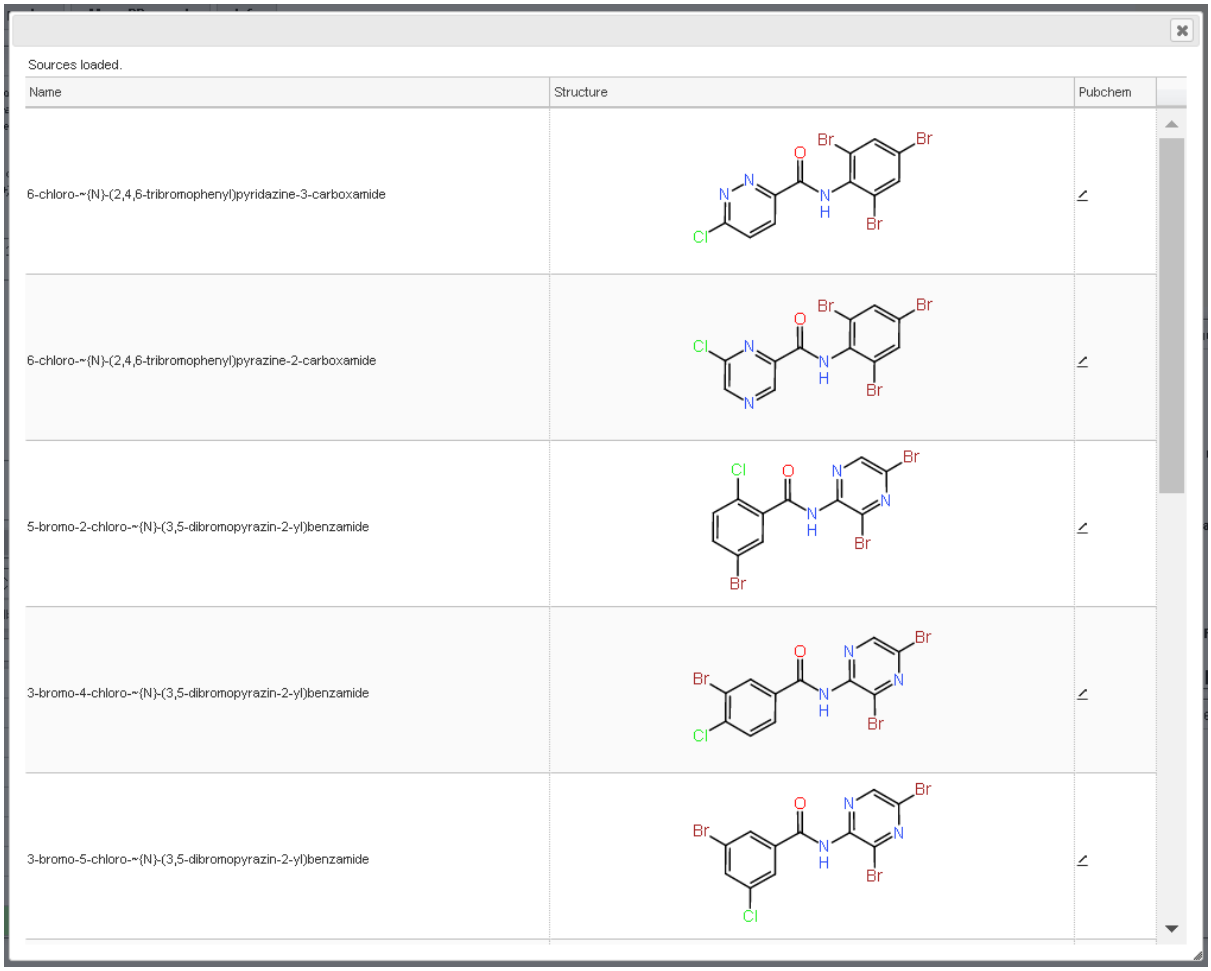


Figure S7. List of compounds, and their respective structures, within PubChem that match the *m/z* and halogenation profile of the unknown feature obtained using the tools provided by cheminfo.org.

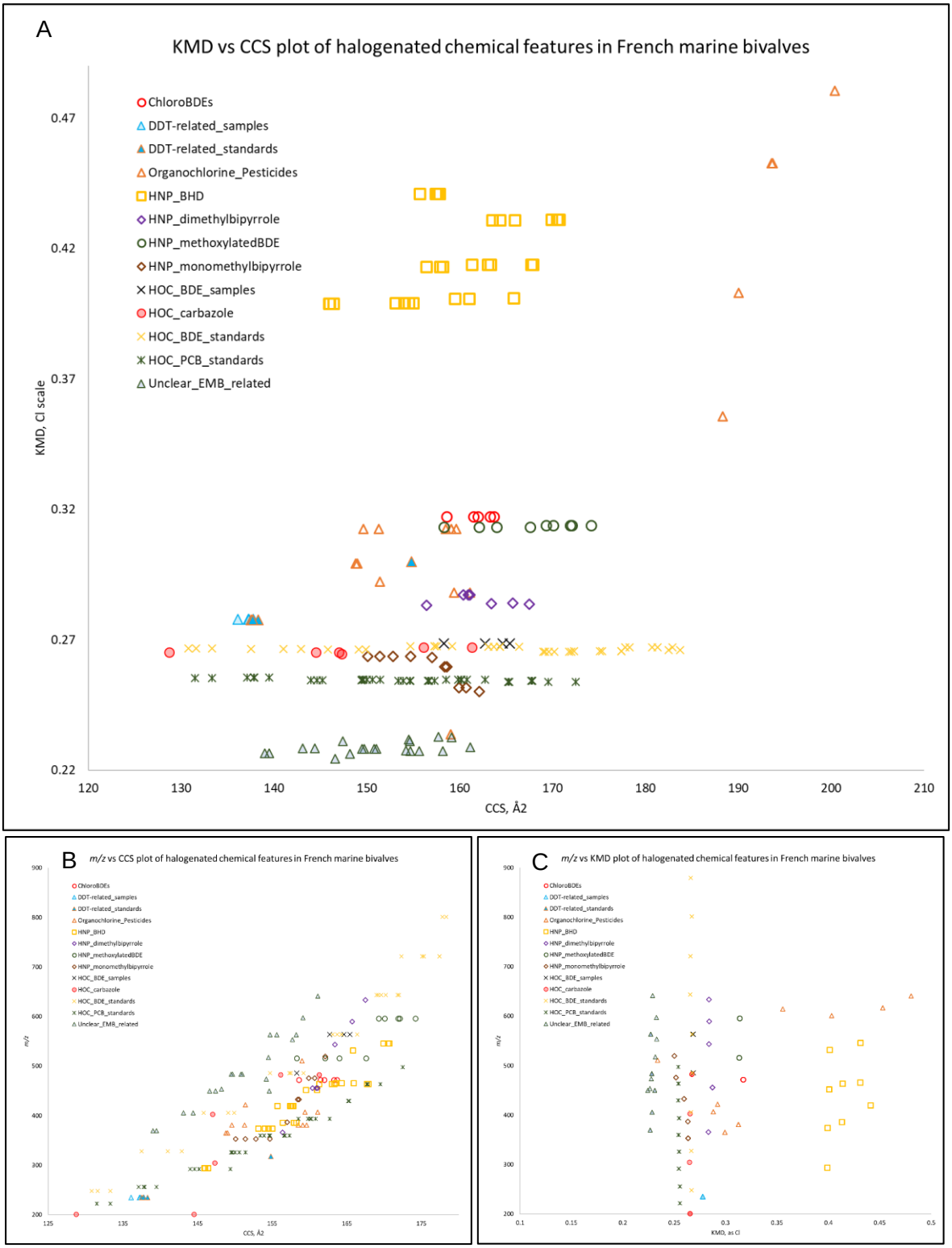


Figure S8. KMD vs CCS plot (A) allows better visualization and differentiation of structurally-related features compared to *m/z* vs CCS (B) and *m/z* vs KMD (C).

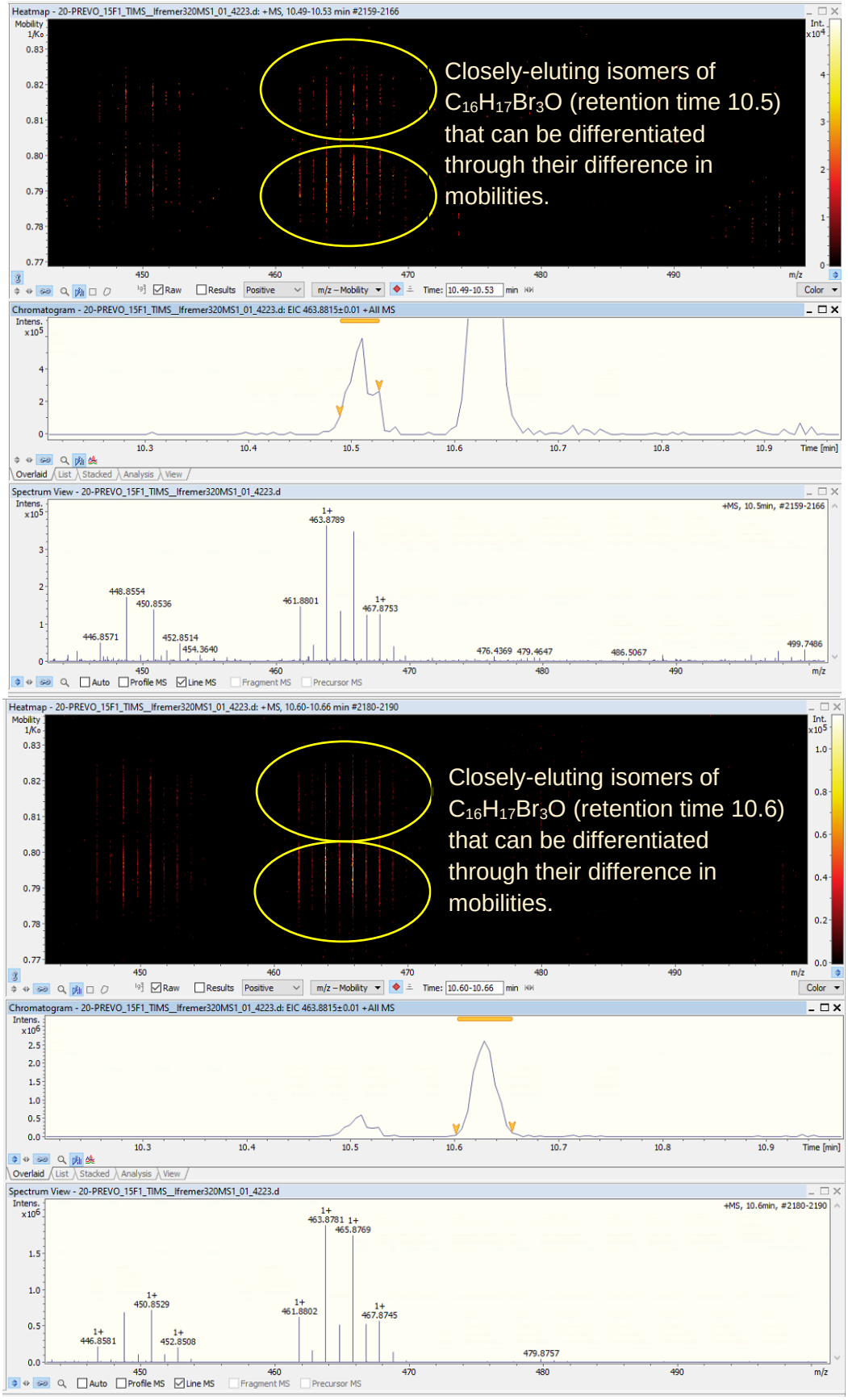


Figure S9. Four different isomers of $C_{16}H_{17}Br_3O$ were identified by virtue of their different mobilities (highlighted by the yellow ovals).

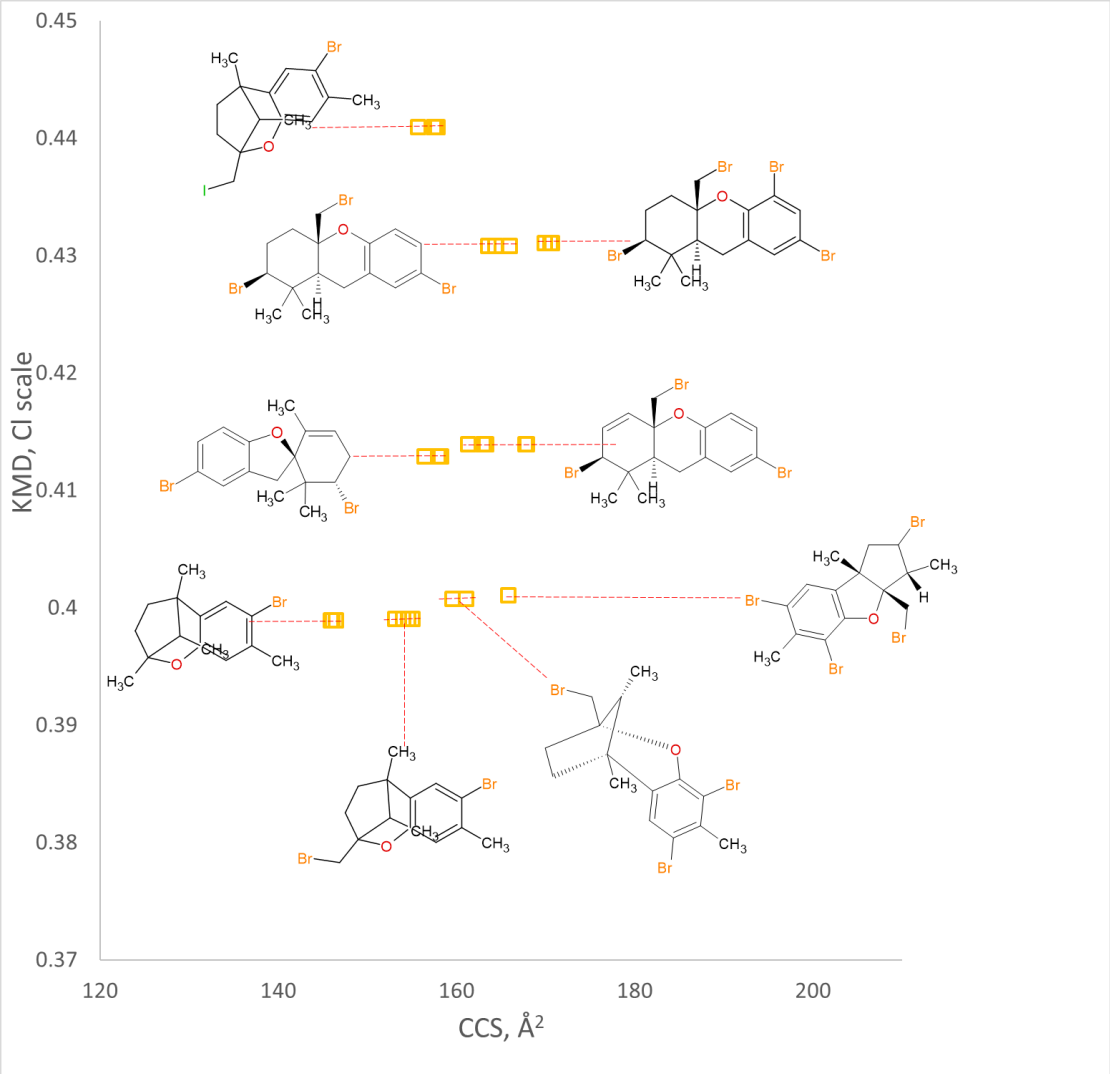


Figure S10. Zoomed in KMD vs CCS plot showing the BHD-related features and their representative structures.

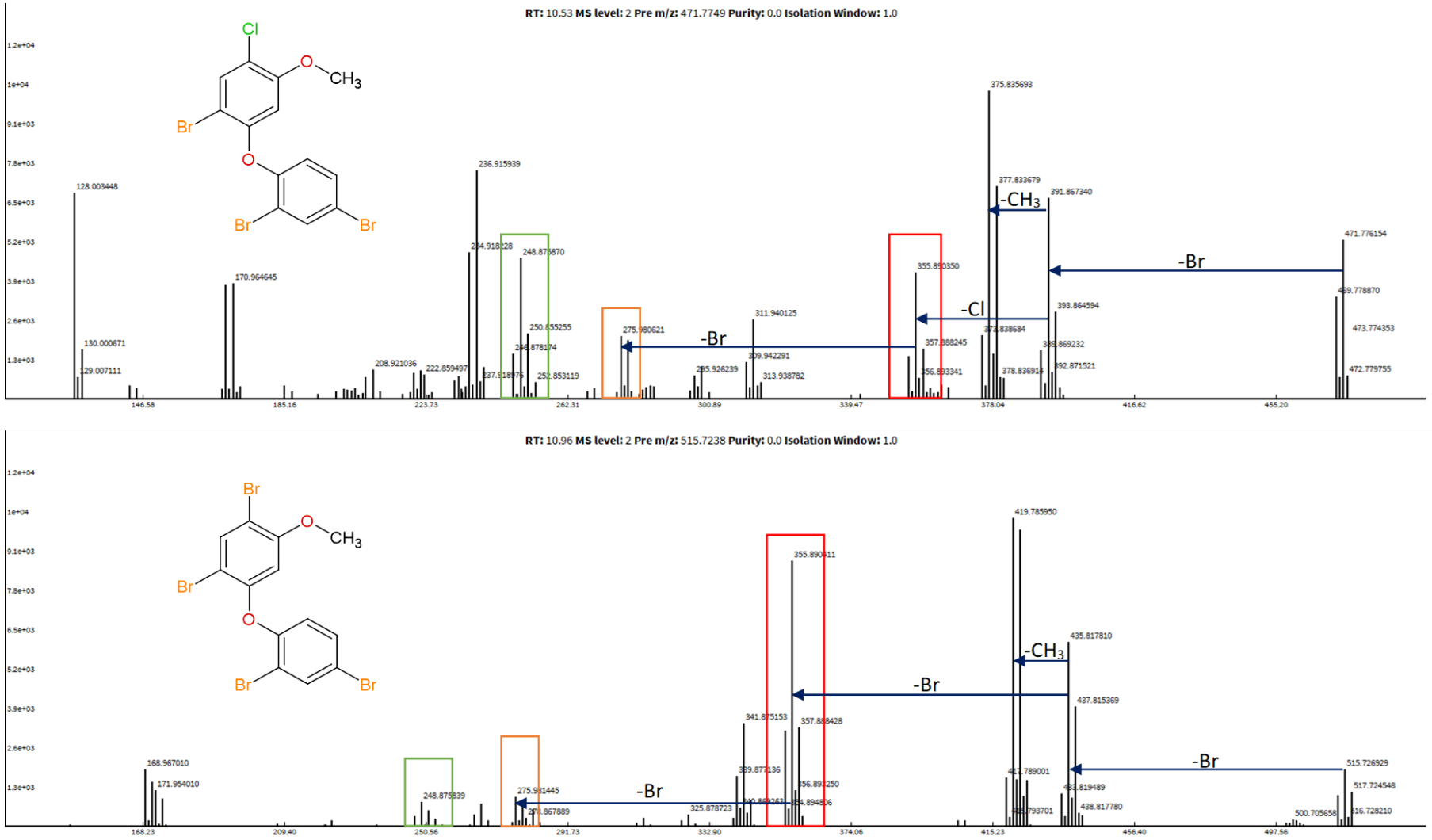
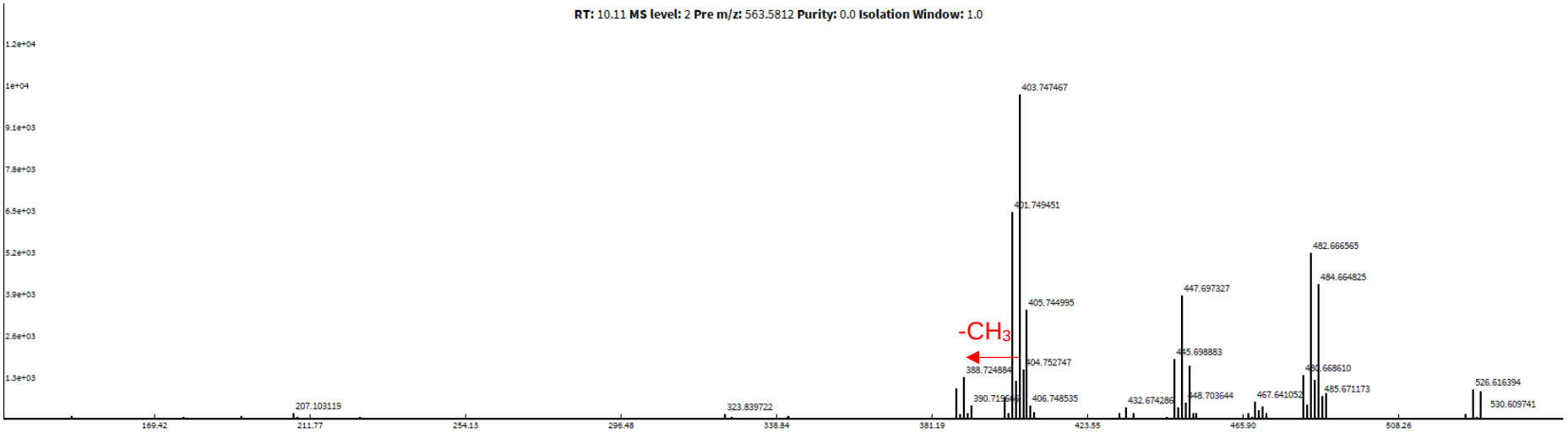
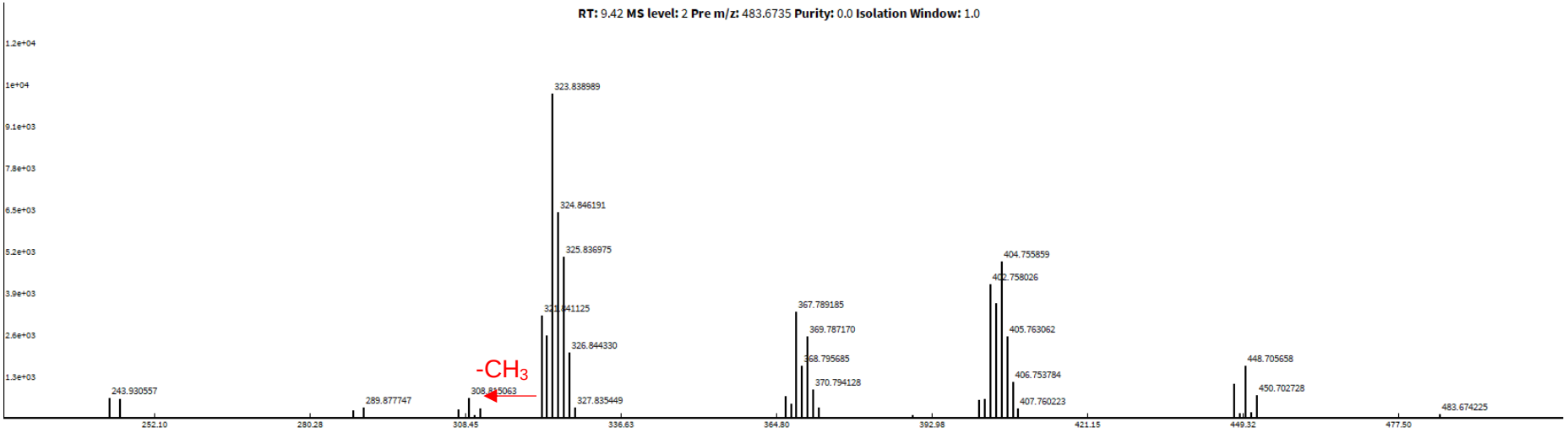
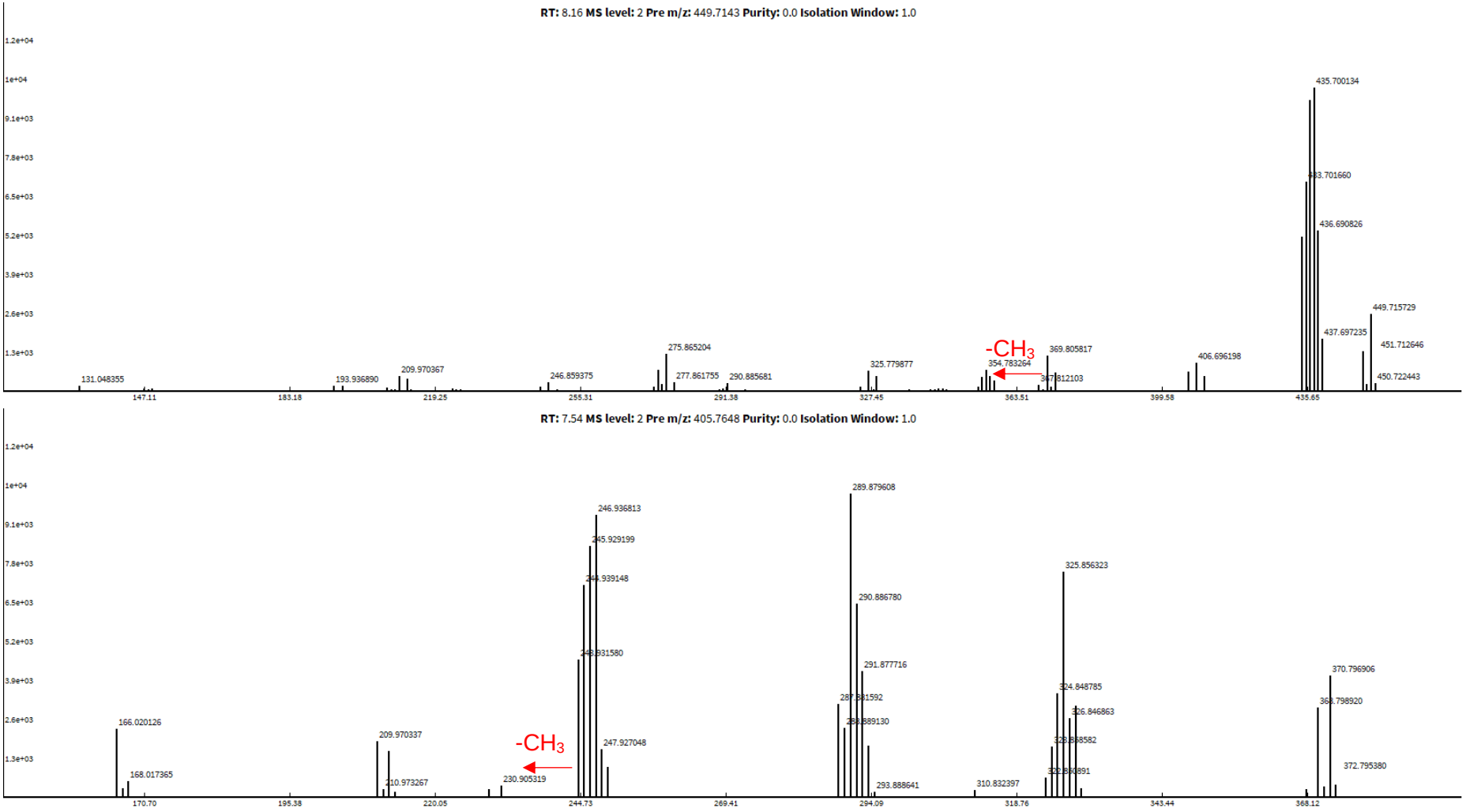


Figure S11. Juxtaposed MSMS spectra of a known methoxylated brominated diphenylether (below) versus the proposed mixed bromochlorodiphenylether (above). The two molecules show similar fragmentation patterns through the consecutive loss of halogens. Additionally., the two molecules possess the same fragments which are enclosed in matching colored boxes.


RT: 10.11 MS level: 2 Pre m/z: 563.5812 Purity: 0.0 Isolation Window: 1.0
1.2e+04
1e+04
9.1e+03
7.8e+03
6.5e+03
5.2e+03
3.9e+03
2.6e+03
1.3e+03
403.747467
401.749451
482.666565
484.664825
447.697327
405.744995
-CH3
404.752747
445.698883
388.724884
480.668610
485.671173
526.616394
530.609741
390.719666
406.748535
432.674286
448.703644
467.641052
207.103119
323.839722
169.42
211.77
254.13
296.48
338.84
381.19
423.55
465.90
508.26


RT: 9.42 MS level: 2 Pre m/z: 483.6735 Purity: 0.0 Isolation Window: 1.0
1.2e+04
1e+04
9.1e+03
7.8e+03
6.5e+03
5.2e+03
3.9e+03
2.6e+03
1.3e+03
323.838989
324.846191
325.836975
404.755859
402.758026
321.841125
367.789185
369.787170
405.763062
326.844330
368.795685
-CH3
448.705658
406.753784
370.794128
450.702728
243.930557
289.877747
308.845063
327.835449
407.760223
483.674225
252.10
280.28
308.45
336.63
364.80
392.98
421.15
449.32
477.50


RT: 8.16 MS level: 2 Pre m/z: 449.7143 Purity: 0.0 Isolation Window: 1.0
1.2e+04
1e+04
9.1e+03
7.8e+03
6.5e+03
5.2e+03
3.9e+03
2.6e+03
1.3e+03
131.048355
193.936890
209.970367
246.859375
275.865204
277.861755
290.885681
325.779877
354.783264
-CH3
369.805817
367.812103
406.696198
435.700134
433.701660
436.690826
437.697235
449.715729
451.712646
450.722443
147.11
183.18
219.25
255.31
291.38
327.45
363.51
399.58
435.65
RT: 7.54 MS level: 2 Pre m/z: 405.7648 Purity: 0.0 Isolation Window: 1.0
1.2e+04
1e+04
9.1e+03
7.8e+03
6.5e+03
5.2e+03
3.9e+03
2.6e+03
1.3e+03
166.020126
168.017365
209.970337
210.973267
230.905319
-CH3
246.936813
245.929199
244.939148
248.931580
247.927048
289.879608
290.886780
291.877716
287.881592
293.889130
293.888641
310.832397
325.856323
324.848785
326.846863
328.858582
329.860891
370.796906
368.798920
372.795380
170.70
195.38
220.05
244.73
269.41
294.09
318.76
343.44
368.12

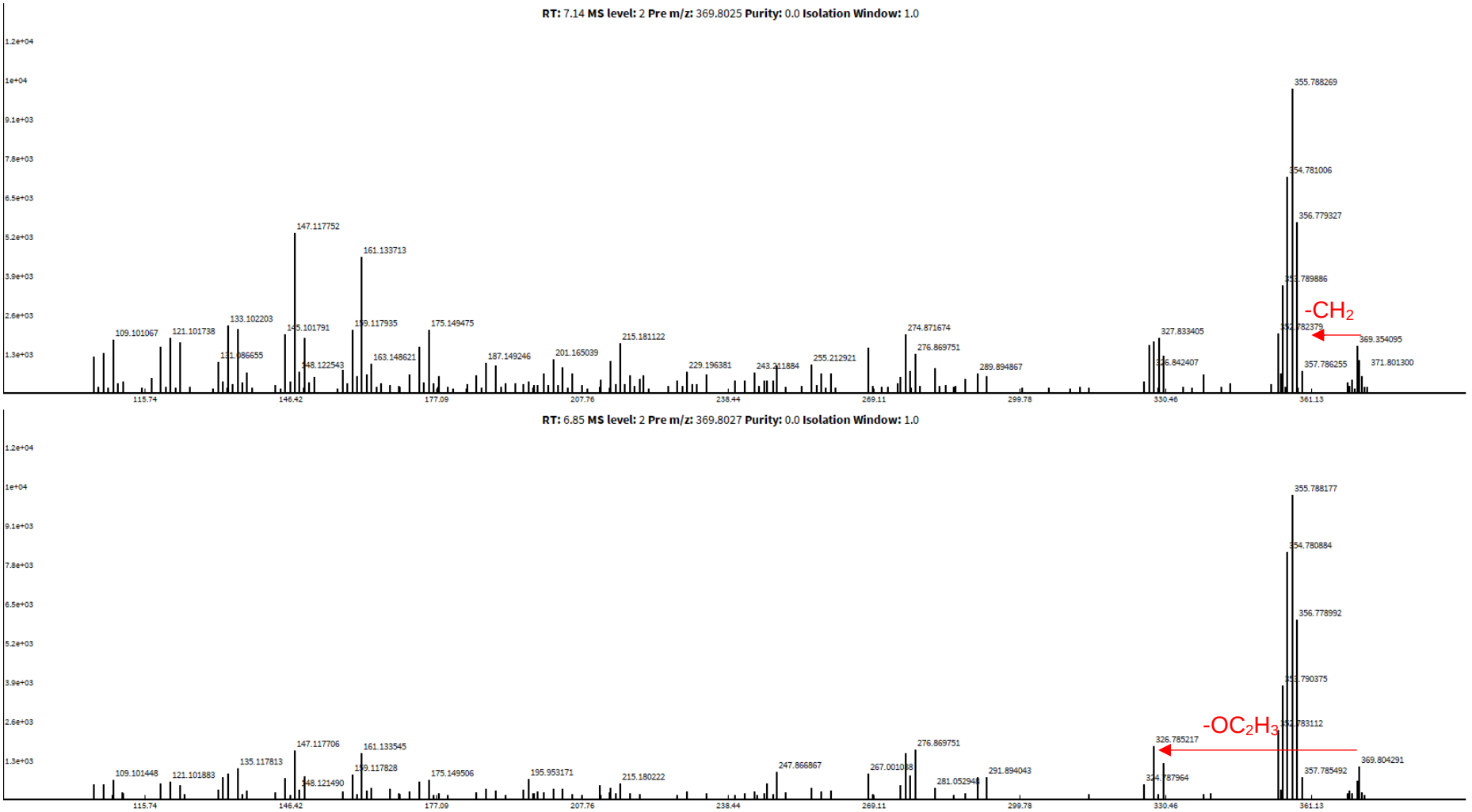


Figure S12. Example fragmentation spectra of some EMB-related features showing the loss of methyl group supporting the proposed structural scaffold.

References

(1) Sun, J.; Liu, J.; Liu, Q.; Qu, G.; Ruan, T.; Jiang, G. Sample Preparation Method for the Speciation of Polybrominated Diphenyl Ethers and Their Methoxylated and Hydroxylated Analogues in Diverse Environmental Matrices. *Talanta* **2012**, *88*, 669–676. https://doi.org/10.1016/j.talanta.2011.11.059.

(2) Koelmel, J. P.; Xie, H.; Price, E. J.; Lin, E. Z.; Manz, K. E.; Stelben, P.; Paige, M. K.; Papazian, S.; Okeme, J.; Jones, D. P.; Barupal, D.; Bowden, J. A.; Rostkowski, P.; Pennell, K. D.; Nikiforov, V.; Wang, T.; Hu, X.; Lai, Y.; Miller, G. W.; Walker, D. I.; Martin, J. W.; Godri Pollitt, K. J. An Actionable Annotation Scoring Framework for Gas Chromatography-High-Resolution Mass Spectrometry. *Exposome* **2022**, *2* (1), osac007. https://doi.org/10.1093/exposome/osac007.