SUPPORTING INFORMATION I

Appendix Ia. **Data/chemical analysis & quality control.**

Appendix Ib. **Details on tentative identification.**

Appendix Ia. Data/chemical analysis & quality control.

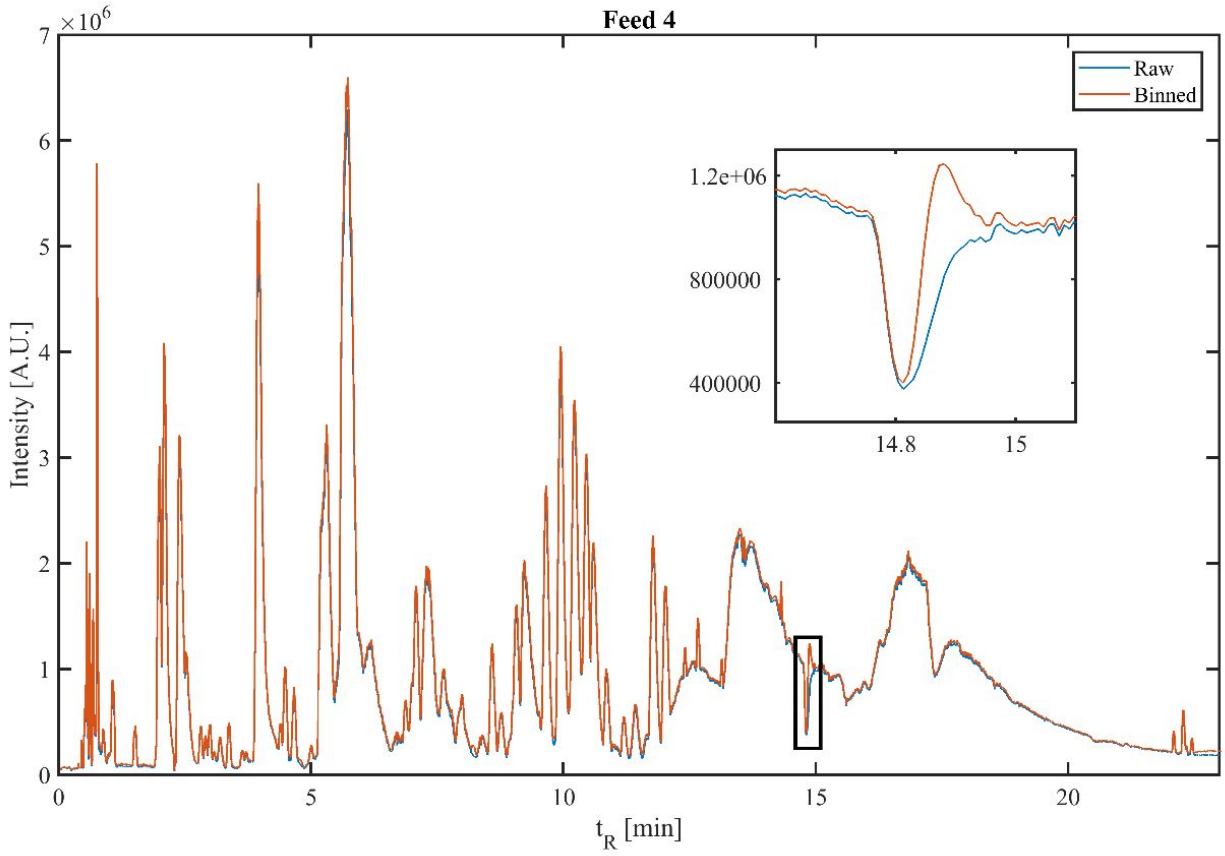
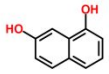
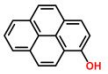
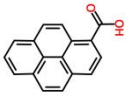


Figure SIa.1. Base peak chromatogram of Feed 4. Raw file in blue and binned to nominal mass in orange. The cut-out displays an example of the minor differences between processing binned and raw data.

Table SIa.1. OPACs used in preliminary study [1].

Compound	Formula	Mass [g/mol]	Structure
1,7-Dihydroxynaphthalene	C10H8O2	160.0524	
1-Naphthoic acid	C11H8O2	172.0524	
1-Hydroxynaphthalene	C10H8O	144.0575	
1-Hydroxypyrene	C16H10O	218.0732	
1-Pyrenecarboxylic acid	C17H10O2	246.0681	

1. Lübeck, J. S. Supercritical fluid chromatography for the fractionation of oxygenated polycyclic aromatic compounds from a coal tar middle distillate. *University of Copenhagen, Copenhagen, Denmark,* **2016**.

Table SIa.2. OptimCOW parameters [2].

Parameter	Value
Initial search space	
Segment length	50 – 200 scans (0.42 – 1.67 min)
Slack	1 – 3 scans (0.5 – 1.5 s)
Number of optimizations from grid maxima	3
Maximum number of optimization steps	50
Maximum absolute correction (as % of run length)	5
Number of grid points in [] axis	
Segment length	5
Slack	3
Activate automatic determination of pad size	1

2. Skov, T.; van den Berg, F.; Tomasi, G.; Bro, R., Automated alignment of chromatographic data. *Journal of Chemometrics* **2006**, 20, (11-12), 484-497.

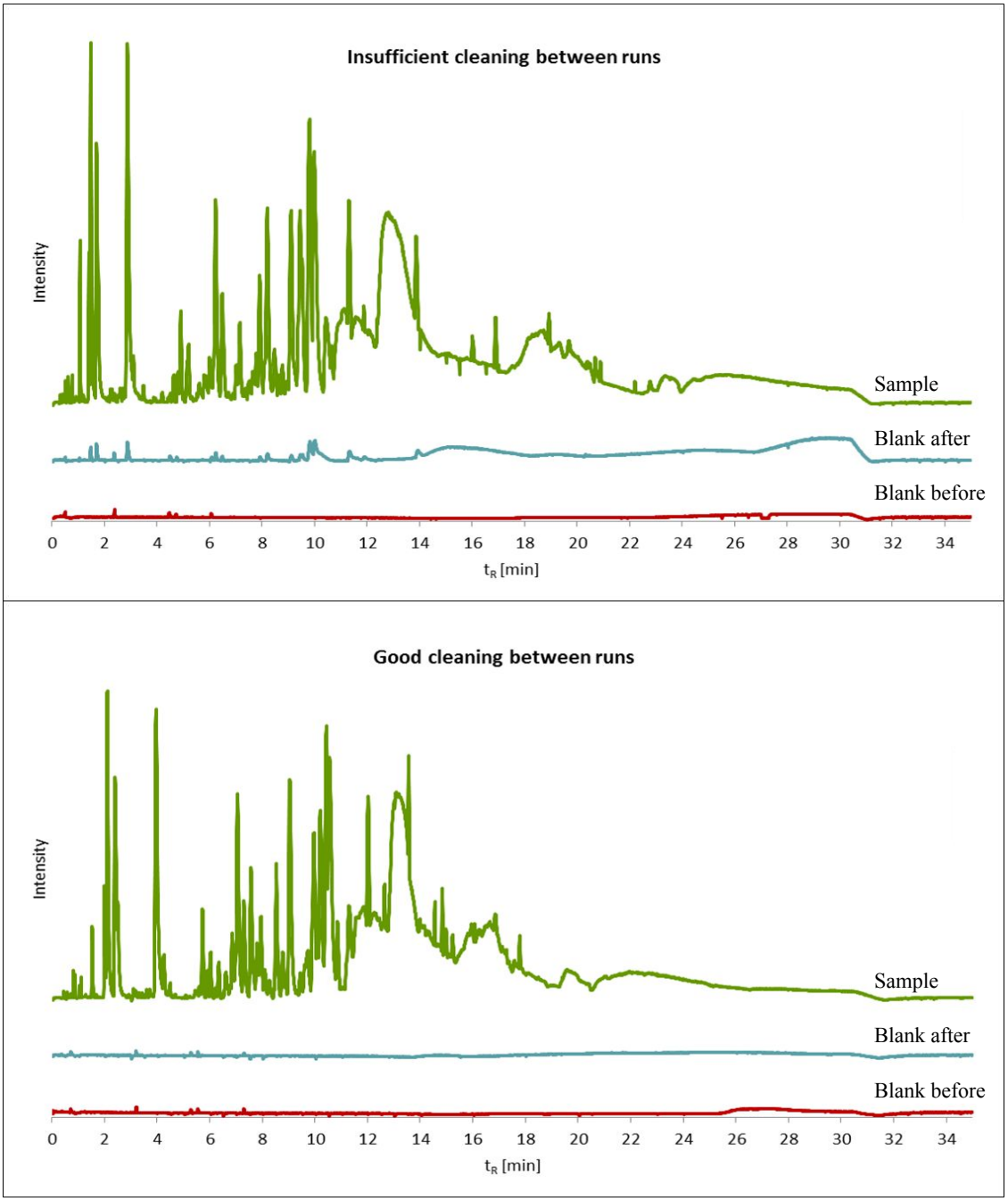


Figure SIa.2. BPCs of samples (diethyl ether extracts) and two methanol blanks before (top) and after (bottom) optimization of the washing solvents.

Table SIa.3. Sequence setup for SFC-HRMS analysis. Wash and Clean* are both blanks (i.e. dichloromethane:methanol, 50:50).

1	Wash	22	Wash	43	Wash	64	Wash
2	Clean	23	Clean	44	Clean	65	Clean
3	**QC**	24	**2A**	45	**1B**	66	**QC**
4	Wash	25	Wash	46	Wash	67	Wash
5	Clean	26	Clean	47	Clean	68	Clean
6	**Method Blank**	27	**5A**	48	**Method Blank**	69	**2C**
7	Wash	28	Wash	49	Wash	70	Wash
8	Clean	29	Clean	50	Clean	71	Clean
9	**4A**	30	**Method Blank**	51	**2B**	72	**1C**
10	Wash	31	Wash	52	Wash	73	Wash
11	Clean	32	Clean	53	Clean	74	Clean
12	**Method Blank**	33	**3B**	54	**Method Blank**	75	**Method Blank**
13	Wash	34	Wash	55	Wash	76	Wash
14	Clean	35	Clean	56	Clean	77	Clean
15	**1A**	36	**4B**	57	**4C**	78	**QC**
16	Wash	37	Wash	58	Wash	79	Wash
17	Clean	38	Clean	59	Clean	80	Clean
18	**QC**	39	**5B**	60	**3C**	81	**Method Blank**
19	Wash	40	Wash	61	Wash	82	Wash
20	Clean	41	Clean	62	Clean	83	Clean
21	**3A**	42	**QC**	63	**5C**		

*) *Clean* is a shorter run at lower flow rate (0.5 mL/min, and to higher proportion of ethanol (75%B), gradient program: 1. Initial, 1.5 mL/min, 99:1 (A:B); 2. 2.0 min, 0.5 mL/min, 75:25; 3. 4.0 min, 0.5 mL/min, 25:75; 4. 16.0 min, 0.5 mL/min, 25:75; 5. 16.5 min, 0.5 mL/min, 99:1; 6. 18.0 min; 0.5 mL/min; 7. 19.0 min; 1.5 mL/min, 99:1.

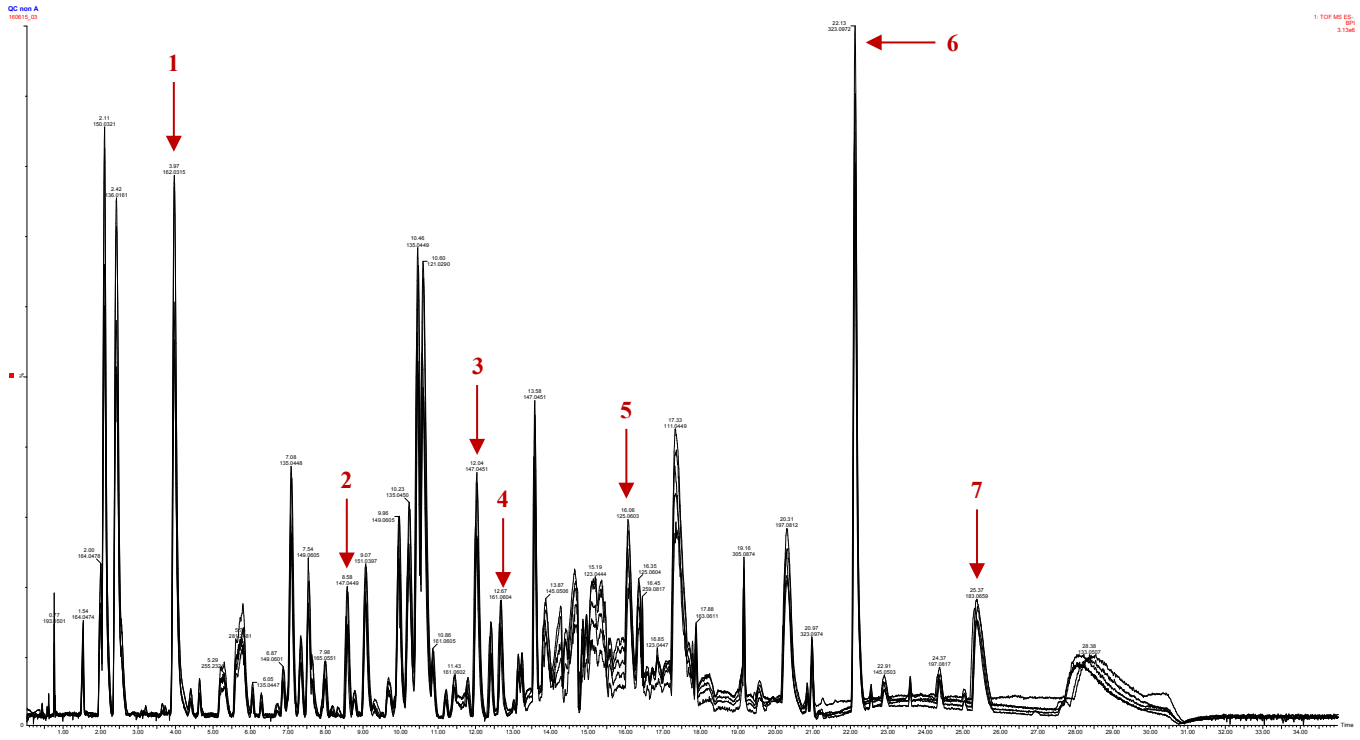


Figure SIa.3. BPCs for five replicate analysis of the QC sample. Red arrows and numbers indicate the peaks chosen to evaluate repeatability. 1) m/z 162.0315, 2) m/z 147.0449, 3) m/z 147.0451, 4) m/z 161.0601, 5) m/z 125.0603, 6) m/z 323.0972 and 7) m/z 183.0659.

Table SIa.4. Seven chosen peaks (m/z in parenthesis) to evaluate repeatability of QC samples. The figures are the mean value of five replicate measurements; the relative standard deviation (RSD) is given in parenthesis.

	Peak [mean (RSD)]						
	1 (162.0315)	**2 (147.0449)**	**3 (147.0451)**	**4 (161.0601)**	**5 (125.0603)**	**6 (323.0972)**	**7 (183.0659)**
t_R [min]	3.97 (0.1%)	8.58 (0.2%)	12.04 (0.1%)	12.64 (0.1%)	16.06 (0.1%)	22.13 (0.02%)	25.31 (0.1%)
Intensity (×10^5)	20.3 (18.9%)	5.30 (15.6%)	9.69 (14.2%)	4.95 (13.8%)	7.99 (14.0%)	27.0 (10.3%)	5.20 (8.2%)
Noise (×10^4)	7.26 (8.0%)	7.02 (23.2%)	7.94 (31.9%)	6.92 (12.6%)	3.02 (28.0%)	9.14 (21.7%)	10.2 (26.4%)
S/N	27.78 (11.8%)	7.66 (8.6%)	12.78 (20.6%)	7.16 (5.3%)	2.74 (15.4%)	36.43 (20.1%)	5.32 (18.1%)

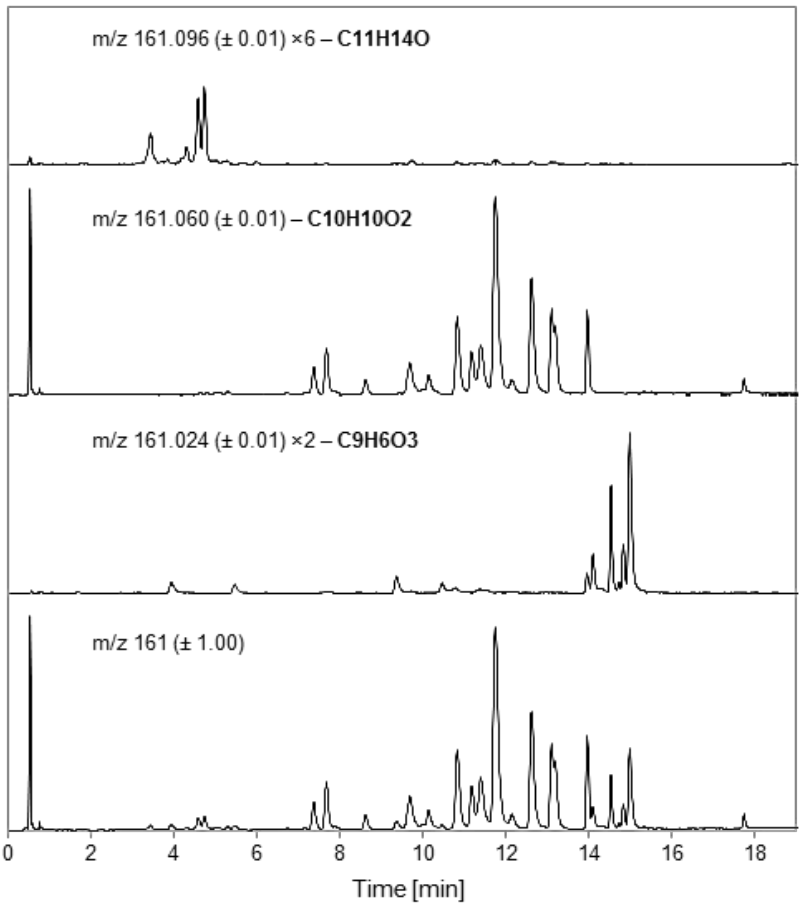


Figure SIa.4. Peak distribution within m/z 161 in Feed 2, mass window ranges and suggested molecular formulae. Intensities for C11H14O and C9H6O3 were enhanced by a factor of ×6 and ×2, respectively.

Appendix Ib. Details on tentative identification.

Table SIb.1. Candidates of the 24 regions of interest from the pixel-based analysis with fragmentation information.

Peak No.	t_R [min]	Measured mass [M-H]$^-$ [m/z]*	Largest fragment [mz]	Fragment type	Theoretical mass [M-H]$^-$ [m/z]	Mass accuracy [ppm]	Mass formula	Suggested compounds**	Similarity score [%]	InChiKeys
1	0.56	161.0605†	133.0652	[M-H-CO]$^-$	161.0603	1.2	C10H10O2	3-(3-Hydroxy-2-methylphenyl)prop-2-enal	62.8	VVMLEXQTELDKEO-HWKANZROSA-N
2	2.00	179.0710	164.0476†	[M-H-CH3]$^-$	179.0708	1.1	C10H12O3	Propioguaiacone	70.7	FBGXENQFSMMBNY-UHFFFAOYSA-N
	2.10	165.0553	150.0319†	[M-H-CH3]$^-$	165.0552	0.6	C9H10O3	Acetovanillone	77.9	DFYRUELUNQRZTB-UHFFFAOYSA-N
3	2.41	151.0398	136.0166†	[M-H-CH3]$^-$	151.0395	2.0	C8H8O3	Iso-vanillin	76.5	MWOOGOJBHIARFG-UHFFFAOYSA-N
	2.52	193.0524†	178.0267	[M-H-CH3]$^-$	193.0501	1.6	C10H10O4	1-(3-Hydroxy-4-methoxyphenyl)propane-1,2-dione	67.2	CUSKWZIYVGDDOW-UHFFFAOYSA-N
	2.82	135.0811	121.0655†	-	135.0811	0.0	C9H12O	3-Propylphenol	64.9	MPWGZBWDLMDIHO-UHFFFAOYSA-N
4	2.92	135.0814†	-	-	135.0811	2.2	C9H12O	2-Propylphenol	52.7	LCHYEKKJCUJAKN-UHFFFAOYSA-N
	3.04	193.0867	149.0969†	[M-H-CO2]$^-$	193.0865	-2.6	C11H14O3	2,6-Diethylphenol	60.6	METWAQRCMRWDAW-UHFFFAOYSA-N
5	3.96	177.0555	162.0320†	[M-H-CH3]$^-$	177.0552	1.7	C10H10O3	Coniferyl aldehyde	81.9	DKZBBWMURDFHNE-NSCUHMNNSA-N
6	4.41	133.0658†	-	-	133.0653	3.8	C9H10O	2-Propenylphenol	68.1	WHGXZPQWZJUGEP-UHFFFAOYSA-N
	4.50	191.0709	147.0813†	[M-H-CO2]$^-$	191.0708	0.5	C11H12O3	4-Methyl-3-prop-1-en-2-yloxybenzoic acid	66.0	UDLGNBLVJXODLP-UHFFFAOYSA-N
7	4.70	191.0708	147.0810†	[M-H-CO2]$^-$	191.0708	0.0	C11H12O3	3-Methyl-5-prop-2-enoxybenzoic acid	62.2	RHWSMUOOAQBLAI-UHFFFAOYSA-N
8	5.29	255.2322†	-	-	255.2324	-0.8	C16H32O2	Palmitic acid	48.7	IPCSVZSSVZVIGE-UHFFFAOYSA-N
	5.44	269.2479	-	-	269.2481	-0.7	C17H34O2	Margaric acid	41.1	KEMQGTRYUADPNZ-UHFFFAOYSA-N
	5.50	283.2636	-	-	283.2637	-0.4	C18H36O2	(E,5S,8S)-Octadec-6-ene-5,8-diol	39.6	MJCOSMVJAWZLTG-MFXBQNAZSA-N
	5.69	281.2484†	-	-	281.2481	1.1	C18H34O2	7-Octadecanoic acid	57.1	RVUCYJXFCAVHNC-QXMHVHEDSA-N
9	5.90	279.2329†	-	-	279.2324	1.8	C18H32O2	Linoleic acid	83.2	OYHQOLUKZRVURQ-HZJYTTRNSA-N
10	7.06	135.0451†	-	-	135.0446	3.7	C8H8O2	4-(1-Hydroxyvinyl)phenol	62.0	LYECQNYXCRATFL-UHFFFAOYSA-N
11	7.89/ 7.96/ 8.35	301.2163†		-	301.2168	-1.7	C20H30O2	(+)-Pisiferol/ (E)-7-[(1S,5Z)-5-[(E)-oct-2-enylidene]cyclopent-2-en-1-yl]hept-5-enoic acid/ 2-Butyl-10-phenyldec-6-enoic acid	46.5/ 45.7/ 39.8	NKGGFTFDYGTUSL-ICSRJNTNSA-N/ SDRIYNVSOZHFNZ-IBGZPJMESA-N/ INUACFUKPNQDHM-SNAWJCMRSA-N
12	8.98	227.1074†	212.0827	[M-H-CH3]$^-$	227.1072	0.9	C15H16O2	4-[2-(4-Methoxyphenyl)ethyl]phenol	56.2	WJIQXMNDMJNIJG-UHFFFAOYSA-N
13	9.18	299.2013†		-	299.2011	3.3	C20H28O2	1-[3-(3,3-dimethylcyclohexyl)phenyl]-cyclopentane-1-carboxylic acid	48.9	ABRNAYSNGWMWOY-UHFFFAOYSA-N
14	9.69	163.0763†	-	-	163.0759	2.5	C10H12O2	1-(3-Hydroxy-4-methylphenyl)propan-2-one	60.0	UUFNCKJLESFHLY-UHFFFAOYSA-N
15	9.94/ 10.17	149.0606†	-	-	149.0603	2.0	C9H10O2	4-Allylcatechol/ 2-Methoxy-4-vinylphenol	63.0/ 64.7	FHEHIXJLCWUPCZ-UHFFFAOYSA-N / YOMSJEATGXXYPX-UHFFFAOYSA-N

Table SIb.1. Continues.

Peak No.	t_R [min]	Measured mass [M-H]$^-$ [m/z]*	Largest fragment [mz]	Fragment type	Theoretical mass [M-H]$^-$ [m/z]	Mass accuracy [ppm]	Mass formula	Suggested compounds**	Similarity score [%]	InChiKeys
16a	10.24	135.0451†	-	-	135.0446	3.7	C8H8O2	3-Vinylcatechol	63.7	PAPAWUGOOGNNAR-UHFFFAOYSA-N
16b	10.47	135.0448†	-	-	135.0446	1.5	C8H8O2	3-Hydroxyacetophenone	64.9	TXFPEBPIARQUIG-UHFFFAOYSA-N
17	10.60	121.0291†	-	-	121.0290	0.8	C7H6O2	3-Hydroxybenzaldehyde	80.3	IAVREABSGIHHMO-UHFFFAOYSA-N
18	10.85	161.0604†	133.0654	[M-H-CO]$^-$	161.0603	0.6	C10H10O2	2,3-Bis(ethenyl)benzene-1,4-diol	59.9	RFAMPQQYSXSAOD-UHFFFAOYSA-N
19	11.79	161.0602†	-	-	161.0603	-0.6	C10H10O2	3-(3-Hydroxyphenyl)but-1-en-1-one	60.8	CELCCBLEZBOUPD-UHFFFAOYSA-N
21	12.40	147.0447†	-	-	147.0446	0.7	C9H8O2	p-Coumaraldehyde	71.3	CJXMVKYNVIGQBS-OWOJBTEDSA-N
22	12.88	225.0916†	210.068	[M-H-CH3]$^-$	225.0916	0.0	C15H14O2	4-[1-(4-Methoxyphenyl)vinyl]phenol	64.8	BOVUWEPBZSXMQH-UHFFFAOYSA-N
23	14.33	249.1285†	233.0955	[M-CH4-H]$^-$	249.1279	2.4	C18H18O	4-Prop-2-enyl-2-(3-prop-2-enylphenyl)phenol	59.2	USUQHQHINSUZTG-UHFFFAOYSA-N
24	14.82	287.0926†	269.0821	[M-H2O-H]$^-$	287.0919	3.1	C16H16O5	1-(2,4-dihydroxyphenyl)-2-(2,4-dimethoxyphenyl) ethanone	71.7	ZQJGSCCWQJKINT-UHFFFAOYSA-N

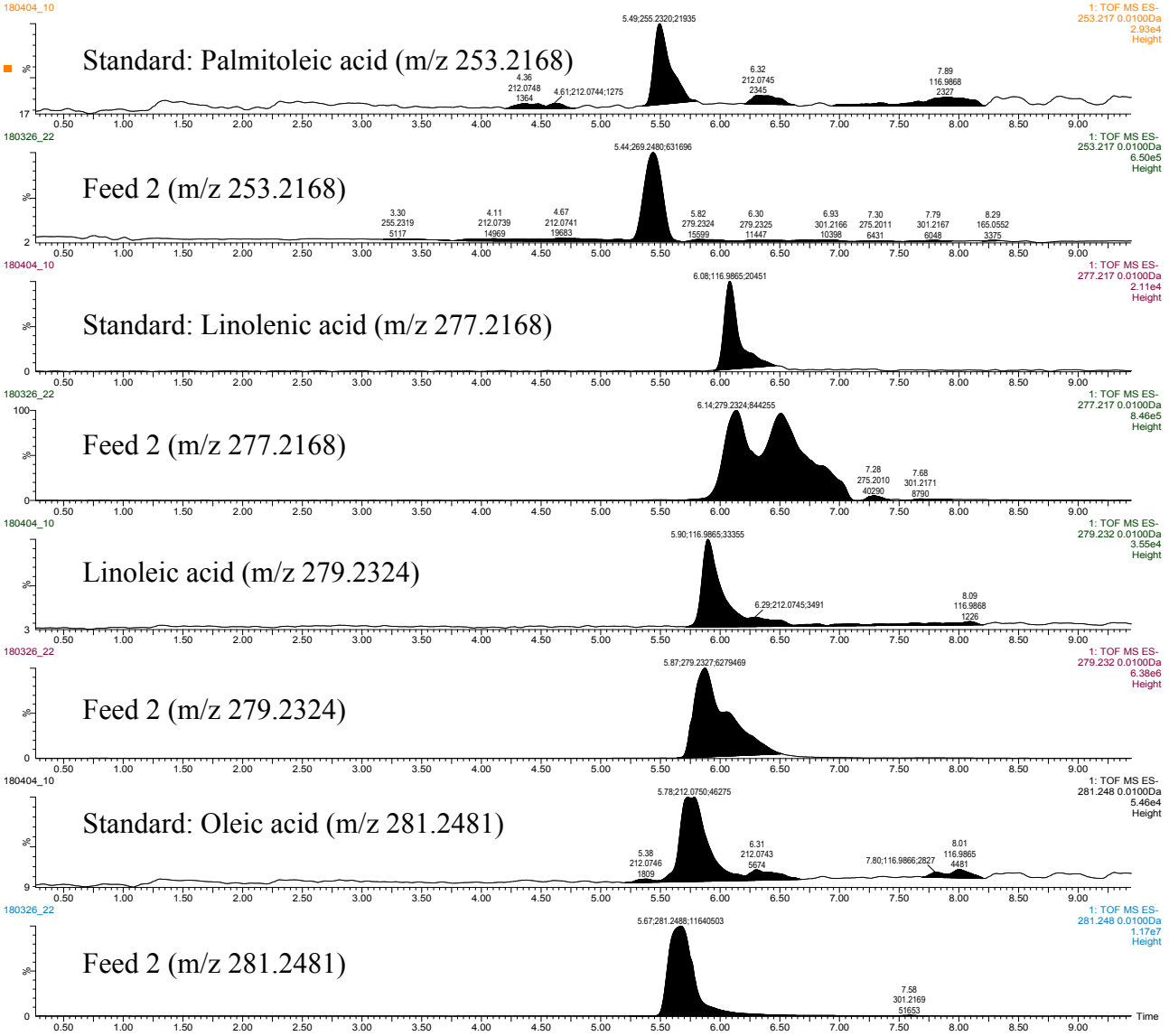


Figure SIb.1. Standard fatty acid analysis (0.066 µg/mL) compared with extracted ion chromatograms from Feed 2.

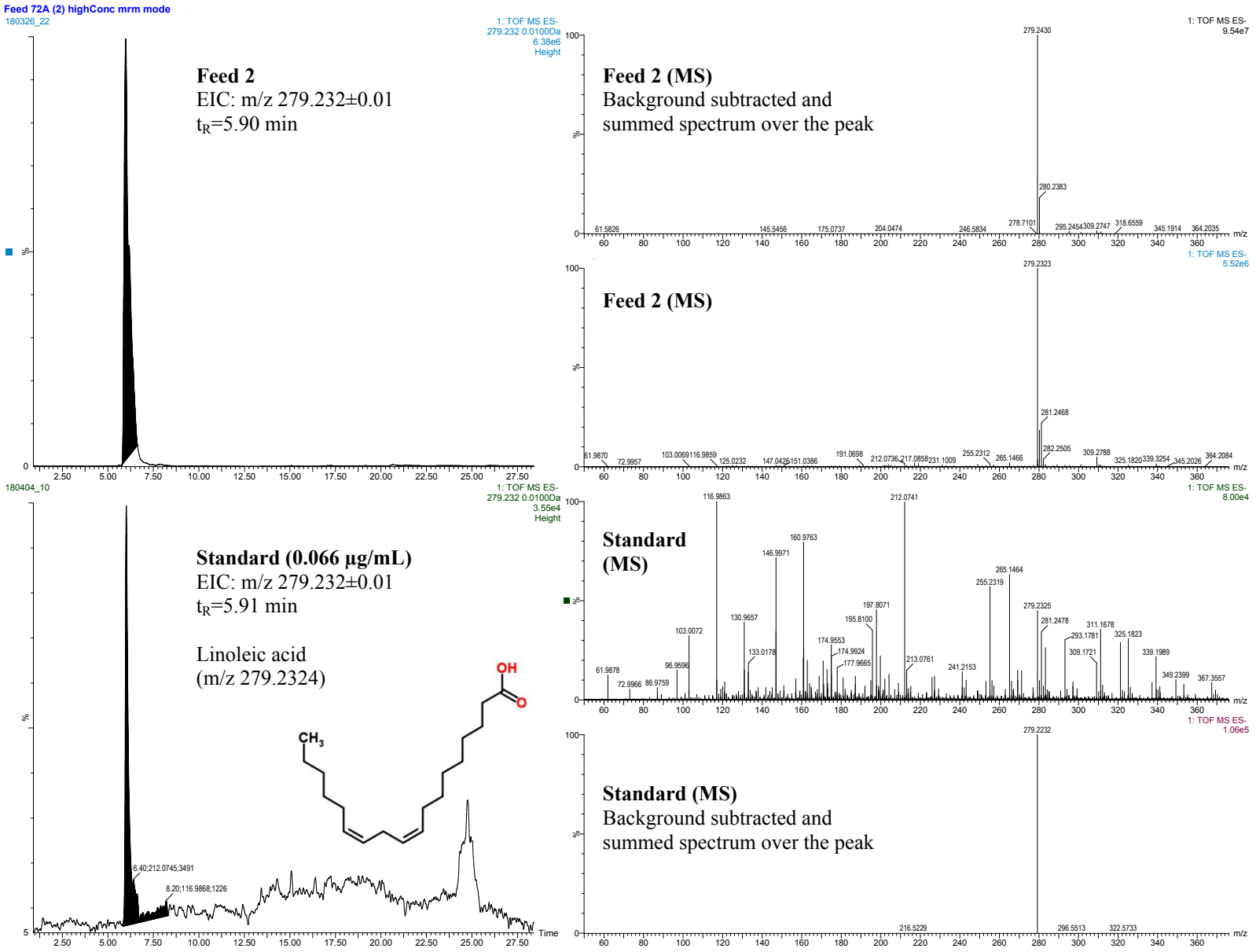


Figure SIb.2. Comparison of linoleic acid in Feed 2 (top) and standard (bottom). Chromatograms (left) and spectra (right).

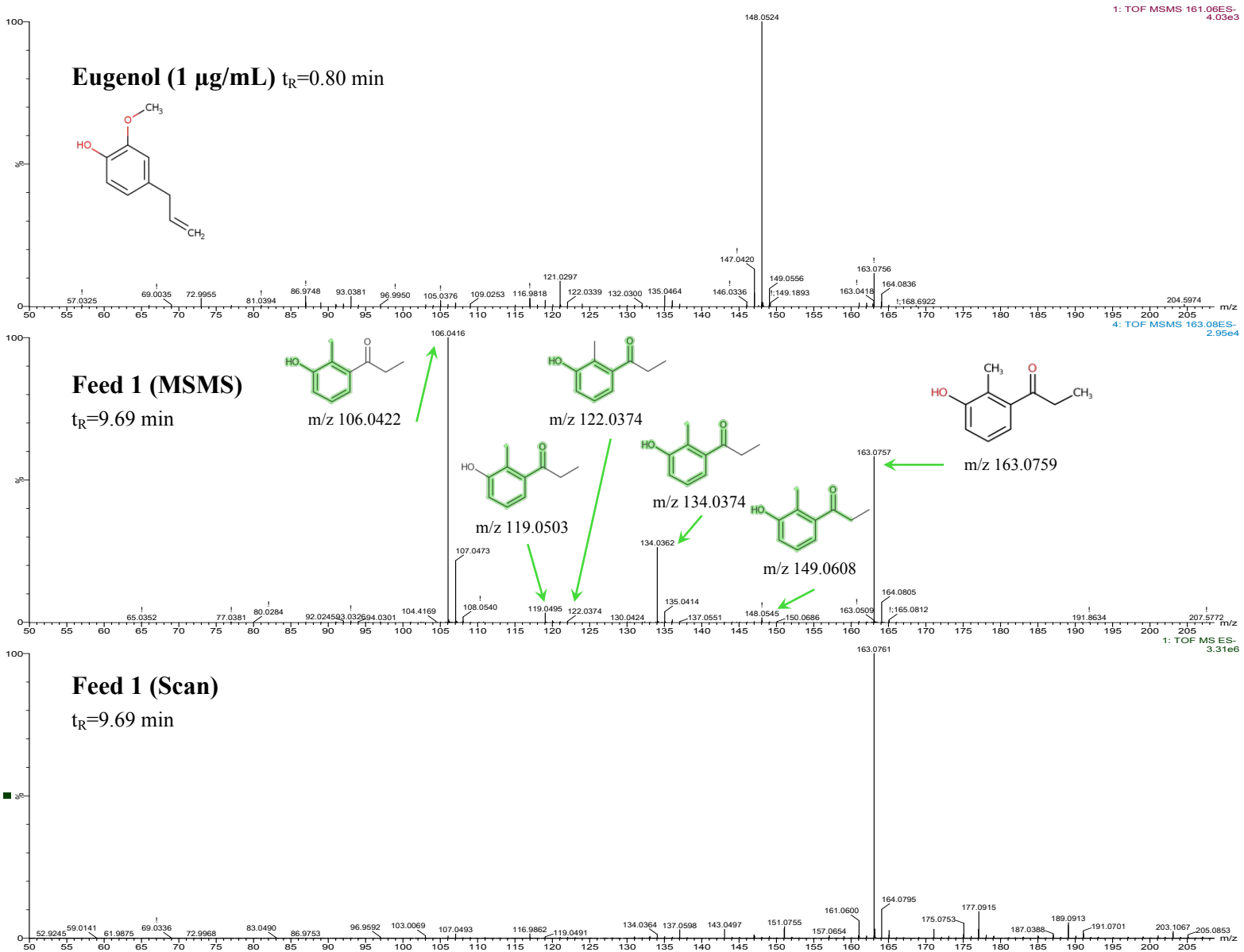


Figure SIb.3. Comparison of MS/MS spectra of Eugenol with Feed 1 (MS/MS and MS Scan).

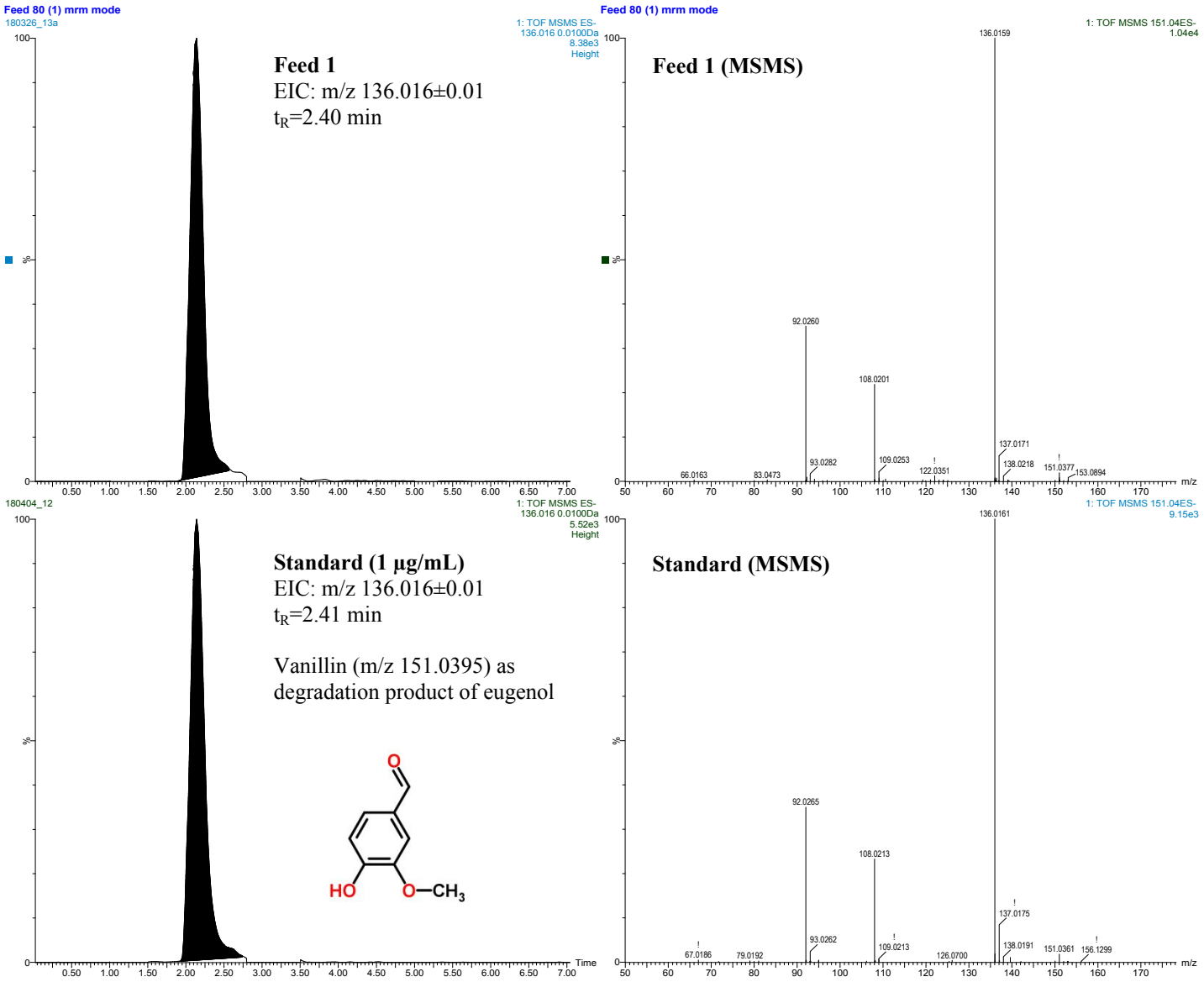


Figure SIb.4. Comparison of vanillin in Feed 1 (top) and standard (bottom). Chromatograms (left) and spectra (right).

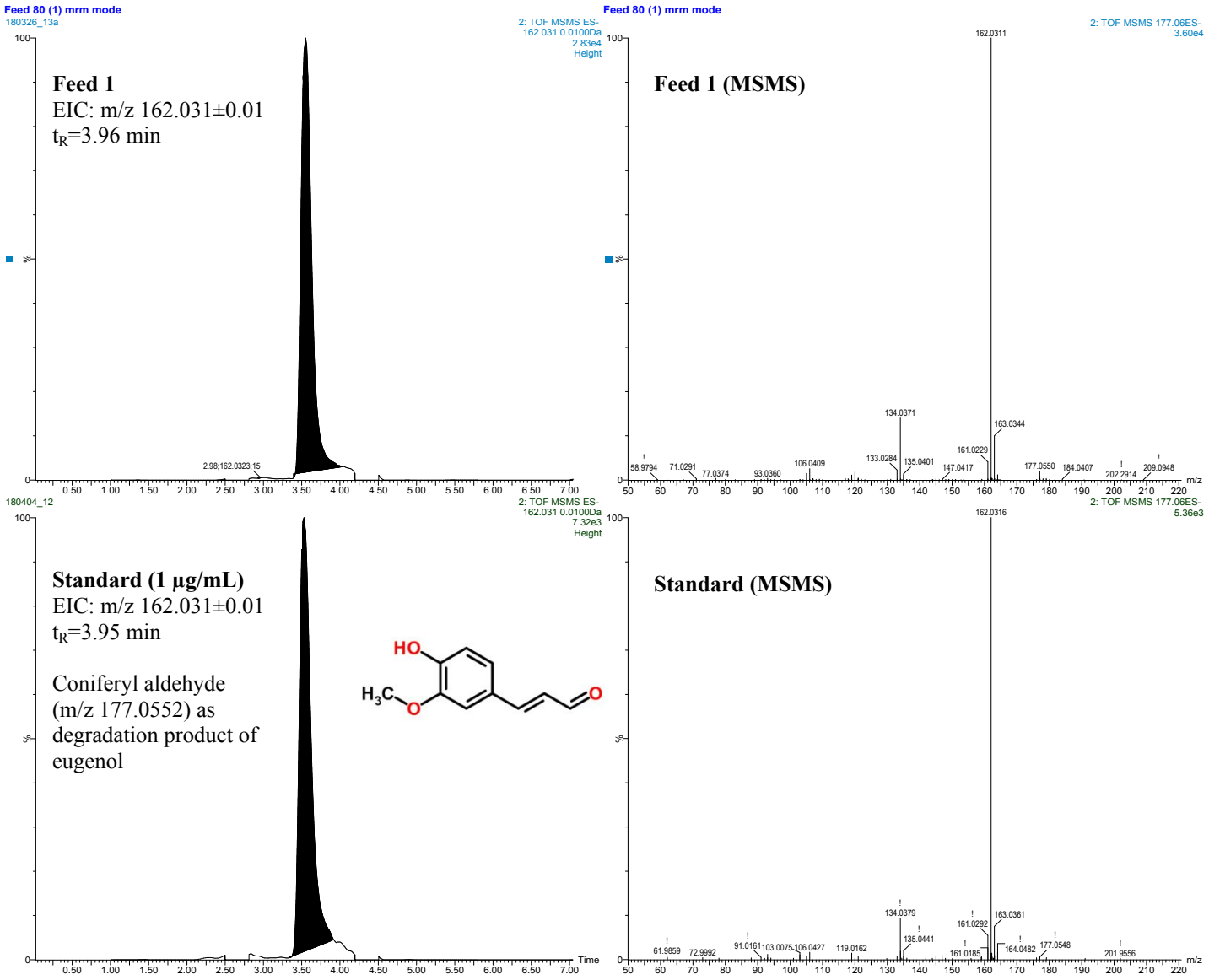


Figure SIb.5. Comparison of coniferyl aldehyde in Feed 1 (top) and standard (bottom). Chromatograms (left) and spectra (right).