

"Supplementary Data"

**Nonlinear Mixture-wise Expansion Approach to Underdetermined Blind Separation of
Nonnegative Dependent Sources**

Ivica Kopriva^{†}, Ivanka Jerić^{‡*} and Lidija Brkljačić[‡]*

Ruder Bošković Institute, Bijenička cesta 54, HR-10000, Zagreb, Croatia

[†]Division of Laser and Atomic Research and Development

[‡]Division of Organic Chemistry and Biochemistry

^{†*} Phone: +385.1.4571.286. Fax: +385.1.4680.104. E-mail: ikopriva@irb.hr.

^{‡*} Phone: +385.1.4560.998. Fax: +385.1.4680.195. E-mail: ijeric@irb.hr.

Table of contents:

Subject	Page
Table S-1. Volumes of components and solvent used for the preparation of mixtures X_1 to X_5 .	3
Table S-2. Normalized correlation coefficients for analytes C_1 to C_{10} shown in Figure S-1. Cross-correlation coefficients greater than 0.1 are in bold font. Analytes C_4 and C_7 have practically same mass spectra. Analytes C_4 and C_5 , C_5 and C_7 , C_3 and C_9 , C_2 and C_3 as well as C_2 and C_9 are significantly correlated.	4
Table S-3. Maximal normalized correlation coefficients between analytes C_1 to C_{10} and components extracted by the NMU and NMF_L0 algorithms applied on LMM (1). Columns from left to right: factorization methods; correlation coefficients. The asterisk signs denotes analytes in the library associated with the same extracted component. In addition to highly correlated analytes C_4 and C_7 other, less correlated analytes, were also associated with the same extracted component.	5
Table S-4. Maximal normalized correlation coefficients between analytes C_1 to C_{10} and components extracted by proposed dimensionality expansion method. Columns from left to right: combinations of mixture spectra; correlation coefficients; type of nonlinear mapping used in dimensionality expansion transform. The star '*' denotes analytes in the library associated with the same extracted component. As expected highly correlated analytes C_4 and C_7 were associated with the same extracted component.	6 - 10
Figure S-1. Mass spectra of compounds (analytes) C_1 - C_{10} .	11-15
Figure S-2. Mass spectra of mixtures X_1 to X_5 .	16-18
Figure S-3. Analytes C_1 to C_{10} extracted from mixtures X_1 to X_5 by NPT-NMU algorithm. Dimensionality expansion transform is based on Gaussian kernel with variance $\sigma^2=1$. Hard thresholding with $\tau=10^{-7}$ has been used before NMU decomposition in reproducible kernel Hilbert space.	19-23

Table S-1. Volumes of analytes and solvent used for the preparation of mixtures **X**₁ to **X**₅.

	Volume of components (μL)										Total volume (μL) of components	Volume (μL) of solvent (10% MeOH)
	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈	C ₉	C ₁₀		
X ₁	10	50	40	55	45	20	35	30	25	15	325	625
X ₂	50	45	25	35	30	20	15	10	40	55	325	625
X ₃	10	45	40	50	55	15	35	25	30	20	325	625
X ₄	20	50	35	55	45	30	25	10	40	15	325	625
X ₅	55	10	20	15	25	50	35	40	45	30	325	625

Table S-2. Normalized correlation coefficients for analytes $\mathbf{C}_1$ to $\mathbf{C}_{10}$ shown in Figure S-1.

Cross-correlation coefficients greater than 0.1 are in bold font. Analytes $\mathbf{C}_4$ and $\mathbf{C}_7$ have practically same mass spectra. Analytes $\mathbf{C}_4$ and $\mathbf{C}_5$, $\mathbf{C}_5$ and $\mathbf{C}_7$, $\mathbf{C}_3$ and $\mathbf{C}_9$, $\mathbf{C}_2$ and $\mathbf{C}_3$ as well as

$\mathbf{C}_2$ and $\mathbf{C}_9$ are significantly correlated. $c_{nm} = \frac{\langle \mathbf{s}_n, \mathbf{s}_m \rangle}{\|\mathbf{s}_n\| \|\mathbf{s}_m\|}$, $n, m = 1, \dots, 10$.

entry	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_3$	$\mathbf{C}_4$	$\mathbf{C}_5$	$\mathbf{C}_6$	$\mathbf{C}_7$	$\mathbf{C}_8$	$\mathbf{C}_9$	$\mathbf{C}_{10}$
$\mathbf{C}_1$	1.0000	0.0037	0.0576	0.0001	0.0005	0.0014	0.0004	0.0001	0.0001	0.0002
$\mathbf{C}_2$	0.0037	1.0000	0.1540	0.0012	0.0003	0.0733	0.0017	0.0009	0.1362	0.0001
$\mathbf{C}_3$	0.0576	0.1540	1.0000	0.0018	0.0021	0.0099	0.0022	0.0197	0.3077	0.0003
$\mathbf{C}_4$	0.0001	0.0012	0.0018	1.0000	0.4775	0.0037	0.9543	0.0015	0.0004	0.0000
$\mathbf{C}_5$	0.0005	0.0003	0.0021	0.4775	1.0000	0.0028	0.3930	0.0054	0.0005	0.0000
$\mathbf{C}_6$	0.0014	0.0733	0.0099	0.0037	0.0028	1.0000	0.0025	0.0023	0.0001	0.0026
$\mathbf{C}_7$	0.0004	0.0017	0.0022	0.9543	0.3930	0.0025	1.0000	0.0037	0.0003	0.0000
$\mathbf{C}_8$	0.0001	0.0009	0.0197	0.0015	0.0054	0.0023	0.0037	1.0000	0.0011	0.0037
$\mathbf{C}_9$	0.0001	0.1362	0.3077	0.0004	0.0005	0.0001	0.0003	0.0011	1.0000	0.0106
$\mathbf{C}_{10}$	0.0002	0.0001	0.0003	0.0000	0.0000	0.0026	0.0000	0.0037	0.0106	1.0000

Table S-3. Maximal normalized correlation coefficients between analytes C_1 to C_{10} and components extracted by the NMU and NMF_L0 algorithms applied on LMM (1) composed of five mixtures X_1 to X_5 . Mixture matrix has been scaled by $\arg \max_{n,t} \{X_{nt}\}_{n,t=1}^{N,T}$. Columns from left to right: factorization method; normalized correlation coefficients: $c_{mm} = \langle s_m, \hat{s}_m \rangle / \|s_m\| \|\hat{s}_m\|$, $m=1,...,10$. The NMU algorithm was run with number of sources set to $M=T=2901$ and the best matching results are shown in the table. The NMF_L0 algorithm was run with the following parameter setup: reverse sparse nonnegative least square (rsNNLS) sparse coder and alternating nonnegative least square (ANLS) for dictionary update stage. NMF_L0 algorithm was cross-validated for values of $K \in \{3,4,5,6\}$. Number of sources was set to $M=300$ and number of iterations was set to 200. The best matching results are shown in the table. The asterisk sign denotes analytes in the library associated with the same extracted component. The best result is highlighted.

	C_{11}	C_{22}	C_{33}	C_{44}	C_{55}	C_{66}	C_{77}	C_{88}	C_{99}	$C_{10;10}$
NMU	0.0405 *	0.6922	0.6492 *	0.9846 x	0.4499 x	0.3641 +	0.9738 x	0.9956	0.9087 +	0.7334
NMF_L0, K=3, Update Err=2.3x10 ⁻³	0.8566	0.7141	0.8772	0.7779 x	0.4348 x	0.9766	0.7498	0.5721	0.8558	0.7735
NMF_L0 K=4, Update Err=3.1x10 ⁻⁴	0.9602	0.8799	0.7435	0.8650 x	0.4716 x	0.7639	0.7036 x	0.7438	0.9567	0.9070
NMF_L0, K=5, Update Err=7x10 ⁻⁶	0.7269	0.9567	0.7448	0.8595	0.5616	0.9922	0.7117	0.6401	0.9924	0.9880
NMF_L0 K=6, Update Err=1.4x10 ⁻⁶	0.3895	0.7619	0.8639	0.8646 x	0.6038	0.9747	0.7042 x	0.7181	0.9573	0.9075

Table S-4. Maximal normalized correlation coefficients between analytes $\mathbf{C}_1$ to $\mathbf{C}_{10}$ and components extracted by proposed dimensionality expansion method: $c_{mm} = \langle \mathbf{s}_m, \hat{\mathbf{s}}_m \rangle / \|\mathbf{s}_m\| / \|\hat{\mathbf{s}}_m\|$, $m=1, \dots, 10$. Columns from left to right: combinations of mixture spectra; correlation coefficients; information of variance, σ^2 , of the Gaussian kernel, type of factorization used with dimensionality expansion transform, type of scaling of the mixing matrix before mapping: $\mathbf{x}_{\max}$ refers to $\arg \max_{n,t} \{\mathbf{X}_{nt}\}_{n,t=1}^{N,T}$ and $\|\mathbf{x}\|_1$ refers to $\arg \max_t \{\|\mathbf{x}_{:,t}\|_1\}_{t=1}^T$. The NMF_L0 algorithm was run with the following parameter setup: reverse sparse nonnegative least square (rsNNLS) sparse coder and alternating nonnegative least square (ANLS) for dictionary update stage. The best result is obtained for number of overlapped sources equal to $Q=50$. Number of sources was set to $P=200$ and number of iterations was set to 100. The star '*' denotes analytes in the library associated with the same extracted component. The best result is highlighted.

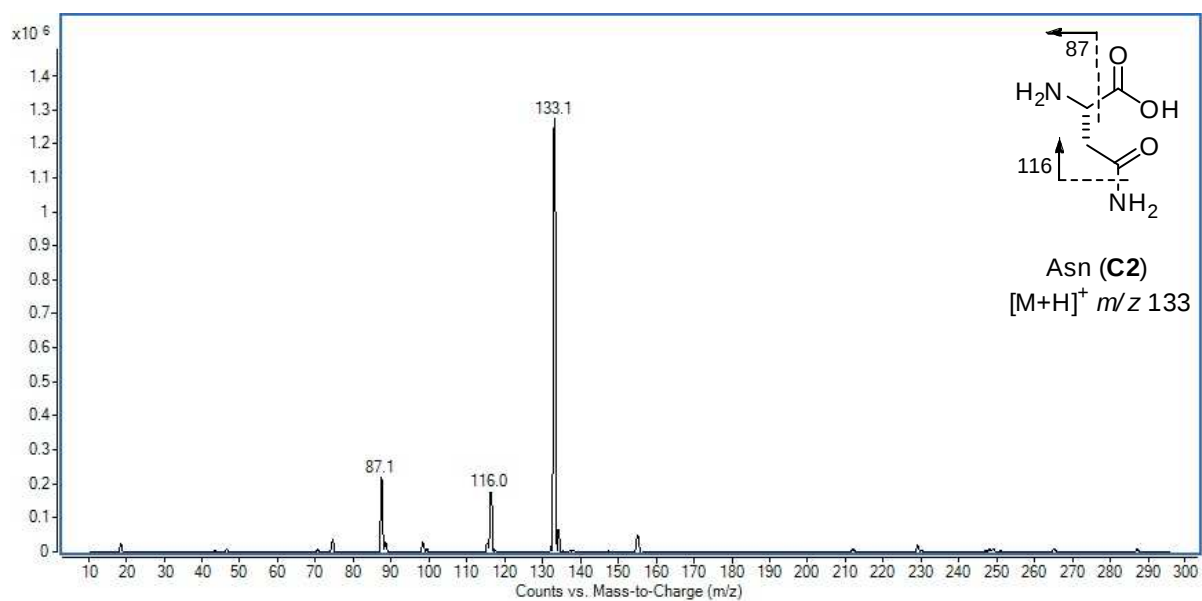
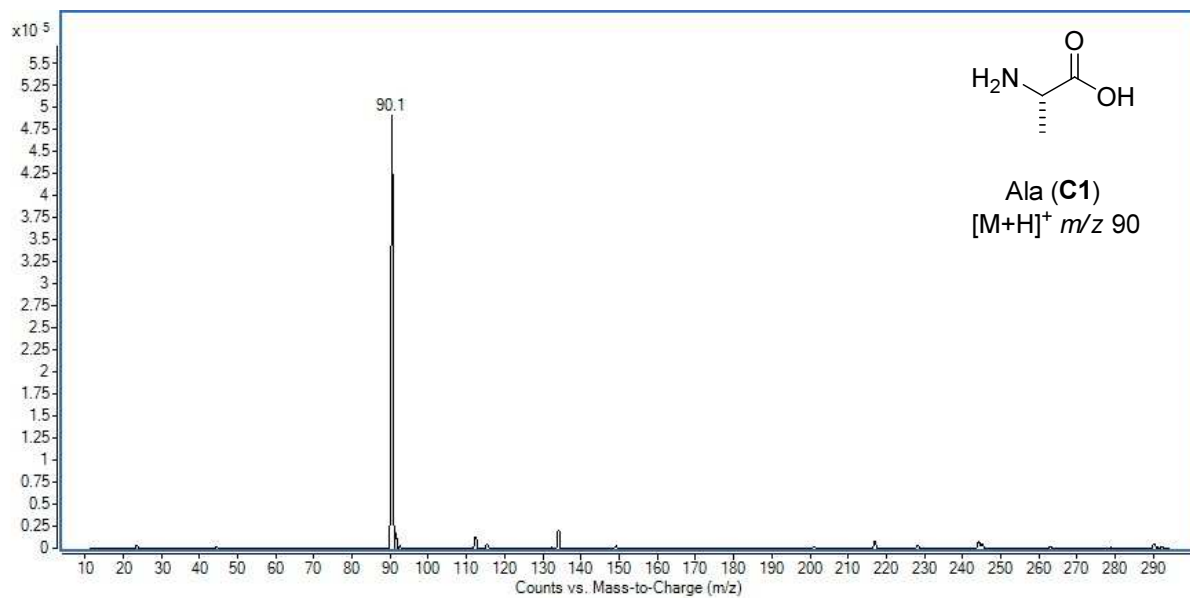
	$\mathbf{C}_{11}$	$\mathbf{C}_{22}$	$\mathbf{C}_{33}$	$\mathbf{C}_{44}$	$\mathbf{C}_{55}$	$\mathbf{C}_{66}$	$\mathbf{C}_{77}$	$\mathbf{C}_{88}$	$\mathbf{C}_{99}$	$\mathbf{C}_{10;10}$	Mapping factorization
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.6114	0.9182	0.6351	0.7619	0.5606	0.6339	0.7784	0.8124	0.8812	0.8609	$\sigma^2=1e0$ NMF_L0, $\mathbf{x}_{\max}$ Err= 2.7×10^{-7}
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.8974	0.9370	0.7808	0.9816 x	0.6956	0.9844	0.9684 x	0.9869	0.9194	0.9398	$\sigma^2=1e0$ NMU, $\mathbf{x}_{\max}$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.8792	0.9370	0.9160	0.9816 x	0.6994	0.9844	0.9684 x	0.9869	0.9194	0.9398	$\sigma^2=1e0$ NMU, $\mathbf{x}_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.7593	0.9253	0.6862	0.9753 x	0.5761	0.9732	0.9599 x	0.9954	0.8130	0.9684	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.7222	0.9253	0.6862	0.9753	0.5621	0.9732	0.9599	0.9954	0.8130	0.9684	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$

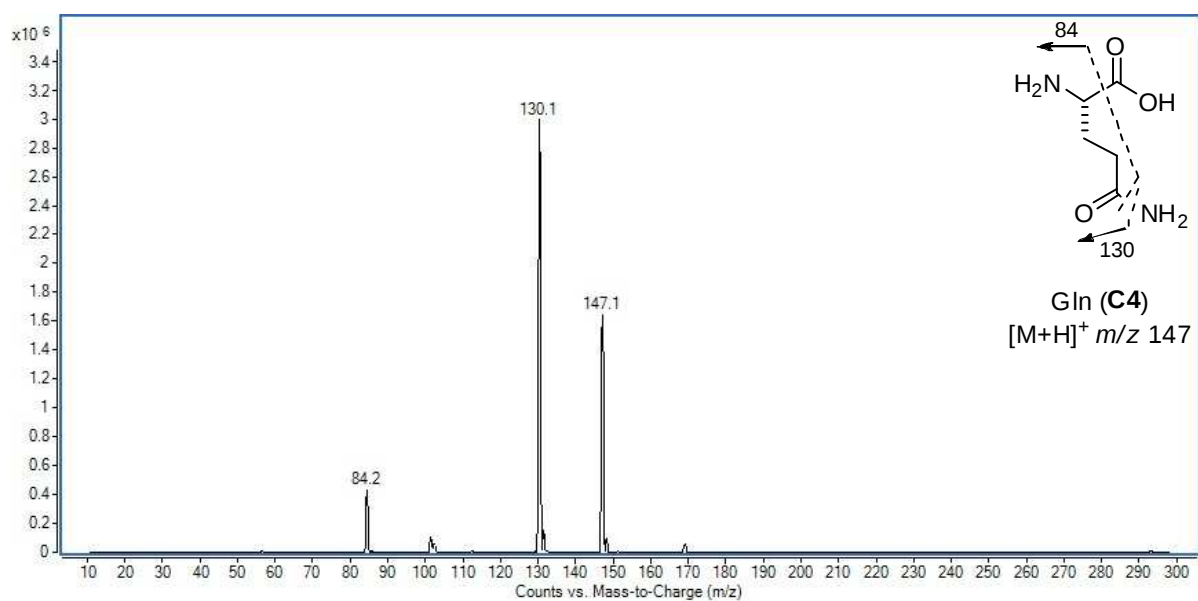
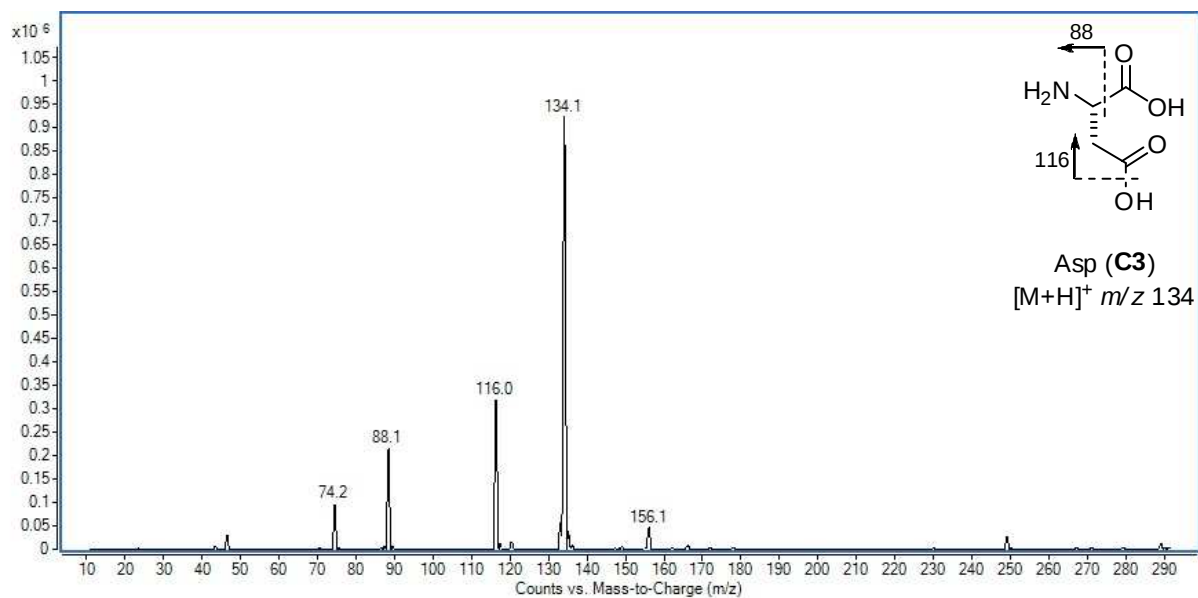
											Denoising hard $\tau=1e-7$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.6249	0.8867	0.8983	0.9052 x	0.7466	0.9255	0.9448 x	0.9069	0.9163	0.9423	$\sigma^2=1e-1$ NMU, $\mathbf{x}_{\max}$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.6970	0.8867	0.9107	0.9052 x	0.6604	0.9255	0.9448 x	0.9069	0.9163	0.9423	$\sigma^2=1e-1$ NMU, $\mathbf{x}_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.9349	0.9596	0.7721	0.9559 x	0.6859	0.9435	0.9449 x	0.9895	0.9142	0.9553	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.8113	0.9596	0.6537	0.9559 x	0.5736	0.9435	0.9449 x	0.9895	0.9142	0.9553	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.9200	0.9218	0.7584	0.9757 x	0.4975	0.7507	0.9608 x	0.9953	0.8295	0.9158	$\sigma^2=1e1$ NMU, $\mathbf{x}_{\max}$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.8577	0.9218	0.7821	0.9757 x	0.5818	0.7990	0.9608 x	0.9953	0.8295	0.9158	$\sigma^2=1e1$ NMU, $\mathbf{x}_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.7083	0.9527	0.6815	0.9746	0.6144	0.9804	0.9584	0.9952	0.8815	0.9894	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_1$ to $\mathbf{X}_5$	0.8143	0.9527	0.6786	0.9746 x	0.6018	0.9804	0.9584 x	0.9952	0.8815	0.9894	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,3,4,5)}$	0.8681	0.8484	0.9145	0.9008 x	0.4864	0.9160	0.8993 x	0.9826	0.8746	0.9359	$\sigma^2=1e0$ NMU, $\mathbf{x}_{\max}$
$\mathbf{X}_{(1,3,4,5)}$	0.8486	0.8484	0.9142	0.9008 x	0.6107	0.9160	0.8993 x	0.9826	0.8746	0.9359	$\sigma^2=1e0$ NMU, $\mathbf{x}_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,3,4,5)}$	0.7637	0.9287	0.7860	0.9713 x	0.4970	0.9546	0.9634 x	0.9938	0.8811	0.8685	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_{(1,3,4,5)}$	0.6711	0.9287	0.8485	0.9713	0.6126	0.9546	0.9634	0.9938	0.8811	0.8685	$\sigma^2=1e0$

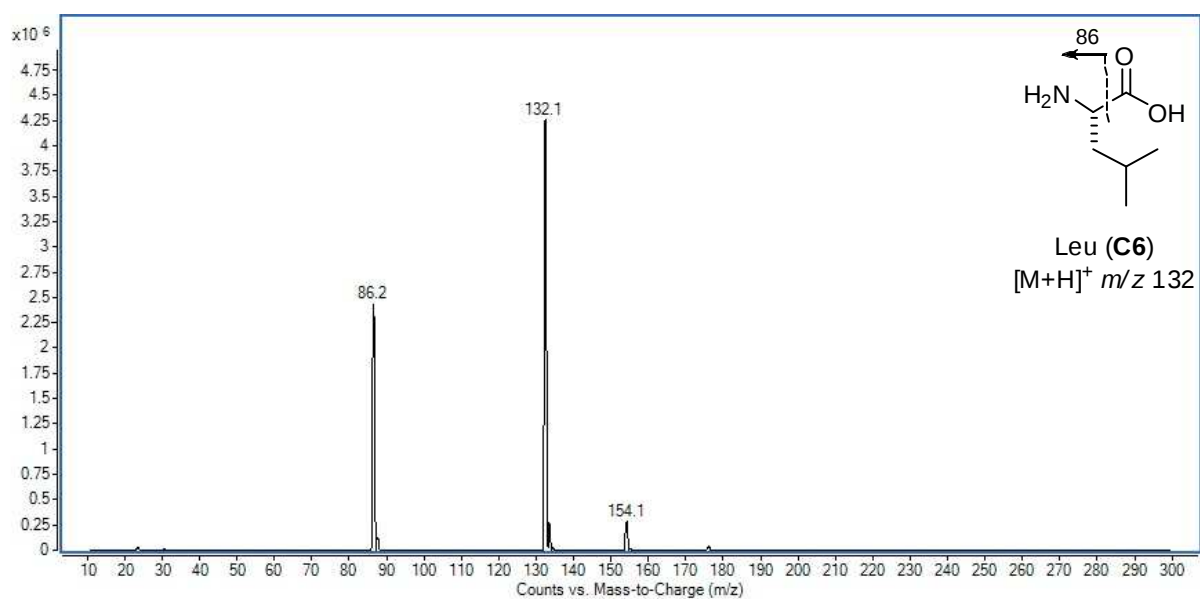
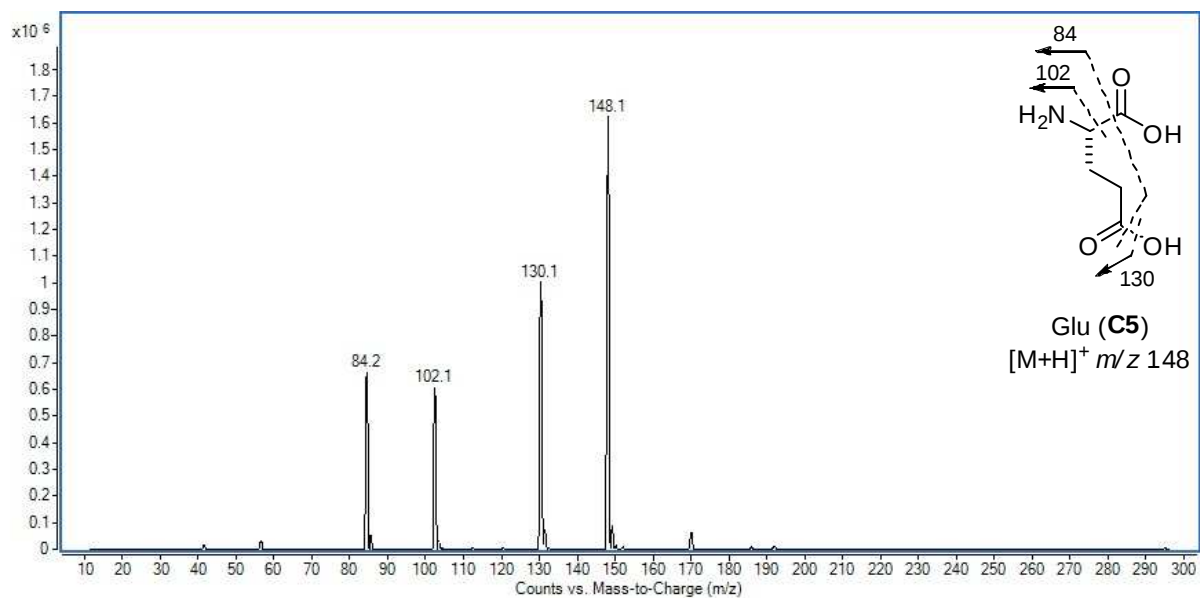
				x							NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,3,4,5)}$	0.6975	0.9446	0.8921	0.9057 x	0.6080	0.8391	0.9424 x	0.8997	0.9209	0.7865	$\sigma^2=1e-1$ NMU, $x_{\max}$
$\mathbf{X}_{(1,3,4,5)}$	0.7005	0.9446	0.8919	0.9057 x	0.6570	0.8391	0.9424	0.8997	0.9209	0.7865	$\sigma^2=1e-1$ NMU, $x_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,3,4,5)}$	0.7928	0.8822	0.8927	0.8965 x	0.6274	0.8860	0.8950 x	0.9822	0.9080	0.9262	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_{(1,3,4,5)}$	0.8348	0.8822	0.8926	0.8965 x	0.6275	0.8860	0.8950 x	0.9822	0.9080	0.9262	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,3,4,5)}$	0.6685	0.8787	0.8537	0.9713 x	0.6439	0.9492	0.9633	0.9938	0.8833	0.9556	$\sigma^2=1e1$ NMU, $x_{\max}$
$\mathbf{X}_{(1,3,4,5)}$	0.6432	0.8787	0.8395	0.9713 x	0.6169	0.9492	0.9633 x	0.9938	0.8833	0.9556	$\sigma^2=1e1$ NMU, $x_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,3,4,5)}$	0.7941	0.9357	0.7805	0.9712 x	0.5039	0.9044	0.9630 x	0.9936	0.9363	0.8031	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_{(1,3,4,5)}$	0.7997	0.9357	0.7805	0.9712 x	0.5515	0.9044	0.9630 x	0.9936	0.9363	0.8031	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,2,3)}$	0.6017	0.8855	0.7136	0.7474 x	0.5781	0.7958	0.7830	0.9671	0.9413	0.9826	$\sigma^2=1e0$ NMU, $x_{\max}$
$\mathbf{X}_{(1,2,3)}$	0.7194	0.8855	0.6495	0.7474 x	0.5863	0.7958	0.7830 x	0.9671	0.9413	0.9826	$\sigma^2=1e0$ NMU, $x_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(1,2,3)}$	0.6309	0.8224	0.5365	0.8612 x	0.7049	0.6451	0.9231 x	0.9686	0.8400	0.9015	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$

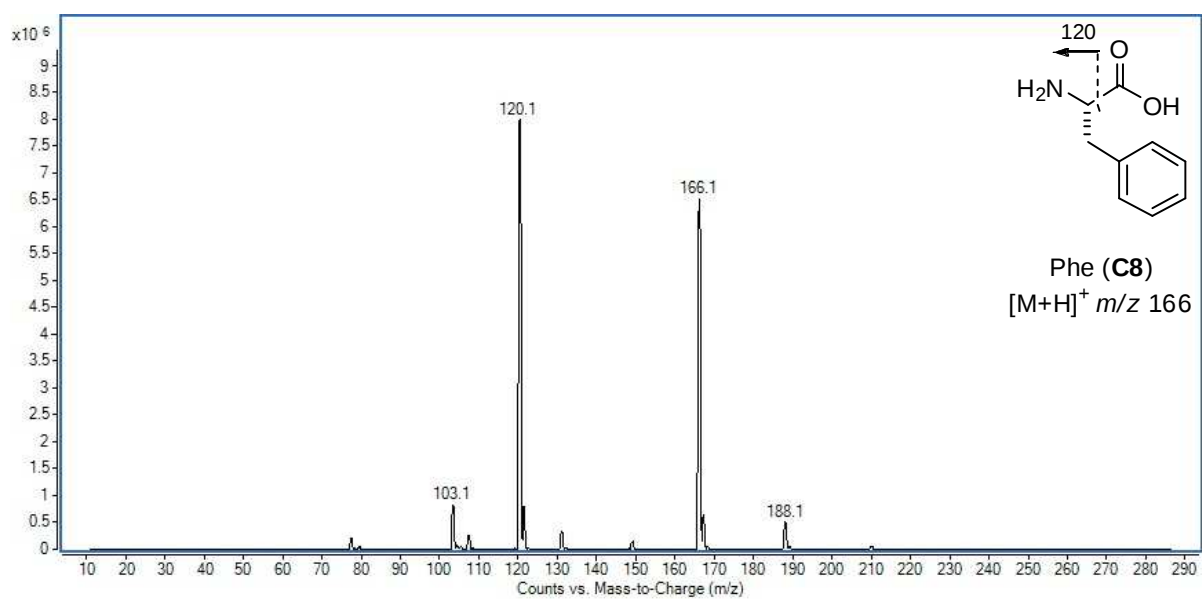
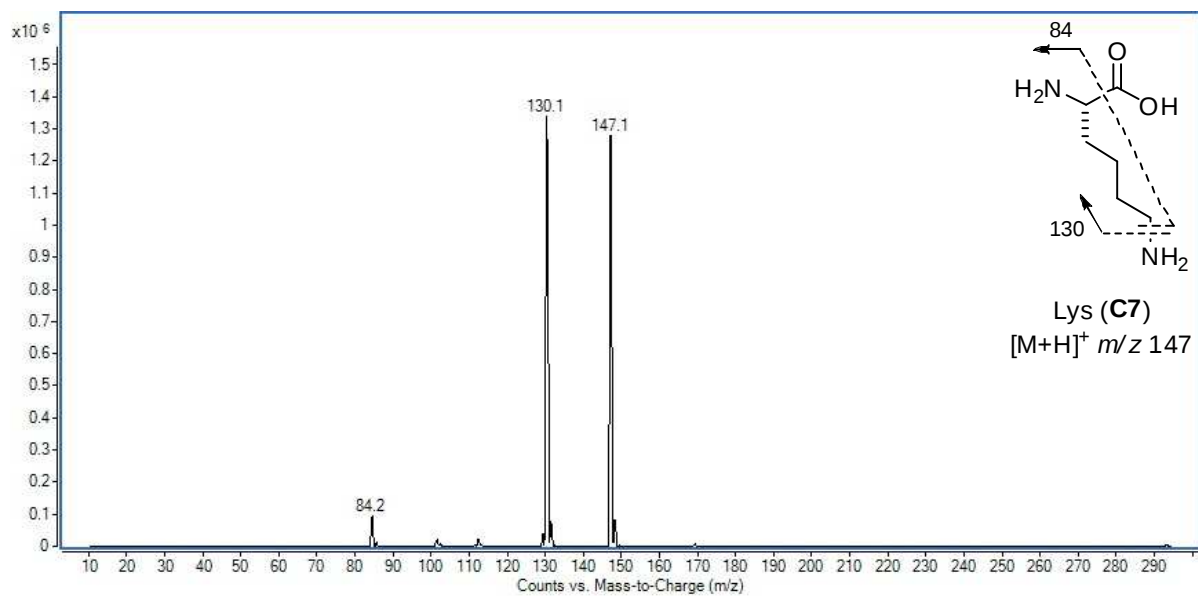
$X_{(1,2,3)}$	0.5509	0.8224	0.4973	0.8612 x	0.7050	0.7498	0.9231 x	0.9686	0.8400	0.9015	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$X_{(1,2,3)}$	0.6191	0.9029	0.7897	0.7435 x	0.6542	0.7157	0.7525 x	0.8516	0.8639	0.8379	$\sigma^2=1e-1$ NMU, $x_{\max}$
$X_{(1,2,3)}$	0.6492	0.9029	0.6715	0.7435 x	0.6539	0.7157	0.7525 x	0.8516	0.8639	0.8379	$\sigma^2=1e-1$ NMU, $x_{\max}$ Denoising hard $\tau=1e-7$
$X_{(1,2,3)}$	0.5934	0.8308	0.5123	0.8180 x	0.5791	0.7640	0.8597 x	0.9571	0.9267	0.9825	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$
$X_{(1,2,3)}$	0.5449	0.8308	0.5434	0.8180 x	0.5751	0.7640	0.8597 x	0.9572	0.9267	0.9825	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$X_{(1,2,3)}$	0.7234	0.8487	0.5026	0.8990 x	0.5603	0.8382	0.9217 x	0.9685	0.8288	0.8713	$\sigma^2=1e1$ NMU, $x_{\max}$
$X_{(1,2,3)}$	0.6652	0.8487	0.5449	0.8990 x	0.5932	0.8387	0.9217 x	0.9685	0.8288	0.8713	$\sigma^2=1e1$ NMU, $x_{\max}$ Denoising hard $\tau=1e-7$
$X_{(1,2,3)}$	0.6157	0.8866	0.4924	0.9547 x	0.6303	0.7579	0.9684 x	0.9676	0.9070	0.9905	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$
$X_{(1,2,3)}$	0.5855	0.8866	0.4924	0.9547 x	0.6303	0.7578	0.9684 x	0.9676	0.9070	0.9905	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$X_{(3,4)}$	0.5554	0.9107	0.6707	0.7453 x	0.6418	0.9166	0.7916 x	0.8869	0.8169	0.7617	$\sigma^2=1e0$ NMU, $x_{\max}$
$X_{(3,4)}$	0.4517	0.9107	0.6707	0.7453 x	0.6418	0.9166	0.7916 x	0.8869	0.8169	0.7621	$\sigma^2=1e0$ NMU, $x_{\max}$ Denoising hard $\tau=1e-7$

$\mathbf{X}_{(3,4)}$	0.4873	0.7182	0.7689	0.7984 x	0.4657	0.9578	0.8249 x	0.9209	0.8092	0.7197	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_{(3,4)}$	0.5377	0.7179	0.7689	0.7984 x	0.4457	0.9578	0.8249 x	0.9209	0.8092	0.7197	$\sigma^2=1e0$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(3,4)}$	0.5309	0.7552	0.7321	0.8030 x	0.6653	0.8359	0.8349 x	0.9235	0.8051	0.7011	$\sigma^2=1e-1$ NMU, $\mathbf{x}_{\max}$
$\mathbf{X}_{(3,4)}$	0.5631	0.7552	0.7613	0.8030 x	0.6653	0.8359	0.8349 x	0.9235	0.8051	0.7011	$\sigma^2=1e-1$ NMU, $\mathbf{x}_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(3,4)}$	0.5104	0.8479	0.6889	0.7308 x	0.5734	0.9386	0.7621 x	0.9318	0.7998	0.8085	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_{(3,4)}$	0.6232	0.8479	0.6889	0.7308 x	0.6461	0.9386	0.7621 x	0.9318	0.7998	0.8085	$\sigma^2=1e-1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(3,4)}$	0.5874	0.6050	0.6381	0.7070	0.6301	0.9400	0.7626	0.9448	0.7206	0.6451	Gauss, $\sigma^2=1e1$ NMU, $\mathbf{x}_{\max}$
$\mathbf{X}_{(3,4)}$	0.4638	0.6050	0.7164	0.7070	0.6329	0.9400	0.7626	0.9448	0.7206	0.6451	Gauss, $\sigma^2=1e1$ NMU, $\mathbf{x}_{\max}$ Denoising hard $\tau=1e-7$
$\mathbf{X}_{(3,4)}$	0.4973	0.7649	0.5762	0.7217 x	0.5866	0.9333	0.7067 x	0.9506	0.6622	0.6451	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$
$\mathbf{X}_{(3,4)}$	0.4692	0.6605	0.5177	0.7217 x	0.5879	0.9333	0.7067	0.9506	0.6622	0.6594	$\sigma^2=1e1$ NMU, $\ \mathbf{x}\ _1$ Denoising hard $\tau=1e-7$









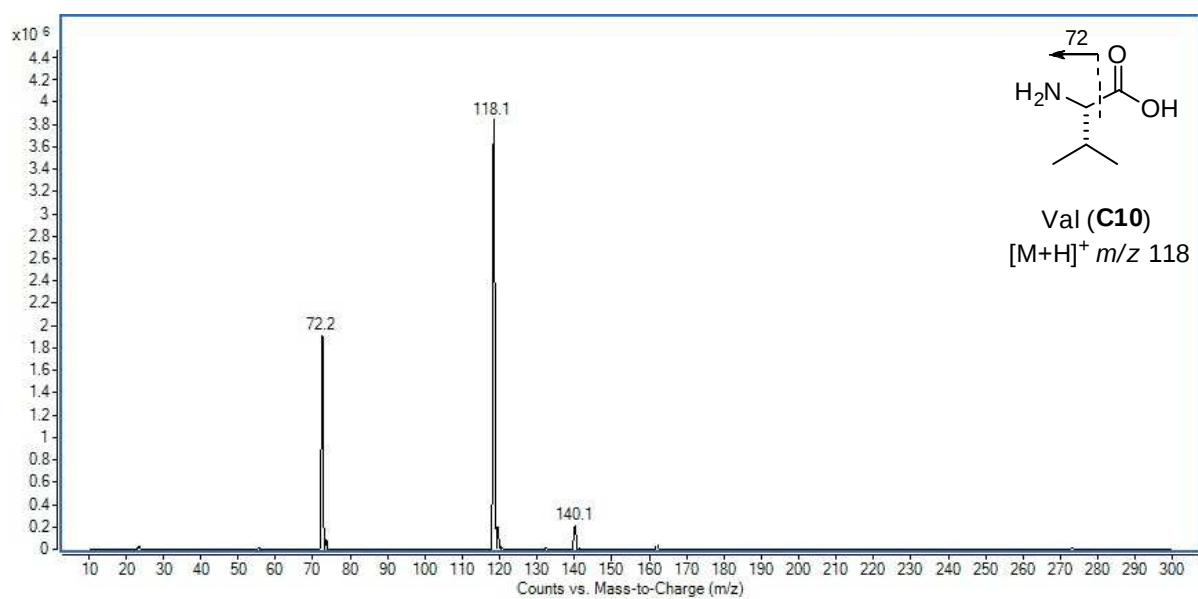
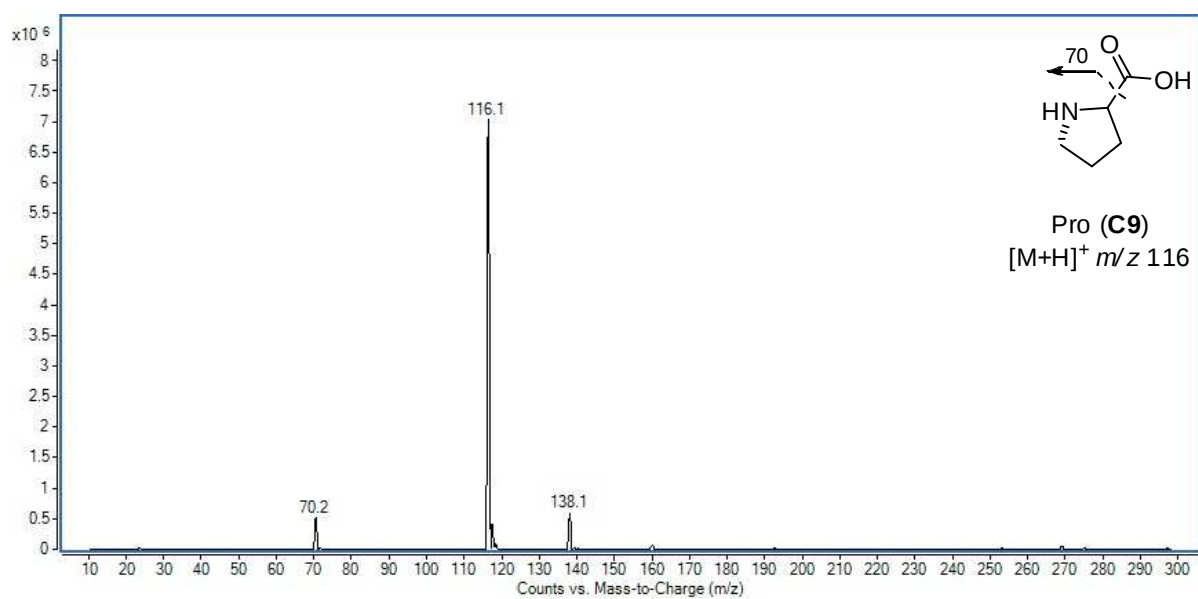
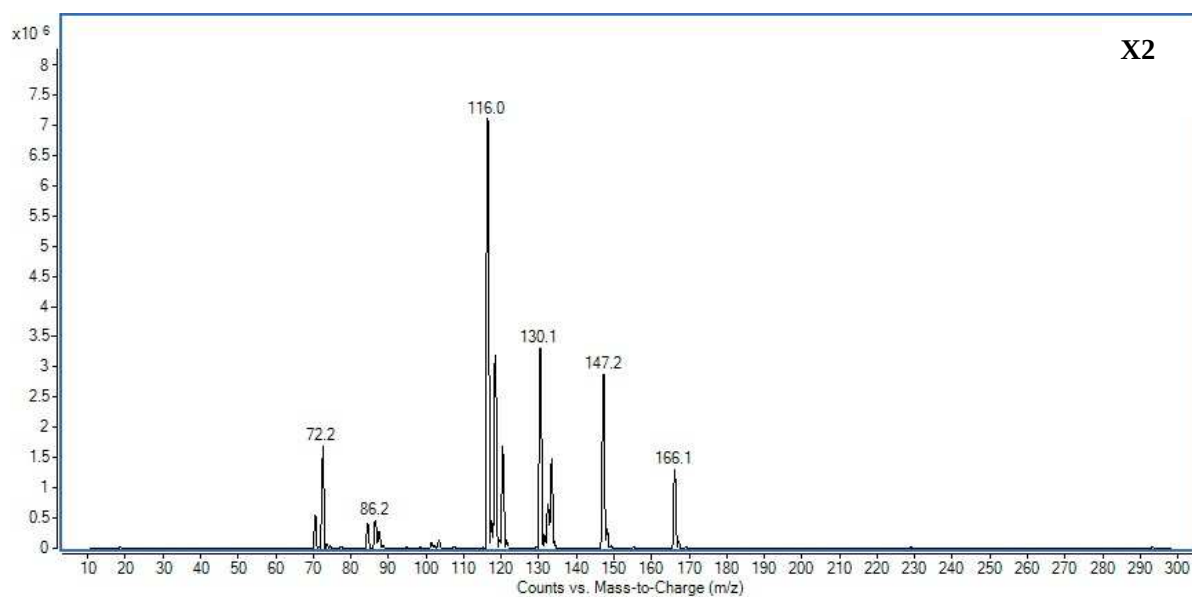
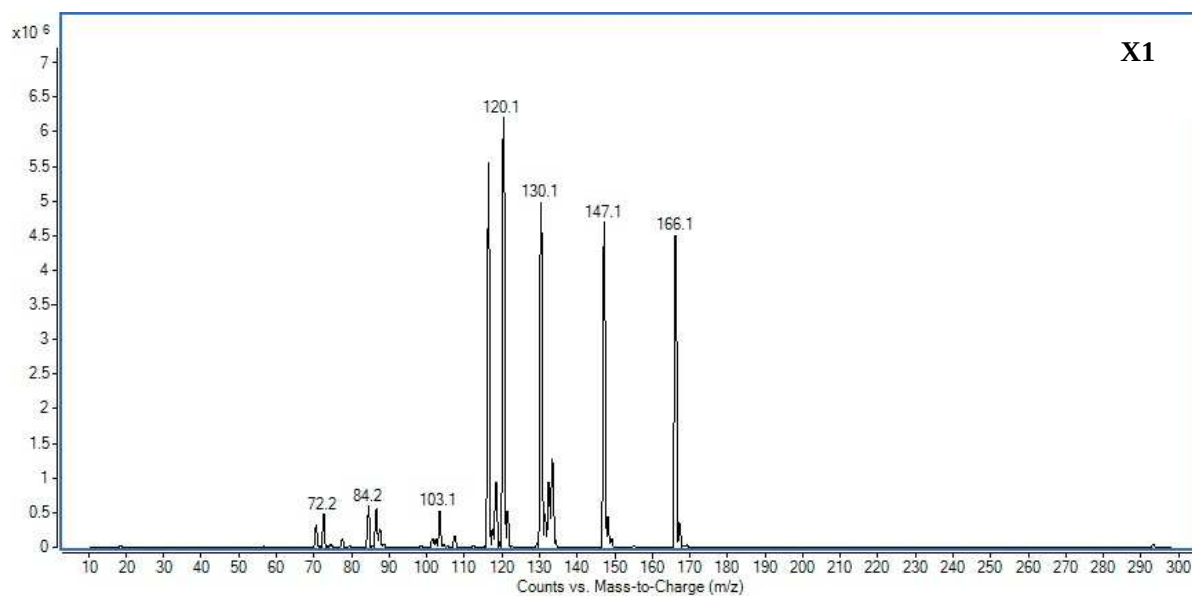
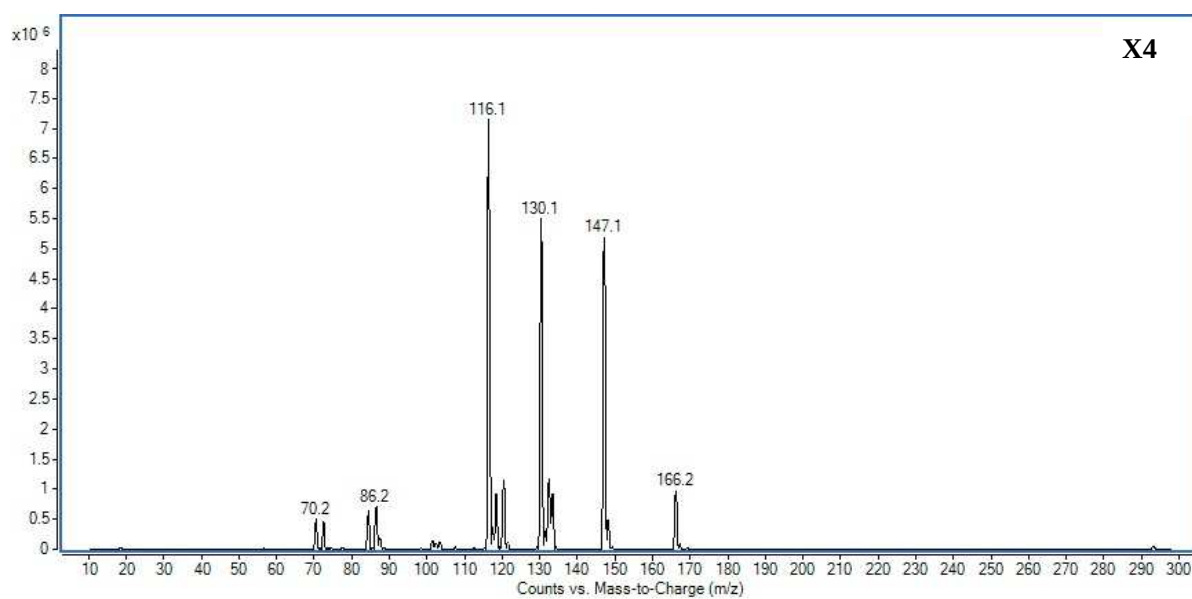
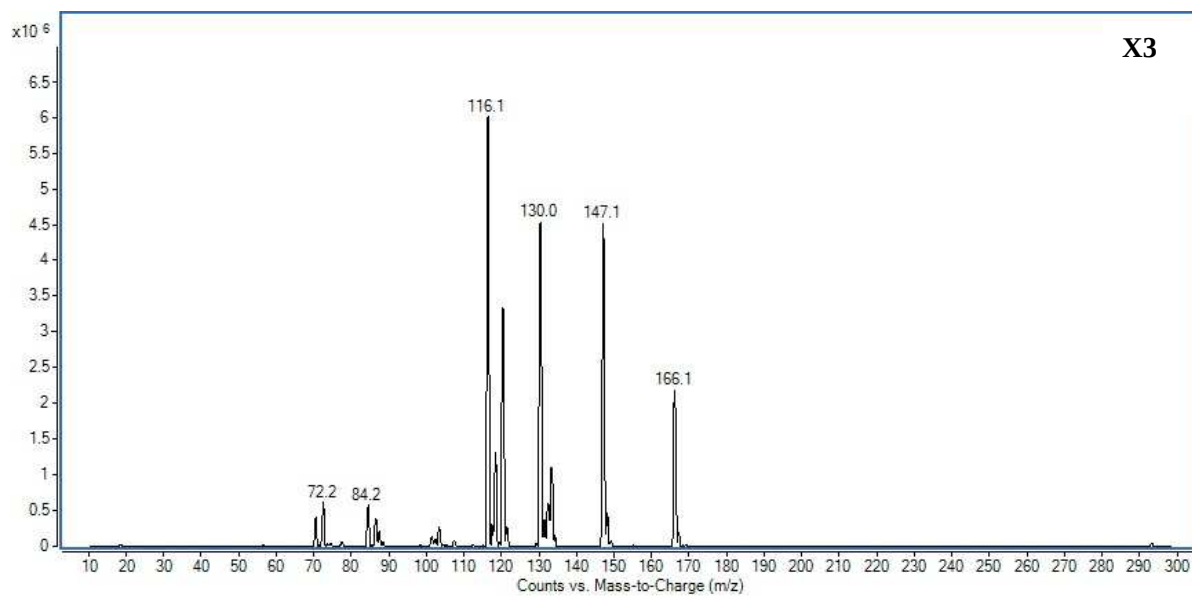


Figure S-1.





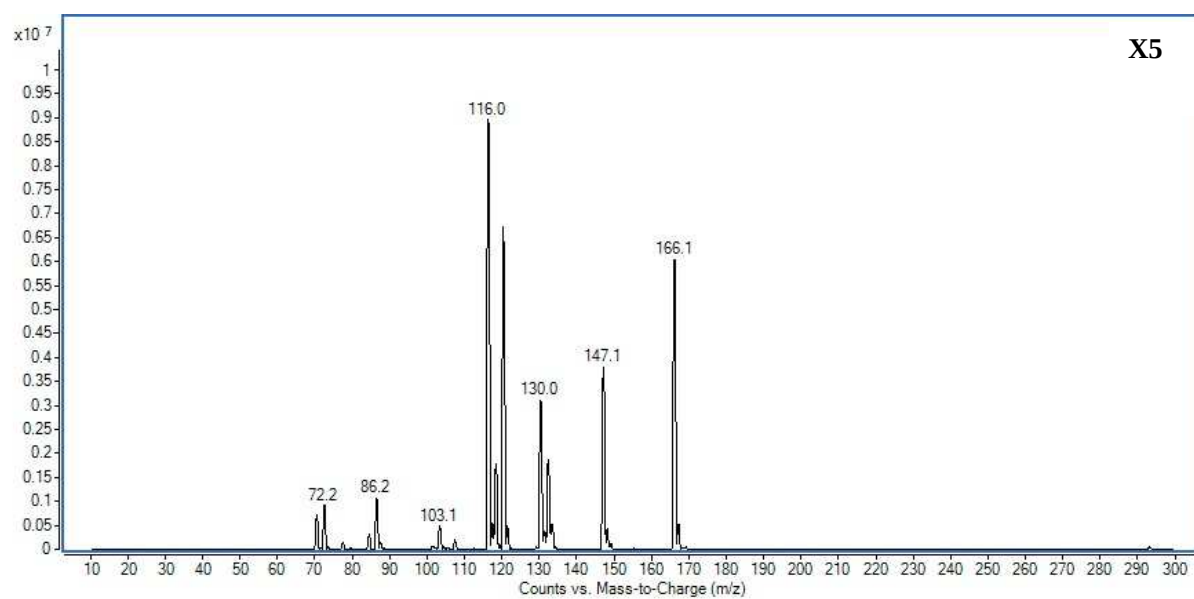
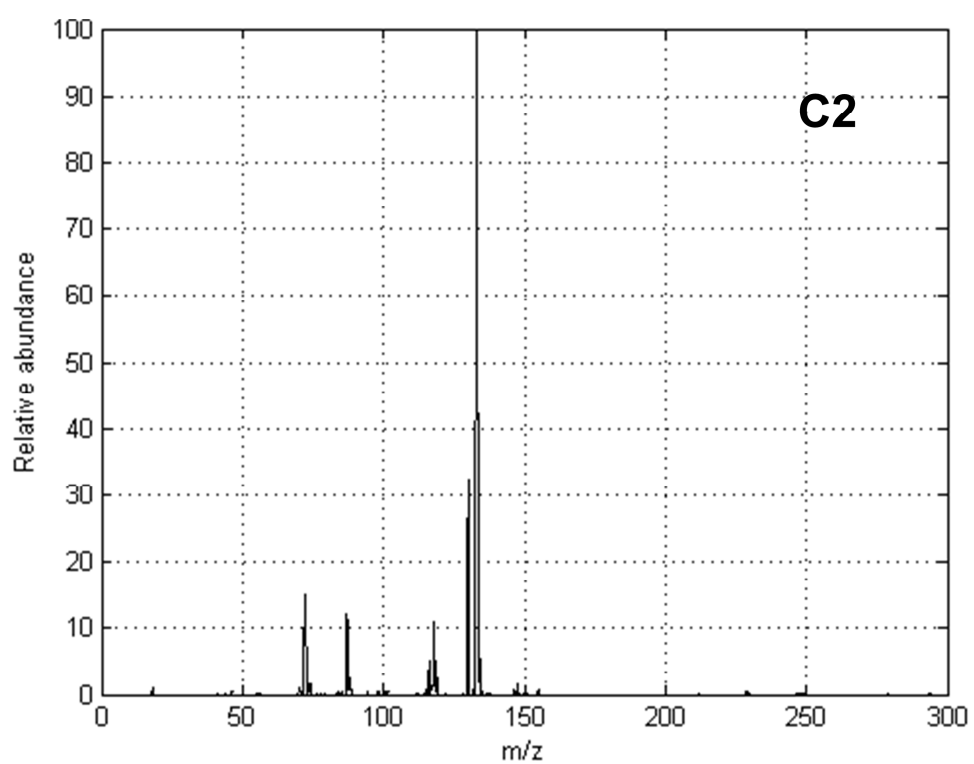
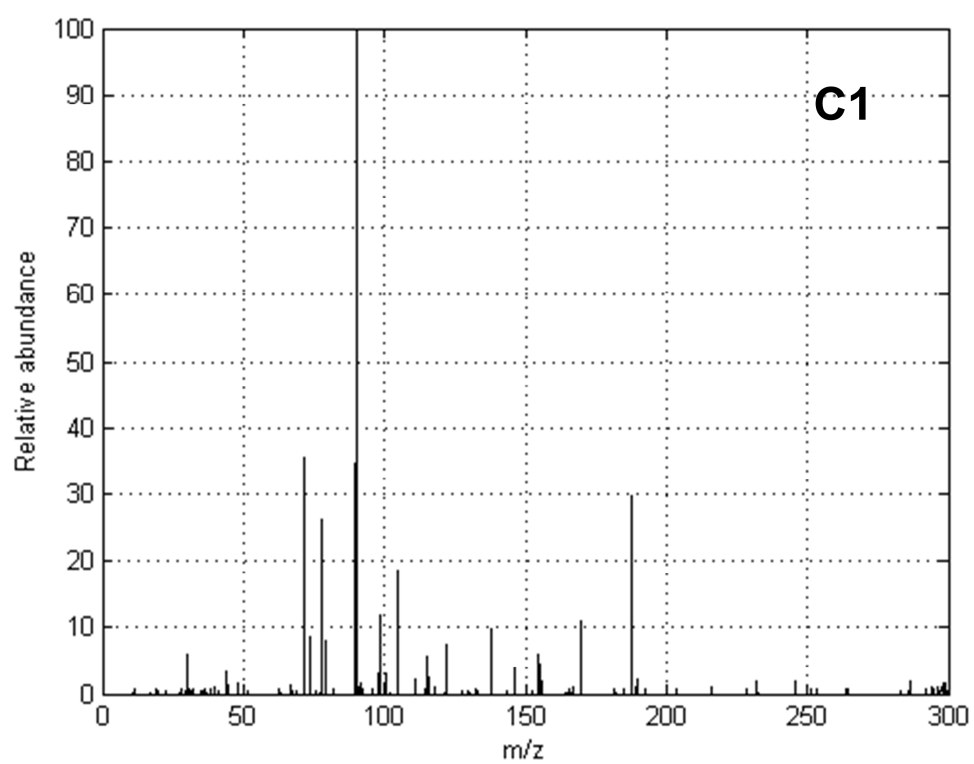
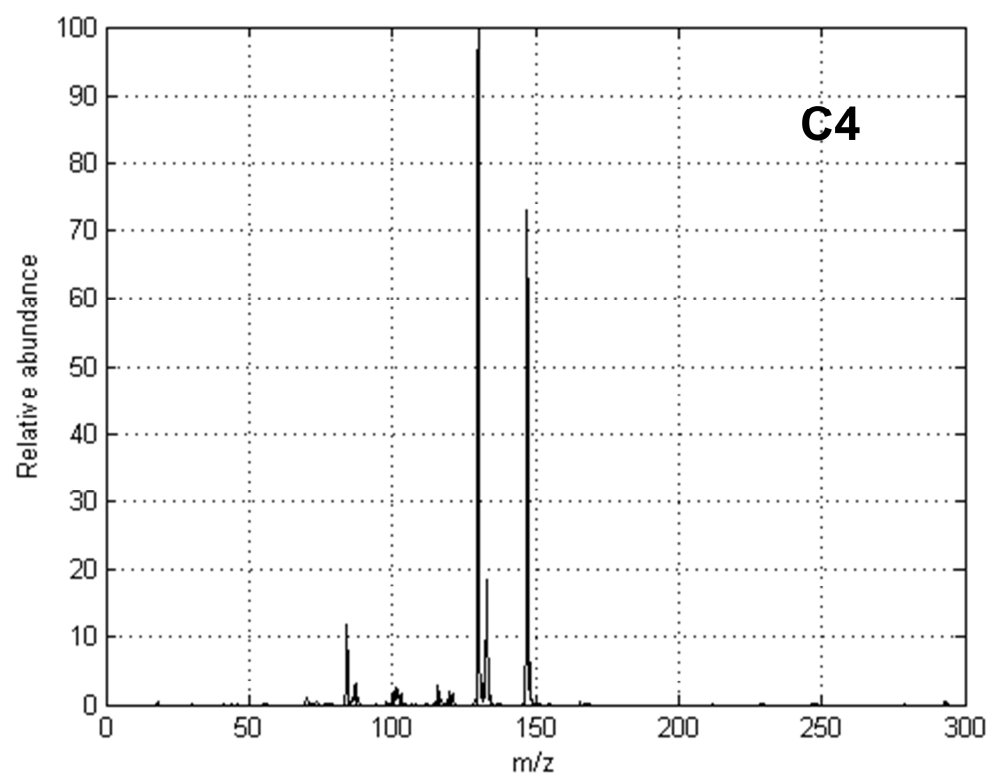
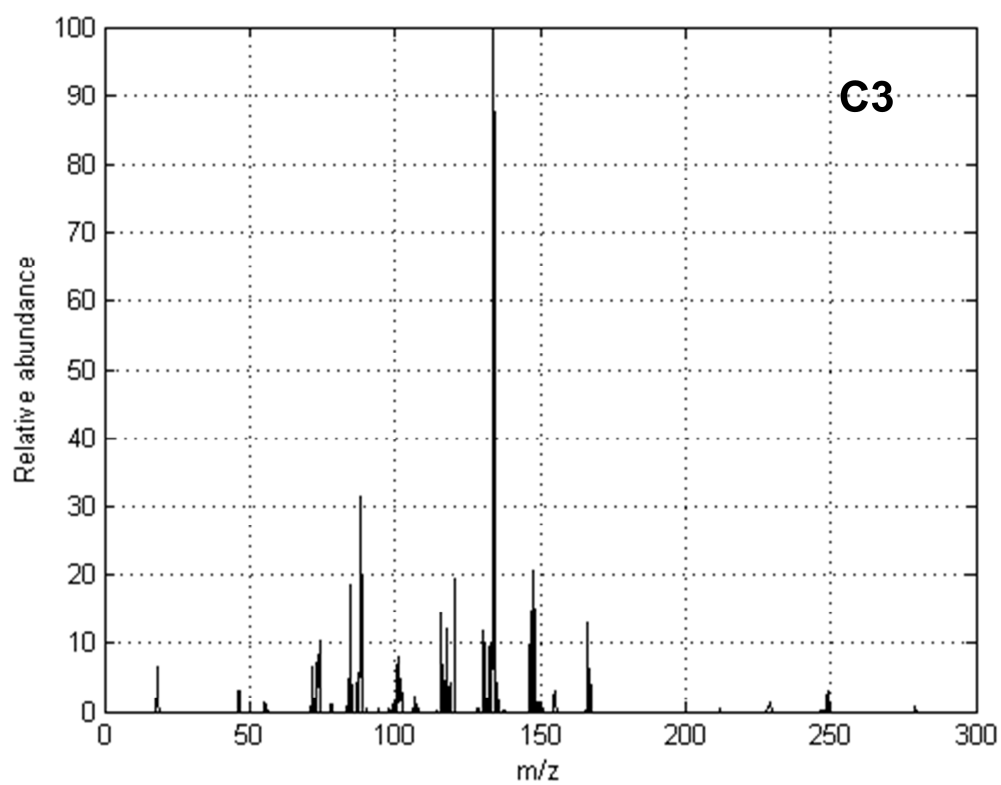
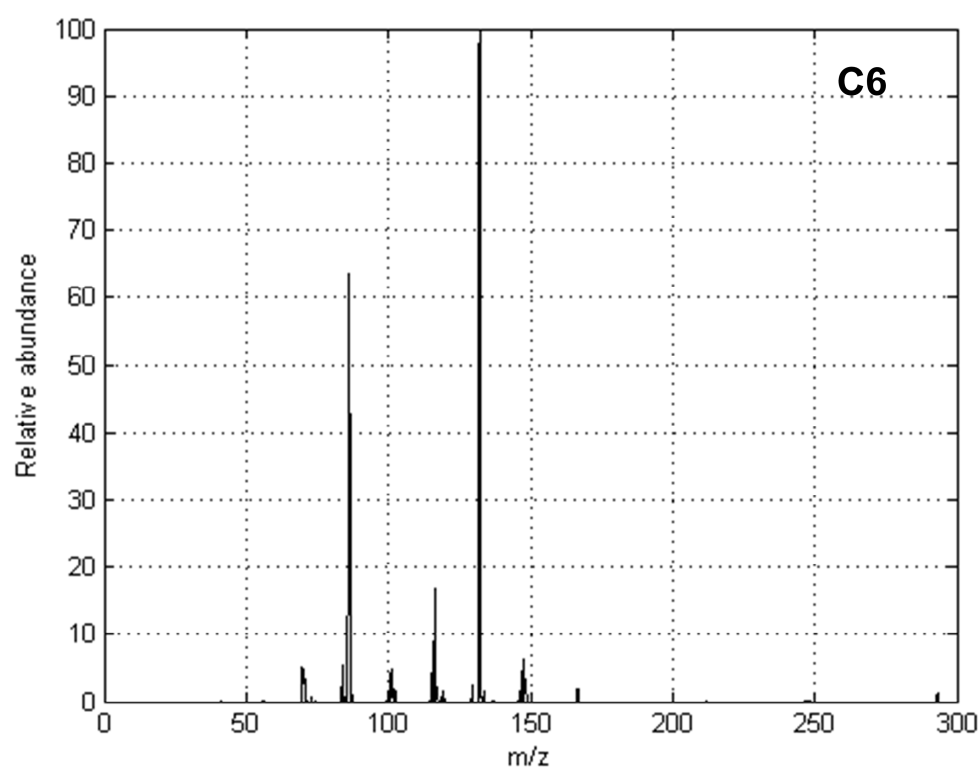
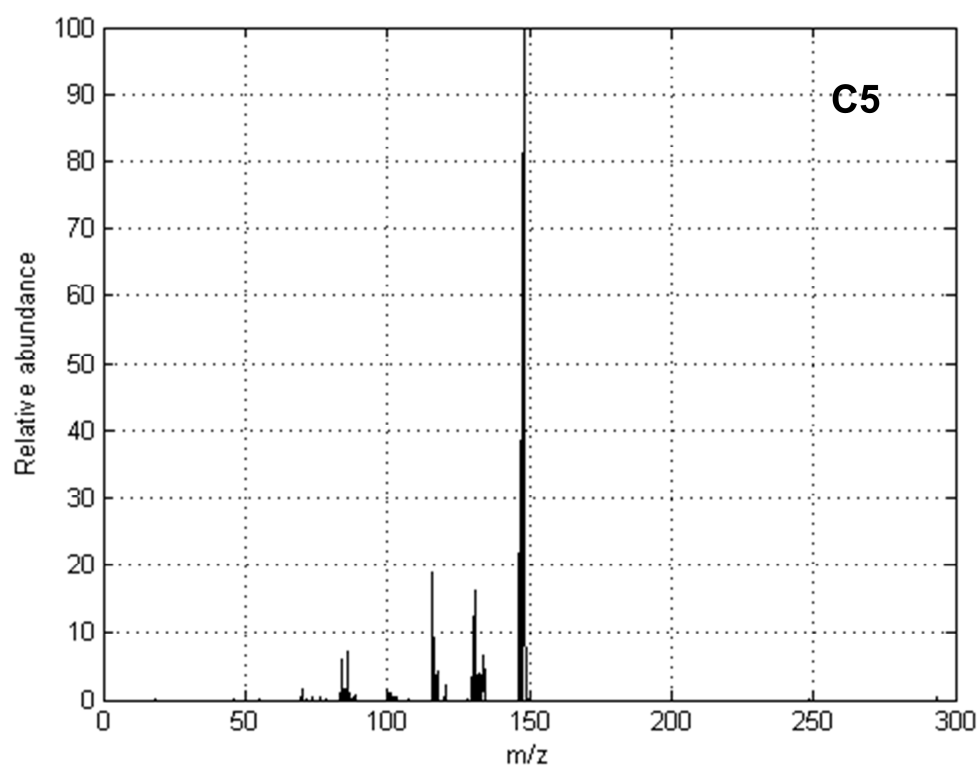
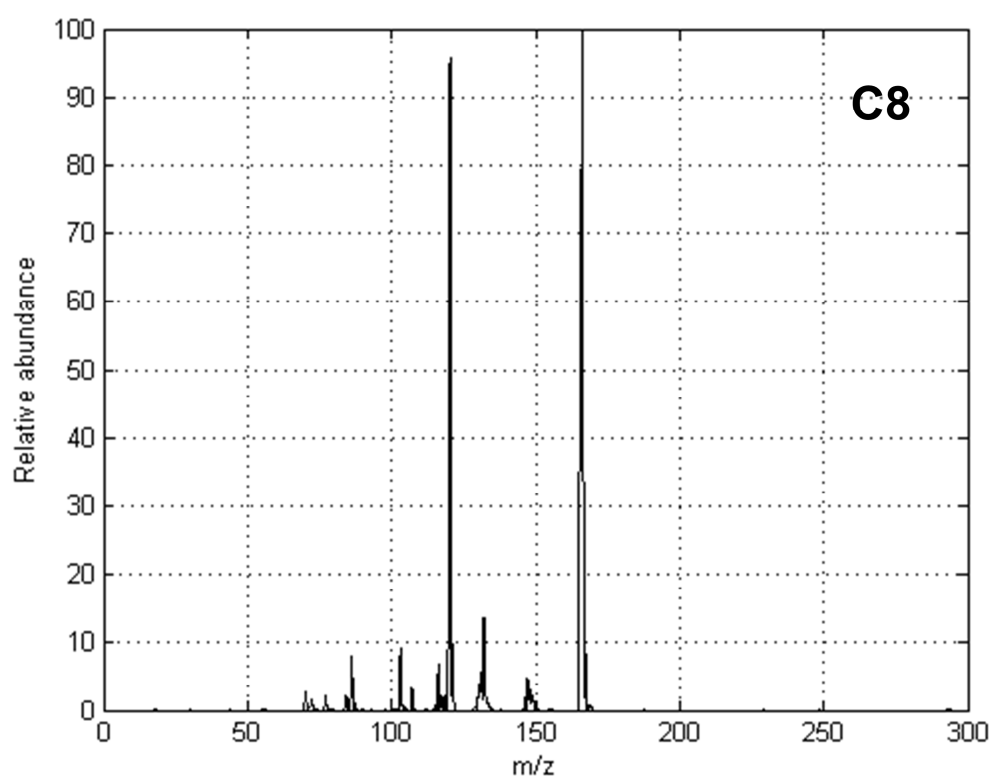
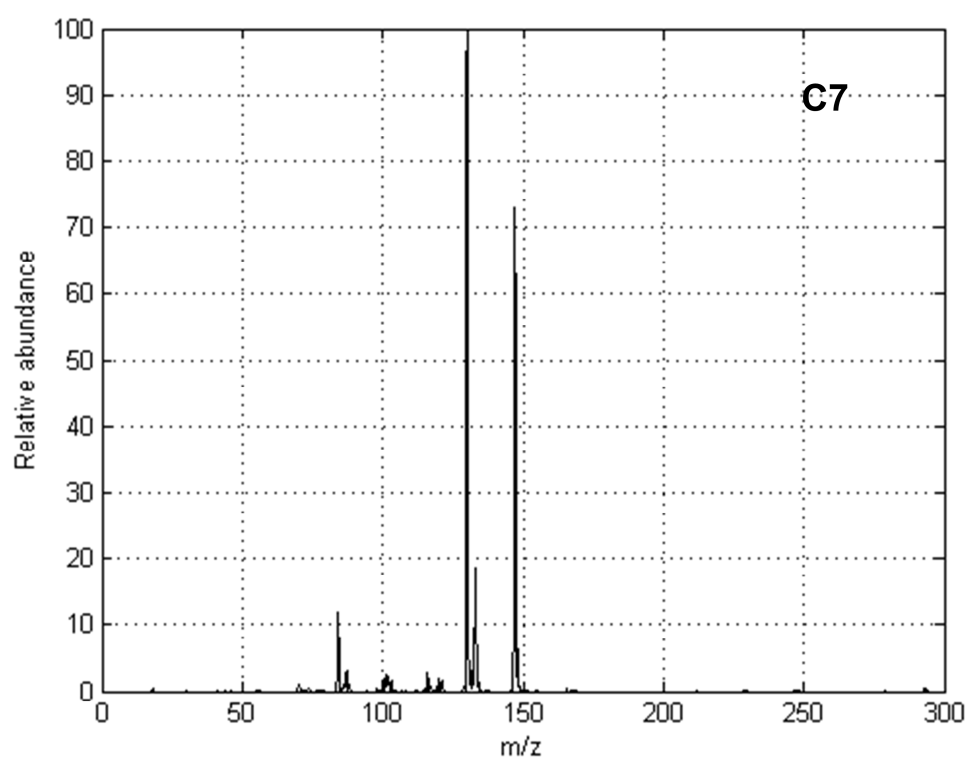


Figure S-2.









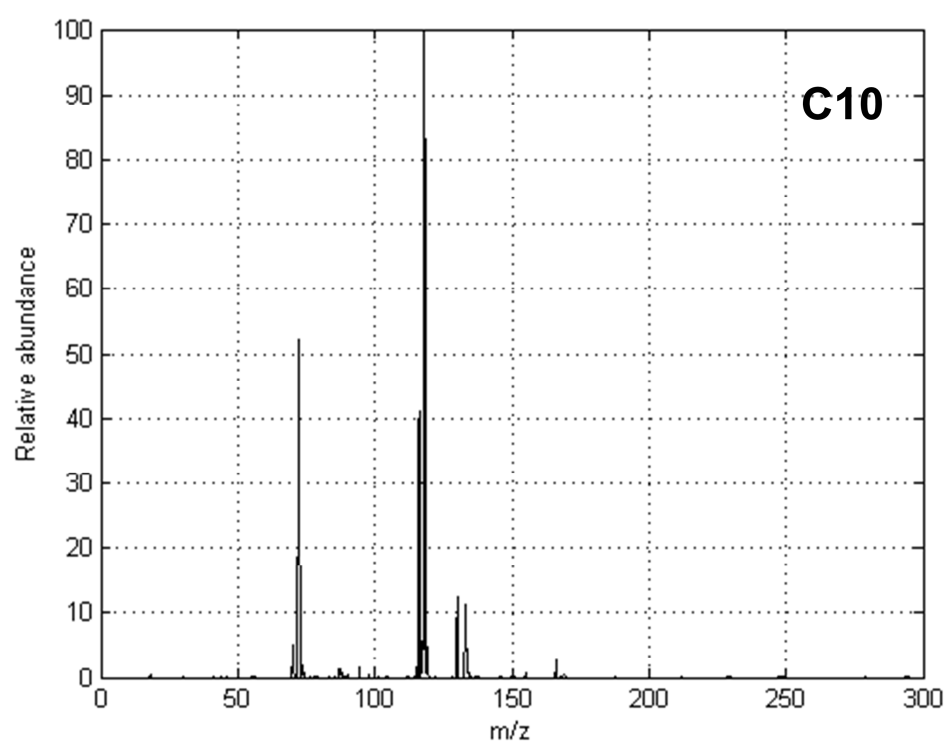
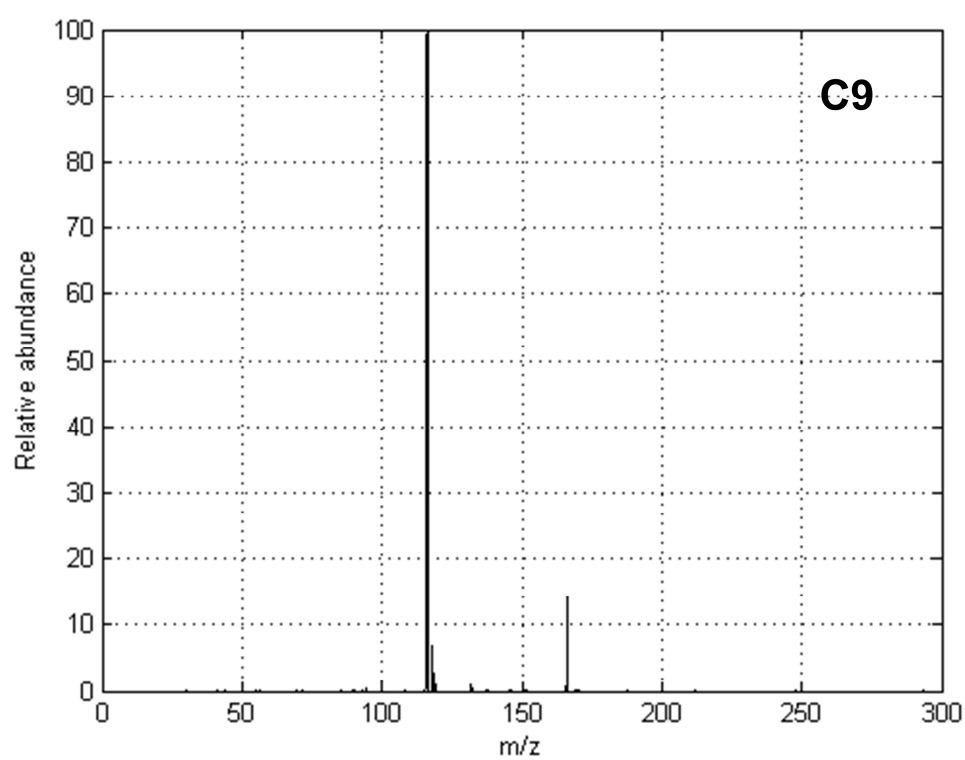


Figure S-3.