Supporting Information

Structural mass spectrometry of tissue extracts to distinguish cancerous and non-cancerous breast diseases

Kelly M. Hines,[1,2,3] Billy R. Ballard,[4] Dana R. Marshall,[3, 4] and John A. McLean[1,2,3*]*

[1]Department of Chemistry, Vanderbilt University, Nashville, TN 37235; [2]Vanderbilt Institute of Chemical Biology, Vanderbilt University, Nashville, TN 37232; [3]Vanderbilt Institute for Integrative Biosystems Research and Education, Vanderbilt University, Nashville, TN 37235; [4]Department of Pathology, Anatomy and Cell Biology, Meharry Medical College, Nashville, TN 37208.

* Authors to whom correspondence should be addressed. For D. R. Marshall: tel, 615.327.6549; fax, 615.327.6409; e-mail, dmarshall@mmc.edu. For J. A. McLean: tel, 615.322.1195; fax, 615.343.1234; e-mail, john.a.mclean@vanderbilt.edu.

Abstract: Breast cancer is well-known to broadly impact cellular metabolism and aberrant metabolism in breast cancer tumors has been widely studied by both targeted and untargeted analyses to characterize the affected metabolic pathways. In this work, we utilize ultra-performance liquid chromatography (UPLC) in tandem with ion mobility-mass spectrometry (IM-MS), which provides chromatographic, structural, and mass information, to characterize the aberrant metabolism associated with breast diseases such as cancer. In a double-blind analysis of matched control (n=3) and disease tissues (n=3), tissues were homogenized, polar metabolites were extracted, and the extracts were characterized by UPLC-IM-MS/MS. Principle component analysis revealed a strong separation between disease tissues, with one diseased tissue clustering with the control tissues along PC1 and two others separated along PC2. Using post-ion mobility MS/MS spectra acquired by data-independent acquisition, the features giving rise to the observed grouping were determined to be biomolecules associated with aggressive breast cancer tumors, including glutathione, oxidized glutathione, thymosins β4 and β10, and choline-containing species. Pathology reports revealed the outlier of the disease tissues to be a benign fibroadenoma, whereas the other disease tissues represented highly metabolic benign and aggressive tumors. This IM-MS-based workflow bridges the transition from untargeted metabolomic profiling to tentative identifications of key descriptive molecular features using data acquired in one analysis, with additional experiments performed only for validation. The ability to resolve cancerous and non-cancerous tissues at the biomolecular level demonstrates UPLC-IM-MS/MS as a robust and sensitive platform for metabolomic profiling of tissues.

Table S.1. Weights of Tissues used for Polar Metabolite Extraction

Sample Name	Tissue Weight (mg, wet)
1C	44.93
2D	50.57
3C	40.69
4D	48.99
5C	52.09
6D	48.80
Average	**47.68**
Standard Deviation	**± 3.81**

Method for Data Processing with XCMS

The R package can be downloaded from http://www.r-project.org/. After installing R, XCMS may be downloaded using the following code in an active R session:

```
source("http://bioconductor.org/biocLite.R")
biocLite("xcms", dep=T)
```

Additional details may be found on the XCMS Bioconductor website (http://www.bio conductor.org/packages/release/bioc/html/xcms.html). Prior to processing data with XCMS, .raw files must be converted to .mzXML using ProteoWizard (http://proteo wizard.sourceforge.net/) MSConvert and the "sortByScanTime" function.

XCMS Method

```
xset <-xcmsSet()
xset
xset <-group(xset)
xset2 <-retcor (xset, method="obiwarp", plottype="deviation")
xset2 <-group(xset2, bw=10)
xset3 <-fillPeaks(xset2)
xset3
reporttab <-diffreport(xset3, "Group_1", "Group_2", "File_Name", [Number of EICs, boxplots, to
        generate])
```

The aligned data in the diffreport was normalized such that all intensities within a sample summed to 10,000. The data was then transposed such that the sample and file names were in a column instead of a row, and this dataset was imported into Umetrics Extended Statistics for statistical analysis.

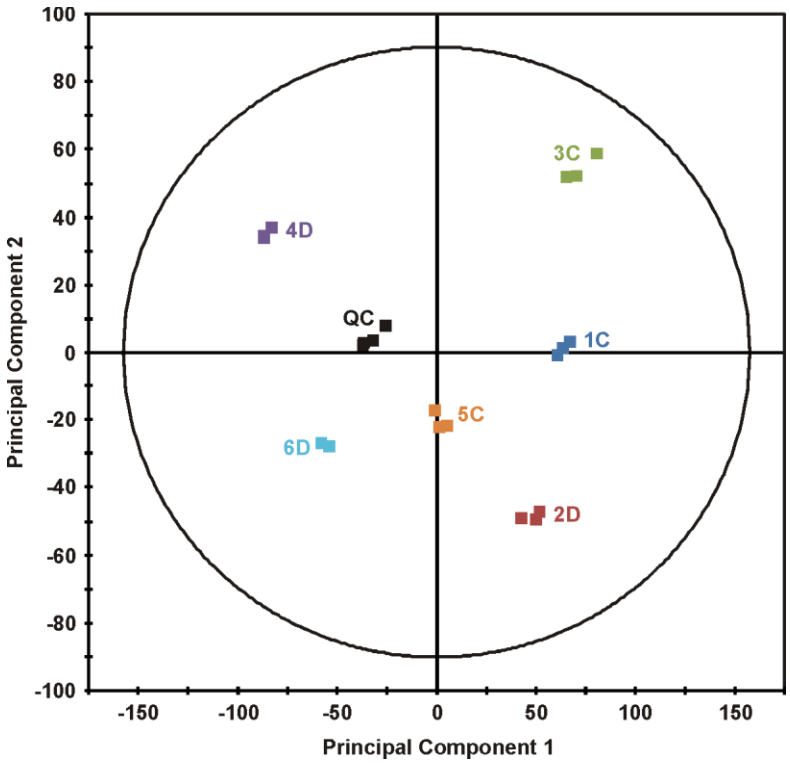

Figure S.1. Principal component analysis score plot for breast cancer tissue dataset including quality control (QC) samples for diagnostic purposes. QC samples group together well, indicating good reproducibility throughout the queue. QC samples cluster near the center of the PCA indicating they are representative of all samples.

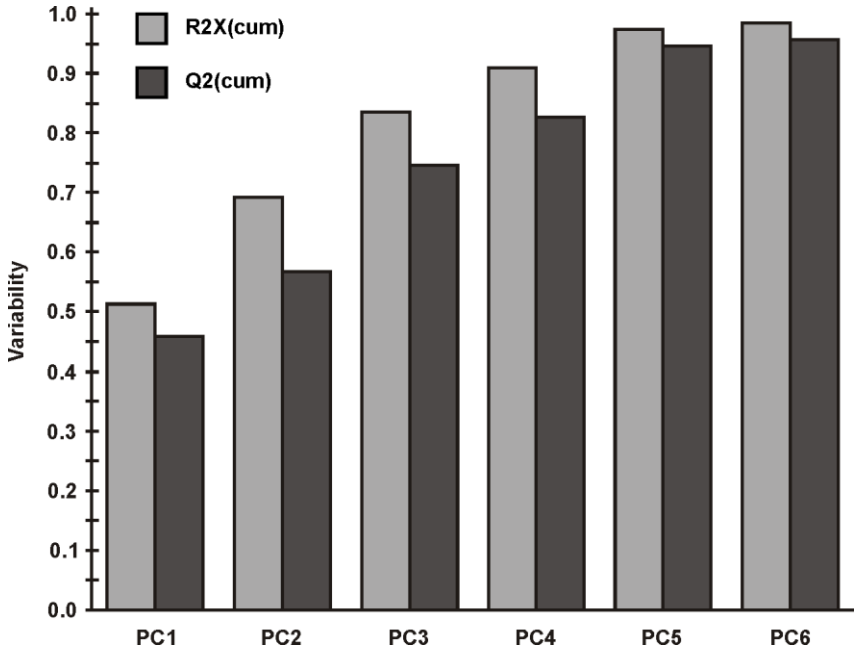

Figure S.2. Goodness of fit for PCA analysis of breast tissue extracts. Model parameters R2X (light grey) and Q2 (dark grey) describe the cumulative variability extracted by including each consecutive principal component (PC1-6). For example, including PC1 and PC2 extracts approximately 70% (R2X) of the variability in the dataset.

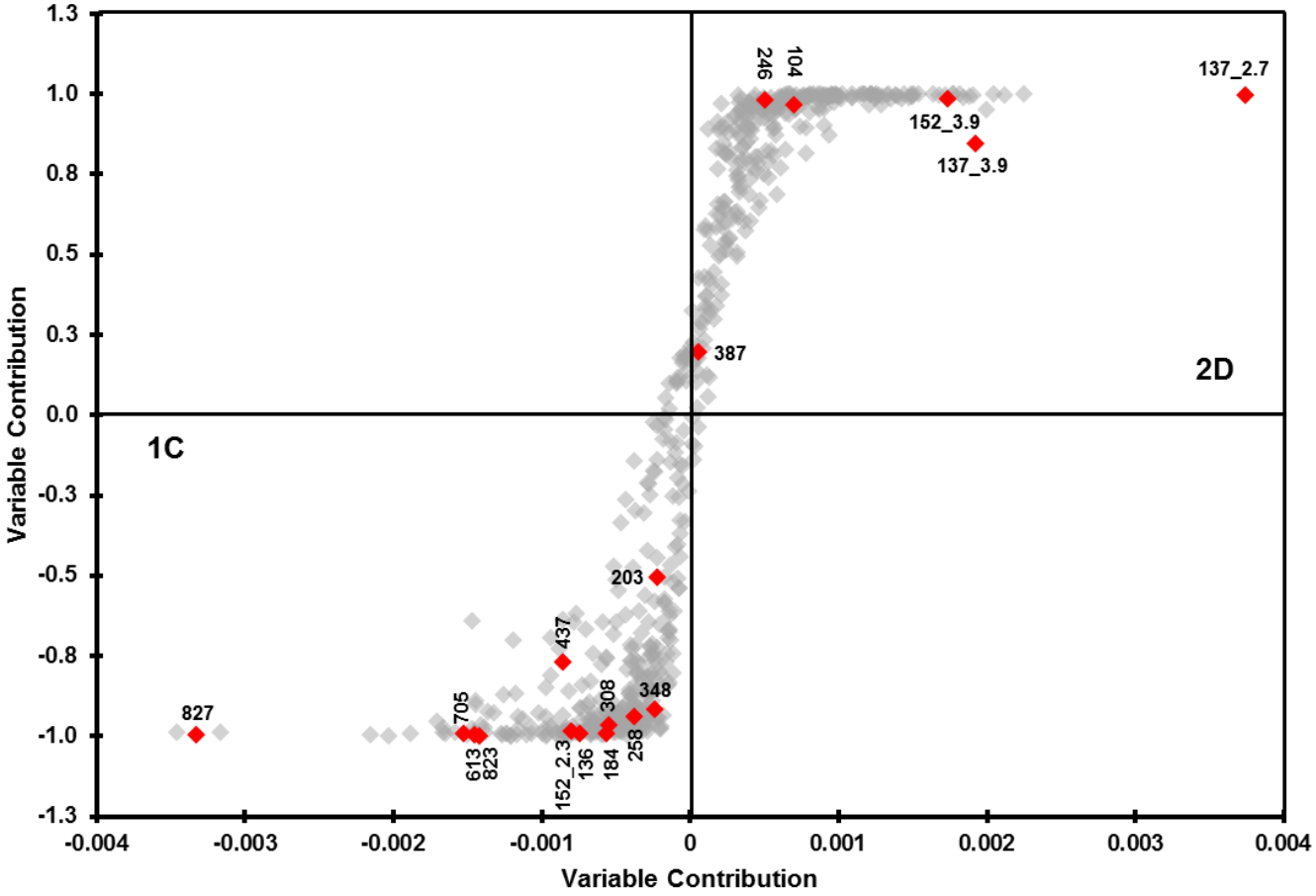

Figure S.3. Orthogonal partial least squares-discriminant analysis for samples 1C (lower left) and 2D (upper right) indicates the molecular features differentially expressed between the two samples. Features approaching the upper right corner are more abundant in and more unique to sample 2D, whereas features approaching the lower left corner are more abundant in and more unique to sample 1C. The 18 features highlighted in Table 1 are indicated with red diamonds.

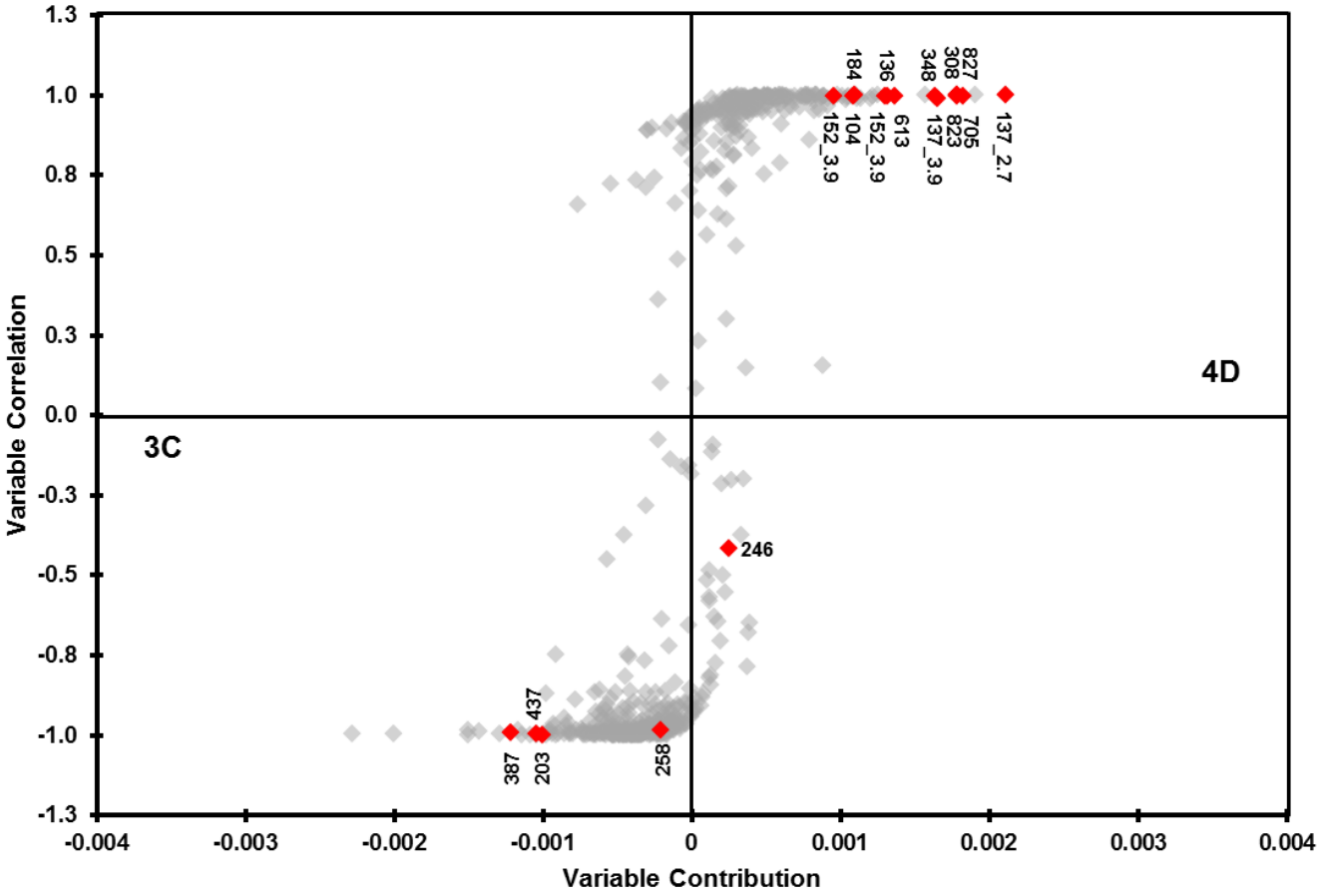

Figure S.4. Orthogonal partial least squares-discriminant analysis for samples 3C (lower left) and 4D (upper right) indicates the molecular features differentially expressed between the two samples. Features approaching the upper right corner are more abundant in and more unique to sample 4D, whereas features approaching the lower left corner are more abundant in and more unique to sample 3C. The 18 features highlighted in Table 1 are indicated with red diamonds.

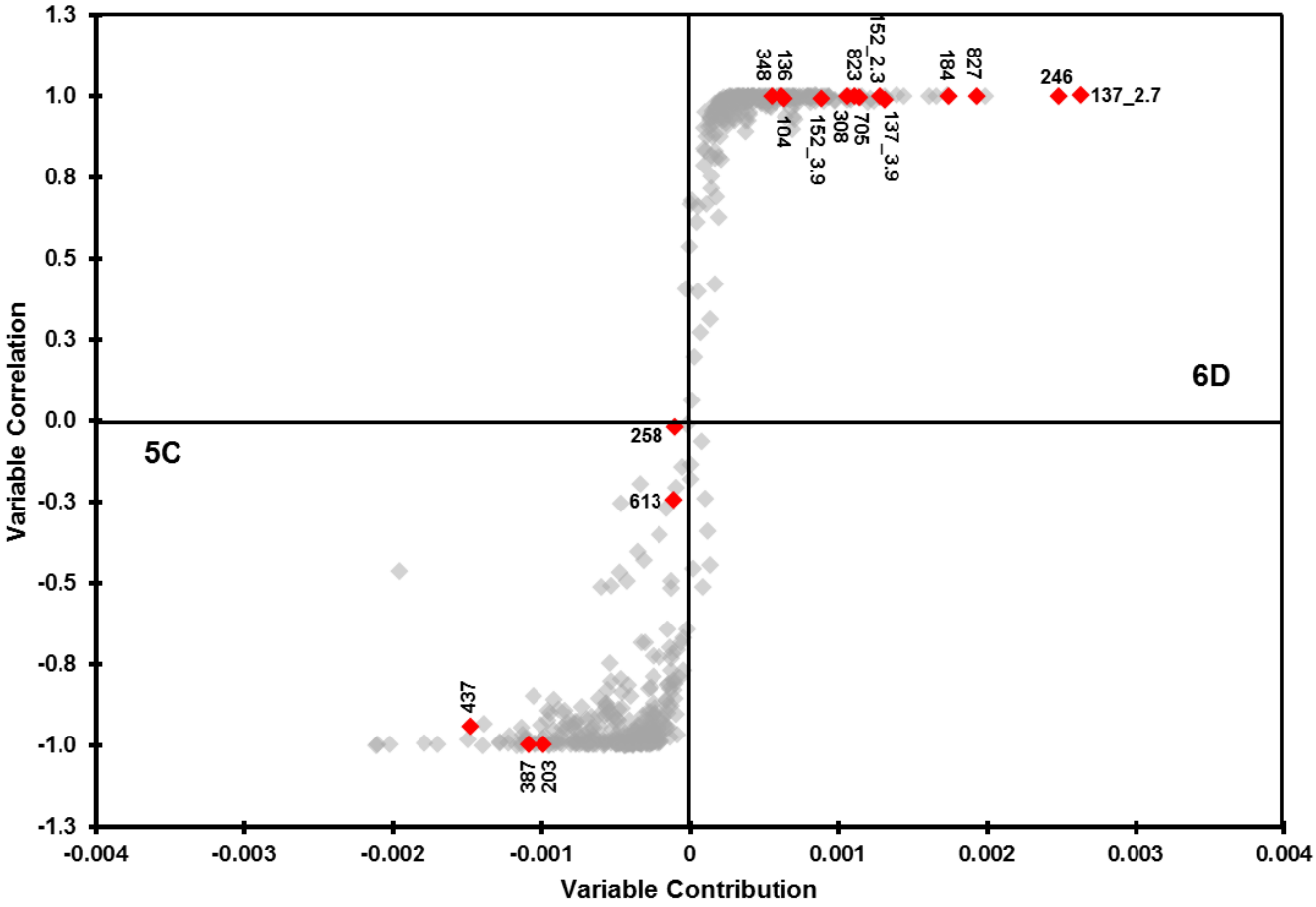

Figure S.5. Orthogonal partial least squares-discriminant analysis for samples 5C (lower left) and 6D (upper right) indicates the molecular features differentially expressed between the two samples. Features approaching the upper right corner are more abundant in and more unique to sample 6D, whereas features approaching the lower left corner are more abundant in and more unique to sample 5C. The 18 features highlighted in Table 1 are indicated with red diamonds.

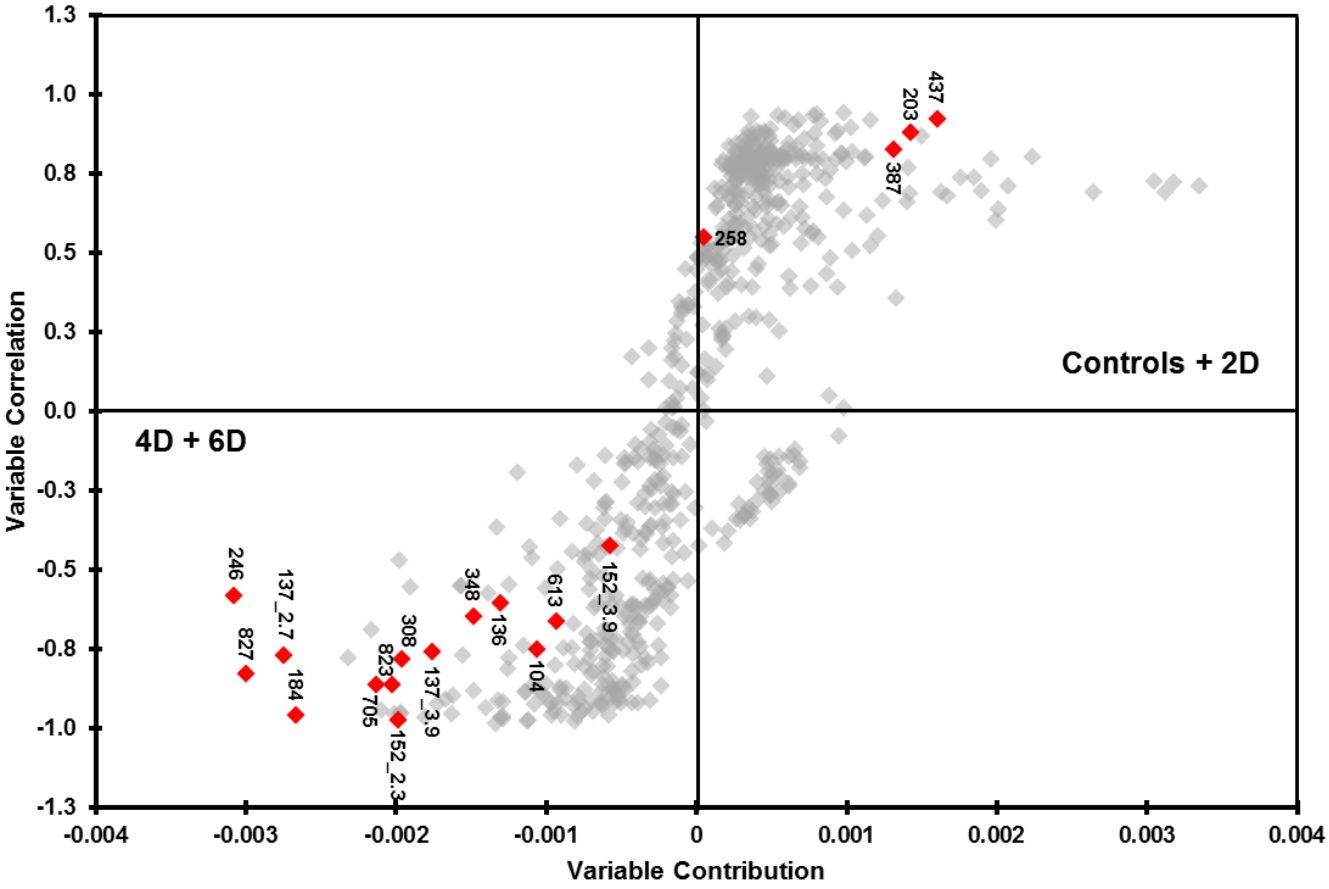

Figure S.6. Orthogonal partial least squares-discriminant analysis for samples 4D and 6D (lower left) and Controls and 2T (upper right) indicates the molecular features differentially expressed between the two samples. Features approaching the upper right corner are more abundant in and more unique to samples 2D, 1C, 3C and 5C, whereas features approaching the lower left corner are more abundant in and more unique to samples 4T and 6T. The 18 features highlighted in Table 1 are indicated with red diamonds.

Data Used to Prepare Figure 2c and Table 1

Data from XCMS peak aligned output. All comparisons are made between matched disease and control pairs. Fold-changes were calculated by dividing the larger of the two tissues (control or diabetic) by the smaller to obtain numbers ≥1. The student's t-test with two-tails, equal variance, and α=0.05 were used to determine significance. The Bonferroni correction has been applied to all p-values to account for multiple testing.

Table S.2. Normalized abundances, fold-changes and *p*-values for Feature #1.

Feature 1 ***m/z* 258.11, 1.53 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	20.62	21.12	0.37	8.1	4.2E-04
1C_02	21.23				
1C_03	21.51				
2D_01	3.01	2.60	0.30		
2D_02	2.44				
2D_03	2.34				
3C_01	57.87	59.18	2.02	3.7	9.0E-04
3C_02	57.65				
3C_03	62.03				
4D_01	222.92	217.99	4.51		
4D_02	219.03				
4D_03	212.03				
5C_01	10.41	9.93	0.49	6.0	1.9E-02
5C_02	10.13				
5C_03	9.26				
6D_01	58.93	60.05	3.33		
6D_02	56.65				
6D_03	64.57				

Table S.3. Normalized abundances, fold-changes and *p*-values for Feature #2.

Feature 2 ***m/z* 348.07, 2.36 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 **1C_02** **1C_03**	6.64 6.86 6.80	6.77	0.09	5.7	2.1E-03
2D_01 **2D_02** **2D_03**	1.01 1.46 1.11	1.19	0.19		
3C_01 **3C_02** **3C_03**	44.55 43.20 49.76	45.84	2.82	3.2	4.8E-03
4D_01 **4D_02** **4D_03**	152.84 145.58 143.86	147.42	3.89		
5C_01 **5C_02** **5C_03**	1.40 1.43 0.91	1.25	0.24	10.5	7.3E-04
6D_01 **6D_02** **6D_03**	13.42 12.81 13.25	13.16	0.26		

Table S.4. Normalized abundances, fold-changes and *p*-values for Feature #3.

Feature 3 ***m/z* 136.06, 2.36 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 **1C_02** **1C_03**	10.41 9.18 9.50	9.70	0.52	3.9	3.4E-02
2D_01 **2D_02** **2D_03**	2.44 2.26 2.74	2.48	0.20		
3C_01 **3C_02** **3C_03**	46.06 47.64 51.60	48.43	2.33	2.4	4.7E-03
4D_01 **4D_02** **4D_03**	116.81 113.81 111.86	114.16	2.04		
5C_01 **5C_02** **5C_03**	2.71 2.47 1.93	2.37	0.33	5.2	2.8E-03
6D_01 **6D_02** **6D_03**	12.69 12.09 12.29	12.36	0.25		

Table S.5. Normalized abundances, fold-changes and *p*-values for Feature #4.

Feature 4 ***m/z* 308.09, 2.45 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 6.01 **1C_02** 4.26 **1C_03** 4.33	6.01 4.26 4.33	4.86	0.81	7.8	1.2
2D_01 0.65 **2D_02** 0.60 **2D_03** 0.62	0.65 0.60 0.62	0.62	0.02		
3C_01 35.38 **3C_02** 37.38 **3C_03** 38.53	35.38 37.38 38.53	37.10	1.30	4.8	8.6E-03
4D_01 186.32 **4D_02** 176.02 **4D_03** 167.92	186.32 176.02 167.92	176.75	7.53		
5C_01 8.66 **5C_02** 9.94 **5C_03** 8.77	8.66 9.94 8.77	9.12	0.58	4.6	6.4E-03
6D_01 42.75 **6D_02** 43.53 **6D_03** 39.88	42.75 43.53 39.88	42.05	1.57		

Table S.6. Normalized abundances, fold-changes and *p*-values for Feature #5.

Feature 5 ***m/z* 613.01, 3.79 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 27.39 **1C_02** 27.43 **1C_03** 30.04	27.39 27.43 30.04	28.29	1.24	20.1	4.4E-03
2D_01 1.36 **2D_02** 1.45 **2D_03** 1.42	1.36 1.45 1.42	1.41	0.04		
3C_01 5.66 **3C_02** 5.44 **3C_03** 5.74	5.66 5.44 5.74	5.61	0.12	15.6	1.0E-02
4D_01 94.05 **4D_02** 85.30 **4D_03** 83.34	94.05 85.30 83.34	87.57	4.66		
5C_01 17.34 **5C_02** 15.74 **5C_03** 17.13	17.34 15.74 17.13	16.74	0.71	1.0	4.2E+02
6D_01 16.07 **6D_02** 16.33 **6D_03** 16.96	16.07 16.33 16.96	16.45	0.37		

Table S.7. Normalized abundances, fold-changes and *p*-values for Feature #6.

Feature 6 ***m/z* 104.17, 1.49 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	21.28	21.11	0.18	1.3	1.3
1C_02	21.19				
1C_03	20.87				
2D_01	27.71	27.26	1.20		
2D_02	28.45				
2D_03	25.63				
3C_01	34.93	36.19	1.29	2.3	1.5E-02
3C_02	35.69				
3C_03	37.96				
4D_01	84.99	81.55	2.53		
4D_02	80.70				
4D_03	78.97				
5C_01	23.26	24.44	0.85	1.4	9.3E-02
5C_02	24.85				
5C_03	25.21				
6D_01	34.34	34.21	0.48		
6D_02	33.57				
6D_03	34.72				

Table S.8. Normalized abundances, fold-changes and *p*-values for Feature #7.

Feature 7 ***m/z* 705.95, 4.64 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	37.77	35.65	2.76	7.4	6.5E-02
1C_02	31.75				
1C_03	37.44				
2D_01	4.54	4.84	0.50		
2D_02	4.44				
2D_03	5.55				
3C_01	12.85	8.67	3.49	22.1	4.7E-04
3C_02	8.88				
3C_03	4.30				
4D_01	187.43	192.01	3.35		
4D_02	193.29				
4D_03	195.32				
5C_01	38.26	34.42	2.75	2.3	1.9E-02
5C_02	31.97				
5C_03	33.03				
6D_01	79.72	79.76	1.29		
6D_02	81.37				
6D_03	78.20				

Table S.9. Normalized abundances, fold-changes and *p*-values for Feature #8.

Feature 8 ***m/z* 823.32, 4.64 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	30.43	29.82	0.46	6.8	6.3E-04
1C_02	29.31				
1C_03	29.73				
2D_01	4.53	4.40	0.55		
2D_02	3.66				
2D_03	5.00				
3C_01	7.15	5.75	1.89	29.8	1.9E-04
3C_02	7.03				
3C_03	3.08				
4D_01	175.74	171.70	2.93		
4D_02	170.48				
4D_03	168.89				
5C_01	30.50	28.01	1.78	2.5	1.1E-02
5C_02	27.13				
5C_03	26.41				
6D_01	69.81	69.05	1.57		
6D_02	70.48				
6D_03	66.87				

Table S.10. Normalized abundances, fold-changes and *p*-values for Feature #9.

Feature 9 ***m/z* 827.82, 4.58 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	159.75	157.54	8.67	11.8	1.5E-02
1C_02	145.99				
1C_03	166.89				
2D_01	12.75	13.37	2.43		
2D_02	10.76				
2D_03	16.60				
3C_01	38.04	25.60	13.36	10.4	9.7E-03
3C_02	31.69				
3C_03	7.06				
4D_01	268.45	266.03	2.25		
4D_02	263.04				
4D_03	266.60				
5C_01	130.72	123.40	6.12	2.0	1.1E-02
5C_02	115.75				
5C_03	123.75				
6D_01	250.24	245.34	3.69		
6D_02	241.32				
6D_03	244.45				

Table S.11. Normalized abundances, fold-changes and *p*-values for Feature #10.

Feature 10 *m/z* **152.11, 2.30 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	12.37	10.96	1.01	4.4	2.2E-01
1C_02	10.01				
1C_03	10.51				
2D_01	2.08	2.49	0.29		
2D_02	2.67				
2D_03	2.72				
3C_01	4.54	4.92	0.27	17.1	3.2E-03
3C_02	5.05				
3C_03	5.16				
4D_01	88.62	83.88	3.35		
4D_02	81.48				
4D_03	81.53				
5C_01	18.37	18.77	0.36	3.6	2.5E-04
5C_02	19.25				
5C_03	18.70				
6D_01	65.54	66.73	1.02		
6D_02	68.03				
6D_03	66.62				

Table S.12. Normalized abundances, fold-changes and *p*-values for Feature #11.

Feature 11 *m/z* **184.11, 1.53 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 6.14 **1C_02** 5.74 **1C_03** 5.59		5.82	0.23	3.5	4.0E-02
2D_01 1.98 **2D_02** 1.58 **2D_03** 1.39		1.65	0.25		
3C_01 18.82 **3C_02** 20.35 **3C_03** 20.93		20.03	0.89	3.7	2.2E-03
4D_01 76.97 **4D_02** 75.48 **4D_03** 72.35		74.93	1.93		
5C_01 5.75 **5C_02** 6.87 **5C_03** 5.63		6.08	0.56	15.8	8.4E-05
6D_01 98.02 **6D_02** 95.57 **6D_03** 94.59		96.06	1.44		

Table S.13. Normalized abundances, fold-changes and *p*-values for Feature #12.

Feature 12 ***m/z* 137.10, 3.94 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	160.13	148.58	13.20	1.3	2.2E+01
1C_02	130.10				
1C_03	155.52				
2D_01	206.12	189.00	12.16		
2D_02	181.94				
2D_03	178.94				
3C_01	99.71	106.03	4.81	2.0	1.4E-01
3C_02	106.99				
3C_03	111.38				
4D_01	218.17	206.81	10.10		
4D_02	208.63				
4D_03	193.64				
5C_01	148.78	139.06	8.24	1.5	2.0E-01
5C_02	139.75				
5C_03	128.65				
6D_01	206.67	209.79	2.36		
6D_02	212.36				
6D_03	210.33				

Table S.14. Normalized abundances, fold-changes and *p*-values for Feature #13.

Feature 13 ***m/z* 246.19, 4.99 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 3.69 **1C_02** 3.26 **1C_03** 3.91		3.62	0.27	1.9	4.5E-01
2D_01 6.63 **2D_02** 7.34 **2D_03** 6.43		6.80	0.39		
3C_01 5.64 **3C_02** 3.97 **3C_03** 5.54		5.05	0.76	1.1	2.7E+02
4D_01 4.67 **4D_02** 4.29 **4D_03** 4.70		4.55	0.19		
5C_01 3.18 **5C_02** 3.32 **5C_03** 2.79		3.10	0.22	67.8	2.9E-06
6D_01 209.56 **6D_02** 208.15 **6D_03** 211.83		209.85	1.52		

Table S.15. Normalized abundances, fold-changes and *p*-values for Feature #14.

Feature 14 m/z 137.10, 2.68 min					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 **1C_02** **1C_03**	72.95 81.97 78.58	77.83	3.72	3.3	1.2E-02
2D_01 **2D_02** **2D_03**	254.93 268.00 243.99	255.64	9.81		
3C_01 **3C_02** **3C_03**	71.24 73.19 83.20	75.87	5.24	3.5	8.7E-04
4D_01 **4D_02** **4D_03**	265.46 268.27 261.99	265.24	2.57		
5C_01 **5C_02** **5C_03**	114.19 119.30 115.82	116.44	2.13	2.6	3.4E-03
6D_01 **6D_02** **6D_03**	308.29 293.77 290.96	297.68	7.59		

Table S.16. Normalized abundances, fold-changes and *p*-values for Feature #15.

Feature 15 ***m/z* 152.11, 3.93 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	34.02	31.56	3.44	2.1	3.0E-01
1C_02	26.69				
1C_03	33.96				
2D_01	70.91	66.60	3.24		
2D_02	65.79				
2D_03	63.11				
3C_01	7.31	7.61	0.57	5.7	2.8E-02
3C_02	7.11				
3C_03	8.41				
4D_01	46.82	43.29	2.56		
4D_02	42.24				
4D_03	40.82				
5C_01	41.50	39.44	1.91	1.7	1.1E-01
5C_02	39.91				
5C_03	36.89				
6D_01	65.51	67.29	2.22		
6D_02	70.43				
6D_03	65.94				

Table S.17. Normalized abundances, fold-changes and *p*-values for Feature #16.

Feature 16 *m/z* **203.09, 1.48 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 87.57 **1C_02** 79.66 **1C_03** 87.08		84.77	3.62	1.0	2.0E+02
2D_01 84.01 **2D_02** 81.81 **2D_03** 76.83		80.88	3.01		
3C_01 71.68 **3C_02** 71.84 **3C_03** 75.53		73.02	1.78	7.3	6.3E-04
4D_01 10.02 **4D_02** 9.67 **4D_03** 10.22		9.97	0.23		
5C_01 39.69 **5C_02** 41.82 **5C_03** 38.84		40.12	1.25	3.5	5.0E-03
6D_01 12.17 **6D_02** 10.88 **6D_03** 11.12		11.39	0.56		

Table S.18. Normalized abundances, fold-changes and *p*-values for Feature #17.

Feature 17 ***m/z* 437.13, 9.31 min**					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 **1C_02** **1C_03**	119.37 111.26 105.22	111.95	5.80	1.1	4.6E+01
2D_01 **2D_02** **2D_03**	96.68 105.65 94.86	99.06	4.72		
3C_01 **3C_02** **3C_03**	125.05 113.68 120.62	119.78	4.68	3.2	3.4E-02
4D_01 **4D_02** **4D_03**	43.04 32.73 35.18	36.98	4.40		
5C_01 **5C_02** **5C_03**	99.13 81.75 76.93	85.94	9.54	1.9	3.3
6D_01 **6D_02** **6D_03**	50.84 44.27 43.61	46.24	3.26		

Table S.19. Normalized abundances, fold-changes and *p*-values for Feature #18.

Feature 18 m/z 387.15					
	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01	86.43	86.41	0.58	1.0	4.7E+02
1C_02	87.11				
1C_03	85.69				
2D_01	82.64	87.65	4.54		
2D_02	93.62				
2D_03	86.69				
3C_01	120.37	130.88	9.03	7.0	4.2E-02
3C_02	129.84				
3C_03	142.42				
4D_01	16.64	18.62	1.65		
4D_02	18.53				
4D_03	20.67				
5C_01	62.81	60.72	1.49	2.1	5.5E-03
5C_02	59.46				
5C_03	59.89				
6D_01	29.30	29.61	0.31		
6D_02	30.03				
6D_03	29.50				

Expression of Thymosins β4 and β10 by Peak Areas for 8+ Species

The region containing thymosins β4 and β10 was mobility-extracted from the IM-MS spectra of each sample (18 total) using a defined selection rule similar to that shown in Figure 1. Total abundances of the 6+ charge states of thymosins β4 and β10 were determined by summing the peak areas in the regions of the mobility-extracted mass spectrum centered about the isotopic distribution for the 8+ species, *m/z* 827.5-829.5 and *m/z* 823.0-825.0, respectively. All comparisons are made between matched disease and control pairs. Fold-changes were calculated by dividing the larger of the two tissues (control or diabetic) by the smaller to obtain numbers ≥1. The student's t-test with two-tails, equal variance, and $\alpha=0.05$ was used to determine significance. The Bonferroni correction was applied to all p-values to account for multiple testing.

Table S.20. Abundances of Thymosin β4 (6+) from IM-Extracted MS Peak Areas.

	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 **1C_02** **1C_03**	206039 208877 198867	204594	4212	3.0	2.9E-02
2D_01 **2D_02** **2D_03**	60145 61216 80429	67263	9320		
3C_01 **3C_02** **3C_03**	60755 47597 16387	41580	18606	38.1	4.5E-04
4D_01 **4D_02** **4D_03**	1631322 1544283 1579135	1584913	35768		
5C_01 **5C_02** **5C_03**	346773 308806 283154	312911	26134	3.8	2.2E-03
6D_01 **6D_02** **6D_03**	1217249 1184695 1163585	1188510	22074		

Table S.21. Abundances of Thymosin β10 (6+) from IM-Extracted MS Peak Areas.

	Norm. Int.	**Average**	**Standard Dev.**	**Fold-change**	***p*-value (corrected)**
1C_01 **1C_02** **1C_03**	45840 48722 42673	45745	2470	1.3	2.3
2D_01 **2D_02** **2D_03**	34868 35239 34954	35020	159		
3C_01 **3C_02** **3C_03**	16016 15619 8612	13416	3401	76.0	8.9E-04
4D_01 **4D_02** **4D_03**	1062637 1000684 993786	1019036	30959		
5C_01 **5C_02** **5C_03**	96636 77735 65772	80048	12706	4.8	5.3E-03
6D_01 **6D_02** **6D_03**	389273 389153 373514	383980	7401		

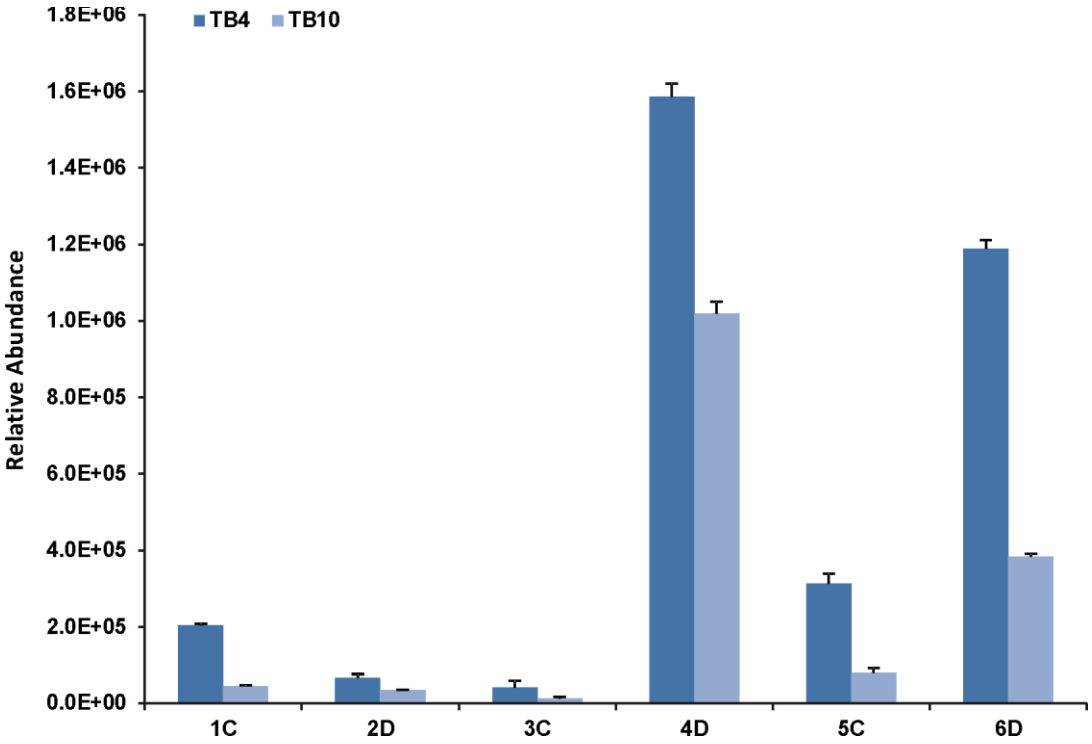

Figure S.7. Expression of thymosins β4 and β10 across the breast tissue samples. Relative abundances were calculated from the peaks areas of the isotopic distribution for the 6+ charge state of thymosins β4 and β10.

MS/MS Spectra for Features in Table 1

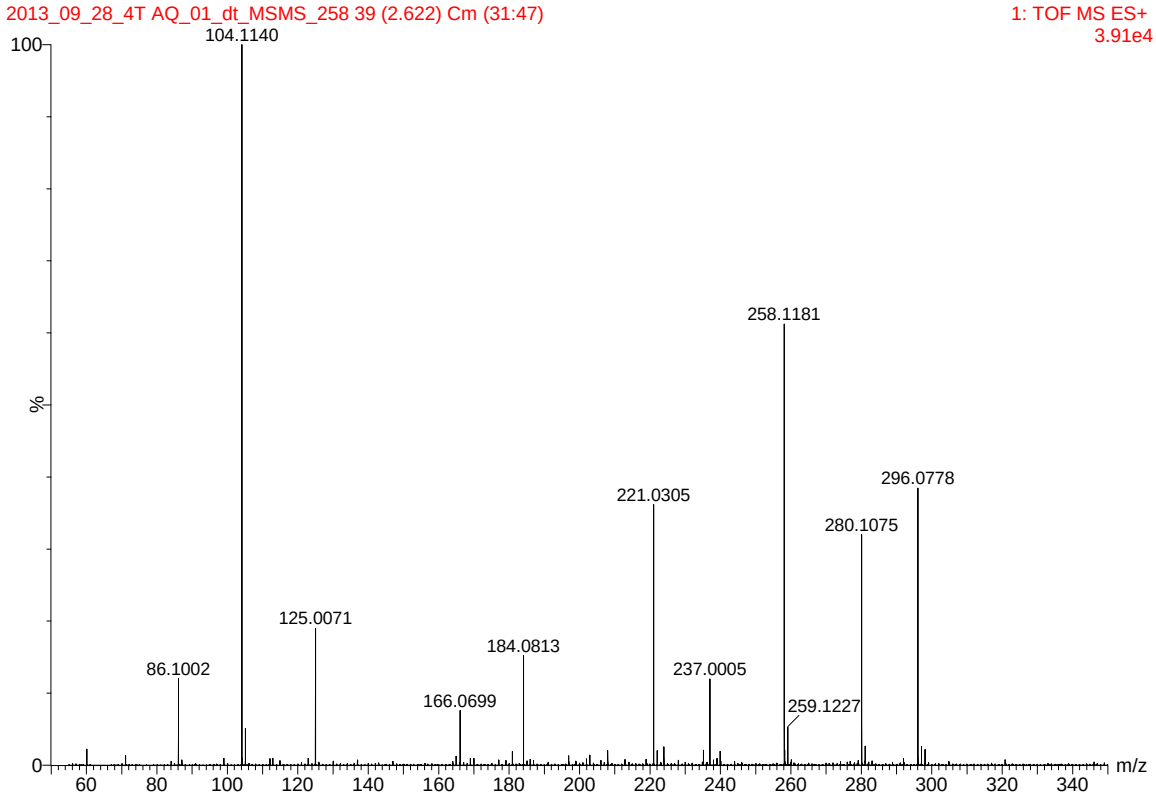

Figure S.8. Mobility- and chromatography-extracted (2.62 *ms*; 1.51 min) DIA MS/MS spectrum of Feature #1, glycerophosphocholine (*m/z* 258.13).

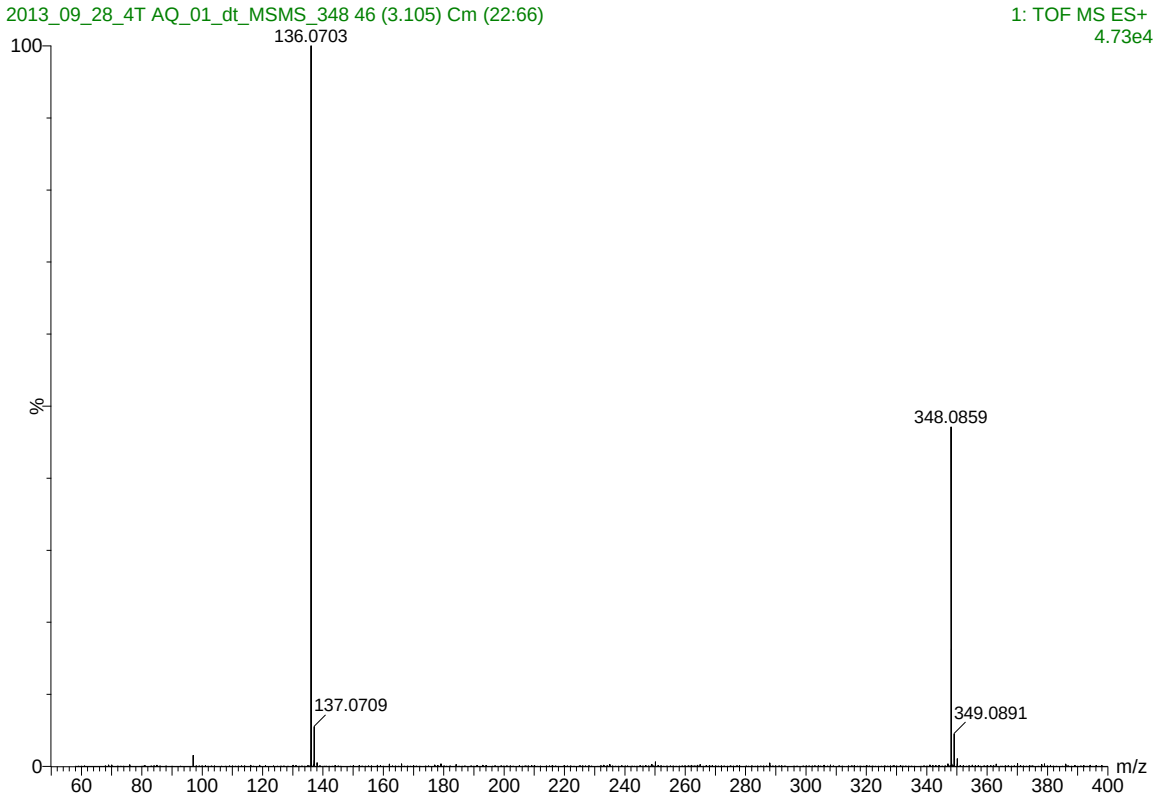

Figure S.9. Mobility- and chromatography-extracted (3.11 *ms*; 2.33 min) DIA MS/MS spectrum of Feature #2, adenosine monophosphate (*m/z* 348.09).

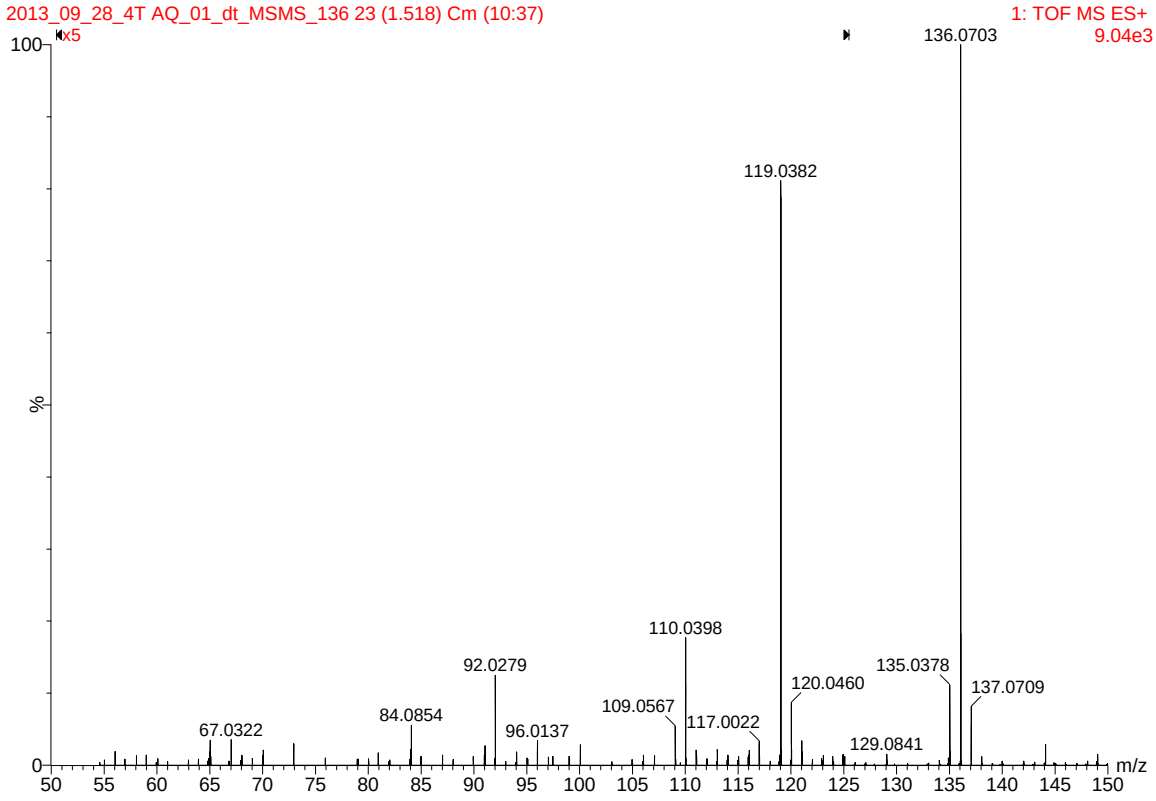

Figure S.10. Mobility- and chromatography-extracted (1.52 *ms*; 2.33) DIA MS/MS spectrum of Feature #3, adenine/fragment of adenosine monophosphate (*m/z* 136.07). The region *m/z* 50-125 is shown at 5X magnification for clarity.

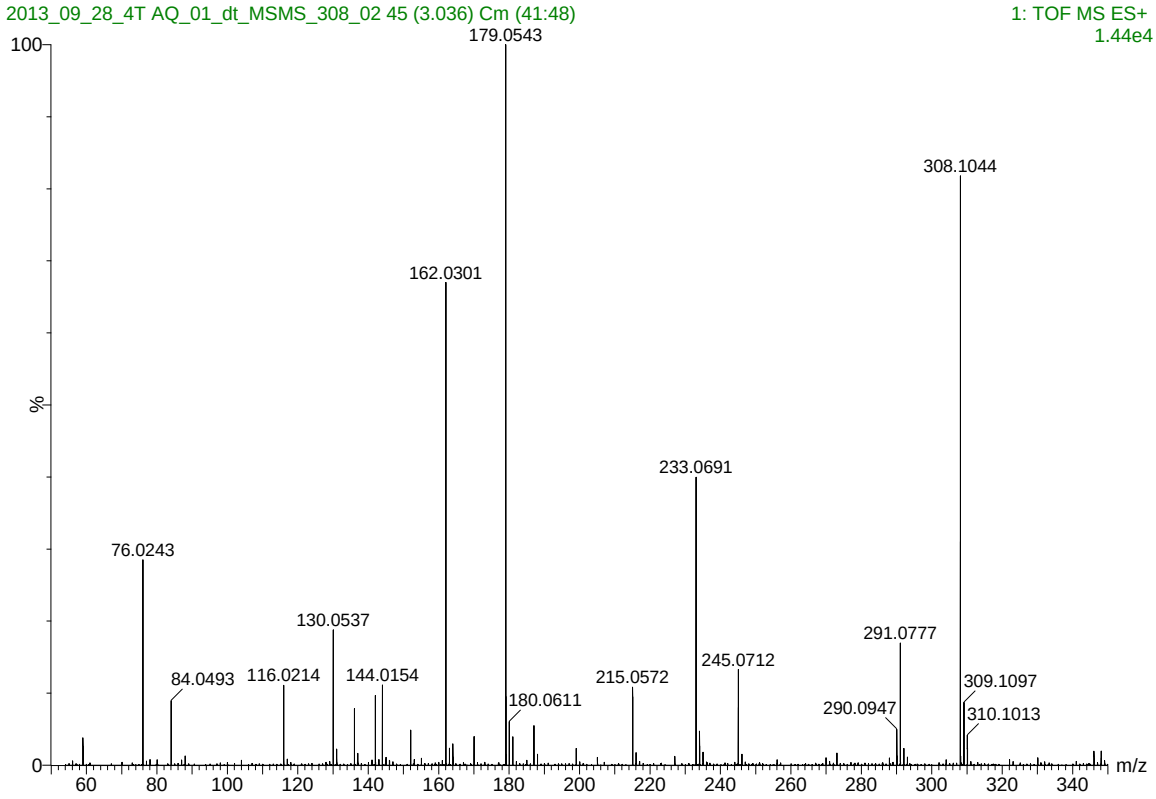

Figure S.11. Mobility- and chromatography-extracted (3.04 *ms*; 2.45 min) DIA MS/MS spectrum of Feature #4, glutathione (*m/z* 308.10).

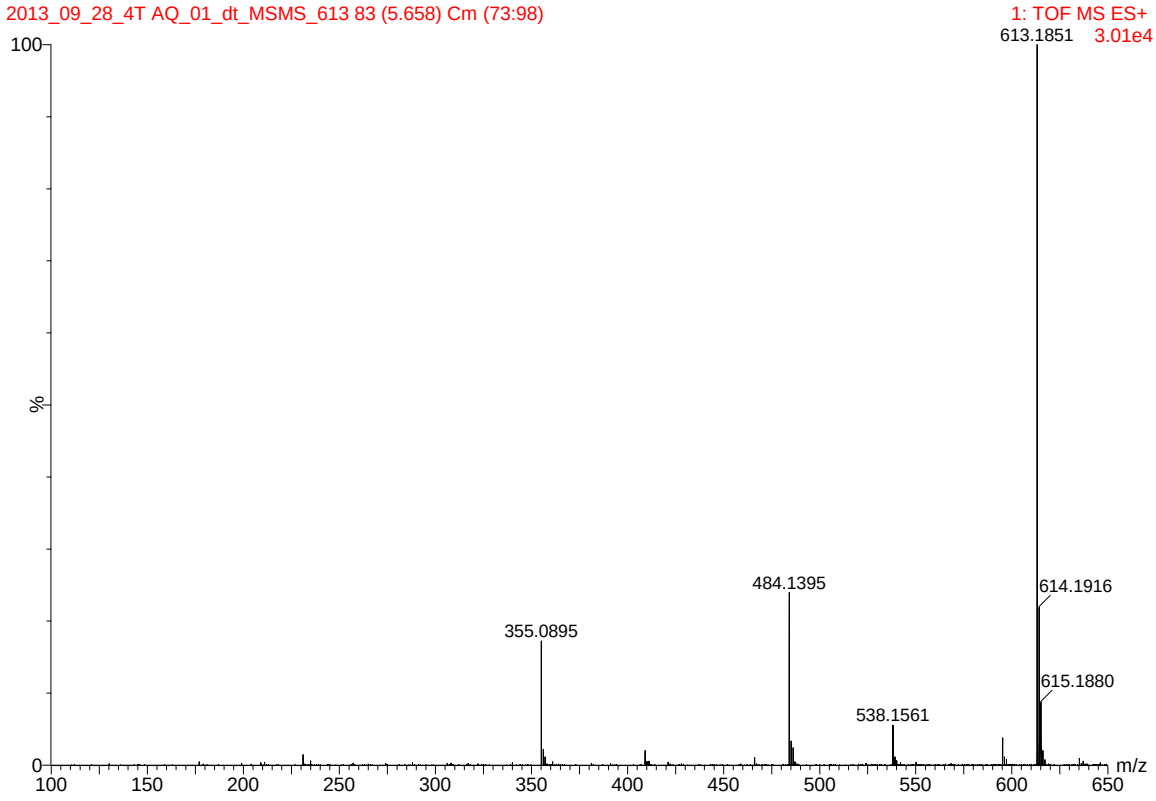

Figure S.12. Mobility- and chromatography-extracted (5.66 *ms*; 3.76 min) DIA MS/MS spectrum of Feature #5, oxidized glutathione (*m/z* 613.19).

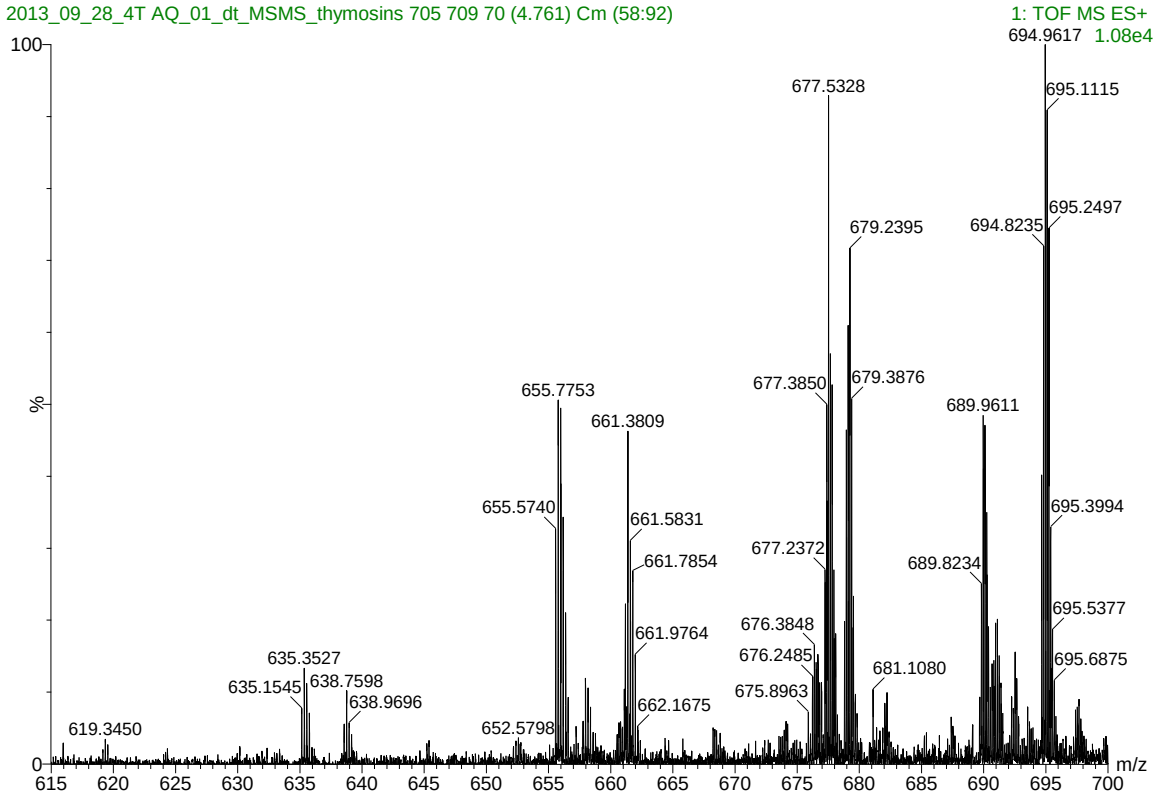

Figure S.13. Mobility- and chromatography-extracted (4.76 *ms*; 4.56 min) DIA MS/MS spectrum of Feature #7-9, thymosins β4 and β10 (*m/z* 705.94, 823.44, 827.76).

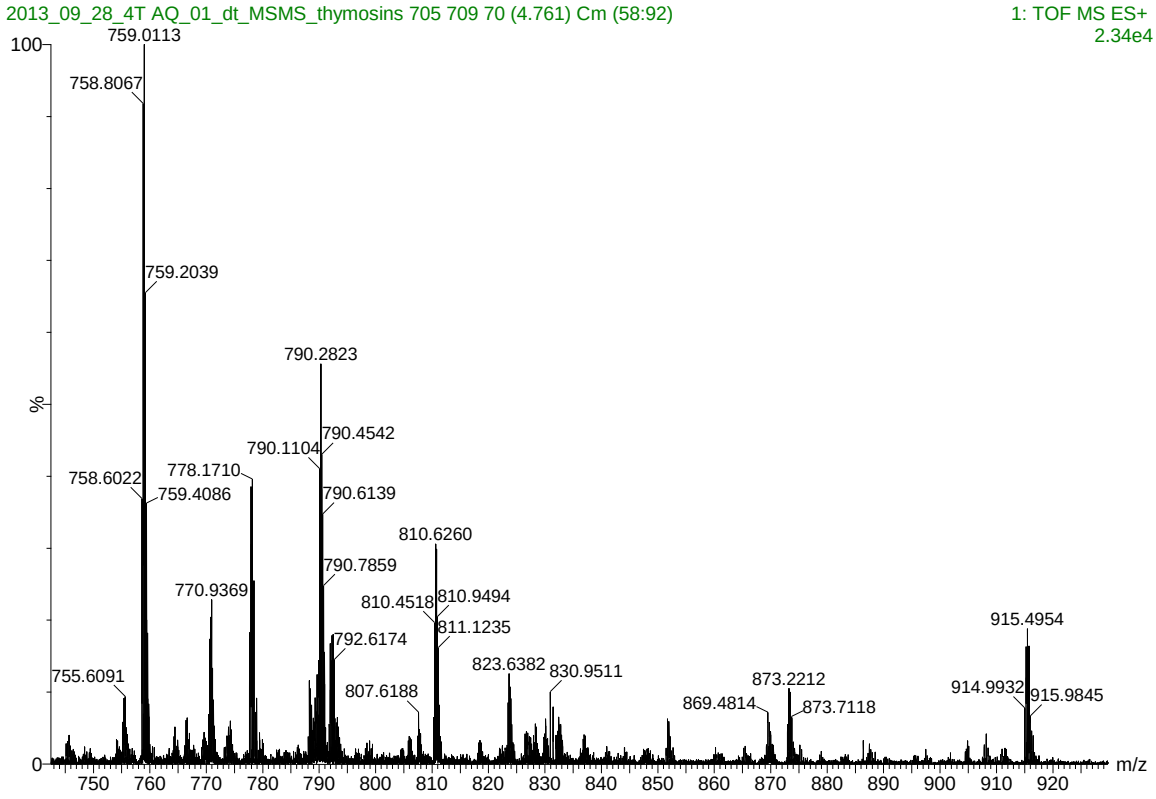

Figure S.14. Mobility- and chromatography-extracted (4.76 *ms*; 4.56 min) DIA MS/MS spectrum of Features #7-9, thymosins β4 and β10 (*m/z* 705.94, 823.44, 827.76).

Table S.21. Backbone cleavages observed for Thymosin β4 and β10

Thymosin β4 Fragmentation				Thymosin β10 Fragmentation		
m/z	**charge**	**fragment**		***m/z***	**charge**	**fragment**
619.1161	8+	$MH\text{-}H_2O^{+8}$		635.1545	5+	$y_{27}\text{-}NH_3^{+5}$
635.3738	5+	$b_{27}\text{-}H_2O^{+5}$		638.5611	5+	y_{27}^{+5}
655.9865	5+	a_{28}^{+5}		660.7968	5+	$y_{28}\text{-}NH_3^{+5}$
661.5927	5+	b_{28}^{+5}		677.2372	7+	$b_{41}+H_2O^{+7}$
676.5419	7+	b_{41}^{+7}		689.6857	7+	y_{42}^{+7}
679.2603	7+	$b_{41}+H_2O^{+7}$				
694.9823	7+	b_{42}^{+7}				
745.4326	4+	b_{25}^{+4}				
773.9556	4+	b_{26}^{+4}				
792.1327	6+	$b_{41}+H_2O^{+6}$				
810.8069	6+	b_{42}^{+6}				

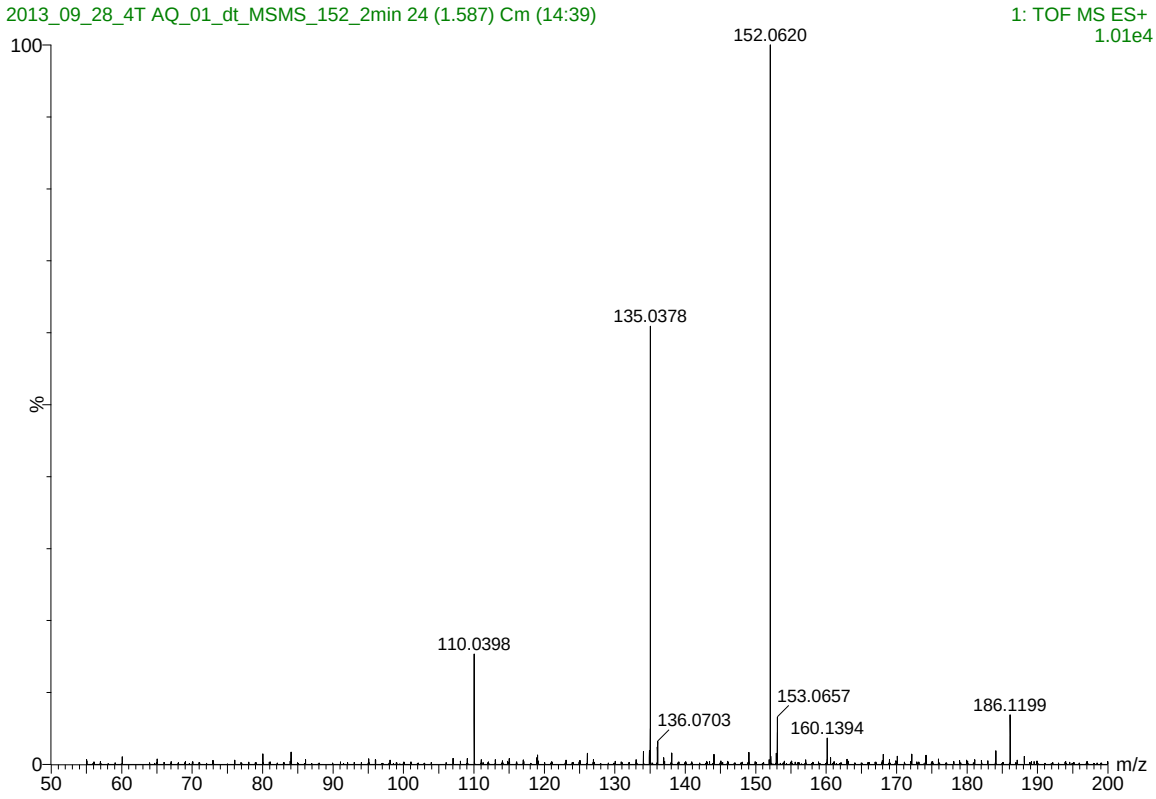

Figure S.15. Mobility- and chromatography-extracted (1.59 *ms*; 2.26 min) DIA MS/MS spectrum of Feature #10 (*m/z* 152.06). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: 2-hydroxyadenine ($C_5H_5N_5O$; cLogP: 0.56); 8-hydroxyadenine ($C_5H_5N_5O$; cLogP: 0.48); and guanine ($C_5H_5N_5O$; cLogP: -1.16). Also observed at 3.93 min in chromatogram, suggesting this particular species is guanine due to CLogP values.

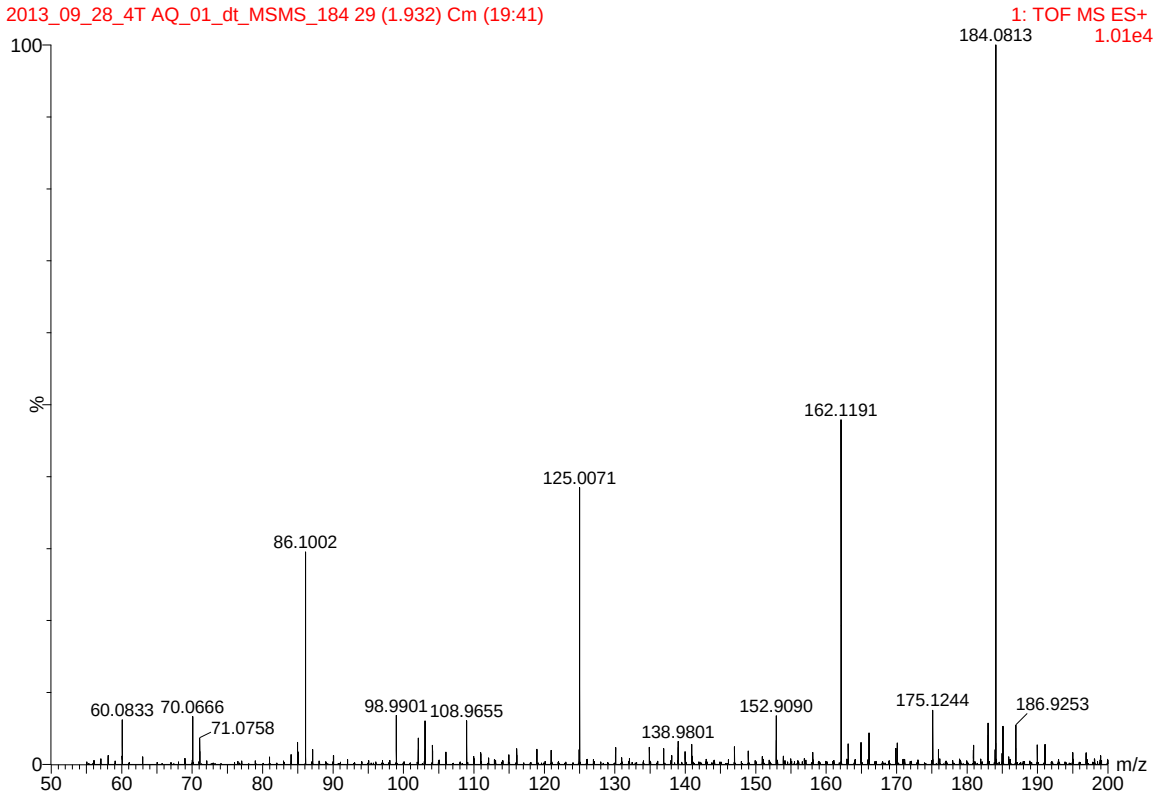

Figure S.16. Mobility- and chromatography-extracted (1.93 *ms*; 1.51 min) DIA MS/MS spectrum of Feature #11, phosphocholine (*m/z* 184.08).

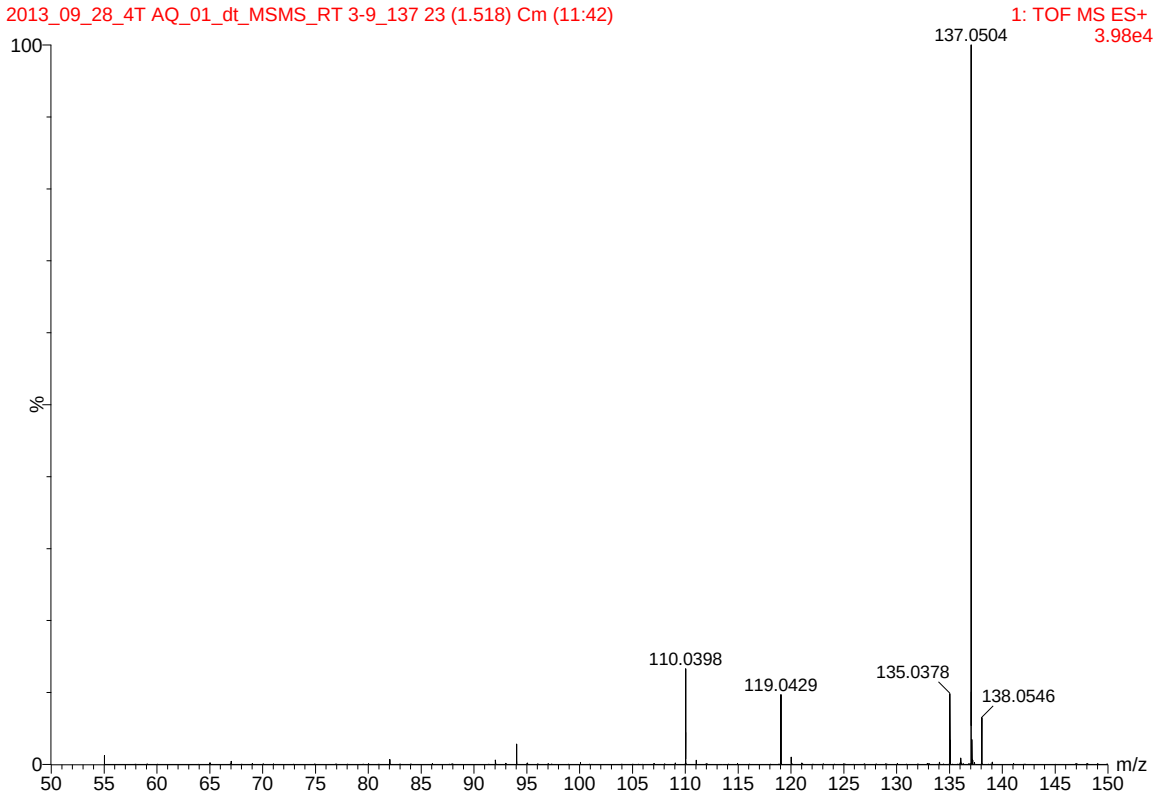

Figure S.17. Mobility- and chromatography-extracted (1.52 *ms*; 3.93 min) DIA MS/MS spectrum of Feature #12 (*m/z* 137.05). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: hypoxanthine ($C_5H_4N_4O$; CLogP: 0.49); allopurinol ($C_5H_4N_4O$; CLogP: -0.74). Also observed at 2.67 min in chromatogram, suggesting this particular species is hypoxanthine due to CLogP values. Signals at *m/z* 137, 119, 110 and 94 correspond to Feature #12, while *m/z* 135 and 110 correspond to feature #15 (*m/z* 152).

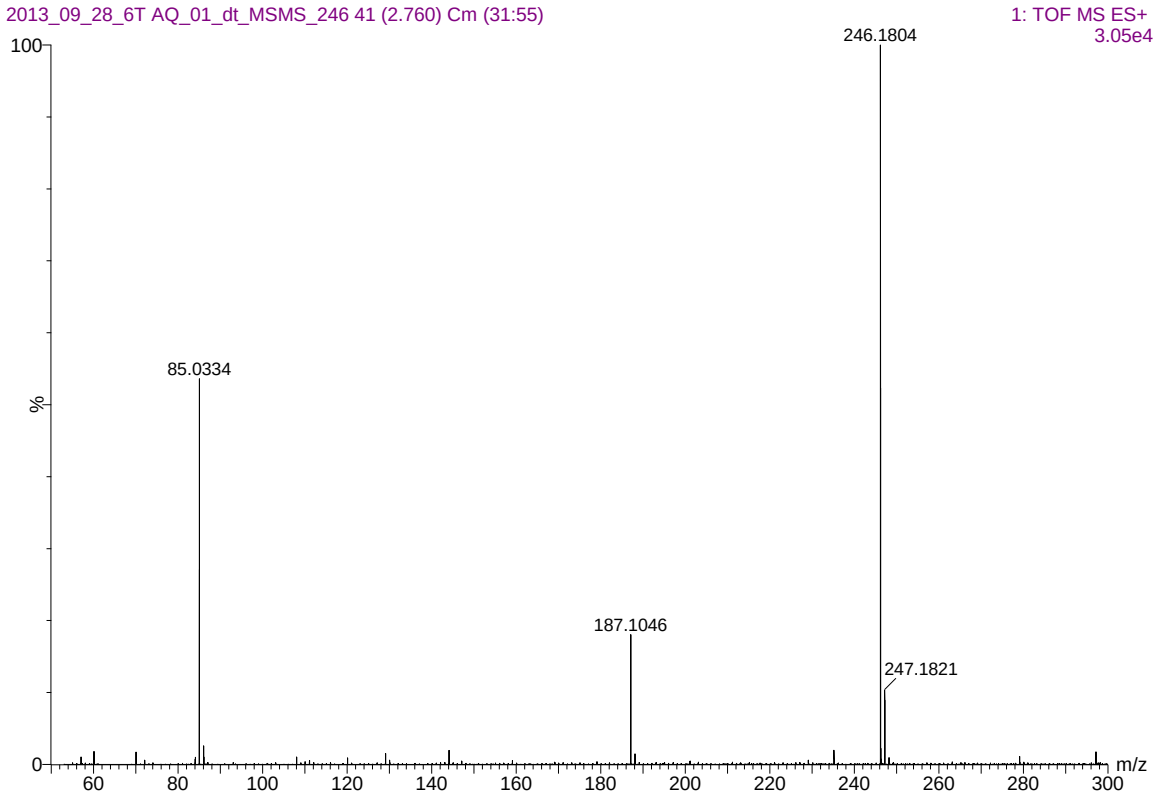

Figure S.18. Mobility- and chromatography-extracted (2.76 *ms*; 5.01 min) DIA MS/MS spectrum of Feature #13 (*m/z* 246.18). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: 2-methylbutyroylcarnitine ($C_{12}H_{23}NO_4$); pivaloylcarnitine ($C_{12}H_{23}NO_4$).

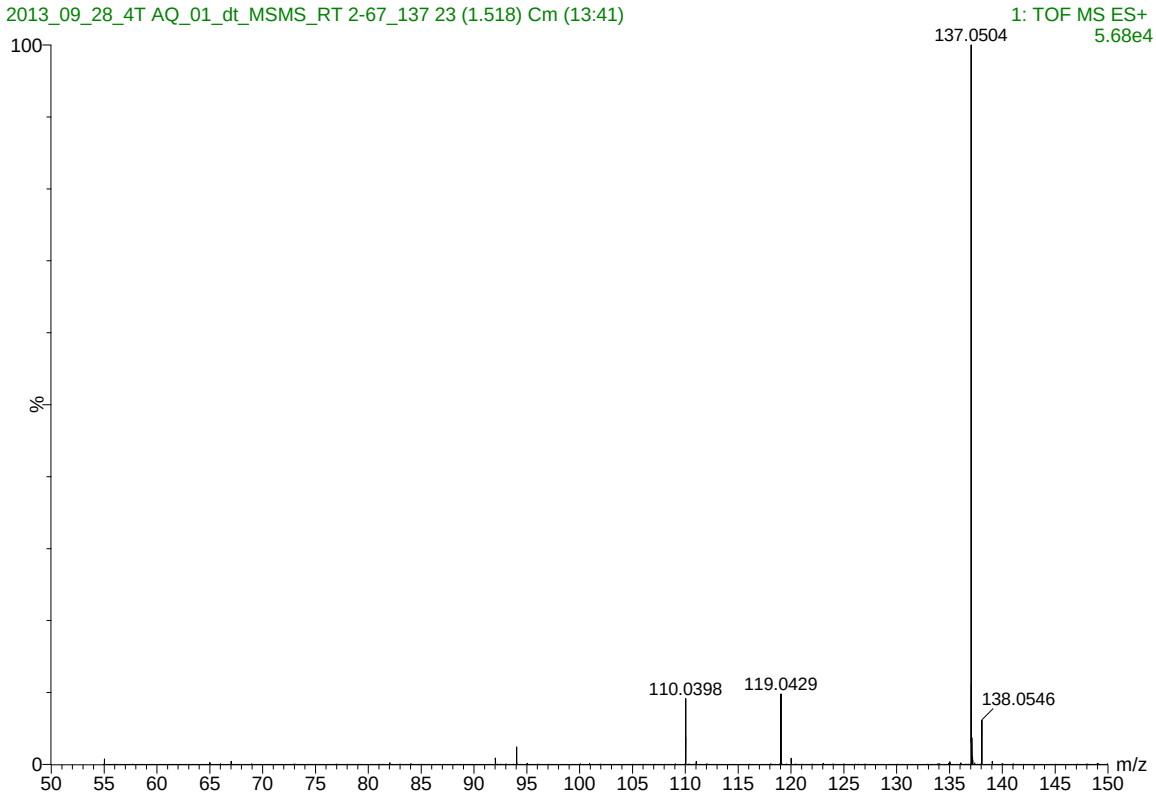

Figure S.19. Mobility- and chromatography-extracted (1.52 *ms*; 2.67 min) DIA MS/MS spectrum of Feature #14 (*m/z* 137.05). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: hypoxanthine ($C_5H_4N_4O$; CLogP: 0.49); allopurinol ($C_5H_4N_4O$; CLogP: -0.74). Also observed at 3.93 min in chromatogram, suggesting this particular species is allopurinol due to CLogP values.

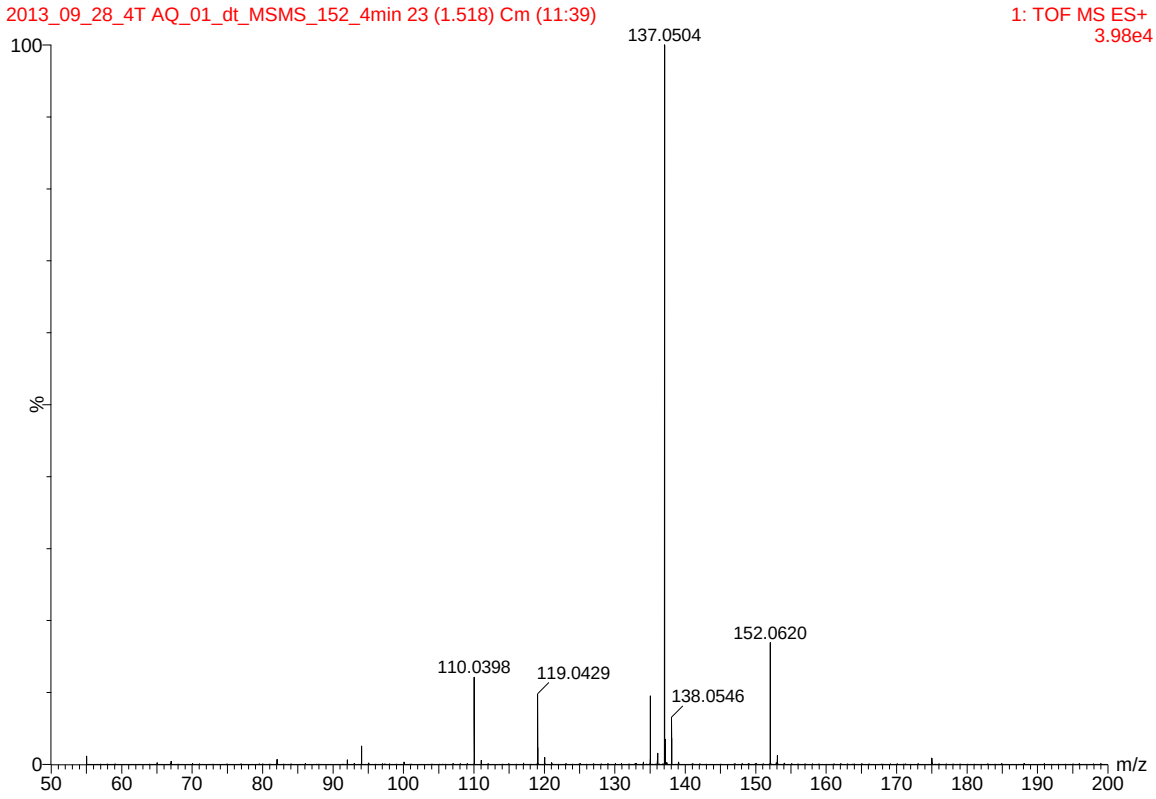

Figure S.20. Mobility- and chromatography-extracted (1.52 *ms*; 3.93 min) DIA MS/MS spectrum of Feature #15 (*m/z* 152.06). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: 2-hydroxyadenine ($C_5H_5N_5O$; cLogP: 0.56); 8-hydroxyadenine ($C_5H_5N_5O$; cLogP: 0.48); and guanine ($C_5H_5N_5O$; cLogP: -1.16). Also observed at 2.67 min in chromatogram, suggesting this particular species is hydroxyadenine due to CLogP values. Signals at *m/z* 137, 119, 110 and 94 correspond to Feature #12, while *m/z* 135 and 110 correspond to feature #15.

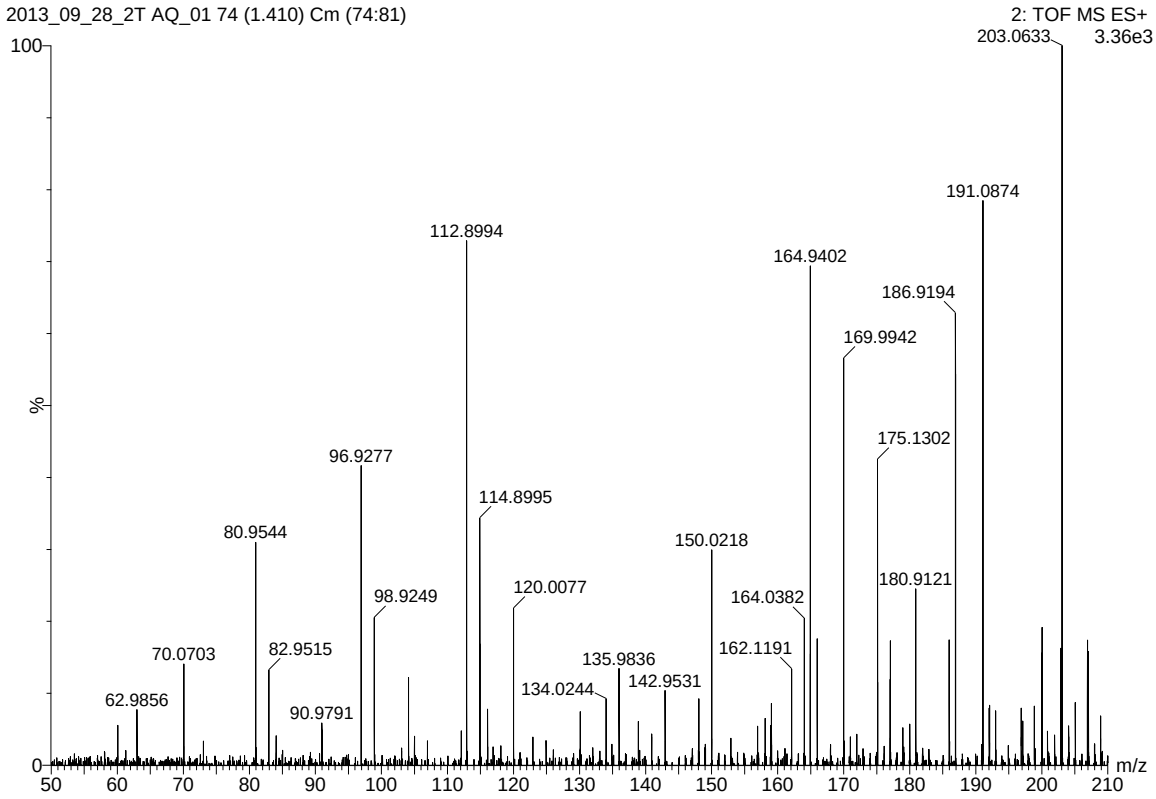

Figure S.21. DIA MS/MS spectrum of Feature #16, a sodiated monosaccharide (*m/z* 203.05) at 1.5 min.

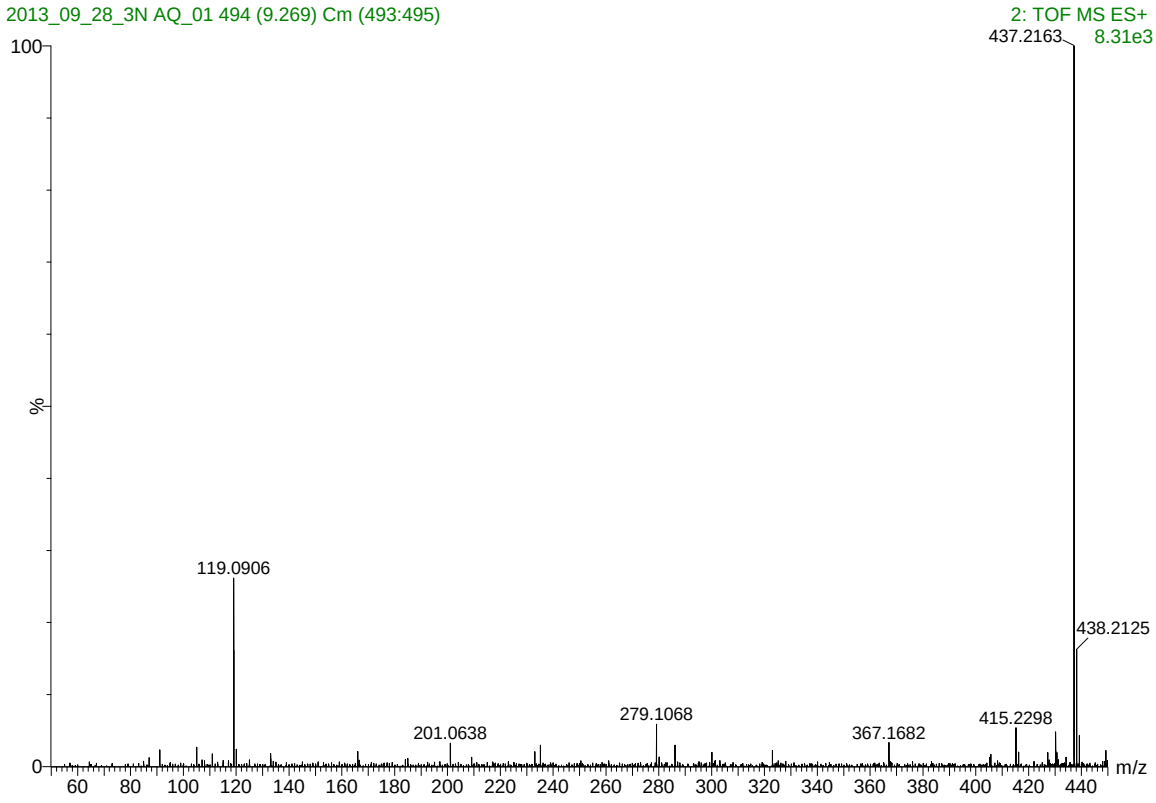

Figure S.22. DIA MS/MS spectrum of Feature #17 (*m/z* 437.19) at 9.31 min.

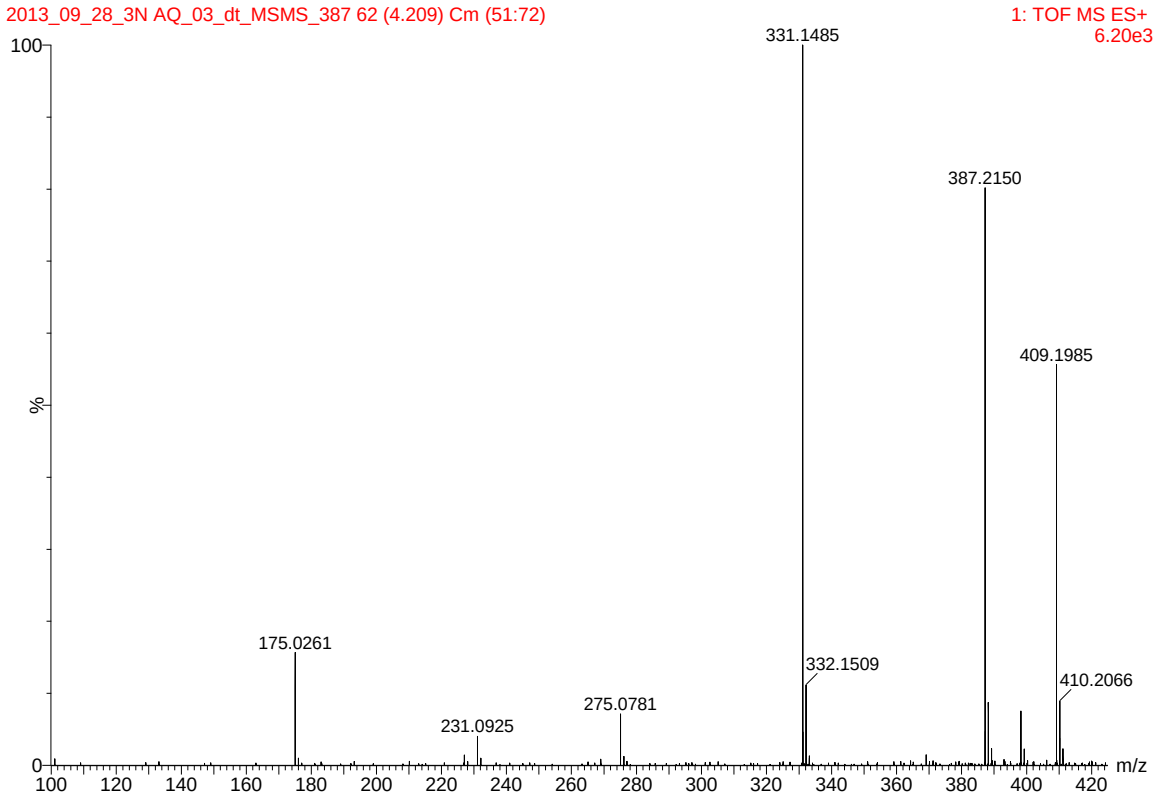

Figure S.23. Mobility- and chromatography-extracted (4.21 *ms*; 9.14 min) DIA MS/MS spectrum of Feature #18, phosphocholine (*m/z* 387.22).

MS/MS Spectra for Standards

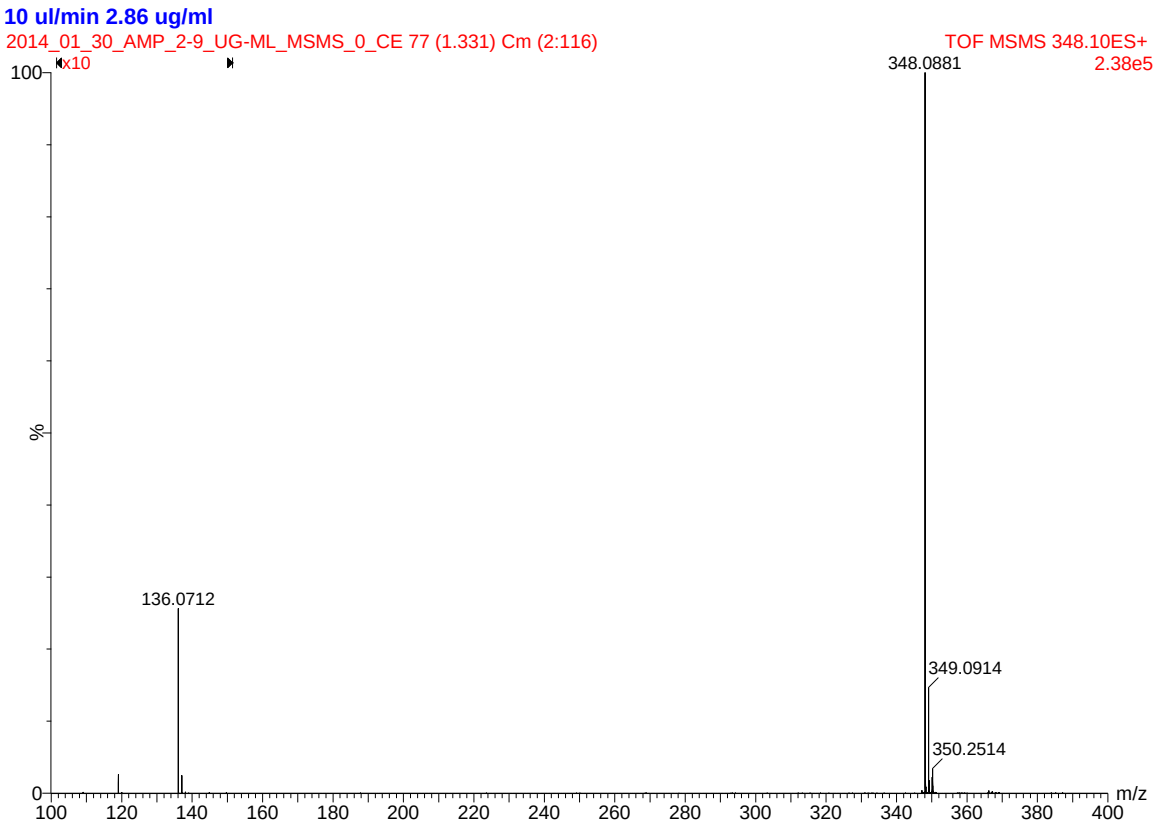

Figure S.24. MS/MS spectrum for 2.86 µg/ml standard of adenosine 5'-monosphosphate with 0 V collision energy. The area from *m/z* 100-150 is shown at 10X magnification for clarity.

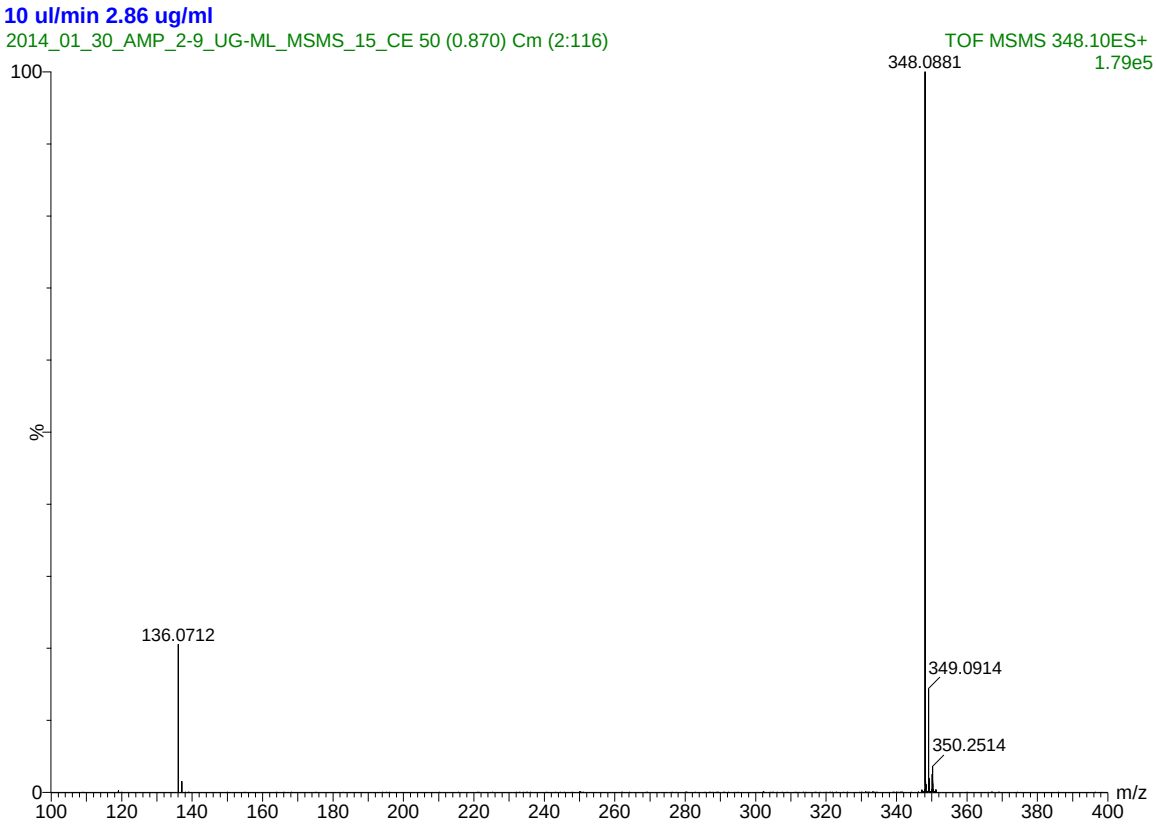

Figure S.25. MS/MS spectrum for 2.86 µg/ml standard of adenosine 5'-monosphosphate (*m/z* 348.09) with 15 V collision energy.

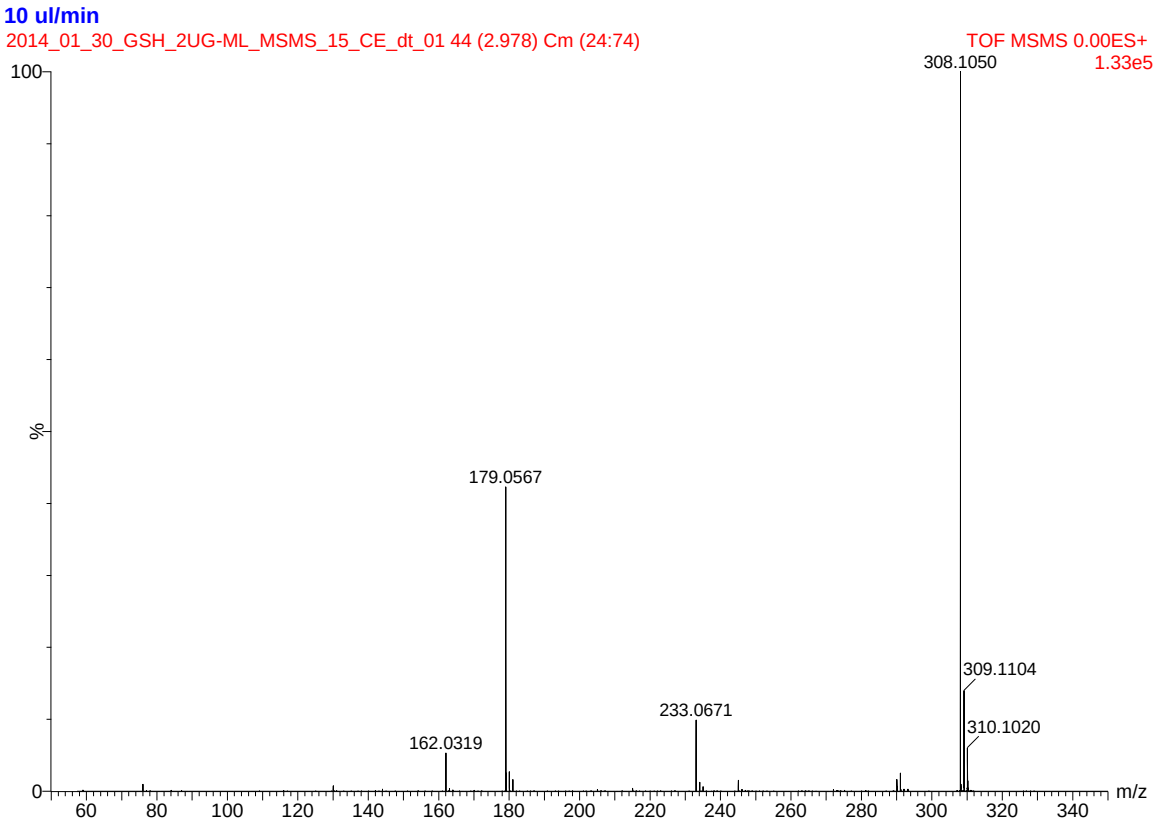

Figure S.26. Mobility-extracted (2.98 *ms*) MS/MS spectrum for 2.12 µg/ml standard of reduced glutathione (*m/z* 308.11) with 15 V collision energy.

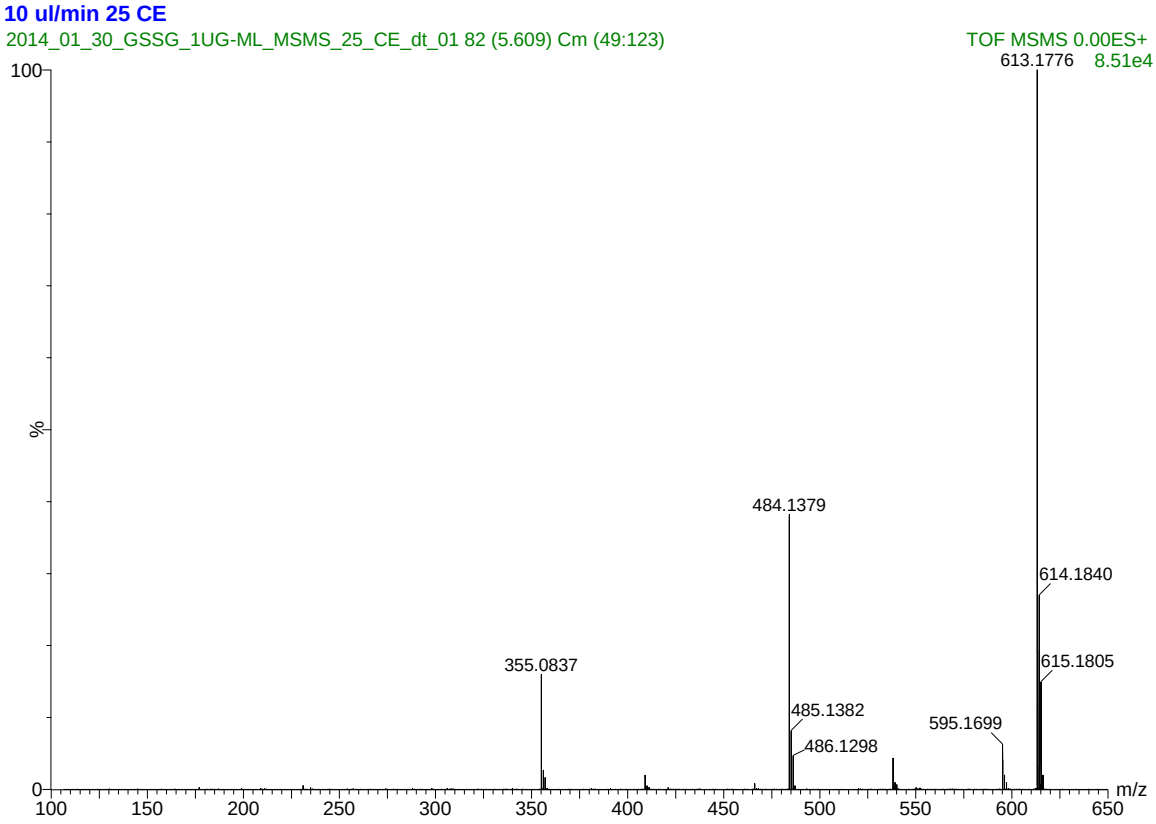

Figure S.27. Mobility-extracted (5.61 *ms*) MS/MS spectrum for 1.1 µg/ml standard of oxidized glutathione (*m/z* 613.18) with 25 V collision energy.

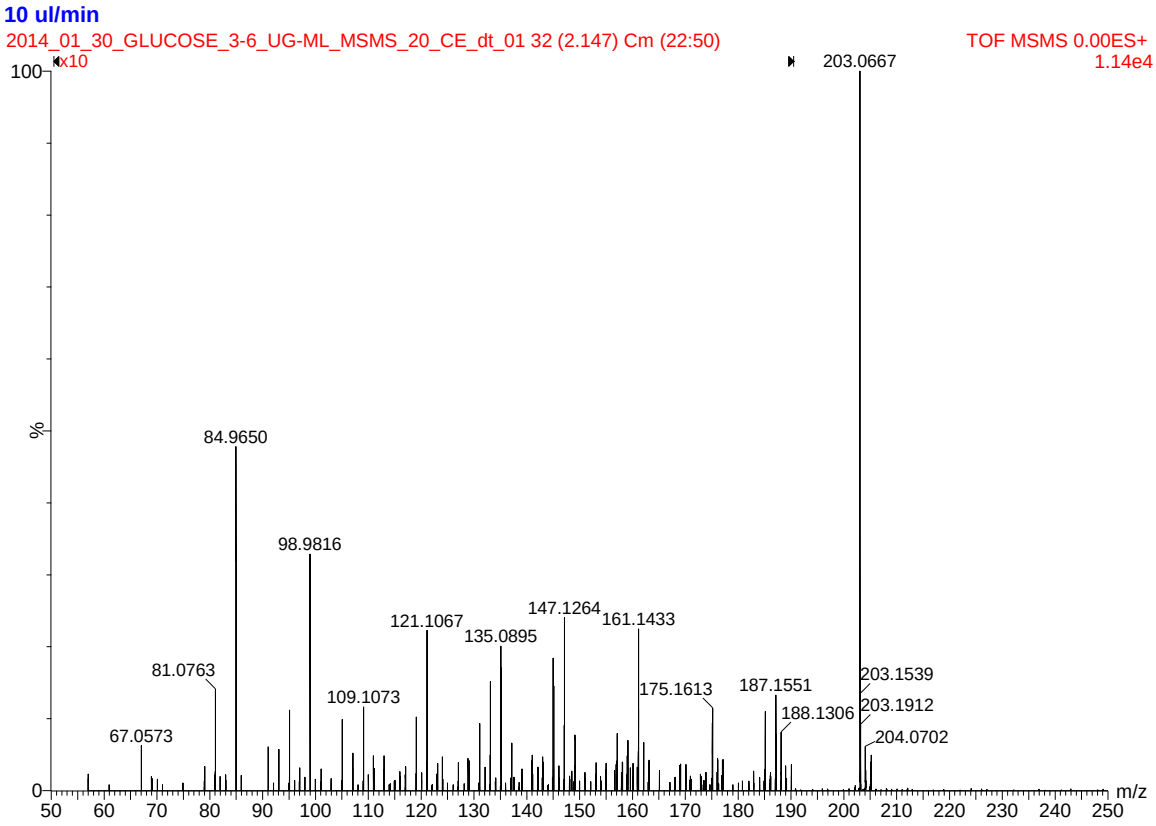

Figure S.28. Mobility-extracted (2.15 *ms*) MS/MS spectrum for 3.6 μg/ml standard of α-D-glucose (*m/z* 203.07) with 20 V collision energy. The region *m/z* 50-190 is shown at 10X magnification for clarity.