

SUPPORTING INFORMATION

An adductomics pipeline for untargeted analysis of modifications to Cys34 of human serum albumin

Hasmik Grigoryan¹, William Edmands¹, Sixin S. Lu¹, Yukiko Yano¹, Luca Regazzoni¹, Anthony T. Iavarone², Evan R. Williams³ and Stephen M. Rappaport^{*1}

¹ School of Public Health, University of California, Berkeley, California 94720-7356, United States

² California Institute for Quantitative Biosciences, University of California, Berkeley, California 94720, United States

³ College of Chemistry, University of California, Berkeley, California 94720, United States

*To whom correspondence should be addressed: Prof. S. M. Rappaport, School of Public Health, University of California, Berkeley, CA 94720-7356, USA. Tel: 1-510-642-4355; Fax: 1-510-642-0427; E-mail: srappaport@berkeley.edu.

Table S1. Characteristics and estimated intake of dietary fat across subjects represented by 34 pooled plasma samples.

Characteristics of Pooled Samples ^a	Race			Gender			Smoking Status		
	Black (B)	White (W)	P-value ^c	Male (M)	Female (F)	P-value ^c	Smoker (S)	Non-Smoker (NS)	P-value ^c
Number of subjects	18 (9M, 9F; 9S, 9NS)	17 (9M, 8F; 10S, 7NS)		18 (9B, 9W; 10S, 8NS)	17 (9B, 8W; 9S, 8NS)		19 (10M, 9F; 9B, 10W)	16 (8M, 8F; 9B, 7W)	
Age (y)	26 ± 2	25 ± 2	0.148	26 ± 3	25 ± 2	0.627	26 ± 3	25 ± 2	0.516
BMI (kg/m ²)	28.0 ± 3.56	24.1 ± 1.64	2.20E-04	25.7 ± 2.21	26.6 ± 4.35	0.41	26.9 ± 3.94	25.2 ± 2.42	0.143
Animal Fat (g/day) b	59.7 ± 21.82	39.2 ± 9.84	1.38E-03	54.3 ± 22.47	45.0 ± 15.78	0.168	57.9 ± 22.03	40.1 ± 11.01	5.91E-03
Vegetable Fat (g/day) b	39.2 ± 14.5	27.3 ± 5.67	3.77E-03	35.0 ± 13.0	31.7 ± 12.1	0.431	38.9 ± 14.0	26.9 ± 6.17	3.34E-03

^a Between 4 and 6 individual plasma specimens were pooled by race, gender and smoking status.

^b Dietary fats and fatty acids were compiled from standardized food frequency questionnaires applied to individual subjects and averaged for pooled-plasma specimens in the current investigation.

^c Based on Student's *t*-test.

Table S2. Evidence used to annotate putative T3 adducts. Accurate masses for 32 adducts led to reasonable elemental compositions added to the thiol of Cys34 in the T3 peptide within three ppm of theoretical values. Annotations of adducts are based on calculated accurate masses, MS2 spectra, literature and database searches and include several classes of modifications. Monoisotopic masses and chromatographic elution times of eleven adducts were confirmed by their synthetic standards. Twenty one adducts were previously reported in the literature as adducts of Cys34, glutathione or other thiols. Database references contain accession numbers from Unimod (www.unimod.org), UniProt (www.uniprot.org) and PubChem (<https://pubchem.ncbi.nlm.nih.gov>).

Adduct	Annotation	Synthetic Standard	Database references	Published references *	Proteins reported modified	Notes
A1	Cys34→Gly		UNIMOD #552	PMID: 15063314, PMID: 21075673	serum transthyretin, papain	
A2	Cys34→Dehydroalanine		UNIMOD #368	PMID: 15063314, PMID: 18265430, 1 PMID: 18722427, PMID: 17450134,	serum transthyretin, serum albumin, synthetic peptides	
A3	Cys34→Oxoalanine/Formylglycine		UNIMOD #402	PMID: 25514000, PMID: 21075673, PMID: 19575405	recombinant proteins, papain, serum albumin	
A4	Not Cys34 adduct					$\gamma 7^{2+}$ or/and $\gamma 7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A5	Not Cys34 adduct					$\gamma 7^{2+}$ or/and $\gamma 7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A6	T3 Labile adduct			1	synthetic peptides	
A7	Unmodified T3	+				
A8	T3 Dimer	+				
A9	Cys34-Gln crosslink		UniProt: KW-0882	PMID: 22591159, PMID: 20729215, PMID: 24220033	serum albumin, pilus adhesin from <i>Streptococcus pyogenes</i>	
A10	Methylation (not Cys34)		UNIMOD #34			
A11	Cyanide adduct		UNIMOD #438	PMID: 24916017, PMID: 21366342, PMID: 21148632	Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH), serum albumin	

A12	Cys34 Sulfinic acid		+ UNIMOD #425	PMID: 24916017, PMID: 21148632, PMID: 24416365	serum albumin, Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH)
A13	Not Cys34 adduct				$\gamma 7^{2+}$ or/and $\gamma 7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A14	Ethylene oxide adduct			PMID: 25407640	recombinant human serum albumin
A15	Cys34 Sulfonic acid		+ UNIMOD #345	PMID: 24916017, PMID: 21148632, PMID: 24416365	serum albumin, Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH)
A16	Not Cys34 adduct				$\gamma 7^{2+}$ or/and $\gamma 7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A17	Not Cys34 adduct				$\gamma 7^{2+}$ or/and $\gamma 7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A18	Acrylonitrile adduct			PMID: 10563837, PMID: 22231508, PMID: 5656374	Glutathione, wheat gluten, bovine serum albumin
A19	Na adduct of A12		UNIMOD #30	PMID: 16444685	synthetic peptides
A20	Methylisocyanate adduct			PMID: 2302207	glutathione
A21	Methylvinylketone adduct			PMID: 25087629, PMID: 1897958, PMID: 22084934	Glutathione, Glyceraldehyde-3- phosphate
A22	S-Mercaptoacetamide		CID: 12961		Dehydrogenase (GAPDH)
A23	S-Mercaptoacetic acid		CID: 1133		
A24	S-Cys (-H2O)		+ UNIMOD #23	PMID: 25800200	synthetic peptides
A25	Unknown				
A26	Unknown				
A27	Unknown				
A28	S-hCys (-H2O)		UNIMOD #23	PMID: 25800200 PMID: 24416365,	synthetic peptides human serum albumin
A29	S-Cys			PMID: 23215783, PMID: 15791656, PMID: 18624771	
A30	S-Cys (NH2→OH)			2	
A31	Unknown				
A32	S-hCys		+ UNIMOD #312	PMID: 17285228,	serum albumin,

				UNIMOD #1271	PMID: 17760510, PMID: 16395265, PMID: 18624771, PMID: 26658763	metallothionein, serum
A33	S-hCys		+			
A34	Unknown					
A35	Na adduct of A29				PMID: 16444685	
A36	Not Cys34 adduct					$y7^{2+}$ or/and $y7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A37	K adduct of A29					
A38	S-Cys-Gly (-H ₂ O)			UNIMOD #23	PMID: 25800200	synthetic peptides
A39	S-(N-acetyl)Cys				PMID: 15499198	human serum albumin
A40	S-Cys-Gly	+			PMID: 18624771, PMID: 1632485	human serum albumin
A41	Not Cys34 adduct					$y7^{2+}$ or/and $y7^+$ were not detected in MS2. $b14^+$ was detected in MS2
A42	S- γ -Gln-Cys				PMID: 24738021	human serum albumin
A43	S-Glutathione	+		UNIMOD #55	PMID: 18624771, PMID: 1632485	suman serum albumin

* Citations:

- 1 Jianyong Wu-PhD thesis 2009/ Tandem Mass Spectrometric Analysis of Protein and Peptide Adducts of Lipid Peroxidation-Derived Aldehydes
- 2 <http://www.ionsource.com/Card/Deamidation/mono0001.htm>

**Table S3. Covariance parameters for mixed models
(performed with natural logarithms of peak-area ratios) and estimated CVs.**

CovParm' refers to the covariance parameter for the mixed model.

ID' refers to the between-specimen variance component

Residual' refers to the within-specimen variance component (error variance)

CV' refers to the coefficient variation estimated from the within-specimen variance component

Variable is $\ln(\text{peak area adduct} * 1000 / \text{peak area housekeeping peptide})$

Adduct	CovParm	Subject	Estimate	CV
1	ID	ID	0.007431	
1	Residual		0.1788	0.442472
2	ID	ID	0	
2	Residual		0.1288	0.37076
3	ID	ID	0.2064	
3	Residual		0.2041	0.475837
4	ID	ID	0	
4	Residual		0.3211	0.61534
5	ID	ID	0.01576	
5	Residual		0.1574	0.412873
6	ID	ID	0.02108	
6	Residual		0.08073	0.289962
7	ID	ID	0.09397	
7	Residual		0.1066	0.335394
8	ID	ID	0.09089	
8	Residual		1.0419	1.354473
9	ID	ID	0.02888	
9	Residual		0.07329	0.275758
10	ID	ID	0.2159	
10	Residual		0.2709	0.557803
11	ID	ID	0.2552	
11	Residual		0.1288	0.37076
12	ID	ID	0.02255	
12	Residual		0.08476	0.297415
13	ID	ID	0	
13	Residual		0.101	0.326001
14	ID	ID	0.1335	
14	Residual		0.07254	0.274291
15	ID	ID	0.08514	
15	Residual		0.09789	0.320689
16	ID	ID	0.3442	
16	Residual		0.396	0.697043
17	ID	ID	0.007284	
17	Residual		0.2534	0.537027
18	ID	ID	0.2221	
18	Residual		0.09907	0.322713

19	ID	ID	0.1866	
19	Residual		0.05752	0.243324
20	ID	ID	0.000045	
20	Residual		0.006651	0.081689
21	ID	ID	0.01044	
21	Residual		0.1809	0.445304
22	ID	ID	0.01105	
22	Residual		0.09973	0.32384
23	ID	ID	0.05202	
23	Residual		0.07186	0.272956
24	ID	ID	0	
24	Residual		0.1009	0.325831
25	ID	ID	0.3888	
25	Residual		0.3133	0.606574
26	ID	ID	0.144	
26	Residual		0.08373	0.295525
27	ID	ID	0.1256	
27	Residual		0.07608	0.281156
28	ID	ID	0.06713	
28	Residual		0.1579	0.413581
29	ID	ID	0.00438	
29	Residual		0.05193	0.230872
30	ID	ID	0.04683	
30	Residual		0.06135	0.251537
31	ID	ID	0.07844	
31	Residual		0.1426	0.391495
32	ID	ID	0	
32	Residual		0.6284	0.935205
33	ID	ID	0.0389	
33	Residual		0.06382	0.256711
34	ID	ID	0.1894	
34	Residual		0.1172	0.352625
35	ID	ID	0.02206	
35	Residual		0.2132	0.487475
36	ID	ID	0.1587	
36	Residual		0.3216	0.6159
37	ID	ID	0.01711	
37	Residual		0.2098	0.483147
38	ID	ID	0.09685	
38	Residual		0.2378	0.518127
39	ID	ID	0.1692	
39	Residual		0.1002	0.324641
40	ID	ID	0.01503	
40	Residual		0.05735	0.242954
41	ID	ID	0.05769	
41	Residual		0.1852	0.451064
42	ID	ID	0	

42 Residual	0.07828	0.285351
43 ID	0.05674	ID
43 Residual	0.09891	0.322439

**Table S4. P-values for 2-sided Wilcoxon (exact) tests
for smoke, race, gender**

smoke	Bonferroni adjusted P-value (0.05/34=0.00147)	
Adduct	P-value	0=Not Sig,1=Sig
A1	0.08286	0
A3	0.97082	0
A5	0.22808	0
A6	0.25613	0
A7	0.86416	0
A8	0.16747	0
A9	0.01375	0
A10	0.01354	0
A11	0.02502	0
A12+A19	0.00058	1
A14	0	1
A15	0.00503	0
A16	0.10119	0
A17	0.17666	0
A18	0	1
A20	1	0
A21	0.60121	0
A22	0.46658	0
A23	0.58361	0
A25	0.95189	0
A26	0.28442	0
A27	0.80984	0
A28	0.5601	0
A29+A35+A37	0.0012	1
A30	0.65622	0
A31	0.06044	0
A33	0.31896	0
A34	0.03956	0
A36	0.09281	0
A38	0.60741	0
A39	0.0063	0
A40	0.28641	0
A41	0.68065	0
A43	0.73213	0

race	P-value
Adduct	
A1	0.9188
A3	0.56038
A5	0.27429

A6	0.04103	0
A7	0.20538	0
A8	0.94581	0
A9	0.97289	0
A10	0.65157	0
A11	0.51772	0
A12+A19	0.37528	0
A14	0.88672	0
A15	0.25941	0
A16	0.09604	0
A17	0.60227	0
A18	0.66206	0
A20	0.72495	0
A21	0.75323	0
A22	0.86706	0
A23	0.13089	0
A25	0.97924	0
A26	0.19656	0
A27	0.16865	0
A28	0.03434	0
A29+A35+37	0.78566	0
A30	0.89187	0
A31	0.20538	0
A33	0.47446	0
A34	0.00539	0
A36	0.25506	0
A38	0.07886	0
A39	0.03413	0
A40	0.56289	0
A41	0.39323	0
A43	0.20538	0

gender

Adduct	P-value	
A1	0.29683	0
A3	0.15742	0
A5	0.22382	0
A6	0.77225	0
A7	0.67023	0
A8	0.57382	0
A9	0.42243	0
A10	0.91272	0
A11	0.93218	0
A12+A19	0.85147	0
A14	0.57991	0
A15	0.4632	0
A16	0.32941	0

A17	0.23015	0
A18	0.62231	0
A20	1	0
A21	0.87173	0
A22	0.40176	0
A23	0.52815	0
A25	0.73816	0
A26	0.30581	0
A27	0.31587	0
A28	0.23716	0
A29+A35+A37	0.98642	0
A30	0.22382	0
A31	0.87825	0
A33	0.82486	0
A34	0.40284	0
A36	0.74854	0
A38	0.93218	0
A39	0.52742	0
A40	0.38379	0
A41	0.31228	0
A43	0.28112	0

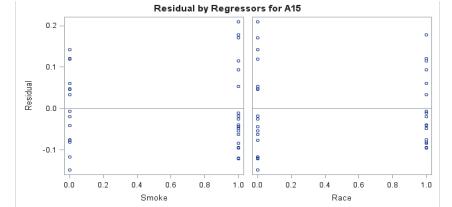
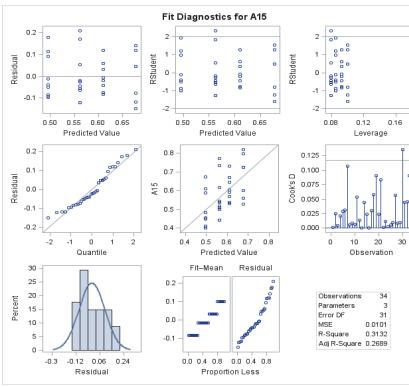
Adduct 15
The REG Procedure
Model: MODEL1
Dependent Variable: A15

Number of Observations Read 34
Number of Observations Used 34

Analysis of Variance
Source DF Sum of Squares Mean Square F Value Pr > F
Model 2 0.14228 0.07114 7.07 0.003
Error 31 0.31201 0.01006
Corrected Total 33 0.4543

Root MSE 0.0932 R-Square 0.3132
Dependent Mean 0.57945 Adj R-Sq 0.2689
Coeff Var 17.31361

Parameter Estimates
Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I
Intercept 1 0.6775 0.03176 21.33 <.0001 0.61272 0.74229
Smoke 1 -0.11554 0.03471 -3.33 0.0023 0.22958 -0.18634 -0.04474
Race 1 -0.06697 0.03447 -1.94 0.0612 0.08361 -0.13727 0.00334



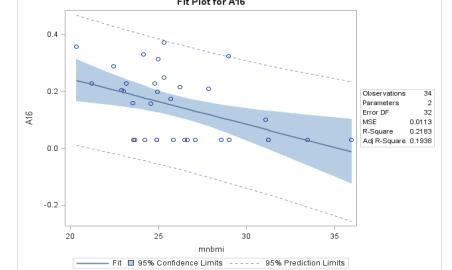
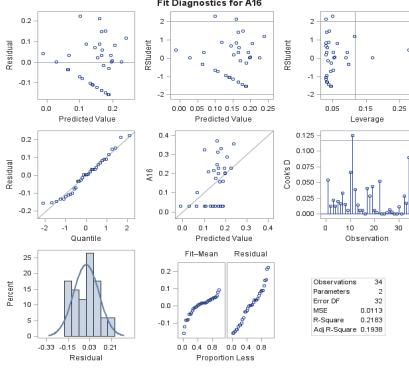
Adduct 16
The REG Procedure
Model: MODEL1
Dependent Variable: A16

Number of Observations Read 34
Number of Observations Used 34

Analysis of Variance
Source DF Sum of Squares Mean Square F Value Pr > F
Model 1 0.10132 0.10132 8.93 0.0053
Error 32 0.36291 0.01134
Corrected Total 33 0.46423

Root MSE 0.10649 R-Square 0.2183
Dependent Mean 0.14675 Adj R-Sq 0.1938
Coeff Var 72.56817

Parameter Estimates
Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I
Intercept 1 0.5663 0.14155 4 0.0003 -0.27797 0.85462
mmbmi 1 -0.01607 0.00538 -2.99 0.0053 0.21825 -0.02703 -0.00512



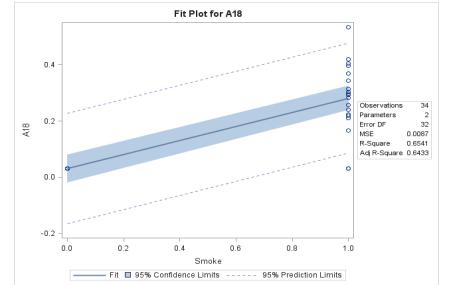
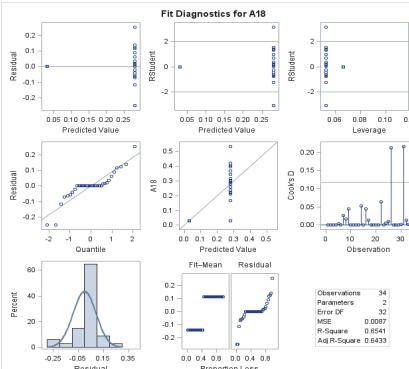
Adduct 18
The REG Procedure
Model: MODEL1
Dependent Variable: A18

Number of Observations Read 34
Number of Observations Used 34

Analysis of Variance
Source DF Sum of Squares Mean Square F Value Pr > F
Model 1 0.5281 0.5281 60.5 <.0001
Error 32 0.27931 0.00873
Corrected Total 33 0.80741

Root MSE 0.09043 R-Square 0.6541
Dependent Mean 0.17027 Adj R-Sq 0.6433
Coeff Var 54.87067

Parameter Estimates
Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I
Intercept 1 0.03 0.02412 1.24 0.2227 -0.01914 0.07914
Smoke 1 0.251 0.03227 7.78 <.0001 0.65407 0.18527 0.31673



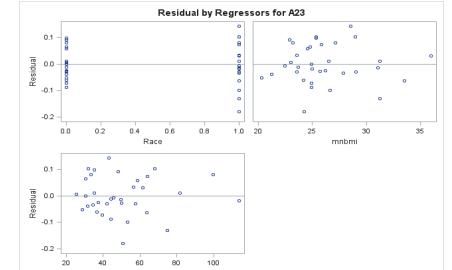
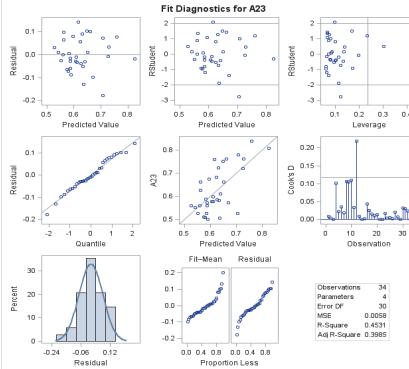
Adduct 23
The REG Procedure
Model: MODEL1
Dependent Variable: A23

Number of Observations Read 34
Number of Observations Used 34

Analysis of Variance
Source DF Sum of Squares Mean Square F Value Pr > F
Model 3 0.14486 0.04829 8.29 0.0004
Error 30 0.17482 0.00583
Corrected Total 33 0.31969

Root MSE 0.07634 R-Square 0.4531
Dependent Mean 0.02052 Adj R-Sq 0.3985
Coeff Var 12.18444

Parameter Estimates
Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I
Intercept 1 0.92356 0.11772 7.85 <.0001 0.68215 1.16397
Race 1 0.09256 0.03569 2.59 0.0145 0.12743 0.19668 0.16543
mmbmi 1 -0.01714 0.00478 -3.58 0.0012 0.20112 -0.02691 -0.00737
mnafat 1 0.00208 0.000796 2.61 0.0138 0.12459 0.000456 0.00371



Adduct 38
 The REG Procedure
 Model: MODEL1
 Dependent Variable: A38

 Number of Observations Read 34
 Number of Observations Used 34

 Analysis of Variance
 Source DF Sum of Squares Mean Square F Value Pr > F

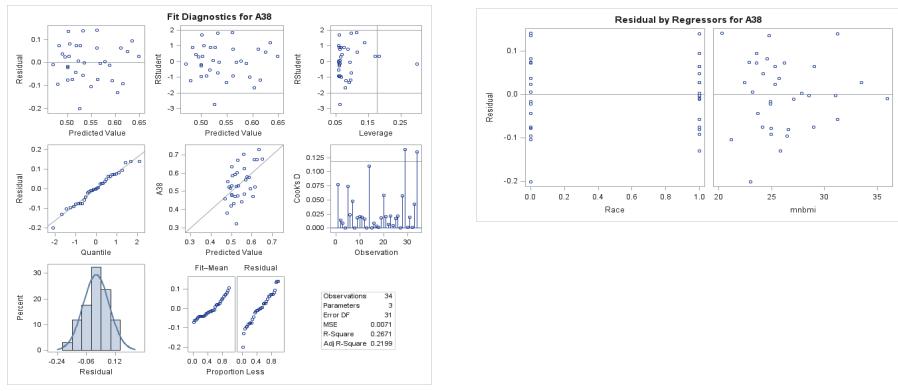
 Model 2 0.07886 0.03993 5.65 0.0081
 Error 31 0.21907 0.00707
 Corrected Total 33 0.29892

Root MSE 0.08406 R-Square 0.2671

Dependent Mean 0.3423 Adj R-Sq 0.2199

Coeff Var 15.50149

 Parameter Estimates
 Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I

 Intercept 1 0.83171 0.12795 6.5 <.0001 0.57075 1.09267
 Race 1 0.11643 0.03559 3.27 0.0062 0.11436 0.04386 0.18801
 mnbmi 1 -0.01332 0.00524 -2.54 0.0162 0.15278 -0.024 -0.00263

Adduct 39
 The REG Procedure
 Model: MODEL1
 Dependent Variable: A39

 Number of Observations Read 33
 Number of Observations Used 33

 Analysis of Variance
 Source DF Sum of Squares Mean Square F Value Pr > F

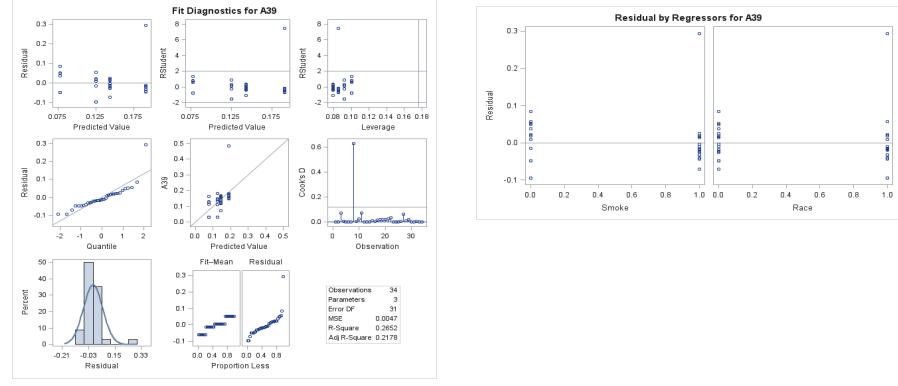
 Model 2 0.02828 0.01141 6.8 0.0037
 Error 30 0.05034 0.00168
 Corrected Total 32 0.07316

Root MSE 0.04996 R-Square 0.3119

Dependent Mean 0.12784 Adj R-Sq 0.266

Coeff Var 32.04275

 Parameter Estimates
 Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I

 Intercept 1 0.08858 0.01305 6.79 <.0001 0.06194 0.11523 0.09831
 Smoke 1 0.04754 0.01438 3.31 0.0025 0.02274 0.01818 0.0769 0.00323
 Race 1 0.02748 0.01432 1.92 0.0646 0.08444 -0.00177 0.05674

Adduct 40
 The REG Procedure
 Model: MODEL1
 Dependent Variable: A40

Number of Observations Read 34

Number of Observations Used 34

 Analysis of Variance
 Source DF Sum of Squares Mean Square F Value Pr > F

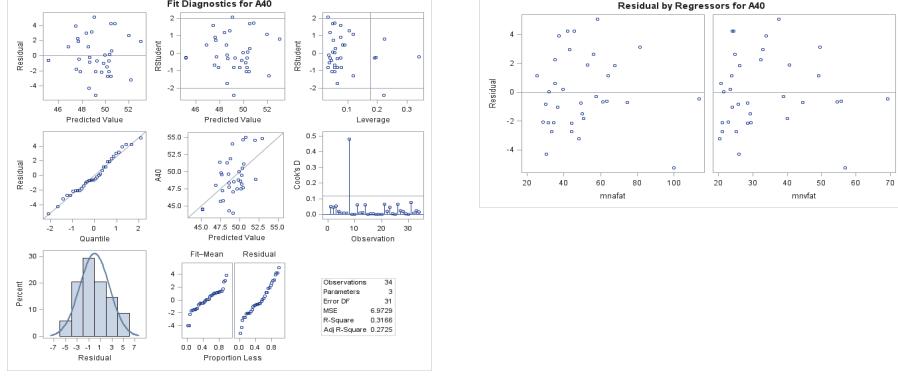
 Model 2 100.1216 50.6008 7.18 0.0027
 Error 31 216.1606 6.97292
 Corrected Total 33 316.2822

Root MSE 2.64063 R-Square 0.3166

Dependent Mean 49.14659 Adj R-Sq 0.2725

Coeff Var 5.37297

 Parameter Estimates
 Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I

 Intercept 4 54.7579 4.24261 20.22 <.0001 0.4879004 54.45209
 mnfat 1 0.11918 0.04203 2.84 0.008 0.00689 0.03346 0.2049
 mnvat 1 -0.24992 0.06668 -3.75 0.0007 0.30967 -0.38592 -0.11392

Adduct 43
 The REG Procedure
 Model: MODEL1
 Dependent Variable: A43

Number of Observations Read 34

Number of Observations Used 34

 Analysis of Variance
 Source DF Sum of Squares Mean Square F Value Pr > F

 Model 1 0.75137 0.75137 9.11 0.005
 Error 32 2.63901 0.08247
 Corrected Total 33 3.39038

Root MSE 0.28717 R-Square 0.2216

Dependent Mean 2.16099 Adj R-Sq 0.1973

Coeff Var 13.28904

 Parameter Estimates
 Variable DF Parameter Estimate Standard Error t Value Pr > |t| Squared Corr Type I

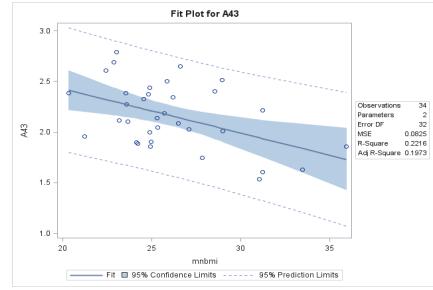
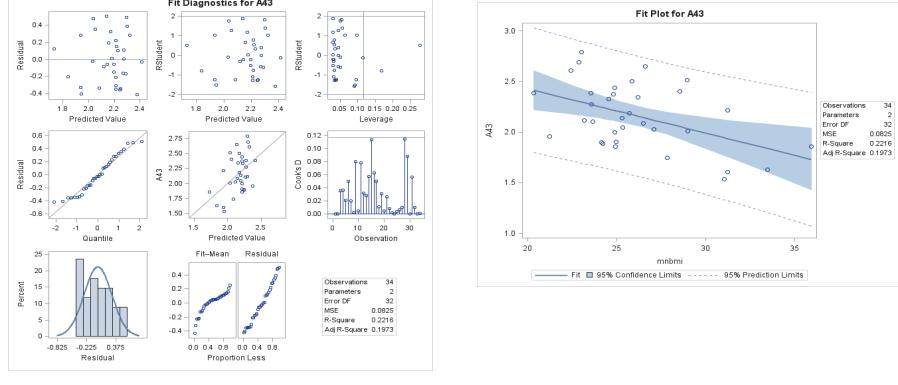
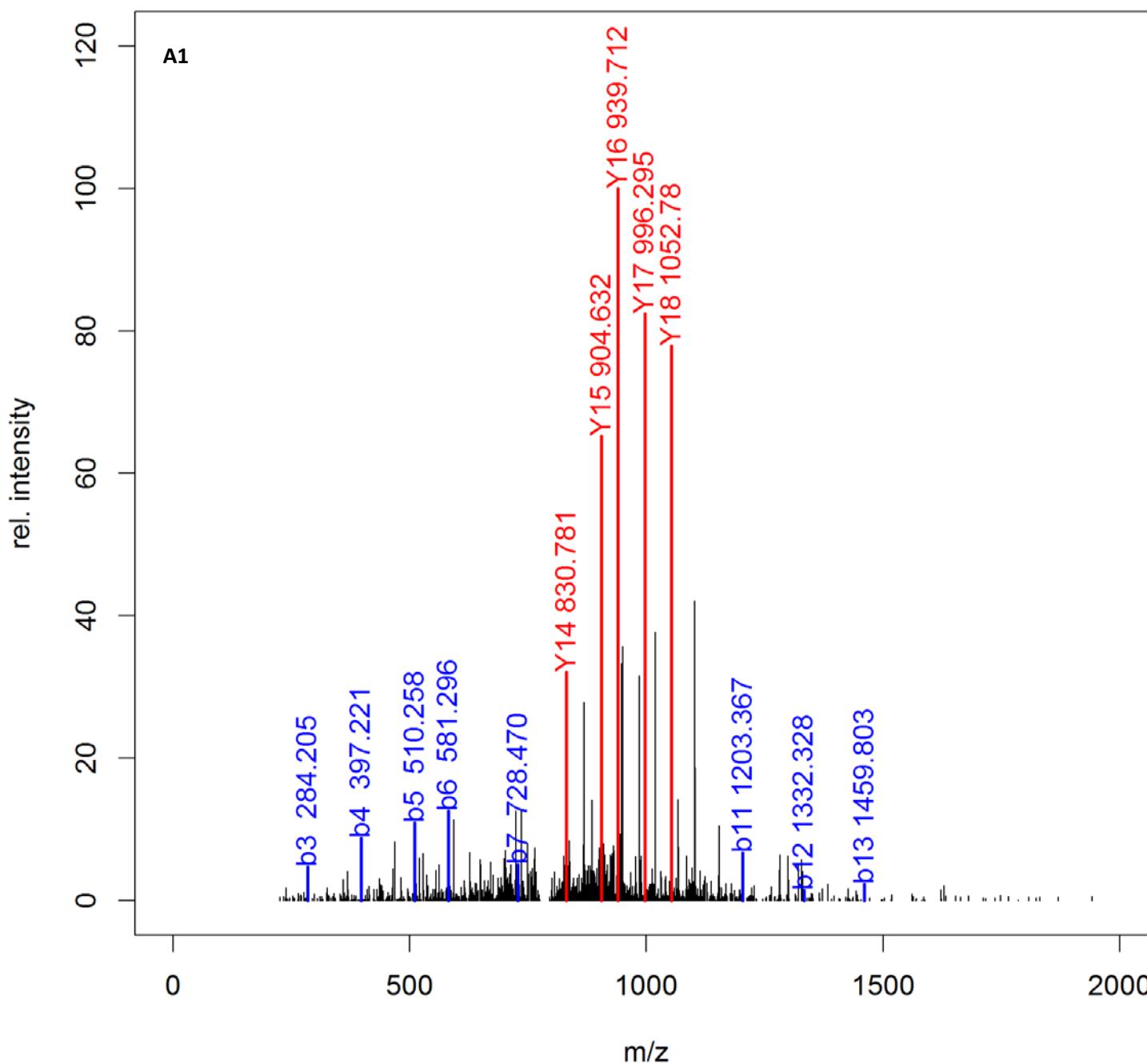
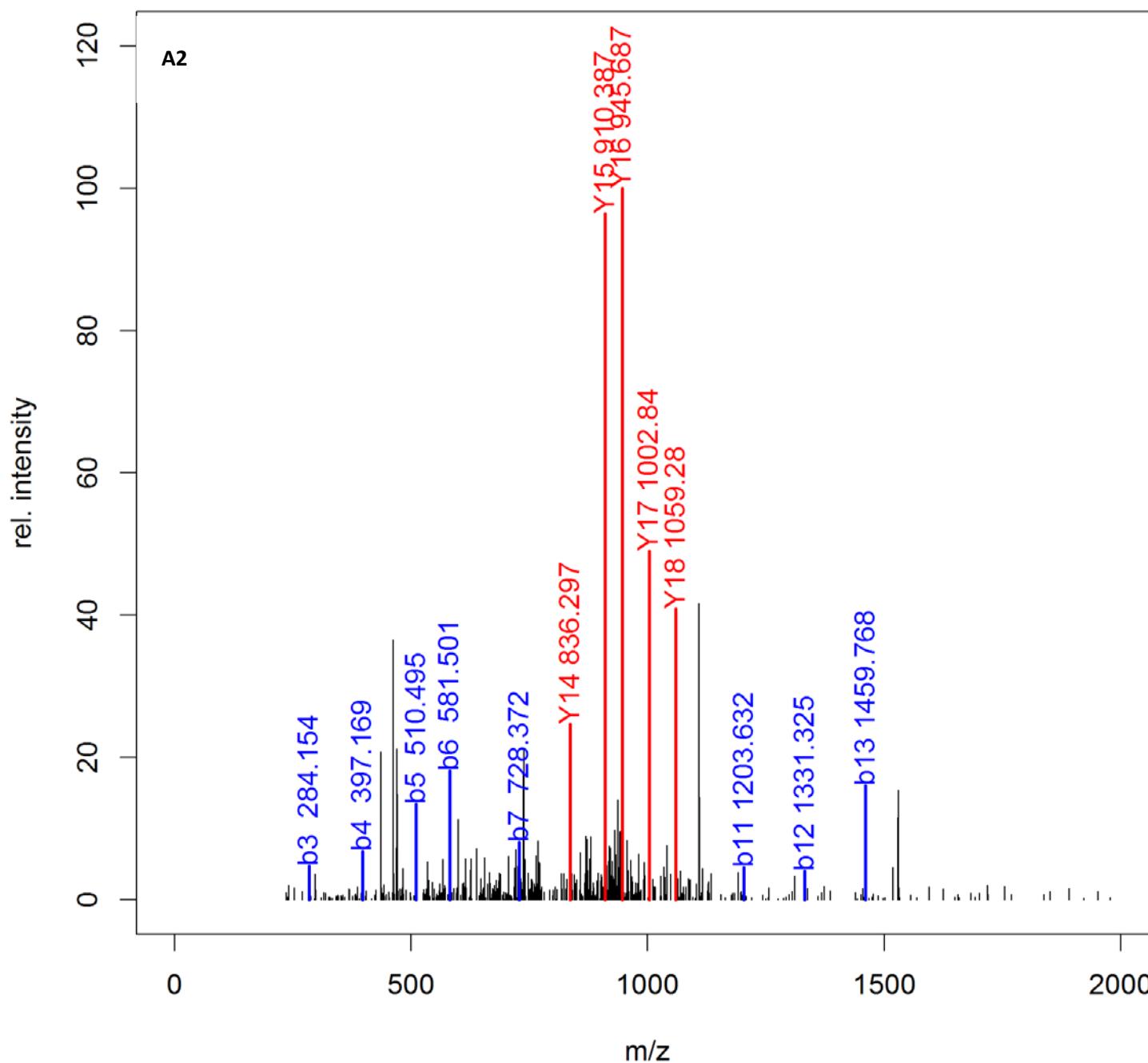
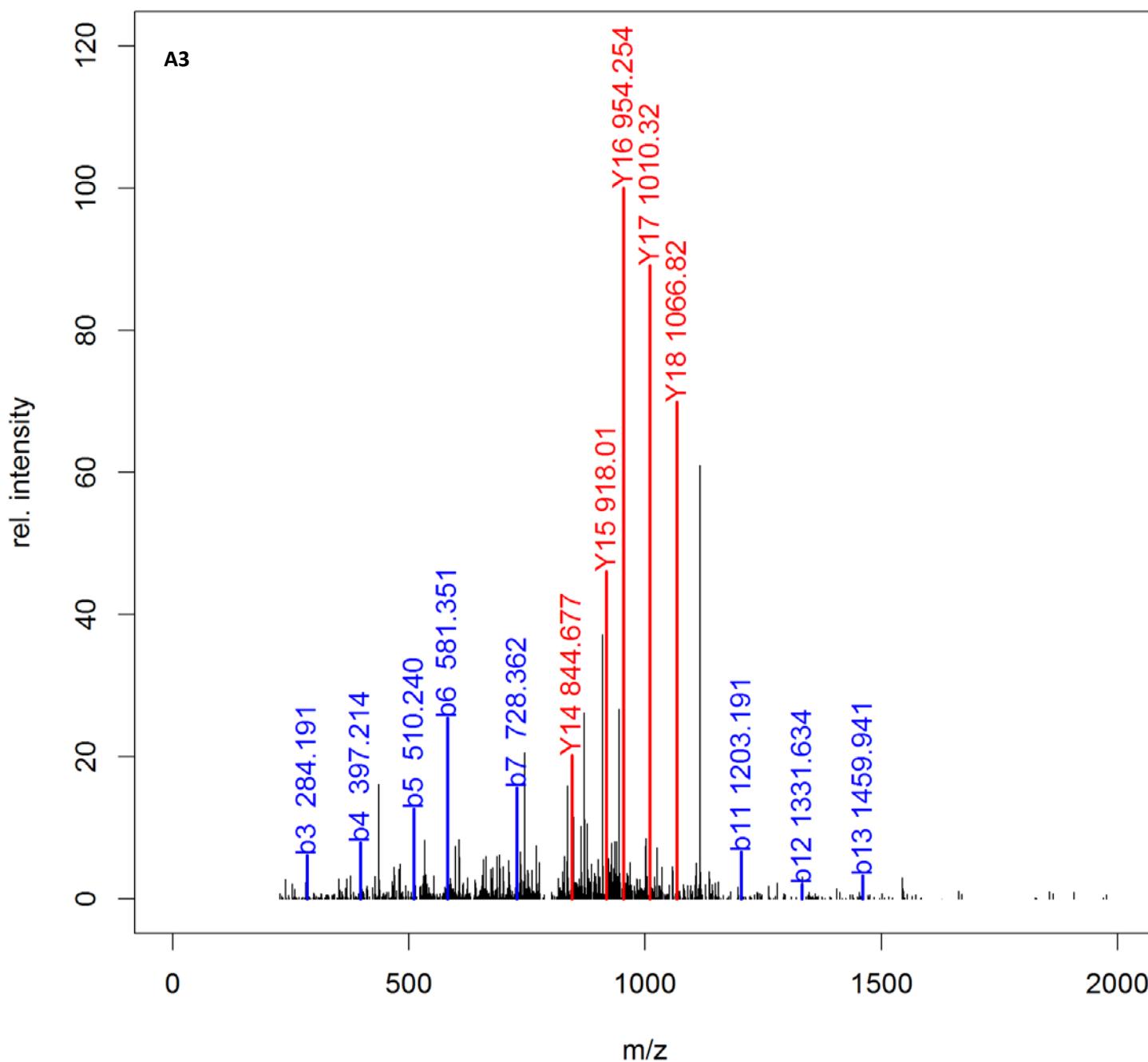
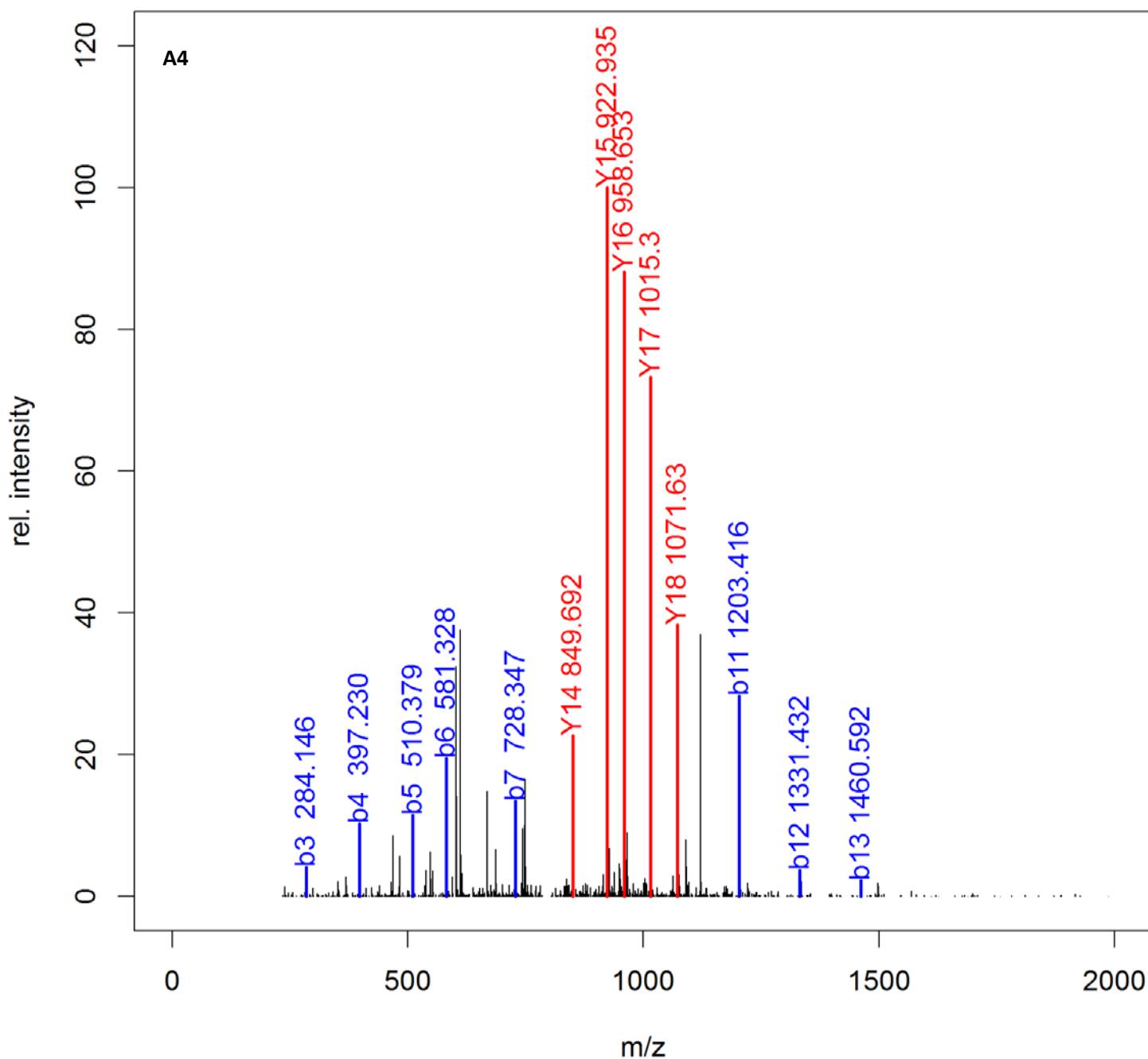
 Intercept 1 3.30349 0.3817 8.65 <.0001 2.52599 4.08099
 mnbmi 1 -0.04377 0.0145 -3.02 0.005 0.22162 -0.07331 -0.01423


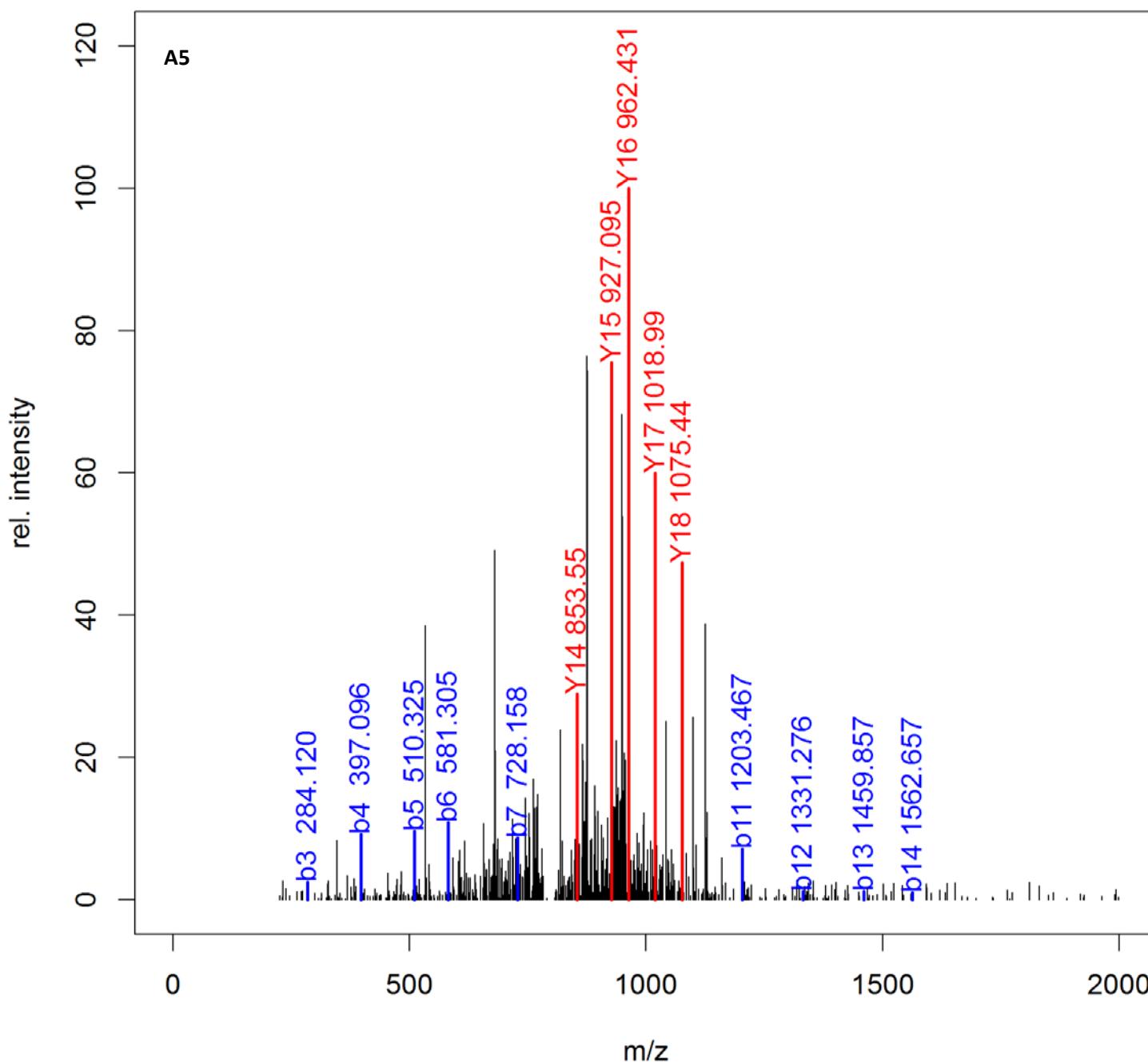
Figure S1 – The MS2 spectra of T3 and its related adducts coded from A1 to A43. The headline in each figure indicates the raw file name_sample ID_Precursor scan #_MIM_RT_Tandem scan #.

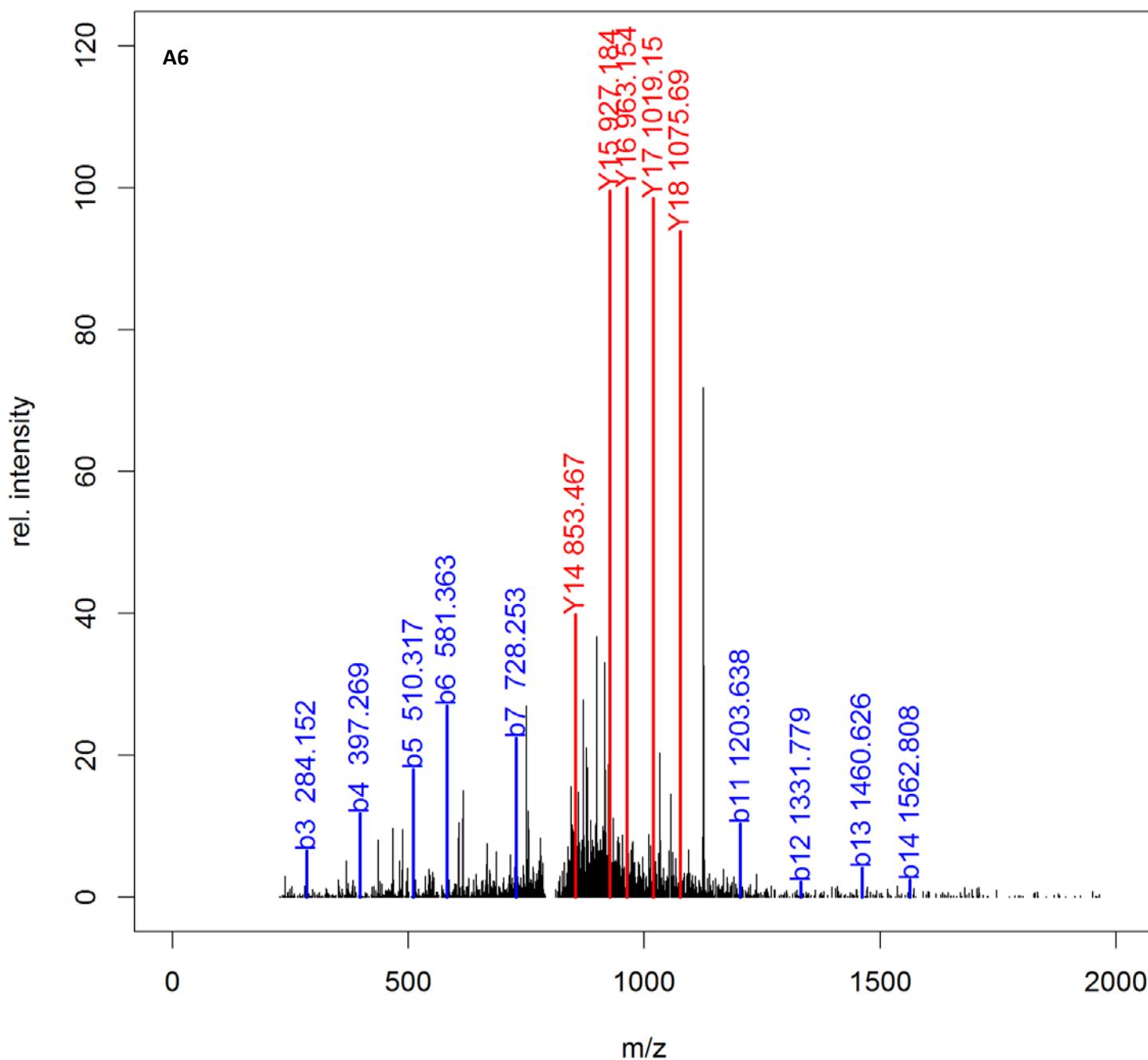


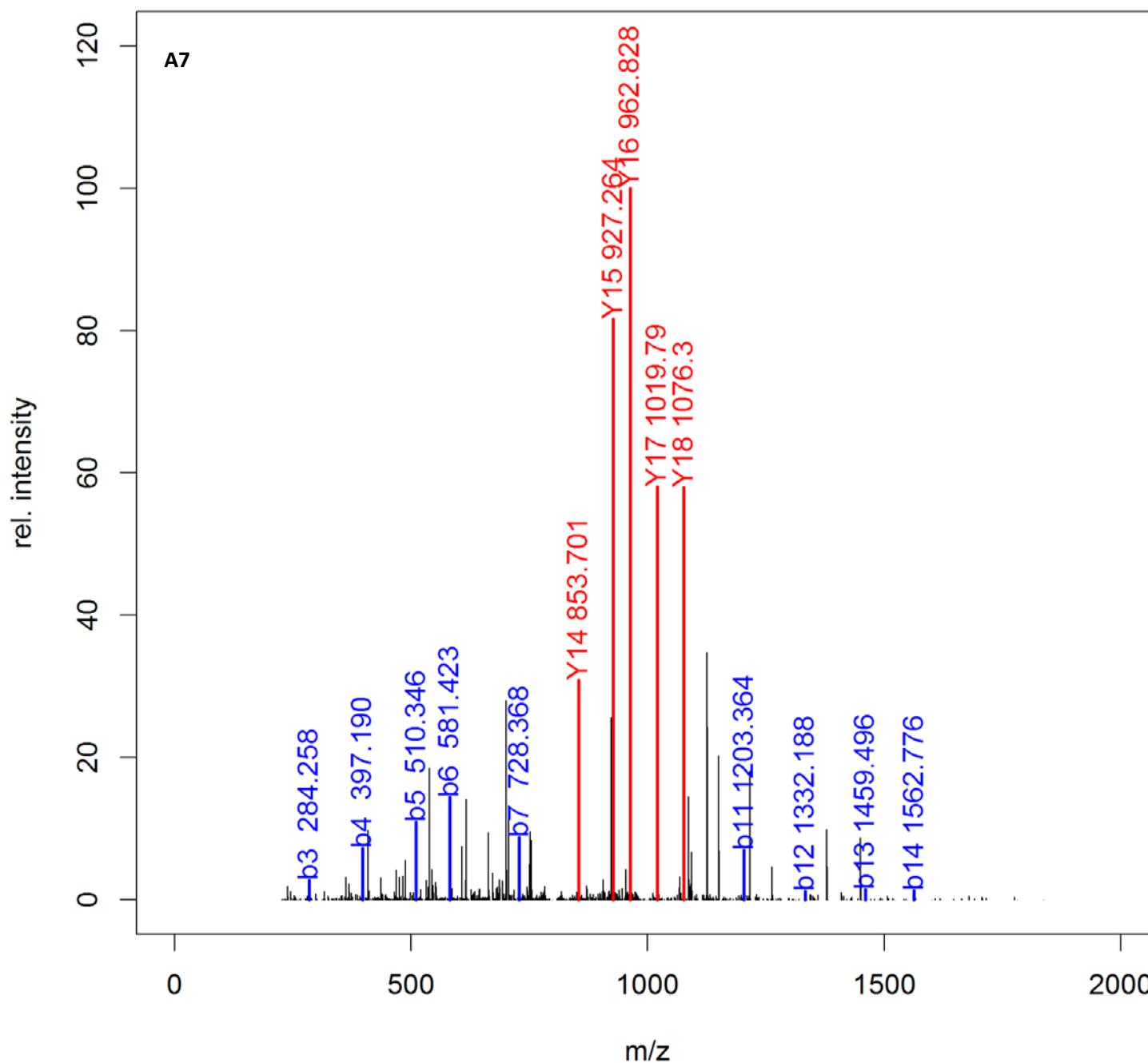


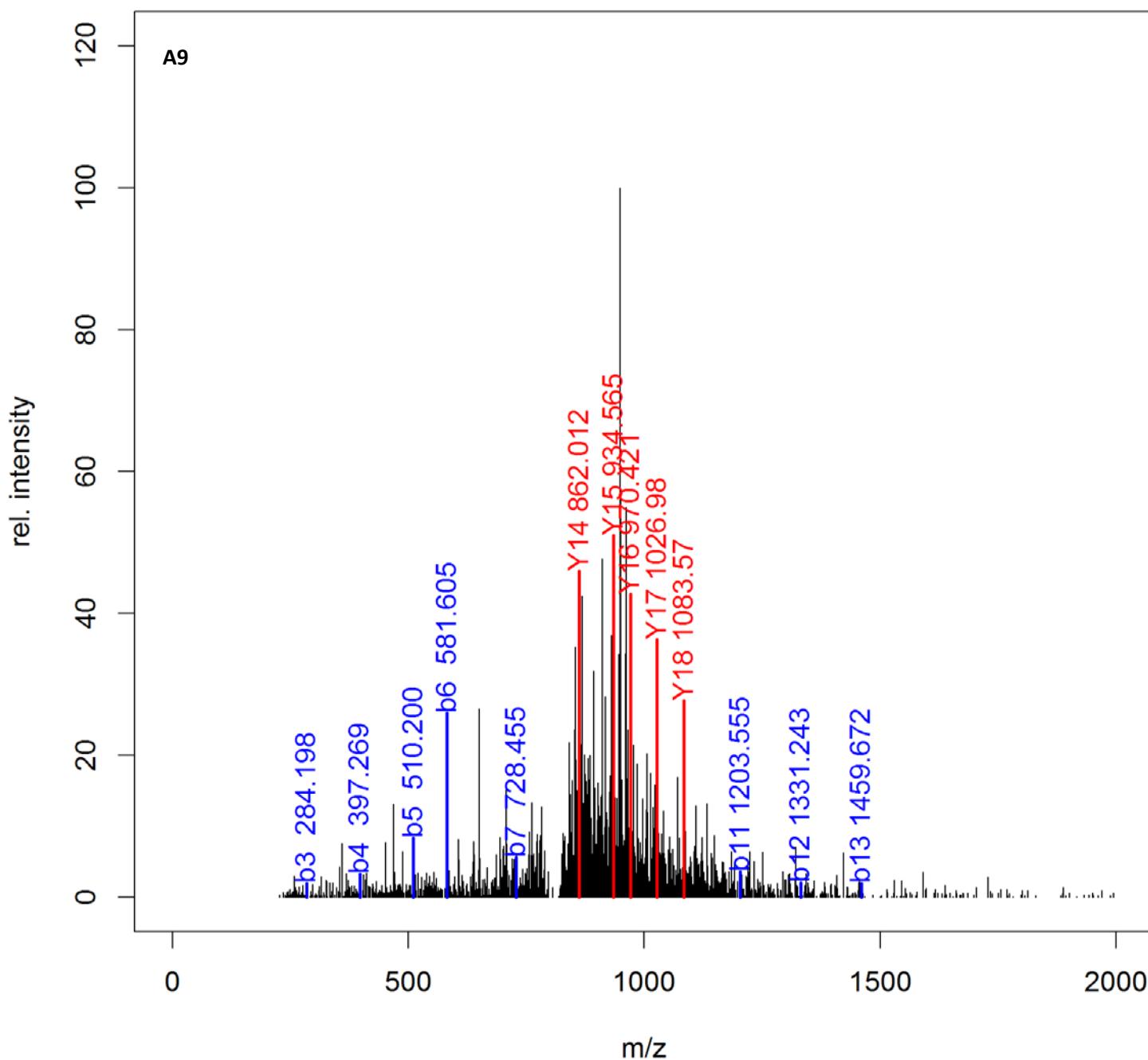


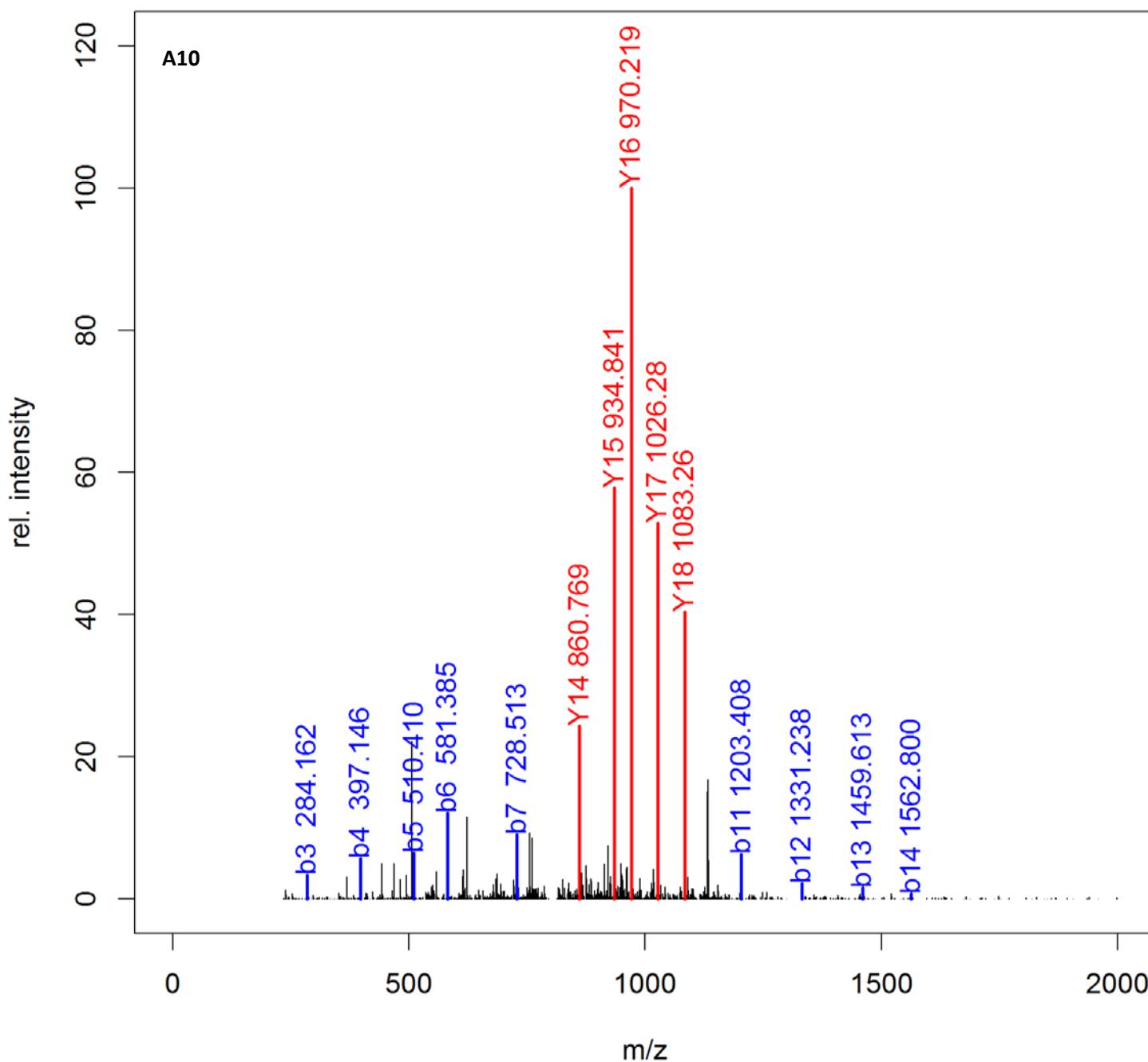


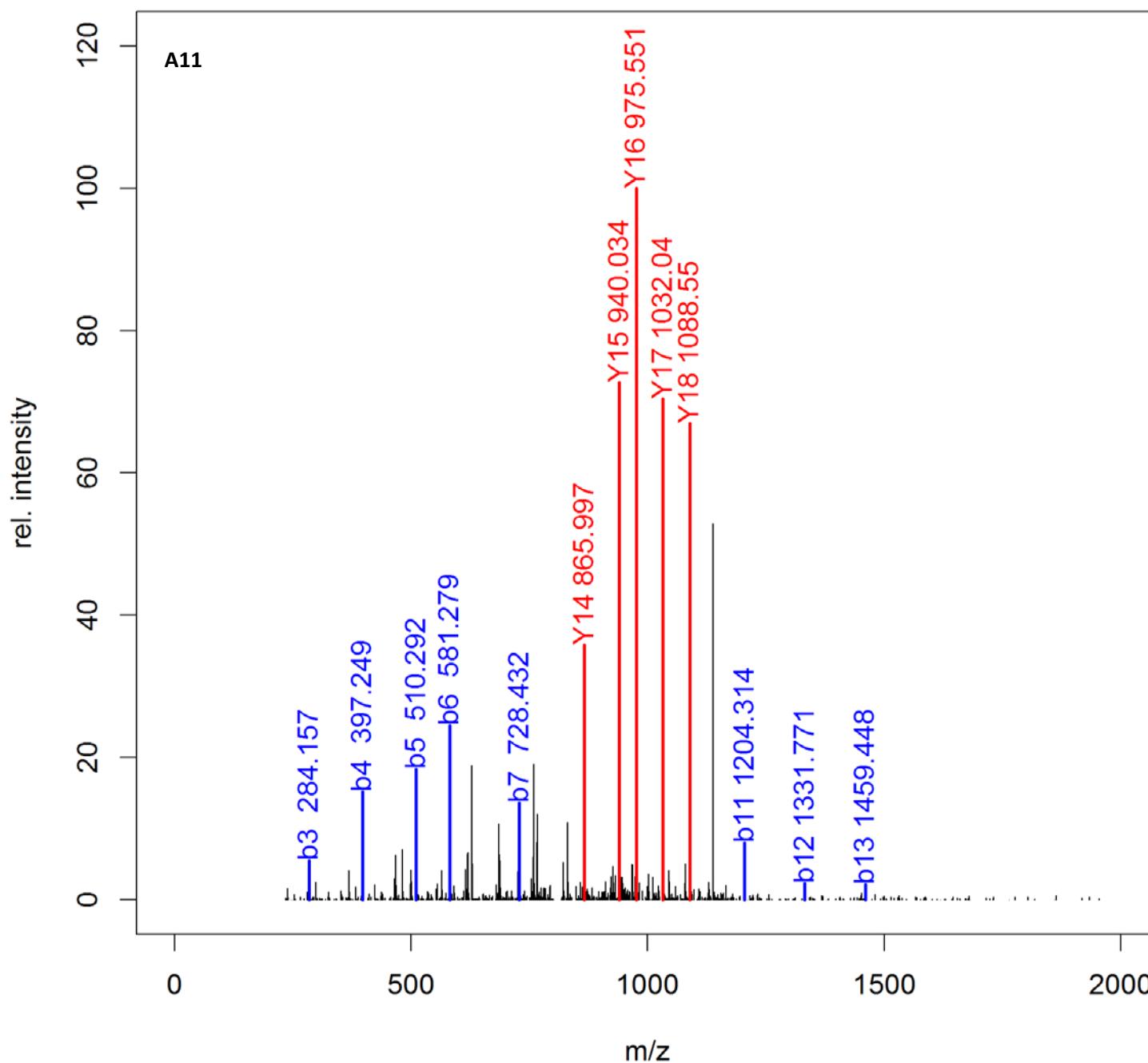


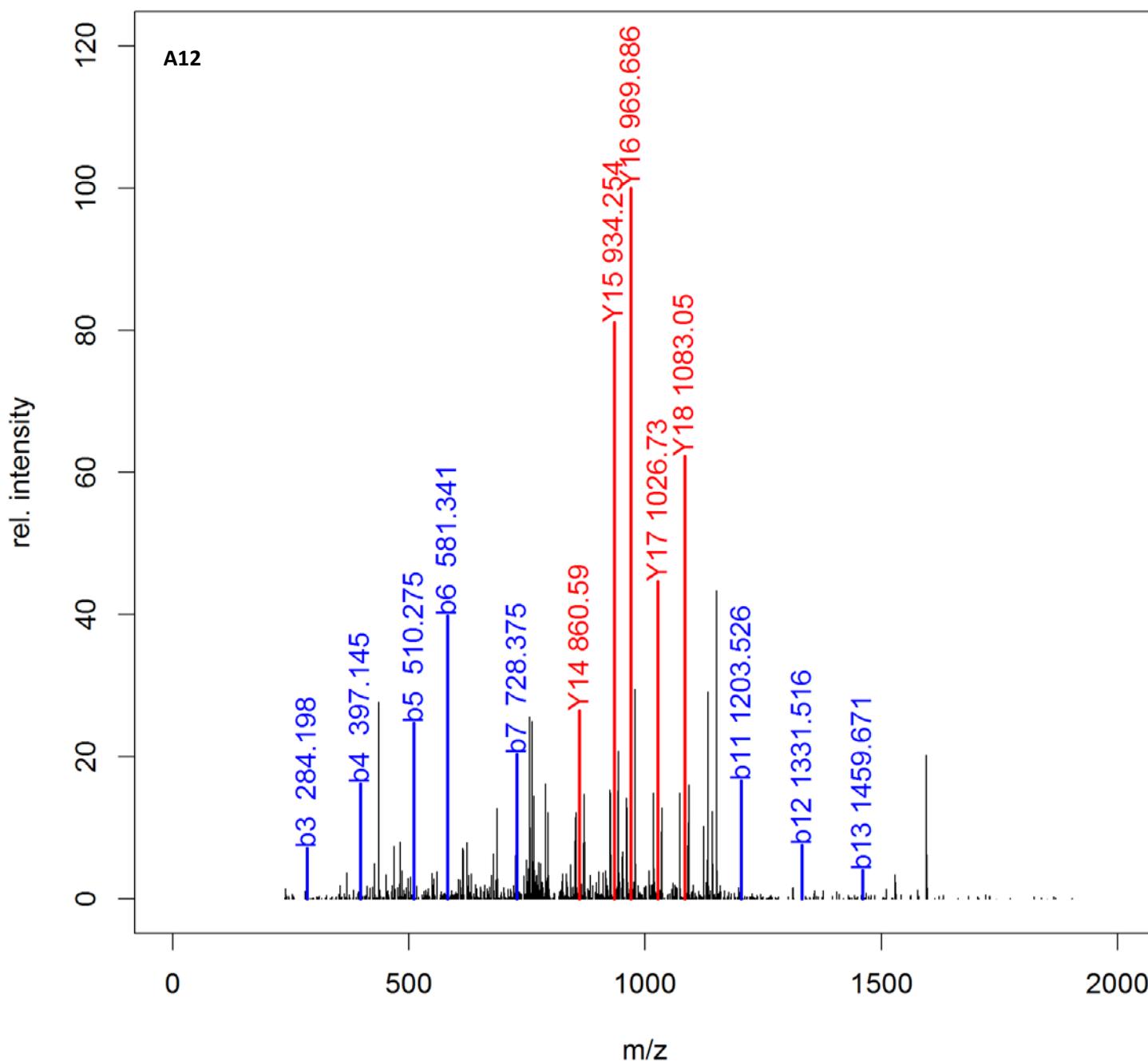


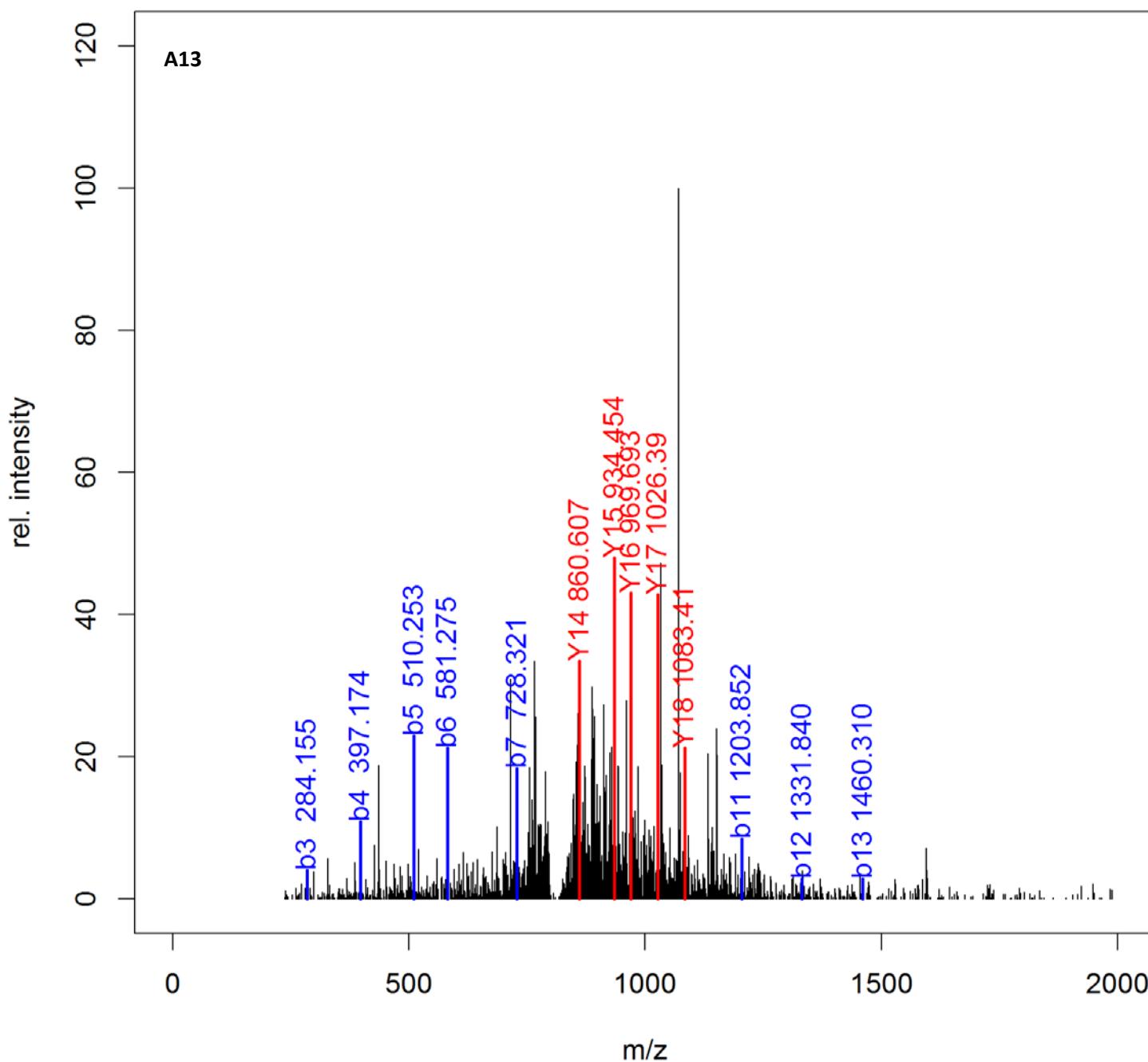


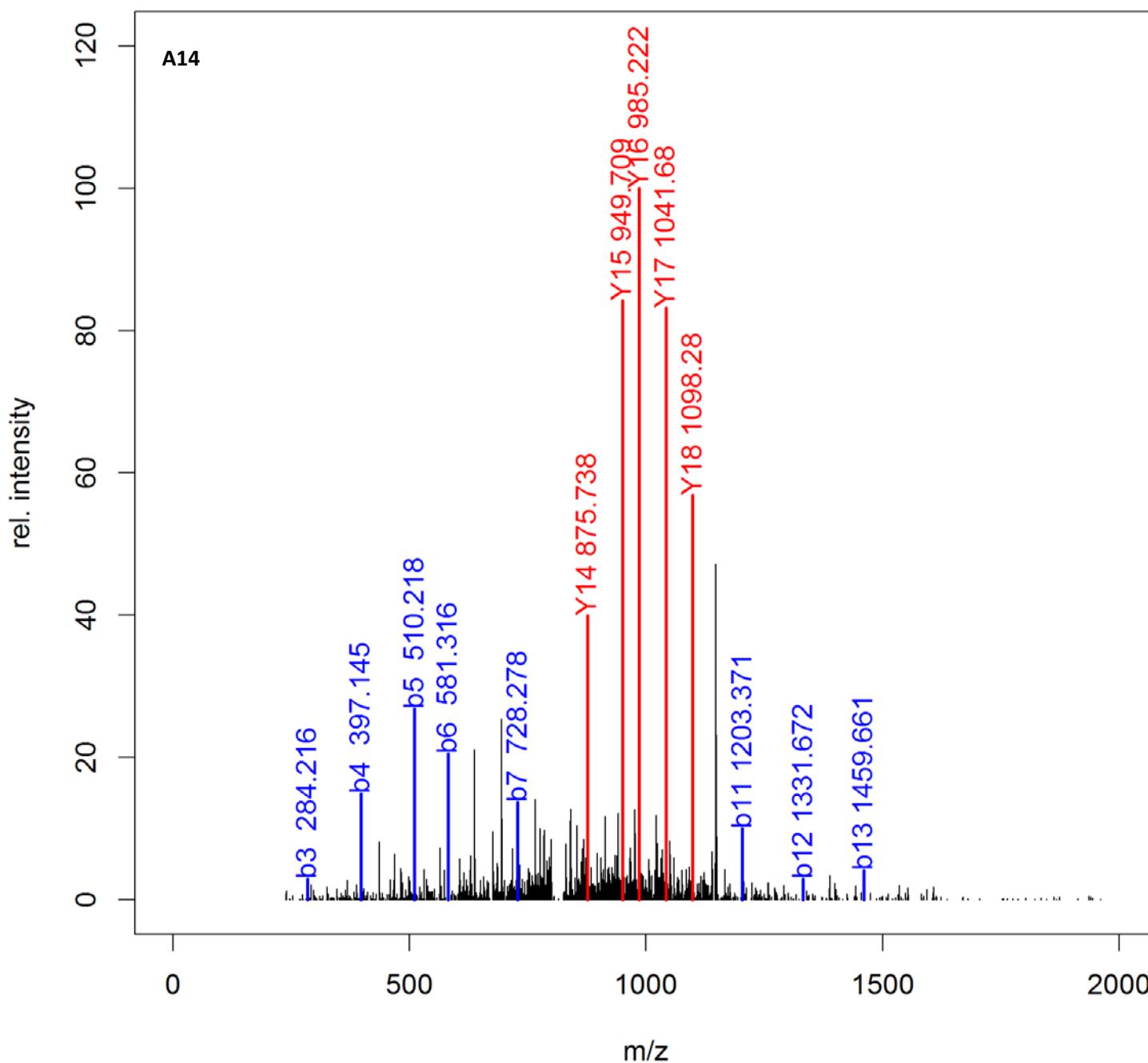


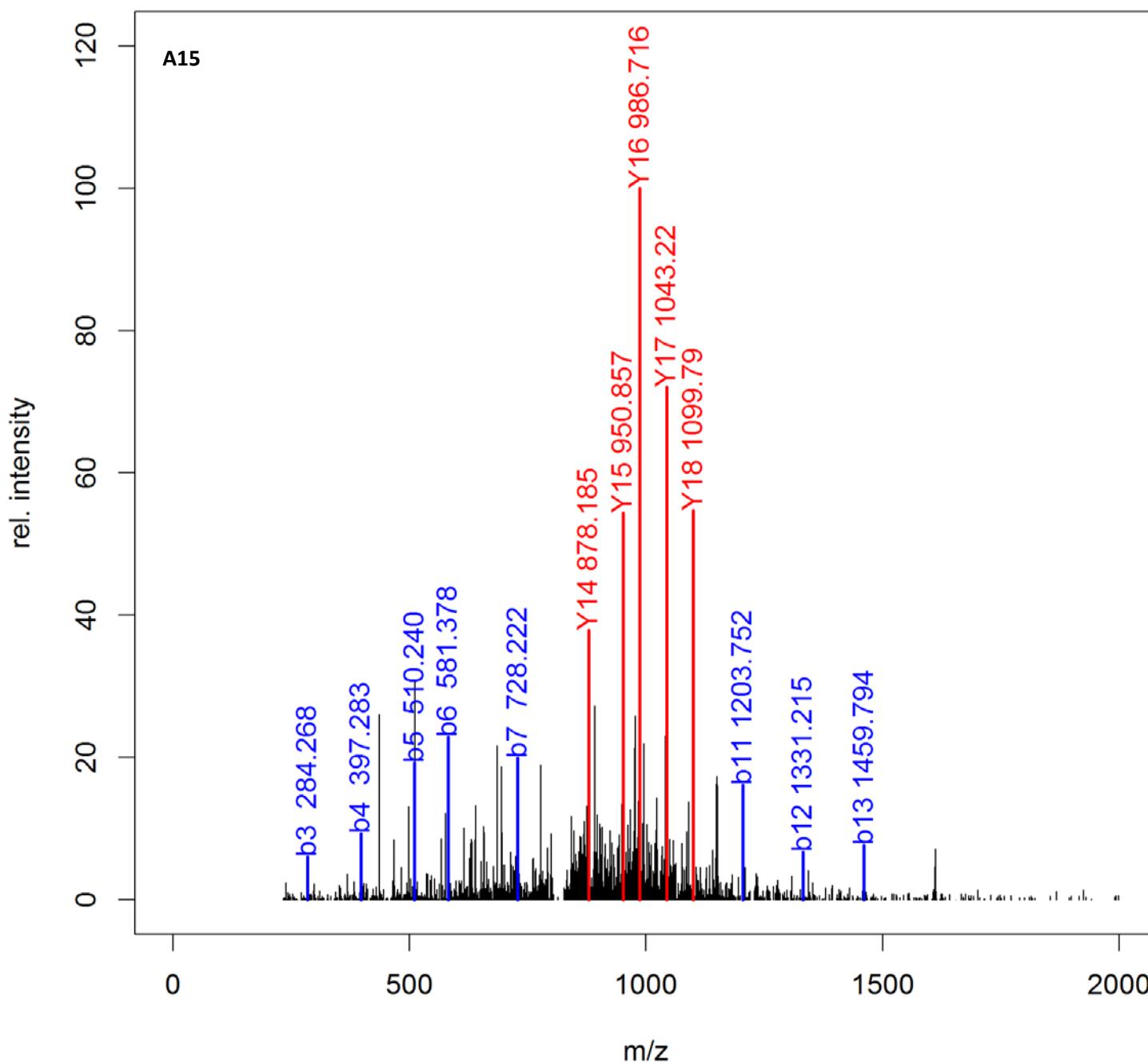


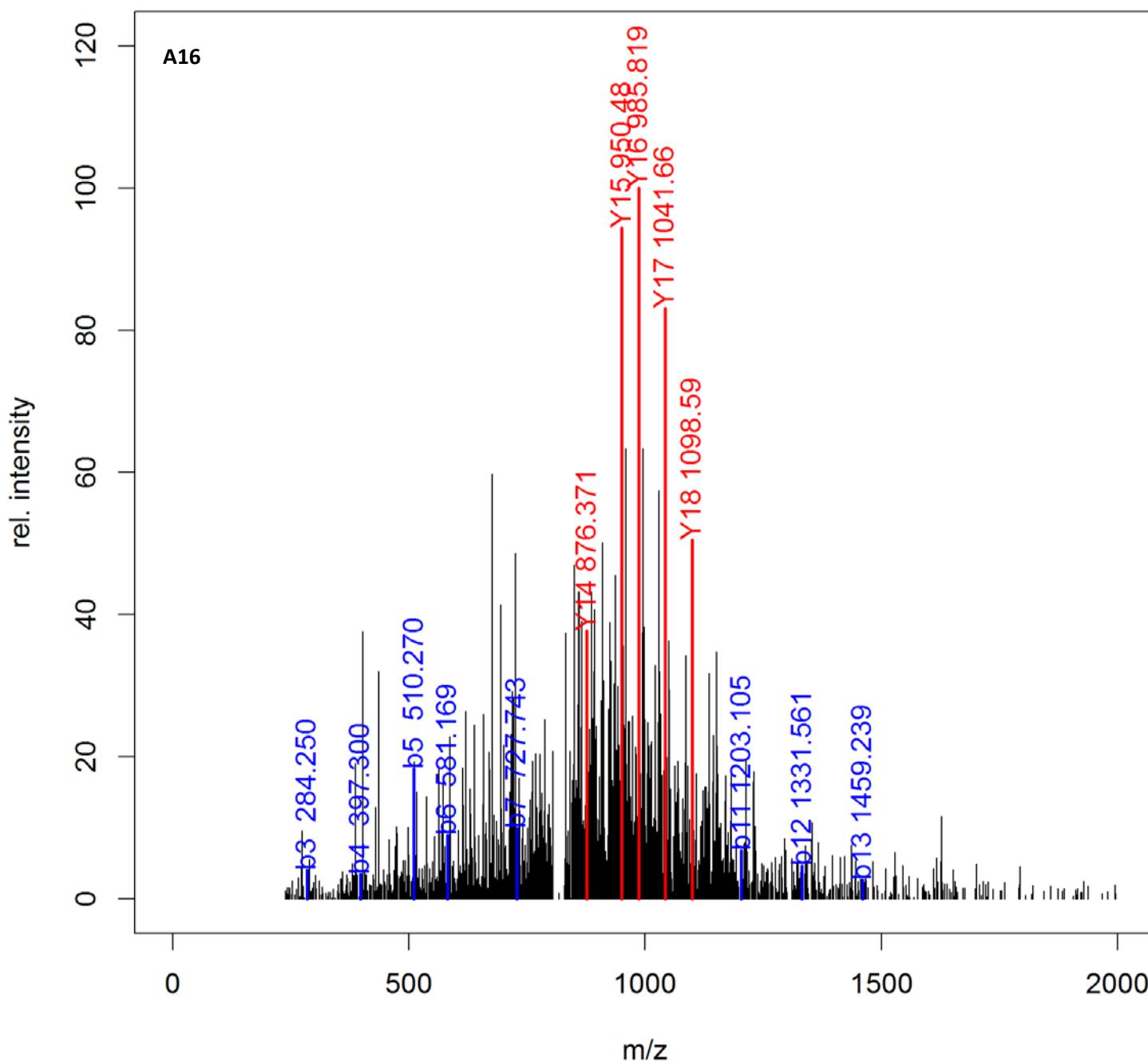


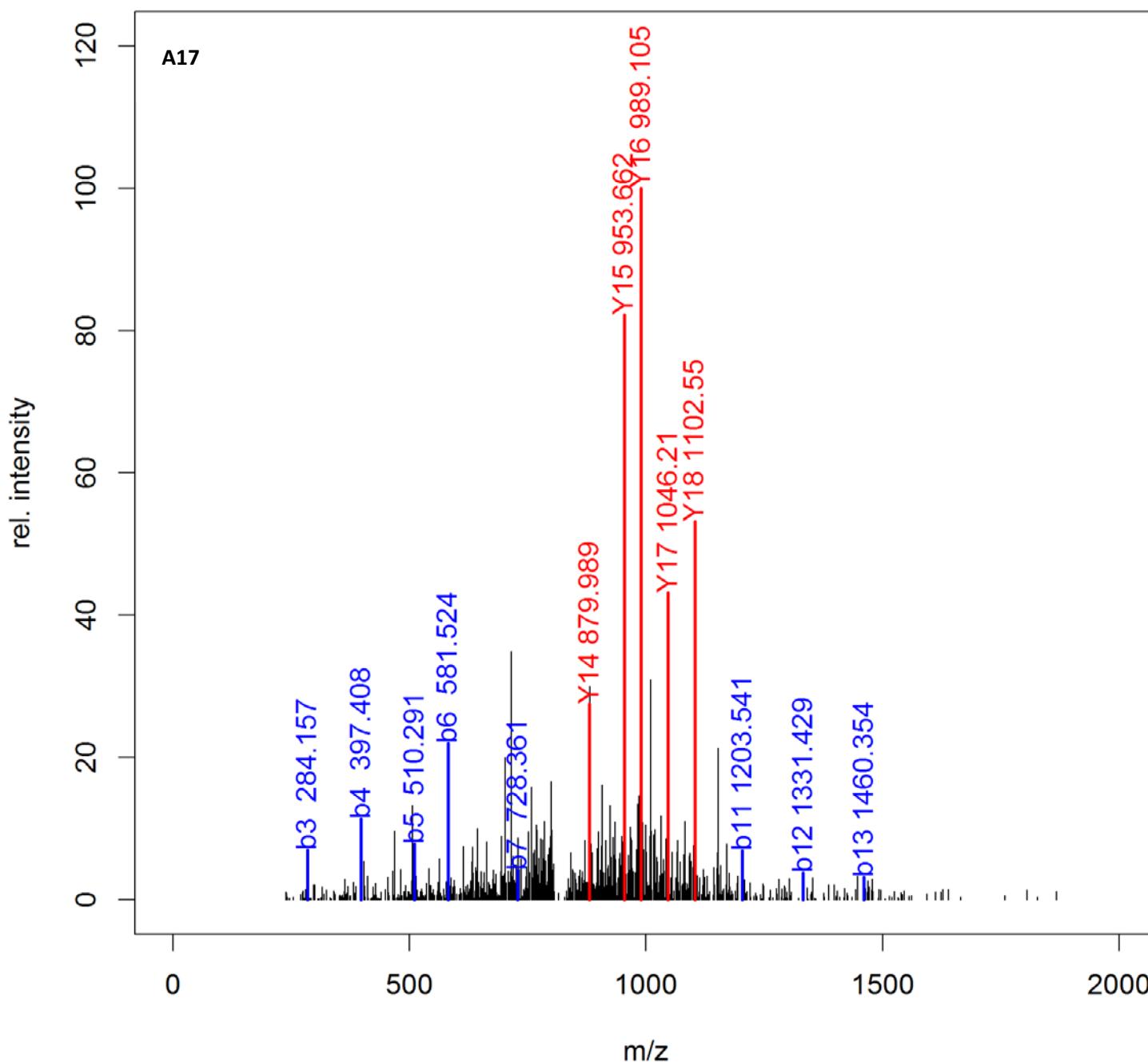


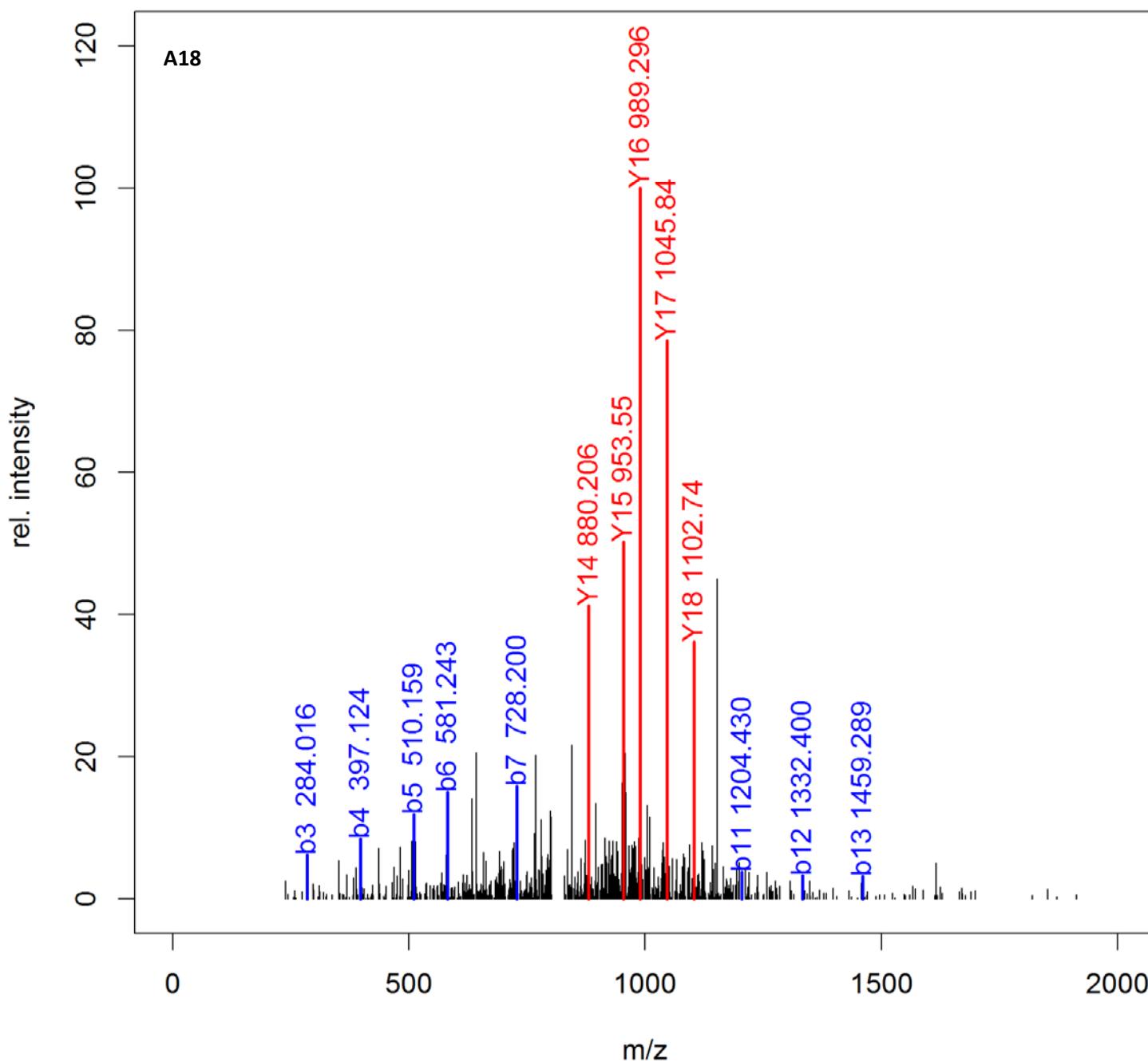


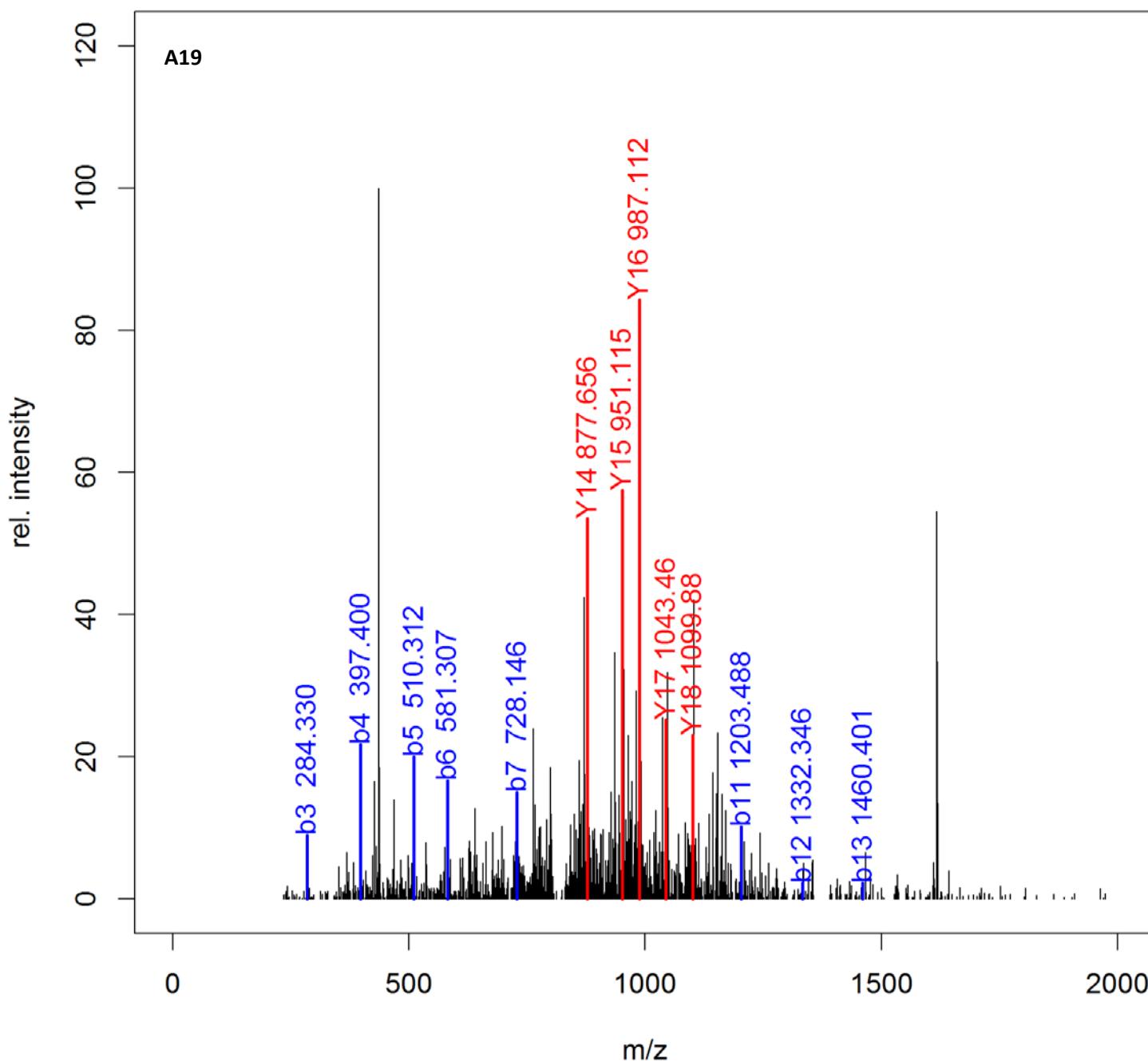


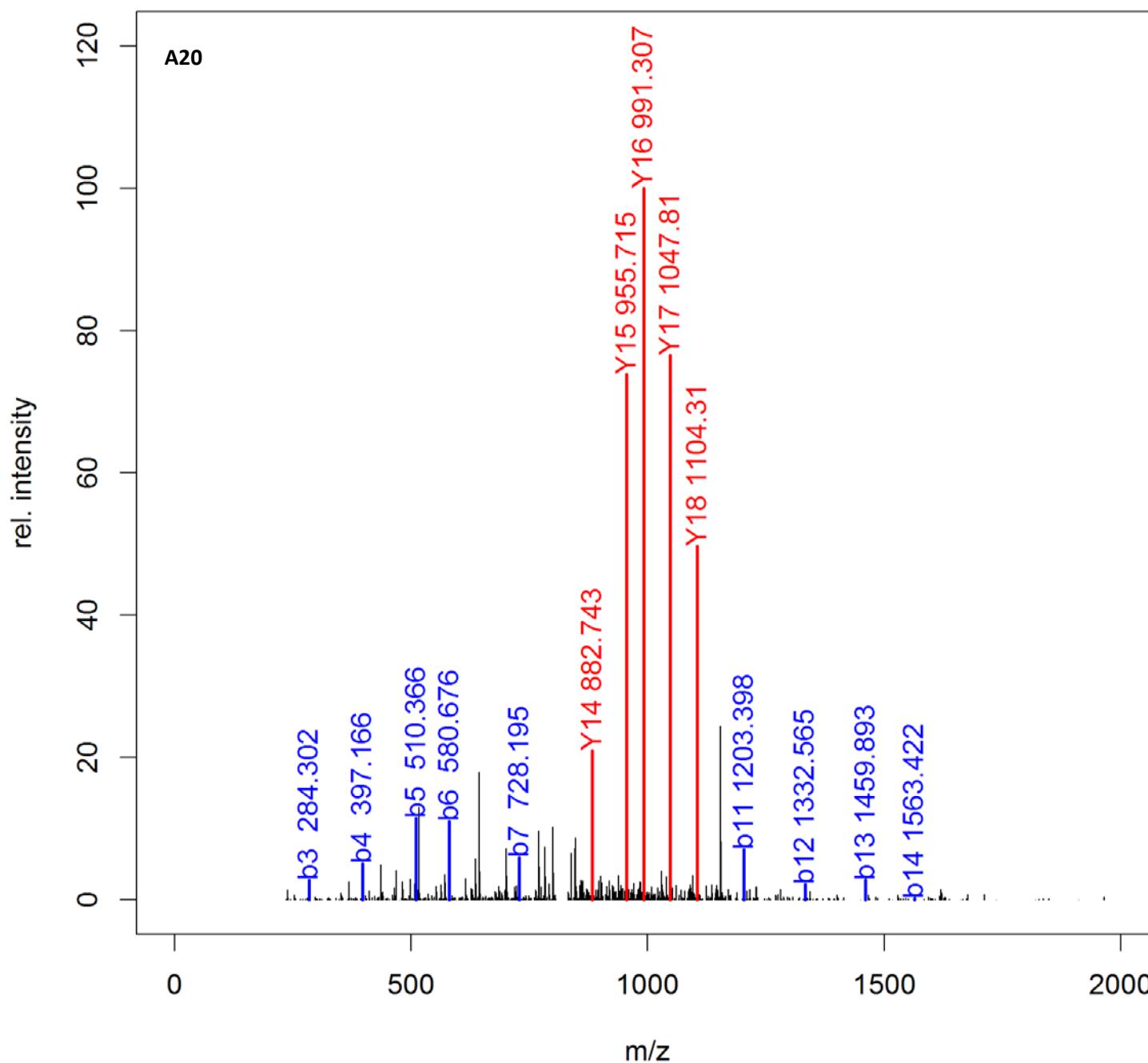


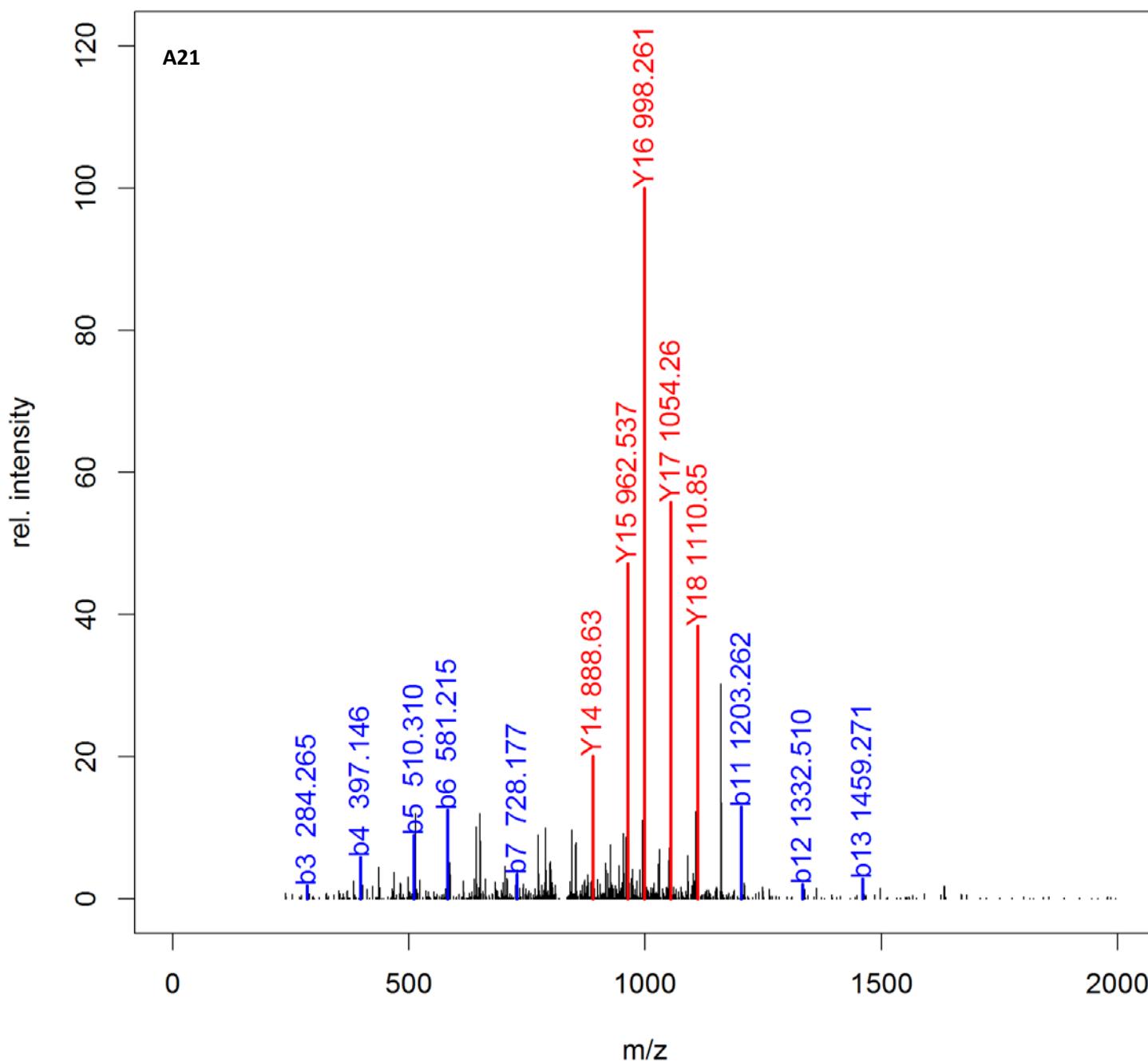


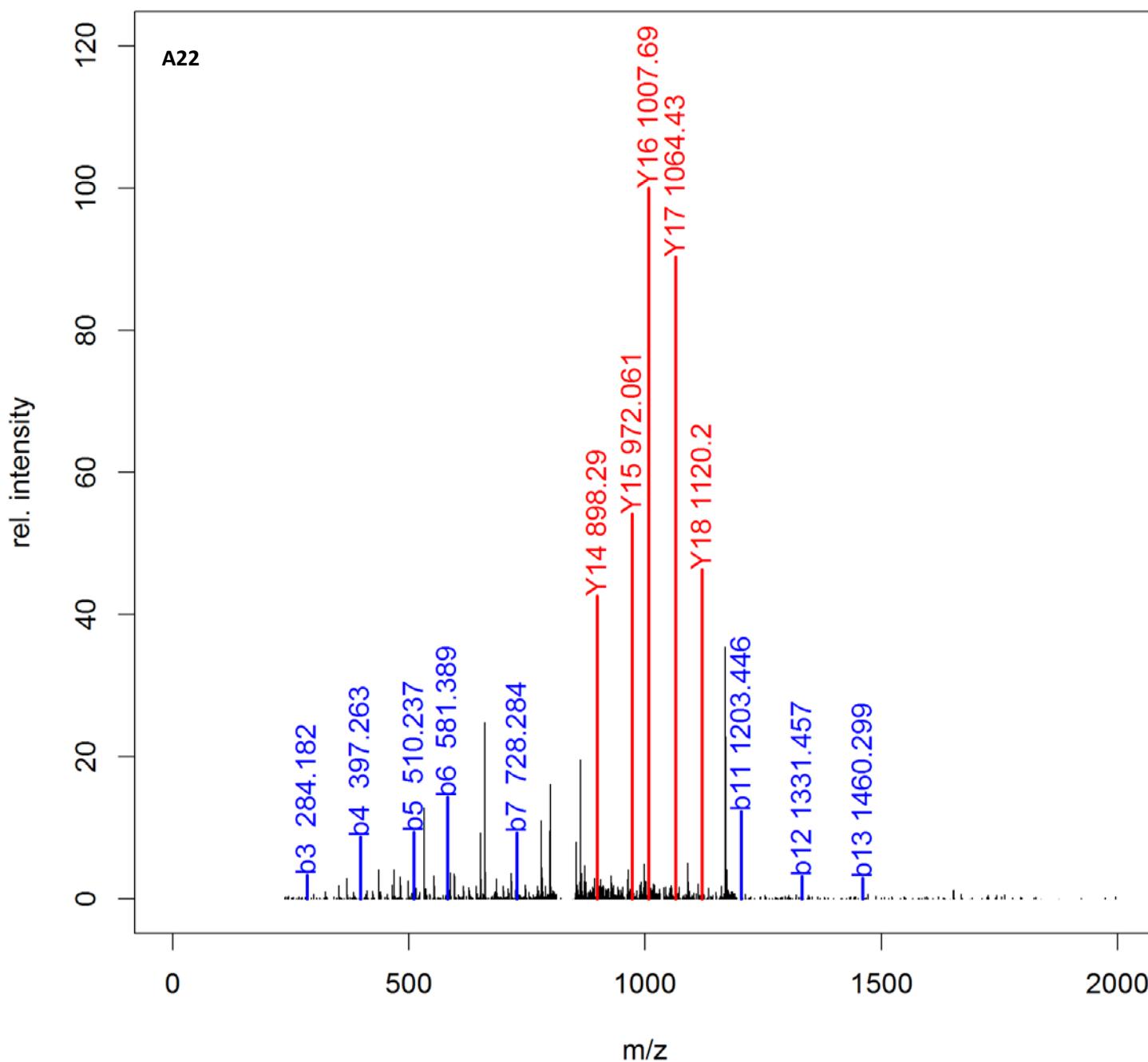


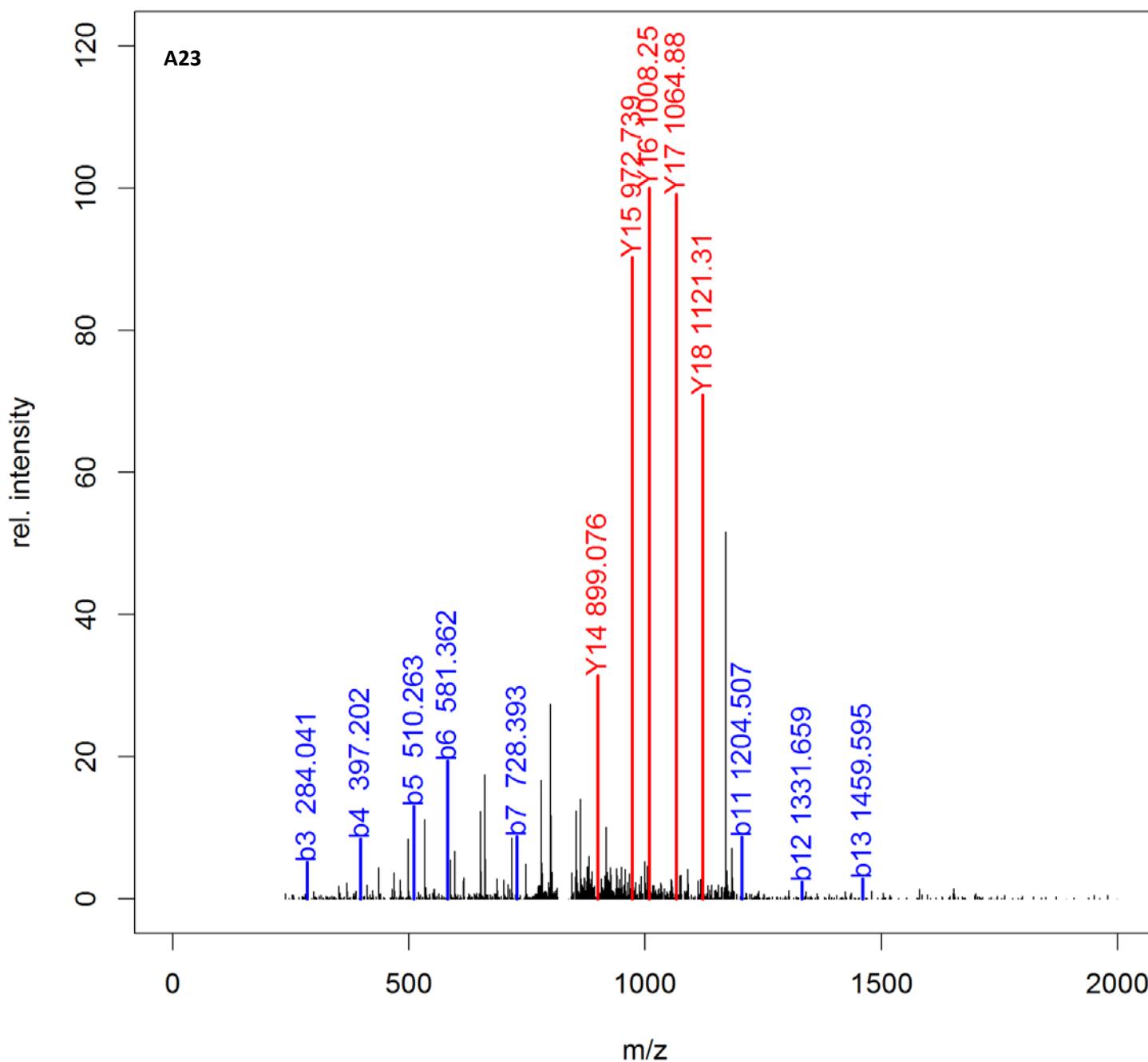


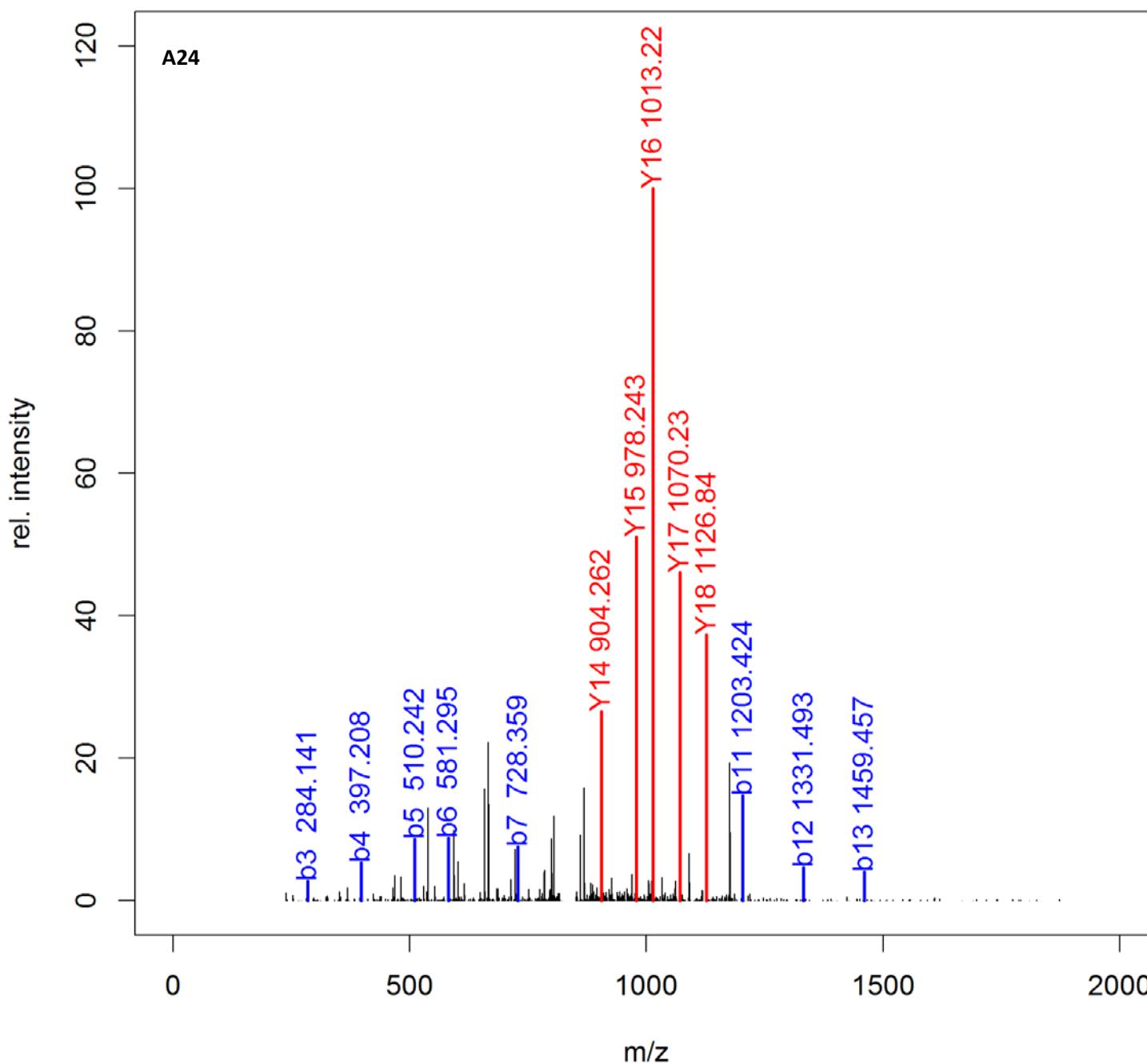


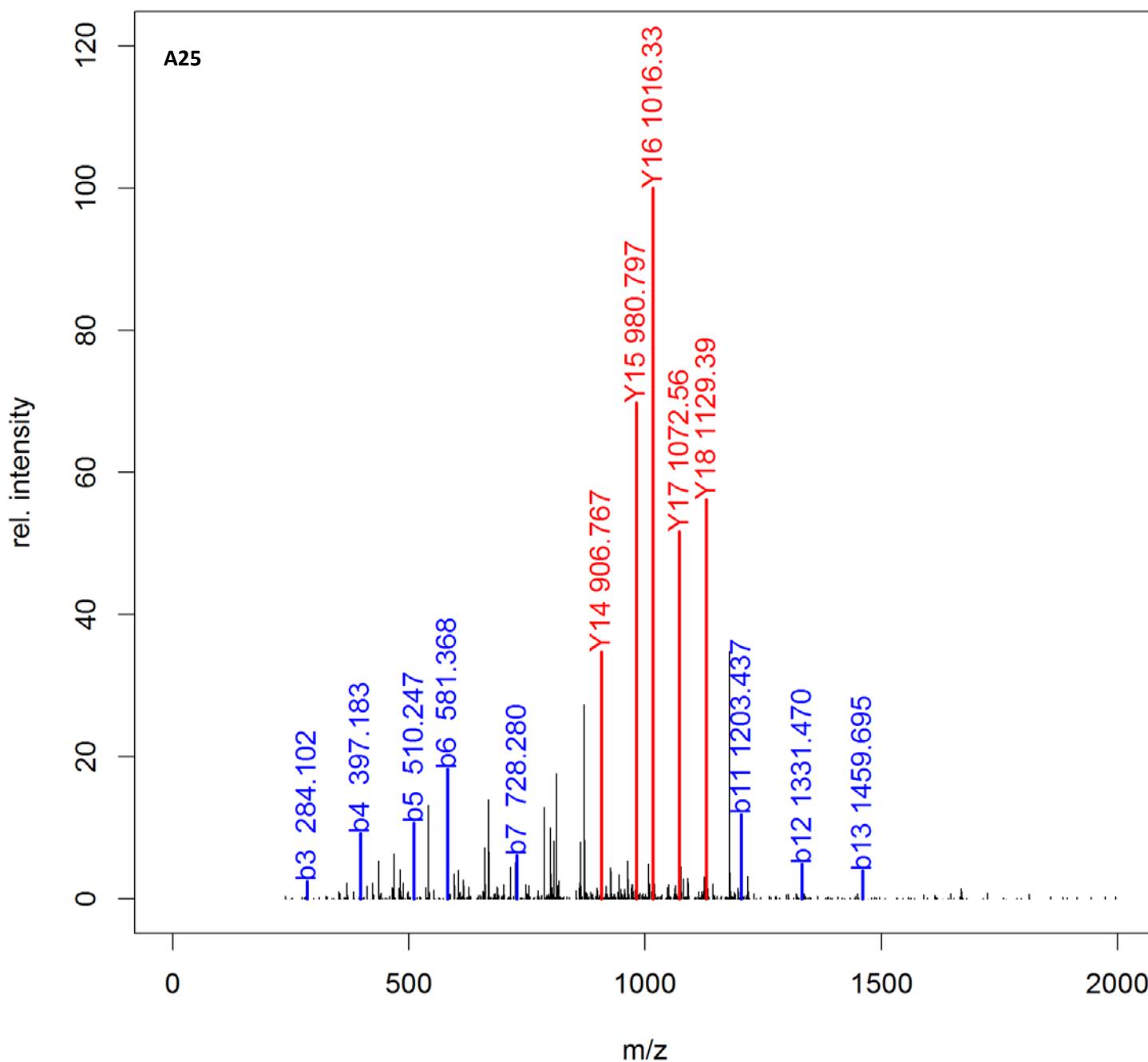


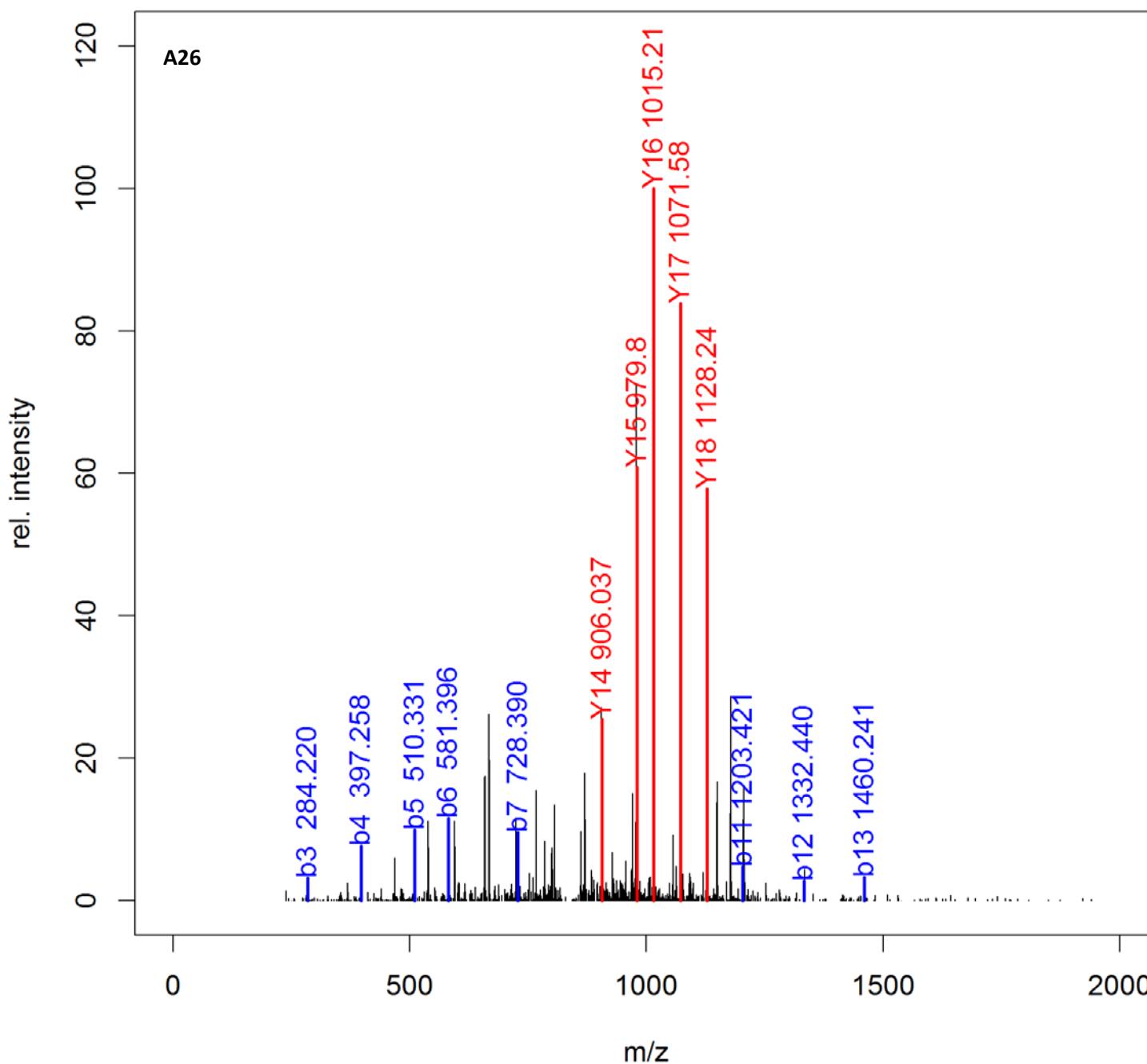


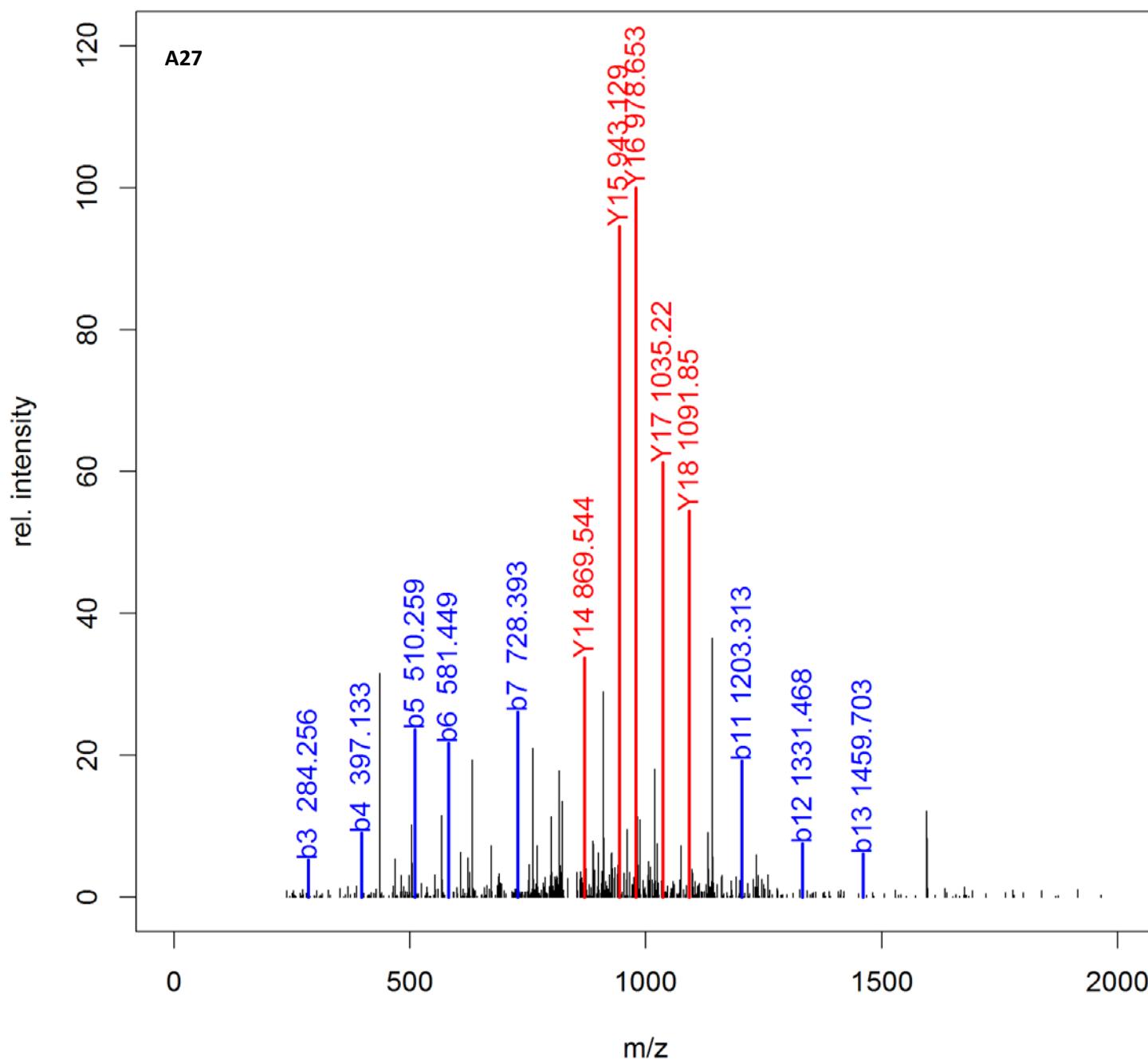


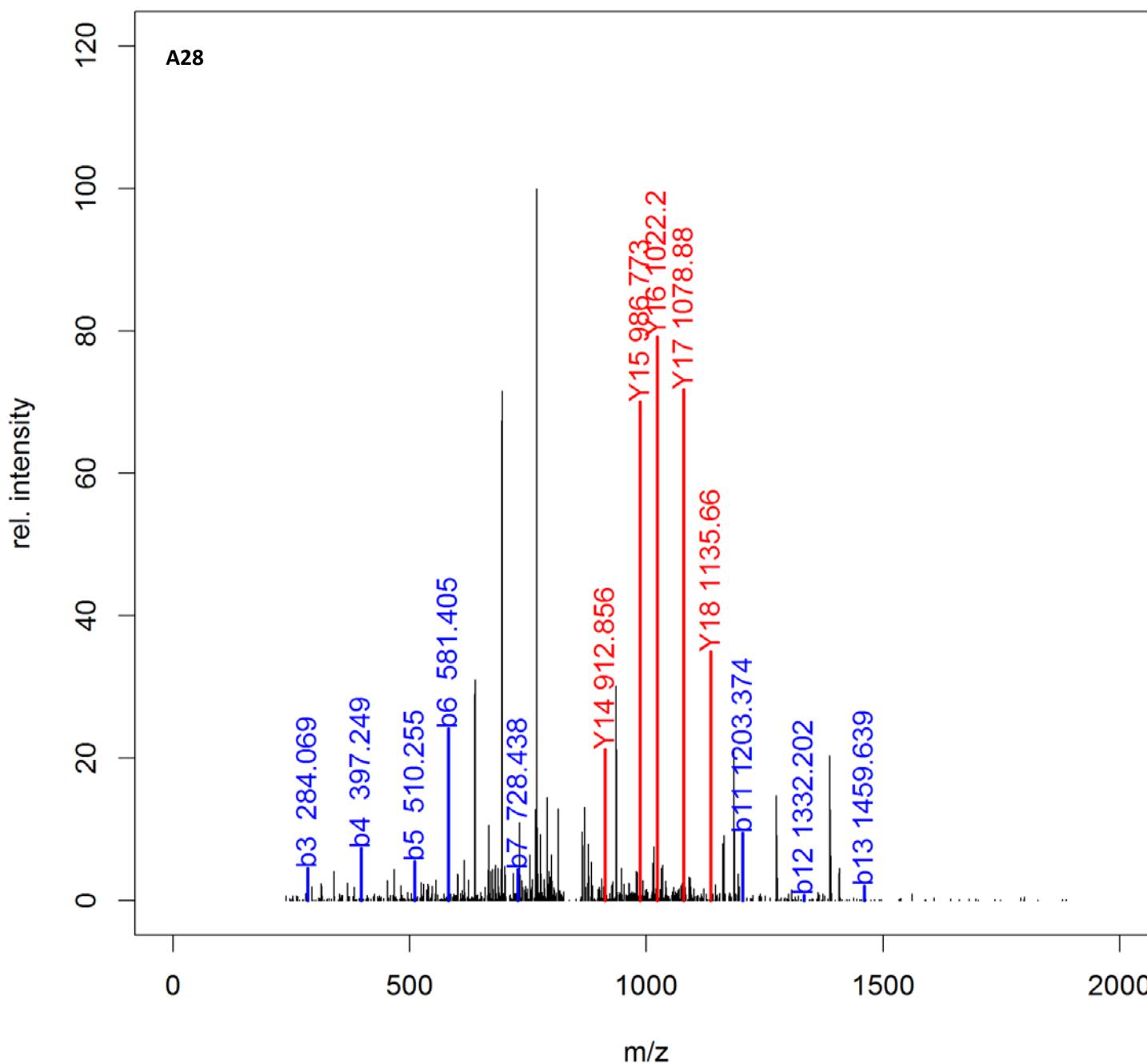


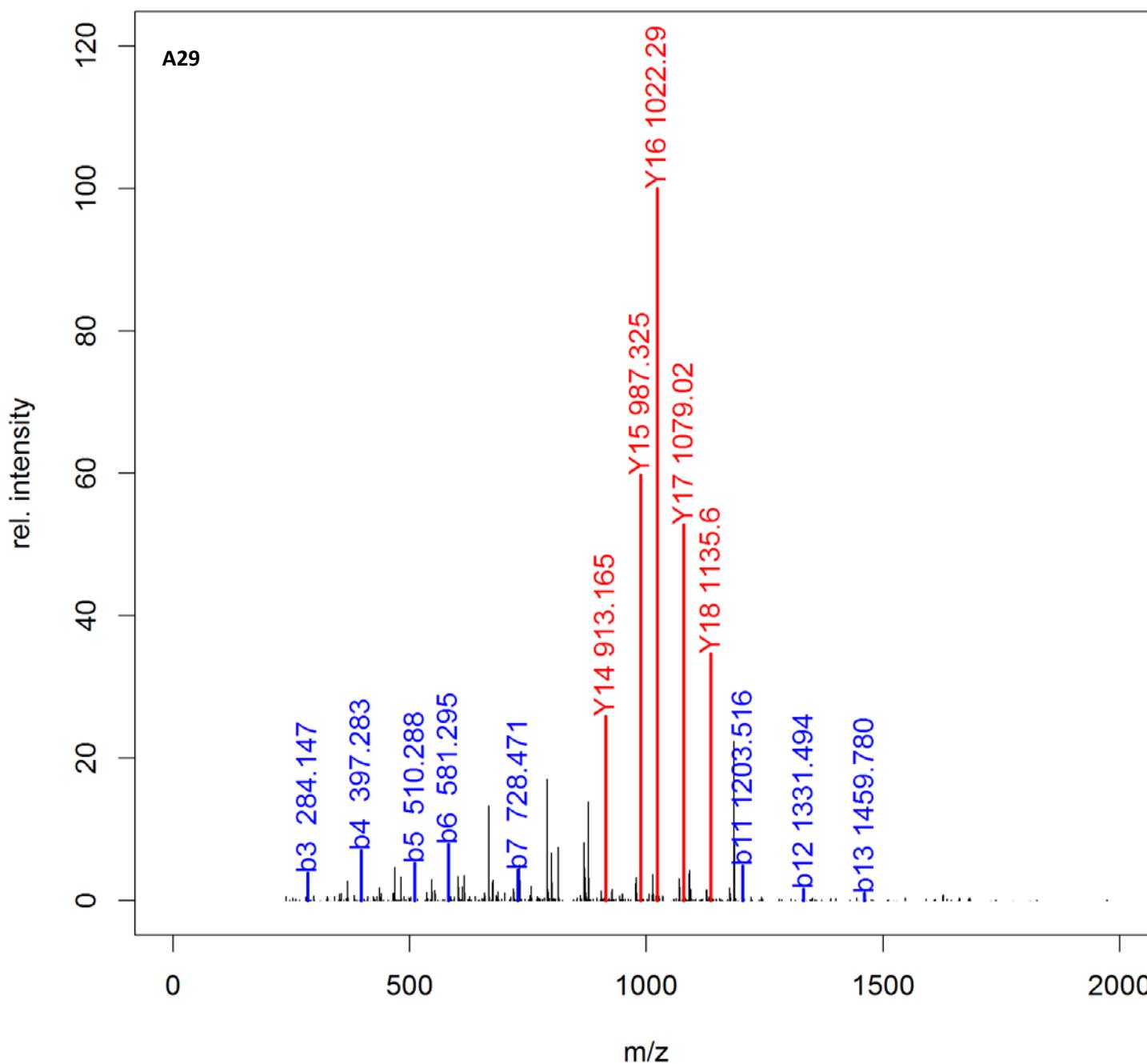


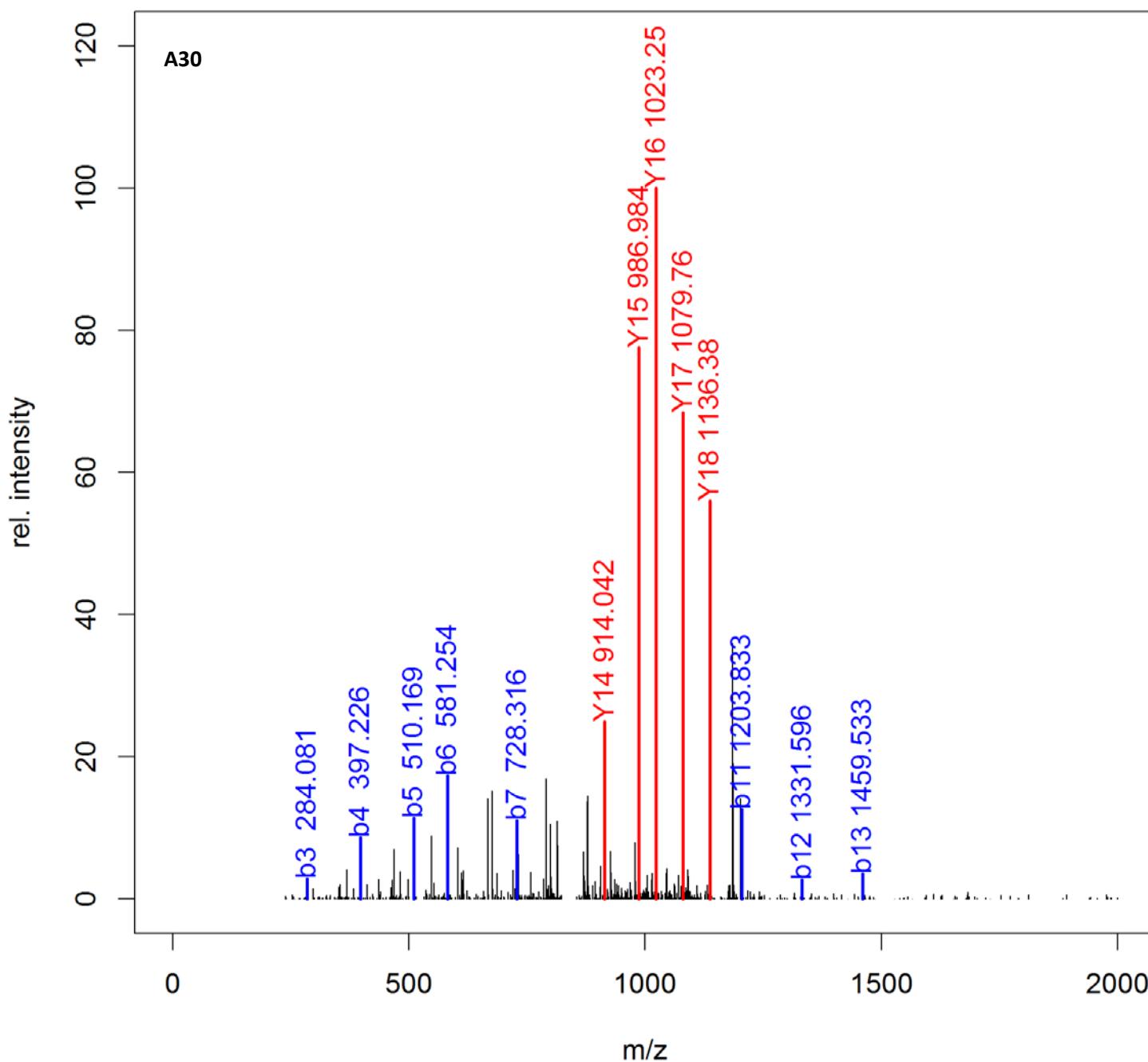


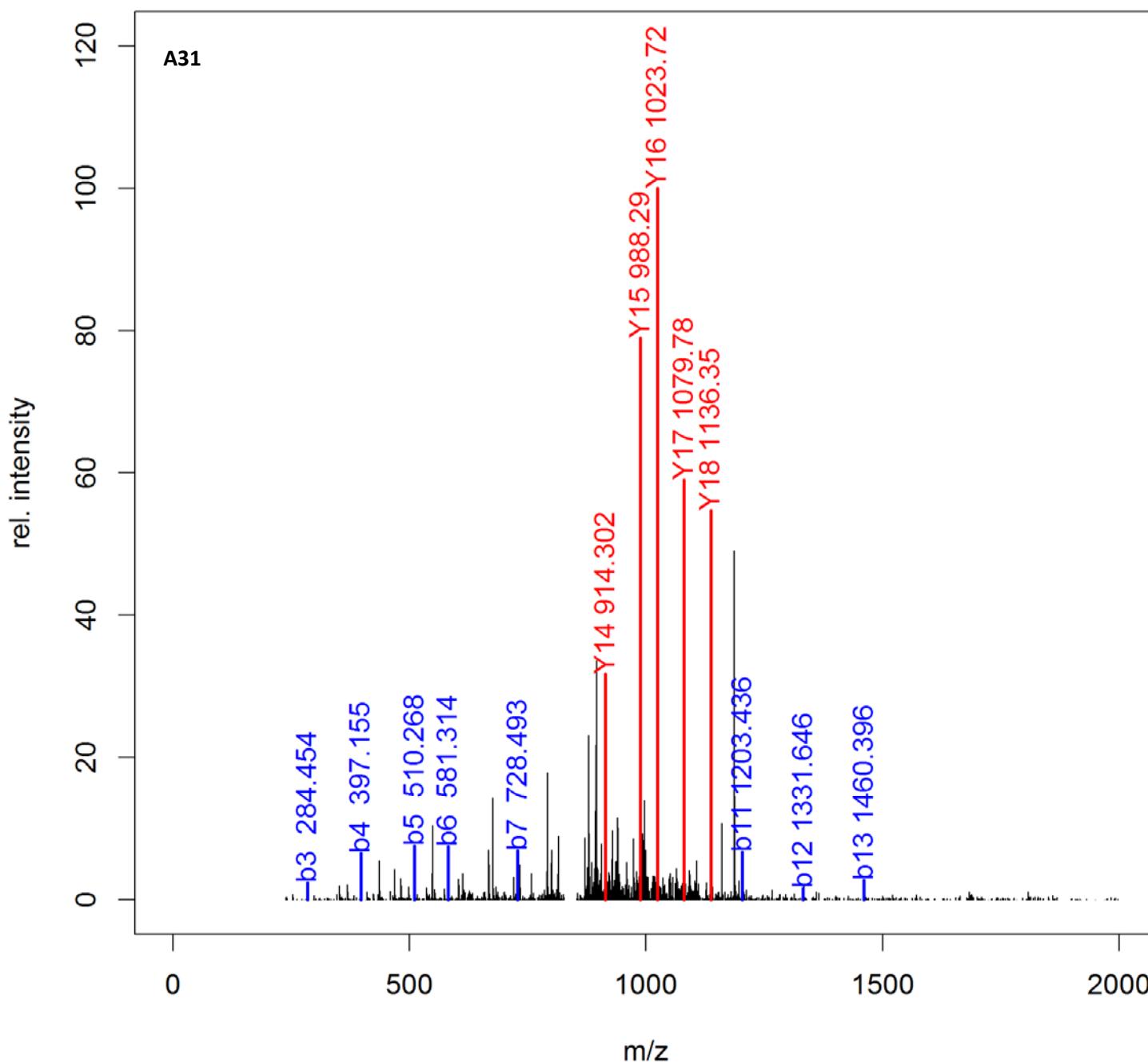


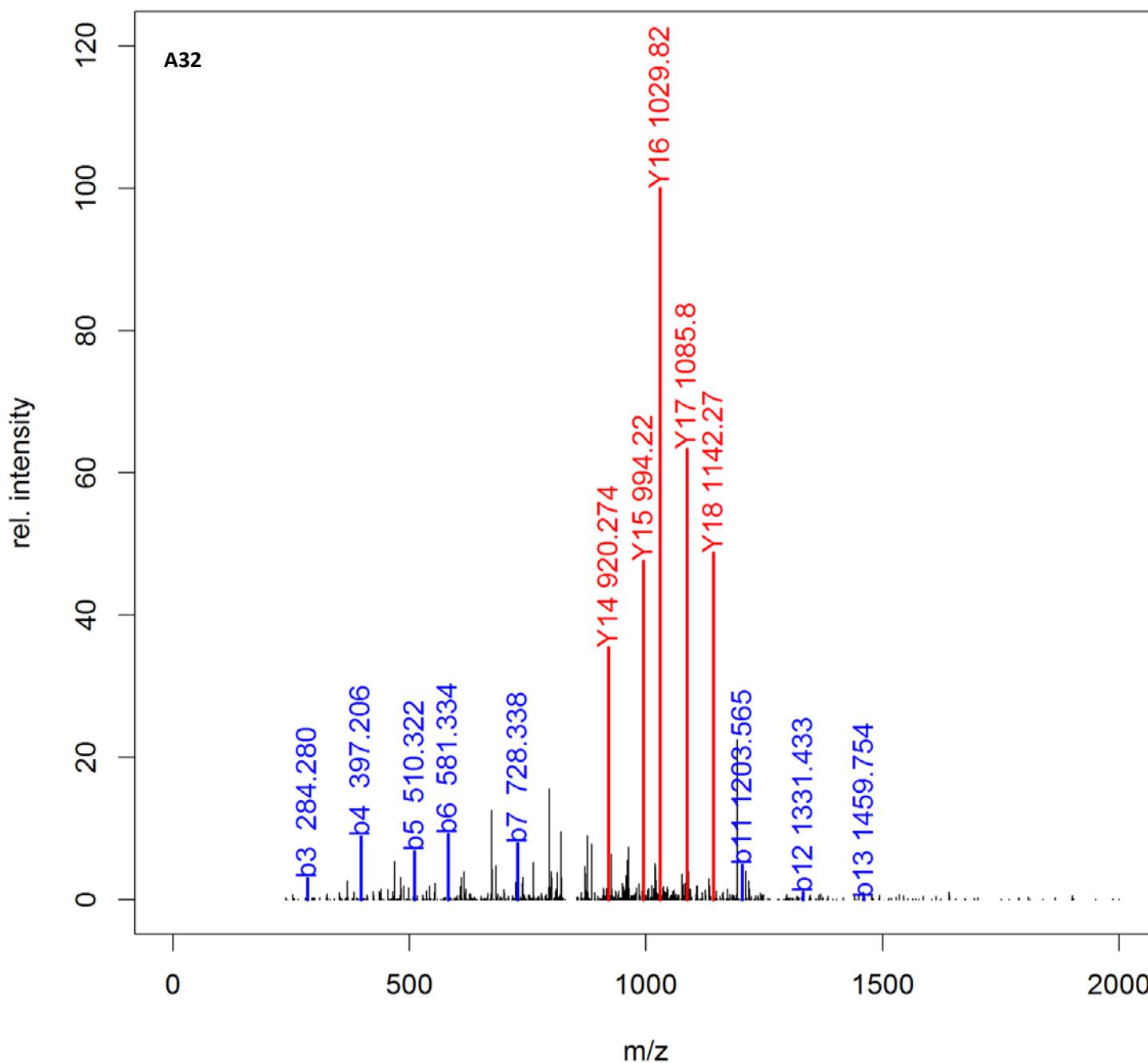


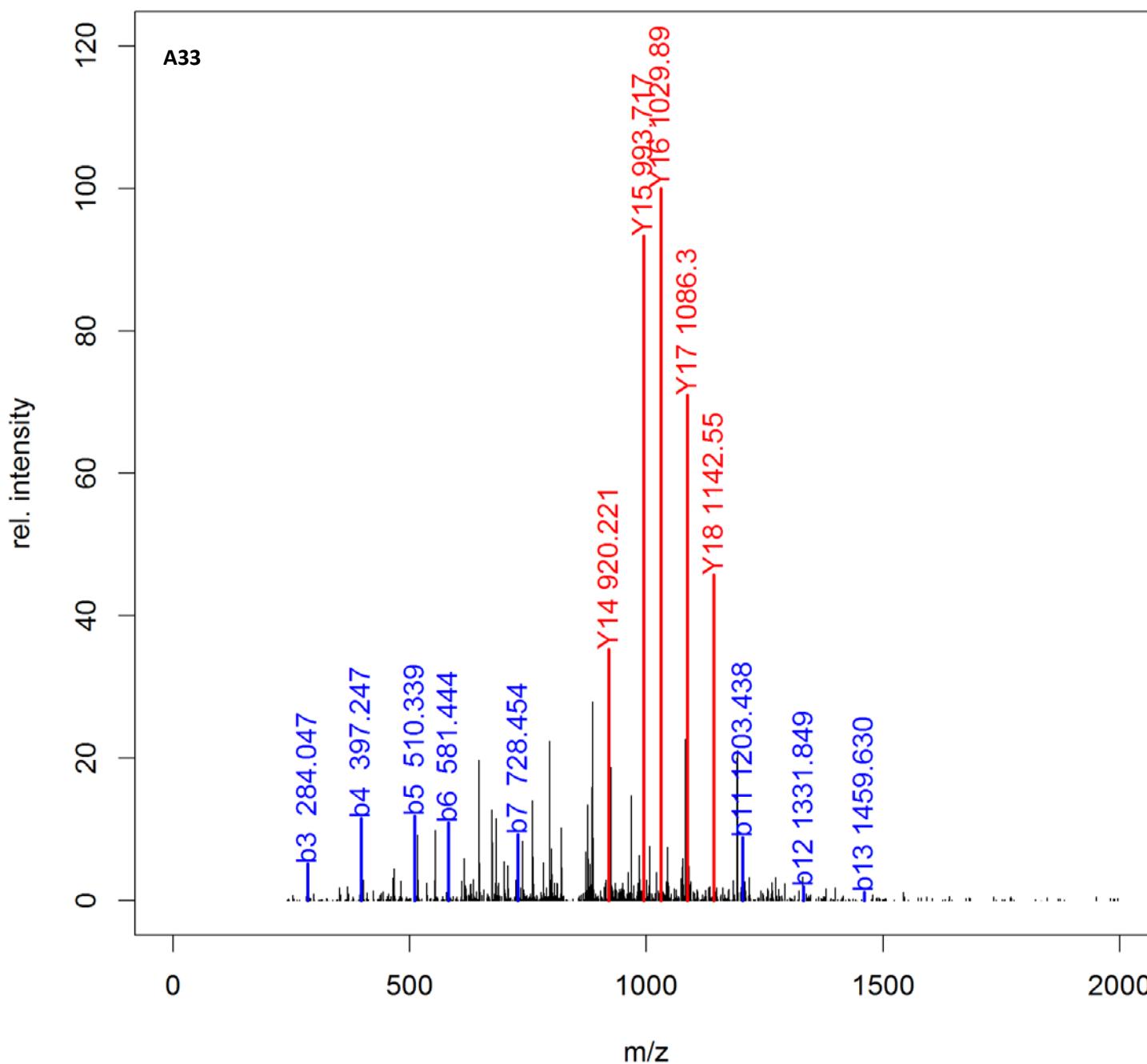


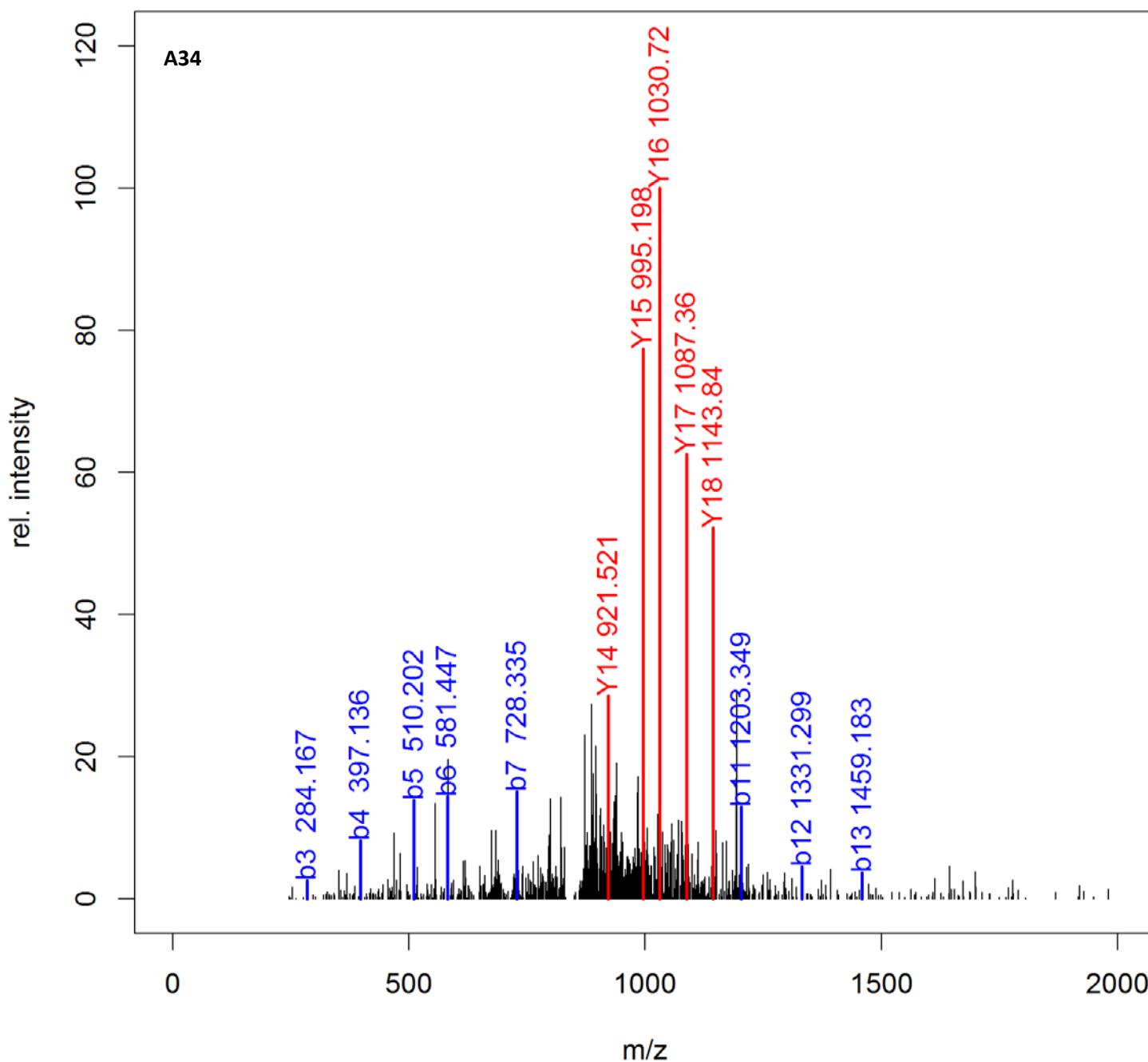


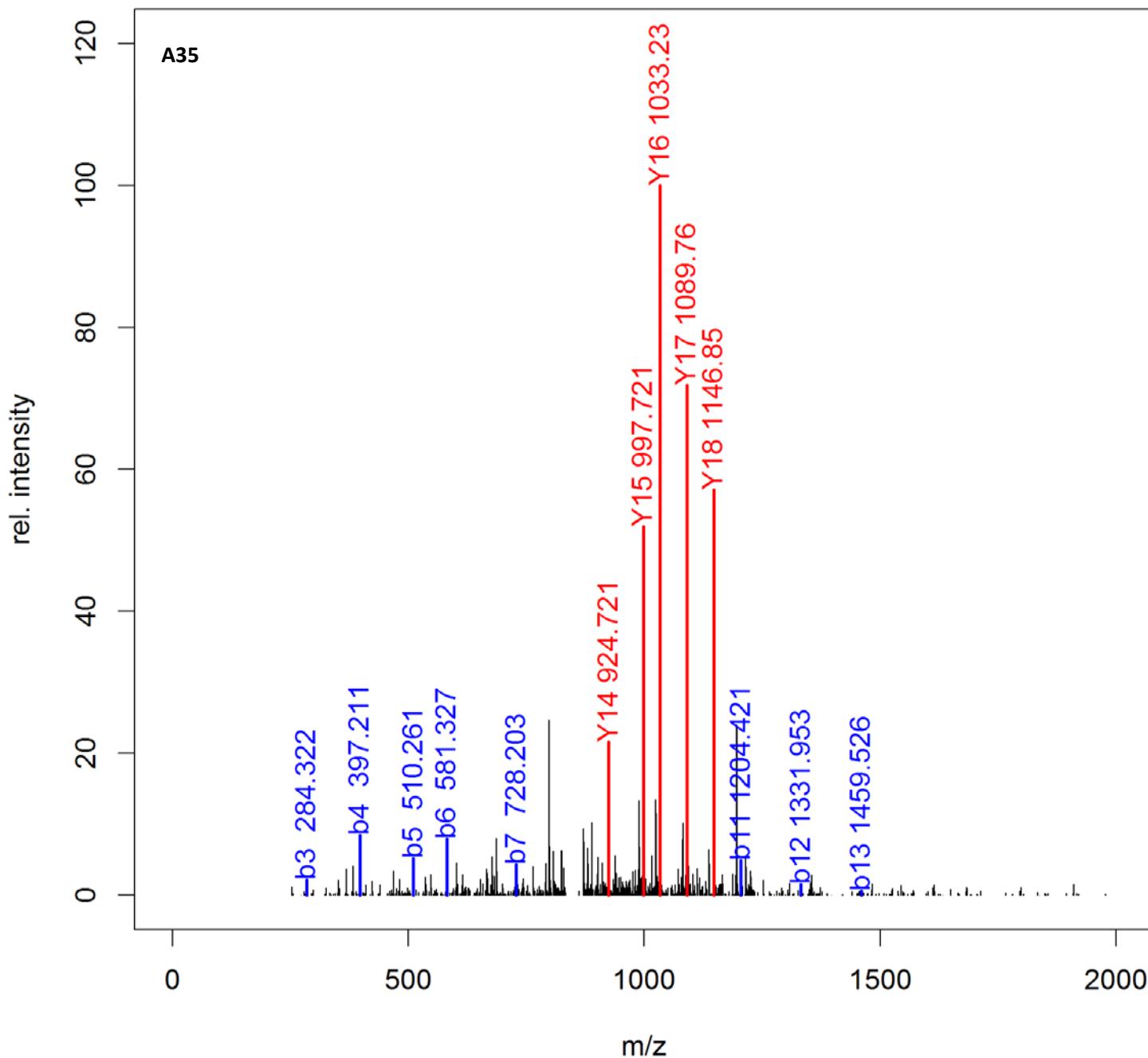


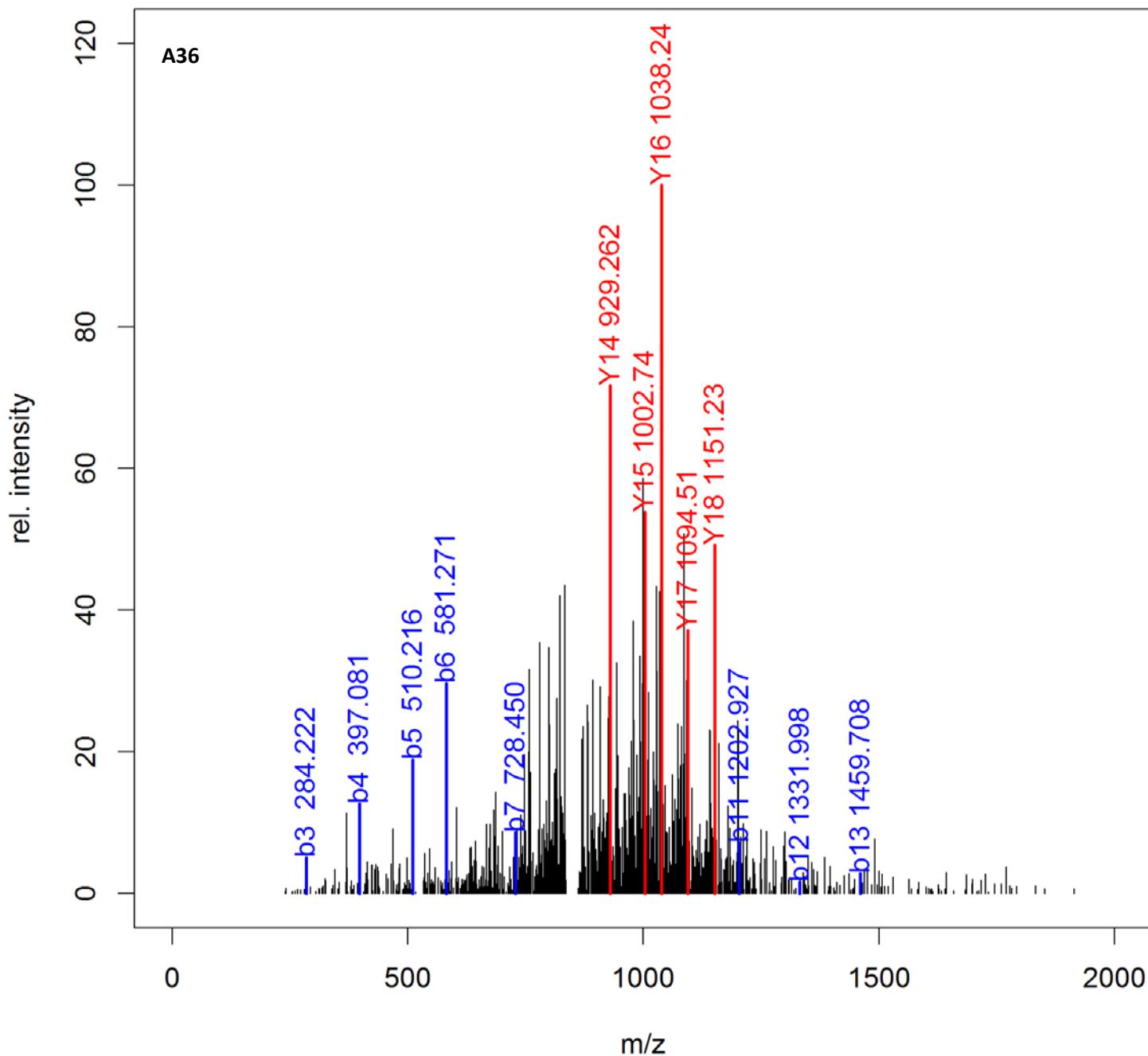


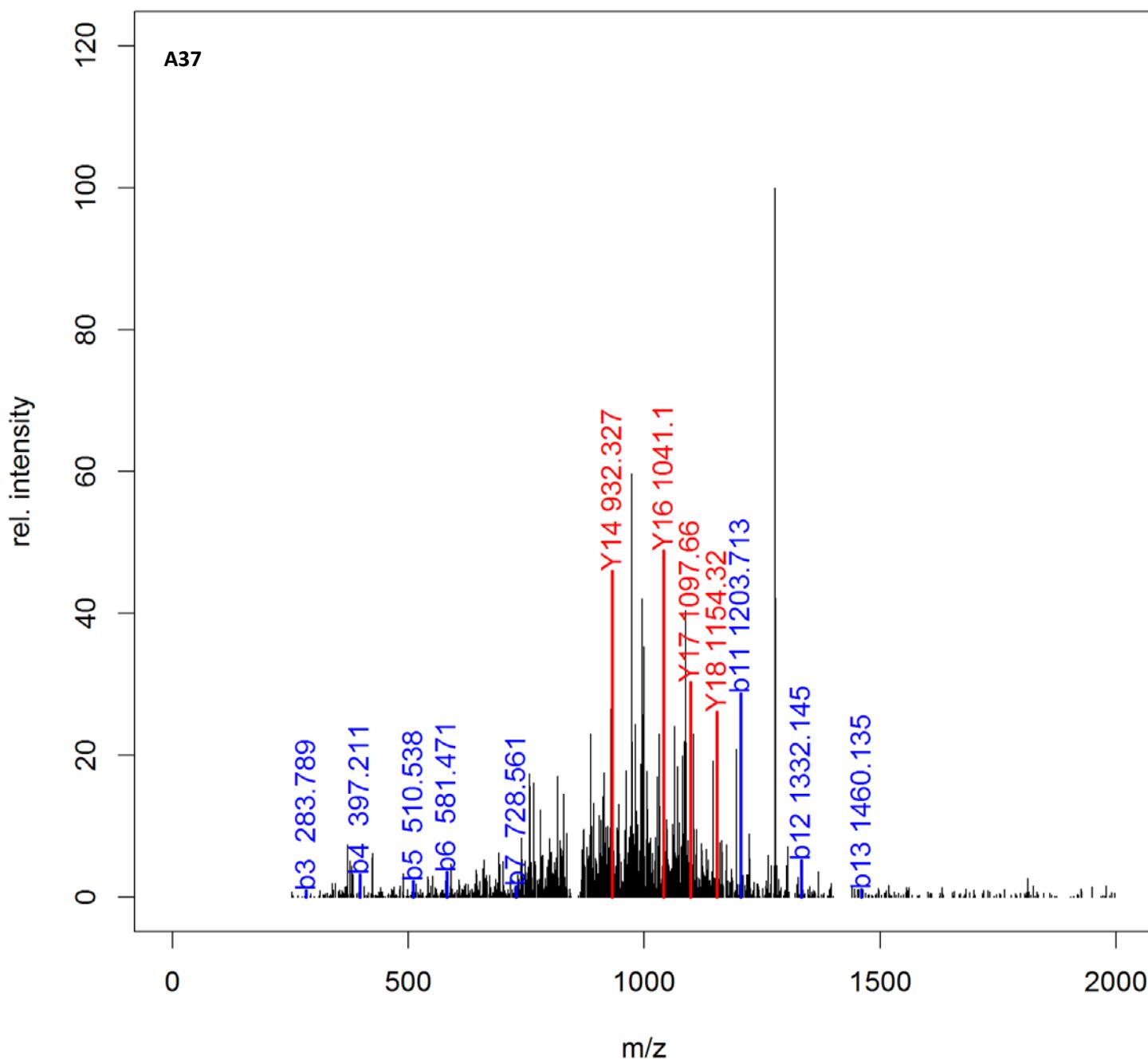


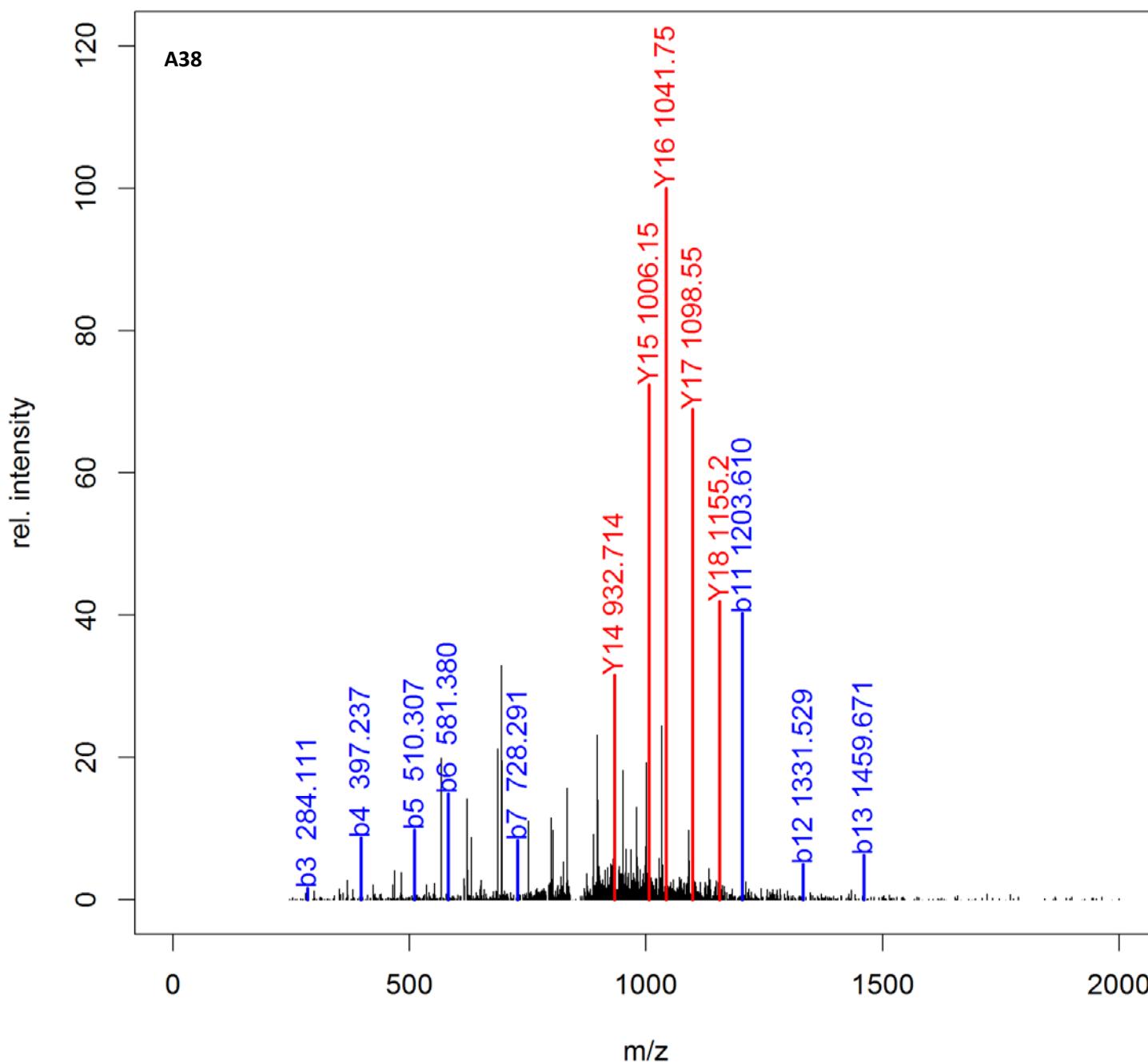


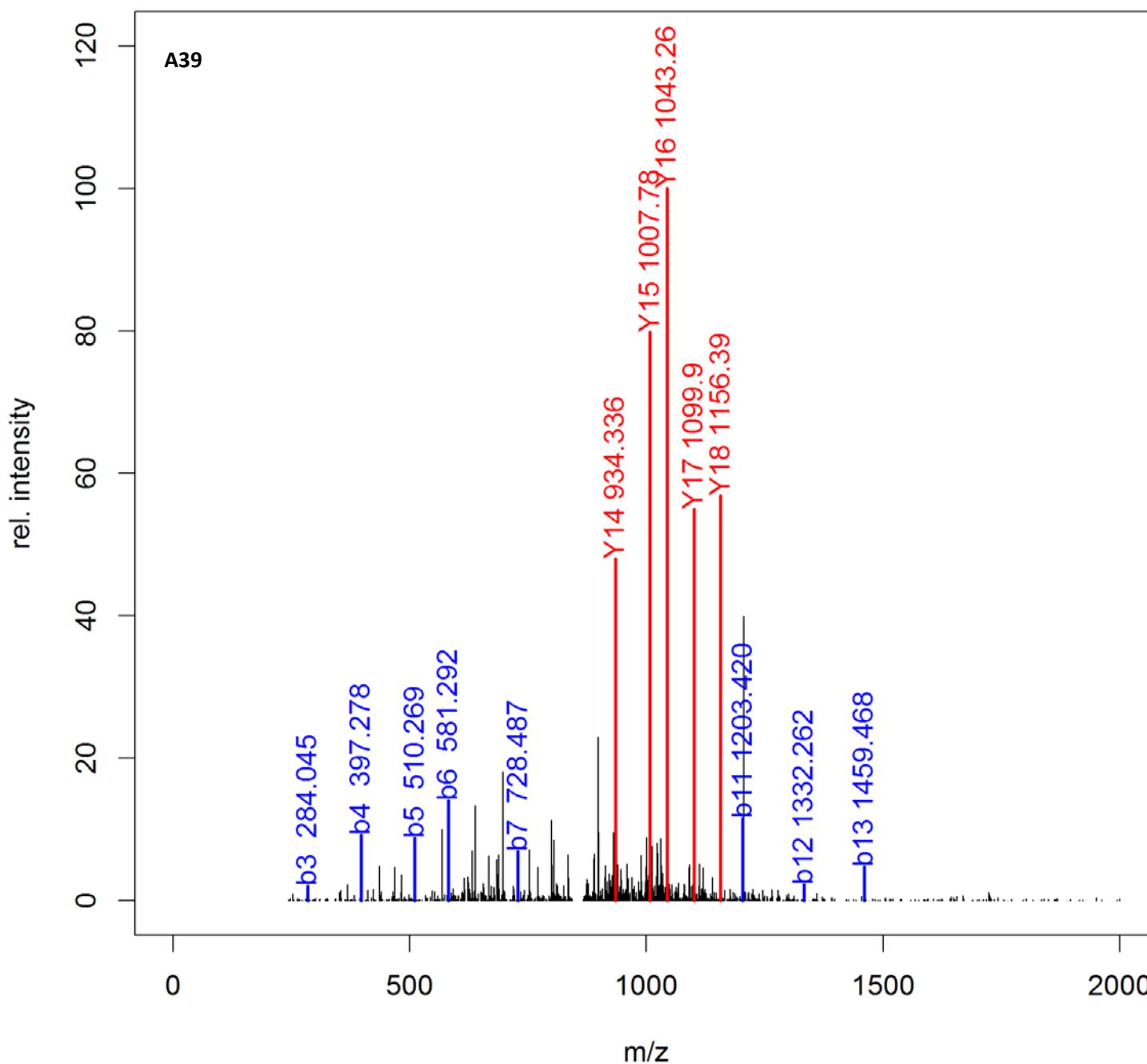


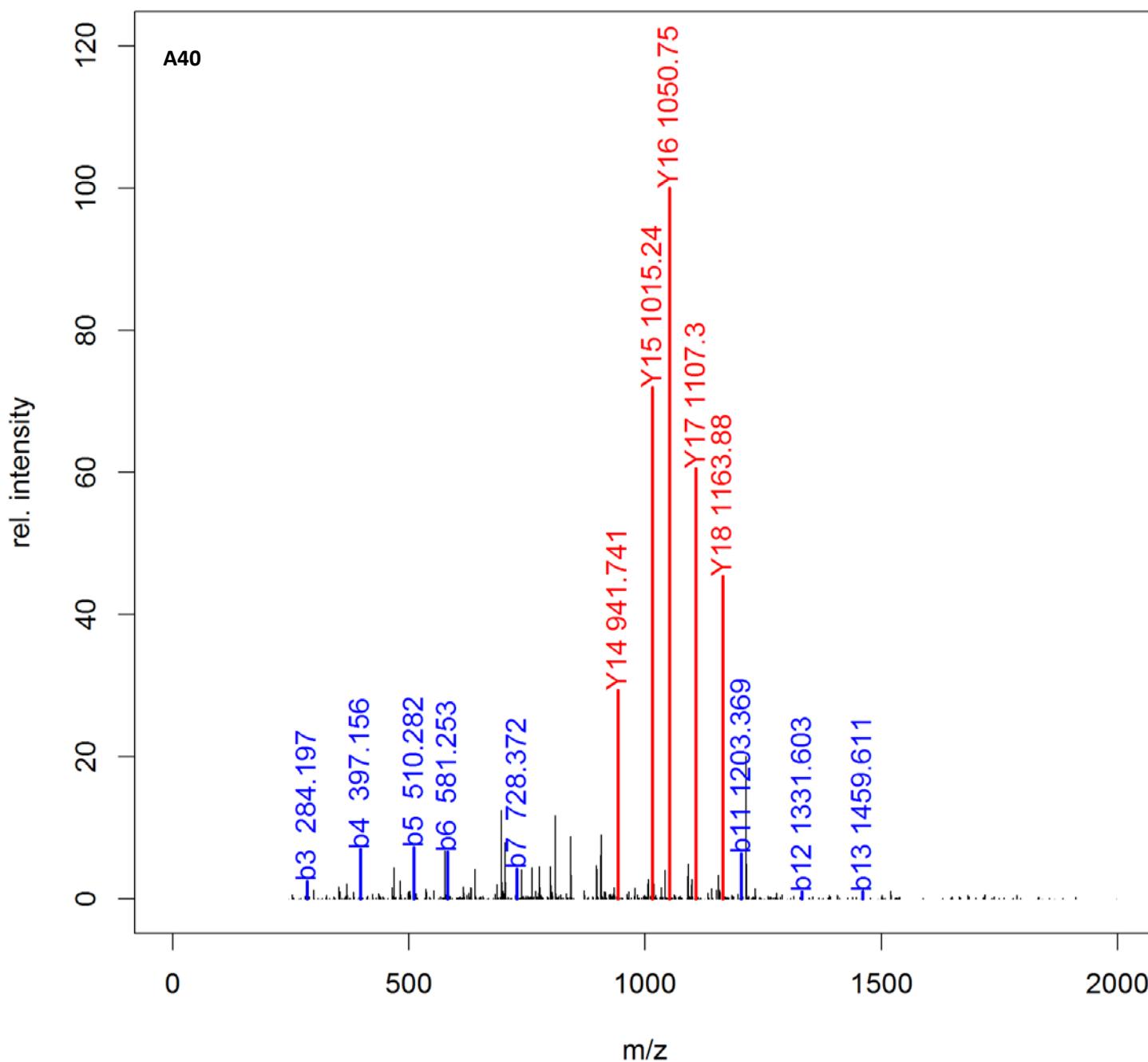


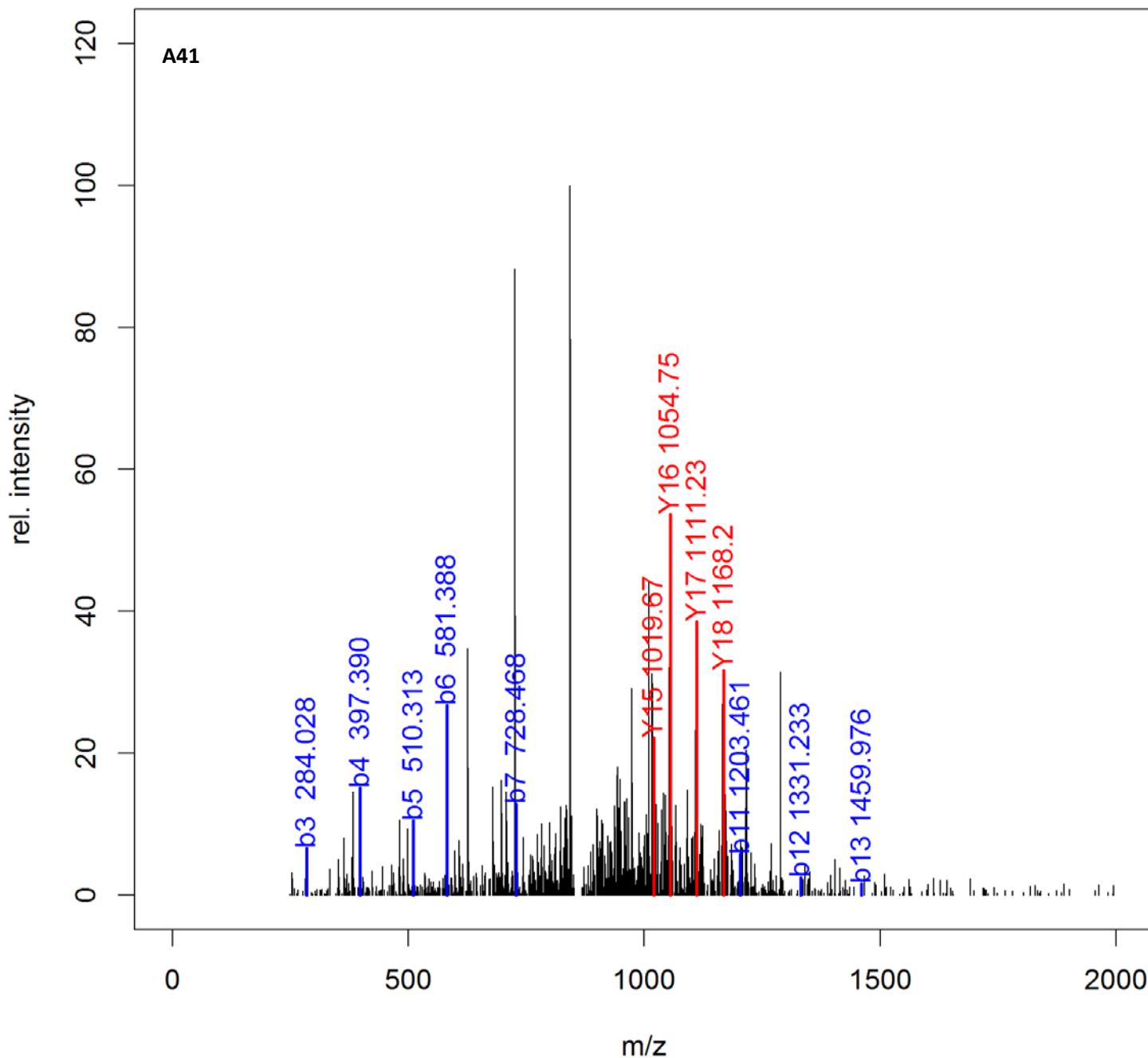


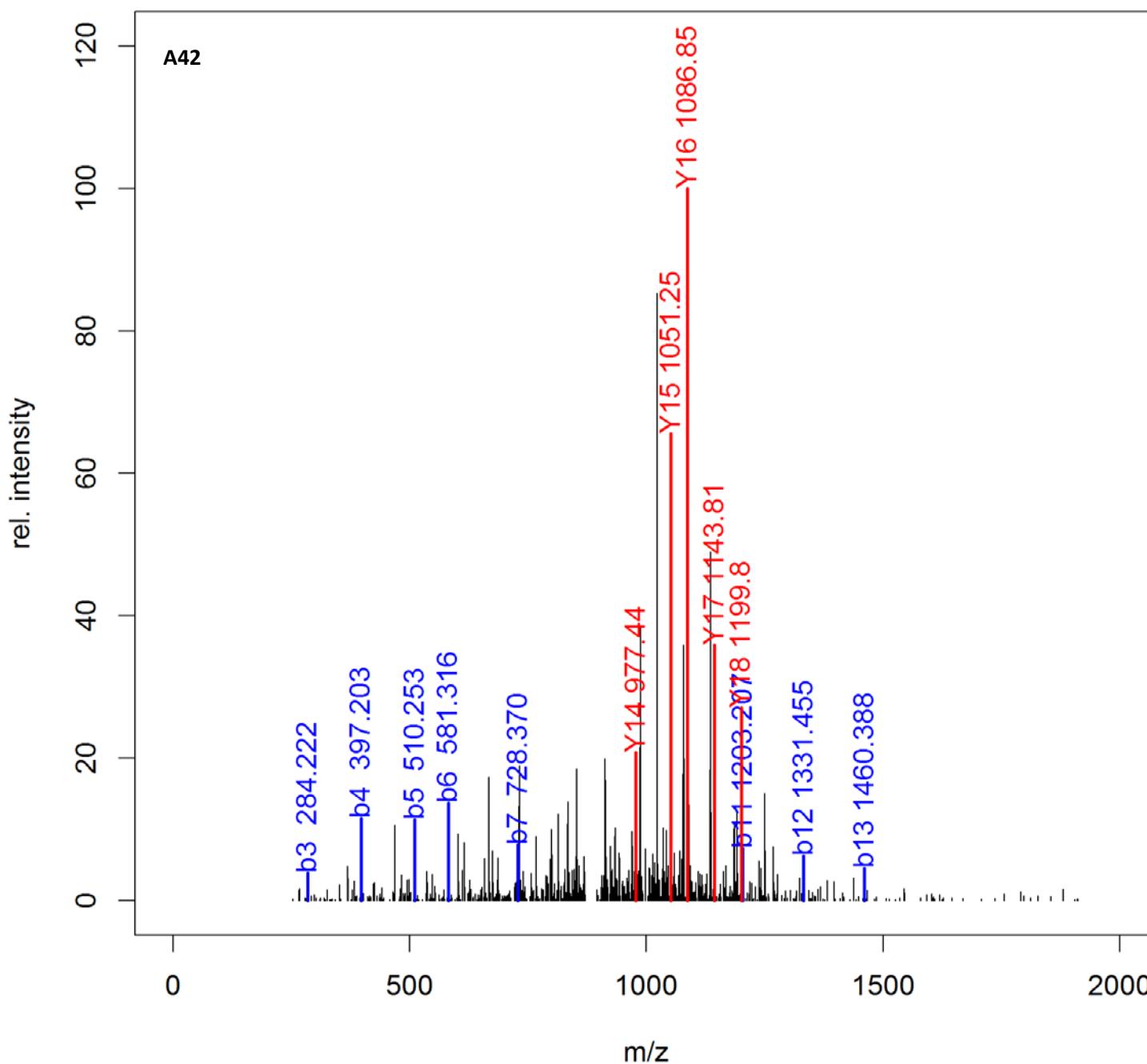












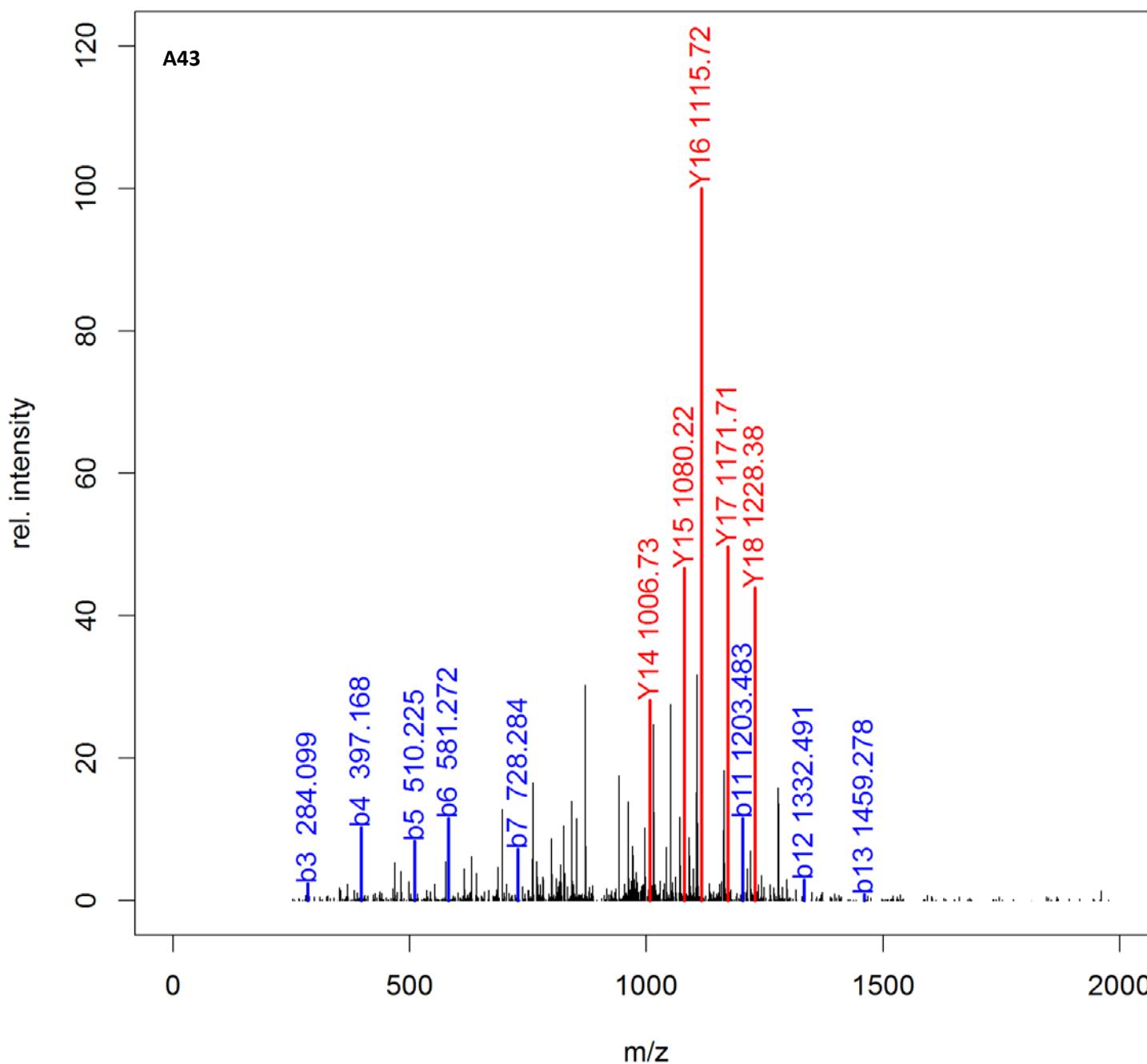


Figure S2. Selected ion chromatograms (SICs) of T3 peptides detected in validation samples. SICs are shown in order of increasing retention time. For clarity, each SIC was selected from a plasma sample with a large S/N. Annotations are given in Table 1 of the text.

