Supporting Information

Allosteric Interaction of Nucleotides and $tRNA^{ala}$ with *E. coli* Alanyl-tRNA Synthetase

John David Dignam[*a], Jingshu Guo[b], Wendell P. Griffith[b], Nichola C. Garbett[c], Amanda Holloway[a] and Timothy Mueser[b]

Analytical gel filtration – Different forms of alanyl-tRNA synthetase, wild type enzyme (ARS75), two deletion constructs (ARS461 and ARS699) and otRNA^{ala} modified forms, were examined by gel filtration on a column of Superose 12 as shown in Figure S1.

Sedimentation velocity (SV) and sedimentation equilibrium (SE) studies - SV was employed to determine the number of sedimenting species and their solution structure. SV profiles analyzed by Sedfit show a major sedimentation peak ~ 7 S representing 93 – 96 % of the optical signal; additional minor sedimentation peaks are evident ~ 4 S (1 – 2 % of the signal) and ~ 9 S (2 – 3 % of the signal) (Figure S2). For all sedimenting species, decreasing *s*-values with increasing sample loading concentration were observed, indicative of molecular asymmetry or other non-ideal behavior [1, 2]. Sedfit allows calculation of a distribution of molecular masses from the sedimentation distribution with the use of an estimated weight-average shape factor, or frictional ratio, f/f_o. While the application of a weight-average f/f_o value is inappropriate for samples with a distribution of sedimenting species, this tool was used to provide an initial qualitative assessment of sample molecular weight based on the observation of one major sedimentation peak. The molecular weight profile showed the opposite trend to *s*-values where the trend was for increased molecular weight with higher sample loading concentration, an observation consistent with a self-associating system. Molecular weights for the major sedimenting species were in the range of 156,000 – 180,000, intermediate between those expected for monomer and dimer species, suggestive of a reaction boundary for a monomer-dimer equilibrium. The estimated f/f_o value from Sedfit analysis of 1.5-1.7 was indicative of significant molecular asymmetry; the slower sedimentation of a less compact solution form

would imply a lower molecular weight species than would be anticipated based solely on the known protein molecular weight. The observation of significant molecular asymmetry would suggest caution in the interpretation of molecular weight distributions determined from SV analysis. In light of decreasing *s*-values with increased sample loading concentration it was necessary to extrapolate the observed sedimentation coefficients to infinite dilution (i.e. zero sample concentration) to yield an $s^0_{20,w}$ value of 7.13 S in the absence of shape effects (Figure S3). Model-independent analysis of g(*s**) distributions using DCDT+ revealed a single broad peak centered ~ 7 S, consistent with the major sedimenting species observed with Sedfit (Figure S4). Extrapolation of the determined $s^*_{20,w}$ values to infinite dilution yielded an $s^0_{20,w}$ value of 7.21 S (Figure S5). Attempts to fit a single species model to these data resulted in reasonable fits with molecular weights of 162,000 – 184,000, which increased with sample loading concentration. Systematic deviations in the fit were evident especially on the high S tail of the g(*s**) distribution; use of multiple species models did not improve the fit.

Analysis of molecular weight is more appropriately performed using SE analysis, particularly given the asymmetric non-ideality observed from SV experiments. Experimental rotor speeds encompassed those appropriate for both monomer and dimer species on the basis of calculated reduced apparent molecular weight (σ) values. All rotor speeds presented conditions appropriate for the observation of dimer with the highest rotor speeds suitable for monomer. Transformation of the raw data into the form of the natural logarithm of the equilibrium sample concentration as a function of radial position as ln(*c*) vs. $r^2/2$, following correction for an experimentally determined y_{offset} value, allowed an assessment of sample homogeneity, non-ideality and association (Figure S6). The concentration distributions were generally well described by a linear fit with molecular weights in the range 143,000 – 195,000. There was evidence of non-

ideal behavior, particularly at higher loading concentrations and rotor speeds. Comparison of the raw data to that expected for single monomer and dimer components revealed that the sample behaved closest to that expected for a dimer, with no evidence of higher order species. There was a trend towards lower molecular weight with increasing rotor speed for a given sample loading concentration, perhaps indicative of the more suitable conditions for the observation of monomer species. However, at each rotor speed sample molecular weight decreased with an increase in sample loading concentration, demonstrating non-ideal behavior.

SV and SE data were subjected to additional model fitting and global analysis using SedAnal and Beckman Coulter Optima XL-A/XL-I Data Analysis software v6.03. Based on the observations of molecular weights slightly less than that of a dimer, the data were fit using a single species dimer model and an associating monomer-dimer model. Global fitting of SV data to a single species model using SedAnal were consistent with analysis using DCDT+ and Sedfit; an acceptable fit (std dev 8 x 10^{-3}) was obtained with a minimized molecular weight (188,000) close to a dimer, but with some systematic deviations in the residuals. A modest improvement (std dev 7 x 10^{-3}) was obtained with a monomer-dimer model with some reduction in the amount of systematic deviations in the residuals, returning an acceptable minimized monomer molecular weight (92,000) and an equilibrium constant indicating predominately dimer species (K_{eq} 9 x 10^5 M^{-1}). Global fitting of SE data gave poor fits to single species and monomer-dimer models; other models, such as monomer-dimer-tetramer and two non-interacting species, gave no improvement in fit. Based on previous observations of significant non-ideality, fitting was restricted to individual data sets and the lowest sample loading concentration where effects of non-ideality will be lowest. On the basis of the standard deviation of the fit, no distinction could be made between the single species and monomer-dimer models; fitting to the monomer-dimer

model indicated predominately dimer species. There is little contribution of a monomer species and the data are best described by a single species dimer model with molecular weights of 178,000 – 189,000 for the three rotor speeds (Figure 1, main manuscript).

Summary of sedimentation studies - Under the conditions examined, the sedimentation velocity profile is dominated by one major sedimenting peak ~ 7 S. The behavior of observed *s*-values as a function of sample loading concentration indicates significant sample non-ideality, believed to be molecular asymmetry. Extrapolation to infinite dilution yields an $s^0_{20,w}$ value of ~ 7.2 S and reveals a frictional ratio of ~ 1.6. Additional minor sedimenting species were observed at ~ 4 S and ~ 9 S representing approximately 5 % of the total sample signal. Sedimentation equilibrium analysis revealed non-ideal behavior with molecular weights in the range of 143,000 – 195,000. The non-ideality of the system made global fitting of the data a challenge and it was not possible to select a definitive model to successfully describe the entirety of the observed sedimentation behavior. The most likely models were a single species (dimer) model and an associating monomer-dimer model. Model fitting was restricted to individual sets of SE data for the lowest sample loading concentration where the least sample non-ideality would be observed, in order to attempt to distinguish between these models. Although these models could not be distinguished on the basis of the standard deviation of the fit, the results revealed little presence of monomer. The data, therefore, appear best described by the presence of predominately dimer species affected by significant non-ideality, with little contribution of monomer species and no compelling evidence of a significant amount of higher order species.

References

1. Cantor, C.R. and Schimmel, P. R., Chapter11-2: Analysis of sedimentation measurements, in *Biophysical Chemistry, Part II: Techniques for the study of biological structure and function*, 1980, San Francisco, CA: W. H. Freeman and Company, pp. 612-614.

2. van Holde, K.E., Johnson, C. W. and Ho, P.S., Chapter 5.2: Sedimentation, in *Physical Biochemistry*. 1998, Upper Saddle River, New Jersey: Prentice Hall, Inc., pp. 194-207.

Figures

Figure S1. Analytical Superose 12 gel filtration on alanyl-tRNA synthetase – Samples (0.5 ml) of ARS875, ARS461, ARS699, otRNA^{ala}-modified ARS875 and otRNA^{ala}-modified ARS461 (circles with X) were applied to Superose 12 and eluted at 0.5 ml per minute. Standards (solid symbols) were thyroglobulin (1), immunoglobulin G (2), human serum albumin (3), ovalbumin (4) and bovine myoglobin (5) and were run separately under the same conditions. The buffer was 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT and 5 mM $MgCl_2$.

Figure S2. Sedimentation coefficient distributions as determined by the program Sedfit. Data for three different loading concentrations of A_{280} ~ 0.2, 0.5 and 1.0 are shown by the black, red and green lines, respectively. Note: *s*-values shown are those determined under the experimental conditions, s_{exp}; these are corrected to $s_{20,w}$ values in the subsequent analysis.

Figure S3. Extrapolation of $s_{20,w}$ values determined by Sedfit to infinite dilution.

Figure S4. Sedimentation coefficient distributions as determined by the program DCDT+. Data for three different loading concentrations of A_{280} ~ 0.2, 0.5 and 1.0 are shown by the black, red and green lines, respectively. Data are corrected to $s^*_{20,w}$ values.

Figure S5. Extrapolation of $s^*_{20,w}$ values determined by DCDT+ to infinite dilution.

Figure S6. Plot of ln(c) versus $r^2/2$ determined from sedimentation equilibrium data. Data for three different loading concentrations of A_{280}~ 0.2, 0.5 and 1.0 are shown by the red, green and blue squares, respectively. Predicted slopes are shown for monomer (dashed line) and dimer (dotted line) species.

Figure S7. Anion exchange chromatography of tRNA^{ala}-peptide – The tRNA^{ala}-peptides derived from the digestion of otRNA^{ala}-modified alanyl-tRNA synthetase were applied to a MonoQ anion exchange column and eluted with a gradient of NaCl as described in Experimental Procedures. **Inset**: urea-acrylamide gel electrophoresis (12% acrylamide gel) of tRNA peptides; lane 1, 0.5 μg of tRNA^{ala}; lane 2, unfractionated tRNA^{ala}-peptides derived from digestion of alanyl-tRNA synthetase with trypsin; lane 3, purified tRNA^{ala}-peptides from the main peak of OD from the anion exchange column; lane 4, 1 μg of tRNA^{ala}.

Figure S8. Mass spectrometry sequencing of tRNA^{ala} peptides – Mass spectrometric sequencing of tRNA^{ala}-peptides from otRNA^{ala}-modified alanyl-tRNA synthetase. Data obtain for the four peptides correspond to residues 71-86 (a), 514-536 (b), 576-592 (c) and 623-639 (d), 638-648 (e) and 640-652 (f). Tables of ions follow the mass spectra. The top spectrum in panel

f is for the peptide with the sugar and the base; the bottom panel shows the spectrum of the peptide without the base. Both are presented because the spectrum with the base did not show as many b or y ion peaks.

Figure S1

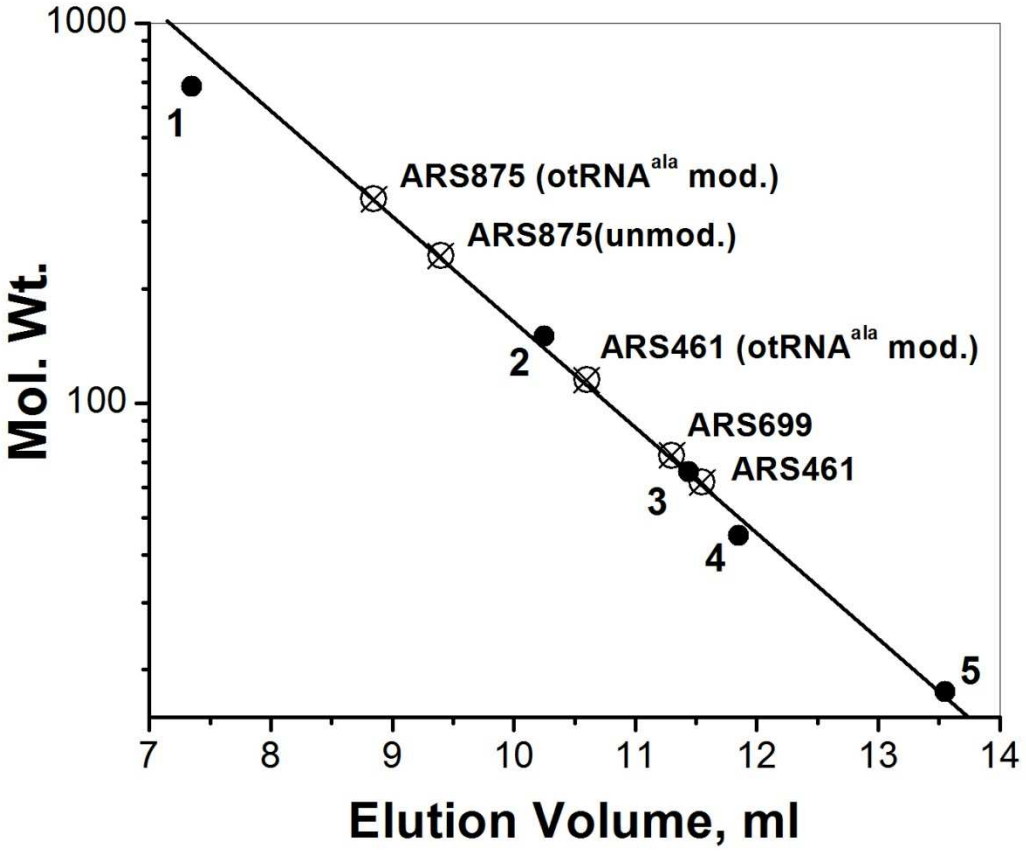

1000
100
Mol. Wt.
1
2
3
4
5
ARS875 (otRNA^{ala} mod.)
ARS875(unmod.)
ARS461 (otRNA^{ala} mod.)
ARS699
ARS461
7
8
9
10
11
12
13
14
Elution Volume, ml

Figure S2

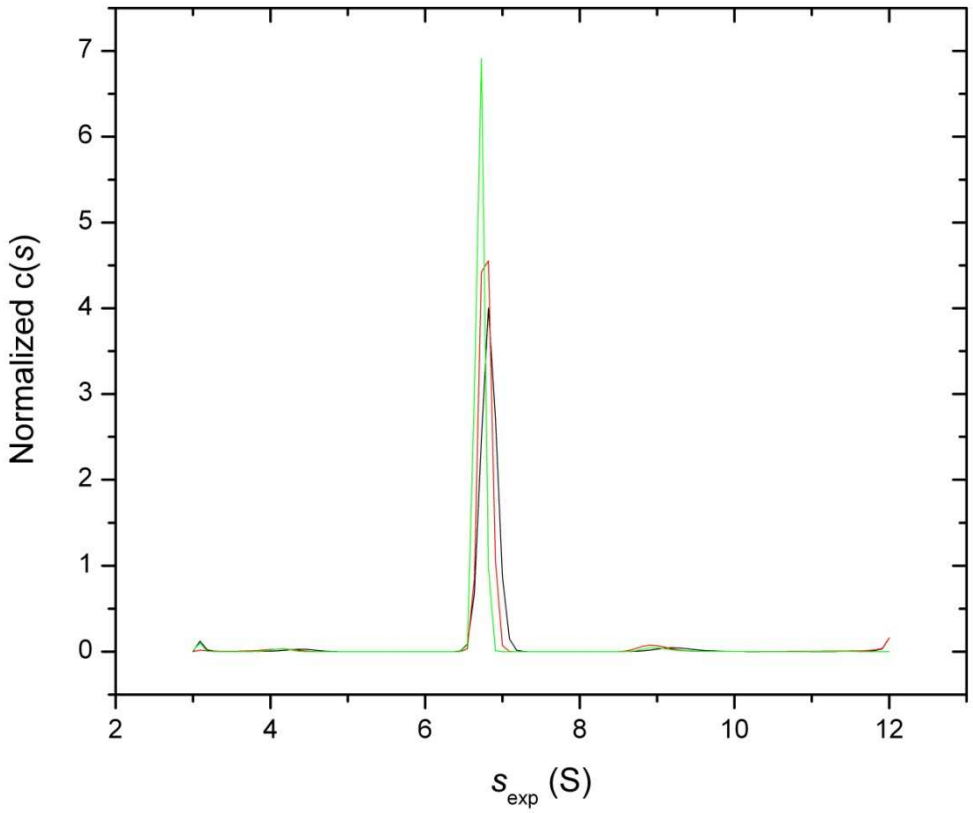


Figure S3

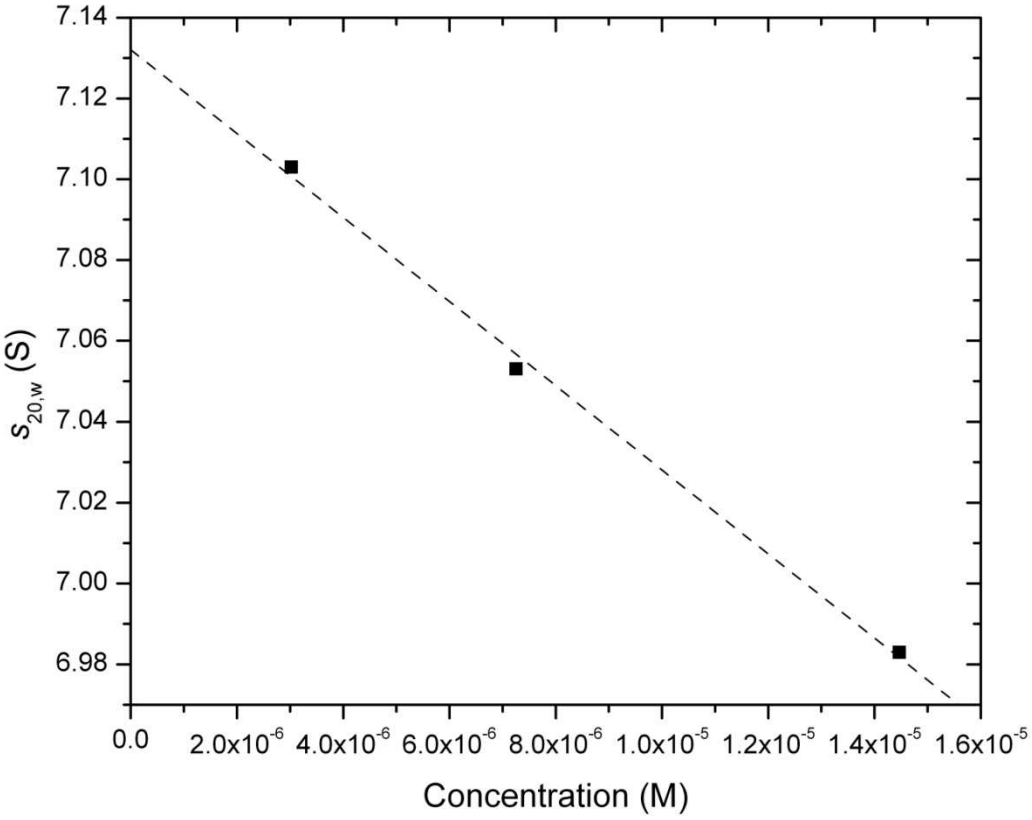

7.14
7.12
7.10
7.08
7.06
7.04
7.02
7.00
6.98
$s_{20,w}$ (S)
0.0
2.0x10^{-6}
4.0x10^{-6}
6.0x10^{-6}
8.0x10^{-6}
1.0x10^{-5}
1.2x10^{-5}
1.4x10^{-5}
1.6x10^{-5}
Concentration (M)

Figure S4

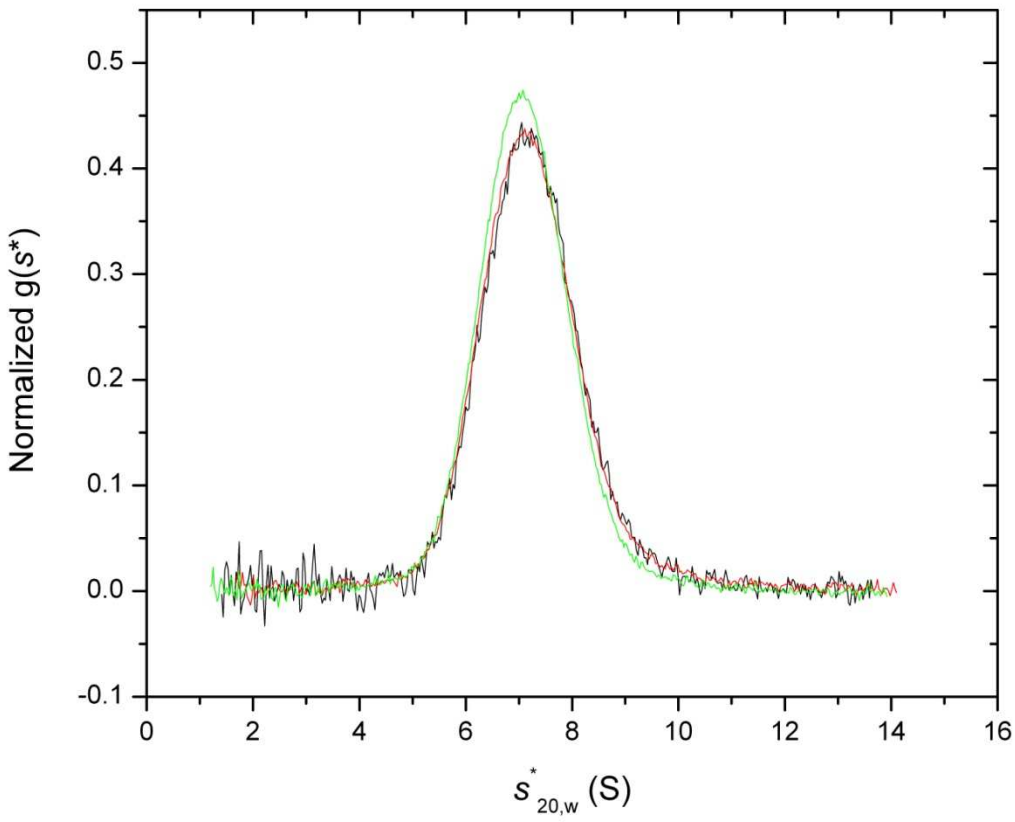


Figure S5

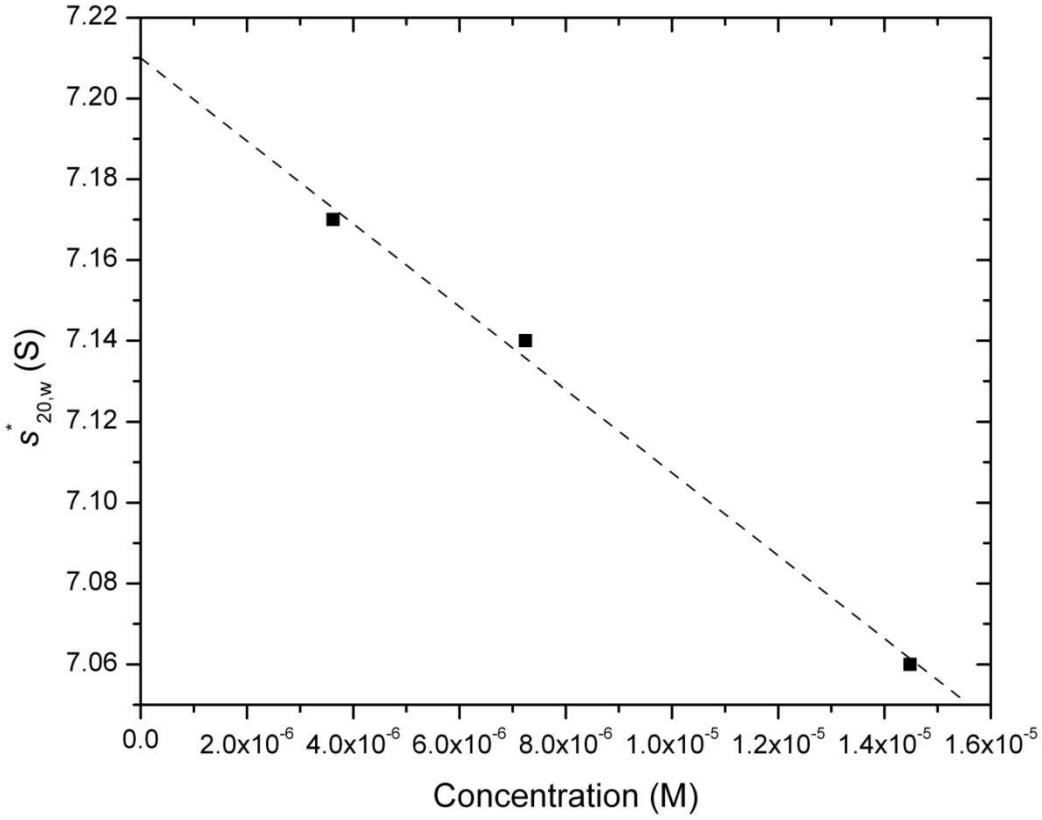

7.22
7.20
7.18
7.16
7.14
7.12
7.10
7.08
7.06
$s^*_{20,w}$ (S)
0.0
2.0x10^{-6}
4.0x10^{-6}
6.0x10^{-6}
8.0x10^{-6}
1.0x10^{-5}
1.2x10^{-5}
1.4x10^{-5}
1.6x10^{-5}
Concentration (M)

Figure S6

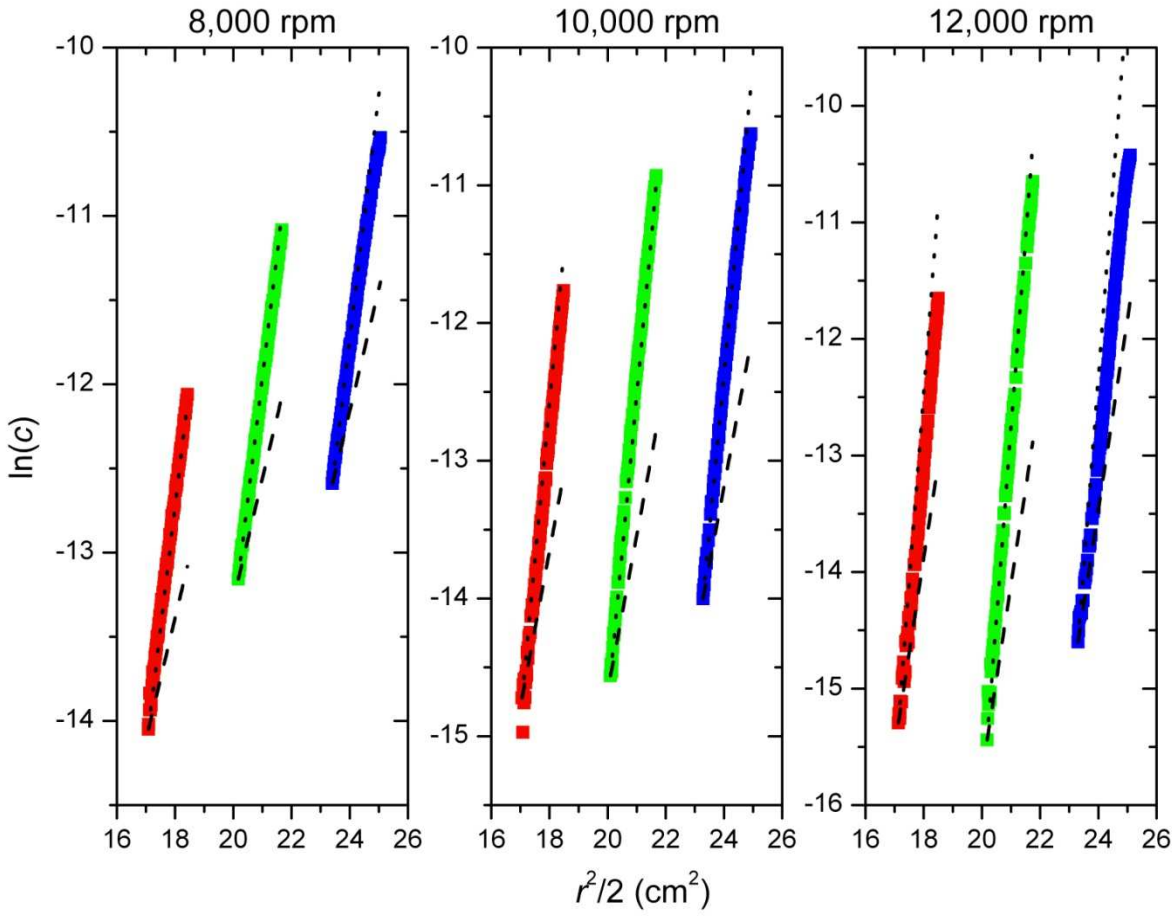


Figure S7

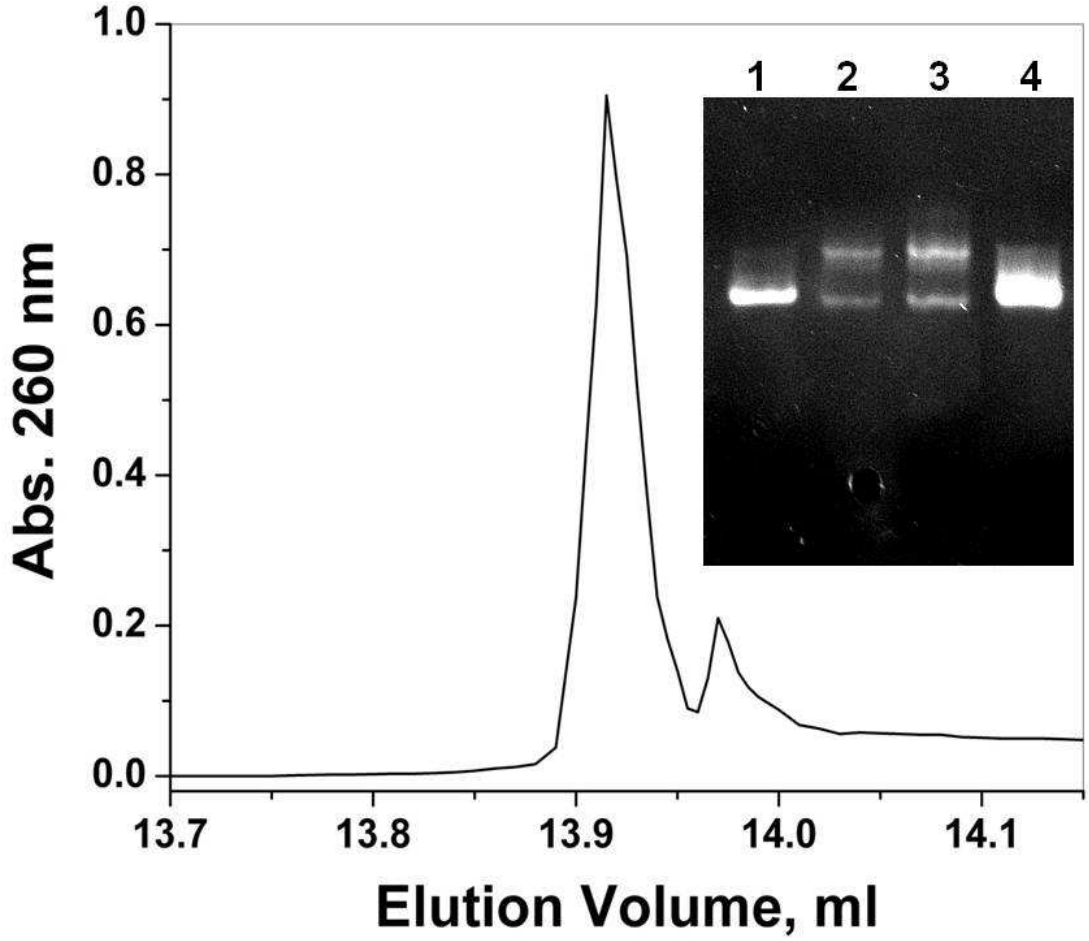
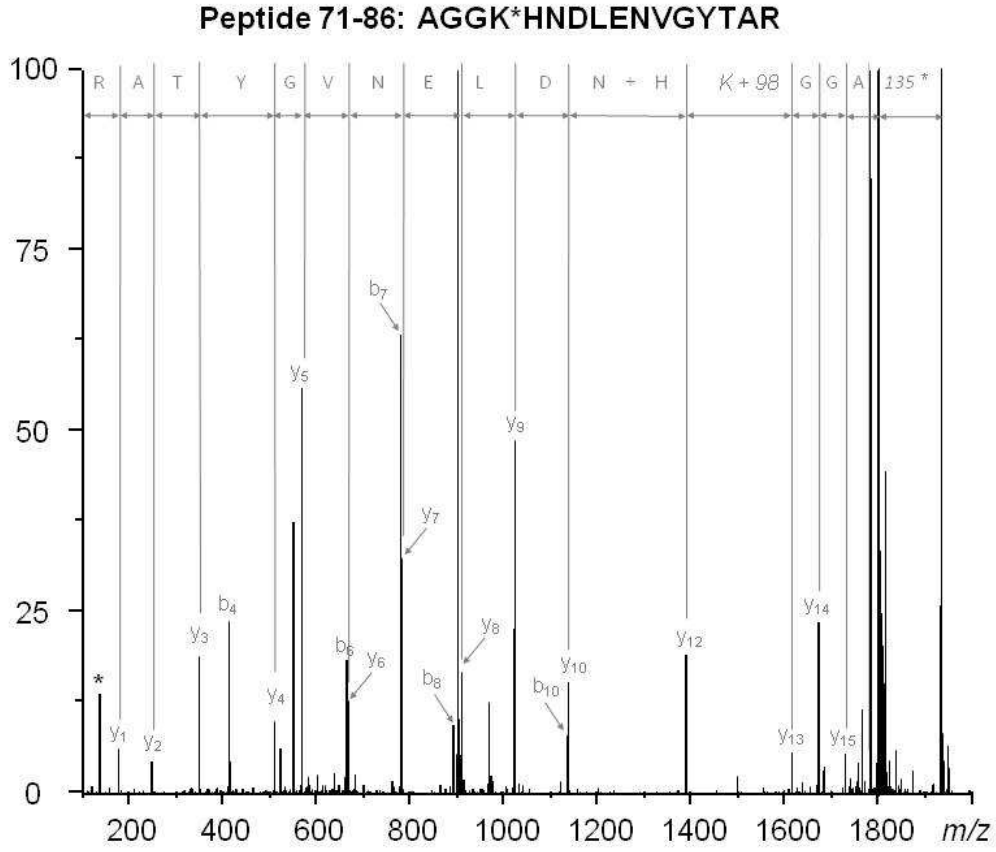
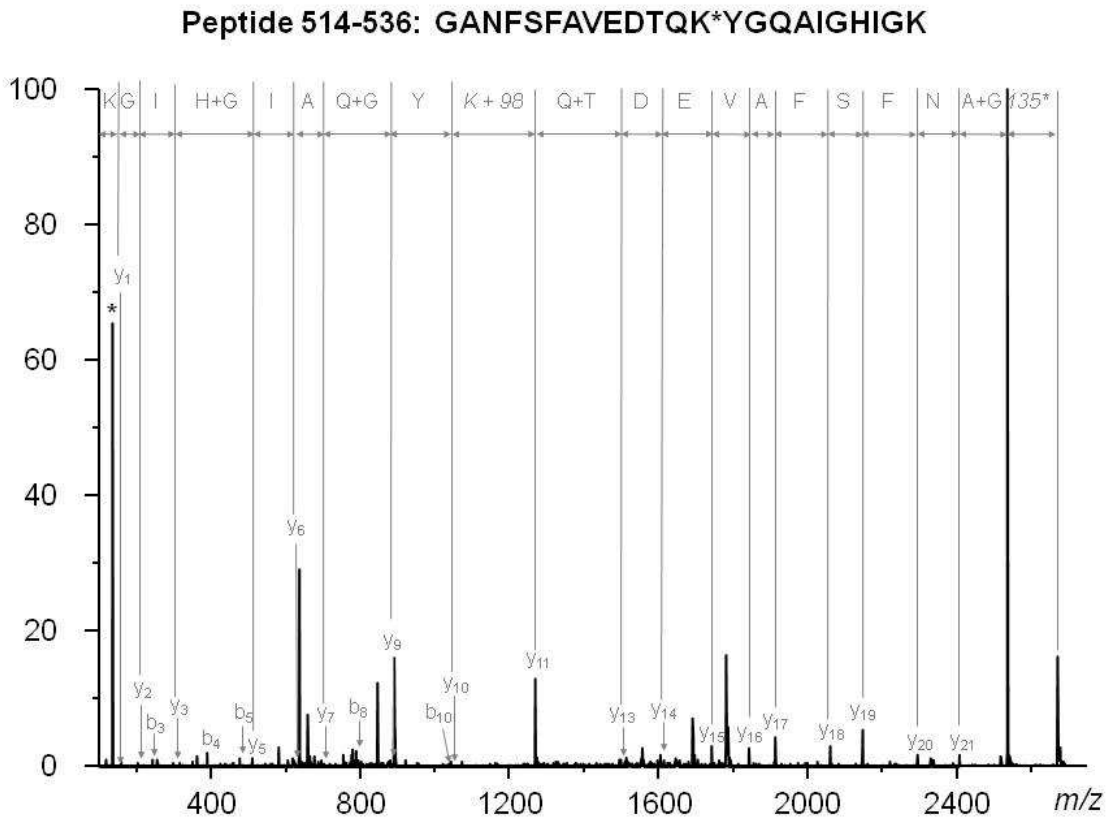

1.0
0.8
0.6
0.4
0.2
0.0
Abs. 260 nm
13.7
13.8
13.9
14.0
14.1
Elution Volume, ml
1
2
3
4

Figure S8a


Peptide 71-86: AGGK*HNDLENVGYTAR
R A T Y G V N E L D N + H K + 98 G G A 135 *
100
75
50
25
0
200 400 600 800 1000 1200 1400 1600 1800 m/z
y1 y2 y3 b4 y4 y5 b6 y6 b7 y7 b8 y8 y9 b10 y10 y12 y13 y14 y15

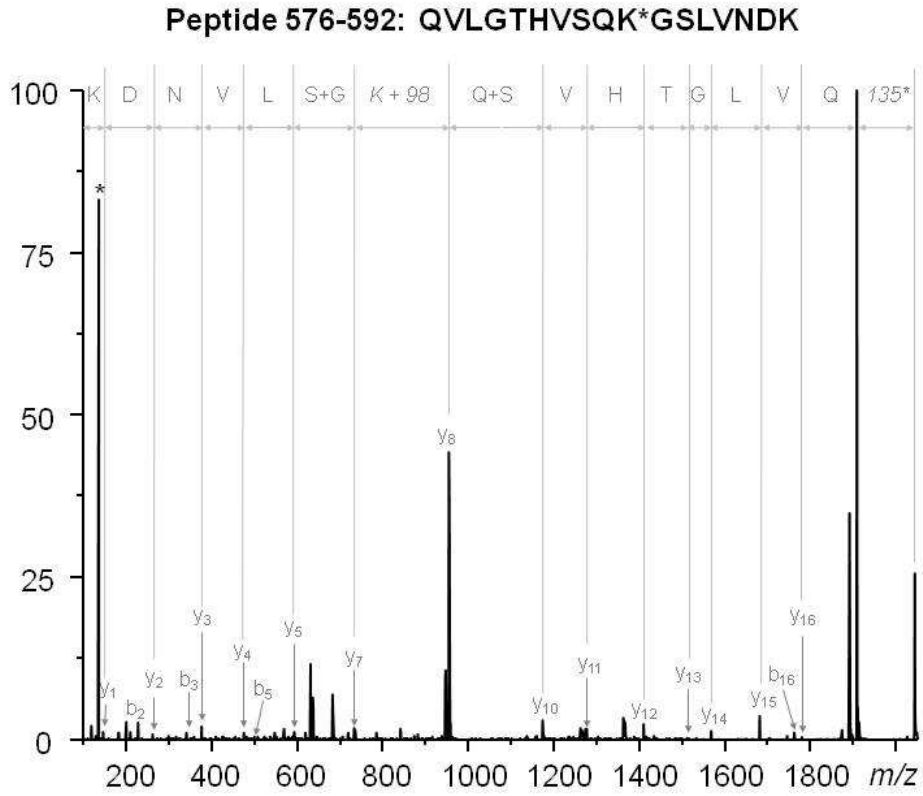
Peptide 71-86: AGGK*HNDLENVGYTAR						
b-ion m/z calculated	b-ion m/z measured	Residue position	Sequence	Residue position	y-ions m/z measured	y-ions m/z calculated
---	---	1	A	16	---	---
129.07	---	2	G	15	1728.89	1728.83
186.08	---	3	G	14	1671.87	1671.81
412.22	412.24	4	K*	13	1614.85	1614.79
549.28	549.31	5	H	12	1388.71	1388.66
663.32	663.35	6	N	11	---	1251.60
778.35	778.39	7	D	10	1137.61	1137.55
891.43	891.54	8	L	9	1022.58	1022.53
1020.47	1020.53	9	E	8	909.49	909.44
1134.51	1134.56	10	N	7	780.44	780.40
1233.59	---	11	V	6	666.39	666.36
1290.61	---	12	G	5	567.32	567.29
1453.67	---	13	Y	4	510.29	510.27
1554.72	---	14	T	3	347.22	347.20
1625.76	---	15	A	2	246.17	246.16
---	---	16	R	1	175.13	175.12
---	---	1	A	16	---	---
All fragment m/z values were calculated for K*, which is the modified lysyl residue after loss of the labile base. See Figure 7 (main text) for the structure of the modified lysine.						

Figure S8b


Peptide 514-536: GANFSFAVEDTQK*YGQAIGHIGK
K G I H+G I A Q+G Y K+98 Q+T D E V A F S F N A+G 135*
100
80
60
40
20
0
400
800
1200
1600
2000
2400
m/z

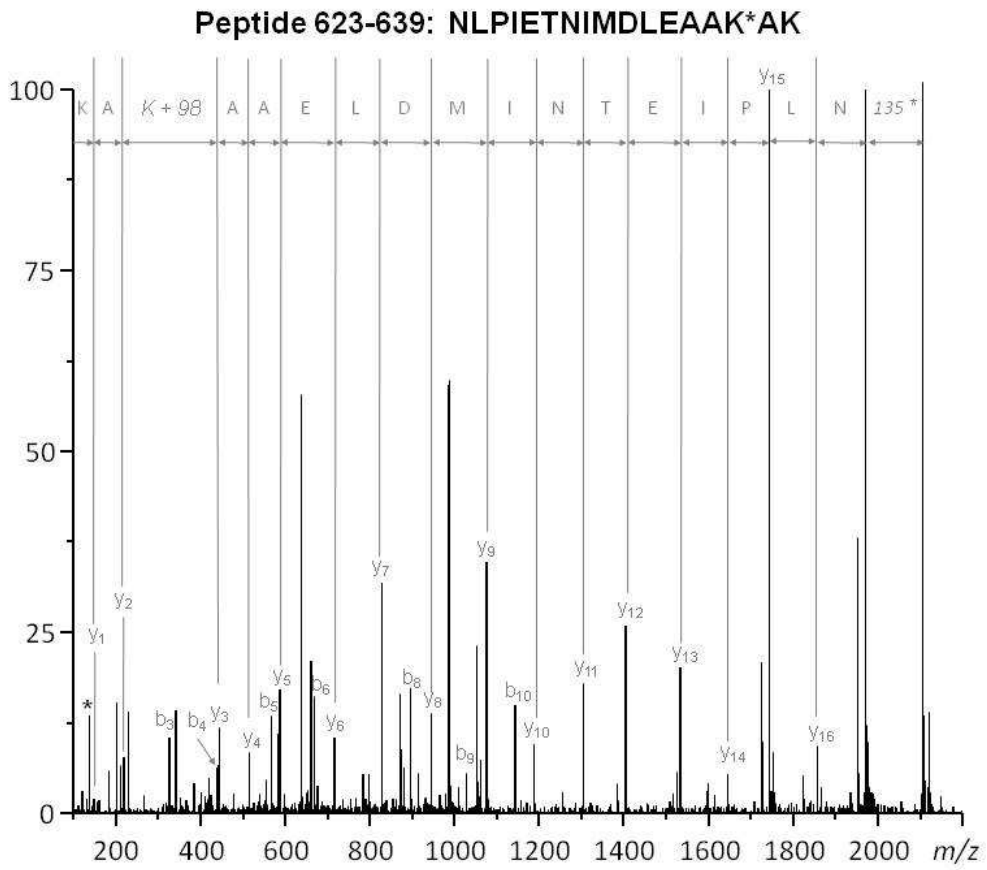
Peptide 514-536: GANFSFAVEDTQK*YGQAIGHIGK						
b-ion m/z calculated	b-ion m/z measured	Residue position	Sequence	Residue position	y-ions m/z measured	y-ions m/z calculated
---	---	1	G	23	---	---
129.07	129.07	2	A	22	---	2479.23
243.11	243.11	3	N	21	2408.11	2408.19
390.18	390.17	4	F	20	2294.08	2294.15
477.21	477.19	5	S	19	2147.01	2147.08
624.28		6	F	18	2060.01	2060.04
695.31	695.32	7	A	17	1912.94	1912.98
794.38	794.35	8	V	16	1841.89	1841.94
923.43	---	9	E	15	1742.83	1742.87
1038.45	1038.39	10	D	14	1613.83	1613.83
1139.50	1139.59	11	T	13	1498.76	1498.80
1267.56	1267.74	12	Q	12	---	1397.75
1493.69	1493.66	13	K*	11	1269.22	1269.70
1656.75	1656.64	14	Y	10	1043.55	1043.56
1713.78	---	15	G	9	880.50	880.50
1841.83	1841.89	16	Q	8	---	823.48
1912.87	1912.94	17	A	7	695.32	695.42
2025.96	2025.88	18	I	6	624.38	624.38
2082.98	2082.80	19	G	5	511.27	511.30
2220.04	2220.01	20	H	4	---	454.28
2333.12	2332.99	21	I	3	317.20	317.22
2390.14	---	22	G	2	204.13	204.13
---	---	23	K	1	147.11	147.11
Fragment m/z values were calculated for K*, which is the modified lysyl residue after loss of the labile base. See Figure 7 (main text) for the structure of the modified lysine.						

Figure S8c


Peptide 576-592: QVLGTHVSQK*GSLVNDK
K D N V L S+G K + 98 Q+S V H T G L V Q 135*
100
75
50
25
0
200 400 600 800 1000 1200 1400 1600 1800 m/z
y1 b2 y2 b3 y3 y4 b5 y5 y7 y8 y10 y11 y12 y13 y14 y15 b16 y16

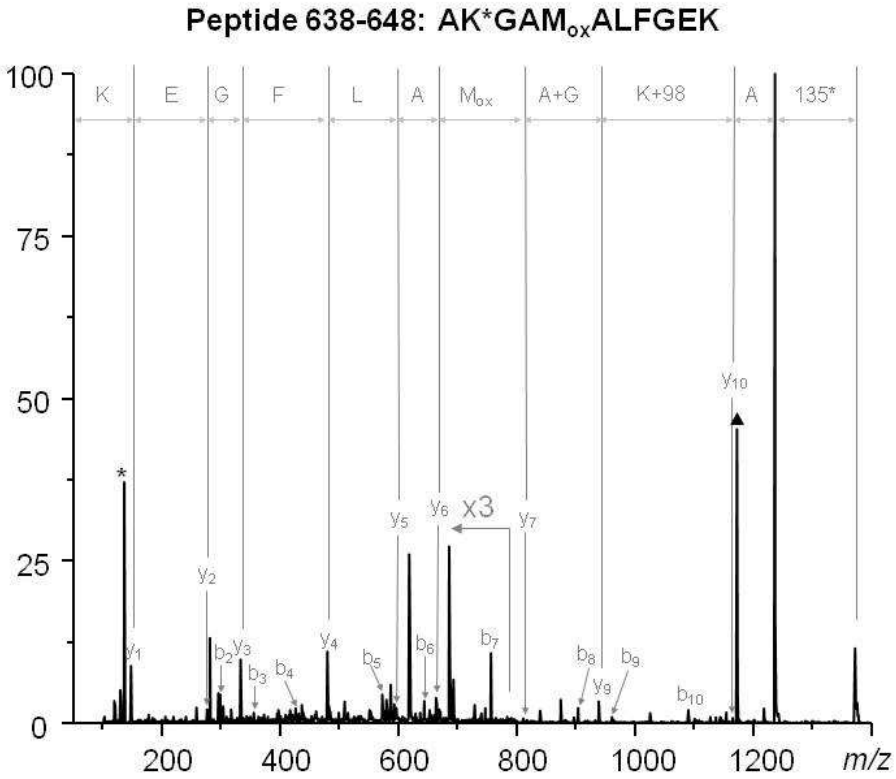
Peptide 576-592: QVLGTHVSQK*GSLVNDK						
b-ion m/z calculated	b-ion m/z measured	Residue position	Sequence	Residue position	y-ions m/z measured	y-ions m/z calculated
---	---	1	Q	17	---	---
228.13	228.13	2	V	16	1779.85	1779.96
341.22	341.21	3	L	15	1680.82	1680.89
398.24	---	4	G	14	1567.73	1567.81
499.29	499.28	5	T	13	1510.74	1510.79
636.35	636.61	6	H	12	1409.69	1409.74
735.41	735.40	7	V	11	1272.62	1272.68
822.45	822.40	8	S	10	1173.55	1173.61
950.51	---	9	Q	9	---	1086.58
1176.64	1176.57	10	K*	8	958.50	958.52
1233.66	1233.57	11	G	7	732.36	732.39
1320.69	---	12	S	6	---	675.37
1433.77	1433.71	13	L	5	588.32	588.34
1532.84	1532.79	14	V	4	475.24	475.25
1646.89	1646.85	15	N	3	376.17	376.18
1761.91	1761.82	16	D	2	262.13	262.14
---	---	17	K	1	147.11	147.11
Fragment m/z values were calculated for K*, which is the modified lysyl residue after loss of the labile base. See Figure 7 (main text) for the structure of the modified lysine.						

Figure S8d


Peptide 623-639: NLPIETNIMDLEAAK*AK
100
75
50
25
0
K A K+98 A A E L D M I N T E I P L N 135*
y15
y1
y2
y3
y4
y5
y6
y7
y8
y9
y10
y11
y12
y13
y14
y16
b3
b4
b5
b6
b8
b9
b10
*
200
400
600
800
1000
1200
1400
1600
1800
2000
m/z

Peptide 623-639: NLPIETNIMDLEAAK*AK						
b-ion m/z calculated	b-ion m/z measured	Residue position	Sequence	Residue position	y-ions m/z measured	y-ions m/z calculated
---	---	1	N	17	---	---
228.13	228.11	2	L	16	1854.80	1854.99
325.19	325.15	3	P	15	1741.72	1741.90
438.27	438.23	4	I	14	1644.74	1644.85
567.31	567.26	5	E	13	1531.65	1531.77
668.36	668.30	6	T	12	1402.59	1402.72
782.40	782.33	7	N	11	1301.56	1301.68
895.49	895.40	8	I	10	1187.53	1187.63
1026.53	1026.43	9	M	9	1074.46	1074.55
1141.56	1141.47	10	D	8	943.42	943.51
1254.64	---	11	L	7	828.41	828.48
1383.68	---	12	E	6	715.34	715.40
1454.72	---	13	A	5	586.30	586.36
1525.76	---	14	A	4	515.27	515.32
1751.89	1751.69	15	K*	3	444.24	444.28
1822.93	1822.77	16	A	2	218.13	218.15
---	---	17	K	1	147.09	147.11
Fragment m/z values were calculated for K*, which is the modified lysyl residue after loss of the labile base. See Figure 7 (main text) for the structure of the modified lysine.						

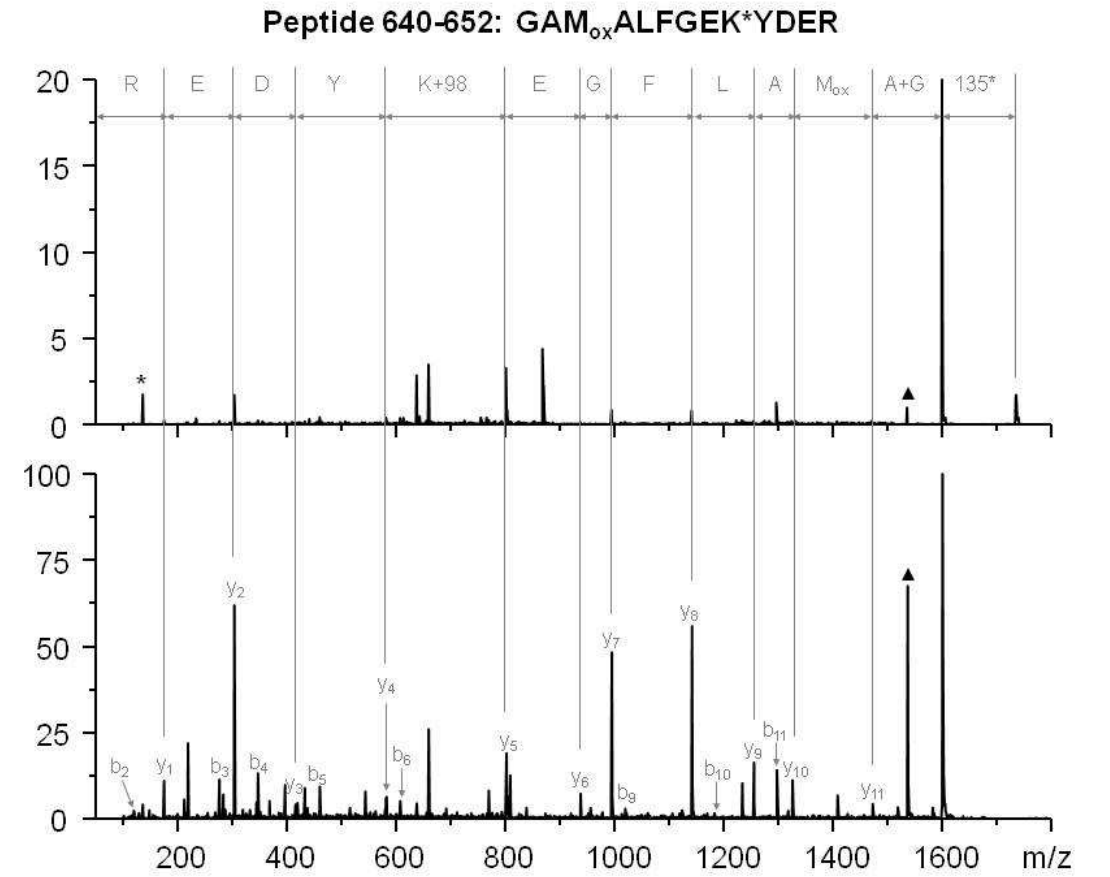
Figure S8e



Solid triangle indicates the loss of the adenine and CH_3SOH from the parent ion.

Peptide 638-648: AK*GAM$_{ox}$ALFGEK						
b-ion m/z calculated	b-ion m/z measured	Residue position	Sequence	Residue position	y-ions m/z measured	y-ions m/z calculated
---	---	1	A	11	---	---
298.18	298.17	2	K*	10	1165.56	1165.59
355.20	355.19	3	G	9	939.42	939.46
426.23	426.23	4	A	8	---	882.44
573.27	573.27	5	M_{ox}	7	811.37	811.40
644.31	644.28	6	A	6	664.35	664.37
757.39	757.37	7	L	5	593.31	593.33
904.46	904.43	8	F	4	480.23	480.25
961.48	961.46	9	G	3	333.17	333.18
1090.52	1090.47	10	E	2	276.14	276.16
---	---	11	K	1	147.11	147.11
Fragment m/z values were calculated for K*, which is the modified lysyl residue after loss of the labile base. See Figure 7 (main text) for the structure of the modified lysine.						

Figure S8f


Peptide 640-652: GAM$_{ox}$ALFGEK*YDER
R
E
D
Y
K+98
E
G
F
L
A
M$_{ox}$
A+G
135*
m/z
Solid triangle indicates the loss of the adenine and CH$_3$SOH from the parent ion.

Peptide 640-652: GAM$_{ox}$ALFGEK*YDER						
b-ion m/z calculated	b-ion m/z measured	Residue position	Sequence	Residue position	y-ions m/z measured	y-ions m/z calculated
---	---	1	G	13	---	---
129.07	129.06	2	A	12	---	1543.71
276.10	276.10	3	M$_{ox}$	11	1472.60	1472.67
347.14	347.13	4	A	10	1325.58	1325.64
460.22	460.22	5	L	9	1254.55	1254.60
607.29	607.28	6	F	8	1141.49	1141.52
664.31	---	7	G	7	994.42	994.45
793.35	793.33	8	E	6	937.40	937.43
1019.49	1019.46	9	K*	5	808.36	808.38
1182.55	1182.55	10	Y	4	582.24	582.25
1297.58	1297.53	11	D	3	419.17	419.19
1426.62	1426.54	12	E	2	304.15	304.16
---	---	13	R	1	175.11	175.12
Fragment m/z values were calculated for K*, which is the modified lysyl residue after loss of the labile base. See Figure 7 (main text) for the structure of the modified lysine.						