# Supporting Information

# Mildly Acidic Conditions Eliminate Deamidation Artifact during Proteolysis: Digestion with Endoprotease Glu-C at pH 4.5

Shanshan Liu, Kevin Ryan Moulton, Jared Robert Auclair, Zhaohui Sunny Zhou*

Barnett Institute of Chemical and Biological Analysis, Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts 02115, United States

*To whom correspondence should be addressed. Email: z.zhou@neu.edu

# Table of Contents

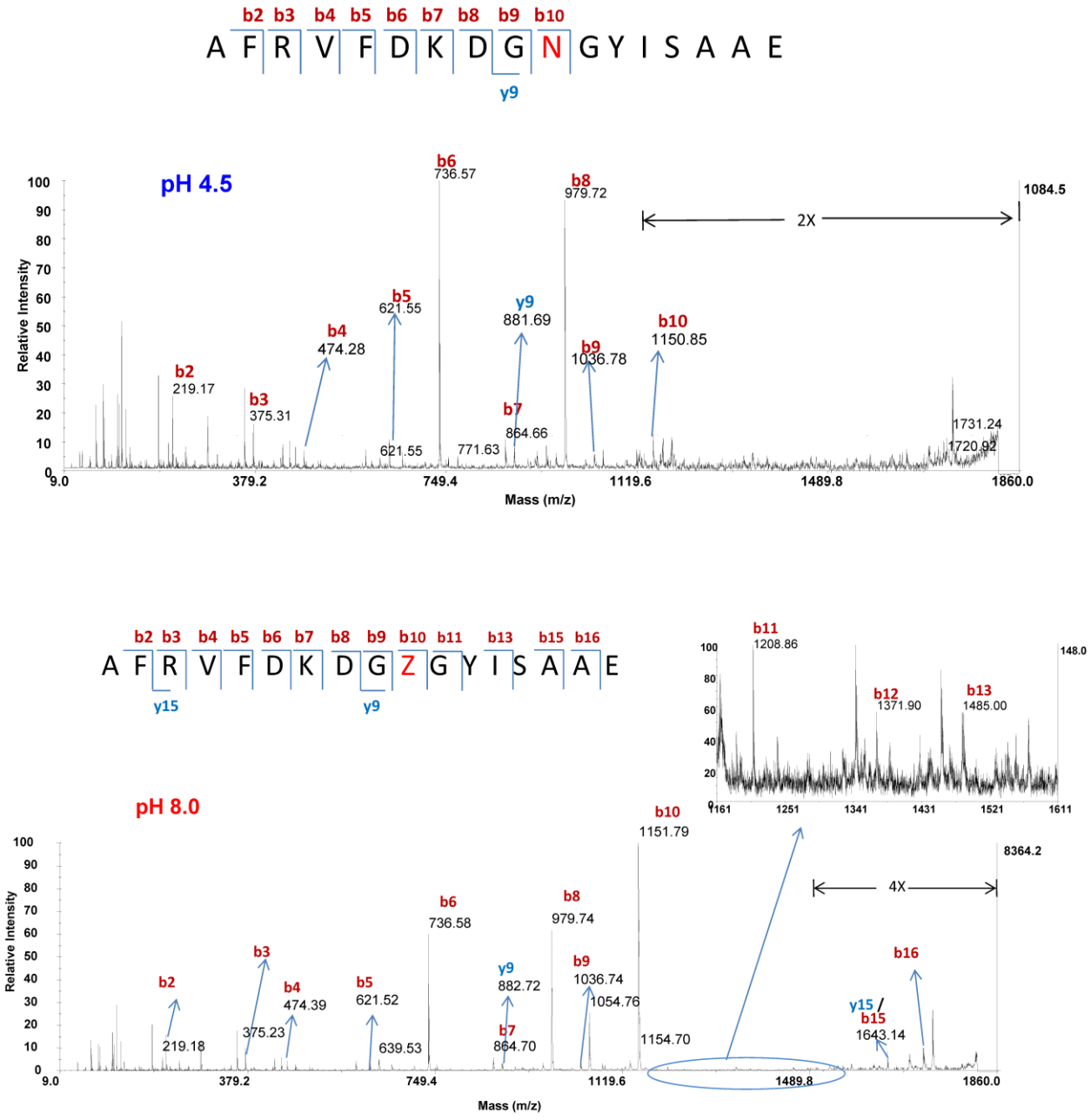


**Fig. S-1** MALDI-TOF/TOF MS/MS spectra of singly charged precursor ion m/z 1859.8 (theoretical m/z 1859.8 for Asn, top) and m/z 1860.8 (theoretical m/z 1859.8 for Asp/isoAsp, bottom) for calmodulin Glu-C calmodulin peptide $^{87}$AFRVFDKDG**NG**YISAAE$^{104}$ after 24 h digestion at pH 4.5 (top) and 8.0 (bottom) in positive ion mode. The letter Z denotes either Asp or isoAsp.

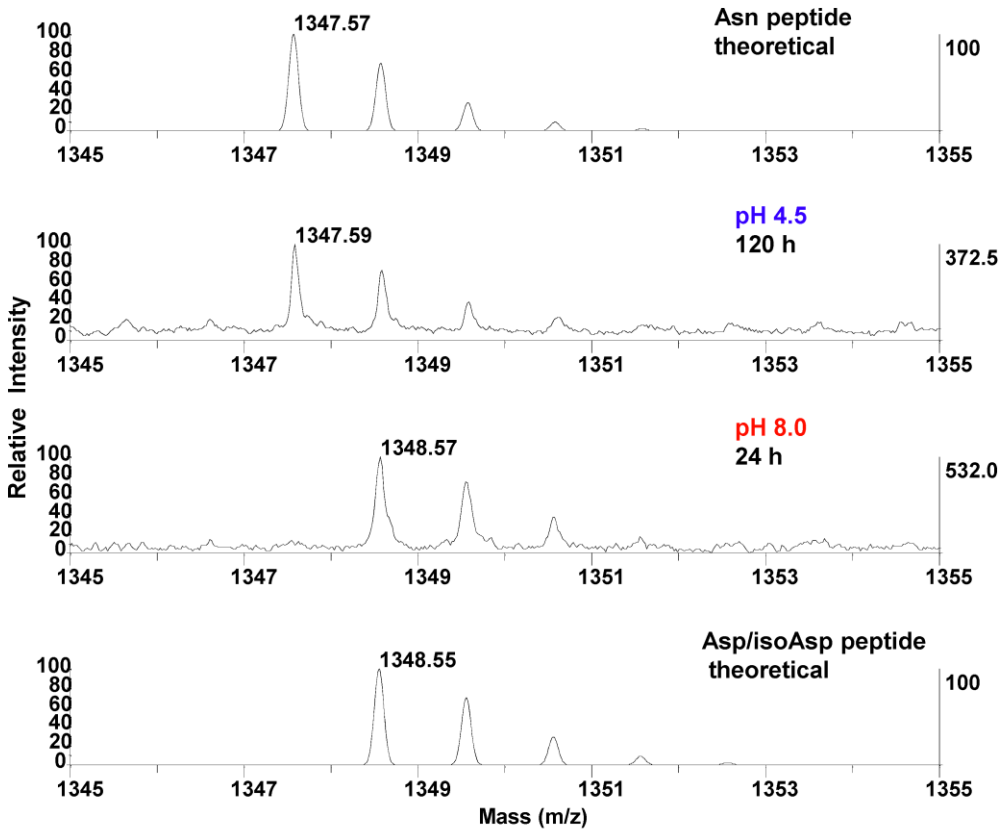


**Fig. S-2** MALDI-TOF MS spectra of calmodulin Glu-C peptide ($^{55}$VDADG**NG**TIDFPE$^{67}$; theoretical m/z 1347.57) in negative ion mode. No observable deamidation was present after 120 h at pH 4.5 in ammonium acetate, while complete deamidation was observed after 24 h at pH 8.0 in ammonium bicarbonate. Theoretical isotopic envelope for Asn and Asp/isoAsp peptides are shown in top and bottom traces respectively.

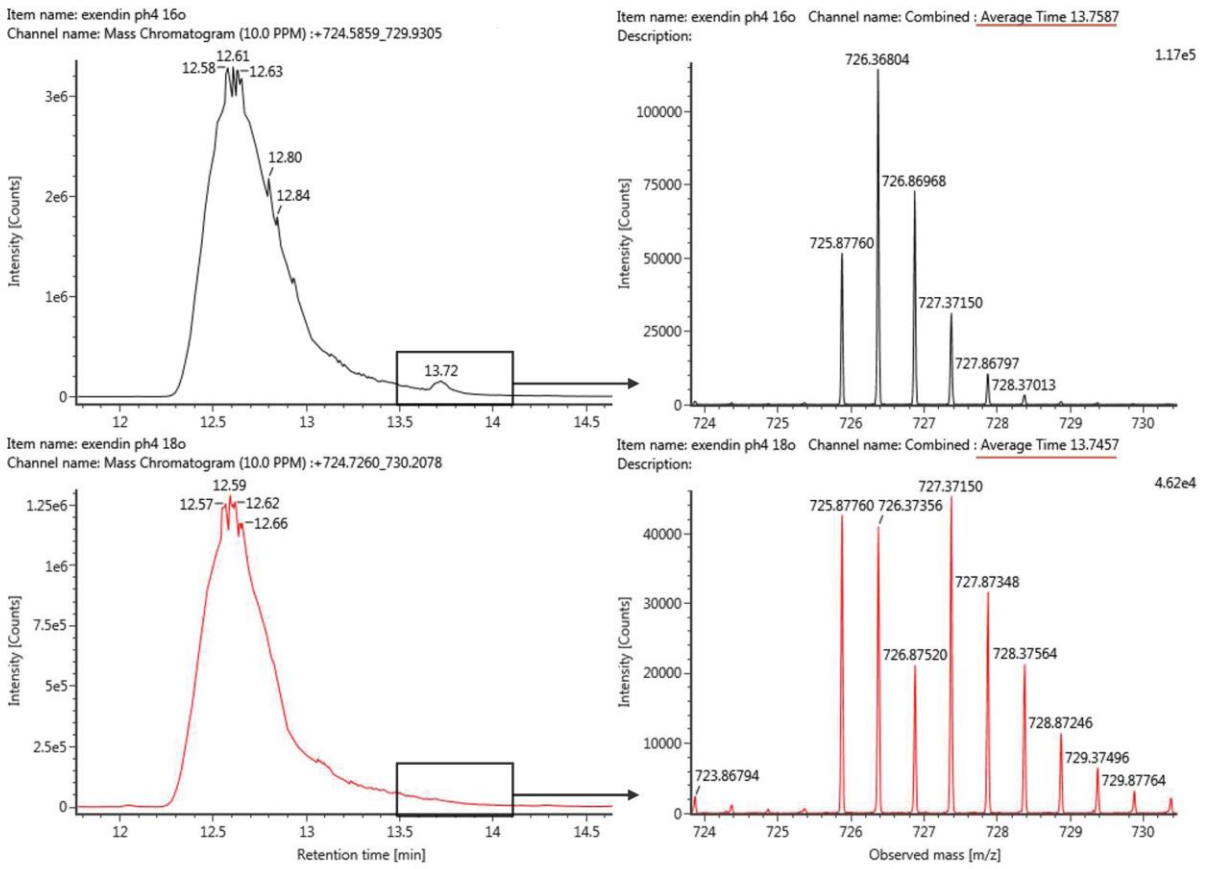


**Fig. S-3** Extracted ion chromatograms (XIC) of doubly charged exenatide Glu-C peptide (left column, $^{25}$WLK**<u>NG</u>**GPSSGAPPPS$^{39}$-$NH_2$, (C-terminal amide), theoretical m/z 725.87 Da) after 48 h proteolysis at pH 4.5 in $^{16}O$ (top) and $^{18}O$ (bottom) water and ESI MS spectra of aspartyl species extracted from XIC (right column). Because only trace amount of deamidation species was generated, considerable peak overlapping in mass spectra was observed. Due to the dominant asparaginyl species, the aspartyl species were only partially resolved by liquid chromatography.

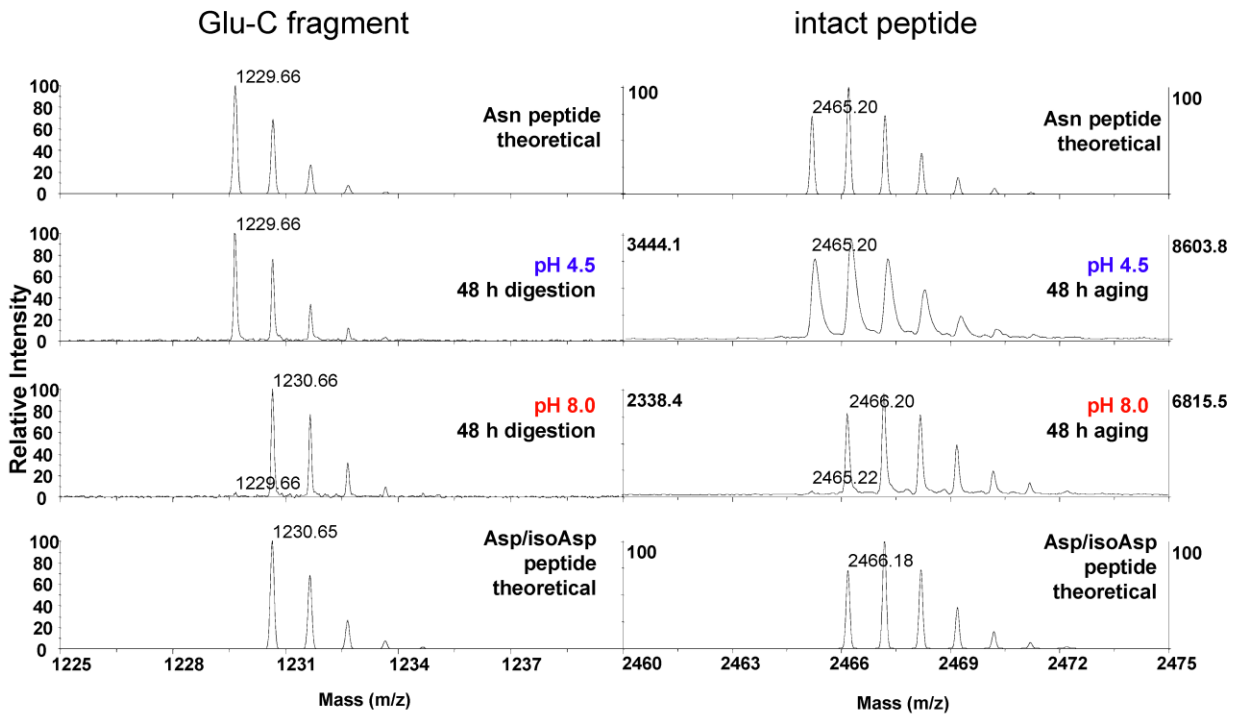


**Fig. S-4** MALDI-TOF MS spectra of ACTH Glu-C peptide (left traces, $^{18}$RPVKVYP**NG**AE$^{28}$; theoretical m/z 1229.66), and ACTH intact peptide after aging (right traces, $^{18}$RPVKVYP**NG**AEDESAEAFPLEF$^{39}$, theoretical m/z 2465.20). At pH 4.5 in ammonium acetate, no deamidation was observed after 48 h digestion and after 48 h aging of the intact ACTH peptide. At pH 8.0 in ammonium bicarbonate, nearly complete deamidation was observed in both the Glu-C digested ACTH peptide and the intact aged ACTH peptide. Theoretical isotopic envelope for Asn and Asp/isoAsp peptides are shown in top and bottom traces respectively.

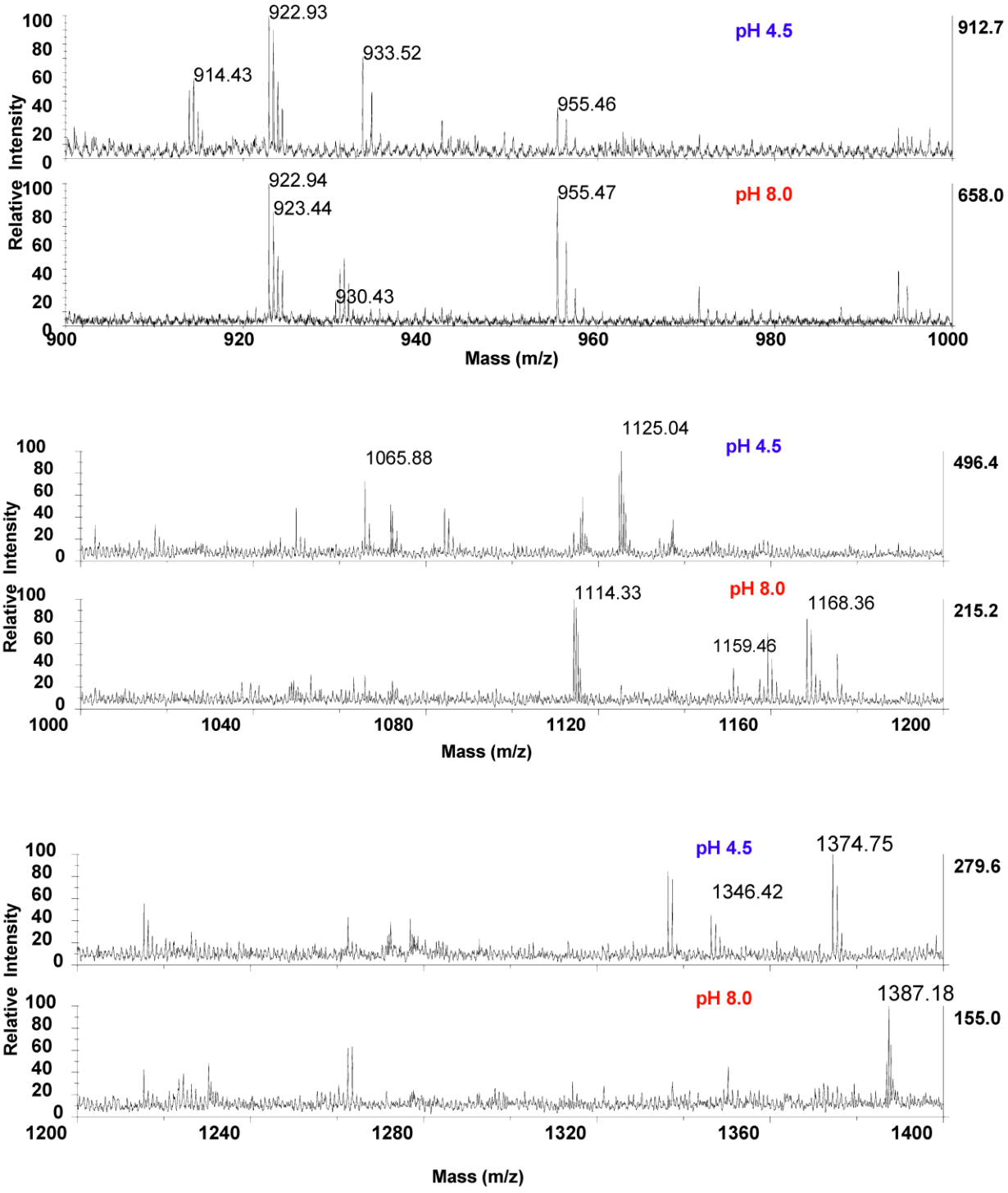


**Fig. S-5** Comparison of MALDI-TOF MS spectra of calmodulin Glu-C digests at pH 4.5 (top) and pH 8 (bottom).

**Fig. S-5 continued**

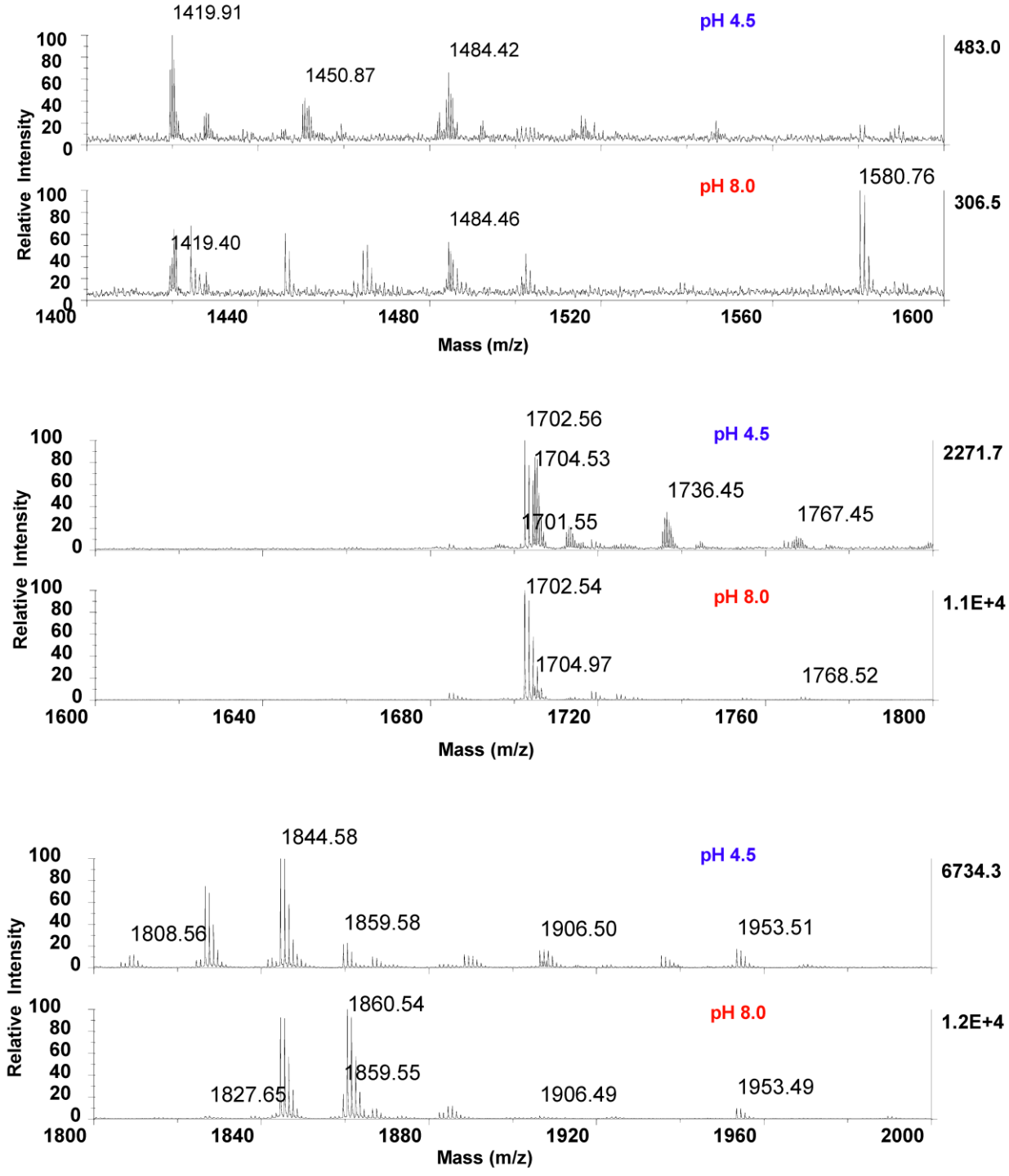

**Fig. S-5 continued**

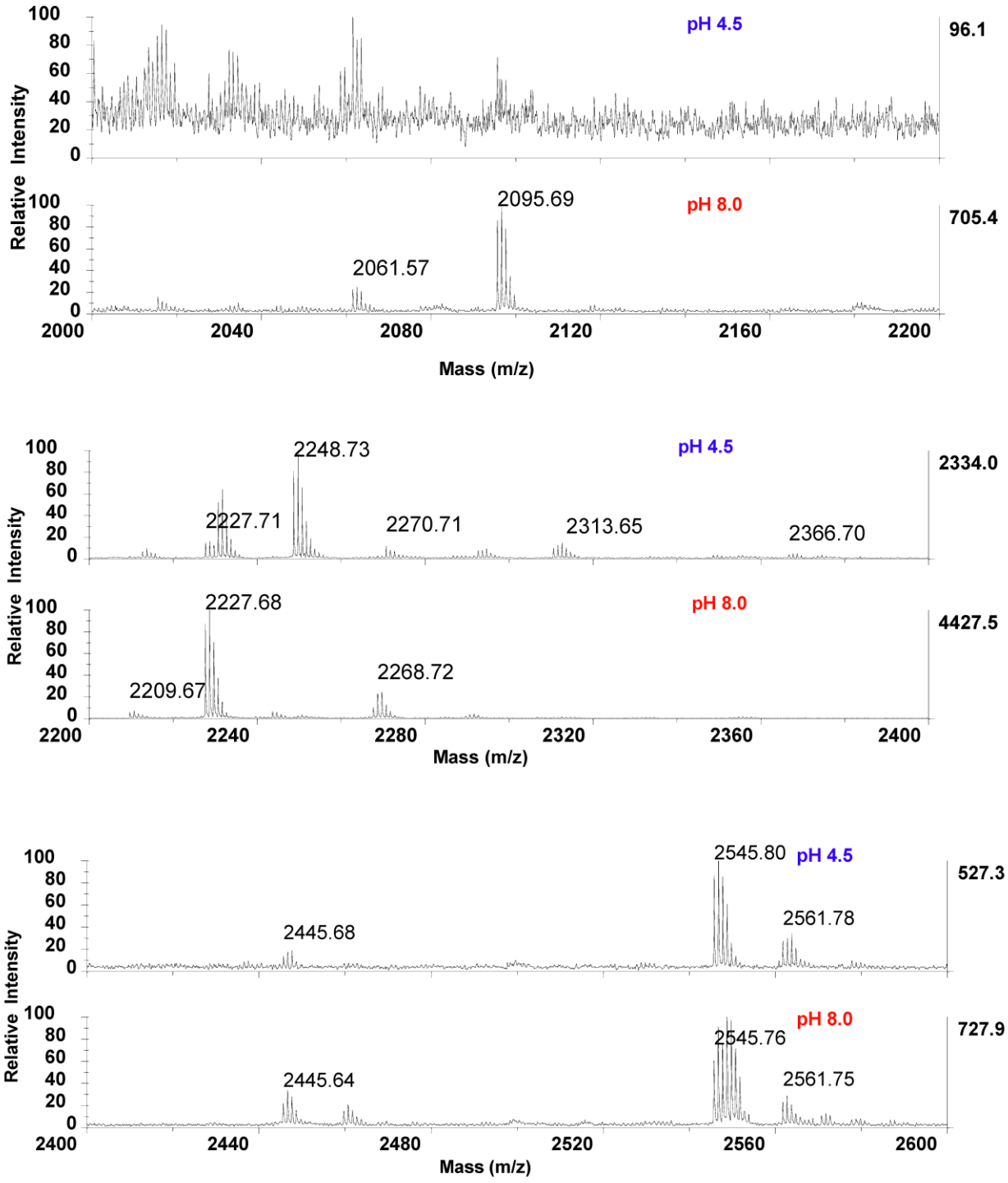

**Fig. S-5 continued**

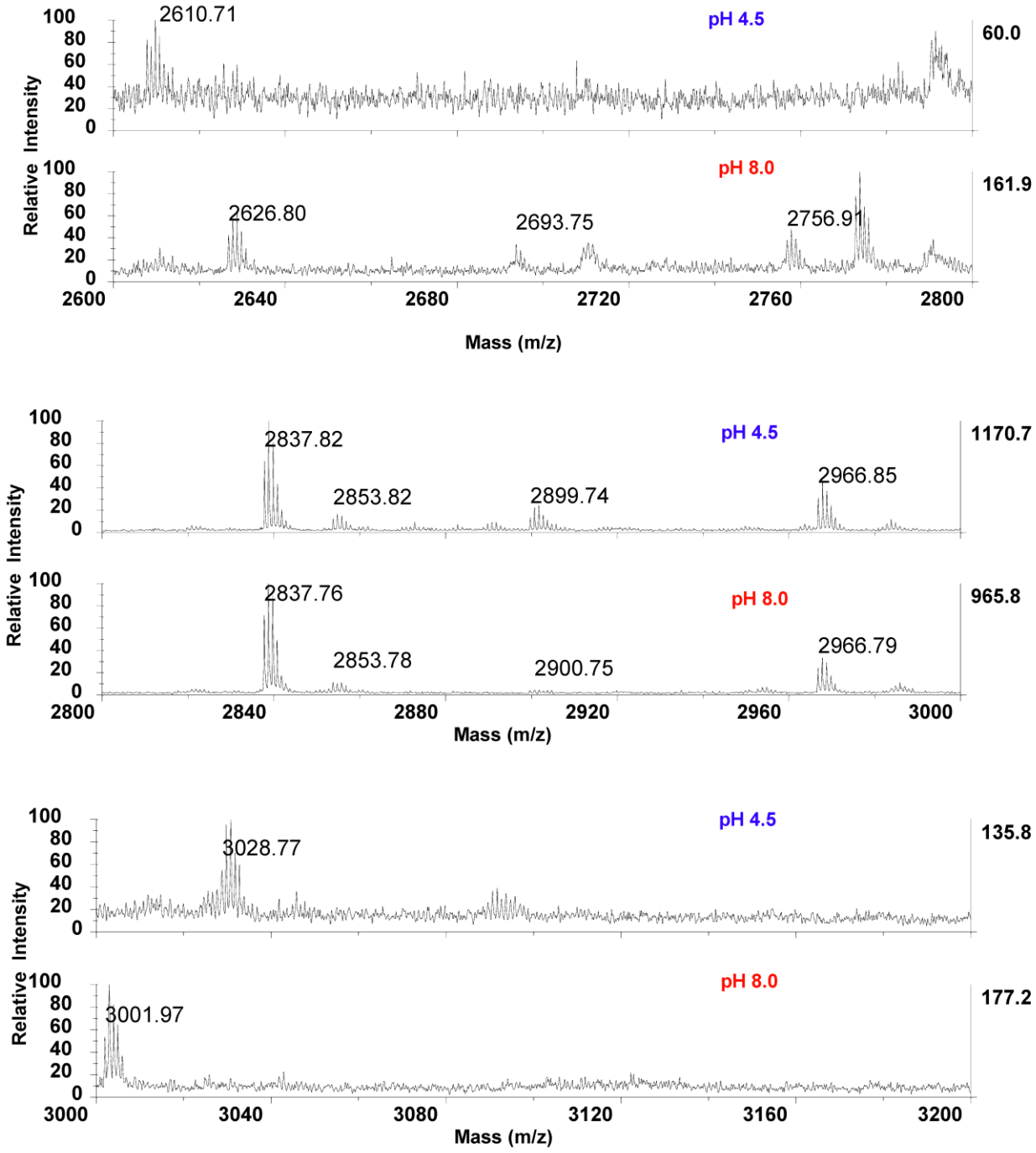

**Fig. S-5 continued**

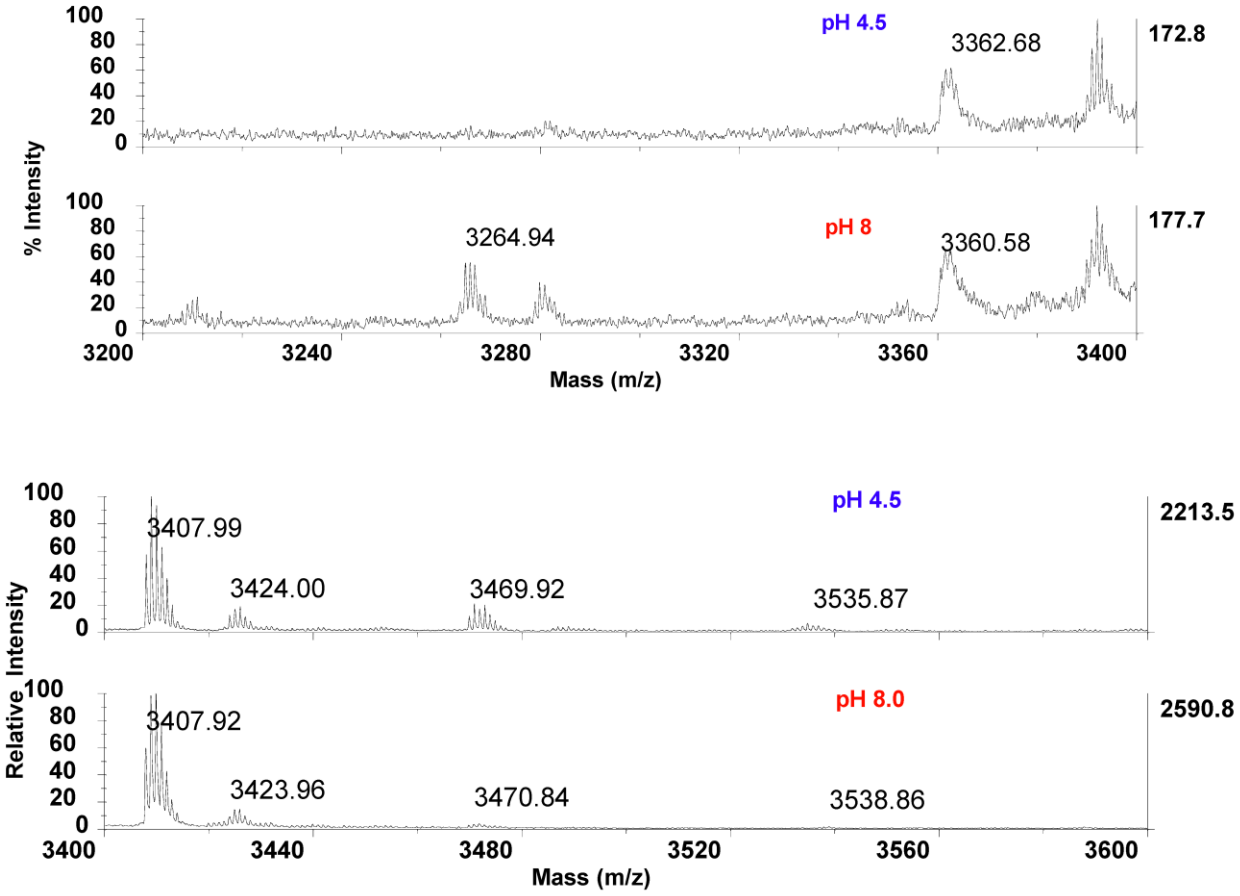

**Table S-1. Matching Peptides for Specific Cleavage at Glutamic acid by Glu-C**

| SEQUENCE | position | theoretical (m/z) | pH 4.5 relative intensity (%) | pH 8.0 relative intensity (%) |
|---|---|---|---|---|
| (E)FVQMMTAK (C-Terminus) | 141-148 | 955.47 | 4.5 | 4.8 |
| (E)EFVQMMTAK (C-Terminus) | 140-148 | 1084.52 | 3.7 | |
| (E)ADIDGDGQVNYEE(F) | 128-140 | 1424.58 | | 1.8 |
| (E)LGTVMRSLGQNPTEAE(L) | 32-45 | 1702.84 | 33 | 91 |
| (E)AFSLFDKDGDGTITTKE(L) | 15-31 | 1844.89 | 100 | 100 |
| Doubly charged | | 922.95 | 13 | 5.3 |
| (E)AFRVFDKDGNGYISAAE(L) | 88-104 | 1859.89 | 22 | 95 |
| Doubly charged | | 930.45 | | 2.4 |
| (E)LRHVMTNLGEKLTDEE(V) | 105-120 | 1884.95 | 3.5 | 11.4 |
| (E)MIREADIDGDGQVNYEE(F) | 124-140 | 1953.85 | 17 | 9.4 |
| (E)FLTMMARKMKDTDSEEE(I) | 68-84 | 2061.93 | | 1.5 |
| (E)LRHVMTNLGEKLTDEEVDE(M) | 105-123 | 2228.09 | 5.8 | 37 |
| Doubly charged | | 1114.54 | | 1.9 |
| (E)FKEAFSLFDKDGDGTITTKE(L) | 12-31 | 2249.10 | 20 | 2.5 |
| Doubly charged | | 1125.05 | 7.4 | |
| (E)FLTMMARKMKDTDSEEEIRE(A) | 68-87 | 2460.16 | | 1.3 |
| GSSHHHHHHSSGLVPRGSHM ADQLTE(E) | His-tag-6 | 2838.30 | 17 | 8.1 |
| Doubly charged | | 1419.65 | 7.4 | 2 |
| (E)LGTVMRSLGQNPTEAELQDM INE(V) | 32-54 | 2546.22 | 7.8 | 6.2 |
| (E)MIREADIDGDGQVNYEEFVQ MMTAK (C-Terminus) | 124-148 | 2890.31 | 1.8 | |
| GSSHHHHHHSSGLVPRGSHMADQLTEE(Q) | His-tag -7 | 2967.34 | 8 | 2.8 |
| (E)VDADGNGTIDFPEFLTMMARKMKDTDSEEE(I) | 55-84 | 3392.50 | 2.7 | 1.5 |
| GSSHHHHHHSSGLVPRGSHMADQLTEEQIAE(F) | His-tag -11 | 3408.56 | 32 | 22 |
| (E)IREAFRVFDKDGNGYISAAELRHVMTNLGE(K) | 85-114 | 3408.71 | | |
| Doubly charged | | 1704.8 | 29 | |
| GSSHHHHHHSSGLVPRGSHMADQLTEEQIAEFKE(A) | His-tag -14 | 3812.77 | 4.5 | 0.6 |

## Table S-2. Peptides for Non-specific Cleavage

The putative peptide sequences were obtained by matching the fragment ions from MALDI-TOF to the sequence of recombinant calmodulin via FindPept tool on ExPASy(Gattiker et al. 2002). (web.expasy.org/findpept/)

| SEQUENCE | position | theoretical (m/z) | pH 4.5 relative intensity (%) | pH 8.0 relative intensity (%) |
|---|---|---|---|---|
| (M)ADQLTEEQ(I) | 1-8 | 933.42 | 9.2 | |
| (E)AELQDMIN(E) | 46-53 | 933.44 | | |
| (T)KELGTVMR(S) | 30-37 | 933.52 | | |
| (E)LGTVMRSLG(Q) | 32-40 | 933.52 | | |
| (S)SHHHHHHSSG(L) | His-tag | 1159.49 | | 1.5 |
| (M)GSSHHHHHHS(S) | His-tag | 1159.49 | | |
| (L)GTVMRSLGQNP(T) | 33-43 | 1159.59 | | |
| (I)REAFRVFDK(D) | 86-94 | 1167.63 | | 1.6 |
| (E)AFRVFDKDGN(G) | 88-97 | 1168.58 | | |
| (L)FDKDGDGTITT(K) | 19-29 | 1169.53 | | |
| (E)FKEAFSLFDKD(G) | 12-22 | 1346.66 | 2.1 | |
| (N)GYISAAELRHVM(T) | 98-109 | 1346.69 | | |
| (Q)NPTEAELQDMIN(E) | 42-53 | 1374.62 | 4.2 | |
| (E)MIREADIDGDGQV(N) | 124-136 | 1418.66 | 7.1 | 1.8 |
| (D)IDGDGQVNYEEFV(Q) | 130-142 | 1484.65 | 4.8 | 1.5 |
| (D)TDSEEEIREAFRV(F) | 79-91 | 1580.76 | | 2.6 |
| (I)READIDGDGQVNYE(E) | 126-139 | 1580.68 | | |
| keratin peaks (KRT1)<br>(E)ISELNRVIQRLRSE(I) | | 1712.98 | 4.4 | |
| keratin peaks (KRT2A)<br>(E)VKAQYEEIAQRSKEE(A) | | 1807.92 | 12 | |
| (S)SGLVPRGSHMADQLTEE(Q) | His-tag-7 | 1826.87 | 70 | |
| Doubly charged | | 913.93 | 7.5 | |
| (A)AELRHVMTNLGEKLTD(E) | 103-117 | 1826.97 | 70 | 2.5 |
| UNKOWN | | 1866.88 | 10.5 | 9.4 |
| (R)SLGQNPTEAELQDMINE(V) | 38-54 | 1888.86 | 12 | |
| (G)DGTITTKELGTVMRSLGQ(N) | 24-41 | 1906.99 | 15.7 | 2.6 |

**Table S-2 continued**

| SEQUENCE | position | theoretical (m/z) | pH 4.5 relative intensity (%) | pH 8.0 relative intensity (%) |
|---|---|---|---|---|
| UNKOWN | | 1935.84 | 11.1 | |
| (E)/QIAEFKEAFSLFDKDGDGTI(T) | 8-27 | 2231.09 | 21 | |
| autolysis of Glu-C | | 2269.14 | | 8.3 |
| autolysis of Glu-C | | 2549.23 | | 6.1 |
| (K)DTDSEEEIREAFRVFDKDGNGYISAAELRH(V) | 78-107 | 3469.63 | 6.9 | |
| (D)SEEEIREAFRVFDKDGNGYISAAELRHVMT(N) | 81-109 | 3469.68 | | |
| Doubly charged | | 1735.75 | 11.6 | |

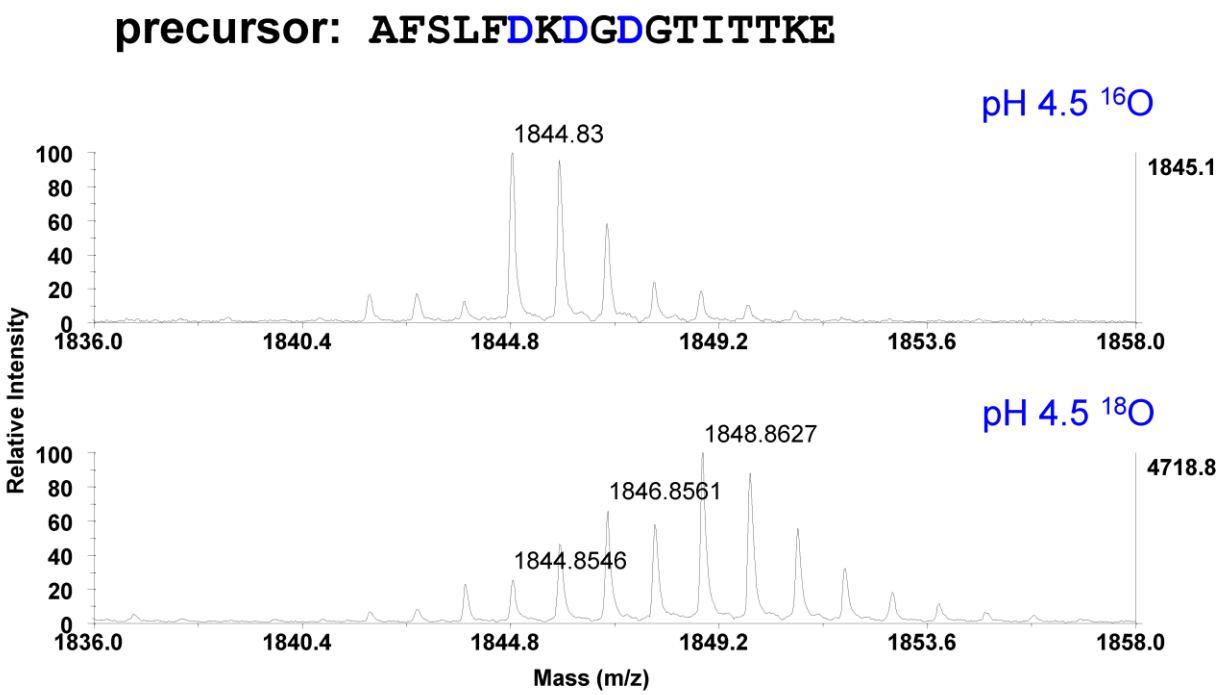


**Fig S-6A** MALDI-TOF MS spectra of calmodulin Glu-C peptide ($^{15}$AFSLFDK**DGDG**TITTKE$^{31}$, theoretical m/z 1844.89) after 96 h digestion at pH 4.5 in $^{16}$O water (top) and $^{18}$O water (bottom); the DG tandem (22-25, underlined in sequence above) is prone to isomerization. Due to varying degrees of $^{18}$O incorporation into the C-terminal, isomerization cannot be determined by MS spectra alone. Tandem MS/MS was therefore acquired to evaluate isomerization, which can be found in Fig S-B.

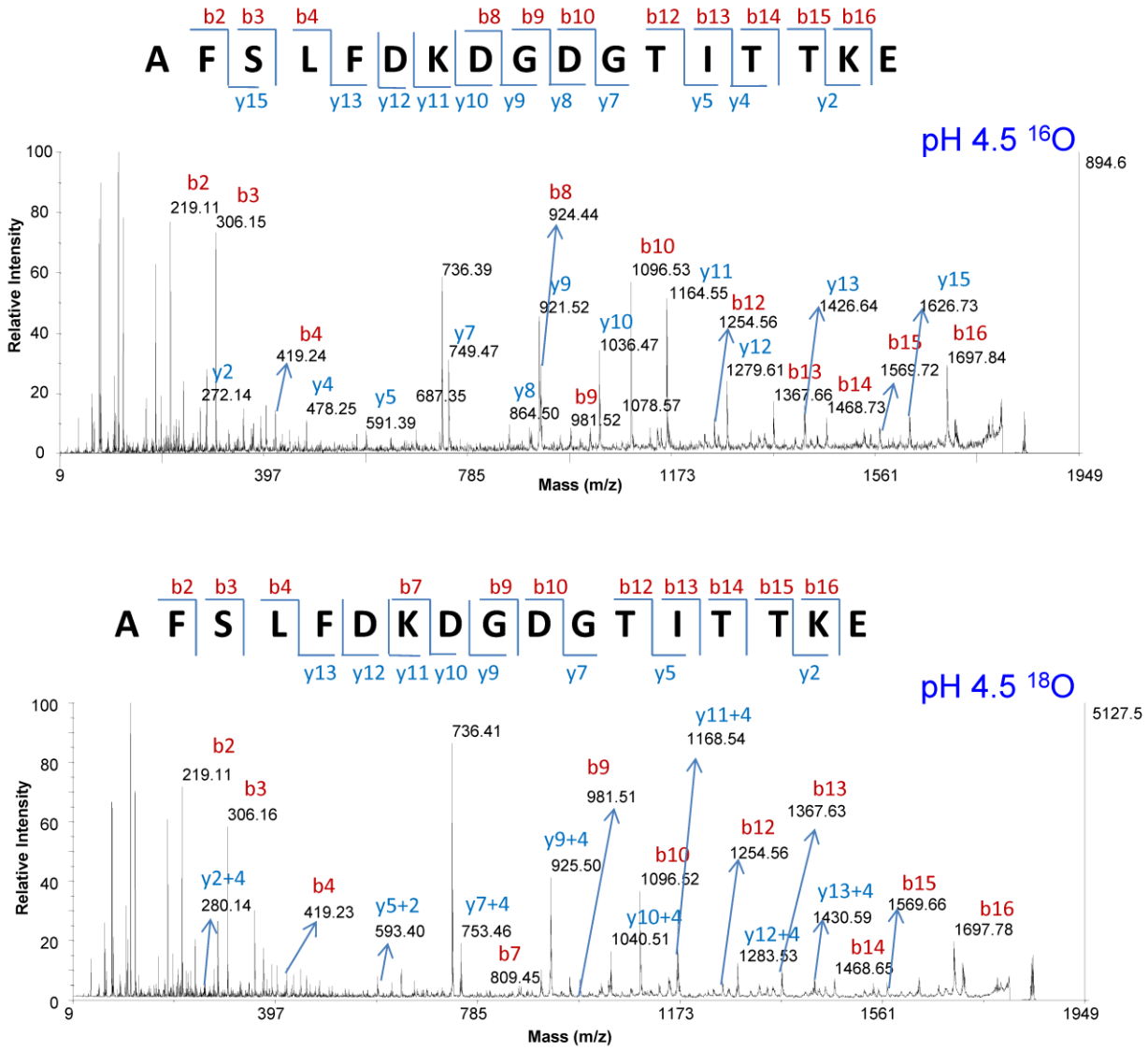


**Fig S-6B** MALDI-TOF/TOF MS/MS spectra of precursor ion m/z 1844.8 for calmodulin Glu-C peptide $^{15}$AFSLFDK**<u>DGDG</u>**TITTKE$^{31}$ after 96 h digestion at pH 4.5 in $^{16}$O water (top) and $^{18}$O water (bottom). No mass shift was detected in the b ion series between $^{16}$O and $^{18}$O proteolysis, while up to a 4 Da mass shift was observed in the y ion series. Therefore these results indicate that isomerization did not occur during proteolysis, and that all $^{18}$O incorporation is from the newly generated C-terminus.

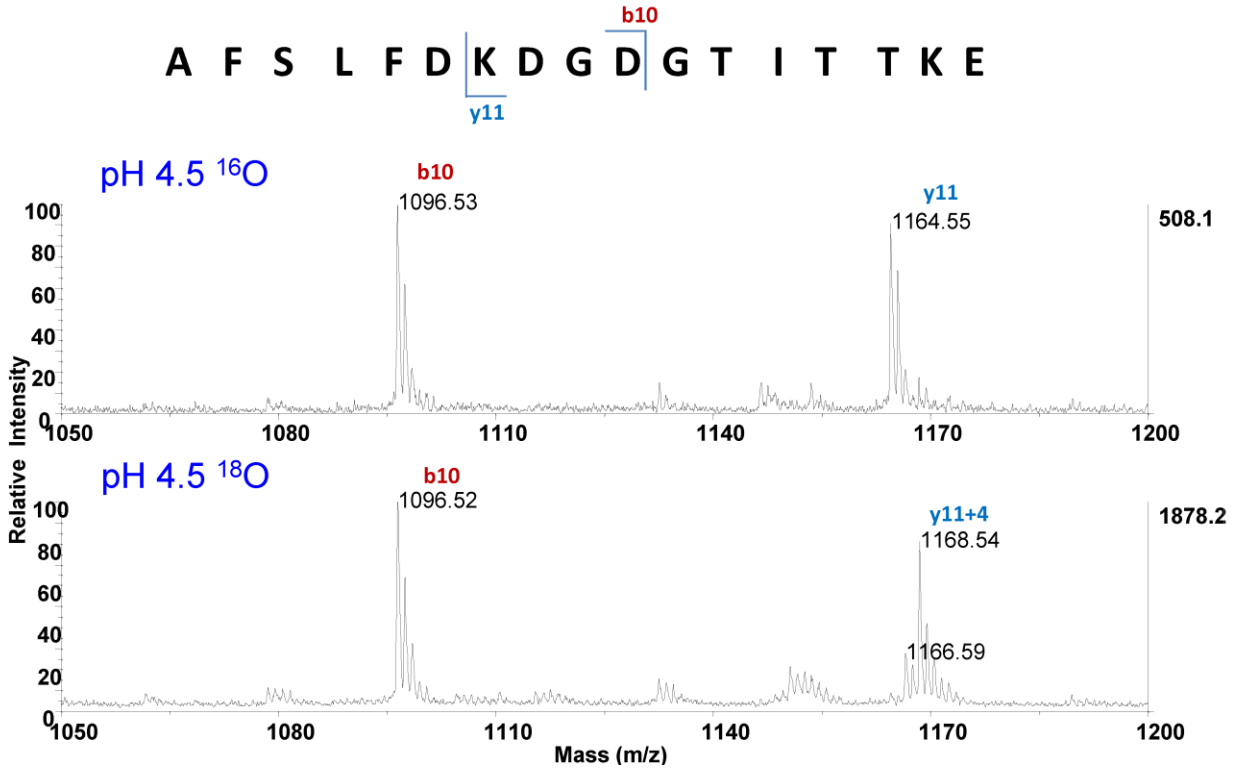


**Fig S-6C** Zoomed-in MALDI-TOF/TOF MS/MS spectra of precursor ion m/z 1844.8 for calmodulin Glu-C peptide $^{15}$AFSLFDK**DGDG**TITTKE$^{31}$ after 96 h digestion at pH 4.5 in $^{16}$O water (top) and $^{18}$O water (bottom); the isotopic envelopes of indicative fragment ions are shown. The b10 ion from $^{18}$O proteolysis shows a nearly identical isotopic envelope compared with $^{16}$O proteolysis, demonstrating that no Asp isomerization occurred during proteolysis. Additionally, the mass shift of the y11 ion (bottom trace) reveals one to two atoms of $^{18}$O incorporation at the C-terminus.

## References

Gattiker A, Bienvenut WV, Bairoch A, Gasteiger E (2002) FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification PROTEOMICS 2:1435-1444 doi:10.1002/1615-9861(200210)2:10<1435::aid-prot1435>3.0.co;2-9