Collisional dissociation of 115 deprotonated peptides of 5-28 amino acids—an updated report on fragment ion types, neutral loss species, and the effect of acidic residues

Mei-Qing Zuo[1,2,3], Rui-Xiang Sun[2,3], Meng-Qiu Dong[2,3]

[1]College of Biological Sciences, China Agricultural University, Beijing, China
[2]National Institute of Biological Sciences, Beijing, China
[3]Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing, China

Correspondence
Meng-Qiu Dong: dongmengqiu@nibs.ac.cn

ABSTRACT

For peptides that are refractory to positive-ion-mode mass spectrometry (MS) analysis, the negative-ion-mode MS analysis is an attractive but insufficiently explored alternative. Based on high-resolution and accurate-mass MS analysis of 115 synthetic peptides of 5-28 amino acids, we confirmed that higher-energy collisional dissociation (HCD) of deprotonated peptides induced abundant neutral losses (NL) from the backbone and the side chains, and updated the ranking list of NLs by abundance. The most abundant fragment ion types are *y*- > *x*-, *z*- > *c*- if the NL ions are included, or *c*- > *y*- > *z*- > *b*- if not. The most frequent side-chain NLs include among others a loss of CH_2O, C_2H_4O, and CH_2N_2 from Ser, Thr, and Arg, respectively. Although NL of CO_2 is common for all peptides, it is markedly enhanced in Asp/Glu-containing peptides. Long peptides and Asp/Glu-rich peptides are prone to carrying multiple negative charges. HCD spectra produced from multiply deprotonated peptides have a lower fraction of sequencing ions (i.e., *a*-, *b*-, *c*-, *x*-, *y*-, *z*- ions) than those produced from 1^- and 2^- precursors. With this knowledge, we predict that for negative-ion-mode MS analysis of highly acidic proteins such as pepsin, which is difficult for positive-ion-mode analysis, Glu-C is a better choice than trypsin in sample digestion.

Keywords: Negative-ion-mode mass spectrometry, deprotonated peptides, peptide anions, negative-ion HCD, collision induced dissociation, neutral loss

1. Introduction

Mass spectrometry (MS)-based proteomics has matured in recent years as manifested by the deep MS analyses of the human proteome in which up to 17,100 protein groups (86.5% of the total by estimation) have been identified (*1, 2*). However, not every protein is readily identifiable, for example, highly acidic proteins tend to escape MS identification (*3*). For proteins that are amenable to MS-based proteomics analysis, the coverage of protein sequence is incomplete. In standard proteomics analysis, which is carried out in the positive ion mode, acidic peptides are harder to identify than hydrophobic peptides containing basic residues (*4*).

Theoretically, negative-ion-mode MS analysis can complement the ubiquitous positive-ion-mode analysis by providing better identification of acidic peptides. Acidic peptides include the ones that are rich in acidic amino acid (aa) residues aspartic acid (Asp or D) or glutamic acid (Glu or E), and the ones that are heavily modified by negatively charged post-translational modifications (e.g., phosphorylation or sulfation). By their physicochemical nature, acidic peptides ionize better in the negative mode than in the positive mode and thus may be better analyzed as peptide anions. Several published studies have shown that negative-ion-mode MS analysis of peptides yields information unattainable in positive-ion-mode MS (*3, 5, 6*).

However, investigations on negative-ion-mode MS of peptides are limited. Bowie and colleagues examined fragmentation of singly-charged peptide anions using mostly collision-induced dissociation (CID) (*7-10*) and summarized their findings in three reviews (*11-13*). Cassady and co-worker studied the fragmentation of several doubly or multiply charged peptide anions by CID (*5*). Using negative electron transfer dissociation (NETD), Coon and co-workers performed proteomics analysis (*3, 6*) while the Brodbelt team did similar analysis using 193-nm ultraviolet photodissociation (UVPD) in the negative ion mode (*14-16*). The numbers of peptides identified in the negative ion mode by either NETD or UVPD are a fraction of those identified in the positive ion mode. This is in part due to a lack of software tools optimized for negative-ion-mode MS data analysis, which in turn has to do with an insufficient understanding of the fragmentation behaviors of peptide anions (or deprotonated peptides, used interchangeably in this paper).

Several studies have shown that fragmentation of deprotonated peptides by either CID or NETD generates an abundance of neutral loss (NL) peaks including a variety of losses from amino acid side chains (*17-22*). For example, threonine, serine, and tryptophan generate characteristic NL of 44.02621 (C_2H_4O), 30.01056 (CH_2O), and 129.05785 (C_9H_7N) Da, respectively (*18, 19*). In a 2019 study, we show that the

number of fragment ions with NL is greater than the number of fragment ions without NL, and the summed intensity of the former is much higher than the summed intensity of the latter (*23*). Therefore, peptide identification in negative-ion-mode MS analysis must take into consideration the NL peaks in MS2 spectra.

However, as the model peptides used in the previous studies are limited in one way or another, characterization of NL of deprotonated peptides is incomplete. In most studies, only a small number of synthetic peptides were analyzed, typically less than ten and occasionally several dozens. One exception is a 2017 study by Liang et al., in which 527 model peptides were analyzed, but they were very short peptides—137 had 3-6 residues and the rest were dipeptides.

In this study, we analyzed the negative-ion-mode MS2 spectra of 115 synthetic peptides of 5-28 residues. The MS2 spectra were generated using higher-energy collisional dissociation (HCD), which is a quadrupole or beam-type CID. We focused on finding the different varieties of NL peaks and the effects of acidic residues on MS analysis of deprotonated peptides. The frequencies of neutral losses associated with different amino acids are updated, as are the occurrence frequency and relative abundance of six ion types (*a-, b-, c-, x-, y-,* and *z-*). The occurrence and intensity of CO_2-loss peaks are associated with the presence of D/E in peptide anions but largely independent of the charge state of the precursors. D/E-rich peptides and long peptides tend to have more negative charges. When a peptide carries more than two negative charges, the percentage of sequencing ions in HCD spectra decreases markedly. This leads to a prediction that Glu-C digestion is better than trypsin digestion for negative ion-mode HCD analysis of highly acidic proteins.

2. Materials and methods

2.1 Materials and sample preparation

Three groups of 12 synthetic peptides (Table 1) were synthesized by Guoping Chemical Ltd. (Hefei, China). Quantified with the extinction coefficients listed in Table 1 and adjusted the peptide concentration using the NanoDrop One (Thermo scientific) at 205 nm to make similar concentrations of all the 4 peptides within the group (*24*). For the 8 synthetic peptides in group 1 and 3 (AAAAAD/E/K/R, GYWGTNLGQPHSLATD), each of them was diluted with 7.5 mM hexylamine (Energy chemical, Shanghai, China), 50% acetonitrile (Thermo Scientific) and sprayed into the mass spectrometer using a syringe pump at the concentrations as table 1 listed. While the 4 peptides in group 2 (IAVNYCMFGD/E/K/R) contain cysteine and exist as disulfide-bonded doubles. Therefore, the peptide solutions of similar concentrations were mixed with equal volume, reduced by 5 mM TCEP, followed by alkylation with 10 mM iodoacetamide, and finally sprayed into mass spectrometer as a mixture.

Besides, we obtained other 115 synthetic peptides as Supplemental Table S1 listed. Among them, 1-37 are used earlier for a different purpose (*25*), 38-55 are gifts from

Dr. Ping Xu from National Center for Protein Sciences (Beijing, China), and 56-115 are synthesized by Scilight Biotechnology (Beijing, China). Peptides are diluted to 10 μM in 7.5 mM hexylamine, 50% acetonitrile before being sprayed into mass spectrometer as previously reported (*23*).

2.2 MS parameters

All experiments were performed on Orbitrap Fusion™ Lumos™ Tribrid™Mass Spectrometer (Thermo Fisher Scientific). For the synthetic peptides, the sample solution was introduced into the electrospray source at a flow rate of 2 μL/min with a syringe pump. The ion spray voltage was 2.5 kV in the negative ion mode. The temperature of ion transfer tube was 320 °C. The MS1 mass ranges were 150 to 2000 m/z and adjusted according to the mass of each peptide, while the MS2 mass ranges were adjusted automatically. Microscans were 3 for both MS1 and MS2. The AGC target was 1E6 for MS1 and 1E5 for MS2, maximum injection time was adjusted automatically both for MS1 and MS2. The resolution was 120K for both MS1 and MS2. The theoretical m/z of each peptide included in the samples—from 1^- to 9^-—were included in the inclusion list, and precursors for fragmentation were selected for fragmentation with a mass tolerance of 20 ppm. Each precursor was fragmented by HCD whose normalized collision energy (NCE) was varied from 10 to 50, with a step length of 5. MS1 spectra were profile while MS2 spectra were centroid.

2.3 Data analysis

In-house python scripts were used to analyze all the MS data. Theoretical *m/z* values of precursors, internal ions caused by twice fragmentation of precursors, a-, b-, c-, x-, y-, z- ions of various charge states, as well as the *m/z* values of neutral loss peaks of all those ions, were calculated and compared with the observed peaks on MS2 spectra. Abundant neutral losses were taken into consideration, including loss of H_2O, NH_3, or CO_2 loss, which are not specific for a particular amino acid, and loss of characteristic amino acid side chains. In total, 66 amino acid side chain neutral losses (NL) were taken into consideration, 65 were reported by Bowie et al. (*11*), Liang et al. (*18*), Rumachik et al. (*19*), and Sun et al.(*26*), and the other one was found in this study (58.005479, $C_2H_2O_2$ from E) (Supplemental Table S2-3). The reported NLs from oxidized methionine (*18*) were ignored. Each observed peak in all MS2 spectra was matched to all the theoretical ions calculated based on peptide sequence with fragment mass tolerance of 10 ppm. The statistics results of peptide-spectrum matches (PSM) were filtered by the relative intensity (intensity of peak/base peak intensity) of peaks higher than 0.01 if not mentioned otherwise.

3. Results

We started out by investigating how an acidic C-terminus (D/E) generated by Glu-C digestion affects MS analysis of deprotonated peptides compared to a basic C-terminus (lysine/arginine, or K/R) generated by trypsin digestion. It has been reported that fragmentation of Glu-C peptides seems to be more extensive than tryptic peptides in NETD analysis (*6*). We thus wondered whether Glu-C digestion might be

advantageous over trypsin digestion for negative HCD analysis of acidic proteins. A total of 12 synthetic peptides in 3 groups (Table 1) were analyzed. Within each group, the four peptides differ only in the C-terminal residue, which is either D, E, K, or R. The length of these peptides varied from 6 to 16 residues, and the 20 common amino acids found in proteins were all included in these peptides.

An equimolar mixture of the four peptides of the same group was sprayed directly into an Orbitrap Fusion Lumos mass spectrometer operated under the negative ion mode. MS1 and HCD MS2 spectra were acquired. For HCD, the optimal normalized collision energy (NCE) of 35 was used (*23*). Fig. 1 shows a summary of the data. Group 1 (AAAAAX) and group 2 (IAVNYC*MFGX) peptides, in which X is D, E, K, or R, generated predominantly 1^- precursor ions and no 3^- precursors. The group 3 peptides (GYWGTNLGQPHSLATX) generated predominantly 2^- precursors and small amounts of 1^- and 3^- precursors (Fig. 1A, B, C, and E). Within each group, the peptide ending in D had the highest total intensity in MS1 whereas the peptide ending in R had the lowest (Fig. 1D). Intriguingly, for peptides sharing the same sequence except the C-terminal E or K, there was no obvious difference in total MS1 intensity. In other words, replacing E with K hardly affected the summed intensity of the precursor ions, but it did decrease the intensity of the highest charge state species (2^- for group 1 and group 2, 3^- for group 3) (Fig. 1A, B, C, and E).

The negative-ion HCD spectra of the 12 peptides displayed widespread NL of CO_2, regardless of the identity of the C-terminal aa (Fig. 2). However, the peptides with a C-terminal D did have more prominent CO_2 loss from precursor ions and fragment ions (Fig. 2A and E). The peptides with a C-terminal K produced the least amount of CO_2 loss (Fig. 2E-F). Comparing the HCD spectra of 1^-, 2^-, and 3^- precursors of the same peptide, we found that NL of CO_2 seemed to be largely independent of the precursor charge state (Fig. 3A-C). Statistical analysis of the HCD spectra of all 1^-, 2^-, and 3^- precursors arrived at the same conclusion, with the summed intensity and the occurrence frequency of the CO_2 loss peaks accounting for ~10% of the total sequencing ions (Fig. 3D, supplemental Fig. S1). Here, the sequencing ions refer to all fragment ions produced after the peptide backbone is cleaved once at any position (i.e., *a-, b-, c-, x-, y-,* and *z-* ions), with or without NLs. This analysis also confirmed that the relative intensity of sequencing ions of 3^- precursors is markedly lower than that of 1^- or 2^- precursors (Fig. 3D-E) (*23*).

The pervasiveness of CO_2 loss was to some extent beyond our expectation. NL of CO_2 from deprotonated peptides upon CID had been reported (*5, 7, 16, 18, 27*), but it was ranked at the 10th (from E) or 14th (from D) position by relative abundance (*18*). In our previous analysis (*23*), only the eight most abundant NLs from aa side chains (*18*), as reported were taken into consideration. Prompted by the unexpectedly extensive loss of CO_2 in the HCD spectra of deprotonated peptides (Fig. 1-3), we repeated the statistical analysis of NL peaks by considering all the previously reported NL species (*11, 18, 19, 26*) on an extended dataset of 115 synthetic peptides of 5-28 aa (Supplemental Table S1). The negative-ion HCD spectra were all acquired at the

optimal NCE of 35, and 22 of the 33 peptides analyzed in our previous study were included (*23*). Table 2 shows a summary of the analysis result. The most abundant NL species include a loss of NH_3 from asparagine (N), a loss of C_2H_4O from threonine (T), and a loss of H_2O or CO_2 from D and E. We also found a previously unnoticed NL of 58.005479 ($C_2H_2O_2$) from E (Table 2 and Fig. 2B).

The relative abundance of the *a-, b-, c-, x-, y-,* and *z-* ions with or without various types of NLs is updated (Fig. 4). As loss of CO_2 does not require the presence of D or E in a peptide (Fig. 2), it is regarded as a common NL, although it is most strongly associated with D and E. The updated result shows again that the majority of the fragment ions have NLs of one form or another, including any of the common NLs (H_2O, NH_3, CO_2) and aa-specific side-chain losses. The four most abundant fragment ion types are *y-* > *x-, z-* > *c-* if the NL ions are included, and *c-* > *y-* > *z-* > *b-* if not (Fig. 4).

The dataset of 115 synthetic peptides allowed a charge-state distribution analysis. As shown in Fig. 5, longer peptides and D/E-rich peptides are more likely to carry multiple negative charges. However, multiply charged peptide anions have a lower fraction of sequencing ions in HCD spectra (Fig. 3E), which does not help for peptide identification.

Using this expanded dataset, we analyzed the positional effect of D/E on HCD MS2 of deprotonated peptides. As shown in Fig. 6, peptides containing only C-terminal D/E produced more and higher-intensity CO_2-loss peaks than D/E-free peptides or peptides containing non-C-terminal D/E (Fig. 6). We further separated the peptides containing only C-terminal D/E into two subgroups by their C-termini, and compared between the subgroups (Supplemental Fig. S2). We did not see higher-intensity CO_2-loss peaks from the peptides ending in D than from those ending in E (Supplemental Fig. S2), as we observed from the small but well-controled 12-peptide data (Fig. 2E-F). Nevertheless, peptides with C-terminal D/E do have extensive CO_2-loss peaks. Therefore, Glu-C peptides have a higher tendency to undergo CO_2 loss than tryptic peptides.

Consistent with the finding based on 12 peptides (Fig. 3D-E), the 115-peptide dataset showed that NL of CO_2 in negative-ion HCD was largely independent of the charge state of precursor ions (Fig. 7A-B). The percentage of CO_2 loss peaks was around 10% either by relative intensity or by occurrence frequency. However, the percentage of sequencing ions, including the ones with NLs, was higher in the HCD spectra of 1^- and 2^- precursors than in the spectra of 3^- and 4^- precursors (Fig. 7C-D). This suggests that peptide sequencing by negative-ion-mode MS should focus mainly on lower-charge state peptide ions. As long peptides and D/E-rich peptides tend to produce higher charge state anions (Fig. 5), relatively short peptides containing no more than a couple of D/E residues are likely suitable for MS sequencing in the negative ion mode.

With this knowledge, we compared the in silico digestion results of an acidic protein pepsin by Glu-C and trypsin. Pepsin is a 385-aa protein with 45 D/E (11.7%) and 15 K/R (3.9%) residues (Supplemental Fig. S3A-B). BSA, a 607-aa protein with 99 D/E (16.3%) and 86 K/R (14.2%), was included as a reference (Supplemental Fig. S3C-D). Allowing no more than two missed cleavage sites, in silico digestion of BSA by Glu-C and trypsin generated 196 and 137 peptides, respectively, of less than 20 aa and no more than 3 D/E residues, which may have a good chance to be detected by negative-ion HCD (Fig. 8A-B). In contrast, the same in silico digestion of pepsin by trypsin produced only 25 theoretical peptides potentially suitable for negative-ion HCD, whereas that by Glu-C produced 69 (Fig. 8C-D). This suggests that Glu-C digestion is worth trying for negative-ion mode MS analysis of acidic proteins.

4. Discussion

Many types of NLs disperse the signal of a fragment ion into different channels and thus decrease the intensity of fragment ions. This reduces the sensitivity of peptide identification in negative-ion-mode MS. How to suppress NL is worth investigating in the future. Methylation of side-chain (D/E) and the C-terminal -COOH is one possibility to reduce CO_2 NL.

Although Glu-C can cleave the peptide bond C-terminal to Asp (*6, 28, 29*), the efficiency of this cleavage is lower than its namesake cleavage—cutting after a Glu residue. Another protease Asp-N cleaves the peptide bond N-terminal to D. The frequency of D, E, K, and R in the 567483 protein sequences of the UniProtKB database is 5.46%, 6.72%, 5.80%, and 5.53%, respectively (UniProtKB/Swiss-Prot protein knowledgebase release 2022_02 statistics, https://web.expasy.org/docs/relnotes/relstat.html). So, the number of potential cleavage sites of Asp-N and Glu-C combined is similar to that of trypsin. However, the number of D/E residues will vary greatly in tryptic peptides. This will not happen in peptides generated by Asp-N and Glu-C digestion. Peptide anions containing multiple D/E residues tend to adopt higher charge states, and higher charge state precursors (e.g. 3^- or above) produce poorer MS2 spectra with lower-intensity sequencing ions (Fig. 7). We thus predict that protein digestion with Asp-N and Glu-C together is likely a good way to prepare samples for negative-ion-mode MS analysis.

The statistical analysis in this study did not look into locally enhanced backbone cleavage by specific aa side chains, such as the cleavage at the N-Cα bond of XXX-Asp/Asn or XXX-Ser/Thr, which has been reported (*7, 11, 22, 25*). Certain amino acid residues including Asp, Glu, and Cys have been shown to influence locally the backbone cleavage of deprotonated peptides (*8, 12, 17, 22, 30, 31*). This waits further analysis.

Previously, we marveled at the observation that the highly basic peptides containing multiple lysine residues, zero acidic residues, and amidated C-termini

ionized well in the negative ion mode and carried up to three negative charges (23). One example is a 16-aa peptide GLFAVIKKVAKVIKKL-NH2, which was first studied by Dr. Bowie (11-13). In this study, our side-by-side comparison of the peptides differing only in the C-terminal residue (Table 1) showed that replacing the C-terminal E with K did not decrease the total intensity of the deprotonated peptides (Fig. 1D), suggesting that lysine has little negative effect on peptide ionization in the negative ion mode. A K-rich and D/E-free peptide, if sufficiently long, is able to carry multiple negative charges (Fig. 5). Further, their 1^- and 2^- precursors can generate beautiful MS2 spectra with abundant sequencing ions (Fig. 7). This is most noticeable if they are devoid of other NL-prone aa residue besides D and E, for example, C, N, Q, S, T, and R (Table 2). The peptide GLFAVIKKVAKVIKKL-NH2 happens to have all the desirable attributes.

Table 1. Twelve synthetic peptides to compare the effects of C-terminal residues

	Sequence	Extinction coefficients	Concentration (μM)	Relative concentration
Group 1	AAAAAD	28.47	6.48	1.00
	AAAAAE	27.68	6.40	0.99
	AAAAAK	27.73	6.71	1.04
	AAAAAR	28.81	7.16	1.11
Group 2	IAVNYC*MFGD	37.67	5.50	1.00
	IAVNYC*MFGE	37.21	6.17	1.12
	IAVNYC*MFGK	37.24	5.71	1.04
	IAVNYC*MFGR	37.50	5.67	1.03
Group 3	GYWGTNLGQPHSLATD	43.23	16.88	1.00
	GYWGTNLGQPHSLATE	42.88	16.54	0.98
	GYWGTNLGQPHSLATK	42.91	17.38	1.03
	GYWGTNLGQPHSLATR	42.99	16.31	0.97

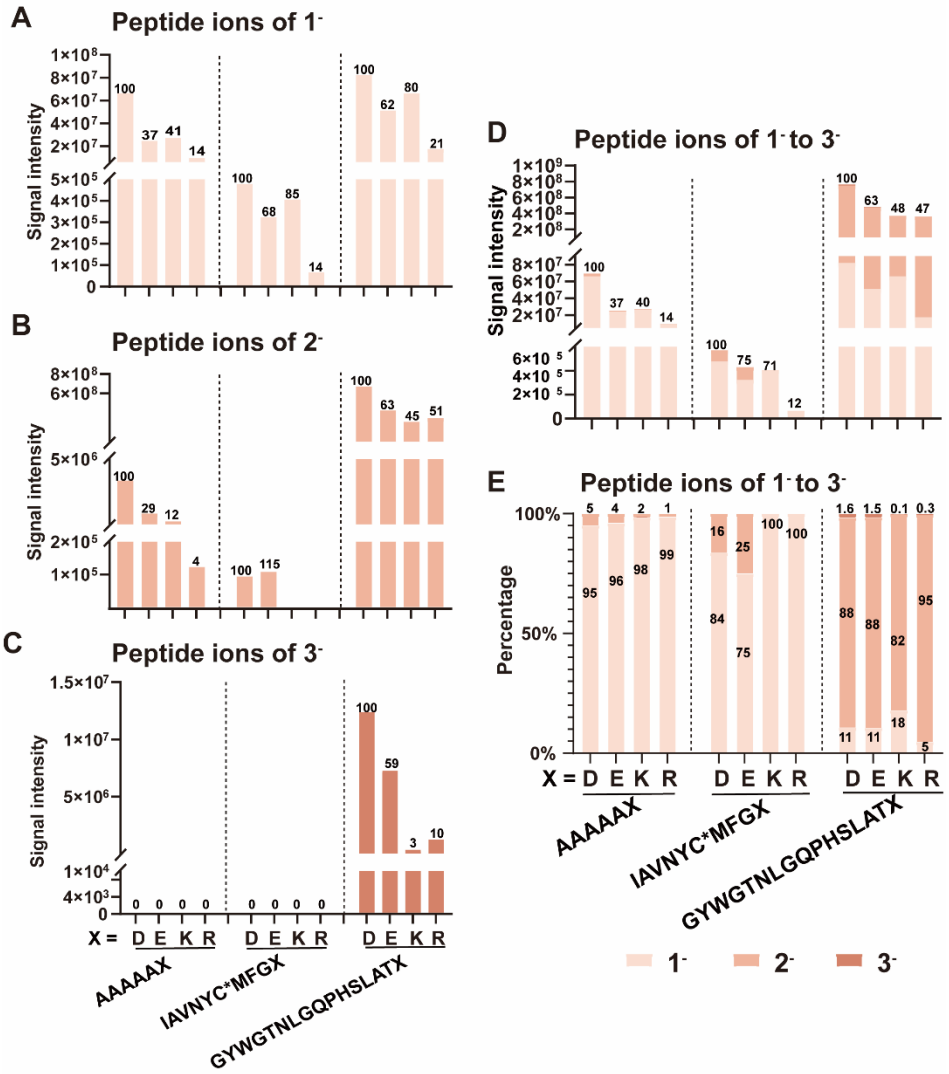

Fig. 1. MS1 intensity of 12 synthetic peptides normalized against the relative concentration shown in Table 1. MS1 intensity of 1^- (A), 2^- (B), 3^- (C), and 1^- to 3^- (D) precursors. (E) Relative MS1 intensities of 1^-, 2^-, and 3^- precursors. The numbers indicate the relative intensity in percentage. *Carbamidomethylated cysteine.

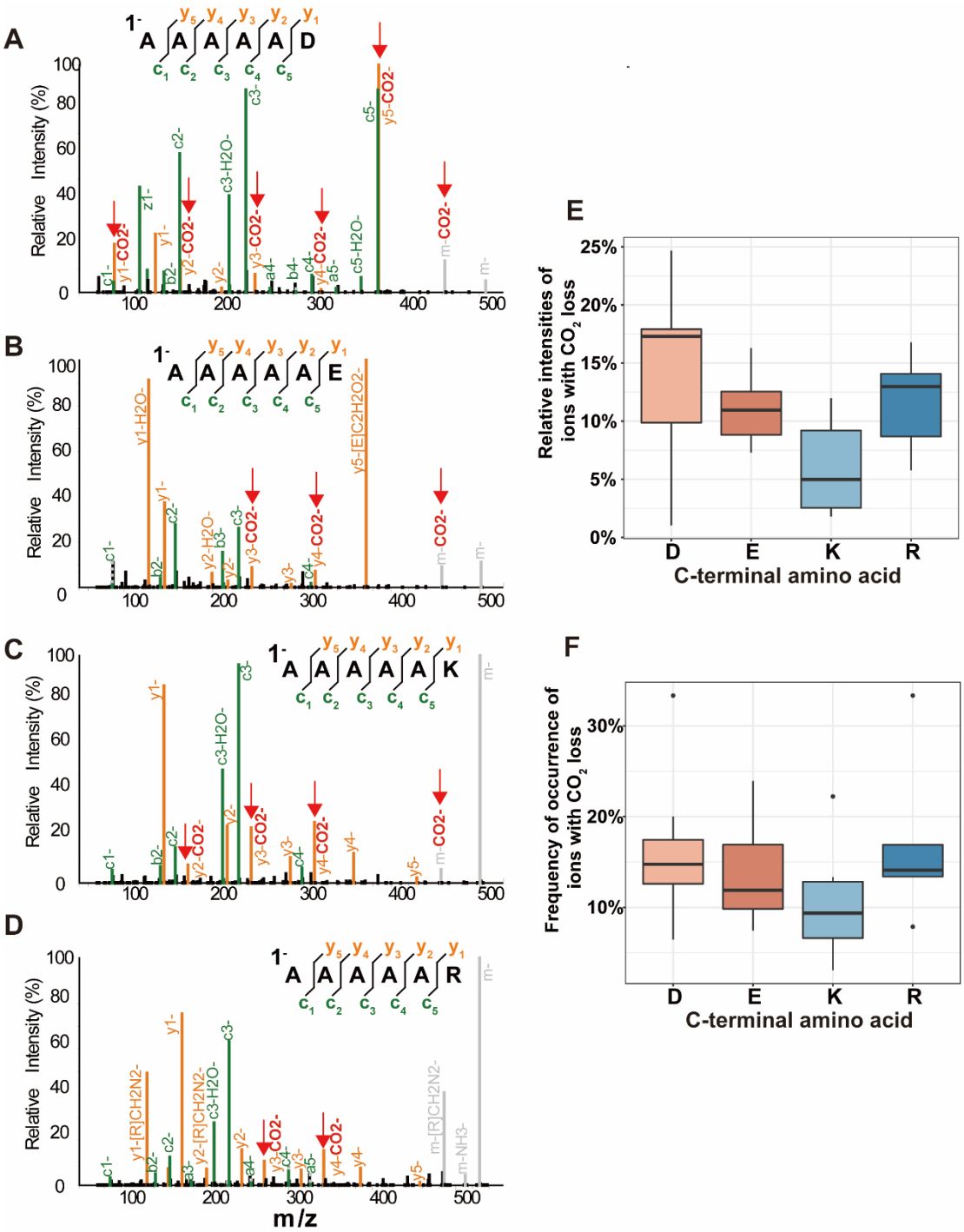

Fig. 2. Frequent CO_2 loss in HCD spectra of deprotonated peptides. HCD spectra of AAAAAD (A), AAAAAE (B), AAAAAK (C), and AAAAAR (D) at NCE 35. Ions with CO_2 loss are highlighted and indicated by red arrows. The relative intensity (E) and frequency of occurrence (F) of ions with CO_2 loss for peptides ending in D, E, K, or R (Three peptides in each category, all charge states combined).

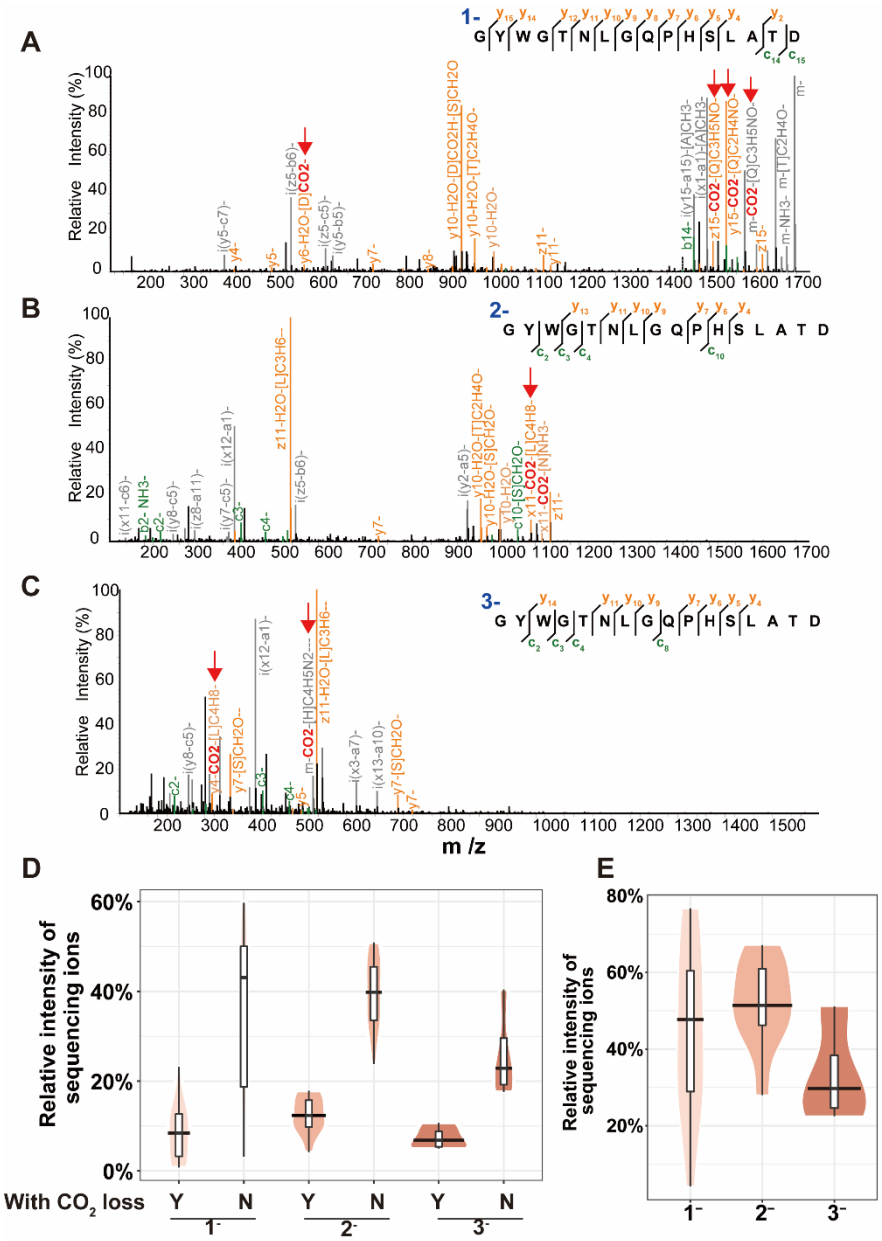

Fig. 3. CO_2 loss in HCD spectra of 1^-, 2^-, and 3^- peptide precursors and the relative intensity of sequencing ions. HCD spectra of GYWGTNLGQPHSLATD from a 1^- (A), 2^- (B), or 3^- (C) precursor. (D) Relative intensity of sequencing ions with or without CO_2 loss. (E) All sequencing ions, based on the HCD spectra of 12 peptides, all charge states included.

Table 2. Side-chain neutral losses observed from 115 peptides

Aa type	Aa	NL[a]	Accurate mass (Da)	Composition	Found in # spectra	# spectra of peptides containing aa in question	Frequency	Average # NL peaks/spectrum
Nonpolar	A	A_15	15.02347510	CH_3	51	176	0.29	12.45
	F	F_77	77.03912516	C_6H_5	17	53	0.32	8.29
		F_91	91.05477522	C_7H_7	25	53	0.47	6.64
		F_92	92.06260026	C_7H_8	19	53	0.36	5.47
	I	I_28	28.03130013	C_2H_4	28	89	0.31	4.43
		I_29	29.03912516	C_2H_5	23	89	0.26	4.87
		I_42	42.04695019	C_3H_6	32	89	0.36	3.66
		I_56	56.06260026	C_4H_8	25	89	0.28	2.96
		I_57	57.07042529	C_4H_9	18	89	0.20	3.78
	L	L_28	28.03130013	C_2H_4	44	194	0.23	2.86
		L_42	42.04695019	C_3H_6	51	194	0.26	3.25
		L_43	43.05477522	C_3H_7	45	194	0.23	3.02
		L_56	56.06260026	C_4H_8	37	194	0.19	2.49
		L_57	57.07042529	C_4H_9	32	194	0.16	1.91
	M	M_46	45.98772075	CH_2S	15	46	0.33	2.00
		M_47	46.99554579	CH_3S	20	46	0.43	3.30
		M_48	48.00337082	CH_4S	11	46	0.24	2.27
		M_61	61.01119585	C_2H_5S	27	46	0.59	4.11
		M_62	62.01902088	C_2H_6S	13	46	0.28	2.00
		M_74	74.01902088	C_3H_6S	9	46	0.20	3.00
		M_75	75.02684591	C_3H_7S	18	46	0.39	2.67
	P	P_28	28.03130013	C_2H_4	13	69	0.19	1.15
		P_42	42.04695019	C_3H_6	17	69	0.25	1.53
	V	V_15	15.02347510	CH_3	36	131	0.27	3.19
		V_28	28.03130013	C_2H_4	26	131	0.20	2.00
		V_42	42.04695019	C_3H_6	22	131	0.17	1.95
		V_43	43.05477522	C_3H_7	22	131	0.17	1.77
	W	W_129	129.0578492	C_9H_7N	13	41	0.32	6.15
		W_130	130.0656743	C_9H_8N	4	41	0.10	1.25
Uncharged polar	C	C_33	32.97989572	HS	9	15	0.60	6.33
		C_34	33.98772075	H_2S	10	15	0.67	14.60
		C_46	45.98772075	CH_2S	6	15	0.40	1.83
		C*_58	58.02928876	C_2H_4NO	0	5	0.00	0.00
		C*_90	90.00135945	C_2H_4NOS	4	5	0.80	3.75
	N	N_17	17.02654910	NH_3	53	57	0.93	13.43
		N_44	44.01363869	CH_2NO	28	57	0.49	4.39
	Q	Q_17	17.02654910	NH_3	60	87	0.69	6.33
		Q_43	43.00581366	CHNO	62	87	0.71	3.02
		Q_58	58.02928876	C_2H_4NO	28	87	0.32	2.57
		Q_71	71.03711379	C_3H_5NO	63	87	0.72	6.10
	S	S_30	30.01056469	CH_2O	51	72	0.71	12.76
	T	T_44	44.02621475	C_2H_4O	124	140	0.89	8.77
	Y	Y_106	106.0418648	C_7H_6O	14	45	0.31	3.64
		Y_107	107.0496898	C_7H_7O	9	45	0.20	3.78
Charged polar	D	D_18	18.01056469	H_2O	160	175	0.91	9.56
		D_44	43.98982924	CO_2	141	175	0.81	6.14
		D_45	44.99765428	CO_2H	57	175	0.33	2.11
	E	E_18	18.01056469	H_2O	151	153	0.99	10.08
		E_44	43.98982924	CO_2	139	153	0.91	4.74
		E_45	44.99765428	CO_2H	31	153	0.20	1.29
		E_58	58.00547931	$C_2H_2O_2$	77	153	0.50	1.86
		E_59	59.01330434	$C_2H_3O_2$	51	153	0.33	3.76
		E_72	72.02112937	$C_3H_4O_2$	123	153	0.80	4.54
	H	H_67	67.02962311	$C_3H_3N_2$	16	80	0.20	2.69
		H_80	80.03744814	$C_4H_4N_2$	25	80	0.31	1.52
		H_81	81.04527317	$C_4H_5N_2$	11	80	0.14	2.55
	K	K_58	58.06567426	C_3H_8N	34	158	0.22	1.76
		K_59	59.07349929	C_3H_9N	50	158	0.32	1.88
		K_71	71.07349929	C_4H_9N	52	158	0.33	2.77
		K_72	72.08132433	$C_4H_{10}N$	24	158	0.15	2.17
	R	R_100	100.0874723	$C_4H_{10}N_3$	39	108	0.36	2.59
		R_42	42.02179807	CH_2N_2	83	108	0.77	9.24
		R_72	72.05617221	$C_2H_6N_3$	39	108	0.36	3.74
		R_86	86.07182227	$C_3H_8N_3$	38	108	0.35	3.92
		R_87	87.07964730	$C_3H_9N_3$	46	108	0.43	3.72
		R_99	99.07964730	$C_4H_9N_3$	55	108	0.51	3.78

[a]NL: Neutral loss, amino acid_round number of NL accurate mass; C*: carbamidomethylated cysteine, the statistical results of NL from C* were only based on the four peptides in group 2 (IAVNYC*MFGD/E/K/R), and other 65 NLs are based on the 115 peptides.

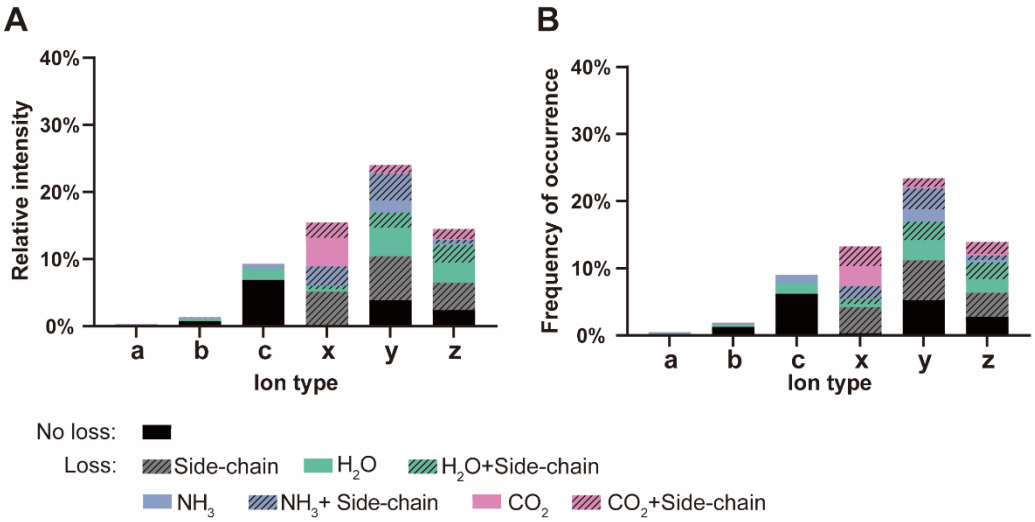

Fig. 4. Relative intensity (A) and frequency of occurrence based on 254 HCD spectra of 115 synthetic peptides. Note: y-H_2O and x-CO_2 have the same mass. As a result, and the peaks matched to the two ion types were counted twice.

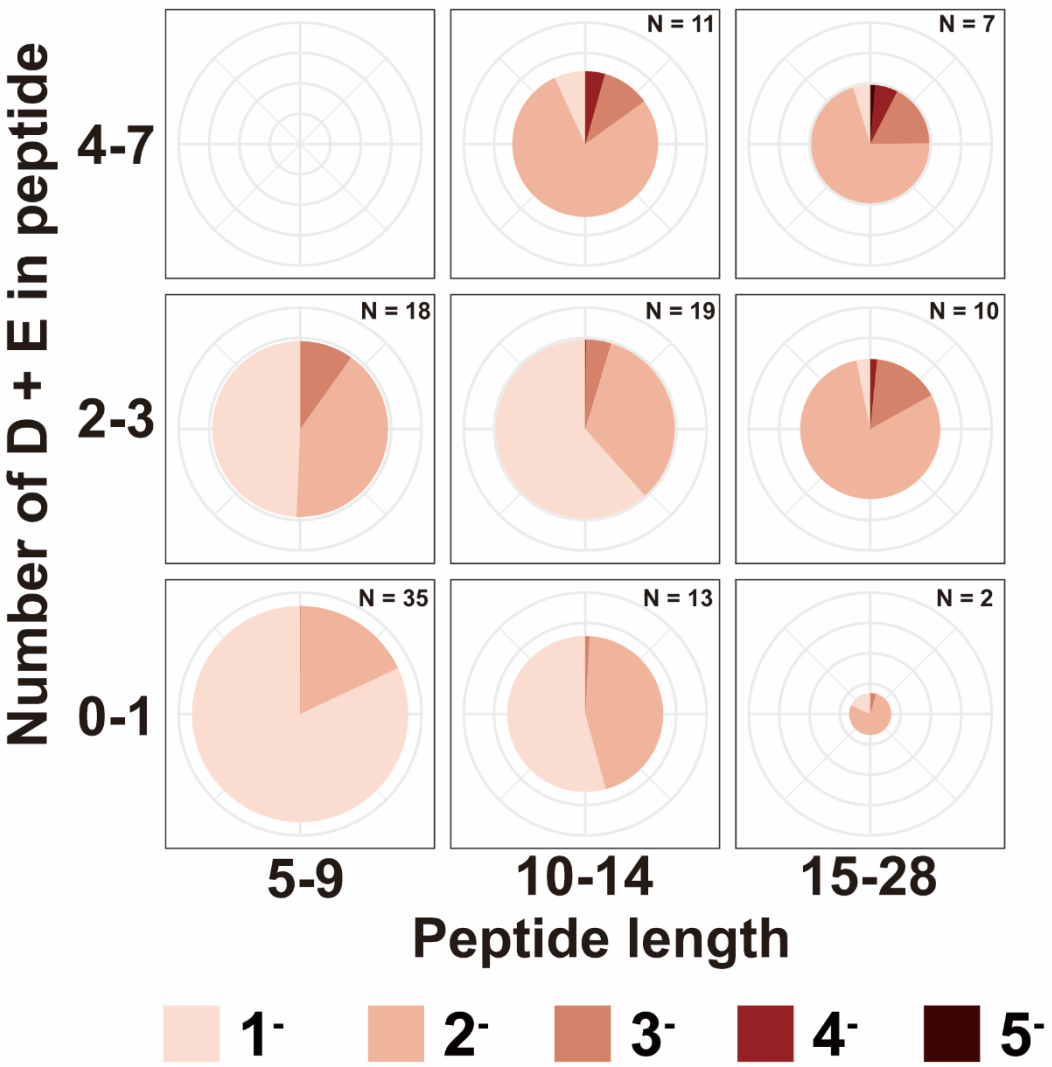

Fig. 5. The length and the number of acidic residues of a peptide influence the charge state of the peptide anion. The statistics is based on the MS1 spectra of 115 synthetic peptides. The mass tolerance is 20 ppm.

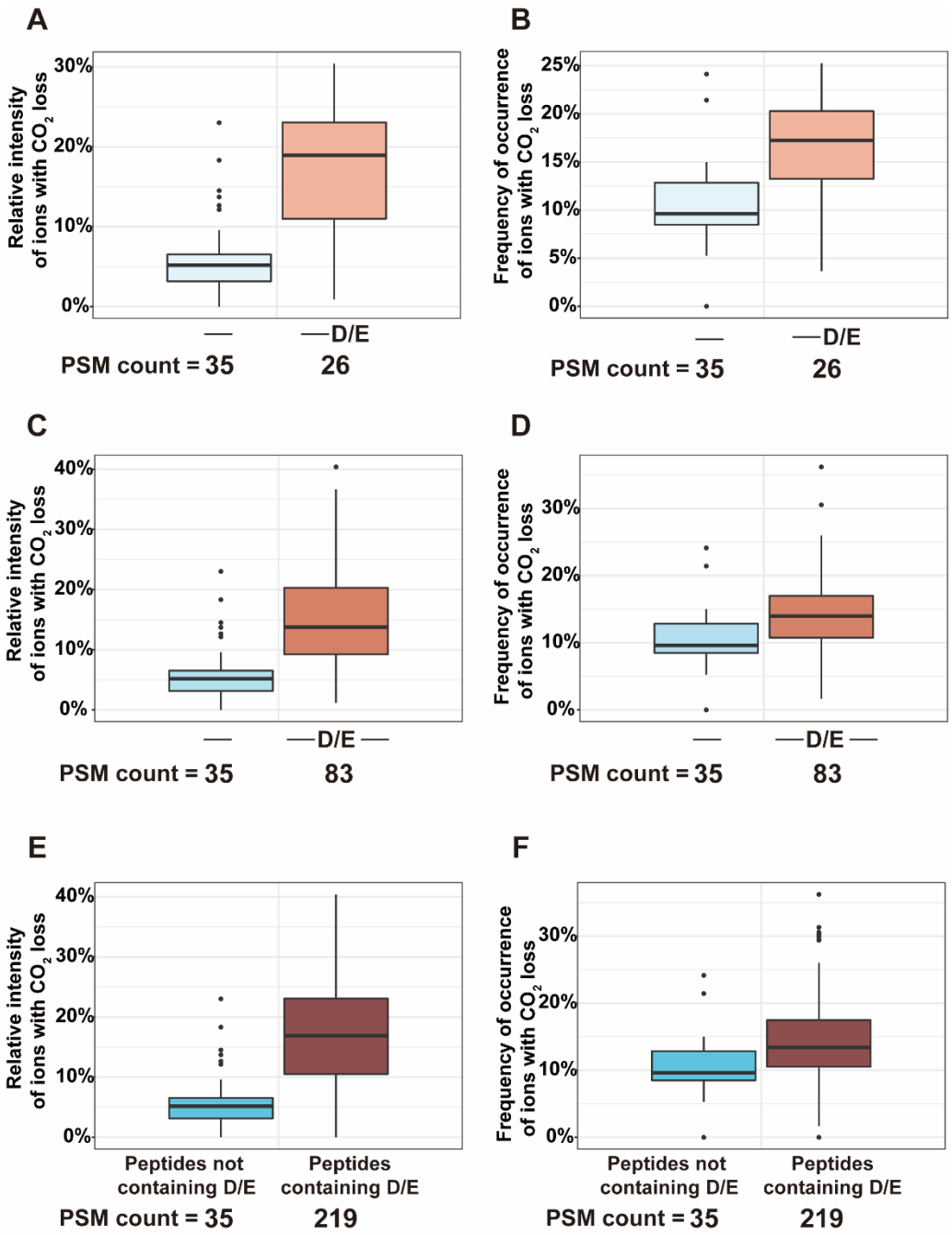

Fig. 6. Relative intensities (A, C, E) and frequency of occurrence (B, D, F) of ions with CO_2 loss. (A, B) peptides without D/E versus peptides with only C-terminal D/E. (C, D) peptides without D/E versus peptides with non-C-terminal D/E. (E, F) D/E free peptides versus D/E containing peptides. Shown is the analysis result of 254 HCD spectra of 115 peptides.

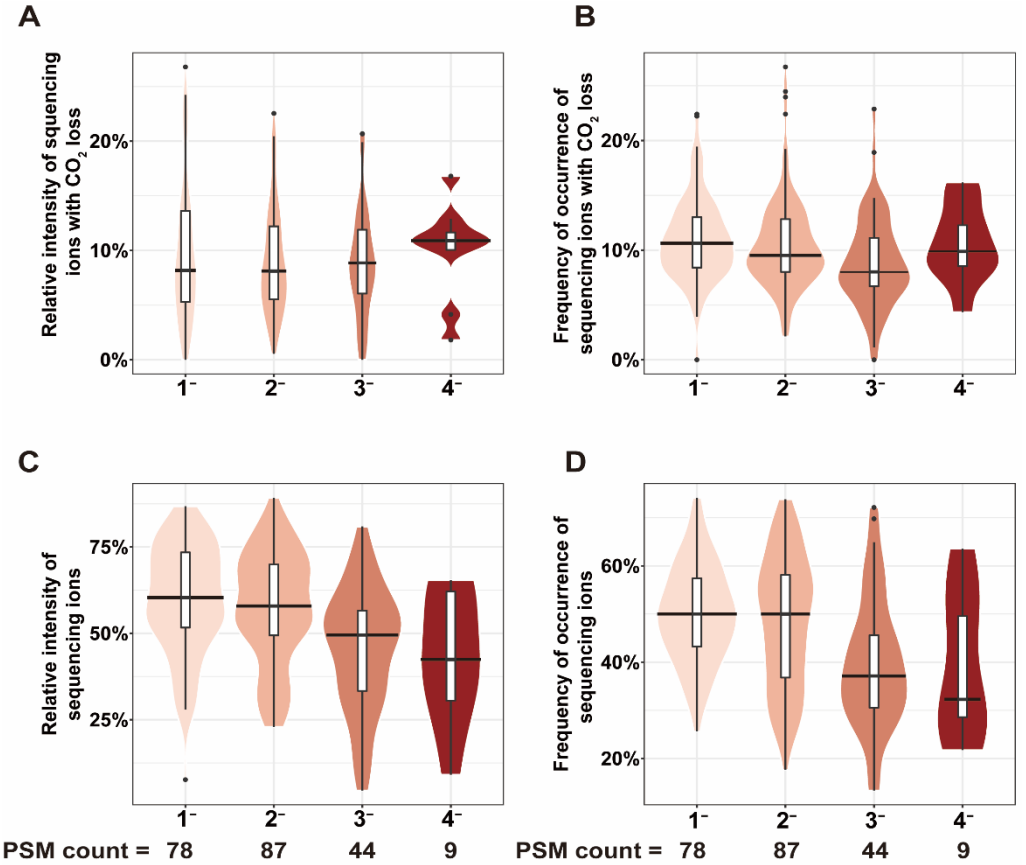

Fig. 7. Relative intensities (A, C) and frequency of occurrence (B, D) of sequencing ions with CO_2 loss (A, B) and all the sequencing ions (C, D). Analysis results of 254 HCD spectra belonging to 115 peptides are shown.

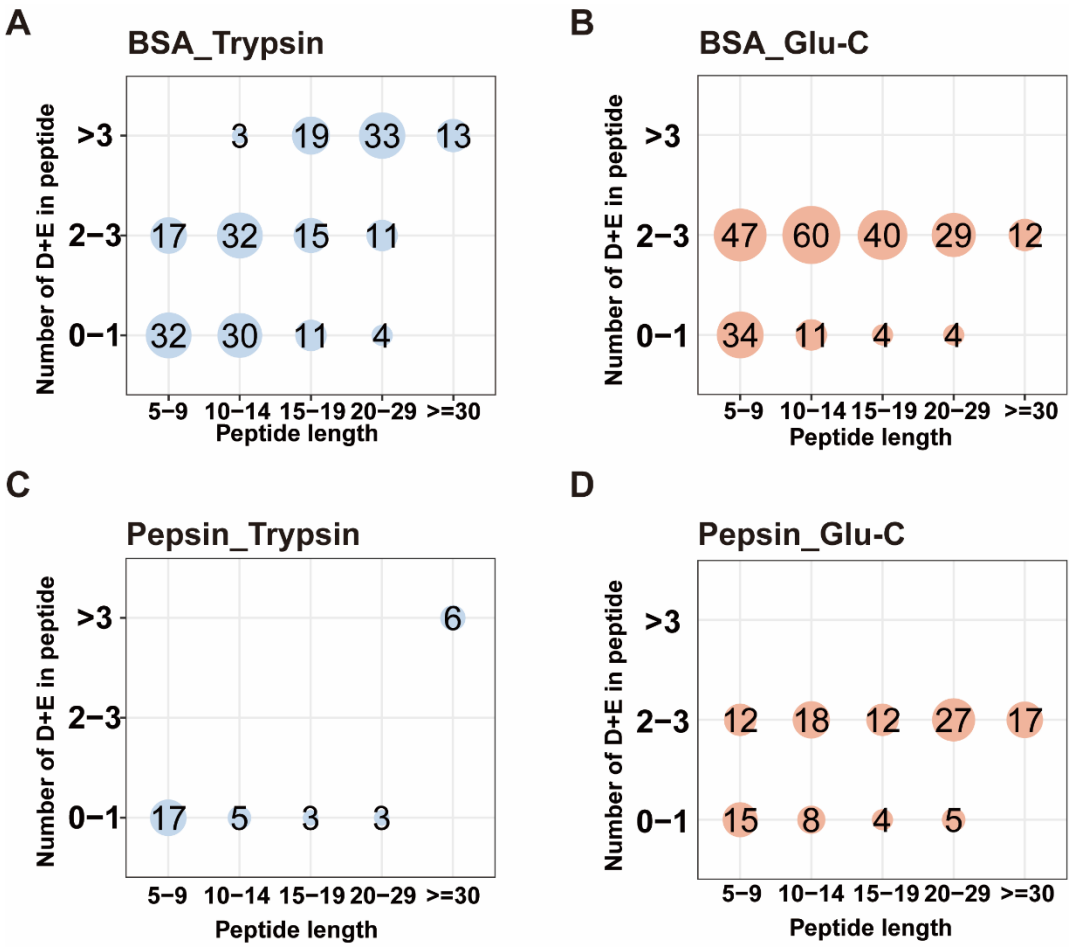

Fig. 8. Peptides from in silico digestion of BSA (A, B) and pepsin (C, D) with trypsin (A, C) or Glu-C (B, D). The number in each circle indicates the number of peptides of the indicated length and containing the indicated number of acidic residues.

Acknowledgments

The authors would like to thank Dr. Ping Xu of the National Center for Protein Sciences (Beijing, China) for providing 18 synthetic peptides, and Peng-Zhi Mao of the Institute of Computing Technology, Chinese Academy of Sciences for modifying pLable software to display peptide anion spectra. This study is supported by National Key Research and Development Program of China (Grant 2020YFF01014505), the Ministry of Science and Technology of China, and Beijing Municipal Science and Technology Commission.

References

1. G. S. Omenn *et al.*, Progress Identifying and Analyzing the Human Proteome: 2021 Metrics from the HUPO Human Proteome Project. *J Proteome Res* **20**, 5227-5240 (2021).
2. Y. Kawashima *et al.*, Single-Shot 10K Proteome Approach: Over 10,000 Protein Identifications by Data-Independent Acquisition-Based Single-Shot Proteomics with Ion Mobility Spectrometry. *J Proteome Res* **21**, 1418-1427 (2022).
3. G. C. McAlister *et al.*, Analysis of the acidic proteome with negative electron-transfer dissociation mass spectrometry. *Anal Chem* **84**, 2875-2882 (2012).
4. P. Mallick *et al.*, Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol* **25**, 125-131 (2007).
5. N. P. Ewing, C. J. Cassady, Dissociation of multiply charged negative ions for hirudin (54–65), fibrinopeptide B, and insulin A (oxidized). *Journal of the American Society for Mass Spectrometry* **12**, 105-116 (2001).
6. N. M. Riley *et al.*, The Negative Mode Proteome with Activated Ion Negative Electron Transfer Dissociation (AI-NETD). *Mol Cell Proteomics* **14**, 2644-2660 (2015).
7. R. J. Waugh, J. H. Bowie, A review of the collision‐induced dissociations of deprotonated dipeptides and tripeptides. An aid to structure determination. *Rapid communications in mass spectrometry* **8**, 169-173 (1994).
8. A. M. Bradford, R. J. Waugh, J. H. Bowie, Characterization of underivatized tetrapeptides by negative‐ion fast‐atom bombardment mass spectrometry. *Rapid communications in mass spectrometry* **9**, 677-685 (1995).
9. M. Eckersley, J. H. Bowie, R. N. Hayes, Collision‐induced dissociations of deprotonated peptides, dipeptides and tripeptides with hydrogen and alkyl α groups. An aid to structure determination. *Organic mass spectrometry* **24**, 597-602 (1989).
10. S. T. Steinborner, J. H. Bowie, A Comparison of the Positive‐and Negative‐ion Mass Spectra of Bio‐active Peptides from the Dorsal Secretion of the Australian Red Tree Frog, Litoria rubella. *Rapid communications in mass spectrometry* **10**, 1243-1247 (1996).
11. J. H. Bowie, C. S. Brinkworth, S. Dua, Collision‐induced fragmentations of the (M‐H)− parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages. *Mass Spectrometry Reviews* **21**, 87-107 (2002).
12. D. Bilusich, J. H. Bowie, Fragmentations of (M–H)− anions of underivatised peptides. Part 2: Characteristic cleavages of Ser and Cys and of disulfides and other post‐translational modifications, together with some unusual internal processes. *Mass spectrometry reviews* **28**, 20-34 (2009).
13. T. Wang *et al.*, Negative ion cleavages of (M–H)− anions of peptides. Part 3. Post‐translational modifications. *Mass Spectrometry Reviews* **37**, 3-21 (2018).
14. S. M. Greer, J. R. Cannon, J. S. Brodbelt, Improvement of shotgun proteomics in the negative mode by carbamylation of peptides and ultraviolet photodissociation mass spectrometry. *Anal Chem* **86**, 12285-12290 (2014).
15. J. A. Madsen, T. S. Kaoud, K. N. Dalby, J. S. Brodbelt, 193-nm photodissociation of singly and multiply charged peptide anions for acidic proteome characterization. *Proteomics* **11**, 1329-1334 (2011).
16. J. A. Madsen *et al.*, High-throughput database search and large-scale negative polarity liquid

chromatography-tandem mass spectrometry with ultraviolet photodissociation for complex proteomic samples. *Mol Cell Proteomics* **12**, 2604-2614 (2013).

17. N. Sugasawa *et al.*, Formation of c- and z-ions due to preferential cleavage at the NC bond of Xxx-Asp/Asn residues in negative-ion CID of peptides. *International Journal of Mass Spectrometry* **383-384**, 38-43 (2015).
18. Y. Liang, P. Neta, X. Yang, S. E. Stein, Collision-induced dissociation of deprotonated peptides. Relative abundance of side-chain neutral losses, residue-specific product ions, and comparison with protonated peptides. *Journal of The American Society for Mass Spectrometry* **29**, 463-469 (2017).
19. N. G. Rumachik *et al.*, Characterizing peptide neutral losses induced by negative electron-transfer dissociation (NETD). *Journal of the American Society for Mass Spectrometry* **23**, 718-727 (2012).
20. A. G. Harrison, Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains. *Journal of the American Society for Mass Spectrometry* **12**, 1-13 (2001).
21. A. G. Harrison, A. B. Young, Fragmentation reactions of deprotonated peptides containing proline. The proline effect. *J Mass Spectrom* **40**, 1173-1186 (2005).
22. A. G. Harrison, A. B. Young, Fragmentation reactions of deprotonated peptides containing aspartic acid. *International Journal of Mass Spectrometry* **255-256**, 111-122 (2006).
23. M.-Q. Zuo, R.-X. Sun, R.-Q. Fang, S.-M. He, M.-Q. Dong, Characterization of collision-induced dissociation of deprotonated peptides of 4–16 amino acids using high-resolution mass spectrometry. *International Journal of Mass Spectrometry* **445**, (2019).
24. N. J. Anthis, G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci* **22**, 851-858 (2013).
25. B. Yang *et al.*, Identification of cross-linked peptides from complex samples. *Nat Methods* **9**, 904-906 (2012).
26. Q. Sun, H. Nelson, T. Ly, B. M. Stoltz, R. R. Julian, Side chain chemistry mediates backbone fragmentation in hydrogen deficient peptide radicals. *J Proteome Res* **8**, 958-966 (2009).
27. S. Guan, J. M. Rabus, P. Maitre, B. J. Bythell, Gas-Phase Dissociation Chemistry of Deprotonated RGD. *J Am Soc Mass Spectrom* **32**, 55-63 (2021).
28. P. Giansanti, L. Tsiatsiani, T. Y. Low, A. J. Heck, Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat Protoc* **11**, 993-1006 (2016).
29. N. M. Riley, J. J. Coon, Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. *Anal Chem* **88**, 74-94 (2016).
30. J. H. Bowie, C. S. Brinkworth, S. Dua, Collision-induced fragmentations of the (M-H)- parent anions of underivatized peptides: an aid to structure determination and some unusual negative ion cleavages. *Mass Spectrom Rev* **21**, 87-107 (2002).
31. A. Kagoshima, K. Sekimoto, M. Takayama, Intramolecular Hydrogen Transfer from the Alpha-Carbon (Calpha) and Backbone Amide Nitrogen (Nb) to Form c- and y-Ions in Negative-Ion CID of Peptides. *J Am Soc Mass Spectrom* **30**, 1592-1600 (2019).