Mass Spectrometry
DOI: 10.5702/massspectrometry.A0051

Vol. 5 (2016), A0051

Original Article

N-Terminal Derivatization with Structures Having High Proton Affinity for Discrimination between Leu and Ile Residues in Peptides by High-Energy Collision-Induced Dissociation

Atsushi Kitanaka[1], Masahiro Miyashita*[,1], Ayumi Kubo[2], Takaya Satoh[2], Michisato Toyoda[3], and Hisashi Miyagawa[1]

[1] *Graduate School of Agriculture, Kyoto University, Kyoto 606–8502, Japan*
[2] *JEOL Ltd., Tokyo 196–8558, Japan*
[3] *Graduate School of Science, Osaka University, Osaka 560–0043, Japan*

De novo sequencing is still essential in the identification of peptides and proteins from unexplored organisms whose sequence information is not available. One of the remaining problems in *de novo* sequencing is discrimination between Leu and Ile residues. The discrimination is possible based on differences in side chain fragmentation between Leu and Ile under high-energy collision-induced dissociation (HE-CID) conditions. However, this is observed only when basic residues, such as Arg and Lys, are present near the N- or C-terminal end. It has been shown that the charge derivatization at the N-terminal end by a quarternary ammonium or phosphonium moiety facilitates the side chain fragmentation by HE-CID. However, the effective backbone fragmentation by low-energy CID (LE-CID) is often hampered in those derivatives with a fixed charge. Previously, we demonstrated that the N-terminal charge derivatization with the structures having high proton affinity induced the preferential formation of b-ions under LE-CID conditions, allowing straightforward interpretation of product ion spectra. In the present study, we further investigated whether the same derivatization approach is also effective for discrimination between Leu and Ile under HE-CID conditions. Consequently, the side chain fragmentation of Leu and Ile residues was most effectively enhanced by the N-terminal derivatization with 4-(guanidinomethyl)benzoic acid among the tested structures. This derivatization approach, which is compatible with both HE- and LE-CID analysis, offers a straightforward and unambiguous *de novo* peptide sequencing method.

Please cite this article as: Mass Spectrom (Tokyo) 2016; 5(1): A0051

Keywords: MALDI-TOF-TOF, ion trap-TOF, *de novo* sequencing, proton affinity

(Received August 31, 2016; Accepted October 5, 2016)

INTRODUCTION

As the number of primary sequences in public databases continue to grow due to the recent availability of extensive genomic information, protein identification can be readily achieved by comparing uninterpreted spectra obtained by tandem mass spectrometry with those constructed based on the amino acid sequences in the database.[1)] However, this methodology is only applicable to the components of major organisms whose genomic information is available. In addition, since many of the peptides or proteins in nature are modified post-translationally, their complete structural determination is difficult only by the database-dependent techniques. Thus, the identification of unknown peptides and proteins from unexplored organisms still strongly relies on *de novo* sequencing techniques, in which the amino acid sequences are deduced by interpretation of their product ion spectra.[2)]

De novo sequencing is based on the mass differences between a series of fragment ions of peptides generated by collision-induced dissociation (CID). One of its difficulties is discrimination between Leu and Ile residues in peptides.[3)] This discrimination is possible by analyzing the differences in dissociation patterns of their branched alkyl side chains to form d- and w-ions under high-energy CID (HE-CID)

* Correspondence to: Masahiro Miyashita, *Graduate School of Agriculture, Kyoto University, Kyoto 606–8502, Japan*, e-mail: miyamasa@kais.kyoto-u.ac.jp
Abbreviations: HE-CID, high-energy collision-induced dissociation; LE-CID, low-energy collision-induced dissociation; DMT-MM, 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride; Gmb, 4-guanidinomethylbenzoic acid; Aba, 4-amidinobenzoic acid; Iza, 4-imidazolecarboxylic acid; Nic, nicotinic acid

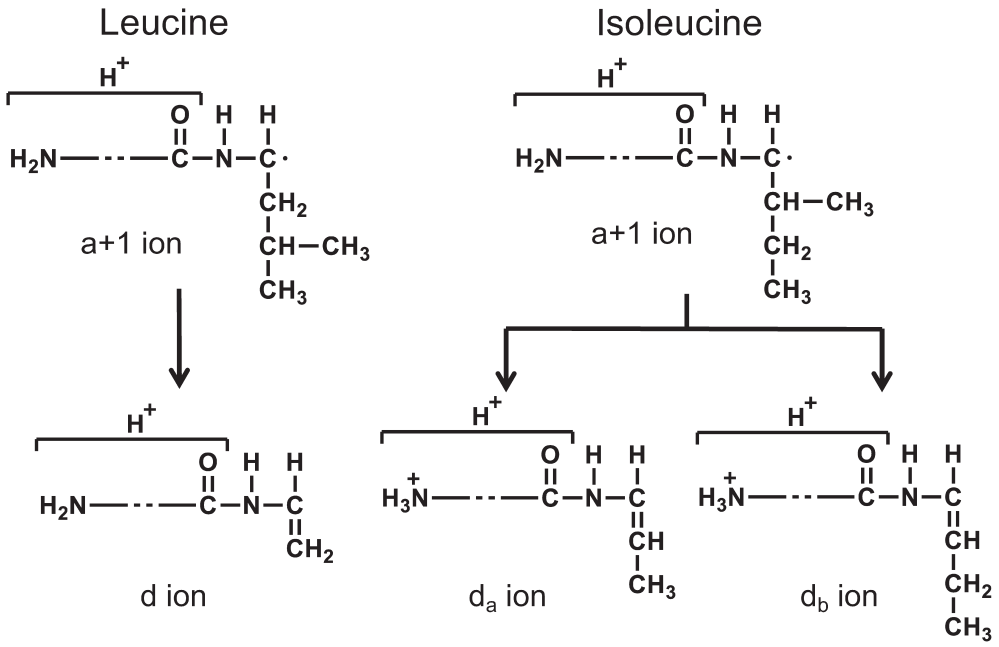


Fig. 1. Difference of side-chain fragmentation between Leu and Ile under HE-CID conditions.[9)]

conditions (Fig. 1). However, since this dissociation occurs through charge-remote fragmentation pathways, the existence of basic residues such as Arg and Lys are required near the N- or C-terminal end.[4)] For the peptides without basic residues, it has been shown that the derivatization of the N-terminal amino group of peptides with the positively-charged moieties, such as quarternary ammonium and phosphonium, effectively enhances the dissociation of side chains of Leu and Ile.[5,6)] Although this derivatization enabled the successful discrimination between Leu and Ile, the relationship between the basicity of the derivatization reagents and the effects on dissociation of Leu/Ile side chains is poorly understood. Meanwhile, the derivatization with a "fixed" charge is generally accompanied by the loss of mobile protons from the N-terminal amino group, which adversely affects the backbone fragmentation of a peptide under low-energy CID (LE-CID) conditions.[7)] Besides, it is difficult to obtain high-resolution product ion spectra by TOF-TOF instruments, which have been widely used for HE-CID analysis in recent years. Therefore, it is practically impossible under their operating conditions to resolve a very small mass difference between Lys and Gln (0.0364 Da) for the discrimination of these two amino acids, another difficulty in *de novo* sequencing. In this case, the use of LE-CID has a definite advantage in obtaining high-resolution product ion spectra by using instruments such as quadrupole-TOF, ion trap-TOF and Orbitrap mass spectrometers. This also indicates that the combined use of HE- and LE-CID analyses is necessary for the complete *de novo* sequencing of wide variety of peptides and proteins. Thus, the N-terminal derivatization approaches compatible with both HE- and LE-CID analyses are highly desirable for this purpose. Although sample quantities sufficient for conducting both analyses are required, it may not practically become a problem due to the high sensitivity of recent instruments.

Previously, we demonstrated that the N-terminal charge derivatization with the structures having high proton affinity induced the preferential formation of b-ions under LE-CID conditions to facilitate the interpretation of product ion spectra.[8)] It is known that a guanidino group of an Arg residue has high proton affinity, which retains a proton on that residue to generate d-ions by charge-remote fragmentation under HE-CID conditions if it exists at or near the N-terminus.[9)] This suggests that the N-terminal derivatization with the structures having high proton affinity may facilitate the side-chain fragmentation to discriminate between Leu and Ile under HE-CID conditions in addition to its improvement effect on fragmentation under LE-CID conditions. In the present study, we examined the effects of this derivatization approach on fragmentation of the alkyl side chains of Leu and Ile residues in peptides under HE-CID conditions. The usefulness of this technique for *de novo* sequencing was also demonstrated using model peptides containing Leu or Ile.

MATERIALS AND METHODS

Chemicals

4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM) was obtained from Kokusan Chemical (Tokyo, Japan). 4-Imidazolecarboxylic acid was purchased from Tokyo Chemical Industry (Tokyo, Japan). 4-(Guanidinomethyl)benzoic acid, nicotinoyloxysuccinimide and 4-amidinobenzoyloxysuccinimide were synthesized as described in the literatures.[8)] The peptides used in this study were synthesized in our laboratory.

Peptide derivatization

For derivatization with 4-amidinobenzoic acid or nicotinic acid, the succinimide ester of each compound was dissolved in phosphate buffer (250 mM, pH 8.5) at a concentration of 2.5 mM, and mixed with peptides (0.25 mM) at room temperature overnight. The derivatized peptides were purified by reversed phase (RP)-HPLC (H_2O/CH_3CN with 0.1% TFA). For derivatization with 4-imidazolecarboxylic acid or 4-(guanidinomethyl)benzoic acid, each compound was

dissolved in DMSO at a concentration of 2.5 mM and mixed with peptides (0.25 mM) and DMT-MM (2.5 mM). The resultant mixture was stirred at room temperature overnight. The derivatized peptides were purified by RP-HPLC (H_2O/CH_3CN with 0.1% TFA).

Mass spectrometric analysis

HE-CID spectra were obtained using a JMS-S3000 SpiralTOF mass spectrometer (JEOL, Tokyo, Japan). The sample dissolved in H_2O containing 0.1% TFA was mixed with a matrix solution containing 30 mg/mL of α-cyano-4-hydroxycinnamic acid and 233 mg/mL of 3-aminoquinoline in methanol, and 1 μL of the solution was spotted on the MALDI sample target. The data acquisition conditions (*i.e.*, the laser power and number of laser irradiations) were optimized to obtain product ion mass spectra of the fragment peaks that had high intensity and high signal-to-noise ratios. All mass spectra were externally calibrated, with the standards deposited on the surface of the metal plate.

LE-CID spectra were obtained using an LCMS-IT-TOF mass spectrometer (Shimadzu, Kyoto, Japan) equipped with an electrospray ion source in the positive mode. HPLC separation was carried out on a micro column, TSKgel ODS-100 V, 150×1.0 mm ID (Tosoh, Tokyo, Japan). The column was eluted at a flow rate of 50 μL/min with a 25 min linear gradient of 5 to 50% CH_3CN in H_2O with 0.1% formic acid.

RESULTS AND DISCUSSION

Figure 2 shows the structures of derivatization reagents used in this study, which contain an imidazole, pyridine, amidine or guanidine moiety. Of these, 4-guanidinomethylbenzoic acid (Gmb) has the highest proton affinity, followed by 4-amidinobenzoic acid (Aba), 4-imidazolecarboxylic acid (Iza) and nicotinic acid (Nic) in descending order. In the present study, we evaluated the effect of these basic moieties in derivatized peptides on the generation of fragment ions, such as d-ions, required for Leu/Ile discrimination under HE-CID conditions using a JMS-S3000 SpiralTOF mass spectrometer. This apparatus has been shown to produce product ions mainly by HE-CID fragmentation pathways without inclusion of the post-source decay ions in the spectra.[10)]

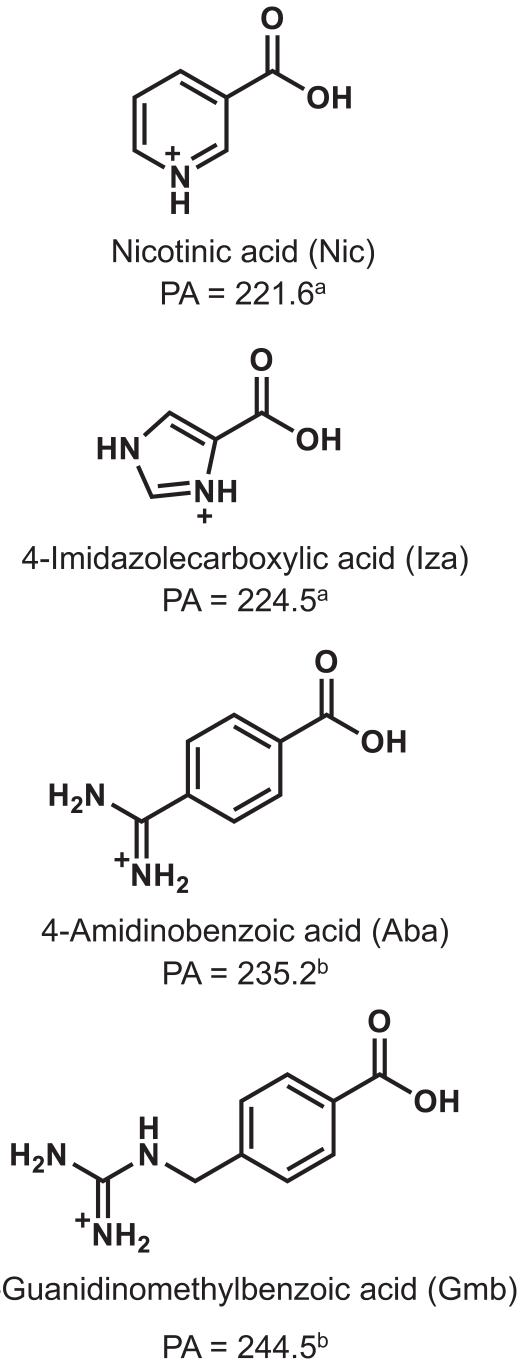


Fig. 2. Chemical structures used for derivatization. PA: Proton affinity (kcal/mol).

[a]Values were calculated as previously described.[8)] [b]Values were taken from the literature.[8)]

Effect of derivatization on d-ion generation

The model peptide (AAGLQIA) in this study was chosen to discriminate Leu and Ile residues in a single peptide. Relative intensities of d_4 and d_{a6} ions necessary for assignment of Leu and Ile, respectively, were compared among derivatized peptides (Fig. 3a). From an Ile residue, d_b-ion also can be generated, but its intensity is often lower than that of d_a-ion.[6)] First, the HE-CID spectrum was obtained without derivatization (Fig. 3b). The intensity of b-ions was relatively high in the product ion spectrum, but the intensity of d-ions (d_4 and d_{a6}) was marginal due to absence of basic residues near the N-terminus, making it difficult to unambiguously determine these residues as Leu or Ile. The similar spectra were obtained when Nic or Iza was used for derivatization (Figs. 3c and 3d). It is known that charge-remote fragmentation facilitated by the "immobilized" proton at the N-terminal basic residue competes with backbone fragmentation triggered by mobile protons attached to amide nitrogen atoms.[11)] Therefore, it is likely that the basicity of Nic or Iza is not high enough to retain protons at the N-terminus to effectively induce charge-remote fragmentation, and that the protons are transferred to amide bonds to generate b- or y-ions. On the other hand, derivatization with Aba and Gmb lead to generation of d-ions with significant intensity, which is sufficient for Leu/Ile discrimination (Figs. 3e and 3f). These results indicate that the proton affinity higher than that of Aba at the N-terminus is necessary for preferential generation of d-ions under these conditions. In particular, derivatization with Gmb gave the highest intensity of d-ions

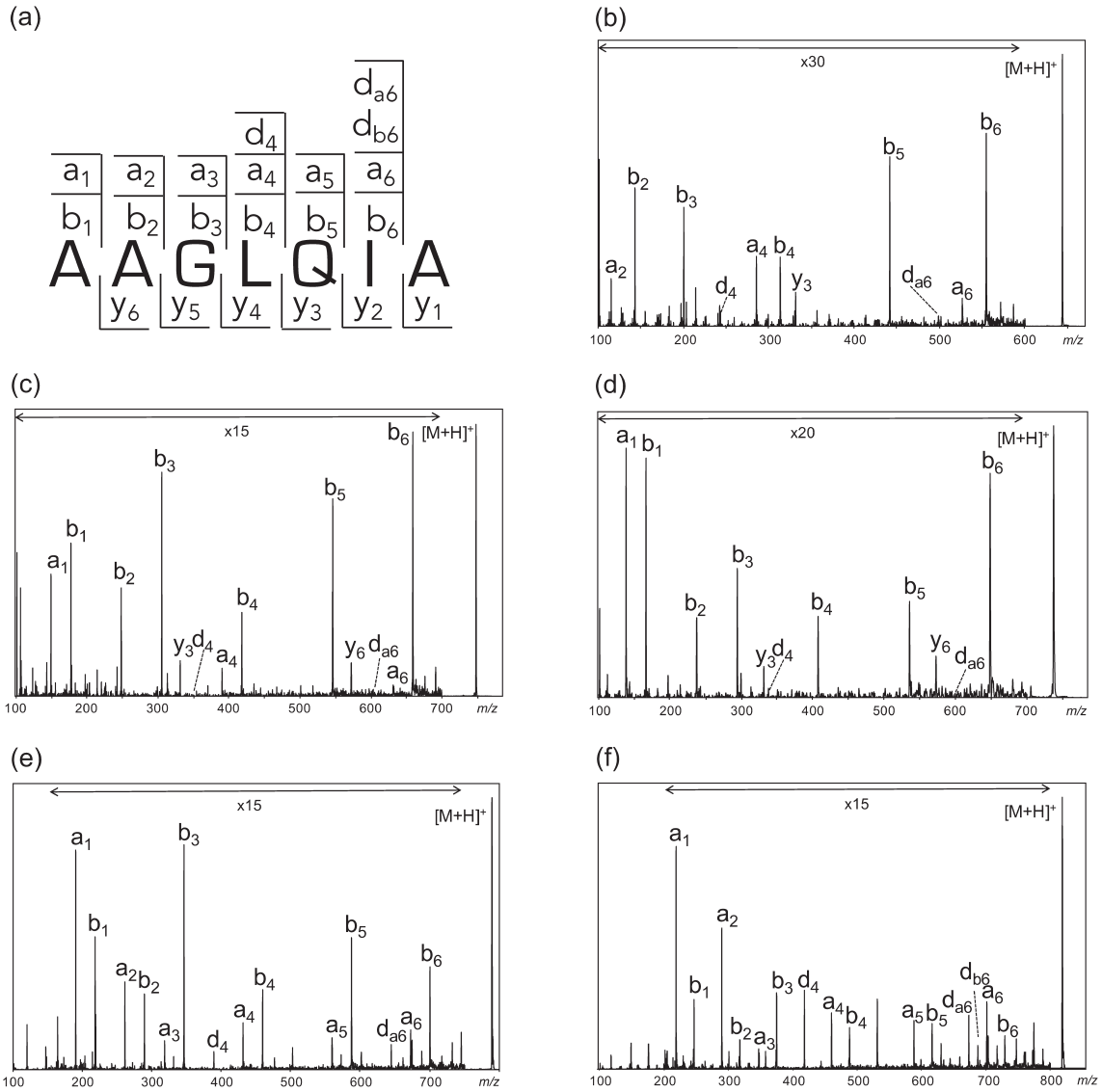


Fig. 3. HE-CID spectra of the model peptide. Possible fragment ions generated from the peptide (a). Product ion spectra of the peptide without derivatization (b), derivatized with Nic (c), Iza (d), Aba (e) and Gmb (f).

among four structures used in this study, in which d_{b6}-ion was also observed. This is consistent with the previous observations that an Arg residue, which has a guanidino group in its side chain, at or near the N-terminus has the most significant influence on generation of d-ions compared to other basic residues such as Lys and His.[4] Since we have shown that Gmb is effective for improving the backbone fragmentation of peptides under LE-CID conditions,[8] derivatization with Gmb was further evaluated for its usefulness in *de novo* sequencing in the following section.

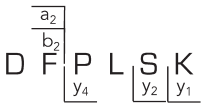
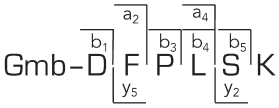
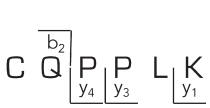
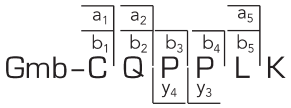
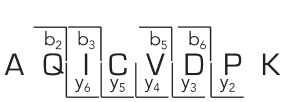
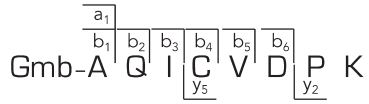
De novo sequencing assisted by Gmb-derivatization

The usefulness of Gmb-derivatization in *de novo* sequencing was evaluated using enzymatic digests of the scorpion toxin LaIT1 as a model peptide[12] (Table 1). In this study, *de novo* sequencing was performed using the following procedure; (1) derivatization with Gmb, (2) HE-CID analysis to estimate the sequence including Leu/Ile discrimination, (3) high-resolution LE-CID analysis to reinforce the sequence data obtained under HE-CID conditions.

First, product ion spectra were obtained for three peptides without Gmb derivatization. The product ion spectra obtained by HE-CID analysis provided insufficient information for sequence determination, especially for peptides 1 and 2 (Figs. 4a, 4c and 4e). This is due to the presence of Pro residues in the sequence, which often yields incomplete dissociation at its C-terminal side. Besides, d-ions were not observed for all three peptides because basic residues were absent in the N-terminal region. In the case of peptide 1 (DFPLSK), w_{a3}-ion, which allows the assignment of 4th residue as Leu, was observed, but its interpretation could be difficult due to the insufficient backbone fragmentation (Fig. 4a). The product ion spectra obtained under LE-CID conditions did not facilitate the sequence determination (Table 1 and Fig. S1). These results indicate that *de novo* sequencing is not possible without derivatization.

Then, the peptides were derivatized with Gmb using DMT-MM as described above. Gmb-derivatization greatly improved the fragmentation efficiency under HE-CID conditions, and d-ions were clearly observed in the spectra as well as backbone fragment ions such as a- and b-ions (Figs. 4b, 4d and 4f). Derivatized peptide 1 (Gmb-DFPLSK) showed all series of a-ions, and Leu/Ile discrimination at the 4th residue was possible by observation of d_4-ions, allowing

Table 1. Fragmentation of LaIT1 digests under LE-CID conditions.

No.	Before derivatization	After derivatization
1		
2		
3		

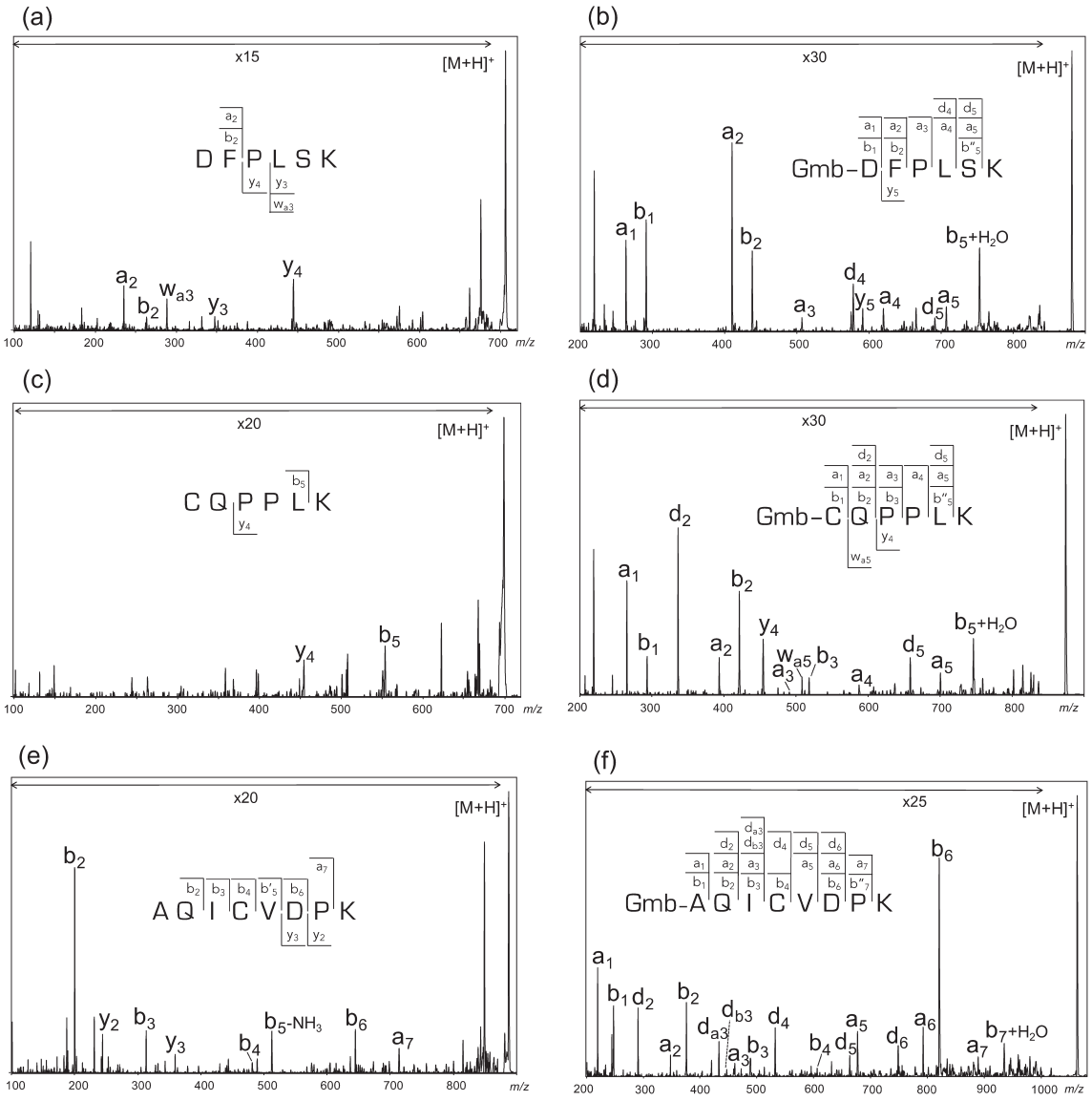


Fig. 4. HE-CID spectra of tryptic digests of LaIT1 without derivatization (a, c and e), and derivatized with Gmb (b, d and f). Single and double quotation marks in b-ions indicate b_n-NH_3 and b_n+H_2O, respectively.

estimation of the full sequence (Fig. 4b). The information of the LE-CID product ion spectrum was sufficient enough to confirm the sequence estimated by HE-CID analysis (Table 1 and Fig. S1). Similarly, the HE-CID spectrum of derivatized peptide 2 (Gmb-CQPPLK) showed all series of a-ions, and the presence of d_5-ion allowed the assignment

of the 5th residue as Leu (Fig. 4d). The estimated sequence was confirmed by the product ion spectrum under LE-CID conditions, in which all series of b-ions were detected (Table 1 and Fig. S1). In the case of derivatized peptide 3 (Gmb-AQICVDPK), most of the backbone fragment ions were observed, and d_{a3}- and d_{b3}-ions, which allowed the assignment of the 3rd residue as Ile, were clearly detected under HE-CID conditions. The assignment of 4th and 5th residues was difficult because a_4-ion was not clearly detected in this spectrum (Fig. 4f). It is possible that a_4-ion easily converted into d_4-ion through elimination of CH_3SH, because d_4-ion was observed at relatively high intensity.[13] On the other hand, the spectrum obtained by the LE-CID analysis, in which all series of b-ions were detected, could resolve this ambiguity to completely determine the sequence (Table 1 and Fig. S1). These results clearly demonstrate that the combination of HE-CID and LE-CID analysis coupled with Gmb-derivatization allows the complete determination of the peptides sequences without Arg residues at the N- or C-terminus.

CONCLUSION

For the discrimination of isomeric Leu and Ile, various methods have been developed for several decades.[6,14–16] Among them, observation of their side chain fragmentation such as d- and w-ions under HE-CID conditions is the most straightforward method. With the advent of TOF-TOF instruments, access to the HE-CID analysis becomes much easier than before, which had been available only by sector instruments. The major drawback of this method is the requirement of basic residues near the N- or C-terminal end of peptides, but there have been few studies addressing this problem. In this study, we attempted to introduce a charged moiety having various proton affinities into the peptide N-terminus, and found that the N-terminal modification with Gmb, which containing a guanidino group, is most effective for enhancing the generation of d-ions under HE-CID conditions. Unlike previous methods using derivatives having a fixed charge, our approach has an additional effect on improvement of fragmentation under LE-CID conditions. The method proposed in this study will facilitate the complete *de novo* sequencing of bioactive peptides from unexplored organisms.

Acknowledgements

This study was partly supported by MEXT KAKENHI (Grant No. 25513001), Japan.

REFERENCES

1) A. I. Nesvizhskii. Protein identification by tandem mass spectrometry and sequence database searching. *Methods Mol. Biol.* 367: 87–119, 2007.
2) K. F. Medzihradszky, R. J. Chalkley. Lessons in *de novo* peptide sequencing by tandem mass spectrometry. *Mass Spectrom. Rev.* 34: 43–63, 2015.
3) P. P. Hurtado, P. B. O'Connor. Differentiation of isomeric amino acid residues in proteins and peptides using mass spectrometry. *Mass Spectrom. Rev.* 31: 609–625, 2012.
4) W. D. van Dongen, H. F. M. Ruijters, H. J. Luinge, W. Heerma, J. Haverkamp. Statistical analysis of mass spectral data obtained from singly protonated peptides under high-energy collision-induced dissociation conditions. *J. Mass Spectrom.* 31: 1156–1162, 1996.
5) J. T. Stults, J. Lai, S. Mccune, R. Wetzel. Simplification of high-energy collision spectra of peptides by amino-terminal derivatization. *Anal. Chem.* 65: 1703–1708, 1993.
6) J. Zaia, K. Biemann. Comparison of charged derivatives for high-energy collision-induced dissociation tandem mass-spectrometry. *J. Am. Soc. Mass Spectrom.* 6: 428–436, 1995.
7) C. G. Gu, G. Tsaprailis, L. Breci, V. H. Wysocki. Selective gas-phase cleavage at the peptide bond terminal to aspartic acid in fixed-charge derivatives of Asp-containing peptides. *Anal. Chem.* 72: 5804–5813, 2000.
8) M. Miyashita, Y. Hanai, H. Awane, T. Yoshikawa, H. Miyagawa. Improving peptide fragmentation by N-terminal derivatization with high proton affinity. *Rapid Commun. Mass Spectrom.* 25: 1130–1140, 2011.
9) R. S. Johnson, S. A. Martin, K. Biemann. Collision-induced fragmentation of $(M+H)^+$ ions of peptides. Side chain specific sequence ions. *Int. J. Mass Spectrom.* 86: 137–154, 1988.
10) T. Satoh, T. Sato, A. Kubo, J. Tamura. Tandem time-of-flight mass spectrometer with high precursor ion selectivity employing spiral ion trajectory and improved offset parabolic reflectron. *J. Am. Soc. Mass Spectrom.* 22: 797–803, 2011.
11) J. E. Vath, K. Biemann. Microderivatization of peptides by placing a fixed positive charge at the N-terminus to modify high-energy collision fragmentation. *Int. J. Mass Spectrom.* 100: 287–299, 1990.
12) N. Matsushita, M. Miyashita, A. Sakai, Y. Nakagawa, H. Miyagawa. Purification and characterization of a novel short-chain insecticidal toxin with two disulfide bridges from the venom of the scorpion *Liocheles australasiae*. *Toxicon* 50: 861–867, 2007.
13) R. A. J. O'Hair, G. E. Reid. Neighboring group *versus cis*-elimination mechanisms for side chain loss from protonated methionine, methionine sulfoxide and their peptides. *Eur. Mass Spectrom.* 5: 325–334, 1999.
14) A. T. Lebedev, E. Damoc, A. A. Makarov, T. Y. Samgina. Discrimination of leucine and isoleucine in peptides sequencing with Orbitrap Fusion mass spectrometer. *Anal. Chem.* 86: 7017–7022, 2014.
15) S. L. Cook, O. L. Collin, G. P. Jackson. Metastable atom-activated dissociation mass spectrometry: Leucine/isoleucine differentiation and ring cleavage of proline residues. *J. Mass Spectrom.* 44: 1211–1223, 2009.
16) A. Armirotti, E. Millo, G. Damonte. How to discriminate between leucine and isoleucine by low energy ESI-TRAP MS^n. *J. Am. Soc. Mass Spectrom.* 18: 57–63, 2007.