Differential $^{14}\text{N}/^{15}\text{N}$-Labeling of Peptides Using N-Terminal Charge Derivatization with a High-Proton Affinity for Straightforward de novo Peptide Sequencing

Yoichiro Nihashi, Masahiro Miyashita, Hiroyuki Awane, and Hisashi Miyagawa

Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

While de novo peptide sequencing is essential in many situations, it remains a difficult task. This is because peptide fragmentation results in complicated and often incomplete product ion spectra. In a previous study, we demonstrated that N-terminal charge derivatization with 4-aminobenzoic acid (Aba) resulted in improved peptide fragmentation under low-energy CID conditions. However, even with this derivatization, some ambiguity exists, due to difficulties in discriminating between N- and C-terminal fragments. In this study, to specifically identify b-ions from complex product ion spectra, the differential $^{14}\text{N}/^{15}\text{N}$-labeling of peptides was performed using Aba derivatization. $^{15}\text{N}$-Labeled Aba was synthesized in the form of a succinimide ester. Peptides were derivatized individually with $^{14}\text{N}$-Aba or $^{15}\text{N}$-Aba and analyzed by ESI-MS/MS using a linear ion trap-Orbitrap hybrid FTMS system. The N-terminal fragments (i.e., b-ions) were then identified based on $m/z$ differences arising from isotope labeling. By comparing the spectra between $^{14}\text{N}$- and $^{15}\text{N}$-Aba derivatized peptides, b-ions could be successfully identified based on the $m/z$ shifts, which provided reliable sequencing results for all of the peptides examined in this study. The method developed in this study allows the easy and reliable de novo sequencing of peptides, which is useful in peptidomics and proteomics studies.

Keywords: de novo sequencing, derivatization, isotope labeling, fragmentation, MS/MS

(Received September 24, 2013; Accepted October 29, 2013)
Differential isotope labeling techniques have been investigated, in attempts to specifically identify a single series of fragments from complex spectra.\textsuperscript{10-12} By introducing isotopes into the N- or C-terminus, either fragment can be specifically identified based on differences in mass caused by isotope labeling. For example, specific C-terminal \textsuperscript{18}O-labeling was accomplished when proteins were enzymatically digested in the presence of H\textsubscript{2}\textsuperscript{18}O.\textsuperscript{13,14} Although this method allows the facile assignment of C-terminal fragments without the need for peptide derivatization, the fragmentation pattern itself is not improved. Furthermore, since enzymatic digestion is essential for this method, its application to peptidomics studies would be difficult, in which intact peptides are generally dealt with. Another differential isotope-labeling method reported previously is N-terminal acetylation using deuterated acetic anhydride.\textsuperscript{10} Unlike \textsuperscript{18}O-labeling, this method is applicable to any sample. However, blocking of the N-terminal amino group results in a decrease in the yield of protonated molecules due to its reduced proton affinity, which could affect the sensitivity of detection of peptides. Although fragmentation efficiency is slightly improved by N-terminal acetylation, the type of product ion formed in both N-acetylated and intact peptides is reported to be similar.\textsuperscript{10} Isotope-labeling using N,N-dimethylation of amino groups was also reported to facilitate \textit{de novo} sequencing. This method not only provides the characteristic isotope signature in the N-terminal fragments but also increases the basicity of the N-terminus.\textsuperscript{15,16} However, because the basicity of the N,N-dimethylamino group is lower than that of an amidino group, its effect on enhancing b-ion intensities could be limited. In addition, this reaction also takes place at Lys side chains, and additional treatments such as masking of the amino group by guanidination\textsuperscript{16} or digestion with a Lys-N protease to produce peptides with Lys residues at the N-terminus\textsuperscript{15} is necessary to avoid the complexity arising from multiple labeling. Thus, when N-terminal modification is employed for differential isotope labeling, its effect on fragmentation must be taken into account, which is associated with the availability of mobile protons.

In this study, we present a novel differential isotope labeling technique using N-terminal derivatization with ABA, in which a stable isotope of nitrogen (\textsuperscript{15}N) is incorporated into the amidino group of ABA. The usefulness of this method for specific identification of b-ions in product ion spectra was demonstrated using high-resolution mass spectrometers, such as an Orbitrap instrument, to establish a reliable and straightforward \textit{de novo} sequencing method.

## EXPERIMENTAL

### Chemicals

4-Aminobenzoylloxysuccinimide (ABA-NHS) was synthesized as previously described.\textsuperscript{9} Angiotensin I was purchased from Calbiochem (Darmstadt, Germany). Other peptides used in this study were synthesized in our laboratory using standard Fmoc solid-phase chemistry. Lysozyme was obtained from Sigma-Aldrich (St. Louis, MO). Acetonitrile and formic acid used for LC-MS/MS analysis were purchased from Kanto Chemical (Tokyo, Japan) and Wako Pure Chemical Industries, Ltd. (Osaka, Japan), respectively. Water used for LC-MS/MS analysis was produced by a Milli-Q gradient system (Millipore, Billerica, MA). Trypsin (mass spectrometry grade) was purchased from Wako Pure Chemical Industries, Ltd. \textsuperscript{15}N-Ammonium chloride (99% \textsuperscript{15}N) was obtained from Cambridge Isotope Laboratories (Andover, MA).

### Synthesis of \textsuperscript{15}N-4-aminobenzoylloxysuccinimide (\textsuperscript{15}N-ABA-NHS)

\textsuperscript{15}N-4-Aminotoluene

\textit{p}-Tolunitrile (7.0 g, 60 mmol) dissolved in 30 mL of methanol was added to the solution of sodium methoxide (0.37 g, 6.6 mmol) in 10 mL of methanol, and reacted for 39 h at room temperature under an atmosphere of argon. To this reaction mixture, \textsuperscript{15}N-ammonium chloride (1.0 g, 20 mmol) was added and the reaction was allowed to proceed for 5 h at room temperature. After evaporation of the solvent, the residue was suspended in diethyl ether, and the supernatant was discarded. This operation was repeated a total of five times to yield \textsuperscript{15}N-4-aminotoluene (2.16 g, 81%).

\textbf{1H-NMR (400 MHz, DMSO-\textit{d}_6)} $\delta$ 2.40 (3H, s), $\delta$ 7.41 (2H, d, $J=$7.5 Hz), $\delta$ 7.74 (2H, d, $J=$7.9 Hz), $\delta$ 8.40 (3H, br s). ESI-MS: [M+H]\textsuperscript{+} (m/z) Calculated: 136.1, Found: 136.0

\textsuperscript{15}N-4-Aminobenzoic acid

\textsuperscript{15}N-Aminobenzoic acid (1.0 g, 7.5 mmol) dissolved in an aqueous solution of sodium hydroxide (0.1 M). Potassium permanganate (2.3 g, 20 mmol) was added to this solution and reacted for 22 h at 40°C. After the reduction of the remaining sodium permanganate with sodium bisulfate, the solution was acidified with 3 M hydrochloric acid. The solution was filtered, and the filtrate was evaporated. The residue was applied to silica gel flash column chromatography eluted with chloroform/methanol/trifluoroacetic acid (90:10:1) to obtain \textsuperscript{15}N-4-aminobenzoic acid (0.27 g, 22%).

\textbf{1H-NMR (400 MHz, DMSO-\textit{d}_6)} $\delta$ 7.91 (2H, d, $J=$8.0 Hz), $\delta$ 8.12 (2H, d, $J=$8.0 Hz), $\delta$ 9.36–9.60 (3H, m), $\delta$ 13.47 (H, s). ESI-MS: [M+H]\textsuperscript{+} (m/z) Calculated: 166.1, Found: 166.0

\textsuperscript{15}N-4-Aminobenzoylloxysuccinimide

\textsuperscript{15}N-4-Aminobenzoylloxysuccinimide (165 mg, 1 mmol) was dissolved in 10 mL of N,N-dimethylformamide (DMF) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) (192 mg, 1 mmol). To the solution, N-hydroxy succinimide (NHS) (115 mg, 1 mmol) were added and reacted for 40 h at room temperature. The solution was evaporated, and the residue was applied to silica gel flush column
Peptide derivatization

For the derivatization of peptides, non-labeled or 15N-labeled 4-amidonbenzoyloxysuccinimide was dissolved in phosphate buffer (final 50 mM, pH 5.9 for peptides 1 and 3, pH 8.2 for peptides 2, 4, and 5) at a final concentration of 10 mM, and mixed with peptides (10 µM) for approximately 20 min at room temperature. The derivatized peptides were desalted using a MonoSpin C18 cartridge. The peptide fragments were dissolved in a 50 mM phosphate buffer (pH 8.2) and mixed with non-labeled or 15N-labeled 4-amidinobenzoyloxysuccinimide for approximately 20 min at room temperature. The derivatized peptides were desalted peptides were desalted using a MonoSpin C18 cartridge and analyzed by LC-MS/MS as shown below.

Protein digestion and derivatization

Lysozyme was reduced with dithiothreitol (10 mM) and its cysteine residues were carbamidomethylated by reaction with iodoacetamide (20 mM). The sample was digested with trypsin at 37°C overnight. The resulting peptide fragments were desalted using a MonoSpin C18 cartridge. The peptide fragments were dissolved in a 50 mM phosphate buffer (pH 8.2) and mixed with non-labeled or 15N-labeled 4-amidinobenzoyloxysuccinimide for approximately 20 min at room temperature. The derivatized peptides were desalted using a MonoSpin C18 cartridge and analyzed by LC-MS/MS as shown below.

Mass spectrometric analysis

LC/MS/MS analysis was performed on an LTQ Orbitrap Velos (Thermo Fisher Scientific, Waltham, MA) equipped with an ESI probe. HPLC separations were performed with a TSK-ODS V100 column (2.1×50 mm for model peptides and 2.1×150 mm for protein digests, TOSOH, Tokyo, Japan) on an Agilent 1200 HPLC system (Agilent Technologies, Santa Clara, CA). The column was eluted with a mixture of 0.1% formic acid in water (solvent A) and acetonitrile (solvent B). For the analysis of derivatized peptides, except for peptide 4, a 3-min linear gradient from 5% to 50% solvent B at a flow rate of 0.32 mL/min was used. For analysis of derivatized peptide 4, a 25-min linear gradient from 5% to 50% solvent B at a flow rate of 0.2 mL/min was used after 5 min of isocratic elution with 5% solvent B. For the analysis of protein digests, a 40-min linear gradient from 2% to 50% solvent B at a flow rate of 0.2 mL/min was used after 2 min of isocratic elution with 2% solvent B. Mass spectra were obtained in a product ion scan mode under low-energy CID conditions using a normalized collision energy of 35%. Doubly charged ions were automatically selected as a precursor ion with an isolation width of 2 Da. The resolution was set to 7,500 or 30,000 at m/z 400. Other mass spectrometric conditions were as follows: spray voltage=3 kV; sheath gas flow=50 and auxiliary gas flow=15; heated capillary temperature=230°C; predictive automatic gain control (AGC) enabled; S-lenz RF level=31 or 67%.

RESULTS AND DISCUSSION

In a previous paper, we demonstrated that charge-derivatization with 4-amidinobenzoic acid (Aba) at the peptide N-terminus significantly improved peptide fragmentmentation by enhancing the intensity of b-ions. However, in some cases, the identification of b-ions was still not very easy due to the overlap of b-ions with other series of ions, such as y-ions, especially when multiply charged ions were selected as a precursor. In the case of angiotensin I, derivatization with Aba permitted the complete series of b-ions to be detected, but several y-ions were observed with a sufficient intensity to complicate the spectrum because two His residues are present in the C-terminal half (Fig. 2a). In the present study, to distinguish N-terminal fragments (b-ions) from C-terminal fragments (y-ions), we utilized isotope-labeled Aba as a derivatization reagent. A stable isotope of nitrogen (15N) was incorporated into an amidino moiety of Aba (15N-Aba, Fig. 1), which provides a mass shift of 0.997 Da for N-terminal fragments when compared with fragments derivatized with non-labeled Aba (14N-Aba). Since this mass shift overlaps with isotope peaks arising from an intrinsic 13C isotope, it is difficult to distinguish between them when peptides derivatized with a mixture of 14N- and 15N-Aba. Therefore, it is necessary to perform MS/MS analysis individually for each derivatized peptide to detect the mass shifts, although it is possible to distinguish incorporated isotopes from natural isotopes using a software tool. The utilization of high-resolution mass spectrometers, such as an Orbitrap mass spectrometer, permits the reliable identification of fragments when ions with similar m/z values are present in product ion spectra. The effectiveness of parallel 14N/15N-Aba derivatization for the unambiguous identification of b-ions was evaluated using several model peptides in the following sections.

Effect of 14N/15N-Aba derivatization on de novo sequencing of angiotensin I

The effect of parallel 14N/15N-Aba derivatization was first evaluated using angiotensin I. This peptide contains three basic amino acid residues that hamper efficient fragmentmentation (Supplemental Fig. 1a). Basic residues such as Arg, His and Lys often suppress the dissociation of amide bonds that are located adjacent to those residues, because the protons required for bond dissociation are repelled by the positive charge of their side chains. In addition, the amide bond C-terminal to Pro is rarely dissociated because of the nature of its structure. These features resulted in a low intensity of several b-ions in the product ion spectrum, even after 14N-Aba derivatization (Fig. 2a). In addition, the intensity of several y-ions was enhanced by the presence of the basic residues. To correctly identify b-ions, the peptides were individually derivatized with 15N-Aba, and the resulting spectrum was compared with that of peptides derivatized with 14N-Aba. Doubly charged ions (m/z 721.87 and
722.37 for 14N- and 15N-Aba, respectively) were chosen as a precursor for the MS/MS analysis, and similar product ion spectra were obtained for both derivatized peptides (Fig. 2a). A complete series of b-ions were observed in both spectra by virtue of Aba-derivatization, as shown in the previous report\(^9\) (Table 1). Closer inspection of the spectra revealed that several fragment ions showed different \(m/z\) values between 14N- and 15N-Aba derivatized peptides. These \(m/z\) shifts (0.997 and 0.499 Da for singly and doubly charged ions, respectively) were detected only for b-ions (Figs. 2b, 2c and Table 1). For example, the y\(_8\) ion was observed without any \(m/z\) shift (\(m/z\) 1025.556 and 1025.555 for both derivatized peptides), but the \(m/z\) value of the b\(_7\) ion was different between 14N- and 15N-Aba derivatized peptides (\(m/z\) 1027.509 and 1028.507 for 14N- and 15N-Aba, respectively, Fig. 2b). The difference of 0.998 Da matched the theoretical value, indicating that these fragments contain the Aba-derivatized N-terminus. Likewise, the b\(_3\) ion showed an \(m/z\) shift whereas the y\(_4\) ion remained unchanged (Fig. 2c). Moreover, high-resolution mass spectra contributed to the reliable assignment of fragments, especially for discriminating between ions with similar \(m/z\) values. For example, Aba derivatization resulted in the formation of the b\(_3\) ion as a doubly-charged ion at \(m/z\) 514.259, which was overlapped with the isotopic ion of singly-charged y\(_1\) at \(m/z\) 514.286 (Fig. 2c). A weak signal at \(m/z\) 514.259 was also observed in the 15N-Aba derivatized peptide, but this is likely due to the remaining 14N-Aba (~3%) in 15N-Aba. Discrimination between these two ions is not possible when a low-resolution mass spectrometer is used, where the \(m/z\) shift of the b\(_3\) ion cannot be detected. Combined with the advantages of high-resolution mass spectrometry, differential 14N/15N labeling of Aba promises to greatly facilitate the unambiguous interpretation of peptide fragmentation, leading to a more reliable de novo sequencing method.

### Evaluation of the effect of 14N/15N-Aba derivatization using various model peptides

To evaluate the usefulness of 14N/15N-Aba derivatization, we performed sequencing of various types of peptides that contain basic residues, which often provide complicated product ion spectra. Peptide 2 (TDVNGDGR), a model for tryptic digests, contains an Arg residue at the C-terminus. As tryptic digests possess an Arg or Lys residue at the C-terminus, y-ions are generally favorably produced without derivatization. However, the assignment of y-ions was not straightforward in the case of peptide 2, because several b-ions were also observed at significant intensities (Supplemental Fig. 1b). After derivatization with 14N-Aba, all of the b-ions were generated, allowing straightforward de novo sequencing (Fig. 3b). Peptide 3 (INSAKDDAAGLAIA) has a relatively long sequence, and contains a Lys residue in the middle of the sequence. Without derivatization, the intensity of most of the b- and y-ions of this peptide were relatively low, making de novo sequencing difficult (Supplemental Fig. 1c). After derivatization with 14N-Aba, all of the b-ions were generated.

### Table 1. List of fragment ions showing mass shifts in the product ion spectra between 14N- and 15N-Aba derivatized angiotensin I.

| m/z 14N-Aba | 262.083 | 418.184 | 465.733 | 514.259 | 517.252 | 517.252 | 587.793 | 656.323 | 680.315 | 793.399 | 930.457 | 1027.509 | 1174.579 |
| m/z 15N-Aba | 263.079 | 419.181 | 466.231 | 514.758 | 518.248 | 588.292 | 656.821 | 681.312 | 794.396 | 931.455 | 1028.507 | 1175.577 |
| Difference | 0.996 | 0.997 | 0.498 | 0.499 | 0.499 | 0.498 | 0.997 | 0.997 | 0.998 | 0.998 | 0.998 | 0.998 | 0.998 | 0.998 |
| Assignment | b\(_1\) | b\(_2\) | b\(_2^+\) | b\(_2^+\) | b\(_3\) | b\(_3^+\) | b\(_3^+\) | b\(_4\) | b\(_4\) | b\(_4\) | b\(_4\) | b\(_5\) | b\(_5\) | b\(_5\) | b\(_5\) | b\(_5\) | b\(_5\) |
but some ambiguity continued to exist as a result of the formation of several y-ions particularly in the high m/z region because of the presence of a Lys residue in the N-terminal part (Fig. 4a). However, a comparison of the spectrum with that obtained using 15N-Aba derivatization allowed b-ions to be unambiguously identified, as shown in Fig. 4b.

Peptide 4 (SFLLRN-NH$_2$) contains an Arg residue in the middle of the sequence, and its C-terminus is amidated. These substructures are often found in naturally occurring bioactive peptides. The underivatized peptide did not provide all of the b- and y-ion series (Supplemental Fig. 1d). After 14N-Aba derivatization, a complete series of b-ions was generated, but the intensity of the y-ions remained high because of the presence of an Arg residue at the C-terminal part. Again, the parallel use of 15N-Aba facilitated the unambiguous identification of b-ions for this peptide (Fig. 5).

Peptide 5 (AAGLQIANRLTTS) also contains an Arg residue in the middle of the sequence, and gave a complicated fragmentation pattern without Aba derivatization (Supplemental Fig. 1e). After 14N-Aba derivatization, all of the b-ions except the b$_8$ ion were observed in the product ion spectrum (Fig. 6). In this case, the presence of an Arg residue at the fourth position from the C-terminus enhanced the intensity of y$_4$ to y$_{11}$ ions. Since amide bond cleavage between the eighth and ninth residues was hampered by the neighboring Arg residue, the b$_8$ ion was not generated in significant intensities. In such cases, however, the corresponding y-ions can compensate for the missing b-ions. By comparing the product ion spectra, the y$_4$ ion, which is complementary to the b$_8$ ion, was identified as a fragment ion without m/z
De novo Sequencing Using Derivatization and Isotope Labeling

shifts between the 14N- and 15N-Aba derivatized peptide 5, and the full sequence was successfully determined. The adverse effects of Arg residues on fragmentation can also be diminished by modifying the guanidino group of the Arg side chain to decrease its proton affinity. 20)

Application to de novo sequencing of proteins

To explore the potential of the method described in this study, we applied it to the de novo sequencing of a protein digest, using lysozyme as a model protein. The protein was digested with trypsin, and the resulting peptide mixture was derivatized individually with light or heavy Aba. Each derivatized peptide mixture was analyzed by LC/MS/MS using the data-dependent scan mode. As a result, seven peptides were detected as ions with mass shifts, and their product ion spectra were obtained (Table 2). Among them, the full sequences of four peptides were easily determined by using only information for the b-ions. For the other three peptides, some of the b-ions were not clearly observed, but their full sequences were successfully determined with the aid of their complimentary y-ions, as in the case of peptide 5. Sequence coverage was not high in this analysis because some digested peptide fragments are too large or too small to permit their analysis MS/MS. Further optimization of sample preparations, such as the utilization of different enzymes and desalting methods, should improve the coverage. Consequently, the method described in this study promises to facilitate the straightforward de novo sequencing of peptides or proteins from organisms with unsequenced genomes.

CONCLUSION

In this study, we investigated de novo peptide sequencing strategies using differential 14N/15N-labeling coupled with N-terminal charge derivatization with Aba. In a previous study, we showed that Aba derivatization confers a well-balanced proton affinity to the peptide N-terminus, which serves as a donor of mobile protons and also an acceptor of protons for the preferential generation of b-ions. 20) In addition to this effect, differential 14N/15N-labeling using Aba
enabled the unambiguous identification of b-ions for all peptides used in this study, which provided more reliable sequence determination. Furthermore, the use of a high-resolution mass spectrometer contributed to the accurate assignment of each fragment ion, particularly when fragment ions with similar \( m/z \) values were present in the spectrum. The method developed in this study allows the easy and reliable de novo sequencing of peptides, which is useful in peptidomics and proteomics studies.

Since the difference in molecular mass between \( ^{15}\text{N} \)- and \( ^{14}\text{N} \)-labeled Aba is only 1Da, MS/MS analysis was performed individually for each derivatized peptide. However, the introduction of two or more isotopes (\( ^{15}\text{N} \) or \( ^{13}\text{C} \)) into Aba would enable a single analysis of peptides derivatized with a mixture of non-labeled and isotope-labeled Aba. In addition, the detection of mass shifts was performed manually in this study, which required some time in terms of assigning b-ions. To further increase the efficiency of de novo sequencing using this method, the development of a software tool for automated detection of mass shifts due to \( ^{15}\text{N} \)-labeling would be required to assign key fragments in a more convenient manner.

**Acknowledgement**

This study was partly supported by MEXT KAKENHI Grant Number 25513001.

**REFERENCES**