High-Throughput Method for N-Terminal Sequencing of Proteins by MALDI Mass Spectrometry

Minoru Yamaguchi,†,‡ Takashi Nakazawa,‡ Hiroki Kuyama,† Takashi Obama,† Eiji Ando,† Taka-aki Okamura,§ Norikazu Ueyama,§ and Shigemi Norioka†

Life Science Laboratory, Shimadzu Corporation, Kyoto 604-8511, Japan, Department of Chemistry, Nara Women’s University, Nara 630-8506, Japan, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, and Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan

A high-throughput method for sequencing of N termini of proteins by using postsource decay (PSD) of matrix-assisted laser desorption/ionization mass spectrometry has been developed. After a protein blotted on the PVDF membrane was successively reduced, S-alkylated, and guanidinated, its N-amino group was coupled to biotinylcysteic acid. The protein was then extracted from the membrane and digested with trypsin. The derivatized N-terminal fragment was then specifically isolated from the tryptic digest with avidin resins, and its de novo sequencing was successfully performed by PSD utilizing a sulfonic acid group introduced to the N terminus.

N-terminal sequencing of proteins is very important not only to confirm N-terminal processing, such as removal of initiator methionine residues and signal peptides and modification of Nα-groups of proteins, but also to identify the target proteins, of which identification is estimated by the routine peptide mass fingerprinting (PMF) method coupled with both mass spectrometry and gene/protein database search. For this purpose, Edman degradation has been widely employed for the N-terminal sequencing of proteins, but this method has limitations in sensitivities and throughput for presently growing proteome research. To overcome such drawbacks in the Edman method, peptide sequencing using postsource decay (PSD) analysis coupled with database search has been used as an alternative method. However, the method developed so far cannot determine the N-terminal sequence of proteins specifically. Furthermore, de novo peptide sequencing with PSD spectra has several problems due to their complexities of the resulting fragmentation patterns of peptides and of the nonestablishment of computer algorithms to interpret exact peptide sequences from the complex spectra. Several chemical derivatization methods have been thus reported to enhance sensitivity and simplify fragmentation patterns observed in PSD spectra for de novo peptide sequencing. Keough and co-workers reported that the addition of sulfonic acid groups to the N terminal of tryptic peptides facilitates de novo peptide sequencing.

Here, we developed a new high-throughput method in combination with specific isolation of the N-terminal peptides from proteins by the biotin–avidin technique and their de novo sequencing by the addition of sulfonic acid groups to the N-terminal of peptides. A general scheme depicting the procedure for the method is shown in Figure 1. Protein was reduced, S-alkylated, and then guanidinated with o-methyisourea to convert ε-amino groups of lysine residues to guanidino groups to protect ε-amino groups against reaction with biotinylcysteic acid (Figure 2). Next, biotinylcysteic acid reacted selectively with the Nα-amino group of the protein. The derivatized protein was digested with

trypsin, and the digests were applied to avidin resins. After washing the resins to remove tryptic peptides except for the N-terminal fragment, the fragment modified with biotinylcysteic acid was eluted from the resins. With effective use of biotinylcysteic acid at the N terminus, the peptide was successfully used for de novo sequencing by PSD. We emphasize that the method developed in this study can support present proteome research by facilitating high-throughput N-terminal analysis of many proteins in a parallel manner.

**EXPERIMENTAL SECTION**

**Chemicals and Reagents.** Cysteic acid, biotin and [2-(N-morpholino)ethanesulfonic acid] (MES) were obtained from Nacalai Tesque, Inc. (Kyoto, Japan). Sulfo-N-hydroxysuccinimidobiotin (sulfo-NHS-biotin), 1-ethyl-3-(dimethylaminopropyl)carbodiimide (EDC), sulfo-N-hydroxysuccinimide (sulfo-NHS), and triethylamine (TEA) were purchased from Pierce (Rockford, IL). o-Methylisourea hemisulfate was obtained from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Bovine serum albumin (BSA), chicken egg white lysozyme (LYZ) and α-cyano-4-hydroxycinnamic acid (CHCA) were purchased from Sigma (St. Louis, MO). Laminin pentapeptide was purchased from Peptide Institute, Inc. (Osaka, Japan). Sequencing grade-modified trypsin was obtained from Promega (Madison, WI). ZipTipC18 was purchased from Millipore Corp. (Bedford, MA). Chicken egg white lysozyme fragment (FESNFNTQATNR) was synthesized using Shimadzu peptide synthesizer PSSM-8 (Kyoto, Japan). All other chemicals were analytical reagent grade and were used without further purification.

**Synthesis of Biotinylcysteic acid (BCA).** Cysteic acid (33 mg) dissolved in 450 µL of distilled water, 34 mg sulfo-NHS-biotin (1.0 equiv) dissolved in 600 µL of distilled water, and 50 µL of EDC were mixed for 2 min prior to use and reacted with 50 µL of sulfo-NHS-biotin in MES buffer (pH 6.0), which was premixed for 2 min prior to use and reacted with MES buffer (pH 6.0). After washing the resins to remove tryptic peptides except for the N-terminal fragment, the fragment modified with biotinylcysteic acid was eluted from the resins. With effective use of biotinylcysteic acid at the N terminus, the peptide was successfully used for de novo sequencing by PSD. We emphasize that the method developed in this study can support present proteome research by facilitating high-throughput N-terminal analysis of many proteins in a parallel manner.

**EXPERIMENTAL SECTION**

**Chemicals and Reagents.** Cysteic acid, biotin and [2-(N-morpholino)ethanesulfonic acid] (MES) were obtained from Nacalai Tesque, Inc. (Kyoto, Japan). Sulfo-N-hydroxysuccinimidobiotin (sulfo-NHS-biotin), 1-ethyl-3-(dimethylaminopropyl)carbodiimide (EDC), sulfo-N-hydroxysuccinimide (sulfo-NHS), and triethylamine (TEA) were purchased from Pierce (Rockford, IL). o-Methylisourea hemisulfate was obtained from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Bovine serum albumin (BSA), chicken egg white lysozyme (LYZ) and α-cyano-4-hydroxycinnamic acid (CHCA) were purchased from Sigma (St. Louis, MO). Laminin pentapeptide was purchased from Peptide Institute, Inc. (Osaka, Japan). Sequencing grade-modified trypsin was obtained from Promega (Madison, WI). ZipTipC18 was purchased from Millipore Corp. (Bedford, MA). Chicken egg white lysozyme fragment (FESNFNTQATNR) was synthesized using Shimadzu peptide synthesizer PSSM-8 (Kyoto, Japan). All other chemicals were analytical reagent grade and were used without further purification.

**Synthesis of Biotinylcysteic acid (BCA).** Cysteic acid (33 mg) dissolved in 450 µL of distilled water, 34 mg sulfo-NHS-biotin (1.0 equiv) dissolved in 600 µL of distilled water, and 50 µL of EDC were mixed for 2 min prior to use and reacted with 50 µL of sulfo-NHS-biotin in MES buffer (pH 6.0), which was premixed for 2 min prior to use and reacted with MES buffer (pH 6.0). After washing the resins to remove tryptic peptides except for the N-terminal fragment, the fragment modified with biotinylcysteic acid was eluted from the resins. With effective use of biotinylcysteic acid at the N terminus, the peptide was successfully used for de novo sequencing by PSD. We emphasize that the method developed in this study can support present proteome research by facilitating high-throughput N-terminal analysis of many proteins in a parallel manner.

**EXPERIMENTAL SECTION**

**Chemicals and Reagents.** Cysteic acid, biotin and [2-(N-morpholino)ethanesulfonic acid] (MES) were obtained from Nacalai Tesque, Inc. (Kyoto, Japan). Sulfo-N-hydroxysuccinimidobiotin (sulfo-NHS-biotin), 1-ethyl-3-(dimethylaminopropyl)carbodiimide (EDC), sulfo-N-hydroxysuccinimide (sulfo-NHS), and triethylamine (TEA) were purchased from Pierce (Rockford, IL). o-Methylisourea hemisulfate was obtained from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Bovine serum albumin (BSA), chicken egg white lysozyme (LYZ) and α-cyano-4-hydroxycinnamic acid (CHCA) were purchased from Sigma (St. Louis, MO). Laminin pentapeptide was purchased from Peptide Institute, Inc. (Osaka, Japan). Sequencing grade-modified trypsin was obtained from Promega (Madison, WI). ZipTipC18 was purchased from Millipore Corp. (Bedford, MA). Chicken egg white lysozyme fragment (FESNFNTQATNR) was synthesized using Shimadzu peptide synthesizer PSSM-8 (Kyoto, Japan). All other chemicals were analytical reagent grade and were used without further purification.
left to stand at 65 °C for 1 h. After reaction, a portion (2 μL, 20 pmol) was loaded onto the C18ZipTip, which was equilibrated with 0.1% trifluoroacetic acid (TFA); the column was washed with 2 × 10 μL of 0.1% TFA; and the derivatized peptide was eluted with 2 μL of 50% acetonitrile/0.1% TFA. For model proteins, BSA or LYZ was used. Each protein was dissolved in distilled water (100 pmol/μL), and 10 μL of each was loaded onto the PVDF membrane, followed by drying. The reactions described below (reduction, S-alkylation, guanidination, and coupling of BCA) were performed on the PVDF membrane. Each protein blotted on the membrane was reduced with 1 mL of a mixture of 10 mM dithiothreitol in 100 mM NH₄HCO₃ and MeCN (8:2 v/v) at 56 °C for 1 h. The solution was removed, and 1 mL of 55 mM iodoacetamide in 100 mM NH₄HCO₃ was added and reacted at room temperature for

Figure 3. MALDI-TOFMS spectrum of reaction mixture of laminin pentapeptide with BCA. (intact peptide: [M + H]+ m/z = 594.36. BCA derivative: [M + H]+ m/z = 971.41).

Figure 4. MALDI-PSD spectra of laminin pentapeptide. The spectra for intact peptide (upper) and for derivatized peptide with BCA (lower) are shown.
45 min. Finally, the membrane was washed with distilled water and dried at 37 °C. The protein on the PVDF membrane was further guanidinated by adding 1 mL of 0.85 M o-methylisourea hemisulfate in 7 M aqueous ammonia/MeCN (8:2 v/v) solution for reaction at 60 °C for 30 min. The PVDF membrane was washed and dried again. The resulting sample was coupled with BCA by adding 900 μL of 0.1 M MES buffer (pH 6.0)/MeCN (8:2 v/v) containing 2 mM BCA, 2 mM EDC, and 5 mM sulfo-NHS to react at 65 °C for 1 h. The PVDF membrane was washed to remove excess coupling reagent and dried. The derivatized protein was extracted from the PVDF membrane by incubating in 200 μL of 1% TFA /MeCN (3:7) solution at 60 °C overnight. To recover the remaining protein, the extraction from the membrane was repeated twice with an incubation for 1 h each. The recovered protein was dried completely using a vacuum centrifuge, and the resultant was digested with trypsin (3 μg) in 20 μL of 100 mM NH₄HCO₃ solution containing 1 M urea and 5 mM CaCl₂ at 37 °C overnight.

Specific Isolation of N-Terminal Peptide Fragments. Avidin resins were regenerated according to the manufacturer’s manual. A 30μL portion of a suspension of avidin resins equilibrated with 100 mM phosphate buffer (pH 7.0) was placed into a microtube (1.5 mL), and the supernatant was carefully removed by centrifugation. The solution of tryptic peptides prepared above was applied to the suspension of avidin resins, and the N-terminal fragment was specifically absorbed to the resins using vortex-mixing at room temperature for 1 h. The avidin resin was washed three times with 1 mL of 0.1 M phosphate buffer (pH 7.0). The absorbed N-terminal peptide fragment was finally eluted twice with 150 μL of 10% aqueous acetic acid and with 150 μL of 0.1% TFA/MeCN (1:1, v/v) solution and then evaporated to dryness. The resulting dried peptide was reconstituted in 10 μL of 0.1 M phosphate buffer.

Mass Spectrometry. Both MALDI-MS and MALDI-PSD spectra were recorded on an AXIMA-CRF-Plus mass spectrometer (Shimadzu/Kratos, Kyoto, Japan) equipped with a 337-nm N₂ laser and an integrated 2-GHz transient recorder. All of the MALDI spectra were acquired in reflectron mode with pulsed extraction and detected positive ions, except for the case in which negative ions were measured for chicken egg white lysozyme fragment. The acceleration voltage was set to 20 kV using a gridless-type electrode. In reflectron mode, each spectrum was obtained by accumulating 250 laser shots, and PSD spectra were acquired by 250–500 laser shots. All of the PSD spectra were obtained by focusing all of the fragment ions in a single spectrum by a curved-

Figure 5. MALDI-TOFMS spectra of synthesized chicken egg white lysozyme fragment (FESNFNTQATNR) derivatized with BCA. The upper spectrum was acquired in positive mode, and the lower was in negative mode. In positive mode, the derivatized peptide peak at m/z 1805.56 shows very low intensity in comparison with the intact peptide peak at m/z 1428.68; however, in negative mode, the intensity of the derivatized peptide at m/z 1803.63 is almost equivalent to that of the intact one at m/z 1426.80.
Either CHCA or DHB was used as the matrix. CHCA was used as a saturated solution in 0.1% TFA containing 50% MeCN, and DHB was prepared as a 10 mg/mL solution in 0.1% TFA containing 30% MeCN. A portion (0.5 μL) of each sample solution and 0.5 μL of the matrix solution were mixed on the MALDI target plate and analyzed after drying. In reflectron mode, LYZ analysis was internally calibrated using autolysis peaks of trypsin, and the other analyses were performed with the peaks of dimeric CHCA and insulin B chain. PSD calibration was performed externally by using selected fragment ions produced from angiotensin II.

RESULTS AND DISCUSSION

Derivatization and de Novo Sequencing of Model Peptides. The introduction of a sulfonic acid group to the N terminus of proteins using 4-sulfophenyl isothiocyanate as a derivatization reagent has been known as an effective method in facilitating de novo sequencing. In this study, we synthesized biotinylcysteic acid for a novel reagent for derivatization at N termini of proteins to facilitate both the specific isolation of N-terminal peptides from proteins and their de novo sequencing.

To examine the effectiveness of the derivatization with BCA for the purpose described above, we attempted to modify the N termini of two peptides having arginine residues at their C termini: Laminin pentapeptide YIGSR-NH2 [M + H]+ at 594.34 and synthesized chicken egg white lysozyme fragment FESN-FNTQATNR [M + H]+ at 1428.64. Figure 3 shows a mass spectrum of the reaction mixture of laminin pentapeptide modified with BCA. At the optimum reaction conditions, the major peak at m/z 971.41, which corresponds to the modified product, was detected over 80%, in comparison with the intensity of intact peptide (m/z 594.36). The PSD spectra of intact laminin pentapeptide and BCA-modified counterpart are compared in Figure 4. The spectra obtained from the intact peptide show a complex fragmentation pattern consisting of various series of ions. The exact interpretation of the spectrum would be difficult. On the other hand, in the spectrum from the peptide derivatized with BCA, y-type ions mainly appeared, and it was possible to interpret directly the sequence of the peptide. In the case of the chicken egg white lysozyme fragment modified with BCA, the intensity of the peptide is very low as compared with the underivatized counterpart (<500-fold) in positive mode; however, in negative mode, an approximately equal intensity of ions was obtained for the derivatized peptide and the underivatized one (Figure 5). These results suggest that the analysis in negative mode is
efficient for screening of the N-terminal peptide fragment of proteins, even when it is contaminated with minor internal fragments. Figure 6 shows the PSD spectrum of intact chicken egg white lysozyme fragment and the modified one with BCA. In the PSD spectrum of the unmodified peptide, complex fragmentations consisting of N- and C-terminal series ions emerged, and those made the interpretation for de novo sequencing more difficult. However, in the PSD spectrum of modified chicken egg white lysozyme fragment with BCA, although the precursor ion was very low in sensitivity when compared with that of the unmodified peptide, y-type ions with sufficient intensity and z-type ions are predominantly observed, and they enable us to interpret the whole sequence of the peptide. These results of model peptides showed the method using BCA as a modification reagent was efficient at facilitating de novo sequencing of peptides having arginine residues at the C termini. Therefore, the introduction of BCA at the N-terminal amino groups of peptides would also be expected as an effective method for de novo sequencing of peptides having homoarginine at the C termini. Considering these reports and the necessity to block ε-amino groups of lysine residues prior to modifying the Nα-amino group of a protein, we employed the guanidination method as a modification of lysine residues.

Specific Isolation of N-Terminal Fragments from Proteins and Their de Novo Sequencing. To confirm effectiveness of the procedure presented here, two proteins (BSA and LYZ) were examined. Each protein was blotted on the PVDF membrane and dried and then subjected to reduction/alkylation, guanidination, derivatization with BCA, and then subjected to direct (unmodified) and modified sequencing. MALDI TOF MS spectra of bovine serum albumin (BSA) and chicken egg white lysozyme (LYZ) were used.

References:

and modification with BCA on the membrane, because in the reaction on the PVDF membrane, removal of excess reagents was possible simply by washing the membrane. The protein was then extracted from the membrane to obtain excellent sequence coverage on trypsin digestion. The extracted protein was digested with trypsin, and the digests were loaded into avidin resins after desalting with C18ZipTip. After washing the resins with the buffer, as described in the Experimental Section, to remove internal tryptic fragments, the N-terminal fragment was eluted from avidin resins as described above and subjected to MALDI-MS and MALDI-PSD analyses. Figure 7a shows the MALDI mass spectrum obtained for trypsin digests from LYZ, of which the ε-amino groups were guanidinated and the N-terminal amino group was modified with BCA. The mass peak corresponding to the N-terminal fragment (BCA-JVFGFR; J = homoarginine) was observed at m/z 1025.50 (calculated mass: [M + H]+ = 1025.47) in the spectrum along with other peaks corresponding to the tryptic fragments of LYZ. After elution of the BCA derivatized peptide from avidin resins, the peak corresponding to the N-terminal peptide was exclusively detected at m/z 1025.45 in the MALDI spectrum (Figure 7b). The result demonstrates the N-terminal peptide fragment of LYZ was specifically isolated in this method. Furthermore, on the PSD analysis of the mass peak at m/z 1025.45, y-ion series were predominantly detected to interpret easily the sequence of the peptide (Figure 7c). The identification of the N-terminal sequence of LYZ was thus carried out smoothly. Figure 8 shows the MALDI-MS spectra at each step in the procedure of this method where BSA was used. The mass peak corresponding to the N-terminal fragment (BCA-DTHJSEIAEFR) could be detected at m/z 1612.78 (calculated mass: [M + H]+ = 1612.70), as shown in Figure 8a, and this N-terminal peptide is enriched with avidin resins, although three other minor peaks still appeared (Figure 8b). Regarding the intensity of these minor peaks, one was matched with the internal fragment (RHPEYAVSLIR; calculated mass: [M + H]+ = 1439.81), whereas the others, at m/z 1649.74 and 1681.85, were unassigned fragments and less than 20% of the main peak. This means that with the careful analysis of the MALDI-PSD spectrum of the peak at m/z 1612.78, interpretation of BCA modification was successfully performed (Figure 8c).

In Figure 8c, y-series and z-series ions were mainly detected. In the product ion corresponding to positions 1–6 from the C-terminal arginine residue, y-series ions (y1–y6) predominantly appear, and in those corresponding to positions 7–10, both z-series ions (Z7–Z10) and y-series ions (y7–y10) were detected. The increase in the intensity of the Z7–Z10 ions in comparison with the y7–y10 ions may relate to the existence of a homoarginine residue at position 7 from the C terminus of the peptide. Therefore, de novo determination of the peptide sequence in this method is not a problem, because its interpretation is possible on the PSD spectrum by considering both y-type and z-type ions, which are the same types of ions containing C-terminal amino acid, even though the intensity between y- and z- ions might be reversed.

CONCLUSIONS

The specific isolation of the N-terminal fragment from a protein and its direct sequencing by MALDI mass spectrometry were carried out through the procedure of the derivatization of the N-terminal group with BCA. Adding that the N-terminal peptide is selectively isolated by modification of N terminus of the protein with BCA, we also demonstrate that peptides having an arginine or a homoarginine residue at the C terminus is facilitated for de novo sequencing by MALDI-PSD and this modification. In this paper, the results for two model proteins are shown, but the method developed here is useful for any proteins and their complexities, of which N termini are not blocked if there are either lysines or arginines at less than dozens of residues from N termini. The manuscript for a method of specific isolation of N-terminal fragments from N-blocked proteins and their subsequent sequence determination by MALDI-PSD is now in preparation. In the near future, we will combine these two methods for the N-terminal sequence analysis of all proteins. Although the sensitivity of the method described here is in picomolar levels so far, the improvement in sensitivity for femtomolar levels could be achieved before long. The issue lies in sample manipulation, and intensive study is underway. The method could, therefore, become useful for high-throughput N-terminal sequencing of proteins with parallel treatment of samples on the PVDF membrane.

ACKNOWLEDGMENT

We thank Dr. Susumu Tsunasawa for organizing and managing this project and Prof. Dr. Nobutaka Fujii, Graduate School of Pharmaceutical Sciences, Kyoto University, for valuable suggestions associated with the synthesis of BCA. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan for S.N.

Received for review August 17, 2004. Accepted October 22, 2004.
AC048776W