Journal of Chromatography B, 877 (2009) 2607–2614

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Review

A new strategy for protein biomarker discovery utilizing 2-nitrobenzenesulfenyl (NBS) reagent and its applications to clinical samples☆

Ei-ichi Matsuo[a,b], Makoto Watanabe[a,b], Hiroki Kuyama[a], Osamu Nishimura[a,*]

[a] Division of Disease Proteomics, Institute for Protein Research, Osaka University, Osaka 565-0871, Japan
[b] Life Science Research Center, Technology Research Laboratory, Shimadzu Corporation, Kyoto, Japan

ARTICLE INFO

Article history:
Received 17 November 2008
Accepted 24 May 2009
Available online 29 May 2009

Keywords:
Quantitative proteome
Biomarker discovery
Stable isotope
NBS method

ABSTRACT

For the purpose of biomarker discovery, we originally developed a novel method for quantitative proteome analysis utilizing both tryptophan-targeted stable isotope tagging and mass spectrometry. The method has now been refined by replacing detergents and an enrichment column and further utilizing a novel matrix that is specifically suitable for tagged peptides. A total analytical system has been constructed by combining this method with HPLC, an automatic spotter, MALDI-TOF MS and analytical software. Clinical tissue samples such as colorectal carcinoma and renal cell carcinoma were analyzed using this system, and the results demonstrated that it is useful for discovering novel biomarker candidates. Here, we review a series of these studies and also discuss future directions for development of this technology.

Contents

1. Introduction

With the advent of whole genome sequencing of human [1–3] and other species [4–6], both transcriptome and proteome analyses have been increasingly performed to discover genes and proteins related to various biological phenomena or diseases [7–9]. Transcriptome analyses, along with the development of microarray systems, have led to the discovery of candidates responsible for diseases and have highlighted features of special cells [10,11]. However, it is often difficult to select effective biomarker proteins from the results of these experiments, because mRNA expression levels do not necessarily correlate with cellular protein abundance [12,13]. In addition, many proteins receive post-translational modifications [14] and/or processing that cannot be predicted only from genome or transcriptome information. Proteome analyses can directly identify a set of proteins whose abundance is altered, and thus this method is well suited to biomarker discovery, although the procedure is somewhat complex and sometimes time-consuming. For many years, two-dimensional gel electrophoresis (2-DE) has been applied to proteome analysis, and biomarkers have been discovered using this technology [15]. However, this approach has some experimental and operational limitations [16]. For example, higher molecular weight proteins, basic proteins, and membrane proteins are difficult to separate effectively. In addition, it is laborious to deal with many samples and it is sometimes difficult to obtain reproducible results. In the last decade, mass spectrometry (MS)-based proteome analysis has become a mainstream method as instrumental and methodological aspects have progressed [17–19], and novel methodologies utilizing stable isotope labeling and MS detection have been developed to perform quantitative proteome analysis

☆ This paper was presented at the 33rd Meeting of the Japanese Society for Biomedical Mass Spectrometry, Tokyo, Japan, 25–26 September 2008.

* Corresponding author. Tel.: +81 6 6879 4320; fax: +81 6 6879 4320.
E-mail address: osamu_nishimura@protein.osaka-u.ac.jp (O. Nishimura).

1570-0232/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi: 10.1016/j.jchromb.2009.05.049

Table 1
Comparison among isotopic-labeling methods for quantitative proteome analysis.

Methods	Labeling	Target	Protein coverage	Peptide coverage	Simplification of analysis	Applicability to samples	Advantage to use MALDI-MS	Isolation of labeled peptides	Data used for quantitation
SILAC	*In vivo*	n.a.[a]	+++[a]	+++[a]	+[a]	+	++	n.a.	MS
cICAT	*In vitro*	Cys residue	++(+)	+	++(+)	++(+)	++	+++	MS
iTRAQ	*In vitro*	Amino group	+++	+++	++	+++	++	n.a.	MS/MS
NBS	*In vitro*	Trp residue	++(+)	+	+++	++(+)	+++	++	MS

Marks indicate as follows: +++; excellent > ++(+) > ++; good > +(+) > +; poor, n.a.; not applicable. All the methods can be combined with other method and thus each evaluation can be changed.

[a] Indicate that the evaluation can be changed depending on which amino acids are isotopically labeled and which protease is used. Shown here is a typical example of using Lys and Arg for labeled amino acid and trypsin for digestion.

[20–23]. This approach has an advantage in that it can be combined with liquid chromatography and automated. Thus it is expected to have great potential for more powerful analyses of complex samples. In our efforts to discover novel biomarkers related to diseases such as cancers, we have developed a novel quantitative proteome method employing this approach [24]. We have now refined our original method [25,26] and constructed an analytical system [27]. Here, we describe the method development, establishment of the analytical system and its applications to clinical samples.

2. NBS method development

A number of methods using stable isotope labeling for quantitative proteome analysis have been developed, and three commonly used methods, as well as ours, are summarized in Table 1. These methods can be roughly classified into two categories: *in vivo* and *in vitro* labeling [21]. *In vivo* labeling techniques utilize stable isotope-labeled nutrients, *e.g.* amino acids for SILAC (stable isotope labeling with amino acids in cell culture [28]), that are metabolically incorporated into cellular proteins. The labeled nutrients are relatively inexpensive and easily used, but these techniques are limited to samples such as cultured cells. On the other hand, *in vitro* labeling techniques utilize stable isotope-labeled reagents that are bound to proteins via a chemical reaction. Therefore, they are applicable to almost all protein samples, including human tissues, and thus are matched to our purpose. In these methods, proteins are generally digested with an enzyme (endopeptidase) that cleaves peptide bonds next to specific residue(s), and then a number of peptides are generated whose lengths are desirable for MS analysis. Since cell or tissue samples are expected to contain thousands of proteins, the digests after enzymatic cleavage include tens of thousands or even more peptides. Therefore, it seems difficult to analyze all of them, although some methods, such as iTRAQ, label and target all peptides present in a given mixture [29]. Theoretically, this type of method can cover all the peptide fragments resulting from protein digests, and thus can be applied even to peptidome analysis. However, it seems advantageous to adopt a strategy where only a specific residue is labeled and the resultant labeled peptides are somehow isolated. Using this type of strategy, only part of the digest is tagged but the labeled peptides are representative of their parent proteins, allowing quantification of protein levels. We have developed a novel *in vitro* labeling method that utilizes tryptophan as a target residue [24]. Because tryptophan is the least abundant amino acid in proteins [30], isolation of tryptophan-labeled peptides reduces the number of analytes and the complexity of the entire analysis. Most proteins (>90%) in *Homo sapiens* contain at least one tryptophan residue [30], so this method is suitable for global proteome analysis. However, there is a limitation of this method: it is not suitable to peptidomic application (and sometimes to other applications), because the coverage of tryptophan labeling becomes less and less as the sequences of targets become shorter and shorter.

Several arylsulfenyl halides are known for their selective reactivity towards the indole ring of tryptophan under acidic conditions [31–33]. Some of these chemicals were tested for their reactivity, and it was found that 2-nitrobenezenesulfenyl chloride (NBSCl) effectively labeled tryptophan residues [24] (Fig. 1). NBSCl also reacts, to some extent, with sulfhydryl groups of cysteine residues. However, the resulting labeled cysteine residues, in which sulfhydryl and NBS groups are linked through disulfide bonds, are all converted to carbamidomethyl cysteins after subsequent reduction and alkylation steps [24] (Fig. 2). Thus, "tryptophan specific" labeling was finally achieved using NBSCl. We prepared a set of "heavy" and "light" NBSCls (referred to as "NBS reagent"; this reagent is commercially available from Shimadzu Corporation as "^{13}CNBS Stable Isotope Labeling Kit-N") that incorporated six ^{13}C and six ^{12}C in their benzene rings, respectively (Fig. 1). The two protein samples were then processed according to the procedure shown in Fig. 2: sample 1 was labeled with a heavy NBS reagent and sample 2 was labeled with a light NBS reagent, leading to a mass difference of 6 Da between sample 1 and sample 2 for all of the tryptophan-containing peptides. The labeled peptides were enriched by taking advantage of the relatively stronger affinity of NBS-labeled tryptophan-containing peptides for Sephadex media (LH-20) [34]. Relative quantitation of the proteins in the two samples was calculated from the intensities of paired peaks having a 6 Da mass difference in the MS spectra; proteins were then identified by a database search using queries based on data from the MS/MS spectra (Fig. 3).

Several feasibility studies were performed, demonstrating that this method is well suited to quantitative proteome analyses [24]. Basic properties of the analyses were evaluated, such as accurate quantitation, simple enrichment of labeled peptides, availability of both MALDI-TOF and ESI-MS analysis, compatibility with MS/MS analysis without any undesirable fragmentation, and co-elution of

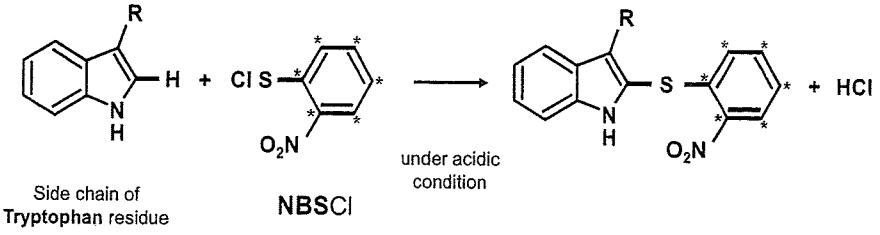


Fig. 1. Structural and reaction formula of the NBS reagent. Asterisk (*) indicates ^{12}C for the light reagent and ^{13}C for the heavy reagent.

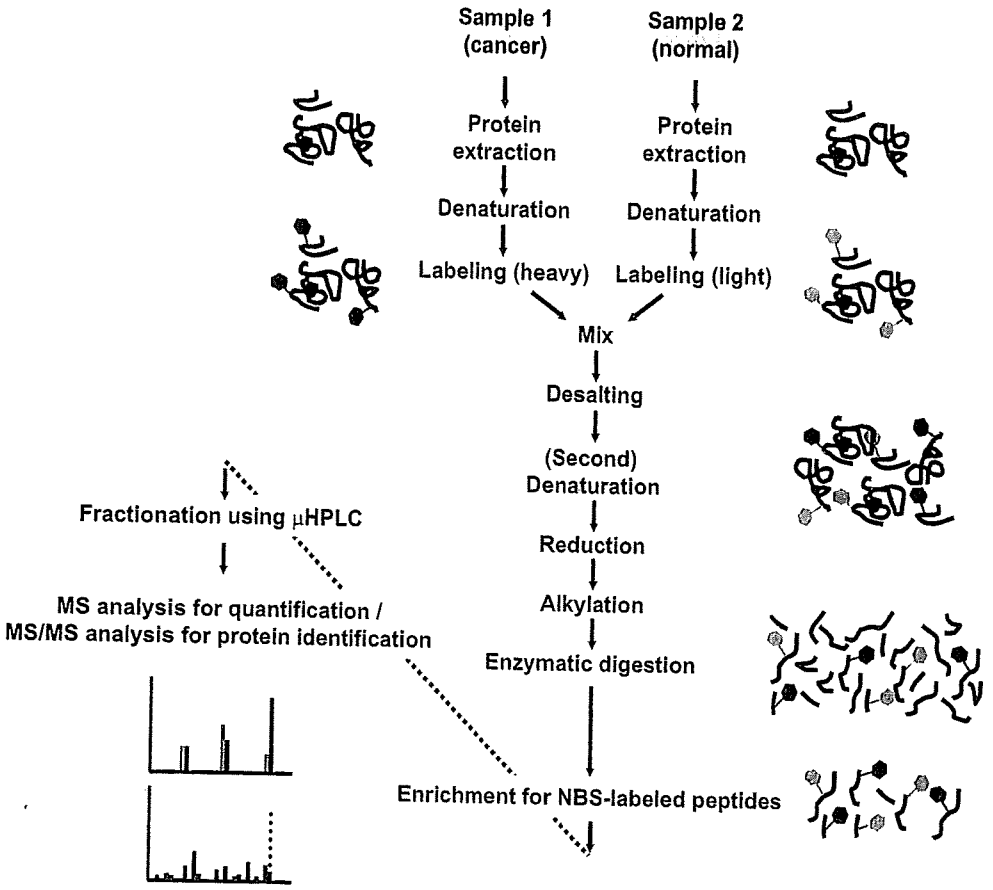


Fig. 2. The NBS method procedure is illustrated. Both proteins and peptides are indicated by black lines. Heavy and light NBS reagents are drawn as blue and red hexagons connected to the peptides with a bar, respectively.

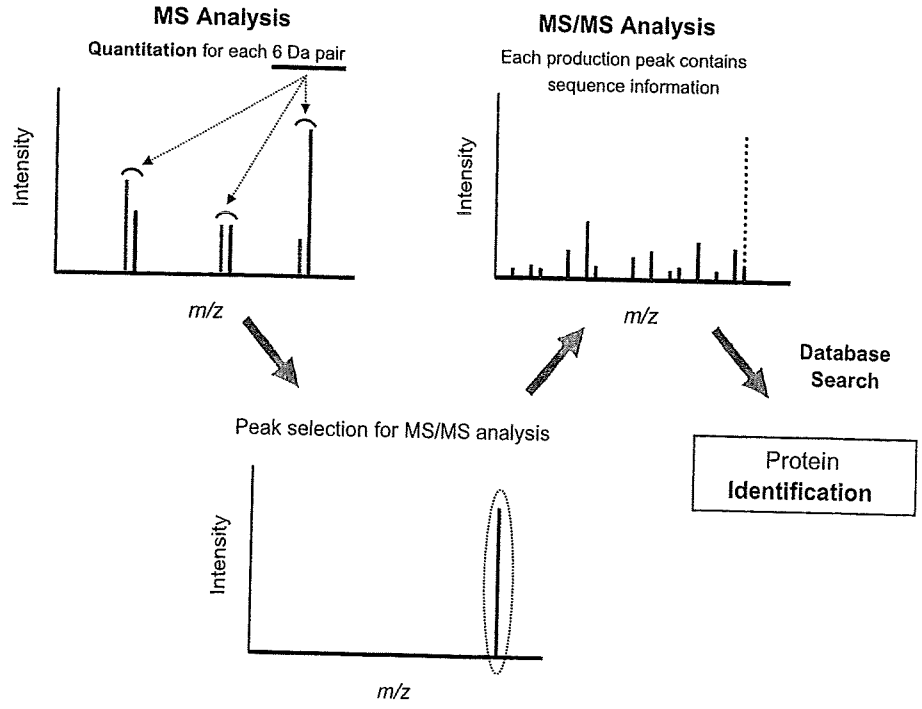


Fig. 3. MS and MS/MS analyses are illustrated. Blue and red lines indicate peaks of heavy and light NBS reagent-labeled peptides, respectively.

heavy and light labeled peptides from a C18 column by RP-LC [24]. The latter two issues were reported problems [35,36] in studies utilizing the original ICAT reagent [37]. Both the simple structure of the NBS reagent and the use of ^{13}C instead of deuterium as a stable isotope element led to solution of these problems.[1] Thus, we have developed a novel method that includes all of the basic characteristics needed for quantitative proteome analysis. We called this method the "NBS method", referring to the abbreviation of the central reagent.

3. Method optimization and improvement

Although the basic methodology was developed as mentioned in the former section, additional refinements were needed to optimize it for practical use, especially in the case of biologically derived samples consisting of complex protein mixtures. The main problems to be addressed were loss of sample, generation of by-products (molecules with mass increases of 57 Da; we assume that this was due to an unexpected alkylation (carbamidomethylation) of a side chain other than the cysteine SH group [25]), and contamination of eluted fractions with unlabeled peptides. We reviewed the entire protocol and decided to optimize the denaturing conditions and the enrichment step [25]. First of all, we used urea or guanidine hydrochloride for the denaturation step, instead of the original protocol's SDS denaturation, because they are compatible with trypsin digestion at relatively high concentrations and can be removed easily. Use of these denaturants at high concentrations is advantageous to keep proteins, including hydrophobic and membrane proteins, soluble and to avoid aggregation and/or proteolysis. Next, we used a phenyl resin instead of a Sephadex LH-20 to enrich labeled peptides, because the NBS-labeled tryptophan side chain is aromatic as well as hydrophobic, and π-electron interactions between the NBS-indole ring moiety and phenyl groups in the media should increase the specificity of the binding.

We then examined which condition is best suited for each step as well as influences of various conditions on downstream steps [25]. We finally established one optimum condition that provided several improvements: almost 100% labeling in less than 10 min, suppression of by-products (+57 Da), at least 80–90% recovery of the labeled peptides with better separation from unlabeled peptides, more accurate quantitation, and reduction of the total operation time [25]. As a result, there was minimal sample loss during the NBS reaction procedure so the sensitivity was preserved. In addition, use of the phenyl column resulted in a somewhat chromatographic separation of labeled peptides, as described below. Comparison of the original and improved protocols showed there was more than a five-fold increase in the number of observed NBS-labeled paired peaks using the improved protocol [25].

4. Discovery of a selective matrix for NBS-labeled peptides

In the NBS method, MS/MS analysis is indispensable for the identification of proteins, and we often use MALDI-IT-TOF-type MS only for this purpose. However, we had a problem here in detecting NBS-labeled peptides by this type of MS. It was generally recommended for this instrument to use a cool matrix such as 2,5-dihydroxybenzoic acid (DHB), to avoid decay of ionized molecules during ion trapping. However, DHB was found to be incompatible with the detection of NBS-labeled peptides [26]. Therefore, we searched for another cool matrix that would be suitable for detection of NBS-labeled peptides by MALDI-IT-TOF MS. After screening benzoic acid derivatives, we found two matrices, α-cyano-3-hydroxycinnamic acid (a structural isomer of conventionally used CHCA (α-cyano-4-hydroxycinnamic acid)) and 3-hydroxy-4-nitrobenzoic acid (HNBA), which satisfied our requirements [26] (Fig. 4). We then made the fortuitous discovery that the HNBA matrix has an additional special property of selectively detecting NBS-labeled peptides in mixtures of labeled and unlabeled peptides [26]. The mechanism behind this selectivity is that the sensitivity of the HNBA matrix in detecting labeled peptides is similar to that of CHCA, whereas its sensitivity for detecting unlabeled peptides is greatly decreased, compared to CHCA. Thus, the HNBA matrix preferentially detects labeled peptides, for example from samples containing unlabeled impurities. This matrix possesses yet another favorable property, in that it suppresses fragmentations (mainly −16 and −32 Da species that result from detachment of oxygen(s)) [26], which were known to occur during the MALDI-TOF MS measurement of nitrobenzene compounds [38,39]. We investigated the mechanisms underlying these phenomena and found that various nitrobenzene compounds showed a similar effect, although the detection sensitivities of these matrices were much lower than that of HNBA [26]. We defined a "selectivity index" to indicate the size ranges over which labeled peptides are detected preferentially over unlabeled peptides by comparing results obtained using a given matrix to results obtained using a conventional CHCA matrix [26]. The selectivity index for the HNBA matrix was as high as 10. Interestingly, that of the original matrix for MALDI-IT-TOF MS, DHB, is about 0.1.

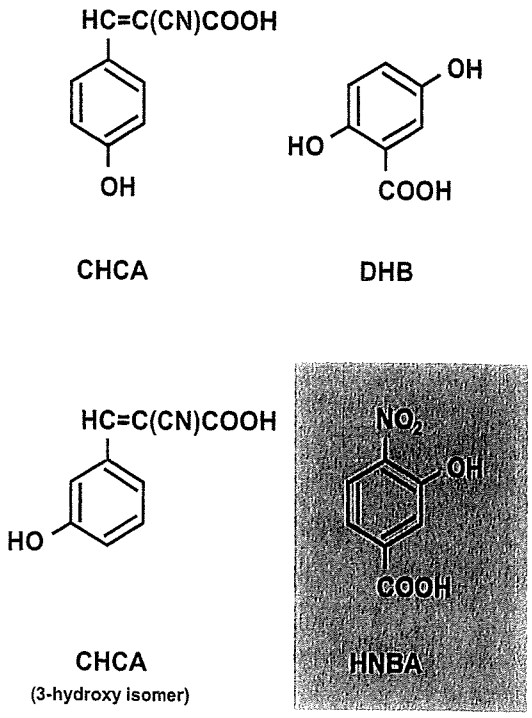


Fig. 4. Structural formulas of the four matrices, CHCA, DHB, CHCA (3-hydroxy isomer) and HNBA.

As mentioned above, the HNBA matrix has quite unique and favorable features for the detection of NBS-labeled peptides. However, it is less usable due to unstable signal detection and rapid signal decay, compared to conventionally used CHCA [26]. These drawbacks were compensated by the combined use of HNBA and CHCA: the usability was greatly increased with a minor loss of selectivity [26]. Surprisingly, in addition to this, the sensitivity in detecting labeled peptides was increased about four-fold [26]. The idea of using two matrices as a co-matrix was very simple, but the practical and beneficial effects have become significant in this

[1] In the case of the ICAT reagent, these issues were already solved by the use of ^{13}C and by removal of the cleavable tag introduced into the improved reagent, cICAT [36].

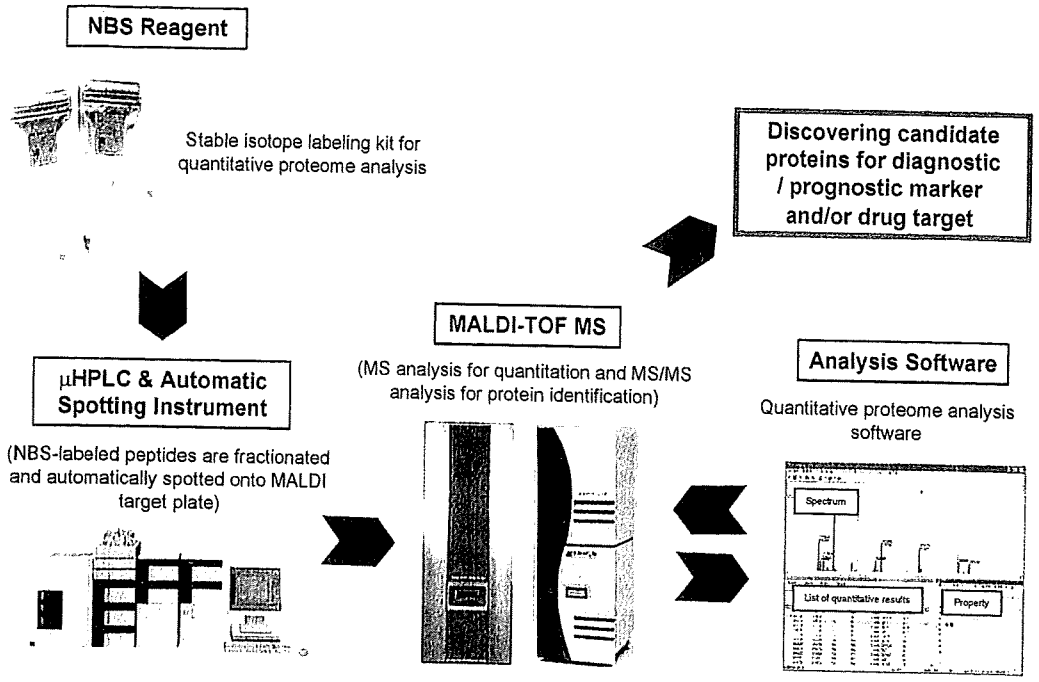


Fig. 5. Total analytical system using the NBS method.

case. In conclusion, discovery of HNBA matrix and its incorporation into an HNBA–CHCA co-matrix system increased the usability and detection sensitivity of target signals, thus increasing the dynamic range of the NBS detection method.

5. Establishment of an analytical system

Aside from the improvement and refinement of the NBS method, construction of an analytical system is also essential to our goal. Although the NBS method reduces much of the sample complexity, typical eluates from the enrichment step still contain many thousands of peptides, and fractionation by HPLC is indispensable for wider coverage and accurate quantitation. Systematic quantitation and protein identification are also desirable.

As mentioned above, the NBS method can be applied to both ESI-MS and MALDI-TOF MS. ESI-MS can be connected "on-line" with HPLC and run cooperatively and automatically. Recent advances in ESI-MS (with regard to both hardware and software) have resulted in fast scan speeds and an improved duty cycle, which can maximize the higher-resolution separation of peptides achieved by HPLC. MALDI-TOF MS can also be connected "off-line" with HPLC, but a droplet spotted onto a well of the MALDI plate is corresponding to a number of ESI-MS scans. From this point of view, LC/on-line ESI-MS-based analysis is superior to LC/off-line MALDI-MS-based analysis. However, in ESI-MS analysis, it is still the case that only part of the eluate is applied to MS or MS/MS analysis, and decisions as to which peaks should be analyzed by MS/MS are made during the continuous flow from HPLC. Sprayed samples cannot be measured again. On the other hand, in MALDI MS analysis, MS/MS analyses can be performed after the HPLC separation and MS analysis, by referring to the quantitation result from MS analysis. In addition, data can be accumulated using a target plate upon which all eluates from HPLC are deposited. This should result in higher sensitivity for protein identification. Here, we preserved all of the advantages of MALDI MS analysis, including the existence of the HNBA matrix, and constructed an LC/off-line MALDI-TOF MS system [27] (Fig. 5). First, an NBS-labeled peptide sample was applied to μHPLC and separated on a capillary ODS (C18) column. Each fraction (drop) of eluate was automatically deposited onto a MALDI target plate by a spotting apparatus, followed by automatic MS analysis. Next, the NBS-labeled paired peaks with significant differences in their intensities were selected and subjected to MS/MS analysis to identify their sequences and parent proteins. Operational conditions throughout these steps were optimized to construct an efficient system. For example, an eluate from a phenyl column was divided into three fractions and each was separately applied to an ODS column [27]. Both ODS and phenyl columns belong to the same reversed-phase but they exploit different interactions. Our results showed that only 10–20% of the peaks overlapped between neighboring fractions. Thus, a simple fractionation on a phenyl column increased the total number of peptides detected. Analytical software was also developed and incorporated into this system. This enables selection of all paired peaks with 6 m/z (and 12 m/z) differences for quantitation, and it is also easy to filter peaks with relatively large differences in their peak intensities. In this way, a total analytical system was constructed and then validated using rat and mouse sera [27]. Three sets of rat and mice sera, each paired samples from normal and diseased animals, were examined in order to evaluate this new analytical system. In all three experiments, 1000–2000 paired peaks were detected, and 32 pairs were selected by the software as differentially expressed protein tags with more than three-fold differences in expression [27]. Less than 100 paired peaks were detected without HPLC separation, clearly demonstrating that the system functions effectively for global proteome analysis [27].

In conclusion, we have constructed an analytical system suitable for the NBS method with off-line LC–MALDI-TOF MS. Using this system, even low-intensity peaks from proteins with a relatively low abundance may be identified and analyzed, for example when they are differentially expressed in two samples.

6. Applications to clinical samples

Our final purpose in developing this quantitative proteome analysis system is to discover novel biomarkers (and drug target proteins), because there is certainly a need to find practical biomarkers for clinical uses, such as early disease detection, diagnosis, prognosis, imaging and so on. There are still no diagnostic markers for many diseases [40], and in other cases specific

Table 2
Summary of the two application studies using human clinical samples.

	Analyzed specimens	Detected paired peaks per specimen	Differential paired peaks selected for MS/MS analysis	Identified peptides	Identified proteins	Reported earlier	Novel
Colorectal carcinoma (CRC)	12	~5000	320	138	128	30	**98**
Renal cell carcinoma (RCC)	14	6000–7000	225	108	92	24[a]	**68**

Identified proteins are classified as "Reported earlier" and "Novel", and the numbers of the latter are highlighted in bold.
[a] Includes two proteins which are discordant in their up/down-regulated states with our results [40].

biomarkers exist but are effective only in advanced disease cases [41]. Here, we applied the analytical system described above to surgically resected specimens from colorectal carcinoma (CRC) [41] and renal cell carcinoma (RCC) [40], and the protein expression profiles of cancerous and normal parts were compared. The results are summarized in Table 2. In both cases, about 200–300 paired peaks were selected as having significantly different expression levels, and as occurring with sufficient frequency among patients, and roughly 100 of these peaks were identified [40,41]. About 20–30% of the identified proteins had been reported in earlier studies, and the remaining 70–80% were newly found to be associated with the corresponding cancer [40,41]. The results showed that this analytical system is reliable as well as quite useful to discover novel biomarker proteins. Compared to earlier proteome studies using a 2D-gel method, higher molecular weight proteins and basic proteins were predominantly identified in our method. More precisely, 17 proteins with molecular weight (MW) larger than 100 kDa and six proteins with p*I* > 10 were identified among 128 proteins in our CRC analysis [41], whereas only two proteins with MW > 100 kDa and no proteins with p*I* > 10 among 168 proteins were reported in two earlier CRC studies using a 2D-gel method [42,43] (Fig. 6). This illustrates the advantage of our method compared to earlier methods. Several proteins were further verified and validated by Western blotting, RT-PCR and immunohistochemical (IHC) staining.

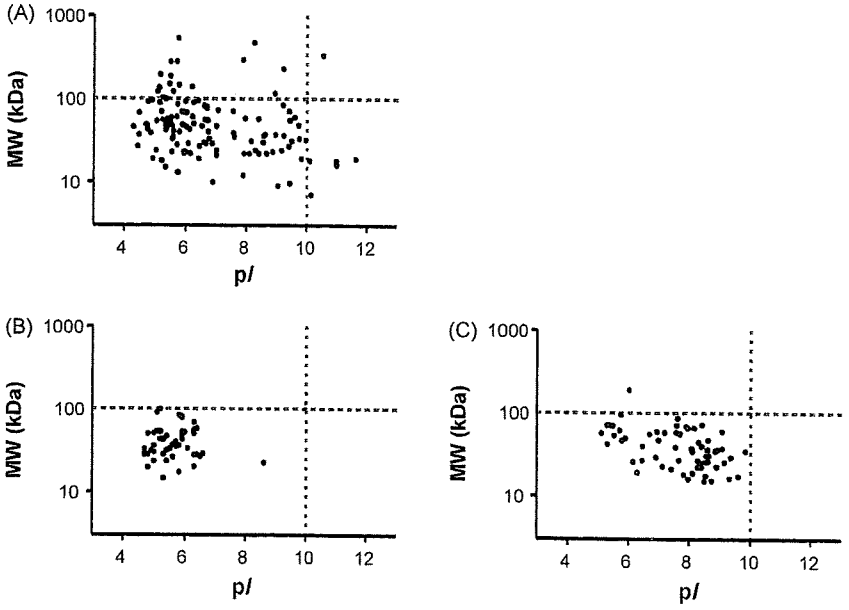


Fig. 6. Proteins identified by each method were plotted according to their predicted p*I* and molecular weight (MW). (A) NBS method [41], (B, C) 2D-gel method [42,43]. *y*-axis is on a logarithmic scale. Areas corresponding to MW larger than 100 kDa and/or p*I* > 10 are highlighted in yellow.

Heavy-labeled

	Sprot	Decoy	False discovery rate
Peptide matches above identity threshold	43	0	0.00 %
Peptide matches above homology or identity threshold	46	0	0.00 %

Light-labeled

	Sprot	Decoy	False discovery rate
Peptide matches above identity threshold	65	0	0.00 %
Peptide matches above homology or identity threshold	66	0	0.00 %

Fig. 7. Confirmation of the reliability of protein identification. The data used for protein identification were divided into two groups (72 for light-labeled and 66 for heavy-labeled) and analyzed again by MASCOT MS/MS Ions Search using both the usual SwissProt database and a decoy database. Search parameters used are as follows: trypsin digestion allowing up to 1 missed cleavage, fixed modifications of NBS (W) (or NBS:13C(6) (W)) and carbamidomethyl (C), variable modifications of oxidation (M), peptide tolerance 0.3 Da, MS/MS tolerance of 0.5 Da, and restriction to peptides with sequences containing one or more tryptophan residues.

Good reproducibility of this system was verified in the above CRC study; correlation coefficients between the first and second experiments were over 0.95 for all 12 specimens [41]. In addition, reduction of analysis complexity by avoiding redundant identification of peptides from the same protein was demonstrated; the number of identified peptides was very close to the number of identified proteins (Table 2). Such accuracy and simplification of analysis are the essence of this method. In terms of protein identification, one may suspect some of the search results, because most of the proteins were identified by only one peptide hit. However, the reliability of the protein identification was confirmed by the absence of false-positive identification ($p < 0.05$) using a decoy database (Fig. 7).

Some differences between the results obtained from the NBS proteome analysis and Western blotting analysis were attributed primarily to differences between the methods [41]; the NBS method is quite precise but only detects tagged peptides derived from whole proteins, whereas Western blotting detects entire proteins, but with less resolution and specificity, depending on the properties of the antibodies utilized. It is necessary to combine NBS analysis with IHC staining, as NBS analysis alone provides no information about the localization of detected proteins. Therefore, combination of the NBS method with other complementary analyses is needed to extract the maximum amount of information from the obtained data. It is important to fractionate samples before NBS analysis to increase the number and dynamic range of detected peptides [41]. This approach has been described in a paper on serum glycoproteome profiling in lung cancer [44]. Here, utilization of the NBS method combined with immunodepletion of six abundant proteins and lectin column selection led to successful detection of interleukin-12 (IL-12), which is an extremely low-abundant protein in serum. The NBS method has also been effectively used in combination with the regular 2-DE method to find drug-responsive proteins using a breast cancer cell line [45]. In this report, fine discrimination and accurate quantitation of two proteins that co-migrated as one spot were demonstrated by using the NBS method. When a conventional 2-DE/MS method was used, it was difficult to evaluate this spot as two proteins and thus the ratio of protein abundance was also reported incorrectly.

7. Conclusions and future aspects

Compared to other quantitative proteome methods, the NBS method has a unique aspect of tryptophan tagging, which is combined with an analytical system and optimized to detect less abundant proteins. We believe that this system has potential to discover novel disease-related proteins; this has already been achieved to a certain degree, as described in the previous section. However, the proteins identified here are just biomarker "candidates". They are now under evaluation for possible use as clinical diagnostic markers, using sera from both cancer patients and healthy volunteers. There are numerous such disease-related candidates awaiting further validation [46]. We must consider by what means and how to validate these biomarker candidates and determine how they will be used. Detailed analysis of each protein identified is needed as well. ELISA (enzyme-linked immunosorbent assay) is one of the most promising and powerful techniques used to screen for biomarker availability [47]. It has been used widely and intensively, but it is time-consuming and costly to establish reliable systems. Development of a novel technique with both sensitivity and specificity, either alone or in combination with other techniques, could offer valuable shortcuts. For example, a multiplex protein detection method with high sensitivity and specificity was used for biomarker validation: it combined oligonucleotide primer-tagged antibodies with real-time PCR and DNA manipulation techniques [48]. Another promising MS measuring technique is multiple reaction monitoring (MRM); this has been increasingly used for data validation and is now being applied to clinical diagnoses [49,50]. The utilization of MS for these types of applications, as well as for biomarker candidate discovery, will open up new possibilities for clinical applications.

In any case, our primary goal is to discover practical biomarkers for clinical applications from the lists of candidate proteins identified using the NBS method. Achievement of this goal will constitute an important contribution to human welfare.

Acknowledgements

We would like to express our appreciation for Dr. Morito Monden, Dr. Ichiro Takemasa, Dr. Katsuhiko Okumura and Dr. Noboru Okamura for providing us with clinical specimens. We also thank Dr. Susumu Tsunasawa, Dr. Susumu Iwasa, Dr. Koichi Tanaka and Dr. Eiji Ando for technical advice and useful discussions throughout this project.

References

[1] International Human Genome Sequencing Consortium, Nature 409 (2001) 860.
[2] J.C. Venter, et al., Science 291 (2001) 1304.
[3] T. Ota, et al., Nat. Genet. 36 (2004) 40.
[4] Mouse Genome Sequence Consortium, Nature 420 (2002) 520.
[5] T.K. Watanabe, et al., Nat. Genet. 22 (1999) 27.
[6] Arabidopsis Genome Initiative, Nature 408 (2000) 796.
[7] J.A. Olson Jr., Surgery 136 (2004) 519.
[8] M. Ohira, S. Oba, Y. Nakamura, E. Isogai, S. Kaneko, A. Nakagawa, T. Hirata, H. Kubo, T. Goto, S. Yamada, Y. Yoshida, M. Fuchioka, S. Ishii, A. Nakagawara, Cancer Cell 7 (2005) 337.
[9] Y. Kuramitsu, K. Nakamura, Proteomics 6 (2006) 5650.
[10] S. Abe, T. Katagiri, A. Saito-Hisaminato, S. Usami, Y. Inoue, T. Tsunoda, Y. Nakamura, Am. J. Hum. Genet. 72 (2003) 73.
[11] K. Takahashi, S. Yamanaka, Cell 126 (2006) 663.
[12] S.P. Gygi, Y. Rochon, B.R. Franza, R. Aebersold, Mol. Cell. Biol. 19 (1999) 1720.
[13] G. Chen, T.G. Gharib, C.-C. Huang, J.M.G. Taylor, D.E. Misek, S.L.R. Kardia, T.J. Giordano, M.D. Iannettoni, M.B. Orringer, S.M. Hanash, D.G. Beer, Mol. Cell. Proteomics 1 (2002) 304.
[14] M. Mann, O.N. Jensen, Nat. Biotechnol. 21 (2003) 255.
[15] H.J. Issaq, T.D. Veenstra, Electrophoresis 28 (2007) 1980.
[16] T. Rabilloud, Proteomics 2 (2002) 3.
[17] B. Domon, R. Aebersold, Science 312 (2006) 212.
[18] R.J. Cotter, W. Griffith, C. Jelinek, J. Chromatogr. B 855 (2007) 2.
[19] C. Fenselau, J. Chromatogr. B 855 (2007) 14.
[20] R. Aebersold, M. Mann, Nature 422 (2003) 198.
[21] S. Sechi, Y. Oda, Curr. Opin. Chem. Biol. 7 (2003) 70.
[22] S. Julka, F. Regnier, J. Proteome Res. 3 (2004) 350.
[23] S.-E. Ong, M. Mann, Nat. Chem. Biol. 1 (2005) 252.
[24] H. Kuyama, M. Watanabe, C. Toda, E. Ando, K. Tanaka, O. Nishimura, Rapid Commun. Mass Spectrom. 17 (2003) 1642.
[25] E. Matsuo, C. Toda, M. Watanabe, T. Iida, T. Masuda, T. Minohata, E. Ando, S. Tsunasawa, O. Nishimura, Rapid Commun. Mass Spectrom. 20 (2006) 31.
[26] E. Matsuo, C. Toda, M. Watanabe, N. Ojima, S. Izumi, K. Tanaka, S. Tsunasawa, O. Nishimura, Proteomics 6 (2006) 2042.
[27] T. Iida, H. Kuyama, M. Watanabe, C. Toda, E. Matsuo, A. Kido, E. Ando, S. Tsunasawa, O. Nishimura, J. Biomol. Tech. 17 (2006) 333.
[28] S.-E. Ong, M. Mann, Nat. Protocols 1 (2006) 2650.
[29] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin, Mol. Cell. Proteomics 3 (2004) 1154.
[30] G. Cagney, S. Amiri, T. Premawaradena, M. Lindo, A. Emili, Proteome Sci. 1 (2003) 5.
[31] E. Scoffone, A. Fontana, R. Rocchi, Biochem. Biophys. Res. Commun. 25 (1966) 170.
[32] E. Scoffone, A. Fontana, R. Rocchi, Biochemistry 7 (1968) 971.
[33] A. Fontana, E. Scoffone, C.A. Benassi, Biochemistry 7 (1968) 980.
[34] T. Sasagawa, K. Titani, K.A. Walsh, Anal. Biochem. 134 (1983) 224.
[35] R. Zhang, C.S. Sioma, S. Wang, F.E. Regnier, Anal. Chem. 73 (2001) 5142.
[36] K.C. Hansen, G. Schimitt-Ulms, R.J. Chalkley, J. Hirsch, M.A. Baldwin, A.L. Burlingame, Mol. Cell. Proteomics 2 (2003) 299.
[37] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[38] A.-S. Petersson, H. Steen, D.E. Kalume, K. Caidahl, P. Roepstorff, J. Mass Spectrom. 36 (2001) 616.
[39] M. Strohalm, M. Kodíček, M. Pechar, Biochem. Biophys. Res. Commun. 312 (2003) 811.
[40] N. Okamura, T. Masuda, A. Gotoh, T. Shirakawa, S. Terao, N. Kaneko, K. Suganuma, M. Watanabe, T. Matsubara, R. Seto, J. Matsumoto, M. Kawakami, M. Yamamori,

T. Nakamura, T. Yagami, T. Sakaeda, M. Fujisawa, O. Nishimura, K. Okumura, Proteomics 8 (2008) 3194.
[41] M. Watanabe, I. Takemasa, N. Kawaguchi, M. Miyake, N. Nishimura, T. Matsubara, E. Matsuo, M. Sekimoto, K. Nagai, N. Matsuura, M. Monden, O. Nishimura, Proteomics Clin. Appl. 2 (2008) 925.
[42] D.B. Friedman, S. Hill, J.W. Keller, N.B. Merchant, S.E. Levy, R.J. Coffey, R.M. Caprioli, Proteomics 4 (2004) 793.
[43] X. Bi, Q. Lin, T.W. Foo, S. Joshi, T. You, H.-M. Shen, C.N. Ong, P.Y. Cheah, K.W. Eu, C.-L. Hew, Mol. Cell. Proteomics 5 (2006) 1119.
[44] K. Ueda, T. Katagiri, T. Shimada, S. Irie, T. Sato, Y. Nakamura, Y. Daigo, J. Proteome Res. 6 (2007) 3475.
[45] K. Ou, D. Kesuma, K. Ganesan, K. Yu, S.Y. Soon, S.Y. Lee, X.P. Goh, M. Hooi, W. Chen, H. Jikuya, T. Ichikawa, H. Kuyama, E. Matsuo, O. Nishimura, P. Tan, J. Proteome Res. 5 (2006) 2194.
[46] M. Polanski, N.L. Anderson, Biomarker Insights 1 (2006) 1.
[47] R.C. Zangar, D.S. Daly, A.M. White, Expert Rev. Proteomics 3 (2006) 37.
[48] S. Fredriksson, W. Dixon, H. Ji, A.C. Koong, M. Mindrinos, R.W. Davis, Nat. Methods 4 (2007) 327.
[49] N. Rifai, M.A. Gillette, S.A. Carr, Nat. Biotechnol. 24 (2006) 971.
[50] G. Duranti, S. Boenzi, C. Rizzo, L. Ravà, V.D. Ciommo, R. Carrozzo, M.C. Meschini, D.W. Johnson, C. Dionisi-Vici, Clin. Chim. Acta 398 (2008) 86.

Proteomics 2009, *9*, 4063–4070 DOI 10.1002/pmic.200900267

RESEARCH ARTICLE

A new approach for detecting C-terminal amidation of proteins and peptides by mass spectrometry in conjunction with chemical derivatization

Hiroki Kuyama[1], *Chihiro Nakajima*[1,2], *Takashi Nakazawa*[3], *Osamu Nishimura*[1] *and Susumu Tsunasawa*[1]

[1] Institute for Protein Research, Osaka University, Suita, Osaka, Japan
[2] Technology Research Laboratory, Shimadzu Corporation, Kyoto, Japan
[3] Department of Chemistry, Nara Women's University, Nara, Japan

We describe a mass spectrometric method for distinguishing between free and modified forms of the C-terminal carboxyl group of peptides and proteins, in combination with chemical approaches for the isolation of C-terminal peptides and site-specific derivatization of the C-terminal carboxyl group. This method could most advantageously be exploited to discriminate between peptides having C-terminal carboxyl groups in the free (COOH) and amide ($CONH_2$) forms by increasing their mass difference from 1 to 14 Da by selectively converting the free carboxyl group into methylamide ($CONHCH_3$). This method has been proven to be applicable to peptides containing aspartic and glutamic acids, because all the carboxyl groups except the C-terminal one are inert to derivatization, according to oxazolone chemistry. The efficiency of the method is illustrated by a comparison of the peaks of processed peptides obtained from a mixture of adrenomedullin, calcitonin, and BSA. Among these components of the mixture, only the C-terminal peptide of BSA exhibited the mass shift of 13 Da upon treatment, eventually unambiguously validating the C-terminal amide structures of adrenomedullin and calcitonin. The possibility of extending this method for the analysis of C-terminal PTMs is also discussed.

Received: April 26, 2009
Revised: May 19, 2009
Accepted: May 27, 2009

Keywords:
Mass spectrometry / Post-translational modification / Protein sequencing / Technology

1 Introduction

One of the most important subjects in proteomics is to identify PTMs characterizing a majority of mature proteins, which are biological entities having various cellular functions that are not anticipated by genomic information. Owing to its rapidly improving performance, mass spectrometry has been the method of choice for protein/peptide analysis. The development of experimental techniques for protein sequencing by mass spectrometry has led to the discovery of a variety of PTMs occurring at the N-termini of proteins, as well as in amino acid side chains. For example, it has been estimated that 30–80% of mature proteins in eukaryotes are subjected to N-terminal acetylation [1, 2]. In contrast, the C-terminal carboxyl group is usually free from such modifications with a very few exceptions, which include amidation in relatively small peptide hormones [3] and derivatization with glycosyl phosphatidylinositol for anchoring the modified proteins to cell membranes [4–6]. In particular, C-terminal amidation is very difficult to detect because the mass difference between the carboxyl and carboxamide groups is only 1 Da. A most unfortunate situation could arise when a protein has asparagine (Asn) or

Correspondence: Dr. Hiroki Kuyama, Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan
E-mail: kuyama@protein.osaka-u.ac.jp
Fax: +81-6-6879-4320

Abbreviations: Asn, asparagine; **Gln**, glutamine; **LysC**, lysyl endopeptidase; **MDPNA**, methanediphosphonic acid

glutamine (Gln) at the C-terminus, whereby it is impossible to discriminate between the corresponding amide forms of aspartic acid (Asp-NH_2) and glutamic acid (Glu-NH_2), unless an additional chemical or enzymatic approach is implemented. In view of such a technical difficulty in detection, it might not be too suspicious to speculate that at least a small proportion of mature proteins could bear C-terminal modification without having been identified as yet.

As is the general case with PTMs, a protein presumed to have a particular type of C-terminal modification can be isolated or enriched by a specifically designed scheme based on either affinity chromatography or enzyme treatment. The RIA is widely used for the detection and quantification of C-terminal amidated peptides, which amount to more than half of the known neural and endocrine peptides [7]. Immunoprecipitation targeting of bioactive peptides with particular amino acid sequences including the C-terminal amide group has made it possible to identify two novel amidated peptides [8]. However, a higher specificity of the method for analyzing the C-terminal modification imposes a narrower range of target proteins, thus severely limiting the applicability of the method. Probably, the most reliable method for distinguishing between free and modified carboxyl groups at the C-terminus of a protein is by a chemical or enzymatic approach.

We suggest here a chemical method for detecting PTMs at the C-termini of peptides and proteins, utilizing a technique for the isolation of the C-terminal peptides [9] in combination with a reaction for selectively derivatizing the C-terminal carboxyl group according to oxazolone chemistry [10, 11]. The scheme for C-terminal derivatization is specifically designed to convert the free C-terminal carboxyl group into methylamide ($CONHCH_3$), increasing the mass of the peptide by 13 Da, while peptides possessing the amide ($CONH_2$) group remain unchanged (Fig. 1). Because an oxazolone is formed solely at the C-terminus, the existence of free carboxyl side chains of Asp and Glu does not affect the performance of the present method. The relatively low yield of the reaction conducted in the manner of the original protocol [11] is improved to the extent that the methylamidated peptide can be detected as a MALDI peak appearing at a higher mass value by 13 Da than that of the original peptide, with more than comparable intensity in the mass spectrum. This should make it easier to distinguish between peptides with the free C-terminal carboxyl group and those with the post-translationally modified C-terminal carboxyl group, based on the presence or absence of an additional MALDI peak.

In this report, we present a new approach for detecting C-terminal amides in proteins and peptides by mass spectrometry.

2 Materials and methods

BSA, acetic anhydride (Ac_2O), 40% methylamine solution (w/w, aqueous), and iodoacetamide were obtained from Sigma-Aldrich (St. Louis, MO, USA). Human adrenomedullin, human calcitonin, VYIHPF, and WMDF-NH_2 were purchased from the Peptide Institute (Osaka, Japan). Tris-(2-carboxyethyl)phosphine hydrochloride and succinimidyloxycarbonylmethyl Tris(2,4,6-trimethoxyphenyl)phosphonium bromide (TMPP-Ac-OSu) were obtained from Fluka (Switzerland). Lysyl endopeptidase (LysC), ACN, 2-propanol, formic acid, pentafluorophenol (Pfp-OH), and TFA were purchased from Wako Pure Chemical Industries (Osaka, Japan). CHCA (high-purity mass-spectrometric grade) was obtained from Shimadzu GLC (Tokyo, Japan). TNVGSEAF-NH_2 and PVTL-OCH_3 were obtained from Bachem AG (Switzerland). HPTFD-NH_2, SFLLRN-OH,

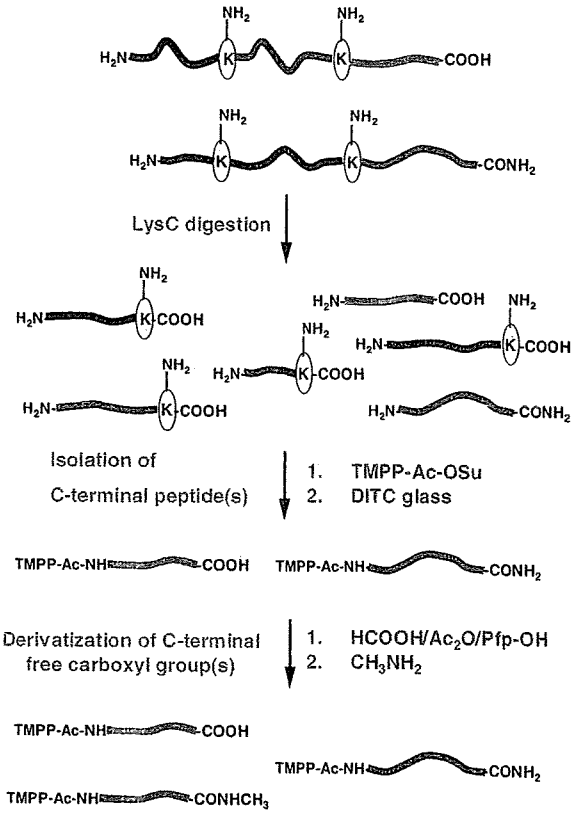


Figure 1. Scheme depicting the isolation of C-terminal peptides and subsequent procedure to distinguish between peptides with C-terminal amide ($CONH_2$) as an example of PTM and those with the ordinary free carboxyl group (COOH). C-Terminal peptides of proteins are prepared, as reported previously [9]. At the last step of the procedure, C-terminally modified peptides are inert to the reaction to derivatize the C-terminal-free carboxyl group, which is selectively converted to the methylamide according to the oxazolone chemistry [11]. The difference in the chemical form of C-termini can be discerned by the presence of or absence of mass shift resulting from the derivatization. Owing to the general incompleteness of the derivatization of the free C-terminal carboxyl group with methylamine, the resulting preparation exhibits a doublet peaks in mass spectrum separated by 13 Da corresponding to the difference in mass of COOH and $CONHCH_3$ groups.

TTSFAE-NH_2, and ALEGSLQ-OH were purchased from Operon Biotechnologies, K.K. (Tokyo, Japan). The *p*-phenylenediisothiocyanate glass used in this study was prepared in house according to the method of Wachter *et al.* [12], using aminopropyl glass (average pore size 170 Å, 200–400 mesh, amine content: 162 μmol/g) obtained from Sigma. Methanediphosphonic acid (MDPNA) was obtained from Tokyo Chemical Industry (Tokyo, Japan). Water used in all the experiments was purified using a MilliQ water purification system. All other chemicals were of analytical reagent grade and were used without further purification.

2.1 C-Terminal amidation

C-Terminal peptides (20 pmol) were prepared from proteins or peptides, as reported previously [9]. The isolated TMPP-Ac-modified C-terminal peptides were purified with ZipTip

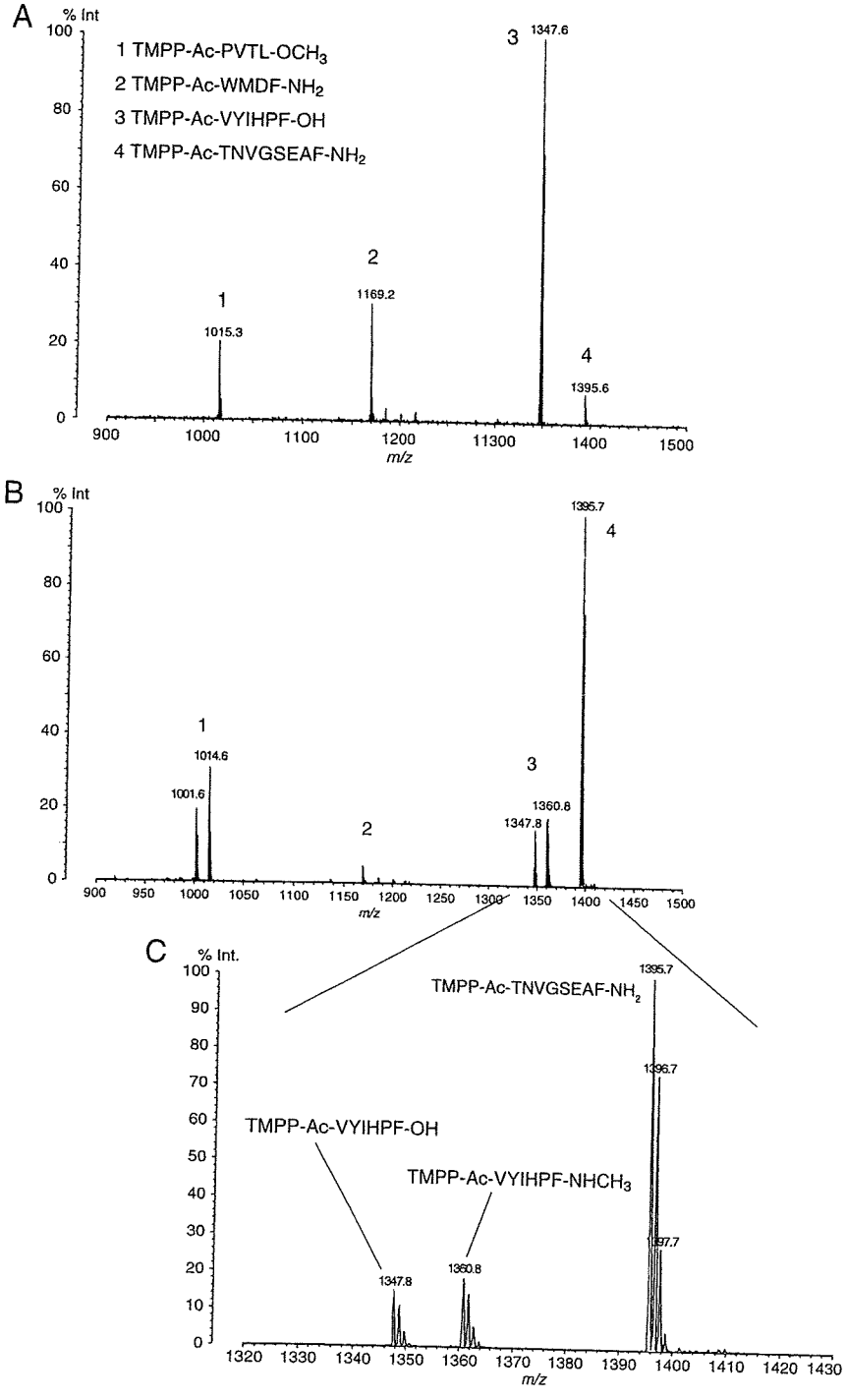


Figure 2. MALDI-TOF mass spectra of a four-peptide mixture (PVTL-OCH_3, WMDF-NH_2, VYIHPF-OH, and TNVGSEAF-NH_2): (A) TMPP-Ac modification; (B) methylamidation; and (C) enlarged view of peptides 3 and 4 after methylamidation: peptide 3 (TMPP-Ac-VYIHPF-OH) having free C-terminal carboxyl group is converted to its methylamide (TMPP-Ac-VYIHPF-$NHCH_3$).

and dried using a vacuum concentrator. Next, 12 μL of HCOOH–Ac_2O–Pfp-OH mixture (1:1:1 in volume) was added to the resulting dried peptides, and the mixture was incubated for 20 min at 60°C, followed by evaporation to dryness using a vacuum concentrator. The dried material was dissolved with 14 μL of a 5:2-mixture of toluene and Pfp-OH, and dried again. Ten microliter of 40% (w/w, aqueous) CH_3NH_2-ACN (1:9) was then added to the resulting peptides; the mixture was subsequently sonicated for 10 min and left standing for 20 min. For mass analysis, an aliquot (1 μL) was acidified with aqueous TFA solution and purified with ZipTip.

2.2 MALDI-TOF MS

MALDI mass spectra were recorded on AXIMA CFR-plus or AXIMA TOF^2 (SHIMADZU/KRATOS, Manchester, UK) reflectron time-of-flight mass spectrometers equipped with a nitrogen laser (337 nm, 3 ns pulse width). All measurements were performed in the positive-ion reflectron mode. The ion acceleration voltage was set to 20 kV, and the reflectron detector was operated at 24 kV. The flight path in the reflectron mode is 240 cm for both instruments. For the MS/MS experiments, CID was carried out using helium at a pressure of ca. 5×10^{-6} mbar in the collision cell.

The matrix used in this experiment was CHCA, which was dissolved to saturation in 50% aqueous ACN containing 0.05% TFA. We used MDPNA, which has been proven useful for MALDI analysis of salt-containing samples, as a matrix additive [13]. MDPNA was used as 1–2% aqueous solution. An aliquot (0.4 μL) of the sample solution was mixed with an equivalent volume of matrix solution and matrix additive solution on the MALDI target plate and analyzed after drying.

The *m/z* values in the spectra were externally calibrated with angiotensin II (human) and ACTH fragment 18–39 (human) using CHCA as a matrix.

3 Results and discussion

3.1 Amidation of model peptides

C-Terminal-specific amidation reaction and its side reaction(s) were investigated using peptides VYIHPF-OH, TNVGSEAF-NH_2, WMDF-NH_2, and PVTL-OCH_3. A mixture of four peptides (20 pmol each) was processed as follows: (i) N^{α}-modification with TMPP-Ac-OSu (Fig. 2A), (ii) selective activation of C-terminal carboxyl group to the active ester through oxazolone, and (iii) amidation with methylamine (Fig. 2B). In principle, this process should cause peptides with the free C-terminal carboxyl group to increase in mass by 13 Da, while causing no change in mass for those modified at the C-termini. Owing to the incomplete C-terminal amidation, a new peak of TMPP-Ac-VYIHPF-$NHCH_3$ arose at the higher mass value by 13 Da, along with the original one for TMPP-Ac-VYIHPF-OH (Fig. 2C). The formation of TMPP-Ac-VYIHPF-$NHCH_3$ was confirmed by *de novo* sequencing of the resulting peak by MS/MS in CID mode (Fig. 3). As expected, the C-terminally amidated peptide (TMPP-Ac-TNVGSEAF-NH_2) remained intact through C-terminal activation and methylamidation (Fig. 2C).

In the mass spectrum of peptide TMPP-Ac-WMDF-NH_2 containing aspartic acid along with the C-terminal amide and having been submitted to the reaction, no peak indicating amidation of the side-chain carboxyl group of aspartic acid appeared. As is the case with aspartic acid, the side-chain carboxyl group of glutamic acid has also proved to be virtually inert to methylamidation in this protocol (TMPP-Ac-TNVGSEAF-NH_2 in Fig. 2C). Thus, the C-terminal and side-chain carboxyl groups have been distinguished successfully.

A few marginal peaks of by-products or artifacts were detected, while this peptide has no C-terminal carboxyl group to be methylamidated (Fig. 2B). Observation of peaks at *m/z* 1185.5 (+16 Da) and *m/z* 1201.6 (+32 Da) is consistent with the possible oxidation of methionine to its sulfoxide and sulfone forms. Although tryptophan is also susceptible to oxidation, the appearance of peaks at *m/z* 1213.6 (+16+28 Da) and *m/z* 1227.5 (+16 × 3 Da) suggests that formylation (+28) might have occurred in preference to oxidation (Fig. 4B). The latter peak could correspond to the peptide incorporating two oxygen atoms in methionine and one in tryptophan. These reactions are predictable [11] and clearly distinguishable from the desired one, allowing for the shift of mass values. Under the present reaction conditions, the α- and ε-amino groups are also formylated. However, the TMPP-Ac derivative of isolated C-terminal peptide is devoid of these amino groups, thus causing no problem [9].

In the peptide TMPP-Ac-PVTL-OCH_3, the C-terminal methyl ester was found to undergo hydrolysis and aminolysis simultaneously to give TMPP-Ac-PVTL-OH with the signal at *m/z* 1001.6 and TMPP-Ac-PVTL-$NHCH_3$ with that at *m/z* 1014.6 as shown in Fig. 4A. However, it is possible that the methyl-amidated peptide occurred *via* a sequential reaction of hydrolysis and amidation is not excluded.

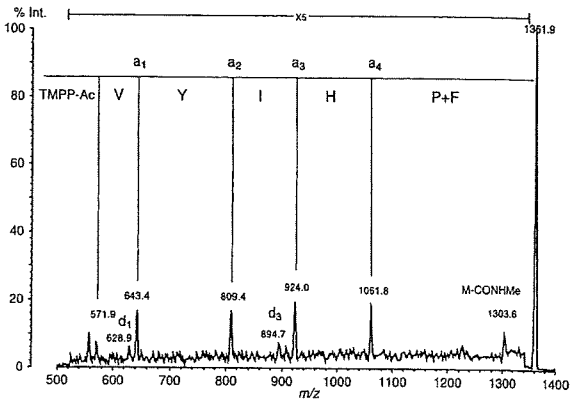


Figure 3. MALDI-CID spectrum of peptide 3 methylamide derivative (TMPP-Ac-VYIHPF-$NHCH_3$).

formylation of amino and hydroxyl groups. If necessary, the mass of methylamine may be altered by isotopically labeling any atom(s) to further accentuate the distinction between the amide and free carboxyl forms of the C-termini, of which masses intrinsically differ as little as 1 Da. With such versatility in the choice of reagent with varied mass values, the present method should be applicable to *de novo* amino acid sequencing of any kind of PTM to occur at the C-termini. Although the additional task of C-terminal derivatization may reduce the throughput of the original protocol [9], this method has a practical feasibility for identifying C-terminal PTMs with high fidelity, which has usually been very difficult to attain.

4 Concluding remarks

The development of a chemical method for detecting C-terminal PTMs of peptides and proteins is described. The method is based on oxazolone chemistry, which enables the specific derivatization of the C-terminal-free carboxyl group to, for example, methylamide by reaction with methylamine, while leaving the side-chain carboxyl groups unchanged. One of the most promising applications of this method is to detect and characterize C-terminally amidated peptides and proteins because the small mass difference of 1 Da between masses of the free carboxyl group and its amide is expanded to 14 Da due to the conversion of the carboxyl group to the corresponding methylamide. The C-terminal amide structures of adrenomedullin and calcitonin are unambiguously discriminated from the free C-terminus of BSA. The efficacy of this method is such that it can determine even the isobaric variations in C-terminal structures corresponding to 132 Da (Asp-NH_2 and Asn-OH) and 146 Da (Glu-NH_2 and Gln-OH). These features suggest that the present method has a wider applicability to determine the existence or nonexistence of PTMs at the C-termini of peptides and proteins.

The authors have declared no conflict of interest.

5 References

[1] Meinnel, T., Peynot, P., Giglione, C., Processed N-termini of mature proteins in higher eukaryotes and their major contribution to dynamic proteomics. *Biochemie* 2005, *87*, 701–712.

[2] Brown, J. L., Roberts, W. K., Evidence that approximately eighty per cent of the soluble proteins from Ehrlich ascites cells are N^{α}-acetylated. *J. Biol. Chem.* 1976, *251*, 1009–1014.

[3] Eipper, B. A., Stoffers, D. A., Mains, R. E., The biosynthesis of neuropeptides: peptide alpha-amidation. *Annu. Rev. Neurosci.* 1992, *15*, 57–85.

[4] Ferguson, M. A., Low, M. G., Cross, G. A., Glycosyl-*sn*-1,2-dimyristylphosphatidylinositol is covalently linked to *Trypanosoma brucei* variant surface glycoprotein. *J. Biol. Chem.* 1985, *260*, 14547–14555.

[5] Paulick, M. G., Bertozzi, C. R., The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. *Biochemistry* 2008, *47*, 6991–7000.

[6] Elortza, F., Mohammed, S., Bunkenborg, J., Foster, L. J. *et al.*, Modification-specific proteomics of plasma membrane proteins: identification and characterization of glycosylphosphatidylinositol-anchored proteins released upon phospholipase D treatment. *J. Proteome Res.* 2006, *5*, 935–943.

[7] Mueller, G. P., Driscoll, W. J., α-Amidated peptides: approaches for analysis. *Methods. Mol. Biol.* 2008, *446*, 67–84.

[8] Yamaguchi, H., Sasaki, K., Satomi, Y., Shimbara, T. *et al.*, Peptidomic identification and biological validation of neuroendocrine regulatory peptide-1 and -2. *J. Biol. Chem.* 2007, *282*, 26354–26360.

[9] Kuyama, H., Shima, K., Sonomura, K., Yamaguchi, M. *et al.*, A simple and highly successful C-terminal sequence analysis of proteins by mass spectrometry. *Proteomics* 2008, *8*, 1539–1550.

[10] Nakazawa, T., Yamaguchi, M., Nishida, K., Kuyama, H. *et al.*, Enhanced responses in matrix-assisted laser desorption/ionization mass spectrometry of peptides derivatized with arginine via a C-terminal oxazolone. *Rapid Commun. Mass Spectrom.* 2004, *18*, 799–807.

[11] Yamaguchi, M., Oka, M., Nishida, K., Ishida, M. *et al.*, Enhancement of MALDI-MS spectra of C-terminal peptides by the modification of proteins *via* an active ester generated *in situ* from an oxazolone. *Anal. Chem.* 2006, *78*, 7861–7869.

[12] Wachter, E., Machleidt, W., Hofner, H., Otto, J., Aminopropyl glass and its *p*-phenylene diisothiocyanate derivative, a new support in solid-phase Edman degradation of peptides and proteins. *FEBS Lett.* 1973, *35*, 97–102.

[13] Kuyama, H., Sonomura, K., Nishimura, O., Sensitive detection of phosphopeptides by matrix-assisted laser desorption/ionization mass spectrometry: use of alkylphosphonic acids as matrix additives. *Rapid Commun. Mass Spectrom.* 2008, *22*, 1109–1116.

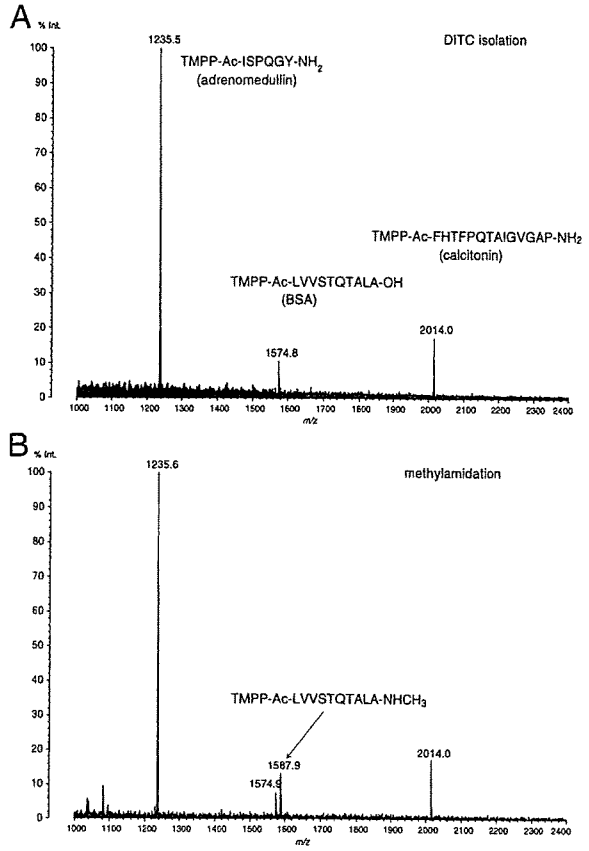


Figure 8. MALDI-TOF mass spectra of protein/peptide mixture: (A) isolated three C-terminal peptides from BSA (*m/z* 1574.8), adrenomedullin (*m/z* 1235.5), and calcitonin (*m/z* 2014.0). (B) Methylamidation converts the C-terminal peptide from BSA to its methylamide form, whereas the other two peptides are intact.

any change in mass, while C-terminal peptide from BSA underwent methylamidation properly to give TMPP-Ac-LVVSTQTALA-$NHCH_3$, the molecular mass of which is higher than that of the C-terminally free precursor by 13 Da. Figure 8B depicts the MALDI mass spectrum of the mixture of TMPP-Ac-ISPQGY-NH_2 from adrenomedullin and TMPP-Ac-FHTFPQTAIGVGAP-NH_2 from calcitonin thus processed. The identity of TMPP-Ac-LVVSTQTALA-$NHCH_3$ derived from BSA was confirmed by tandem mass spectrometry (Fig. 9). Each of these isolated peptides contains a few hydroxyl groups of serine, threonine, or tyrosine, where formylation can occur concomitantly with the activation of the C-terminus to oxazolone in the mixture of formic acid and acetic anhydride. However, no formylation was detected in these peptides, probably because the hydroxyl groups could be restored by aminolysis of the formate during the reaction with methylamine [11].

3.4 Reagents and improvement of protocol for future development

We chose to employ methylamine for amidation of the C-terminal carboxyl group, because it is the simplest alkylamine that is least likely to suffer from any steric hindrance to the reaction with the activated C-terminal carboxyl group. In addition, it can be used in large excess with the peptides to be amidated, due to its high solubility in any solvent. This is particularly important to suppress hydrolysis, which could lower the yield of methylamidation; there is the risk of reporting false-positive results in the mass spectra when the yield of the reaction is very poor or negligibly low. Methylamidation of a carboxylic acid is accompanied by an increase in the molecular mass by 13 Da, making it easier to discriminate from the mass shift caused by other side reactions such as

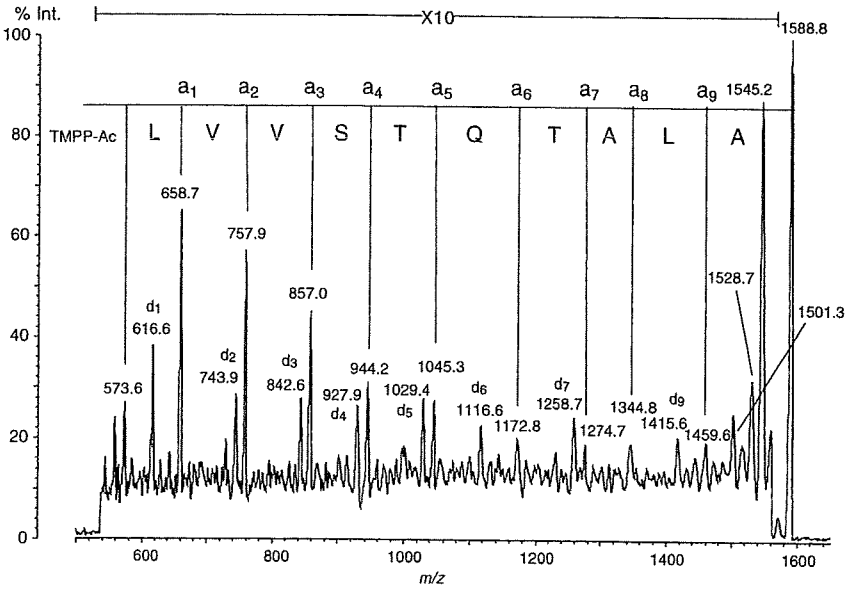


Figure 9. MALDI-CID spectrum of a methylamidated C-terminal peptide from BSA (TMPP-Ac-LVVSTQTALA-$NHCH_3$).

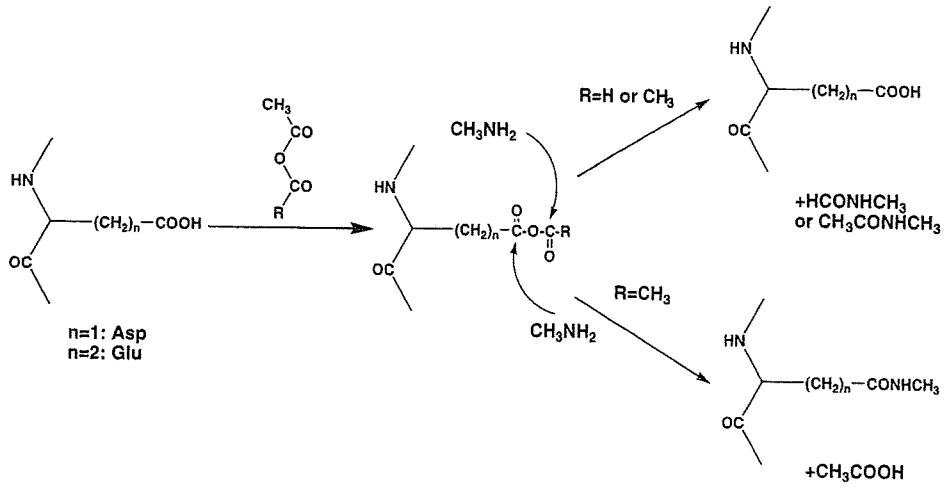


Figure 6. Possible involvement of acetic anhydride (R = CH_3) in the false activation of the side-chain carboxyl group to the reaction with methylamine. The mixed anhydride with acetic acid allows the reaction to proceed in both directions of leaving the carboxyl group free and changing it to the amide form.

twice (at the initial and final steps of C-terminal activation) so that the undesirable hydrolysis of oxazolone or active ester could be suppressed. Although the yield of amidation appeared to become higher by increasing the proportion of Ac_2O in the reagent (*e.g.* HCOOH–Ac_2O–Pfp-OH = 1:1:1, 1:3:1, 1:5:1, *etc.*), the specificity of methylamidation to the C-terminal carboxyl group was adversely affected (data not shown). This was possible because the side chains of aspartic acid and glutamic acid can be activated due to the formation of mixed anhydride with acetic acid, allowing methylamine to react with the activated side-chain carboxyl group as well as the acetyl group of the mixed anhydride (Fig. 6). In the proper condition of 1:1 mixture of formic acid and acetic anhydride, acetic anhydride is almost completely converted to formic acetic anhydride, which then reacts with any carboxyl group to form the mixed anhydride with formic acid in preference to that with acetic acid. It is far less likely that the resulting mixed anhydride with formic acid leads the side-chain carboxyl group to be amidated. Therefore, the modest yield in the range of 60–70% is sufficient to distinguish between the free and modified C-terminal carboxyl groups, according to the appearance of MALDI mass peak(s) in a single spectrum. Given the recognition of the peak of interest, the MS/MS C-terminal amino acid sequencing of the peak would follow for the identification of the peptide and the characterization of C-terminal modification.

3.3 Amidation of the model protein and large peptides

To test the applicability of the method to proteins or large peptides, we tried to isolate and characterize the C-terminal peptides by using a mixture of BSA, human adrenomedullin, and human calcitonin (10 pmol each). Of these samples, adrenomedullin (52 amino acid residues; 6028 Da) has Tyr-NH_2, and calcitonin (32 amino acid residues; 3417 Da) has Pro-NH_2 at the C-termini, distinguished from BSA having

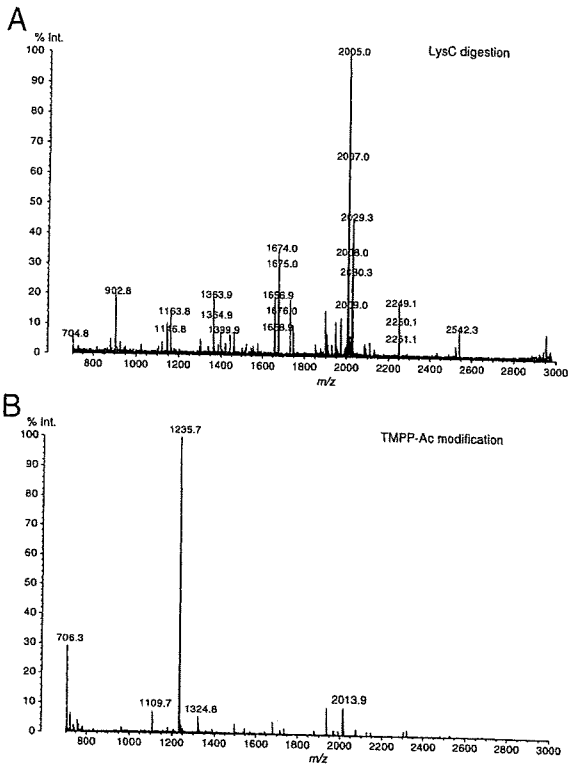


Figure 7. MALDI-TOF mass spectra of the protein/peptide mixture containing BSA, human adrenomedullin, and human calcitonin: (A) LysC digest and (B) TMPP-Ac modification.

the free C-terminal carboxyl group. The mixture was subjected to LysC digestion, followed by N-terminal modification with TMPP-Ac-OSu and isolation of C-terminal peptides by the treatment with *p*-phenylenediisothiocyanate glass, according to the standard protocol [9]. MALDI mass spectra recorded at each step of the procedure are indicated in Fig. 7. Owing to the absence of free carboxyl groups at their C-termini, neither the C-terminal peptide derived from adrenomedullin nor that derived from calcitonin indicated

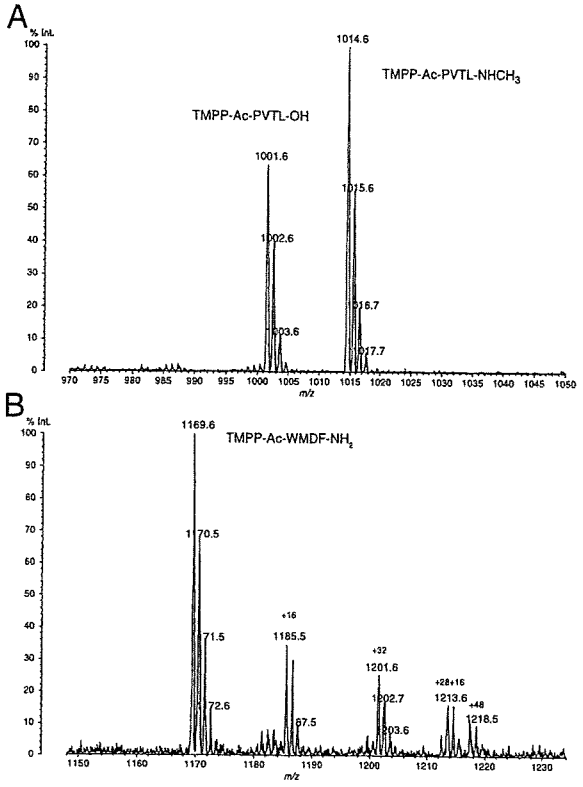


Figure 4. Peptides 1 and 2 after methylamidation: enlarged spectra from Fig. 2B. (A) peptide 1 (TMPP-Ac-PVTL-OCH_3) and (B) peptide 2 (TMPP-Ac-WMDF-NH_2).

Regardless of the course by which the ester has been converted to the amide, any kind of ester could be analyzed in this manner by assuming the same reactions and results.

3.2 Discrimination between the C-terminal and side-chain carboxyl groups

As we presumed in the Introduction, in the unfortunate situation in which the C-terminus of a peptide is Asn-OH/Gln-OH or amidated Asp-NH_2/Glu-NH_2, it is impossible to discriminate between these isobaric residues by mass spectrometry alone. However, the present method can achieve the requisite discrimination by causing a mass shift by +13 Da for peptides having free Asn or Gln at the C-terminus through methylamidation, but no change for those with C-terminal amide. We further investigated the validity of this approach using two pairs of model peptides: HPTFD-NH_2/SFLLRN-OH and TTSFAE-NH_2/ALEGSLQ-OH. As expected, the peptide derivatives TMPP-Ac-HPTFD-NH_2 and TMPP-Ac-TTSFAE-NH_2 containing the amide form of aspartic and glutamic acids at their C-termini remain intact after methylamidation, whereas C-terminal-free TMPP-Ac-ALEGSLQ-OH and TMPP-Ac-SFLLRN-OH are converted to the corresponding methylamides with the increment of mass value by 13 Da (Figs. 5A and B). An additional peak accompanied to

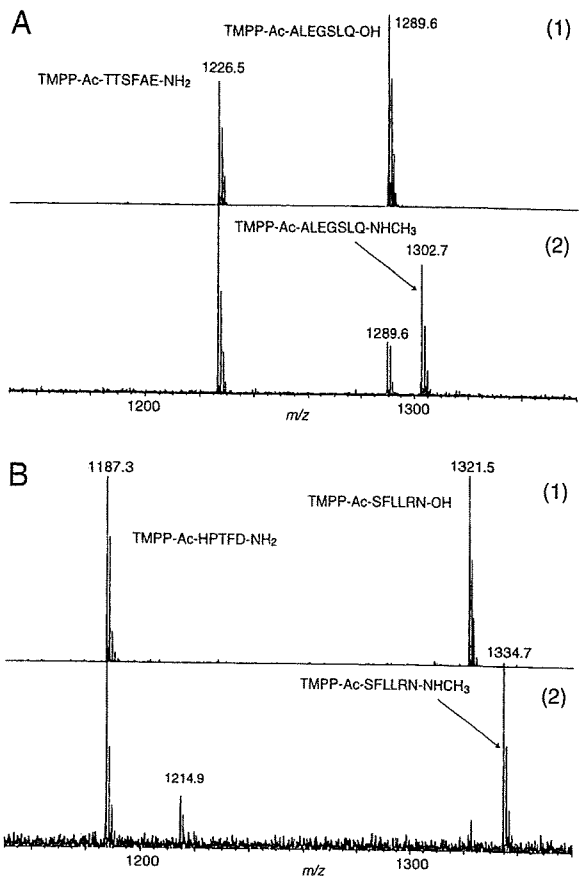


Figure 5. Methylamidation of two pairs of peptides: (A) TMPP-Ac-TTSFAE-NH_2, TMPP-Ac-ALEGSLQ-OH and (B) TMPP-Ac-HPTFD-NH_2, TMPP-Ac-SFLLRN-OH. Peptide having Asn (N) or Gln (Q) at the C-terminus is converted to its methylamide (+13 Da), whereas peptide incorporating the amide form of aspartic acid (D) or glutamic acid (E) is intact.

that of TMPP-Ac-HPTFD-NH_2 (Fig. 5B-2) seems to arise from the formylation (+28 Da) of the histidine or threonine residue during methylamidation. Nevertheless, no confusion due to the side-chain formylation might be imposed on the discrimination of peptides entailing this isobaric ambiguity.

The efficiency of our method for the detection of C-terminal amide and the discrimination between the C-terminal and side-chain carboxyl groups depends on the yield of methylamidation, which was usually incomplete so that the yield was 60–90%, as roughly estimated from the signal intensities of several pairs of peaks separated by 13 Da each (Figs. 2B, 5A-2, 5B-2, and 8B). Fortunately, this incompleteness of the reaction is advantageous for detecting the prospective single peak of C-terminal peptide amides not exhibiting the mass shift of 13 Da, which could otherwise be discerned as the appearance of a pair of peaks that signify the existence of a C-terminal-free carboxyl group. In this study, we slightly modified the original protocol [10, 11] to increase the yield of methylamidation by adding Pfp-OH

Bioorganic & Medicinal Chemistry Letters 19 (2009) 6544–6547

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

A method for terminus proteomics: Selective isolation and labeling of N-terminal peptide from protein through transamination reaction

Kazuhiro Sonomura [a,b], Hiroki Kuyama [a], Ei-ichi Matsuo [a,b], Susumu Tsunasawa [a], Osamu Nishimura [a,*]

[a] Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan
[b] Life Science Research Center, Technology Research Laboratory, Shimadzu Corporation, Kyoto 619-0237, Japan

ARTICLE INFO

Article history:
Received 9 September 2009
Revised 8 October 2009
Accepted 9 October 2009

Keywords:
Transamination
N-terminal labeling
MALDI-TOF MS
N-terminal sequence analysis

ABSTRACT

A novel method for selectively labeling and isolating N-terminal peptide from protein has been developed. An N^{α}-amino group of protein was converted to a carbonyl group through transamination reaction and the resulting carbonyl group was modified with *O*-(4-nitrobenzyl)hydroxylamine (NBHA). After proteolytic digestion using *Grifola frondosa* metalloendopeptidase (LysN), the modified N-terminal peptide remained unbound in the following treatment using amino-reactive *p*-phenylenediisothiocyanate (DITC) glass, whereas peptides other than the N-terminal peptide were effectively scavenged from the supernatant solution. The modified N-terminal peptide was thus successfully isolated and sequenced by matrix-assisted laser desorption/ionization tandem mass spectrometry (MALDI-MS/MS) analysis.

Protein identification using mass spectrometry is a key technology in proteome research. In the conventional approach, such as peptide mass fingerprinting (PMF), target proteins are enzymatically or chemically digested into peptide fragments and measured mass values are compared with theoretical ones calculated from genome or protein databases.[1] Therefore, the exact state of mature proteins is not always determined; the mature protein sequence may often differ from that translated from its DNA sequence due to processes such as splicing and shuffling of mRNA, and/or the various post-translational modifications (PTMs). For conducting further study in depth on each protein identified by conventional proteome analysis, more detailed information about the exact amino acid sequence and PTMs of mature proteins from N-terminus to C-terminus has been highly desired.

Mass spectrometry has become a major tool in protein and peptide analysis and various methods have been developed and applied to actual samples.[2,3] However Edman sequencing is still the only conventional method for determining the N-terminal sequence. This method lacks generality, in that it cannot perform sequencing of N-terminally capped protein. Therefore, methodology using MS for both N- and C-terminal sequence analysis has been intensively studied for general and practical applications.[4–10]

We have been developing methods for isolating N- and C-terminal peptides from a peptide mixture of digested proteins and sequencing the isolated terminal peptides by MS.[8–10] In our review (in press), we proposed that this scientific field be termed 'terminus proteomics.'[11] In this study, we focused on the characteristic reactivity of metal-ion catalyzed transamination reaction,[12] which one of the authors (O.N.) had used for removing the N-terminal methionine residue from recombinant proteins produced from *Escherichia coli*.[13,14] This reaction selectively converts the N^{α}-amino group (N^{α}-NH_2) of peptides or proteins into a carbonyl group, whereas the coexisting N^{ε}-amino group (N^{ε}-NH_2) of lysine residue is kept intact. Reactions that selectively modify N^{α}-NH_2 or N^{ε}-NH_2 have often been utilized to isolate N- or C-terminal peptides from proteolytic digests; however, the selectivity is hardly obtained in general. The high selectivity of transamination reaction for N^{α}-NH_2 encouraged us to investigate its feasibility for isolating terminal peptides and we successfully developed a new method for C-terminal sequencing analysis using this transformation.[10] Here, we report an application of transamination reaction to N-terminal analysis, which was achieved by changing the enzyme used for digestion and by optimizing the procedure.

In the previous C-terminal analysis, proteins were first digested with lysylendopeptidase (LysC). LysC cleaves peptide bonds at the carboxyl side of lysine residues (-Lys-|-Xaa-) to yield mainly peptides having N^{α}-NH_2 and N^{ε}-NH_2 at both ends; however, C-terminal peptides have only N^{α}-NH_2. Subsequent transamination reaction converts the N^{α}-NH_2 to a carbonyl group; thus, incubation with amino-reactive *p*-phenylenediisothiocyanate (DITC) glass effectively scavenges N^{ε}-NH_2-containing peptides (peptides other than the C-terminal one). The recovered C-terminal peptide is open to modification with various nucleophilic reagents, such as hydrazine and hydroxylamine. We used 2,4-dinitrophenylhydrazine (DNPH), which demonstrated signal enhancement in MS analysis.

* Corresponding author. Tel./fax: +81 6 6879 4320.
E-mail address: osamu_nishimura@protein.osaka-u.ac.jp (O. Nishimura).

0960-894X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2009.10.044

In studying the C-terminal analysis employing a transamination reaction, we envisioned extending the method to N-terminal analysis by using *Grifola frondosa* metalloendopeptidase (LysN),[15–17] which cleaves proteins at the amino side of lysine residues, instead of LysC, as described in the previous report.[9] In addition, the protocol was changed so that a sample protein was first subjected to transamination and modification of the resulting carbonyl group, which facilitated the easy purification of the modified protein by simple gel filtration or ultrafiltration. Transamination reaction has been applied to proteins as well as peptides.[12,18–20] Thus, the procedure for N-terminal analysis was optimized (Scheme 1). The procedure consists of the following five steps: (1) transamination of proteins, (2) modification of the resulting carbonyl group through oxime formation, (3) LysN digestion, (4) treatment with DITC glass, and (5) MALDI -MS and MS/MS analysis. In this paper, sample proteins are transaminated and modified with *O*-(4-nitrobenzyl)hydroxylamine (NBHA) at their N-termini.[21] NBHA has a nitrobenzene moiety as well as DNPH, which would be expected to increase sensitivity in MALDI-MS analysis by combining a matrix system developed for sensitive detection of 2-nitrobenzenesulfenyl(NBS)-modified peptides.[22] LysN cleaves peptide bonds at the amino side of lysine residues (-Xaa-|-Lys-). After LysN digestion, the modified N-terminal peptide contains no amino group. Amino-reactive DITC glass scavenges the peptide fragments other than the N-terminal peptide, and the target peptide is left unreacted in the supernatant. Thus recovered N-terminal peptides are analyzed by MALDI-MS and MS/MS.

Two model proteins, bovine α-lactalbumin (α-la) and bovine β-lactoglobulin (β-lg), were chosen for testing the protocol. Proteins were transaminated in 10% pyridine containing 0.2 M glyoxylic acid, 6 mM $CuSO_4$, and 2 M urea. It was reported that the transamination reaction proceeded as well in the presence of urea.[12,14] Reaction time was set to 1 h, based on the results of our preliminary experiments using some peptides incorporating a different residue at their N-termini (data not shown). The reaction mixture was then subjected to a buffer exchange into 100 mM phosphate (pH 6.0) containing 2 M urea by ultrafiltration. Subsequent modification with NBHA was performed in 100 mM phosphate (pH 6.0) containing 50 mM NBHA-HCl and 2 M urea for 2 h at 37 °C. After a buffer exchange into 50 mM $NaHCO_3$ containing 2 M urea by ultrafiltration, disulfide bonds were reduced with tris(2-carboxyethyl)phosphine hydrochloride and the resulting free sulfhydryl groups were alkylated with iodoacetamide. LysN digestion was then performed at an enzyme-to-substrate ratio of 1:40 at room temperature for 15 h.[9] The upper panels in Figure 1 depict MALDI-MS spectra[23] after LysN digestion (Fig. 1a for α-la and Fig. 1c for β-lg), in which the modified N-terminal peptides (indicated by arrows) were detected among the peptide fragments as a sodium adduct ($[M+Na]^+$: 661.3 from α-la and $[M+Na]^+$: 976.5 from β-lg). When peptides containing a nitrobenzene moiety are analyzed by MALDI-MS, they are often accompanied by peaks with a 16 Da decrease probably due to splitting off of the oxygen atom from nitro group.[22] In both spectra these peaks were detected

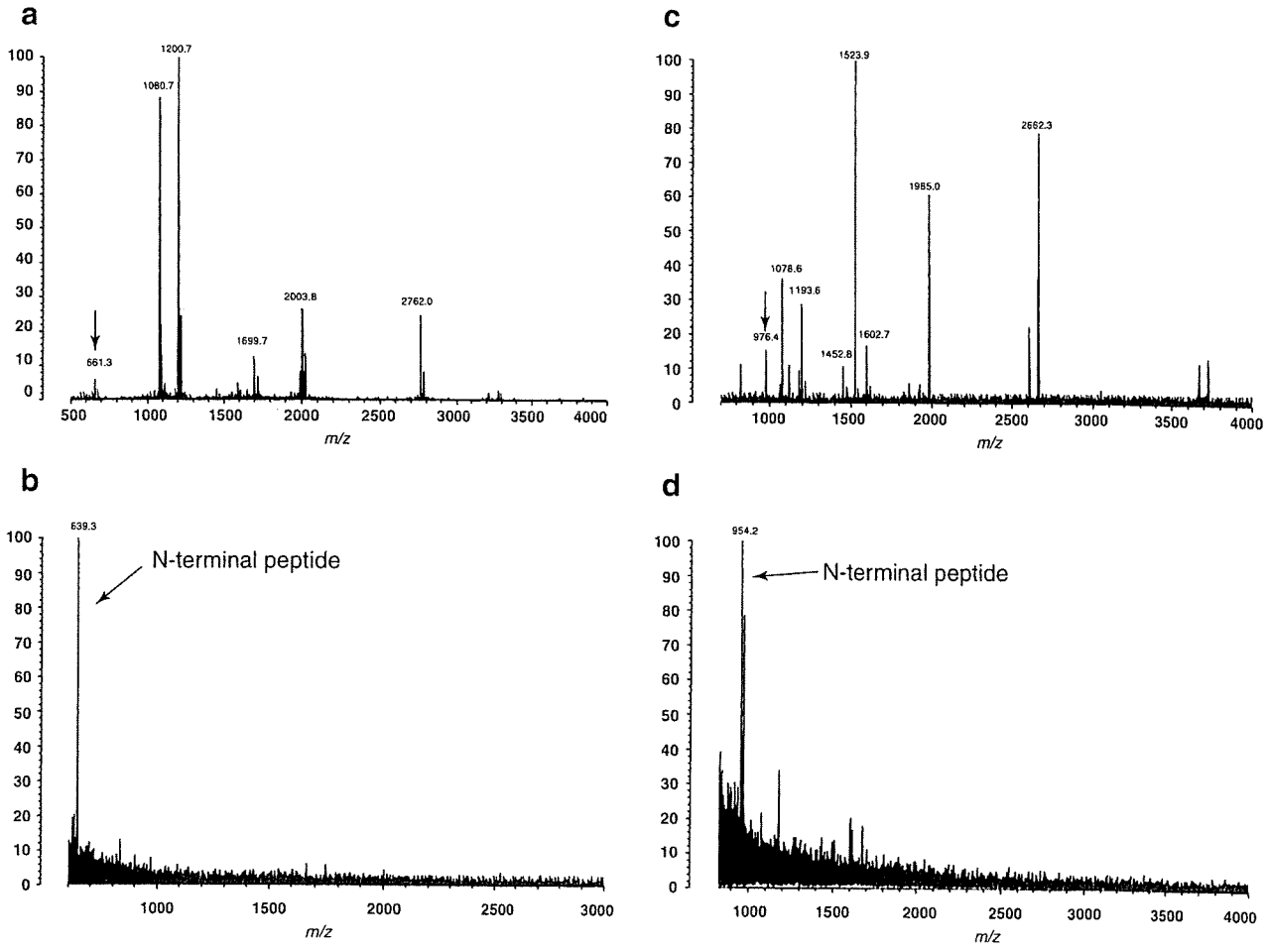


Figure 1. MALDI-MS spectra of peptide fragments of model proteins after LysN digestion ((a) α-lactalbumin, (c) β-lactoglobulin) and after treatment with DITC glass ((b) α-lactalbumin, (d) β-lactoglobulin.). Arrows indicate the modified N-terminal peptides of each protein. An aliquot (3 pmol of digest) was loaded onto the target plate.

and aided in discriminating the modified N-terminal peptides from the other peptide fragments even before the isolation step.

The LysN digest was then treated with DITC glass.[9,24] An aliquot of the digest (30 pmol) was added to pre-washed DITC glass. After 2 h incubation at 60 °C, the supernatant was directly analyzed by MALDI-MS. We employed a mixture of α-cyano-4-hydroxycinnamic acid (CHCA) and 3-hydroxy-4-nitrobenzoic acid (3H4NBA) as a binary matrix system, which was used for sensitive detection of NBS-modified peptides and was shown to be applicable to peptides containing a nitrobenzene moiety.[22,25] The lower panels in Figure 1 illustrate the MALDI-MS spectra after the DITC treatment (Fig. 1b for α-la and Fig. 1d for β-lg). Each N-terminal peptide was singly recovered ($[M+H]^+$: 639.3 from α-la and $[M+H]^+$: 954.5 from β-lg). In this experiment, these modified N-terminal peptides needed 3H4NBA as a matrix for the detection, the usefulness of which was also demonstrated in this case.

The isolated N-terminal peptides were subjected to MALDI-MS/MS analysis in the post-source decay (PSD) mode. The amino acid sequences of the peptides were analyzed using the observed fragment peaks (Fig. 2a for α-la and Fig. 2b for β-lg).

For N-terminally acetylated proteins, the method described here can be employed for N-terminal analysis because the N-terminal acetyl group is not affected in the transamination reaction (data not shown) and the acetylated N-terminal peptide after LysN digestion is free from the amino group. Studies in the case of N-blocked proteins were described in the previous reports.[5,6,9,26]

In the present study, we developed N-terminal analysis of proteins through transamination reaction. The method was applied to two model proteins, and successful isolation and sequencing of the N-terminal peptides were demonstrated. This method can not be used for some proteins (e.g., proteins containing N-terminal proline) because of limited reactivity for transamination.[10,27] However, an α-carbonyl of N-terminal ketoacyl group generated after transamination reaction can react with various types of nucleophilic reagents, and virtually any functionality can be introduced into the N-terminus of the peptide. This flexibility is advantageous for the MALDI-MS analysis of terminal peptides whose sensitivity largely depends on its amino acid composition, size, and modification states.

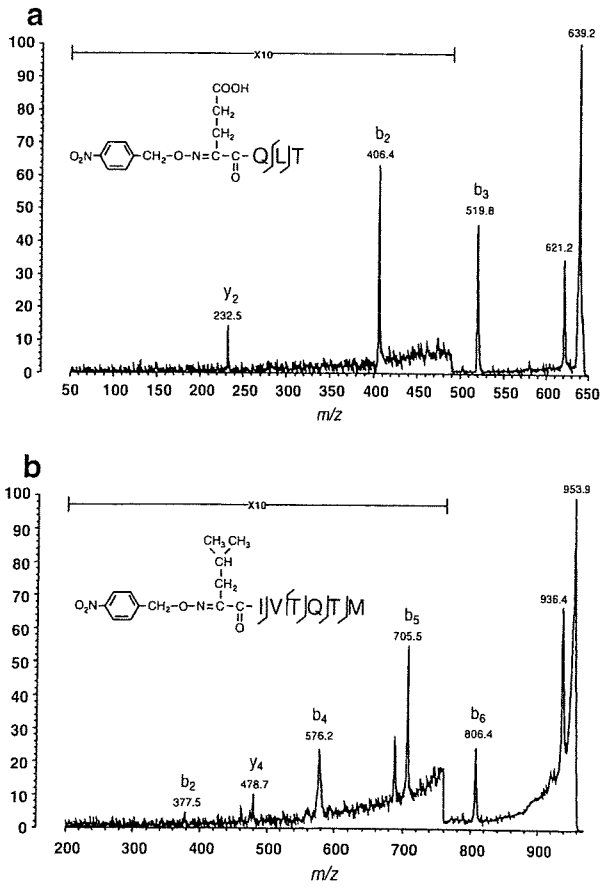
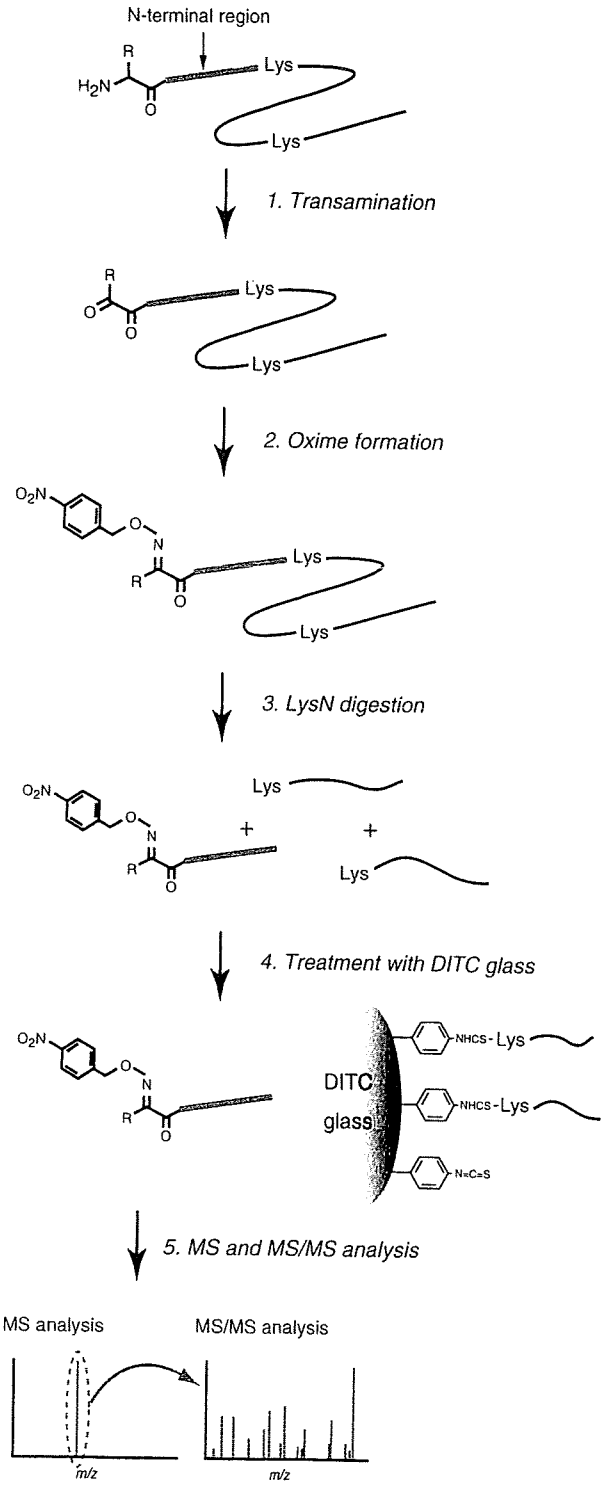


Figure 2. MALDI-MS/MS spectra (PSD mode) of the isolated N-terminal peptides of model proteins. (a) α-lactalbumin, (b) β-lactoglobulin.



Scheme 1. Protocol for selectively labeling and isolating N-terminal peptides.

References and notes

1. Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5011.
2. Mann, M.; Jensen, O. N. *Nat. Biotechnol.* **2003**, *21*, 255.
3. Thingholm, T. E.; Jensen, O. N.; Larsen, M. R. *Proteomics* **2009**, *9*, 1451.
4. Samyn, B.; Sergeant, K.; Castanheira, P.; Faro, C.; Beeumen, J. V. *Nat. Methods* **2005**, *2*, 193.
5. McDonald, L.; Robertson, D. H. L.; Hurst, J. L.; Beynon, R. J. *Nat. Methods* **2005**, *2*, 955.
6. Yamaguchi, M.; Nakayama, D.; Shima, K.; Kuyama, H.; Ando, E.; Okamura, T.; Ueyama, N.; Nakazawa, T.; Norioka, S.; Nishimura, O.; Tsunasawa, S. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 3313.
7. Mahrus, S.; Trinidad, J. C.; Barkan, D. T.; Sali, A.; Burlingame, A. L.; Wells, J. A. *Cell* **2008**, *134*, 866.
8. Kuyama, H.; Shima, K.; Sonomura, K.; Yamaguchi, M.; Ando, E.; Nishimura, O.; Tsunasawa, S. *Proteomics* **2008**, *8*, 1539.
9. Kuyama, H.; Sonomura, K.; Nishimura, O.; Tsunasawa, S. *Anal. Biochem.* **2008**, *380*, 291.
10. Sonomura, K.; Kuyama, H.; Matsuo, E.; Tsunasawa, S.; Nishimura, O. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 611.
11. We once used the term 'terminal proteomics' for the proteomics focusing on N- and/or C-terminal part of protein (Nakazawa et al. *Proteomics* **2008**, *8*, 673). However, we propose after careful consideration the term 'terminus proteomics' instead of former 'terminal proteomics'. The reason is that using 'terminus' removes any negative connotations and avoids misunderstanding or confusion. See the footnote in Kuyama, H.; Nishimura, O.; Tsunasawa, S. *Handbook on Mass Spectrometry: Instrumentation, Data and Analysis, and Applications*; Lang, J. K., Eds.; Nova Science Publishers: New York, in press; Chapter 19 (A New Method for N- and C-terminal de novo Amino Acid Sequencing of Proteins by Mass Spectrometry).
12. Dixon, H. B. F.; Fields, R. *Methods Enzymol.* **1972**, *25*, 409.
13. Nishimura, O.; Suenaga, M.; Ohmae, H.; Tsuji, S.; Fujino, M. *Chem. Commun.* **1998**, 1135.
14. Suenaga, M.; Ohmae, H.; Okutani, N.; Kurokawa, T.; Asano, T.; Yamada, T.; Nishimura, O.; Fujino, M. *J. Chem. Soc., Perkin Trans. 1* **1999**, 3727.
15. Nonaka, T.; Ishikawa, H.; Tsumuraya, Y.; Hashimoto, Y.; Dohmae, N.; Takio, K. *J. Biochem.* **1995**, *118*, 1014.
16. Nonaka, T.; Hashimoto, Y.; Takio, K. *J. Biochem.* **1998**, *124*, 157.
17. Hori, T.; Kumasaka, T.; Yamamoto, M.; Nonaka, T.; Tanaka, N.; Hashimoto, Y.; Ueki, T.; Takio, K. *Acta Crystallogr., Sect. D* **2001**, *57*, 361.
18. Dong, S. Y.; Ma, H. M.; Duan, X. J.; Chen, X. Q.; Li, J. *J. Proteome Res.* **2005**, *4*, 161.
19. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307.
20. Sasaki, T.; Kodama, K.; Suzuki, H.; Fukuzawa, S.; Tachibana, K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4550.
21. We first tried the isolation of N-terminal peptide through condensation with DNPH, but the modified N-terminal peptide was not obtained. Therefore, we changed DNPH for NBHA.
22. Matsuo, E.; Toda, C.; Watanabe, M.; Ojima, N.; Izumi, S.; Tanaka, K.; Tsunasawa, S.; Nishimura, O. *Proteomics* **2006**, *6*, 2042.
23. MALDI-MS spectra and tandem mass spectra (PSD mode) were obtained on an AXIMA-TOF^2 mass spectrometer (Shimadzu, Kyoto, Japan; Kratos, Manchester, UK). All analyses were performed in positive-ion reflectron mode.
24. Wachter, E.; Machleidt, W.; Hofner, H.; Otto, J. *FEBS Lett.* **1973**, *35*, 97.
25. Matrix solution was prepared by dissolving each CHCA and 3H4NBA in 0.1% TFA solution containing 50% acetonitrile at a concentration of 10 mg/ml and combining equal volumes of each solution. In addition, we used methanediphosphonic acid (MDPNA), which was reported to be effective for eliminating alkali metal adduct, as a matrix additive (Kuyama, H.; Sonomura, K.; Nishimura, O. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 1109). Hence, sample solution, matrix solution, and MDPNA solution were mixed on a MALDI target plate and air-dried (Fig. 1b and 1d).
26. Coussot, G.; Hawke, D. H.; Mularz, A.; Koomen, J. M.; Kobayashi, R. *Anal. Biochem.* **2007**, *361*, 302.
27. Scheck, R. A.; Dedeo, M. T.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 11762.