A Method to Enhance a_1 Ions and Application for Peptide Sequencing and Protein Identification

Ning Liu, Wan Chan, Kim-Chung Lee, and Zongwei Cai
Department of Chemistry, Hong Kong Baptist University, Hong Kong SAR, China

A simple and effective method was developed for peptide sequencing and protein identification through the determination of its N-terminal residue. The method of N-terminal carbamidomethylation with iodoacetamide could specifically and remarkably enhance the intensity of a_1 ions in the tandem mass spectra of the peptide derivatives without significantly altering their fragmentation pattern, thus allowing determination of their N-terminal residues. The effectiveness and specificity of the method was demonstrated by confirming and extending sequence interpretation of several model peptides and proteins. The developed method was then applied in the LC-MS/MS analysis of the tryptic digests of myoglobin and a whole protein extract from rat heart tissues. The results from database searches were well validated with the enhancement of a_1 ions in tandem mass spectra and the specificity of protein identification was obtained when the information of N-terminal residues was included in the database search.   (J Am Soc Mass Spectrom 2009, 20, 1214–1223) © 2009 American Society for Mass Spectrometry

Protein identification is a major component of proteomics, with the methods heavily relying on both mass spectrometric peptide sequencing and database searching [1]. Although the application of tandem mass spectrometry for peptide sequencing based on the chemistry of peptide fragmentation in the gas phase has been reported for many years [2–5], various methodologies are still being developed and modified for improving the sensitivity and specificity [6–9]. In common equipment settings, such as tandem quadrupole, quadrupole-time of flight, and ion trap mass spectrometers, a, b, and y ions are the most common peptide fragments observed in low-energy collisions. Both a and b ions are viewed as N-terminal fragments, while y ions belong to C-terminal ones. Theoretically, the sequence of a peptide may be deduced from the analysis of fragment ions. However, in practice, the complete interpretation of product ion spectra is often hard to achieve without the aid of database search. Thus, de novo sequencing often ends up with partial interpretation of some internal sequences in peptides. The most common reason is the absence of b_1 ions in product ion spectra of the protonated peptides due to the difficulty to form the corresponding cyclic intermediates [10, 11], which often posted a challenge for the determination of N-terminal residues by mass spectrometry.

Several derivatization reagents that were commonly used to label the N-terminus were found to be able to promote and stabilize b_1 ions and to facilitate the interpretation process [12–18], some of which focused on the comparative proteomics by using different isotopically labeled reagents. While the reported methods regarding the derivatization at the peptide N-terminus could stabilize the modified b_1 ions to a certain extent, none of them seemed to serve as a simple, sensitive tool to identify the N-terminal residues in proteomics due to the following reasons. The intensity of many derivatized peptides appeared to be reduced in the analysis using positive ion mode mass spectrometry compared with those of underivatized peptides. N-acetylation with acetic anhydride, phenyl isothiocyanate, or 4-sulfophenyl isothiocyanate often resulted in a reduced intensity to some extent. In addition, some derivatization reactions resulted in different fragmentation patterns in product ion spectra [14, 18], making it difficult to gain the sequence information by the comparison of the spectrum of modified peptide with that of unmodified one. A typical example was the N-terminal derivatization with phenyl isothiocyanate, of which the product ion spectrum was often dominated by b_1 and y_{n-1} ions [14]. However, the limited amount of fragmentation often resulted in the limited sequence information. Furthermore, some derivatization regents were not commercially available [17, 18].

While the assignment of b_1 ions in MS/MS spectra has often been used for obtaining information of N-terminal residues, limited reports on the use of a_1 ions to determine the N-terminus of peptides [19] and proteins [20] have been published. The detection frequency

Address reprint requests to Dr. Z. Cai, Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China. E-mail: zwcai@hkbu.edu.hk

and abundance of a ions are often lower than those of b ions. The formation of a_1 ion was believed to occur via the fragmentation of b_2 ion and/or a_2 ion [10]. Because a_1 ion is often intermixed with immonium ions including those from the internal residues that are identical to the N-terminal residue, the determination of the N-terminal residue is difficult. In this study, a simple and effective method to specifically and remarkably enhance a_1 ions in the MS/MS spectra of the peptide derivatives was developed for identifying the N-terminal amino acids of peptides. The method applicability for the enhancement of a_1 ions in mass spectrometric sequencing and protein identification was demonstrated.

Experimental

Materials and Reagents

Dithiothreitol (DTT), iodoacetamide, ammonium bicarbonate, Tris, phenylmethanesulfonyl fluoride (PMSF), $CaCl_2$, and horse heart myoglobin were purchased from Sigma-Aldrich (St. Louis, MO). Sequencing grade trypsin was obtained from Promega (Madison, WI). HPLC grade formic acid was from Panreac (Barcelona, Spain). HPLC-grade acetonitrile was obtained from Tedia (Fairfield, OH). PDAATAAPLR (MW 982.09) was a gift from the Biology Department, Hong Kong University of Science and Technology. Octreotide was obtained from GL Biochem (Shanghai) Ltd. Water was purified from a Milli-Q Ultrapure water system (Millipore, Billerica, MA).

Derivatization of Model Peptides

For the derivatization of PDAATAAPLR, the sample (1 nmol) was dissolved in 10 μL 50% ACN, and then mixed with iodoacetamide (50 μL, 55 mM) in ammonium bicarbonate buffer (pH 7.8, 50 mM). The reaction mixture was allowed to stay at 37 °C in dark place for 60 min. The cyclic octreotide was derivatized under the same conditions. For the reduction of octreotide, a mixture of octreotide (1 nmol, 10 μL 50% ACN), and 1,4-dithioerythritol (50 μl, 10 mM) in ammonium bicarbonate buffer (pH 7.8, 50 mM) was kept at 37 °C for 60 min. The reduced octreotide was then derivatized with the addition of iodoacetamide (50 μL, 55 mM in 50 mM pH 7.8 ammonium bicarbonate buffer) and incubation for additional 60 min in the dark.

Derivatization of Protein Digests

Myoglobin (5 nmol) were first reduced by 10 mM 1,4-dithioerythritol for 60 min at 56 °C, alkylated with 55 mM iodoacetamide for 60 min in the dark at 37 °C, and then digested overnight at 37 °C in dark place with modified trypsin (Promega, sequencing-grade). The reaction was stopped by addition of formic acid to 5% final. The resulting tryptic peptides were then subject to mass spectrometric analysis or stored in −20 °C for future use.

For the preparation of digests of rat heart tissues, hearts were excised from 4 wk-old male Sprague-Dawley rats. After washed in PBS-buffered saline, ventricular tissues were immediately frozen in liquid nitrogen and ground to a fine powder. The powdered tissues were homogenized using a Polytron in solubilization buffer composed of 9.5 M urea, 20 mM Tris (pH 8.0), and 0.1 mM PMSF. The crude extract was then centrifuged at 16,000 g at 10 °C for 1 h. The resultant supernatant was then reduced by adding DTT to 10 mM for 60 min, and carbamidomethylated in 50 mM iodoacetamide at 37 °C for 60 min in the dark. The reduced and alkylated sample was then diluted with 4 volumes of 100 mM ammonium bicarbonate buffer (pH 7.8) containing 1 mM $CaCl_2$, followed by adding modified trypsin (Promega, sequencing-grade) at an estimated enzyme-to-substrate ratio of 1:100 (wt/wt). The mixture was allowed to stand for over 24 h at 37 °C with occasional shaking. After incubation, the reaction was stopped by adding formic acid to 5% final and stored at −20 °C until analysis.

Capillary LC-MS/MS

Capillary LC-MS/MS experiments were conducted on an ABI QSTAR spectrometer coupled with an Agilent 1100 system with a micro-autosampler. Approximate 200 pmol of each of the model peptides and the tryptic digest of myoglobin was injected. For the analysis of the whole protein extract of rat heart tissue, approximate 1.0 mg of digests was used. Samples were injected directly onto a 0.3 mm × 25 cm column packed with C18 material (LC Packings C18 PepMap 100, 5 μm, 100 A; LC Packings, Amsterdam, The Netherlands). Peptides were eluted at a flow rate of 10 μL/min with mobile phase A (99% water, 1% acetonitrile, 0.1% formic acid), to which mobile phase B (99% acetonitrile, 1% water, 0.1% formic acid) was added by a linear gradient (initially 1% B for 2 min, and then increased to 40% B at 100 min, to 95% B at 110 min, and then kept isocratic for 10 min).

Mass spectrometric analysis was performed on an ABI QSTAR spectrometer in information dependent acquisition mode (IDA) (Analyst QS, Applied Biosystems), selecting the four most intense ions for MS/MS analysis. A survey scan of 300 to 2000 Da was collected for 3 s followed by 5 s MS/MS scans of 40 to 1500 Da using the standard rolling collision energy settings.

Database Searching

In the MS/MS ion search mode, mascot generic files were generated by using a script embedded in the Analyst QS 1.1 software (MDS SCIEX) and then used to search against the SwissProt database in the entries of Rattus or Chordata on a local Mascot server with the following parameters: one missed cleavage; peptide tolerance, 0.2 Da; MS/MS tolerance, 0.2 Da; cysteine carbamidomethylation as fixed modification; carbamidomethyl

(N-term), carbamidomethyl (K), and methionine oxidation as variable modifications.

In the sequence query mode, the N-terminal sequence information was combined with the MS/MS data to search against the SwissProt database. In this type of search, the mass list in raw MS/MS data was uploaded in an ions qualifier, whereas the N-terminal sequence information derived from a_1 enhancement was included in a sequence qualifier.

Results and Discussion

Enhancement of a_1 Ions in N-carbamidomethyl Derivatives of Model Peptides

Iodoacetamide, a widely used derivatization reagent in proteomics for the alkylation of sulfhydride groups in peptides and proteins, was used to react with the N-terminal amino groups in peptides, thus leading to formation of carbamidomethyl derivatives at peptide N-termini. These carbamidomethyl derivatives were commonly viewed as undesired by-products formed during iodoacetamide derivatization of reduced cysteine residues in peptides [21, 22]. During the mass spectrometric analysis of these derivatives, however, an enhancement of the modified a_1 ions was observed in the product ion spectra.

Figure 1 exhibits an LC-MS/MS analysis of a model peptide PDAATAAPLR after the reaction with iodoacetamide at 37 °C for 60 min. The chromatographic peak eluted at 30.25 min showed the doubly charged molecular ion of the underivatized peptide at m/z 491.8 (Figure 1a). The other peak eluted at 31.15 min with an

m/z value of 520.3 represents the doubly charged molecular ion of its carbamidomethyl derivative. A difference of 57 Da in the molecule weight was observed between the above two compounds. Figure 1b and c show the product ion spectra recorded following low-energy collision fragmentation of the $[M + 2H]^{2+}$ ions of the peptide PDAATAAPLR and its carbamidomethyl derivative, respectively. In the product ion spectrum of underivatized peptide PDAATAAPLR, several a series ions such as a_1 and a_2 were detected in the low mass range of the spectrum, along with the b series ions (except for b_1 ion). Interpretation of the product ion spectrum corresponding to the peak at 31.15 min confirmed that the compound represented the N-carbamidomethyl derivative of PDAATAAPLR because the m/z values of the detected y series ions were the same as those observed for the underivatized PDAATAAPLR, while the mass values of both a and b series ions were 57 Da higher than those detected in the underivatized peptide. Noticeably, an enhancement of the modified a_1 ion at m/z 127 was observed in the product ion spectrum, whereas the patterns of fragmentation pathway and relative intensity of other series ions were almost the same as those observed in the product ion spectrum of the underivatized peptide. The analytical results from the specific enhancement of a_1 ion with N-terminal modification by iodoacetamide readily provided information of N-terminal amino acid residue of the peptide.

Applicability of the method developed with the enhancement property of the a_1 ion of the N-carbamidomethyl derivative of PDAATAAPLR was extended to other peptides. Figure 2 shows the total ion chromatograph (TIC)
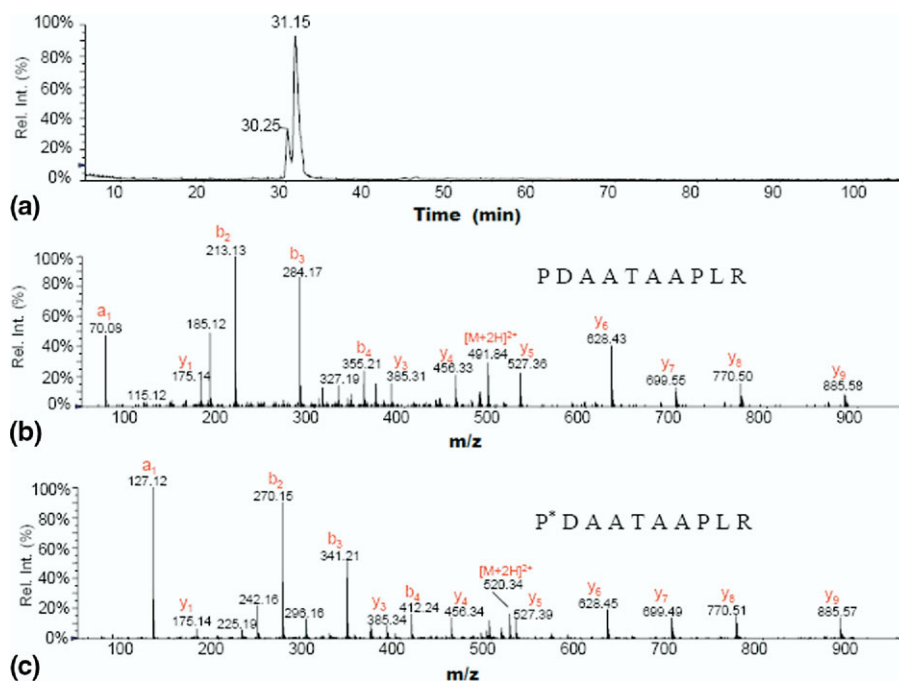

Figure 1. LC-MS/MS analysis of PDAATAAPLR after the reaction with iodoacetamide: (**a**) TIC as well as product ion spectra for (**b**) m/z 491.8 and (**c**) m/z 520.3.

and extracted ion chromatograph (XIC) from the LC-MS/MS analysis of carbamidomethyl derivatives of octreotide that is a cyclic octapeptide with a disulfide bond linked between the two cysteines and with a monoisotopic molecular weight of 1018.4. The detailed structures of octreotide and its reduced form were presented in Figure S1 in supplementary materials, which can be found in the electronic version of this article. The octreotide derivatives were formed from the incubation of octreotide in bicarbonate buffer containing iodoacetamide at 37 °C for 60 min. The XIC chromatograph extracted at m/z 510.2 represented the peak of the doubly charged molecular ion of octreotide (Figure 2b). The product ion spectrum (Figure 3a) of the underivatized octreotide was very simple under low collision energy due to its cyclic structure. The fragment ion at m/z 120 was the immonium ion of the N-terminal Phe residue, whereas the ion at m/z 106 was the y_1 ion of octreotide. The XIC chromatograph in Figure 2c, extracted at m/z 538.7, showed two peaks at 11.43 and 12.02 min, respectively, indicating that two different forms of mono-carbamidomethyl derivatives of octreotide might exist in the reaction mixture. The product ion spectrum (Figure 3b) of the peak eluted at 11.43 min with m/z 538.7 showed a similar fragmentation patterns as that of the underivatized octreotide (Figure 3a), indicating that the carbamidomethyl moiety was not added on the amino group of the N-terminal Phe residue. The other residue with possibility to react with iodoacetamide in this peptide was Lys, which had an ϵ amino group in the side chain. Similar to the underivatized peptide, a_1 ion at m/z 120 was observed as base peak. In the product ion spectrum of the peak eluted at 12.02 min with m/z 538.7 (Figure 3c), on the other hand, the intensive a_1 ion peak of carbamidomethyl derivative

was detected at m/z 177, which was 57 Da higher than the ion at m/z 120. The observation of this ion allowed the confirmation of the N-terminal Phe residue.

Besides the ion peaks at m/z 510.2 and 538.7, a minor doubly charged ion peak at m/z 567.2 was detected from the analysis of the reaction mixture, indicating that the derivatized octreotide might have two carbamidomethyl moieties. The XIC chromatograph (Figure 2d) at m/z 567.2 shows a peak at 11.61 min, which had the same elution time as that of the underivatized octreotide (Figure 2b). The product ion spectrum (Figure 3d) of m/z 567.2 was very similar to that of the peak eluted at 12.02 min with m/z 538.2 (Figure 3c). Again, the detection of the intensive ion peak at m/z 177 indicated that the amino group of N-terminal Phe was modified by iodoacetamide, while the other carbamidomethyl moiety might be introduced at another site, probably in the ϵ amino group of Lys residue.

In addition to the cyclic octreotide, the developed method was also applied for the investigation of the reduced octreotide. The enhancement of the a_1 ion in product ion spectra was again achieved from the carbamidomethyl derivatization with iodoacetamide (Figure 4). Three different reaction products of carbamidomethyl derivatives were detected in the reaction mixture of the reduced octreotide. Results of the MS/MS analyses of the carbamidomethyl derivatives of the reduced octreotide exhibited the series of b and y ions, which provided the sequence information of the peptides.

Figure 4a showed the product ion spectrum of one of three carbamidomethyl derivatives, which had the m/z value at 568.3. The MS/MS analysis and spectrum interpretation allowed the identification of the S-carbamidomethyl derivative resulted from the alkylation at sulfhydride
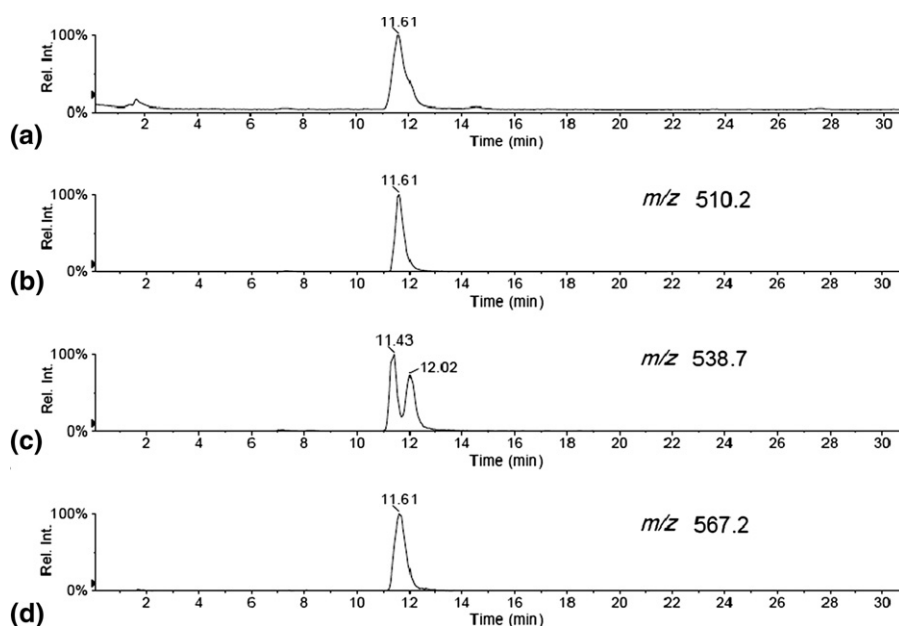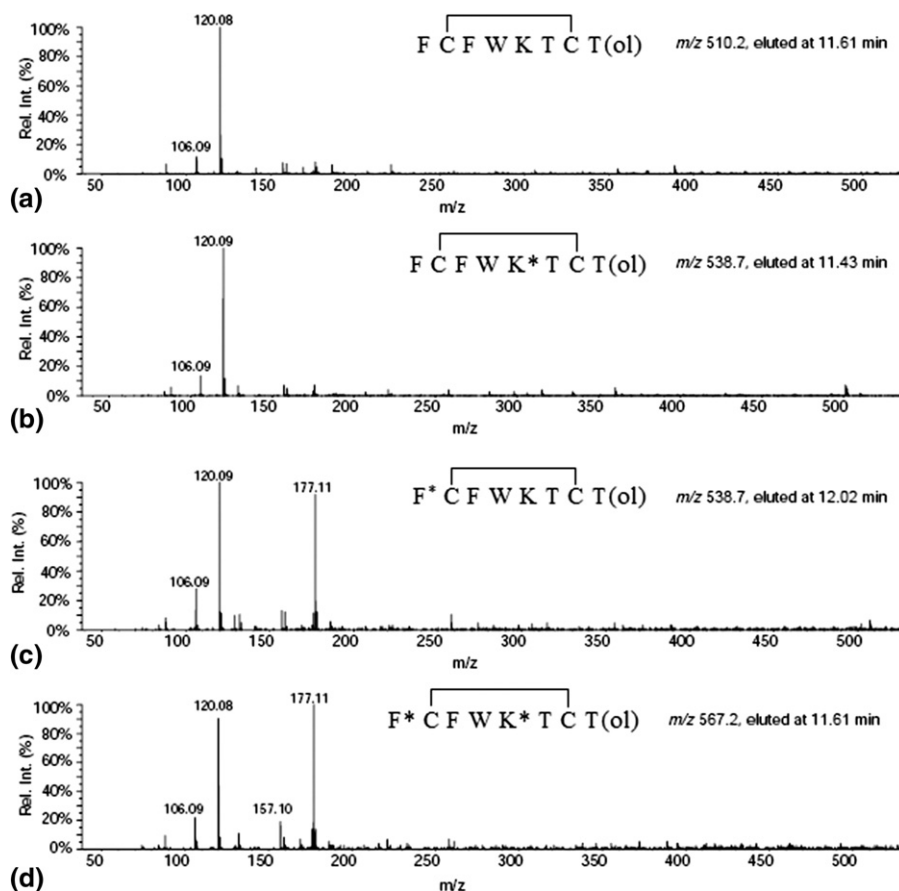

Figure 2. LC-MS/MS analysis of octreotide after the reaction with iodoacetamide: (**a**) TIC as well as XIC for (**b**) m/z 510.2, (**c**) m/z 538.7, and (**d**) m/z 567.2.
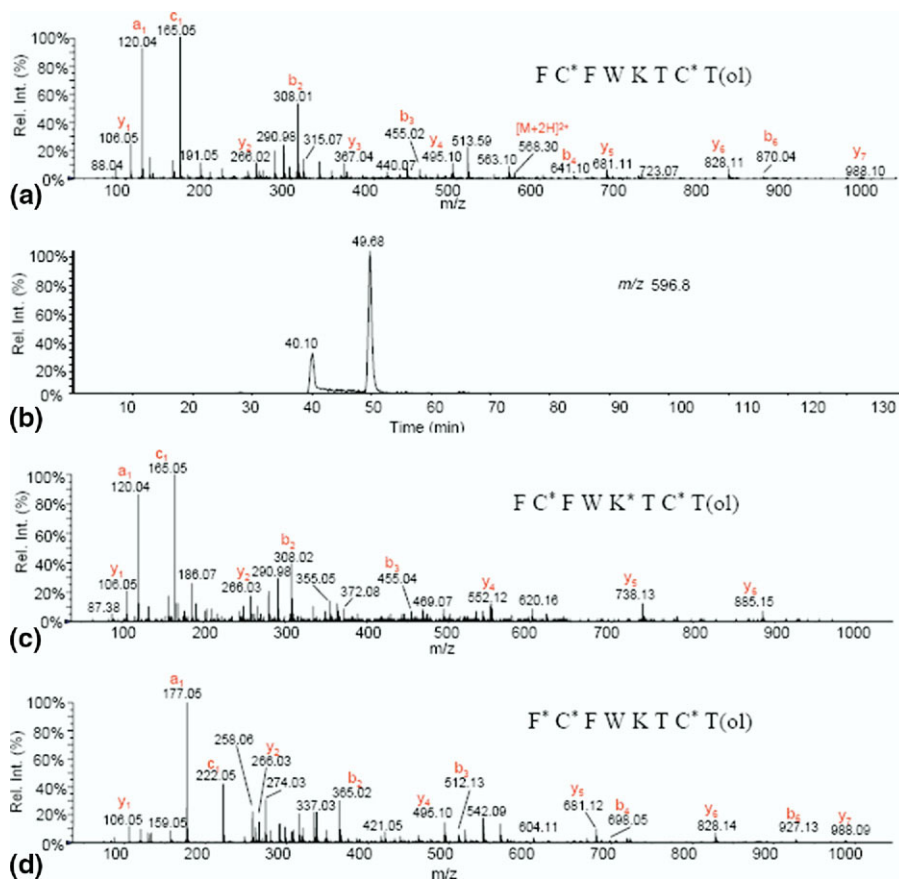
Figure 3. LC-MS/MS analysis of octreotide after the reaction with iodoacetamide: product ion spectrum for (**a**) m/z 510.2 eluted at 11.61 min, (**b**) m/z 538.7 eluted at 11.43 min, (**c**) m/z 538.7 eluted at 12.02 min, and (**d**) m/z 567.2 eluted at 11.61 min.

groups of the two Cys residues. The other two reaction products with the same m/z value of 596.8 but different elution time (Figure 4b) were also investigated. The product ion spectrum of the minor peak eluted at 40.2 min (Figure 4c) indicated that this compound was a carbamidomethyl derivative with two S-carbamidomethylated cysteines and one N-carbamidomethylated lysine. The very similar spectra in the low end ($m/z < 200$) in Figure 4a and c suggested that the N-terminal Phe residues in the two compounds were underivatized. However, in the product ion spectrum of the major peak of m/z 596.8 eluted at 49.6 min, the a_1 ion was greatly enhanced after the N-terminal derivatization with iodoacetamide. The observation of the derivatized a_1 ion at m/z 177 (Figure 4d) suggested that this compound might be an N-terminally carbamidomethylated derivative with additional two S-carbamidomethylated cysteines. Noticeably, c_1 ions at m/z 165 (underivatized) and m/z 222 (derivatized) were observed with high abundance, confirming the N-terminal phenylalanine residue.

The formation of a_1 ions as immonium species has been considered to occur via either direct dissociation of b_2 to a_1 or indirect fragmentation of a_2 to a_1 [10]. It was found in this study that a_1 ion of a peptide could be remarkably enhanced and stabilized when a carbamidomethyl moiety is introduced at the amine of the N-terminus of peptide, which could be easily rationalized as the proton stabilizing effect of amino group in the carbamidomethyl moiety. Under the mild reaction conditions, N-carbamidomethylation by iodoacetamide was inclined to take place at α amino groups in the N-terminal residues. Even though the epsilon amino group of the lysine residue and the sulfhydride group of cysteine might be other liable moieties, the possible modified immonium ions with m/z values of 158.1 (modified immonuim ion of lysine) and 133.1 (modified immonuim ion of cysteine) were not detected under the low-energy collision condition, which allowed the low end of the product ion spectrum unchanged after carbamidomethyl derivatization at internal residues such as lysine and cysteine. Thus, the specificity of enhancement of a_1 ion by carbamidomethyl derivatization at α amino groups in the N-terminal residues was high and useful in mass spectrometric sequencing of peptides.

The carbamidomethyl derivatization with iodoacetamide appeared to be incomplete under the mild conditions used in the present study. The incompleteness of the reaction allowed the direct comparison of the MS/MS spectra of the modified peptides with those of the unmodified peptides, through which the enhanced

Figure 4. LC-MS/MS analysis of the reduced octreotide after the reaction with iodoacetamide: product ion spectrum for (**a**) m/z 568.3 eluted at 46.70 min, (**c**) m/z 596.8 eluted at 40.10 min, (**d**) m/z 596.8 eluted at 49.68 min and XIC for (**b**) m/z 596.8.

a₁ ion could be assigned and therefore the N-terminal amino acid residue could be determined. Specifically, when the peptides were derived from tryptic digestion, the comparison of the spectra was straightforward because the fragment ions at high-end of the MS/MS spectra, most of which belonged to y series ions, remained unchanged after N-terminal carbamidomethylation. It should be noted that the a₁ ion in the MS/MS spectrum of a modified peptide could not be assigned without the comparison with that of the intact peptide because other immonium ions from internal residues were usually intermixed with the enhanced a₁ ion in the low end of MS/MS spectrum.

Applications of Enhancement of a₁ Ions by N-carbamidomethyl Derivatization for Protein Identification

Although the reduction of disulfides by DTT followed by alkylating the sulfhydryl groups with iodoacetamide and the subsequent enzymatic digestion followed by LC-MS/MS analysis have been widely applied for protein identification [23–27], observation of the enhanced a₁ ions from N-terminal carbamidomethylation and its application for peptide sequencing and protein

identification have not been reported. Because the excess iodoacetamide could readily react with the α amino groups of the N-terminal residues, the carbamidomethylation was observed as previously reported [22] and should be included in the combination of variable modifications during the database searching process.

The method was then applied to the analysis of the tryptic digest of myoglobin. The tryptic digest of this model protein was prepared in a manner that was very similar with the commonly used protocol in which iodoacetamide was used to alkylate sulfhydryl groups of Cys after reduction with DTT. The excess iodoacetamide was usually left in the digestion mixture. Moreover, the digestion conditions (37 °C, pH 7.5–8.0 and 16–20 h) favor the reaction of iodoacetamide with amino groups in the peptides.

Table 1 summarized results from the LC-MS/MS analysis of tryptic digest of myoglobin. A total of 17 unique sequences in myoglobin were identified with sequence coverage of 96.7%. For most of the sequences identified, signals of both underivatized peptides and their carbamidomethyl derivatives were detected. Interpretation of the MS/MS spectra of the carbamidomethyl derivatives indicated that most carbamidomethylation took place at N-termini of the peptides, whereas

Table 1. Summary of LC-MS/MS analysis of tryptic digest of myoglobin

| | | m/z (Monoisotopic) and charges | |
Position	Peptide sequence	Unmodified	Carbamidomethyl derivatives
1–16	GLSDGEWQQVLNVWGK	605.96(3+)	936.95(2+)
17–31	VEADIAGHGQEVLIR	803.92(2+), 536.28(3+)	832.40(2+), 555.29(3+)
32–42	LFTGHPETLEK	636.32(2+), 424.56(3+)	664.82(2+)[b], 443.57(3+)[b]
43–45	FDK	409.20(1+)	466.21(1+)
46–47	FK	294.18(1+)	351.19(1+)
48–50	HLK	397.24(1+)	454.26(1+)
51–56	TEAEM[a]K	724.28(1+)	781.32(1+)
57–63	ASEDLKK	395.71(2+)	424.22(2+)
64–77	HGTVVLTALGGILK	689.90(2+)	718.40(2+)[b]
80–96	GHHEAELKPLAQSHATK	618.64, 464.23	637.64, 478.50
99–102	IPIK	470.31(1+)	527.35(1+)
103–118	YLEFISDAIIHVLHSK	942.98(2+), 629.01(3+)	971.50(2+), 648.00(3+)
120–133	HPGDFGADAQGAMTK	751.81(2+), 501.54(3+)	780.31(2+), 520.56(3+)
134–139	ALELFR	748.42(1+), 374.72(2+)	805.43(1+), 403.23(2+)
140–145	NDIAAK	631.32(1+), 316.17(2+)	ND[c]
146–147	YK	310.17(1+)	367.18(1+)
148–153	ELGFQG	650.29(1+)	ND[c]

[a]Oxidized methionine (M).
[b]Both the α amino group of N-terminal residue and the epsilon amino group of Lysine (K) were carbamidomethylated.
[c]Not detected.

the carbamidomethylation at ε amino groups of Lys residues were only found in two sequences, namely peptide 32–42 and peptide 64–77 listed in Table 1. Figure 5 showed the product ion spectra of an intact tryptic peptide (ALELFR) from the doubly charged ion at m/z 374.7 (Figure 5a) and the corresponding carbamidomethyl derivative from the doubly charged ion at m/z 403.2 (Figure 5b). As expected, a modified a$_1$ ion at m/z 101.1 was observed with high abundance in the product ion spectrum of the doubly charged carbamidomethyl derivative at m/z 403.2, confirming that the N-terminal residue in this peptide was Ala. The enhancements of a$_1$ ions were also found in the MS/MS spectra of other N-terminally carbamidomethyl derivatives, allowing the rapid determination of the N-terminal residues of these tryptic peptides with high specificity. It should be pointed out that at high-temperature (e.g., 60 °C) the

reaction would not only produce N-terminal carbamidomethyl derivatives, but also generate several other types of derivatives resulted from carbamidomethylation at Arg, His, or from O-carbamidomethylation at C-terminal carboxylic group [21]. Nielsen et al. also reported the observation of alkylation of lysine with 2 mol iodoacetamide [28]. These unwanted derivatives made the MS/MS spectra interpretation more difficult. The problem, however, was not detected under the mild reaction conditions used in this study. When the concentration of iodoacetamide was 5 mM, the N-terminal derivatization reaction appeared to be very slow. On the other hand, no difference in the reactions was observed when the concentration of iodoacetamide was 55 mM and 100 mM. Therefore, the concentration of iodoacetamide at 55 mM, which has been used routinely in most proteomics labs, was used in this study.
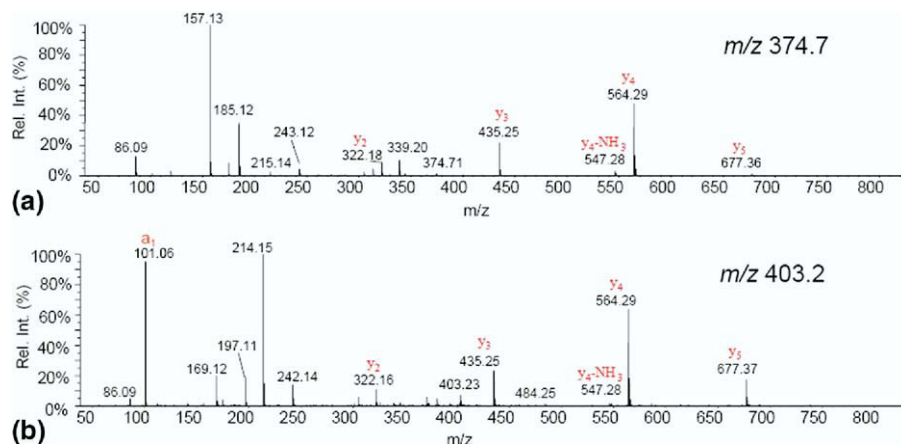
Figure 5. LC-MS/MS analysis of a peptide (ALELFR) and its N-terminal carbamidomethyl derivative from the tryptic digest of myoglobin: product ion spectrum for (a) m/z 374.7 and (b) m/z 403.2.
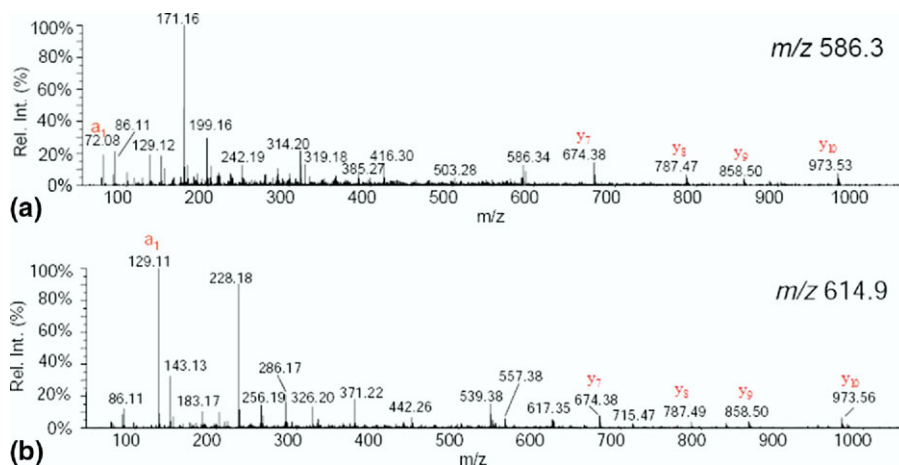
The application of N-terminal carbamidomethylation in protein identification by database searches after the LC-MS/MS analysis of tryptic digest of a whole protein extract from rat heart was subsequently performed. After recorded and converted into the corresponding MGF files, the mass spectrometric data were loaded into an in-house Mascot server to search in the Rattus database with or without variable modifications of carbamidomethyl (N-term) and carbamidomethyl (K). The detailed results for the identification of the first 20 hits are listed in Table S1 of the supplementary materials. Most peptides in the identified proteins were found to undergo N-terminal carbamidomethylation to some extent. An example for the identification of cardiac actin was demonstrated with the result shown in Figure S2 in supplementary materials. A Total of 17 unique peptides were identified, 10 of which were N-terminally carbamidomethylated. Enhancements of a_1 ions were found in the MS/MS spectra of the N-terminal derivatives, allowing determination of N-terminal residues and thus confirming the results from database search.

It should be pointed out that the N-terminal carbamidomethyl derivatization was seldom found in mass spectrometric analysis of samples from in-gel digests, since the iodoacetamide was often washed away before the enzyme was added. However, the N-terminal carbamidomethylation readily occur during the analysis of the samples prepared by in-solution digestion. The excess iodoacetamide from the process of sample preparation would readily react with the N-terminal α amino acid groups, and thus the N-terminal carbamidomethylation should be considered as a variable modification during the database searching process. Even though the inclusion of N-terminal alkylation contributed insignificant gains in peptide identifications [29], the developed method could be applied to validate the identified peptide sequences obtained from ordinary MS/MS ion database search.

Added Specificity of Protein Identification through the Detection of N-Terminal Tags

The term "sequence tag" was first proposed by Mann and Wilm at EMBL [30], which was used as the combination of a few residues of sequence data for protein identification through database search. Sequence tag search algorithm has since been well developed and used as a powerful tool for protein identification though database search [31–33]. As shown in the above examples, the N-terminal information in peptides could be obtained from the detection of the enhanced a_1 ions by N-terminal carbamidomethyl derivatization. This information of N-terminal residue in certain peptides could be viewed as a special sequence tag, which would increase the specificity of protein identification during database searching by narrowing down the candidate list.

As an example, a peptide derived from ATPase was identified when the information of its a_1 ion was included in the database search. Figure 6a and b showed the two product ion spectra from the ions at m/z 586.4 and m/z 614.9, respectively. While molecular weight of the two peptides showed a difference of 57 Da, the same y series ions were observed in the two MS/MS spectra in the high mass end, which suggested that the fragmentations were produced from the same sequence except that one of the peptides was N-terminally carbamidomethylated. The mass list of product ion spectrum of the ion at m/z 586.4 (Figure 6a) was generated by the script built in the Analyst QS, and searched against the SwissProt database in the sequence query mode. A score of 15 for a peptide (VVDALGNAIDGK) derived from ATP synthase was obtained, whereas the individual ions scores indicating identity or extensive homology ($P < 0.05$) were above 19 (Figure S2 in supplementary materials). So the result was considered as negative, or not confident.

Figure 6. LC-MS/MS analysis of a peptide (VVDALGNAIDGK) and its N-terminal carbamidomethyl derivative from the tryptic digest of rat heart tissues: product ion spectrum for (**a**) m/z 586.3 and (**b**) m/z 614.9.

Fortunately, in Figure 6b, a great enhancement of the ion at m/z 129 was observed in the product ion spectrum of the signal at m/z 614.9, indicating that the N-terminal amino acid was valine. Therefore, this N-terminal tag was combined with the above mass list to re-search against the SwissProt database using the same searching parameters as above. The resultant score for the peptide (VVDALGNAIDGK) was 15, which was the same as that obtained without inclusion of the information of the N-terminal amino acid. However, in this case, because the individual ions score indicating the identity or extensive homology ($P < 0.05$) was lowered to 10 due to the inclusion of N-terminal tag, the result turned out to be positive, indicating that peptide (VVDALGNAIDGK) was confidently identified (Figure S3 in supplementary materials).

The sequence query search algorithm was often initiated with the determination of certain sequence tag, which was usually obtained by de novo sequencing. Therefore, the a_1 ion assignment by N-terminal carbamidomethylation was likely suitable for the interpretation of MS/MS data by de novo sequencing or de novo sequencing-based search algorithms (e.g., sequence query in MASCOT).

Conclusions

A simple and effective method to enhance a_1 ions in tandem mass spectrometry analysis with low collision energy was developed by using carbamidomethylation at α amino moieties of N-terminal amino acids in peptides. The N-terminal carbamidomethylation by iodoacetamide was demonstrated to be compatible with the sample preparation protocol in proteomics analysis. The method was effective and specific due to the proton stabilizing effect of amino group in the carbamidomethyl moiety. The developed method was proven to be useful for de novo peptide sequencing with mass spectrometric analysis. The method could also be considered as a major artificial modification when the peptide digests were prepared in an environment with excess iodoacetamide for a long time, similar to the widely used protocol for in-solution digestion. Moreover, the added specificity of protein identification could be achieved when the information of N-terminal residues was included in the database searches. The method may be extended to the application for quantitative proteomics when the stable isotope labeled iodoacetamide is available for the sample preparation.

Acknowledgments

Appendix A.
Supplementary Material

References

1. Nesvizhskii, A. I. Protein Identification by Tandem Mass Spectrometry and Sequence Database Searching. *Methods Mol. Biol.* **2007,** *367,* 87–119.
2. Biemann, K. Contributions of Mass Spectrometry to Peptide and Protein Structure. *Biomed. Environ. Mass Spectrom.* **1988,** *16,* 99–111.
3. Roepstorff, P.; Fohlman, J. Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomed. Mass Spectrom.* **1984,** *11,* 601.
4. Johnson, R. S.; Martin, S. A.; Biemann, K. Collision-Induced Fragmentation of $(M + H)^+$ Ions of Peptides. Side Chain Specific Sequence Ions. *Int. J. Mass Spectrom. Ion Processes* **1988,** *86,* 137–154.
5. Ambihapathy, K.; Yalcin, T.; Leung, H.-W.; Harrison, A. G. Pathways to Immonium Ions in the Fragmentation of Protonated Peptides. *J. Mass Spectrom.* **1997,** *32,* 209–215.
6. Ji, C.; Li, L. Quantitative Proteome Analysis Using Differential Stable Isotopic Labeling and Microbore LC-MALDI MS and MS/MS. *J. Proteome Res.* **2005,** *4,* 734–742.
7. Alves, G.; Wu, W. W.; Wang, G.; Shen, R. F.; Yu, Y. K. Enhancing Peptide Identification Confidence by Combining Search Methods. *J. Proteome Res.* **2008,** *7,* 3102–3113.
8. Sweet, S. M.; Mardakheh, F. K.; Ryan, K. J.; Langton, A. J.; Heath, J. K.; Cooper, H. J. Targeted Online Liquid Chromatography Electron Capture Dissociation Mass Spectrometry for the Localization of Sites of in Vivo Phosphorylation in Human Sprouty2. *Anal. Chem.* **2008,** *80,* 6650–6657.
9. Shen, Y.; Tolić, N.; Hixson, K. K.; Purvine, S. O.; Pasa-Tolić, L.; Qian, W. J.; Adkins, J. N.; Moore, R. J.; Smith, R. D. Proteome-Wide Identification of Proteins and Their Modifications with Decreased Ambiguities and Improved False Discovery Rates Using Unique Sequence Tags. *Anal. Chem.* **2008,** *80,* 1871–1882.
10. El Aribi, H.; Rodriquez, C. F.; Almeida, D. R.; Ling, Y.; Mak, W. W.; Hopkinson, A. C.; Michael Siu, K. W. Elucidation of Fragmentation Mechanisms of Protonated Peptide Ions and Their Products: A Case Study on Glycylglycylglycine Using Density Functional Theory and Threshold Collision-Induced Dissociation. *J. Am. Chem. Soc.* **2003,** *125,* 9229–9236.
11. Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. The Structure and Fragmentation of Bn ($n \geq 3$) ions in peptide spectra. *J. Am. Soc. Mass Spectrom.* **1996,** *7,* 233–242.
12. Keough, T.; Youngquist, R. S.; Lacey, M. P. A Method for High-Sensitivity Peptide Sequencing Using Post-Source Decay Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96,* 7131–7136.
13. Noga, M. J.; Lewandowski, J. J.; Suder, P.; Silberring, J. An enhanced method for peptides sequencing by N-terminal derivatization and MS. *Proteomics* **2005,** *5,* 4367–4375.
14. Summerfield, S. G.; Bolgar, M. S.; Gaskell, S. J. Promotion and Stabilization of b_1 Ions in Peptide Phenylthiocarbamoyl Derivatives: Analogies with Condensed-Phase Chemistry. *J. Mass Spectrom.* **1997,** *32,* 225–231.
15. Marekov, L. N.; Steinert, P. M. Charge Derivatization by 4-SIulfophenyl isothiocyanate Enhances Peptide Sequencing by Post-Source Decay Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **2003,** *38,* 373–377.
16. Yu, Y.; Cui, J.; Wang, X.; Liu, Y.; Yang, P. Studies on Peptide Acetylation for Stable-Isotope Labeling After 1-D PAGE Separation in Quantitative Proteomics. *Proteomics* **2004,** *4,* 3112–3120.
17. Cárdenas, M. S.; van der Heeft, E.; de Jong, A. P. J. M. On-Line Derivatization of Peptides for Improved Sequence Analysis by Microcolumn Liquid Chromatography Coupled with Electrospray Ionization-Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1997,** *11,* 1271–1278.
18. Huang, Z. H.; Wu, J.; Roth, K. D.; Yang, Y.; Gage, D. A.; Watson, J. T. A Picomole-Scale Method for Charge Derivatization of Peptides for Sequence Analysis by Mass Spectrometry. *Anal. Chem.* **1997,** *69,* 137–144.
19. Boja, E. S.; Sokoloski, E. A.; Fales, H. M. Divinyl Sulfone as a Post-Digestion Modifier for Enhancing the a_1 Ion in MS/MS and Post-Source Decay: Potential Applications in Proteomics. *Anal. Chem.* **2004,** *76,* 3958–3970.
20. Hsu, J. L.; Chen, S. H.; Li, D. T.; Shi, F. K. Enhanced a_1 Fragmentation for Dimethylated Proteins and Its Applications for N-Terminal Identification and Comparative Protein Quantition. *J. Proteome Res.* **2007,** *6,* 2376–2383.
21. Yang, Z.; Attygalle, A. B. LC/MS Characterization of Undesired Products Formed During Iodoacetamide Derivatization of Sulfhydryl Groups of Peptides. *J. Mass Spectrom.* **2007,** *42,* 233–243.

22. Boja, E. S.; Fales, H. M. Overalkylation of a Protein Digest with Iodoacetamide. *Anal. Chem.* **2001,** *73,* 3576–3582.
23. Chong, P. K.; Wright, P. C. Identification and Characterization of the *Sulfolobus solfataricus* P2 Proteome. *J. Proteome Res.* **2005,** *4,* 1789–1798.
24. Zhang, J.; Xu, X.; Gao, M.; Yang, P.; Zhang, X. Comparison of 2-D LC and 3-D LC with Post- and Pre-tryptic Digestion SEC Fractionation for Proteome Analysis of Normal Human Liver Tissue. *Proteomics* **2007,** *7,* 500–512.
25. Hoffert, J. D.; Pisitkun, T.; Wang, G.; Shen, R. F.; Knepper, M. A. Quantitative Phosphoproteomics of Vasopressin-Sensitive Renal Cells: Regulation of Aquaporin-2 Phosphorylation at Two Sites. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103,* 7159–7164.
26. Benschop, J. J.; Mohammed, S.; O'Flaherty, M.; Heck, A. J.; Slijper, M.; Menke, F. L. Quantitative Phosphoproteomics of Early Elicitor Signaling in Arabidopsis. Mol. Cell. Proteom. **2007,** *6,* 1198–1214.
27. Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faergeman, N. J.; Mann, M.; Jensen, O. N. Quantitative Phosphoproteomics Applied to the Yeast Pheromone Signaling Pathway. *Mol. Cell. Proteom.* **2005,** *4,* 310–327.
28. Nielsen, M. L.; Vermeulen, M.; Bonaldi, T.; Cox, J.; Moroder, L.; Mann, M. Iodoacetamide-Induced Artifact Mimics Ubiquitination in Mass Spectrometry. *Nat. Methods* **2008,** *5,* 459–460.
29. Creasy, D. M.; Cottrell, J. S. Error Tolerant Searching of Uninterpreted Tandem Mass Spectrometry data. *Proteomics* **2002,** *2,* 1426–1434.
30. Mann, M.; Wilm, M. Error-Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Anal. Chem.* **1994,** *66,* 4390–4399.
31. Chu, I. K.; Cox, D. M.; Guo, X.; Kireeva, I.; Lau, T. C.; McDermott, J. C.; Michael Siu, K. W. Sequencing of Argentinated Peptides by Means of Matrix-Assisted Laser Desorption/Ionization Tandem Mass Spectrometry. *Anal. Chem.* **2002,** *74,* 2072–2082.
32. Sunyaev, S.; Liska, A. J.; Golod, A.; Shevchenko, A.; Shevchenko, A. MultiTag: Multiple Error-Tolerant Sequence Tag Search for the Sequence-Similarity Identification of Proteins by Mass Spectrometry. *Anal. Chem.* **2003,** *75,* 1307–1315.
33. Liska, A. J.; Sunyaev, S.; Shilov, I. N.; Schaeffer, D. A.; Shevchenko, A. Error-Tolerant EST Database Searches by Tandem Mass Spectrometry and MultiTag Software. *Proteomics* **2005,** *5,* 4118–4122.