Table 1
Nucleotide sequences of oligonucleotides used for the plasmid construction

Primer	Sequence
EcoRV-Δpro	5′- GCGCGATATCATGGTCAGATCATCTTCTCGAACC-3′
EcoRV-3A	GCGCGATATCATGGGCGCAGCAGCATCTGCAGCA
EcoRV-3M	GCGCGATATCATGGGCATGGCAGCATCTGCAGCA
EcoRV-3G	GCGCGATATCATGGGCGGTGCAGCATCTGCAGCA
EcoRV-3S	GCGCGATATCATGGGCTCCGCAGCATCTGCAGCA
EcoRV-3C	GCGCGATATCATGGGCTGCGCAGCATCTGCAGCA
EcoRV-3T	GCGCGATATCATGGGCACCGCAGCATCTGCAGCA
EcoRV-3P	GCGCGATATCATGGGCCCCGCAGCATCTGCAGCA
EcoRV-3V	GCGCGATATCATGGGCGTGGCAGCATCTGCAGCA
EcoRV-3D	GCGCGATATCATGGGCGACGCAGCATCTGCAGCA
EcoRV-3N	GCGCGATATCATGGGCAACGCAGCATCTGCAGCA
EcoRV-3L	GCGCGATATCATGGGCCTGGCAGCATCTGCAGCA
EcoRV-3I	GCGCGATATCATGGGCATCGCAGCATCTGCAGCA
EcoRV-3Q	GCGCGATATCATGGGCCAGGCAGCATCTGCAGCA
EcoRV-3E	GCGCGATATCATGGGCGAGGCAGCATCTGCAGCA
EcoRV-3H	GCGCGATATCATGGGCCACGCAGCATCTGCAGCA
EcoRV-3F	GCGCGATATCATGGGCTTCGCAGCATCTGCAGCA
EcoRV-3K	GCGCGATATCATGGGCAAGGCAGCATCTGCAGCA
EcoRV-3Y	GCGCGATATCATGGGCTACGCAGCATCTGCAGCA
EcoRV-3W	GCGCGATATCATGGGCTGGGCAGCATCTGCAGCA
EcoRV-3R	GCGCGATATCATGGGCCGTGCAGCATCTGCAGCA
B1	GCCGGGATCCTAGGGCGAATTGGGTACC
N-EcoRV-tBID	ATATGATATCATGGGCAACCGCAGCAGC
N-EcoRV-tBIDG2A	ATATGATATCATGGCCAACCGCAGCAGC
BID-STR-C	GCGCGGTACCTCACTTTTCAAACTGCGGATGGCTCCAGTCCATCCCATTTCT
N-BID	GCGCGGATCCATGGACTGTGAG
N-tBID	ATATGGATCCATGGGCAACCGCAGC
BID-C	GCGCGAATTCGTCCATCCCATT

Bid, was constructed by utilizing PCR. For this procedure, pcBid–dsRed (Clontech) served as a template and two oligonucleotides (N-BID, BID-C) as primers. After digestion with *Bam*HI and *Eco*RI, the amplified products were subcloned into pcDNA3–FLAG at the *Bam*HI and *Eco*RI sites. Plasmid pcDNA3tBid–FLAG was constructed by a method similar to that used to construct pcDNA3Bid–FLAG using two oligonucleotides (N-tBID, BID-C) as primers.

Plasmids pcDNA3gelsolin–FLAG and pcDNA3tGelsolin–FLAG were constructed as previously described [11]. The DNA sequences of these recombinant cDNAs were confirmed by the dideoxynucleotide chain termination method [18].

In vitro transcription and translation

The cDNAs were cloned into pTD1 vector (Shimadzu) at a site under the control of the T7 promoter. After digestion with *Hin*dIII, T7 polymerase was used to obtain transcripts of these cDNAs. The transcripts were purified by phenol–chloroform extraction and ethanol precipitation before use in the translation reaction. Subsequently, the translation reaction was carried out using the rabbit reticulocyte lysate (Promega) or insect cell-free protein synthesis system (Shimadzu Co.) in the presence of [^{3}H]leucine, [^{3}H]myristic acid, or [^{35}S]methionine under conditions recommended by the manufacturer. For the rabbit reticulocyte lysate system, the mixture (composed of 17.5 μl of rabbit reticulocyte lysate, 0.5 μl of 1 mM leucine-free amino acid mixture, 2.0 μl of [^{3}H]leucine (2 μCi), 1.0 μl of mRNA (2 μg), and 4.0 μl of distilled water) was incubated at 30 °C for 90 min. For the insect cell-free protein synthesis system, the mixture (composed of 12.5 μl of insect cell lysate, 7.5 μl of reaction buffer, 0.5 μl of 1 mM leucine-free amino acid mixture, 2.0 μl of [^{3}H]leucine (2 μCi), and 2.5 μl of mRNA (5 μg)) was incubated at 25 °C for 3 h. The samples were then analyzed by SDS–PAGE and fluorography.

Immunoprecipitation

In vitro translation products of cDNAs coding for TNF mutants were immunoprecipitated with a specific goat anti-human TNF polyclonal antibody (R&D Systems) as described [14].

SDS–PAGE and fluorography

Samples were denatured by boiling for 3 min in SDS–sample buffer followed by analysis by SDS–PAGE on a 12.5% gel. The gel was fixed and stained with Coomassie brilliant blue (CBB) and then soaked in Amplify for 20 min. The gel was dried under vacuum and exposed to X-ray film for an appropriate period.

Transfection of COS-1 cells and detection of N-myristoylated proteins

The simian-virus-40-transformed African green money kidney cell line COS-1 was maintained in Dulbecco's

modified Eagle's medium (DMEM; Gibco BRL) supplemented with 10% fetal calf serum (FCS; Gibco BRL). Cells (2×10^5) were plated onto 35-mm diameter dishes 1 day before transfection. pcDNA3 construct (2 μg) containing cDNA coding for gelsolin, Bid, or their mutants was used to transfect each plate of COS-1 cells along with 4 μl of LipofectAmine (2 mg/ml; Gibco BRL) in 1 ml of serum-free medium. After incubation for 5 h at 37 °C, the cells were refed with serum-containing medium and incubated again at 37 °C for 48 h. The cells were then washed twice with 1 ml of serum-free DMEM and incubated for 4 h at 37 °C in 1 ml of DMEM with 2% FCS containing [^{3}H]myristic acid (100 μCi/ml). Subsequently, the cells were washed three times with Dulbecco's phosphate-buffered saline (DPBS), collected with a cell scraper, and then lysed with 200 μl of RIPA buffer [50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), proteinase inhibitors] on ice for 20 min. After immunoprecipitation with anti-FLAG antibody, the samples were analyzed by SDS–PAGE and fluorography.

Western blotting

The labeled translation products that were obtained using the insect cell lysate or that were present in the total cell lysate of each group of transfected cells were resolved by 12.5% SDS–PAGE and then transferred to an Immobilon-P transfer membrane (Millipore). After blocking with nonfat milk, the membrane was probed with a specific anti-FLAG antibody as described previously [19]. Immunoreactive proteins were specifically detected by incubation with horseradish-peroxidase-conjugated Protein G (Bio–Rad). The membrane was developed with enhanced chemiluminescence Western blotting reagent (Amersham) and exposed to X-ray film (Kodak).

Results

In vitro translation product could easily be detected by metabolic labeling in the insect cell-free protein synthesis system

The structure and N-terminal amino acid sequence of model N-myristoylated proteins analyzed in this study are shown in Fig. 1. When the mRNA coding for MG3A6S–TNF, a model N-myristoylated protein having an N-myristoylation consensus motif at its N terminus, was translated with rabbit reticulocyte lysate in the presence of [^{3}H]leucine and analyzed by SDS–PAGE and fluorography, a faint smeared protein band with a molecular mass of 17 kDa was detected on the fluorogram after 24 h exposure to the X-ray film as shown in Fig. 2B, lane 4. CBB-staining of the SDS–PAGE gel revealed broadening of the protein bands, probably due to the presence of a large amount of proteins with low molecular weight (~17 kDa) in the rabbit reticulocyte lysate (Fig. 2B, lanes 1 and 2). In fact, the pattern of the

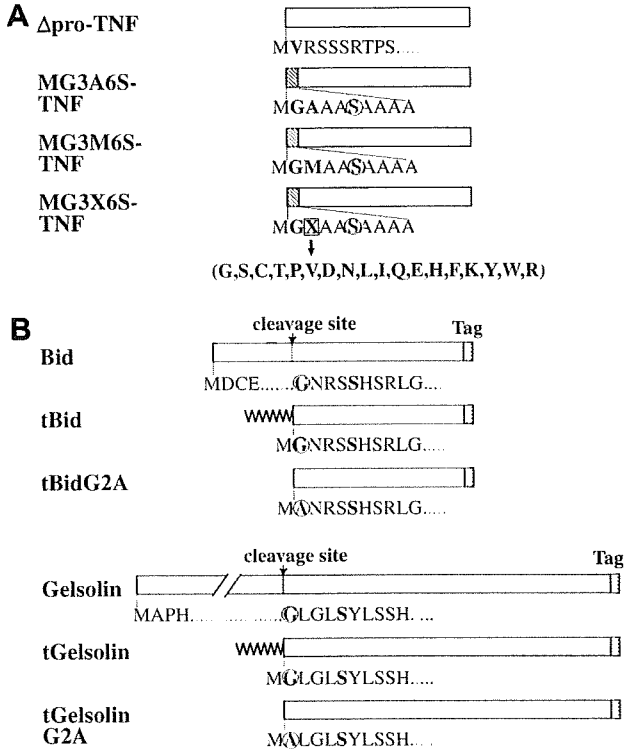

Fig. 1. Structure and N-terminal amino acid sequence of model N-myristoylated proteins analyzed in this study. (A) Model N-myristoylated proteins used to analyze cotranslational protein N-myristoylation. (B) Structure of Bid and gelsolin and their caspase-cleavage products (tBid and tGelsolin) that undergo posttranslational protein N-myristoylation.

protein band detected on the fluorogram became clear when the in vitro translation product was purified by immunoprecipitation using anti-TNF antibody (lane 8). In contrast, when the same experiment was performed using insect cell lysate as the cell-free protein synthesis system, a strong and clear protein band with the same molecular mass was detected without immunoprecipitation as shown in Fig. 2A, lane 4. In accordance with this result, the CBB-staining of the gel revealed the lack of a large amount of proteins with low molecular weight (lanes 1 and 2). As expected, the pattern of the detected protein band was not affected by the immunoprecipitation procedure (lane 8). Thus, the in vitro translation product could easily be detected by metabolic labeling in an insect cell-free protein synthesis system without the need for any purification procedure such as immunoprecipitation.

Detection of cotranslational protein N-myristoylation by metabolic labeling in the insect cell-free protein synthesis system

To determine whether cotranslational protein N-myristoylation could be detected by metabolic labeling in an insect cell-free protein synthesis system, the incorporation of [^{3}H]myristic acid into MG3A6S–TNF was measured. When the mRNA coding for MG3A6S–TNF was

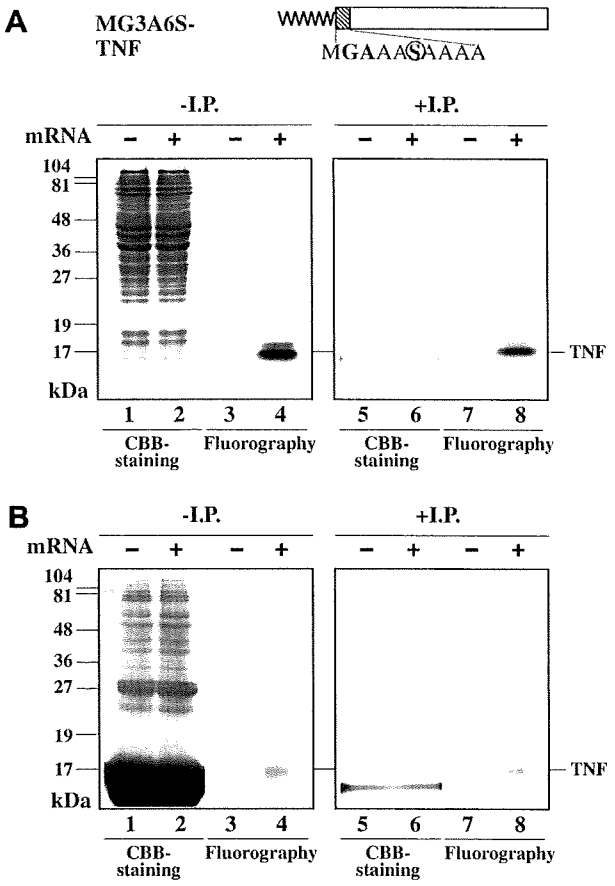

Fig. 2. Detection of in vitro translation product by metabolic labeling in the insect cell-free protein synthesis system. The mRNA coding for MG3A6S–TNF, a model N-myristoylated protein having an N-myristoylation consensus motif at its N terminus, was translated in vitro with insect cell lysate (A) or with rabbit reticulocyte lysate (B) in the presence of [^{3}H]leucine. The labeled translation products were analyzed directly (−I.P.) or following immunoprecipitation with anti-TNF antibody (+I.P.) by SDS–PAGE and fluorography.

translated with the insect cell-free protein synthesis system in the presence of [^{3}H]leucine or [^{3}H]myristic acid, efficient incorporation of both of the ^{3}H-labeled compounds into the translation product was detected on the fluorogram after a brief (10 h) exposure to X-ray film, as shown in Fig. 3B, lanes 2 and 4. In contrast, when the rabbit reticulocyte lysate was used, faint, broad, and retarded protein bands were detected even after 36 h of exposure to the X-ray film (lanes 5 and 7). These results indicate that cotranslational protein N-myristoylation could easily and rapidly be detected by metabolic labeling in an insect cell-free protein synthesis system.

Detection of removal of the initiating Met residue during cotranslational protein N-myristoylation reaction

In the cotranslational protein N-myristoylation reaction, the initiating Met is removed from the N terminus before the attachment of myristic acid to the Gly residue at position 2. The amino acid sequence requirements for removal of initiating Met have been well characterized. The amino acid residue next to the initiating Met is the determinant for this reaction. The seven amino acids (Gly, Ala, Ser, Cys, Thr, Pro, Val) having the smallest radii of gyration are substrates for methionine aminopeptidases, while those with the 13 largest side chains are not [20]. To detect the removal of the initiating Met residue, differential labeling of the in vitro translated protein bands with [^{3}H]leucine, [^{35}S]methionine, and [^{3}H]myristic acid was performed. In this experiment, two other TNF mutants, Δpro–TNF and MG3M6S–TNF, were used to characterize the labeling pattern of the protein bands. The N-terminal sequences of Δpro–, MG3A6S–, and MG3M6S–TNF are shown in Fig. 4. In these three mutants, the amino acid residue next to the initiating Met is Val, Gly, and Gly, respectively. Therefore, it is anticipated that the initiating Met of these three mutants will be removed by methionine aminopeptidase. In Δpro–TNF and MG3A6S–TNF, the initiating Met is the only Met residue in the entire molecule. In MG3M6S–TNF, Met is present at position 3 in addition to the initiating Met. MG3A6S– and MG3M6S–TNF contain an N-myristoylation consensus motif but Δpro–TNF does not. When the mRNA coding for Δpro–TNF was translated, a single protein band with a molecular mass of 17 kDa was labeled with [^{3}H]leucine, as shown in the upper panel of Fig. 4. As expected, neither [^{35}S]methionine nor [^{3}H]myristic acid was incorporated into this mutant. In the case of MG3A6S–TNF, a major protein band with a molecular mass of 17 kDa was labeled with [^{3}H]leucine and [^{3}H]myristic but not with [^{35}S]methionine (middle panel of Fig. 4). When MG3M6S–TNF, which has an intramolecular Met residue, was labeled, the protein band was labeled with all three labeled compounds, as shown in the lower panel of Fig. 4. These results clearly indicated that the initiating Met is removed from the N terminus before the attachment of the myristic acid to the protein. It was also found that the minor protein band present in the in vitro translation products of MG3A6S–TNF is a protein species retaining the initiating Met and is not N-myristoylated. Thus, the removal of the initiating Met residue during the cotranslational protein N-myristoylation reaction could be detected by the labeling of the in vitro translation products with [^{35}S]methionine and [^{3}H]leucine.

Substrate specificity of NMT in insect cell lysate is similar to that in rabbit reticulocyte lysate

To determine whether the substrate specificity of NMT in the insect cell lysate was similar to that in rabbit reticulocyte lysate, the susceptibility of a series of TNF mutants (MG3X6S–TNF) to cotranslational protein N-myristoylation was evaluated by metabolic labeling in the insect cell lysate and rabbit reticulocyte lysate. MG3X6S–TNF is a series of model N-myristoylated proteins in which the Ala residue at position 3 in a model N-myristoylation motif (MGAAASAAA) in MG3A6S–TNF is

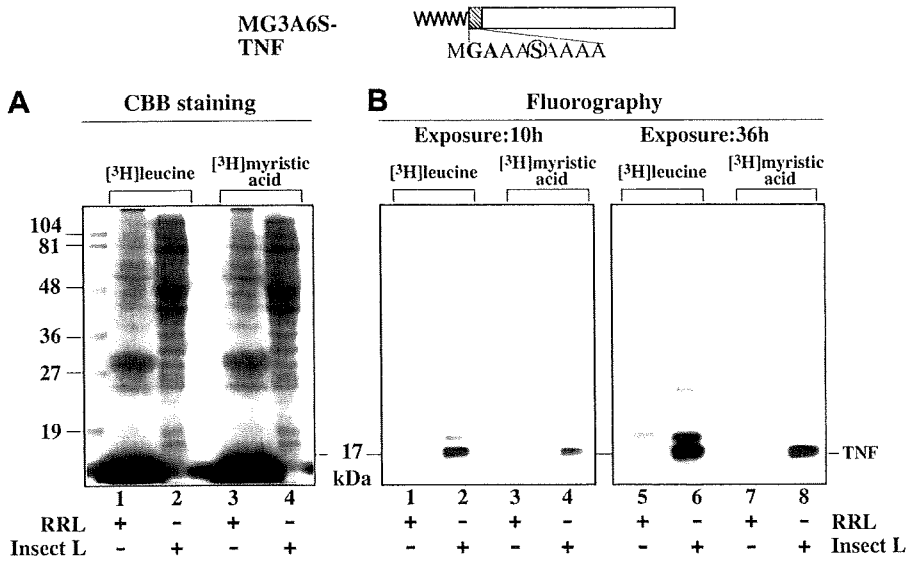

Fig. 3. Detection of cotranslational protein N-myristoylation by metabolic labeling in the insect cell-free protein synthesis system. The mRNA coding for MG3A6S–TNF was translated with insect cell lysate (Insect L) or with rabbit reticulocyte lysate (RRL) in the presence of [^{3}H]leucine or [^{3}H]myristic acid. The labeled translation products were analyzed directly by SDS–PAGE and fluorography. (A) CBB staining; (B) Fluorography.

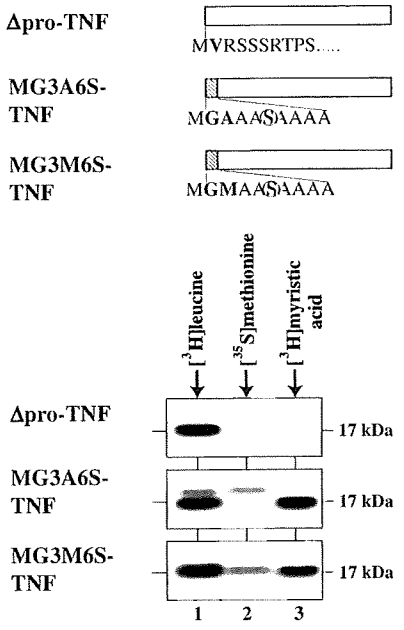

Fig. 4. Detection of the removal of the initiating Met residue during cotranslational protein N-myristoylation reaction. The mRNAs coding for Δpro–, MG3A6S–, and MG3M6S–TNF were translated with insect cell lysate in the presence of [^{3}H]leucine, [^{35}S]methionine, or [^{3}H]myristic acid. The labeled translation products were directly analyzed by SDS–PAGE and fluorography.

replaced with the 19 other amino acids. The results for the 20 amino acids are arranged according to their radius of gyration. As shown in the upper panels of Figs. 5A and B, most of the translation products, except for those of mutants having a Gly or Tyr residue at position 3, were efficiently expressed in both of the in vitro translation systems, as determined by [^{3}H]leucine incorporation. In these two sets of translation products, the mobility shift of the protein band was observed equally with two mutants having an Asp or Glu residue at position 3. As for protein N-myristoylation, exactly the same amino acid requirements were observed at position 3 in the two different cell-free protein synthesis systems: 12 amino acids (Gly, Ala, Ser, Cys, Thr, Val, Asn, Leu, Ile, Gln, His, Met) were permitted at position 3 to direct protein N-myristoylation, as shown in the lower panels of Figs. 5A and B. These results clearly indicated that the substrate specificity of NMT in the insect cell lysate is quite similar to that in rabbit reticulocyte lysate.

Posttranslational protein N-myristoylation could be detected in the insect cell-free protein synthesis system

In addition to cotranslational protein N-myristoylation, it was shown that N-myristoylation can also occur post-translationally as in the case of some caspase substrates. However, the simple and sensitive method to detect this modification has not been established. Recently, we have shown that posttranslational N-myristoylation of caspase substrates could be detected by [^{3}H]myristic acid labeling of cultured mammalian cells in which the cDNA coding for the C-terminal caspase cleavage product having N-myristoylation consensus motif at its N terminus was transfected [10,11]. In this case, however, the cytotoxic N-myristoylated protein such as tBid could not be subjected to this assay system because of the lack of protein expression in the transfected cells. In fact, when the cDNAs coding for Bid–, tBid–, gelsolin– and tGelsolin–FLAG were transfected into COS-1 cells, protein expression was observed with Bid–, gelsolin– and tGelsolin–FLAG but not with tBid–FLAG, as shown in Fig. 6A, lanes 5–8. N-myristoylation of tGelsolin–FLAG expressed in COS-1 cells was clearly detected (Fig. 6A, lane 12).

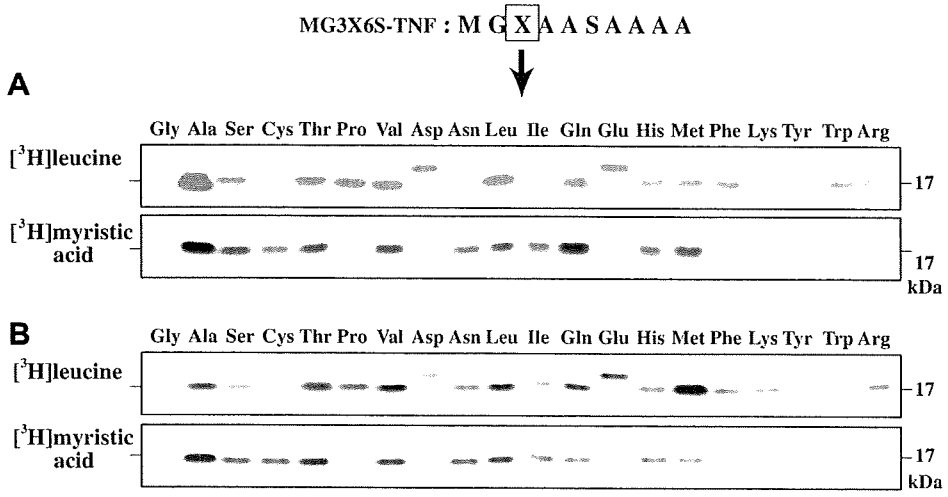

Fig. 5. Similarity of the substrate specificity of NMT in insect cell lysate to that in rabbit reticulocyte lysate. The mRNAs coding for MG3X6S–TNF were translated with rabbit reticulocyte lysate (A) or with insect cell lysate (B) in the presence of [^{3}H]leucine or [^{3}H]myristic acid. The labeled translation products were analyzed directly (B) or following immunoprecipitation with anti-TNF antibody (A) by SDS–PAGE and fluorography. Results for the 20 amino acids are arranged according to their radius of gyration.

Since the cell-free protein synthesis system can synthesize cytotoxic proteins, it might be used to detect the protein modification that occurs on cytotoxic proteins. To determine whether the posttranslational protein N-myristoylation could be detected in the insect cell-free protein synthesis system, incorporation of [^{3}H]myristic acid into tBid and tGelsolin was measured. In this case, the initiating Met was inserted at the N terminus of the newly exposed N terminus of the caspase-cleavage products of Bid and gelsolin and the susceptibility of the N-terminal sequence to protein N-myristoylation was evaluated by the metabolic labeling in the insect cell lysate. As shown in Fig. 6B, lanes 5, 7, 9, and 11, tBid–FLAG was efficiently expressed in the insect cell-free protein synthesis system and its N-myristoylation was successfully detected, as in the case of tGelsolin–FLAG. As expected, when Gly 2 of tBid–FLAG was replaced with Ala (tBidG2A–FLAG), [^{3}H]myristic acid incorporation was completely inhibited (Fig. 6B, lanes 6 and 10). These results clearly indicate that the posttranslational protein N-myristoylation that occurs on the cytotoxic protein such as tBid could be detected in the insect cell-free protein synthesis system.

Discussion

Since all protein synthesis begins at the N terminus, this region provides an initial and important site of cotranslational protein processing. Removal of the initiator methionine and modification of the α-amino group are examples of commonly observed N-terminal modifications. Among them, three cotranslational protein modifications, excision of the initiator methionine, N-acetylation, and N-myristoylation, potentially affect many eukaryotic cytoplasmic proteins and variously correlate with their stability, physiological function, and/or degradation. In particular, protein N-myristoylation plays critical roles in many cellular signal transduction pathways. Proteins destined to become N-myristoylated begin with the sequence Met-Gly. The initiating Met is removed cotranslationally by methionine aminopeptidase and then myristic acid is linked to Gly-2 via an amide bond by NMT. However, not all proteins with an N-terminal glycine are N-myristoylated and the ability to be recognized by NMT depends on the downstream amino acid sequence. In addition, proteins with an N-terminal glycine may also be subject to another cotranslational protein modification, N-acetylation. For postgenomic studies, reliable tools for the prediction of co- and posttranslational modifications would be valuable for functional assignments of functionally unknown proteins. In fact, many protein sequences in the publicly available database are automatically predicted as possibly N-myristoylated proteins by a prediction program [7,8]. Since the predicted results are not necessarily correct, experimental methods to verify the modification are indispensable to confirm the prediction.

In general, N-myristoylation that occurs on a certain gene product is detected by in vivo metabolic labeling of cells transfected with the cDNA coding for the gene product. However, some proteins such as cytotoxic proteins cannot be expressed in cells, and therefore these proteins could not be tested for susceptibility to protein N-myristoylation. In fact, as shown in Fig. 6, in the present study, tBid, a well-known proappoptotic protein that has been shown to be posttranslationally N-myristoyated, could not be expressed in COS-1 cells. In addition to the in vivo metabolic labeling in the transfected cells, in vitro metabolic labeling in cell-free protein synthesis systems has previously been shown to be useful for detecting co- and posttranslational protein N-myristoylation. Cell-free protein synthesis systems are powerful tools for the analysis of co- and posttranslational modification of protein

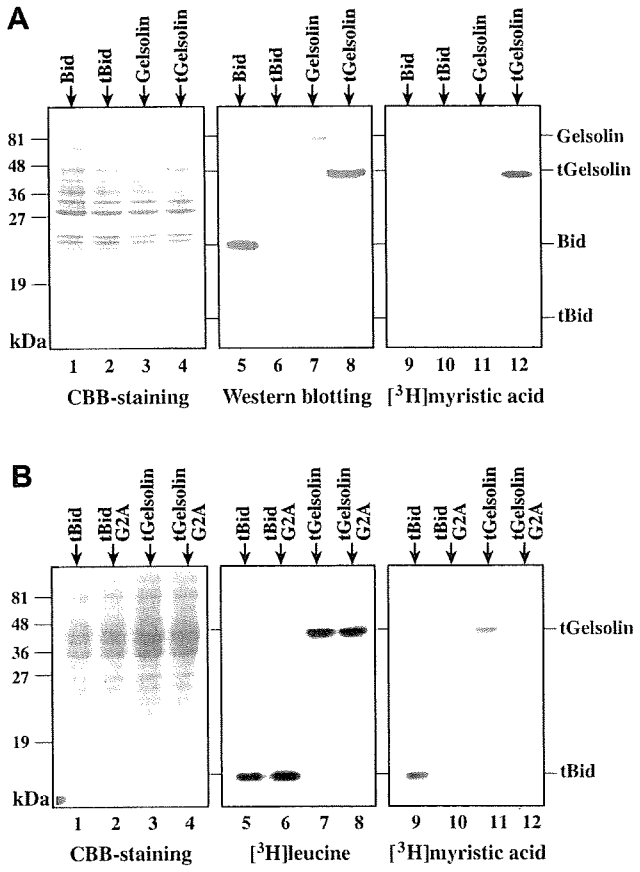

Fig. 6. Detection of posttranslational protein N-myristoylation by metabolic labeling in the insect cell-free protein synthesis system. (A) The cDNAs coding for Bid–, tBid–, gelsolin–, and tGelsolin–FLAG were transfected into COS-1 cells, and the cells were labeled with [^{3}H]myristic acid. Protein expression in the total cell lysate was detected by Western blotting using anti-FLAG antibody. Following immunoprecipitation with anti-FLAG antibody, protein N-myristoylation was detected by SDS–PAGE and fluorography. (B) The mRNAs coding for tBid–, tBidG2A–, tGelsolin–, and tGelsolinG2A–FLAG were translated with insect cell lysate in the presence of [^{3}H]leucine or [^{3}H]myristic acid. The labeled translation products were directly analyzed by SDS–PAGE and fluorography.

because they make it possible to translate exogenous mRNAs with high speed [21] and to synthesize any desired proteins, including cytotoxic proteins. A rabbit reticulocyte lysate cell-free system [22] has been widely utilized for metabolic labeling because it contains all the components involved in N-terminal protein modifications, such as methionine aminopeptidases, NMTs, and NATs [23–25]. However, this cell-free protein synthesis system has several disadvantages for the detection of co- and posttranslational modifications. First, it takes a long time (several days to weeks) to obtain the final results of metabolic labeling, probably because the level of protein expression is low. Second, to obtain clear results of metabolic labeling, purification steps such as immunoprecipitation or purification using an affinity column are required. Since the specific antibodies against most of the gene products derived from the cDNA in the cDNA library are not available, the possibility of being able to use a specific antibody for the metabolic labeling experiment is quite low. For the affinity purification, the addition of an epitope-tag such as His-tag, FLAG-tag, or myc-tag to the 5′ or 3′ end of the cDNA is required. However, most of the commercially available cDNA clones do not contain an epitope-tag. Thus, it would be a great advantage if the protein modification could be detected in the in vitro translation system within a short period of time without any purification step.

Recently, a cell-free protein synthesis system (Transdirect™ *insect cell*) derived from *Spodoptera frugiperda* 21 (Sf21) insect cells has been developed. In the present study, in vitro metabolic labeling of N-myristoylated protein in this newly developed cell-free protein synthesis system was performed and the usefulness of this system for detecting protein N-myristoylation was examined.

The results revealed that cotranslational protein N-myristoylation of a model N-myristoylated protein could easily be detected by metabolic labeling in an insect cell-free protein synthesis system within a short period of time (18 to 24 h) without any purification step, as shown in Figs. 2 and 3. In addition, when the susceptibility of tBid, a posttranslationally N-myristoylated cytotoxic protein that could not be expressed in the transfected cells, to protein N-myristoylation was tested, N-myristoylation was successfully detected by this assay system, as shown in Fig. 6. Thus, metabolic labeling in an insect cell-free protein synthesis system is found to be an easy and effective strategy to detect co- and posttranslational protein N-myristoylation.

Protein N-myristoylation is catalyzed by *N*-myristoyltransferase. The precise substrate specificity of this enzyme has been characterized using purified enzyme and synthetic peptides derived from the N-terminal sequences of known N-myristoylated proteins [1,26,27]. From these studies, it was revealed that, in addition to Gly at position 2, the amino acids at positions 3, 6, and 7 play important roles, in substrate recognition by NMT [1,28,29]. In fact, we have previously shown that the amino acid at position 3 strongly affects the protein N-myristoylation by metabolic labeling of a series of model N-myristoylated proteins in the in vitro translation system derived from rabbit reticulocyte lysate [30]. In the present study, the same experiments were performed using two cell-free protein synthesis systems (insect cell lysate and rabbit reticulocyte lysate) and the N-myristoylation level was compared. The results showed that exactly the same amino acid requirements at position 3 were observed in the two different cell-free protein synthesis systems as shown in Fig. 5. These results clearly indicated that the substrate specificity of NMT in the insect cell lysate is quite similar to that in rabbit reticulacyte lysate.

Thus, metabolic labeling in an insect cell-free protein synthesis system could be an easy and effective strategy to detect co- and posttranslational protein N-myristoylation in a wide variety of eukaryotic cellular proteins.

Acknowledgments

This work was supported in part by Grants-in-Aid for Scientific Research (Nos. 15580080 and 17580080) from the Ministry of Education, Science and Culture of Japan.

References

[1] D.A. Towler, J.I. Gordon, S.P. Adams, L. Glase, The biology and enzymology of eukaryotic protein acylation, Annu. Rev. Biochem. 57 (1988) 69–99.
[2] A.M. Spiegel, P.S. Backlund, J.E. Butrynski, T.L.Z. Jones, W.F. Simonds, The G protein connection: molecular basis of membrane association, Trends Biochem. Sci. 16 (1991) 338–341.
[3] J.A. Boutin, Myristoylation, Cell Signal. 9 (1997) 15–35.
[4] M.D. Resh, Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins, Biochim. Biophys. Acta. 1451 (1999) 1–16.
[5] T.A. Farazi, G. Waksman, J.I. Gordon, The biology and enzymology of protein N-myristoylation, J. Biol. Chem. 276 (2001) 39501–39504.
[6] F. Dyda, D.C. Klein, A.B. Hickman, GCN5-related N-acetyltransferases: a structural overview, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 81–103.
[7] B. Eisenhaber, F. Eisenhaber, S. Maurer-Stroh, G. Neuberger, Prediction of sequence signals for lipid post-translational modifications: Insight from case studies, Proteomics 4 (2004) 1614–1625.
[8] S. Maurer-Stroh, M. Gouda, M. Novatchkova, A. Schleiffer, G. Schneider, F.L. Sirota, M. Wildpaner, N. Hayashi, F. Eisenhaber, MYRbase: analysis of genome-wide glycine myristoylation enlarges the functional spectrum of eukaryotic myristoylated proteins, Genome Biol. 5 (2004) R21.
[9] J. Zha, S. Weiler, K.-J. Oh, M.C. Wei, S.J. Korsmeyer, Posttranslational N-myristoylation of BID as a molecular switch for targeting mitochondria and apoptosis, Science 251 (2000) 1761–1765.
[10] T. Utsumi, N. Sakurai, K. Nakano, R. Ishisaka, C-terminal 15 kDa fragment of cytoskeletal actin is posttranslationally N-myristoylated upon caspase-mediated cleavage and targeted to mitochondria, FEBS Lett. 539 (2003) 37–44.
[11] N. Sakurai, T. Utsumi, Posttranslational N-myristoylation is required for the anti-apoptotic activity of human tGelsolin, the C-terminal caspase cleavage product of human gelsolin, J. Biol. Chem. 281 (2006) 14288–14295.
[12] G.L. Vilas, M.M. Corvi, G.J. Plummer, A.M. Seime, G.R. Lambkin, L.G. Berthiaume, Posttranslational myristoylation of caspase-activated p21-activated protein kinase 2 (PAK2) potentiates late apoptotic events, Proc. Natl. Acad. Sci. USA 103 (2006) 6542–6547.
[13] T. Ezure, T. Suzuki, S. Higashide, E. Shintani, K. Endo, S. Kobayashi, M. Shikata, M. Ito, K. Tanimizu, O. Nishimura, Cell-free protein synthesis system prepared from insect cells by freeze-thawing, Biotechnol. Prog. 22 (2006) 1570–1577.
[14] T. Utsumi, K. Akimaru, Z. Kawabata, A. Levitan, T. Tokunaga, P. Tang, A. Ide, M.C. Hung, J. Klostergaard, Human pro-tumor necrosis factor: molecular determinants of membrane translocation, sorting, and maturation, Mol. Cell. Biol. 15 (1995) 6389–6405.
[15] T. Utsumi, J. Kuranami, E. Tou, A. Ide, K. Akimaru, M.C. Hung, J. Klostergaard, In vitro synthesis of an N-myristoylated fusion protein that binds to the liposomal surface, Arch. Biochem. Biophys. 326 (1996) 179–184.
[16] T. Suzuki, M. Ito, T. Ezure, S. Kobayashi, M. Shikata, K. Tanimizu, O. Nishimura, Performance of expression vector, pTD1, in insect cell-free translation system, J. Biosci Bioeng. 102 (2006) 69–71.
[17] T. Suzuki, M. Ito, T. Ezure, M. Shikata, E. Ando, T. Utsumi, S. Tsunasawa, O. Nishimura, N-terminal protein modifications in an insect cell-free protein synthesis system and their identification by mass spectrometry, Proteomics 6 (2006) 4486–4495.
[18] F. Sanger, S. Nickelen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.
[19] T. Utsumi, A. Levitan, M.C. Hung, J. Klostergaard, Effects of truncation of human pro-tumor necrosis factor transmembrane domain on cellular targeting, J. Biol. Chem. 268 (1993) 9511–9516.
[20] S. Tsunasawa, J.W. Stewart, F. Sherman, Amino-terminal processing of mutant forms of yeast iso-1-cytochrome *c*. The specificities of methionine aminopeptidase and acetyltransferase, J. Biol. Chem. 260 (1985) 5382–5391.
[21] T. Sawasaki, T. Ogasawara, R. Morishita, Y. Endo, A cell-free protein synthesis system for high-throughput proteomics, Proc. Natl. Acad. Sci. USA 99 (2002) 14652–14657.
[22] R.J. Jackson, T. Hunt, Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA, Methods Enzymol. 96 (1983) 50–74.
[23] R.D. Palmiter, J. Gagnon, K.A. Walsh, Ovalbumin: a secreted protein without a transient hydrophobic leader sequence, Proc. Natl. Acad. Sci. USA 75 (1978) 94–98.
[24] A. Mayer, N.R. Siegel, A.L. Schwarts, A. Ciechanover, Degradation of proteins with acetylated amino termini by the ubiquitin system, Science 244 (1989) 1480–1483.
[25] I. Deichaite, L.P. Casson, H.P. Ling, M.D. Resh, In vitro synthesis of $pp60^{v\text{-}src}$: myristoylation in a cell-free system, Mol. Cell. Biol. 8 (1988) 4295–4301.
[26] D.A. Towler, S.P. Adam, S.R. Eubanks, D.S. Towery, E. Jackson-Machelski, L. Glaser, J.I. Gordon, Myristoyl CoA: protein N-myristoyltransferase activities from rat liver and yeast possess overlapping yet distinct peptide substrate specificities, J. Biol. Chem. 263 (1988) 1784–1790.
[27] W.J. Rocque, C.A. McWherter, D.C. Wood, J.I. Gordon, A comparative analysis of the kinetic mechanism and peptide substrate specificity of human and *Saccharomyces cerevisiae* myristoyl CoA: protein N-myristoyltransferase, J. Biol. Chem. 268 (1993) 9964–9971.
[28] D.R. Johnson, R.S. Bhatnagar, L.J. Knoll, J.I. Gordon, Genetic and biochemical studies of protein N-myristoylation, Annu. Rev. Biochem 63 (1994) 869–914.
[29] C.A. McWherter, W.J. Rocque, M.E. Zupec, S.K. Freeman, D.L. Brown, B. Devadas, D.P. Getman, J.A. Sikorski, J.I. Gordon, Scanning alanine mutagenesis and de-peptidization of a Candida albicans myristoyl-CoA: protein N-myristoyltransferase octapeptide substrate reveals three elements critical for molecular recognition, J. Biol. Chem. 272 (1997) 11874–11880.
[30] T. Utsumi, K. Nakano, T. Funakoshi, Y. Kayano, S. Nakao, N. Sakurai, H. Iwata, R. Ishisaka, Vertical-scanning mutagenesis of amino acids in a model N-myristoylation motif reveals the major amino-terminal sequence requirements for protein N-myristoylation, Eur. J. Biochem. 271 (2004) 863–874.

DOI 10.1002/pmic.200700237

RESEARCH ARTICLE

Protein prenylation in an insect cell-free protein synthesis system and identification of products by mass spectrometry

Takashi Suzuki[1], Masaaki Ito[1], Toru Ezure[1], Masamitsu Shikata[1], Eiji Ando[1], Toshihiko Utsumi[2], Susumu Tsunasawa[1] and Osamu Nishimura[1,3]

[1] Life Science Laboratory, Analytical and Measuring Instruments Division, Shimadzu Corporation, Kyoto, Japan
[2] Department of Biological Chemistry, Faculty of Agriculture, Yamaguchi University, Yamaguchi, Japan
[3] Institute for Protein Research, Osaka University, Osaka, Japan

Received: March 8, 2007
Accepted: March 13, 2007

To evaluate the ability of an insect cell-free protein synthesis system to carry out proper protein prenylation, several CAIX (X indicates any C-terminal amino acid) sequences were introduced into the C-terminus of truncated human gelsolin (tGelsolin). Tryptic digests of these mutant proteins were analyzed by MALDI-TOF MS and MALDI-quadrupole-IT-TOF MS. The results indicated that the insect cell-free protein synthesis system possesses both farnesyltransferase (FTase) and geranylgeranyltransferase (GGTase) I, as is the case of the rabbit reticulocyte lysate system. The C-terminal amino acid sequence requirements for protein prenylation in this system showed high similarity to those observed in rat prenyltransferases. In the case of rhoC, which is a natural geranylgeranylated protein, it was found that it could serve as a substrate for both prenyltransferases in the presence of either farnesyl or geranylgeranyl pyrophosphate, whereas geranylgeranylation was only observed when both prenyl pyrophosphates were added to the *in vitro* translation reaction mixture. Thus, a combination of the cell-free protein synthesis system with MS is an effective strategy to analyze protein prenylation.

Keywords:
Insect cell-free protein synthesis system / MALDI-quadrupole-IT-TOF MS / MALDI-TOF MS / protein prenylation

1 Introduction

The functional analysis of proteins in postgenomic studies has been attracting considerable attention, and interest in analyzing PTMs of proteins is increasing. Cell-free protein synthesis systems are assumed to be powerful tools for such studies, because they are capable of translating exogenous mRNAs with high speed [1] and they have the potential to synthesize any desired proteins, including both native proteins and those that are toxic to cells [2]. We developed a cell-free protein synthesis system from *Spodoptera frugiperda* 21 (Sf21) insect cells [3], which are widely used as the host for baculovirus expression systems, and demonstrated by MALDI-TOF MS and MALDI-quadrupole-IT (QIT)-TOF MS analysis that the insect cell-free protein synthesis system could generate N-terminal protein modifications, such as cleavage of the initiator Met, *N*-acetylation, and *N*-myristoylation [4].

Prenylation is one of the important lipid modifications of proteins, and it plays crucial roles in regulating reversible protein–membrane and protein–protein interactions [5, 6]. Farnesyltransferase (FTase) and geranylgeranyltransferase

Correspondence: Takashi Suzuki, Life Science Laboratory, Analytical and Measuring Instruments Division, Shimadzu Corporation, 1 Nishinokyo-Kuwabaracho, Nakagyo-ku, Kyoto 604-8511, Japan
E-mail: t-suzuki@shimadzu.co.jp
Fax: +81-75-823-1364

Abbreviations: FPP, farnesyl pyrophosphate; **FTase,** farnesyltransferase; **GGPP,** geranylgeranyl pyrophosphate; **GGTase,** geranylgeranyltransferase; **Lys-C,** lysyl endopeptidase; **QIT,** quadrupole IT; **Sf21,** *Spodoptera frugiperda* 21; **tGelsolin,** truncated gelsolin

(GGTase) I recognize a C-terminal CAAX motif (C is cysteine, A is usually an aliphatic amino acid, and X is one of a variety of amino acids) and covalently attach a farnesyl group from farnesyl pyrophosphate (FPP) or a geranylgeranyl group from geranylgeranyl pyrophosphate (GGPP) to the free cysteine residue. The C-terminal amino acid residue is the major determinant for the selection of which of the two types of prenyl group is to be attached to the cysteine residue [7, 8]. In recent studies, however, it was revealed that these enzymes have overlapping preferences for their C-terminal amino acid [9–11]. Therefore, it is quite difficult to predict whether uncharacterized CAAX-terminating peptides, with their sequences deduced by sequencing of cDNAs, will be modified with a farnesyl or a geranylgeranyl group.

Metabolic labeling is an effective strategy for the analysis of protein prenylation using a cell-free protein synthesis system [7, 12]. A rabbit reticulocyte lysate cell-free system [13] has been widely utilized for metabolic labeling, because it contains all of the components involved in protein prenylation, such as FTase and GGTase I, *etc.* [14]. Metabolic labeling has a great advantage for determining whether modifications have occurred, in that it is quite simple. However, it cannot be used to identify the exact location of the modification unless mutagenesis studies are undertaken. In this study, we established an effective strategy to analyze protein prenylation by combining an insect cell-free protein synthesis system with MALDI-TOF MS and MALDI-QIT-TOF MS.

2 Materials and methods

2.1 Materials

Transdirect *insect cell*, which is based on the Sf21 extract, is a commercial product of Shimadzu (Kyoto, Japan). Restriction endonucleases and DNA modifying enzymes were purchased from Toyobo (Osaka, Japan) and New England Biolabs (Ipswich, MA, USA). CHCA, 2,5-dihydroxybenzoic acid (DHB), ANTI-FLAG® M2-Agarose from mouse, and FLAG® peptide were purchased from Sigma (St. Louis, MO, USA). GGPP triammonium salt and FPP triammonium salt solution were purchased from MP Biomedicals (OH, USA) and Wako Pure Chemical Industries (Osaka, Japan), respectively. Human cDNA clone rhoC (NM_175744) was purchased from Toyobo.

2.2 Construction of vectors for the analysis of protein prenylation

N-terminal FLAG-tag was introduced into the truncated human gelsolin (tGelsolin) gene by PCR using tGel-FLAG-N (5′-ATGGACTACAAGGATGACGATGACAAGGGCCTGGGCTTGTCCTAC-3′) as the sense primer, tGel-C (5′-GGGGATCCTTAGGCAGCCAGCTCAGCCAT-3′) as the antisense primer, and pcDNA3-tGelsolin-FLAG [15] as the template. The amplified DNA fragment was then treated with T4 polynucleotide kinase. After digestion with *Bam*HI, the amplified fragment was subcloned into the *Eco*RV-*Bam*HI sites of a pTD1 vector [16], and the resulting vector was designated as pTD1-tGelsolin-FLAG. The nucleotide sequence coding for RSHEHHFFCAIL, which contains a geranylgeranylation motif, was inserted upstream of the stop codon of tGelsolin, and this construct was designated as pTD1-tGelsolin-CAIL. PCR was performed using the tGel-CAIL-F primer (5′-TTCTTCTGTGCTATCCTGTAAGGATCCTCTAGAGTCGG-3′), the tGel-CAIL-R primer (5′-ATGGTGCTCGTGGCTCCGGGCAGCCAGCTCAGCCAT-3′), and pTD1-tGelsolin-FLAG as the template. The amplified DNA fragment was then treated with T4 polynucleotide kinase. After the treatment, the DNA fragment was self-ligated and transformed into *Escherichia coli* DH5α. In order to investigate the sequence requirements for protein prenylation, the C-terminal Leu residue of tGelsolin-CAIL was replaced with Ala, Cys, Phe, Met, Gln, or Ser, since these six amino acids have been shown to be preferable as C-terminal amino acids for prenyltransferase substrates [11]. These constructs were generated following the same procedure as pTD1-tGelsolin-CAIL and designated as pTD1-tGelsolin-CAIX (X is the C-terminal amino acid residue). PCR was carried out using the ARSHEH-R primer (5′-GTGCTCGTGGCTCCGGGC-3′), the CAIX primer (5′-CATTTCTTCTGTGCTATCNNSTAAGGAT-3′, where the underlined NNS sequence indicates the codon for one of the six amino acid residues), and pTD1-tGelsolin-CAIL as the template.

The ORF of rhoC [17, 18], which is a well-characterized natural prenylated protein, was also cloned into the pTD1 vector as follows: PCR was performed using the RhoC-FLAG primer (5′-ATGGACTACAAGGATGACGATGACAAGGCTGCAATCCGAAAGAAG-3′), the RhoC-kpn primer (5′-GGGGTACCTCAGAGAATGGGACAGCCC-3′), and rhoC cDNA as the template. The amplified DNA fragment was then treated with T4 polynucleotide kinase. After digestion with *Kpn*I, the amplified fragment was subcloned into the *Eco*RV-*Kpn*I sites of a pTD1 vector, and the resulting vector was designated as pTD1-rhoC-FLAG. The DNA sequences of these recombinant constructs were confirmed by the dideoxynucleotide chain termination method.

2.3 Preparation of mRNAs

The vectors were linearized by *Hin*dIII, then purified by phenol–chloroform extraction and ethanol precipitation. The mRNAs were synthesized and purified as previously described [4].

2.4 Purification of the proteins

Cell-free protein synthesis was carried out at a 1 mL scale using the Transdirect *insect cell* with or without the addition of GGPP or FPP at a final concentration of 50 μM. Reactions were performed by adding prepared mRNA followed by incubation at 25°C for 4 h. After the reaction, 100 μL of 20%

w/v Triton X-100 was added to the reaction mixture, which was then centrifuged at 15 000 rpm for 15 min. The supernatant was desalted using a PD-10 column (GE Healthcare, Piscataway, USA) equilibrated with 50 mM Tris-HCl, pH 8.0, containing 300 mM NaCl and 2% w/v Triton X-100 (buffer A). The void volume was collected and applied to an ANTI-FLAG M2-Agarose column (0.5 mL) equilibrated with buffer A. The column was washed with the same buffer (1.0 mL) and then further washed with Tris-HCl, pH 8.0, containing 300 mM NaCl (4.0 mL, buffer B). The protein was eluted with buffer B containing 100 μg/mL FLAG peptide (2.5 mL). The eluate was concentrated to about 30 μL by ultrafiltration (molecular weight cutoff = 10 kDa). In the case of rhoC protein, purification was performed using the same buffer systems containing 100 μM GTP. The purities and yields of the expressed proteins were estimated by SDS-PAGE using purified tGelsolin [4] as a standard. The concentrates were stored at −20°C until use.

2.5 MS

The affinity-purified proteins (about 0.5–1 μg) were electrophoresed on an SDS-polyacrylamide gel and then stained with CBB R-250. The protein band was reduced and *S*-alkylated with iodoacetamide and then digested overnight with trypsin (Promega, Madison, WI, USA) or lysyl endopeptidase (Lys-C) (Wako Pure Chemical Industries). The tryptic digests were extracted twice using 60% v/v ACN containing 0.1% v/v TFA and 0.1% w/v *n*-octyl-β-D-glucopyranoside (Wako Pure Chemical Industries). In the case of Lys-C digests, peptides were extracted twice using 50% ACN containing 0.1% TFA. The extracted solution was dried, then dissolved into 10 μL of 50% v/v ACN containing 0.1% v/v TFA. The sample was mixed with CHCA or DHB solution (10 mg/mL in 50% v/v ACN containing 0.1% v/v TFA). The mass spectra and MS/MS spectra were acquired in reflectron positive ion mode with an AXIMA-CFR-plus MALDI-TOF MS instrument and an AXIMA-QIT MALDI-QIT-TOF hybrid mass spectrometer (Shimadzu/Kratos, Manchester, UK), respectively.

3 Results

3.1 Preparation of prenylated proteins using an insect cell-free protein synthesis system

To evaluate the performance of this system in carrying out proper protein prenylation in an insect cell-free protein synthesis system, Transdirect *insect cell*, CAAX sequences such as – CAIL (geranylgeranylation) and – CAIS (farnesylation) were introduced into the C-terminus of tGelsolin [15]. N-terminal FLAG-tagged tGelsolin mutants were expressed using the insect cell-free protein synthesis system with or without the addition of GGPP or FPP. Cell-free synthesized proteins were purified by affinity chromatography. For the translation reaction in the absence of prenyl pyrophosphate, each purified tGelsolin mutant protein was detected as a main band with an apparent molecular mass of 46 kDa (about 70–80% purity) (Fig. 1; lanes 1 and 3). On the other hand, for the reactions in the presence of FPP or GGPP, a protein having 50 kDa and some extra bands were also detected in addition to the 46 kDa protein band (Fig. 1: lanes 2 and 4). The 50 kDa protein was identified as β-tubulin by PMF (data not shown), but we cannot explain the reason why β-tubulin was coeluted in the purification step.

These protein bands were reduced and *S*-alkylated and then digested with trypsin. The tryptic digests were analyzed by MALDI-TOF MS, and the MS spectra produced from these samples were almost identical (Fig. 2A). In the case of the tGelsolin-CAIL mutant, which was assumed to be geranylgeranylated, when GGPP was added to the *in vitro* translation reaction mixture, a peak corresponding to the geranylgeranylated C-terminal tryptic peptide (theoretical monoisotopic mass value = 1612.87) was clearly observed at the *m/z* value of 1612.81, whereas the carbamidomethylated C-terminal peptide (theoretical monoisotopic mass value = 1397.64) was not detected at all (Fig. 2B-b). When GGPP was not added, only a peak equivalent to the carbamidomethylated C-terminal peptide was detected at the *m/z* value of 1397.56 (Fig. 2B-a). The peptide peak at *m/z* 1612.87 was subjected to MS/MS analysis and was identified as the C-terminal tryptic fragment in which the cysteine residue in

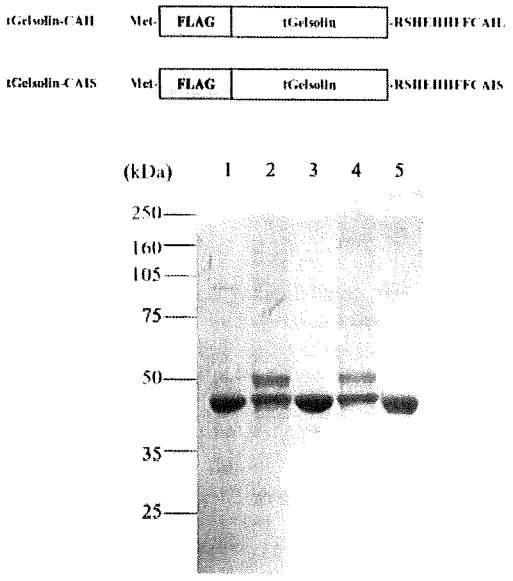

Figure 1. SDS-PAGE of the affinity-purified tGelsolin mutant proteins. The mRNAs transcribed from the pTD1-tGelsolin-CAIL (lanes 1 and 2) and pTD1-tGelsolin-CAIS (lanes 3 and 4) were individually translated using the insect cell-free protein synthesis system either with (lane 2) or without (lane 1) the addition of GGPP, and either with (lane 4) or without (lane 3) the addition of FPP, respectively. Two microliters (lanes 1 and 3) and 6 μL (lanes 2 and 4) of the concentrates were electrophoresed on a 10% SDS-PAGE. Lane 5: purified tGelsolin (1 μg) [4].

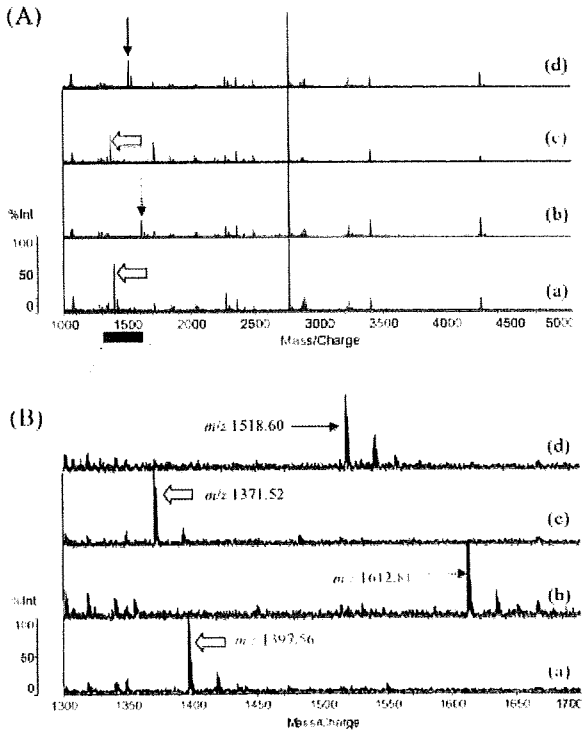

Figure 2. MALDI-mass spectra showing tryptic digests of the tGelsolin-CAIL and tGelsolin-CAIS mutant proteins. Tryptic digests of the tGelsolin-CAIL and tGelsolin-CAIS mutant proteins translated either without (a) or with (b) the addition of GGPP, and either without (c) or with (d) the addition of FPP were analyzed by MALDI-TOF MS. The acquired profiles focused on a mass/charge range from 1000 to 5000 (A) and from 1300 to 1700 (B). Open boxed, solid, and dotted arrows indicate peaks corresponding to the C-terminal tryptic peptides containing a carbamidomethylated, farnesylated, or geranylgeranylated cysteine in the CAAX motif, respectively.

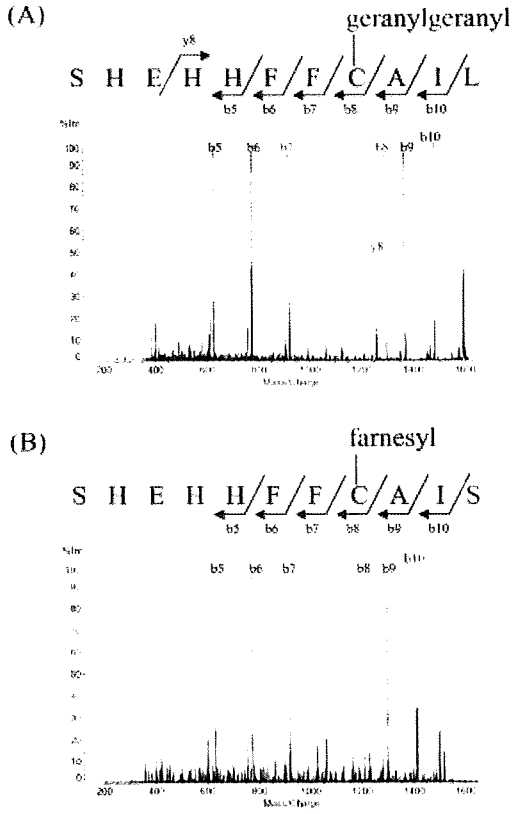

Figure 3. MALDI-MS/MS spectra of the probably prenylated C-terminal peptides from the tGelsolin-CAIL and tGelsolin-CAIS mutant proteins. MS/MS analyses were performed for the peaks at *m/z* 1612.81 observed for the tGelsolin-CAIL (A) and at *m/z* 1518.60 for the tGelsolin-CAIS (B). The observed fragment ions are indicated in the sequences shown.

the CAAX motif was geranylgeranylated (Fig. 3A). In the case of the tGelsolin-CAIS mutant, which was predicted to be farnesylated, a probable peak corresponding to the farnesylated C-terminal tryptic peptide (theoretical monoisotopic mass value = 1518.76) was detected at the *m/z* value of 1518.60 only when FPP was added to the *in vitro* translation reaction mixture (Fig. 2B-d). On the other hand, in the absence of FPP, a peak corresponding to the carbamidomethylated C-terminal tryptic peptide (theoretical monoisotopic mass value = 1371.59) was observed at the *m/z* value of 1371.52, but none was seen at the position for the farnesylated peptide (Fig. 2B-c). The peptide peak at *m/z* 1518.60 was identified as the farnesylated C-terminal peptide fragment by MS/MS analysis (Fig. 3B). These results clearly indicated that the Sf21 extract contains FTase and GGTase I, as was found for the rabbit reticulocyte lysate [14]. Furthermore, the results suggest that protein prenylation could be controlled by the addition of prenyl pyrophosphate to the reaction mixture of the Transdirect *insect cell*.

3.2 Effect of the C-terminal amino acid residue on the protein prenylation reaction

To evaluate the effect of the C-terminal amino acid residue on the protein prenylation reaction, seven tGelsolin-CAIX mutants (X = A, C, F, L, M, Q, and S, which are preferred C-terminal amino acid residues for FTase or/and GGTase I) were constructed. *In vitro* translation reactions were carried out in the presence of FPP or GGPP using mRNAs transcribed from pTD1-tGelsolin-CAIX as the template. All the mutant proteins were successfully expressed (data not shown). The susceptibility to protein prenylation was analyzed by the same method described above. Tables 1 and 2 summarize the calculated and observed C-terminal tryptic digests of tGelsolin-CAIX proteins generated in the presence of FPP and GGPP in the translation reaction mixture, respectively.

When FPP was added to the *in vitro* translation reaction mixture, peaks corresponding to the farnesylated C-terminal tryptic peptide were observed in the MS spectra for all

Table 1. Calculated and observed monoisotopic mass values for the tryptic C-terminal peptide of tGelsolin-CAIX proteins obtained after FPP was added to the *in vitro* translation reaction mixture

X residue	Calculated mass value			Observed mass value		
	CAM[a)]	FAR[b)]	GERA[c)]	CAM	FAR	GERA
A	1355.59	1502.76	1570.82	ND[d)]	1502.82	ND
C	1444.59	1591.75	1659.82	ND	1591.78	ND
F	1431.63	1578.79	1646.86	1431.64	1578.71	ND
L	1397.64	1544.81	1612.87	1397.65	1544.85	ND
M	1415.60	1562.76	1630.83	ND	1562.66	ND
Q	1412.62	1559.78	1627.85	ND	1559.73	ND
S	1371.59	1518.76	1586.82	ND	1518.60	ND

a) Carbamidomethylated.
b) Farnesylated.
c) Geranylgeranylated.
d) Not detected.

Table 2. Calculated and observed monoisotopic mass values for the tryptic C-terminal peptide of tGelsolin-CAIX proteins after GGPP was added to the *in vitro* translation reaction mixture

X residue	Calculated mass value			Observed mass value		
	CAM[a)]	FAR[b)]	GERA[c)]	CAM	FAR	GERA
A	1355.59	1502.76	1570.82	1355.53	ND[d)]	ND
C	1444.59	1591.75	1659.82	1444.51	ND	ND
F	1431.63	1578.79	1646.86	1431.62	ND	1646.81
L	1397.64	1544.81	1612.87	ND	ND	1612.81
M	1415.60	1562.76	1630.83	ND	ND	1630.80
Q	1412.62	1559.78	1627.85	1412.68	ND	ND
S	1371.59	1518.76	1586.82	1371.52	ND	ND

a) Carbamidomethylated.
b) Farnesylated.
c) Geranylgeranylated.
d) Not detected.

mutant proteins. Carbamidomethylated C-terminal tryptic peptides were not detected at all in the tGelsolin-CAIA, -CAIC, -CAIM, -CAIQ, and -CAIS mutant proteins (Table 1 and Fig. 4A). In the case of the tGelsolin-CAIF mutant protein, a farnesylated C-terminal peptide was predominantly detected in comparison with the carbamidomethylated peptide. On the other hand, the intensity of a peak corresponding to the carbamidomethylated C-terminal peptide was stronger than that of the farnesylated one in the tGelsolin-CAIL mutant protein (Fig. 4A). We confirmed that these mutant proteins were farnesylated after the addition of FPP to the *in vitro* translation reaction mixture by MS/MS analyses (data not shown). These results suggested that the tGelsolin-CAIX mutants containing C-terminals A, C, (F), M, Q, or S were effectively farnesylated, and these C-terminal amino acid preferences were quite similar to those observed for a rat FTase [11].

When *in vitro* translation reactions were performed in the presence of GGPP, peaks equivalent to the geranylgeranylated C-terminal tryptic peptide were observed in the tGelsolin-CAIF, -CAIM, and -CAIL mutant proteins (Table 2 and Fig. 4B), although a peak corresponding to the carbamidomethylated peptide in the tGelsolin-CAIF mutant protein was also detected. Furthermore, geranylgeranylation of the tGelsolin-CAIF, -CAIM, and -CAIL mutant proteins at their C-termini was confirmed by MS/MS analyses (data not shown). However, in the case of tGelsolin-CAIA, -CAIC, -CAIQ, and -CAIS mutant proteins, geranylgeranylated C-terminal tryptic peptides could not be detected at all (Table 2 and Fig. 4B). These results suggested that the C-terminal (F), L, and M mutants of tGelsolin-CAIX were effectively geranylgeranylated. The C-terminal amino acid sequence requirements for protein geranylgeranylation in the insect cell-free protein synthesis system showed high similarity to

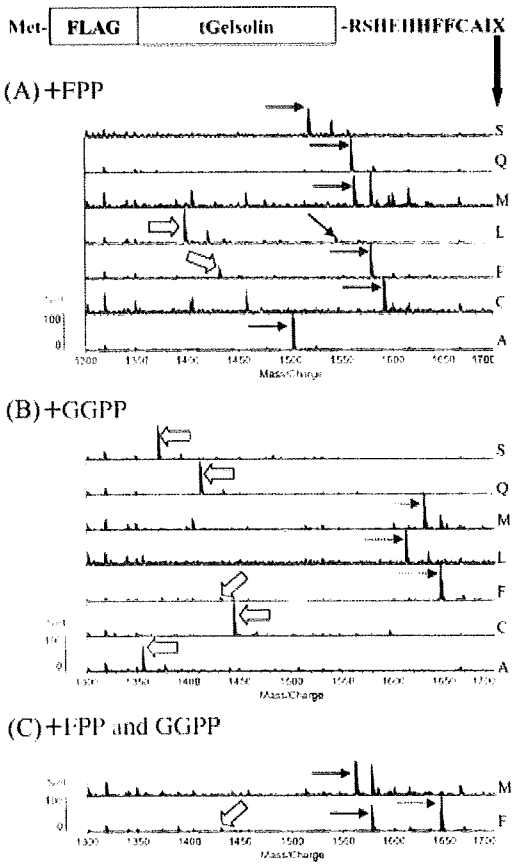

Figure 4. Effect of the C-terminal amino acid residue on the type of protein prenylation reaction. The mRNAs encoding tGelsolin-CAIX (X = A, C, F, M, L, Q, and S) were translated in the presence of FPP (A) or GGPP (B). In the case of tGelsolin-CAIF and tGelsolin-CAIM constructs, these mRNAs were also translated with the addition of both FPP and GGPP (25 μM each) (C). Tryptic digests of affinity-purified mutant proteins were analyzed as described in Section 2. Open boxed, solid, and dotted arrows indicate peaks corresponding to the C-terminal tryptic peptides containing a carbamidomethylated, farnesylated, or geranylgeranylated cysteine in the CAAX motif, respectively.

those observed for a rat GGTase I, except for a C-terminal cysteine, in which case the rat GGTase I could catalyze the TKCVIC hexapeptide [11].

We also performed cell-free protein synthesis of the tGelsolin-CAIF and -CAIM mutants in the presence of both FPP and GGPP (final 25 or 50 μM each), because these mutant proteins were modified with a farnesyl and geranylgeranyl group when either of the prenyl pyrophosphates was added to the reaction mixture. MS spectra of the tryptic digests from each protein were almost identical, regardless of the concentration of added prenyl pyrophosphates (data not shown). A mass spectrum of tryptic digests of tGelsolin-CAIM mutant protein suggested that the mutant protein was predominantly farnesylated but was not geranylgeranylated (Fig. 4C). On the other hand, in MS of the tGelsolin-CAIF mutant protein, we observed peaks corresponding to both farnesylated and geranylgeranylated C-terminal tryptic peptide in the presence of FPP and GGPP (Fig. 4C). These results indicate that the tGelsolin-CAIM mutant protein is the favored substrate for FTase, rather than GGTase I, and that the tGelsolin-CAIF mutant protein can serve as a substrate for both prenyltransferases. These findings were in good agreement with a previous report [19].

3.3 Analysis of prenylation occurring on the rhoC protein

To evaluate whether our strategy to analyze protein prenylation is applicable to naturally prenylated proteins, we chose human rhoC as a model protein, because geranylgeranylation of rhoC has been established by *in vivo* metabolic labeling [18]. N-terminal FLAG-tagged rhoC was synthesized using the insect cell-free protein synthesis system with the addition of FPP and/or GGPP or without the addition of prenyl pyrophosphate. When GGPP was added to the reaction mixture, the resulting affinity-purified proteins migrated more slowly than those obtained by translation without GGPP (Fig. 5). Lys-C digests from these protein bands were analyzed by MALDI-TOF MS. MS spectra of these samples were almost identical and showed that these bands were rhoC (Fig. 6A). When the cysteine residue in the CAAX motif was modified by a carbamidomethyl, farnesyl, or geranylgeranyl group, the calculated monoisotopic mass values of the

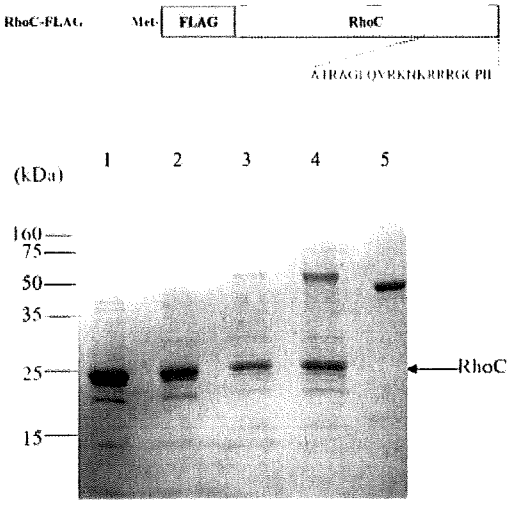

Figure 5. SDS-PAGE of the affinity-purified rhoC proteins. The mRNA transcribed from the pTD1-rhoC-FLAG was translated using the insect cell-free protein synthesis system without the addition of prenyl pyrophosphate (lane 1) or with the addition of FPP (lane 2), GGPP (lane 3), or FPP and GGPP (25 μM each) (lane 4). Three microliters (lanes 1 and 2) and 6 μL (lanes 3 and 4) of the concentrates were electrophoresed on a 15% SDS-PAGE. Lane 5: purified tGelsolin (1 μg) [4].

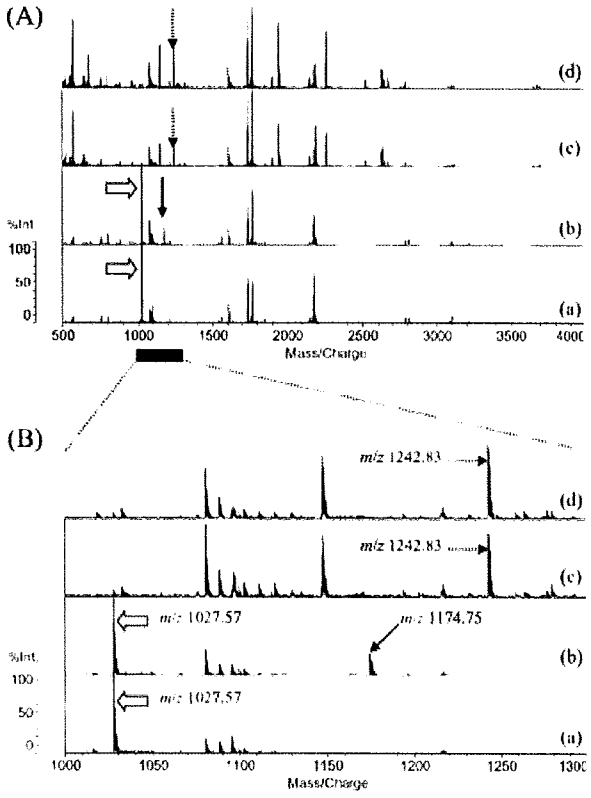

Figure 6. MALDI-mass spectra of Lys-C digests of the rhoC proteins. The mRNA transcribed from pTD1-rhoC-FLAG was translated without the addition of prenyl pyrophosphate (a) or with the addition of FPP (b), GGPP (c), or FPP and GGPP (25 μM each) (d). The acquired profiles focused on a mass/charge range from 500 to 4000 (A) and from 1000 to 1300 (B). Open boxed, solid, and dotted arrows indicate peaks corresponding to the C-terminal Lys-C digests containing a carbamidomethylated, farnesylated, or geranylgeranylated cysteine in the CAAX motif, respectively.

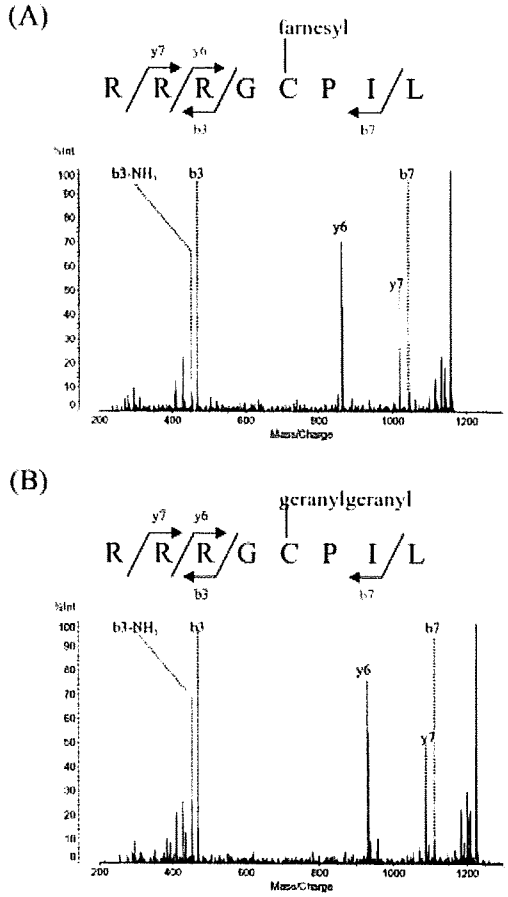

Figure 7. MALDI-MS/MS spectra of the probably prenylated C-terminal peptides from the rhoC proteins. MS/MS analyses were performed for the peaks at *m/z* 1174.75 (A) and *m/z* 1242.83 (B) that were detected in Fig. 6. The observed fragment ions are indicated in the sequences shown.

C-terminal peptides obtained by Lys-C digestion of the rhoC protein were 1027.59, 1174.76, and 1242.82, respectively. In the translation reaction without prenyl pyrophosphate, only a peak equivalent to the carbamidomethylated C-terminal peptide was observed at *m/z* 1027.57 (Fig. 6B-a). This result indicated that the rhoC protein, as in the case of tGelsolin-CAIX mutant proteins, was not modified with a prenyl group without the addition of prenyl pyrophosphate to the insect cell-free protein synthesis system. When FPP was added to the translation reaction mixture, two peaks corresponding to the carbamidomethylated and farnesylated C-terminal peptides were observed at *m/z* 1027.57 and 1174.75, respectively. This result suggested that the rhoC protein could serve as a substrate for FTase, and it was partially farnesylated in the insect cell-free protein synthesis system. A peak equivalent to the geranylgeranylated C-terminal peptide was detected at *m/z* 1242.83 when GGPP or both of the two prenyl pyrophosphates were added to the translation reaction mixture, but the carbamidomethylated peak was barely detectable (Fig. 6B-c and B-d). These results strongly indicated that the rhoC protein was a better substrate for GGTase I than for FTase. The two peaks that probably corresponded to prenylated peptides, at *m/z* 1174.75 and 1242.83, were identified as the C-terminal Lys-C peptides containing farnesyl and geranylgeranyl groups, respectively, by MS/MS analyses (Fig. 7).

4 Discussion

FTase and GGTase I recognize the C-terminal CAAX motif in proteins. Small GTP-binding proteins [7, 10, 12, 14, 18], nuclear lamins [20], and γ-subunits of heterotrimeric G proteins [21, 22] have been identified as substrates for these prenyltransferases. Protein prenylation plays a key role in the functions of these proteins. It has been thought that the C-terminal amino acid residue in the CAAX motif would be a

key element directing which kinds of prenyl groups are attached to the cysteine residue [7, 8, 20, 21]. MS analyses were performed in order to evaluate the ability of an insect cell-free protein synthesis system, Transdirect *insect cell*, to generate proper protein prenylation, and to establish an effective strategy to analyze the PTM in detail. We found that the insect cell-free protein synthesis system, as is the case with the rabbit reticulocyte lysate system, possesses both prenyltransferases. Furthermore, the substrate specificities of prenyltransferases in the insect cell-free protein synthesis system showed a high similarity to those in mammalian prenyltransferases.

It has been shown that FPP is converted to GGPP in cholesterol synthesis pathways in eukaryotic cells and in rabbit reticulocyte lysate [14]. In the insect cell-free protein synthesis system, we could not detect an MS peak equivalent to the geranylgeranylated C-terminal peptide in the tGelsolin-CAIX and rhoC proteins after FPP was added to the *in vitro* translation reaction mixture. This might have been due to the addition of excess FPP, because protein prenylation could be detected by metabolic labeling using [^{3}H]mevalonic acid in the insect cell-free protein synthesis system (data not shown).

In naturally prenylated proteins having the C-terminal CAAX motif, the three amino acid peptide (AAX) of the CAAX motif is cleaved by an endoprotease [23, 24] after prenylation. The newly exposed carboxyl group of the prenylated cysteine residue is then methylated on its α-carboxyl group by a methyltransferase [25, 26]. These enzymes are specifically localized in microsomes. Our results, however, indicated that the C-terminal proteolytic cleavage and carboxyl methylation did not occur in the cell-free synthesized proteins. We think that adding microsomal membranes to the insect cell-free protein synthesis system, however, could generate these PTMs, because *N*-linked glycosylation and signal peptide cleavage occurred in the presence of canine pancreatic microsomal membranes in our system (Utsumi, T. *et al.*, unpublished results).

As described above, the X residue in the CAAX motif is a major determinant for substrate specificities of prenyltransferases. However, recent studies have shown that these enzyme substrate specificities are more complex. For example, N-ras, K-ras4A, and K-ras4B, of which each have a methionine residue at their C-terminus, were effective substrates for both prenyltransferases *in vitro* [10] and *in vivo* [27]. Similarly, rhoB protein, which has a C-terminal leucine residue, was found to be modified with both farnesyl and geranylgeranyl groups *in vitro* and *in vivo* [18]. Our present results show that mutant proteins of tGelsolin and rhoC protein having C-terminal F, M, and L residues were recognized by both prenyltransferases. Thus, it is impossible to predict which prenyl group(s) will attach to the cysteine residue in the CAAX motif based on the C-terminal residue, and this must be determined experimentally. On the other hand, it might be possible to predict whether target proteins containing the CAAX motif are modified *in vivo* by the farnesyl and/or geranylgeranyl group by the addition of both prenyl pyrophosphates to an insect *in vitro* translation reaction mixture, because tGelsolin-CAIM and rhoC proteins were selectively modified with the farnesyl and geranylgeranyl groups, respectively, in the presence of both prenyl pyrophosphates.

It has been shown that metabolic labeling using [^{3}H]mevalonic acid or [^{3}H]mevalonolactone is an effective method to analyze protein prenylation. Prenyl group structures can be identified by GC [21], gel-permeation chromatography [7, 14], or by HPLC [12, 18, 28] after releasing the protein-bound lipids with Raney nickel or methyl iodide. However, it is necessary to construct a mutant (typically serine) of the cysteine residue in the CAAX motif to identify the exact location of the modification. Mass spectrometric analyses not only provide information on the lipid structures, but they also give the exact location of the modification. Thus, the combination of an insect cell-free protein synthesis system and MS could be an effective strategy to accurately characterize protein prenylation.

We are grateful to Mr. Shinichiro Kobayashi, Dr. Masaki Yamada, and Mr. Daisuke Nakayama, Shimadzu Corporation, for helpful discussions of mass spectrometric analyses.

5 References

[1] Sawasaki, T., Ogasawara, T., Morishita, R., Endo, Y., A cell-free protein synthesis system for high-throughput proteomics. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 14652–14657.

[2] Sakurai, N., Moriya, K., Suzuki, T., Sofuku, K. *et al.*, Detection of co- and post-translational protein N-myristoylation by metabolic labeling in an insect cell-free protein synthesis system. *Anal. Biochem.* 2007, *362*, 236–244.

[3] Ezure, T., Suzuki, T., Higashide, S., Shintani, E. *et al.*, Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol. Prog.* 2006, *22*, 1570–1577.

[4] Suzuki, T., Ito, M., Ezure, T., Shikata, M. *et al.*, N-terminal protein modifications in an insect cell-free protein synthesis system and their identification by mass spectrometry. *Proteomics* 2006, *6*, 4486–4495.

[5] Casey, P. J., Lipid modifications of G proteins. *Curr. Opin. Cell Biol.* 1994, *6*, 219–225.

[6] Marshall, C. J., Protein prenylation: A mediator of protein-protein interactions. *Science* 1993, *259*, 1865–1866.

[7] Kinsella, B. T., Erdman, R. A., Maltese, W. A., Posttranslational modification of Ha-ras p21 by farnesyl versus geranylgeranyl isoprenoids is determined by the COOH-terminal amino acid. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 8934–8938.

[8] Yokoyama, K., Goodwin, G. W., Ghomashchi, F., Glomset, J. A. *et al.*, A protein geranylgeranyltransferase from bovine brain: Implications for protein prenylation specificity. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 5302–5306.

[9] Yokoyama, K., Zimmerman, K., Scholten, J., Gelb, M. H., Differential prenyl pyrophosphate binding to mammalian protein geranylgeranyltransferase-I and protein farnesyltransferase and its consequence on the specificity of protein prenylation. *J. Biol. Chem.* 1997, *272*, 3944–3952.

[10] Zhang, F. L., Kirschmeier, P., Carr, D., James, L. *et al.*, Characterization of Ha-ras, N-ras, Ki-ras4A and Ki-ras4B as *in vitro* substrates for farnesyl protein transferase and geranylgeranyl protein transferase type I. *J. Biol. Chem.* 1997, *272*, 10232–10239.

[11] Hartman, H. L., Hicks, K. A., Fierke, C. A., Peptide specificity of protein prenyltransferases is determined mainly by reactivity rather than binding affinity. *Biochemistry* 2005, *44*, 15314–15324.

[12] Farrell, F. X., Yamamoto, K., Lapetina, E. G., Prenyl group identification of rap2 proteins: A ras superfamily member other than ras that is farnesylated. *Biochem. J.* 1993, *289*, 349–355.

[13] Jackson, R. J., Hunt, T., Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods Enzymol.* 1983, *96*, 50–74.

[14] Kinsella, B. T., Erdman, R. A., Maltese, W. A., Carboxyl-terminal isoprenylation of *ras*-related GTP-binding proteins encoded by *rac*1, *rac*2, and *ral*A. *J. Biol. Chem.* 1991, *266*, 9786–9794.

[15] Sakurai, N., Utsumi, T., Posttranslational *N*-myristoylation is required for the anti-apoptotic activity of human tGelsolin, the C-terminal caspase-cleavage product of human gelsolin. *J. Biol. Chem.* 2006, *281*, 14288–14295.

[16] Suzuki, T., Ito, M., Ezure, T., Kobayashi, S. *et al.*, Performance of expression vector, pTD1, for insect cell-free translation system. *J. Biosci. Bioeng.* 2006, *102*, 69–71.

[17] Chardin, P., Madaule, P., Tavitian, A., Coding sequence of human rho cDNAs clone 6 and clone 9. *Nucleic Acids Res.* 1988, *16*, 2717.

[18] Adamson, P., Marshall, C. J., Hall, A., Tilbrook, P. A., Post-translational modifications of $p21^{rho}$ proteins. *J. Biol. Chem.* 1992, *267*, 20033–20038.

[19] Reid, T. S., Terry, K. L., Casey, P. J., Beese, L. S., Crystallographic analysis of CaaX prenyltransferases complexed with substrates defines rules of protein substrate selectivity. *J. Mol. Biol.* 2004, *343*, 417–433.

[20] Vorburger, K., Kitten, G. T., Nigg, E. A., Modification of nuclear lamin proteins by a mevalonic acid derivative occurs in reticulocyte lysates and requires the cysteine residue of the C-terminal CXXM motif. *EMBO J.* 1989, *8*, 4007–4013.

[21] Lai, R. K., Perez-sala, D., Canada, F. J., Rando, R. R., The γ subunit of transducin is farnesylated. *Proc. Natl. Acad. Sci. USA* 1990, *87*, 7673–7677.

[22] Kasai, H., Satomi, Y., Fukuda, Y., Takao, T., Top-down analysis of protein isoprenylation by electrospray ionization hybrid quadrupole time-of-flight tandem mass spectrometry; the mouse Tγ protein. *Rapid Commun. Mass Spectrom.* 2005, *19*, 269–274.

[23] Ma, Y. T., Rando, R. R., A microsomal endoprotease that specifically cleaves isoprenylated peptides. *Proc. Natl. Acad. Sci. USA* 1992, *89*, 6275–6279.

[24] Jang, G. F., Yokoyama, K., Gelb, M. H., A prenylated protein-specific endoprotease in rat liver microsomes that produces a carboxyl-terminal tripeptide. *Biochemistry* 1993, *32*, 9500–9507.

[25] Hrycyna, C. A., Sapperstein, S. K., Clarke, S., Michaelis, S., The *Saccharomyces cerevisiae STE14* gene encodes a methyltransferase that mediates C-terminal methylation of a-factor and RAS proteins. *EMBO J.* 1991, *10*, 1699–1709.

[26] Romano, J. D., Schmidt, W. K., Michaelis, S., The *Saccharomyces cerevisiae* prenylcysteine carboxyl methyltransferase Ste14p is in the endoplasmic reticulum membrane. *Mol. Biol. Cell* 1998, *9*, 2231–2247.

[27] Rowell, C. A., Kowalczyk, J. J., Lewis, M. D., Garcia, A. M., Direct demonstration of geranylgeranylation and farnesylation of Ki-ras *in vivo*. *J. Biol. Chem.* 1997, *272*, 14093–14097.

[28] Whitten, M. E., Yokoyama, K., Schieltz, D., Ghomashchi, F. *et al.*, Structural analysis of protein prenyl groups and associated C-terminal modifications. *Methods Enzymol.* 2000, *316*, 436–451.

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2007; **21**: 3329–3336
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rcm.3215

RCM

Specific isolation of N-terminal fragments from proteins and their high-fidelity *de novo* sequencing

Minoru Yamaguchi[1]*, Takashi Obama[1], Hiroki Kuyama[2], Daisuke Nakayama[1], Eiji Ando[1], Taka-aki Okamura[3], Norikazu Ueyama[3], Takashi Nakazawa[4], Shigemi Norioka[5], Osamu Nishimura[1,2] and Susumu Tsunasawa[1,2]

[1]Life Science Laboratory, Shimadzu Corporation, Kyoto 604-8511, Japan
[2]Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan
[3]Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan
[4]Department of Chemistry, Nara Women's University, Nara 630-8506, Japan
[5]Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan

Received 29 April 2007; Revised 23 July 2007; Accepted 12 August 2007

A new method to determine N-terminal amino acid sequences of multiple proteins at low pmol level by a parallel processing has been developed. The method contains the following five steps: (1) reduction, S-alkylation and guanidination for targeted proteins; (2) coupling with sulfosuccinimidyl-2-(biotinamido)ethyl-1,3-dithiopropionate(sulfo-NHS-SS-biotin) to N^{α}-amino groups of proteins; (3) digestion of the modified proteins by an appropriate protease; (4) specific isolation of N-terminal fragments of proteins by affinity capture using the biotin-avidin system; (5) *de novo* sequence analysis of peptides by MALDI-TOF-/MALDI-TOF-PSD mass spectrometry with effective utilization of the CAF (chemically assisted fragmentation) method.[1] This method is also effective for N-terminal sequencing of each protein in a mixture of several proteins, and for sequencing components of a multiprotein complex. It is expected to become an essential proteomics tool for identifying proteins, especially when used in combination with a C-terminal sequencing method.[2,3]

Mass spectrometric analysis has become a typical means to identify proteins in proteomics studies. Many proteins separated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) or by liquid chromatography (LC) have been analyzed by effectively using various types of mass spectrometers in combination with genome and protein databases. These analyses are extremely useful, but neither peptide mass finger printing (PMF)[4–7] nor MS/MS ion search analyses[8,9] have sufficient power to resolve the complete primary structures of proteins acting in living cells in their mature forms. Such analyses often fail to identify residues modified post-translationally, or to detect sequence polymorphisms, etc. Determination of N- and C-terminal sequences of proteins has also been overlooked in spite of their importance to function, for example, in dynamic proteome analysis that focuses on turnover of key proteins by the N-end rule.[10,11] This is because the methods to identify proteins by mass spectrometry often do not cover the full length of the target proteins. Moreover, recent information from genome analyses indicates that the number of proteins exceeds by far the number of genes in the genome. This diversity arises from the fact that distinct proteins can be generated even from a single gene as a result of alternative splicing of primary transcripts.[12] Genomic databases do not offer information on such alternative splicings, wherein exons can be shuffled to create different proteins from a single gene, and therefore both high-fidelity and high-throughput protein analysis is desirable. Therefore, as a tool for high-fidelity identification of targeted proteins, it is very important to establish an easy method for N-terminal sequence analysis.

Recently, several methods for isolation of N-terminal peptides from proteins especially and for sequencing them have been reported.[13–15] Gevaert *et al.* developed a method to sort N-terminal peptides by combined fractional diagonal chromatography and determine their sequence by LC/tandem mass spectrometric (MS/MS) analysis and database searching.[13] McDonald *et al.* reported a method for isolating N-terminal peptides by combining an acetylating and biotin-avidin system.[15] Undoubtedly, sequencing of N-termini of proteins by utilizing both LC/MS/MS and database search is suitable for large-scale protein analysis in which a highly complex mixture of samples in a living organism by cooperation with its genome database is analyzed. However, these methods will need multiple LC/MS/MS analyses that require time when many independent candidate proteins exist. On the other hand, in the case that several target proteins separated on 2D-PAGE are

*Correspondence to: M. Yamaguchi, Life Science Laboratory, Shimadzu Corporation, Kyoto 604-8511, Japan.
E-mail: yamagu@shimadzu.co.jp
Contract/grant sponsor: Grant-in-Aid for Scientific Research; contract/grant number: A-15201043.

analyzed, use of matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is easier for both operating and maintaining the analytical system. Therefore, this instrument has been widely used also among many researchers who are not mass analysis specialists.

Here, we have developed a method involving specific isolation of the N-terminal peptides from proteins separated by 1D/2D-PAGE followed by their *de novo* sequencing by MALDI-TOF-/MALDI-TOF-PSD analysis. This method is effective for analysis at low pmol amounts of sample and high throughput by performing parallel treatment of samples. In addition, the method is highly reliable for sequencing and fully overcomes the inherent drawbacks of other techniques, i.e. the requirements for high purity of the sample protein and the low-throughput performance of the conventional Edman method. The basic principles of the method including introduction of both biotin and sulfonic acid groups to the N-terminus of a protein have already been reported, but, in the study reported in this paper, we have developed a more simplified method using sulfo-NHS-SS-biotin instead of biotinyl cysteic acid (BCA) as described above. A general scheme depicting the procedure is shown in Fig. 1. In this method, trypsin is generally used to produce the N-terminal peptides from proteins because the existence of arginine or homoarginine at their C-termini is an effective CAF (chemically assisted fragmentation) method. However, in cases where tryptic peptides caused from the N-termini are too large or too small, which is disadvantageous for *de novo* sequence analysis, another protease such as GluC or AspN can be used. Especially in cases where the genome sequences are already known, the N-terminal amino acid sequence of the target protein can be more easily determined by mass spectrometry analysis, because it is possible to determine its amino acid sequence by using only the *m/z* value without further *de novo* sequencing using post-source decay (PSD). We also report the optimization of the method and its application to proteins prepared in-gel by 1D/2D-PAGE.

EXPERIMENTAL

Chemicals and reagents

Ribonuclease A (RNase A: bovine pancreas), cytochrome *c* (*Saccharomyces cerevisiae*), bovine serum albumin (BSA), glyceraldehyde-3-phosphate dehydrogenase (G3P), α-cyano-4-hydroxycinnamic acid (CHCA) and lyophilized *Escherichia coli* (K-12) extracts were from Sigma (St. Louis, MO, USA). Sequencing-grade modified trypsin (TPCK-trypsin) and SoftLink™ soft release avidin resin were from Promega (Madison, WI, USA). *O*-Methylisourea hemisulfate was from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Sulfo-NHS-SS-biotin was obtained from Pierce Co. (Rockford, IL, USA) and ZipTip C18 was purchased from Millipore Corp. (Bedford, MA, USA). All other chemicals were analytical reagent grade and were used without further purification.

Specific isolation of N-terminal peptides

A series of chemical modifications of proteins, including S-alkylation, guanidination and N-terminal derivatization,

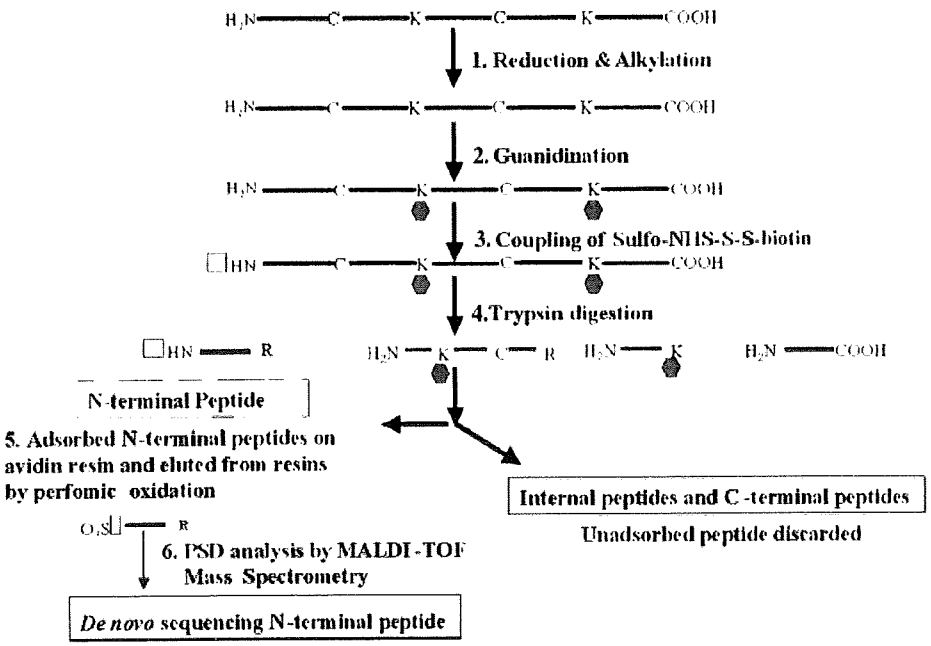

Figure 1. The protocol is summarized for isolation of N-terminal peptides. The method consists of the following five steps: (1) reduction, S-alkylation and (2) guanidination of targeted proteins; (3) coupling of sulfo-NHS-SS-biotin to N^{α}-amino groups of proteins; (4) digestion of the modified proteins by trypsin; (5) specific isolation of N-terminal fragments of proteins by affinity capture using a biotin-avidin system; (6) *de novo* sequence analysis of peptides by MALDI-TOF-/MALDI-TOF-PSD-MS with effective utilization of the CAF method.

DOI: 10.1002/rcm

was usually carried out in polyacrylamide gels: Protein samples separated by 1D or 2D sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were stained with Coomassie Brilliant Blue and the corresponding protein spots were excised. The gel pieces in microtubes were washed successively with 100 μL of 100 mM ammonium bicarbonate solution and acetonitrile and dried in a SpeedVac evaporator. The dried gel pieces were incubated in 100 μL of 10 mM dithiothreitol/100 mM ammonium bicarbonate solution for 60 min at 56°C. After centrifugation the reagent solution was discarded, then replaced with 100 μL of freshly prepared 55 mM iodoacetamide/100 mM ammonium bicarbonate solution. This was incubated for 45 min at room temperature under protection from light. The gel piece was then washed with 100 mM ammonium bicarbonate solution with acetonitrile in triplicate and completely dried in a vacuum concentrator. A mixture of 5 mg *O*-methylisourea hemisulfate and 30 μL of 7 M NH_4OH was then added and incubated at 65°C for 10 min, then washed with 0.2 M phosphate buffer (pH 7.2), 100 mM ammonium bicarbonate solution and acetonitrile in duplicate, then dried in a SpeedVac evaporator. Next, 30 μL of 0.2 M phosphate buffer (pH 7.6) containing 0.2 mg of sulfo-NHS-SS-biotin were added and incubated at 37°C for 1 h. The gel piece was washed with 100 mM ammonium bicarbonate solution followed by acetonitrile in triplicate and dried in a SpeedVac evaporator. After addition of trypsin and 0.1% n-octyl beta-D-glucopyranoside, followed by addition of 15 μL of 50 mM ammonium bicarbonate solution containing 0.1% n-octyl beta-D-glucopyranoside, it was incubated at 37°C for 4 h. The resulting tryptic peptides were extracted twice with 40 μL 50% acetonitrile/0.1% trifluoroacetic acid (TFA) solution. The recovered solution was finally dissolved in 0.1 M phosphate buffer (pH 7.2)

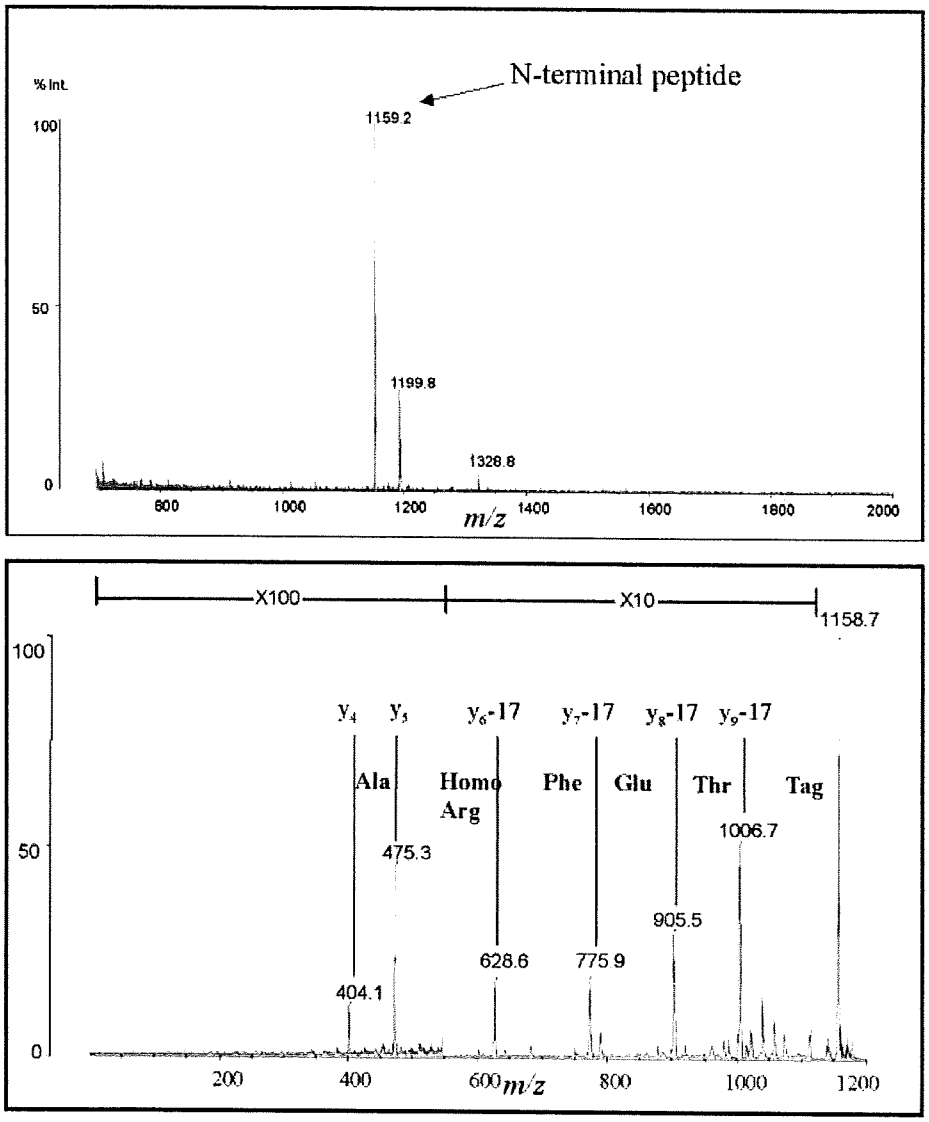

Figure 2. MALDI spectra corresponding to each N-terminal sequence obtained for cytochrome *c*. MALDI-TOF MS spectrum of the N-terminal peptide specifically isolated from the trypsin digests is shown (top). MALDI-PSD spectrum of the N-terminal peptide prepared as described (bottom).

DOI: 10.1002/rcm

containing 0.01% SDS and incubated at 95°C for 30 min. After cooling the sample to room temperature, 10 μL of a 50% suspension of avidin beads in 0.1 M phosphate buffer (pH 7.2) were added and gently mixed at room temperature. The beads were thoroughly washed with 40 μL of 0.1 M phosphate buffer containing 0.01% SDS and then with 120 μL of 0.1 M phosphate buffer (pH 7.2) to remove all internal tryptic peptides. After the supernatant containing a suspension of beads had been removed by gentle centrifugation, 50 μL of freshly prepared performic acid (950 μL of 99% formic acid and 50 μL of hydrogen peroxide solution was mixed and left to stand at room temperature for 2 h) were added to the beads, incubated at 4°C for 1 h, and the supernatant of the bead suspension was collected. The resulting sulfonated N-terminal fragment was recovered by washing the beads with 30 μL of 50% acetonitrile/0.1% TFA solution followed by 150 μL of distilled water, and evaporated in a SpeedVac evaporator.

Two-dimensional electrophoresis (2DE) of an *Escherichia coli* extract

A crude cell extract of *Escherichia coli* strain K-12 was prepared using the ProteoPrep™ sample extraction kit (Sigma) according to the manufacturer's protocol. Samples containing 500 μg of protein dissolved in rehydration buffer containing 0.2% carrier ampholyte (Amersham Pharmacia Biotech, Bucks, UK) were rehydrated on an Immobiline dry strip gel of pH 5.0–8.0 (Amersham Pharmacia Biotech) and isoelectric focusing was performed. After performing 2DE, separated proteins were detected by Coomassie Brilliant

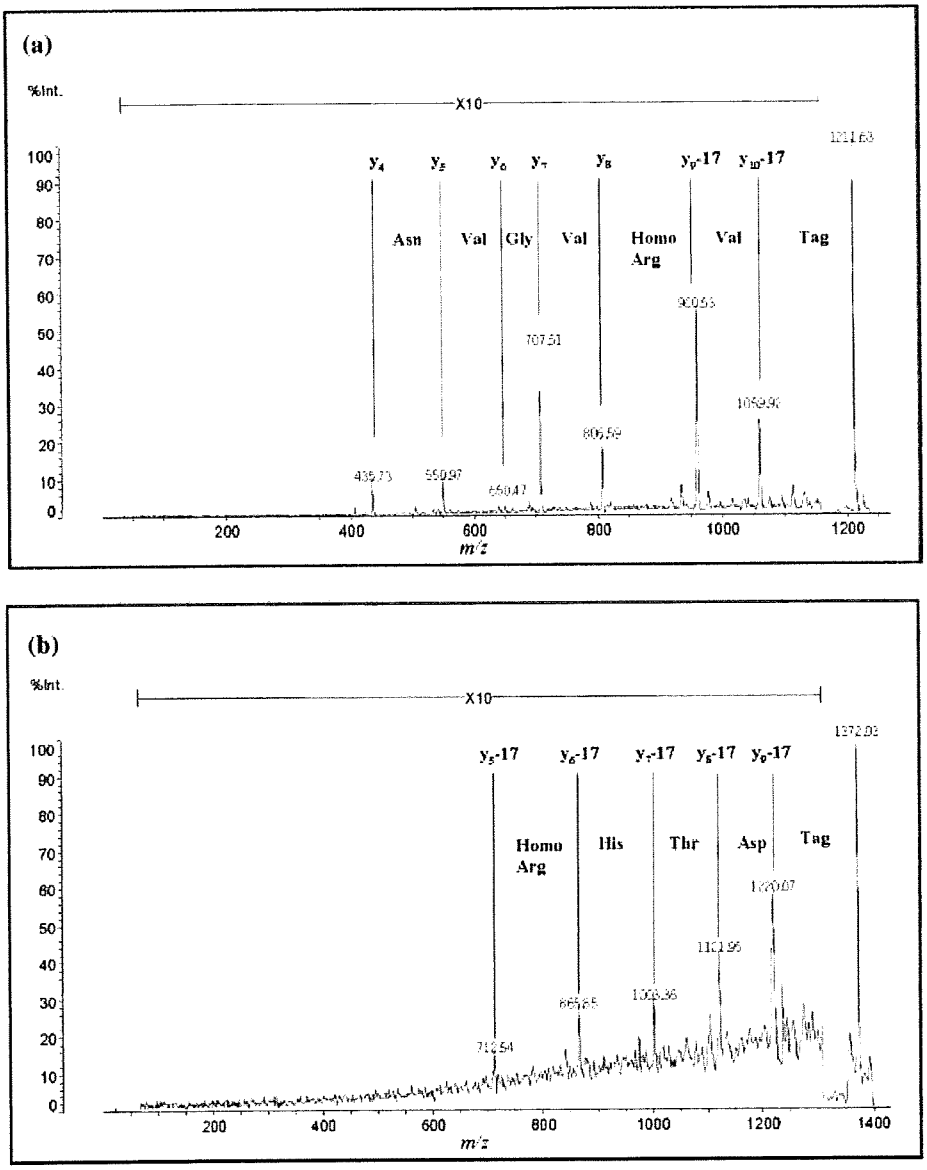

Figure 3. MALDI-PSD spectra of the N-terminal peptides obtained from G3P (a) and BSA (b) when 1 pmol of protein was separated by SDS-PAGE.

DOI: 10.1002/rcm