FEBS Open Bio 5 (2015) 844–851

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

Production and characterization of neurosecretory protein GM using *Escherichia coli* and Chinese Hamster Ovary cells

Keiko Masuda, Megumi Furumitsu, Shusuke Taniuchi, Eiko Iwakoshi-Ukena, Kazuyoshi Ukena *

Section of Behavioral Sciences, Graduate School of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan

ARTICLE INFO

Article history:
Received 12 August 2015
Revised 6 October 2015
Accepted 9 October 2015

Keywords:
Neurosecretory protein GM
Neurosecretory protein GL
Disulfide bond
Free cysteine residue
Escherichia coli
Chinese Hamster Ovary cell

ABSTRACT

Neurosecretory protein GL (NPGL) **and** ***neurosecretory protein GM (NPGM)*** **are paralogs recently discovered in birds and in mammals. The post-translational products of *NPGL* and of *NPGM* genes include a signal peptide sequence, a glycine amidation signal, and a dibasic amino acid cleavage site. This suggests that the mature forms of NPGL and of NPGM are small proteins secreted in the hypothalamus and containing an amidated C-terminus. However, endogenous NPGL and NPGM have not yet been identified. Chicken NPGL and NPGM have two highly conserved Cys residues that are likely to form a disulfide bond, while mammalian NPGM has one additional Cys residue located between the two conserved Cys residues and the correct disulfide bond pattern is unclear. In this study, we prepared rat NPGM to elucidate the structure of its mature form. We first expressed the predicted mature NPGM, containing an extra C-terminal Gly, in *Escherichia coli* SHuffle cells, which are engineered to promote the formation of native disulfide bridges in recombinant proteins. We observed the presence of a disulfide bond between the N-terminal Cys residue and the second Cys residue, while the C-terminal Cys residue was free. Secondly, we transfected a construct containing the entire NPGM open reading frame into Chinese Hamster Ovary cells, and observed that NPGM was cleaved immediately after the signal peptide and that it was secreted into the medium. Furthermore, the protein presented a disulfide bond at the same location observed in recombinant NPGM.**

1. Introduction

Using a cDNA subtraction screen, we recently identified a novel cDNA highly expressed in the chicken hypothalamus [1]. A genome database search revealed the existence of a paralog gene; both are widely conserved in vertebrates [1]. The predicted post-translational products of the two genes from chicken, human, and rat are shown schematically in Fig. 1A. Each precursor protein includes a signal peptide at the N-terminus, a glycine amidation signal, and a dibasic amino acid cleavage site. This architecture suggests that a small 80–88 amino acid protein containing an amidated C-terminus is secreted as a mature form. The predicted mature proteins have characteristic C-terminal sequences, Gly-Leu-NH_2 or Gly-Met-NH_2, which suggested their name: neurosecretory protein GL (NPGL) and neurosecretory protein GM (NPGM), respectively [1]. In chicks, subcutaneous infusions of rat NPGL, which shares 83% identity with chicken NPGL, caused an increase of the body weight without affecting food intake [1]. The function of NPGM is still unknown.

NPGL and NPGM have highly conserved Cys residues that likely form an intramolecular disulfide bond. Disulfide bonds are important for the biological activity and for the stability of a protein [2]. For example, insulin has one intra-molecular and two inter-molecular disulfide bonds that stabilize its secondary structure and contribute to its overall hydrophobicity; when any one of these bonds is lost, the receptor binding activity decrease below 0.02% [3]. It is noteworthy that the number of Cys residues in mammalian NPGL and NPGM are different. NPGL contains two Cys residues, which likely form a disulfide bond. On the other hand, mammalian NPGM has two Cys residues at the conserved positions and one additional Cys residue, namely Cys^1, Cys^2, and Cys^3 (Fig. 1A). These three Cys residues can potentially give rise to three different disulfide bonds, Cys^1–Cys^2, Cys^1–Cys^3, or Cys^2–Cys^3; from NPGM primary structure is however impossible to claim which one will form in the native protein.

Abbreviations: CHO cells, Chinese Hamster Ovary cells; *E. coli*, *Escherichia coli*; NPGL, neurosecretory protein GL; NPGM, neurosecretory protein GM; SP, signal peptide; TF, trigger factor

* Corresponding author. Tel.: +81 82 424 6571; fax: +81 82 424 0759.
E-mail address: ukena@hiroshima-u.ac.jp (K. Ukena).

http://dx.doi.org/10.1016/j.fob.2015.10.002
2211-5463/

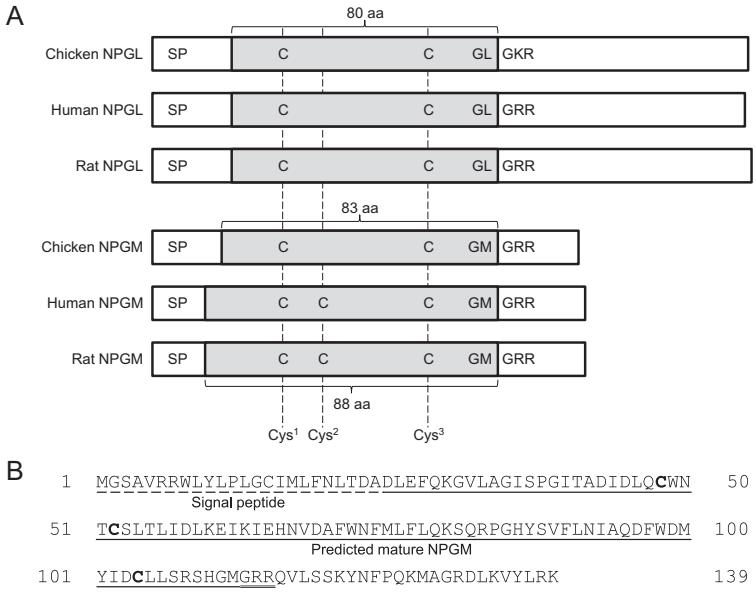

Fig. 1. Primary structure of NPGL and NPGM precursor proteins. (A) Schematic alignment of NPGL and NPGM precursor proteins derived from chicken, human, and rat genomic DNA sequences. The amino acid sequences include a signal peptide (SP), two or three Cys residues (C), sequences characteristic of each species (GL or GM), a C-terminal amidation signal and a dibasic processing site (GKR or GRR). The predicted mature peptide regions are painted gray and their amino acid lengths are indicated. The positions of the conserved Cys residues, named Cys^1 to Cys^3 for descriptive purpose, are indicated by dash line. (B) Amino acid sequence of rat NPGM precursor protein. The signal peptide and the mature NPGM are underlined with a dash line and a solid line, respectively. The C-terminal amidation signal and the dibasic processing site are underlined with a double line. The three Cys residues are highlighted in bold.

To determine the structure of the mature protein and the location of the disulfide bond, we produced rat NPGM in *Escherichia coli* (*E. coli*) and in Chinese Hamster Ovary (CHO) cells. The amino acid sequence of the precursor protein of rat NPGM is shown in Fig. 1B. First, we expressed the predicted mature NPGM in *E. coli* and we determined the location of the disulfide bond in the recombinant protein using protease digestion. Recombinant NPGM was also used as an antigen for raising specific antibody. Secondly, a construct containing the entire NPGM open reading frame was transfected into CHO cells to determine whether NPGM was secreted into the culture medium. Finally, the structure of the secreted NPGM and the location of the disulfide bond were analyzed.

2. Materials and methods

2.1. RNA and cDNA preparation

Male Wistar rats (7 weeks old) were purchased from a commercial company (Kyudo, Saga, Japan), housed on a 12:12 light–dark cycle in a room maintained at 23 ± 2 °C with *ad libitum* access to food and tap water. All animal procedures were performed according to the Guide for the Care and Use of Laboratory Animals prepared by Hiroshima University (Higashi-Hiroshima, Japan).

Rats were sacrificed by decapitation. The medial basal hypothalamus was dissected and snap-frozen in liquid nitrogen for further RNA processing. Total RNA was extracted from the medial basal hypothalamus using the TRIzol reagent (Life technologies, Carlsbad, CA, USA) followed by the isolation of poly(A)+ RNA with Oligotex-(dT) 30 (Takara Bio, Shiga, Japan). The first-strand of cDNA was synthesized from the mRNA using a ReverTra Ace qPCR RT Kit (TOYOBO, Osaka, Japan).

2.2. Construction of the NPGM-Gly expression plasmid

The cDNA encoding NPGM was amplified with a forward primer (5′- GCCGCATATGGACTTGGAATTTCAGAAAGG -3′) containing the *Nde* I site (underlined) and a reverse primer (5'- CGGAATT**CTA**GCCCATCCCATGAGACCTTGAAA -3') containing the *EcoR* I site (underlined), stop codon (bold), and the codon encoding the amidating donor residue, Gly (squared). PCR amplifications were carried out with the Ex Taq polymerase (Takara Bio) using the following program: 95 °C for 20 s, 40 cycles at 95 °C for 20 s, at 55 °C for 20 s, and at 72 °C for 20 s. Additional elongation was performed at 72 °C for 10 min for TA cloning. The insert was ligated into the pGEM-T easy vector (Promega, Madison, WI, USA) using Ligation high (TOYOBO), to produce pGEM-NPGM-Gly plasmid. *E. coli* DH5α cells (Nippon Gene, Tokyo, Japan) were transformed with the plasmid and grown overnight at 37 °C on an LB agar plate containing 50 μg/ml of ampicillin. The colonies were then grown in fresh LB medium containing ampicillin at 37 °C overnight. The amplified plasmids were extracted using NucleoSpin Plasmid (MACHEREY–NAGEL, Düren, Germany).

The pGEM-NPGM-Gly plasmid and pCold TF DNA vector (Takara Bio) were digested separately with *Nde* I and *EcoR* I, and ligated using Ligation high (TOYOBO) to produce pCold-NPGM-Gly plasmid. The plasmid was propagated as described above. The sequence of the insert was confirmed using ABI Prism 310 Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA).

2.3. Expression of recombinant His_6-TF tagged NPGM-Gly

The pCold-NPGM-Gly plasmid was transformed into *E. coli* BL21 strain (GE Healthcare, Little Chalfont, UK) or *E. coli* SHuffle strain (New England Biolabs, Ipswich, MA, USA). The transformants were selected on LB agar plates containing 50 μg/ml of ampicillin and grown at 37 °C overnight. The colonies were then grown in LB medium containing ampicillin at 37 °C overnight. An aliquot of the pre-culture solution was diluted with 200 ml of fresh LB medium and incubated at 37 °C. When the cells reached an optical density $(OD)_{600}$ of 0.5, the culture was refrigerated at 15 °C for 30 min. The culture solution was added with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.1 mM

and continued with shaking at 15 °C for 24 h. Cells were collected by centrifugation and frozen at −80 °C until further use.

The expression of the target protein in soluble and in insoluble fractions was confirmed by SDS–PAGE. An aliquot of the cell pellet was lysed with BugBuster Protein Extraction Reagent (Novagen, Madison, WI, USA) and separated by centrifugation. Each sample was dissolved in SDS sample buffer and subjected to 15% SDS–PAGE. The gels were stained with 2D-SILVER STAIN II (Cosmobio, Tokyo, Japan).

2.4. Affinity purification and tag removal

The cell pellet was fully resuspended with BugBuster Protein Extraction Reagent (Novagen) and centrifuged. The supernatant was purified by Ni–NTA agarose (Qiagen, Venlo, Netherlands). The His_6-TF-NPGM-Gly protein was allowed to bind to the resin for 1 h at 4 °C before elution with the elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 100 mM imidazole, pH 8.0).

The elution solution was dialyzed at 4 °C overnight in a Factor Xa reaction buffer (20 mM Tris–HCl, 100 mM NaCl, 2 mM $CaCl_2$, pH 8.0). The protein solution was further added with 2 μg of Factor Xa protease (New England Biolabs) per 1 mg of protein, and incubated at 20 °C for 16 h. The digestion reaction was stopped with 1 mM phenylmethylsulfonyl fluoride (PMSF). The cleaved His_6-TF-tag was removed by centrifugation after adding acetone at a final concentration of 40%. The supernatant was purified by reverse-phase high-performance liquid chromatography (RP-HPLC) using the TSKgel CN-80Ts column (4.6 × 250 mm; Tosoh, Tokyo, Japan) with a linear gradient of 35–55% acetonitrile (CH_3CN) containing 0.1% TFA over 40 min with a flow rate of 0.5 ml/min. The elution was monitored at 220 nm. The collected peak was evaporated and lyophilized. The mass value was analyzed by electrospray ionization mass spectrometry (ESI-MS) on the LTQ Orbitrap XL and Xcalibur (Thermo Fisher Scientific, Waltham, MA, USA).

2.5. Analysis of disulfide bond of NPGM-Gly expressed in E. coli

0.1 mg of NPGM-Gly was dissolved in 500 μl of a buffer (50 mM iodoacetamide, 50 mM sodium phosphate, pH 7.0) and incubated at room temperature for 2 h under the quenching condition. The alkylated NPGM-Gly was purified by RP-HPLC using a C8 column (10 × 150 mm; YMC, Kyoto, Japan) with a linear gradient of 40–60% CH_3CN containing 0.1% TFA over 40 min with a flow rate of 0.5 ml/min. The elution was monitored at 220 nm. The collected peak was evaporated.

Non-alkylated or alkylated NPGM-Gly was dissolved in 200 μl of 50 mM sodium phosphate buffer (pH 7.0) containing 0.5 μg of endoproteinase Asp-N (Roche Diagnostics, Basel, Switzerland) and incubated at 37 °C overnight. The mixture was purified similarly by RP-HPLC with a linear gradient of 10–60% CH_3CN containing 0.1% TFA over 50 min. The collected peaks were analyzed by ESI-MS.

2.6. C-terminal amidation of recombinant NPGM-Gly

Amidation of NPGM-Gly was performed using peptidylglycine α-amidating enzyme (R&D systems, Minneapolis, MN, USA) as previously published [4]. 1 mg of NPGM-Gly was dissolved in 2 ml of buffer (0.2 M Tris–HCl, pH 7.0, 0.5% Nonidet P-40, 5 mM ascorbic acid, 20 μM $CuSO_4$, 0.1 mg/ml catalase, 2 μg/ml enzyme) and incubated at 37 °C overnight. The mixture was purified by RP-HPLC using a C8 column (4.6 × 150 mm; YMC) with a linear gradient of 40–60% CH_3CN containing 0.1% TFA over 40 min with a flow rate of 0.5 ml/min. The elution was monitored at 220 nm. The collected peaks were analyzed by ESI-MS.

2.7. Production of antibody against NPGM

Antisera were raised following our published method [5] using recombinant NPGM-Gly as antigen. The antigen solution was mixed with Freund's complete adjuvant and injected into a rabbit. After the booster injection, blood was collected and the antibody purified. The specificity of the antibody was confirmed by dot-immunoblot analysis. Rat NPGL [6] and NPGM were blotted on a nitrocellulose membrane and probed with the antibody against NPGM (1:250 dilution) and revealed with an alkaline phosphatase-conjugated secondary antibody (1:1000 dilution).

2.8. Establishment of NPGM-expressing CHO cell line

The full-length open reading frame of NPGM was amplified from cDNA of the medial basal hypothalamus and inserted into pcDNA3.1 (Life technologies), a mammalian expression vector. CHO (*dehydrofolate reductase*$^-$; *dhfr*$^-$) cells were transfected with the NPGM expression vector and dhfr expression vector (pSV2-*dhfr*), using X-treme GENE 9 DNA transfection reagent (Roche Diagnostics) in 6-well plates. The transfected cells were selected in nucleoside-free alpha-MEM (Life technologies) containing 10% fetal bovine serum and 1 mg/ml G418. Gene amplification was then performed by addition of 5–500 nM methotrexate for 6 weeks.

Expression of NPGM was confirmed by immunocytochemistry. The cells transfected with a control vector (mock) or with the NPGM expression vector were re-plated and probed with the antibody against NPGM (1:1000 dilution) and incubated with rhodamine-conjugated goat anti-rabbit IgG (1:1000 dilution).

2.9. Characterization of NPGM expressed by CHO cells

The stably transfected cells were cultured in the serum-free medium Hyclone CDM4CHO SH30557 (Thermo Fisher Scientific). 400 ml of the culture medium was collected for a week. The medium was boiled, cooled and, after addition of acetic acid at final concentration of 5%, centrifuged. The supernatant was fractionated every 2 min by RP-HPLC using a C18 column (TSK gel ODS-80Ts, 4.6 × 150 mm; Tosoh) with a linear gradient of 20–60% CH_3CN containing 0.1% TFA over 40 min with a flow rate of 0.5 ml/min. The elution was monitored at 220 nm. Aliquots of the collected fractions were analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) on the AXIMA-CFR plus (Shimadzu, Kyoto, Japan). The rest of fractions were evaporated, dissolved in SDS sample buffer, and subjected to 15% SDS–PAGE. After transfer onto polyvinylidene fluoride membrane (Immobilon-P; MerckMillipore, Darmstadt, Germany), the blot was probed with the anti-NPGM antibody (1:1000 dilution) and incubated with horseradish peroxidase-conjugated donkey anti-rabbit IgG (GE Healthcare). The protein bands were detected by ECL Western Blotting Detection Reagents (Promega).

Fraction 40 was analyzed in more detail. The culture medium was collected and fractionated by RP-HPLC as described above. Fraction 40 was dissolved in SDS sample buffer, and subjected to 15% SDS–PAGE. The gels were stained with Silver Stain MS kit (Wako Pure Chemical Industries, Osaka, Japan) and the band running at a similar size to recombinant NPGM-Gly was cut out. In-gel digestion was performed with trypsin at 35 °C overnight. The digested products were analyzed by MALDI-TOF-MS and ESI-MS/MS. In a different experiment, fraction 40 was digested with Asp-N at 37 °C overnight. The digested products were purified by RP-HPLC, and the collected peaks were analyzed by MALDI-TOF-MS.

3. Results

3.1. Expression of NPGM in E. coli

To obtain soluble recombinant NPGM at high purity, we cloned it into the multiple cloning site of the pCold TF DNA vector. This plasmid encodes the cold shock inducible *cspA* promoter, the translational enhancer element (TEE), which promotes the translation of the target gene, the poly histidine tag (His_6) for affinity purification of the expressed protein, the bacterial chaperone trigger factor (TF), which promotes the folding of the nascent peptide chains, and the cleavage site of Factor Xa protease for removal of His_6-TF tag from the target protein (Fig. 2A). A cDNA encoding the predicted mature NPGM with an extra Gly residue at the C-terminus was used for cloning; the Gly serves as an amide donor for the post-translational *in vitro* amidation, which is not naturally carried out by *E. coli*. *E. coli* transformed with this expression plasmid produces His_6-TF-NPGM-Gly, with a molecular mass of approximately 62 kDa (His_6-TF tag = 52 kDa, NPGM-Gly = 10 kDa). Cleavage of the His_6-TF tag by Factor Xa leaves a His-Met sequence, derived from the *Nde* I restriction site, at the N-terminus of NPGM-Gly (Fig. 2A).

Using the above expression vector, we transformed *E. coli* BL21 strain, which is most widely used, and the *E. coli* SHuffle strain, which is modified to assist folding of the expressed proteins and correct disulfide bonds formation. Most of the His_6-TF-NPGM-Gly precipitated into inclusion bodies when expressed in BL21 strain, whereas soluble protein was obtained using the SHuffle strain (Fig. 2B). After expression, His_6-TF-NPGM-Gly was purified on a Ni^{2+} column, and the His_6-TF tag was removed with Factor Xa. NPGM-Gly was then purified by RP-HPLC to high purity (Fig. 2C). ESI-MS indicated that the main HPLC peak contained multiple ions correspondent to the bridged NPGM-Gly species, whose calculated mass value is 10491.13 $[M+H]^+$ (Fig. 2D). The minor HPLC peak just after the main peak (Fig. 2C) corresponded to a degradation product of NPGM, which lacked the 5 C-terminal amino acids due to non-specific digestion by Factor Xa (data not shown). The yield of NPGM-Gly was 0.4 mg/L.

Specific antibodies against recombinant NPGM-Gly were raised in rabbit and the specificity was checked by dot-immunoblot analysis (Fig. 3).

3.2. Analysis of the disulfide bond pattern in recombinant NPGM-Gly

The location of the disulfide bond in recombinant NPGM-Gly was determined by alkylation of free Cys residue and protease digestion. The alkylation reaction was performed using iodoacetamide. Non-alkylated and alkylated NPGM-Gly were separately digested using protease Asp-N, which cleaves peptide chains at the N-terminal side of an aspartic acid residue. Each reaction solution was purified by RP-HPLC and the collected peaks were assigned by ESI-MS analysis (Fig. 4A and B). In the digestion of non-alkylated NPGM-Gly, "DCLLSRSHGMG" fragment was eluted at 22 min on RP-HPLC (Fig. 4A). In contrast, in the digestion of alkylated NPGM-Gly, the peak of "DCLLSRSHGMG" fragment disappeared and alkylated "DCLLSRSHGMG" fragment appeared at 21 min on RP-HPCL (Fig. 4A). Bridged "DLQCWNTCSLTLI" fragment was obtained in both experiments (Fig. 4A). These results showed that Cys^1 and Cys^2 were involved in disulfide bond formation and Cys^3 remained free. Alkylated "HMDLEFQKGVLAGISPGITA" fragment was also appeared in the digestion of alkylated NPGM-Gly (Fig. 4A). Some residue was presumably alkylated due to over-alkylation reaction.

3.3. C-terminal amidation of NPGM-Gly

The C-terminus of recombinant NPGM-Gly was amidated using an amidating enzyme. The reaction yielded three products, as it can be seen from the RP-HPLC profile reported in Fig. 5A. Further

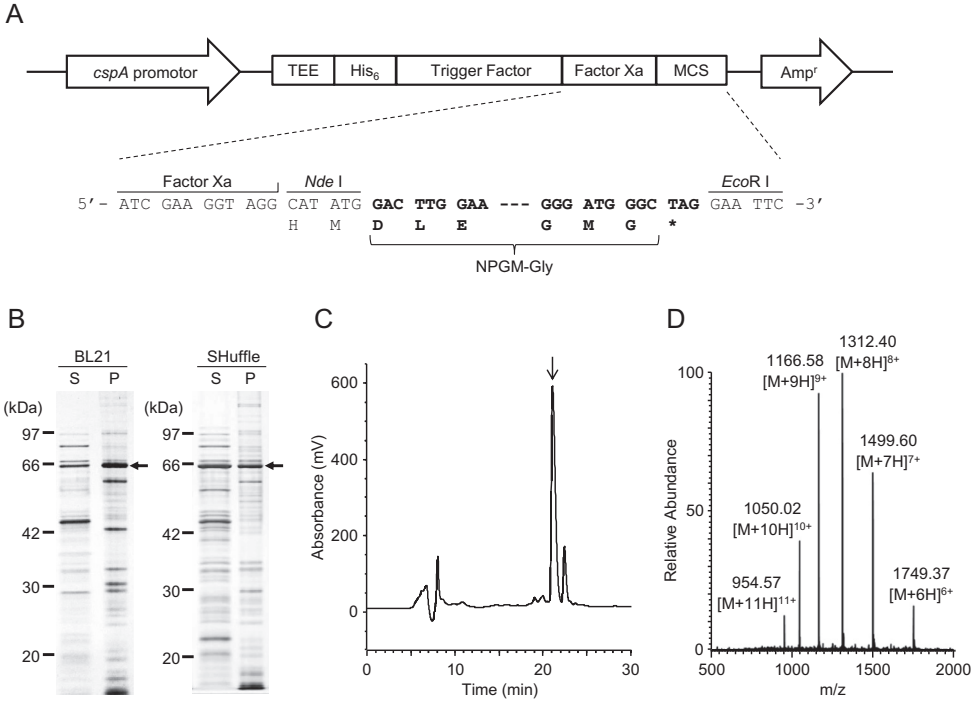

Fig. 2. Expression of NPGM-Gly in *E. coli*. (A) Construction of the expression plasmid vector. pCold TF DNA vector has a cold inducible *cspA* promoter, a translation-enhancing element (TEE), a hexa-histidine tag (His_6), a trigger factor acting as a solubility tag, a Factor Xa cleavage site, the multiple cloning site (MCS), and the ampicillin resistance gene (Amp^r). The cDNA encoding the predicted mature form of NPGM and the Gly amidation signal was inserted into the MCS. (B) Expression of His_6-TF tagged NPGM-Gly in soluble (S) and pellet (P) fractions was confirmed by SDS–PAGE and silver stain. Arrows indicate the target fusion protein. (C) Tag-cleaved NPGM-Gly was purified by RP-HPLC. (D) Purified NPGM-Gly was analyzed by ESI-MS.

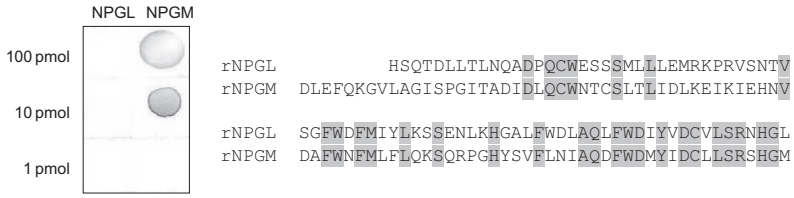

Fig. 3. Confirmation of the specificity of the anti-NPGM antibody by dot immunoblot assay. Sequence alignment of NPGL and NPGM is represented on the right side, and conserved amino acids are highlighted in gray.

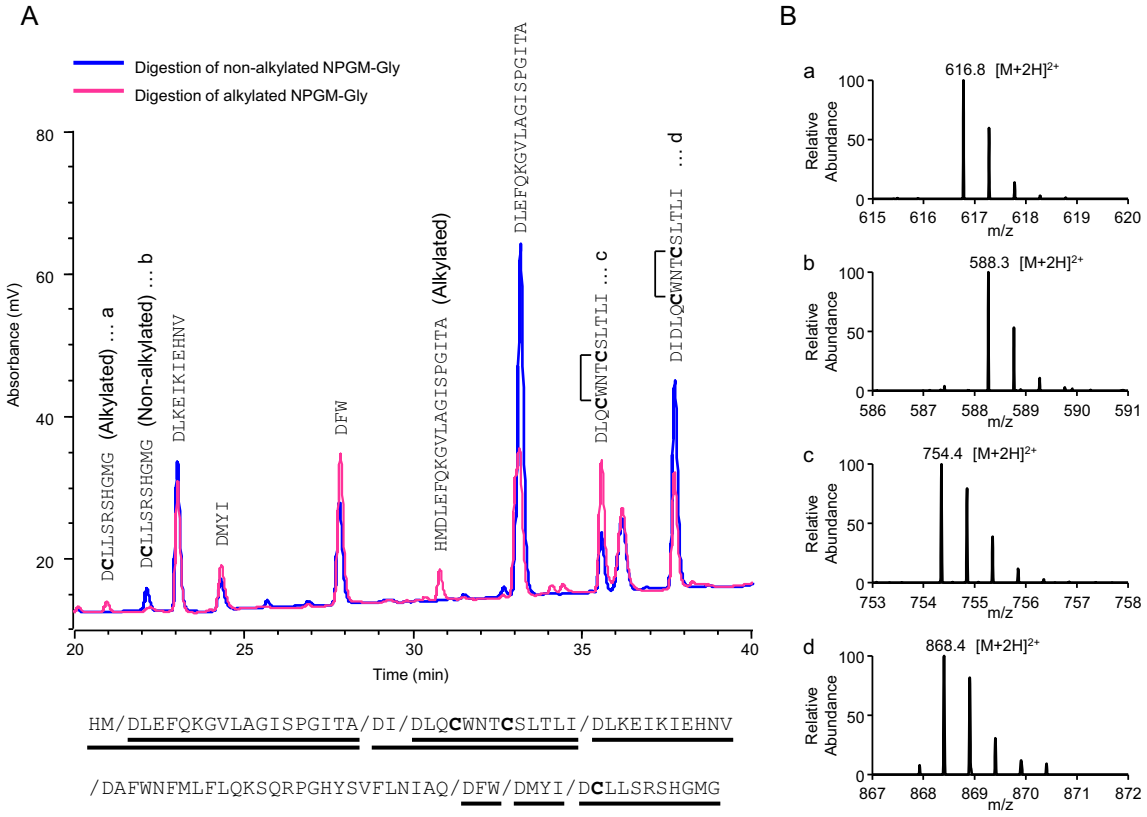

Fig. 4. Analysis of the disulfide bond location in NPGM-Gly expressed in *E. coli*. (A) Non-alkylated and alkylated NPGM-Gly were separately digested by Asp-N protease and purified by RP-HPLC. Each peptide fragment was assigned by ESI-MS analysis. (B) The mass spectra of the peptide fragments containing Cys residues; alkylated "DCLLSRSHGMG" (calcd. $[M+2H]^{2+}$ = 616.8), non-alkylated "DCLLSRSHGMG" (calcd. $[M+2H]^{2+}$ = 588.3), bridged "DLQCWNTCSLTLI" (calcd. $[M+2H]^{2+}$ = 754.4), and bridged "DIDLQCWNTCSLTLI" (calcd. $[M+2H]^{2+}$ = 868.4).

separation of the species on RP-HPLC was difficult to achieve. ESI-MS indicated that peaks 1 and 2 contained a species with mass value of 10433.16 $[M+H]^+$, corresponding to NPGM-NH_2, and byproducts with mass values of around 10415.15 $[M+H]^+$, corresponding to dehydrated NPGM-NH_2 (Fig. 5B). The byproduct from peak 3 had a mass value of 10465.10 $[M+H]^+$, which is the expected mass of NPGM-NH_2 oxidized in two positions, probably Met-oxidized product (Fig. 5B).

3.4. Characterization of NPGM secreted from CHO cells

We transfected the plasmid DNA encoding the NPGM precursor into CHO cells, and confirmed protein expression by immunocytochemistry (Fig. 6A). The cell culture medium was collected and fractionated every 2 min by RP-HPLC; two intense peaks eluted in correspondence of fractions 38 and 40 (Fig. 6B). These fractions were analyzed by western blotting using NPGM-Gly expressed in *E. coli* as a positive control (Fig. 6C). Fraction 40 contained a species that run at a similar size to the one of recombinant NPGM-Gly, while fraction 38 contained a larger species (Fig. 6C). MALDI-TOF-MS indicated that these species were the C-terminal 2∼6 or 24 amino acid residues-elongated forms of predicted mature NPGM, respectively.

Further analyses were performed on fraction 40. First, we run a sample from this fraction on an SDS gel and we digested the excised band with trypsin, which cleaves peptide chains at the C-terminal side of lysine or arginine. MALDI-TOF-MS indicated that the resulting proteolytic products had mass values of 779.39 and 2152.07 $[M+H]^+$, which corresponded to the values calculated for the "DLEFQK" and for the "IEHNVDAFWNFMLFLQK" sequence, respectively (Fig. 6F). The "DLEFQK" sequence was confirmed by ESI-MS/MS (Fig. 6D). Because of the substrate specificity of trypsin, (Fig. 6F), this result suggests that the "DLEFQK" region represents the N-terminus of mature NPGM. In a second experiment, fraction 40 was digested using Asp-N. The two resulting fragments were analyzed by MALDI-TOF-MS, and were identified as "DLEFQKGVL AGISPGITA" and as "DLQCWNTCSLTLI" (Fig. 6F). In the latter case, the mass value of 1529.9 $[M+Na]^+$ was corresponded to the calculated value for a bridged form of the fragment, and this value was increased of 2 mass units by reduction of the fragment with

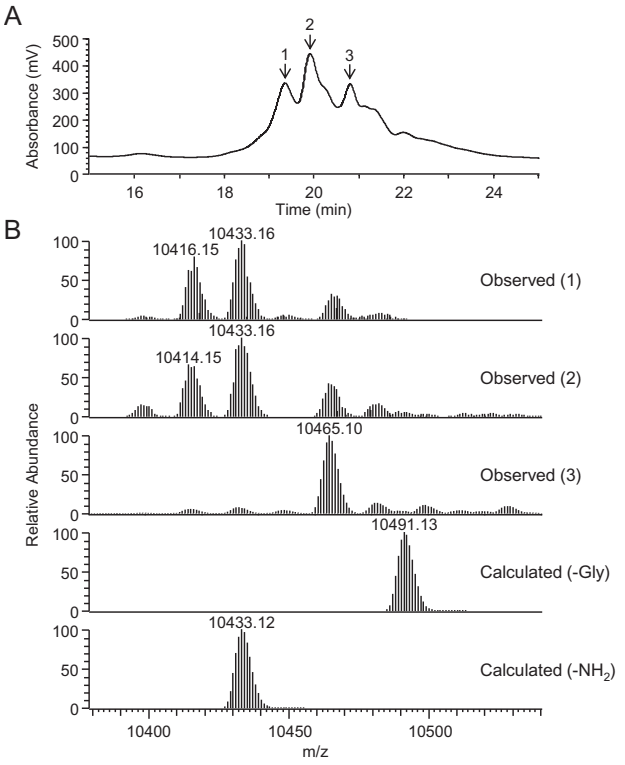

Fig. 5. C-terminal amidation of NPGM-Gly. (A) Amidated NPGM-Gly eluted in three different peaks when analyzed with RP-HPLC species. (B) The multiply charged ions observed from peak 1 to 3 were converted into $[M+H]^+$ ions, and compared with the calculated pattern of $[M+H]^+$ ion peaks of recombinant NPGM-Gly and NPGM-NH_2.

tris[2-carboxyethyl]phosphine hydrochloride (Fig. 6E). This result indicates that Cys^1 and Cys^2 in the fragment form a disulfide bond.

4. Discussion

While searching for novel neuropeptides and peptide hormone precursors in the chicken hypothalamus, we found a cDNA encoding a novel small secretory protein, NPGL [1]. We have previously showed that subcutaneous infusion of NPGL increased body weight gain in chicks [1]. In addition, we have also discovered the presence of a paralogous protein in vertebrates, NPGM. In mammals, NPGM contains three Cys residues and the exact structure of the endogenous mature form remains unclear, especially regarding the disulfide bond pattern. In the present study, we characterized mature NPGM from rat, using *E. coli* and CHO cells.

In a preliminary experiment, we expressed recombinant NPGM as a glutathione-S-transferase (GST) fusion protein, using a conventional *E. coli* expression system. However, we could only obtain insoluble protein that precipitated in inclusion bodies. In the present study, NPGM-Gly was efficiently expressed in soluble form by cold-shock induction in the *E. coli* SHuffle strain, using the cold shock expression vector pCold TF DNA. The vector contains the trigger factor (TF) fused to the target protein as a solubility tag; TF is a bacterial chaperone that interacts with nascent proteins and facilitates their folding [7,8]. The effectiveness of the TF-fusion approach has been reported in several cases [9,10]. At low induction temperatures, the expression of target proteins lying downstream of the *cspA* promoter is facilitated, whereas expression of most endogenous proteins is inhibited. This system usually leads to better results in terms of protein solubility, yield, conformational

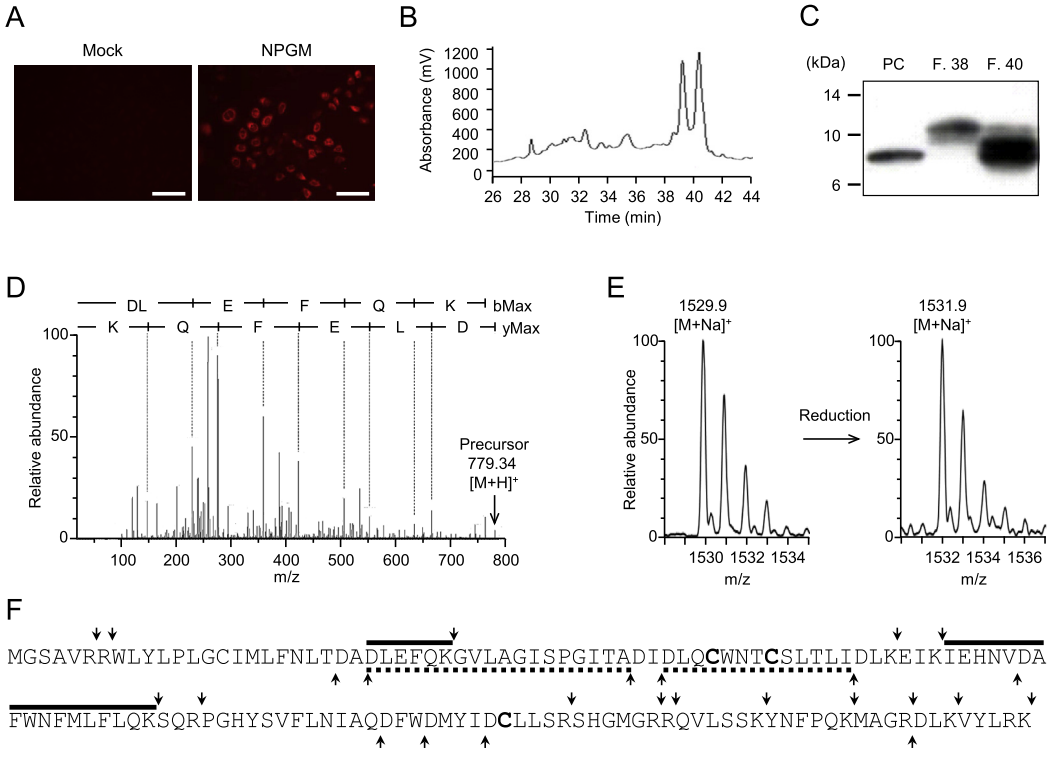

Fig. 6. Expression of NPGM in CHO cells and characterization of the secreted protein. (A) Confirmation of NPGM expression in CHO cells by immunocytochemistry. Scale bars = 100 μm. (B) The cell culture medium was fractionated by RP-HPLC. (C) Fractions 38 (F. 38) and 40 (F. 40) were analyzed by western blot using an antibody against NPGM. NPGM-Gly expressed in *E. coli* was used as positive control (PC). (D) Fraction 40 was in-gel digested with trypsin. One of the obtained fragments was sequenced by ESI-MS/MS. (E) Fraction 40 was digested with Asp-N protease. MALDI-TOF-MS indicated that one of the fragments had mass value of 1529.9 $[M+Na]^+$ corresponding to the bridged "DLQCWNTCSLTLI" sequence. This value was increased of 2 mass units by sample reduction. (F) Trypsin and Asp-N protease digestion sites are indicated by downward or upward arrows, respectively. The fragments obtained from each digestion are indicated by solid or dotted lines. The three Cys residues are highlighted in bold.

quality, and activity levels [11]. In our experiments, TF-fused NPGM-Gly was insoluble when expressed in *E. coli* BL21 strain, while good yields of soluble protein were obtained using *E. coli* SHuffle strain. This strain lacks cytoplasmic thioredoxin reductase (trxB) and glutaredoxin reductase (gor), and contains the cytoplasmic disulfide bond isomerase DsbC, which promotes native disulfide bonds formation and correct protein folding [12]. We characterized the pattern of the intramolecular disulfide bond in recombinant NPGM, as described later in the text.

The presence of a Gly–Arg–Arg sequence in the NPGM precursor (Fig. 1) suggests that the C-terminus is amidated. We performed amidation of NPGM-Gly *in vitro*, since the reaction does not occur in *E. coli*. After the amidating enzyme has been isolated from several animals [13,14], recombinant amidating enzyme has been used for amidation of various peptides such as mastoparan-X [15], human adrenomedullin [16], and crustacean hyperglycemic hormone [4,17,18]. In this study, amidation of the NPGM-Gly C-terminus resulted in unavoidable side reactions of dehydration and of oxidation, which prevented us from carrying out further investigations, such as physiological analyses. To improve the yield and purity of amidated NPGM, we need to use other methods. For example, we could use the transacylation reaction by carboxypeptidase Y [19,20] or the intein fusion protein method, which do not require a recombinant enzyme [21]. Moreover, we have recently succeeded in the synthesizing NPGL through microwave-assisted solid-phase peptide synthesis (SPPS), using the rink-amide resin [6]. NPGM containing a C-terminal amidation will be synthesized in a similar way in the near future. Currently, we are using microwave-assisted SPPS to prepare NPGM with good yield and purity.

We expected that NPGM was a secretory small protein because of the presence of a putative signal sequence at the N-terminal region of the protein precursor. In this study, we showed that NPGM is indeed a secretory protein using CHO cells. The N-terminus of the secreted NPGM is located just after the signal peptide, as expected. The C-terminus of the secreted NPGM was not amidated and contained an additional region of several amino acids. We believed that the C-terminus was uncleaved because of the nature of the secretory pathway in CHO cells. In fact, it has been reported that prorenin, which contains a Lys–Arg site, was also uncleaved in CHO cells, which have only the constitutive secretory pathway, while it was converted to active renin in AtT-20 cells, which have regulated secretory pathways [22]. Although we chose CHO cells for their manageability and productivity, it would be better to use endocrine cells such as AtT-20 cells, to demonstrate the presence of the mature form of a C-terminal amidated NPGM. Besides carrying out *in vitro* analyses, we would like to show the endogenous mature form in the rat hypothalamus. We are currently working to achieve this goal, using the specific antibodies against NPGM that we raised in this study.

Judging from the conservation of the Cys residues between NPGL and NPGM, we speculated that Cys^1 and Cys^3 of mammalian NPGM are also involved in the formation of a disulfide bond (Fig. 1). However, rat NPGM expressed in *E. coli* was bridged between Cys^1 and Cys^2. The location of the disulfide bond did not change when NPGM was reduced and bridged again using a glutathione or DMSO/HCl. NPGM secreted from mammalian CHO cells displayed the same bridge between Cys^1 and Cys^2. These results confirmed that Cys^1 and Cys^2 are involved in a disulfide bond while Cys^3 remains free. It has been shown that many proteins contain a free Cys residue. Human interleukin-2 has a free Cys residue that is not necessary for the interaction with the receptor [23]. Similarly the free Cys residues in herpes simplex virus type 1 glycoprotein D [24] and in human granulocyte colony stimulating factor [25], are not necessary for their activities and are located in inaccessible positions, in the transmembrane region and in the interior of the protein, respectively. It has been reported that elimination of a free Cys residue in human interferon-β and in Drosophila acetylcholinesterase improves their stability [26,27]. Furthermore, it is known that disulfide isomerization driven by a free Cys residue is required to activate phage endolysin [28,29]. As the function of NPGM is unknown at the present time, the contribution of the disulfide bond and of the free Cys residue to its activity are still obscure. It is possible that NPGM forms a dimer, although we think that this is unlikely. The dimerization product of NPGM has not been observed.

In conclusion, in this study we produced recombinant rat NPGM in *E. coli* and in mammalian cells and showed that it is a small secretory protein, with two Cys residues, Cys^1 and Cys^2, involved in the formation of a disulfide bond and a third free Cys. In the future, we will undertake studies to elucidate the physiological function of NPGM and to investigate the importance of the disulfide bond, of the free Cys residue, and of the C-terminal amidation for protein activity.

Author contributions

K.U. planned experiments. K.M., M.F., S.T., E.I.-U., and K.U. performed experiments and analyzed data. K.M. and K.U. wrote the paper.

Conflict of interest

None declared.

Acknowledgements

Mass spectrometry was outsourced to APRO life Science Institute (Tokushima, Japan) or performed at the Natural Science Center for Basic Research and Development (N-BARD), Hiroshima University. We thank Ms. Tomoko Amimoto (N-BARD), Hiroshima University. This work was supported by MEXT/JSPS KAKENHI grants (22687004 and 26291066 to K.U. and 25440171 to E.I.-U.), Grant-in-Aid for JSPS Fellows (14J05738 to K.M.), the Toray Science Foundation (K.U.), and the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry (K.U.).

References

[1] Ukena, K., Iwakoshi-Ukena, E., Taniuchi, S., Bessho, Y., Maejima, S., Masuda, K., Shikano, K., Kondo, K., Furumitsu, M. and Tachibana, T. (2014) Identification of a cDNA encoding a novel small secretory protein, neurosecretory protein GL, in the chicken hypothalamic infundibulum. Biochem. Biophys. Res. Commun. 446, 298–303.

[2] Zhang, L., Chou, C.P. and Moo-Young, M. (2011) Disulfide bond formation and its impact on the biological activity and stability of recombinant therapeutic proteins produced by *Escherichia coli* expression system. Biotechnol. Adv. 29, 923–929.

[3] Chang, S.G., Choi, K.D., Jang, S.H. and Shin, H.C. (2003) Role of disulfide bonds in the structure and activity of human insulin. Mol. Cells 16, 323–330.

[4] Nagai, C., Asazuma, H., Nagata, S., Ohira, T. and Nagasawa, H. (2009) A convenient method for preparation of biologically active recombinant CHH of the kuruma prawn, *Marsupenaeus japonicus*, using the bacterial expression system. Peptides 30, 507–517.

[5] Ukena, K., Tachibana, T., Iwakoshi-Ukena, E., Saito, Y., Minakata, H., Kawaguchi, R., Osugi, T., Tobari, Y., Leprince, J., Vaudry, H. and Tsutsui, K. (2010) Identification, localization, and function of a novel avian hypothalamic neuropeptide, 26RFa, and its cognate receptor, G protein-coupled receptor-103. Endocrinology 151, 2255–2264.

[6] Masuda, K., Ooyama, H., Shikano, K., Kondo, K., Furumitsu, M., Iwakoshi-Ukena, E. and Ukena, K. (2015) Microwave-assisted solid-phase peptide synthesis of neurosecretory protein GL composed of 80 amino acid residues. J. Pept. Sci. 21, 454–460.

[7] Hartl, F.U. and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295, 1852–1858.

[8] Maier, T., Ferbitz, L., Deuerling, E. and Ban, N. (2005) A cradle for new proteins: trigger factor at the ribosome. Curr. Opin. Struct. Biol. 15, 204–212.

[9] Nishihara, K., Kanemori, M., Yanagi, H. and Yura, T. (2000) Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. Appl. Environ. Microbiol. 66, 884–889.

[10] Bird, L.E. (2011) High throughput construction and small scale expression screening of multi-tag vectors in *Escherichia coli*. Methods 55, 29–37.
[11] Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front. Microbiol. 5, 172.
[12] Lobstein, J., Emrich, C.A., Jeans, C., Faulkner, M., Riggs, P. and Berkmen, M. (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb. Cell Fact. 11, 56.
[13] Mizuno, K., Sakata, J., Kojima, M., Kangawa, K. and Matsuo, H. (1986) Peptide C-terminal alpha-amidating enzyme purified to homogeneity from *Xenopus laevis* skin. Biochem. Biophys. Res. Commun. 137, 984–991.
[14] Murthy, A.S., Mains, R.E. and Eipper, B.A. (1986) Purification and characterization of peptidylglycine α-amidating monooxygenase from bovine neurointermediate pituitary. J. Biol. Chem. 261, 1815–1822.
[15] Kohno, T., Kusunoki, H., Sato, K. and Wakamatsu, K. (1998) A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: an application to the production of mastoparan-X uniformly enriched with ^{15}N and $^{15}N/^{13}C$. J. Biomol. NMR 12, 109–121.
[16] Mitsuda, Y., Takimoto, A., Kamitani, S., Kitamura, K., Sakata, T. and Mitsushima, K. (2002) Large-scale production of functional human adrenomedullin: expression, cleavage, amidation, and purification. Protein Expr. Purif. 25, 448–455.
[17] Katayama, H., Ohira, T., Aida, K. and Nagasawa, H. (2002) Significance of a carboxyl-terminal amide moiety in the folding and biological activity of crustacean hyperglycemic hormone. Peptides 23, 1537–1546.
[18] Mosco, A., Edomi, P., Guarnaccia, C., Lorenzon, S., Pongor, S., Ferrero, E.A. and Giulianini, P.G. (2008) Functional aspects of cHH C-terminal amidation in crayfish species. Regul. Pept. 147, 88–95.
[19] Breddam, K., Widmer, F. and Meldal, M. (1991) Amidation of growth hormone releasing factor (1–29) by serine carboxypeptidase catalysed transpeptidation. Int. J. Pept. Protein Res. 37, 153–160.
[20] Zhang, Z.Z., Yang, S.S., Dou, H., Mao, J.F. and Li, K.S. (2004) Expression, purification, and C-terminal amidation of recombinant human glucagon-like peptide-1. Protein Expr. Purif. 36, 292–299.
[21] Cottingham, I.R., Millar, A., Emslie, E., Colman, A., Schnieke, A.E. and McKee, C. (2001) A method for the amidation of recombinant peptides expressed as intein fusion proteins in *Escherichia coli*. Nat. Biotechnol. 19, 974–977.
[22] Hosaka, M., Nagahama, M., Kim, W.S., Watanabe, T., Hatsuzawa, K., Ikemizu, J., Murakami, K. and Nakayama, K. (1991) Arg-*X*-Lys/Arg–Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. J. Biol. Chem. 266, 12127–12130.
[23] Wang, A., Lu, S.D. and Mark, D.F. (1984) Site-specific mutagenesis of the human interleukin-2 gene: structure-function analysis of the cysteine residues. Science 224, 1431–1433.
[24] Wilcox, W.C., Long, D., Sodora, D.L., Eisenberg, R.J. and Cohen, G.H. (1988) The contribution of cysteine residues to antigenicity and extent of processing of herpes simplex virus type 1 glycoprotein D. J. Virol. 62, 1941–1947.
[25] Lu, H.S., Boone, T.C., Souza, L.M. and Lai, P.H. (1989) Disulfide and secondary structures of recombinant human granulocyte colony stimulating factor. Arch. Biochem. Biophys. 268, 81–92.
[26] Mark, D.F., Lu, S.D., Creasey, A.A., Yamamoto, R. and Lin, L.S. (1984) Site-specific mutagenesis of the human fibroblast interferon gene. Proc. Natl. Acad. Sci. U.S.A. 81, 5662–5666.
[27] Fremaux, I., Mazères, S., Brisson-Lougarre, A., Arnaud, M., Ladurantie, C. and Fournier, D. (2002) Improvement of Drosophila acetylcholinesterase stability by elimination of a free cysteine. BMC Biochem. 3, 21.
[28] Xu, M., Arulandu, A., Struck, D.K., Swanson, S., Sacchettini, J.C. and Young, R. (2005) Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. Science 307, 113–117.
[29] Kuty, G.F., Xu, M., Struck, D.K., Summer, E.J. and Young, R. (2010) Regulation of a phage endolysin by disulfide caging. J. Bacteriol. 192, 5682–5687.