Protein Expression and Purification 84 (2012) 38–46

Contents lists available at SciVerse ScienceDirect

Protein Expression and Purification

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

Electrostatic interaction-induced inclusion body formation of glucagon-like peptide-1 fused with ubiquitin and cationic tag

Sung-Gun Kim [a], Won-Ki Min [b], Yong-Taek Rho [a], Jin-Ho Seo [b,*]

[a] Department of Biomedical Science, Youngdong University, Chungbuk 370-701, Republic of Korea
[b] Department of Agricultural Biotechnology and Center for Agricultural Biomaterials, Seoul National University, Seoul 151-742, Republic of Korea

ARTICLE INFO

Article history:
Received 31 March 2012
and in revised form 14 April 2012
Available online 28 April 2012

Keywords:
Ubiquitin
6-lysine tag
Glucagon-like peptide-1
Inclusion body
Electrostatic interaction

ABSTRACT

In an attempt to produce glucagon-like peptide-1 (GLP-1) using recombinant *Escherichia coli*, ubiquitin (Ub) as a fusion partner was fused to GLP-1 with the 6-lysine tag (K6) for simple purification. Despite the high solubility of ubiquitin, the fusion protein K6UbGLP-1 was expressed mainly as insoluble inclusion bodies in *E. coli*. In order to elucidate this phenomenon, various N- and C-terminal truncates and GLP-1 mutants of K6UbGLP-1 were constructed and analyzed for their characteristics by various biochemical and biophysical methods. The experiment results obtained in this study clearly demonstrated that the insoluble aggregation of K6UbGLP-1 was attributed to the electrostatic interaction between the N-terminal 6-lysine tag and the C-terminal GLP-1 before the completion of folding which might be one of the reasons for protein misfolding frequently observed in many foreign proteins introduced with charged amino acid residues such as the His tag and the protease recognition sites. The application of a cation exchanger for neutralizing the positive charge of the 6-lysine tag in solid-phase refolding of K6UbGLP-1 successfully suppressed the electrostatic interaction-driven aggregation even at a high protein concentration, resulting in properly folded K6UbGLP-1 for GLP-1 production.

Introduction

Even though recombinant *Escherichia coli* has been used extensively for the production of heterologous proteins for the biotechnology industry, many proteins fail to fold into their native conformation which usually results in the accumulation of insoluble aggregates known as inclusion body [1]. A number of studies suggest that the inclusion body formation is attributed to improper intra- and/or inter-molecular interactions in folding intermediates of nascent polypeptides that are overexpressed in *E. coli* [2,3]. Other possible mechanisms for inclusion body formation have been proposed: the aggregation of native proteins of limited solubility and the aggregation of the unfolded state [4].

Fusion of an affinity tag such as His tag either at the N- or C-terminus of a target protein gene has been widely used for simple purification using affinity chromatography [5]. Several studies have shown that the fusion of an affinity tag results in positive effects such as increased protein stability and solubility, improved purification and crystallization [5,6]. In other cases, the introduction of an affinity tag or a protease cleavage site to a target protein negatively affected the function and structure of the product protein [6,7]. It has been frequently reported that fusion of the His tag leads to protein misfolding, resulting in inclusion body formation or loss of biological function of the target protein [8,9]. However, the exact mechanism by which the His tag fusion causes this misfolding is not clearly understood.

Ubiquitin is a highly conserved 76 amino acids protein in eukaryotes. It functions as a chaperone in ribosome biogenesis and as a marker for targeting proteins to proteasomes [10]. A number of studies have reported that the fusion of ubiquitin to the N-terminal of a target protein gene improved both the quality and quantity of the fusion proteins expressed in recombinant *E. coli* [11]. Two-dimensional NMR studies showed that the energetics of the ubiquitin folding pathway is highly favored and the protein completes a major folding phase in 10 ms [12]. Although the inner core of ubiquitin is highly hydrophobic, its outer surface is quite hydrophilic [13]. These unique properties of ubiquitin presumably promote proper folding of the fused protein, and prevent formation of inclusion body in recombinant *E. coli* [11]. In a general ubiquitin fusion system, six consecutive histidine residues known as the His tag were fused at the N-terminal of ubiquitin for purification *via* nickel affinity chromatography. In our previous study, for production of glucagon-like peptide-1 (GLP-1) that is a pharmaceutical peptide for treating the type 2 diabetes mellitus, ubiquitin along with the 6-lysine tag as a purification tag were fused with GLP-1 of 31 residues at its N-terminal, resulting in the fusion protein K6UbGLP-1 [14]. Despite the high solubility of ubiquitin and the

* Corresponding author. Fax: +82 2 873 5095.
E-mail address: jhseo94@snu.ac.kr (J.-H. Seo).

1046-5928/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.pep.2012.04.010

relatively low molecular weight of the attached protein (GLP-1), K6UbGLP-1 was found mainly in the insoluble fraction as inclusion bodies in *E. coli*.

In this study, various truncates and mutants of K6UbGLP-1 were constructed to decipher the reason for the inclusion body formation of K6UbGLP-1 in *E. coli*. Through various *in vivo* and *in vitro* analyses, the reason turned out to be the electrostatic interaction between the N-terminal 6-lysine tag and the C-terminal GLP-1. Furthermore, this study has shown that neutralization of the positive charge in the 6-lysine tag was able to suppress the electrostatic interaction-driven aggregation in a refolding process and thus allowed K6UbGLP-1 to fold into a proper conformation.

Materials and methods

Bacterial strains and culture conditions

Bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* DH5α is used for genetic manipulation. All plasmids constructed in this study were transformed into *E. coli* BL21 (DE3) for protein expression. The recombinant *E. coli* BL21(DE3) cells harboring each expression plasmid were cultured at 37 °C in a 500-ml baffled flask (Nalgene, NY, USA) containing Luria–Bertani medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) supplemented with 50 mg/L kanamycin. Expression of each gene was induced at the logarithmic growth phase (OD_{600} = 0.5–1.0) by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cell growth was monitored by measuring the optical density of the culture broth at 600 nm using a spectrophotometer (UV-1601, Shimadzu, Kyoto, Japan).

Construction of expression plasmids

The oligonucleotides used in this study are listed in Table 2. The 6-lysine tagged ubiquitin (K6Ub) gene was amplified from pAP-K6UbGLP-1 using the PCR primers of K6Ub-NdeI-F and Ub-BamHI-R, and was then inserted into a pAP vector in order to construct plasmid pAPK6Ub. To generate plasmid pAPUbGLP-1, the UbGLP-1 gene was amplified from pAPK6UbGLP-1 with the PCR primers of Ub-NdeI-F and GLP1-BamHI-R, and then cloned into the pAP vector. The schematic structures of K6UbGLP-1, K6Ub and UbGLP-1 are shown in Fig. 1A. The primers R6Ub-NdeI-F, H6Ub-NdeI-F, D6Ub-NdeI-F and S6Ub-NdeI-F were used to construct plasmids pAPR6UbGLP-1, pAPH6UbGLP-1, pAPD6UbGLP-1 and pAPS6UbGLP-1, respectively, with the same procedure as done for pAPUbGLP-1. The methods for synthesis of the GLP-1QN gene and for construction of pAPK6UbGLP-1QN were the same as those used for pAPK6UbGLP-1 [14] except for the two overlapping oligonucleotides of GLP-1QN-F and GLP-1QN-R. The amino acid sequences of GLP-1QN were shown in Fig 1B.

SDS–PAGE and Western blotting analysis

For sample preparation, protein concentrations were determined using the protein assay kit (Bio-Rad, Hercules, CA) according to the manufacturer's instructions. The samples were analyzed by SDS–PAGE with 14% or 16% T. Proteins were visualized by staining the gels with Coomassie brilliant blue R-250. The dried SDS–PAGE gels were analyzed by densitometry software (TotalLab 1.01, Nonlinear Dynamics Ltd., Newcastle, UK). The expression of K6UbGLP-1 was confirmed by Western blot analysis with a mouse monoclonal antibody against ubiquitin (Sigma, MO, USA) and a HRP-conjugated goat anti-mouse IgG secondary antibody (Bio-Rad, Hercules, CA). The bands of K6UbGLP-1 which specifically reacted with the antibodies were visualized by the Opti-4CN™ substrate and a detection kit (Bio-Rad, Hercules, CA).

In vivo solubility

A recombinant *E. coli* BL21(DE3) strain harboring each plasmid was grown at 37 °C and induced by addition of 1.0 mM IPTG for gene expression. After a 4-h induction, an appropriate volume of the culture broth was harvested by centrifugation at 13,000 rpm for 1 min, and the cell pellet was resuspended in 50 mM sodium phosphate buffer (pH 7.0), followed by cell disruption using an ultrasonic processor (Cole-Parmer, Niles, IL). The total cell lysate was separated into the soluble and the insoluble fractions by centrifugation at 14,000 rpm for 20 min, and then followed by SDS–PAGE analysis.

Electrostatic binding to cation exchanger

ÄKTA FPLC (GE Healthcare, Piscataway, NJ), an automated chromatography system, was used to analyze the electrostatic binding of K6UbGLP-1 and K6Ub on a cation exchanger. Two milliliters of the protein sample were injected into a HiTrap SP FF column (1 mL) packed with SP Sepharose Fast Flow (GE Healthcare, Piscataway, NJ) as a cation exchanger. The 50 mM sodium phosphate buffer (pH 7.0) was fed at a flow rate of 1 mL/min. After washing with 4 column volumes of the buffer, stepwise gradient of the buffer containing 1 M NaCl was applied to elute the proteins bound to the cation exchanger.

Table 1
Bacterial strains and plasmids used in this study.

Strain, plasmid	Relevant characteristics or construction	Source
Strains		
E. coli DH5α	F^- (Φ80Δ*lac*ZΔM15) *rel*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17(rK^-, mK+) *sup*E44 Δ(lacZYA-argF) U169	Invitrogen
E. coli BL21(DE3)	F^- *ompT hsd*SB(rB^-mB^-) *gal dcm* (DE3)	Novagen
Plasmids		
pUC18K6Ub	2.9 kb, Amp^r, 6-lysine tag and ubiquitin	AP technology
pAP	5.4 kb, Kan^r, *tac* promoter	AP technology
pAPK6UbGLP-1	5.8 kb, Kan^r, 6-lysine tagged ubiquitin fused to glucagon-like peptide-1, *tac* promoter	[14]
pAPK6UbGLP-1QN	5.8 kb, Kan^r, 6-lysine tagged ubiquitin fused to glucagon-like peptide-1QN, *tac* promoter	This study
pAPK6Ub	5.7 kb, Kan^r, 6-lysine tagged ubiquitin, *tac* promoter	This study
pAPUbGLP-1	5.8 kb, Kan^r, ubiquitin fused to glucagon-like peptide-1, *tac* promoter	This study
pAPR6UbGLP-1	5.8 kb, Kan^r, 6-arginine tagged ubiquitin fused to glucagons-like peptide-1, *tac* promoter	This study
pAPH6UbGLP-1	5.8 kb, Kan^r, 6-histidine tagged ubiquitin fused to glucagons-like peptide-1, *tac* promoter	This study
pAPD6UbGLP-1	5.8 kb, Kan^r, 6-aspartate tagged ubiquitin fused to glucagons-like peptide-1, *tac* promoter	This study
pAPS6UbGLP-1	5.8 kb, Kan^r, 6-serine tagged ubiquitin fused to glucagons-like peptide-1, *tac* promoter	This study

Table 2
Oligonucleotides used in this study.

Oligonucleotide	Nucleotide sequence
K6Ub-NdeI-F	GGAATTCCATATGAAGAAAAAAAAAGAAAAAGCAGATTTTCGTCAAGACT
Ub-NdeI-F	GAATTCCATATGCAGATTTTCGTCAAGACT
GLP-1-BamHI -R	CGCGGATCCTCATTAGCCGCGGCC
Ub-BamHI-R	CGCGGATCCTCATTAGCCACCTCTTAGCCTTAGCACAAGATG
GLP-1-F	CTTGTGCTAAGGCTAAGAGGTGGCCATGCGGAAGGCACCTTTACCAGCGATGTGAGCAGCTATCTGGAAGGCCAGGCG
GLP-1-R	CGCGGATCCTCATTAGCCGCGGCCTTTCACCAGCCACGCAATAAATTCTTTCGCCGCCTGGCCTTCCAGATAGCTGCT
GLP-1QN-F	CTTGTGCTAAGGCTAAGAGGTGGCCATGCGCAGGGCACCTTTACCAGCAACGTGAGCAGCTATCTGCAGGGCCAGGCG
GLP-1QN-R	CGCGGATCCTCATTAGCCACGGCCTTTCACCAGCCACGCAATAAACTGTTTCGCCGCCTGGCCCTGCAGATAGCTGCT
R6Ub-NdeI-F	GAATTCCATATGCGTCGTCGCCGTCGTCGCCAGATTTTCGTCAAGACT
H6Ub-NdeI-F	GAATTCCATATGCATCATCATCATCATCATCAGATTTTCGTCAAGACT
D6Ub-NdeI-F	GAATTCCATATGGATGATGACGATGATGACCAGATTTTCGTCAAGACT
S6Ub-NdeI-F	GAATTCCATATGAGCAGCTCTAGCAGCTCTCAGATTTTCGTCAAGACT

The underlined sequences correspond to the specific recognition site of the restriction enzymes as described in Materials and methods Section.

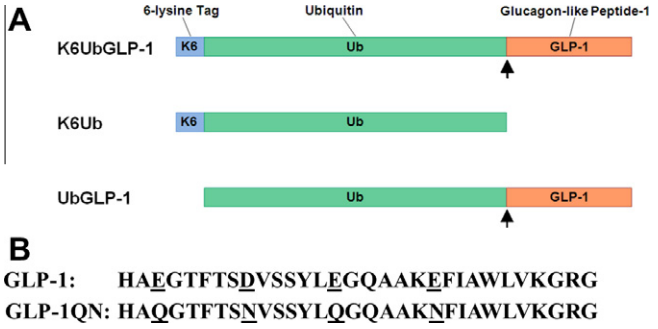

Fig. 1. Schematic structures of K6UbGLP-1, K6Ub and UbGLP-1, and protein sequences of GLP-1 and GLP-1QN. (A) Schematic structure of K6UbGLP-1, K6Ub and UbGLP-1. The lengths of the boxes are drawn to reflect the amino acid length. K6, Ub and GLP-1 are abbreviated as the 6-lysine tag, ubiquitin and glucagon-like peptide-1, respectively. The arrow indicates the cleavage site by deubiquitinating enzymes. (B) Protein sequences of GLP-1 and GLP-1QN. Glu (E) and Asp (D) (underlined) in GLP-1 are substituted with Gln (Q) and Asn (N) in GLP-1QN, respectively, in order to eliminate their negative charge.

Solid-phase refolding

The solid-phase refolding for K6UbGLP-1 was carried out in a batch mode. The cells were harvested by centrifugation at 6000 rpm and 4 °C for 10 min after a batch culture of the recombinant *E. coli* strain harboring plasmid pAPK6UbGLP-1. The cell pellets (0.97 g of dry cell weight) were then resuspended in 50 ml of buffer A (50 mM sodium phosphate buffer, pH 7.0). After cell disruption by an ultrasonic processor (Cole-Parmer, IL, USA), the insoluble pellet was prepared by centrifugation at 12,000 rpm and 4 °C for 10 min followed by washing with 20 mL of buffer B (50 mM sodium phosphate buffer, pH 7.0, 2% Triton X-100, 20 mM EDTA, 0.5 M NaCl). The washed insoluble pellets were defined as K6UbGLP-1 inclusion body. It was denatured in 20 mL of buffer C (buffer A containing 8 M urea) for more than 2 h. Insoluble cell debris mixed with the denatured K6UbGLP-1 was removed by centrifugation at 12,000 rpm and 4 °C for 20 min. The denatured K6UbGLP-1 was added to 0.5 mL of SP Sepharose Fast Flow medium (GE Healthcare, Piscataway, NJ) as a cation exchanger in a 2.0 ml of microtube that had been pre-equilibrated with buffer C. The K6UbGLP-1 immobilized on the cation exchanger was refolded by exchanging buffer C with A. The refolded K6UbGLP-1 was eluted from the cation exchanger by adding 0.5 mL of buffer D (buffer A containing 2 M NaCl).

Cleavability by deubiquitinating enzymes

For the cleavage test with deubiquitinating enzymes, both ubiquitin-specific protease 1 (UBP1) and yeast ubiquitin C-terminal hydrolase 1 (YUH1) used in this study were produced in recombinant *E. coli* cells [15,16]. Since both UBP1 and YUH1 have the 6-histidine tag at their N-termini, they were purified *via* nickel affinity chromatography using Ni-nitrilotriacetate resin (Qiagen, Studio City, CA). The purified deubiquitinating enzymes were added to be 0.1–0.3 mg/mL in the samples of K6UbGLP-1 and K6UbGLP-1QN. After the cleavage reactions at 37 °C for 12 h, the samples were analyzed on SDS–PAGE and their cleavages were quantified by the densitometry program.

Gel filtration chromatography

Analytical gel filtration chromatography was performed on an ÄKTA FPLC system using a Superdex 75 HR 10/30 column (GE Healthcare, Piscataway, NJ). The refolded K6UbGLP-1 was loaded on the column that was equilibrated with 50 mM sodium phosphate buffer (pH 7.0) containing 150 mM NaCl and separated at a flow rate of 0.5 ml/min at ambient temperature. Marker proteins for column calibration were purchased from Sigma (bovine serum albumin, 67 kDa; ovalbumin, 43 kDa; carbonic anhydrase, 29 kDa; cytochrome c, 12.4 kDa).

Circular dichroism analysis

Far-UV circular dichroism (CD) measurements were carried out on an automated AVIV model 400 CD spectrometer (Aviv Biomedical, Lakewood, NJ). The spectra were measured over a wavelength range of 260–190 nm, using a 1-mm pathlength. The protein was desalted using a HiTrap™ desalting column equilibrated with 50 mM sodium phosphate buffer (pH 7.0), and then adjusted to be 0.2 mg/mL protein concentration for the CD analysis. The temperature was maintained at 25 °C. The background CD spectrum of the buffer was subtracted from that of the protein sample to obtain the spectrum of K6UbGLP-1.

Results

Misfolding of K6UbGLP-1 expressed in E. coli

K6UbGLP-1 was expressed mainly as insoluble inclusion bodies in recombinant *E. coli*, but approximately 20% of K6UbGLP-1 was found in the soluble fraction (Fig. 2A). This result was reproduced

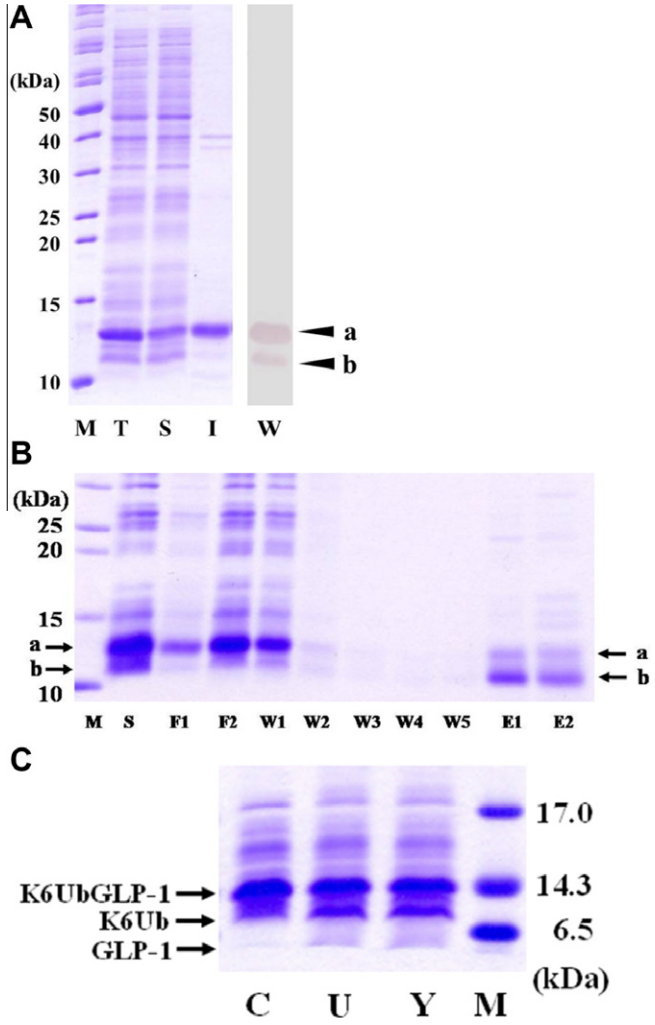

Fig. 2. *In vivo* solubility, electrostatic binding to cation exchanger, and cleavability by deubiquitinating enzymes for K6UbGLP-1. (A) *In vivo* solubility. After a 4-h induction, the recombinant *E. coli* cells expressing K6UbGLP-1 were harvested and disrupted, and their lysates were fractionated into total cell lysate (T), soluble fraction (S) and insoluble fraction (I). Lane M denotes a protein size marker. The arrows 'a' and 'b' indicate the positions of the intact K6UbGLP-1 and the truncated K6UbGLP-1, respectively. The lane 'W' represents the Western blotting analysis using the mouse monoclonal antibody against ubiquitin for the soluble fraction. (B) Electrostatic binding to cation exchanger. The positions of the intact K6UbGLP-1 and the truncated K6UbGLP-1 are indicated by the arrows 'a' and 'b', respectively. Lanes: M, marker proteins; S, sample loaded, F1 and F2, flow-through fractions; W1–W5, wash fractions; E1 and E2, elution fractions. (C) Cleavability by deubiquitinating enzymes. The soluble fraction was treated with UBP1 (U) and YUH1 (Y) at 37 °C for 12 h. The lane 'C' denotes the soluble fraction without treatment of the deubiquitinating enzyme as a control. K6Ub and GLP-1 indicated by the arrows represent the cleavage products of K6UbGLP-1 by UBP1 or YUH1.

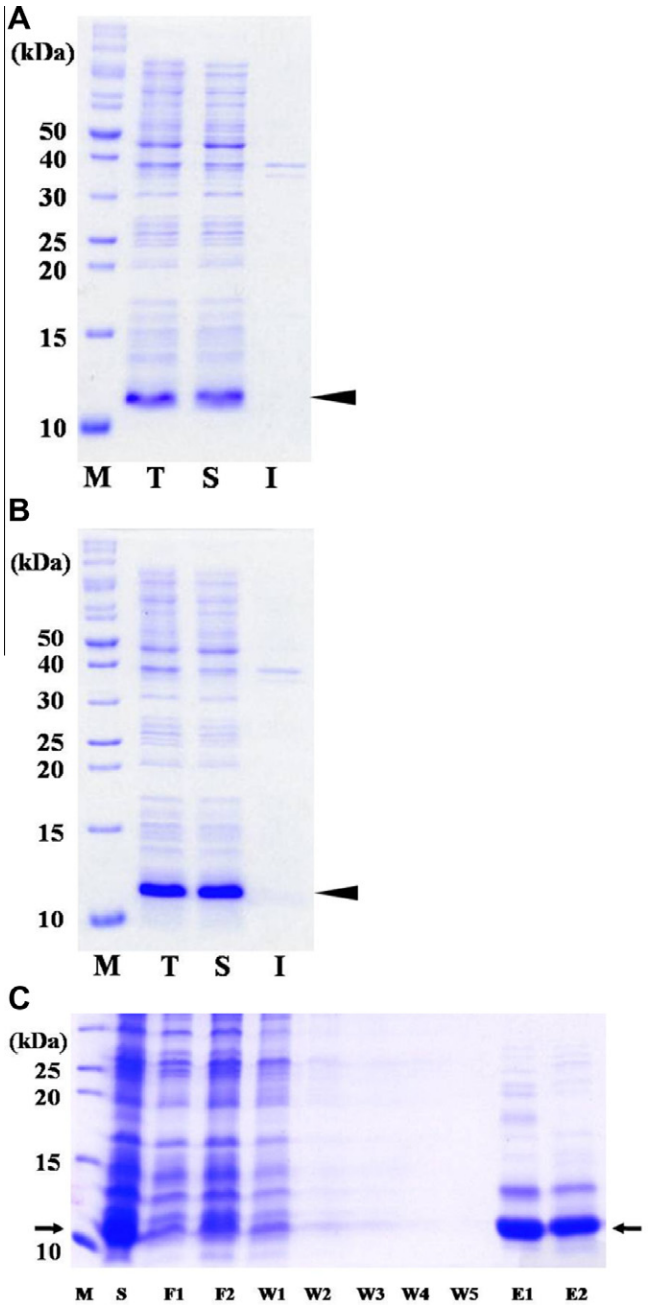

Fig. 3. *In vivo* solubility and electrostatic binding to cation exchanger for K6Ub and UbGLP-1. (A) and (B) *In vivo* solubility for K6Ub and UbGLP-1. After a 4-h induction, the recombinant *E. coli* cells expressing K6Ub or UbGLP-1 were harvested and disrupted, and their lysates were fractionated into total cell lysate (T), soluble fraction (S) and insoluble fraction (I). Lane M denotes a protein size marker. The arrow heads indicate the positions of K6Ub and UbGLP-1. (C) Electrostatic binding to cation exchanger for K6Ub. The position of K6Ub is indicated by the arrow. Lanes: M, marker proteins; S, sample loaded, F1 and F2, flow-through fractions; W1–W5, wash fractions; E1 and E2, elution fractions.

in accordance with our previous result [14]. Besides the major band of K6UbGLP-1 (Fig. 2A, arrow '*a*'), of which the apparent molecular weight was consistent with the theoretical molecular weight (12.7 kDa), a minor band that is 2–3 kDa smaller than the major band was observed in the soluble fraction (Fig. 2A, arrow '*b*'). In order to validate whether the protein of the minor band is derived from the intact K6UbGLP-1, the soluble fraction was subjected to a Western blotting analysis using a mouse monoclonal antibody against ubiquitin. The minor band as well as the major band of K6UbGLP-1 was visualized, indicating that the minor band is a truncated protein of the intact K6UbGLP-1 (Fig. 2A, lane '*W*').

As the 6-lysine tag in K6UbGLP-1 is designed to bind to a cation exchanger *via* electrostatic interaction, cation exchange chromatography was attempted to purify K6UbGLP-1 in the soluble fraction. As shown in Fig. 2B, most of the intact K6UbGLP-1 protein was observed not in the elution fractions but in the flow-through fractions, indicating that it was not bound to the cation exchanger. This result suggests that the 6-lysine tag of the soluble K6UbGLP-1 was not displayed binding to the cation exchanger and was likely buried inside the soluble K6UbGLP-1, or its positive charge was neutralized by intra- or inter-molecular interactions. On the other hand, the truncated protein (Fig. 2A, arrow '*b*') of K6UbGLP-1 in the soluble fraction

was mostly bound to the cation exchanger and then eluted by increasing NaCl concentration which indicates electrostatic binding of the N-terminal 6-lysine tag at the truncated protein to the cation exchanger (Fig. 3B, *arrow 'b'*). These results provide evidence that the truncated region in K6UbGLP-1 is located at its C-terminus, since the electrostatic binding of the truncated K6UbGLP-1 to the cation exchanger is ascribed to the presence of the 6-lysine tag at its N-terminus.

Interaction between 6-lysine tag and GLP-1 causes insoluble aggregation

Since the truncated K6UbGLP-1 missing the C-terminal 2–3 kDa region was not only expressed mainly as soluble form, but also electrostatically bound to the cation exchanger, it could be hypothesized that the C-terminal 2–3 kDa region caused the misfolding of K6UbGLP-1 as observed in *E. coli* (Fig. 2A) and cation exchange chromatography (Fig. 2B). In order to verify this hypothesis, the gene of the C-terminal GLP-1 was removed from the gene coding for K6UbGLP-1. As expected, the GLP-1 truncated form (K6Ub) was completely soluble when expressed in *E. coli* (Fig. 3A) which is contrary to the result of K6UbGLP-1. To see the electrostatic binding of the 6-lysine tag of K6Ub to the cation exchanger, cation exchange chromatography for K6Ub was performed in the same manner of K6UbGLP-1. In contrast to the result of the soluble K6UbGLP-1 (Fig. 2C), most of the K6Ub protein was bound to the cation exchanger and eluted at high NaCl concentrations (Fig. 3C), for which its purification profile was nearly the same as that of the truncated protein of K6UbGLP-1 (Fig. 2B). Therefore, these results clearly indicate that the truncation of GLP-1 in K6UbGLP-1 allows the protein to fold into a proper conformation displaying the 6-lysine tag, and subsequently to bind to the cation exchanger.

Considering the results for the soluble expression and electrostatic binding of K6Ub to the cation exchanger, misfolding of K6UbGLP-1 might be caused by improper interactions between the C-terminal GLP-1 region and the 6-lysine tag or ubiquitin during the protein folding process. To assess this assumption, either the gene of ubiquitin or the 6-lysine tag was deleted in the gene of K6UbGLP-1, resulting in the 6-lysine tagged GLP-1 (K6GLP-1) or the ubiquitin-GLP-1 (UbGLP-1). When the ubiquitin truncated form (K6GLP-1) was expressed in *E. coli*, the protein band of K6GLP-1 was not detectable on SDS–PAGE analysis, indicating that the small K6GLP-1 peptide of 5 kDa is presumably degraded in *E. coli* (data not shown). Meanwhile, the SDS–PAGE analysis of *E. coli* cells expressing the 6-lysine tag truncated form (UbGLP-1) shows that UbGLP-1 folded into a soluble conformation in the cell (Fig. 3B). Consequently, the soluble expressions of both K6Ub and UbGLP-1 suggest that the molecular interaction between the 6-lysine tag and GLP-1 caused the misfolding of ubiquitin, resulting in the insoluble aggregation of K6UbGLP-1 as inclusion bodies in recombinant *E. coli*.

Insoluble aggregation by electrostatic interaction between 6-lysine tag and GLP-1

The molecular interaction of GLP-1 with the 6-lysine tag carrying a positive charge caused the insoluble aggregation of K6UbGLP-1 in the cell so that the positive charge of the 6-lysine tag is expected to be involved in the insoluble aggregation *via* electrostatic interaction with GLP-1. To verify this hypothesis, the 6-lysine tag in K6UbGLP-1 was replaced with cationic (6-arginine and 6-histidine), anionic (6-aspartate) and nonionic (6-serine) amino acid tags and then their *in vivo* solubilities were analyzed by the SDS–PAGE (Fig. 4). Consistent with the result of K6UbGLP-1, the 6-arginine tagged UbGLP-1 (R6UbGLP-1) was found mainly

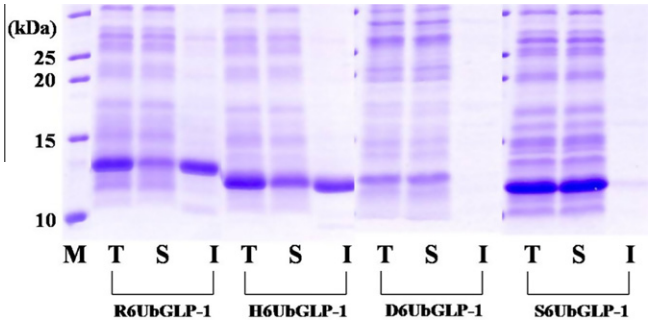

Fig. 4. *In vivo* solubility for R6UbGLP-1, H6UbGLP-1, D6UbGLP-1 and S6UbGLP-1. After a 4-h induction, the recombinant *E. coli* cells expressing each protein were harvested and disrupted, and their lysates were fractionated into total cell lysate (T), soluble fraction (S) and insoluble fraction (I). Lane M denotes a protein size marker.

in the insoluble fractions. The fusion of the 6-histidine tag, which has been widely used in protein purification *via* nickel affinity chromatography, to UbGLP-1 (H6UbGLP-1) also led to the insoluble aggregation in accordance with the results of both K6UbGLP-1 and R6UbGLP-1. Despite the fact that the histidine residue is 10% protonated at pH 7, the positive charge of the six-consecutive histidine residues of H6UbGLP-1 was enough to cause the insoluble aggregation *via* the electrostatic interaction with GLP-1. On the contrary, both the anionic 6-aspartate tagged UbGLP-1 (D6UbGLP-1) and the nonionic 6-serine tagged UbGLP-1 (S6UbGLP-1) were completely soluble in *E. coli*. These results again support that the positive charge of the 6-lysine tag is involved in the electrostatic interaction with GLP-1, which leads to the misfolding of K6UbGLP-1.

Since the estimated pI value of GLP-1 is 5.53, the 6-lysine tag is able to interact electrostatically with GLP-1 at the intracellular pH (~7.0) of *E. coli*. This acidic pI value of GLP-1 is attributed to the presence of three glutamates and one aspartate residues in GLP-1. The four anionic residues in GLP-1 may be involved in the electrostatic interaction with the 6-lysine tag. In order to prove this possibility, the negative charge of GLP-1 was eliminated by substituting glutamate with glutamine and aspartate with asparagine, resulting in GLP-1QN (Fig. 1B). The GLP-1QN fused to the 6-lysine tagged ubiquitin (K6UbGLP-1QN) was expressed in *E. coli* in the same manner of K6UbGLP-1. It was found mostly in the soluble fraction (Fig. 5A), which is contrary to the result of K6UbGLP-1. This result clearly shows that the elimination of the negative charge in GLP-1 allowed the protein to fold into a soluble form. Taken together with the above results, it is obvious that the electrostatic interaction between the 6-lysine tag and GLP-1 caused the insoluble aggregation of K6UbGLP-1 in *E. coli*.

Using cation exchanger to neutralize positive charge of 6-lysine tag in refolding of K6UbGLP-1

In order to verify whether the electrostatic interaction-driven aggregation of K6UbGLP-1 observed in *E. coli* is reproducible in the *in vitro* condition, protein refolding using dialysis was attempted at various protein concentrations (Fig. 6A). When the refolded protein sample was fractionated into the soluble and insoluble fractions after refolding, any soluble protein was not detectable but it was found to form insoluble aggregates even at a low protein concentration of 0.25 mg/mL which is in consistent with the *in vivo* result. However, when NaCl was supplemented in the refolding buffer, K6UbGLP-1 was found in the soluble fraction (Fig. 6B). The amount of K6UbGLP-1 in the soluble fraction was increased as a function of NaCl concentrations. Particularly,

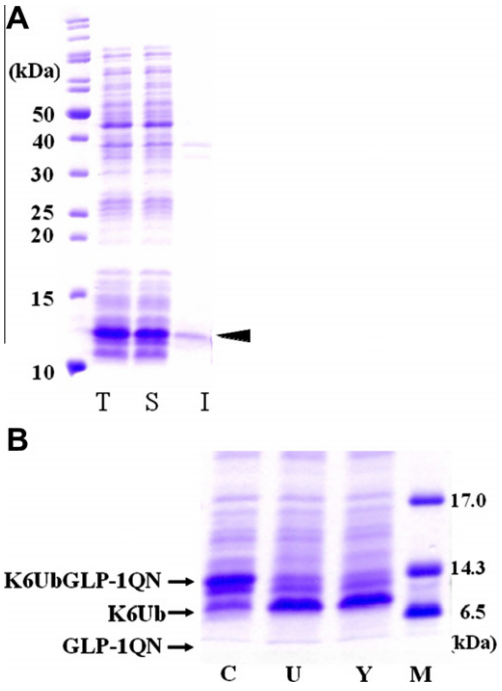

Fig. 5. *In vivo* solubility and cleavability by deubiquitinating enzymes for K6UbGLP-1QN. (A) *In vivo* solubility. After a 4-h induction, the recombinant *E. coli* cells expressing K6UbGLP-1QN were harvested and disrupted, and then the lysate was fractionated into total cell lysate (T), soluble fraction (S) and insoluble fraction (I). Lane M denotes a protein size marker. The arrow head indicates the position of K6UbGLP-1QN. (B) Cleavability by deubiquitinating enzymes. The soluble fraction was treated with UBP1 (U) and YUH1 (Y) at 37 °C for 12 h. The lane 'C' denotes the soluble fraction without treatment of the deubiquitinating enzyme as a control. K6Ub and GLP-1QN indicated by the arrows represent the cleavage products of K6UbGLP-1 by UBP1 or YUH1.

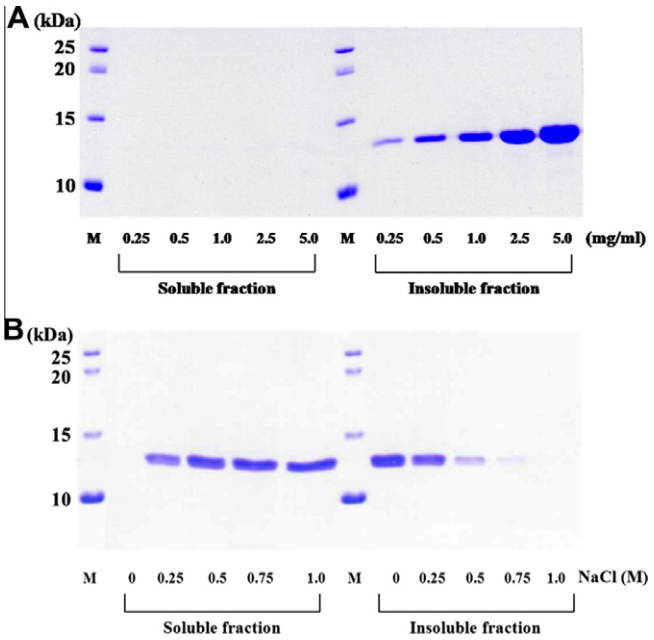

Fig. 6. *In vitro* refolding for inclusion body of K6UbGLP-1. (A) *In vitro* refolding in the absence of NaCl. The inclusion body of K6UbGLP-1 was solubilized in 50 mM sodium phosphate buffer (pH 7.0) containing 8 M urea. The unfolded K6UbGLP-1 at different concentrations (0.25–5.0 mg/ml) in the buffer containing 8 M urea was dialyzed against the refolding buffer (50 mM sodium phosphate buffer, pH 7.0) at 4 °C for overnight. (B) *In vitro* refolding in the presence of NaCl. The inclusion body of K6UbGLP-1 was solubilized in 50 mM sodium phosphate buffer (pH 7.0) containing 8 M urea. The unfolded K6UbGLP-1 of 5.83 mg/ml in the buffer containing 8 M urea was diluted 10-fold into the refolding buffers (50 mM sodium phosphate buffer, pH 7.0) of various NaCl concentrations (0, 0.25, 0.5, 0.75, 1.0 M) at 4 °C. The refolded protein samples were separated into the soluble and the insoluble fractions by centrifuging at 14,000 rpm for 20 min, and subjected to the SDS–PAGE analysis.

the supplementation of 1 M NaCl in the refolding buffer was able to prevent dramatically the insoluble aggregation of K6UbGLP-1, suggesting that the neutralization of the positive charge in the 6-lysine tag by the salt supplementation abolishes the electrostatic-driven aggregation of K6UbGLP-1.

In our previous study [14], K6UbGLP-1 was successfully refolded with a yield above 90% even at 5.2 mg/mL *via* solid-phase refolding using a cation exchanger. Furthermore, the refolded K6UbGLP-1 was efficiently cleaved to be K6Ub and GLP-1 by a deubiquitinating enzyme (UBP1) indicating that it adopted a proper conformation for GLP-1 production. Based on the results in this study, this was clearly due to the neutralization of the positive charge in the 6-lysine tag being electrostatically bound to the cation exchanger. In order to examine whether the application of the cation exchanger to neutralize the positive charge of the 6-lysine tag rendered K6UbGLP-1 folded into a correct conformation, the inclusion body of K6UbGLP-1 was refolded in the same manner as the previous study and then subjected to several biophysical and biochemical analyses. The unfolded K6UbGLP-1 inclusion bodies in 8 M urea were bound to the cation exchanger at a concentration of 3.43 mg protein/ml resin in which the positive charge of the 6-lysine tag was neutralized by anionic functional groups of the cation exchanger upon electrostatic immobilization to it. The following removal of urea allowed the unfolded K6UbGLP-1 to refold. The refolded K6UbGLP-1 was eluted from the cation exchanger by increasing NaCl concentration. The recovery yield of the soluble K6UbGLP-1 after the elution was determined to be 92%. However, the elution buffer contained 1 M NaCl which might prevent the electrostatic interaction-driven aggregation of K6UbGLP-1 after elution. To exclude this possibility, the refolded K6UbGLP-1 was desalted by using a chromatographic desalting column. The protein recovery yield after the desalting was close to 100% without any detectable protein aggregate in the desalted sample fractions (data not shown). Furthermore, when the desalted protein was subjected to a cation exchange chromatography, it was able to bind to the cation exchanger without any detectable protein in the flow-through fractions, and then eluted by increasing NaCl concentration (Fig. 7A) indicating an electrostatic binding of the 6-lysine tag to the cation exchanger. This purification profile was the same as those of K6Ub (Fig. 3C) and the truncated K6UbGLP-1 (Fig. 2B, arrow '*b*').

In analytical gel filtration chromatography, the refolded K6UbGLP-1 showed a single peak with an apparent molecular mass of 13.5 kDa (Fig. 7B) which is consistent with its theoretical molecular weight (12.7 kDa). There was no detectable soluble aggregate that is generally eluted at void volume indicating that K6UbGLP-1 was homogeneously folded as a monomer. The far-UV CD spectrum of the refolded K6UbGLP-1 exhibited distinct double minima at around 208 and 222 nm (Fig. 7C) which are indicative of a protein with high content of α-helix, and is similar to that of the 6-histidine tagged ubiquitin [17]. In thermal unfolding analysis using the CD spectrometry at 222 nm, a significant change in the CD signal was not observed in a range of 25–90 °C (data not shown), indicating that the refolded K6UbGLP-1 also had a highly stable structure as observed in ubiquitin by itself [18].

Cleavability by deubiquitinating enzymes

The accurate cleavage of the peptide bond between ubiquitin and GLP-1 is required in order to obtain the authentic GLP-1 from

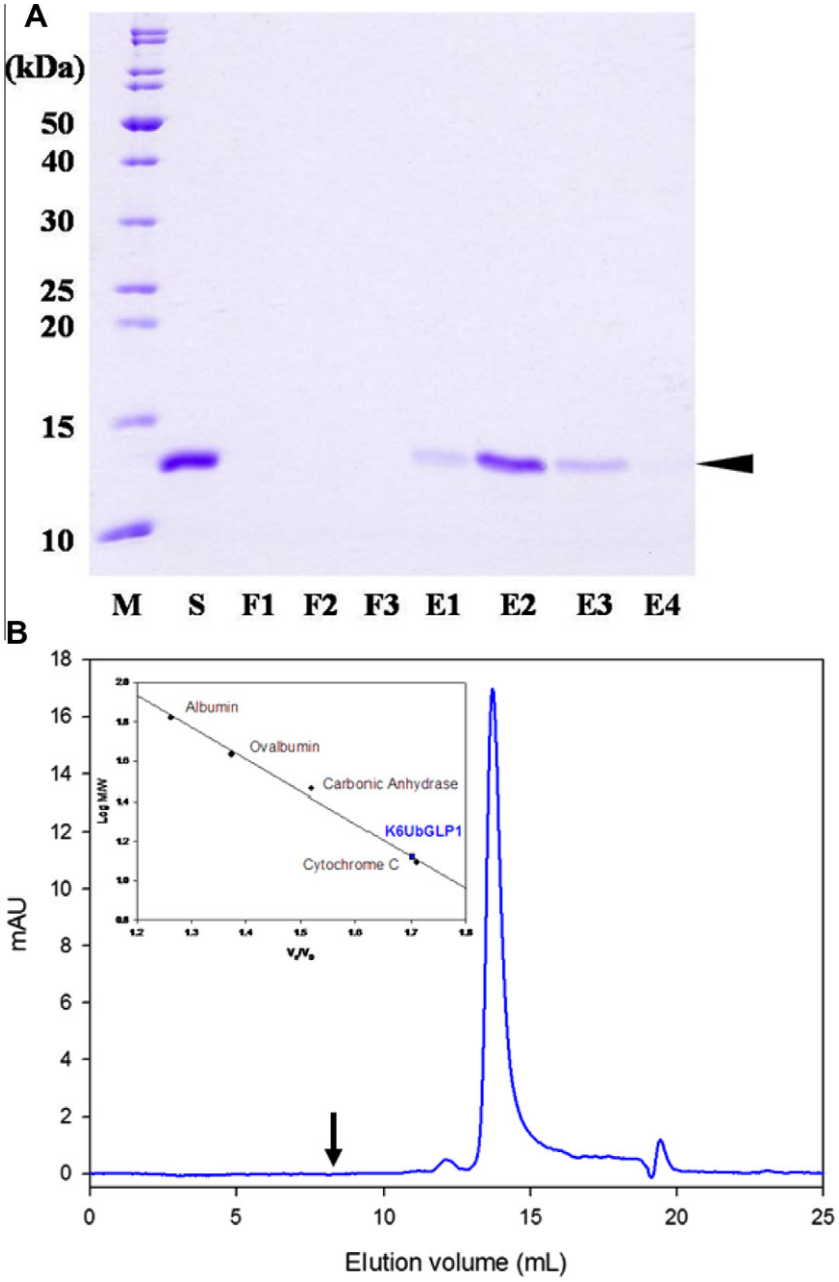

Fig. 7. Analyses for K6UbGLP-1 refolded by solid-phase refolding using cation exchanger. (A) Electrostatic binding to cation exchanger. The position of K6UbGLP-1 is indicated by the arrow. Lanes: M, marker proteins; S, sample loaded, F1–F3, flow-through fractions; E1–E4, elution fractions. (B) Analytical gel filtration chromatography. The apparent molecular weight of K6UbGLP-1 was estimated by gel filtration chromatography with molecular standards (inset). The column volume of the gel filtration column (Superdex 75 HR 10/30) was 24 mL. The void volume (V_0) determined with blue dextran was 8.06 (indicated by the arrow). (C) Far-UV circular dichroism spectrum. The CD signal was converted to the mean residue ellipticity. The mean residue ellipticity is plotted *versus* the wavelength. (D) Cleavability by deubiquitinating enzymes. The refolded K6UbGLP-1 was treated by UBP1 (U) and YUH1 (Y) at 37 °C for 12 h. The lane 'C' denotes the refolded K6UbGLP-1 without treatment of the deubiquitinating enzyme as a control. K6Ub and GLP-1 indicated by the arrows represent the cleavage products of K6UbGLP-1 by UBP1 or YUH1.

K6UbGLP-1. Since deubiquitinating enzymes recognize the ubiquitin molecule and precisely cleave the peptide bond between ubiquitin and GLP-1, the cleavability by deubiquitinating enzymes provides information whether the protein adopts a proper conformation to be recognized by the deubiquitinating enzymes such as ubiquitin-specific protease 1 (UBP1) and yeast ubiquitin C-terminal hydrolase 1 (YUH1) which cleave precisely at the C-terminal glycine residue of ubiquitin. The K6UbGLP-1 in the soluble fraction (Fig. 2A) was treated by UBP1 and YUH1. Intriguingly, the SDS–PAGE analysis showed that the soluble K6UbGLP-1 remained uncleaved except for the minor cleavage (Fig. 2C) which indicates that it did not adopt a proper conformation to be recognized by UBP1 and YUH1. On the other hand, both UBP1 and YUH1 were able to efficiently cleave K6UbGLP-1QN in the soluble fraction (Fig. 5B) which indicates that K6UbGLP-1QN folded into a correct conformation and was thus recognized by the deubiquitinating enzymes. Both the deubiquitinating enzymes cleaved the refolded K6UbGLP-1 with over 80% cleavage yield (Fig. 7D) which is contrary to the low cleavage yield of the soluble K6UbGLP-1 (Fig. 2C). Therefore, these results clearly demonstrated that the solid-phase refolding using the cation exchanger enabled K6UbGLP-1 to fold into a proper conformation for GLP-1 production. The

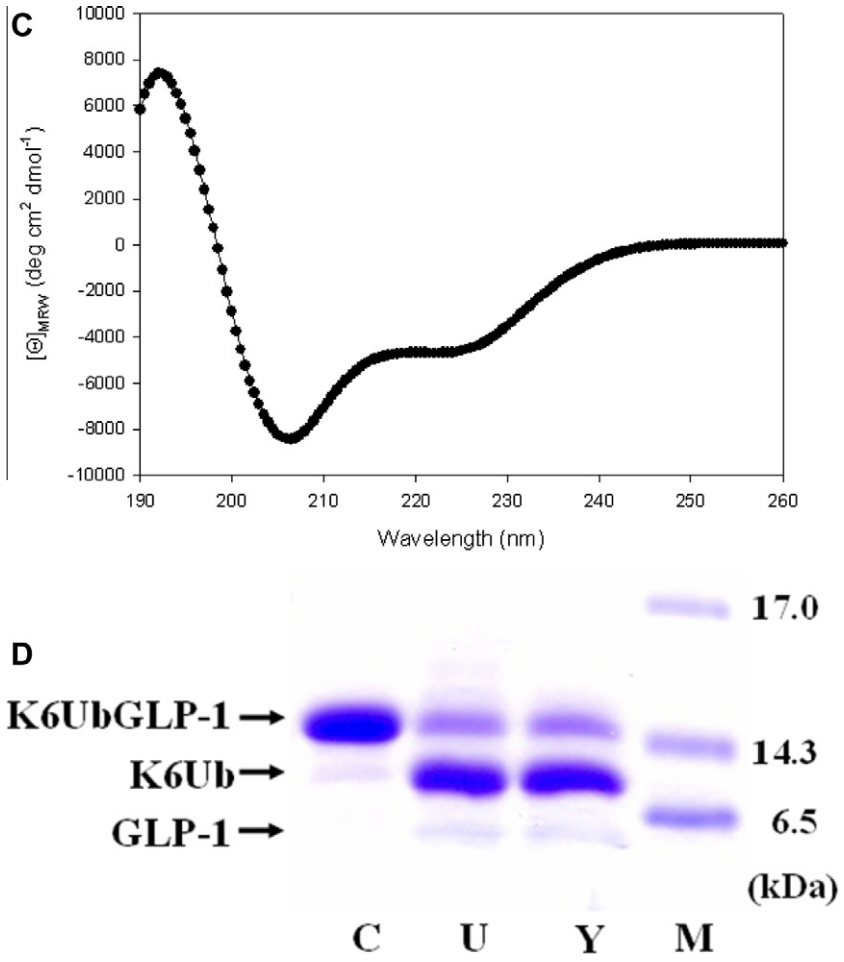

Fig. 7 (*continued*)

recombinant GLP-1 released from the refolded K6UbGLP-1 by YUH1 cleavage showed the identity with the chemically synthesized one in the physiochemical analyses including molecular weight determination, N-terminal sequencing and reverse-phase HPLC (data not shown) of which all the results were the same as those of our previous study [14].

Discussion

Since ubiquitin is a highly stable and soluble protein and thus enhances the solubility of a target protein when fused to it, the ubiquitin fusion system has been developed and applied to the production of a number of proteins using recombinant *E. coli* [11,19–21]. However, despite the high solubility of ubiquitin, K6UbGLP-1 was aggregated as insoluble inclusion bodies in *E. coli*. Our results demonstrated that the insoluble aggregation observed in both *in vivo* and *in vitro* conditions was caused by the electrostatic interactions between the N-terminal 6-lysine tag and the C-terminal GLP-1. Misfolding and aggregation caused by intra- or intermolecular electrostatic interaction are frequently found in various proteins especially amyloid-forming proteins [22,23]. Furthermore, a fragment of Alzheimer's β-amyloid peptide coaggregated with some soluble proteins from different species *via* the electrostatic interaction between the two molecules [24].

In this study, the application of a cation exchanger successfully suppressed the electrostatic interaction-driven aggregation by neutralizing the positive charge of the 6-lysine tag in *in vitro* refolding for K6UbGLP-1. Furthermore, the simultaneous immobilization to the cation exchanger likely prevented intermolecular interactions during the refolding process which enabled refolding of K6UbGLP-1 into a proper conformation even at 3.43 g/L of protein concentration. The electrostatic interaction did not cause the insoluble aggregation after the completion of K6UbGLP-1 refolding (Fig. 7). This is due to ubiquitin with a highly structured state (T_m: 90–100 °C) [12], and its outer surface being very hydrophilic, allowing high concentrations (50 mg/mL) of the protein to be solubilized in an aqueous environment [13]. In addition, any intermolecular multimerization of the refolded K6UbGLP-1 by the electrostatic interaction between the N-terminal 6-lysine tag and the C-terminal GLP-1 was not observed in the gel filtration analysis, suggesting that the electrostatic interaction is transient during the folding process of K6UbGLP-1. Therefore, it seems that the electrostatic interaction takes place at an early step before the completion of folding. Indeed, the major reason for protein aggregation is the intermolecular interaction between hydrophobic surfaces in partially folded intermediates [4]. Ubiquitin folds in accordance with a nucleation–condensation mechanism. In the early steps of ubiquitin folding, its N-terminal region contributes to the formation of a folding nucleus comprising of N-terminal β-hairpin and α-helix [25]. Considering the folding mechanism of ubiquitin, it is expected that the N-terminal 6-lysine tag electrostatically interacts with the C-terminal GLP-1 upon the onset of K6UbGLP-1 folding. This in turn disturbs the contribution of the N-terminal region to the formation of the folding nucleus and a stable tertiary structure resulting in accumulation of the misfolded folding intermediate as the insoluble aggregate.

The refolded K6UbGLP-1 adopted a correct conformation for being recognized by the deubiquitinating enzymes. The mechanisms for substrate binding and cleavage of YUH1 were uncovered by its crystal structure [26]. YUH1 has an active-site-crossover loop that passes directly over the active site. The binding of YUH1 to ubiquitin requires the ubiquitin C-terminal adduct of a substrate polypeptide to pass through this loop. The internal diameter of the loop is only approximately 15 Å which is too small to allow the passage of folded substrates larger than a simple helix. Taking into consideration that the diameter of α-helix ranges from 10 to 15 Å, it seems that GLP-1 of the refolded K6UbGLP-1 adopts an α-helix or a random coil which is consistent to the NMR structure of GLP-1 showing a single-stranded α-helix or a flexible random coil in different trifluoroethanol concentrations [27]. Furthermore, the incomplete cleavage displayed in Fig 7D was likely attributed to the helical dimer formation of GLP-1 [28] in which the diameter of the GLP-1 dimer is larger than that of the loop in YUH1.

In conclusion, this work has revealed that the reason for the inclusion body formation of K6UbGLP-1 in *E. coli*. The electrostatic interaction between the 6-lysine tag and GLP-1 before the completion of folding led to an off-pathway aggregation of K6UbGLP-1. Elimination either of the positive charge in the 6-lysine tag or the negative charge in GLP-1 abolished the electrostatic interaction-driven aggregation. The solid-phase refolding using a cation exchanger fulfilled both the neutralization of the positive charge in the 6-lysine tag and the prevention of intermolecular interaction which resulted in productive folding of K6UbGLP-1 into a proper conformation for GLP-1 production.

Acknowledgments

This work was supported by a Grant from the Next-Generation BioGreen 21 Program (No. PJ008033), Rural Development Administration, Republic of Korea.

References

[1] L.F. Vallejo, U. Rinas, Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins, Microb. Cell Fact. 3 (2004) 11.
[2] A. Villaverde, M.M. Carrio, Protein aggregation in recombinant bacteria: biological role of inclusion bodies, Biotechnol. Lett. 25 (2003) 1385–1395.
[3] S. Ventura, A. Villaverde, Protein quality in bacterial inclusion bodies, Trends Biotechnol. 24 (2006) 179–185.
[4] A.L. Fink, Protein aggregation: folding aggregates, inclusion bodies and amyloid, Fold. Des. 3 (1998) R9–23.
[5] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 60 (2003) 523–533.
[6] A.K. Mohanty, M.C. Wiener, Membrane protein expression and production: effects of polyhistidine tag length and position, Protein Expr. Purif. 33 (2004) 311–325.
[7] M. Kurz, N.P. Cowieson, G. Robin, D.A. Hume, J.L. Martin, B. Kobe, P. Listwan, Incorporating a TEV cleavage site reduces the solubility of nine recombinant mouse proteins, Protein Expr. Purif. 50 (2006) 68–73.
[8] C.M. Halliwell, G. Morgan, C.P. Ou, A.E. Cass, Introduction of a (poly)histidine tag in L-lactate dehydrogenase produces a mixture of active and inactive molecules, Anal. Biochem. 295 (2001) 257–261.
[9] E.A. Woestenenk, M. Hammarstrom, S. van den Berg, T. Hard, H. Berglund, His tag effect on solubility of human proteins produced in *Escherichia coli*: a comparison between four expression vectors, J. Struct. Funct. Genomics 5 (2004) 217–229.
[10] A. Hershko, A. Ciechanover, The ubiquitin system, Annu. Rev. Biochem. 67 (1998) 425–479.
[11] T.R. Butt, S. Jonnalagadda, B.P. Monia, E.J. Sternberg, J.A. Marsh, J.M. Stadel, D.J. Ecker, S.T. Crooke, Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*, Proc. Nat. Acad. Sci. U.S.A. 86 (1989) 2540–2544.
[12] S.E. Jackson, Ubiquitin: a small protein folding paradigm, Org. Biomol. Chem. 4 (2006) 1845–1853.
[13] M. Rechsteiner, Osmotic lysis of pinosomes, Methods Enzymol. 149 (1987) 42–48.
[14] S.G. Kim, S.Y. Shin, Y.C. Park, C.S. Shin, J.H. Seo, Production and solid-phase refolding of human glucagon-like peptide-1 using recombinant *Escherichia coli*, Protein Expr. Purif. 78 (2011) 197–203.
[15] H.A. Yu, S.G. Kim, E.J. Kim, W.J. Lee, D.O. Kim, K. Park, Y.C. Park, J.H. Seo, Characterization of ubiquitin C-terminal hydrolase 1 (YUH1) from *Saccharomyces cerevisiae* expressed in recombinant *Escherichia coli*, Protein Expr. Purif. 56 (2007) 20–26.
[16] K.I. Na, M.D. Kim, W.K. Min, J.A. Kim, W.J. Lee, D.O. Kim, K.M. Park, J.H. Seo, Expression and purification of ubiquitin-specific protease (UBP1) of *Saccharomyces cerevisiae* in recombinant *Escherichia coli*, Biotechnol. Bioprocess Eng. 10 (2005) 599–602.
[17] M.D. Finucane, D.N. Woolfson, Core-directed protein design. II. Rescue of a multiply mutated and destabilized variant of ubiquitin, Biochemistry 38 (1999) 14118.
[18] P.L. Wintrode, G.I. Makhatadze, P.L. Privalov, Thermodynamics of ubiquitin unfolding, Proteins 18 (1994) 246–253.
[19] R.T. Baker, Protein expression using ubiquitin fusion and cleavage, Curr. Opin. Biotechnol. 7 (1996) 541–546.
[20] C. Lee, S.G. Lee, S.T. Takahashi, B.G. Kim, The soluble expression of the human renin binding protein using fusion partners: a comparison of ubiquitin, thioredoxin, maltose binding protein and NusA, Biotechnol. Bioprocess. Eng. 8 (2003) 89–93.
[21] A. Pilon, P. Yost, T. Chase, G. Lohnas, W. Bentley, High-level expression and efficient recovery of ubiquitin fusion proteins from *Escherichia coli*, Biotechnol. Prog. 12 (1996) 331–337.
[22] F. Chiti, M. Calamai, N. Taddei, M. Stefani, G. Ramponi, C.M. Dobson, Studies of the aggregation of mutant proteins in vitro provide insights into the genetics of amyloid diseases, Proc. Nat. Acad. Sci. U.S.A. 99 (2002) 16419–16426.
[23] M.P. Morrissey, E.I. Shakhnovich, Evidence for the role of PrPC helix 1 in the hydrophilic seeding of prion aggregates, Proc. Nat. Acad. Sci. U.S.A. 96 (1999) 11293–11298.
[24] T. Konno, Amyloid-induced aggregation and precipitation of soluble proteins: A electrostatic contribution of the Alzheimer's β(25–35) amyloid fibril, Biochemistry 40 (2001) 2148–2154.
[25] H.M. Went, S.E. Jackson, Ubiquitin folds through a highly polarized transition state, Protein Eng. Des. Sel. 18 (2005) 229–237.
[26] S.C. Johnston, S.M. Riddle, R.E. Cohen, C.P. Hill, Structural basis for the specificity of ubiquitin C-terminal hydrolases, EMBO J. 18 (1999) 3877–3887.
[27] X.Q. Chang, D. Keller, S. Bjorn, J.J. Led, Structure and folding of glucagon-like peptide-1-(7–36)-amide in aqueous trifluoroethanol studied by NMR spectroscopy, Magn. Reson. Chem. 39 (2001) 477–483.
[28] X. Chang, D. Keller, S.I. O'Donoghue, J.J. Led, NMR studies of the aggregation of glucagon-like peptide-1: formation of a symmetric helical dimer, FEBS Lett. 515 (2002) 165–170.