

Original Article

Over-expression of recombinant proteins with N-terminal His-tag via subcellular uneven distribution in *Escherichia coli*

Won-Ji Park, Sung-Hwan You, Hyoung-An Choi, Yeon-Jin Chu, and Geun-Joong Kim*

Department of Biological Sciences, College of Natural Sciences, Chonnam National University, Gwangju 500-757, Republic of Korea

*Correspondence address. Tel: +82-62-530-3403; Fax: +82-62-530-3409; E-mail: gjkim@chonnam.ac.kr

Received 31 December 2014; Accepted 18 March 2015

Abstract

Specific tags with defined amino acid residues are widely used to purify or probe target proteins. Interestingly, the tagging system occasionally results in an increase of the recombinant protein expression *in vivo*. Here, we systematically examined this phenomenon using a poly-histidine (His)-tag fused to N- or C-terminal region of green, red, and blue fluorescent proteins by quantification and uneven distribution in cytoplasm of *Escherichia coli*. This effect was further supported by the distinct over-expression of several unrelated proteins, such as esterase, neopullulanase, and chloramphenicol acetyltransferase, tagging with the same tag. These results suggest that a poly-His-tag placed at N-terminal region can induce over-expression of recombinant protein via subcellular uneven distribution *in vivo*.

Key words: His-tag, over-expression, uneven distribution, immunofluorescence

Introduction

It has become clear that bacterial cytoplasm is not homogeneous but a differentiated medium in which various macromolecules, such as RNAs and proteins, are not freely diffused to particular subcellular regions [1–4]. Uneven distribution of various molecules, referred specifically to as subcellular localization, is common in eukaryotic cells [1]. The unevenly distributed localization is essential for the proper function of proteins in the processes such as cytokinesis, replication, and a wide array of signal transduction pathways [3,5]. This is further reinforced by recent studies showing that RNAs are also localized at certain sites within cells [6]. For example, transfer-messenger RNA is spatially concentrated by organizing into a cytoskeletal structure to avoid ribonuclease R. Therefore, it has been recognized that subcellular localization is not only important for specific functions, but also in regulating concentrations of biomolecules through interactions with factors for quality control [3,6]. Although some progress has been made in the understanding of unevenly distributed phenomena (subcellular localization) in bacterial cytoplasm, the structures or sequence

elements that dictate the uneven distributions of proteins and RNAs remain to be further elucidated.

Functional purification is prerequisite to define the activity and structure of the resulting recombinant protein. In this respect, a poly-histidine (His) peptide tag is commonly used for affinity purification after a prototype study has been reported [7]. This tag provides clear advantages due to its relatively small size and strong binding to metal ion-immobilized resins, which ensures that protein activity is rarely impaired, although there are also reports of cases showing that it may affect the structure or activity of some proteins [8,9]. Therefore, researches of His-tag have mainly focused on the solubility and yield of the target proteins from affinity purification procedures [10–13]. Interestingly, it has been reported that the orientation of His-tags, either N- or C-terminal of the target proteins, affected the expression levels of recombinant proteins, although the underlying mechanism has not been clearly elucidated [14–16].

Recently, we have observed the distinct pattern of uneven distribution of recombinant mBFP protein with an N-terminal His-tag [17].

This pattern was observed to be closely associated with the over-expression in *Escherichia coli* under no induction conditions. The protein mBFP is an NADPH-dependent fluorescence protein derived from a metagenome; thus, there is no possibility of finding a functionally interacting partner in *E. coli* [17]. In this study, we provide evidence that the His-tag is responsible for the uneven distribution (subcellular localization) and over-expression, particularly when it is tagged to N-terminal region of recombinant proteins.

Materials and Methods

Strains, plasmids, and culture conditions

The *E. coli* strain used in this study was XL1-Blue (Stratagene, La Jolla, USA). The genes of two fluorescent proteins, *Discosoma* sp. red fluorescent protein (DsRed) and green fluorescent protein variant optimized for excitation by UV light (GFPuv), were obtained from the commercial vectors pDsRed2-N1 and pGFPuv (Clontech, Mountain View, USA), respectively. The gene of blue fluorescent protein (mBFP) was from pBGR-Blue [17]. The chloramphenicol acetyltransferase (CAT) gene was derived from the cloning vector pACYC184 (Promega, Madison, USA). The genes encoding an esterase 1767 and a neopullulanase (BCN, *Bacillus cereus* neopullulanase) were obtained from the corresponding genomic DNA by PCR [18,19]. The cells were grown at 37°C in LB media supplemented with 50 µg/ml ampicillin. When required, the expression of recombinant protein was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) under defined conditions.

Construction of recombinant plasmids

Two expression vectors, pQE30 (QIAGEN, Heidelberg, Germany) and pET-22b(+) (Novagen, Darmstadt, Germany), containing His-tag at their N- or C-terminal region of multi-cloning site, were used to construct the His-tagged proteins. The genes (*dsred*, *gfpuv*, *mbfp*, and *cat*) encoding the target proteins with His-tag were amplified by PCR using a set of primers (Table 1) and then digested with restriction enzymes (*SphI/HindIII* for pQE30 and *EcoRI/HindIII* for pET-22b(+)), respectively. These genes were ligated with expression vectors and then transformed into *E. coli* cells. The resulting constructs were confirmed by DNA sequencing.

For a clear comparison of the His-tag effects under the same expression conditions, all constructs were further subcloned into the pTrc99A expression vector using the *EcoRI/HindIII*-digested (encoding each protein with an N-terminal His-tag) or PCR-amplified (encoding each protein with a C-terminal His-tag) DNA fragments. The construct pTrc99A-His-CAT was prepared by replacing the gene encoding DsRed of pTrc99A-His-DsRed with the gene encoding CAT. All genes without any tags were also cloned into the same vector as controls. The additional two His-tag constructs, His-BCN and His-1767, were also prepared to compare the distribution pattern in cytoplasm with those of fluorescent proteins under the same conditions. Therefore, these two genes were subcloned into the expression vector pTrc99A by PCR using a set of primers containing an appropriate restriction site (Table 1).

Fluorescent cell imaging and intensity

The recombinant *E. coli* cells were cultivated at 37°C in LB media and induced with 0.1 mM IPTG at an optical density (600 nm) of 0.5, and then incubated for an appropriate time. To monitor the uneven distribution (localization) patterns under a low level of expressed proteins, the cells were incubated without inducer at 37°C for 8 h, or further

production of proteins in the cells was blocked with 50 µg/ml tetracycline and 170 µg/ml chloramphenicol after 25 min of induction with 0.1 mM IPTG. All cells were subject to image analyses in the following manner. An aliquot of cultured cells was withdrawn to investigate the time course cell images. These cells were diluted with LB media or phosphate-buffered saline to an optical density of 1.0 and placed on slide glasses. After air drying for 20 min at room temperature, the cells were observed using a BX43 fluorescence microscope (Olympus, Tokyo, Japan) equipped with reflected fluorescence system and ×100 differential interference contrast (DIC) or phase-contrast objectives. The excitation wavelength for the filter mounted on the BX3-URA universal reflected illuminator (Olympus) was 340–390 nm (U-FUW) for mBFP, 460–495 nm (U-FBW) for GFPuv, and 530–550 nm (U-FGW) for DsRed and immunofluorescent proteins. The image was captured by a DP71 digital camera (Olympus) with a 500 ms exposure time for the DIC and fluorescence images. Images were processed using software (DP-BSW; TOMORO Digital Image Package, Tokyo, Japan). The relative fluorescence was also measured with harvested cells in 96-well plates using a fluorescence spectrophotometer (Infinite M200; Tecan, Männedorf, Switzerland).

Protein expression analyses

The recombinant *E. coli* cells were cultured at 37°C in LB media for comparative analyses of the expression level of the protein with or without the His-tag. After 8 h of culture without induction, the cells were appropriately concentrated (1.5 mg/ml) and re-suspended in a buffer (50 mM Tris-HCl, pH 6.8) and then sonicated for 10 s twice at 4°C. After centrifugation at 14,000 g for 20 min, an aliquot of cell lysates (20 µl) was subject to 11% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) analyses according to a general procedure. The expression level of proteins in soluble fraction was determined according to the previous report [20]. Band intensity of expressed proteins was quantitatively analyzed using ImageJ software (ver. 1.44p) and also confirmed by western blotting using anti-His-tag and anti-mBFP antibody according to a general procedure of manufacturer (Santa Cruz Biotech, Santa Cruz, USA).

MIC and zymogram assay of proteins

The clone expressing CAT with a His-tag was cultured for 14 h at 37°C, and the effect of the increased expression level by the His-tag on antibiotic resistance of the host was evaluated. In brief, an aliquot of harvested cells (10⁸ cells) was mixed with fresh LB media and spread on the solid plates containing LB media. Chloramphenicol was dropped on 6 mm filter paper in a 10-fold concentration series from 10² to 10⁵ µg/ml. After incubation for 16 h, the diameter of the growth inhibition zone was measured. Clones expressing CAT with a C-terminal His-tag or no tag were assayed under the same conditions as controls. The activity staining (zymogram) of esterase was conducted according to the general procedure of a previous report [18]. In brief, cell lysates were mixed with 0.2 volume of native sample buffer and resolved on a 11% PAGE. After gel electrophoresis, half of the separating gel was washed twice with 50 mM Tris-HCl (20 ml) buffer and then soaked into the same buffer containing 4 mg of α-naphthyl acetate dissolved in ethoxyethanol (0.5 ml). The activity band was developed by dropping Fast Blue RR (2 mg/ml) solution.

Indirect immunofluorescence staining

Immunofluorescence images of the cells were determined as described elsewhere with slight modifications [21], using rabbit anti-His-tag

Table 1. Primers and vectors used in this study

Primer ^a	Sequence (5' → 3')	Construct
H-DsRedf	ATAGCATGCCCTCCTCCGAGAAC	pQE30-DsRed
H-DsRedr ^b	ATAAAGCTTCTACAGGAACAGGTGG	
H-GFPf	ATAGCATGCAGTAAAGGAGAAGAACTT	pQE30-GFPuv
H-GFP _r ^b	TTTAAGCTTTTATTTGTAGAGCTCATCC	
H-mBFPf	ATAGCATGCCAGAATCTGAACG	pQE30-mBFP
H-mBFP _r ^b	TTAAGCTTCA AGCGGCGAAG CCGCCGT	
H-BCNf	AAGGATCCATGTTTAAAGAAGCTATTTATCATAGGCC	pQE30-BCN
H-BCNr	AACTGCAGTTACCATTTTAATATGGAGAATCCATATGG	
H-1767f	CACGGATCCGTGCAGATTCAGGGTACA	pQE30-1767
H-1767r	CAAGCTTTTACAGACAACCGGCCAATA	
DsRed-Hf ^c	TATCATATGGAATTCGCCTCCTCCGAGAACGTCA	pET-22b(+)-DsRed
DsRed-Hr	ATACTCGAGAAGCTTCAGGAACAGGTGGTGGCGG	
GFP-Hf ^c	GAGGAATTCAGTAAAGGAGAAGAAGCTTT	pET-22b(+)-GFPuv
GFP-Hr	TATAAGCTTGTAGAGCTCATCCATGC	
mBFP-Hf ^c	ATAGAATTCCAGAATCTGAACGGC	pET-22b(+)-mBFP
mBFP-Hr	ATAAAGCTTAGCGGCGAAGCCGCC	
PET _r -1	TTCGGGCTTTTGTAGCAGCCCGGGTCAGTGGTGGTG	pTrc99A-DsRed- and GFPuv-His
PET _r -2	TTCGGGCTTTTGTAGCAGCCCTGCAGTCAGTGGTGGT G	pTrc99A-mBFP- and CAT-His
DsRedf	AGACCATGGCCTCCTCCGAGAACGTCATCACCG	pTrc99A-DsRed
GFPuvf	ATGGAATTCATGAGTAAAGGAGAAGAAGCT	pTrc99A-GFPuv
mBFPf	ATACATATGCAGAATCTGAACGGCAAAGTGCG	pTrc99A-mBFP
CATf	ATAGGATCCGAGAAAAAATCACTGGATATACCACCGTT	pTrc99A-His-CAT and pTrc99A-CAT
CAT _r	TATAAGCTTACGCCCCGCCCTGCCA	
CAT-Hr	ATACTCGAGCGCCCCGCCCTGCCA	pET-22b(+)-CAT
BCNf	CGCCCATGGGAATGTTTAAAGAAGCTATTTATCATAGGCCG	pTrc99A-BCN
H-BCN2f	TAAACATGGGACATCACCATCACCATCAGGATCC	pTrc99A-His-BCN
1767f	GCGGAATTCATGGTGCAGATTTCAGGGTCATTA	pTrc99A-1767
1767r	TATAAGCTTTTACAGACAACCGGC	

Restriction sites are underlined.

^aThe name of primer denotes target gene's name with His-tag located.

^bThese reverse primers were also used in PCR for pTrc99A-gene constructs.

^cThese forward primers were also used in PCR for pTrc99A-gene-His constructs.

antibody (1:75; Santa Cruz Biotech) and rhodamine-conjugated secondary antibody (1:200, goat anti-rabbit IgG; Millipore).

Statistical analysis

All experiments were conducted in triplicate and the representative ones were shown in all the figures.

Results

Distinct over-expression of proteins with N-terminal His-tag under no induction condition

Blue fluorescent protein mBFP is a member of the short-chain dehydrogenase family proteins isolated from a metagenomic DNA library. The mBFP contains two signature sequences, an NADPH binding motif and a Rossmann fold, for distinct structures and emits fluorescence only upon association with NADPH [17]. In order to further evaluate its potential as a novel reporter protein, we constructed an expression vector carrying His-tag at N-terminus of mBFP (pQE30-mBFP), and observed unevenly distributed blue fluorescence loci in the transformed *E. coli* cells, particularly at the end poles or membranes of cells (Fig. 1A). Most interestingly, the expression level of mBFP was dramatically increased to over 25% of total cellular protein even in the absence of the inducer IPTG (Fig. 1B). An evenly distributed pattern of blue fluorescence was observed with mBFP fused to other peptide tags such as Myc. No localization patterns were seen in the cells expressing

mBFP alone or that fused to maltose-binding protein (MBP). When His-tag was fused to the N-terminal region of MBP-mBFP, the unevenly distributed fluorescence signal was partly observed, although image was somewhat vague presumably due to abnormal folding and/or low stability of the resulting fusion protein His-MBP-mBFP with a calculated molecular mass of ~70 kDa (data not shown). This might be due to a high oligomerization tendency of mBFP [17]. Taken together, these results provide a clue that His-tag might function at the N-terminal region of recombinant protein as a subcellular uneven distribution signal and leads to an over-expression of target proteins.

This possibility was further tested using two fluorescent proteins DsRed and GFPuv, which could be readily imaged at the single cell level. mBFP was included as a positive control. The genes for these fluorescent proteins were amplified by PCR and subcloned into the pQE30 and pET-22b(+) expression vectors, carrying innate His-tag at N- and C-terminal region of the multiple cloning site, respectively. These constructs including the His-tag were further subcloned into pTrc99A expression vector to compare the expression level and localization patterns of all proteins under the same conditions. The *E. coli* cells carrying DsRed and GFPuv constructs (incorporated gene with His-tag or no tag in the same vector pTrc99A) were streaked on LB plate and incubated for 24 h at 37°C. As expected, cells harboring the gene with the N-terminal His-tag showed distinct fluorescence even in solid plates without any inducer (Fig. 2A). At the colony level, the increase in fluorescence was dramatic and thus readily discriminated from cells expressing proteins with the C-terminal His-tag or

no tag. Subsequently, we quantified fluorescence proteins from all constructs on SDS-PAGE using the cultivated cells for 8 h without any inducer. Cell lysates were divided into soluble and insoluble fractions after centrifugation at 15,000 g for 30 min. As shown in Fig. 2B, the expression level of the fluorescent protein with the N-terminal His-tag was significantly higher than those of constructs with C-terminal His-tag or no tag. The N-terminal His-tagged fluorescent proteins were

accumulated up to $12.2\% \pm 2.4\%$ (DsRed), $10.4\% \pm 2.2\%$ (GFPuv), and $23.8\% \pm 1.9\%$ (mBFP) of total cell protein. These were 5.36, 5.38, and 4.86 times more than that without tags. Most intriguingly, almost all proteins expressed with the N-terminal His-tag were found in soluble fractions. These results provided evidence that the proteins were not over-expressed as inclusion bodies. In the contrast to that of N-terminal His-tagged mBFP, the expression level of C-terminal

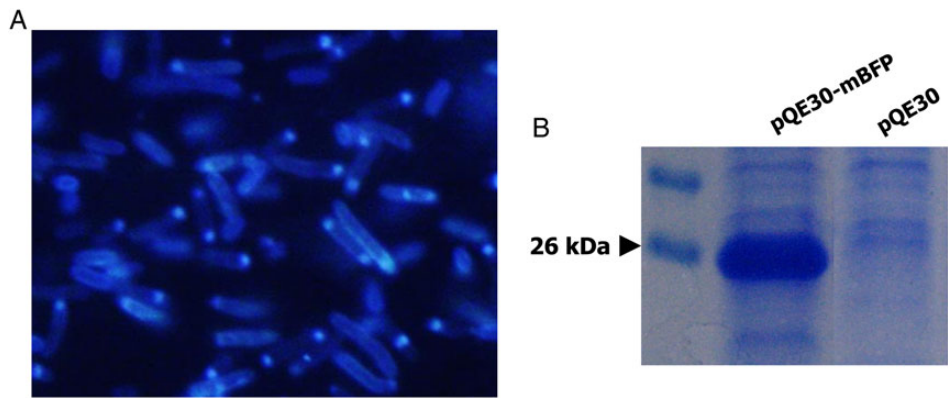


Figure 1. Fluorescence microscopy and SDS-PAGE analysis of mBFP When expressed with an N-terminal His-tag, mBFP was unevenly distributed and mainly localized at the end pole or membrane (A) The expression level of the resulting protein increased to be ~25% of total cell protein (B) when harvested at optical density of 1.5 without induction. The expression level of mBFP was analyzed on a 11% SDS-PAGE gel using cells harboring an expression vector pQE30 with innate N-terminal His-tag (RGSHHHHH).

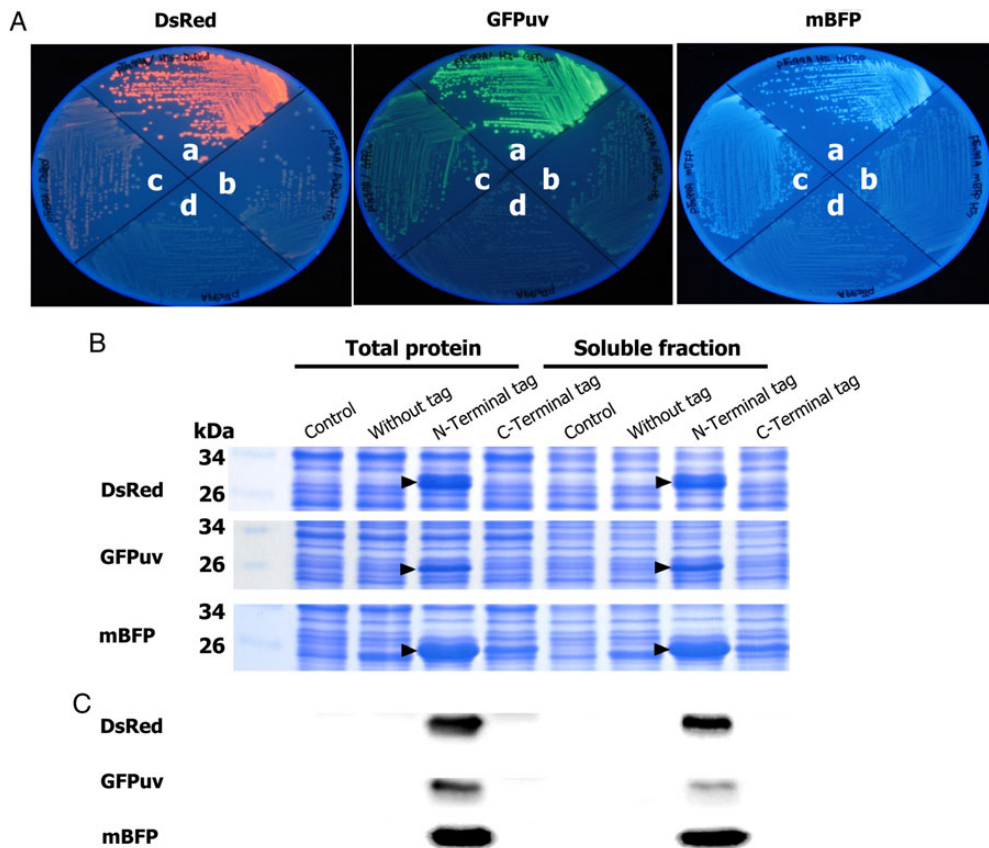


Figure 2. Fluorescent intensities and expression levels of fluorescent proteins with His-tag (A) Clones expressing N-terminal (a), C-terminal (b), and no-tagged (c) fluorescent proteins, and the clone transformed with an empty vector (d). The images were captured after 24 h cultivation in LB-agar plates. (B) SDS-PAGE (11%) was conducted using the culture in LB media for 8 h without any inducer. (C) Western blotting was conducted using anti-His-tag antibody according to a provided procedure by the manufacturer. Bands corresponding to the calculated molecular masses of proteins are indicated by arrows.

His-tagged protein was severely unstable and often lower than that of non-tagged proteins. Western blotting using anti-His-tag antibody further supported these results (Fig. 2C).

Uneven distribution of proteins with N-terminal His-tag *in vivo*

We then examined the fluorescent images of tagged proteins in the cells cultivated for 8–12 h at 37°C to address whether or not the His-tag of the constructs dictated uneven distribution in cytoplasm. Expectedly, many fluorescence images from N-terminal His-tagged proteins showed unevenly distributed and concentrated at the membrane or end poles of cells. However, the localized signals of proteins expressed in cells harboring the expression vector pTrc99A had more unclear boundary image than proteins expressed in cells harboring the expression vector pQE-30. This could be partly due to lower quantum yields and shorter excitation/emission wavelengths of GFPuv (400/509 nm) and mBFP (350/450 nm) than those of DsRed (488/583 nm). It is also well known that the slow maturing DsRed requires a prolonged cultivation time (9–12 h) for visible images after seeding bacterial cells. Thus, hyper-level expressed proteins,

particularly in the case of protein with an N-terminal His-tag, could generate an unclear localization boundary image due to super-saturation of protein expression. Therefore, further attempts to confirm the function of His-tag as a localized distribution signal at an unsaturated level were made by blocking translation with 50 µg/ml tetracycline and 170 µg/ml chloramphenicol after 25 min of the induction with 0.1 mM IPTG. IPTG induction permitted a detectable amount of protein expression in a controllable manner and did not disturb the localized distribution. Consequently, overall fluorescent intensities of these constructs were weaker, but images with more distinct localization of DsRed and GFPuv with the N-terminal His-tag were obtained (Fig. 3). All proteins with N-terminal His-tags showed similar fluorescent loci as those of translation-unblocked cells. About 98% of cells expressing DsRed exhibited unevenly distributed patterns of fluorescent loci, whereas relatively lower frequencies of unevenly distributed patterns were observed for GFPuv (32%–50%) and mBFP (45%–60%). However, fluorescence of the C-terminal- or non-tagged proteins was evenly distributed throughout the entire cytoplasmic region. The fluorescence intensities of C-terminal tagged proteins were relatively lower than those of proteins with an N-terminal His-tag. Especially, the fluorescence of C-terminal His-tagged DsRed

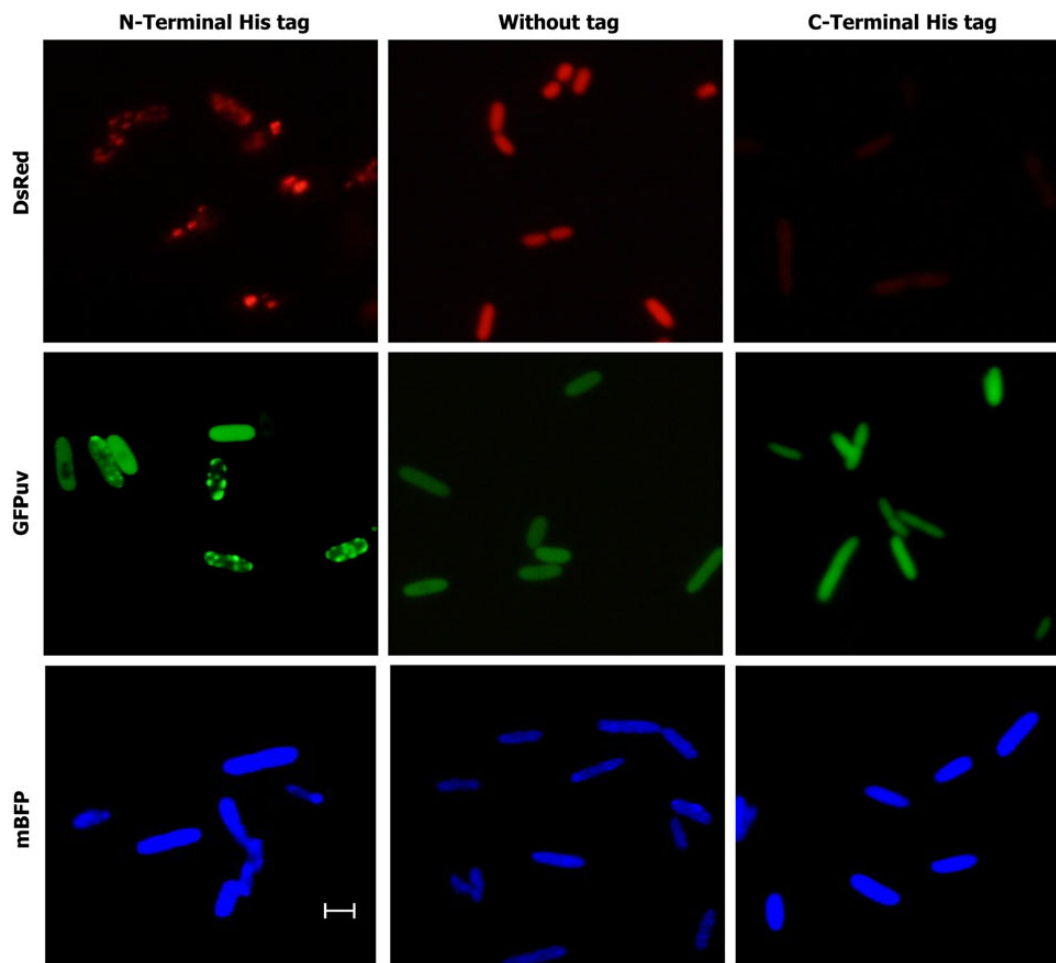


Figure 3. Subcellular localization of His-tagged fluorescent proteins All fluorescent proteins were expressed under the same conditions using pTrc99A as the expression vector. Cells expressing proteins with an N-terminal His-tag were induced with 0.1 mM IPTG for 25 min and blocked from further expression by treatment with 170 µg/ml chloramphenicol and 50 µg/ml tetracycline. Cell images were captured after further incubation for 9 h (His-DsRed) and 2 h (His-GFPuv and His-mBFP), according to their maturation time. To induce an expression suitable for imaging, cells expressing proteins with a C-terminal or no tag were induced with 1 mM IPTG for 1 h and treated with the same amount of antibiotics to block further expression.

was quite low and thus hardly detected in this experiment even when induced with 1 mM IPTG (Fig. 3). Therefore, no dictated roles of C-terminal His-tag for subcellular uneven distribution were observed *in vivo*. We observed the similar results with other *E. coli* strains (DH5 α , Top10, and BL21), suggesting that the localized uneven distribution with N-terminal His-tag is not strain-specific.

Finally, to provide further evidences for localized uneven distribution of expressed proteins with N-terminal His-tag, structurally unrelated proteins including CAT (chloramphenicol acetyl transferase, 26 kDa, trimer), BCN (neopullulanase, 61 kDa, dimer), and enzyme 1767 (esterase, 41 kDa, monomer) were arbitrary selected and cloned into the expression vector pTrc99A as those of fluorescent proteins

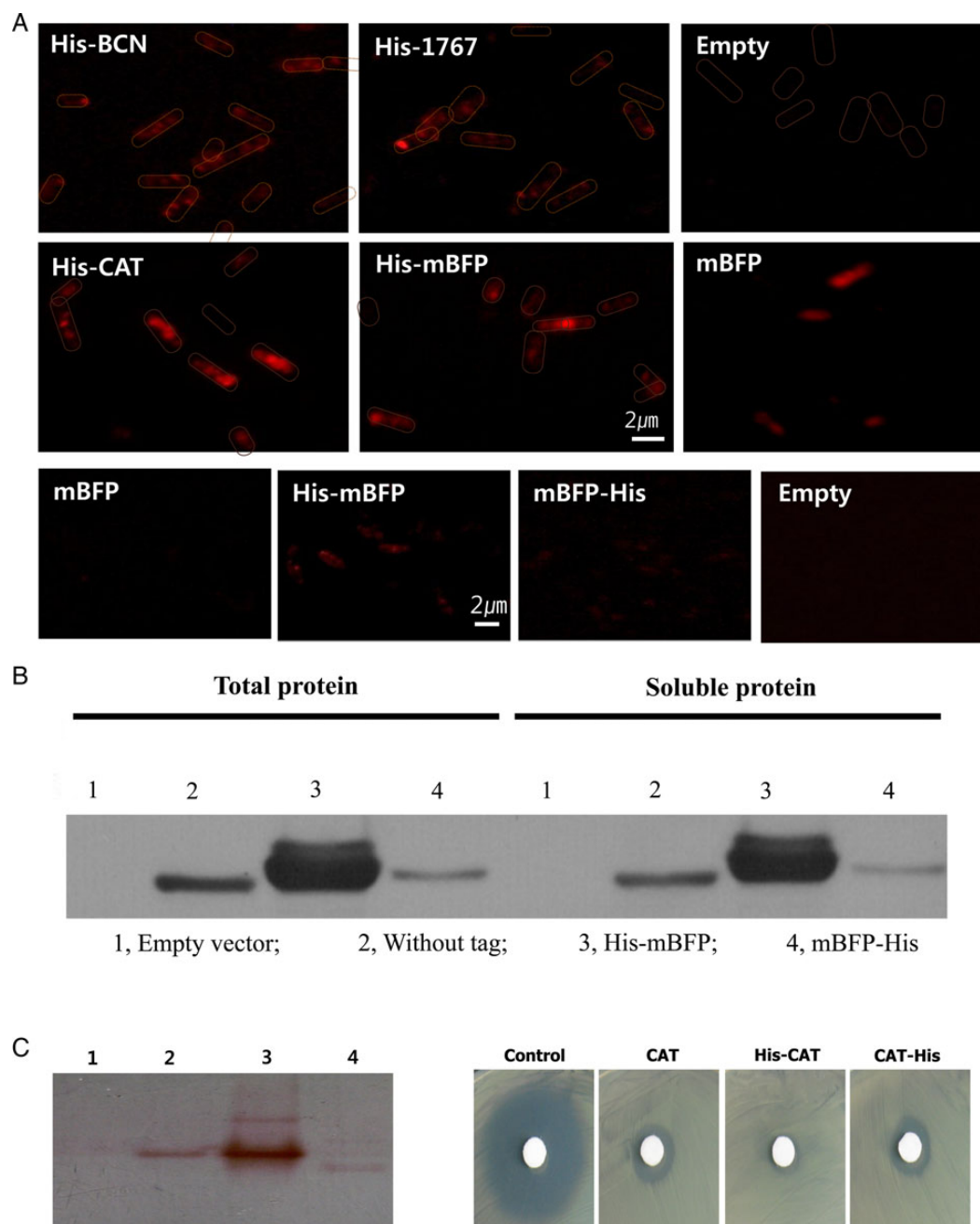


Figure 4. His-tag effect on uneven distribution and activity of chloramphenicol acetyltransferase (CAT), neopullulanase (BCN), and esterase 1767 (A) Distinct uneven distribution of proteins with N-terminal His-tag was observed by immunofluorescence staining using anti-His-tag antibody. As a negative control, unlocalized patterns of mBFP (without His-tag) using anti-mBFP antibody were shown in this figure. All constructs were expressed under the same conditions using the expression vector pTrc99A. (B) Comparison of distribution pattern of proteins with N- and C-terminal His-tag, or without tag under the same culture conditions. Right panel, light microscopic image; left panel, immunofluorescent image, using anti-His-tag antibody. Western blot was also conducted using anti-mBFP antibody as a control for protein expression under the same condition. (C) Zymogram assay for esterase 1767 (left panel) and MIC assay for CAT (right panel). For zymogram assay, an aliquot (10 μ g) of the crude enzymes was analyzed on a 11% native-PAGE and then stained with α -naphthyl acetate and Fast Blue RR. Lane 1, empty vector; lane 2, 1767; lane 3, His-1767; lane 4, 1767-His. MIC assay showed differences in inhibitory zone against 6 mm paper impregnated with the same concentration of chloramphenicol (10^5 μ g/ml).

described earlier. Subsequently, the recombinant cells harboring each construct were cultivated in LB media at 37°C for 8 h, and then immunostained with anti-His-tag antibody. We used the same construct of mBFP as a positive control. As expected, immunofluorescence staining results indicated that all proteins expressed with N-terminal His-tag were unevenly distributed in cytoplasm and thus localized as discrete loci near the cell poles or membranes under the no-induction conditions (Fig. 4A). As observed in the cases of three fluorescent proteins, the majority of cells (62%–85%) displayed distinct localization patterns. These results further provided evidence that the expressed proteins with N-terminal His-tag were unevenly distributed in subcellular regions consistently. Although the expressed proteins without tags could not be detected with antibodies due to the absence of His-tag, immunofluorescence of mBFP using a specific antibody (raised against mBFP) was evenly distributed throughout the cells, as identical to that of blue fluorescence of mBFP protein (Figs. 3 and 4A). The immunofluorescence loci of C-terminal His-tagged proteins could not be easily detected under the same culture conditions due to their low expression levels, as typically observed in the case of mBFP (Fig. 4B). Western blot analysis also supported these results (Fig. 4B, lower panel).

The increased activity also showed that the proteins with N-terminal His-tag were over-expressed than those of no-tags and C-terminal His-tag (Fig. 4C). As shown in the zymogram analyses, the expression level of esterase 1767 was distinctly increased in the cells harboring a construct with N-terminal His-tag. Maximum inhibitory concentration (MIC) assay also showed that the viability of cells expressing CAT with N-terminal His-tag was clearly increased. These cells overgrew on solid media supplemented with 10^5 µg/ml chloramphenicol as expected. The increased activity in zymogram assay of His-BCN (1.2–1.4 folds higher than that of no-tag) was relatively low compared with those of the His-CAT and His-1767. Quantification analyses using the band intensity of the expressed proteins on SDS-PAGE showed that the His-tagged proteins at N-terminal regions accumulated *in vivo* to high level up to $15.7\% \pm 2.6\%$ (CAT), $10.2\% \pm 2.2\%$ (esterase 1767), and $6.5\% \pm 1.7\%$ (BCN) of total cell proteins, which were 6.2, 4.8, and 1.3 times more than those of no-tagged proteins when cultivated at 37°C for 8 h under no-induction conditions (data not shown).

Discussion

We have provided plausible evidence for subcellular uneven distribution of proteins tagged with poly-His at N-terminus, accompanied by an increase in expression level under the no-induction conditions. When induced with IPTG, the differences in expression level between protein with N-terminal His-tag and no-tag (or C-terminal His-tag) were decreased, especially when induced with a relatively high concentration (>0.1 mM) of IPTG for a prolonged time (>90 min). However, the localizing loci of unevenly distributed protein in cytoplasm were consistently observed, although the boundaries of fluorescence loci were occasionally ambiguous due to hyper-level expression of N-terminal His-tagged proteins under the induction conditions. There are two possible mechanisms for uneven distribution of proteins with N-terminal His-tag: the localization signal could be either the His-tag RNA or poly-His of expressed proteins. The latter case is presumed to be more reasonable due to the different localization patterns and fairly concerted effects (high avidity of three or four His-tags of oligomeric proteins could allow for better uneven distribution) by the innate oligomeric state. In fact, we observed more distinct localization patterns with tetrameric (DsRed and mBFP) and trimeric proteins (CAT) than that of monomeric (GFPuv and esterase) and dimeric

(BCN) proteins. Relatively low localization signal with high molecular weight proteins (BCN and esterase) also further support the second model. However, the exact mechanisms and/or *in vivo* interacting partner for over-expression of unevenly distributed proteins with N-terminal His-tag are unknown under the no-induction conditions, although there is a possibility that the quality control, such as chaperon partitioning or degradation of product (RNA or protein), was avoided by localization. Further studies are needed to elucidate the role or mechanism of the His-tag in uneven distribution and thus subcellular localization. For example, different tags composed of the synonymous His codon may also help to confirm whether localization can be addressed at the transcription or translation level.

The signal sequence prediction programs PSORTb [22] and SignalP [23] revealed that the poly-His-tag used here is not a known signal peptide in bacteria but partly matches RcnRA, a protein localized in the cytoplasmic membrane with a 13 histidine-rich stretch that is functionally related to the nickel/cobalt efflux system in *E. coli*. It is likely that the tag used in this study also interacts with a related protein or charged molecules (cationic metal-rich) through electrostatic interactions. In this context, a recently reported organelle in bacteria, similar to acidocalcisomes of unicellular eukaryotes [24], could be a possible interacting partner. The acidocalcisomes are phosphorus, magnesium, potassium, and calcium storage compartments and, thus, could have an innate ability to bind histidine. Demonstrating the existence of an acidocalcisome counterpart in *E. coli* is an intriguing subject for future studies. Although a detailed mechanism remains to be elucidated, His-tag may protect tagged proteins in a subcellular region from cellular machinery responsible for the quality control of RNAs or proteins.

Funding

This work was supported by a grant from the Intelligent Synthetic Biology Center of the Global Frontier Project of the Korean Ministry of Education, Science and Technology (No. 2014M3A6A8063669).

References

- Harold FM. Molecules into cells: specifying spatial architecture. *Microbiol Mol Biol Rev* 2005, 69: 544–564.
- Shapiro L, Losick R. Dynamic spatial regulation in the bacterial cell. *Cell* 2000, 100: 89–98.
- Iniesta AA, McGrath PT, Reisenauer A, McAdams HH, Shapiro L. A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc Natl Acad Sci USA* 2006, 103: 10935–10940.
- Taghbalout A, Rothfield L. RNaseE and the other constituents of the RNA degradosome are components of the bacterial cytoskeleton. *Proc Natl Acad Sci USA* 2007, 104: 1667–1672.
- Montero Llopis P, Jackson AF, Sliusarenko O, Surovtsev I, Heinritz J, Emonet T, Jacobs-Wagner C. Spatial organization of the flow of genetic information in bacteria. *Nature* 2010, 466: 77–81.
- Russell JH, Keiler KC. Subcellular localization of a bacterial regulatory RNA. *Proc Natl Acad Sci USA* 2009, 106: 16405–16409.
- Hochuli E, Bannwarth W, Döbeli H, Gentz R, Stüber D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Nat Biotechnol* 1988, 6: 1321–1325.
- Horchani H, Sabrina L, Régine L, Sayari A, Gargouri Y, Verger R. Heterologous expression and N-terminal His-tagging processes affect the catalytic properties of staphylococcal lipases: a monolayer study. *J Colloid Interface Sci* 2010, 350: 586–594.
- Horchani H, Ouertani S, Gargouri Y, Sayari A. The N-terminal His-tag and the recombination process affect the biochemical properties of

- Staphylococcus aureus* lipase produced in *Escherichia coli*. *J Mol Catal B Enzym* 2009, 61: 194–201.
10. Letourneur O, Ottone S, Delaunay V, Bastide MC, Foussadier A. Molecular cloning, overexpression in *Escherichia coli*, and purification of 6x his-tagged C-terminal domain of *Clostridium difficile* toxins A and B. *Protein Expr Purif* 2003, 31: 276–285.
 11. Zakalskiy AE, Zakalska OM, Rzhepetskiy YA, Potocka N, Stasyk OV, Horak D, Gonchar MV. Overexpression of (His)(6)-tagged human arginase I in *Saccharomyces cerevisiae* and enzyme purification using metal affinity chromatography. *Protein Expr Purif* 2012, 81: 63–68.
 12. Job V, Molla G, Pilone MS, Pollegioni L. Overexpression of a recombinant wild-type and His-tagged *Bacillus subtilis* glycine oxidase in *Escherichia coli*. *Eur J Biochem* 2002, 269: 1456–1463.
 13. Pagliato L, Negri A, Nonnis S, Taverna F, Sangiorgio L, Ronchi S, Tedeschi G. Prion protein from *Xenopus laevis*: overexpression in *Escherichia coli* of the His-tagged protein and production of polyclonal antibodies. *Protein Expr Purif* 2006, 46: 489–494.
 14. Mason AB, He QY, Halbrooks PJ, Everse SJ, Gumerov DR, Kaltashov IA, Smith VC, et al. Differential effect of a His tag at the N- and C-termini: functional studies with recombinant human serum transferrin. *Biochemistry* 2002, 41: 9448–9454.
 15. Xu CG, Fan XJ, Fu YJ, Liang AH. Effect of location of the His-tag on the production of soluble and functional *Buthus martensii* Karsch insect toxin. *Protein Expr Purif* 2008, 59: 103–109.
 16. Woestenak EA, Hammarstrom M, van den Berg S, Hard T, Berglund H. His tag effect on solubility of human proteins produced in *Escherichia coli*: a comparison between four expression vectors. *J Struct Funct Genomics* 2004, 5: 217–229.
 17. Hwang CS, Choi ES, Han SS, Kim GJ. Screening of a highly soluble and oxygen-independent blue fluorescent protein from metagenome. *Biochem Biophys Res Commun* 2012, 419: 676–681.
 18. Kim JH, Choi GS, Kim SB, Kim WH, Lee JY, Ryu YW, Kim GJ. Enhanced thermostability and tolerance of high substrate concentration of an esterase by directed evolution. *J Mol Catal B Enzym* 2004, 27: 169–175.
 19. Ivanova N, Sorokin A, Anderson I, Galleron N, Candelon B, Kapatral V, Bhattacharyya A, et al. Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* 2003, 423: 87–91.
 20. Kim GJ, Lee DE, Kim HS. Functional expression and characterization of the two cyclic amidohydrolase enzymes, allantoinase and a novel phenylhydantoinase, from *Escherichia coli*. *J Bacteriol* 2000, 182: 7021–7028.
 21. Carbajal D, Morano KA, Morano LD. Indirect immunofluorescence microscopy for direct detection of *Xylella fastidiosa* in xylem sap. *Curr Microbiol* 2004, 49: 372–375.
 22. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 2010, 26: 1608–1615.
 23. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011, 8: 785–786.
 24. Seufferheld M, Vieira MC, Ruiz FA, Rodrigues CO, Moreno SN, Docampo R. Identification of organelles in bacteria similar to acidocalcisomes of unicellular eukaryotes. *J Biol Chem* 2003, 278: 29971–29978.