

RESEARCH

Open Access



The signal peptide of Cry1Ia can improve the expression of eGFP or mCherry in *Escherichia coli* and *Bacillus thuringiensis* and enhance the host's fluorescent intensity

Jianhua Gao^{1*} , Hongmei Qian¹, Xiaoqin Guo¹, Yi Mi¹, Junpei Guo¹, Juanli Zhao¹, Chao Xu², Ting Zheng², Ming Duan³, Zhongwei Tang¹, Chaoyang Lin², Zhicheng Shen², Yiwei Jiang⁴ and Xingchun Wang^{1*}

Abstract

Background: The signal peptides (SPs) of secretory proteins are frequently used or modified to guide recombinant proteins outside the cytoplasm of prokaryotic cells. In the periplasmic space and extracellular environment, recombinant proteins are kept away from the intracellular proteases and often they can fold correctly and efficiently. Consequently, expression levels of the recombinant protein can be enhanced by the presence of a SP. However, little attention has been paid to the use of SPs with low translocation efficiency for recombinant protein production. In this paper, the function of the signal peptide of *Bacillus thuringiensis* (Bt) Cry1Ia toxin (Iasp), which is speculated to be a weak translocation signal, on regulation of protein expression was investigated using fluorescent proteins as reporters.

Results: When fused to the N-terminal of eGFP or mCherry, the Iasp can improve the expression of the fluorescent proteins and as a consequence enhance the fluorescent intensity of both *Escherichia coli* and Bt host cells. Real-time quantitative PCR analysis revealed the higher transcript levels of *legfp* over those of *egfp* gene in *E. coli* TG1 cells. By immunoblot analysis and confocal microscope observation, lower translocation efficiency of IeGFP was demonstrated. The novel fluorescent fusion protein IeGFP was then used to compare the relative strengths of *cry1Ia* (P_i) and *cry1Ac* (P_{ac}) gene promoters in Bt strain, the latter promoter proving the stronger. The eGFP reporter, by contrast, cannot indicate unambiguously the regulation pattern of P_i at the same level of sensitivity. The fluorescent signals of *E. coli* and Bt cells expressing the Iasp fused mCherry (ImCherry) were also enhanced. Importantly, the Iasp can also enhanced the expression of two difficult-to-express proteins, matrix metalloprotease-13 (MMP13) and myostatin (growth differentiating factor-8, GDF8) in *E. coli* BL21-star (DE3) strain.

Conclusions: We identified the positive effects of a weak signal peptide, Iasp, on the expression of fluorescent proteins and other recombinant proteins in bacteria. The produced IeGFP and ImCherry can be used as novel fluorescent protein variants in prokaryotic cells. The results suggested the potential application of Iasp as a novel fusion tag for improving the recombinant protein expression.

Keywords: Signal peptide, Cry1Ia, Fusion tag, Fluorescent proteins, Fluorescent intensity, Expression level

*Correspondence: gaojh_edu@163.com; jhgao@sxau.edu.cn;

wxingchun@163.com; wxingchun@sxau.edu.cn

¹ College of Life Sciences, Shanxi Agricultural University, Taigu 030801, China

Full list of author information is available at the end of the article

Introduction

GFP is the first fluorescent protein (FP) discovered in *Aequorea victoria* and has been widely used as a molecular reporter (reviewed in [1–3]). However, in bacteria



expression systems, such as *Escherichia coli*, GFP cannot fluoresce efficiently due to poor three-dimensional (3D) structure formation [4–6]. When fusing with other proteins of interest, GFP may even interfere with the folding of its partner in bacteria [7]. The phenomenon was attributed to the high contact order topology of GFP and the lack of more advanced co-translational folding mechanisms in bacteria [7, 8]. Although the chromophore of GFP contains only three contiguous amino acids (Ser65-Tyr66-Gly67), its formation depends on the native β -barrel structure of the whole molecule [1, 9]. The final β -barrel structure of GFP needs contacts among distant residues in the primary sequence, leading to the low folding rates. Thus, when highly expressed in prokaryotic cells, the GFP peptides aggregate in the crowded cytosol before they can correctly fold.

Strategies have been developed to improve the folding efficiency and fluorescing of GFP in bacteria. Point mutations were showed to improve the fluorescence and/or the folding of GFP. Out of these reported mutants, the enhanced GFP (eGFP) [10] and the superfolder GFP (sfGFP) [11] are widely used. Besides the modification of the GFP molecule per se, general methods to improve the expression of recombinant proteins have also been used [12–17]. For instance, the productive folding of N-terminal partner in fusion proteins with GFP could prevent the aggregation of the whole proteins [18]. The popular solubility enhancer partners maltose-binding protein (MBP) and N-utilization substance (NusA) can significantly enhance the solubility of GFP at low temperature [19], but the positive effects disappeared at high temperature [7]. Recently, several novel fusion tags, including Fh8 (an 8-kDa calcium-binding protein extracted from *Fasciola hepatica*), LEA-like peptide (a hydrophilic late embryogenesis abundant peptide), SKIK (Ser-Lys-Ile-Lys) peptide, OsTDX (*Oryza sativa* tetratricopeptide domain-containing thioredoxin), S1v1 (a self-assembling amphipathic peptide) and NT11 (the first 11 amino acid residues of a duplicated carbonic anhydrase), were reported can improve the expression of fluorescent proteins in bacteria [17, 20–24]. Chaperones, such as the group containing the DnaK and co-chaperones DnaJ and GrpE (K/J/E), can help the folding of the corresponding aggregation-prone partner and sequentially enhance the fluorescence of GFP in fusion protein [25]. A newly identified molecular chaperone Spheroplast Protein Y (Spy) can be used as a remarkable solubility enhancing fusion tag and its tandem pattern enable it to enhance the soluble expression of large-size proteins [26]. Codon usage and mRNA-structure also affect the fluorescing of GFP [27]. Siller et al. [28] used mutant ribosomes to reduce polypeptide elongation rates, thus facilitated the folding efficiency of recombinant proteins, especially from eukaryotes. This

generalized reduction strategy did not adversely impact the folding of the endogenous bacterial proteome.

Generally, the signal peptides (SPs) or their mutants are used to guide the recombinant proteins into periplasmic space or out of the cell for improving the structure formation or the expression level (reviewed in [12–14]). But the GFP protein cannot fold into its fluorescent form upon translocation via the SecYEG-translocon due to the post translocation folding mechanism [15, 29, 30]. Unlike the SecYEG-translocon, the twin-arginine translocation system (Tat system) only guides the fully folded and/or co-factor incorporated proteins translocation [31, 32]. As expected, the fluorescent signal of GFP molecules translocated into periplasmic space by the signal peptide of torA (trimethylamine N-oxide reductase) was observed [33]. The signal peptide of torA was also reported to be an inclusion body tag that can aggregate the recombinant proteins in *E. coli* cytosol [34]. It was not only a novel application of SPs, but also prompted us to pay more attention to the diversity of SPs. In this study, we used a novel fusion tag, the signal peptide of CryIIa (Iasp), to improve the expression of GFP in *E. coli* and *Bacillus thuringiensis* (Bt).

CryII toxins of Bt belong to a special branch of the Cry protein family (reviewed in [35, 36]). The *cryII* genes generally locate at approximately 500 bp downstream of the *cryIA* genes. Since there is not a classic promoter structure, the regulation of *cryII* genes was a mystery until their transcripts were identified by Tounsi and Jaoua [37–39]. Another special characteristic of CryII protein is the loss of the classic long C-terminal peptide of CryI toxins, which plays an important role in folding the N-terminal three domains into parasporal crystal. The absence of the C-terminal part in CryII proteins suggested a different expression pattern. In fact, Kostichka et al. [40] identified CryIIa4 protein in supernatant fluids of Bt AB88 cultures and predicted its secretion signal peptide (N-terminal 45 amino acids) which can be removed by unknown peptidase(s) after translocation. Although the CryIIa was showed to be a secretory protein, some of the expressed products were observed in the cell pellets of the natural host. The results suggested that the translocation capacity of the signal peptide of CryIIa was not high or the corresponding secretory pathway was easily saturated by the overexpressed targets [41, 42]. At present, the related secretory pathway of CryIIa is still unclear, but it could be excluded from the Tat system due to lack of the classic conserved sequence (S/T-RRXFLK), especially the double-arginine motif, in signal sequence [43–45].

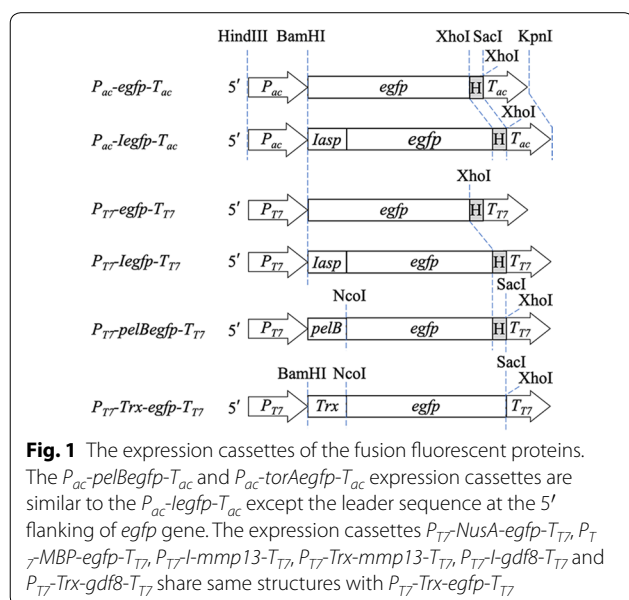
The effects of the Iasp on the eGFP and mCherry proteins in *E. coli* and Bt cells were evaluated. The results showed that the N-terminal location of Iasp in the fusion fluorescent proteins improved the expression of eGFP

and mCherry and consequently enhanced the fluorescent intensity of host cells. The produced IeGFP and ImCherry can be used as ideal reporters in prokaryotic cells. Moreover, the production yields of other difficult-to-express recombinant proteins, such as matrix metalloprotease-13 (MMP13) and myostatin (growth differentiating factor-8, GDF8) in *E. coli* BL21-star (DE3) strain [19], were dramatically enhanced after fusing with the Iasp. These results indicated that the Iasp can be used as a novel fusion tag for enhancing recombinant protein expression.

Results

Fluorescent intensity detection of IeGFP in *E. coli*

The expression cassettes and the corresponding plasmids and strains used in this study were showed in Fig. 1 and Table 1. As a constitutive promoter in *E. coli*, the promoter of *cryIAC* gene (P_{ac}) can regulate the expression of its downstream gene uninterruptedly after inoculation [46]. As expected, the expression of both IeGFP and eGFP was detected at 4 h after inoculation by immunoblot analysis (Fig. 2a). The expression products cannot be identified easily by Coomassie bright blue staining SDS-PAGE analysis (data not shown). The migration of intact IeGFP proteins was consistent with the calculated molecular weight (33.1 kDa). Interestingly, only a small quantity of IeGFP proteins were cleaved to give the same molecular weight product (27.9 kDa) as eGFP. The more intense color of cell pellets of strain TAc-IeGFP clearly



distinguished them from pellets of strain TAc-eGFP (Fig. 2b).

The fluorescent signals of TAc-IeGFP and TAc-eGFP strains were also monitored. An increasing trend of the fluorescent intensities for both TAc-IeGFP and TAc-eGFP was observed over time (Fig. 2c and Additional file 1: Table S1). Importantly, the fluorescent intensities of TAc-IeGFP strain were approximately 1.5, 2.4, 2.1 or 2.0-fold higher than TAc-eGFP strain at 8, 10, 12 or 24 h after inoculation, respectively.

Fluorescent intensity detection of IeGFP in Bt strain

The expression and resulting fluorescent signals of IeGFP and eGFP in Bt were investigated. In Bt cells, the expression of IeGFP controlled by P_{ac} were detected at 9 h after inoculation in the collected cells (Fig. 3a and b). During 12 to 36 h, the degradation of IeGFP protein was observed. The main degradation product was an approximately 26 kDa peptide (Fig. 3a, b and Additional file 1: Figure S1A). From 48 to 72 h, the intact band of IeGFP protein accumulated continuously and almost half of them were cleaved into the 26 kDa band. The eGFP protein shown a similar expression pattern with IeGFP (Fig. 3b). The truncated peptide of eGFP was also about 26 kDa, the same as the degradation product of IeGFP in vivo. IeGFP protein could not be detected in the supernatant of cell culture of BAc-IeGFP strain until 60 h after inoculation. Surprisingly, eGFP protein also appeared in the supernatant of the BAc-eGFP strain at the same time. The main immune signal in the supernatant samples were also about 26 kDa which was very similar to the truncated IeGFP and eGFP inside the cell.

With increasing incubation time, the fluorescent intensity of the BAc-IeGFP and BAc-eGFP cell increased during the first 48 h, consistent with the immunoblot analysis (Fig. 3a–c and Additional file 1: Table S2). Thereafter, whilst the protein continued to accumulate, it was accompanied by a decline in fluorescent intensity. It was notable that the fluorescent intensities of BAc-IeGFP kept more than three times higher than that of BAc-eGFP throughout the detection period (Fig. 3c and Additional file 1: Table S2). Interestingly, the fluorescent signals of IeGFP or eGFP in the supernatant were detected at same time after 48 h of inoculation at high sensitivity (Fig. 3d and Additional file 1: Table S3). Similarly, the enhanced fluorescent intensity of the BAc-IeGFP strain was readily observable in the altered pigmentation of cell pellets (Fig. 3e).

The expression pattern of IeGFP and eGFP in *E. coli* TG1 and BL21-star (DE3) strain, and Bt BMB171 strain were compared by western blot analysis (Additional file 1: Figure S1A). The result revealed the stability of eGFP expressed either in *E. coli* TG1 or in BL21-star (DE3) strains. The *pelB* signal peptide can guide almost all of eGFP (*pelB*-eGFP) to the periplasmic space of *E.*

Table 1 Plasmids and strains used and constructed in this work

Plasmid	Characteristics	Source
pHT304	<i>Amp^r</i> , <i>Erm^r</i> , <i>Bacillus-E. coli</i> shuttle vector	[64]
p304ΔSacl	pHT304 without Sacl recognition site	This work
pAc-leGFP	p304ΔSacl, <i>P_{ac}-legfp-T_{ac}</i> expression cassette	This work
pAc-eGFP	p304ΔSacl, <i>P_{ac}-egfp-T_{ac}</i> expression cassette	This work
pAc-pelBeGFP	p304ΔSacl, <i>P_{ac}-pelBeGfp-T_{ac}</i> expression cassette	This work
pAc-torAeGFP	p304ΔSacl, <i>P_{ac}-torAegfp-T_{ac}</i> expression cassette	This work
pAc-eGFPI	p304ΔSacl, <i>P_{ac}-egfpl-T_{ac}</i> expression cassette	This work
pl-leGFP	p304ΔSacl, <i>P_l-legfp-T_{ac}</i> expression cassette	This work
pl-eGFP	p304ΔSacl, <i>P_l-egfp-T_{ac}</i> expression cassette	This work
pAc-lmCherry	p304ΔSacl, <i>P_{ac}-lmCherry-T_{ac}</i> expression cassette	This work
pAc-mCherry	p304ΔSacl, <i>P_{ac}-mCherry-T_{ac}</i> expression cassette	This work
pET28aDel	<i>Kan^r</i> , pET28a without the sequence between RBS and BamHI	[61]
pET28aDel2	<i>Kan^r</i> , add the GATATA at the 5' of BamHI site of pET28aDel	This work
p28aD-leGFP	pET28aDel, <i>P_{T7}-legfp-T_{T7}</i> expression cassette	This work
p28aD-eGFP	pET28aDel, <i>P_{T7}-egfp-T_{T7}</i> expression cassette	This work
p22b-eGFP	pET22b, <i>P_{T7}-pelBeGfp-T_{T7}</i> expression cassette	This work
pCMV-N-mCherry	<i>mCherry</i> gene	Beyotime bio-technology, China
p28aD2-T-eGFP	pET28aDel2, <i>P_{T7}-Trx-egfp-T_{T7}</i> expression cassette	This work
p28aD2-N-eGFP	pET28aDel2, <i>P_{T7}-NusA-egfp-T_{T7}</i> expression cassette	This work
p28aD2-M-eGFP	pET28aDel2, <i>P_{T7}-MBP-egfp-T_{T7}</i> expression cassette	This work
p28aD2-MMP13	pET28aDel2, <i>P_{T7}-mmp13-T_{T7}</i> expression cassette	This work
p28aD2-GDF8	pET28aDel2, <i>P_{T7}-gdf8-T_{T7}</i> expression cassette	This work
p28aD2-l-MMP13	pET28aDel2, <i>P_{T7}-l-mmp13-T_{T7}</i> expression cassette	This work
p28aD2-T-MMP13	pET28aDel2, <i>P_{T7}-Trx-mmp13-T_{T7}</i> expression cassette	This work
p28aD2-l-GDF8	pET28aDel2, <i>P_{T7}-l-gdf8-T_{T7}</i> expression cassette	This work
p28aD2-T-GDF8	pET28aDel2, <i>P_{T7}-Trx-gdf8-T_{T7}</i> expression cassette	This work
p28aD-Cry11a	pET28aDel, <i>P_{T7}-cry11a-T_{T7}</i> expression cassette	This work
p28aD-Cry11aD44	pET28aDel, <i>P_{T7}-cry11aD44-T_{T7}</i> expression cassette	This work
p28aD-Cry11aD73	pET28aDel, <i>P_{T7}-cry11aD73-T_{T7}</i> expression cassette	This work
pAc-Cry11a	p304ΔSacl, <i>P_{ac}-cry11a-T_{ac}</i> expression cassette	This work
pAc-Cry11aD44	p304ΔSacl, <i>P_{ac}-cry11aD44-T_{ac}</i> expression cassette	This work
pAc-Cry11aD73	p304ΔSacl, <i>P_{ac}-cry11aD73-T_{ac}</i> expression cassette	This work
Strain	Genotype	
<i>E. coli</i>		
TG1	<i>K12Δ(lac-pro)</i> , <i>supE</i> , <i>thi</i> , <i>hsdD5</i> , <i>F⁻ [traD36 proAB⁺ lac^r lacZΔM15]</i>	[65]
BL21-star (DE3)	<i>F⁻ ompT hsdS_B (r_B⁻, m_B⁻) galdcmrne131</i> (DE3)	Invitrogen
T304	TG1 transformed with pHT304	This work
TAc-leGFP	TG1 transformed with pAc-leGFP	This work
TAc-eGFP	TG1 transformed with pAc-eGFP	This work
TAc-pelBeGFP	TG1 transformed with pAc-pelBeGFP	This work
TAc-torAeGFP	TG1 transformed with pAc-torAeGFP	This work
TAc-eGFPI	TG1 transformed with pAc-eGFPI	This work
TAc-lmCherry	TG1 transformed with pAc-lmCherry	This work
TAc-mCherry	TG1 transformed with pAc-mCherry	This work
BL28aD	BL21-star (DE3) transformed with pET28aDel	This work
BL28-leGFP	BL21-star (DE3) transformed with p28aD-leGFP	This work
BL28-eGFP	BL21-star (DE3) transformed with p28aD-eGFP	This work

Table 1 (continued)

Strain	Genotype	
BL22b	BL21-star (DE3) transformed with pET22b	This work
BL22-eGFP	BL21-star (DE3) transformed with p22b-eGFP	This work
BL28D2-T-eGFP	BL21-star (DE3) transformed with p28aD2-T-eGFP	This work
BL28D2-N-eGFP	BL21-star (DE3) transformed with p28aD2-N-eGFP	This work
BL28D2-M-eGFP	BL21-star (DE3) transformed with p28aD2-M-eGFP	This work
BL28D2-MMP13	BL21-star (DE3) transformed with p28aD2-MMP13	This work
BL28D2-GDF8	BL21-star (DE3) transformed with p28aD2-GDF8	This work
BL28D2-I-MMP13	BL21-star (DE3) transformed with p28aD2-I-MMP13	This work
BL28D2-T-MMP13	BL21-star (DE3) transformed with p28aD2-T-MMP13	This work
BL28D2-I-GDF8	BL21-star (DE3) transformed with p28aD2-I-GDF8	This work
BL28D2-T-GDF8	BL21-star (DE3) transformed with p28aD2-T-GDF8	This work
BL28-Cry1Ia	BL21-star (DE3) transformed with p28aD-Cry1Ia	This work
BL28-Cry1IaD44	BL21-star (DE3) transformed with p28aD-Cry1Ia44	This work
BL28-Cry1IaD73	BL21-star (DE3) transformed with p28aD-Cry1Ia73	This work
<i>B. thuringiensis</i>		
BMB171	acrystalliferous mutant strain of YBT-1463	[66]
B304	BMB171 transformed with pHT304	This work
BAc-IeGFP	BMB171 transformed with pAc-IeGFP	This work
BAc-eGFP	BMB171 transformed with pAc-eGFP	This work
BI-IeGFP	BMB171 transformed with pl-IeGFP	This work
BI-eGFP	BMB171 transformed with pl-eGFP	This work
BAc-ImCherry	BMB171 transformed with pAc-ImCherry	This work
BAc-mCherry	BMB171 transformed with pAc-mCherry	This work
BAc-Cry1Ia	BMB171 transformed with pAc-Cry1Ia	This work
BAc-Cry1IaD44	BMB171 transformed with pAc-Cry1IaD44	This work
BAc-Cry1IaD73	BMB171 transformed with pAc-Cry1IaD73	This work

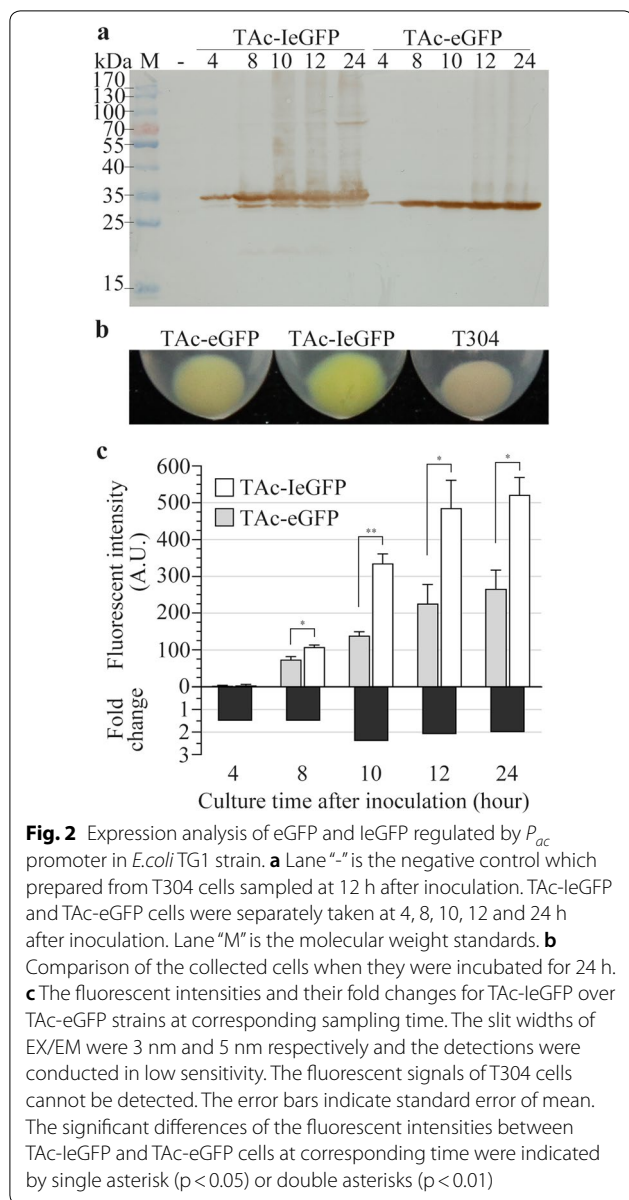
coli and then the leader peptide was removed. However, most of IeGFP proteins remained intact and only a fraction of them were processed, especially when its expression was regulated by the *T7* promoter. In Bt strain, both eGFP and IeGFP protein were cut to a slightly lower molecular weight band (26 kDa) in cells or in supernatant of cell culture.

Iasp enhanced the expression level of recombinant proteins

To confirm the effect of Iasp on improving the expression of recombinant proteins, the eGFP and IeGFP was expressed in *E. coli* BL21-star (DE3) strain, respectively. As a result, the *T7* promoter regulated the expression of IeGFP at higher level than eGFP at different temperatures, especially at 16 °C (Fig. 4). The expression level of eGFP was very low after 4 h induction at 16 °C, but the product can keep soluble (Additional file 1: Figure S2A). At the same condition, about half of the IeGFP products kept soluble. Nevertheless, the amount of this soluble fraction was higher than the total production of eGFP. It was notable that the expression level of

IeGFP was similar to the NusA and MBP guided eGFP (N-eGFP and M-eGFP) and slightly higher than the Trx (thioredoxin) guided eGFP (T-eGFP). All of the fusion tags tested can produce considerable fraction of soluble fluorescent proteins in this study. For GDF8 and MMP13 proteins, their expressions were hard to identified by Coomassie brilliant blue stained SDS-PAGE analysis. However, with the assistance of Iasp or Trx, the expression levels were dramatically enhanced (I-MMP13, T-MMP13, I-GDF8 and T-GDF8) in *E. coli* BL21-star (DE3) strain (Fig. 4b and c). The difference was that almost all of I-MMP13, T-MMP13, I-GDF8 and T-GDF8 proteins were expressed as the inclusion bodies at the same induction condition (Additional file 1: Figure S2B and C).

The effect of the Iasp on its original protein Cry1Ia was also investigated. In Bt cells, the accumulation of the Cry1Ia (81.2 kDa) regulated by the *P_{ac}* promoter at 72 h after inoculation can be detected by SDS-PAGE (Additional file 1: Figure S3A). When removed the Iasp from the Cry1Ia, more accumulation of the truncated protein (Cry1IaD44, 76.2 kDa) was observed in Bt cells.



Interestingly, Cry1IaD44 showed different degradation pattern compared to the intact Cry1Ia. The main degradation products of Cry1IaD44 (14 kDa) were soluble in the alkali buffer (50 mM Na_2CO_3 , 10 mM dithiothreitol, pH 10.5) (Additional file 1: Figure S3B). Similar degraded fragment was observed in the expression product of a further truncated variant Cry1IaD73 (73.3 kDa), losing the N-terminal 73 amino acids of Cry1Ia. The result suggested the Iasp or its associated chaperone can prevent the attacking of specific protease(s). Interestingly, the different processing pattern observed in Bt cells was not detected in *E. coli* strain (Additional file 1: Figure S3C).

Effect of Iasp encoding sequence on the transcript level of *legfp* gene

To identify the effect of leGFP encoding sequence on the fusion gene, the transcription levels of the *Iegfp* and *egfp* genes in *E. coli* TAc-leGFP and TAc-eGFP strains were compared using real-time quantitative PCR. As shown in Table 2, the transcription level of *Iegfp* gene in TAc-leGFP strain was approximately 7 or 13 times higher than *egfp* gene in TAc-eGFP at 8 h or 12 h after inoculation. But interestingly, the wide difference in mRNA level between the *Iegfp* and *egfp* genes only resulted in the double protein level indicated by fluorescent intensity in Fig. 1. In addition, the structure of the 5' leader sequence of the target mRNA plays an important role in affecting the translational efficiency, especially the nucleotides near the start codon (-4 to +37) [27]. We predicted the secondary structures of *Iegfp* and *egfp* genes by RNA-structure software [47]. The result showed that the calculated folding energy of the nucleotides near the start codon (-4 to +37) of *Iegfp* was about -0.6 kcal/mol, which implied less thermodynamic stability in this section than that of *egfp* gene (-5 kcal/mol, Fig. 5).

Comparison the Iasp with the pelB and torA signal peptide

The expression of leGFP in *E. coli* was compared with the signal peptide of pelB (pectate lyase B) or torA guided eGFP (pelB-eGFP and torA-eGFP). The signal peptides of pelB or torA can guide the target protein into periplasm of *E. coli* by two representative inner membrane spanning transporters (SecYEG translocon or the twin-arginine translocation system) respectively [31]. Another fusion protein eGFP-I was also constructed to identify the function of Iasp when located at the C-terminal of eGFP. The expression of these proteins was monitored at five time points (4, 8, 10, 12 and 24 h after inoculation). The western blot analysis showed that all of these proteins can expressed under regulation of the constitutive promoter P_{ac} (Fig. 6a). The expression level of leGFP after 10 h of incubation was higher than pelB-eGFP and torA-eGFP. But when the Iasp was located at C-terminal of eGFP (eGFP-I), the expression level decreased. Interestingly, the observed band of eGFP-I was larger than the expected molecular weight (33.1 kDa). Both of the signal peptides pelB and torA did not improve the expression level of eGFP but might translocate eGFP efficiently because the observed bands of pelB-eGFP and torA-eGFP were equal to the eGFP in molecular weight. The result indicated that both signal peptides were cleaved by corresponding peptidase after translocation. The growth and the fluorescent intensities of these strains were also monitored and the results were consistent

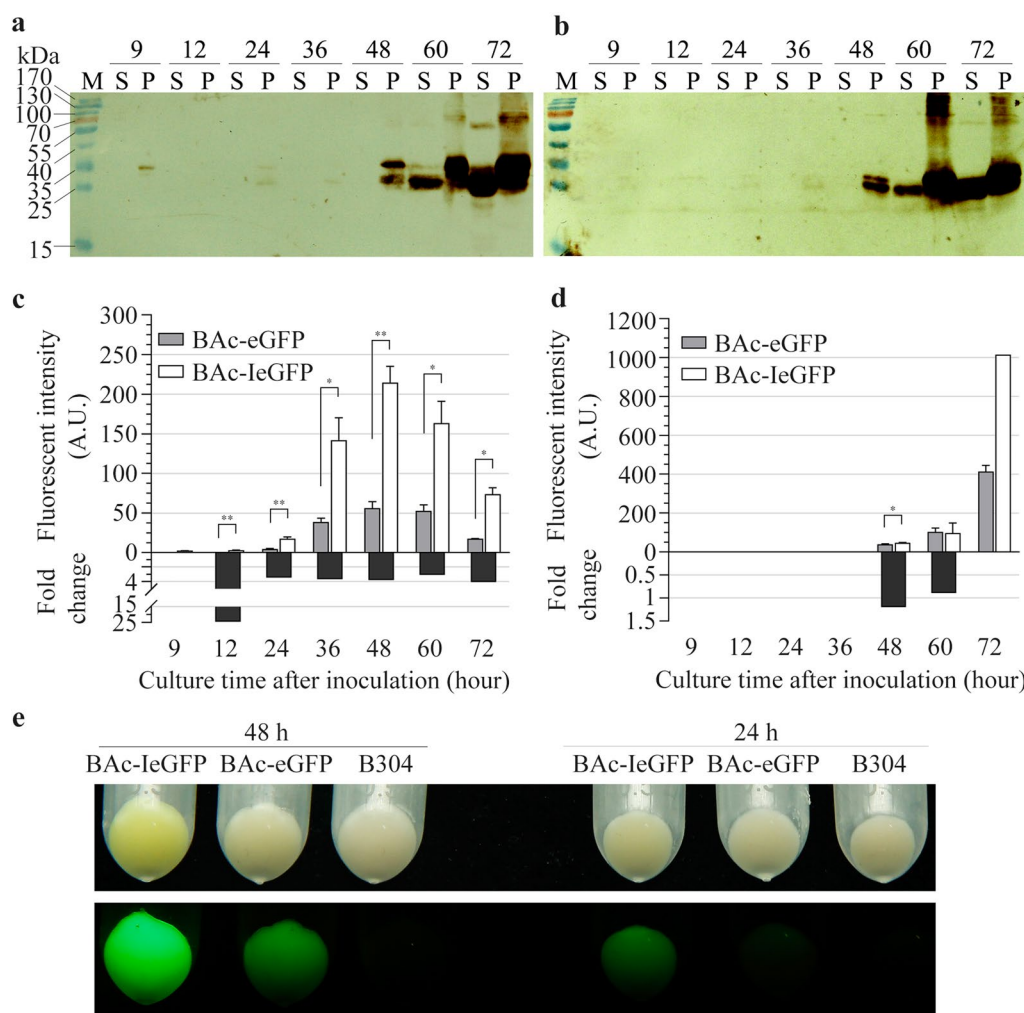
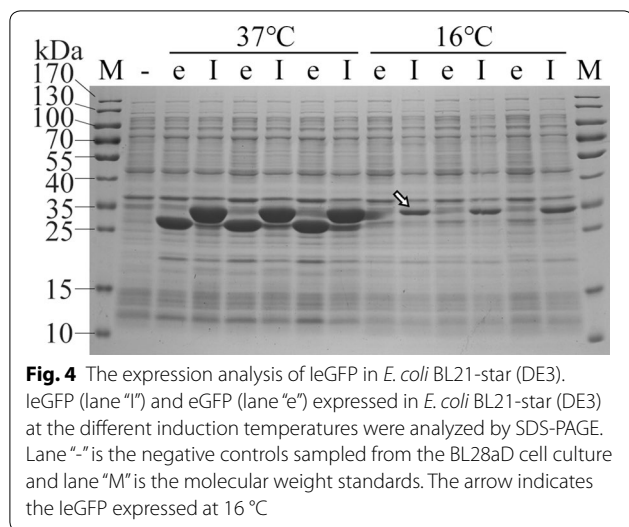


Fig. 3 Expression analysis of eGFP and leGFP regulated by the P_{ac} promoter in Bt strain. **a, b** The expression of leGFP (panel **a**) and eGFP (panel **b**) in Bt at different times (9, 12, 24, 36, 48, 60 and 72 h after inoculation) were analyzed respectively by western blot. "S" represents the supernatant of cell culture and "P" represents the resuspended cells by PBS buffer. Lane "M" is the molecular weight standards. **c, d** The fluorescent intensities of the harvested cells (panel **c**) or the supernatants of cell cultures (panel **d**) and their fold changes for BAc-leGFP over BAc-eGFP strain at corresponding culture time. The slit widths of EX/EM were both 3 nm and the detections were conducted in low sensitivity for the resuspended cells and in high sensitivity for the supernatant. The fluorescent signals of B304 cells cannot be detected and the fluorescent intensity of the 72 h supernatant of BAc-leGFP strain cell culture was beyond the limit (1000 A.U.). The error bars indicate standard error of mean. The significant differences of the fluorescent intensities between BAc-leGFP and BAc-eGFP strains at corresponding time were indicated by single asterisk ($p < 0.05$) or double asterisks ($p < 0.01$). **e** Comparison of the collected cells before (up) and after excitement (down) when they were incubated for 24 h and 48 h. The cells were excited by blue light using Luyor 3415RG hand held lamp

with the immunoblot analysis except the strains expressing pelB-eGFP and torA-eGFP respectively (Fig. 6b, c and Additional file 1: Table S4). The expression of the pelB-eGFP negatively affected the growth of the *E. coli* cells and the expression products did not fluoresce normally. The expression level of torA-eGFP was lower than pelB-eGFP observed by western blot analysis, but its fluorescent signal was detected after 12 h of incubation.

Subcellular location of leGFP in *E. coli*

According to the confocal images, the fluorescent signals were observed evenly in the cytoplasm for cells expressing eGFP, leGFP and eGFP-I (Fig. 7). The fluorescent signal of pelB-eGFP, by contrast, was not detected in the host cells. For cells expressing torA-eGFP, the polar localization was observed.



Verification of the *cry11a* gene promoter regulation using leGFP

To confirm the better performance of leGFP, the *legfp* and *gfp* genes were placed under control of the *cry11a* promoter (P_i) and expressed in the Bt strain. The expression patterns of both proteins were identified and compared by the detection of fluorescent intensity. The result showed that the P_i promoter regulated the *legfp* gene expression in a similar way to that of P_{ac} over time (Fig. 8 and Additional file 1: Tables S5 and S6). But the expression level of P_i regulated genes was lower because the fluorescent signal of leGFP and eGFP in BI-leGFP and BI-eGFP cells was only detected on the 3/3 nm (the slit widths of EX/EM) in high sensitivity. According to the specification, the sensitivity of the spectrofluorophotometer (RF 5301PC) at high mode is about 50 times in comparison to that at low mode. In fact, the color of the cell pellets of BI-leGFP and BI-eGFP were difficult

to distinguish from the negative control cells by visible observation (B304 strain, data not shown).

Verification of the function of lasp on mCherry protein

To identify the effect on other FPs in vivo, the lasp was used to guide the expression of mCherry. The new fusion protein ImCherry was also expressed in *E. coli* TG1 strain (TAc-ImCherry) and the Bt BMB171 strain (BAc-ImCherry), respectively, under control of P_{ac} promoter. Similar results to those with leGFP were observed. The fluorescent signals of the TAc-ImCherry strain were more intensive than the corresponding control strain TAc-mCherry (Fig. 9a and Additional file 1: Table S7). The 24 h-after-inoculation cells of TAc-ImCherry strain were distinguished easily with the TAc-mCherry strain (Fig. 9b). Most of ImCherry proteins were expressed in the intact molecules (33.0 kDa) and only a fraction of them were digested into a 28 kDa band which was equal to the individually expressed mCherry protein (27.8 kDa, Fig. 9c).

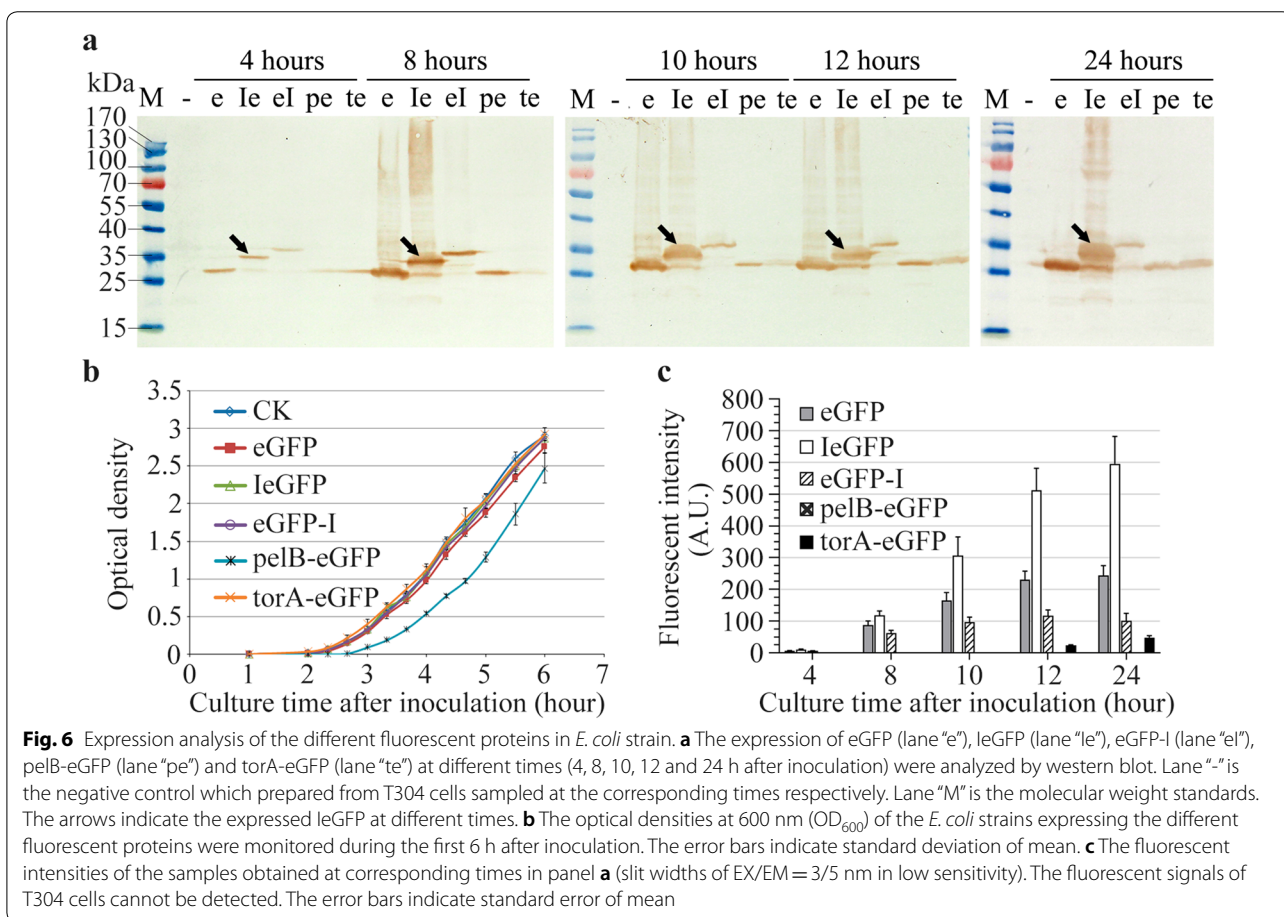
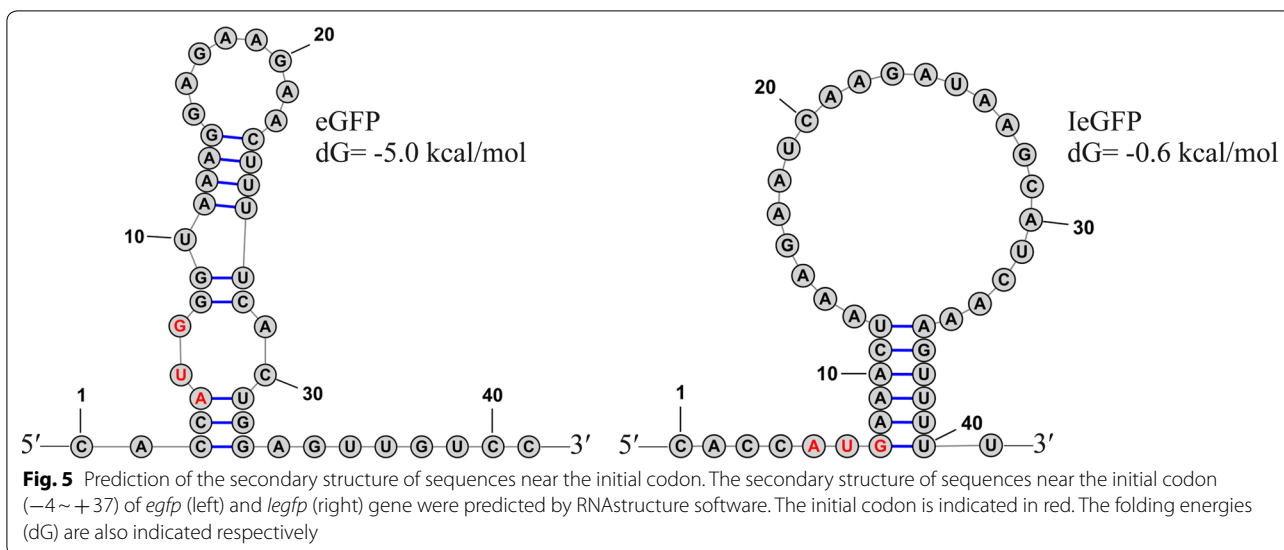
In Bt cells, the fluorescent signal of mCherry cannot be detected during the cultivation period tested whereas the ImCherry performed well (Fig. 9c and e, and Additional file 1: Tables S8 and S9). The significant accumulation of ImCherry protein was also observed after 60 h of incubation, consistent with the fluorescent signal variation (Fig. 9f). The expression of the individual mCherry cannot be detected by western blot analysis even in 72 h-after-inoculation cells (data not shown). These results implied lower expression levels of the mCherry and ImCherry than the eGFP and leGFP which might be attributed to the non-optimal codon used in the *mCherry* gene. Due to lower expression level in Bt strain, only weak violet color was observed for the BAc-ImCherry strain after 60 h inoculation (data not shown).

Interestingly, the immune signals of ImCherry were also detected in the supernatant of cell culture after 60 h

Table 2 Comparison of *legfp* and *egfp* gene in mRNA level

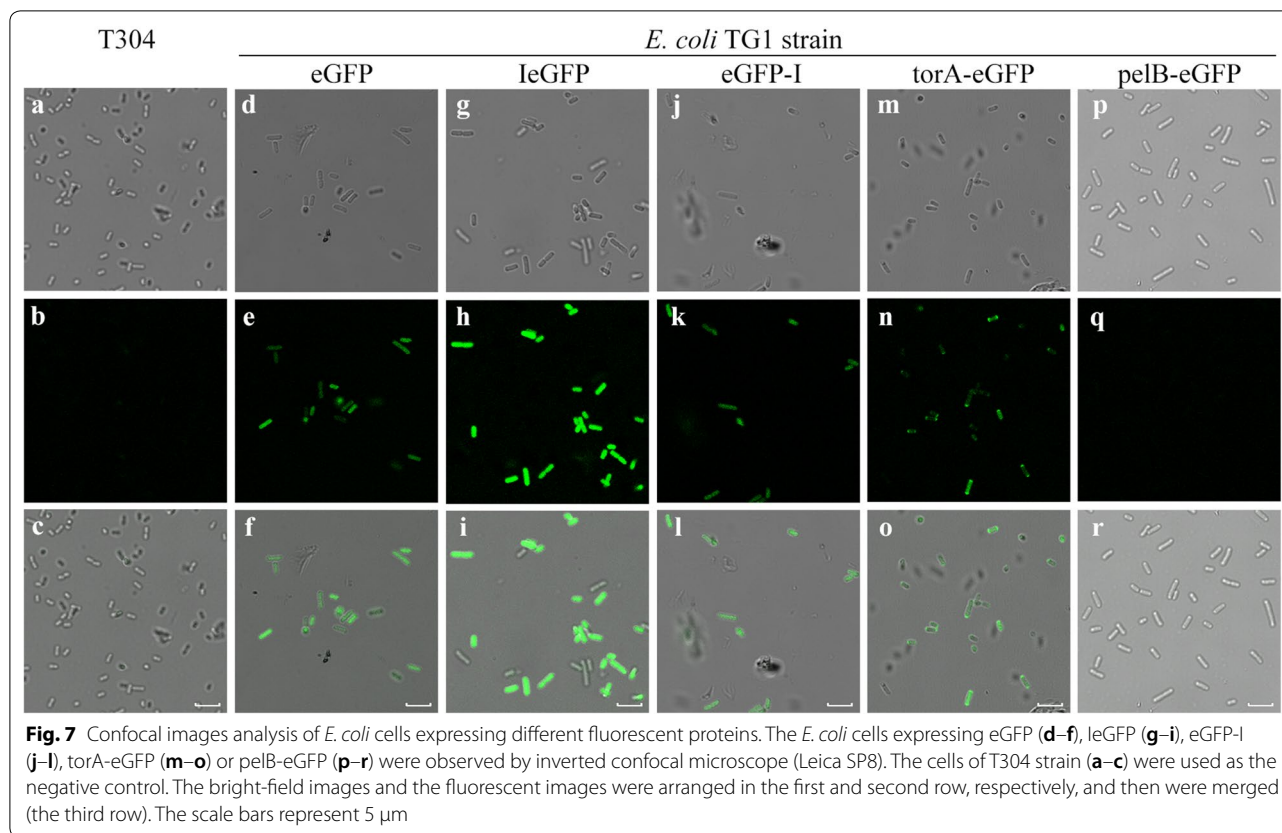
No.	TAc-eGFP				TAc-leGFP				Fold change
	Mean Ct of <i>hcat</i> gene	Mean Ct of <i>egfp</i> gene	Mean of $2^{-\Delta Ct}$	Variable coefficient (%)	Mean Ct of <i>hcat</i> gene	Mean Ct of <i>legfp</i> gene	Mean of $2^{-\Delta Ct}$	Variable coefficient (%)	
8 h									
1	26.91 ± 0.159	13.82 ± 0.131	989.466 ± 1018.333	10.289	30.18 ± 0.162	14.11 ± 0.118	75,603.881 ± 9462.533	12.516	7.639
2	27.12 ± 0.024	13.78 ± 0.200			30.42 ± 0.031	14.30 ± 0.153			
3	27.40 ± 0.076	14.03 ± 0.476			27.94 ± 0.033	11.54 ± 0.292			
12 h									
1	26.59 ± 0.136	14.73 ± 0.203	2945.768 ± 980.636	33.290	27.29 ± 0.088	11.86 ± 0.079	38,480.124 ± 14,483.546	37.639	13.639
2	27.21 ± 0.111	15.53 ± 0.353			26.68 ± 0.128	12.25 ± 0.120			
3	26.30 ± 0.058	15.46 ± 0.199			25.55 ± 0.086	9.97 ± 0.065			

The precision of the measures was indicated by the standard deviations



of incubation and the corresponding bands were lower than the degraded products inside the cells (Fig. 9f and

Additional file 1: Figure S1B). The band detected in the supernatant of cell culture of BAc-ImCherry strain was



equal to the mCherry expressed in *E. coli* TG1 strain (Additional file 1: Figure S1B).

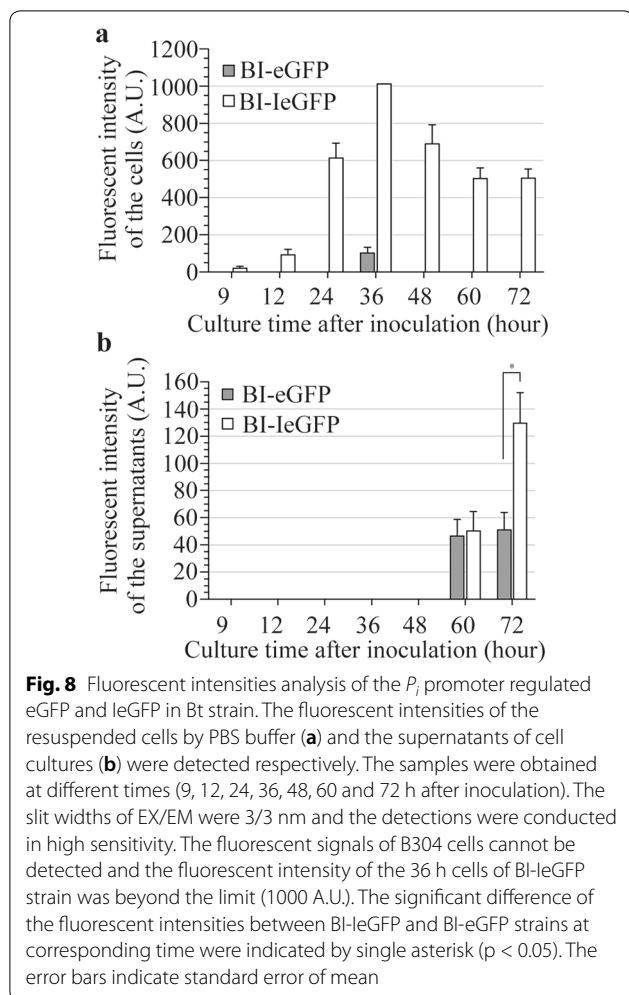
Discussion

In this study, the predicted secretion signal peptide of Cry1Ia (Iasp) was fused to the N-terminal of eGFP and mCherry to produce novel fluorescent proteins, IeGFP and ImCherry. These fusion fluorescent proteins were not constructed for translocation across the membrane after expression, but to identify the positive effect of the Iasp on protein expression in the cytoplasm of prokaryotic cells.

Iasp can improve the expression level of eGFP and sequentially enhance the fluorescent signal of the tested prokaryotic cells. Under regulation of P_{ac} promoter, the fluorescent intensity of TAc-IeGFP strain was significantly higher than that of TAc-eGFP strain and so did the corresponding Bt strains. In *E. coli* cells tested, the expression of fluorescent proteins regulated by P_{ac} cannot be identified easily by SDS-PAGE analysis that suggested a weaker activity of P_{ac} than the $T7$ promoter. But it was sufficient to alter the pigmentation of the host cells. Similarly, Iasp could also enhance the expression level of mCherry in the tested bacteria and the fluorescent intensity of the host cells. These results indicated

that Iasp can improve the expression level of fluorescent proteins. The fusion fluorescent proteins, such as IeGFP and ImCherry constructed in this study, can accumulate inside host cells and therefore be used as novel fluorescent protein variants.

Iasp encoding sequence can affect the stability of mRNA and/or improve the translation efficiency. The observed high levels of transcripts of *Iegfp* gene were probably responsible for the enhanced expression level of IeGFP. In addition, Kudla et al. [27] reported that the structure of the 5' leader sequence of the target mRNA played an important role in influencing the translational efficiency, especially the nucleotides near the initial codon (−4 to +37). The authors fused a 28-codon leader sequence with low secondary structure into the 5' of different GFP encoding sequences. These fusion genes produced uniformly high expression levels. The strong correlation between mRNA folding energy and expression level suggests that the tightly folded messages obstruct translation initiation and thereby reduce protein synthesis [27, 48]. According to the prediction results of RNAstructure software [47], the participation of Iasp encoding sequence decreased the thermodynamic stability of the secondary structure near the initial codon (−4 to +37). Interestingly, in Kudla's report, the *gfp* gene



variants resulting high fluorescent intensities showed the similar folding energy level with the *egfp* construct used in this study (about -5.0 kcal/mol). This implicated that Iasp encoding sequence could optimize the secondary structure of this region further and sequentially improve the translational efficiency. The result was also consistent with the report of Boël G et al. [49]. By high-throughput protein-expression analysis, the authors found that several factors, including the folding of mRNA head and the codon usage near this region, affected the translation efficiency in *E. coli*. The codon usage of 5' coding sequence (CDS) does not only regulate the secondary structure of mRNA, but also influences the efficient ribosome docking. Therefore, Iasp can also be used as a fusion tag to improve the expression level of other recombinant proteins, such as two difficult-to-express proteins MMP13 and GDF8 tested in this study.

The interactions between the nascent peptide and the chaperones would alleviate the aggregation tendency of the highly expressed peptides and provide them more

opportunities or directly assist them in folding into the correct 3D structure. For instance, co-expression of some chaperones and co-chaperones can prevent misfolding of eukaryotic proteins in prokaryotic system [50, 51]. Zhang et al. [25] reported that the combination of DnaK, DnaJ and GrpE (K/J/E) helped the folding of N-terminal nascent peptide and improved the solubility of the GFP partner in fusion protein. Consequently, the fluorescence of the fusion protein was enhanced in *E. coli*. Another chaperones, such as the novel Spy [26], or chaperone-interacted partner proteins, such as the classic MBP and NusA [17], can be used in the fusion proteins to enhance their solubility. For secretory proteins, the signal peptides also interact with corresponding chaperones which not only transfer the nascent peptides to the translocation machinery but also prevent aggregation. It was proven that alternating signal peptides or proper modification improved the expression-secretion efficiency of recombinant proteins simultaneously [52, 53]. However, some the signal peptides utilizing the SecYEG translocon, which only translocate the unfolded peptides, were reported to negatively affect the stability and folding of target proteins. For instance, the signal peptide of MBP (malE) and pelB led the thermodynamic destabilization of the C-terminal mature MBP or thioredoxin (Trx) due to the hydrophobic interactions between the signal peptides and the unfolded peptides. The interactions were thought necessary to hold the unfolded state before translocation [54–57]. Although the related secretory pathway of the Cry11a is unclear and the components participated in the machinery still keep uncovered, Iasp would interact with the corresponding chaperone(s) in its secretory pathway. In this study, soluble expression of IeGFP regulated by $T7$ promoter was observed and the production yield of soluble fraction was higher than that of eGFP. The solubility enhancing effect of Iasp on eGFP was similar to the well-known fusion tags NusA, MBP and Trx. However, both Iasp or Trx fused MMP13 (I-MMP13 or T-MMP13) and GDF8 (I-GDF8 or T-GDF8) were expressed in insoluble form. The expression result of T-MMP13 and T-GDF8 was consistent with the previous report [19]. Therefore, the fitness of solubility enhancing effect of Iasp on diverse recombinant proteins need to be investigated further. Several hypotheses have been suggested to explain the mechanisms of solubility enhancement by fusion tags, including formation of micelle-like structures, attracting chaperones, exerting an intrinsic chaperon-like activity or net charged (reviewed in [17]). As a signal peptide, the interaction of Iasp with the corresponding chaperon(s) would be supposed and the deletion mutation on its original protein Cry11a might provide a clue. After removal of the Iasp, the truncated Cry11a proteins (Cry11aD44) accumulated inside the Bt cells and showed different

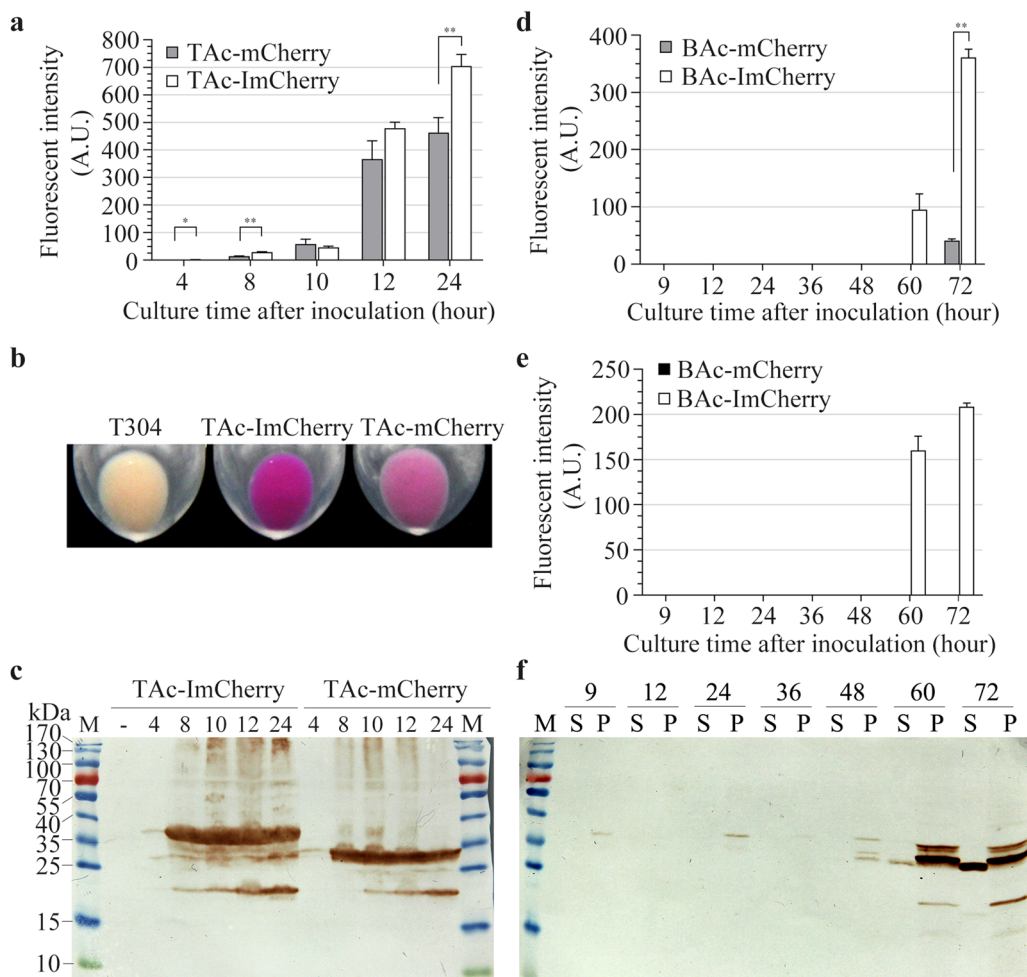


Fig. 9 Expression analysis of mCherry and ImCherry in *E. coli* TG1 and Bt BMB171 strain. **a** The fluorescent intensities of the harvested TG1 cells expressing mCherry or ImCherry were monitored at different times (4, 8, 10, 12 and 24 h after inoculation). The slit widths of EX/EM were 1.5 nm and 3 nm, and the detections were conducted in high sensitivity. The fluorescent signals of B304 cells cannot be detected. The error bars indicate standard error of mean. The significant differences of the fluorescent intensities between TAc-ImCherry and TAc-mCherry cells at corresponding time were indicated by single asterisk ($p < 0.05$) or double asterisks ($p < 0.01$). **b** Comparison of the collected *E. coli* TG1 cells when they were incubated for 24 h. **c** Western blot analysis of the expression of mCherry and ImCherry in *E. coli* TG1 strain. “M” lane is negative control which prepared from T304 cells sampled at 12 h after inoculation. TAcImCherry and TAc-mCherry cells were separately taken at 4, 8, 10, 12 and 24 h after inoculation. “M” represents the molecular weight standards. **d, e** The fluorescent intensities of the supernatants of cell cultures (panel D, slit widths of EX/EM = 3/5 nm in high sensitivity) or the harvested Bt cells (e slit widths of EX/EM = 3/3 nm in high sensitivity) at corresponding culture time. The fluorescent signals of B304 cells cannot be detected. The error bars indicate standard error of mean. The significant differences of the fluorescent intensities between BAc-ImCherry and BAc-mCherry cells at corresponding time were indicated by double asterisks ($p < 0.01$). **f** Western blot analysis of the expression of ImCherry in Bt BMB171 strain. “S” represents the supernatant of cell culture and the “P” means the resuspended cells by PBS buffer. “M” represents the molecular weight standards

degradation profile compared to the intact Cry1Ia. Interestingly, in *E. coli* BL21-star (DE3) strain, no visible difference between them were observed by SDS-PAGE analysis besides the expected bands of Cry1Ia, Cry1Ia44 and Cry1Ia73. The result might be attribute to the different protease constitution between *E. coli* and Bt cells, and the phenomenon implied the effect of Iasp on preventing

the Cry1Ia protein from attacking by some proteases directly or indirectly.

Iasp cannot translocate the eGFP efficiently in *E. coli*. By confocal microscopy, most of IeGFP proteins accumulated in the cytosol of *E. coli* cells that was consistent with the western blot analysis. According to the previous reports, the saturation of corresponding secretory pathway led to accumulation of recombinant proteins

in cytosol [41, 42]. Although a fraction of IeGFP molecules were detected as a smaller band that probably represented the molecules translocated into periplasm, no fluorescent signal in this location were captured by confocal microscopy. Similarly, the expression of pelB-eGFP was detected by western blot analysis but no fluorescent signal was observed either in the cytoplasm or periplasmic space. Western blot analysis showed that almost all of expressed pelB-eGFP proteins were digested into a 27 kDa band that was equal to the individually expressed eGFP. Therefore, we concluded the efficiently translocation of eGFP guided by pelB signal peptide regulated by P_{ac} . However, the high translocation efficiency of pelB-eGFP would overload the SecYEG-translocon and result in the growth retardation of TAc-pelBeGFP strain. Out of the known GFP variants, the sfGFP is the only one that can fold into its fluorescent form upon translocation via the SecYEG-translocon [15, 29, 30]. The post translocation folding mechanism was probably responsible for the observation for *E. coli* cells expressing pelB-eGFP. The Iasp would guide some eGFP peptide into periplasm via an analogous pathway besides that most of IeGFP peptides could not be captured by the corresponding chaperone(s) in their secretory pathway for unknown reasons, and then folded into the fluorescent form in the cytoplasm. The fluorescent signal of torA-eGFP can be detected albeit weak. The signal peptide of torA can translocate the target proteins into periplasm by Tat system. Unlike the SecYEG-translocon, the Tat system only guides the fully folded and/or co-factor incorporated proteins translocation [31, 32]. This different translocation mechanism explained the completely opposite fluorescing ability of pelB-eGFP and torA-eGFP in vivo. The polarization of fluorescence of torA-eGFP observed by confocal microscopy indicated the periplasm localization and the phenomenon resulted from the plasmolysis of *E. coli* cells in response to the osmotic up-shock when resuspended in PBS buffer [33]. In addition, when the Iasp was located at the C-terminal of eGFP (eGFP-I), its translocation function and effects on enhancing fluorescent signal were both lost.

In Bt cells, the IeGFP could not be translocated out of cells or in low efficiency because its immunoblot or fluorescent signal emerged at the same time point as the eGFP. The result indicated that the target proteins would not only be released by corresponding secretory pathway but also by the lysis of mother cells. Kostichka et al. [40] reported the Cry1Ia protein in natural host AB88 was not detected in pure crystal samples isolated from a 52 h culture but was detected in 26 h pellets of cells. Therefore, we speculated that some soluble Cry1Ia proteins stayed in the mother cells of Bt strain may release directly into supernatant during the spore maturation. The release

of the soluble components would also be one of factors leading to the decline of fluorescent intensity of the resuspended precipitates of the BAc-eGFP and BAc-IeGFP cell cultures during last 12 h, especially at the 72 h after inoculation. The dramatically increased fluorescent intensity of 72 h supernatant of both BAc-eGFP and BAc-IeGFP could also support the speculation. Moreover, during the sporulation period, the P_{ac} promoter can accelerate the synthesis of the recombinant proteins that would prompt the formation of insoluble inclusion bodies. In this study, the rapid accumulations of eGFP and IeGFP after 48 h incubation were indicated by the immunoblot analysis. Therefore, the formation of inclusion bodies would be the second factor leading to the inconsistency between the significant decreasing of fluorescent intensity and the rapid accumulation of the protein products. The third factor would be the pH fluctuation in Bt cells at the latter sporulation stage. In *Bacillus subtilis*, the fluorescent intensity of a fusion proteins β -GFP, including the β subunit of prokaryotic RNA polymerase located at the N-terminal of GFP, was dramatically weakened during the sporulation stage V-VI by the decreased pH inside mother cells [58]. The period represented the last time before the mature spore release. The mCherry is more tolerant to a drop in pH [58]. Nevertheless, the fluorescent intensity of the resuspended BAc-IeGFP cells kept 3.5- to 4.1-fold times to that of BAc-eGFP during the whole culture period except the first 9 h, which were higher than the differences between the TAc-IeGFP and TAc-eGFP (1.5- to 2.4-fold).

We also noticed the degradation of eGFP and IeGFP proteins in the Bt strains. In the cells at 9 or 12 h after incubation, the intact IeGFP or eGFP proteins were observed. But during the 12 to 36 h after inoculation, part of them were digested by unknown protease(s). After 48 h, the recombinant proteins accumulated rapidly and approximately half of them were processed by these protease(s). The genus *Bacillus* can produce diverse proteases and Bt is also an excellent source of proteases [59]. Generally, the overall protease activity of Bt strains follows a slow increase in the initial 12 h and then gradually increased to the peak at around 30 h after inoculation [59]. The protease activity variation would be partly responsible for the degradation process of the eGFP and IeGFP retained in Bt cells.

In supernatant of cell culture of corresponding Bt strains, the 26 kDa fragment, slightly smaller than the intact eGFP (27.9 kDa), was the main product. The N-terminal analysis revealed that the 26 kDa fragment started with GKGEELFT amino acids and therefore another cleavage site(s) would locate at the C-terminal of eGFP protein. The deficiency of the C-terminal of eGFP proteins would partially account for the weaker fluorescent

intensity of eGFP and IeGFP in supernatant compared to the collected cells. For example, with loss of the C-terminal 11 amino acids, the expressed superfolder GFP 1-10 was insoluble and cannot fluoresce [60]. In addition, the immune signals of the ImCherry protein in the supernatant of cell culture and inside Bt cells were different in molecular weight. The result proved that there were different proteases responding to the degradation of ImCherry and IeGFP inside and outside the Bt cell. The intact ImCherry proteins expressed by the prokaryotic cells tested was slightly larger than its calculated molecular weight (33.0 kDa). This may be attributed to the unknown post-translational modification(s).

At present, little attention has been paid to the use of SPs with low translocation efficiency for recombinant protein production. The *torA* signal sequence was proven to direct expression of recombinant proteins in inclusion bodies even if the target was a well-soluble protein in *E. coli* [34]. This small inclusion body tag can help to express the toxic or unstable recombinant proteins. But the signal peptide of *torA* and *Iasp* would belong to different secretory pathways because the signal sequences of Tat system share the classic conserved sequence (S/T-RRXFLK), especially the double-arginine motif, which is not existed in *Iasp* [43–45]. More importantly, *Iasp* can guide the expression of eGFP at higher level than *torA* signal peptide under control of P_{ac} promoter in this study.

In conclusion, we identified the capacity of the SP of Cry1Ia protein (*Iasp*) in improving the expression of eGFP and mCherry proteins in *E. coli* and *Bt* cells. The insertion of encoding sequence of *Iasp* optimized the secondary structure near the initial codon and then facilitated the docking of ribosome. The more efficient recruitment of ribosome would account for the stability of mRNA and the rapid synthesis of the fusion fluorescent proteins. Meanwhile, *Iasp* can keep a considerable fraction of the fusion fluorescent proteins soluble inside the cells that guarantee the correct folding of fluorescent proteins. Therefore, the higher production of the soluble IeGFP and ImCherry generated the more intense fluorescent signal inside bacterial cells and these fusion proteins could be used as ideal candidates of fluorescent protein variants. Moreover, the function of *Iasp* on improving the expression of the recombinant proteins should also be highlighted and worthy of further investigation. This study also hinted the versatility of some signal peptides.

Materials and methods

Bacterial strains and growth conditions

All plasmids and strains used in this work are detailed in Table 1. Unless otherwise noted, *E. coli* cells were incubated in Luria–Bertani medium (LB medium) with 50 µg/mL ampicillin for pMD19 and pHT304 derived

vectors, or 50 µg/mL kanamycin for pET28a derived vectors at 37 °C with 200 rpm shaking. Bt strains grew in LB medium (containing 25 µg/mL erythromycin if harboring the pHT304 plasmid or its derivatives) at 28.5 °C with 200 rpm shaking.

Construction of expression vectors

pET expression constructions The *egfp* gene was amplified with primers EGFP-F and I/EGFP-R (Table 3) by PCR method and was TA-cloned into pMD19 vector (Takara, Tokyo, Japan). After sequence analysis, the BamHI/XhoI fragment of *egfp* gene was prepared and inserted into the corresponding site of the pET28aDel plasmid [61]. This vector is designated as p28aD-eGFP. The *Iegfp* gene was constructed by fusing *Iasp* encoding sequence (primers IEGFP-F and IEGFP-fuR) with *egfp* gene (primers IEGFP-fuF and I/EGFP-R) by overlapping PCR using primers IEGFP-F and I/EGFP-R (Table 3). The PCR product of *Iegfp* gene was cloned into pMD19 vector. After sequencing, the BamHI/XhoI restriction fragment of *Iegfp* gene was inserted into the corresponding site of the pET28aDel plasmid to form p28aD-IeGFP plasmid. The p22b-eGFP expression vector was constructed by ligating NcoI/SacI fragment of pET22b with same restriction fragment of *egfp* from the pAc-eGFP plasmid described below. The encoding sequences of matrix metalloprotease-13 (MMP13, genbank accession number AAP78940.1) and myostatin (growth differentiating factor-8, GDF8, genbank accession number 5F3B_C) genes and their fusion version I-MMP13, T-MMP13 (Trx-MMP13), T-GDF8 and I-GDF8 with *Iasp* (I) or *Txr* (T) were synthesized after codon optimization and then separately cloned into the BamHI/SacI site of the pET28aDel2 plasmid. The NcoI recognition site was added at the initial codon of the *mmp13* and *gdf8* genes. The pET28aDel2 was obtained by replacing the sequences between XbaI and BamHI of pET28a with 5'-AATAATTTTGTTTAACTTTAA GAAGGAGATATA. The T-eGFP encoding sequence was synthesized and inserted into the BamHI/SacI site of the pET28aDel2 plasmid. The resulted plasmid was designated as p28aD2-T-eGFP. The p28aD2-M-eGFP and p28aD2-N-eGFP plasmids were constructed by replacing the *trx* sequence located between BamHI/NcoI site with the MBP and NusA encoding sequences, respectively. All of these expression vectors were separately transformed into BL21-star (DE3) strain by calcium chloride (CaCl₂) transformation. Both the eGFP and IeGFP expressed by pET vectors described above were tagged by the His-tag at the C terminal.

pHT304 expression constructions The pHT304 plasmid was linearized by SacI and blunted by T4 DNA polymerase (Thermo Fisher scientific, Grand Island, NY), and

Table 3 Primers used in this work

Primer	Sequence (5'-3')
EGFP-F	ACGCGGATCCACCATGGGTAAGGAGAAGAACTTTTCACCTG
IEGFP-F	ACGCGGATCCACCATGAACTAAAGAATCAAGAT
IEGFP-fuF	ATGAAGATTGTTTGAATGATGGGTAAGGAGAAGAACT
IEGFP-fuR	AGTTCTTCTCTTACCATCATTTTCAAACAATCTTCAT
I/EGFP-R	GACTCTCGAGTTTGTATAGTTTCATCCATGCCA
304I/EGFP-R	GAGCTCTTAGTGGTGGTGGTGGTGGTGGT
Pac-F	GCC <u>AAGCTT</u> CAGGTAATGGTTCTAACATGT
Pac-R	GGTGGATCCCTCCATCTCTTTTATTAAGAT
Tac-F	GGATCCATCATC <u>GAGCTC</u> GAGGCAAACCTCAGGTTTAAATATCGT
Tac-R	GGTACCCTGACAAATTGATTGAATGAT
cry1Ia-F	ATAAAAGAGATGGAGGGATCCACCATGAACTAAAGAATCAAGATAAG
cry1IaD44-F	ATAAAAGAGATGGAGGGATCCACCATGCTGAGTATGAAAATGTAGAG
cry1IaD73-F	ATAAAAGAGATGGAGGGATCCACCATGCTAGGCGTCTCTTTTGCAGGACAAG
cry1Ia-R	AACCTGAGTTTGCCTCGAGCTCTACATGTTACGCTCAATATGGAGT
Pi-F	CGCC <u>AAGCTT</u> GTATTATAGGTGTTTGAAGT
Pi-R	TGGTGGATCCCTCCACTTATACTATTATTTAAT
pelBlaz-F	CTTAATAAAAGAGATGGAGGGATCCACCATGAAATACCTGCTGCCGACC
pelBlaz-R	GAAAAGTTCTTCTCTTTACCCATGGTGGATCTGGCCATCGCCGGCTGGGCAGCGAGGA
torAlaz-F	CTTAATAAAAGAGATGGAGGGATCCACCATGAACTAAAGAATCAAGATAAG
torAlaz-R	GAAAAGTTCTTCTCTTTACCCATGGTGGATCTCGCTTGCAGGCGAGTGCACGCTCG
egfpI-F	ATGGATGAACTATACAACTCGAGCACCACCACCACCACATGAACTAAAGAATCA
egfpI-R	TATTTAACTGAGTTTGCCTCGAGCTCTTACATTTTCAAACAATCTTCATGAT
mCherry-F	ACGCGGATCCACCATGGTGGAGCAAGGGCGAGGAGGAT
ImC-fuF	ATCATGAAGATTGTTTGAATGATGGTGGAGCAAGGGCGAG
ImC-fuR	CTCGCCCTTGCCTACCATCATTTTCAAACAATCTTCATGAT
304I/mC-R	TGGTGGATCCCTCCACTTATACTATTATTTAAT
Hcat-F	CATCGGTCAACGGTACCA
Hcat-R	TGGCACTGCTGACACTTC
Hcat-Probe	FAM-AAGCCAATCATCACCAGCATCAGCCA-BHQ1
gfp-CF	GCCCTGTCCTTTTACCAGA
gfp-CR	CATCCATGCCATGTGAATCC
gfp-Probe	FAM-CCATTACCTGTCCACACAATCTGCCCT-BHQ1

Restriction sites were underline

then cyclized by ligase to produce p304Δ*SacI* plasmid. Both the *egfp* (primers GFP-F and 304I/EGFP-R) and *legfp* (primers IEGFP-F and 304I/EGFP-R) gene were re-amplified by PCR method from the p28aD-eGFP and p28aD-leGFP plasmids, respectively. Each PCR product contained the intact coding sequences (CDS) of the corresponding gene. The promoter (P_{ac} , primers Pac-F and Pac-R) and terminator (T_{ac} , primers Tac-F and Tac-R) of the *cryIaC* gene were obtained by PCR method. The whole DNA extracted from a Bt strain preserved in the lab was used as the template. After sequencing, all of the P_{ac} fragment (HindIII/BamHI), the T_{ac} fragment (BamHI/KpnI) and the *egfp* gene or *legfp* gene (BamHI/SacI) were inserted one by one into the corresponding

restriction sites of the p304Δ*SacI* plasmid. The plasmids were named pAc-eGFP or pAc-leGFP. The *cry1Ia* gene (primers cry1Ia-F and cry1Ia-R) and its truncated variants *cry1IaD44* (primers cry1IaD44-F and cry1Ia-R) and *cry1IaD73* (primers cry1IaD73-F and cry1Ia-R) were amplified by PCR method using the Bt strain harboring a *cry1Ia* gene. These amplicons were cloned into the BamHI/SacI site of the pAc-eGFP, respectively, to obtain the pAc-Cry1Ia, pAc-Cry1IaD44 and pAc-Cry1IaD73 plasmids. The *cry1Ia*, *cry1Ia44* and *cry1Ia73* were also cloned into the BamHI/SacI site of pET28aDel plasmid for expression in the *E. coli* BL21-star (DE3) strain. The promoter of *cry1Ia* gene (P_i) was obtained by PCR method using primers Pi-F and Pi-R and then was used

to replace the P_{ac} sequence of pAc-eGFP or pAc-IeGFP with HindIII and BamHI. The resulting plasmids were designated as pI-eGFP and pI-IeGFP respectively.

The pAc-torAeGFP plasmid were constructed by inserting the *torA* fragment (the *torA* signal peptide encoding sequence) to the BamHI site of pAc-eGFP vector by ClonExpress II One Step Cloning Kit (Vazyme Biotech, Nanjing, China). The *torA* fragment was obtained by PCR method using torAlaz-F and torAlaz-R primers with the genomic DNA of *E. coli* TG1 strain as the template. The pAc-pelBeGFP plasmid was constructed by the same method except that the *pelB* fragment (the *pelB* signal peptide encoding sequence) was amplified by pelBlaz-F and pelBlaz-R primer with the pET22b plasmid as the template. Similarly, the amplicon of *Iasp* encoding sequence amplified by egfp-F and egfp-R primers was recombined into the XhoI site of pAc-eGFP vector to produce the pAc-eGFPI plasmid.

The plasmid pAc-mCherry was constructed by replacing the *egfp* gene in pAc-eGFP plasmid with *mcherry* fragment amplified using mCherry-F and 304I/mC-R primers. The template for the PCR reaction was pCMV-N-mCherry plasmid (Beyotime biotechnology, Beijing, China) which is used for protein expression in mammalian cells. The *ImCherry* gene was constructed by fusing *Iasp* encoding sequence (primers IEGFP-F and ImC-fuR) with *mCherry* gene (primers ImC-fuF and 304I/mC-R) by overlapping PCR using primers IEGFP-F and 304I/mC-R and then was inserted into the BamHI/SacI sites of the pAc-eGFP plasmid. The resulting plasmid was designated as pAc-ImCherry.

All of the pHT304, pAc-eGFP, pAc-IeGFP, pAc-Cry1Ia, pAc-Cry1IaD44 and pAc-Cry1IaD73, pI-eGFP, pI-IeGFP, pAc-pelBeGFP, pAc-torAeGFP, pAc-eGFPI, pAc-mCherry and pAc-ImCherry plasmids were transformed separately into calcium chloride (CaCl_2) treated competent cells of *E. coli* TG1 strain and then into Bt BMB 171 strain by Bio-rad (Hercules, CA) MicroPulser™ Electroporator. The electroporation was performed using the 0.2 cm cuvette with 2.5 kilovolt pulse according to the manufacturer's instructions.

Proteins expression and samples preparation

E. coli TG1 strain The single colony of each *E. coli* TG1 strain harboring p304ΔSaI or its derived vectors described above was inoculated into 3 mL LB medium respectively and incubated overnight. The optical density at 600 nm (OD_{600}) of cell cultures for all strains were adjusted to be equal and then 6 μL cell culture of each strain was collected and inoculated into 6 mL fresh LB medium. Unless otherwise noted, five repeats were inoculated for each strain. At 4, 8, 10, 12 or 24 h after inoculation, 1 mL of cell cultures from each repeat for each

strain were taken and mixed. Out of the 5 mL mixture for each sample, 2 mL were separated and used to analyze the protein expression by western blot and the remaining 3 mL mixture were used for optical density and/or fluorescence detection. The analysis was replicated 3 times.

Bt BMB 171 strain The samples were prepared as same as the *E. coli* TG1 strains described above with the exception that each strains was sampled at 9, 12, 24, 36, 48, 60 or 72 h after inoculation. The cells and supernatant of the 5 mL mixture were separated by centrifugation. Then they were used for proteins expression analysis and fluorescence investigation. To express Cry1Ia and its truncated variants in Bt cells, the BAc-Cry1Ia, BAc-Cry1IaD44 and BAc-Cry1IaD73 strains were cultivated for 72 h and the total proteins in cells was analyzed by SDS-PAGE. In addition, the precipitates of BAc-Cry1Ia, BAc-Cry1IaD44 and BAc-Cry1IaD73 were resuspended by alkali buffer (50 mM Na_2CO_3 , 10 mM dithiothreitol, pH 10.5). After 1 h incubation at 37 °C, the soluble and insoluble components were separated by centrifugation at 12,000 rpm and analyzed by SDS-PAGE.

E. coli BL21-star (DE3) strain The single colony of each BL21-star (DE3) strain was inoculated into 3 mL LB medium respectively and incubated at 37 °C with 250 rpm shaking overnight. Then 50 μL cell culture for each strain was pipetted into 50 mL fresh LB medium and continuously cultivated in the same condition. When the OD_{600} of the cell culture reached 0.8, the recombinant protein expression was induced by 1 mM IPTG (isopropyl β -D-1-thiogalactopyranoside) with 250 rpm shaking for 4 h. After induction, cells for each strain were harvested respectively by centrifuging at 10,000 rpm for 10 min and were washed by deionized water for three times. The cells were resuspended in 20 mL ice-cold LE buffer (50 mmol/L Na_2HPO_4 , 0.3 mol/L NaCl, pH 8.0). Two milliliters of the cell suspension for each strain was pipetted into a new tube for protein analysis. The remaining cells were ultrasonic-treated in ice-cold condition. The lysates were centrifuged at 12,000 rpm at 4 °C for 15 min and then the supernatant and the resuspended precipitate by same volume of LE buffer as the supernatant of each strain was prepared for SDS-PAGE analysis.

Fluorescence detection

The fluorescent signals were detected by Shimadzu RF530PC Spectrofluorophotometer (Kyoto, Japan). The excitation wavelength (EX) was 488 nm and the fluorescent intensities were recorded at 511 nm (emission wavelength, EM) for eGFP and IeGFP proteins. The EX/EM for detecting the mCherry were 565 nm and 603 nm, respectively. The fluorescent intensity is an arbitrary unit (A.U.). Before recording, the emission spectrum of each strain was scanned from 450 nm to 650 nm to confirm

the emission peak was at or very close to the EM. The slit widths of EX and EM and the sensitivity of detection were indicated independently. The sensitivity at high mode is about 50 times higher than that at low mode according to the specification of RF 5301PC. For Bt strains, both the supernatant of the cell culture after centrifugation and the cells resuspended by fresh LB medium were monitored. For each *E. coli* strain, the whole cell culture was the loading sample. Each treatment for either *E. coli* or Bt strains was replicated three times. Unless otherwise noted, in each replication every record represented the value of mixture from 5 parallel tubes.

Proteins analysis by SDS-PAGE and western blot

Unless otherwise noted, the collected cells of each sample from 2 mL culture were resuspended by 100 μ L PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 2 mM KH₂PO₄, pH 7.4) and the corresponding supernatant of Bt strains was concentrated to 200 μ L by ultrafiltration using Amicon Ultra-0.5 Centrifugal Filters (3 kDa, Millipore, MA, USA). All of the protein samples were mixed with one-fourth volume of 5 \times SDS gel-loading buffer respectively, and boiled for 5 min. After centrifuged at 12,000 rpm for 5 min, they were loaded onto gels for separation by SDS-PAGE and analyzed after Coomassie bright blue staining.

For western blot analysis, the separated proteins were transformed onto the nitrocellulose membrane without staining and incubated with rabbit antiserum against IeGFP protein and the horseradish peroxidase-conjugated goat antirabbit IgG (H+L) antibody (MultiSciences, HangZhou, China) successively. For detecting mCherry and ImCherry proteins, the primary antibody was the mCherry-tag monoclonal antibody (MultiSciences, HangZhou, China) and the secondary antibody was the horseradish peroxidase-conjugated goat anti-mouse IgG (H+L) antibody (MultiSciences, HangZhou, China). The target bands were visualized using 3,3'-diaminobenzidine (DAB, Sigma, St. Louis, MO).

mRNA analysis of *legfp* and *egfp* gene

The transcription levels of *legfp* and *egfp* gene were compared by real-time quantitative PCR method. First, the total RNAs of TAC-eGFP and TAC-IeGFP cells collected at different times were extracted using RNAiso Plus (Takara Biomedical Technology, Beijing, China). Second, the obtained RNA samples were reverse-transcribed to cDNA by PrimeScript RT reagent kit with gDNA Eraser (Takara Biomedical Technology, Beijing, China). Finally, the cDNAs were used for real-time PCR (Takara Premix Ex Taq for Probe qPCR) to quantify the transcription level of *egfp* and *legfp* genes. The *Hcat* gene of *E. coli* was used as the reference gene [62]. The primers used

for real-time PCR were listed in Table 2. The differential expression analysis of *egfp* and *legfp* genes were evaluated by 2^{- Δ Ct} method [63].

The secondary structures of nucleotides near the initial codon (-4~+37 nt) for both *legfp* and *egfp* genes were predicted by RNAstructure software [47]. The structure with lowest folding energy for each sequence were chosen for comparison.

Fluorescence localization

The cells were prepared by Schlegel's method with modification [15]. Briefly, cells corresponding to 1 A600 unit were harvested (4000 \times g, 2 min) and washed by 1 mL PBS buffer twice. Subsequently, 600 μ L fixing solution (2% Paraformaldehyde, 2.5% Glutaraldehyde in PBS) was added and cells were incubated for 45 min at room temperature (RT). Subsequently, cells were washed three times with PBS and resuspended in 100 μ L PBS. Three microliters of the cell suspension was mounted on a glass slide. Fluorescence images of cells were obtained using inverted confocal microscope (Leica SP8) with a 60 \times times oil immersion objective (Leica). The resulting images were recorded with the LAS-AF-Lite software (Leica).

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12934-020-01371-8>.

Additional file 1: Table S1. The fluorescent intensity data of TAC-eGFP and TAC-IeGFP cell cultures. **Table S2.** The fluorescent intensity data of the resuspended BAC-eGFP and BAC-IeGFP cells. **Table S3.** The fluorescent intensity data of the supernatant of BAC-eGFP and BAC-IeGFP cell cultures. **Table S4.** The fluorescent intensity data of five *E. coli* strains expressing eGFP and its variants. **Table S5.** The fluorescent intensity data of BI-eGFP and BI-IeGFP cells. **Table S6.** The fluorescent intensity data of the supernatants of BI-eGFP and BI-IeGFP cell cultures. **Table S7.** The fluorescent intensity data of TAC-mCherry and TAC-ImCherry cell cultures. **Table S8.** The fluorescent intensity data of supernatants of BAC-mCherry and BAC-ImCherry cell cultures. **Table S9.** The fluorescent intensity data of BAC-ImCherry cells. **Figure S1.** Comparison of fluorescent proteins expressed in different strains. **Figure S2.** Solubility analysis of three recombinant proteins. **Figure S3.** SDS-PAGE analysis of Cry1Ia and its truncated variants.

Acknowledgements

We wish to thank Dr. Jianhua Lv of college of life sciences in Shanxi Agricultural University for her kindly gift of the *egfp* gene, and we also appreciate Dr. Rupert Fray of the University of Nottingham for revising this manuscript.

Authors' contributions

J Gao and XW conceived of the presented idea and planned the experiments. J Gao, HQ, XG, YM, JGuo, XW, CL and ZS analyzed the data. JGao wrote the manuscript with support from XW, CL, ZS, YJ, HQ, XG, YM, J Guo and JZ constructed the vectors and analyzed the protein expression. J Gao, HQ, XG, YM and J Guo investigated of fluorescent intensity with support from MD and TZ. J Gao, HQ, XG and MD took photographs with confocal microscope. CX and TZ prepared the polyclonal antibodies of IeGFP. All authors read and approved the final manuscript.

Funding

This work was supported by Project 31601690 supported by National Natural Science Foundation of China.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ College of Life Sciences, Shanxi Agricultural University, Taigu 030801, China.

² State Key Laboratory of Rice Biology, Institute of Insect Sciences, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China. ³ Experimental Teaching Center, Shanxi Agricultural University, Taigu 030801, China. ⁴ Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA.

Received: 26 August 2019 Accepted: 16 May 2020

Published online: 24 May 2020

References

- Zimmer M. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem Rev*. 2002;102:759–81.
- Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. *Nat Methods*. 2005;2:905–9.
- Snapp EL. Fluorescent proteins: a cell biologist's user guide. *Trends Cell Biol*. 2009;19:649–55.
- Cramer A, Whitehorn EA, Tate E, Stemmer WPC. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. 1996;14:315–9.
- Fukuda H, Arai M, Kuwajima K. Folding of green fluorescent protein and the cycle3 mutant. *Biochemistry*. 2000;39:12025–32.
- Wang JD, Herman C, Tipton KA, Gross CA, Weissman JS. Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell*. 2002;111:1027–39.
- Chang H-C, Kaiser CM, Hartl FU, Barral JM. De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *J Mol Biol*. 2005;353:397–409.
- Plaxco KW, Simons KT, Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol*. 1998;277:985–94.
- Tsien RY. The green fluorescent protein. *Annu Rev Biochem*. 1998;67:509–44.
- Zhang G, Gurtu V, Kain SR. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem Biophys Res Commun*. 1996;227:707–11.
- Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*. 2006;24:79–88.
- Gupta SK, Shukla P. Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. *Crit Rev Biotechnol*. 2016;36:1089–98.
- Overton TW. Recombinant protein production in bacterial hosts. *Drug Discov Today*. 2014;19:590–601.
- Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;5:1–17.
- Schlegel S, Rujas E, Ytterberg AJ, Zubarev RA, Luirink J, de Gier J-W. Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. *Microb Cell Fact*. 2013;12:24–35.
- Sørensen HP, Mortensen KK. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact*. 2005;4:1–8.
- Costa S, Almeida A, Castro A, Domingues L. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Front Microbiol*. 2014;5:63.
- Waldo GS, Standish BM, Berendzen J, Terwilliger TC. Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol*. 1999;17:691–5.
- Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X, Butt TR. Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein Sci*. 2006;15:182–9.
- Ikeno S, Haruyama T. Boost protein expression through co-expression of LEA-like peptide in *Escherichia coli*. *PLoS ONE*. 2013. <https://doi.org/10.1371/journal.pone.0082824>.
- Nguyen TKM, Ki MR, Son RG, Pack SP. The NT11, a novel fusion tag for enhancing protein expression in *Escherichia coli*. *Appl Microbiol Biotechnol*. 2019;103:2205–16.
- Ojima-Kato T, Nagai S, Nakano H. N-terminal SKIK peptide tag markedly improves expression of difficult-to-express proteins in *Escherichia coli* and *Saccharomyces cerevisiae*. *J Biosci Bioeng*. 2017;123:540–6.
- Xiao W, Jiang L, Wang W, Wang R, Fan J. Evaluation of rice tetracopeptide domain-containing thioredoxin as a novel solubility-enhancing fusion tag in *Escherichia coli*. *J Biosci Bioeng*. 2018;125:160–7.
- Zhao W, Liu L, Du G, Liu S. A multifunctional tag with the ability to benefit the expression, purification, thermostability and activity of recombinant proteins. *J Biotechnol*. 2018;283:1–10.
- Zhang A, Cantor EJ, Barshevsky T, Chong S. Productive interaction of chaperones with substrate protein domains allows correct folding of the downstream GFP domain. *Gene*. 2005;350:25–31.
- Ruan A, Ren C, Quan S. Conversion of the molecular chaperone spy into a novel fusion tag to enhance recombinant protein expression. *J Biotechnol*. 2019;307:131–8.
- Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science*. 2009;324:255–8.
- Siller E, DeZwaan DC, Anderson JF, Freeman BC, Barral JM. Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. *J Mol Biol*. 2010;396:1310–8.
- Dinh T, Bernhardt TG. Using superfolder green fluorescent protein for periplasmic protein localization studies. *J Bacteriol*. 2011;193:4984–7.
- Aronson DE, Costantini LM, Snapp EL. Superfolder GFP is fluorescent in oxidizing environments when targeted via the Sec translocon. *Traffic*. 2011;12:543–8.
- Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol*. 2015;13:343.
- Berks BC. The twin-arginine protein translocation pathway. *Annu Rev Biochem*. 2015;84:843–64.
- Santini C-L, Bernadac A, Zhang M, Chanal A, Ize B, Blanco C, Wu L-F. Translocation of Jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem*. 2001;276:8159–64.
- Jong WSP, Vikström D, Houben D, van den Berg van Saparoea HB, de Gier J-W, Luirink J. Application of an *E. coli* signal sequence as a versatile inclusion body tag. *Microb Cell Fact*. 2017;16:50.
- Adalat R, Saleem F, Crickmore N, Naz S, Shakoori A. *In vivo* crystallization of three-tomain cry toxins. *Toxins*. 2017;9:80–92.
- Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol Mol Biol Rev*. 1998;62:775–806.
- Gleave AP, Williams R, Hedges RJ. Screening by polymerase chain reaction of *Bacillus thuringiensis* serotypes for the presence of cryV-like insecticidal protein genes and characterization of a cryV gene cloned from *B. thuringiensis* subsp. *kurstaki*. *Appl Environ Microbiol*. 1993;59:1683–7.
- Shin BS, Park SH, Choi SK, Koo BT, Lee ST, Kim JI. Distribution of cryV-type insecticidal protein genes in *Bacillus thuringiensis* and cloning of cryV-type genes from *Bacillus thuringiensis* subsp. *kurstaki* and *Bacillus thuringiensis* subsp. *entomocidus*. *Appl Environ Microbiol*. 1995;61:2402–7.
- Tounsi S, Jaoua S. Identification of a promoter for the crystal protein-encoding gene cry11a from *Bacillus thuringiensis* subsp. *kurstaki*. *FEMS Microbiol Lett*. 2002;208:215–8.
- Kostichka K, Warren GW, Mullins M, Mullins AD, Palekar NV, Craig JA, Koziel MG, Estruch JJ. Cloning of a cryV-type insecticidal protein gene from *Bacillus thuringiensis*: the cryV-encoded protein is expressed early in stationary phase. *J Bacteriol*. 1996;178:2141–4.

41. DeLisa MP, Lee P, Palmer T, Georgiou G. Phage shock protein PspA of *Escherichia coli* relieves saturation of protein export via the Tat pathway. *J Bacteriol*. 2004;186:366–73.
42. Rothman JE, Orci L. Molecular dissection of the secretory pathway. *Nature*. 1992;355:409–15.
43. Wilson JW. Bacterial Protein Secretion Mechanisms. In: Nickerson CA, Schurr MJ, editors. *Molecular paradigms of infectious disease: a bacterial perspective*[M]. Boston: Springer; 2006. p. 274–320.
44. Freudl R. Leaving home ain't easy: protein export systems in Gram-positive bacteria. *Res Microbiol*. 2013;164:664–74.
45. Goosens VJ, Monteferrante CG, van Dijl JM. The Tat system of Gram-positive bacteria. *Biochimica et Biophysica Acta (BBA) Molecular Cell Research*. 2014;1843:1698–706.
46. Schnepf HE, Wong HC, Whiteley HR. Expression of a cloned *Bacillus thuringiensis* crystal protein gene in *Escherichia coli*. *J Bacteriol*. 1987;169:4110–8.
47. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinform*. 2010;11:129.
48. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*. 2005;361:13–37.
49. Boël G, Letso R, Neely H, Price WN, Wong K-H, Su M, Luff JD, Valecha M, Everett JK, Acton TB, Xiao R, Montelione GT, Aalberts DP, Hunt JF. Codon influence on protein expression in *E. coli* correlates with mRNA levels. *Nature*. 2016;529:358.
50. Georgiou G, Valax P. Expression of correctly folded proteins in *Escherichia coli*. *Curr Opin Biotechnol*. 1996;7:190–7.
51. Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol*. 1999;10:411–21.
52. Borrero J, Jiménez JJ, Gútiérrez L, Herranz C, Cintas LM, Hernández PE. Protein expression vector and secretion signal peptide optimization to drive the production, secretion, and functional expression of the bacteriocin enterocin A in lactic acid bacteria. *J Biotechnol*. 2011;156:76–86.
53. Klatt S, Konthur Z. Secretory signal peptide modification for optimized antibody-fragment expression-secretion in *Leishmania tarentolae*. *Microb Cell Fact*. 2012;11:1–10.
54. Beena K, Udgaonkar JB, Varadarajan R. Effect of signal peptide on the stability and folding kinetics of maltose binding protein. *Biochemistry*. 2004;43:3608–19.
55. Krishnan B, Kulothungan SR, Patra AK, Udgaonkar JB, Varadarajan R. SecB-mediated protein export need not occur via kinetic partitioning. *J Mol Biol*. 2009;385:1243–56.
56. Kulothungan SR, Das M, Johnson M, Ganesh C, Varadarajan R. Effect of crowding agents, signal peptide, and chaperone SecB on the folding and aggregation of *E. coli* maltose binding protein. *Langmuir*. 2009;25:6637–48.
57. Singh P, Sharma L, Kulothungan SR, Adkar BV, Prajapati RS, Ali PSS, Krishnan B, Varadarajan R. Effect of signal peptide on stability and folding of *Escherichia coli* Thioredoxin. *PLoS ONE*. 2013;8:e63442.
58. Doherty GP, Bailey K, Lewis PJ. Stage-specific fluorescence intensity of GFP and mCherry during sporulation in *Bacillus Subtilis*. *BMC Res Notes*. 2010;3:303–10.
59. Brar SK, Verma M, Tyagi RD, Surampalli RY, Barnabé S, Valéro JR. *Bacillus thuringiensis* proteases: production and role in growth, sporulation and synergism. *Process Biochem*. 2007;42:773–90.
60. Cabantous S, Terwilliger TC, Waldo GS. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol*. 2005;23:102–7.
61. Gao J, Zhang Y, Zhao Q, Lin C, Xu X, Shen Z. Transgenic rice expressing a fusion protein of Cry1Ab and Cry9Aa confers resistance to a broad spectrum of lepidopteran pests. *Crop Sci*. 2011;51:2535–43.
62. Zhou K, Zhou L, Lim Q, Zou R, Stephanopoulos G, Too HP. Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. *BMC Mol Biol*. 2011;12:18.
63. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc*. 2008;3:1101.
64. Arantes O, Lereclus D. Construction of cloning vectors for *Bacillus thuringiensis*. *Gene*. 1991;108:115–9.
65. Gibson TJ. Studies on the Epstein-Barr virus genome. Cambridge: University of Cambridge; 1984.
66. Li L, Yang C, Liu Z, Li F, Yu Z. Screening of acrySTALLIFEROUS mutants from *Bacillus thuringiensis* and their transformation properties. *Wei Sheng Wu Xue Bao Acta Microbiologica Sinica*. 2000;40:85–90 (In Chinese).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

