J Ind Microbiol Biotechnol (2014) 41:1435–1442
DOI 10.1007/s10295-014-1482-8

SIMB
Society for Industrial Microbiology and Biotechnology

Effect of codon-optimized *E. coli* signal peptides on recombinant *Bacillus stearothermophilus* maltogenic amylase periplasmic localization, yield and activity

Shalaka Samant · Gunja Gupta · Subbulakshmi Karthikeyan · Saiful F. Haq · Ayyappan Nair · Ganesh Sambasivam · Sunilkumar Sukumaran

Received: 13 March 2014 / Accepted: 28 June 2014 / Published online: 20 July 2014

Abstract Recombinant proteins can be targeted to the *Escherichia coli* periplasm by fusing them to signal peptides. The popular pET vectors facilitate fusion of target proteins to the PelB signal. A systematic comparison of the PelB signal with native *E. coli* signal peptides for recombinant protein expression and periplasmic localization is not reported. We chose the *Bacillus stearothermophilus* maltogenic amylase (MA), an industrial enzyme widely used in the baking and brewing industry, as a model protein and analyzed the competence of seven, codon-optimized, *E. coli* signal sequences to translocate MA to the *E. coli* periplasm compared to PelB. MA fusions to three of the signals facilitated enhanced periplasmic localization of MA compared to the PelB fusion. Interestingly, these three fusions showed greatly improved MA yields and between 18- and 50-fold improved amylase activities compared to the PelB fusion. Previously, non-optimal codon usage in native *E. coli* signal peptide sequences has been reported to be important for protein stability and activity. Our results suggest that *E. coli* signal peptides with optimal codon usage could also be beneficial for heterologous protein secretion to the periplasm. Moreover, such fusions could even enhance activity rather than diminish it. This effect, to our knowledge has not been previously documented. In addition, the seven vector platform reported here could also be used as a screen to identify the best signal peptide partner for other recombinant targets of interest.

Keywords Maltogenic amylase · Periplasmic localization · *E. coli* · Signal peptides · Codon-optimization

S. Samant (✉) · G. Gupta · S. Karthikeyan · S. F. Haq · A. Nair · G. Sambasivam · S. Sukumaran
Anthem Biosciences Pvt. Ltd, #49, Bommasandra Industrial Area Phase-I, Hosur Road, Bangalore 560099, India
e-mail: shalaka.s@anthembio.com

Present Address:
G. Gupta
Molecular Lab, Apollo Hospitals, Lab block, Fourth Floor, Greams Lane, Off Greams Road, Opp. Krishna Deep Chambers, Chennai 600006, India

Present Address:
S. Karthikeyan
Center for Pharmaceutical Biotechnology MC870 College of Pharmacy, University of Illinois at Chicago, 3150-MBRB, 900 South Ashland Avenue, Chicago, IL 60607-7173, USA

Introduction

In spite of the several *Escherichia coli* expression systems available for the expression of heterologous recombinant proteins [8, 9, 13, 15], protein misfolding and aggregation, proteolysis, and recombinant protein toxicity continue to pose challenges [8, 12]. One approach to circumvent this problem is to export the desired protein to the periplasmic space by fusing it to signal peptides. This affords several advantages such as avoiding cytoplasmic proteases, facilitating proper folding and ease of purification [20].

Periplasmic and outer membrane proteins in *E. coli* are synthesized in the cytoplasm as precursors that contain a short specific amino acid sequence (signal sequence) that allows proteins to be exported across the cytoplasmic membrane [10]. During transport of proteins out of the cytoplasm, the signal sequence is cleaved by a signal peptidase to yield the mature protein product [16]. The major mechanism driving the post-translational translocation of unfolded proteins in *E. coli* is the sec-dependent type II secretion pathway [5, 22]. Co-translational transport through the sec translocon is mediated via a signal recognition particle (SRP) that recognizes hydrophobic stretches in the N-terminal signal peptide [23]. Another,

sec-independent translocation mechanism is mediated through the twin arginine transport (TAT) pathway which transports proteins containing a twin arginine signal sequence in a folded conformation [19].

The efficiency of translocation of a given protein is tightly linked to its compatibility with the signal sequence used and does not guarantee localization to the periplasm [5]. Since the SEC pathway requires the use of chaperone proteins [known to be substrate (protein)-specific] [2] to effect translocation, many heterologous proteins when expressed in *E. coli* with SEC signal sequences cannot be exported due to lack of recognition by the host chaperones. For SEC-based translocation, the chaperones must retain the substrate protein in a partially unfolded state, a phenomenon that is observed only in a certain class of proteins. It has also been reported that for secretion using the TAT pathway, some proteins cannot be exported in a fully folded state due to steric interference [14]. Co-translational SRP-dependent transport has also been known to be hindered by the nature of the protein sequence [14]. A series of charged or hydrophobic amino acid residues could block the fusion protein from being translocated due to strong protein–protein interactions. The choice of signal sequence for efficient secretion should therefore be individually tested for each new protein. This warrants the development of a platform of vectors with signal sequences that are known to target proteins through both the SEC and TAT systems. This platform could act as a screen for identifying the best signal peptide partner for efficient periplasmic localization.

Maltogenic amylase is an exo-acting amylolytic enzyme that catalyzes the hydrolysis of alpha-1,4-glucosidic linkages in amylose, amylopectin and related glucose polymers with wide applications in the baking, starch and brewing industry. Previously recombinant *Bacillus stearothermophilus* maltogenic amylase (MA) has been expressed in *E. coli* [3]. However, the reported, elaborate enzyme purification process involving sonication for cell lysis resulted in the recovery of only 3.65 % of the total maltogenic amylase activity.

We hypothesized that by analyzing various native *E. coli* signal peptide–MA fusions, we should be able to identify an efficient recombinant expression system for this enzyme that facilitates periplasmic localization thereby reducing the complexity of the purification. Signal sequences of native *E. coli* proteins are reported to exhibit a preference for non-optimal codons and altering this bias affects protein stability [23]. In an effort to characterize a novel set of signal peptides for facilitating periplasmic translocation of MA, we modified seven native *E. coli* signal peptides to encode codons with the highest frequency of occurrence in *E. coli* protein sequences. The aim of the present study was to compare the efficiency of these codon-optimized signal peptides in translocating MA to the *E. coli* periplasm to that of the popular PelB signal which directs proteins to the *E. coli* periplasm through the SEC pathway.

Materials and methods

Strains, vectors, reagents and enzymes

E. coli strains DH5α (Novagen, USA) and BL21 DE3 (Novagen, USA) were used as hosts for cloning and recombinant protein overexpression, respectively. Restriction enzymes *Nco*I, *BamH*I, *Nde*I and *Bgl*II were purchased from NEB (New England Biolabs). PET20b+ vector from Novagen was used as the backbone for construction of the platform of seven signal peptide–MA fusions.

Construction of signal peptide–MA fusions and transformation of BL21 DE3 cells

For construction of seven different vectors with different signal peptide–MA fusions, the maltogenic amylase gene from *B. stearothermophilus* was amplified using the following primers MANcoI fwd primer: 5′-gatcgtaccatgggaAT GAGCAGTTCCGCAAGCGT-3′ and MABglII rev primer: 5′-gatcgtacagatctTCTAGACTAGTTTTGCCACG-3′. This PCR product was digested with *Nco*I and *Bgl*II enzymes and cloned into the *Nco*I and *BamH*I digested pET20b+ vector. The resulting ligation mix was transformed into DH5α *E. coli* cells. The resulting plasmid was sequenced to verify that the coding sequence of maltogenic amylase was correct and in frame with the pET20b+ translation start site. The plasmid was then digested with *Nde*I and *Nco*I and the PelB signal peptide was removed by gel elution. Seven signal peptides carrying codon-optimized versions of native *E. coli* signal sequences that are known to follow different transport pathways (sequences listed in Table 1) were synthesized with *Nde*I and *Nco*I overhangs and cloned into the digested pET vector to yield the plasmids pAEV01–pAEV07. The resulting seven plasmids were transformed into DH5α *E. coli* cells and the recombinant *E. coli* cells were selected on ampicillin containing LB plates and screened by a colony PCR method. Plasmids isolated from positive DH5α *E. coli* cells were validated by restriction digests with *Nde*I and *Nco*I and sequence verified using T7 fwd primer (5′-taatacgactcactataggg-3′).

Induction of MA expression and SDS-PAGE analysis

To determine the expression levels of MA from the seven constructs compared to the commercial vector pET20b+ that carried a PelB–MA fusion, the validated constructs were transformed into BL21 DE3 *E. coli*. All eight BL21 DE3 strains including the control pET20b+ construct as

Table 1 Codon-optimized nucleotide sequences of signal peptide used in this study and the putative secretory pathway they follow

Plasmid name	Putative secretion pathway	Signal peptide name	Signal peptide nucleotide sequence
pAEV01	SEC	MalE	atgaaaattaaaaccggcgcgcgcattctggcgctgagcgcgctgaccaccatgatgtt-tagcgctagcgcgctggcc
pAEV02	SEC	OmpA	atgaaaaaaaccgcgattgcgattgcggtggcgctggcgggctttgcgaccgtggcgcaggcc
pAEV03	TAT	TorA	atgaacaacaacgatctgtttcaggcgagccgccgccgctttctggcgcagctgggcggcctgaccgtg-gcgggcatgctggggcccagcctgctgaccccgcgccgcgcgaccgcggcgcaggcc
pAEV04	SRP	DsbA	atgaaaaaaatttggctggcgctggcgggcctggtgctggcgtttagcgctagcgcc
pAEV05	TAT + SEC	FhuD	atgagcggcctgccgctgattagccgccgccgcctgctgaccgcgatggcgctgagcccgctgctgtg-gcagatgaacaccgcgcatgcc
pAEV06	TAT + SEC	YcdO	atgaccattaactttcgccgcaacgcgctgcagctgagcgtggcggcgctgtttagcagcgcgtttatggc-gaacgcc
pAEV07	TAT + SEC	MdoD	atggatcgccgccgctttattaaaggcagcatggcgatggcggcggtgtgcggcaccagcggcatt-gctagcctgtttagccaggcggcgtttgcc

well as the seven novel plasmids were grown in minimal medium supplemented with glucose as the carbon source and 100 μg mL^{-1} ampicillin. Minimal media contained 12.8 g Na_2HPO_4 l^{-1}, 2 g KH_2PO_4 l^{-1}, 0.5 g NaCl l^{-1}, 1 g NH_4Cl l^{-1}, 0.5 g $MgSO_4$ l^{-1}, 4 g dextrose monohydrate l^{-1}, 0.0147 g $CaCl_2$ l^{-1}. The cultures were incubated overnight at 37 °C in a shaker incubator. Following overnight incubation the cultures were diluted 1:100 into a fresh 250 ml flask with 50 ml yeast extract media containing 100 μg mL^{-1} ampicillin. Yeast extract media contained 11 g dextrose monohydrate l^{-1}, 3.6 g NH_2SO_4 l^{-1}, 7.2 g $MgSO_4$ l^{-1}, 20 g yeast extract l^{-1}, K_2HPO_4 10 g L^{-1} supplemented with trace salts. The cultures were incubated at 37 °C with shaking. Cell growth was monitored spectrophotometrically at 600 nm (OD_{600}) by measuring the absorption of the culture. 1 mM IPTG was added to the culture when the OD_{600} reached 0.6. The culture was then incubated at 30 °C for 16 h with shaking. The induced and uninduced cultures were pelleted by centrifugation at 12,000*g* for 5 min. The pellet was resuspended in sample buffer containing 10 mM NaCl, pH 5. This pellet was sonicated (10 s pulse, 10 s off for 5 min) to release the soluble protein, cell debris was pelleted out by centrifugation at 12,000*g* for 5 min and the supernatant was collected. Total protein amounts in the sonicated supernatant were determined by using Bradford's reagent and 10 μg total protein corresponding to each sample was analyzed on a 12 % SDS-PAGE gel followed by Coomassie blue staining. Similarly, to determine the effect of varying temperature on the expression of MA, induction was carried out by adding 1 mM IPTG to the cultures and induction temperature was maintained at 26 °C.

Localization of MA

To determine the intracellular localization of MA as a result of fusion to the various signal peptides used in this study, analysis of various cellular fractions was conducted as described earlier [18]. Briefly, BL21 DE3 *E. coli* carrying four different signal peptide–MA fusions, i.e., DsbA, MalE, FhuD and PelB were induced with 1 mM IPTG and incubated at 30 °C for 16 h. A BL21 DE3 strain carrying empty pET20b+ vector was included as the control. 2 mL of the induced culture was centrifuged at 17,000*g* for 10 min. The pellet was resuspended in 100 μl of periplasmic buffer I (20 % sucrose, 1 mM EDTA and 100 mM Tris–HCl, pH 8.0). This mix was incubated on ice for 30 min. The cells were pelleted at 17,000*g*, 10 min, 4 °C. The supernatant represented the periplasmic fraction. The remaining pellet was resuspended in 100 μl of periplasting buffer II (50 mM $MgCl_2$) and incubated on ice for 20 min. The cells were pelleted at 17,000*g*, 10 min, 4 °C. The supernatant represented the osmotic fraction. The osmotic shock fraction and periplasmic fraction were combined and loaded on a 12 % SDS-PAGE gel and protein detected by Coomassie staining. This is referred to as the periplasmic fraction of the cells.

MA activity assay

Sonicated supernatant samples corresponding to all eight MA–signal peptide fusions including the PelB–MA fusion were prepared as described under induction of MA expression and SDS-PAGE analysis in materials and methods. The sonicated supernatant was also analyzed for maltogenic amylase activity using the glucose oxidase method. The method is based on the ability of the enzyme to hydrolyze maltotriose to maltose and glucose. Briefly, 500 μl of substrate solution (20 mg mL^{-1} maltotriose, prepared in 0.1 M citrate buffer, pH 5.0, pre-warmed to 37 °C) to an equal volume of enzyme solution also pre-warmed to 37 °C. The resulting solution was mixed thoroughly and transferred to a water bath maintained at 37 °C. After 30 min of incubation GOPOD reagent (0.008 g of glucose oxidase in 100 mL of ortho-dianisidine peroxidase reagent)

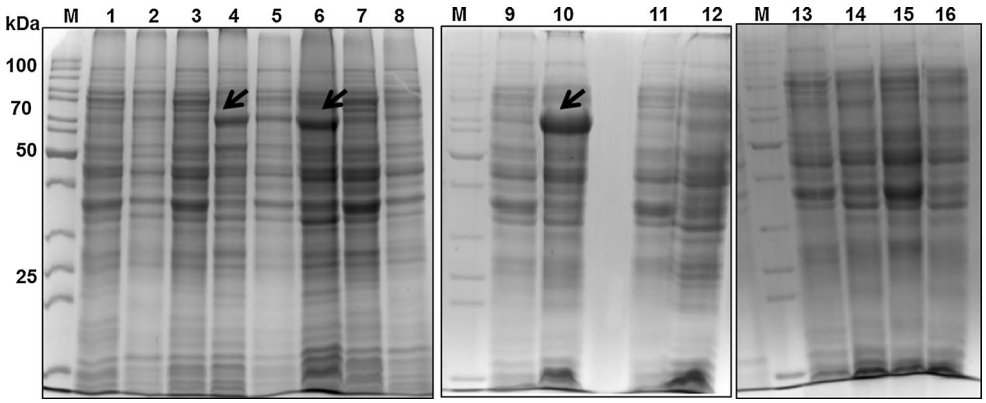

Fig. 1 Codon-optimized DsbA, MalE and FhuD signals are optimal signal peptide partners for the inducible expression of MA. SDS-PAGE analysis of *E. coli* BL21 DE3 cultures carrying different signal peptide–MA fusion constructs induced with 1 mM IPTG for 16 h at 30 °C. 10 µg of total protein was loaded on a 12 % SDS-PAGE gel. *Lane M*: molecular weight marker, *1* uninduced PelB–MA fusion, *2* induced PelB–MA fusion, *3* uninduced FhuD–MA fusion, *4* induced FhuD–MA fusion, *5* uninduced MalE–MA fusion, *6* induced MalE–MA fusion, *7* uninduced TorA–MA fusion, *8* induced TorA–MA fusion, *9* uninduced DsbA–MA fusion, *10* induced DsbA–MA fusion, *11* uninduced MdoD–MA fusion, *12* induced MdoD–MA fusion, *13* uninduced YcdO–MA fusion, *14* induced YcdO–MA fusion, *15* uninduced OmpA–MA fusion, *16* induced OmpA–MA fusion. The position of maltogenic amylase (MA) in the strains that showed significant induction is indicated by an *arrowhead*

was added and the mixture was incubated for an additional 30 min for color development. The reaction was stopped by adding 0.1 N HCl and the OD read at 420 nm. Three independent induction experiments were carried out and the MA activity assay was used to determine fold induction.

Results

Fusing recombinant proteins to signal peptides for facilitating periplasmic localization is a well-known strategy [20]. However, the efficiencies of various signal peptides in translocating recombinant targets are not very well characterized. *B. stearothermophilus* maltogenic amylase (MA) is a widely used enzyme in the baking and brewing industry. Previously MA yields from *E. coli* expression systems have not been promising [3]. We used MA as a model protein to evaluate the efficiency of seven different *E. coli* signal peptides in transporting MA to the *E. coli* periplasm.

Codon-optimized DsbA, MalE and FhuD signals are optimal signal peptide partners for the inducible expression of MA

Seven native *E. coli* signal peptides known to target proteins to the periplasm via each one of the three secretory mechanisms in *E. coli*, i.e., post-translational (SEC), TAT and co-translational (SRP) were chosen (Table 1). Three signal peptides that are hypothesized to target the proteins both through the SEC and TAT pathway (i.e., FhuD, YcdO and MdoD) were also included (Table 1). The native sequences of these peptides were altered to include codons that occur with the highest frequency in the *E. coli* genome. To determine if inducible expression of MA was possible using the MA fusions to these seven peptides, an SDS-PAGE analysis was carried out. We observed that unlike what was expected not all signal peptide fusions were compatible for inducible expression of MA. Following induction with 1 mM IPTG and growth for 16 h at 30 °C, strains carrying pAEV01 (MalE–MA fusion), pAEV04 (DsbA–MA fusion) and pAEV05 (FhuD–MA fusion) constructs revealed expression of a 70 kDa protein, the expected size of maltogenic amylase, on a 12 % SDS-PAGE gel. Induction of maltogenic amylase protein levels in these three strains was significantly higher compared to any other fusion (Fig. 1). Interestingly, no induction of maltogenic amylase that could be detected by Coomassie staining was observed in strains carrying pAEV02 (OmpA–MA fusion), pAEV03 (TorA–MA fusion), pAEV06 (YcdO–MA fusion) and pAEV07 (MdoD–MA fusion) as well as PelB–MA fusion (Fig. 1). Our observations suggest that codon-optimized MalE, FhuD and DsbA signal peptides represent the most favorable partners for MA production with the DsbA–MA fusion being the best.

MA expression pattern is not altered when fused to the codon-optimized MalE, FhuD and DsbA signals following temperature change post-induction

Protein expression can be enhanced by lowering the temperature that the culture is incubated at post-induction. To determine whether such a change in temperature could alter the expression pattern of MA, we induced the BL21 DE3 cultures with 1 mM IPTG for 16 h at 26 °C. Again only

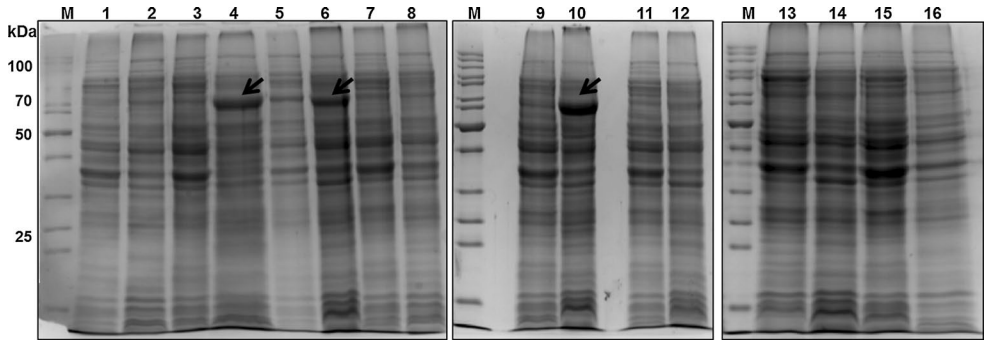

Fig. 2 MA expression pattern is not altered when fused to the codon-optimized MalE, FhuD and DsbA signals following temperature change post-induction: SDS-PAGE analysis of *E. coli* BL21 DE3 cultures carrying different signal peptide–MA fusion constructs induced with 1 mM IPTG for 16 h at 26 °C. 10 µg of total protein was loaded on a 12 % SDS-PAGE gel. *Lane M*: molecular weight marker, *1* uninduced PelB–MA fusion, *2* induced PelB–MA fusion, *3* uninduced FhuD–MA fusion, *4* induced FhuD–MA fusion, *5* uninduced MalE–MA fusion, *6* induced MalE–MA fusion, *7* uninduced TorA–MA fusion, *8* uninduced TorA–MA fusion, *9* uninduced DsbA–MA fusion, *10* induced DsbA–MA fusion, *11* uninduced MdoD–MA fusion, *12* induced MdoD–MA fusion, *13* uninduced YcdO–MA fusion, *14* induced YcdO–MA fusion, *15* uninduced OmpA–MA fusion, *16* induced OmpA–MA fusion. The position of maltogenic amylase (MA) in the strains that showed significant induction is indicated by an *arrowhead*

strains carrying pAEV01 (MalE–MA fusion), pAEV04 (DsbA–MA fusion) and pAEV05 (FhuD–MA fusion) constructs revealed expression of the 70 kDa maltogenic amylase, on a 12 % SDS-PAGE gel. DsbA–MA expression levels were the highest post-induction and no inducible MA expression was seen in the remaining fusions (Fig. 2). This data indicates that depending on the signal peptide partner, MA production yields can vary greatly. In fact fusion with a non-preferred signal can even result in the complete absence of the amylase in the culture pellet post-induction when detected by Coomassie staining of the gel. However, once a suitable signal peptide fusion has been established varying temperature post-induction does not significantly change the production pattern of MA.

Codon-optimized FhuD, MalE and DsbA signals translocate maltogenic amylase more efficiently to the *E. coli* periplasm compared to the PelB signal

BL21 DE3 *E. coli* carrying four different signal peptide–MA fusions, i.e., DsbA, MalE, FhuD and PelB were induced with 1 mM IPTG and incubated at 30 °C for 16 h. A BL21 DE3 strain carrying empty pET20b+ vector was included as the control. Analysis of the subcellular fractions separated using a fractionation protocol as described in "Materials and methods" was subsequently carried out to determine the localization of MA. We observed that upon fusion of MA to the MalE, FhuD and DsbA signal peptides, a significant amount of the protein is translocated to the periplasm compared to when MA is fused to the PelB signal (Fig. 3). In fact, MA was the predominant species in the periplasmic fraction when fused to MalE, FhuD and DsbA.

DsbA–MA, FhuD–MA and MalE–MA fusions show higher amylolytic activity compared to PelB–MA

To determine how various signal peptides affect the amylolytic activity of MA, the sonicated supernatant from all eight signal peptide–MA fusion carrying strains following induction with 1 mM IPTG and incubation at 30 °C was used for determination of maltogenic amylase activity using the glucose oxidase method [6]. An 18-, 30- and 50-fold higher induction, respectively, in terms of maltogenic amylase activity in the sonicated supernatant fraction was observed in the strains harboring the pAEV05 (FhuD–MA fusion), pAEV01 (MalE–MA fusion) and pAEV04 (DsbA–MA fusion) constructs compared to PelB–MA (Fig. 4). The strains harboring OmpA–MA, TorA–MA, MdoD–MA and YcdO–MA fusions did not show significant improvement in amylolytic activity over that of PelB–MA (Fig. 4). Results indicate that a much higher amount of functional maltogenic amylase is getting targeted into the periplasmic space compared to the parent plasmid in the case of strains harboring pAEV01, pAEV04 and pAEV05. MalE, FhuD and DsbA signal peptides thus appear to be far more efficient than the PelB signal sequence for periplasmic localization of recombinant maltogenic amylase in *E. coli*.

Discussion

Although *E. coli* is the most common heterologous host for recombinant protein expression, very few proteins are secreted by *E. coli* to the periplasmic space. Accumulation of recombinant proteins in the periplasmic space is

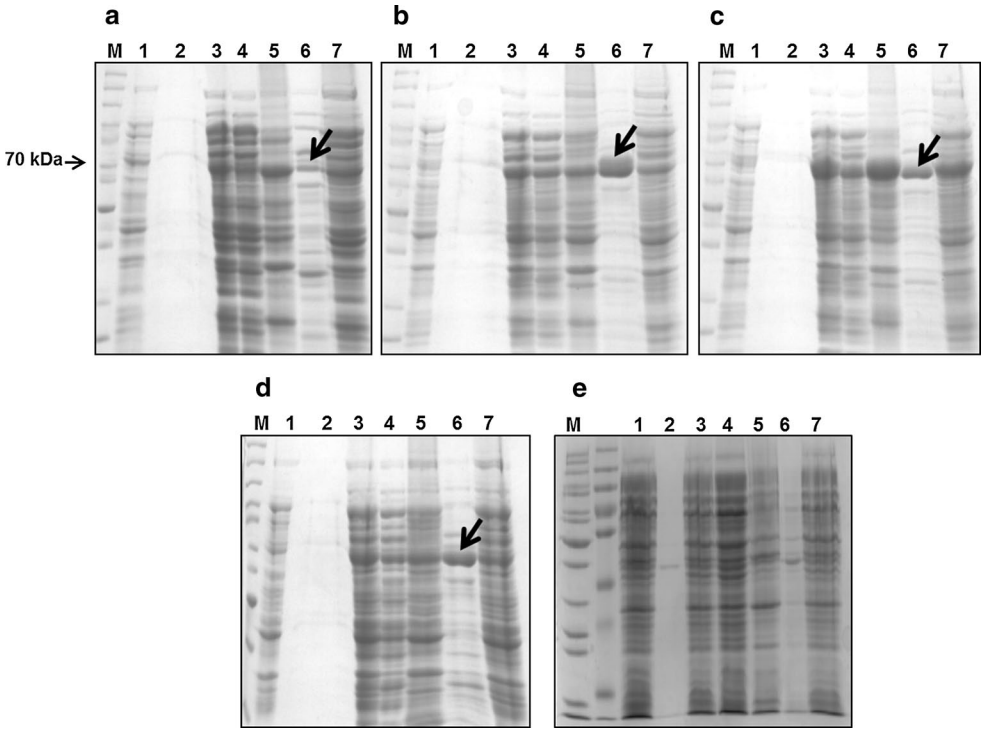

Fig. 3 Codon-optimized FhuD, MalE and DsbA signals translocate maltogenic amylase more efficiently to the *E. coli* periplasm compared to the PelB signal: cellular fractions corresponding to different cellular compartments of *E. coli* were loaded on a 12 % SDS-PAGE gel and proteins were visualized by Coomassie blue staining. **a** PelB–MA fusion, **b** DsbA–MA fusion, **c** FhuD–MA fusion, **d** MalE–MA fusion, **e** empty pET20b+ vector control. *M* protein molecular weight marker, *1* uninduced culture pellet, *2* induced culture supernatant, *3* induced culture pellet, *4* sonicated supernatant, *5* sonicated pellet, *6* periplasmic fraction, *7* cytoplasmic fraction. *Arrowheads* indicate position of MA on the gels

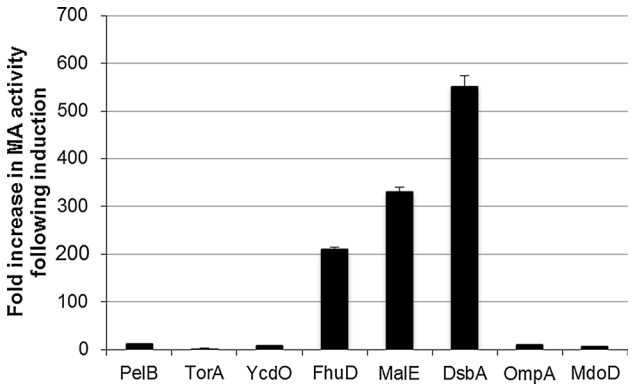

Fig. 4 DsbA–MA, FhuD–MA and MalE–MA fusions show higher amylolytic activity compared to PelB–MA: the amylolytic activity of maltogenic amylase was determined by the glucose oxidase method as described in "Materials and methods". Fold change in activity following induction of MA expression with 1 mM IPTG at 30 °C compared to an uninduced culture was plotted. Results of triplicate experiments are shown

usually accomplished by fusing an N-terminal sequence, commonly referred to as a signal/leader sequence, to the gene of interest [14]. Most commonly, SEC-dependent or TAT-dependent pathways are used for the translocation of fusion proteins through the inner membrane. However, a single export mechanism cannot translocate all recombinant proteins with the same efficiency. Thus, there is a need to analyze, incorporate and exploit new signal sequences in expression vectors to facilitate the selection of an appropriate export pathway which is most efficient for the production, effective processing and secretion of the desired protein.

We describe the generation of novel vectors that carry seven different codon-optimized signal peptides whose native versions are known to be targeted to different secretory pathways in *E. coli*. These vectors were used for expressing maltogenic amylase (MA) from *B. stearothermophilus* in *E. coli*. The expression of MA when fused to these seven signals was compared to that with PelB, a common fusion obtained upon using the popular pET20b+ vector for expression. We observed that following induction most signal-peptide MA fusions, including PelB, failed to yield any MA (as detected by Coomassie staining of PAGE gel) and MA production was observed only following fusion to its preferred partners, i.e., MalE, DsbA and FhuD. Previously, target proteins have been fused to signal peptide sequences (e.g., PelB, OmpA, DsbA, TorA and MalE) that follow different secretion pathways [1, 4, 5, 11, 17] to facilitate localization to the periplasm. However, not all of these signal peptide fusions were compatible for expression of MA. MA has not been previously expressed as a signal peptide fusion and this is the first report of it being directed to the periplasm as a result of fusion to certain preferred signals. We hypothesize that our vector platform could be used to similarly screen for favorable signal peptide partners for other recombinant proteins, potentially even toxic ones. Interestingly, only when the codon-optimized TorA and YcdO signals reported here were fused to another

heterologous protein, extracellular metalloprotease of *Serratia marcescens*, were we able to get BL21 DE3 transformants. Fusions to other signals yielded no transformants indicating the toxic nature of the metalloprotease (data not shown). This observation further underlines the strength of the platform described here.

Earlier attempts at purifying MA required sonication of cell cultures to isolate soluble protein [3]. We demonstrate that the DsbA, FhuD and MalE–MA fusions localize to the periplasm and MA can be purified merely by using an osmotic shock method. This greatly simplifies the purification process for recombinant MA expressed in *E. coli*. In addition, we also report that the amylase activity of MA was also enhanced when fused to these signals compared to PelB.

Recently, the effect of native *E. coli* signal peptides on homologous expression of *E. coli* thioredoxin has been reported [21]. The authors describe that the MalE signal was more destabilizing for the activity and folding of *E. coli* thioredoxin compared to the PelB signal. We, however, show that the contrast is true for the heterologous protein of choice, MA. This enzyme shows several fold increased activity as a MalE fusion when compared to the PelB fusion. The popular pET vector series are often the first choice for sub-cloning genes for overexpression in *E. coli*. Three of our codon-optimized signal peptides, FhuD, MalE and DsbA are capable of improving yield, activity and periplasmic localization of MA compared to the commercial PelB fusion. Significantly, we do not observe a bias for a particular secretion pathway while expressing MA. We observe that all three signals, i.e., post-translational (SEC)-specific (MalE), TAT + SEC signal peptide (FhuD) and co-translational (SRP)-specific (DsbA) are capable of efficiently translocating MA to the periplasmic space at the same time improving protein yield and activity over a commercial expression system. This improved activity could be attributed to better folding of protein targeted to the periplasm as a result of fusion to these three peptides.

Codons with a low usage and low abundance of corresponding tRNA are defined as non-optimal codons. Earlier reports demonstrate a striking abundance of non-optimal codons in the signal sequences of secretory proteins exported via the SEC and SRP pathways in *E. coli* [25, 26]. In fact when non-optimal codons in the signal sequence of maltose binding protein (MBP) were optimized, the authors observed a 20-fold reduction in maltose binding protein levels and affects folding of MBP [25]. The authors suggest a role for non-optimal codon usage in secretion by slowing the rate of translation across the N-terminal signal sequence to facilitate proper folding of the secreted protein [24]. We demonstrate that codon-optimized versions of the native signal peptides in *E. coli* are capable of not only effectively transporting recombinant MA across inner membranes, but also improving its yield over a commercial vector. The use of optimized codons also does not seem to affect folding since the amylolytic activity of MA was unaffected. In fact, the optimized signal peptide fusions showed better MA activity than a PelB fusion in the popular pET20b+ vector. The role of signal peptides in facilitating protein targeting to the periplasm is fairly well documented [7]. However, it is not very well known whether these signals also affect the recombinant protein's yield and activity. This to our knowledge is the first report demonstrating improved heterologous protein production and activity in *E. coli* using codon-optimized *E. coli* signal sequences.

To conclude, our findings highlight the need to identify the optimal signal peptide partner using a platform of vectors with different signal sequences to arrive at the most efficient combination. Identification of such a combination can not only ease the purification process, but also improve the functional properties of the target protein. The tough task of predicting an ideal signal peptide partner for new heterologous recombinants can be addressed by screening an array of several signal peptide fusions as we describe here.

References

1. Angelini S, Moreno R, Gouffi K, Santini C, Yamagishi A, Berenguer J, Wu L (2001) Export of *Thermus thermophilus* alkaline phosphatase via the twin-arginine translocation pathway in *Escherichia coli*. FEBS Lett 506:103–107
2. Baneyx F, Muhacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol 22:1399–1408
3. Cha HJ, Yoon HG, Kim YW, Lee HS, Kim JW, Kweon KS, Oh BH, Park KH (1998) Molecular and enzymatic characterization of a maltogenic amylase that hydrolyzes and transglycosylates acarbose. Eur J Biochem 253:251–262
4. Choi JH, Jeong KJ, Kim SC, Lee SY (2000) Efficient secretory production of alkaline phosphatase by high cell density culture of recombinant *Escherichia coli* using the Bacillus sp. endoxylanase signal sequence. Appl Microbiol Biotechnol 53:640–645
5. Choi JH, Lee SY (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl Microbiol Biotechnol 64:625–635
6. Diderichsen et al. (1986) Preparation of a maltogenic amylase enzyme United States Patent 4598048
7. Fekkes P, Driessen AJ (1999) Protein targeting to the bacterial cytoplasmic membrane. Microbiol Mol Biol Rev 63:161–173
8. Georgiou G, Segatori L (2005) Preparative expression of secreted proteins in bacteria: status report and future prospects. Curr Opin Biotechnol 5:538–545
9. Gold L (1990) Expression of heterologous proteins in *Escherichia coli*. Method Enzymol 185:11–14
10. Izard JW, Kendall DA (1994) Signal peptides: exquisitely designed transport promoters. Mol Microbiol 13:765–773
11. Lee SY, Jeong KJ (2006) US20067070989B2
12. Makrides SC (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiol Rev 60:512–538
13. Marino MH (1989) Expression systems for heterologous protein production. BioPharmacology 2:18–33

14. Mergulhão FJ, Summers DK, Monteiro GA (2005) Recombinant protein secretion in *Escherichia coli*. Biotechnol Adv 23:177–202
15. Olins PO, Lee SC (1993) Recent advances in heterologous gene expression in *Escherichia coli*. Curr Opin Biotechnol 4:520–525
16. Pugsley AP (1993) The complete general secretory pathway in gram-negative bacteria. Microbiol Rev 57:50–108
17. Qian ZG, Xia XX, Choi JH, Lee SY (2008) Proteome-based identification of fusion partner for high-level extracellular production of recombinant proteins in *Escherichia coli*. Biotechnol Bioeng 101:587–601
18. Rouet R, Lowe D, Dudgeon K, Roome B, Schofield P, Langley D, Andrews J, Whitfeld P, Jermutus L, Christ D (2012) Expression of high-affinity human antibody fragments in bacteria. Nat Protoc 7:364–373
19. Sargent F, Gohlke U, de Leeuw E, Stanley NR, Palmer T, Saibil HR, Berks BC (2001) Purified components of the *Escherichia coli* Tat protein transport system form a double layered ring structure. Eur J Biochem 268:3361–3367
20. Shokri A, Sanden AM, Larsson G (2003) Cell and process design for targeting of recombinant protein. Appl Microbiol Biotechnol 60:654–664
21. Singh P, Sharma L, Kulothangan RS, Adkar BV, Prajapati RS et al (2013) Effect of signal peptide on stability and folding of *Escherichia coli* thioredoxin. PLoS ONE 8(5):e63442
22. Stuart RA, Neupert W (2000) Making membranes in bacteria. Nature 406:575–577
23. Valent QA, Scotti PA, High S, de Gier JW, von Hejine G, Lentzen G, Wintermeyer W, Oudega B, Luirink J (1998) The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. EMBO J 17(9):2504–2512
24. Zalucki YM, Beachem IR, Jennings MP (2009) Biased codon usage in signal peptides: a role in protein export. Trends Microbiol 17:146–150
25. Zalucki YM, Jones CE, Ng PS, Schulz BL, Jennings MP (2010) Signal sequence non-optimal codons are required for the correct folding of mature maltose binding protein. Biochem Biophys Acta 1798:1244–1249
26. Zalucki YM, Shafer WM, Jennings MP (2011) Directed evolution of efficient secretion in the SRP-dependent export of TolB. Biochem Biophys Acta 1808:2544–2550