Available online at www.sciencedirect.com
ScienceDirect
Protein Expression and Purification 58 (2008) 23–31
Protein Expression & Purification
www.elsevier.com/locate/yprep

# Identification of an alternative translation initiation site for the *Pantoea ananatis* lycopene cyclase (*crtY*) gene in *E. coli* and its evolutionary conservation

Sung-Woo Kim [a,c,1], Woo-Hyuk Jung [a], Ji-Myung Ryu [a], Jae-Bum Kim [a], Hyung-Wook Jang [a], Young-Bae Jo [a,2], Joon-Ki Jung [b,*], Jung-Hoe Kim [c]

[a] *AceBiotech Co., Ltd., #114 Bio-Venture Center, KRIBB, Yuseong, Daejeon 305-806, Republic of Korea*
[b] *Division of Biotechnology R&BD, KRIBB, Yuseong, Daejeon 305-806, Republic of Korea*
[c] *Department of Biological Sciences, KAIST, Yuseong, Daejeon 305-701, Republic of Korea*

Received 22 September 2007, and in revised form 5 November 2007
Available online 19 November 2007

**Abstract**

Previous sequence analyses of the lycopene cyclase gene (*crtY*) from *Pantoea ananatis* revealed that translation of its protein product in *Escherichia coli* began at the ATG start codon. We found, however, that this enzyme could also be produced in *E. coli* without the ATG start codon present. Results of experiments using *crtY* mutants revealed that a GTG (Val) sequence, located in-frame and 24 bp downstream of the ATG, could act as a potential start codon. Additionally, a point-mutated GTA (Val), replaced from alternative GTG start codon, also displayed its potential as a start codon although the strength as a translation initiation codon was considerably weak. This finding suggests that non-ATG codons, especially one base pairing with the anticodon (3′-UAC-5′) in fMet-tRNA, might be also able to function as start codon in translation process. Furthermore, amino acid sequence alignment of lycopene cyclases from different sources suggested that a Val residue located within the N-terminus of these enzymes might be used as an alternative translation initiation site. In particular, presence of a conserved Asp, located in-frame and 12 bp upstream of potential start codon, supports this assumption in view of the fact that Asp (GAT or GAC) can function as part of the Shine–Dalgano sequence (AGGAGG).


*Keywords:* Alternative translation; GTG start codon; Lycopene cyclase; *Pantoea ananatis*; Translational reinitiation

Carotenoids are natural pigments that play important biological roles, including antioxidant activity, cancer prevention, and enhancement of immune responses [1–3]. To date, at least 700 carotenoids have been identified but most of these are biosynthetic intermediates that accumulate only in trace amounts, making it very difficult to extract sufficient material for their purification and application. Of those carotenoids, β-carotene ($C_{40}$) and lycopene ($C_{40}$) are industrially important carotenoids used as feed supplements for boosting vitamin A and antioxidant intake, respectively. Therefore, many researchers have focused on increasing carotenoid yields using metabolically engineered organisms.

Since the nucleotide sequences of carotenoid biosynthesis genes in the genus *Pantoea* have been identified [4,5], many studies were published describing production of carotenoids in non-carotenogenic microorganisms, including *Escherichia coli* [6,7], *Saccaromyces cerevisiae* [8,9], *Candida utilis* [10], etc. The genes involved in carotenoid biosynthesis in *Pantoea* species encompass six open reading frames, specifically *crtE*, *crtX*, *crtY*, *crtI*, *crtB*, and *crtZ*. The sequential reactions governed by these genes products

* Corresponding author. Fax: +82 42 860 4516.
*E-mail address:* jkjung@kribb.re.kr (J.-K. Jung).
[1] Present address: The Therapeutic Antibody Center (TTAC), KRIBB, Yuseong, Daejeon 305-806, Republic of Korea.
[2] Present address: 416, Cooperative Laboratory Center, Pukyong National University #599-1, Daeyeon 3, Nam-gu, Busan 608-737, Republic of Korea.

1046-5928/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2007.11.004

result in biosynthesis of various carotenoids [11]. For example, lycopene ($C_{40}$), an acylic isomer of β-carotene ($C_{40}$), is a red pigment generated via reactions catalyzed by the proteins encoded by *crtE*, *crtB*, and *crtI*. The lycopene cyclase (*crtY*) enzyme converts lycopene into β-carotene, a bright yellow pigment; subsequent reactions catalyzed by the *crtZ* and *crtX* gene products lead to zeaxanthin and zeaxanthin-β-D-diglucoside synthesis, respectively.

Most of the end products that accumulate in various organisms, including a variety of plant tissues, are cyclic carotenoids bearing either two β-ionone rings or one β-ionone ring and one ε-ionone ring [12]. The enzyme involved in these cyclization reactions is lycopene cyclase. Genes encoding lycopene cyclase have been cloned and sequenced from various microorganisms, *Pantoea ananatis* (old name: *Erwinia uredovora*) [5], *Pantoea agglormerans* (old name: *Erwinia herbicola*) [13,14], *Synechococcus* [15,16], *Blakeslea trispora* [17], *Streptomyces griceus* [18], *Erythrobacter longus* [19], etc. All of these microbial lycopene cyclases are comprised of approximately 400 amino acids with a molecular weight of ~43 kDa [20], whereas the equivalent enzymes from plants contain an additional N-terminal transit sequence of 100 amino acids. However, the similarity shared between these lycopene cyclase proteins is considerably low, with the exception of distinct conserved regions among the different amino acid sequence [21].

We constructed synthetic operons to enable synthesis of lycopene and β-carotene and used these to successfully produce both substances in *E. coli* (Fig. 1). However, we observed a significant difference, approximately 3-fold, between the yields of lycopene and β-carotene obtained via this approach (unpublished). The two lines (pTB-EIB vs. pTB-EIBY) harboring these operons differ only at the lycopene cyclase (*crtY*) gene in the β-carotene-producing vectors (Fig. 1A). To determine if inserting an additional *crtY* gene increased carotenoid production in this system,

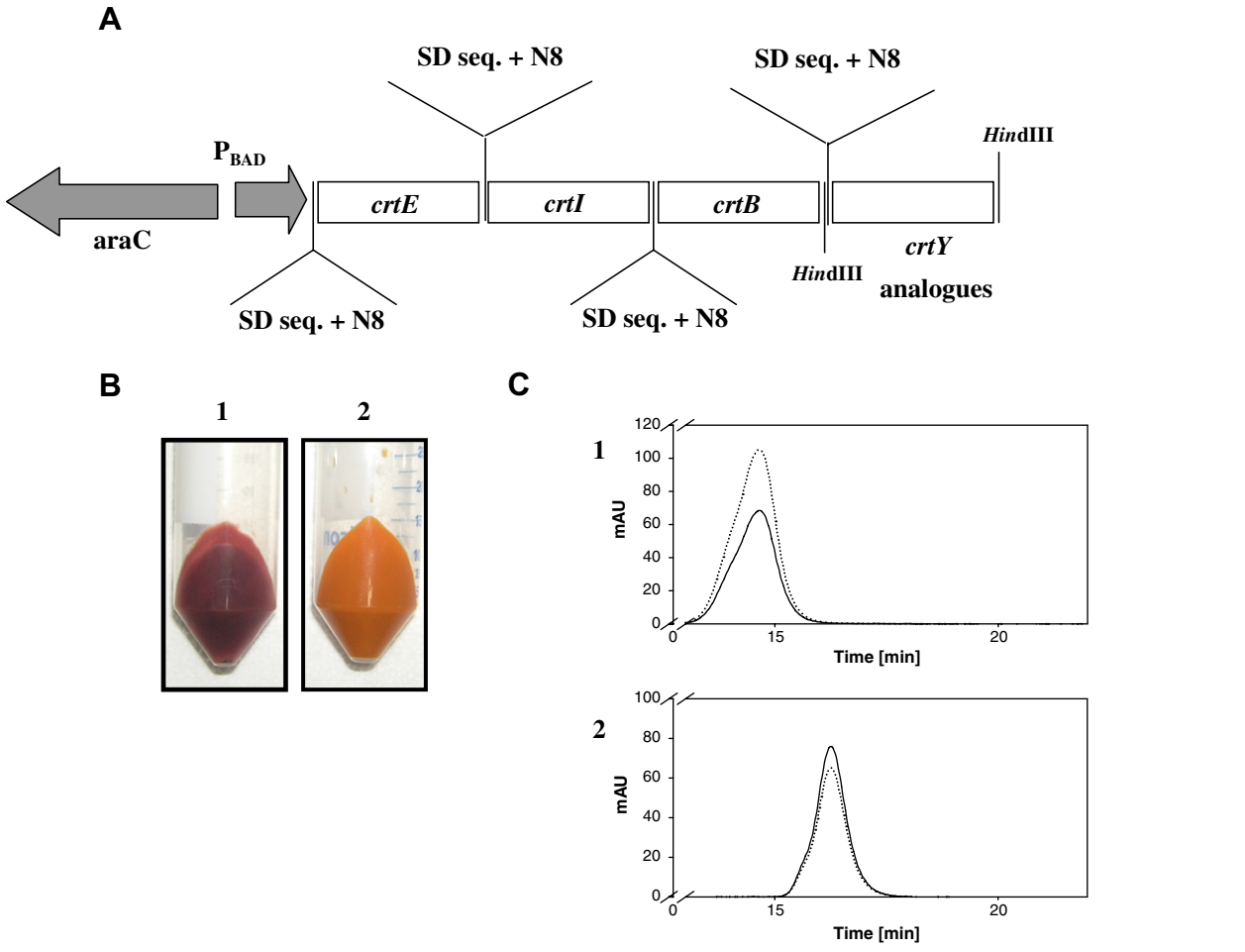


Fig. 1. Detailed descriptions of synthetic *crt* operon (A), cell harvests producing carotenoids (B), and HPLC profiles of standard carotenoids (C). In (A), SD seq. + N8: TAAGGAGGATTACAAT. The expression system was designed to express all genes of interest at the same level by inserting an identical SD sequence between each gene. In (B), 1, recombinant *E. coli* producing lycopene; 2, recombinant *E. coli* producing β-carotene. In (C), 1, Lycopene standard ($\lambda_{max} = 470$, Sigma # L9879); 2, β-carotene standard ($\lambda_{max} = 450$, Sigma # C9750). Lycopene and β-carotene standards were eluted at about 14 and 17 min, respectively. Dotted line and solid line indicate the values at $\lambda = 470$ and $\lambda = 450$, respectively.

several *crtY* mutants were constructed using site-directed mutagenesis. While characterizing the role(s) of *crtY* using these mutants, we unexpectedly found an alternative translation start codon of *crtY* in *E. coli*. To date, it has been known that lycopene cyclase from *P. ananatis* is translated from ATG start codon [5,20]. In this study, we introduced various *crtY* mutants into *E. coli* to pinpoint this potential start codon, through which we ultimately found an internal GTG translation initiation codon. Furthermore, by comparing the amino acid sequences of lycopene cyclases across multiple organisms, we identified this potential translation start codon among them.

## Materials and methods

### *Bacterial strains, vectors, and culture conditions*

*Escherichia coli* DH5α was used for basic gene manipulations and production of lycopene or β-carotene. A cluster of genes (*crtE*, *crtI*, *crtB*, and *crtY*) involved in carotenoid biosynthesis has been cloned from *P. ananatis* (ATCC(R) 19321). The vector constructs and primers used in this study are described in Table 1. Recombinant cells harboring vector constructs were cultivated in 2 × YT (1.6% (w/v) bacto tryptone, 1% (w/v) yeast extract, and 0.5% (w/v) NaCl) medium supplemented with ampicillin (100 μg/ml) at 180 rpm and 37 °C for 24 h. Recombinant cell cultures exhibited red (lycopene) or yellow pigments (β-carotene) on their cell membrane without any induction [22]. These strains therefore could be easily divided into lycopene- or β-carotene-producing bacteria (Fig. 1B).

### *Construction of carotenoid-producing vectors containing crtY mutants*

Site-directed mutagenesis was carried out to construct a series of *crtY* mutants. Seven vectors were generated in stepwise fashion for identification of a potential translation initiation codon within *crtY*. The *i-MaxII* DNA polymerase (Cat. No. 25261, iNtRON, Korea) enzyme and T-vec crtY template were used for site-directed mutagenesis. The PCR conditions were as follows: after heating at 94 °C for 5 min, the following cycle was repeated 15 times: 94 °C, 45 s; 55 °C, 45 s; 72 °C, 7 min, and finally heated at 72 °C for 7 min. Each PCR product acquired using the designed primers was digested using *Dpn*I at 37 °C for 1 h and then transformed into *E. coli* DH5α. After sequence analysis, *Hin*dIII-digested fragments from the correct clones were integrated into dephosphorylated pTB-EIB (*Hin*dIII-digested). Seven constructs were transferred into *E. coli* DH5α and each 100 transformant then cultured in a test tube. We then counted the cell harvests producing lycopene or β-carotene after centrifugation.

### *Extraction and quantification of lycopene and β-carotene*

To quantify lycopene and β-carotene, these compounds were first extracted from the recombinant cells using acetone. Specifically, an aliquot of acetone was added to the harvested cells. After resuspending the harvested cells by vortexing, the samples were incubated at 55 °C for 15 min and then centrifuged at 13,000*g*. The supernatants were then transferred into dark glass vials and the cell pel-

Table 1
List of vectors, primers, and templates used in this study

| Vector name | Descriptions [features of *crtY* mutants] | Reference |
|---|---|---|
| T-vec crtY | pGEM-T vec harboring *P. ananatis crtY*. | [22] |
| pTB-EIB | Lycopene-producing vector. | [22] |
| pTB-EIBY | pTB-EIB harboring *crtY* from T-vec crtY. | [22] |
| pTB-EIBY($M_1$) | pTB-EIB harboring *crtY* ($M_1$). [Met1 (ATG) → Ile1 (ATC)] | This study |
| pTB-EIBY($M_1M_2$) | pTB-EIB harboring *crtY* ($M_1M_2$). [Met1 (ATG) and Met30 (ATG) → Ile1 (ATC) and Ile30 (ATC)] | This study |
| pTB-EIBYstop1 | pTB-EIB harboring *crtY* (stop1). [Tyr5 (TAT) → TAA] | This study |
| pTB-EIBYstop2 | pTB-EIB harboring *crtY* (stop2). [Trp40 (TGG) → TGA] | This study |
| pTB-EIBYstop13 | pTB-EIB harboring *crtY* (stop13). [Tyr5 (TAT) and Gln24 (CAG) → TAA and TAG, respectively] | This study |
| pTB-EIBY(V) | pTB-EIB harboring *crtY* (V). [Val10 (GTG) → Val10 (GTA)] | This study |
| pTB-EIBY(L) | pTB-EIB harboring *crtY* (L). [Val10 (GTG) → Leu10 (TTG)] | This study |
| **Primer name** | **Sequences (5′ → 3′)** | **Products (templates)** |
| F-($M_1$) | GAGGATTACAATATCCAACCGCATTATGAT | T-vec crtY($M_1$) |
| R-($M_1$) | ATCATAATGCGGTTGGATATTGTAATCCTC | (T-vec crtY) |
| F-($M_2$) | CAACCTGATATTCGTATTTTGCTTATC | T-vec crtY($M_1M_2$) |
| R-($M_2$) | GATAAGCAAAATACGAATATCAGGTTG | (T-vec crtY($M_1$)) |
| F-(stop1) | ATCCAACCGCATTCCGATCTGATTCTCGTG | T-vec crtYstop1 |
| R-(stop1) | CACGAGAATCAGATCTTAATGCGGTTGCAT | (T-vec crtY) |
| F-(stop2) | GAATCATACGTGATCATTTCACCACGATG | T-vec crtYstop2 |
| R-(stop2) | CATCGTGGTGAAATGATCACGTATGATTC | (T-vec crtY) |
| F-(stop13) | TGCGTCTTTAGCAGCAGCAACCTGATATG | T-vec crtYstop13 |
| R-(stop13) | CATATCAGGTTGCTGCTGCTAAAGACGCA | (T-vec crtYstop1) |
| F-(V) | GATCTGATTCTCGTAGGGGCTGGACTC | T-vec crtY(V) |
| R-(V) | GAGTCCAGCCCCTACGAGAATCAGATC | (T-vec crtY) |
| F-(L) | GATCTGATTCTCTTGGGGGCTGGACTC | T-vec crtY(L) |
| R-(L) | GAGTCCAGCCCCCACGAGAATCAGATC | (T-vec crtY) |

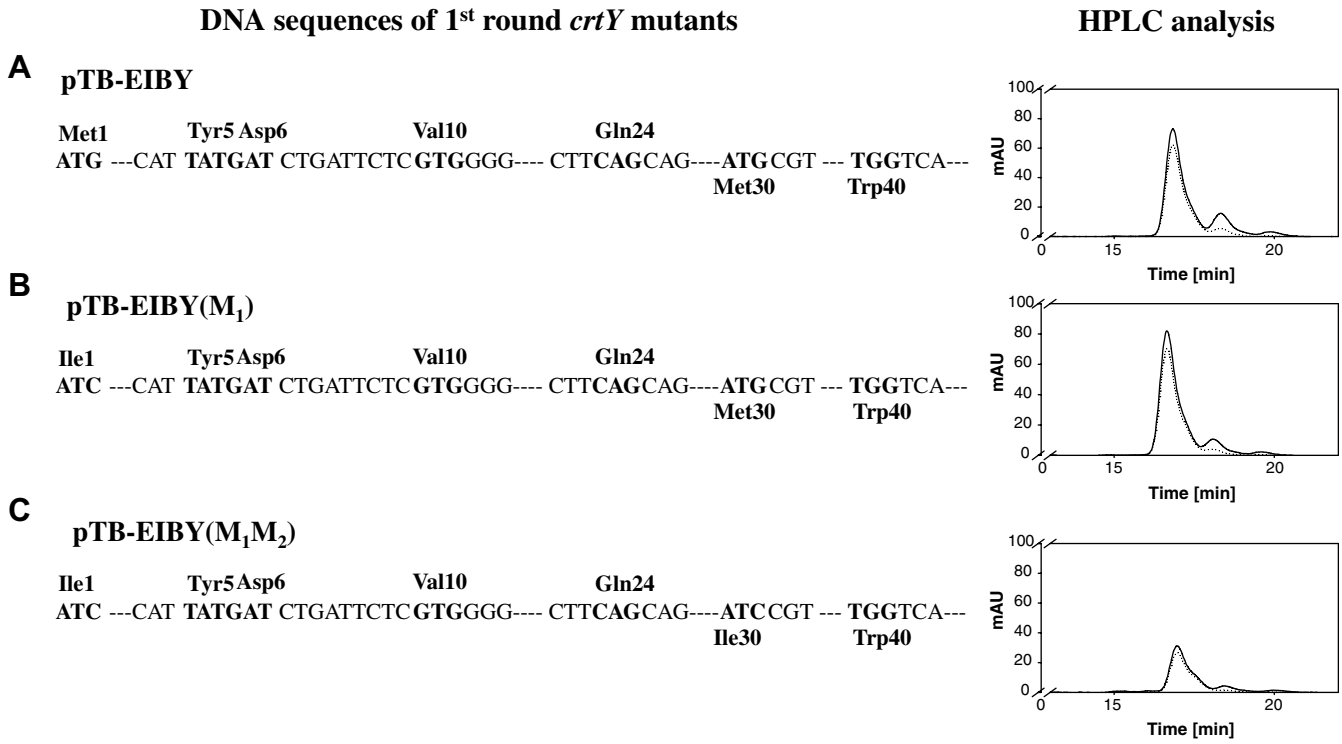


Fig. 2. The nucleotide sequences of 1st round *crtY* mutants and HPLC analyses characterizing cell extracts from recombinant *E. coli* harboring the mutants-inserted synthetic operons. (A) pTB-EIBY (native *crtY* from *P. ananatis*), (B) pTB-EIBY($M_1$), and (C) pTB-EIBY($M_1M_2$). Bold sequence letters are key features in this study. In HPLC profiles, dotted line and solid line indicate the values at $\lambda = 470$ and $\lambda = 450$, respectively.

lets treated with additional acetone to completely extract the remaining carotenoids. This process was repeated until no more carotenoids remained in the cell pellets. To quantify the carotenoids, we followed previously described methods using a HPLC (ÄKTA Basic, Amersharm, Sweden) [22]. To calculate dry cell weight (DCW),[3] culture broth was harvested by centrifuging at 13,000*g* and the resulting pellets washed twice with distilled water. Dry cell weight was then measured by weighing the dried cells after 2 days of incubation at 105 °C.

### *HPLC analysis*

To monitor changes in carotenoid profiles of cell extracts dependant upon the vector constructs harbored by each sample, HPLC analysis was performed using a Shiseido C18 5 μm column (Cat. No. 90104) with an ethanol/methanol (70:30, by vol.) mobile phase, run at a rate of 1 ml/min.

## Results

### *Carotenoid production of 1st round expression vectors, including crtY mutants by replacing ATG with ATC*

Our strategy involved replacing ATG translation initiation codon with something inert for construction of an untranslatable *crtY*. There are eight methionine residues in the *P. ananatis* lycopene cyclase protein, yet apparently only two of these residing at the N-terminus are potential translation start codon(s). Considering that the other methionines are close to the C-terminus of lycopene cyclase, it is unlikely that our *crtY* mutant, with over approximately one third of its lycopene cyclase N-terminus deleted, functions properly. Therefore, we preferentially focused upon site-directed mutagenesis of the first (Met1) and the second methionine (Met30) in lycopene cyclase. As a result, we generated pTB-EIBY ($M_1$) and pTB-EIBY ($M_1M_2$) in which the first and the second methionine were appropriately replaced with ATC (Ile) (Fig. 2B and C). However, introduction of these constructs unexpectedly still resulted in lycopene cyclase activity, with β-carotene produced as a result. According to HPLC profiles in Fig. 2, similarly to control (pTB-EIBY), those mutants displayed a major peak at about 17 min and a minor peak at about 18 min, which indicate a β-carotene and a cis isomer of β-carotene, respectively. Lycopene cyclase is able to convert lycopene as substrate efficiently via the intermediate γ-carotene to all-trans β-carotene and a cis isomer [20], therefore these results demonstrated that lycopene was completely converted into all-trans β-carotene and a cis isomer by lycopene cyclase.

Considering the point that weaker pairing (two rather than three base pairs) with fMet-tRNA is part of the reason that translation is less efficient when an alternative start codon replaces ATG [23], it is possible that ATC was used

[3] *Abbreviations used:* DCW, dry cell weight; RBS, ribosome binding site; SD, Shine–Dalgano.

instead of ATG in Fig. 2. In addition, according to previous reports which performed quantitative analysis of ribosome binding sites in *E. coli*, both ATA and ACG can function as initiation codons under some growth conditions [24]. In this sense, our result strongly suggested two types of possibility: the first is ATC might play a role as a translation initiation codon and the corresponding gene could be expressed. The second is an internal start codon within *crtY* might function in the absence of the ATG start.

*Carotenoid production of 2nd round expression vectors, including untranslatable crtY mutants via addition of a stop codon*

In order to confirm whether ATC functioned as a start codon or not in our experiments, three vectors, including pTB-EIBYstop1, pTB-EIBYstop2, and pTB-EIBYstop13, were developed (Fig. 3) and tested. Results using these constructs revealed that recombinant *E. coli* (pTB-EIBYstop1) still produced β-carotene, whereas *E. coli* harboring pTB-EIBYstop2 or pTB-EIBYstop13 produced lycopene only (Fig. 3). β-Carotene production in *E. coli* (pTB-EIBYstop1) suggested that the potential translation initiation site should be downstream of Tyr5, implying that ATC didn't function as a start codon in the pTB-EIBY($M_1$) and pTB-EIBY($M_1M_2$) constructs previously tested. Furthermore, lycopene production in *E. coli* (pTB-EIBYstop2) implied that the potential translation initiation site should be upstream of Trp40. Additionally, the result obtained using *E. coli* (pTB-EIBYstop13) revealed that Met30 is not the alternative initiation site and further narrowed the range where the translation initiation site may sit within the *crtY* mutant sequence. Considering these data, we identified a potentially internal translation start codon residing between Tyr5 and Gln24 of *crtY*.

*Identification of a potential translation start site within crtY*

Assuming that a potential translation initiation codon, which is not an ATG start codon, exists between Tyr5 and Gln24 within the *crtY* sequence, we further dissected this sector of the gene. Using this approach, we found a candidate internal start codon, namely Val10 (GTG). We then constructed two varieties of Val10-derived mutants, pTB-EIBY(V) and pTB-EIBY(L) which were modified version of pTB-EIBYstop1 (Fig. 4). According to Fig. 4A, *E. coli* harboring pTB-EIBY(V), where GTG (Val) was replaced with GTA (Val), produced a considerable amount of lycopene as well as β-carotene. Although it was assumed that pTB-EIBY(V) should still retain lycopene cyclase activity, the enzyme generated from a base replacement was not sufficient to convert all lycopene into β-carotene. This result implied that GTA could be used as a start codon in addition to Val10 (GTG), as a potential translation initiation site. Another result of substituting TTG for GTG in pTB-EIBY(L) clearly indicated that Val10 is an alternative translation start site. We confirmed that *E. coli* containing pTB-EIBY(L), where GTG (Val) was replaced with TTG (Leu), only produced lycopene (Fig. 4B). However, the result with TTG was unexpected

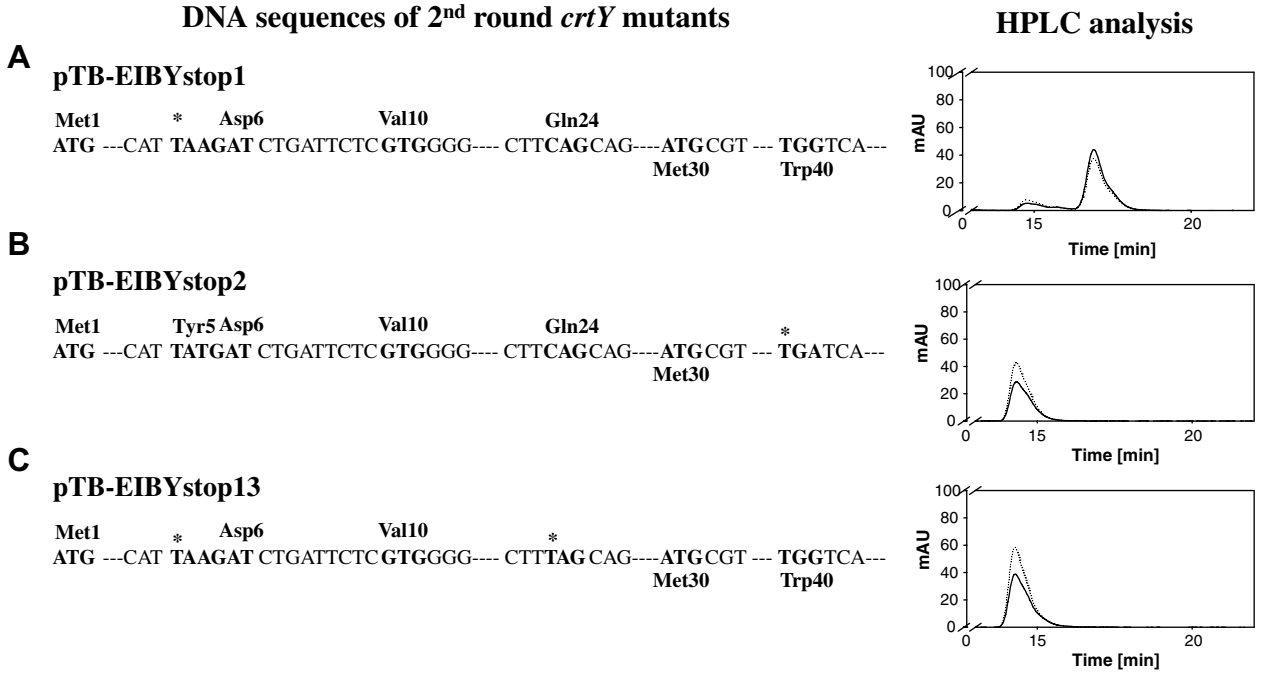


Fig. 3. The nucleotide sequences of 2nd round *crtY* mutants and HPLC analyses characterizing cell extracts from recombinant *E. coli* harboring the mutants-inserted synthetic operons. (A) pTB-EIBYstop1, (B) pTB-EIBYstop2, (C) pTB-EIBYstop13, *stop codon. Bold sequence letters are key features in this study. In HPLC profiles, dotted line and solid line indicate the values at $\lambda = 470$ and $\lambda = 450$, respectively.

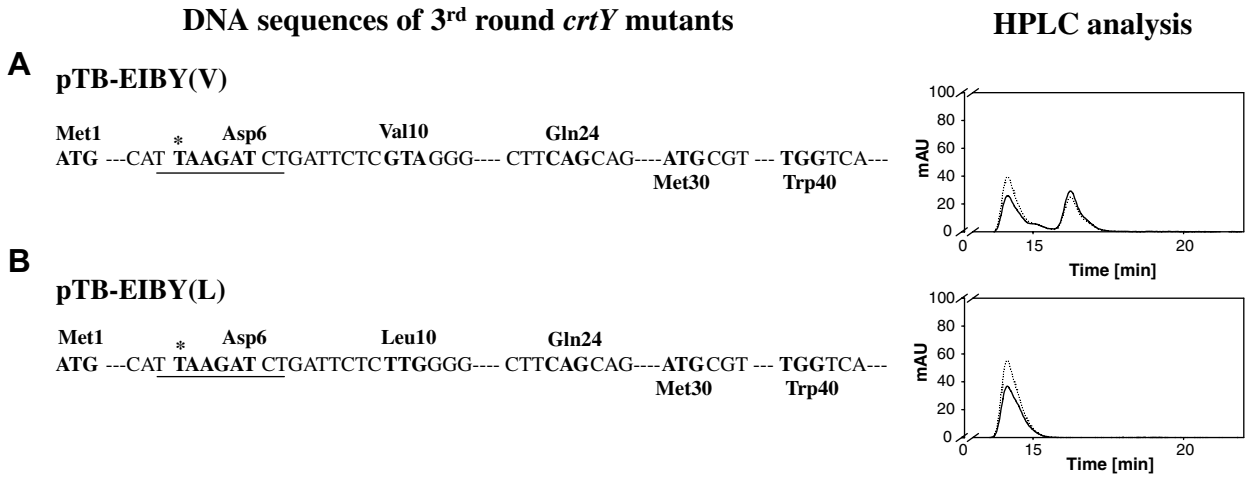


Fig. 4. The nucleotide sequences of 3rd round *crtY* mutants and HPLC analyses characterizing cell extracts from recombinant *E. coli* harboring the mutants-inserted synthetic operons. (A) pTB-EIBY(V), (B) pTB-EIBY(L), *stop codon. Bold sequence letters are key features in this study. The underline indicates a potential ribosome binding site [RBS]. In HPLC profiles, dotted line and solid line indicate the values at $\lambda = 470$ and $\lambda = 450$, respectively.

since this codon is also a valid albeit weak start codon at least in *E. coli*.

### *Analysis of carotenoid production from expression vectors containing crtY mutants*

Results of carotenoid yields and final products among *crtY* mutant-harboring carotenoid expression vectors are summarized in Table 2. Carotenoid yields from most of vectors except pTB-EIBY($M_1$) decreased up to 45% considering control (pTB-EIBY); however, we found that untranslatable *crtY* mutants within pTB-EIBYstop2, pTB-EIBYstop13, pTB-EIBY(V), and pTB-EIBY(L), affected the increase (~2-fold) of lycopene yields considering control (pTB-EIB, data not shown). From these results, we assumed that the introduction of untranslatable *crtY* following *crtEIB* operon consequently affected mRNA secondary structure, enhancing ribosome binding on *crtEIB* or its stability. As a result, increased translation efficiency or mRNA stability enhanced lycopene production. In addition, we observed one interesting result, which was the difference of carotenoid yield among constructs, especially between 1st and 2nd or 1st and 3rd round constructs (Table 2). It is notable that even if 2nd and 3rd round constructs both contain an insertion of stop codon, those constructs still can produce carotenoids. This observation could be explained by the phenomenon of translational reinitiation, in which a start site is exposed for translation dependent upon a prior initiation on the sequence, in addition to a termination event in the vicinity of the new start site [25]. Considering 2nd and 3rd round constructs had these requisites for translational reinitiation (Figs. 3 and 4), the decrease of carotenoid yield in 2nd or 3rd round constructs is understandable. Furthermore, of translational reinitiation constructs—pTB-EIBYstop1, pTB-EIBY(V), and pTB-EIBY(L)—pTB-EIBYstop1 harboring the GTG as translational reinitiation codon showed the best translational efficiency of *crtY* (β-carotene [mg/g DCW] in Table

Table 2
Summary of production results from various *crtY* mutants

| Vector name | Carotenoids (mg/g DCW) | | Number of colored pellet after centrifugation of cultures | | |
|---|---|---|---|---|---|
| | Lycopene | β-Carotene | Red | Mixed[a] | Yellow |
| pTB-EIBY | — | 1.3 ± 0.1 | — | — | 100 |
| 1st round | | | | | |
| pTB-EIBY($M_1$) | — | 1.2 ± 0.2 | 1 | — | 99 |
| pTB-EIBY($M_1M_2$) | — | 0.8 ± 0.3 | — | — | 100 |
| 2nd round | | | | | |
| pTB-EIBYstop1 | <0.1 | 0.6 ± 0.1 | 5 | — | 95 |
| pTB-EIBYstop2 | 0.7 ± 0.1 | — | 100 | — | — |
| pTB-EIBYstop13 | 0.8 ± 0.2 | <0.1 | 100 | — | — |
| 3rd round | | | | | |
| pTB-EIBY(V) | 0.6 ± 0.1 | 0.3 ± 0.1 | 46 | 36 | 18 |
| pTB-EIBY(L) | 0.8 ± 0.1 | — | 98 | — | 2 |

[a] Mixed means the pellet displaying both colors.

2). Though the GTG is not a standard start codon (ATG), considering it is known to valid start codon and more frequently used than GTA or TTG codon in *E. coli*, this is a reasonable result. However, it was interesting that the GTA codon which is rarely used as start codon showed better translational efficiency of *crtY* than TTG, a valid but minor initiation codon.

Another notable finding in Table 2 is that transformants containing the same vector constructs exhibited differential carotenoid production. In the case of pTB-EIBY($M_1$), pTB-EIBYstop1, pTB-EIBY(V), and pTB-EIBY(L), we confirmed differential carotenoid production when we cultured those transformants in tubes and counted after centrifugation. These results imply that a potential alternative translation process occurred elsewhere within the *crtY* sequence, and as a result lycopene cyclase activity existed with various spectrum depending on alternative start codon strength. Therefore, reduced lycopene cyclase activity via this alternative process may account for the observed differential accumulation of carotenoids. On the basis of this, we could also easily investigate this alternative start codon strength by counting the culture harvests which were either colored yellow, red or both colors. Particularly, the number of yellow pellets in Table 2 clearly displayed their relative strength as alternative start codons.

In addition to identifying an alternative translation initiation site, from these studies we concluded that nine amino acids at the lycopene cyclase N-terminus did not significantly contribute to cyclization of lycopene because the result of consistent β-carotene yields between recombinant cell lines harboring pTB-EIBY($M_1$) and pTB-EIBY indicated that an alternative GTG start codon instead of the normal ATG could be adequate to produce β-carotene in our expression system (Table 2).

### *Amino acid sequence alignments of lycopene cyclases*

We carried out amino acid sequence alignments to assess feasibility of a functional GTG start codon in lycopene cyclases from other microorganisms. Fig. 5 illustrates that a highly conserved region at the N-termini of these lycopene cyclases. In particular, the results indicated that *Streptomyces griseus*, *P. ananatis*, *P. agglomerans*, and *E. longus* all contained conserved Asp, Val, and Gly residues.

## Discussion

It is well established that in prokaryotes alternative start codons are used. In this research, we unexpectedly found an alternative start codon of *P. ananatis crtY* in *E. coli* in the process of constructing untranslatable *crtY*. To date, it has been known that lycopene cyclase from *P. ananatis* is translated from ATG start codon [5,20]. More interestingly, potential of GTA as a translation initiation codon in Fig. 4A is meaningful in that a weak pairing, especially one base pairing with anticodon (3′-UAC-5′) in fMet-tRNA, could be used as a start codon in *E. coli* and worth considering that many genes starting from this weak pairing could be discovered. In our opinion, this finding has never been reported yet.

As mentioned in "Introduction", increasing the yields of lycopene by using these *crtY* mutants was the one of key purposes for this study. According to Table 2, untranslatable *crtY* mutants exhibited positive effects for enhancing lycopene production considering pTB-EIB (average lycopene yield of pTB-EIB is 0.45 mg/g DCW, data not shown). Nevertheless, there were still the difference of carotenoid yields between lycopene and β-carotene. Recently, a research group has reported similar results with ours [26]. In the present study, we were unable to completely resolve this enigma. It is likely that another factor influences the difference of carotenoid yields between lycopene and β-carotene.

On the other hand, HPLC profiles in Figs. 3 and 4 presented a diagnostic for low vs high levels of lycopene cyclase. For example, a low level gives a higher lycopene peak and a smaller β-carotene peak; conversely, a high level gives little or no lycopene and a large β-carotene peak.

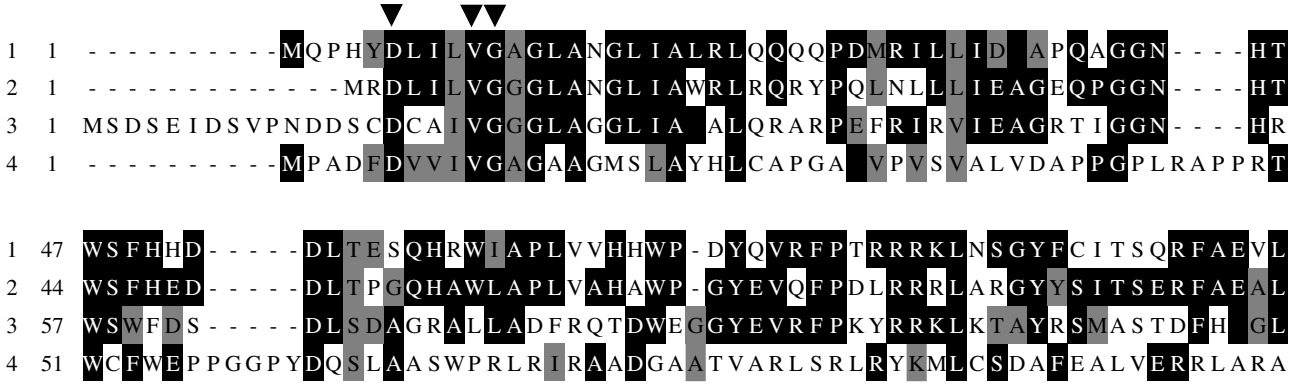


Fig. 5. A partial presentation of the amino acid sequence alignments of lycopene cyclases from various microorganisms. Alignments were generated using CLUSTAL W 1.8 comparing amino acid sequences of lycopene cyclases from various microorganisms. 1, *P. ananatis* (GenBank Accession No. BAA14126); 2, *P. agglomerans* (GenBank Accession No. AAA64980); 3, *E. longus* (GenBank Accession No. BAA20275); 4, *S. girseus* (GenBank Accession No. AAG28705). Arrows indicate key features of the alternative translation initiation process.

The difference of lycopene cyclase expression level seems to be caused by translational reinitiation efficiency depending on start codon of translational reinitiation site. Similar results are found in Table 2, which demonstrated the presence of transformants accumulating different carotenoids, despite their harboring of the same constructs. This phenomenon could be also explained by decreased lycopene cyclase activity resulting from alternative start codons, i.e., GTG, GTA, TTG, in *E. coli*. Because lycopene cyclase activities are various in a colony by colony, even though transformants contain the same construct, depending on translational reinitiation efficiency induced by those weak start codons, such a phenomenon is understandable considering that insufficient lycopene cyclase can't totally convert lycopene into β-carotene. In this point of view, although we didn't observe differential carotenoid production in transformants harboring pTB-EIBY($M_1M_2$) in Table 2, they have a potential to express differential accumulation of carotenoids as long as they have alternative translation process.

The amino acid sequence alignment data suggested that lycopene cyclases from other microorganisms might also possess an alternative translation initiation site (Fig. 5). The results indicated that all microorganisms contained highly conserved region at the N-termini of these lycopene cyclases, i.e., Asp (GAT for *P. ananatis* and *P. agglomerans*, GAC for *E. longus* and *S. griseus*), Val (GTG for *S. griseus* and *P. ananatis*, GTC for *P. agglomerans*, and GTT for *E. longus*) and Gly (GGG for *P. ananatis*, and GGC for others). Considering a weak base pairing could be used as a start codon like GTA in Fig. 4B, *P. agglomerans* and *E. longus* might be also in agreement with our result. Previously published studies support these features even though the researches were carried out in eukaryotes; for instance Kozak has reported that recognition of ATG as well as alternative initiator codons are augmented by a G in position +4 [27]. Also, Boeck and Kolakofsky [28] reported that GTG could be used efficiently for ribosomal initiation when the codon following was GAT. According to our findings, the codon following a potential translation initiation codon is GGG (Gly); this result perfectly agreed with these previous reports.

Another notable result presented in Fig. 5 is the existence of a conserved Asp residue common to all of these microorganisms. We propose that this Asp residue may play a key role in interactions with the 16S rRNA. Because this Asp residue is located within a potential ribosome binding site (RBS) region, and the codons of Asp (GAT or GAC) are a part of Shine–Dalgano (SD) sequence (AGGAGG), this particular residue may serve a critical function in the alternative translation process. If this assumption is true, the presence of these alternative translation processes in multiple organisms might be evolutionarily conserved as auxiliary measures for guaranteeing a cyclic carotenoid pool which may function by providing protection against photooxidative damages.

Future work is required to test this theory, involving extensive site-directed mutagenesis covering pertinent sites within these various lycopene cyclase enzymes. Other worthwhile studies could include elucidating functional differences between native and mutant *crtY*.

## Acknowledgment

This study was supported by a Grant from the Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (A020547).

## References

[1] S.T. Mayne, Beta-carotene, carotenoids, and disease prevention in humans, FASEB J. 10 (1996) 690–701.
[2] G.A. Armstrong, Genetics of eubacterial carotenoid biosynthesis: a colorful tale, Annu. Rev. Microbiol. 51 (1997) 629–659.
[3] S.K. Clinton, Lycopene: chemistry, biology, and implications for human health and disease, Nutr. Rev. 56 (1998) 35–51.
[4] K.L. Perry, T.A. Simonitch, K.J. Harrison-Lavoie, S.T. Liu, Cloning and regulation of *Erwinia herbicola* pigment genes, J. Bacteriol. 168 (1986) 607–612.
[5] N. Misawa, M. Nakagawa, K. Kobayashi, S. Yamano, Y. Izawa, K. Nakamura, K. Harashima, Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*, J. Bacteriol. 172 (1990) 6704–6712.
[6] G. Sandmann, M. Albrecht, G. Schnurr, O. Knorzer, P. Boger, The biotechnological potential and design of novel carotenoids by gene combination in *Escherichia coli*, Trends Biotechnol. 17 (1999) 233–237.
[7] S.W. Kim, J.D. Keasling, Metabolic engineering of the non-mevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production, Biotechnol. Bioeng. 72 (2001) 408–415.
[8] N. Misawa, H. Shimada, Metabolic engineering for the production of carotenoids in non-carotenogenic bacteria and yeasts, J. Biotechnol. 59 (1998) 169–181.
[9] P.C. Lee, C. Schmidit-Dannert, Metabolic engineering towards biotechnological production of carotenoids in microorganisms, Appl. Microbiol. Biotechnol. 60 (2002) 1–11.
[10] Y. Miura, K. Kondo, T. Saito, H. Shimada, P.D. Fraser, N. Misawa, Production of the carotenoids lycopene, β-carotene, and astaxanthin in the food yeast *Candida utilis*, Appl. Environ. Microbiol. 64 (1998) 1226–1229.
[11] G. Sandmann, Carotenoid biosynthesis and biotechnological application, Arch. Biochem. Biophys. 385 (2001) 4–12.
[12] G.A. Armstrong, Genetics of eubacterial carotenoid biosynthesis: a colorful tale, Annu. Rev. Microbiol. 51 (1997) 629–659.
[13] B.S. Hundle, M. Alberti, V. Nievelstein, P. Beyer, G.A. Armstrong, D.H. Burke, J.E. Hearst, Functional assignment of *Erwinia herbicola* Eho10 carotenoid genes expressed in *Escherichia coli*, Mol. Gen. Genet. 245 (1994) 406–416.
[14] K.Y. To, E.M. Lai, L.Y. Lee, T.P. Lin, C.H. Hung, C.L. Chen, Y.S. Chang, S.T. Liu, Analysis of the gene cluster encoding carotenoid biosynthesis in *Erwinia herbicola* Eho13, Microbiology 140 (1994) 331–339.
[15] F.X. Cunningham Jr., D. Chamovitz, N. Misawa, E. Gantt, J. Hirschberg, Cloning and functional expression in *Escherichia coli* of a cyanobacterial gene for lycopene cyclase, the enzyme that catalyzes the biosynthesis of β-carotene, FEBS Lett. 328 (1993) 130–138.
[16] F.X. Cunningham Jr., B. Pogson, Z. Sun, K. McDonald, A.D. DellaPenna, E. Gantt, Functional analysis of the β and ε lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation, Plant Cell 8 (1996) 1613–1626.
[17] M. Rodríguez-Sáiz, B. Paz, J.L. de la Fuente, M.J. López-Nieto, W. Cabri, J.L. Barredo, *Blakeslea trispora* genes for carotene biosynthesis, Appl. Environ. Microbiol. 70 (2004) 5589–5594.

[18] G. Schumann, H. Nürnberger, G. Sandmann, H. Krügel, Activation and analysis of cryptic *crt* genes for carotenoid biosynthesis from *Streptomyces griseus*, Mol. Gen. Genet. 252 (1996) 658–666.
[19] H. Matsumura, H. Takeyama, E. Kusakabe, J.G. Burgess, T. Matsunaga, Cloning, sequencing and expressing the carotenoid biosynthesis genes, lycopene cyclase and phytoene desaturase, from the aerobic photosynthetic bacterium *Erythrobacter longus* sp. strain Och101 in *Escherichia coli*, Gene 189 (1997) 169–174.
[20] G. Schnurr, N. Misawa, G. Sandmann, Expression, purification and properties of lycopene cyclase from *Erwinia uredovora*, Biochem. J. 315 (1996) 869–874.
[21] P. Krubasik, G. Sandmann, Molecular evolution of lycopene cyclase involved in the formation of carotenoids with ionone end groups, Biochem. Soc. Trans. 28 (2000) 806–810.
[22] S.W. Kim, J.B. Kim, W.H. Jung, J.H. Kim, J.K. Jung, Overproduction of β-carotene from metabolically engineered *Escherichia coli*, Biotechnol. Lett. 28 (2006) 897–904.
[23] J.K. Sussman, E.L. Simons, R.W. Simons, *Escherichia coli* translation initiation factor 3 discriminates the initiation codon *in vivo*, Mol. Microbiol. 21 (1996) 347–360.
[24] S. Shinedling, M. Gayle, D. Pribnow, L. Gold, Mutations affecting translation of the bacteriophage T4 *rIIB* gene cloned in *Escherichia coli*, Mol. Gen. Genet. 207 (1987) 224–232.
[25] L. Song, W. Mandecki, E. Goldman, Expression of non-open reading frames isolated from phage display due to translation reinitiation, FASEB J. 17 (2003) 1674–1681.
[26] S.H. Yoon, H.M. Park, J.E. Kim, S.H. Lee, M.S. Choi, J.Y. Kim, D.K. Oh, J.D. Keasling, S.W. Kim, Increased beta-carotene production in recombinant *Escherichia coli* harboring an engineered isoprenoid precursor pathway with mevalonate addition, Biotechnol. Prog. 23 (2007) 599–605.
[27] M. Kozak, Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6, EMBO J. 16 (1997) 2482–2492.
[28] R. Boeck, D. Kolakofsky, Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis, EMBO J. 13 (1994) 3608–3617.