ELSEVIER

Gene 189 (1997) 169–174

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

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*

Haruo Matsumura [b], Haruko Takeyama [a], Etsuko Kusakabe [b], J. Grant Burgess [a], Tadashi Matsunaga [a,*]

[a] *Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan*
[b] *Asahi Chemical Industry Co., Ltd., 2-1 Samejima, Fuji-city, Shizuoka 416, Japan*

Received 1 March 1996; received in revised form 19 June 1996; accepted 25 June 1996; Received by J.R. Kinghorn

Abstract

Two genes which encode the enzymes lycopene cyclase and phytoene desaturase in the aerobic photosynthetic bacterium *Erythrobacter longus* sp. strain Och101 have been cloned and sequenced. The gene for lycopene cyclase, designated *crtY*, was expressed in a strain of *Escherichia coli* which contained the *crtE*, *B*, *I* and *Z* genes encoding geranylgeranyl pyrophosphate synthase, phytoene synthase, phytoene desaturase, and β-carotene hydroxylase, respectively. As a result, zeaxanthin production was observed in *E. coli* transformants. In addition, expression of the *E. longus* gene *crtI* for phytoene desaturase in *E. coli* containing *crtE* and *B* resulted in the accumulation of lycopene in transformants. Zeaxanthin and lycopene were also determined by mass spectrum. Nucleotide sequence similarities between *E. longus crtY* gene and other microbial lycopene cyclase genes are 40.2% (*Erwinia herbicola*), 37.4% (*Erwinia uredovora*) and 22.9% (*Synechococcus* sp.), and those between phytoene desaturase genes are 50.3% (*E. herbicola*), 54.7% (*E. uredovora*) and 39.6% (*Rhodobacter capsulatus*). © 1997 Elsevier Science B.V.

Keywords: β-Carotene; Lycopene; *crtI*; *crtY*

1. Introduction

Carotenoids are an important group of natural pigments which are widely distributed in living organisms. They can function as protective compounds against photooxidative damage, as light harvesting pigments and as pigments which impact color to the living tissues. Certain cyclic carotenoids, such as β-carotene, are precursors of vitamin A in animals and are of current interest as nutritional factors important for cancer prevention (Lambert et al., 1990).

Genes which encode the biosynthetic enzymes for carotenoids have been cloned and sequenced from the photosynthetic bacterium *Rhodobacter capsulatus* (Armstrong et al., 1989), and the phytopathogenic bacteria *Erwinia herbicola* and *Erwinia uredovora* (Hundle et al., 1993; Misawa et al., 1990). However,

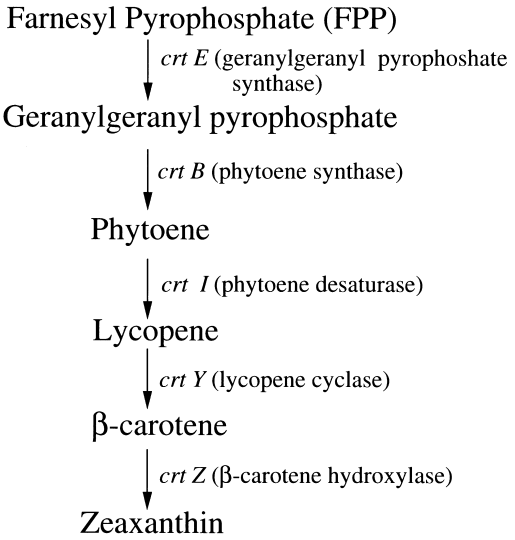

Fig. 1. Scheme of the carotenoid biosynthesis pathway.

* Corresponding author. Tel. +81 423 887020; Fax +81 423 857713; e-mail: tmatsuna@cc.tuat.ac.jp

PII S0378-1119(96)00788-3

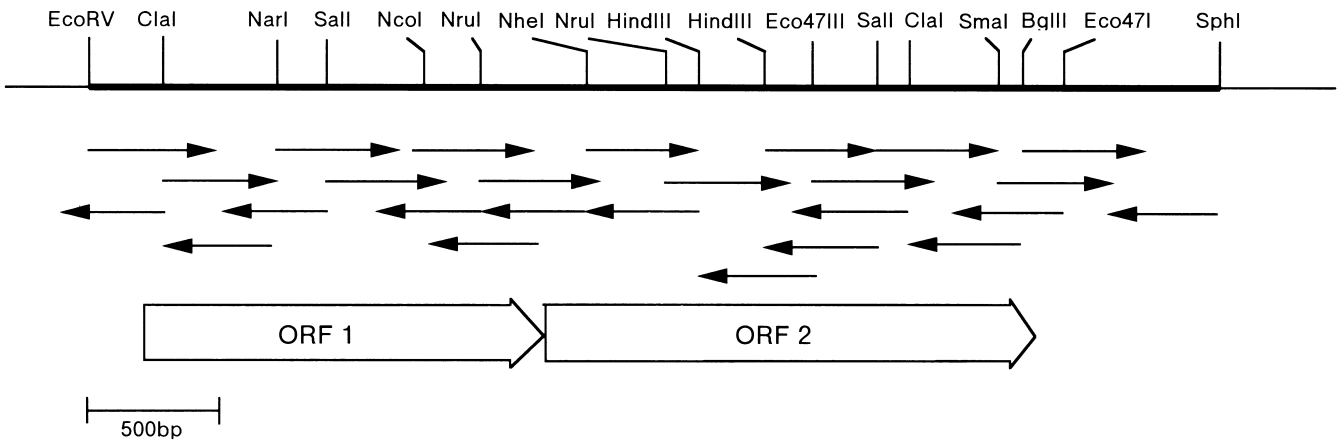

Fig. 2. Cloning and subcloning of DNA fragments containing the *crtY* and *crtI* regions from *E. longus*. Partial restriction map is indicated above the bar. Horizontal arrows below the bar indicate the determined sequences.

```
GATATCGCCATGCCACAGCTTCACATATCCGCATTCACAGATCGGCGGGCGGTTCATGCC      60
AAGCAGGATACCAACGGTCGCCAGTGCGATAATGATGGAGACGATAATCGTCATCCGGGT     120
CGGTTTGAGTAATGCCATGCGCAAAATCCTGAATCAAAAGCGGTAAACGCGCAAGAGTAT     180
TCGAGGAATACCCACTGTAGCGTTGCCCCCACACTGCGTGTGTCCTAGGGCCTTTACGCA     240

GATGAGCGACTCAGAAATCGATAGCGTCCCCAATGACGATAGTTGCGACTGCGCAATCGT     300
 M  S  D  S  E  I  D  S  V  P  N  D  D  S  C  D  C  A  I  V
 crtY
TGGCGGCGGACTTGCTGGCGGGTTGATTGCGCTTGCGCTCCAACGTGCGCGGCCCGAATT     360
 G  G  G  L  A  G  G  L  I  A  L  A  L  Q  R  A  R  P  E  F

TCGCATCCGCGTGATCGAGGCAGGGCGCACCATCGGCGGCAATCACCGGTGGAGCTGGTT     420
 R  I  R  V  I  E  A  G  R  T  I  G  G  N  H  R  W  S  W  F

TGACAGCGACCTCTCCGACGCCGGCCGTGCGCTACTTGCCGACTTTCGCCAGACCGATTG     480
 D  S  D  L  S  D  A  G  R  A  L  L  A  D  F  R  Q  T  D  W

GGAGGGCGGATACGAGGTGCGCTTTCCCAAATATCGCCGCAAGCTGAAGACCGCCTATCG     540
 E  G  G  Y  E  V  R  F  P  K  Y  R  R  K  L  K  T  A  Y  R

CTCGATGGCATCGACCGATTTCCACGAAGGGCTTTTGCGCGCTCTGCCCGAAGGATCGGT     600
 S  M  A  S  T  D  F  H  E  G  L  L  R  A  L  P  E  G  S  V

AATCCTGGGGCGCAAAGCGGTGGGTTTGGACGCGCGCGGCGTGGATTTGGCGCCGTCGCA     660
 I  L  G  R  K  A  V  G  L  D  A  R  G  V  D  L  A  P  S  Q

ATATGGCCCCGCAACCCGCATCAACGCGCGCAGTGTCATCGACTGCCGCAGCTTCAAACC     720
 Y  G  P  A  T  R  I  N  A  R  S  V  I  D  C  R  S  F  K  P

AAGCGCGCATCTCAAGGGCGGCTGGCAGGTGTTTCTTGGCCGACATATGCGGCTGCAAGA     780
 S  A  H  L  K  G  G  W  Q  V  F  L  G  R  H  M  R  L  Q  E

ACCGCACGGGGTGGAAAATCCGGTCATCATGGACGCAACCGTCGACCAGCTTGCGCCGCA     840
 P  H  G  V  E  N  P  V  I  M  D  A  T  V  D  Q  L  A  P  H

CGGTAATGGCGGTTCATACCGGTTCGTCTATGTTCTTCCCTTGGGAAGCCACGATGTCTT     900
 G  N  G  G  S  Y  R  F  V  Y  V  L  P  L  G  S  H  D  V  F

TATCGAAGACACCTATTACGCCGATGACCCGCTGCTTGACCGCAATGCCTTGTCGGGCCG     960
 I  E  D  T  Y  Y  A  D  D  P  L  L  D  R  N  A  L  S  G  R

GATAGACCAATATGCCCGCGCAAACGGTTGGGAGAACGGCACGCCCGTTCATCACGAAGC    1020
 I  D  Q  Y  A  R  A  N  G  W  E  N  G  T  P  V  H  H  E  A

AGGCGTCTTGCCGGTCCTGACGGGCGGCGATTTTTCCGCCTATCAGGACGAAGTGCGCAT    1080
 G  V  L  P  V  L  T  G  G  D  F  S  A  Y  Q  D  E  V  R  I

TCCCGGCGTTGCCATTGCGGGCGCGCGCGGCGGGTTTACCCATCCGCTGACCAGCTACAC    1140
 P  G  V  A  I  A  G  A  R  G  G  F  T  H  P  L  T  S  Y  T

CATGTGCGTCGCGGTCGAAAACGCGCTTGCCATGGCCGAGCAACCTGACCTCTCGGGCGA    1200
 M  C  V  A  V  E  N  A  L  A  M  A  E  Q  P  D  L  S  G  E

GCAATTGGCGGCCTTTTTTGACAGCCGCGCACGCCGCCATTGGTCAAAGACGGGATACTA    1260
 Q  L  A  A  F  F  D  S  R  A  R  R  H  W  S  K  T  G  Y  Y

CCGGCTGCTTGCGCGTTTCTTGTTCTTCGCGGCCAAGCCGGAGAAGCGCGTCAAGGTGTT    1320
 R  L  L  A  R  F  L  F  F  A  A  K  P  E  K  R  V  K  V  F

CCAACGCTTTTACGGACTTCGCGAAGGGTTGATCGAGCGGTTCTATGCCGCGCGCTCAAA    1380
 Q  R  F  Y  G  L  R  E  G  L  I  E  R  F  Y  A  A  R  S  N

CACCTTCGATAAGGTGCGCGTCCTATGGGGGGAGCCCCCCGTAGCTATACACTCGGCCAT    1440
 T  F  D  K  V  R  V  L  W  G  E  P  P  V  A  I  H  S  A  I

CCTGGCCATGTTCAAATCGGGTCCGGCGCTCAAGTCGGAAAAATCCGACAGGGGGGTCGC    1500
 L  A  M  F  K  S  G  P  A  L  K  S  E  K  S  D  R  G  V  A

TCAGGCGGCGCTCGATGAAGAATTGCAAACGGAGAAAAGGCCATGAACGCCGATCAAAAC    1560
 Q  A  A  L  D  E  E  L  Q  T  E  K  R  P  ...
                                          M  N  A  D  Q  N
                                          crtI

ATCGCTACAGGGCTCAACTTTGCGCCAGCCAATACTGGCGAGCGCGGCATTAATCCGGTG    1620
 I  A  T  G  L  N  F  A  P  A  N  T  G  E  R  G  I  N  P  V

ATCGCCGAAAAATACAAAGGCCGCACCGCCTGTGTGATCGGTTCCGGTTTTGGCGGCTTG    1680
 I  A  E  K  Y  K  G  R  T  A  C  V  I  G  S  G  F  G  G  L
                          *  *  *  *  *  *  *  *  *  *  *  *
```

Fig. 3. Nucleotide and amino-acid sequence of *E. longus* sp. Och101 carotenoid synthetic genes. The ORF for *crtY* is from nt 242 to 1546. The ORF for *crtI* is from nt 1553 to 3123. Start codons are underlined. Conserved domains at the N- and C-terminals of *crtI* from other organisms are indicated by asterisks. The nucleotide sequences has been submitted to the GenBank®/EMBL Data Bank with accession numbers for *crtY*: D83513, and for *crtI*: D83514.

genes which encode lycopene cyclase (Fig. 1), a key enzyme which converts the acyclic carotenoid lycopene

```
GCGCTAGCACTGCGGCTGCAATCGCATGGCATTCAAACGACCATCGTCGAAGCGCGCGAC  1740
A  L  A  L  R  L  Q  S  H  G  I  Q  T  T  I  V  E  A  R  D
*  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

AAGCCCGGTGGCCGCGCCTATTTCTGGGAAAAAGACGGCTTTACCTTCGATGCTGGCCCC  1800
K  P  G  G  R  A  Y  F  W  E  K  D  G  F  T  F  D  A  G  P

ACGGTCATCACCGACCCGCCGTGTTTGAAAGAACTGTGGGAGCTGACCGGCCACGACATT  1860
T  V  I  T  D  P  P  C  L  K  E  L  W  E  L  T  G  H  D  I

TCCGAAGATGTCGAGCTGATGAAGGTTCACCCTTTCTACCGCCTCAACTGGCCCGATGGC  1920
S  E  D  V  E  L  M  K  V  H  P  F  Y  R  L  N  W  P  D  G

ACAAACTTCGATTATTCGAACGTTGATGAGGAATTGAACGCCGAAATCGCGAAGCTCAAT  1980
T  N  F  D  Y  S  N  V  D  E  E  L  N  A  E  I  A  K  L  N

CCTGACGATGTGATCGGCTATCAAAAATTCCTCGAATATTCGGCGCGCGTGCACGAGGAA  2040
P  D  D  V  I  G  Y  Q  K  F  L  E  Y  S  A  R  V  H  E  E

GGCTATGTGAAGCTTGGCACGGTGCCGTTCCTCGATTTCAAGTCGATGCTGAAAGCCGCC  2100
G  Y  V  K  L  G  T  V  P  F  L  D  F  K  S  M  L  K  A  A

CCTGCCCTTGTTAAAGAGCGCGCATGGCGCAGCGTTTACGATATGGTCTCAAGCTACATC  2160
P  A  L  V  K  E  R  A  W  R  S  V  Y  D  M  V  S  S  Y  I

AAGGATGAGCGCCTGCGCGAAGCGTTCAGCTTCCACACGCTGCTTGTCGGCGGCTCGCCG  2220
K  D  E  R  L  R  E  A  F  S  F  H  T  L  L  V  G  G  S  P

ATGAAGACCAGCGCCATTTATGCGTTGATCCACAAGCTTGAAAAAGACGGCGGTGTCTGG  2280
M  K  T  S  A  I  Y  A  L  I  H  K  L  E  K  D  G  G  V  W

TGGGCGCGCGGCGGGACCAACCGGTTGATCGCCGGAATGGTGCGCCATTTTGAACGCCTC  2340
W  A  R  G  G  T  N  R  L  I  A  G  M  V  R  H  F  E  R  L

GGCGGCACGATGCGCATCGGCGATCCGGTGGTTCAGGTCCACACCCAAGGGACCAAAGCG  2400
G  G  T  M  R  I  G  D  P  V  V  Q  V  H  T  Q  G  T  K  A

ACCGAGGTTGAAACGAAGAGCGGTTGGAAAGAGCGCTTTGACGCGGTGTGTTCAAACGCC  2460
T  E  V  E  T  K  S  G  W  K  E  R  F  D  A  V  C  S  N  A

GACATCATGCACTCTTACAAGGAACTTCTGGGCGAATCCGACCGTGGCAGAAAATACGCT  2520
D  I  M  H  S  Y  K  E  L  L  G  E  S  D  R  G  R  K  Y  A

AAGTCATTGGCTCGCAAAAGCTATTCGCCTTCGCTATTCGTCGTACACTTTGGGCTTGAG  2580
K  S  L  A  R  K  S  Y  S  P  S  L  F  V  V  H  F  G  L  E

GGGTCGTGGCCCGGTATTGCCCACCACATGATCCTGTTTGGCCCACGTTACAAGGAACTG  2640
G  S  W  P  G  I  A  H  H  M  I  L  F  G  P  R  Y  K  E  L

GTCGACGACATCTACAAGCACGGCGTTCTGCCGCAGGATTTTTCGATCTATCTTCACCAC  2700
V  D  D  I  Y  K  H  G  V  L  P  Q  D  F  S  I  Y  L  H  H

CCGACCGTCACCGACCCATCGATGGCGCCCAAGGGCATGAGCACATTCTACGCGCTTGTC  2760
P  T  V  T  D  P  S  M  A  P  K  G  M  S  T  F  Y  A  L  V

CCCGTCGCCCACCTTGGCAAGATGCCGATTGATTGGGACGTCGAAGGACCCAAGTTTGAA  2820
P  V  A  H  L  G  K  M  P  I  D  W  D  V  E  G  P  K  F  E

AAGGCGATTTTGGACGAGATCGGTCGCCGCCTGATCCCCGACATCCACGACCGGATCGTC  2880
K  A  I  L  D  E  I  G  R  R  L  I  P  D  I  H  D  R  I  V

ACCAAATTCAGCTACGCACCAAAGGACTTTCAGGCAGACCTCAACGCCCATATGGGCAGC  2940
T  K  F  S  Y  A  P  K  D  F  Q  A  D  L  N  A  H  M  G  S

GCGTTCAGCCTTGAGACGGTCCTGTGGCAAAGCGCCTACATGCGCGGCCACAACCGCGAC  3000
A  F  S  L  E  T  V  L  W  Q  S  A  Y  M  R  G  H  N  R  D
                          *  *  *  *  *  *  *  *  *

GATGTGATCGACAATTTCTACCTCGTGGGCGCAGGGACACACCCGGGCGCTGGTATCCCC  3060
D  V  I  D  N  F  Y  L  V  G  A  G  T  H  P  G  A  G  I  P
*  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

GGAGTGGTCGGTAGCGCGAAGGCAACGGCGGGGCTGATGCTTGAAGATCTGTCGGTCAAA  3120
G  V  V  G  S  A  K  A  T  A  G  L  M  L  E  D  L  S  V  K
*  *  *  *  *  *  *  *  *  *  *

TAATCGGGTATGATGTCCCTTTTACTCGCCGCCCAGCTCGCTGCGACAGTCGCAAATCTG  3180
...

TCGCACACGGTGGCGACACCGGAAAATCAGAGCGAAAGTTCGCAAGCGTTGGAAGAGCCG  3240
GCGAACGGCAGCGCT                                               3255
        Eco47III
```

Fig. 3. (*Continued*)

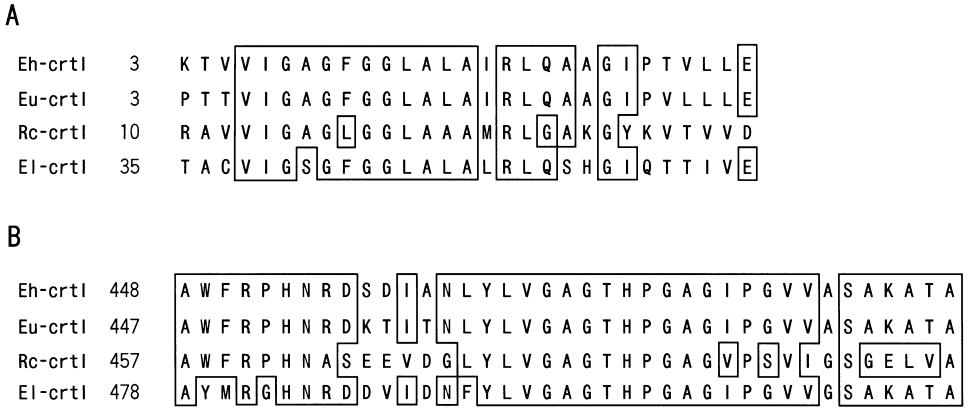

Fig. 4. Sequence of the conserved N-terminal (**A**) and C-terminal (**B**) regions in phytoene desaturases from *E. herbicola* (Eh-crtI, Armstrong et al., 1990), *E. uredovora* (Eu-crtI, Misawa et al., 1990), *Rhodobacter capsulatus* (Rc-crtI, Bartley and Scolnik, 1989) and *E. longus* (El-crtI, this work). Conserved aas are boxed.

into the cyclic carotenoid β-carotene, have not been isolated from photosynthetic bacteria. In the oxygenic cyanobacteria, where β-carotene is an essential component of the photosystem, a lycopene cyclase gene has recently been characterized (Cunningham et al., 1993, 1994).

Here we describe the cloning and sequencing of genes encoding phytoene desaturase and lycopene cyclase from the aerobic photosynthetic bacterium *Erythrobacter longus* in an effort to understand the organization and expression of the carotenoid biosynthesis genes in this organism. *E. longus* has an interesting carotenoid composition and has been found to produce about 20 different kinds of carotenoids such as β-carotene and monocyclic carotenoids such as rubixanthan (Takaichi et al., 1990).

2. Results and discussion

2.1. Cloning of the lycopene cyclase and phytoene desaturase genes

A gene library was constructed from chromosomal *Erythrobacter longus* DNA which was partially digested with *Sau*3AI and size fractionated to give fragments which were 6–20 kb in size. The DNA was ligated into the *Bam*HI site of the vector λEMBL3. The resulting gene library contained about 1.5×10^5 (95%) individual clones. To clone the phytoene desaturase genes from *E. longus*, a 1 kb *Bgl*II-*Pst*I fragment containing the phytoene desaturase gene from *Erwinia herbicola* was used as a probe for heterologous hybridization experiments. This fragment encodes the C-terminal region of the *E. herbicola crtI* gene which is highly conserved in other bacterial *crtI* genes. This fragment also encodes part of the N-terminal region of the phytoene synthase gene from *E. herbicola*.

Table 1
Comparison of codon usage among the carotenoid synthesis genes from *E. longus*, *R. capsulatus* and *E. herbicola*

Amino acid	Codon	Codon usage		
		E. longus	*R. capsulatus*	*E. herbicola*
Ile	ATT	1.0	0.3	1.6
	ATC	3.4	3.7	1.9
	ATA	0.2	–	0.3
Leu	TTA	–	–	0.2
	TTG	1.9	0.8	0.4
	CTT	2.3	1.9	1.2
	CTC	1.5	1.2	2.0
	CTA	0.4	0.0	0.5
	CTG	2.6	6.1	6.5
Val	GTT	1.1	0.7	0.6
	GTC	3.1	3.2	1.4
	GTA	0.3	<0.1	0.9
	ATA	2.2	3.8	2.8

–, not observed.
[a] *crtY* and *crtI* (this work; total 962 codons = 100%).
[b] *crtA*, *crtB*, *crtC*, *crtD*, *crtE*, *crtF*, *crtI* and *crtK* (Armstrong et al., 1989; toal 3038 codons = 100%).
[c] *crtB*, *crtE*, *crtI*, *crtZ*, *crtY* and *crtX* (Hundle et al., 1993).

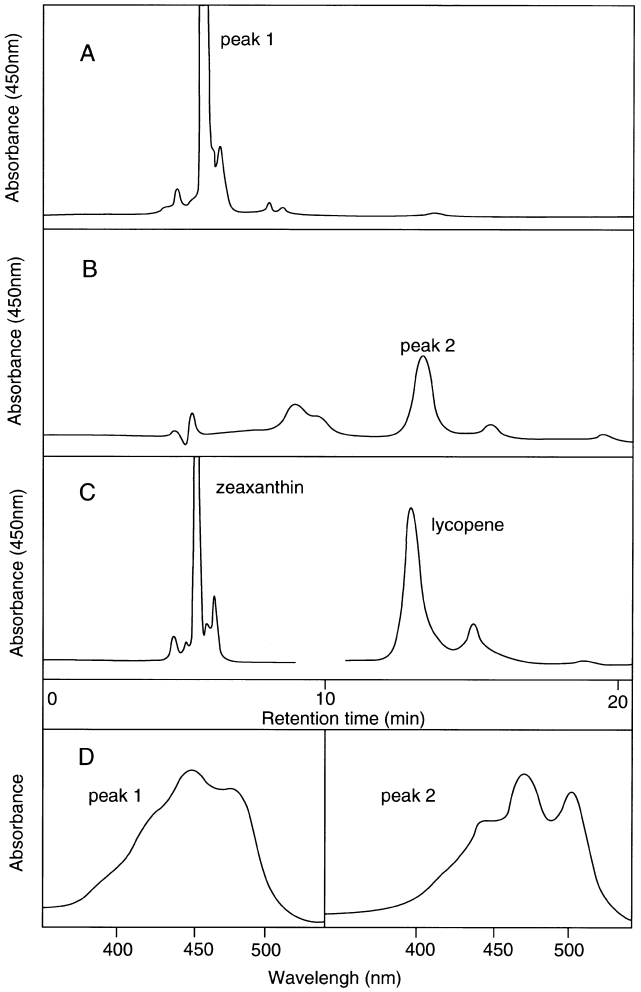

Fig. 5. HPLC analysis of carotenoids in the transformed *E. coli* having (**A**) pAYL and (**B**) pAYP. (**C**) gives traces of standard carotenoids, zeaxanthin and lycopene. (**D**) shows the spectra of the elution peaks.

Hybridization was carried out in 5 × SSC/30% formamide at 37°C for 16 h. The filters were washed twice in 2 × SSC at room temperature, twice in 0.1 × SSC at 42°C. Then positive clones were obtained after screening about 1.1×10^3 plaques. Restriction mapping of purified DNA from the positive clones showed that a common 4 kb *Sal*I fragment was present (Fig. 2). One clone was selected for further analysis and subcloned into M13 for sequencing.

2.2. Nucleotide sequence of the Erythrobacter longus crtY *and* crtI *gene*

The complete nucleotide sequence (3255 bp) of the *Eco*RV-*Eco47* III fragment was determined and found to contain two open reading frames 1305 and 1581 bp in length (Fig. 3). The smaller ORF, which begins at nt 242 and ends at nt 1546, appears to encode a 435 aa polypeptide, and nucleotide sequence similarities between this smaller ORF and the lycopene cyclase

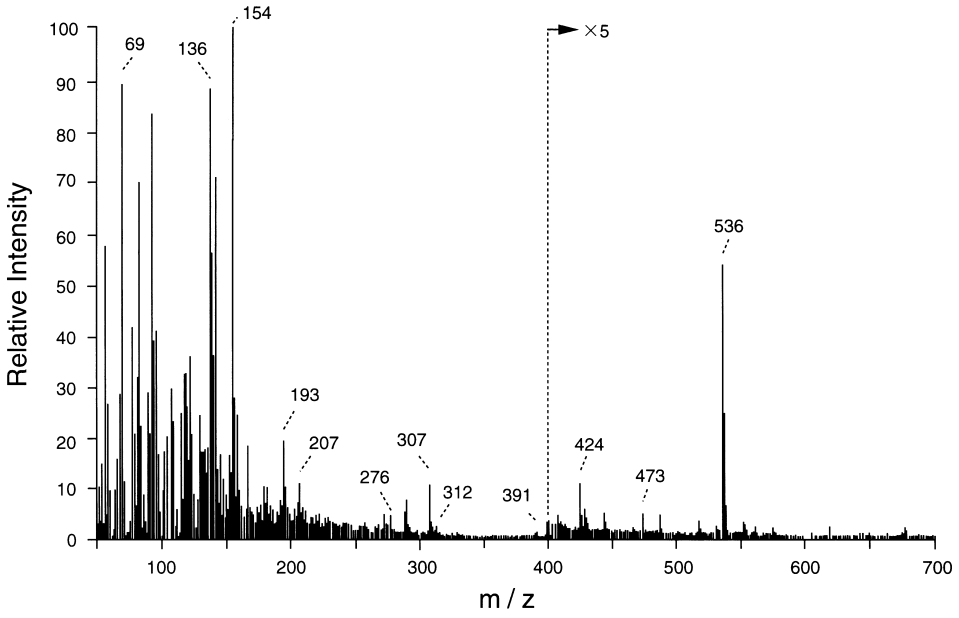

Fig. 6. Mass spectrum of the pigment extracted from *E. coli* pAYP.

genes from *Erwinia herbicola*, *Erwinia uredovora* and *Synechococcus* sp. are 40.2%, 37.4% and 22.9%, respectively.

The larger ORF starts from nt 1543 and ends at nt 3123 and encodes a 527 aa polypeptide with nucleotide sequence similarities to the phytoene desaturase genes of *E. herbicola*, *E. uredovora* and *Rhodobacter capsulatus* of 50.3%, 54.7% and 39.6%, respectively. Although the overall similarity is moderate, there are two highly conserved regions at the N- and C-terminal regions (Fig. 4).

A strong bias against codons having an A and T in the third position in the *R. capsulatus* genes was observed (Youvan et al., 1984). Among the 3979 codons from *R. capsulatus crt* genes, the triplets ATA, CTA, GTA and TTA are never used (Armstrong et al., 1989). A similar bias was observed in the *Erythrobacter crtI* and *crtY* genes. Especially no triplet TTA was observed in *E. longus* whereas in *E. herbicola crt* genes TTA was used 5 times in 2413 codons (Table 1).

2.3. *Expression of the* Erythrobacter longus crtY *and* crtI *genes in* Escherichia coli

In order to confirm that the ORFs did encode the enzymes lycopene cyclase and phytoene desaturase, the genes were cloned and expressed in *E. coli*. A 2.0 kb *Eco*RV-*Hin*dIII fragment containing the lycopene cyclase gene and a 2.6 kb *Nco*I-*Eco47* III fragment containing the phytoene desaturase gene were cloned into the vector pACYC184 and designated pAYL and pAYP respectively. The plasmid pAYL that contained *Erythrobacter* lycopene cyclase gene was transformed into *E. coli* HB101 which contained the *E. herbicola crtE*, *crtB, crtI* and *crtZ* genes and produced lycopene. This *E. coli* might produce zeaxanthin through lycopene when the pAYL plasmid was introduced into the cell.

Pigments were extracted from transformants with methanol, partitioned into *n*-hexane. The *n*-hexane phase was evaporated to dryness. After resolution into *n*-hexane, carotenoids were analyzed by HPLC using a 25 cm Vydac 218TP column with acetonitrile/methanol (90/10; v/v), and a flow rate of 0.7 ml/min. Zeaxanthin was observed at Rf = 5 min as expected demonstrating the expression of the lycopene cyclase activity (Fig. 5). But the transformant colony was yellow pigmented very slightly. On the other hand, transformant which contained both *E. herbicola crtY* gene in plasmid pACYC184 and *crtE*, *crtB* and *crtZ* was pigmented strongly.

Similarly, pAYP that contained *Erythrobacter* phytoene desaturase gene was transformed into *E. coli* HB101 which contained the *E. herbicola crtE* and *crtB* genes and produced phytoene. Lycopene was observed in the transformed *E. coli* as expected at Rt = 12 min (Fig. 5). The amount of produced lycopene in *Erythrobacter* phytoene desaturase gene contained in the transformant was half of that in the *E. herbicola* phytoene desaturase gene contained in the transformant (2 mg/g dry cell weight).

Further confirmation of products, lycopene and zeaxanthin, positive-ion FAB mass spectra of products were obtained on a JEOL JMS-SX102 double focusing mass spectrometer with *m*-nitrobenzylalcohol as a matrix. Xenon fast atoms at 6 kV were used for FAB ionization. The accelerating voltage was 10 keV, and the instrument resolution was 1500 for all measurements. The mass spectrum of isolated pigment from *E. coli* pAYP gave

the corresponding molecular ion (M+ at 536) of lycopene (Fig. 6). Similarly, molecular ion (M+ at 569) of zeaxanthin was observed in mass spectrum of the extracted pigment from *E. coli* pAYL.

Transformant which contained both *Erythrobacter crtY* gene and *E. herbicola crtB*, *crtE*, *crtI* and *crtZ* genes produce less pigment than transformant which contained *E. herbicola crtY* gene in pACYC and *E. herbicola crtB*, *crtE*, *crtI* and *crtZ* genes. These results show that *E. longus crt* genes expresses weakly in *E. coli*. Photosynthetic pigments (Bchl and Crt) do not accumulate in *E. coli* strains harboring the *R. capsulatus* photosynthetic gene cluster carried on pRS404 (Marrs, 1981). Although codon usage of *E. longus crt* genes was also biased against codons having an A or T in the third position, the degree seems to be intermediate between *R. capsulatus* and *E. herbicola* (Table 1). *E. longus* was classified in photosynthetic purple bacteria but separately categorized from *R. capsulatus* (Woese, 1987). This phylogenetic relationship may be proportional with the bias level of codon usage among taxa. The weak expression of *E. longus crt* genes in *E. coli* may be related to the biased codon usage.

3. Conclusions

(1) Two carotenoid biosynthetic genes, *crtY* and *crtI*, encoding lycopene cyclase and phytoene desaturase respectively have been cloned from the aerobic photosynthetic bacterium *Erythrobacter longus*.
(2) The nucleotide sequence similarities of *E. longus crtY* gene are 40.2%, 37.4% and 22.9% to the corresponding genes in *E. herbicola*, *E. uredovora* and *Synechococcus* sp., respectively.
(3) The nucleotide sequence similarities of *E. longus crtI* gene is 50.3%, 54.7% and 39.6% to the corresponding genes in *E. herbicola*, *E. uredovora* and *R. capsulatus*, respectively.
(4) Two highly conserved regions at the N- and C-terminal ends of the *crtI* gene have been identified.
(5) The *E. longus crtY* and *crtI* genes were successfully expressed in *E. coli* in which the corresponding carotenoid pigment had accumulated.

References

Armstrong, G.A., Alberti, M., Leach, F. and Hearst, J.E. (1989) Nucleotide sequence, organization and nature of the protein products of the carotenoid biosynthesis gene cluster of *Rhodobacter capsulatus*. Mol. Gen. Genet. 216, 254–268.

Armstrong, G.A., Alberti, M. and Hearst, J.E. (1990) Conserved enzymes mediate the early reactions of carotenoid biosynthesis in nonphotosynthetic and photosynthetic prokaryotes. Proc. Natl. Acad Sci. USA 87, 9975–9979.

Bartley, G.E. and Scolnik, P.A. (1989) Carotenoid biosynthesis in photosynthetic bacteria: genetic characterization of the *Rhodobacter capsulatus* CrtI protein. J. Biol. Chem. 264, 13109–13113.

Cunningham Jr., F.K., Chamovitz, D., Misawa, N., Gantt, E. and Hirschberg, J. (1993) 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, 130–138.

Cunningham Jr., F.X., Sun, Z., Chamovitz, D. Hirschberg, J. and Gannt, E. (1994) Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium *Synechococcus* sp. strain PCC7942. Plant Cell 6, 1107–1121.

Hundle, B.S., O'Brien, D.A., Beyer, P., Kleinig, H. and Hearst, J.E. (1993) In vitro expression and activity of lycopene cyclase and β-carotene hydroxylase from *Erwinia herbicola*. FEBS Lett. 315, 329–334.

Lambert, L.A., Koch, W.H., Warner, W.G. and Kornhauser, A. (1990) Antitumor activity in skin of SKh and Sencar mice by two dietary β-carotene formulations. Nutr. Cancer 13, 213–221.

Marrs, B. (1981) Mobilization of the genes for photosynthesis from *Rhodopseudomonas capsulata* by a promiscuous plasmid. J. Bacteriol. 146, 1003–1012.

Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K. and Harashima, K. (1990) Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. J. Bacteriol. 172, 6704–6712.

Takaichi, S., Shimada, K. and Ishidsu, J. (1990) Carotenoids from the aerobic photosynthetic bacterium, *Erythrobacter longus*: β-carotene and its hydroxy derivatives. Arch. Microbiol. 153, 118–122.

Woese, C.R. (1987) Bacterial evolution. Microbiol. Rev. 51, 221–271.

Youvan, D.C., Bylina, E.J., Alberti, M., Begusch, H. and Hearst, J.E. (1984) Nucleotide and deduced polypeptide sequences of the photosynthetic reaction center, B870 antenna and flanking polypeptides from *Rhodopseudomonas capsulata*. Cell 37, 949–957.