ELSEVIER

Gene 297 (2002) 129–139

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Genetics of isoprenoid biosynthesis in *Paracoccus zeaxanthinifaciens*

Markus Hümbelin[a,*], Ann Thomas[a,1], Junying Lin[a,2], Juan Li[a,3], Jan Jore[b], Alan Berry[a]

[a]*Biotechnology Department, Roche Vitamins AG, CH-4070 Basel, Switzerland*
[b]*TNO Voeding, Utrechtseweg 48, P.O. Box 360, NL-3700 AJ Zeist, The Netherlands*

Received 29 May 2002; received in revised form 22 July 2002; accepted 1 August 2002
Received by A.M. Campbell

Abstract

The genes coding for all the enzymes involved in the conversion of acetyl-CoA to farnesyl diphosphate (FPP) in the zeaxanthin-producing bacterium *Paracoccus zeaxanthinifaciens* were cloned and characterized. Two genes encoding enzymes catalysing the condensation of two acetyl-CoA molecules to acetoacetyl-CoA were found. The six enzymes involved in the conversion of acetyl-CoA and acetoacetyl-CoA to isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) are grouped in an operon, designated the mevalonate operon. The gene encoding the enzyme catalysing two consecutive condensations, IPP and DMAPP to geranyl diphosphate (GPP) and IPP and GPP to FPP, is not clustered with any other gene encoding an enzyme of the isoprenoid pathway. Genes encoding enzymes involved in the biosynthesis of poly-hydroxyalkanoate and non-carotenoid isoprenoids found in *P. zeaxanthinifaciens* are also presented.

Keywords: Mevalonate; Biosynthesis; Pathway; Carotenoid; Zeaxanthin

1. Introduction

Isoprenoids are a large group of natural products that serve many physiological functions. In mammalian cells isoprenoids are synthesized or ingested with the diet and function as membrane components, steroid hormones, antioxidants and radical scavengers, vitamins and cofactors and substrates for post-translational modifications. In plants, isoprenoids and molecules with prenyl side chains function as photosynthetic pigments, membrane components and phytohormones, to name just the most important. All isoprenoids are synthesized from the five-carbon compound isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP).

Two pathways exist for IPP biosynthesis (Fig. 1). The pathway of IPP biosynthesis from acetyl-CoA via mevalonate was discovered in the late 1950s (for a review see Bloch, 1992). Although evidence that *Escherichia coli* and some other bacteria use a different pathway was published as early as 1965 (Raman et al., 1965), it was not until recently that an alternate pathway was discovered (Rohmer et al., 1996). 1-deoxyxylulose 5-phosphate (DXP) is the substrate for the first committed reaction of this non-mevalonate pathway, the DXP pathway. The final steps of the DXP pathway were only elucidated recently (Rohdich et al., 2002). While most eubacteria seem to utilize the DXP pathway, a number of recent reports indicate that several bacterial genera possess genes of the mevalonate pathway. Some *Streptomyces* strains utilize both pathways (Seto et al., 1996; Takahashi et al., 1999; Kuzuyama et al., 2000) and a gene cluster encoding the enzymes of the mevalonate pathway was cloned from *Streptomyces* sp. strain CL190 (accession number AB037666) and *Streptomyces griseolosporeus* (accession number AB037907). Mevalonate pathway gene clusters were also found in other Gram-positive bacteria including

Abbreviations: CoA, coenzyme A; FPP, farnesyl diphosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl diphosphate; DXP, 1-deoxyxylulose 5-phosphate; FMN, flavin mononucleotide; NAD(P)H, nicotinamide-adenine dinucleotide (phosphate), reduced form; i^6A37, N^6-(Δ^2-isopentenyl)-adenosine-37; ms^2i^6A37, 2-methylthio-N^6-(Δ^2-isopentenyl)-adenosine-37; UPP, undecaprenyl diphosphate; RNase, ribonuclease; DNase, deoxyribonuclease; EDTA, ethylenediaminetetraacetic acid; PCR, polymerase chain reaction; PHA, polyhydroxyalkanoate; bp, base pair(s); HMG, 3-hydroxy-3-methlyglutaryl; FARM, first aspartate-rich motif; SARM, second aspartate-rich motif; CLD, chain-length determination; DPP, decaprenyl diphosphate

* Corresponding author. Tel.: +41-61-687-3078; fax: +41-61-687-1847.
E-mail address: markus.huembelin@roche.com (M. Hümbelin).

[1] Present address: Drug Metabolism and Pharmacokinetics, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA.

[2] Present address: Allergy Department, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA.

[3] Present address: EluSys Therapeutics Inc., 10 Bloomfield Avenue, Pine Brook, NJ 07058, USA.

0378-1119/02/$ - see front matter
PII: S0378-1119(02)00877-6

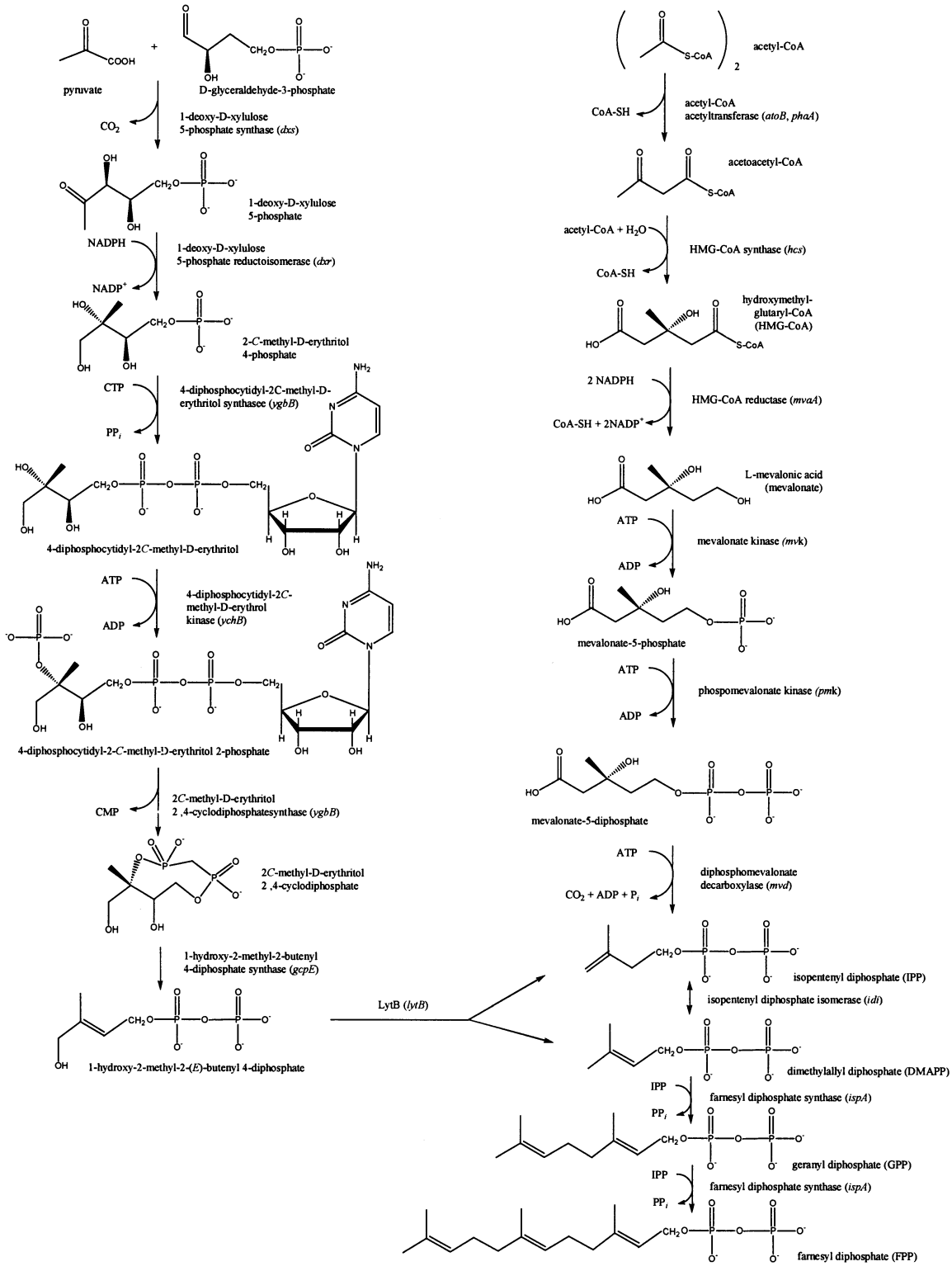

Fig. 1. Biosynthesis of FPP. The biosynthesis of IPP and DMAPP via the DXP pathway (left) and the mevalonate pathway (right) and the condensation of DMAPP with IPP to GPP and GPP with IPP to FPP is shown. Cofactors, names of the compounds and enzymes are indicated. Gene designations are given in parentheses next to the enzyme names.

several species of *Streptococcus*, *Enterococcus* and *Staphylococcus* (Wilding et al., 2000). In addition, a cluster of genes encoding the enzymes for the conversion of mevalonate to IPP and DMAPP was identified from *Lactobacillus helveticus* (Smeds et al., 2001). The only published mevalonate gene cluster from a Gram-negative organism, comprising

all genes necessary for the conversion of acetoacetyl-CoA to IPP and DMAPP, was found through the sequencing of the *Borrelia burgdorferi* genome by homology to eukaryotic and archaeal proteins (Fraser et al., 1997).

The isomerization of IPP to DMAPP is catalysed by the enzyme IPP isomerase (Fig. 1). Until the end of 2000, several *idi* genes, encoding IPP isomerase, were known from eukaryotic organisms, but only two eubacterial genes (from *Rhodobacter capsulatus* and *E. coli*) and no archaeal genes were known (Smit and Mushegian, 2000). The *idi* gene from *E. coli* was reported to be non-essential and it was found that DMAPP and IPP are synthesized separately from a common precursor in organisms that utilize the DXP pathway (Fig. 1). Recently, the *idi* gene from *Streptomyces* sp. strain CL190 was identified (Kaneda et al., 2001). This IPP isomerase, designated type 2, is different from the eukaryotic IPP isomerase (type 1) based on lack of sequence similarity and by its requirement for the cofactors FMN and NAD(P)H. Homologues of type 2 IPP isomerase are found, among others, in a carotenoid gene cluster of *Pantoea agglomerans* (formerly *Erwinia herbicola*) (Hundle et al., 1994) and in mevalonate gene clusters of several Gram-positive cocci (Wilding et al., 2000) and archaea (Smit and Mushegian, 2000).

Biosynthesis of geranyl diphosphate (GPP) and farnesyl diphosphate (FPP) follows the same chemistry in all organisms (Fig. 1). GPP is formed by condensation of one IPP and one DMAPP molecule catalysed by the enzyme GPP synthase. One molecule each of GPP and IPP are then condensed to form FPP by FPP synthase (Fig. 1). In many organisms, the latter enzyme catalyses both reactions. The bacterial FPP synthase is encoded by the *ispA* gene.

Considering the great number of isoprenoids found in nature, the number of isoprenoid compounds synthesized by many bacteria is relatively small. DMAPP is used for the modification of tRNA by tRNA prenyltransferase, encoded by the *miaA* gene. This enzyme condenses DMAPP with the adenosine residue 37 of most tRNA species, which bind to codons starting with uridine. In *E. coli* and other bacteria, the product, N^6-(Δ^2-isopentenyl)-adenosine-37 (i^6A37), is subsequently methylthiolated by the *miaB* and *miaC* gene products to form 2-methylthio-N^6-(Δ^2-isopentenyl)-adenosine-37 (ms^2i^6A37). In some bacteria the residue is further hypermodified by hydroxylation (reviewed in Björk, 1996). Undecaprenyl diphosphate (UPP), which is essential for bacterial cell wall biosynthesis, is synthesized from FPP by sequential addition of eight IPP units, catalysed by the enzyme UPP synthase (Fujihashi et al., 2001). Quinones, which are essential for respiration, also have prenyl moieties, which are synthesized from FPP and IPP. The length of the prenyl moiety is dependent on the quinone(s) synthesized by a given bacterium.

Carotenoids are C-40 isoprenoid compounds with a number of commercial applications. Schocher and Wiss (Schocher and Wiss, 1975) reported the isolation of a Gram-negative marine bacterium (originally designated *Flavobacterium* ATCC 21588) that produced the yellow carotenoid zeaxanthin (3,3′-dihydroxy-β-carotene), which has application in poultry pigmentation and in the prevention of age-related macular degeneration in humans. Goodwin (1972) showed that cell-free extracts of the *Flavobacterium* incorporated labeled mevalonate into zeaxanthin, suggesting that IPP biosynthesis, and hence zeaxanthin biosynthesis, occurred via the mevalonate pathway in this organism. Recently, we reported the taxonomic reclassification of the zeaxanthin-producing '*Flavobacterium*' as a new *Paracoccus* species, *P. zeaxanthinifaciens* (Berry et al., in press), and confirmed the earlier findings that IPP biosynthesis occurs exclusively via the mevalonate pathway in this organism (Eisenreich et al., 2002). Another zeaxanthin-producing bacterium, *Spingobacterium multivorum* (formerly *Flavobacterium multivorum*) was recently shown to utilize the DXP pathway (Rosa-Putra et al., 2001). As part of our continued effort to understand zeaxanthin biosynthesis in *P. zeaxanthinifaciens*, we now report the cloning and sequence of all genes encoding the enzymes involved in the biosynthesis of FPP from acetyl-CoA in this organism. In addition, we present the sequences of genes involved in the biosynthesis of poly-hydroxyalkanoate and non-carotenoid isoprenoid compounds.

2. Materials and methods

2.1. Bacteria and culture conditions

High zeaxanthin producing mutant strains R1534 and R114, derived from *Paracoccus zeaxanthinifaciens* (ATCC 21588) by random mutagenesis and screening (Berry et al., in press), were obtained from the culture collection of F. Hoffmann–La Roche. The bacteria were grown at 28 °C in F-medium (10 g/l tryptone, 10 g/l yeast extract, 30 g/l NaCl, 10 g/l D-glucose, 5 g/l $MgSO_4 \cdot 7H_2O$, pH 7.0).

2.2. Isolation of genomic DNA

Cultures of *P. zeaxanthinifaciens* strain R114 (600 ml) were centrifuged for 10 min at $10,000 \times g$ at 4 °C and the pellet was washed once with 200 ml lysis buffer (0.1 M NaCl, 50 mM EDTA, 10 mM Tris–HCl, pH 7.5) and once with 100 ml lysis buffer. The final pellet was resuspended in 20 ml lysis buffer containing 50 mg lysozyme and 1 mg RNase A (DNase free). After incubation for 15 min at 37 °C, 1.5 ml 20% sodium *N*-lauroyl-sarcosinate and 2.25 mg proteinase K were added. After incubation at 50 °C for 30–60 min, the lysate was extracted with one volume buffer-saturated phenol, pH 7.5–7.8 (LifeTechnologies, Rockville, MD, USA) by gentle but thorough mixing. The emulsion was centrifuged for 20 min at $30,000 \times g$ and the aqueous phase was re-extracted with phenol. The phases were separated as before and the aqueous phase was extracted twice with 1 volume phenol/chloroform (1:1). At this step centrifugation for 20 min at $3200 \times g$ in a swinging bucket rotor was sufficient for phase separation. After a final extraction

with one volume chloroform, 0.1 volume 3 M sodium-acetate (pH 5.2) was added and the solution was overlaid with 2 volumes ice-cold ethanol. The precipitated DNA was spooled with a glass rod, soaked in 70% ethanol for 5 min, rinsed with chloroform and then air-dried for 5–10 min. The DNA was resuspended overnight in 5 ml TE (10 mM Tris–HCl (pH 7.5), 1 mM EDTA). If the solution was yellow due to traces of zeaxanthin, the organic extractions, precipitation and spooling were repeated as above.

2.3. Polymerase chain reaction (PCR)

Oligonucleotides were purchased from LifeTechnologies (Rockville, MD, USA). PCR was performed in a GeneAmp PCR system 9700 (PE Applied Biosystems, Foster City, CA, USA) using the GC-rich PCR system (Roche Molecular Biochemicals, Mannheim, Germany) according to the manufacturers instructions. Typically, $MgCl_2$ concentration was 1.5 mM and the resolution solution was added to 1 M final concentration. PCR fragments were cloned in the vector pCR 2.1-TOPO (Invitrogen, Carlsbad, CA, USA).

2.4. DNA labeling and detection

The PCR DIG Probe Synthesis Kit and the DIG Luminescent Detection Kit were used for DNA labeling and detection, respectively (both obtained from Roche Molecular Biochemicals).

2.5. DNA sequencing

Sequencing reactions were performed using the BigDye DNA sequencing kit (PE Applied Biosystems) according to the manufacturer's instructions. Sequencing reactions were purified on DyeEx spin columns (Qiagen, Hilden, Germany) and fragment separation and detection was done with an ABI Prism 310 Genetic Analyser (PE Applied Biosystems).

2.6. λ-Library

A custom-made library with partially *Sau*3AI digested *P. zeaxanthinifaciens* strain R114 DNA in lambda FIX II was purchased from Stratagene (La Jolla, CA, USA). λ-DNA was isolated using the Qiagen Lambda Kit following the manufacturer's instructions.

2.7. Sequence analysis

Sequence similarities were determined using the GCG gap program (Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, WI, USA) with the optional parameter endweight (all other parameters were default).

2.8. Genome sequence

Sequencing of the *P. zeaxanthinifaciens* genome was performed under contract by Integrated Genomics, Chicago, IL, USA.

3. Results and discussion

3.1. The condensation of two acetyl-CoA molecules to acetoacetyl-CoA

Acetoacetyl-CoA can be synthesized by the condensation of two molecules of acetyl-CoA by the enzyme acetyl-CoA acetyltransferase, also known as acetoacetyl-CoA thiolase. In *P. zeaxanthinifaciens*, there are at least two acetyl-CoA acetyltransferases, which are encoded by the two genes *atoB* and *phaA*. The *atoB* gene is found 165 nucleotides upstream of *crtE*, which is part of the carotenoid gene cluster (Pasamontes et al., 1997) (Table 1). Acetoacetyl-CoA is also the substrate for poly-hydroxyalkanoate (PHA) biosynthesis. In many bacteria the genes involved in PHA biosynthesis are grouped in operons (Madison and Huisman, 1999). In *Paracoccus denitrificans* the *phaA* and *phaB* genes, encoding the acetyl-CoA acetyltransferase and acetoacetyl-CoA reductase, respectively, are clustered in an operon whereas *phaC*, the gene encoding the last enzyme in the pathway, poly(3-hydroxyalkanoate) synthase, is not part of this operon. PCR fragments containing parts of *phaA* from *P. zeaxanthinifaciens* strain R1534 and *phaC* from *P. zeaxanthinifaciens* strain R114 were obtained with primers based on the *P. denitrificans phaA* and *phaC* sequences. The fragments were used to screen a *P. zeaxanthinifaciens* strain R114 λ-library and several positive clones were isolated and characterized. As in *P. denitrificans*, *phaA* and *phaB* are clustered in *P. zeaxanthinifaciens* (Table 1) whereas *phaC* is found at a different locus. The clustering of *phaA* and *phaB* suggests that the two genes are expressed together when the cell produces poly-hydroxyalkanoates. On the other hand, a putative transcriptional stop signal was found between the *P. zeaxanthinifaciens phaA* and *phaB* genes, which is absent in *P. denitrificans*. Thus, the expression of the two genes might not be coupled in *P. zeaxanthinifaciens* strain R114. A mutant strain with an interrupted *phaA* gene had about 10% of the acetyl-CoA acetyltransferase activity compared to its parent *P. zeaxanthinifaciens* strain R114 and produced only small amounts of zeaxanthin (unpublished observation). This suggests that PhaA is responsible for most acetyl-CoA acetyltransferase activity and that AtoB contributes only marginally to the total acetyl-CoA acetyltransferase activity in *P. zeaxanthinifaciens* strain R114. If PhaA is the major acetyl-CoA acetyltransferase in *P. zeaxanthinifaciens* strain R114 its expression must not to be linked exclusively to PHA biosynthesis. This might explain the above mentioned putative transcriptional stop signal between *phaA* and *phaB*.

3.2. The genes and proteins of the mevalonate pathway from P. zeaxanthinifaciens strain R114

As mentioned above, we recently confirmed that *P. zeaxanthinifaciens* uses the mevalonate pathway for IPP biosynthesis (Eisenreich et al., 2002). With this information

Table 1
Sequence accession numbers of isoprenoid biosynthesis genes and gene clusters in *P. zeaxanthinifaciens*

Accession no.	Gene	Enzyme	EC no.	Reaction catalysed
AAV40146[b]	*atoB*	Acetyl-CoA acetyltransferase	2.3.1.9	Acetyl-CoA to acetoacetyl-CoA[a]
AJ431695[c]	*phaA*	Acetyl-CoA acetyltransferase	2.3.1.9	Acetyl-CoA to acetoacetyl-CoA
	phaB	Acetoacetyl-CoA reductase	1.1.1.36	Acetoacetyl-CoA to 3-hydroxybutyryl-CoA
AJ431696[c]	*hcs*	HMG-CoA synthase	4.1.3.5	Acetoacetyl-CoA to HMG-CoA[a]
	mvaA	HMG-CoA reductase	1.1.1.34	HMG-CoA to mevalonate[a]
	mvk	Mevalonate kinase	2.7.1.36	Mevalonate to phosphomevalonate[a]
	pmk	Phosphomevalonate kinase	2.7.4.2	Phosphomevalonate to mevalonate diphosphate[a]
	mvd	Mevalonate diphosphate decarboxylase	4.1.1.33	Mevalonate diphosphate to IPP[a]
	idi	IPP isomerase	5.3.3.2	IPP to DMAPP[a]
AJ431697[c]	*ispA*	GPP synthase	2.5.1.1	IPP + DMAPP to GPP[a]
		FPP synthase	2.5.1.10	IPP + GPP to FPP[a]
U62808[c]	*crtE*	GGPP synthase	2.5.1.29	IPP + FPP to GGPP
	crtB	Phytoene synthase	2.5.1.32	2 GGPP to phytoene
	crtY	Phytoene desaturase	1.3.99.-	Phytoene to lycopene
	crtI	Lycopene cyclase	5.3.1.–	Lycopene to β-carotene
	crtZ	β-Carotene hydroxylase		β-Carotene to zeaxanthin
AJ431695[c]	*ddsA*	Decaprenyl diphosphate synthase	2.5.1.29	7 IPP + FPP to decaprenyl diphosphate
AJ431698[c]	*uppS*	Undecaprenyl diphosphate synthase	2.5.1.31	8 IPP + FPP to undecaprenyl diphosphate
	miaA	tRNA prenyltransferase	2.5.1.8	See text

[b] In GeneSeq database of patented nucleic acids, position 1180–2355.
[a] See Fig. 1 for details.
[c] In GenBank, EMBL, DDBJ.

in hand, we set out to clone the genes encoding the enzymes of the mevalonate pathway in *P. zeaxanthinifaciens*. Inspection of the GenBank database revealed that one of the enzymes of the mevalonate pathway, mevalonate diphosphate decarboxylase (Mvd), contains highly conserved regions spanning several amino acids. Three such regions were chosen from an alignment of bacterial mevalonate diphosphate decarboxylases and oligonucleotides were designed using the preferred codon usage found in the carotenoid gene cluster of *P. zeaxanthinifaciens* strain R1534 (Pasamontes et al., 1997). The PCR product obtained with such a primer combination was cloned and sequenced (hatched box in Fig. 2). Using the cloned fragment as a probe a 0.94 kb *Bam*HI–*Sal*I fragment was cloned (box C in Fig. 2). In parallel, a λ-library (see Section 2) was screened using the *mvd*-PCR fragment as a probe. DNA was isolated from two positive λ-clones and cut with *Bam*HI and *Sal*I or *Eco*RI and *Sal*I. A number of the restriction fragments was isolated and cloned in the vector pUC19. Several of the fragments contained sequences with extensive similarities to genes encoding proteins of the mevalonate pathway (boxes A, B, D, L, M in Fig. 2). The clones connecting these individual sequences (Fig. 2, boxes 26-1, 22-1, 14-1, 4-3, 49-1 and mvd11-3) were obtained by PCR with primers derived from the sequences of the cloned restriction fragments using the DNA of the λ-clones as template.

The arrangement of the mevalonate pathway genes in *P. zeaxanthinifaciens* strain R114 is unique when compared to known mevalonate gene clusters of other bacteria (Fig. 3). Besides *P. zeaxanthinifaciens* strain R114 only *Borrelia burgdorferi* (accession number AE000783), *Streptomyces* sp. strain CL190 (accession number AB037666) and *Streptomyces griseolosporeus* (accession number AB037907) have all mevalonate genes in a single operon. In *Streptococcus pyogenes* all mevalonate genes are clustered in a

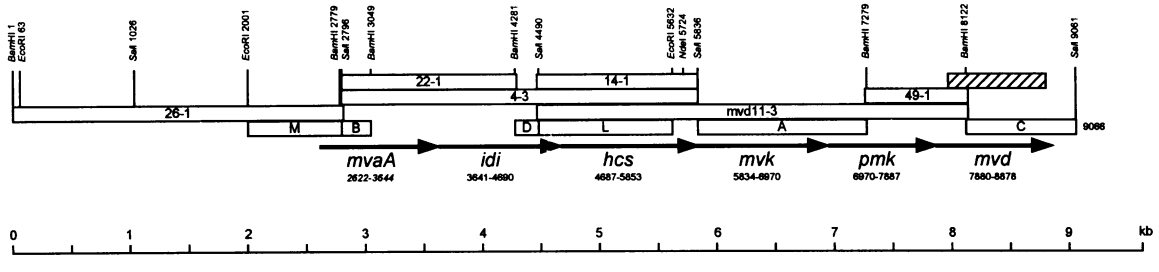

Fig. 2. The mevalonate operon from *P. zeaxanthinifaciens*. The cloned restriction fragments and PCR fragments are represented as boxes. The hatched box represents the original PCR fragment obtained with degenerate primers. Arrows indicate the mevalonate genes and the position of the first nucleotide of the start codon and the last nucleotide of the stop codon is given below each gene designation. Sites for the restriction endonucleases *Eco*RI, *Bam*HI, *Sal*I and *Nde*I are shown.

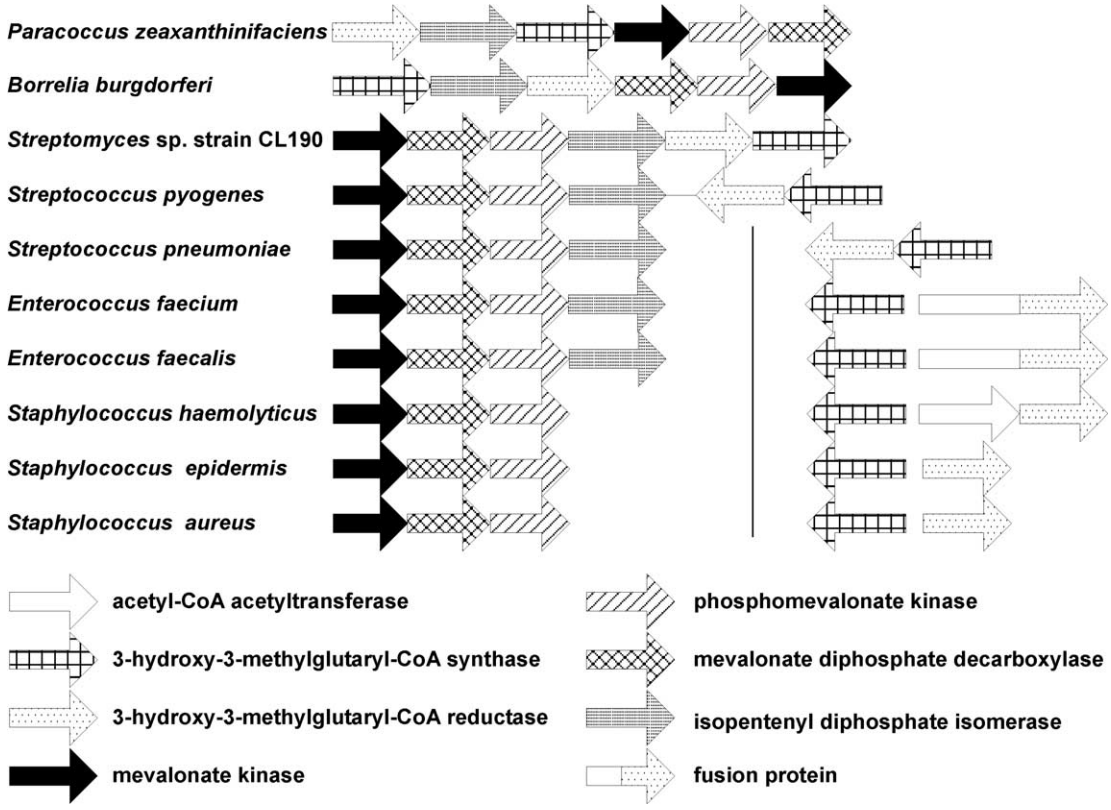

Fig. 3. Bacterial mevalonate gene clusters. Arrangement of the genes in known bacterial mevalonate gene clusters. Patterned arrows represent genes. The drawing is not to scale. The vertical bar in the lower six species indicates that the genes on the left and the right are not linked. The gene arrangements in the mevalonate clusters from *Streptomyces* sp. strain CL190 and *S. griseolosporeus* are identical.

single locus but they are grouped in two operons. All other species have two loci with the two kinase genes and the mevalonate diphosphate decarboxylase gene grouped in one operon and the HMG-CoA synthase and HMG-CoA reductase genes on a second locus, either forming an operon (in *Streptococcus pneumoniae*) or as separate transcription units. All species except the members of *Staphylococcus* have an additional gene linked with the mevalonate gene cluster, which was recently identified as an IPP isomerase gene (*idi*) in *Streptomyces* sp. strain CL190 (Kaneda et al., 2001). The two *Enterococcus* species and *Staphylococcus haemolyticus* have an acetyl-CoA acetyltransferase gene linked with the HMG-CoA reductase gene. In the *Enterococcus* species the two genes are fused.

The genes of the mevalonate operon from *P. zeaxanthinifaciens* were identified by similarity of their products to proteins in general databases. An alignment of the HMG-CoA reductase from *P. zeaxanthinifaciens* and from three *Streptomyces* species is shown in Fig. 4. There are two classes of HMG-CoA reductases (Bochar et al., 1999; Boucher and Doolittle, 2000). Eubacterial HMG-CoA reductases are generally of class II, whereas class I enzymes are found in eukaryotes and archaea. The *Streptomyces* and the *P. zeaxanthinifaciens* HMG-CoA reductases together with the enzyme from *Vibrio cholerae* (Boucher and Doolittle, 2000) are the only eubacterial HMG-CoA reductases of class I known so far. *V. cholerae* utilizes the DXP pathway for IPP and DMAPP biosynthesis and the HMG-CoA reductase is the only enzyme of the mevalonate pathway found so far in this organism (Boucher and Doolittle, 2000).

The first reaction in IPP biosynthesis catalysed by an enzyme encoded in the mevalonate operon is the condensation of acetyl-CoA and acetoacetyl-CoA to hydroxymethylglutaryl-CoA, catalysed by HMG-CoA synthase (Fig. 1). The known orthologues of the HMG-CoA synthase from *P. zeaxanthinifaciens* with $>40\%$ identical residues are all from eubacteria (Table 2). All amino acids that had been identified earlier to be identical in eubacterial and eukaryotic HMG-CoA synthases (Wilding et al., 2000), are also found in the enzyme from *P. zeaxanthinifaciens* strain R114, except for a conserved change of a lysine to an arginine. The HMG-CoA synthase from the archaea *Halobacterium* sp. NRC-1 (accession number Q9HPJ0) is the closest non-eubacterial orthologue with 38% identical residues. Two proteins from *Myxococcus xanthus*, Tac and Taf (accession numbers Q9XB06 and Q9XB03, respectively) and a protein from *Bacillus subtilis*, PksG, (accession number P40830), share substantial sequence similarities with the *P. zeaxanthinifaciens* HMG-CoA synthase (34–36% identical residues). These proteins are involved in

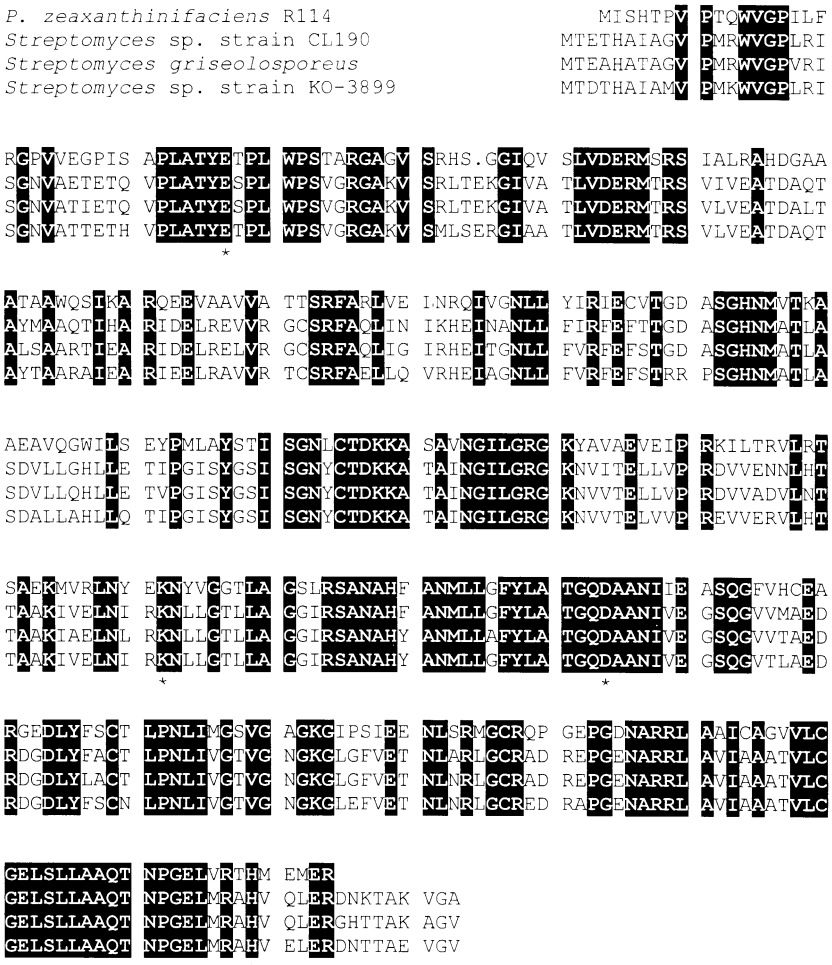

Fig. 4. Bacterial class I HMG-CoA reductases. Alignment of HMG-CoA reductases from *P. zeaxanthinifaciens* strain R114, *Streptomyces* sp. strain CL190 (accession number Q9Z9N4), *S. griseolosporeus* (Q9ZNH1) and *Streptomyces* sp. strain KO-3899 (Q9ZNH0). Amino acid residues that are the same in all four strains are shown in white on black background. Asterisks indicate amino acids proposed to function in catalysis (Tabernero et al., 1999).

Table 2
Similarities between mevalonate pathway enzymes from *P. zeaxanthinifaciens* and bacterial orthologues[a]

	Streptomyces sp. strain CL109	*Streptomyces griseolosporeus*	*Staphylococcus epidermis*	*Staphylococcus aureus*	*Staphylococcus haemolyticus*	*Listeria innocua*	*Enterococcus faecalis*	*Enterococcus faecium*	*Streptococcus pyogenes*	*Streptococcus pneumoniae*	*Lactococcus lactis*
HMG-CoA synthase	49% Q9KWG1	51% Q9KWF5	45% Q9FD76	45% Q9FD87	43% Q9FD82	44% Q92BU0	43% Q9FD71	41% Q9FD66	42% Q9FD61	41% Q9FD56	41% Q9CFA9
Mevalonate diphosphate decarboxylase	45% Q9KWG4	49% Q9KWF8	38% Q9FD73	38% Q9FD84	38% Q9FD78	42% Q92FU2	41% Q9FD68	42% Q9FD63	35% Q9FD58	37% Q9FD53	38% Q9CIF7
Mevalonate kinase	30% Q9KWG5	31% Q9KWF9	25% Q9FD74	26% Q9FD85	27% Q9FD79	22% Q92FU3	29% Q9FD69	26% Q9FD64	28% Q9FD59	29% Q9FD54	25% Q9CIF8
Phosphomevalonate kinase	36% Q9KWG3	37% Q9KWF7	26% Q9FD72	26% Q9FD83	28% Q9FD77	27% Q92FU1	30% Q9FD67	33% Q9FD62	27% Q9FD57	27% Q9FD52	26% Q9CIF6

[a] The upper number indicates % identical residues, and database accession numbers are given below.

polyketide biosynthesis and their similarity to HMG-CoA synthases was explained by their proposed function as enzymes with transacylase and condensing activities (Paitan et al., 1999). Another *Myxococcus* species, *M. fulvus*, has been found to utilize the mevalonate pathway (Rosa-Putra et al., 1998). Since it is likely that *M. xanthus* also uses the mevalonate pathway, an HMG-CoA synthase is expected to be found in this organism unless either Tac or Taf is able to catalyse the condensation of acetyl-CoA and acetoacetyl-CoA to HMG-CoA.

The closest orthologues of the mevalonate diphosphate decarboxylase from *P. zeaxanthinifaciens* are all of eubacterial origin (Table 2). However, when the introduction of gaps is not limited, the *P. zeaxanthinifaciens* enzyme can be aligned with a number of eukaryotic orthologues resulting in high percentages of identical residues (35–40%) (not shown). All conserved amino acids identified earlier in mevalonate diphosphate decarboxylases from eubacteria and eukaryotes (Wilding et al., 2000), are also found in the enzyme from *P. zeaxanthinifaciens* strain R114, except for a different spacing (four residues instead of three) between the arginine and the glycine in the conserved region AXXXSXSXXRXXXG. No mevalonate diphosphate decarboxylase orthologues are found in archaea (Smit and Mushegian, 2000).

There is much less sequence similarity among the bacterial mevalonate kinases (Mvks) and phosphomevalonate kinases (Pmks) than among the bacterial orthologues of the other enzymes of the mevalonate pathway (Table 2). The Mvk from *P. zeaxanthinifaciens* strain R114 has a 37-amino-acid insert in the amino-terminal region, which is lacking in other Mvks. Together with the bacterial Mvks some archaeal enzymes, e.g. from *Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum*, *Halobacterium* sp. strain NRC-1 and *Pyrococcus abyssi*, and Mvks from the plants *Oryza sativa* (rice), *Hevea brasiliensis* (para rubber tree) and *Arabidopsis thaliana* are among the most similar to the Mvk from *P. zeaxanthinifaciens* strain R114. The conserved amino acids identified earlier in Mvks from eubacteria and eukaryotes (Wilding et al., 2000) are also found in the enzyme from *P. zeaxanthinifaciens* strain R114, except for the serine-glycine in the third homology block and the glycine-glycine in the fourth homology block.

The similarity among bacterial Pmks is even weaker than the similarity among the bacterial Mvks (Table 2). The conserved amino acids identified earlier in Pmks from eubacteria and eukaryotes (Wilding et al., 2000) are also found in the enzyme from *P. zeaxanthinifaciens* strain R114, except for the tyrosine in the first homology block. In addition, there is a conserved lysine to arginine substitution in the second homology block. The proteins most similar to the Pmk from *P. zeaxanthinifaciens* strain R114 are Mvks from archaea, e.g. *Halobacterium* sp. strain NRC-1, *Aeropyrum pernix*, *Pyrococcus horikoshii*, *M. thermoautotrophicum*, *P. abyssi* and *A. fulgidus*. The fact that no Pmks are found in archaea (Smit and Mushegian, 2000), suggests that the same kinase might perform both phosphorylations.

3.3. The genes and proteins involved in the biosynthesis of FPP from IPP

In bacteria utilizing the mevalonate pathway, except for *Staphylococci*, the gene encoding IPP isomerase (*idi*) is found in operons with genes encoding enzymes of the mevalonate pathway (Fig. 3). An alignment of Idi from *P. zeaxanthinifaciens* strain R114 with the best matches found in the EMBL database is shown in Fig. 5. The first fourteen sequences are from eubacteria and the next eight sequences are from archaea but this order does not reflect the degree of similarity. Interestingly, one eukaryotic species, the protozoan parasite *Leishmania major*, also has a protein that is highly similar. This is unexpected because all known Idis from other eukaryotes are of a different type, designated type 1 (Kaneda et al., 2001).

In *P. zeaxanthinifaciens*, as in many eubacteria, the GPP synthase and FPP synthase reactions are carried out by the same enzyme, which is encoded by the *ispA* gene (Table 1). An alignment with some of the most similar orthologues is shown in Fig. 6. The closest homologues are the enzymes from various *Rhodobacter* species, *R. capsulatus* (68% identity), *R. sphaeroides* (68%) and *R. sulfidophilus* (61%). Prenyltransferases have two conserved aspartate-rich motifs FARM (first aspartate-rich motif) and SARM (second aspartate-rich motif). The eubacterial enzymes (type II FPP synthases) possess an aromatic amino acid (preferentially a tyrosine) at the fifth position before the FARM motif within the so called chain-length determination (CLD) region, which comprises the FARM motif plus the five amino acids preceding it (for a review see Wang and Ohnuma, 1999). FARM and SARM and a CLD region characteristic for eubacterial FPP synthases are found the *P. zeaxanthinifaciens* enzyme and its orthologues (Fig. 6).

3.4. Isoprenoid end-products synthesized by P. zeaxanthinifaciens

The carotenoid zeaxanthin is responsible for the characteristic yellow to orange color of *P. zeaxanthinifaciens*. The isolation and characterization of a gene cluster encoding all the proteins involved in zeaxanthin biosynthesis from FPP has been described earlier (Table 1) (Pasamontes et al., 1997) and will not be discussed further here. However, genes encoding enzymes involved in the synthesis of other isoprenoids in *P. zeaxanthinifaciens* are of interest in the context of the present work. Undecaprenyl diphosphate (UPP), which is involved in bacterial cell wall biosynthesis, is synthesized from FPP by sequential addition of eight IPP units, catalysed by the enzyme UPP synthase. The gene encoding UPP synthase was found by sequence similarity in the partially completed genome sequence of *P. zeaxanthinifaciens* strain R114 (Table 1). The *ddsA* gene, encoding decaprenyl diphosphate (DPP) synthase, was found fortuitously when analysing the *phaA–phaB* gene cluster (Table 1), and is located downstream of *phaB*. DPP is formed by

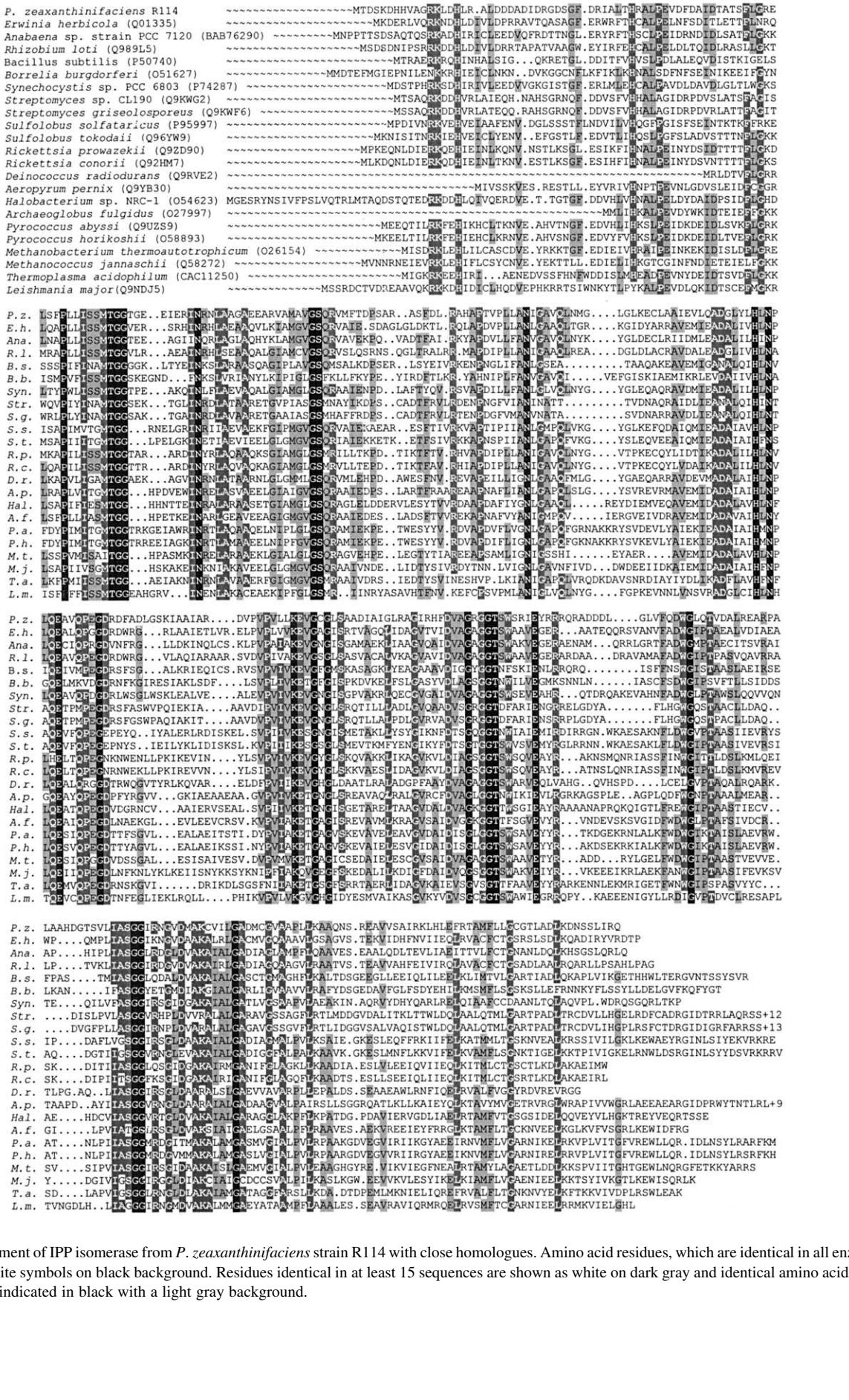

Fig. 5. Alignment of IPP isomerase from *P. zeaxanthinifaciens* strain R114 with close homologues. Amino acid residues, which are identical in all enzymes, are shown as white symbols on black background. Residues identical in at least 15 sequences are shown as white on dark gray and identical amino acids in 10–14 proteins are indicated in black with a light gray background.

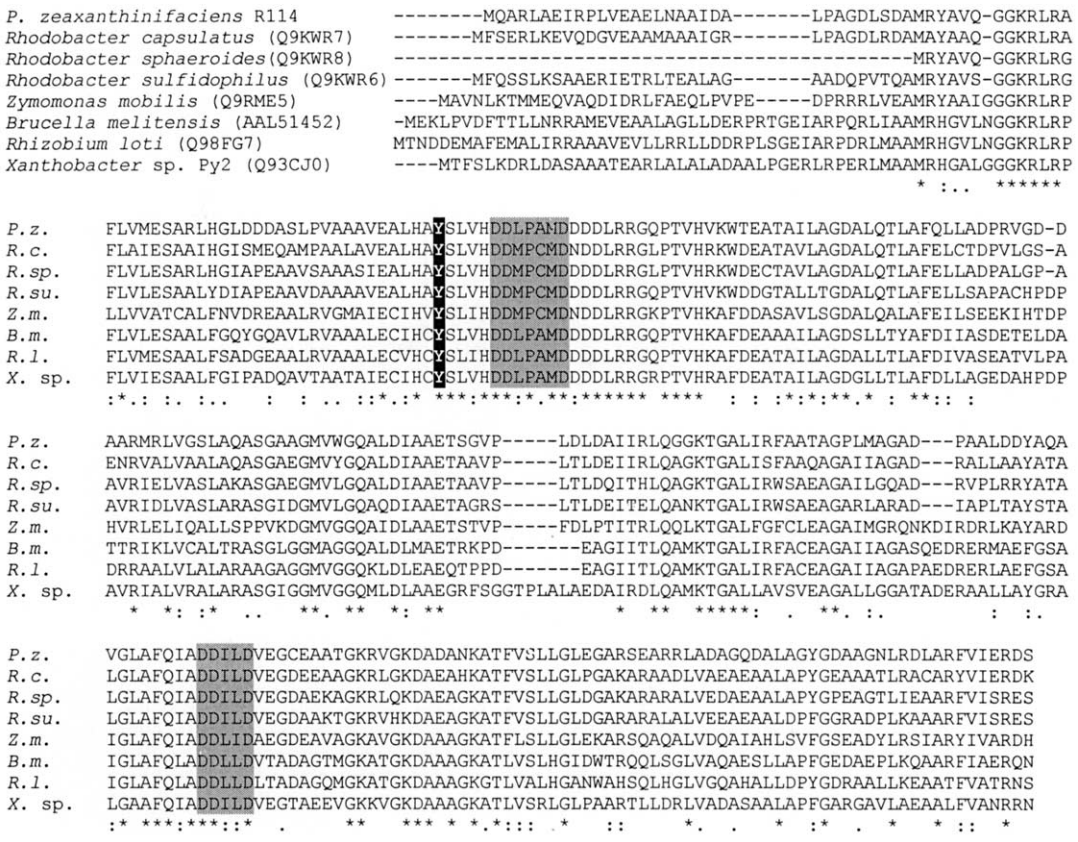

Fig. 6. Alignment of IspA from *P. zeaxanthinifaciens* strain R114 with some orthologues. Clustal W alignment (Thompson et al., 1994). The symbols are: (*) identical residue in all sequences, (:) strongly conserved residues in all sequences, (.) conserved residues in all sequences. The tyrosine in the CLD region, which is characteristic for eubacterial FPP synthases, is shown in white on black background and the FARM (DDxxxxD) and SARM (DDxxD) motifs are shaded in gray.

sequential addition of seven IPP units to FPP and is utilized by members of the genus *Paracoccus* for synthesis of ubiquinone (coenzyme Q-10) (Baj, 2000). Finally, the gene encoding tRNA prenyltransferase was also found by sequence similarity in the genome sequence of *P. zeaxanthinifaciens* strain R114 (Table 1). The enzyme modifies tRNA by condensing DMAPP with an adenosine residue, forming N^6-(Δ^2-isopentenyl)-adenosine-37 (i^6A37).

At present, the only Gram-negative eubacteria in which genes encoding enzymes of the mevalonate pathway have been identified are *B. burgdorferi* (Fraser et al., 1997) and *P. zeaxanthinifaciens*. As more bacteria are examined for the two different pathways of IPP biosynthesis, the distribution of the mevalonate and DXP pathways will become better defined. Furthermore, identification of the genes involved in IPP biosynthesis lends itself to the metabolic engineering of microorganisms with improved production of commercially important isoprenoid compounds.

Acknowledgements

We thank Talia Biran for some initial work in the cloning of the *ispA* gene.

References

Baj, J., 2000. Taxonomy of the genus *Paracoccus*. Acta Microbiol. Pol. 49, 185–200.

Berry, A., Janssens, D., Hümbelin, M., Jore, J.P.M., Hoste, B., Cleenwerck, I., Vancanneyt, M., Bretzel, W., Mayer, A.F., Lopez-Ulibarri, R., Shanmugam, B., Swings, J., Pasamontes, L., 2002. *Paracoccus zeaxanthinifaciens* sp. nov., a zeaxanthin-producing species of *Paracoccus*. Int. J. Syst. Evol. Microbiol. in press (on-line publication: http://www.sgm.ac.uk/IJSEM/PiP/ijsem02368.pdf).

Björk, G.R., 1996. Stable RNA modification. In: Neidhardt, F.C., Curtiss, R., Ingraham, J.L., Reznikoff, W.S., Riley, M., Schaechter, M., Umbarger, H.E. (Eds.). *Escherichia coli* and *Salmonella*, ASM Press, Washington, DC, pp. 861–886.

Bloch, K., 1992. Sterol molecule: structure, biosynthesis, and function. Steroids 57, 378–383.

Bochar, D.A., Stauffacher, C.V., Rodwell, V.W., 1999. Sequence comparisons reveal two classes of 3-hydroxy-3-methylglutaryl Coenzyme A reductase. Mol. Genet. Metab. 66, 122–127.

Boucher, Y., Doolittle, W.F., 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. Mol. Microbiol. 37, 703–716.

Eisenreich, W., Bacher, A., Berry, A., Bretzel, W., Hümbelin, M., Lopez-Ulibarri, R., Mayer, A.F., Yeliseev, A., 2002. Biosynthesis of zeaxanthin via mevalonate in *Paracoccus* species strain PTA-3335. A product-based retrobiosynthetic study. J. Org. Chem. 67, 871–875.

Fraser, C.M., et al., 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature 390, 553–555.

Fujihashi, M., Zhang, Y.-W., Higuchi, Y., Li, X.-Y., Koyama, T., Miki, K., 2001. Crystal structure of *cis*-prenyl chain elongating enzyme, undecaprenyl diphosphate synthase. Proc. Natl. Acad. Sci. USA 98, 4337–4342.

Goodwin, T.W., 1972. Recent developments in the biosynthesis of carotenoids. Biochem. Soc. Symp. 35, 233–244.

Hundle, B., Albertini, M., Nievelstein, V., Beyer, P., Kleinig, H., Armstrong, G.A., Burke, D.H., Hearst, J.E., 1994. Functional assignment of *Erwinia herbicola* Eho10 carotenoid genes expressed in *Escherichia coli*. Mol. Gen. Genet. 245, 406–416.

Kaneda, K., Kuzuyama, T., Takagi, M., Hayakawa, Y., Seto, H., 2001. An unusual isopentenyl diphosphate isomerase found in the mevalonate pathway gene cluster from *Streptomyces* sp. strain CL190. Proc. Natl. Acad. Sci. USA 98, 932–937.

Kuzuyama, T., Takagi, M., Takahashi, S., Seto, H., 2000. Cloning and characterization of 1-deoxy-D-xylulose 5-phosphate synthase from *Streptomyces* sp. strain CL190, which uses both the mevalonate and nonmevalonate pathways for isopentenyl diphosphate biosynthesis. J. Bacteriol. 182, 891–897.

Madison, L.L., Huisman, G.W., 1999. Metabolic engineering of poly(3-hydroxyalkanoates): from DNA to plastic. Microbiol. Mol. Biol. Rev. 63, 21–53.

Paitan, Y., Orr, E., Ron, E.Z., Rosenberg, E., 1999. Genetic and functional analysis of genes required for the post-modification of the polyketide antibiotic TA of *Myxococcus xanthus*. Microbiology 145, 3059–3067.

Pasamontes, L., Hug, D., Tessier, M., Hohmann, H.-P., Schierle, J., 1997. Isolation and characterization of the carotenoid biosynthesis genes of *Flavobacterium* sp. strain R1534. Gene 185, 35–41.

Raman, T.S., Sharma, B.V.S., Jayaraman, J., Ramasarma, T., 1965. Biosynthesis of coenzyme Q in microorganisms. Arch. Biochem. Biophys. 110, 75–84.

Rohdich, F., Hecht, S., Gärtner, K., Adam, P., Krieger, C., Amslinger, S., Arigoni, D., Bacher, A., Eisenreich, W., 2002. Studies on the nonmevalonate terpene biosynthetic pathway: metabolic role of IspH (LytB) protein. Proc. Natl. Acad. Sci. USA 99, 1158–1163.

Rohmer, M., Seemann, M., Horbach, S., Bringer-Meyer, S., Sahm, H., 1996. Glyceraldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative non-mevalonate pathway for terpenoid biosynthesis. J. Am. Chem. Soc. 118, 2564–2566.

Rosa-Putra, S., Disch, A., Bravo, J.-M., Rohmer, M., 1998. Distribution of mevalonate and glyceraldehyde 3-phosphate/pyruvate routes for isoprenoid biosynthesis in some Gram-negative bacteria and mycobacteria. FEMS Microbiol. Lett. 164, 169–175.

Rosa-Putra, S., Hemmerlin, A., Epperson, J., Bach, T.J., Guerra, L.H., Rohmer, M., 2001. Zeaxanthin and menaquinone-7 biosynthesis in *Sphingobacterium multivorum* via the methylerythritol phosphate pathway. FEMS Microbiol. Lett. 204, 347–353.

Schocher, A.J., Wiss, O., 1975. Process for the Manufacture of Zeaxanthin, Hoffmann–La Roche, USA.

Seto, H., Watanabe, H., Furihata, K., 1996. Simultaneous operation of the mevalonate and non-mevalonate pathways in the biosynthesis of isopentenyl diphosphate in *Streptomyces aeriouvifer*. Tetrahedron Lett. 37, 7979–7982.

Smeds, A., Joutsjoki, T., Palva, A., 2001. Identification of a gene cluster for the mevalonate pathway in *Lactobacillus helveticus*. DNA Seq. 12, 187–190.

Smit, A., Mushegian, A., 2000. Biosynthesis of isoprenoids via mevalonate in archaea: the lost pathway. Genome Res. 10, 1468–1484.

Tabernero, L., Bochar, D.A., Rodwell, V.W., Stauffacher, C.V., 1999. Substrate-induced closure of the flap domain in the ternary complex structures provides insights into the mechanism of catalysis by 3-hydroxy-3-methylglutaryl-CoA reductase. Proc. Natl. Acad. Sci. USA 96, 7167–7171.

Takahashi, S., Kuzuyama, T., Seto, H., 1999. Purification, characterization, and cloning of a eubacterial 3-hydroxy-3-methylglutaryl coenzyme A reductase, a key enzyme involved in biosynthesis of terpenoids. J. Bacteriol. 181, 1256–1263.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Wang, K., Ohnuma, S.-I., 1999. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. Trends Biochem. Sci. 24, 445–451.

Wilding, E.I., Brown, J.R., Bryant, A.P., Chalker, A.F., Holmes, D.J., Ingraham, K.A., Iordanescu, S., So, C.Y., Rosenberg, M., Gwynn, M.N., 2000. Identification, evolution, and essentiality of the mevalonate pathway for isopentenyl diphosphate biosynthesis in Gram-positive cocci. J. Bacteriol. 182, 4319–4327.