Cytidine 5′-triphosphate-dependent biosynthesis of isoprenoids: YgbP protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-*C*-methylerythritol

Felix Rohdich*, Juraithip Wungsintaweekul†, Monika Fellermeier†, Silvia Sagner†, Stefan Herz*, Klaus Kis*, Wolfgang Eisenreich*, Adelbert Bacher*, and Meinhart H. Zenk†‡

†Lehrstuhl für Pharmazeutische Biologie, Universität München, Karlstrasse 29, D-80333 Munich, Germany; and *Lehrstuhl für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

Contributed by Meinhart H. Zenk, August 13, 1999

2-*C*-methylerythritol 4-phosphate has been established recently as an intermediate of the deoxyxylulose phosphate pathway used for biosynthesis of terpenoids in plants and in many microorganisms. We show that an enzyme isolated from cell extract of *Escherichia coli* converts 2-*C*-methylerythritol 4-phosphate into 4-diphosphocytidyl-2-*C*-methylerythritol by reaction with CTP. The enzyme is specified by the hitherto unannotated ORF *ygbP* of *E. coli*. The cognate protein was obtained in pure form from a recombinant hyperexpression strain of *E. coli* harboring a plasmid with the *ygbP* gene under the control of a T5 promoter and lac operator. By using the recombinant enzyme, 4-diphosphocytidyl-[2-^{14}C]2-*C*-methylerythritol was prepared from [2-^{14}C]2-*C*-methylerythritol 4-phosphate. The radiolabeled 4-diphosphocytidyl-2-*C*-methylerythritol was shown to be efficiently incorporated into carotenoids by isolated chromoplasts of *Capsicum annuum.* The *E. coli ygbP* gene appears to be part of a small operon also comprising the unannotated *ygbB* gene. Genes with similarity to *ygbP* and *ygbB* are present in the genomes of many microorganisms, and their occurrence appears to be correlated with that of the deoxyxylulose pathway of terpenoid biosynthesis. Moreover, several microorganisms have genes specifying putative fusion proteins with *ygbP* and *ygbB* domains, suggesting that both the YgbP protein and the YgbB protein are involved in the deoxyxylulose pathway. A gene from *Arabidopsis thaliana* with similarity to *ygbP* carries a putative plastid import sequence, which is well in line with the assumed localization of the deoxyxylulose pathway in the plastid compartment of plants.

Terpenes are one of the largest groups of natural products with important representatives in all taxonomic groups. The more than 30,000 naturally occurring terpenes comprise physiologically important compounds such as vitamins A and D, cholesterol, steroid hormones, chlorophylls, and carotenoids, to mention just a few (1).

The pathway of terpenoid biosynthesis in animal and yeast cells via mevalonate has been established by the pioneering studies of Bloch, Cornforth, Lynen, and coworkers (for review, see refs. 2–5). More specifically, mevalonate was shown to be assembled from three molecules of acetyl CoA. A sequence of dehydrogenation, phosphorylation, and dehydration steps affords isopentenyl pyrophosphate (IPP) from mevalonate. IPP can be converted into dimethylallyl pyrophosphate (DMAPP), and IPP and DMAPP jointly serve as universal precursors of terpene biosynthesis.

Rather recently, evidence for the existence of an alternative isoprenoid biosynthetic pathway emerged from independent studies in the research groups of Rohmer and Arigoni (for review, see refs. 6–8), who found that the isotope labeling patterns observed in studies on certain eubacterial and plant terpenoids could not be explained in terms of the mevalonate pathway. Arigoni and coworkers subsequently showed that 1-deoxyxylulose, or a derivative thereof, serves as an intermediate of the novel pathway (9). More recent studies showed the formation of 1-deoxyxylulose 5-phosphate (**3**, Fig. 1) from one molecule each of glyceraldehyde 3-phosphate (**2**) and pyruvate (**1**) (9–11) by an enzyme specified by the *dxs* gene (12, 13). 1-Deoxyxylulose 5-phosphate can be further converted into 2-*C*-methylerythritol 4-phosphate (**4**) by a reductoisomerase specified by the *dxr* gene (Fig. 1) (14, 15).

This paper shows that 2-*C*-methylerythritol 4-phosphate is converted into the respective cytidyl pyrophosphate derivative by an enzyme specified by the unannotated gene *ygbP* from *Escherichia coli*.

Experimental Procedures

Materials. [5-^{3}H]Cytidine 5′-triphosphate, NH_4 salt (24.0 Ci/mmol) was purchased from Amersham Pharmacia Biotech, [α-^{32}P]CTP, tetra-triethylammonium salt (>3,000 Ci/mmol) was from ICN, and [2-^{14}C]pyruvate (17.5 mCi/mmol) and [γ-^{32}P]adenosine 5′-triphosphate (7,000 Ci/mmol) were purchased from NEN.

Sepharose Q FF, Phenyl Sepharose 6FF, Red-Sepharose CL-6B Source 15Q, and Superdex 75 were purchased from Amersham Pharmacia Biotech. Cibacron blue 3GA type 3000 CL was obtained from Sigma.

Oligonucleotides were custom synthesized by MWG Biotech, Ebersberg, Germany.

The synthesis of 2-*C*-methylerythritol 4-phosphate, [2-^{14}C]2-*C*-methylerythritol 4-phosphate, ribitol 5-phosphate, and erythritol 4-phosphate will be reported elsewhere (A.B., W.E., S.H., K.K., F.R., J.W. and M.H.Z., unpublished work).

Preparation of [γ-^{32}P]CTP. A reaction mixture containing 50 mM Tris·hydrochloride, pH 7.6, 10 mM $MgCl_2$, 0.5 mM cytidine 5′-diphosphate, 0.07 μM [γ-^{32}P]ATP (7,000 Ci/mmol), and 1 unit of nucleoside 5′-diphosphate kinase was incubated at 25°C. After 1 h, the enzyme was removed by ultrafiltration with centrifugal filter tubes with a 30-kDa cutoff (Eppendorf).

Partial Purification of 4-Diphosphocytidyl-2-*C*-methylerythritol Synthase. *E. coli* strain DH5α (16) was grown at 37°C in minimal medium with aeration by using a 300-l fermentor. Cells were harvested by centrifugation and were stored at −20°C. Frozen cell mass (400 g) was thawed in 1,200 ml of 50 mM Tris·hydrochloride, pH 8.0, containing 5 mM $MgCl_2$, 1 mM dithioerythritol, and 0.02% sodium azide (buffer A). The suspension was treated with 240 mg of lysozyme and 12 mg of DNase I for 60 min at 37°C. The cells were broken by ultrasonic treatment, and the suspension was centrifuged in a Sorval RC 5B Plus (13,000 rpm, 60 min.). The supernatant was loaded on top of a Sepharose Q FF column (4.6 × 24 cm) at a flow rate of 5

‡To whom reprint requests should be addressed. E-mail: zenk@lrz.uni-muenchen.de.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

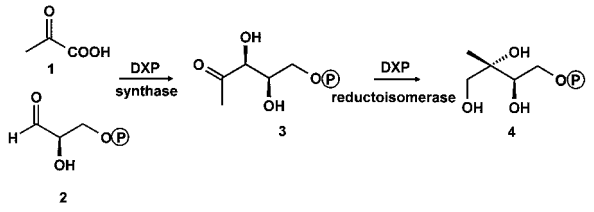

Fig. 1. The deoxyxylulose phosphate pathway of isoprenoid biosynthesis.

ml/min. The column was washed with 300 ml of buffer A and was subsequently developed with a linear gradient of 0–1.0 M NaCl in buffer A (total volume, 600 ml). The effluent was monitored photometrically (280 nm). Fractions were combined. Saturated ammonium sulfate solution was added to a final concentration of 1 M. The solution was placed on top of a Phenyl Sepharose 6FF column (5.6 × 10 cm), which was then developed with a linear gradient of 1.0–0 M ammonium sulfate in buffer A (total volume, 200 ml). The flow rate was 5 ml/min. Fractions were combined and concentrated to 30 ml by ultrafiltration (Amicon UF-10 membrane). The solution was dialyzed against buffer A and applied to a column of Cibacron blue 3GA (type 3000 CL; 2.6 × 10 cm; flow rate, 3 ml/min), which was then developed with buffer A. Fractions were combined and concentrated by ultrafiltration (Amicon UF-10 membrane).

Construction of Expression Plasmid pNCOygbP. The ORF of the putative *ygbP* gene of *E. coli* was amplified by PCR by using the oligonucleotides 5′-AAATTAACCATGGCAACCACTCATTTGG-3′and 5′-TTGGGCCTGCAGCGCCAAAGG-3′ as primers and chromosomal *E. coli* DNA as template. The primers introduced *Nco*I and *Pst*I restriction sites. The amplificate was digested with the restriction endonucleases *Nco*I and *Pst*I. The fragment was ligated into the plasmid expression vector pNCO113 (17), which had been prepared with the same restriction enzymes. Electroporation of the ligation mixture into *E. coli* XL1-blue cells (Stratagene) (18) afforded the recombinant strain XL1 pNCOygbP.

Purification of Recombinant 4-Diphosphocytidyl-2-*C*-methylerythritol Synthase. Cells of the recombinant *E. coli* strain XL1 pNCOygbP (5 g) were suspended in 50 ml of buffer A. Cell extract was prepared as described above. The crude extract was loaded on top of a Sepharose Q FF column (2.6 × 10 cm). The column was developed with a linear gradient of 0–0.5 M NaCl in 300 ml of buffer A (flow rate, 4 ml/min). Fractions were combined, dialyzed, and loaded on top of a Red Sepharose CL-6B column (1.6 × 8 cm; flow rate, 2 ml/min), which was developed with buffer A. The effluent was loaded on top of a Source 15 Q column (volume, 20 ml), which was developed with a linear gradient of 0–0.5 M NaCl in 250 ml of buffer A (flow rate, 3 ml/min). Fractions containing diphosphocytidyl-2-*C*-methylerythritol synthase, as judged by SDS/PAGE, were combined.

Estimation of Molecular Mass. The molecular mass of the native 4-diphosphocytidyl-2-*C*-methylerythritol synthase was estimated by gel filtration with a Superdex 75 column (2.6 × 60 cm) equilibrated in 50 mM Tris·hydrochloride, pH 8.0, containing 100 mM sodium chloride, 1 mM dithioerythritol, and 0.02% sodium azide, at a flow rate of 2 ml/min.

Radiochemical Assay of 4-Diphosphocytidyl-2-*C*-methylerythritol Synthase. Assay mixtures containing 100 mM Tris·hydrochloride, pH 8.0, 20 mM sodium fluoride, 10 mM $MgCl_2$, 100 μM CTP, 11.4 μM [2-^{14}C]2-*C*-methylerythritol 4-phosphate (17.5 mCi/mmol), and diphosphocytidyl-2-*C*-methylerythritol synthase were incubated at 37°C for 20 min. The reaction was terminated by the addition of methanol (20 μl). The mixture was centrifuged. Aliquots were spotted on Sil-NHR thin layer plates (Macherey & Nagel), which were developed with a mixture of *N*-propanol/ethyl acetate/H_2O (6:1:3, vol/vol). Radioactivity was monitored with a PhosphorImager (Storm 860, Molecular Dynamics). The R_f value of 4-diphosphocytidyl-2-*C*-methylerythritol was 0.36.

Photometric Assay of 4-Diphosphocytidyl-2-*C*-methylerythritol Synthase. In this assay, the inorganic pyrophosphate formed by 2-*C*-methylerythritol 4-phosphate synthase is consumed in a cascade of reactions conducive to the reduction of $NADP^+$ (19). Reaction mixtures contained 50 mM Tris·hydrochloride, pH 8.0, 5 mM $MgCl_2$, 1 mM DTT, 200 μM 2-*C*-methylerythritol 4-phosphate, 200 μM CTP, 1 μM glucose-1,6-biphosphate, 500 μM UDP-glucose, 174 μM $NADP^+$, 0.125 units of UDP-glucose pyrophosphorylase, 0.16 units of phosphoglucomutase, 1 unit of glucose 6-phosphate dehydrogenase, and 10 μl of a solution containing 4-diphosphocytidyl-2-*C*-methylerythritol synthase in a total volume of 1 ml. The reaction was monitored photometrically at 340 nm.

Sequence Determination. DNA sequencing was performed by the automated dideoxynucleotide method by using a 377 Prism sequencer from Perkin–Elmer. N-terminal peptide sequences were determined by using a PE Biosystems Model 492 (Weiterstadt, Germany).

Preparation of 4-Diphosphocytidyl-2-*C*-methylerythritol. A solution containing 100 mM Tris·hydrochloride, pH 8.0, 10 mM $MgCl_2$, 46 mM CTP, 46 mM 2-*C*-methyl[2-^{14}C]erythritol 4-phosphate (3.7 μCi/mmol), and 225 μg of recombinant 4-diphosphocytidyl-2-*C*-methylerythritol synthase from *E. coli* in a total volume of 0.7 ml was incubated at 37°C for 1 h. The reaction was monitored by ^{31}P-NMR. The product was purified by HPLC by using a column of Nucleosil 10SB (4.6 × 250 mm; eluent, 0.1 M ammonium formate containing 40% (vol/vol) methanol; flow rate, 1 ml/min). The effluent was monitored by using a diode array photometer from J&M TIDAS, Aalen, Germany, and a radiomonitor from Berthold, Wildbad, Germany. The retention volume of 4-diphosphocytidyl-2-*C*-methylerythritol was 30 ml. Fractions were collected and lyophylized (yield, 7 μmol).

The use of 2-*C*-methyl[2-^{14}C]erythritol 4-phosphate (17.5 mCi/mmol) afforded 4-diphosphocytidyl-2-*C*-methyl[2-^{14}C]erythritol.

NMR Spectroscopy. ^{1}H NMR and ^{1}H decoupled ^{13}C NMR spectra were recorded by using an AVANCE DRX 500 spectrometer from Bruker, Karlsruhe, Germany. The chemical shifts were referenced to external trimethylsilylpropane sulfonate. Two-dimensional correlation experiments (gradient-enhanced double quantum-filtered correlated spectroscopy, heteronuclear multiple quantum correlation) were performed by using XWIN-NMR software from Bruker. ^{1}H decoupled ^{31}P NMR spectra were recorded by using an AC 250 spectrometer from Bruker. Chemical shifts were referenced to external 85% (vol/vol) H_3PO_4.

Preparation of Chromoplasts and Enzyme Assays. Chromoplasts were isolated from *Capsicum annuum* and incubated with radiolabeled substrates as described previously (20). Enzyme assays with cell extract of *E. coli* or chromoplasts from *C. annuum* were analyzed by using published procedures (21).

Results

Recent evidence implicates 2-*C*-methylerythritol 4-phosphate as an intermediate in the deoxyxylulose pathway of terpenoid biosynthesis (14, 15, 20). In search for downstream intermediates

BIOCHEMISTRY

Table 1. Utilization of nucleotide 5′-triphosphates by YgbP protein

	Relative enzyme activity %	
	Wild-type protein*	Recombinant protein†
CTP	100	100
UTP	30	0
GTP	20	8
ATP	20	0
ITP	17	0

*Partial purified; enzyme activity was measured by the radiochemical assay (see *Experimental Procedures*).

†Enzyme activity was measured by the photometric assay (see *Experimental Procedures*).

of this pathway, we incubated radiolabeled 2-*C*-methylerythritol 4-phosphate with *E. coli* cell extracts. Aliquots of the reaction mixture were analyzed by thin-layer chromatography monitored by a PhosphorImager as described in *Experimental Procedures*. A radioactive product with a R_f value of 0.36 was observed when the reaction mixture contained ATP. We assumed tentatively that this compound might be an intermediate of the deoxyxylulose pathway.

An enzyme fraction catalyzing the formation of the compound was partially purified by a sequence of three chromatographic steps as described in *Experimental Procedures*. The purification was accompanied by a severe reduction of the total activity, thus suggesting that a low molecular weight compound required for enzyme action had been lost in the purification procedure. Subsequent experiments showed that CTP could serve more efficiently than ATP as a substrate for the partially purified enzyme (Table 1).

We also found that radioactivity from [α-^{32}P]CTP but not from [γ-^{32}P]CTP was incorporated into the enzyme product obtained with the partially purified *E. coli* enzyme.

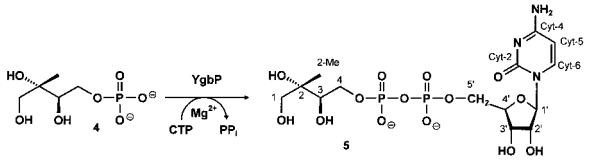

Fig. 2. Reaction catalyzed by YgbP protein (4-diphosphocytidyl-2-*C*-methylerythritol synthase).

To unequivocally determine the structure of the new metabolite, the product obtained by incubation of 2-*C*-methylerythritol 4-phosphate and CTP with partially purified enzyme was isolated chromatographically and was analyzed by ^{13}C, ^{1}H, and ^{31}P NMR spectroscopy (Table 2). The ^{31}P NMR spectrum was characterized by two signals at −7.2 ppm and −7.8 ppm (doublets with ^{31}P^{31}P coupling constants of 20 Hz). This spectroscopic signature was tentatively attributed to a pyrophosphate motif that is also reflected in the ^{13}C NMR spectrum where 4 of 14 signals showed ^{31}P ^{13}C coupling with coupling constants in the range of 5 to 9 Hz. Two-dimensional correlated spectroscopy and heteronuclear multiple quantum correlation experiments identified the spin networks of cytidine and 2-*C*-methylerythritol motifs. On the basis of these data, the structure of the enzyme product was assigned as 4-diphosphocytidyl-2-*C*-methylerythritol (**5**, Fig. 2).

A database search with the ENTREZ browser at the National Center for Biotechnology Information with cytidine 5′-diphosphate (CDP) and pyrophosphorylase as key words retrieved the *acs1* gene from a serotype-specific DNA region of *Haemophilus influenzae* (19). This gene specifies a bifunctional ribulose 5-phosphate reductase/CDP-ribitol pyrophosphorylase. More specifically, the N-terminal domain of this enzyme had been shown to catalyze the formation of CDP-ribitol from ribitol 5-phosphate and CTP.

The 5′ moiety of the *acs1* gene is similar to the unannotated *ygbP* gene of *H. influenzae.* Subsequent database searches start-

Table 2. NMR data of 4-diphosphocytidyl-2C-methylerythritol

	Chemical shifts, ppm			Coupling constants, Hz			
Position	^{1}H	^{13}C	^{31}P	J_{HH}	J_{PH}	J_{PC}	J_{PP}
1	3.36 (d, 1)	66.24 (s)‡		11.7 (1*)			
1*	3.48 (d, 1)			11.7 (1)			
2		73.76 (s)					
2-Methyl	1.02 (s, 3)	18.13 (s)					
3	3.72 (dd, 1)	73.27 (d)		8.4 (4), 2.7 (4*)		7.5	
4	3.85 (ddd, 1)	66.87 (d)		11.0 (4*), 8.3 (3)	6.8	5.7	
4*	4.10 (ddd, 1)			11.0 (4), 2.7 (3)	6.1		
1′	5.68 (d, 1)	89.25 (s)		4.1 (2′)			
2′	4.24 (m, 1)	74.21 (s)					
3′	4.21 (m, 1)	69.09 (s)					
4′	4.17 (m, 1)	82.83 (d)				9.1	
5′	4.10 (m, 1)	64.41 (d)				5.5	
5′*	4.17 (m, 1)						
Cyt-2		163.87 (s)					
Cyt-4		170.51 (s)					
Cyt-5	6.09 (d, 1)	95.99 (s)		7.8 (Cyt-6)			
Cyt-6	7.96 (d, 1)	142.46 (s)		7.8 (Cyt-5)			
P			−7.2 (d)¶				19.6
P*			−7.8 (d)				20.4

*Diastereotopic H position of the index carbon atom.

†Referenced to external trimethylsilylpropane sulfonate. The multiplicities and the relative integral values are given in parentheses.

‡Referenced to external trimethylsilylpropane sulfonate. The multiplicities of the ^{1}H decoupled ^{13}C NMR signals are indicated in parentheses.

§Coupling partners as analyzed from two-dimensional correlated spectroscopy experiments are given in parentheses.

¶Referenced to external 85% orthophosphoric acid. The multiplicites of the ^{1}H decoupled ^{31}P NMR signals are given in parentheses.

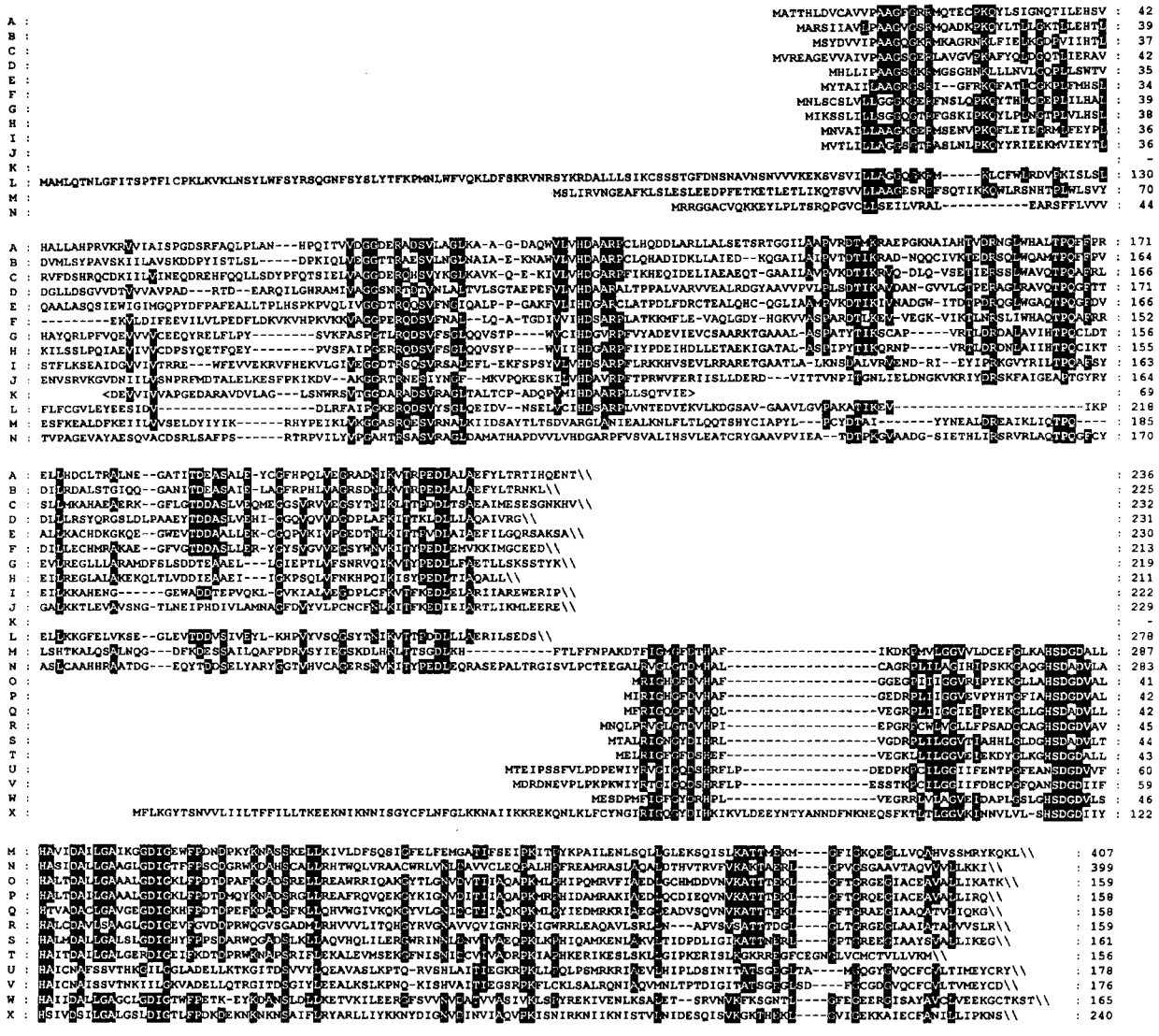

Fig. 3. Deduced amino acid sequences of *ygbP* and *ygbB* genes. Residues conserved in more than 50% of the sequences are shown in inverse contrast. A–L, putative YgbP proteins; M and N, bifunctional YgbP/YgbB proteins; O-X, YgbB proteins. A, *E. coli*; B, *H. influenzae*; C, *B. subtilis*; D, *M. tuberculosis*; E, *Synechocystis* sp. PC6803; F, *Aquifex aeolicus*; G, *Chlamydia trachomatis*; H, *Chlamydia pneumonia*; I, *Thermotoga maritima*; J, *P. horikoshii*; K, *P. falciparum*-(fragment); L, *A. thaliana*; M, *H. pylori*; N, *T. pallidum*; O, *E. coli*; P, *H. influenzae*; Q, *B. subtilis*; R, *M. tuberculosis*; S, *Synechocystis* sp. PC6803; T, *A. aeolicus*; U, *C. trachomatis*; V, *C. pneumonia*; W, *T. maritima*; X, *P. falciparum*.

ing with the unannotated *ygbP* gene of *H. influenzae* uncovered a relatively large number of similar genes from various organisms (Fig. 3, Table 3), including *E. coli*, *Bacillus subtilis*, *Arabidopsis thaliana* (GenBank accession no. AC004136) and *Plasmodium falciparum* (contig number ID_M9Fe7.p1t, fragment). Remarkably, the occurrence of these putative *ygbP* homologs appeared to correlate with the occurrence of the deoxyxylulose pathway. Thus, genes with similarity to *ygbP* were present only in plants and in certain eubacteria, but not in archaea, intracellular parasitic bacteria, yeast, or *Caenorhabditis elegans* (Table 3). It was also noteworthy that the *Arabidopsis* gene encompassed a putative plastid type leader sequence.

An unannotated ORF designated *ygbB* is located downstream from the *ygbP* gene of *E. coli.* This gene is likely to be cotranscribed with the *ygbP* gene and belongs to an unknown gene family. In *E. coli*, *H. influenzae*, *B. subtilis*, *Mycobacterium tuberculosis* (Table 3) genes with similarity to *ygbP* resp. *ygbB* are closely adjacent on the bacterial chromosome. The genomes of *Helicobacter pylori* and *Treponema pallidum* encompass genes specifying putative bifunctional proteins with *ygbP* as well as *ygbB* domains (Fig. 3). These findings suggest that the YgbP and YgbB proteins are involved in the same metabolic pathway.

On the basis of circumstantial evidence, we constructed an *E. coli* strain carrying the *ygbP* gene of *E. coli* on plasmid pNCOygbP under the control of a T5 promoter and a lac operator. On induction with isopropyl-1-thio-β-D-thiogalactopyranoside, the recombinant strain produced large amounts of a polypeptide with an apparent molecular mass of 26 kDa as judged by SDS/PAGE, whereas gel permeation chromatography under nondenaturating conditions suggested an approximate mass of 50 kDa. These data suggest tentatively that the native protein is a dimer. N-terminal sequence determination confirmed the predicted sequence and showed that the start methionine is removed by posttranslational processing.

Cell extracts of the YgbP hyperexpression strain catalyzed the formation of 4-diphosphocytidyl-2-*C*-methylerythritol at high rate. The recombinant protein was obtained in pure form (Fig. 4) and had a specific activity of 23 μmol mg^{-1} min^{-1}. The K_M

BIOCHEMISTRY

Table 3. Deoxyxylulose pathway genes in complete genomes

Organism	Accession nos.			
	*dxs**	*dxr*†	*ygbP*‡	*ygbB*
Bacteria				
E. coli§	AF035440	AB013300	AE000358	AE000358
H. influenzae§	U32822	U32763	U32750	U32750
A. aeolicus§	AE000712	AE000688	AE000734	AE000715
Synechocystis sp. PCC6803¶	D90903	D64000	D90914	D90906
B. subtilis‖	D84432	Z99112	Z99101	Z99101
T. maritima§	AE001815.1	AE001754.1	AE001792.1	AE001738.1
M. tuberculosis‖	Z96072	Z74024	Z92774	Z92774
T. pallidum§	AE001253	AE001235	AE001227	AE001227
H. pylori§	AE001468	AE000541.1	AE001373	AE001474
C. pneumoniae§	AE001686	AE001617	AE001642	AE001639
C. trachomatis§	AE001306	AE001281	AE001320	AE001317
M. genitalium	—	—	—	—
R. powazekii	—	—	—	—
B. burgdorferi	—	—	—	—
Archaea				
P. horikoshii¶	—	—	AE000002	—
Aeropyrum pernix	—	—	—	—
Archeoglobus fulgidus	—	—	—	—
Methanobacterium thermoautotrophicum	—	—	—	—
Methanococcus jannaschii	—	—	—	—
Eukaryotes				
C. elegans	—	—	—	—
Saccharomyces cerevisiae	—	—	—	—

*Specifying 1-deoxyxylulose-5-phosphate synthase
†Specifying 1-deoxyxylulose-5-phosphate reductoisomerase
‡Specifying 4-diphosphocytidyl-2C-methylerythritol synthase
§GenBank database
¶Database of Japan
‖European Molecular Biology Laboratory database

values for 2-*C*-methylerythritol 4-phosphate resp. CTP were 3.14 μM and 131 μM. Ribitol 5-phosphate, erythritol 4-phosphate, ATP, UTP, and ITP could not be used as substrates; about 8% of activity was observed when CTP was replaced by GTP (Table 1). The pH optimum was at 8.3. The enzyme was catalytically active in the presence of Mg^{2+}, Mn^{2+}, or Co^{2+}. Other divalent cations such as Cu^{2+}, Ni^{2+}, Ca^{2+}, Fe^{2+}, or Zn^{2+} could not serve as cofactors.

The recombinant enzyme was used to generate μmol amounts of the product, 4-diphosphocytidyl-2-*C*-methylerythritol. NMR analysis confirmed the identity of this material with that produced by the partially purified wild-type protein.

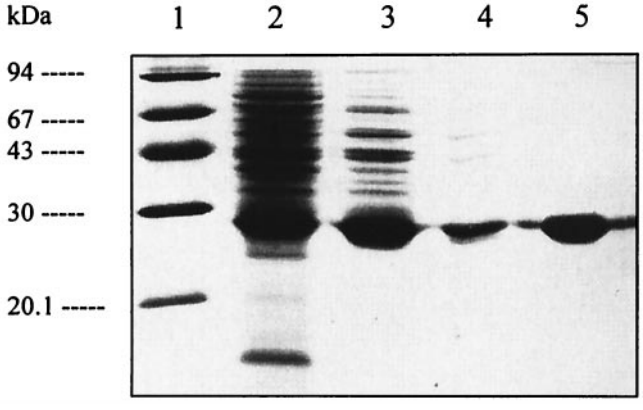

Fig. 4. SDS/PAGE. Lane 1, molecular mass markers; lane 2, crude cell extract of the recombinant *E. coli* strain expressing YgbP protein; lane 3, YgbP protein after Sepharose Q chromatography; lane 4, YgbP protein after Red Sepharose chromatography; lane 5, YgbP protein after Source 15Q chromatography.

Incubation of 11.4 μM [2-^{14}C]4-diphosphocytidyl-2-*C*-methylerythritol (specific activity, 17.5 mCi/mmol) with a chromoplast preparation of *C. annuum* resulted in incorporation of 40% of the proffered radioactivity into the carotenoid fraction. β-Carotene isolated from the carotenoid mixture had a specific radioactivity of 0.38 μCi/μmol.

Discussion

We have shown that the *ygbP* gene of *E. coli* can be expressed to high levels in a homologous hyperexpression strain. The N-terminal methionine residue of the protein is removed by posttranslational processing. The resulting polypeptide has a mass of 26 kDa. Gel permeation chromatography experiments suggest tentatively that the native protein is a homodimer.

Our data show that YgbP protein catalyzes the formation of 4-diphosphocytidyl-2-*C*-methylerythritol from 2-*C*-methylerythritol 4-phosphate and CTP. Notably, the pure recombinant protein has a higher specificity with respect to the nucleotide triphosphate substrate as compared with the partially purified enzyme from *E. coli* wild strain extract (Table 1). This may be caused by the presence of contaminating proteins and/or low molecular weight compounds in the partially purified wild-type protein fraction. The enzyme is highly specific with regard to the branched-chain 2-*C*-methylerythritol 4-phosphate as substrate. Erythritol 4-phosphate and ribitol 5-phosphate were unable to serve as substrates.

The enzyme product, 4-diphosphocytidyl-2-*C*-methylerythritol, can serve as precursor for the biosynthesis of carotenoids by chromoplasts from *C. annuum.* More specifically, 40% of the

proffered radioactivity was incorporated into the fraction of lipophilic compounds. The low relative specific activity of isolated β-carotene (2% as compared with that of the proffered 4-diphosphocytidyl-2-*C*-methylerythritol) is not surprising because the material formed *de novo* was diluted by the large amount of preformed unlabeled carotene present in the chromoplasts.

It could be argued that the radioactivity from 4-diphosphocytidyl 2-*C*-methylerythritol may have been diverted to the terpenoid fraction of the chromoplasts after hydrolytic cleavage affording 2-*C*-methylerythritol 4-phosphate, which had been shown earlier to serve as a precursor for terpenoids in chromoplasts (20). This argument cannot be ruled out conclusively on the basis of the present data. However, the involvement of YgbP and YgbB proteins in the deoxyxylulose pathway of terpenoid biosynthesis is supported by comparative whole genome analysis. In the microbial genomes databases that are currently in the public domain, the distribution of genes with similarity of *dxs* (specifying 1-deoxyxylulose 5-phosphate synthase), *dxr* (specifying 1-deoxyxylulose 5-phosphate reductoisomerase), *ygbP*, and *ygbB* follow the same pattern (Table 3) with one possible exception (see below). Moreover, the occurrence of these genes is orthogonal to the genes of the mevalonate pathway of terpenoid biosynthesis, except that *Mycoplasma genitalium*, *Rickettsia prowazekii*, and *Borrelia burgdorferi* have neither mevalonoid nor deoxyxylulose pathway genes. *Pyrococcus horikoshii* (Table 3) has a putative *ygbP* homolog, but no homologs of *dxs*, *dxr*, or *ygbB*. Notably, *P. horikoshii* has a eubacterial (as opposed to archaeal) type riboflavin synthase (22), although the microorganism has been assigned as archaebacterium.

The *Arabidopsis* homolog of the *ygbP* gene and the *Plasmodium* homolog (GenBank accession no. AE001394) of the *ygbB* gene specify putative leader sequences in line with the assumed location of the deoxyxylulose pathway enzymes in organelles.

These data indicate that the product of the *ygbP* gene, 4-diphosphocytidyl-2-*C*-methylerythritol, serves as an intermediate in the deoxyxylulose pathway of isoprenoid biosynthesis. The specific metabolic role of the YgbB protein in this pathway remains to be determined.

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 369). We thank F. Wendling, A. Werner, and H. Hofner for help with the preparation of the manuscript and Dr. P. Köhler for peptide sequencing.

1. Saccettini, J. C. & Poulter, C. D. (1997) *Science* **277,** 1788–1789.
2. Qureshi, N. & Porter, J. W. (1981) *Biosynthesis of Isoprenoid Compounds,* eds. Porter, J. W. & Spurgeon, S. L. (Wiley, New York), Vol. 1, pp. 47–94.
3. Bloch, K. (1992) *Steroids* **57,** 378–382.
4. Bach, T. J. (1995) *Lipids* **30,** 191–202.
5. Bochar D. A., Friesen, J. A., Stauffacher, C. V. & Rodwell, V. W. (1999) *Comprehensive Natural Product Chemistry*, eds. Barton, D. & Nakanishi, K. (Pergamon, Oxford), Vol. 2, pp. 15–44.
6. Eisenreich, W., Schwarz, M., Cartayrade, A., Arigoni, D., Zenk, M. & Bacher, A. (1998) *Chem. Biol.* **5,** R221–R233.
7. Rohmer, M. (1998) *Prog. Drug Res.* **50,** 135–154.
8. Rohmer, M. (1999) *Comprehensive Natural Products Chemistry*, eds. Barton, D. & Nakanishi, K. (Pergamon, Oxford), Vol. 2, pp. 45–68.
9. Broers, S. T. J. (1994) Ph.D. thesis (Eidgenössiche Technische Hochschule, Zürich, Switzerland).
10. Schwarz, M. K. (1994) Ph.D. thesis (Eidgenössiche Technische Hochschule, Zürich, Switzerland).
11. Rohmer, M., Seemann, M., Horbach, S., Bringer-Meyer, S. & Sahm, H. (1996) *J. Am. Chem. Soc.* **118,** 2564–2566.
12. Sprenger, G. A., Schörken, U., Wiegert, T., Grolle, S., deGraaf, A. A., Taylor, S. V., Begley, T. P., Bringer-Meyer, S. & Sahm, H. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12857–12862.
13. Lois, L. M., Campos, N., Putra, S. R., Danielsen, K., Rohmer, M. & Boronat, A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2105–2110.
14. Takahashi, S., Kuzuyama, T., Watanabe, H. & Seto, H. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9879–9884.
15. Lange, B. M. & Croteau, R. (1999) *Arch. Biochem. Biophys.* **365,** 170–174.
16. Woodcock, D. M., Crowser, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989) *Nucleic Acids Res.* **17,** 3469–1575.
17. Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987) *BioTechniques* **5,** 376–379.
18. Stüber, D., Matile, H. & Garotta, G. (1990) *Immunol. Methods,* eds. Lefkovits, F. & Pernis, P. (Academic, San Diego), Vol. IV, pp. 121–152.
19. Follens, A., Veigha-Da-Cunha, M., Merckx, R., Van Schaftingen, E. & Van Eldere J. (1999) *J. Bacteriol.* **181,** 2001–2007.
20. Fellermeier, M., Kis, K., Sagner, S., Maier, U., Bacher, A. & Zenk, M. H. (1999) *Tetrahedron Lett.* **40,** 2743–2746.
21. McCaskill, D. & Croteau, R. (1993) *Anal. Biochem.* **215,** 142–149.
22. Eberhardt, S., Korn, S., Lottspeich, F. & Bacher, A. (1997) *J. Bacteriol.* **179,** 2938–2943.