J. Biochem. 108, 995-1000 (1990)

Cloning and Nucleotide Sequence of the *ispA* Gene Responsible for Farnesyl Diphosphate Synthase Activity in *Escherichia coli*[1]

Shingo Fujisaki,* Hiroshi Hara, Yukinobu Nishimura,** Kensuke Horiuchi,** and Tokuzo Nishino*****

**Department of Chemistry, Faculty of General Education, Gifu University, Gifu, Gifu 501-11; **National Institute of Genetics, Mishima, Shizuoka 411; and ***Department of Biochemistry and Engineering, Faculty of Engineering, Tohoku University, Sendai, Miyagi 980*

Received for publication, July 30, 1990

The molecular cloning and the determination of the nucleotide sequence of the *ispA* gene responsible for farnesyl diphosphate (FPP) synthase [EC 2.5.1.1] activity in *Escherichia coli* are described. *E. coli ispA* strains have temperature-sensitive FPP synthase, and the defective gene is located at about min 10 on the chromosome. The wild-type *ispA* gene was subcloned from a λ phage clone containing the chromosomal fragment around min 10, picked up from the aligned genomic library of Kohara *et al.* [Kohara, Y., Akiyama, K., & Isono, K. (1987) *Cell* 50, 495-508]. The cloned gene was identified as the *ispA* gene by the recovery and amplification of FPP synthase activity in an *ispA* strain. A 1,452-nucleotide sequence of the cloned fragment was determined. This sequence specifies two open reading frames, ORF-1 and ORF-2, encoding proteins with the expected molecular weights of 8,951 and 32,158, respectively. A part of the deduced amino acid sequence of ORF-2 showed similarity to the sequences of eucaryotic FPP synthases and of *crtE* product of a photosynthetic bacterium. The plasmid carrying ORF-2 downstream of the *lac* promoter complemented the defect of FPP synthase activity of the *ispA* mutant, showing that the product encoded by ORF-2 is the *ispA* product. The maxicell analysis indicated that a protein of molecular weight 36,000, approximately consistent with the molecular weight of the deduced ORF-2-encoded protein, is the gene product.

A variety of isoprenoid compounds whose physiological roles differ from each other are synthesized by various organisms (*1*). *Escherichia coli* contains isoprenoids such as isopentenyl tRNA, isoprenoid quinones, and sugar carrier lipids (*2, 3*). In addition to the above compounds, other isoprenoids such as sterols and farnesyl proteins are found in eucaryotic cells (*4*). A branch point of the synthetic pathway of various isoprenoids is a reaction utilizing farnesyl diphosphate (FPP) or both isopentenyl diphosphate (IPP) and FPP, and the concentrations of IPP and FPP are considered to influence the rates of synthesis of various isoprenoids (*5*).

FPP synthase catalyzes the condensation of IPP with dimethylallyl diphosphate or geranyl diphosphate to yield FPP. Recently, cDNAs or a genomic clone encoding FPP synthase of rat, human, and yeast have been isolated and sequenced (*6-9*). In rat and human, transcription of these genes was shown to be affected by cellular cholesterol level (*6, 10*), as was that of the genes of low density lipoprotein receptor (*11*), 3-hydroxy-3-methylglutaryl CoA reductase (*12*), and 3-hydroxy-3-methylglutaryl CoA synthase (*13*). FPP synthase of procaryotes is considered to have different roles and properties from the enzyme of eucaryotes, because procaryotic cells do not contain sterols. We previously isolated the *E. coli ispA* mutant which produces the temperature-sensitive FPP synthase (*2*). In order to elucidate the roles and properties of the procaryotic FPP synthase, we cloned the *ispA* gene and determined the nucleotide sequence.

[1] This study was carried out under the NIG Cooperative Research Program ('89-20), and was supported in part by grants from the Ministry of Education, Science and Culture of Japan.

Abbreviations: FPP, farnesyl diphosphate; IPP, isopentenyl diphosphate; ORF, open reading frame.

MATERIALS AND METHODS

Enzymes and Chemicals—Restriction enzymes and other DNA-modifying enzymes were purchased from Takara Shuzo and Toyobo. [α-^{32}P]dCTP (>800 Ci/mmol), L-[^{35}S]methionine (1,000 Ci/mmol), and [1-^{14}C]IPP (53 Ci/mol) were from Amersham. IPP and dimethylallyl diphosphate were synthesized by phosphorylation of the corresponding prenols (*14*).

Bacterial Strains, Phages, and Plasmids—*E. coli* strain GP407 (*ispA recA*) (*2*) was used as the *ispA* mutant. Strains JM109 (*15*), MV1184 (*16*) were employed for the cloning of the *ispA* locus. Strain CSR603 (*uvrA recA phr*) (*17*) was used for the maxicell technique. Phage λ2H5, from the *E. coli* genomic library of Kohara *et al.* (*18*), was used as a source of the *ispA* gene. Phage M13KO7 (*16*) was used as the helper phage for the isolation of single-stranded DNA from plasmid. Plasmids pUC19 (*15*), pUC118, and pUC119 (*16*) were used as the vectors. Other plasmids used in this work are described in Fig. 1 and in an appropriate experimental section. Plasmid pHP1 was constructed from the 3.5 kb *Hin*dIII-*Pst*I fragment of λ2H5 inserted between the *Hin*dIII and *Pst*I sites of pUC19. Plasmids

pSAP1, pSMP1 and pHX1 were constructed by subcloning the 2.9 kb *Sal*I-*Pst*I fragment, 2.7 kb *Sma*I-*Pst*I fragment, and 2.8 kb *Hin*dIII-*Xho*I fragment, respectively, between the corresponding sites of the pUC119. pR1, pR3, and pR6 were deletions of pHX1 produced by exonuclease III. pR10 and pR11 were constructed from the 1.5 kb *Eco*RI-*Sal*I fragment of pR1 inserted at the same site of pUC118 and pUC119, respectively. pR30 and pR31 were derived from pR3, and pR60 and pR61 from pR6, in the same way as pR10 and pR11.

Media and Growth Conditions—The LB and LB agar media (*19*) were used for the growth of bacteria. Glucose or thymine or both were added when needed. 2×YT (*19*) was used for the growth of MV1184 cells harboring a plasmid for the preparation of single-stranded DNA. Bacteria were grown at 37°C.

Measurement of FPP Synthase Activity—FPP synthase activity was assayed in the crude extract as described previously (*2*) with slight modifications. The reaction mixture contained, in a final volume of 0.1 ml, 0.5 μmol of $MgCl_2$, 0.01 μmol of $MnCl_2$, 1.0 μmol of 2-mercaptoethanol, 5.0 μmol of potassium phosphate buffer (pH 7.5), 0.9 nmol of [^{14}C]IPP (diluted to the specific radioactivity of 25 Ci/mol), 4.5 nmol of dimethylallyl diphosphate, and enzyme (50 μg of protein). After incubation, 20 μl of 6 M HCl was added to stop the enzyme reaction. Hydrolysis and extraction of the reaction products were carried out as described previously (*20*).

DNA Sequence Analysis—After complete digestion of plasmids pSAP1 and pHX1 with restriction enzymes *Bam*HI and *Sac*I, exonuclease III deletions were prepared according to the method of Henikoff (*21*). The isolated deletions were independently introduced into MV1184 cells, and their single-stranded DNAs were prepared after infection with phase M13KO7 as described (*16*). Nucleotide sequence analysis was performed by the dideoxy chain termination method of Sanger *et al.* (*22*). Computer analysis and comparison of DNA sequence were performed, using DNASIS DNA sequence analysis software (Hitachi Software Engineering).

Maxicell Experiment—Plasmid-encoded polypeptides were detected by the maxicell technique of Sancar *et al.* (*23*). Strain CSR603 harboring a plasmid to be analyzed was grown in M9 medium (*19*) containing 1% casamino acids to 2×10^8 cells per ml and irradiated with UV light. Cells were transferred to a tube and incubated overnight in the presence of 100 μg of D-cycloserine per ml. Cells were collected, washed, and then suspended in half of the original volume of Hershay medium (*24*). They were incubated for 1 h at 37°C. To 1 ml of the suspension was added 30 μCi of L-[^{35}S]methionine, and the mixture was further incubated for 1 h. The labeled cells were collected, suspended in 0.1 ml of the sample buffer for SDS-PAGE, and heated for 4 min at 100°C. The proteins were precipitated with cold 5% trichloroacetic acid, solubilized in the sample buffer, and fractionated by SDS-PAGE (*25*). Labeled proteins were located by fluorography.

RESULTS

Cloning of ispA Locus—The *ispA* mutant has temperature-sensitive FPP synthase, and the defective gene was mapped at min 10 between *tsx* and *lon* loci on the *E. coli* chromosome (*2*). To localize precisely the *ispA* locus, several chromosomal fragments of *Eco*RI or *Pst*I digestion from phage λ2H5 were subcloned into the same sites of

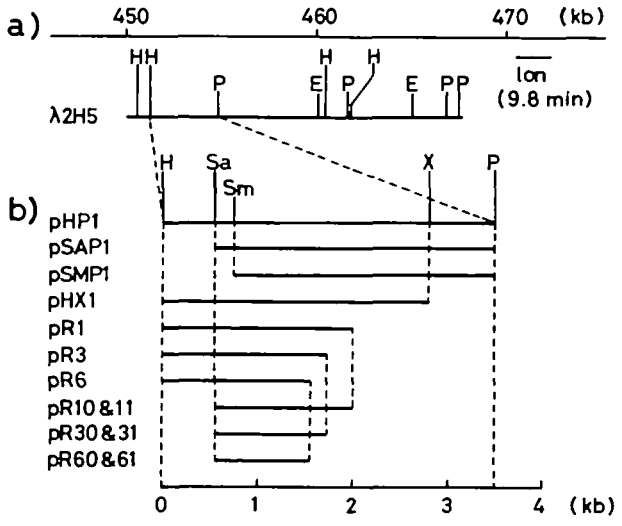

Fig. 1. **Restriction maps of the chromosomal DNA insert.** (a) A fragment inserted in λ2H5. Scales are expressed in kilobase (kb) coordinates as described by Kohara *et al.* (*18*), and the position of the *lon* gene on the chromosome is shown. (b) Fragments subcloned in plasmids. Restriction enzymes: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; Sa, *Sal*I; Sm, *Sma*I; X, *Xho*I.

TABLE I. **Increase in FPP synthase activity in the crude extracts of the transformants derived from JM109.** Enzyme reaction was carried out for 10 min at 30°C as described in "MATERIALS AND METHODS." Activity was expressed as picomoles of IPP incorporated into acid-labile products.

Plasmids	Activity (pmol/min/mg protein)	Ratio
None	35	1.0
pP1	387	11.1
pHP1	390	11.1
pSAP1	426	12.2
pSMP1	53	1.5
pHX1	429	12.3
pR1	582	16.6
pR3	38	1.1
pR6	22	0.6

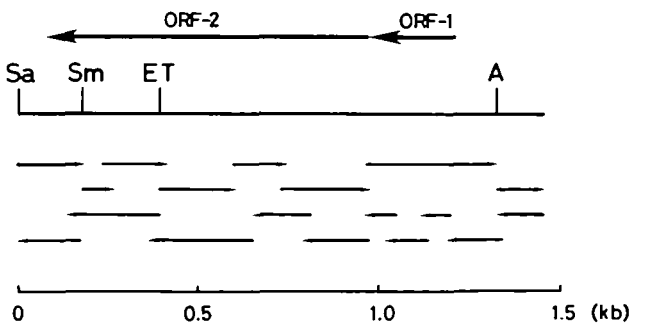

Fig. 2. **Sequencing strategy.** The smaller arrows below the restriction map indicate the direction and extent of each sequence determination. The larger arrows above the restriction map indicate the positions and orientations of ORF-1 and ORF-2. Restriction enzymes: Sa, *Sal*I; Sm, *Sma*I; ET, *Eco*T14I; A, *Acc*I.

pUC19. λ2H5 is a clone in the aligned genomic library (*18*) and contains a 17.7 kb chromosomal fragment separated from the *lon* gene by about 2.5 kb (Fig. 1). One of the resulting plasmids, pP1, which contained a 7.2 kb *Pst*I fragment consisting of the left end of the chromosomal fragment and a part of the left arm of the phage vector, complemented the defect of FPP synthase activity of an *ispA* strain, GP407. Furthermore, JM109 (*ispA*$^+$) harboring this plasmid overproduced the enzyme activity by about 10 times when compared with the strain carrying no plasmid (Table I). We succeeded in cloning a gene that produced the FPP synthase activity, regarding approximately 10-fold overproduction of the enzyme activity in the transformant as the criterion of the expression of the cloned gene. Plasmids pHP1, pSAP1, pHX1, and pR1 elevated the enzyme activity, but pSMP1, pR3, and pR6 did

```
   1 GGTGATTTCCGGGAACAATTTAATGATAAACTTCATGGCGGCAATGGTTCGTTGGCAAGC   60

  61 CTTAAGCGACTTGTATAGGGAAAAATACAGCAGCCCACACCTGCGGCTGCATCCAGGCGC  120

 121 GGAAGTATACCACTAACATCGCTTTGCTGTGCACATCACCTTACCATTGCGCGTTATTTG  180

 181 CTATTTGCCCTGAGTCCGTTACCATGACGGGGCGAAAAATATTGAGAGTCAGACATTCAT  240

 241 TATGCCGAAGAAAAATGAGGCGCCCGCCAGCTTTGAAAAGGCGCTGAGCGAGCTGGAACAG 301
   1  MetProLysLysAsnGluAlaProAlaSerPheGluLysAlaLeuSerGluLeuGluGln  20

 302  ATTGTAACCCGTCTGGAAAGTGGCGACCTGCCGCTGGAAGAGGCGCTGAACGAGTTCGAA 361
  21  IleValThrArgLeuGluSerGlyAspLeuProLeuGluGluAlaLeuAsnGluPheGlu  40

 362  CGCGGCGTGCAGCTGGCACGTCAGGGGCAGGCCAAATTACAACAAGCCGAACAGCGCGTA 421
  41  ArgGlyValGlnLeuAlaArgGlnGlyGlnAlaLysLeuGlnGlnAlaGluGlnArgVal  60

 422  CAAATTCTGCTGTCTGACAATGAAGACGCCTCTCTAACCCCTTTTACACCGGACAATGAG 481
  61  GlnIleLeuLeuSerAspAsnGluAspAlaSerLeuThrProPheThrProAspAsnGlu  80

 482  TAATGGACTTTCCGCAGCAACTCGAAGCCTGCGTTAAGCAGGCCAACCAGGCGCTGAGC 540
   1  **MetAspPheProGlnGlnLeuGluAlaCysValLysGlnAlaAsnGlnAlaLeuSer  19

 541 CGTTTTATCGCCCCACTGCCCTTTCAGAACACTCCCGTGGTCGAAACCATGCAGTATGGC 600
  20 ArgPheIleAlaProLeuProPheGlnAsnThrProValValGluThrMetGlnTyrGly  39

 601 GCATTATTAGGTGGTAAGCGCCTGCGACCTTTCCTGGTTTATGCCACCGGTCATATGTTC 660
  40 AlaLeuLeuGlyGlyLysArgLeuArgProPheLeuValTyrAlaThrGlyHisMetPhe  59

 661 GGCGTTAGCACAAACACGCTGGACGCACCCGCTGCCGCCGTTGAGTGTATCCACGCTTAC 720
  60 GlyValSerThrAsnThrLeuAspAlaProAlaAlaAlaValGluCysIleHisAlaTyr  79

 721 TCATTAATTCATGATGATTTACCGGCAATGGATGATGACGATCTGCGTCGCGGTTTGCCA 780
  80 SerLeuIleHisAspAspLeuProAlaMetAspAspAspAspLeuArgArgGlyLeuPro  99

 781 ACCTGCCATGTGAAGTTTGGCGAAGCAAACGCGATTCTCGCTGGCGACGCTTTACAAACG 840
 100 ThrCysHisValLysPheGlyGluAlaAsnAlaIleLeuAlaGlyAspAlaLeuGlnThr 119

 841 CTGGCGTTCTCGATTTTAAGCGATGCCGATATGCCGGAAGTGTCGGACCGCGACAGAATT 900
 120 LeuAlaPheSerIleLeuSerAspAlaAspMetProGluValSerAspArgAspArgIle 139

 901 TCGATGATTTCTGAACTGGCGAGCGCCAGTGGTATTGCCGGAATGTGCGGTGGTCAGGCA 960
 140 SerMetIleSerGluLeuAlaSerAlaSerGlyIleAlaGlyMetCysGlyGlyGlnAla 159

 961 TTAGATTTAGACGCGGAAGGCAAACACGTACCTCTGGACGCGCTTGAGCGTATTCATCGT 1020
 160 LeuAspLeuAspAlaGluGlyLysHisValProLeuAspAlaLeuGluArgIleHisArg 179

1021 CATAAAACCGGCGCATTGATTCGCGCCGCCGTTCGCCTTGGTGCATTAAGCGCCGGAGAT 1080
 180 HisLysThrGlyAlaLeuIleArgAlaAlaValArgLeuGlyAlaLeuSerAlaGlyAsp 199

1081 AAAGGACGTCGTGCTCTGCCGGTACTCGACAAGTATGCAGAGAGCATCGGCCTTGCCTTC 1140
 200 LysGlyArgArgAlaLeuProValLeuAspLysTyrAlaGluSerIleGlyLeuAlaPhe 219

1141 CAGGTTCAGGATGACATCCTGGATGTGGTGGGAGATACTGCAACGTTGGGAAAACGCCAG 1200
 220 GlnValGlnAspAspIleLeuAspValValGlyAspThrAlaThrLeuGlyLysArgGln 239

1201 GGTGCCGACCAGCAACTTGGTAAAAGTACCTACCCTGCACTTCTGGGTCTTGAGCAAGCC 1260
 240 GlyAlaAspGlnGlnLeuGlyLysSerThrTyrProAlaLeuLeuGlyLeuGluGlnAla 259

1261 CGGAAGAAAGCCCGGGATCTGATCGACGATGCCCGTCAGTCGCTGAAACAACTGGCTGAA 1320
 260 ArgLysLysAlaArgAspLeuIleAspAspAlaArgGlnSerLeuLysGlnLeuAlaGlu 279

1321 CAGTCACTCGATACCTCGGCACTGGAAGCGCTAGCGGACTACATCATCCAGCGTAATAAA 1380
 280 GlnSerLeuAspThrSerAlaLeuGluAlaLeuAlaAspTyrIleIleGlnArgAsnLys 299

1381 TAAACAATAAGTATTAATAGGCCCCTGATGAGTTTTGATATTGCCAAATACCCGACCCTG 1440
      ***

1441 GCACTGGTCGAC                                                 1452
```

Fig. 3. **Nucleotide and deduced amino acid sequence of the *ispA* locus.** The DNA sequence of the sense strand of the *ispA* locus is shown. The nucleotides are numbered from the end of pR1 generated by digestion with exonuclease III. The ribosome-binding sites for ORF-1 and *ispA* (ORF-2) are boxed. The −10 and −35 regions of promoters recognized by σ^{70} and σ^{32} are underlined.

not (Table I). Thus, we concluded that the *ispA* gene was contained in the 1.5 kb fragment between the *Sal*I site and the right end of the chromosomal insert of pR1 (Fig. 1). This fragment was subcloned to construct plasmids pR10 and pR11, in which the orientations of the 1.5 kb chromosomal insert to the *lac* promoter of the vector are opposite to each other. Both plasmids enabled cells to overproduce the enzyme activity (not shown). So we determined the nucleotide sequence of the 1.5 kb chromosomal insert of pR10.

Nucleotide Sequence—The nucleotide sequence of the 1.5 kb chromosomal fragment was determined by the dideoxy chain termination method. Figure 2 illustrates the sequencing strategy used. The ordered fragments were obtained as described in "MATERIALS AND METHODS" or by digestion with the restriction enzymes described in Fig. 2. The complete nucleotide sequence is shown in Fig. 3.

There are two ORFs (designated as ORF-1 and ORF-2) in one strand and one ORF in the other strand. ORF-1 is located at nucleotides 242-481 and encodes a protein of the expected molecular weight of 8,950. ORF-1 possesses a potential ribosome binding site (*26*), GAG, at nucleotides 226-228. ORF-2 is located at nucleotides 484 to 1,380, and therefore overlaps with the termination codon of ORF-1 by one base pair. It encodes a protein with the expected molecular weight of 32,158, and possesses a potential ribosome binding site, GGA, at nucleotides 472-474. As described later, ORF-2 was identified as the *ispA* gene.

TABLE II. **Complementation of the defect of FPP synthase activity of *ispA* strains by the plasmid carrying ORF-2.** Cells were grown in 30 ml of LB containing glucose to an A_{660} of 0.9. Isopropyl β-D-thiogalactopyranoside was added to a final concentration of 1 mM, and after 2 h the cells were collected to prepare the crude extract. Incubation was carried out for 5 min for the reaction with the enzyme of the strains harboring pR60, and for 30 min for the reaction with the enzyme of the other strains at the indicated temperature. Activity was expressed in the same way as in Table I. The values shown represent the averages of two measurements with the ranges in parentheses.

Host	Plasmid	Activity (pmol min mg protein)	
		30°C	40°C
JM109	pUC118	27.6 (±2.1)	14.2 (±1.1)
JM109	pR60	516 (±44)	458 (±109)
JM109	pR61	17.5 (±0.4)	13.5 (±0.4)
GP407	pUC118	8.0 (±0.4)	0.4 (±0)
GP407	pR60	458 (±44)	407 (±15)
GP407	pR61	10.2 (±1.1)	0.4 (±0.4)

These two ORFs are transcribed counterclockwise on the chromosome. Sequences similar to the −10 and −35 consensus sequences recognized by σ^{70} (*27*) were found at nucleotides 200-205 (TACCAT) and 178-183 (TTGCTA), respectively, and sequences similar to consensus sequences recognized by σ^{32} (*28*) were found at nucleotides 459-468 (CCCTTTTAC) and 437-445 (GACAATGAA) (the underlined bases correspond to the consensus sequences).

Identification of ORF-2 as ispA Gene—Since the plasmid pSMP1 which lacked a part of ORF-2 did not enable cells to overproduce the enzyme activity (Table I), ORF-2 seemed to be *ispA* gene. To confirm this conclusion, plasmids pR60 and pR61, which consisted of 976 bp of the chromosomal insert (nucleotides 477-1452 in Fig. 3) and so contained the entire region of ORF-2, were assayed for their ability to complement an *ispA* mutation. The defect of FPP synthase activity at 40°C of strain GP407 (*ispA*) was complemented by pR60, in which the direction of ORF-2 is the same as that of the *lac* promoter upstream of ORF-2. Furthermore GP407 harboring pR60 overproduced the FPP synthase activity to nearly the same extent as JM109 (*ispA*$^+$) harboring pR60 (Table II). The plasmid pR61 in which the orientation of the chromosomal insert with respect to the *lac* promoter is opposite from that in pR60 neither complemented an *ispA* mutation nor enabled cells to overproduce the enzyme activity. From these results, it was concluded that ORF-2 under the control of the *lac* promoter was sufficient to complement the *ispA* mutation, and that ORF-2 is the *ispA* gene.

Radiolabeling of Plasmid-Encoded Proteins—The poly-

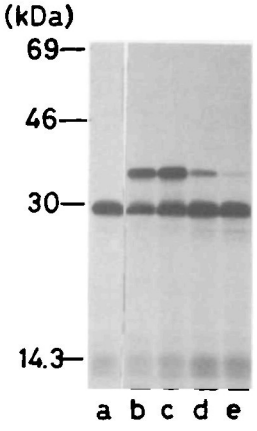

Fig. 4. **Radiolabeling of plasmid-encoded proteins in maxicells.** Proteins directed by plasmid DNA were radiolabeled with [^{35}S]methionine and analyzed on a 10% SDS polyacrylamide gel as described in "MATERIALS AND METHODS." Molecular weight standards (indicated on the left) are as follows: bovine serum albumin (69,000), ovalbumin (46,000), carbonic anhydrase (30,000), and lysozyme (14,300). Strains used were transformants of CSR603, and the plasmids used were as follows: a, pUC118; b, pR10; c, pR11; d, pR30; e, pR31.

```
( a )
crtE product          200IGSAFQIADDLKDALMSAEAMGK222
                        :: ::::  ::  :       ::
ispA product          215IGLAFQVQDDILDVVGDTATLGK237
                        :  :: ::: ::  ::      ::
FPP synthase (Rat)    235MGEFFQIQDDYLDLFGDPSVTGK257
                        :  :: ::: ::  :       ::
FPP synthase (Yeast)  232LGEYFQIQDDYLDCFGTPEQIGK254

( b )
ispA product  67DAPAAAVECIHAYSLIHDDLPAMDDDDLRRGLPTCHVKFGEANAILAGDAL117
              :: : : :  :  :: :::::: :  : ::: :  :     :  : :::: :
crtE product  63DAAAVALELMHCASLVHDDLPAFDNADIRRGKPSLHKAYNEPLAVLAGDSL113

ispA product  81KTGALIRAAVRLGALSAG98
                :::::  ::   ::  ::
crtE product  70KTGALFIAATQMGAIAAG87
```

Fig. 5. **Comparison of amino acid sequences of the *ispA* product, the *crtE* product, and the FPP synthases.** The deduced amino acid sequences of the *ispA* product, the *crtE* product; and the FPP synthases are from Fig. 3, Refs. *29*, *6*, and *8*, respectively. Identical residues are marked with colons.

peptide directed by the *ispA* gene was studied by the maxicell technique. Strain CSR603 harboring pR10 and pR11 carrying the whole 1,452 bp insert gave a [^{35}S]methionine-labeled protein band with a molecular weight of approximately 36,000 in addition to the 30 kDa band of β-lactamase encoded by pUC vectors (Fig. 4). A faint band corresponding to the molecular weight of 36,000 was also observed with plasmids pR30 and pR31, which consisted of 1,200 bp of the chromosomal insert (nucleotides 253-1,452 in Fig. 3) and thus lacked a part of ORF-1 and the putative promoter sequences located upstream of ORF-1. These results show that the protein with the apparent molecular weight of 36,000 corresponds to the *ispA* product encoded by ORF-2 whose deduced molecular weight is 32,158. Since the 36,000 protein was also directed by pR31 in which the direction of ORF-2 is opposite to that of the *lac* promoter, there should be weak promoter-like sequences after nucleotide 253.

Although analyses by 13 and 15% SDS-polyacrylamide gel were performed in order to detect the product encoded by ORF-1, no significant band with a molecular weight of 8,000-10,000 was found.

Amino Acid Sequence Homology—The deduced amino acid sequence of the ORF-2-encoded protein was compared with those of the FPP synthases from eucaryotes (*6-9*). Conservation of a fragment of the *ispA* product (residues 215 to 237) among FPP synthases was revealed (Fig. 5a). In addition, the amino acid sequence of the *crtE* gene product of a photosynthetic bacterium *Rhodobacter capsulatus* (*29*), which was considered to be responsible for the conversion of prephytoene diphosphate to phytoene (*30*), was found to have homology with the sequence in Fig. 5a: 12 (52%), 11 (48%), and 10 (43%) amino acids in a 23-residue segment of the *ispA* product were identical to those in the corresponding segment of FPP synthase of rat and yeast, and the *crtE* product. The sequence of human FPP synthase is identical to that of the rat enzyme in this region. The sequence of the *ispA* product revealed two other regions of similarity to that of the *crtE* product. The first region (residues 67 to 117) exhibited 55% identity over 51 residues, and the second region (residues 181 to 198) exhibited 61% identity over 18 residues (Fig. 5b).

DISCUSSION

The precise location of the *ispA* gene, which had been identified as the defective gene of the mutant strain containing temperature-sensitive FPP synthase, on the *E. coli* chromosome was determined through the success in this work in cloning the *ispA* gene. It was shown that ORF-2 located 18 kb counterclockwise from the *lon* gene is the *ispA* gene. The multicopy plasmid containing ORF-2 under the control of the *lac* promoter not only complemented the defect of FPP synthase activity in the *ispA* strain, but also enabled the cells to overproduce the FPP synthase activity to the same extent irrespective of the *ispA* locus on the chromosome of the host. These facts suggest that the *ispA* gene is the structural gene of the FPP synthase.

However, no amino acid sequence similarity was found between the putative substrate binding site (*6, 31*) in the eucaryotic FPP synthases (residues 173-203 in the rat enzyme) and the *ispA* product, though there was sequence homology in the region of 23 residues near their C-termini (residues 215-237 in the *ispA* product and 235-257 in the rat enzyme). This may indicate that the region of homology near the C-terminus is important in the enzyme function. The *crtE* product, which shows striking homology with the *ispA* product, is considered to be responsible for synthesis of phytoene from prephytoene diphosphate (*30*). From the standpoint of the reaction mechanism, this reaction is similar to the reaction of FPP synthase with respect to the formation of the carbonium cation intermediate by the cleavage of the C-O bond of the substrate (*32, 33*). To elucidate the relation between the sequence homology and the similarity of reaction mechanism, it will be necessary to purify and characterize the proteins encoded by these genes.

ORF-1 and ORF-2 seem to be cotranscribed, taking into account the characteristics of the plasmids with serial deletions and the locations of the promoter-like sequence and the putative ribosome-binding site. Although the maxicell experiment showed that the product of ORF-2 gave a band of molecular weight of 36,000, the product encoded by ORF-1 was not detected. Since ORF-1 does not contain the methionine codon ATG except for the initiation codon, the product encoded by ORF-1 might not be detected by labeling with [^{35}S]methionine, the procedure used in

	ORF-1	ORF-2		ORF-1	ORF-2		ORF-1	ORF-2		ORF-1	ORF-2
TTT Phe	2	4	TCT Ser	2	1	TAT Tyr	0	3	TGT Cys	0	1
TTC Phe	1	4	TCC Ser	0	0	TAC Tyr	0	3	TGC Cys	0	3
TTA Leu	1	9	TCA Ser	0	2	TAA End	1	1	TGA End	0	0
TTG Leu	0	3	TCG Ser	0	5	TAG End	0	0	TGG Trp	0	0
CTT Leu	0	6	CCT Pro	1	3	CAT His	0	5	CGT Arg	2	8
CTC Leu	0	4	CCC Pro	1	3	CAC His	0	2	CGC Arg	2	6
CTA Leu	1	1	CCA Pro	0	2	CAA Gln	3	5	CGA Arg	0	1
CTG Leu	9	16	CCG Pro	3	4	CAG Gln	5	13	CGG Arg	0	2
ATT Ile	2	8	ACT Thr	0	2	AAT Asn	3	1	AGT Ser	1	2
ATC Ile	0	7	ACC Thr	2	6	AAC Asn	1	4	AGC Ser	2	6
ATA Ile	0	0	ACA Thr	1	1	AAA Lys	2	8	AGA Arg	0	1
ATG Met	1	7	ACG Thr	0	3	AAG Lys	2	5	AGG Arg	0	0
GTT Val	0	6	GCT Ala	0	6	GAT Asp	0	15	GGT Gly	0	11
GTC Val	0	1	GCC Ala	4	17	GAC Asp	4	13	GGC Gly	2	7
GTA Val	2	2	GCA Ala	1	11	GAA Glu	7	8	GGA Gly	0	5
GTG Val	1	5	GCG Ala	3	8	GAG Glu	5	4	GGG Gly	1	0

Fig. 6. **Codon usage in ORF-1 and ORF-2.**

this work. The pattern of codon usage in both ORF-1 and ORF-2 is rather typical for *E. coli* genes (Fig. 6) (*34*). Thus, the use of rare codons is unlikely to limit translation of these polypeptides. Further work is needed to clarify whether the region of ORF-1 is actually the coding region or the untranslated regulatory region. Concerning the above question, it is interesting that pR31, which lacks the promoter region upstream of ORF-1, directs weak expression of the product encoded by ORF-2, indicating the existence of a promoter-like sequence in pR31. Actually there is a sequence which resembles heat-shock promoter (*28*) in the region of ORF-1 (Fig. 3). Studies on regulation of *ispA* expression and its relation to the physiology of the cell are in progress.

Note Added in Proof—The sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank Nucleotide Sequence Databases under the accession number D00694.

We thank Dr. Y. Kohara, National Institute of Genetics, for providing the phage clones. We thank Dr. K. Izui, Faculty of Science, Kyoto University, and Dr N. Fujita, National Institute of Genetics, for valuable discussions. Thanks are also due to Dr. K. Okada, Faculty of General Education, Gifu University, for his interest and encouragement.

REFERENCES

1. Spurgeon, S.D. & Porter, J.W. (1981) in *Biosynthesis of Isoprenoid Compounds* (Porter, J.W. & Spurgeon, S.D., eds.) pp. 1-46, John Wiley and Sons, New York
2. Fujisaki, S., Nishino, T., Katsuki, H., Hara, H., Nishimura, Y., & Hirota, Y. (1989) *J. Bacteriol.* **171**, 5654-5658
3. Sherman, M.M., Petersen, L.A., & Poulter, C.D. (1989) *J. Bacteriol.* **171**, 3619-3628
4. Goldstein, J.L. & Brown, M.S. (1990) *Nature* **343**, 425-430
5. Bruenger, E. & Rilling, H.C. (1988) *Anal. Biochem.* **173**, 321-327
6. Clerke, C.F., Tanaka, R.D., Svenson, K., Wamsley, M., Fogelman, A.M., & Edwards, P.A. (1987) *Mol. Cell. Biol.* **7**, 3138-3146
7. Sheares, B.T., White, S.S., Molowa, D.T., Chan, K., Ding, V.D.-H., Kroon, P.A., Bostedor, R.G., & Karkas, J.D. (1989) *Biochemistry* **28**, 8129-8135
8. Anderson, M.S., Yarger, J.D., Burck, C.L., & Poulter, C.D. (1989) *J. Biol. Chem.* **264**, 19176-19184
9. Wilkin, D.J., Kutsunai, S.Y., & Edwards, P.A. (1990) *J. Biol. Chem.* **265**, 4607-4614
10. Rosser, D.S.E., Ashby, M.N., Ellis, J.L., & Edwards, P.A. (1989) *J. Biol. Chem.* **264**, 12653-12656
11. Dawson, P.A., Hofmann, S.L., van der Westhuyzen, D.R., Südhof, T.C., Brown, M.S., & Goldstein, J.L. (1988) *J. Biol. Chem.* **263**, 3372-3379
12. Osborn, T.F., Gil, G., Goldstein, J.L., & Brown, M.S. (1988) *J. Biol. Chem.* **263**, 3380-3387
13. Smith, J.R., Osborn, T.F., Brown, M.S., Goldstein, J.L., & Gil, G. (1988) *J. Biol. Chem.* **263**, 18480-18487
14. Kandutsch, A.A., Paulus, H., Levin, E., & Bloch, K. (1964) *J. Biol. Chem.* **239**, 2507-2515
15. Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene* **33**, 103-119
16. Vieira, J. & Messing, J. (1987) *Methods Enzymol.* **153**, 3-11
17. Sancar, A., Hack, A.M., & Rupp, W.D. (1979) *J. Bacteriol.* **137**, 692-693
18. Kohara, Y., Akiyama, K., & Isono, K. (1987) *Cell* **50**, 495-508
19. Miller, J.H. (1972) in *Experiments in Molecular Genetics* pp. 431-435, Cold Spring Harbor Lab., New York
20. Fujisaki, S., Nishino, T., & Katsuki, H. (1986) *J. Biochem.* **99**, 1327-1337
21. Henikoff, S. (1984) *Gene* **28**, 351-359
22. Sanger, F., Nicklen, S., & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467
23. Sancar, A., Wharton, R.P., Seltzer, S., Kacinski, B.M., Clarke, N.D., & Rupp, W.D. (1981) *J. Mol. Biol.* **148**, 45-62
24. Worcel, A. & Burgi, E. (1974) *J. Mol. Biol.* **82**, 91-105
25. Laemmli, U.K. (1970) *Nature* **227**, 680-685
26. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1342-1346
27. Hawley, D.K. & McClure, W.R. (1983) *Nucleic Acids Res.* **11**, 2237-2255
28. Craig, E.A. (1985) *CRC Crit. Rev. Biochem.* **18**, 239-280
29. Armstrong, G.A., Alberti, M., Leach, F., & Hearst, J.E. (1989) *Mol. Gen. Genet.* **216**, 254-268
30. Armstrong, G.A., Schmidt, A., Sandmann, G., & Hearst, J.E. (1990) *J. Biol. Chem.* **265**, 8329-8338
31. Brems, D.N., Bruenger, E., & Rilling, H.C. (1981) *Biochemistry* **20**, 3711-3718
32. Poulter, C.D. & Rilling, H.C. (1981) in *Biosynthesis of Isoprenoid Compounds* (Porter, J.W. & Spurgeon, S.D., eds.) pp. 161-224, John Wiley and Sons, New York
33. Poulter, C.D. & Hughes, J.M. (1977) *J. Am. Chem. Soc.* **99**, 3830-3837
34. Konigsberg, W. & Godson, G.N. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 687-691