Available online at www.sciencedirect.com
SCIENCE DIRECT®
Metabolic Engineering 7 (2005) 18–26
METABOLIC ENGINEERING
www.elsevier.com/locate/ymben

# Directed evolution of *Escherichia coli* farnesyl diphosphate synthase (IspA) reveals novel structural determinants of chain length specificity

Pyung Cheon Lee, Ralf Petri, Benjamin N. Mijts, Kevin T. Watts, Claudia Schmidt-Dannert*

*Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108, USA*

Received 26 February 2004; accepted 26 May 2004
Available online 7 January 2005

**Abstract**

Directed evolution of farnesyl diphosphate (FPP, $C_{15}$) synthase (IspA) of *Escherichia coli* was carried out by error-prone PCR with a color complementation screen utilizing C40 carotenoid pathway enzymes. This allowed IspA mutants with enhanced production of the C40 carotenoid precursor geranylgeranyl diphosphate (GGPP, $C_{20}$) to be readily identified. Analysis of these mutants was carried out in order to better understand the mechanisms of product chain length specificity in this enzyme. The 12 evolved clones having enhanced $C_{20}$ GGPP production have characteristic mutations in the conserved regions of prenyl diphosphate synthases (designated regions I through VII). Some of these mutations (I76T, Y79S, Y79H, C75Y, H83Y, and H83Q) are found near or before the conserved first aspartate rich motif (FARM), which is involved in the mechanism for chain elongation reaction of all prenyl synthases. Molecular modeling suggested a mechanism for chain length determination for these mutations including substitutions at the 1st and 9th amino acids upstream of the FARM that have not been reported previously. In addition, a mutation on a helix adjacent to the FARM within the substrate-binding pocket (D115G) suggests a novel mechanism for chain length determination. One mutant IspA clone carries a mutation of C155G at the 2nd amino acid upstream of conserved region IV (GQxxDL), which was recently found to be an important region controlling the chain elongation of a Type III GGPP synthase. One IspA clone carries mutations (T234A and T249I) near the conserved second aspartate rich motif (SARM). As a verification of the in vivo activity of the mutant clones (represented as $C_{40}$ carotenoid formation), we confirmed the product distribution of wild-type and mutant IspA using an in vitro assay.


*Keywords:* Prenyl diphosphate synthase; FPP synthase; In vitro evolution; Carotenoid

## 1. Introduction

Prenyl diphosphate synthases are classified as a *cis*- and *trans*-type according to the stereochemistry of double bonds present in the final products. *Trans*-prenyl diphosphate synthases can be further classified into short ($C_{15}$–$C_{25}$), medium ($C_{30}$–$C_{35}$), and long-chain ($C_{40}$–$C_{50}$) synthases based on the chain lengths of the final product. Over the past decades many types of *trans*-diphosphate synthases, especially farnesyl diphosphate (FPP, $C_{15}$) and geranylgeranyl diphosphate (GGPP, $C_{20}$) synthases from different sources, have been studied structurally and enzymatically. The short chain prenyl diphosphate synthases catalyze the sequential condensation reactions of isopentenyl diphosphate (IPP, $C_5$) with different lengths of allylic diphosphates to produce linear isoprenoid derivatives: geranyl diphosphate (GPP, $C_{10}$), farnesyl diphosphate (FPP, $C_{15}$), geranylgeranyl diphosphate (GGPP, $C_{20}$) or farnesylgeranyl diphosphate (FGPP, $C_{25}$). The isoprenoids of different chain lengths are then transformed or modified

*Corresponding author.
*E-mail address:* schmi232@tc.umn.edu (C. Schmidt-Dannert).

1096-7176/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2004.05.003

into bioactive compounds such as steroids, quinones, prenylated proteins, lipid components and carotenoids (for a review see: Liang et al., 2002).

Even though chain lengths of the final products of these enzymes vary from $C_{15}$ to $C_{25}$, short chain *trans*-prenyl diphosphate synthases have similar tertiary structures and precisely control a definite chain length for their final product. The alignment of deduced amino acid sequences of these enzymes shows that they have two highly conserved aspartate-rich motifs (DDxxD): the first aspartate-rich motif (FARM) and second aspartate-rich motif (SARM) as well as other five conserved regions (Wang and Ohnuma, 1999). The aspartate-rich motifs are known to be involved in substrate binding and catalysis with the cofactor metal ion $Mg^{2+}$. Based on known protein structures within this family, a number of short-chain prenyl-diphosphate synthases have been subjected to mutational analysis which has indicated that amino acids before the FARM, especially the 11th, 8th and 5th positions, are of importance in determining of the chain lengths of isoprenoids (Kharel and Koyama, 2003). Mutations of this kind have been utilized to specifically alter the chain-lengths of the products of prenyl-diphosphate enzyme catalysis. Chemical random mutagenesis of the *Bacillus stearothermophilus* FPP synthase indicated a number of sites important in chain length specificity, particularly a site upstream of the FARM (Ohnuma et al., 1996).

Carotenoids are one of the major derivatives of isoprenoids found in nature. Generally, the conjugated double bonds of the hydrocarbon backbone confer their unique color properties and the number of backbone $C_5$ isoprenoid units classifies carotenoids into $C_{30}$, $C_{40}$, and $C_{50}$ subfamilies (Lee et al., 2003). In vivo synthesis of $C_{40}$ carotenoids proceeds by sequential elongation, condensation and desaturation reactions (Fig. 1): (1) sequential elongations of $C_5$ DMADP (dimethylallyl diphosphate) with $C_5$ IPP (isopentyl diphosphate) to produce $C_{15}$ FPP by FPP synthase (*ispA* in *Escherichia coli*), (2) further elongation of $C_{15}$ FPP with $C_5$ IPP by GGPP synthase (*crtE*) to produce $C_{20}$ GGPP, (3) the head to head condensation of $2 \times C_{20}$ GGPP by phytoene synthase (*crtB*) to produce $C_{40}$ phytoene and (4) the sequential desaturations involved in phytoene desaturase (*crtI*) to produce the colored $C_{40}$ carotenoid lycopene.

To better understand the mechanisms of product chain length specificity of *E. coli* FPP synthase (IspA), we attempted to direct the evolution of *E. coli* FPP synthase to GGPP synthase by error-prone PCR. To monitor the altered product profiles of the evolved FPP synthases in vivo, we used a visual screening method where the products of the evolved IspA catalysis feed suitable substrates into an engineered $C_{40}$ carotenoid biosynthetic pathway. In this screening system, positively evolved FPP synthase with an increased product chain-length ($C_{15}$–$C_{20}$) generate colored, carotenoid producing colonies that are selected. As this random mutagenesis approach does not rely on the known enzyme mechanism, it may suggest novel structural regions that are important in enzyme function that have not been identified in previous studies. Molecular modeling studies were utilized to provide insight into specific structural mechanisms of chain-length determination.

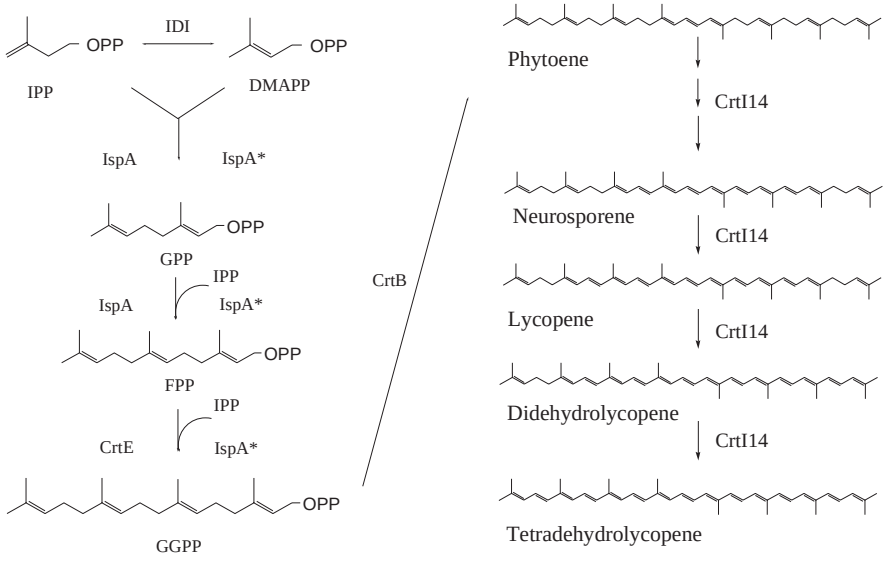


Fig. 1. Isoprenoid and carotenoid biosynthetic pathways. The enzymes involved in the above pathway are isopentenyl diphosphate (IPP) isomerase (IDI); geranyl diphosphate synthase (IspA); geranylgeranyl diphosphate (GGPP) synthase (CrtE); phytoene synthase (CrtB); in vitro evolved phytoene desaturase (CrtI14).

## 2. Experimental procedures

### 2.1. Genes, plasmids, and culture conditions

The *ispA* gene encoding farnesyl diphosphate (FPP) synthase (Genbank Acc. No. D00694) was amplified from the genomic DNA of *E. coli* JM109 using a 5′ primer containing at its 5′ end a *Xba*l site followed by an optimized Shine-Dalgarno sequence (underlined) and a start codon (bold) (5′- AGGAGGATTACAAA**ATG**-3′) and a 3′ primer containing at its 5′ end a *Eco*RI site. The PCR product was then digested with restriction enzymes and cloned into the corresponding sites of plasmid pUCmod to facilitate constitutive expression from a modified *lac*-promoter (Lee et al., 2003). For assembly of *crtB* and *crtI14*, genes encoding phytoene synthase (*crtB*) or mutant phytoene desaturase (*crtI14*) were subcloned from pUCmod into the *Sal*I site (*crtB*) and *Hin*dIII (*crtI14*) sites of pACmod by amplification of the genes together with the modified constitutive *lac*-promoter, using primers that introduce the corresponding restriction enzyme sites at both ends, to give pAC_crtB_crtI14 where *crtB* and *crtI14* have the same orientation as the disrupted tetracycline resistance gene (Table 1). For carotenoid production, recombinant *E. coli* JM109 were cultivated for 48 hr in the dark at 28 °C in Luria-Bertani (LB) medium (100 or 500 ml medium) supplemented with the appropriate selective antibiotics chloramphenicol (50 mg/ml) and/or carbenicillin (100 mg/ml).

### 2.2. Error-prone PCR mutagenesis

The *ispA* gene together with the modified constitutive *lac*-promoter was amplified under optimized mutagenic PCR conditions with PCR primers (5′-CCGACTGGAAAGCGGG-3′ and 5′-ACAAGCCCGTCAGGG-3′) flanking the gene and promoter. The PCR reaction mix consisted of 1 × Promega $Mg^{2+}$ free thermophilic buffer (Promega, Madison, WI), 10 ng/ml template plasmid, 1 μM each primer, 2.5 U *Taq* DNA polymerase, 2.5 mM $MgCl_2$, 25 μM $MnCl_2$, 0.3 mM dGTP/dATP and 0.25 mM dTTP/dCTP. PCR was carried out with a program of 95 °C for 4 min followed by 32 cycles of 94 °C for 1 min, 50 °C for 40 s and 72 °C for 1 min and finally 72 °C for 7 min. The 0.9-kb PCR product was purified using a QIAquick gel extraction kit (Qiagen, Valencia, CA) followed by digestion with the restriction enzymes *Xba*I and *Eco*RI. The DNA fragments were ligated into the corresponding sites of the pUCmod vector (Lee et al., 2003) and electrotransformed into competent *E. coli* JM109 [pAC_crtB_crtI14]. Transformants were plated on LB plates supplemented with 100 μg/ml carbenicillin and 50 μg/ml chloramphenicol. After 18 h of incubation at 30 °C in the dark, colonies were replicated using a nitrocellulose membrane and

Table 1
Genes and plasmids used in this study
(A)

| Gene | Enzyme | Typical Reaction | Acc. # or Ref. |
|---|---|---|---|
| *crtE* | GGPP synthase | Head-to-head condensation of IPP + FPP | D90087 |
| *crtB* | Phytoene synthase | Head-to-head condensation of 2 GGPP | D90087 |
| *crtI* | Phytoene desaturase | Introduction of 4 desaturations in phytoene | D90087 |
| *crtI14* | Mutant phytoene desaturase | Introduction of 6 desaturations in phytoene | Schmidt-Dannert et al. (2000) |
| *ispA* | FPP synthase | Sequential head-to-head condensations of IPP + DMADP to GPP and IPP + GPP to FPP | AB025791 |

(B)

| Plasmid | Properties | Ref. |
|---|---|---|
| pUCmod | Constitutive expression vector modified from pUC19, $Ap^R$ | Schmidt-Dannert et al. (2000) |
| pACmod | Cloning vector modified from pACYC184 for constitutive expression of *crt*-genes subcloned (including their individual promoters) from pUCmod, $Cm^R$ | Schmidt-Dannert et al. (2000) |
| pUC_crtE | pUCmod expressing *crt*E | Schmidt-Dannert et al. (2000) |
| pUC_crtB | pUCmod expressing *crt*B | Schmidt-Dannert et al. (2000) |
| pUC_ispA | pUCmod expressing FPP synthase (IspA) | This study |
| pUC_ispA_M1-27 series | pUCmod expressing mutant IspA_M1-27 | This study |
| pUC_ispA_W_$His_6$ | pUCmod expressing His-tagged wild-type IspA | This study |
| pUC_ispA_M4_$His_6$ | pUCmod expressing His-tagged mutant IspA_M4 | This study |
| pUC_ispA_M25_$His_6$ | pUCmod expressing His-tagged mutant IspA_M25 | This study |
| pUC_ispA_M26_$His_6$ | pUCmod expressing His-tagged mutant IspA_M26 | This study |
| pAC_crtB_crtI14 | pACmod expressing *crt*B and *crt*I14 | This study |
| pAC_crtE_crtB_crtI14 | pACmod expressing *crt*E, *crt*B and *crt*I14 | Lee et al. (2003) |

transferred onto fresh LB plates containing the same antibiotics. Colonies were screened visually for color variants after additional 24 h incubation at room temperature. Overnight cultures (3 ml LB) were grown from selected colonies for analysis of carotenoid synthesis. Mutations of IspA sequence were confirmed by DNA sequencing.

### 2.3. Purification of wild-type and mutant FPP synthases

N-terminal 6 × His-tags were added to wild-type and mutant IspA clones. *E. coli* cells [pUC_ispA_W_$His_6$], [pUC_ispA_M4_$His_6$], [pUC_ispA_M25_$His_6$], [pUC_ispA_M26_$His_6$] or [pUCmod] were grown in 100 ml of LB medium with 100 μg/ml carbenicillin at an $OD_{600}$ of 1.8. The cells were harvested by centrifugation, washed with 50 mM phosphate buffer (pH 8.0) containing 50 mM NaCl. 10 units of DNase was added into 5 ml of the cell suspension and incubated at room temperature for 15 min. Cells were then sonicated on ice and cell debris was removed by centrifugation at 13,000*g* for 30 min at 4 °C. The supernatant was loaded onto a pre-equilibrated Ni-agarose column (5 ml; Qiagen) and washed with a washing buffer A (pH 8.0, 50 mM phosphate, 300 mM NaCl) and buffer B (pH 8.0, 50 mM phosphate, 300 mM NaCl, 20 mM imidazole). The proteins were eluted in buffer C (pH 8.0, 50 mM phosphate, 300 mM NaCl, 300 mM imidazole). Fractions containing the recombinant protein (5 ml) were immediately pooled and were concentrated and then diluted repeatedly with 50 mM Tris pH 8.0 in a Centricon YM-10 centrifugal filter (Millipore) and finally concentrated to 15 mg/ml. Proteins were analyzed by SDS-PAGE with Coomassie Brilliant Blue staining.

### 2.4. Prenyltransferase assay

Prenyltransferase activity was assayed in a 200 μl reaction volume of 50 mM Tris/HCl Buffer (pH 8.0), 5 mM 2-mercaptoethanol, 5 mM $MgCl_2$, 50 μM allylic substrates (DMADP, GPP and FPP), 50 μM IPP and 50 μg/ml of the purified recombinant enzyme. This reaction mixture was incubated at 30 °C for 2 h or 10 h. After filtration, the reaction mixtures were directly applied to a Zorbax RX-C18 column (4.6 × 150 mm, 5 μm; Agilent Technologies) without further purification steps and eluted under gradient conditions with acetonitrile and 25 mM $NH_4HCO_3$ at a flow rate of 1 ml/min (Zhang and Poulter, 1993) using an Agilent 1100 HPLC system equipped with a photodiode array detector. A linear gradient was used for the separation of phosphorylated isoprenoids (0–5 min, 25 mM $NH_4HCO_3$; 5–30 min, 0–100% acetonitrile) and the eluent was monitored at 214 nm.

The reaction samples were further analyzed on a Finnigan LCQ mass spectrometer (Thermo Finnigan) equipped with a Zorbax RX-C18 column (4.6 × 250 mm, 5 μm; Agilent Technologies). Molecular ions $(M–H^+)^-$ of phosphorylated isoprenoids were detected on negative ion electron spray ionization (ESI) mode. Retention times and fragmentation patterns were compared with those of authentic samples of FPP and GGPP (Sigma-Aldrich).

### 2.5. Analysis of carotenoids

Wet cells from a 100 ml culture (~250 mg) were repeatedly extracted at 4 °C with a total volume of 30 ml or 40 ml acetone until all visible pigments were extracted. After centrifugation (4 °C, 6000 rpm), the colored supernatants were further purified as described (Lee et al., 2003). A 5–10 μl aliquot of the crude extracts were analyzed on Whatman silica gel 60 plates (4.5 μm particle size, 200 μm thickness) using hexane and chloroform (100:5) or applied to the Zorbax SB-C18 column and eluted under isocratic conditions with a solvent system (80% acetonitrile, 15% methanol and 5% 2-propanol) at a flow rate of 1 ml/min using an Agilent 1100 HPLC system equipped with an photodiode array detector. Carotenoids were identified by a combination of HPLC retention times, absorption spectra and mass spectra. Mass spectra were monitored on a LCQ mass spectrometer equipped with atmosphere pressure chemical ionization (APCI) interface (Lee et al., 2003).

### 2.6. Alignment and modeling

#### 2.6.1. Alignment and modeling

The recently solved structure of IspA (Hosfield et al., 2004) was obtained from the PDB database (PDB database entry 1RQJ) and used for mapping of the mutations using the SWISS-PDB software (Schwede et al., 2003). Energy minimization of the mutations were performed in vacuo using the GROMOS96 43B1 parameter set without reaction field as implemented in the software (van Gunsteren et al., 1996).

## 3. Results and discussion

### 3.1. Validation of screening method

In order to screen and select mutant FPP synthases producing GGPP, we used a screening method based on color complementation adapted from Ohnuma et al. (1996). Theoretically, if $C_{20}$ GGPP is sufficiently supplied (by CrtE in the native pathway), CrtB and CrtI14 will catalyze the sequential condensation of $2 \times C_{20}$ GGPP to phytoene and the desaturation of phytoene to produce colored carotenoids (Fig. 1). The in vitro evolved carotenoid desaturase CrtI14 was selected

for the screening procedure rather than wild-type CrtI as the red carotenoid tetradehydrolycopene was more readily observed on a white filter background compared to yellow/orange lycopene. Although expression of CrtB and CrtI14 in *E. coli* cells [pAC_crtB_crtI14] produced no carotenoids, trace amounts of carotenoids in the form of a very weak red phenotype were observed when pUC_IspA was co-expressed in this strain. This suggests that very small amounts of $C_{20}$ GGPP are being produced in addition to the main product $C_{15}$ FPP. Formation of very small amounts of non-specific products of smaller or greater chain lengths have been observed during in vitro assays of some prenyl phosphate synthase enzymes (Ohnuma et al., 1992; Pan et al., 2000). These trace levels of carotenoid production in the presence of wild-type IspA did not interfere with the assay procedure as only mutants with significantly enhanced GGPP production levels produced a distinguishable colored phenotype.

### 3.2. Directed evolution of farnesyl diphosphate (FPP) synthase by random PCR mutagenesis

The *ispA* gene encoding FPP synthase was amplified under the mutagenic PCR conditions and the mutated *ispA* gene library was transformed as pUC_ispA into *E. coli* [pAC_crtB_crtI14]. Among ~5000 clones visually screened, 28 distinguishable clones with a strong red phenotype were isolated and the plasmid insert DNA sequenced. Based on the deduced amino acid sequence of the isolated 28 mutant clones (designated as IspA_M1 to M28), 12 IspA mutants were identified to have a unique mutation in their amino acid sequence (Table 2): two single point mutants (S140T and H83Q), nine double mutants (D2G, C75Y, I76T, Y79H, H83Y, H83Q, D115G, S140T, C155G, and Q276R) and one multi-point mutant (H57R, Y79S, S140 T, T234A, T249I).

### 3.3. In vivo activity of GGPP synthase of the IspA mutants

The availability of precursors is strongly correlated to the production level of final products. In particular, increasing the level of the precursor pool enhances carotenoid production (Lee and Schmidt-Dannert, 2002). Therefore, we examined indirectly the in vivo GGPP synthase activity of the mutant IspA clones by measuring carotenoid production based on the assumption of the conversion of all $C_{20}$ GGPP into carotenoids. Plasmid DNA from wild-type IspA and each mutant was purified, retransformed into *E. coli* cells [pAC_CrtB_CrtI14] and the carotenoid production analyzed. Organic extracts of the 12 mutant IspA clones were analyzed by HPLC and LC-mass spectrometry with a positive APCI interface, all mutant clones produced lycopene ($\lambda_{max}$: 449, 475, 507 nm; $[M+H^+]^+ = 537.4$) and tetradehydrolycopene ($\lambda_{max}$: 460, 493, 516 nm; $[M+H^+]^+ = 533.4$) at a detectable amount except the wild-type IspA clone (Fig. 2) and mutant M25. Based on these results it appears that mutant M25 was a false positive that was incorrectly identified in screening. The production of carotenoids in the remaining selected mutants at levels similar to those observed in the presence of the wild-type CrtE gene strongly indicates that these IspA mutants produced significant amounts of GGPP ($C_{20}$) (Fig. 3).

### 3.4. In vitro activity of wild-type and mutant FPP synthase

As carotenoid production gives no indication of actual product distribution of each mutant enzyme, we

Table 2
Mutant FPP synthases generated by random mutagenesis

| Mutant | Base mutation[a] | Amino acid mutation | Carotenoid production[b] |
|---|---|---|---|
| IspA M1 | A170G, A236C, T418A, (T531C), A700G, C746T | H57R, Y79S, S140T, T234A, T249I | ++++ |
| IspA M3 | T235C, T418A | Y79H, S140T | ++++ |
| IspA M4 | G224A, T418A | C75Y, S140T | ++++ |
| IspA M8 | (T147C), (T285C), (C327T), A344G, A827G | D115G, Q276R | ++ |
| IspA M9 | (T306C), A344G, (T372A), (A396G), T418A | D115G, S140T | +++ |
| IspA M10 | C247T, (T285C), T418A | H83Y, S140T | ++ |
| IspA M12 | (T159C), A344G, T418A, (T583C) | D115G, S140T | +++ |
| IspA M21 | T227C, T418A | I76T, S140T | ++++ |
| IspA M22 | A5G, T463G, (T576C) | D2G, C155G | ++++ |
| IspA M25 | T418A | S140T | – |
| IspA M26 | T249G | H83Q | +++ |
| IspA M27 | T249A, T418A | H83Q, S140T | ++++ |

[a]Parentheses represent silent mutation.
[b]++++; highest and –; background.

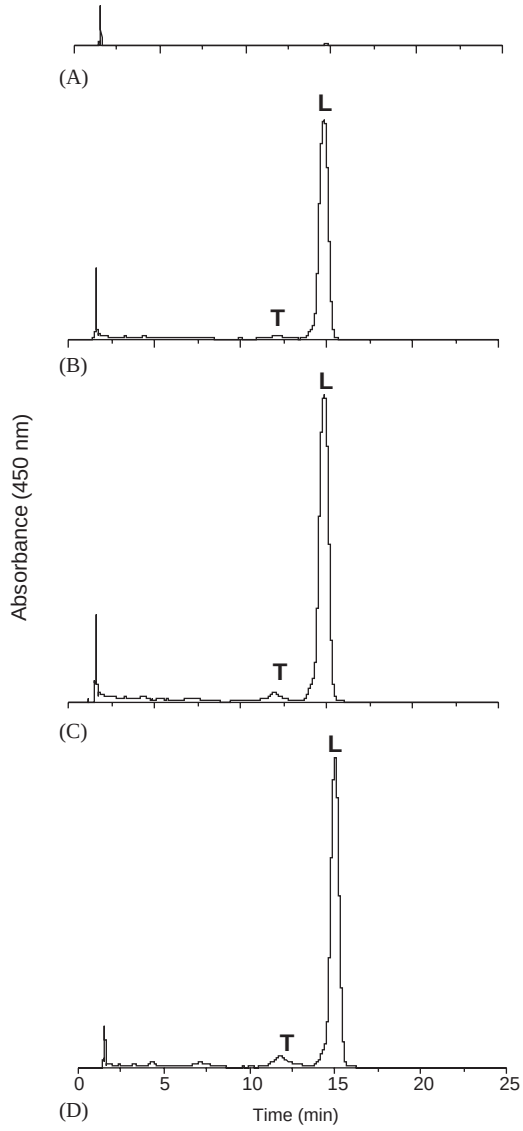


Fig. 2. HPLC analysis of carotenoid extracts of *E. coli* transformants expressing wild-type and selected in vitro evolved FPP synthases. HPLC analysis of the carotenoid extracts of *E. coli* transformants having (A) pAC_crtB_crtI14/pUC_ispA, (B) pAC_crtB_crtI14/pUC_ispA_M1, (C) pAC_crtB_crtI14/pUC_ispA_M4, and (D) pAC_crtB_crtI14/pUC_ispA_M26. Carotenoid peaks identified by retention time and spectra are indicated as L (lycopene) and T (tetradehydrolycopene).

examined the product distribution of wild-type IspA mutants IspA_M4, IspA_M25 and IspA_M26 using an in vitro assay. These enzymes were fused with 6× His-tags at the N-terminus and purified by Ni-affinity chromatography. Purified wild-type IspA in a reaction containing 50 μM DMAPP and 50 μM IPP as substrates, produced only $C_{15}$ FPP as detected by C18 reverse-phase HPLC (Fig. 4A). Negative ESI LC-MS clearly showed the molecular ion of $C_{15}$ FPP ($[M–H^+]^- = 381.1$) without any corresponding ions of shorter and/or longer chain length products (i.e., $C_{20}$ GGPP). Similarly, the purified mutant IspA_M4 produced in vitro only $C_{20}$ GGPP ($[M–H^+]^- = 449.1$) as a final product without any intermediates (i.e., $C_{15}$ FPP) and longer chain length products (i.e., $C_{25}$ FGPP) (Fig. 4B). Mutant IspA_M25 generated a product profile identical to that of wild-type IspA confirming the classification of this as a false positive clone. Mutant IspA_M26 produced a mixed population of $C_{15}$ FPP and $C_{20}$ GGPP but based on the results above this appears to be sufficient for carotenoid production in vivo.

### 3.5. Sequence analysis and molecular modeling of IspA mutants

In order to better understand the mechanisms by which the observed amino acid substitutions influence product chain length, structural maps of IspA mutants were constructed based on the recently published X-ray structure of IspA (Hosfield et al., 2004). As inferred from the sequence alignment, IspA contains a two amino acid insertion (proline and alanine) in the FARM motif compared to other farnesyl transferases such as avian FPS. This insertion is a signature sequence that characterizes IspA as a true type II FPP synthase according to the chain-length determination model developed by Wang (Wang and Ohnuma, 1999).

The mutation S140T was found among all mutants except three (M8, M22, and M26) suggesting that this specific mutation was introduced at an early stage of error-prone PCR. However, mutant IspA_M25 carries the single mutation S140T and retransformation of this clone indicated that no significant levels of carotenoids were produced. Since the mutation S140T was found to have no effect on the product spectrum, mutants 3, 4, 9, 10, 12, 21, and 27 can be considered as single mutants and the observed effects can be attributed to single amino acid exchanges. Nevertheless, the obtained mutants share a certain pattern of mutations which occur repeatedly: Y79, C75, H83, and D115. These mutations altered the product spectrum most effectively and they are located in close vicinity to the FARM region. Of interest is the substitution of Tyr79 to smaller residues like histidine or serine in mutants IspA_M1 and IspA_M3. Tyr79 is the 5th amino acid upstream of the FARM region and various studies have demonstrated that this particular residue limits the final length of the product by blocking the product-binding channel and alters product specificity in related enzymes (Wang and Ohnuma, 1999). When this amino acid is replaced by a smaller one the binding channel can host a larger isoprenoid chain like GGPP and larger products are obtained. Beside the 5th position, mutations of residues 4, 8, and 11 positions upstream of the FARM region, respectively, were found to limit the product length. Most previous reports about chain-length determination have focused on the exchange of aromatic residues in the

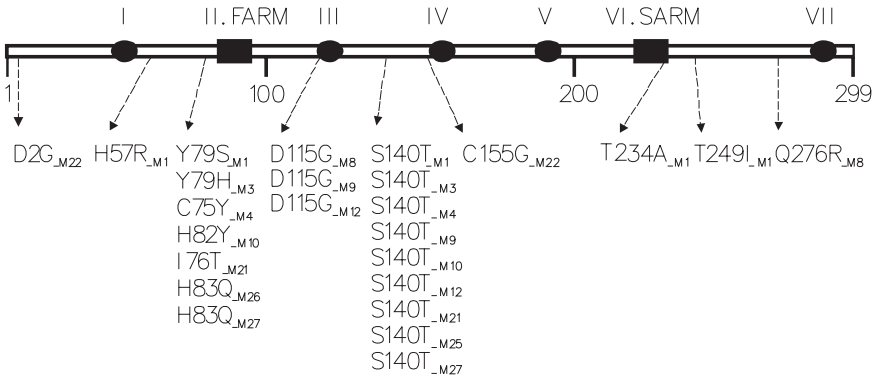


Fig. 3. Conserved regions I through VII of prenyl diphosphate synthases and mutated amino acids of the mutant IspA. The red boxes represent first aspartate-rich motif (FARM) in the conserved region II and second aspartate-rich motif (SARM) in the region VI of the enzyme and the circles indicate the other conserved regions (for details see Wang and Ohnuma, 1999).

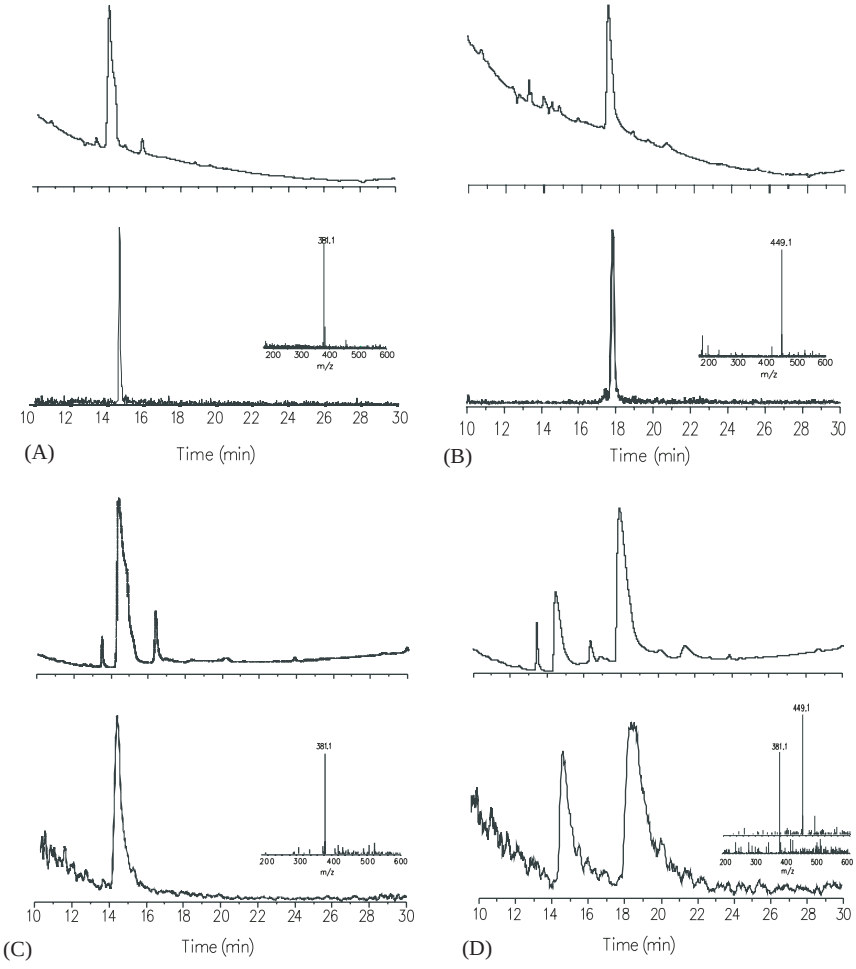


Fig. 4. HPLC and LC-mass chromatogram of final products from in vitro assay of purified wild-type IspA (A) and mutant IspA (B–D). HPLC chromatogram (upper panel) and selected negative ion chromatogram (lower panel) of the final products from in vitro assay of the purified (A) wild-type IspA and (B) mutant ispA_M4 (C) mutant ispA_M25 (D) mutant ispA_M 26: FPP $(M–H^+)^-$ at $m/e = 381.1$; GGPP $(M–H^+)^-$ at $m/e = 449.1$.

4th and 5th positions upstream of the FARM region (Wang and Ohnuma, 1999).

In our study, we found the first evidence that the residues at position 1 and 9 N-terminal of the FARM motif (His83 and Cys75, respectively) can influence the product spectrum in the same way, as these mutants yielded GGPP even in the presence of Tyr79. Intriguingly, the residue His83, which is located in Helix D one

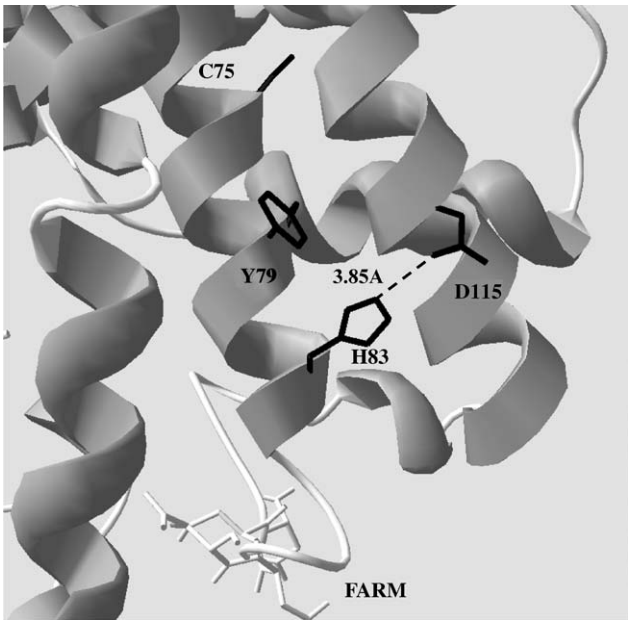


Fig. 5. Structural mapping of IspA. The three dimensional structure of the IspA catalytic region is shown with mutated residues discussed in the text highlighted.

turn below Tyr79, and the residue Cys75, which is located one turn above Tyr79, were found to be mutated to bulky tyrosines (Fig. 5, mutants 10 and 4, respectively). This observation was unexpected insofar as the substrate-binding channel is already limited in space by Tyr79 and the occupation with additional aromatic residues was expected to limit the product length rather than enhance it. However, the orientation of these large aromatic side chains in close vicinity to each other indicates a strong interaction of the benzyl rings, which might lead to a bending of the side chain of Tyr79 out of the way of the elongating isoprenoid chain (Fig. 5). It cannot be excluded that structural rearrangement provoked by the large side chains cause the observed effect. However, the same effect was observed in mutant IspA_M26 (H83Q) where the basic histidine was mutated to glutamine enabling a hydrophilic interaction between Y79 and Q83.

In addition to mutations near the FARM region, mutation D115G within conserved region III was also prevalent, occurring in mutants IspA_M8, M9 and M12. This particular residue is located opposite of Y79 in helix E that runs parallel to helix D (Fig. 5) and forms an adjacent side of the substrate-binding pocket. The replacement of an acidic aspartate residue by a small glycine might allow the sidechain of Y79 to rearrange and to open the way for further chain elongation. Mutations within conserved region III, on the substrate binding helix adjacent to the FARM region therefore provide a novel mechanism influencing prenyl phosphate synthase chain length specificity. However, the spatial vicinity of the basic histidine residue H83 and the acidic aspartate residue D115 may suggest the formation of a hydrogen bonding network or a salt bridge between the two residues, which fixes the corresponding helices in their position. A mutation in either of these residues and a possible disruption of this interaction may allow more flexibility of the two helices and therefore open the way for longer isoprenoids. These observations enhance the proposed picture of possible sites for rational evolution of isoprenoid synthase genes as the large aromatic residue blocking the product binding channel can be removed by interaction with side chains of surrounding residues.

Beside the mutations located close to the active site, three other interesting mutants were selected by our screening approach. Mutant 1 contains five amino acid exchanges including the Y79S mutation. Based on the observation with mutant IspA_M3 (Y79 H) this is very likely the effective mutation that alters the product spectrum. However, the other mutations may contribute to the GGPP activity as T234 for example is located next to the SARM region. The same is true for mutant IspA_M21 containing an I76T mutation located in spatial proximity to the SARM region as well.

The mutant IspA_M22 (D2G and C155G) carries a mutation of C155G at 2nd amino acid upstream of conserved region IV ($_{157}$GQALDL$_{162}$). Recently the 2nd amino acid upstream the region IV of prenyl diphosphate synthases was found to be an important factor controlling the chain elongation of Type III GGPP synthase (Hemmi et al., 2003) where the replacement of His to Ala (H139A) at the 2nd amino acid caused the *S. cerevisiae* GGPP synthase to produce longer products. The same region was identified by chemical mutagenesis in a FPP synthase from *Bacillus stearothermophilus* (Ohnuma et al., 1996). In a similar way, the mutant IspA_M22 might produce longer $C_{20}$ GGPP rather than $C_{15}$ FPP. However, the structural mechanism by which this mutation functions remains unclear.

## 4. Conclusion

Directed evolution and an appropriate screening/selection method are regarded as powerful tools for generating desirable mutants from which sequence alternation can be correlated with function. Even though rational design based on structural modeling and site-directed mutagenesis is very effective and informative, a directed evolution approach does not rely on assumptions about enzyme activity drawn from known structures and can therefore provide novel insights into a catalytic mechanism of interest. In this study, in vitro evolved FPP synthases with GGPP synthase activity were generated by mutagenic PCR and screened based on in vivo activity of the enzymes.

Without utilizing protein structural information available for *E. coli* IspA, we created mutant clones that (1) confirm the existing importance of amino acids before FARM of prenyl diphosphate synthase, (2) confirm the significance of other conserved regions such as SARM and (3) suggest novel approaches to engineer the function of prenyl phosphate synthase enzymes. The high proportion of mutants with the desired gain in function (28 from 5000 colonies screened) suggests this approach would be suitable for additional screening to generate a fine map of residues controlling product chain length in this enzyme family and gain additional insights into this important mechanism.

## Acknowledgments

This research was supported by grants from the David and Lucile Packard Foundation and Defense Advanced Research Projects Agency (DARPA-BIOS N66001-02-1-8928).

## References

Hemmi, H., Noike, M., Nakayama, T., Nishino, T., 2003. An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase. Eur. J. Biochem. 270, 2186–2194.

Hosfield, D.J., Zhang, Y., Dougan, D.R., Broun, A., Tari, L.W., Swanson, R.V., Finn, J., 2004. Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. J. Biol. Chem. 279, 8526–8529 Epub 2003 Dec 12.

Kharel, Y., Koyama, T., 2003. Molecular analysis of cis-prenyl chain elongating enzymes. Nat. Prod. Rep. 20, 111–118.

Lee, P.C., Schmidt-Dannert, C., 2002. Metabolic engineering towards biotechnological production of carotenoids in microorganisms. Appl. Microbiol. Biotechnol. 60, 1–11.

Lee, P.C., Momen, A.Z., Mijts, B.N., Schmidt-Dannert, C., 2003. Biosynthesis of structurally novel carotenoids in *E. coli*. Chem. Biol. 10, 453–462.

Liang, P.H., Ko, T.P., Wang, A.H., 2002. Structure, mechanism and function of prenyltransferases. Eur. J. Biochem. 269, 3339–3354.

Ohnuma, S., Koyama, T., Ogura, K., 1992. Chain length distribution of the products formed in solanesyl diphosphate synthase reaction. J. Biochem. 112, 743–749.

Ohnuma, S., Nakazawa, T., Hemmi, H., Hallberg, A.M., Koyama, T., Ogura, K., Nishino, T., 1996. Conversion from farnesyl diphosphate synthase to geranyl geranylgeranyl diphosphate synthase by random chemical mutagenesis. J. Biol. Chem. 271, 10087–10095.

Pan, J.J., Yang, L.W., Liang, P.H., 2000. Effect of site-directed mutagenesis of the conserved aspartate and glutamate on *E. coli* undecaprenyl pyrophosphate synthase catalysis. Biochemistry 39, 13856–13861.

Schmidt-Dannert, C., Umeno, D., Arnold, F.H., 2000. Molecular breeding of carotenoid biosynthetic pathways. Nat. Biotechnol. 18, 750–753.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. 31, 3381–3385.

van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P., Tironi, I.G., 1996. Biomolecular Simulation: The GROMOS96 manual and user guide. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, 1-1042.

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

Zhang, D., Poulter, C.D., 1993. Analysis and purification of phosphorylated isoprenoids by reversed-phase HPLC. Anal. Biochem. 213, 356–361.