NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Cloning and structure-function analysis of farnesyl pyrophosphate synthase and geranylgeranyl pyrophosphate synthase from *Lupinus albus*

Susan-Marie Aitken

A thesis

in the Department of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

December 1995

ISBN 0-612-10818-X

ABSTRACT

Cloning and structure-function analysis of farnesyl pyrophosphate synthase and geranylgeranyl pyrophosphate synthase from *Lupinus albus*

Susan-Marie Aitken

Two distinct farnesyl pyrophosphate synthase (FPS) clones (pFPS1 and pFPS3) and one geranylgeranyl pyrophosphate synthase (GGPS) clone (pGGPS1) were cloned from a 10-day *Lupinus albus* seedling root cDNA library. The pFPS3 and pGGPS1 clones show only 45% similarity and 22% identity in amino acid sequence. However, they contain 5 small regions of conserved amino acids. pFPS3 and pGGPS1 were predicted by secondary structure prediction to have similar structures to each other and to the crystal structure of avian FPS.

Truncated pFPS3, chimeric pGGPS1-pFPS3, and site-directed pGGPS1 clones were designed in order to study the structure-function relationship of these two enzymes. The activities of constructs and mutants were drastically altered. The truncated pFPS3 clones showed similar profiles to pFPS3. The pG1F3a3Ia chimeric, of full-length pGGPS1 with one terminal helix transferred from pFPS3, and the K201E and K201R mutants of pGGPS1 demonstrated modified allylic substrate preference profiles, having decreased geranyl pyrophosphate and increased farnesyl pyrophosphate affinity.

To Mom and Dad

ACKOWLEDGMENTS

I would like to thank Dr. P. Gulick for his support and enthusiasm throughout the course of my studies. Also for giving me the opportunity to explore new avenues of research.

I would also like to thank the members of my committee, Dr. R. Ibrahim, Dr. A. English, and Dr. J. Turnbull for their insight, support, and confidence. Also thanks to Dr. Ibrahim and Dr. Turnbull for the time I spent in each of their labs.

I would like to thank all the postdoctoral fellows with whom I have had the opportunity to work, Dr. J. Grandmaison, Dr. A. Gauthier, Dr. S. Ananvoranich, Dr. S. Gubatz, Dr. J. Saleeba, and Dr. S. Attucci. Their friendship, guidance, expertise, and reality checks is much appreciated.

I would also like to thank my fellow students and labmates whose diverse personalities made my stay interesting: Marc Pelletier, Vivian Saridakis, Dinesh Christendat, Dave Bakker, Avi Stevi, Maria Koutromanis, Colin Reiger, Hubert Gagnon, Wei Shen, Amani Abu-Indres, and many others.

Finally, I would like to thank my parents. Their love, support, patience, and frozen dinners made this all possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BLAST N	nucleic acid similarity search at NCBI
BLAST P	amino acid sequence similarity search at NCBI
BLAST X	1 - 6 reading frame translation and similarity search with amino acid sequences at NCBI
cDNA	Complementary deoxyribonucleic acid
CF	Chou-Fasman method of protein secondary structure prediction
D	Dalton
DEAE	Diethylaminoethyl
DMAPP	Dimethylallyl pyrophosphate
DTT	1,4-Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
FOH	Farnesol
FPP	Farnesyl pyrophosphate
FPS	Farnesyl pyrophosphate synthase
GGOH	Geranylgeraniol
GGPP	Geranylgeranyl pyrophosphate
GGPS	Geranylgeranyl pyrophosphate synthase
GOH	Geraniol
GOR	Garnier-Osguthorpe-Robson method of protein secondary structure prediction
GPP	Geranyl pyrophosphate

IOH	Isopentenyl alcohol
IPP	Isopentenyl pyrophosphate
IPTG	Isopropyl B-D-thioglucoside
LB	Luria broth (media)
mRNA	Messenger ribonucleic acid
NCBI	National Center for Bioinformatics
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PHD	protein-predict method of protein secondary structure prediction
SDS	Sodium dodecyl sulfate
TBS	Tris-buffered saline
TLC	Thin layer chromatography
Tris	Tris-[hydroxymethyl] aminomethane
UTR	Untranslated region

A. LITERATURE REVIEW

A.1 Introduction

Farnesyl pyrophosphate synthase (FPS) and geranylgeranyl pyrophosphate synthase (GGPS) catalyze the sequential condensation of two or three molecules of isopentenyl pyrophosphate (IPP) with allylic pyrophosphates to produce isoprenoid compounds of 15 and 20 carbon lengths, respectively. Both enzymes catalyze the condensation of IPP with dimethyl allyl pyrophosphate (DMAPP) to produce geranyl pyrophosphate (GPP), as well as a second condensation between IPP and GPP to produce farnesyl pyrophosphate (FPP). GGPS can efficiently catalyze an additional condensation, between FPP and IPP, to produce the 20 carbon compound geranylgeranyl pyrophosphate (GGPP), whereas FPS has only weak activity for this reaction. FPP and GGPP are key intermediates in the synthesis of a vast and complex array of naturally occurring isoprenoid metabolites (Poulter and Rilling, 1981). Isoprenoid compounds are also referred to as terpenoids, a term having its origins in the turpentine oils originally used to isolate and study a variety of these compounds in plants (Chappell, 1995). The isoprenoid family is one of the most diverse groups of naturally occurring compounds with over 23 000 known members (Chen et al., 1994; Tarshis et al., 1994). The pathway shown in Figure 1 demonstrates that isoprenoid compounds are formed from acetyl CoA *via* mevalonate. The isoprenoid biosynthetic pathway is thought to be ubiquitous, since at least portions of it have been discovered in all organisms that have been examined (Poulter and Rilling, 1978). Therefore, the pathway shown (Fig. 1) encompasses a wide range of prokaryotic and eukaryotic organisms including bacteria, yeast, animals, and

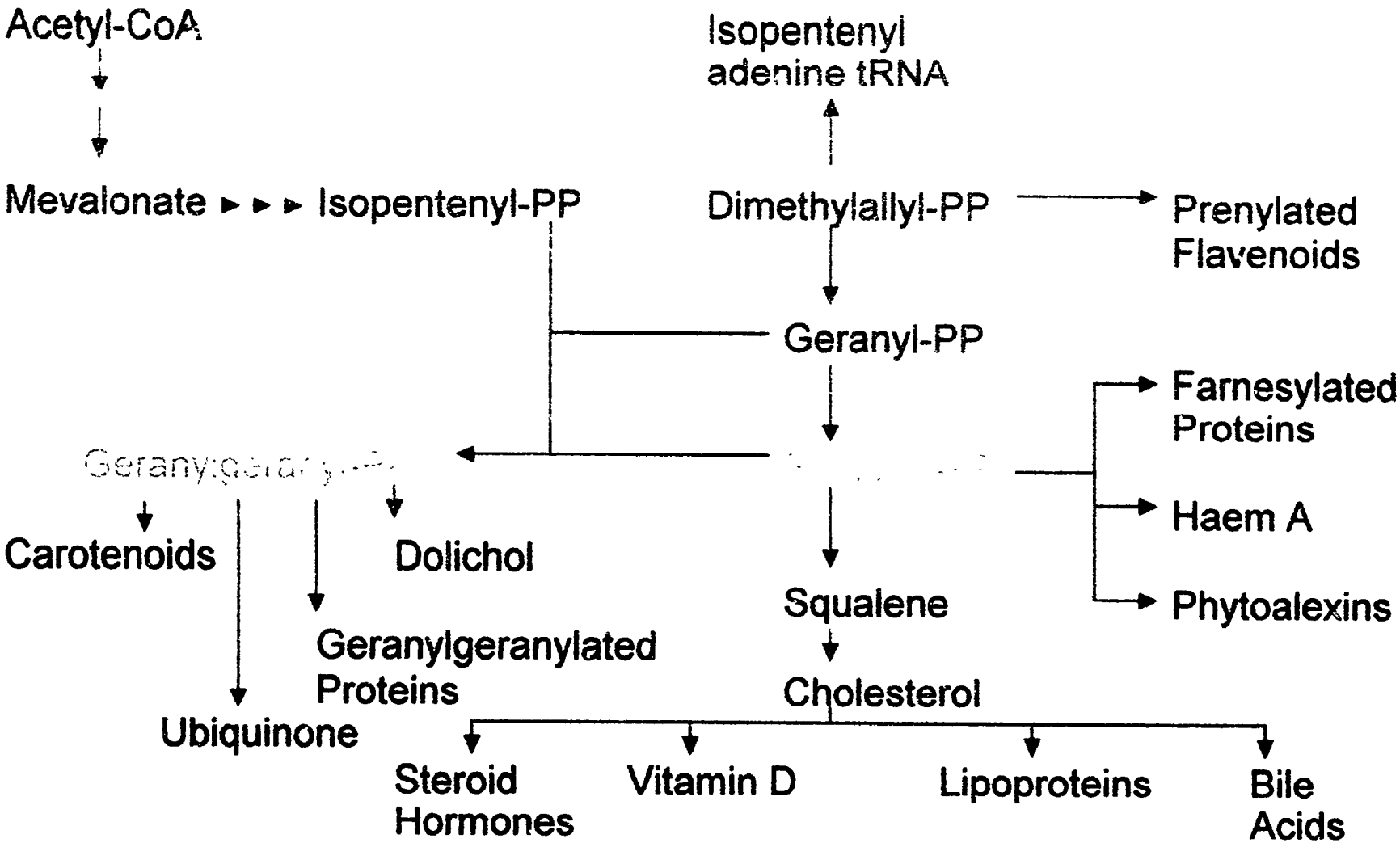


Figure 1. General isoprenoid biosynthetic pathway.

plants, though not all functions are carried out by any given species. The figure also indicates the multitude of compounds produced via this pathway; these include photosynthetic pigments (chlorophyll and carotenoids), mitochondrial electron transport components (ubiquinone and cytochrome), sterols (cholesterol, steroid hormones, gibberellins and abscisic acid), phytoalexins and phytoanticipins, dolichols, prenylated proteins, and a number of other compounds (Green and West, 1974; Ashby et al., 1990; Poulter and Rilling, 1981; Tarshis et al., 1994; Chen et al., 1994; Song and Poulter, 1994).

In plant species isoprenoids can be divided into the classes of primary and secondary metabolites. The primary metabolites include sterols, carotenoids, growth regulators such as gibberellins and abscisic acid, and the polyprenol substituents of dolichols, quinones, and proteins. The secondary metabolites include a wide variety of compounds derived from monoterpenes (C10), sesquiterpenes (C15), and diterpenes (C20). Unlike the primary metabolites, these compounds are not essential for viability, but mediate interactions between plants and their environment and act as protective agents and in unknown roles (Chappell, 1995).

Farnesyl pyrophosphate synthase and geranylgeranyl pyrophosphate synthase are branch-point enzymes in isoprenoid biosynthesis, initiating, respectively, pathways for C15 and C20 based compounds. Due to their central role in isoprenoid biosynthesis, FPS and GGPS have been studied in a number of animal, bacterial, yeast, and plant species. FPS has been examined most intensively for its role in cholesterol biosynthesis and it has been targeted as an enzyme whose inhibition could achieve controlled reduction

of cholesterol in mammalian systems (Spear et al., 1992; Popjak, 1971). This enzyme has been purified from a number of animal sources including human, rat, porcine, and avian liver (Eberhardt and Rilling, 1975; Reed and Rilling, 1975; Yeh and Rilling, 1977; Barnard and Popjak, 1981) and genes encoding FPS have been cloned from human (Wilkin et al., 1990; Sheares et al., 1989), rat (Teruya et al., 1990; Spear et al., 1992), *E. coli* (Fujisaki et al., 1990), *S. cerevisiae* (Anderson et al., 1989), *B. stearothermophilus* (Koyama et al., 1993), and recently from three plant sources, *Arabidopsis thaliana* (Delorme et al., 1994), *Zea mays* (Li and Larkin, 1995), and *Lupinus albus* (Attucci et al., 1995a, Attucci et al., 1995b). GGPS has been studied mainly for its role in carotenoid biosynthesis (Ericsson et al., 1993). Therefore, it has been purified in a number of plant, bacterial, and fungal species. Clones of genes encoding GGPS have been isolated from various bacterial and fungal species including *Erwinia herbicola* (Math et al., 1992), *Neurospora crassa* (Carattoli et al., 1991), *Erwinia uredovora* (Misawa et al., 1990), and recently in three plant species: *Capsicum annuum* (Camara et al., 1991), *Arabidopsis thaliana* (Bartley and Scolnik, 1994), and *Lupinus albus* (Aitken et al., 1995).

The reactions catalyzed by FPS and GGPS are shown in Figure 2. Both of these enzymes are capable of catalyzing more than one prenyl transfer. FPS catalyzes the transfer of a prenyl donor, either dimethylallyl pyrophosphate (DMAPP) or geranyl pyrophosphate (GPP), to a prenyl acceptor, isopentenyl pyrophosphate (IPP), to produce farnesyl pyrophosphate (FPP) (reactions 2 and 3) as an end product. GGPS catalyzes the transfer of a prenyl donor, DMAPP, GPP, or FPP, to a prenyl acceptor, IPP to produce

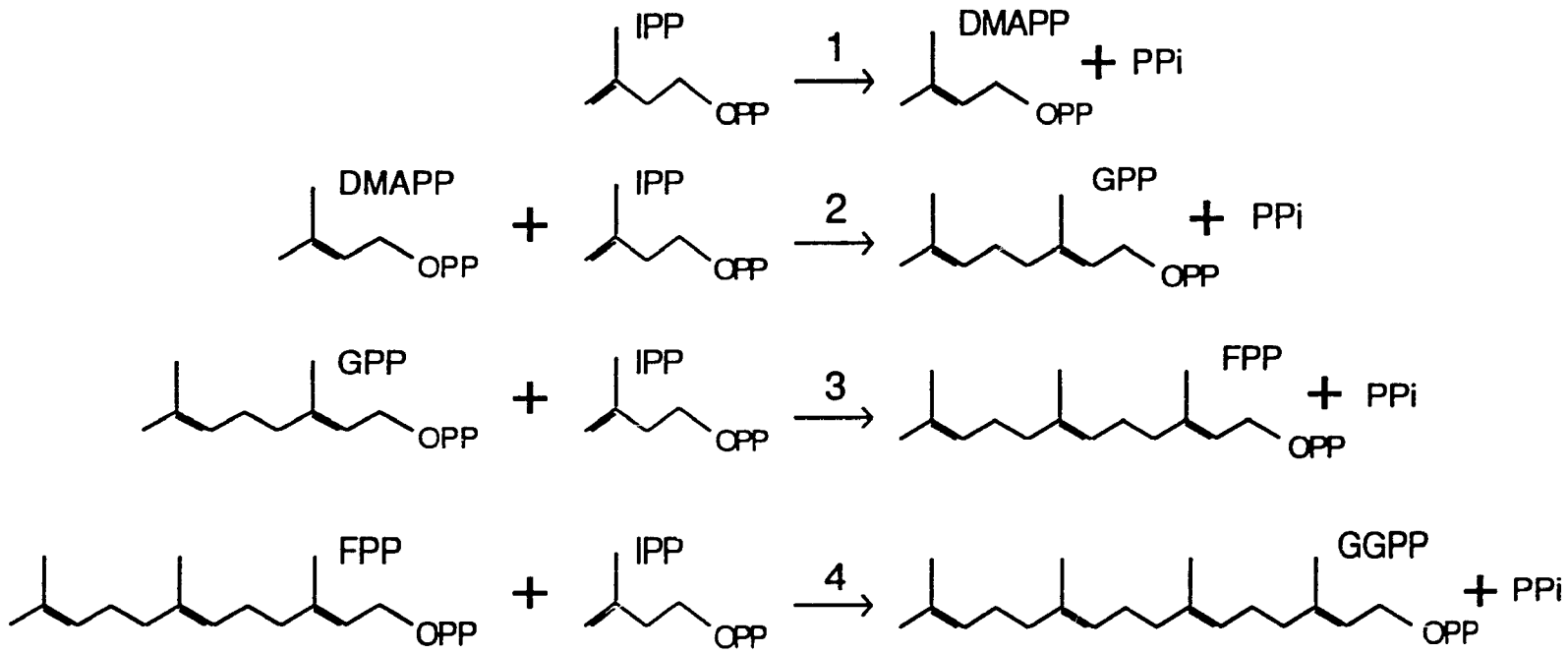


Figure 2. Reactions catalyzed by FPS (2 and 3) and GGPS (2, 3, and 4).

geranylgeranyl pyrophosphate (GGPP) (reactions 2, 3, and 4) as an end product.

FPS and GGPS belong to the class of prenyltransferases which catalyze the head-to-tail condensation reactions in which both the donor and acceptor molecules are prenyl pyrophosphates. They are distinguished from two other classes of prenyltransferases, one of which catalyze head-to-head condensations of polyprenyl compounds and includes such enzymes as squalene synthase and phytoene synthase of the sterol and carotenoid biosynthetic pathways, respectively. The third class of prenyltransferases are those which transfer either a mono or polyprenyl moiety to a nonprenyl acceptor molecule. An example of this group are the protein prenyltransferases (Marrero et al., 1992; Poulter and Rilling, 1978). Cyclases may be considered to be a fourth class of prenyltransferases, catalyzing an intramolecular prenyl transfer. In condensations between two isoprenoids, the homoallylic substrate (IPP) is numbered normally and the allylic substrate is numbered with a prime annotation to distinguish between the two (Poulter and Rilling, 1981).

A.2 Stereochemistry and mechanism of the reaction

The cryptic stereochemistry involved in the chain elongation reactions of isoprenoid biosynthesis provides an indication of the reaction mechanism and the orientation of the substrates with respect to one another (Ito et al., 1987; Poulter and Rilling, 1981). Cornforth et al. (1966b) used (1R)-[1-^{2}H]DMAPP to determine that there is inversion of configuration at the C1' position during condensation. They also used (2R)-[2-^{3}H]IPP and (2S)-[2-^{3}H]IPP, to show that it is the pro-R hydrogen of IPP that is

eliminated in the final step of the condensation. The direction of attack on the double bond of IPP (Fig. 3) was determined to be at the *si*-face by using (Z)-[4-^{2}H]IPP followed by determination of the stereochemistry of the 4- and 8-positions of FPP (Cornforth et al., 1966a). Ito et al. (1987) provided evidence that the *si*-addition-pro-R-elimination stereochemistry of the prenyltransferase reaction observed with FPS (Cornforth et al., 1966b; Cornforth et al., 1966a) would seem to be conserved among (all-E) prenyltransferases from both prokaryotic and eukaryotic species.

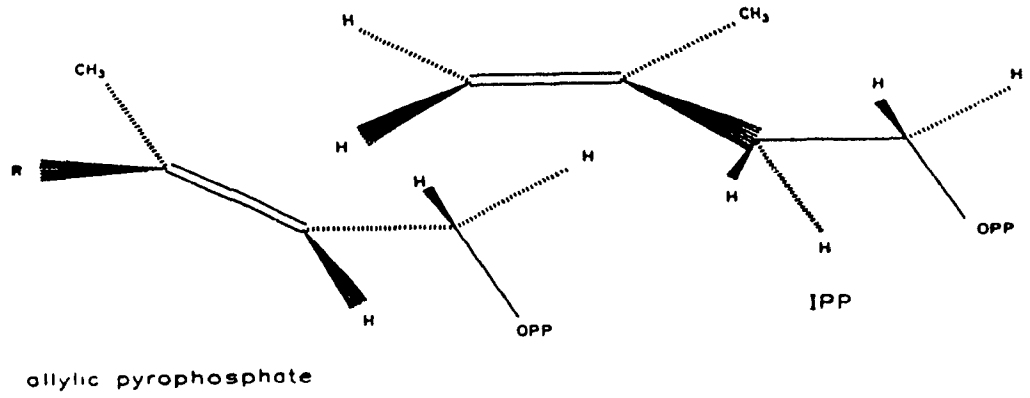


Figure 3. Stereochemistry of the prenyl transfer reaction.

An ionization-condensation-elimination (I-C-E) mechanism was proposed for the prenyl transfer reaction based on the knowledge that allylic pyrophosphates are easily ionized. This mechanism involved an allylic carbonium ion intermediate (Poulter and

Rilling, 1981). Since the stereochemical data indicated there is inversion at the C1'-position upon condensation, and that the allyl group adds and the proton leaves from the same side of IPP, Cornforth et al. (1966a) proposed a two-step condensation-elimination mechanism in which an electron-donating X-group was suggested in order to maintain the stereochemistry. Thus there were two possible reaction mechanisms for FPS: (1) the condensation-elimination mechanism, also referred to as the X-group mechanism and (2) the ionization-condensation-elimination mechanism (I-C-E).

It was possible to conclude from the stereochemical data, the ability of the enzyme to hydrolyse the allylic substrate in the absence of IPP (Poulter and Rilling, 1976), the inability to affinity label the X-group (Poulter et al., 1978), the presence of resonance stabilization, and the rate reductions observed for the allylic fluoro-analogues, that the mechanism of FPS is ionization-condensation-elimination (Poulter and Rilling, 1976; Poulter and Rilling, 1978; Poulter and Satterwhite, 1977). However, while the I-C-E mechanism does not require the presence of an X-group, neither does it preclude it.

A.3 Substrate binding

Substrate binding studies with avian liver FPS showed that the allylic pyrophosphates (DMAPP, GPP, and FPP) compete for the same site on the enzyme and bind at 2 mol of allylic substrate per mol of enzyme (Reed and Rilling, 1976). This is consistent with the presence of one allylic binding site per monomer. They determined the relative binding affinity for the allylic substrate to be: GPP > FPP > DMAPP. FPS

is capable of binding 4 mol of IPP per mol of enzyme, in the presence of divalent cation and absence of allylic substrates, although this is reduced to 2 mol in the presence of a nonreactive allylic analogue. This indicates that IPP can bind to both the allylic and the homoallylic sites and is displaced from the allylic site in the presence of the allylic substrate (Reed and Rilling, 1976). Inorganic pyrophosphate is capable of binding to both sites on the enzyme, however it requires the presence of divalent cations. Saito and Rilling (1979) used inorganic pyrophosphate as a probe to investigate the binding of FPP. They found that FPP reduced the binding of inorganic pyrophosphate from 4:1 to 0-0.5:1 in contrast with IPP, which only reduced its binding to 2:1. They propose that the allylic binding site is "small and cup-shaped", requiring the folding of the allylic substrate at C6' and causing the hydrocarbon-moiety of the allylic substrate to protrude into the homoallylic binding site if it is longer than the natural substrates (Saito and Rilling, 1979).

Saito and Rilling (1979) also found that inorganic pyrophosphate facilitates the hydrolysis of allylic pyrophosphates by acting as an IPP analogue and allowing water to act as the prenyl acceptor. Two possible mechanisms have been proposed to explain how inorganic pyrophosphate facilitates hydrolysis: (1) it enhances the binding of water to the binding site of the hydrocarbon-moiety of IPP, or (2) it induces a conformation change in the enzyme that is required for catalysis (Poulter and Rilling, 1981).

Although divalent cations, Mg^{++} and/or Mn^{++}, are required for prenyl transfer and enhance binding, they are not required for binding to occur. King and Rilling (1977) determined that in their absence both the allylic and homoallylic substrates bind and

while there is no competition between DMAPP or GPP and IPP, there is competition between FPP and IPP. At high concentrations of IPP, in the absence of divalent cation, 2 mol of IPP and 2 mol of FPP are bound per mol of enzyme. 2-fluorofarnesyl pyrophosphate can compete with IPP for binding in the presence of divalent cation indicating that FPP can either occupy the allylic site or span both sites. This provides support for the suggestion of Reed and Rilling (1976) that FPP was capable of binding to the allylic site even though it is the normal end-product of the enzyme (King and Rilling, 1977).

King and Rilling (1977) showed that, while divalent cations do not bind in the absence of substrate, 2 divalent cations are bound to the enzyme when either or both of the substrates are bound. The "hot-trap" experiments of Laskovics and Poulter (1981) showed that both substrates bind as their magnesium salts. Therefore, the most rational explanation for divalent cation binding would seem to be that they interact with both substrates, possibly by bridging between the pyrophosphate moieties. This would allow the substrates to be fixed in relation to each other and would position the pyrophosphate moiety of the allylic substrate in such a way that while leaving it could aid in the elimination of the *pro*-R hydrogen (King and Rilling, 1977). Therefore, when the substrates are positioned for the formation of a trans double bond, only the *pro*-R hydrogen is accessible to the leaving pyrophosphate. The difference between the E- and Z- prenyltransferases is in the binding of IPP in relation to the allylic pyrophosphate (Popjak, 1971; King and Rilling, 1977; Poulter and Rilling, 1978; Poulter and Rilling, 1981).

A.4 Enzyme mechanism

Holloway and Popjak (1967) and Popjak et al. (1969) used porcine liver FPS to conduct product inhibition studies in an effort to determine the kinetic mechanism. The results showed that two mechanisms were possible, either a ping-pong mechanism (with IPP binding, PPi being released, GPP binding, and FPP being released), or an ordered-sequential mechanism. The ping-pong mechanism was discarded because it required the IPP bind and PPi be released before GPP bound and thus requiring the pyrophosphate moiety to be released from IPP instead of the allylic substrate. This was extremely unlikely considering the previously determined mechanism of the uncatalyzed reaction. Therefore, the ordered-sequential mechanism was accepted with GPP adding before IPP, and FPP being released before PPi.

It was later shown that product inhibition data were insufficient to distinguish between the ping-pong and ordered-sequential mechanisms because FPP and PPi do not add in the same way as generated on condensation. Experiments with fluorinated dead-end inhibitors were used to distinguish between these two mechanisms and supported the sequential mechanism. However, it was still not possible to distinguish between ordered and random sequential mechanisms because analogues of IPP did not bind specifically to the homoallylic site (Holloway and Popjak, 1967; Popjak et al., 1969a; Reed and Rilling, 1976; King and Rilling, 1977; Saito and Rilling, 1979).

Laskovics and Poulter (1981) using "hot-trap" experiments, found that 5 μM Mg^{++}-IPP was able to trap Mg^{++}-GPP, while concentrations of Mg^{++}-GPP as high as 2 mM were not able to trap Mg^{++}-IPP. They concluded that the reaction is ordered with

GPP adding before IPP.

Laskovics and Poulter (1981) also determined that the rate constants for addition of IPP and GPP and for dissociation of FPP, although not controlled by diffusion alone, are lower than the expected diffusion-controlled limit. This is an indication that a conformational change in the enzyme is associated with binding and dissociation. This is supported by the finding that the initial rate of the reaction is approximately 50 times faster than the steady-state rate. Therefore, product release is the rate limiting step in the reaction catalyzed by FPS and this is explained by the requirement for a conformational change to allow the products to be released (Laskovics and Poulter, 1981; Poulter and Rilling, 1981).

The reaction proceeds as follows:

$$E + GPP \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\blacklozenge GPP \underset{k_{-2}}{\overset{k_2\ [IPP]}{\rightleftharpoons}} E\blacklozenge GPP\blacklozenge IPP \xrightarrow{k_3} E\blacklozenge FPP\blacklozenge PPi \xrightarrow{k_4} E + FPP + PPi$$

with the following rate constants as determined by King and Rilling (1977) for the avian liver FPS:

k_1	$2 \times 10^6\ M^{-1}s^{-1}$
k_{-1}	$1.4\ s^{-1}$
k_2	$5 \times 10^6\ M^{-1}s^{-1}$
k_{-2}	$5.0\ s^{-1}$
k_3	$4.7\ s^{-1}$
k_4	$0.1\ s^{-1}$

The Michaelis constants for both IPP and GPP for FPS and for DMAPP, GPP, FPP, and IPP for GGPS have been determined for enzyme from a number of sources and some are listed in Table 1 and Table 2, respectively.

Table 1. Michaelis constants for FPS

source	K_m (μM) GPP	IPP
human liver	0.44	0.94
rat	1.0	2.3
avian liver	0.5	0.5
pig liver	0.58	0.31
E. coli	8.7	4.7
yeast	6.7	3.4
B. stearothermophilus	5.8	4.7

human liver FPS (Barnard and Popjak, 1981); *Bacillus stearothermophilus* FPS (Koyama et al., 1994a); rat FPS (Joly and Edwards, 1993); avian liver FPS (Reed and Rilling, 1976); pig liver FPS (Yeh and Rilling, 1977); *Saccharomyces cerevisiae* FPS (Song and Poulter, 1994).

Table 2. Michaelis constants for GGPS

source	K_m (μM) DMAPP	GPP	FPP	IPP
C. annuum	0.95	1.0	1.2	3.0
rat	n.a.	n.a.	0.6	3.5
P. blakesleanus	n.a.	60	9	n.a.
bovine brain	33	0.8	0.74	2.0
E. herbicola	1800	9	11	36
M. theroformicicum	16.8	12.6	4.7	30.8
S. cerevisiae	4	8	14	n.a.

Capsicum annuum GGPS (Dogbo and Camara, 1987); *Saccharomyces cerevisiae* GGPS (Eberhardt and Rilling, 1975); rat GGPS (Ericsson et al., 1993); *Phycomyces blakesleanus* GGPS (Lutke-Brinkhaus and Rilling, 1988); bovine brain GGPS (Sagami et al., 1994); *Methanobacterium thermoformicicum* GGPS (Tachibana et al., 1993); *Erwinia herbicola* GGPS (Wiedmann et al., 1993).

A.5 Substrate analogues

FPS has been intensively studied because of its role in the biosynthesis of cholesterol and thereby its potential implication in atherosclerosis; many substrate analogues for this enzyme have been synthesized. Substrate analogue work has also been done with GGPS, but to a lesser extent than with FPS. The majority of substrate analogues fall into two categories: (1) those used to determine the mechanism of the condensation, and (2) those intended for probing the substrate specificity of the enzyme (Poulter and Rilling, 1981). The substrate analogues designed to determine the mechanism of the reaction have been discussed above.

Binding studies with substrate analogues elucidate a number of characteristics concerning the binding specificity of these enzymes. The pyrophosphate moiety is essential for binding and catalysis for both the allylic and homoallylic substrates. The monophosphate and alcohol derivatives of the natural pyrophosphorylated substrates are inactive. Whereas, the alcohols do not bind strongly, the monophosphates are inhibitors (Popjak et al., 1969, Holloway and Popjak et al., 1969, Popjak, 1971). In addition, the phosphonate and phosphonophosphate analogues of DMAPP and GPP were strong inhibitors of FPS (Parker et al., 1978).

The C2'-C3' double bond of the allylic substrate is another essential feature. Popjak et al. (1969b) showed that citronellyl pyrophosphate, the 2,3-dihydro analogue of GPP was inactive as a substrate, but acted as an effective inhibitor. It is possible that the double bond is essential due to its role in the resonance stabilization of the allylic carbocation following elimination of the pyrophosphate moiety.

There is little flexibility for substrate binding with reference to the C2'-position. A hydrogen is normally present at this position and the only exception observed are the 2-fluoro analogues, although these are the poor substrates (Ogura et al., 1970; Poulter et al., 1978; King and Rilling, 1977).

The presence, length, and geometry of the alkyl substitutions at the 3'-position of the allyl substrate have been the focus of intense investigation due to the ability of these enzymes to accept allylic substrates of varying lengths. Using FPS from porcine liver (Popjak et al., 1969c; Nishino et al., 1973) and pumpkin (Ogura et al., 1970; Nishino et al., 1970; Nishino et al., 1971; Nishino et al., 1973) the maximum lengths of the E- and Z-alkyl substitutions were probed (Fig. 4). The Z-series findings differed depending on the source of the enzyme. The pumpkin FPS was capable of accepting only the subsitution of an ethyl group for the methyl group normally at this position (Ogura et al., 1970; Nishino et al, 1970; Nishino et al., 1971; Nishino et al., 1973). In contrast, the porcine liver FPS the Z-methyl group can be replaced by linear chains up to butyl in length (Popjak et al., 1969; Nishino et al., 1973).

The FPS from both pumpkin and porcine liver were both much more flexible in their ability to accept variation at the E-position. The longest trans analogue accepted by FPS from both sources was E-3-methyl-2-dodecenyl-PP, which is a C13 analogue, resulting in a C18 product, the longest product produced (Nishino et al., 1970; Nishino et al., 1973). Nishino et al. (1970) also studied a series of 7,8-dihydrogeranyl-PP analogues (3,7-dimethyl-2-octenyl-PP (n=3)), in which the distance between the two methyl substitutions (n) was varied. They found that very low activity was obtained

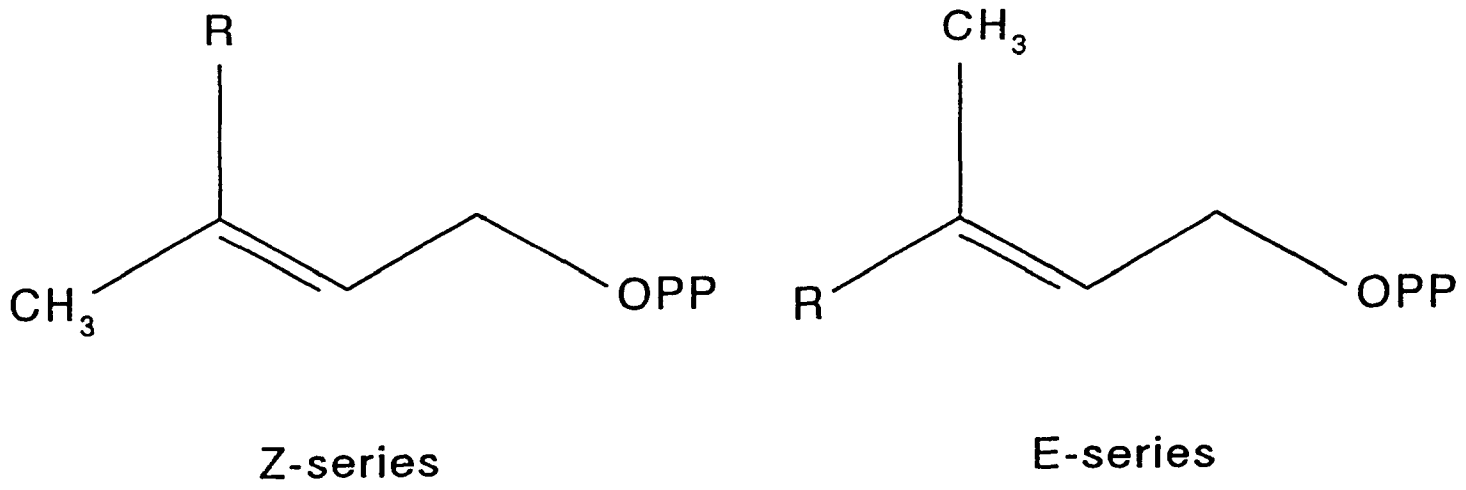


Figure 4. *Cis* and *trans* series substrate analogues.

when this distance was less than 3 carbons. This indicates that FPS is less tolerant of branching close to the 3-position.

Nishino et al. (1971), using pumpkin FPS performed an elegant experiment to determine the importance of geometry on substrate specificity. They compared the activities of various C7 and C8 allylic analogues, which varied only in the degree of cis or trans nature of the substitutions at the 3-position. They found that within both the C7 and C8 series the K_ms of the analogues decreased throughout the series with increasing trans nature. While these findings reinforced previous work (Ogura et al., 1970; Nishino et al., 1970) that had determined the Z-series length limit for pumpkin FPS, they were also able to show that the relationship is not simple because the cyclic C7 and C8 compounds (incorporating both the E- and Z-substitutions of the 3'-position) showed the highest V_{max} of the series in both cases and the highest K_m for the C7 and the lowest K_m for the C8 cyclic analogue.

Shinka et al., (1975) compared the substrate and substrate analogue specificities of GGPS from pumpkin and *Micrococcus lysodeikticus*. They found that the substrate specificities differed between the two, with that of the bacterial GGPS being DMAPP > FPP > GPP and the plant enzyme being GPP > DMAPP > FPP. While both enzymes accepted a broad range of cis and trans series substrate analogues, the bacterial enzyme preferred some of them to GPP, while this was not true of the plant enzyme. They also showed that the longest product was a C21 compound produced by the addition of 2 molecules of IPP to the C11 analogue 3-methyl-2-dodecenyl-PP. This product is two carbons longer than the natural final product (GGPP). In contrast, FPS is capable

of producing a C18 product, that is 4 carbons longer than its natural final product (FPP). The C18 product of FPS is the same length as GGPP.

Studies of analogues of IPP have shown that an alkyl substitution is required at position three and only an ethyl group is tolerated in the place of the methyl group normally present (Popjak et al., 1969). Ogura et al. (1974) made a series of IPP analogues with differing lengths of linear carbon chains between positions C1 and C3 in the natural substrate. Only the analogue possessing an extra methylene was active, however, the C3-C4 double bond of the product was in the cis conformation rather than the trans form normally produced. This suggests that the conformation of the analogue was such that the angle between the C2-C3 and the C4-C5 bonds was less than 90°, as compared to IPP in which it is close to 180°.

The substrate analogue work described above showed that FPS and GGPS are tolerant of some modifications in the structure of the hydrocarbon-moiety of both the allylic and the homoallylic substrates. This information was used by Brems and Rilling (1979) in the design of a photoreactive substrate analogue for labelling the catalytic site of the enzyme. Among 3 photoreactive analogues, o-azidophenethyl pyrophosphate had the highest affinity. The labelling of avian liver FPS was monitored by loss of enzyme activity; over 90% of activity was lost upon irradiation in the presence of the photoreactive analogue.

Brems and Rilling (1979) determined that the o-azidophenethyl pyrophosphate was covalently bound to the enzyme with a stoichiometry of 2 mol per mol of enzyme, or one label per catalytic site. The results of protection studies showed that FPP gave complete

protection, IPP gave partial protection, and GPP did not protect at all against labelling. This indicates that o-azidophenethyl pyrophosphate specifically labels the homoallylic site of FPS, although labelled enzyme looses its ability to bind either substrate. After CNBr digestion of the photolabelled avian liver FPS, Brems et al. (1981) isolated a 30 amino acid fragment containing 80% of the label. They found that 16 of the 30 amino acids in the fragment were labelled, although only a fraction of each one was modified overall. Of these labelled amino acids, one arginine and one alanine were most extensively labelled. These results indicate that the label was distributed nonselectively. Without structural information little could be deduced about the nature of the catalytic site from the photoaffinity labelling.

The identity of the amino acids at the active site was also probed by modifying agents specific for certain amino acids. A study by Holloway and Popjak (1967) showed that iodoacetamide, N-ethylmaleimide, and p-hydroxymercuribenzoate inhibited activity and suggested that the sulfhydryl groups of one or more cysteine residues may be involved in catalysis or maintenance of structural integrity.

Barnard and Popjak (1980) demonstrated that the arginine-specific agent, phenylglyoxal, caused a biphasic inactivation of porcine liver FPS. This was interpreted as an indication that enzyme inactivation results from modification of 2 arginine residues per subunit. Protection data showed that while DMAPP and IPP gave little or no protection, GPP and divalent cations were both capable of protection against inactivation by phenylglyoxal. The results of the protection studies indicate that it is likely that it is the allylic site that is modified by phenylglyoxal.

A.6 Structure

The number of sources from which prenyltransferases have been cloned has allowed multiple sequence alignments to be used in order to identify conserved sequences. In 1994 Chen et al. aligned the amino acid sequences of six FPS, six GGPS, and one hexaprenyl pyrophosphate synthase (HPS). The results of these alignments showed that pairwise percentage identities seemed to follow an evolutionary trend with prenyltransferases from closely related species being more similar than those from more distantly related species. Although there was little conservation of primary amino acid sequence, all of the genes contained five small, completely conserved regions. Two of these conserved regions, domains II and V, are aspartate- and arginine-rich. They proposed that these prenyltransferases had diverged from a common ancestral enzyme. They also proposed that if this evolutionary theory were correct it would be reasonable to assume that in addition to the five regions of conserved primary sequence, the gross topological features of this ancestral enzyme would have been conserved.

Based on amino acid composition Chen et al. (1994) also predicted that these enzymes were either entirely α-helical or α-helical-dominated. They then used both the Chou-Fasman (CF) and Garnier-Osguthorpe-Robson (GOR) methods to predict an average structure of eight to ten α-helices and two to four short β-sheets. High α-helical content was predicted for this family of enzymes based both on amino acid composition and on secondary structure prediction.

Tarshis et al. (1994) produced the first crystal structure of a prenyltransferase enzyme using cloned avian liver FPS. This enzyme has an entirely α-helix and loop

structure (Fig. 5). The protein is comprised of a single domain of ten major α-helices arranged in three planes. The second (helices J, C, D, and E) and third (helices F, G, H, and I) planes, are nearly parallel and stack against each other. The smaller first plane, which is comprised of only helices A and B is parallel to the two larger planes with the two helices running perpendicular to the helices of the other planes.

The crystal structure also verified that FPS exists as a homodimer. Dimerization occurs, in an end-on-end manner, along the plane of the two major planes of α-helices and elongates each of these from four to eight helices (Fig. 6). Helices E and F form the dimerization interface and the structure is secured on either side by the perpendicular helices A and B from each monomer. The resulting homodimeric structure is a long, flat protein with approximate dimensions of 65X65X35 Å. There are extensive hydrophobic interactions both between the antiparallel α-helices of the major planes and at the dimerization interface.

On the lower end of the protein the loops between the helices are short (Fig. 5). They are all no longer than nine amino acids in length and contain amino acids that are often found in tight turns, such as glycine and proline. In constrast, the loops at the top of the protein are much longer, ranging between thirteen and sixty amino acids in length and, with the exception of the loop between helices H and I, they all have an extended structure. The loop between helices H and I is remarkable because it contains three small α-helices (α1, α2, and α3) that form a spiral at the top of the protein. This loop is thought to have some flexibility of movement and interacts with the core of the protein only between α-3 and helix I and the F-G loop.

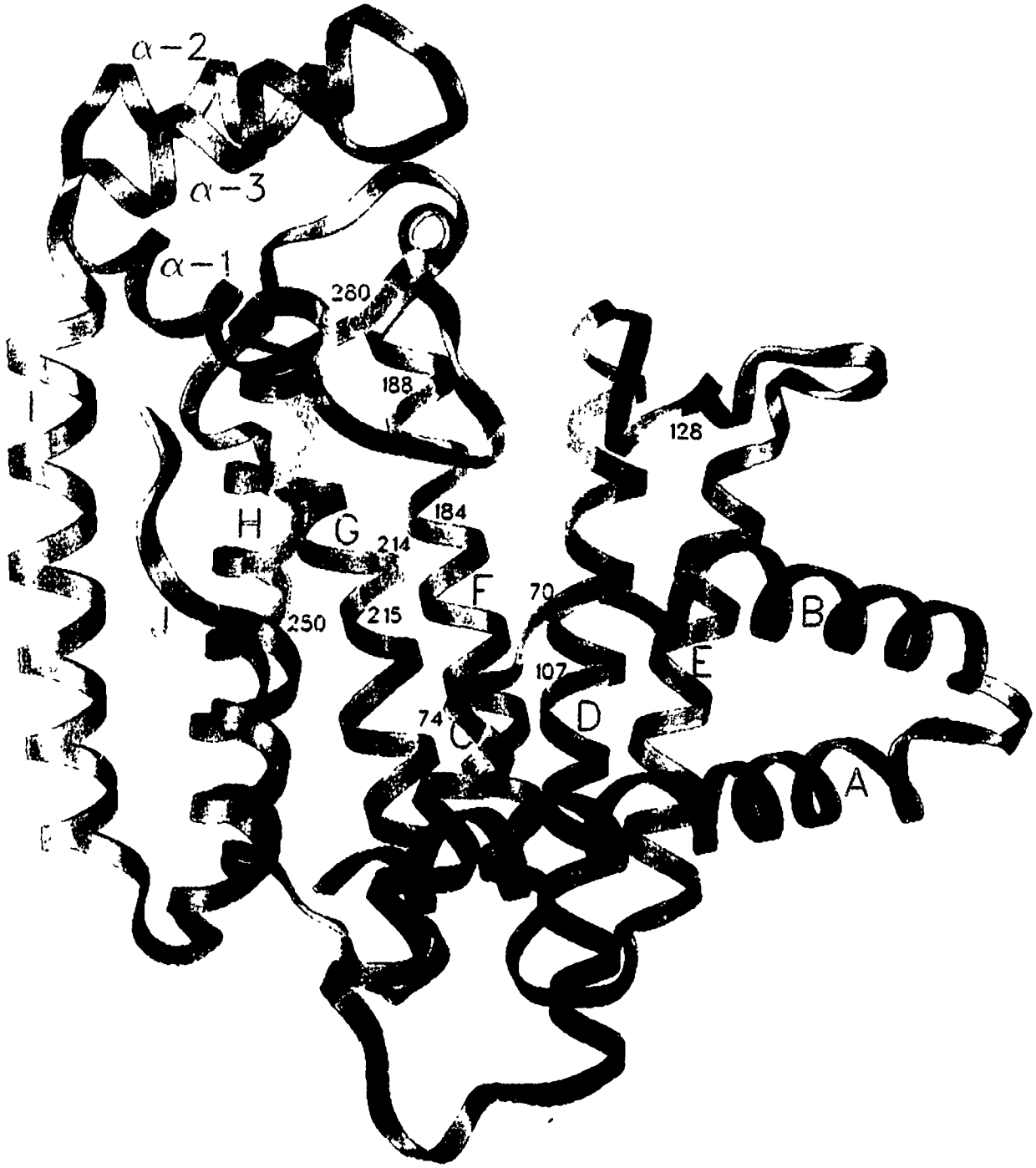


Figure 5. Structure of the avian liver FPS monomer (Tarshis et. al., 1995).

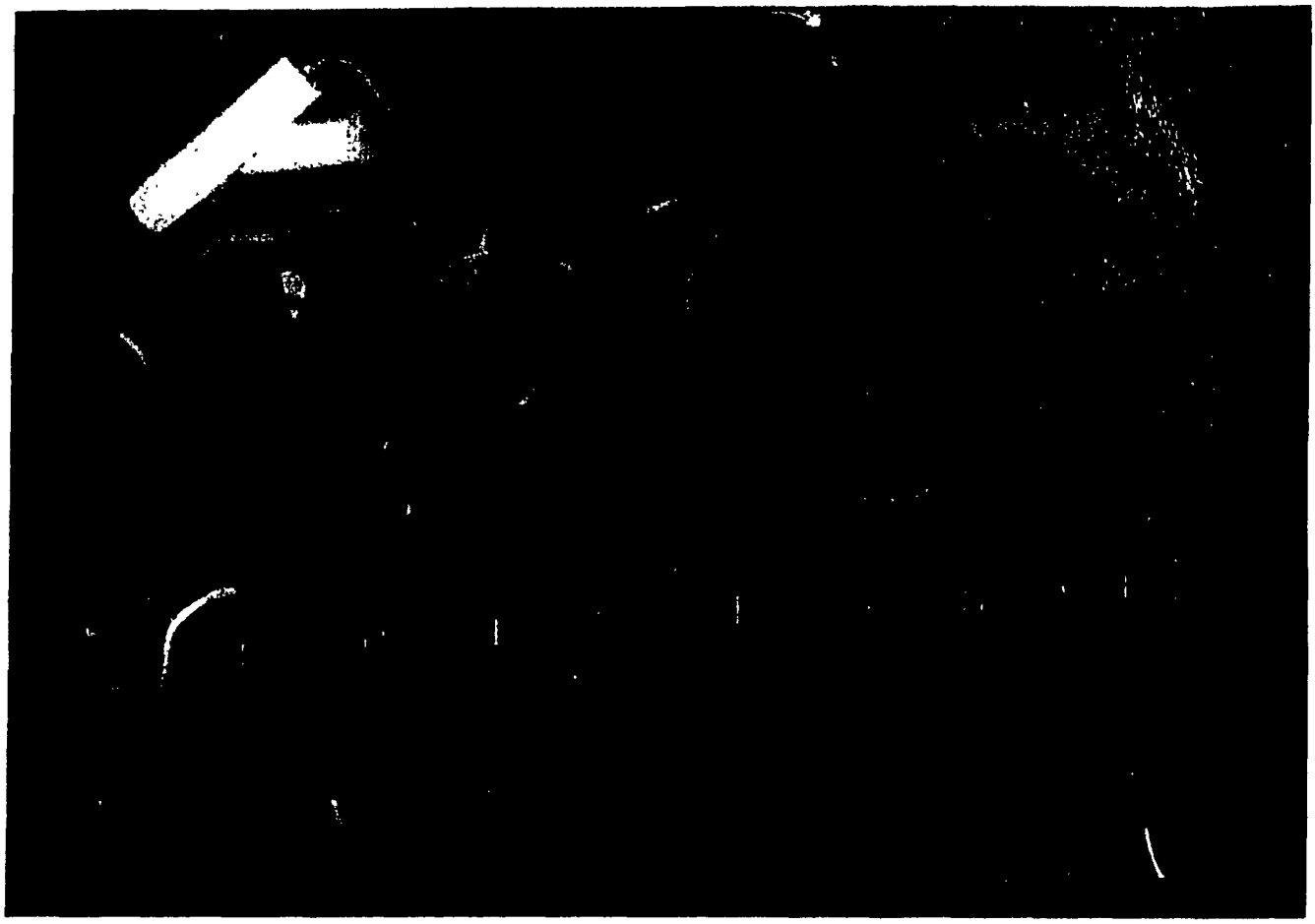


Figure 6. Structure of the avian liver FPS dimer (Tarshis et. al., 1995).

Between the two major planes of helices, the core helices surround a large cleft that measures approximately 14X18X18 Å. The crystal structure shows that the floor of the cleft is formed by helices J and C, the back wall by helices F and G, the sides by helices D and H, and the roof by the D-E and H-I loops.

One very important feature shown by the crystal structure is the placement of the conserved residues in relation to each other and to the cleft. All six of the aspartates of conserved domains II and V are positioned in such a way that their carboxylate side chains are pointing into the cleft. These conserved aspartates of domains II and V are located at the carboxy-termini of helices D and H, respectively, so that they oppose each other across the cleft. Evidence for the role of these conserved aspartate residues is also presented, as crystals of samarium heavy atom derivatives of FPS showed that there are two samarium-binding sites, one at each of the two aspartate-rich regions on either side of the cleft, approximately 10 Å apart. The conserved arginine and lysine residues of domains II and V, as well as most of the other conserved amino acids, were also shown to be oriented in such a way that their side-chains pointed into the cleft.

Tarshis et al. (1994) also crystallized FPS in the presence of IPP. They found that two molecules of IPP bound per monomer and that their diphosphate moieties were located adjacent to the aspartate-rich regions on either side of the cleft. This confirms what was previously known from the binding studies (Reed and Rilling, 1976; King and Rilling, 1977; Saito and Rilling, 1979), that two molecules of IPP can bind per active site in the absence of the allylic substrate. Tarshis et al. (1994) stated that IPP and DMAPP could be placed in the regions occupied by IPP in the crystal structure in a

manner that would allow them to form GPP, but did not discuss what this orientation would be or give any additional discussion on the subject.

Tarshis et al. (1994) also observed that the location of the two IPP molecules in the crystal structure leaves both the hydrocarbon tails and the carbocation that would be formed exposed to water. Therefore, they suggested that there is likely a conformational change that accompanies binding. Such a change has also been suggested by others, based on kinetic data, which showed that binding is slower than would be expected if it was only controlled by diffusion (Laskovics and Poulter, 1981). Tarshis et al. (1994) suggest that the 264-288 (in the avian FPS sequence) loop is flexible based on its weak electron density and that it has three glycine residues (264, 270, 273) which would aid in movement.

A.7 Site-directed mutagenesis

Since clones for genes encoding FPS have become available there have been a number of studies based on site-directed mutagenesis of this enzyme. Most of this work has been carried out by the labs of Edwards and Poulter at University of Utah and by the lab of Ogura at Tohoku University in Japan. Targets for mutagenesis have been chosen from previous binding studies and from the analysis of multiple sequence alignments. To date, most of this work has focused on the conserved aspartate, arginine, and lysine residues of domains II and V and on the positively charged carboxy-terminus. The conserved aspartates of domains II and V were the first targets.

Marrero et al. (1992) separately changed the first and third aspartates in domain

V to glutamates. They found that substitution of the third aspartate for glutamate resulted in a mutant that was similar to the wild-type in kinetic properties. However, the mutant in which the first aspartate of domain V had been replaced by glutamate, while having a $K_{m[GPP]}$ similar to the wild-type, had a $K_{m[IPP]}$ 23-fold that of the wild-type and a V_{max} 1/90th that of the wild-type enzyme. They concluded that according to the ordered-sequential mechanism previously discussed (Laskovics and Poulter, 1981; Laskovics et al., 1979) it was possible that either the rate of the reaction or the rate of product release were being affected. They suggested that both were possible, but that it was more likely that it was the rate of the reaction that was being reduced and that it had become the rate-limiting step. This would indicate that the first aspartate in domain V is important for catalysis. Marrero et al. (1992) went on to conclude from their data and from the conservation of domain V among a wide variety of other types of prenyltransferases, including those transferring a prenyl group to a nonprenyl acceptor molecule, that domain V is likely to be the allylic binding site.

Joly and Edwards (1993) separately replaced all three of the aspartates of domain II (D103E, D104E, D107E) and the second aspartate of domain V (D244E) with glutamate. They also replaced the two arginines of domain II (R112K, R113K) and the arginine that was highly modified (R192K) by photoaffinity labelling (Brems and Rilling, 1981; Brems et al., 1979) by lysine. They found that their mutants fell into two categories: (1) those which had kinetic properties similar to the wild-type enzyme, D103E and R192K, and (2) those which had K_m values for both IPP and GPP similar to the wild-type enzyme, but V_{max} values approximately 1/1000th that of wild-type, D104E,

D107E, R112K, and R113K. An exception was D244E, which had K_m values similar to wild-type and only a 7-fold reduction in V_{max}. They proposed a similar explanation to that of Marrero et al. (1992) in that D104E, D107E, R112K, and R113K all have lower values of V_{max} due to a decrease in the rate constant for prenyl transfer and therefore these residues must play important roles in catalysis. They also suggested that as in the case of the third aspartate in domain V (Marrero et al., 1992), the second aspartate (D244E) in domain V does not play as important a role in catalysis as the other conserved aspartate residues.

The results of mutagenesis of the two conserved arginines (R112K and R113K) in domain II suggesting that these two residues are important in catalysis in in agreement with previous work with phenylglyoxal. Barnard and Popjak (1980) showed that pig liver FPS was inactivated by phenylglyoxal, which is a arginine modifying agent. They also found that the inactivation was biphasic and proposed that this was due to inactivation of two important arginine residues in each active site. They suggested that these two arginines might be important in binding. The results of Joly and Edwards (1993) show that they are important catalytic residues.

Joly and Edwards (1993) also mutated the proposed "active site" arginine (R192K) that had been extensively labelled in photoaffinity labelling (Brems and Rilling, 1981; Brems et al., 1979) in order to determine if this residue was involved in binding or catalysis. The only effect of the R192K mutation was a 65% decrease in V_{max}. This evidence provides further support to the theory that the two arginine residues modified by phenylglyoxal (Barnard and Popjak, 1980) were those of domain II.

The results of Marrero et al. (1992) and of Joly and Edwards (1993) show that many of the conserved aspartate and arginine residues of domains II and V play crucial roles in the catalysis of this enzyme.

Following the conservative site-directed mutagenesis of domains II and V (Marrero et al., 1992; Joly and Edwards, 1993), Song and Poulter (1994) made nonconservative mutants of the conserved aspartates of domain II and V in the yeast FPS by replacing the aspartates with alanine (D100A, D101A, D104A, D240A, D241A and D244A). They also replaced the two conserved arginines of domain II with glutamine (R109Q and R110Q). They found that all mutants had activities ranging from 10^{-5} to 10^{-7} that of the wild-type enzyme, except for the last aspartate in domain V (D244A). The more drastic inactivation caused by these mutations as compared to the conservative mutations of Marrero et al. (1992) and Joly and Edwards (1993) indicates the important functional groups are the carboxylate moieties and that there is sufficient flexibility to allow some activity when aspartate is replaced by glutamate.

Song and Poulter (1994) also investigated the importance of the charged carboxy-terminus of FPS. They added an EEF epitope to facilitate purification of the mutant enzymes and compared the activity of the EEF-tagged enzyme to the wild type. They found that activity was reduced 12-fold and $K_{m[IPP]}$ was increased 14-fold, while $K_{m[GPP]}$ was comparable to the wild-type enzyme. They suggested that, following the proposal of Laskovics and Poulter (1981), that there is a conformational change on binding, the carboxy terminus may be part of a flap that binds to IPP and moves to seal the active site from water.

Koyama et al. (1994a) also noted that many prenyltransferases have a positively charged carboxy-terminus. They made three separate mutants of the three carboxy-terminal residues of FPS from *B. stearothermophilus* (R295V, D296G, and H297L). They found that the activities of all three mutants were comparable to the wild-type enzyme, although the $K_{m[IPP]}$ values were approximately double the wild type. They postulated that the carboxy terminus may either be located near the IPP binding site or in a location that allows it to affect IPP binding.

Koyama et al. (1994b) did further site-directed mutagenesis on FPS from *B. stearothermophilus* to investigate whether cysteine residues are involved in catalysis. This had previously been suggested by other workers from observations that FPS could be inactivated by p-(chloromercuri)benzoic acid (PCMB), which is a thiol-specific reagent (Holloway and Popjak, 1967; Popjak, 1971). Also, many groups had reported two forms of FPS and suggested that this might be due to oxidation or reduction of cysteine residues (Koyama et al., 1977; Yeh and Rilling, 1977; Barnard et al , 1978). There are two cysteines in *B. stearothermophilus* FPS, one at position 73, which is conserved among other FPS, and one at position 289, which is not conserved. They separately mutated both of these cysteines to serine and to phenylalanine (C73S, C73F, C289S and C289F) and made a double mutant replacing both cysteines with serine (C73S-C289S). All five of the mutants had activity comparable to the wild-type enzyme, which indicated that neither of these cysteines were involved in catalysis. However, the *B. stearothermophilus* FPS is a thermostable enzyme and these researchers noted that the C73F mutation had the effect of reducing the thermostability of the enzyme by 15°C.

They also observed that the C289F mutant had a $K_{m[IPP]}$ 10-fold greater than the wild-type. They suggested that the conserved C73 is likely involved in stabilization of the enzyme and that C289 may be close to the IPP binding site. This leads to the conclusion that the inactivation of FPS by PCMB occurs by introduction of a bulky group rather than a loss of an active-site cysteine. Koyama et al. (1994b) observed that the double mutant (C73S-C289S) was also present in two peaks when separated by anion exchange chromatography, as had previously been observed for FPS (Koyama et al., 1977; Yeh and Rilling, 1977; Barnard et al., 1978). They also noted that the double mutant could be Cu^{II}-oxidized and reduced (thiol reducing agents) to give bands at 31kDa and 32,5kDa, respectively on SDS-PAGE. Therefore, the two forms of FPS do not result from oxidation and reduction of cysteines.

Much of the site directed mutagenesis focused on the active site of FPS has dealt with the conserved aspartates and asparagines involved in catalysis. However, in view of the substrate analogue work which determined the structural requirements of the alkyl moieties of the substrates (Nishino et al., 1970; Nishino et al., 1972; Nishino et al., 1973; Ogura et al., 1970; Ogura et al., 1974; Shinka et al., 1975; Popjak et al., 1969), Koyama et al. (1995) hypothesized that some hydrophobic amino acids would be required to ensure that the substrates are bound productively. They mutagenized the conserved phenylalanine-glutamine motif within domain V. They separately made nonconservative mutants of both of these residues (F220A and Q221E) and found that F220A and Q221E had 10^{-5} and 10^{-3} the activity of the wild-type enzyme. The two mutants had $K_{m[IPP]}$ and $K_{m[GPP]}$ similar to that of the wild type enzyme, but the $K_{m[DMAPP]}$ of F220A was

approximately 25-fold larger and that of Q221E was slightly larger than the wild-type enzyme. Although, an increase in the $K_{m[IPP]}$ following site-directed mutagenesis has been observed by other workers (Koyama et al., 1994; Koyama et al., 1994b; Song and Poulter, 1994), this was the first demonstration of an increase in the K_m of one of the allylic substrates. Work with substrate analogues, described above, (Ogura et al., 1970) had shown that a methyl group at the 3'-position of the allylic substrate was required for activity. Koyama et al. (1995) suggest that the phenylalanine of the FQ motif interacts with the methyl group at the 3'-position and that the glutamine increases the binding affinity by hydrogen bonding with the oxygen at the 1'- position of the allylic substrate. Therefore, the FQ motif is involved in maintaining productive binding of the allylic substrate by hydrophobic and hydrogen bonding interactions. It is possible that the $K_{m[GPP]}$ does not change, while that of DMAPP does, because the affinity of GPP is dominated by interactions with its longer hydrocarbon tail, thus allowing it to be bound productively to both of the mutants.

Blanchard and Karst (1993) isolated a mutant of yeast FPS which excretes the ten carbon prenyl alcohols geraniol and linalool into the media. These two compounds are normally not synthesized by yeast and the mutant was found to have a single base pair substitution in the FPS gene converting a lysine to a glutamate. This lysine is one of the conserved residues of domain IV and is in the region of the avian FPS enzyme which was photoaffinity labelled with o-azidophenethyl pyrophosphate (Brems and Rilling, 1981; Brems et al., 1979). This mutation suppresses the elongation pr[illegible]s at the C10 to C15 step and changes the GPP and FPP product ratio of the enzyme from 25:75 for

the wild type enzyme to 70:30 for the mutant.

A.8. Aim of work

The aims of this work were two-fold: (1) cloning of FPS and GGPS from *Lupinus albus*, and (2) examination of the structure-function relationship between these two enzymes with regard to product specificity. The similarity in their secondary structures as well as in the reactions that FPS and GGPS catalyze make them amenable to this type of study. Investigation of the this relationship involved secondary structure prediction, creation of truncated FPS clones and chimeric GGPS/FPS clones, and GGPS site-directed mutants.

B. MATERIALS AND METHODS

B.1 Chemicals

The Erase-A-Base DNA deletion kit was obtained from Promega (Madison, WI); the Sequenase Version 1.0 sequencing kit from USB (Cleveland, OH.); Western blotting kit from BioRad (Mississauga, ONT); pTrc-His Xpress expression vectors from Invitrogen (San Diego, CA); Ni-NTA nickel affinity resin from Qiagen (Chatsworth, CA); nylon membranes for library screening and colony hybridization from Amersham Canada (Oakville, ONT); Exassist helper phage from Stratagene (La Jolla, Ca.); prenyl pyrophosphates from American Radiolabeled Chemicals, Inc. (St. Louis, MO); prenyl alcohol standards from Fluka (New York, NY); nitrocellulose membrane for Western blotting from BioRad (Mississauga, ONT); DNA modifying enzymes (ligase, phosphorylase, and Klenow fragment of DNA polymerase) and restriction endonucleases were from BioCan (Mississauga, Ont.). pfu DNA polymerase was from Stratagene (La Jolla, Ca.).

B.2 Molecular biology

B.2.1 Library screening and sequencing

A Lambda phage cDNA library constructed from mRNA of 10-day *Lupinus albus* seedling roots was screened. 4 X 10^5 plaques were screened using cDNA probes corresponding to internal portions of these genes. The probes were generated by a polymerase chain reaction (PCR) strategy (Attucci et al., 1995). Positive plaques were purified by two additional rounds of plating and hybridization. Plasmids containing the

cDNA clones were isolated from phage clones by *in vivo* excision of the Bluescript plasmid using Exassist helper phage (Stratagene).

One full-length GGPS (*pGGPS1*) and three full-length FPS (*pFPS1* and *pFPS2*, and pFPS3) clones were identified by partial DNA sequencing and selected for further characterization. A series of nested deletions were generated for each of the three clones using the Erase-A-Base kit (Promega). Clones were sequenced using the Sequenase version 1.0 kit (USB).

B.2.2 Polymerase chain reaction (PCR)

The cDNA clones were subcloned into the pTrc-His expression vector (Invitrogen). Primers were designed to create restriction sites in the 5' end of the clone to facilitate subcloning in the proper reading frame. PCR primers were also designed to modify the 3' end of the coding sequence of the clones. These were used to either introduce stop codons to eliminate portions of the carboxyl end of the protein or to facilitate the creation of FPS/GGPS chimeric clones. Care was taken that the amino acid changes resulting from the introduction of restriction sites should be conservative and limited in number, in order to minimize changes to the coding sequence.

Reactions were carried out in 100 μL volumes and contained 10 ng of Bluescript plasmid DNA containing the cDNA clone, 1 μM of each primer, 200 μM of each dNTP, 2.5 units of *pfu* DNA polymerase (Stratagene), and reaction buffer (containing 20 mM Tris-HCl (pH 8.2), 10 mM KCl, 6 mM $(NH_4)_2SO_4$, 1.5-2.0 mM $MgCl_2$, and 0.1% Triton X-100). Reactions were overlayed with 100 μL of mineral oil.

A thermocycle apparatus (Interscience, Ont.) was used to carry out PCR amplification. The following PCR program was employed: initial denaturation at 95°C (4 min), followed by 30 amplification cycles of denaturation, 95°C (1 min), annealing, 55°C (1 min), extension, 72°C (2 min), followed by a final extension step of 72°C (5 min).

Aliquots of PCR reactions were analyzed by agarose gel electrophoresis. Reaction products were phenol/chloroform extracted and ethanol precipitated.

pGK201E and pGK201R site-directed mutant clones were obtained from Dr. J. Saleeba in the pBluescript vector.

Table 3. Oligonucleotide primers used in PCR amplification

Primer	Position	Restriction site Introduced	Constructs
5F3	67-87	BamHI	pF3, pF3A3I,pF3IJ
5G1	133-157	BclI	pG1, pG1F3HA1, pG1F3A3IA, pG1F3A3IB
F3HA1	777-806	KpnI	pG1F3HA1
F3IJ2	1033-1062	XhoI	pF3IJ, pG1F3A3IA
F3A3I	947-976	XhoI	pG1F3A3IA, pG1F3A3IB
F3A3I2	951-984	XhoI	pF3A3I
G1HA1	868-899	KpnI	pG1F3HA1
G1A3I	1051-1083	XhoI	pG1F3A3IA, pG1F3A3IB

B.2.3 Subcloning of cDNA inserts into expression vector pTrc-His (A,B,C)

Purified PCR products were digested with the appropriate restriction enzymes (table 3), phenol/chloroform extracted, and ethanol precipitated. pTrc-His plasmid vector was digested with compatible restriction enzymes to carry out forced orientation

Table 4. Nucleotide and Amino Acid Sequences of Oligonucleotide primers

Primer	Nucleotide Sequence	Original Amino Acid Sequence Modified Amino Acid Sequence
I. primers introducing restrictions sites at the 5' end of the gene		
5F3	5'-gtgtaaatggcggatccaagg	V*MADLR V*MADPR
5G1	5'-gcaatgatcactaaagaagataccg	AMLTKEDT AMITKEDT
II. primers introducing restriction sites at the 3' end of the gene		
F3HA1	5'-ggattgctttggggtacctgaaacaattgg	DCFGAPETIG DCFGVPETIG
F3A3I	5'-cgctgtatgacgagctcgagcttcagggtg	ALYDELNLQGV ALYDELELQGV
G1HA1	5'-cccaattcctttggtaccttagtaacatcaag	LDVTKSSKELG LDVTKVPKELG
G1A3I	5'-atttgggctatactcgaggtaattagctaatgc	ALANYIAYSPN ALANYLEYSPN
III. primers introducing stop codons into the 3' end of the gene		
F3A3I2	5'-cgtaaatacaccctcgagattaagctagtcatac	YDELNLQGVFTE YD*LNLEGVFTE
F3IJ2	5'-caatagagctcgagctgcttagctaggatg	HPSKAVQALL HPS*AARALL
IV. primers for site-directed mutagenesis at K201 of pGGPS1		
GK201E	5'-tctagaagtgctgctgtttcatgaagatgaat	IHLHKTAALLE IHLHETAALLE
GK201R	5'-cttctagaagtgctgctgttctatgaagatgaa	IHLHKTAALLE IHLHRTAALLE

cloning of the insert. The pTrc-His B vector was used for the pFPS3 clone and truncated clones derived from it. The pTrc-His C vector was used for the pGGPS1 clone, and the chimeric and site-directed mutant clones derived from it, and the site-directed pGGPS1 mutants. Ligations were carried out in a 10 μL volume containing 1.5 μL of digested vector (50 ng/μL) and 2.5 μL of digested PCR product (50 ng/μL). Ligation products were transformed into Top10 cells using heat shock.

Transformants were streaked onto gridded LB plates for screening by colony hybridization and grown overnight at 37°C on ampicillin. Colonies were transferred to nylon membranes by overlaying the plate with the membrane for one minute. Membranes were transferred to a bed of 3MM Whatman paper in denaturing solution (0.5 M NaOH, 1,5 M NaCl) for 7 minutes followed by transfer to neutralizing solution (0.5 M Tris-HCl, pH 7.2, 1.5 M NaCl) for 3 minutes. The neutralizing step was repeated with fresh solution and the membranes were briefly washed with 2X SSC. Membranes were subsequently air-dried and baked at 80°C for 2 hours.

Membranes were prehybridized 4 to 6 hours at 65°C in hybridization buffer (5X Denhardts, 5X SSPE, 0.5% SDS, X salmon sperm DNA) followed by hybridization at the same temperature in fresh hybridization buffer, containing ^{32}P-labeled cDNA inserts obtained from clones in pBluescript vectors. Membranes were washed twice at room temperature (1X SSC, 0.1% SDS) and twice at 65°C for 15 minutes per wash and exposed overnight to Fuji RX film with an intensifying screen at -80°C. Proper orientation and reading frame of the clones in pTrc-His was confirmed by DNA sequencing of the ligation junctions.

B.3 Computer Analysis

Sequence alignments, nucleotide sequence fragment assembly, open reading frame (ORF) analysis, restriction site mapping, and protein secondary structure predictions (Chou-Fasman (CF) and Garnier-Osguthorpe-Robson (GOR)) were carried out using the Sequence Software package of the Genetic Computer Group. DNA and protein sequence similarity searches were performed at NCBI using BLAST N, BLAST X, and BLAST P. Protein secondary structure prediction was also done using the Protein-predict method (Rost and Sander, 1994; Rost et. al., unpublished).

B.4 Enzymology

B.4.1 Prenyltransferase assay and product analysis

Prenyltransferase enzyme assays were carried out in a 100 μL reaction mixture containing 50 mM Tris-HCl pH 7.6, 2 mM DTT, 0.2 mM $MnCl_2$, 10 mM iodoacetamide, 50 mM NaF, 20 μM/0.05 μCi [^{14}C]IPP (55 mCi/mmol), and 20 μM DMAPP, GPP, or FPP (American Radiolabeled Chemicals). Reactions were incubated for 30 minutes at 37°C, stopped by addition of 300 μL of 4 N HCL in methanol, and further incubated at 37° for 15 minutes. Reaction products were extracted with 400 μL of chloroform and quantified by scintillation counting in 5 mL of Ecolite scintillation cocktail (ICN).

For product analysis separate reactions were carried out under the same conditions, except alkaline phosphatase treatment was used in place of acid hydrolysis. The reaction mixture was adjusted to pH 9.0 with NaOH, 10 units of calf intestinal

alkaline phosphatase (BioCan Scientific) was added and the mixture was incubated for 3 hours at 37°C with additional incubation overnight at room temperature. The alcohol products were extracted with 500 μL of chloroform. The chloroform phase was back extracted with 500 μL of water and then evaporated. The residue was resuspended in 20μL of methanol containing the standard alcohols and chromatographed on TLC silica gel F_{254} plates (Merck) with benzene/ethyl acetate (9:1). The TLC plate was then air-dried for 30 minutes and exposed for 2 days to a CS filter on the phosphoimaging apparatus (BioRad). The alcohol standards (dimethylallyl alcohol, E-geraniol, E,E-farnesol, and E,E,E-geranylgeraniol) (Fluka) were visualized under UV light following exposure to iodine vapour.

B.4.2 Protein expression and purification

The pFPS3 protein was soluble when expressed in *E. coli* culture and was characterized differently than other proteins. Due to the possible formation of inclusion bodies by some of the other clones, two purification protocols were followed, depending on whether the overexpressed protein was soluble or in inclusion bodies. For purification of the pFPS3 protein, 10 mL of LB medium containing 50 μg/mL ampicillin was inoculated with a single colony containing pFPS3 and grown overnight with shaking at 37°C. Prewarmed LB medium (1L), containing 50 μg/mL ampicillin was inoculated with 10 mL of overnight culture, grown for 3 hours at 37°C, induced with 1 mM IPTG and incubation continued for 4 hours. Cells were collected by centrifugation for 15 minutes at 4000Xg and pellets were stored at -80°C. The pellets were resuspended in 25 mL of

native lysis buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 14 mM β-mercaptoethanol, 5 mM PMSF, and 2 μg/mL leupeptin) containing 1 mg/mL lysozyme and incubated on ice for 30 minutes, followed by 3 rounds of freeze-thaw and sonication. The lysate was then treated with 10 μg/mL each of DNase and RNase (BioCan), incubated on ice for 15 minutes, and centrifuged at 10000Xg for 20 minutes to remove cell debris. The supernatant was mixed with 8 mL of a 50% slurry of previously equilibrated Ni-NTA nickel resin (Qiagen) and rotated overnight at 4°C. The suspension was loaded into a 1.6 cm diameter column (BioRad) and washed with 100 mL of pH 8.0 native loading buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 14 mM β-mercaptoethanol), followed by 100 mL of pH 6.0 native wash buffer (50 mM sodium phosphate pH 6.0, 300 mM NaCl, 14 mM β-mercaptoethanol). Protein was eluted with a 80 mL linear gradient of 0 to 500 mM imidazole in pH 6 wash buffer. Fractions were analyzed by SDS-PAGE and pooled. Buffer was exchanged for DEAE loading buffer (50 mM Tris-HCl pH 7.6, 14 mM β-mercaptoethanol) with a PD10 desalting column (BioRad). The eluent was applied to a DEAE column. The column was washed with at least 500 mL of DEAE loading buffer and eluted with an 80 mL linear gradient of NaCl (0 to 500 mM) in DEAE loading buffer. The fractions eluting from the DEAE column were analyzed by SDS-PAGE.

For the other constructs, smaller scale preparations were made for comparative purposes. Fifty mL cultures were grown and expressed in the same manner as previously described for the expression of pFPS3. Cell pellets were resuspended and lysed in 2.0 mL of lysis buffer. After centrifugation of the lysate for 20 minutes at

10000Xg, the pelleted insoluble fraction was saved for later analysis and the supernatent was added to 0.5 mL of a 50% slurry of equilibrated Ni-NTA resin (Qiagen) and mixed by rotation for 1 hour at 4°C. Resin was loaded into a 1,6 cm diameter column (BioRad) and washed consecutively with 1.0 mL each of pH 8.0 native loading buffer, pH 6.0 native wash buffer, and pH 6.0 native wash buffer containing 50 mM imidazole. Protein was eluted in 1.0 mL of pH 6.0 native wash buffer containing 250 mM imidazole. These eluted fractions were used for comparative analysis by enzymatic assays, for product analysis, and for SDS-PAGE and immunoblotting.

The pelleted insoluble fraction of these 50 mL cultures were solubilized in 0.5 mL of denaturing buffer (6M guanidine hydrochloride, 100 μM sodium phosphate, 10 μM Tris-HCl, pH 8.0). Samples of the solubilized pellet of each construct were run on SDS-PAGE and immunoblotted to determine if the expressed protein was present in a insoluble form.

C.4.3 Immunoblotting

For both the soluble and the insoluble fractions two identical SDS-PAGE were run to determine if the protein was being expressed and whether it was present in inclusion bodies. One gel was Coomassie stained; the other gel was soaked in transfer buffer (192 mM Glycine, 25mM Tris pH 8.0, 20% methanol) then transferred to a nitrocellulose membrane (BioRad) in a BioRad semi-dry electroblotting apparatus for 30 minutes (15 V, 250 mA). The nitrocellulose membrane was stained with Ponceau red stain for 2 minutes followed by washing with water. Molecular weight standards were marked and the membrane washed briefly in 1X TBS (20 mM Tris pH 7.5, 150 mM NaCl) to

remove the stain.

All subsequent procedures were done with shaking with the exception of the colour development step. The membrane was blocked with 2% defatted skim milk blocking agent (BioRad) in 1X TBS for 30 minutes. The membrane was incubated with a 1:250 dilution of the FPS antibody (Attucci et. al., 1995b) in 1X TBS for 1 hour. It was then washed three times for 10 minutes each, the first two times in 1X TBS (containing 0.5% Tween-20) and the third time in 1X TBS alone. Incubation with the secondary antibody (goat anti-rabbit conjugated to alkaline phosphatase, (BioRad) at a dilution 1:2000 (in 1X TBS) was for 30 minutes, then the membrane washed another three times as before. Colour development was done according to the manufacturer's instructions (BioRad, Western Blotting Kit), except that the volume was reduced from 100 mL to 20 mL. Colour development was stopped by washing with distilled water.

C. RESULTS

C.1 Molecular cloning

C.1.1 Isolation of *L. albus* FPS and GGPS cDNA clones

Approximately 4 X 10^5 plaques of the 10-day lupin seedling root cDNA library were screened separately with nucleic acid probes for FPS and GGPS, respectively. Two clones for GGPS and 7 for FPS were initially purified and characterized by restriction digestion and end sequencing. The two GGPS clones (pGGPS1 and pGGPS2) had identical restriction maps, sequence, were the same length, and had very high similarity to other GGPS clones in the GenBank data base. Three FPS clones had very high similarity to other FPS clones in the GenBank data base and were designated pFPS1, pFPS2, and pFPS3. These clones had similar sequence and length, except one, pFPS2, which was 50 bp longer than the other two. The other four clones selected in the FPS screening were unrelated to FPS and had inserts ranging from 900 to 1500 bp in length.

Clones containing GGPS and FPS inserts in the pBluescript vector were tested for activity. pFPS1 was the only clone in the proper reading frame and the only clone which exhibited prenyltransferase activity.

Partial sequencing indicated that pGGPS1 and pGGPS2 were identical and pGGPS1 was selected for full sequencing. All three FPS clones were selected for further sequencing. However, it became apparent that pFPS2 was identical to pFPS3 with the exception of a 50 bp insert near the 5' end. This insert has a splice site consensus sequence on it's 5' end. Since this clone was otherwise identical to pFPS3 it was assumed to contain an unspliced intron and was not further sequenced.

C.1.2 Nucleotide sequence and deduced amino acid sequence

The complete nucleotide sequence was determined for both strands of pGGPS1 (Fig. 7), pFPS1 (Fig. 8), and pFPS3 (Fig. 9). pGGPS1 had very high sequence similarity to clones for GGPS in the GenBank data base. GGPS from the plant *Arabidopsis thaliana* was the most similar sequence, having 85% and 72% amino acid similarity and identity, respectively, to pGGPS1 from *Lupinus albus*. pFPS1 and pFPS3 also had high degrees of similarity to previously reported FPS sequences in GenBank. These include numerous sequences from plant, animal, yeast, and bacterial sources. The most similar FPS sequence was also from *Arabidopsis thaliana*, which had 91% and 81% similarity and identity, respectively in amino acid sequence to both pFPS1 and pFPS3 from *Lupinus albus*.

The presumed 5' UTR of the pGGPS1 clone (Fig. 7) is quite long and contains no stop codons in frame with the ORF. The initial methionine in Figure 7 was placed to align with other reported GGPS sequences. However, it is possible that the ORF extends upstream to the start of the clone, and that this region encodes a signal peptide. This region contains no potential start codons, and would encode 46 amino acids. It is also of interest that there is a stretch of glutamates in the ORF just after the start codon in this clone. In the *Arabidopsis thaliana* GGPS clone there is a stretch of histidines within its signal peptide. A frame shift of one bp could change a stretch of glutamates into a stretch of lysines, which is also a positively charged amino acid.

```
   1 CACAAAACCACTAGATTCAGATCTTCAAGTTTTCAAATCTTCCATAGCTTT
  52 ACAAAGTTACCCATTTCTTTTATAAGCATCGAAAAACCAAAAAGGTCACAA
 103 CTTTCATATTTTTCTTCTTTCTCTATCTCTGCAATGCTCACTAAAGAAGAT
                                       M  L  T  K  E  D
 154 ACCGTTAAAGACAAAGAAGAAGAAGAAGAAGAAGAAGAGAAACCCAATTCC
     T  V  K  D  K  E  E  E  E  E  E  E  E  K  P  R  F
 205 AATTTCAACTTGTACATGGTTGAGAAATCACGTTCAGTGAACCAAGCTTTA
     N  F  N  L  Y  M  V  E  K  S  R  S  V  N  Q  A  L
 256 AACGACGCCGTTTCGTTACGTGAGCCACATAAAATTCACGAAGCTATGCGG
     N  D  A  V  S  L  R  E  P  H  K  I  H  E  A  M  R
 307 TATTCTCTACTCGCCGGCGGGAAGCGCGTGAGGCCGGTGTTATGCATTGCA
     Y  S  L  L  A  G  G  K  R  V  R  P  V  L  C  I  A
 358 GCCTGTGAAGTTGTCGGCGGCAATGAATCTACGGCGATGGCCGCCGCATGT
     A  C  E  V  V  G  G  N  E  S  T  A  M  A  A  A  C
 409 TCAATTGAGATGATACACACTATGTCCCTTATTCATGATGATCTCCCCTGC
     S  I  E  M  I  H  T  M  S  L  I  H  D  D  L  P  C
 460 ATGGACAACGACGATCTTCGCCGAGGGAAACCCACAAACCACAAGGTTTTC
     M  D  N  D  D  L  R  R  G  K  P  T  N  H  K  V  F
 511 GGCGAAAATATCGCTGTCCTTGCAGGGGATGCATTATTGGCATTTGCCTTT
     G  E  N  I  A  V  L  A  G  D  A  L  L  A  F  A  F
 562 GAGCATATCGCCGTTTCAACCTCCGGTGTGTCACCGGAGAGAATTATCGGC
     E  H  I  A  V  S  T  S  G  V  S  P  E  R  I  I  G
 613 GCAATCGGGGAGCTTGCGAAATCGATCGGAACTGAAGGCCTTGTTGCAGGA
     A  I  G  E  L  A  K  S  I  G  T  E  G  L  V  A  G
 664 CAAGTAGTGGATATTAATTCAGAAGGGTTATGTGATATTGGGTTAGAGAAA
     Q  V  V  D  I  N  S  E  G  L  C  D  I  G  L  E  K
 715 TTGGAATTTATTCATCTTCATAAAACAGCAGCACTTCTAGAAGGTTCTGTT
     L  E  F  I  H  L  H  K  T  A  A  L  L  E  G  S  V
 766 GTTGTTGGAGCAATTCTTGGTGGTGGGTGTAATGAAGAGGTTGAAAAGTTA
     V  V  G  A  I  L  G  G  G  C  N  E  E  V  E  K  L
 817 AGGATGTTTGCTAGGTACATAGGATTGATGTTTCAGGTTGTTGATGATGTT
     R  M  F  A  R  Y  I  G  L  M  F  Q  V  V  D  D  V
 868 CTTGATGTTACTAAGTCTTCAAAGGAATTGGGAAAAACTGCAGGGAAGGAT
     L  D  V  T  K  S  S  K  E  L  G  K  T  A  G  K  D
 919 TTGGTGGCTGATAAGGTTACTTACCCTAAGCTTTTAGGGATTGAGAAATCT
     L  V  A  D  K  V  T  Y  P  K  L  L  G  I  E  K  S
 970 AATGAGTTTGCTCAGAAATTGAATAGGGATGCACAAGAACAGCTTTCTGGT
     N  E  F  A  Q  K  L  N  R  D  A  Q  E  Q  L  S  G
1021 TTTGATCCAGTTAAGGTTGCTCCTTTGATTGCATTAGCTAATTACATTGCT
     F  D  P  V  K  V  A  P  L  I  A  L  A  N  Y  I  A
1072 TATAGCCCAAATTAAACTCTCTTGTGTAACAATTTGTTGTTTAGAAAAGGT
     Y  S  P  N  *
1123 TAAATTTATGGCAATTATTGAGTTTACTTATTACCAAAAAAAAAAAAAAAA
1174 A
```

Figure 7. Nucleotide and deduced amino acid sequence of pGGPS1.

```
   1 TTCTTCTCTCTCTCTCTCTCTGTTTTTCAAAGCATTGTCTTTTTTTGTT
  52 GTTTCAATGGCAGATCCCAAATCTAGTTTCTTGAATGTCTATTCCATTCTC
        M  A  D  P  K  S  S  F  L  N  V  Y  S  I  L
 103 AAATCTGAGCTCCTTCAAGACCCTGCTTTCGAATTCTCCACTGATTCTCGT
     K  S  E  L  L  Q  D  P  A  F  E  F  S  T  D  S  R
 154 CAATGGGTCGAAAGGATGCTGGACTACAATGTGCCTGGAGGAAAGCTGAAC
     Q  W  V  E  R  M  L  D  Y  N  V  P  G  G  K  L  N
 205 CGTGGACTGTCGGTTATTGACAGCTACAAATTGTTAAAAGATGGACAGGAA
     R  G  L  S  V  I  D  S  Y  K  L  L  K  D  G  Q  E
 256 TTAAATGATGAAGAAATTTTCCTTGCTAGTGCTCTTGGTTGGTGCATTGAA
     L  N  D  E  E  I  F  L  A  S  A  L  G  W  C  I  E
 307 TGGCTTCAGGCATATTTCCTTGTTCTTGATGACATCATGGATAACTCTCAC
     W  L  Q  A  Y  F  L  V  L  D  D  I  M  D  N  S  H
 358 ACACGGCGTGGTCACCCGTGCTGGTTTAGAGTACCCAAGGTTGGAATGATT
     T  R  R  G  H  P  C  W  F  R  V  P  K  V  G  M  I
 409 GCACCAAACGATGGGGTGGTTCTACGAAACCATATTCCTCGTATCCTTAAG
     A  P  N  D  G  V  V  L  R  N  H  I  P  R  I  L  K
 460 AAACACTTCAGGGGAAAGCCATACTATGTTGATCTGCTTGATCTGTTTAAT
     K  H  F  R  G  K  P  Y  Y  V  D  L  L  D  L  F  N
 511 GAGGTCGAGTTTCAGACTGCTTCAGGACAGATGATAGATCTGATCACTACA
     E  V  E  F  Q  T  A  S  G  Q  M  I  D  L  I  T  T
 662 CTGGAAGGAGAAAAAGATCTGTCTAAATACACATTATCACTGCACCGTCGT
     L  E  G  E  K  D  L  S  K  Y  T  L  S  L  H  R  R
 613 ATTGTTCAGTACAAGACTGCCTATTATTCATTTTACCTCCCAGTTGCATGT
     I  V  Q  Y  K  T  A  Y  Y  S  F  Y  L  P  V  A  C
 764 GCATTGCTCATGGTGGGTGAGAATCTTGATAACCATACTGATGTGAAAAAC
     A  L  L  M  V  G  E  N  L  D  N  H  T  D  V  K  N
 715 ATTCTTGTTGAGATGGGAACATACTTTCAAGTGCAGGATGATTATTTGGAT
     I  L  V  E  M  G  T  Y  F  Q  V  Q  D  D  Y  L  D
 866 TGCTTTGGTGCTCCTGAAACAATTGGAAAGATAGGTACAGATATTGAAGAT
     C  F  G  A  P  E  T  I  G  K  I  G  T  D  I  E  D
 817 TTTAAGTGCTCTTGGTTGGTTGTGAAAGCATTGGAACTTAGCAATGAAGAA
     F  K  C  S  W  L  V  V  K  A  L  E  L  S  N  E  E
 968 CAAAAGAAAGTTTTATATGAGAATTATGGGAAGCCAGATCCTGCAAATGTT
     Q  K  K  V  L  Y  E  N  Y  G  K  P  D  P  A  N  V
 919 GCTAAAGTGAAGACCCTATATAATGAGCTAAATCTTGAGGGTGCGTATGCG
     A  K  V  K  T  L  Y  N  E  L  N  L  E  G  A  Y  A
 970 GATTACGAGAGCAAGAGCTATGAGAAACTTGTAACCTGCATTGAAGGTCAT
     D  Y  E  S  K  S  Y  E  K  L  V  T  C  I  E  G  H
1021 CCTAGCAAAGCAGTTCAAGGTGTATTGAAGTCGTTTTGGGCTAAAATTTAC
     P  S  K  A  V  Q  G  V  L  K  S  F  W  A  K  I  Y
1072 AAGAGGCAGAAGTAGAGTAACTCCAAGGCCTAAATGGTTGAAGAAGCATTA
     K  R  Q  K  *
1123 GAATCTAGAGTCCTGGTTTGGTTTTTATCTAGAGTTCTATCATGTTTCATG
1174 ACTTTATCTTTTTTTTACTTGTGGGTAATCTCTATTAGTCTTGTTGTATTGT
1225 TTTTTCTTTTTTTACTTAAGATGTTGTATTTTATGTTGTAAACTTCTTCTCT
1276 TACTTGCTTGCCTTAGTAATTCTAATGATCTTTCTGATCTTTGTTAAAAAA
1327 AAAAAAAAAAAAAAAAA
```

Figure 8. Nucleotide and deduced amino acid sequence of pFPS1.

```
   1 TACATATACATATATAAGAAATGCAATAACCATATTAGTTTGTGTTGTAAC
  52 GTTGAATCTGTTTGTGTGTAAATGGCAGATCTAAGGTCAACTTTCTTGAAT
                          M  A  D  L  R  S  T  F  L  N
 103 GTGTATTCGGTTCTGAAATCAGAGCTCCTTCATGACCCAGCTTTTGAATTC
     V  Y  S  V  L  K  S  E  L  L  H  D  P  A  F  E  F
 154 TCTCCTGATTCTCGTCAATGGCTCGACCGGATGCTGGACTACAATGTGCCT
     S  P  D  S  R  Q  W  L  D  R  M  L  D  Y  N  V  P
 205 GGAGGAAAGCTAAACCGTGGACTGTCAGTTATTGACAGCTACAGATTGTTG
     G  G  K  L  N  R  G  L  S  V  I  D  S  Y  R  L  L
 256 AAAGATGGACATGAATTAAACGATGATGAAATTTTTCTTGCTAGTGCTCTT
     K  D  G  H  E  L  N  D  D  E  I  F  L  A  S  A  L
 307 GGTTGGTGTATTGAATGGCTTCAGGCATATTTCCTTGTTCTTGATGACATT
     G  W  C  I  E  W  L  Q  A  Y  F  L  V  L  D  D  I
 358 ATGGATAACTCCCACACACGGCGTGGTCAGCCATGTTGGTTCAGAGTACCC
     M  D  N  S  H  T  R  R  G  Q  P  C  W  F  R  V  P
 409 AAGGTTGGAATGATTGCAGCAAATGATGGGGTGCTGCTACGGAACCATATT
     K  V  G  M  I  A  A  N  D  G  V  L  L  R  N  H  I
 460 CCTCGTATCCTTAAGAAACACTTCAGGGGAAAACCTTATTATGCTGATCTT
     P  R  I  L  K  K  H  F  R  G  K  P  Y  Y  A  D  L
 511 CTTGATCTGTTTAATGAGGTTGAGTTTCAAACTGCTTCAGGGCAGATGATA
     L  D  L  F  N  E  V  E  F  Q  T  A  S  G  Q  M  I
 562 GATCTGATCACTACACTGGAAGGAGAAAAAGACCTGTCTAAATACACATTA
     D  L  I  T  T  L  E  G  E  K  D  L  S  K  Y  T  L
 613 TCACTACACCGCCGTATTGTTCAGTACAAGACTGCCTATTATTCGTTTTAC
     S  L  H  R  R  I  V  Q  Y  K  T  A  Y  Y  S  F  Y
 664 CTCCCAGTCGCATGTGCATTACTCATGGTGGGTGAGAATCTTGATAACCAT
     L  P  V  A  C  A  L  L  M  V  G  E  N  L  D  N  H
 715 ATTGACGTGAAAAACATTCTTGTTGATATGGGAACGTACTTTCAAGTACAG
     I  D  V  K  N  I  L  V  D  M  G  T  Y  F  Q  V  Q
 866 GATGATTATTTGGATTGCTTTGGTGCTCCTGAAACAATTGGAAAGATAGGT
     D  D  Y  L  D  C  F  G  A  P  E  T  I  G  K  I  G
 817 ACAGATATTGAAGATTTTAAGTGCTCTTGGTTGGTCGTGAAAGCATTGGAA
     T  D  I  E  D  F  K  C  S  W  L  V  V  K  A  L  E
 868 CTTAGCAATGATGAACAGAAGAAAGTTTTATATGATAACTATGGGAAACCA
     L  S  N  D  E  Q  K  K  V  L  Y  D  N  Y  G  K  P
 919 GATCCAGCAAATGTTGCTAAAGTGAAGGCGCTGTATGACGAGCTTAATCTT
     D  P  A  N  V  A  K  V  K  A  L  Y  D  E  L  N  L
 970 CAGGGTGTATTTACGGAGTATGAGAGCAAGAGTTATGAGAAGCTTGTAACC
     Q  G  V  I  T  E  Y  E  S  K  S  Y  E  K  L  V  T
1021 TCCATTGAAGCTCATCCTAGCAAAGCAGTTCAAGCTCTATTGAAGTCCTTT
     S  I  E  A  H  P  S  K  A  V  Q  A  L  L  K  S  F
1072 TTGGGTAAAATTTACAAGAGGCAGAAATAGAGTAACTGTAAGGTGAAGACC
     L  G  K  I  Y  K  R  Q  K  *
1123 AAAATGGTTGAAGATTCCTTAGAATCTAGATCCTTGATTTCTATCTAGAGC
1175 TGTGTCTTTTTTATAAAAATATTGGTCTTGTATTGTTTTTCTTATTACTTT
1225 TGGGTGATAAGATGTTTGTATTTTATCTTGTAAACTTTCTGGTGATACCTG
1277 TTCTGCTATTATAGTAACTATTTTCTTGGAAAAAAAAAAAAAAAAA
```

Figure 9. Nucleotide and deduced amino acid sequence of pFPS3.

```
pFPS1   MADLRSTFLNVYSVLKSELLHDPAFEFSPDSRQWLDRMLDYNVPGGKLNR 50
        ||| :|.||||||:||||||:|||||||.||||::|||||||||||||||
pFPS3   MADPKSSFLNVYSILKSELLQDPAFEFSTDSRQWVERMLDYNVPGGKLNR 50

pFPS1   GLSVIDSYRLLKDGHELNDDEIFLASALGWCIEWLQAYFLVLDDIMDNSH100
        ||||||||:|||||:||||:||||||||||||||||||||||||||||||
pFPS3   GLSVIDSYKLLKDGQELNDEEIFLASALGWCIEWLQAYFLVLDDIMDNSH100

pFPS1   TRRGQPCWFRVPKVGMIAANDGVLLRNHIPRILKKHFRGKPYYADLLDLF150
        ||||:|||||||||||||:||||:||||||||||||||||.||||||
pFPS3   TRRGHPCWFRVPKVGMIAPNDGVVLRNHIPRILKKHFRGKPYYVDLLDLF150

pFPS1   NEVEFQTASGQMIDLITTLEGEKDLSKYTLSLHRRIVQYKTAYYSFYLPV200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
pFPS3   NEVEFQTASGQMIDLITTLEGEKDLSKYTLSLHRRIVQYKTAYYSFYLPV200

pFPS1   ACALLMVGENLDNHIDVKNILVDMGTYFQVQDDYLDCFGAPETIGKIGTD250
        ||||||||||||||.|||||||:||||||||||||||||||||||||||
pFPS3   ACALLMVGENLDNHTDVKNILVEMGTYFQVQDDYLDCFGAPETIGKIGTD250

pFPS1   IEDFKCSWLVVKALELSNDEQKKVLYDNYGKPDPANVAKVKALYDELNLQ300
        |||||||||||||||||:|||||||:||||||||||||.||:|||||:
pFPS3   IEDFKCSWLVVKALELSNEEQKKVLYENYGKPDPANVAKVKTLYNELNLE300

pFPS1   GVFTEYESKSYEKLVTSIEAHPSKAVQALLKSFLGKIYKRQK 342
        |.:.:|||||||||||:||:||||||::||||::|||||||
pFPS3   GAYADYESKSYEKLVTCIEGHPSKAVQGVLKSFWAKIYKRQK 342
```

Figure 10. Alignment of the deduced amino acid sequences of pFPS1 and pFPS3. The alignment was made using the GCG package GAP program. (|) indicates the presence of an identical residue in both sequences, (:) represents a conserved residue, and (.) represents a moderately conserved residue.

```
pFPS1   ADPKSSF.LNVYSILKSELLQDPAFE.FSTDSRQWVERMLDYN.VPGGKLN  49
pGGPS1  EEEKPRFNFNLYMVEKSRSVNQALNDAVSLREPHKIHEAMRYSLLAGGKRV  67
pFPS3   ..ADLRSTFLNVYSVLKSELLHDPAFEFSPDSRQWLDRMLDYN.VPGGKLN  49

pFPS1   RGLSVIDSYKLLKDGQELNDEEIFLASALGWCIEWLQAYFLVLDDI..MDN  98
pGGPS1  RPVLCIAACEVVGGNES.......TAMAAACSIEMIHTMSLIHDDLPCMDN 111
pFPS3   RGLSVIDSYRLLKDGHELNDDEIFLASALGWCIEWLQAYFLVLDDI..MDN  98

pFPS1   SHTRRGHPCWFRV..PKVGMIAPNDGVVLRNHIPRILKKHFRGKPYYVDLL 147
pGGPS1  DDLRRGKPTNHKVFGENIAVLAGDALLAFAFEHIAVSTSGVSPERIIGAIG 162
pFPS3   SHTRRGQPCWFRV..PKVGMIAANDGVLLRNHIPRILKKHFRGKPYYADLL 147

pFPS1   DLFNEVEFQT.ASGQMIDLITTLEGEKDLSKYTLSLHRRIVQYKTAYYSFY 197
pGGPS1  ELAKSIGTEGLVAGQVVDINSEGLCDIGLEKLEF.....IHLHKTAAL.LE 207
pFPS3   DLFNEVEFQT.ASGQMIDLITTLEGEKDLSKYTLSLHRRIVQYKTAYYSFY 197

pFPS1   LPVACALLMVGENLDNHTDVKNILVEMGTYFQVQDDYLDCFGAPETIGKI. 247
pGGPS1  GSVVVGAILGGGCNEEVEKLRMFARYIGLMFQVVDDVLDVTKSSKELGKTA 258
pFPS3   LPVACALLMVGENLDNHIDVKNILVDMGTYFQVQDDYLDCFGAPETIGKI. 247

pFPS1   GTDIEDFKCSWLVVKALELSNEEQKKV...LYENYGKPDPANVAKVKTLYN 295
pGGPS1  GKDLVADKVTYPKLLGIEKSNEFAQKLNRDAQEQLSGFDPVKVAPLIALAN 309
pFPS3   GTDIEDFKCSWLVVKALELSNDEQKKV...LYDNYGKPDPANVAKVKALYD 295

pFPS1   ELNLEGAYADYESKSYEKLVTCIEGHPSKAVQGVLKSFWAKIYKRQK 342
pGGPS1  YIAYSPN..................................... 316
pFPS3   ELNLQGVFTEYESKSYEKLVTSIEAHPSKAVQALLKSFLGKIYKRQK 342
```

Figure. 11 Alignment of the deduced amino acid sequences of pFPS1, pGGPS1, and pFPS3. The alignment was made using the GCG package GAP program. (|) represents an identical residue, (:) represents a conservative residue, and (.) represents a moderately conservative residue. Gaps are introduced for optimization of alignment.

C.2 Sequence analysis and secondary structure prediction

C.2.1 Sequence Alignments

Alignment of amino acid sequences (Fig. 10) of pFPS1 and pFPS3 show 97% similarity and 90% identity. The genes are 85% identical in nucleic acid sequence. pGGPS1 had 45% similarity and 22% identity in amino acid sequence with both pFPS1 and pFPS3 (Fig. 11). The nucleic acid sequence of pGGPS1 was also aligned with both pFPS1 and pFPS3 and showed 38% and 39% identity with pFPS1 and pFPS3, respectively.

The five conserved regions (Chen et al., 1994) are conserved in both FPS clones and in GGPS1 and these are highlighted by the diamond regions in Fig. 12.

C.2.2 Phylogenetic analysis

The multiple sequence alignment (Fig. 12) shows that there are many conserved amino acids both among FPSs and the GGPSs as well as amino acids which are conserved between these two groups of enzymes. The alignment suggests that these genes have arisen from a common ancestral gene. The relative degree of sequence similarity among 20 clones for GGPS and FPS from a variety of species is shown in a dendrogram in Figure 13. While the FPS and GGPS genes sort into two groups it is of interest that there are two bacterial FPS genes that are found in the GGPS cluster.

```
                                                                                         GGK  R
 Fps7     1 .......... .......MNG DQNSDVYAQE KQDFVQHFSQ IVRVLTEDEM FHPEIGDAIA RLKEVLEYNA I♦♦♦Y+♦G+T
 Fps8     1 .......... .......MNG DQKLDVHNQE KQNFIQHFSQ IVKVLTEDEL GHPEKGDAIT RIKEVLEYNT V♦♦♦Y+♦G+T
Fps10     1 .......... .......... MLSPVVVERE REEFVGPPPQ IVRDLTEDGI GHPEVGDAVA RLKEVLQYNA P♦♦♦C+♦G+T
 Fps5     1 .......... .......... .......MAD LRSTFLNVYS VLKSELLHDP AFEFSPDSRQ WLDRMLDYNV P♦♦♦LN♦GLS
 Fps6     1 .......... .......... .......MAD PKSSFLNVYS ILKSELLQDP AFEFSTDSRQ WVERMLDYNV P♦♦♦L+♦G+S
 Fps3     1 .......... .......... ......METD LKSTFLNVYS VLKSDLLHDP SFEFTNESRL WVDRMLDYNV R♦♦♦L+♦G+S
 Fps9     1 .......... .........M AAGGNGAGGD TRAAFARIYK TLKEELLTDP AFEFTEESRQ WIDRMVDYNV L♦♦♦C+♦G+S
Fps11     1 .......... .......... .......... .......... .......... .......... ......NYNT P♦♦♦LN♦GLS
 Fps4     1 .......... .......... ......MSDN RAQFLEVFPS LVQELRDILA GYGMPEEAIE WYEKSLNYNT P♦♦♦L+♦G+S
Ggps6     1 .......... .......... ..MVSGSKAG VSPHREIEVM RQSIDDHLAG LLPETDSQDI VSLAMREGVM AP♦♦RI♦PLL
Ggps7     1 .......... .......... .MTVCAKKHV HLTRDAAEQL LADIDRRLDQ LLPVEGERDV VGAAMREGAL AP♦♦RI♦PM*
Ggps3    45 KRTVSSSSSI VSSSVVTKED NLRQSEPS.. SFDFMSYIIT KAELVNKALD SAVPLREPLK IHEAMSYSLL A♦♦♦RV♦PV*
Ggps5     1 .........M LTKEDTVKDK EEEEEEEEKP RFNFNLYMVE KSRSVNQALN DAVSLREPHK IHEAMRYSLL A♦♦♦RV♦PV*
Ggps4    45 NKKRPFSPLT VSAIATTKED ERIEAAQTEE PFNFKIYVTE KAISVNKALD EAIIVKEPHV IHEAMRYSLL A♦♦♦RV♦PM*
 Fps1     1 .......... .......... .......MAQ LSVEQFLNEQ KQAVETALSR YIERLEGPAK LKKAMAYSLE A♦♦♦RI♦P+*
 Fps2     1 .......... .......... .........M DFPQQLEACV KQANQALSRF IAPLPFQNTP VVETMQYGAL L♦♦♦RL♦PF*
Ggps1     1 .......... .......... .......... ....MAFEQR IEAAMAAAIA RGQGSEAPSK LATALDYAVT P♦♦ARI♦PT*
Ggps2     1 .......... .......... .......... ..MSYFNYFN EIVNSVNDII KSYISGDVPK LYEASYHLFT S♦♦♦RL♦PLI
Ggps8     1 ........MA LTLPDAQPPT GLLPLEQAWL QLVQTEVETS LAELFELPDE AGLDVRWTQA LTQARAYT.L RPA♦RL♦PA*
Ggps9    66 QAQTRKRKAS VAQISMLPTS FSPYTMAPQP PQPPPNPDRF ATEDFFSPSR RTWSEEKEKV LTGPYDYLNG HP♦♦DI♦SQM


                                          E      L   DD   MD      RRG
 Fps7    64 VVVAFRELVE ...PRKQDAD SLQRAWTVGW CV♦WLQAFf♦ VA♦♦I..♦♦S SLT♦♦♦QTCW YQKPGVGLDA INDANLLEAC
 Fps8    64 +VQTFQE+VE ...PRKQDAE SLQRALTV++ +V♦L+++F+♦ +L♦♦+..♦♦S +H+♦♦♦QI++ YQKPGI+LD+ I++ALLLEAA
Fps10    61 +VAAYRE+SG ...PGQKDAE SLRCALAV++ +I♦L+++F+♦ +A♦♦+..♦♦Q +L+♦♦♦QL++ YKKEGV+LD+ I++SPLLESS
 Fps5    54 +IDSYRL+KD ...GHELNDD EIFLASAL++ +I♦W+++Y+♦ +L♦♦+..♦♦N +H+♦♦♦QP++ FRVPKV+MI+ A++GVLLRNH
 Fps6    54 +IDSYKL+KD ...GQELNDE EIFLASAL++ +I♦W+++Y+♦ +L♦♦+..♦♦N +H+♦♦♦HP++ FRVPKV+MI+ P++GVVLRNH
 Fps3    55 +VDSFKL+KQ ...GNDLTEQ EVFLSCAL++ +I♦W+++Y+♦ +L♦♦+..♦♦N +V+♦♦♦QP++ FRVPQV+MV+ I++GILLRNH
 Fps9    62 +VDSYKL+KG ...ADALGEE ETFLACTL++ +I♦W+++F+♦ +L♦♦+..♦♦D +H+♦♦♦QP++ FRVPQV+LI+ A++GIILRNH
Fps11    15 +VDTYAI+SN .KTVEQLGQE EYEKVAIL++ +I♦L+++Y+♦ +A♦♦M..♦♦K +I+♦♦♦QP++ YKVPEV+EI+ I++AFMLEAA
 Fps4    55 +VDTYAL+KG YKSVSELSAE EYKKVAIL++ +I♦L+++Y+♦ +A♦♦M..♦♦Q +I+♦♦♦QP++ YKVENV+DI+ I++AFMLEGA
Ggps6    59 MLLAARDLRY QGSMPTLLDL AC........ AV♦LTHTAS♦ ML♦♦MPC♦♦N AEL♦♦♦QPTT HKKFGESVAI LASVGLLSKA
Ggps7    60 LLLTARDLGC AVSHDGLLDL AC........ AV♦MV*AA*♦ IL♦♦M**♦♦D AKL♦♦♦RPTI *SHYG*HV** *AAVA*LSKA
Ggps3   123 CIAACELVG. .GEESTAMPA RC........ AV♦MI*TM*♦ IH♦♦L**♦♦N DDL♦♦♦KPTN *KVFG*DV*V *AGDA*LSFS
Ggps5    72 CIAACEVVG. .GNESTAMAA AC........ SI♦MI*TM*♦ IH♦♦L**♦♦N DDL♦♦♦KPTN *KVFG*NI*V *AGDA*LAFA
Ggps4   125 CLAACELVG. .GNQENAMAA AC........ AV♦MI*TM*♦ IH♦♦L**♦♦N DDL♦♦♦KPTN *KIYG*DV*V *AGDS*LAFA
 Fps1    54 LLSTVRA+G. .KDPAVGLPV AC........ AI♦MI*TY*♦ IH♦♦L*S♦♦N DDL♦♦♦KPTN *KVFG*AM** *AGDG*LTYA
 Fps2    52 +YATGHMFG. .VSTNTLDAP AA........ AV♦CI*+Y*♦ IH♦♦L*A♦♦D DDL♦♦♦LPTC *VKFG*AN** *AGDA*QTLA
Ggps1    47 LLSVATRCG. DSRPALSDAA EV........ AL♦LI*CA*♦ VH♦♦L**F♦D AEI♦♦♦KPTV *RAYS*PL** *TGDS*IVMG
Ggps2    49 LTISSDLFG. .GQRERAYYA GA........ AI♦VL*TFT♦ VH♦♦I..♦♦Q DNI♦♦♦LPTV *VKYGLPL** *AGDL*HAKA
Ggps8    72 VMAGHCLARG SAVVPS.... ...GLWRFAA GL*LL*TFL♦ IH♦♦VA..♦Q AEL♦♦♦APPL *RMLAPGRAG EDLAVVVGDH
Ggps9   146 VKAFDQWL.. .DVPSE.... ...SLEVITK VISML*TA*♦ LV♦♦VE..♦N SVL♦♦♦FPVA *SIF.....G IPQTINTSNY


                                                   GQ   D                         KT
 Fps7   141 IYRLLKLYCR EQPYYLN+IE LFLQSSYQTE I......♦♦T L♦LLTAPQGN VDLVRFTEKR YKSIVKY♦♦A FYSFYLPIAA
 Fps8   141 IYRL+KFYC+ EQ+++LN+LE ++LQSSYQTE I......♦♦T L♦LI+APQGQ V++GRYTEKR YKS++KY♦♦+ F++++++I+A
Fps10   138 VYRV+KKYC+ QR+++VH+LE ++LQTAYQTE L......♦♦M L♦LI+APVSK V++SHFSEER YKA++KY♦♦+ F++++++V+A
 Fps5   131 IPRI+KKHF+ GK+++ADLLD ++NEVEFQTA S......♦♦M I♦LI+TLEGE K++SKYTLSL HRR++QY♦♦+ Y++++++V+C
 Fps6   131 IPRI+KKHF+ GK+++VD+LD ++NEVEFQTA S......♦♦M I♦LI+TLEGE K++SKYTLSL HRR++QY♦♦+ Y++++++V+C
 Fps3   132 IHRI+KKHF+ DK+++VD+VD ++NEVELQTA C......♦♦M I♦LI+TFEGE K++SKYSLSI HRR++QH♦♦+ Y++++++V+C
 Fps9   139 ISRI+RRHFK GK+++AD+LD ++NEVEFKTA S......♦♦L L♦LI+TFEGE K++TKYNITV HGR++QY♦♦+ Y++++++V+C
Fps11    83 IYKL+KSHF+ NEK++IDITE ++HEVTFQTE L......♦♦L M♦LI+APEDK V++SKFSLKK HSF++TF♦♦+ Y++++++V+L
 Fps4   132 IYCL+KKHF+ TE+++VD+LE ++HDVTFQTE L......♦♦L L♦LI+APEDK V++SKFSLEK HSF++IF♦♦+ Y+++++AV+L
Ggps6   131 FGLIAATGD. ..LPGERRAQ AVNELSTAVG VQ.GLVL♦♦F R♦LNDAAL.. ....DRTPDA ILSTNHL♦♦G IL.FSAMLQI
Ggps7   132 *GVIADADG. ..LTPLAKNR AVSE*SNAI* MQ.*LVQ♦♦F K♦LSEGDK.. ....PRSAEA ILMTNHF♦♦S T*.FCASMQM
Ggps3   193 *EHLASATSS DVVSPV*VVR AVGE*AKAI* TE.*LVA♦♦V V♦ISSEGLDL ...NDVGLEH LEFIHLH♦♦A A*.LEASAVL
Ggps5   132 *EHIAVSTSG ..VSPE*IIG AIGE*AKSI* TE.*LVA♦♦V V♦INSEG..L ...CDIGLEK LEFIHLH♦♦A A*.LEGSVVV
Ggps4   195 *EHIVNSTAG ..VTPS*IVG AVAE*AKSI* TE.*LVA♦♦V A♦IKCTGN.. ...ASVSLET LEFIHVH♦♦A A*.LESSVVL
 Fps1   124 *QLITEIDDE RI+PSV*+.R +IER*AKAA* PE.*MVA♦♦A A♦MEGEGK.. ....TLTLSE LEY+HRH♦♦G KM.LQYSVHA
 Fps2   122 *SILSDADMP EVSDRD*I.S MISE*ASAS* IA.*MCG♦♦A L♦LDAEGK.. ....HVPLDA LER+HRH♦♦G A*.IRAAVRL
Ggps1   118 *EVLAGAAAD R...PQ*ALQ LVTA*AVRT* MPM*ICA♦♦G WESESQ.... ........IN LSAYHRA♦♦G A*.FIAATQM
Ggps2   119 *QLLTQALRG ..LPSETIIK AFDIFTRSII I...ISE♦♦A V♦MEFEDR.. ...IDIKEQE YLDMISR♦♦A A*.FSASSSI
Ggps8   143 LFARALEVML GSG....LTC VAGVVQYYL* VSGHTAA♦♦Y L♦L...DLGR APLAEVTLFQ TLRVAHL♦♦A RY.GFCAPLV
Ggps9   209 VYFYALQEL. ...QKLKNPK AVSIFSEEL. ..LNLHR♦♦G M♦L...FWRD TLTCPTEDDY LEMV.SN♦♦G GL.F....RL
```

```
                                    G  FQ   DD LD         GK    D     K
 Fps7  215 AMYMAGIDGE KEHANAKKIL LEM♦EF♦♦IQ ♦♦Q♦♦LFGDP SVT♦♦.IGT♦ IQDN♦CSWLV VQCLQRATPE QYQILKENYG
 Fps8  215 +MYMA+IDGE KEHANALKI+ LEM♦EF♦♦I+ ♦♦+++LF+D+ SVT♦♦.V+++ +Q+N♦+++L+ VQC+LRATPQ +RQI+EE+++
Fps10  212 +MYMV+IDSK EEHENAKAI+ LEM♦EY♦♦I+ ♦♦+++CF+D+ ALT♦♦.V+++ +Q+N♦+++K+ VQC+QRVTPE +RQL+ED+++
 Fps5  205 +LLMV+.ENL DNHIDVKNI+ VDM♦TY♦♦V+ ♦♦+++CF+A+ ETI♦♦.I+++ +E+F♦+++L+ VKA+ELSNDE +KKV+YD+++
 Fps6  205 +LLMV+.ENL DNHIDVKNI+ VEM♦TY♦♦V+ ♦♦+++CF+A+ ETI♦♦.I+++ +E+F♦+++L+ VKA+ELSNEE +KKV+YD+++
 Fps3  206 +LLMA+.ENL ENHIDVKNV+ VDM♦IY♦♦V+ ♦♦+++CFAD+ ETL♦♦.I+++ +E+F♦+++L+ VKA+ERCSEE +TKI+YE+++
 Fps9  213 +LLLS+.ENL DNYGDVENI+ VEM♦TY♦♦V+ ♦♦+++CY+D+ EFI♦♦.I+++ +E+Y♦+++L+ VQA+ERADES +KRI+FE+++
Fps11  157 +MYVA+ITDE KDLKQARDV+ IPL♦EY♦♦I+ ♦♦+++CFGT+ EQI♦♦.I+++ +Q+N♦+++VI NKA+ELASAE +RKT+DE+++
 Fps4  206 +MFAA+ITDS KDLKQASDV+ IPL♦EY♦♦I+ ♦♦F♦♦CF+K+ EDI♦♦.I+++ +Q+N♦+++VI NVA+KNATKE +RDI+DE+++
Ggps6  200 VALASAS.SP STRETLHAFA LDF♦QA♦♦LL ♦♦LR♦DHPET ...♦♦..DRN KDAG♦STLVN RLGADAARQK LREHIDSADK
Ggps7  201 ASLVANA.SS EARDCLHRFS LDL♦QA♦♦LL ♦♦LT♦GMTDT ...♦♦..DSN QDAG♦STLVN LL*PRAVEER LRQHLQL*SE
Ggps3  268 GALVGGG.SD DEIERLRKFA RCI♦LL♦♦VV ♦♦I♦♦VTKSS KEL♦♦TAGK♦ LIAD♦LTYPK IM*LEKSREF AEKLNRE*RD
Ggps5  203 GALLGGG.CN EEVEKLRMFA RYI♦LM♦♦VV ♦♦V♦♦VTKSS KEL♦♦TAGK♦ LVAD♦VTYPK LL*IEKSNEF AQKLNRD*QE
Ggps4  266 GALLGGG.TN VEVEKLRRFA RCI♦LL♦♦VV ♦♦I♦♦VTKSS EEL♦♦TAGK♦ LVVD♦TTYPK LL*LEKAKEF AAELNRE*KQ
 Fps1  195 GALIG+ADAR QT.RELDEFA AHL♦LA♦♦IR ♦♦I♦♦IE+AE EKI♦♦PVGS♦ QSNN♦ATYPA LLS+AGAKEK LAFHIEA*Q.
 Fps2  193 GALSA+DKGR RALPVLDKYA ESI♦LA♦♦VQ ♦♦I♦♦VV+DT ATL♦♦RQGA♦ QQLG♦STYPA LL*+EQARKK ARDLIDD*RQ
Ggps1  182 GAIAAG.... YEAEPWEELG ARI♦EA♦♦VA ♦♦LR♦ALCDA ETL♦♦PAGQ♦ EIHARPSAVR EY*VEGAAKG LKDILGG*IA
Ggps2  179 GALIAGANDN .DVRLMSDFG TNL♦IA♦♦IV ♦♦I♦GLTADE KEL♦♦PVFS♦ IREG♦KTILV IKTLELCKED EKKIVLK*LG
Ggps8  214 CAAMLGGA.S SGLVEELERV GRHVLAY♦LR ♦♦L♦GLFWAL PVEQ♦DAAAL ARARALVESC GGRAACERMV VRASRAARRS
Ggps9  273 GIKLMQAESR SPVDCVP.LV NII♦LI♦♦IA ♦♦YHNLWNRE YTAN♦GMCE♦ LTEG♦FSFTV IHSIRSNPSN MQLLNILKQK

 Fps7  294 QKEAEKVARV KALYEELDLP AVFLQYEEDS YSHIMALIEQ Y..AAPLPPA VFLGLARKIY KRR.K..... .....
 Fps8  294 QKDPEK++R+ KA++E+LDLR SVFFK+EEDS +NRLKSL++G C..SAPLPPS IFLELAN+++ ++R.+..... .....
Fps10  291 RKEPEK++K+ KE++E+VGMR AAFQQ+EESS +RRLQEL++K H..SNRLPKE IFLGLAQ+++ ++Q.+..... .....
 Fps5  283 KPDPAN++K+ KA++D+LNLQ GVFTE+ESKS +EKLVTS++A HPSKA..VQA LLKSFLG+++ ++Q.+..... .....
 Fps6  283 KPDPAN++K+ KT++N+LNLE GAYAD+ESKS +EKLVTC++G HPSKA..VQG VLKSFWA+++ ++Q.+..... .....
 Fps3  284 KTDPSN++K+ KD++K+LDLE GVFME+ESKS +EKLTGA++G HQSKA..IQA VLKSFLA+++ ++Q.+..... .....
 Fps9  291 KKDPAC++K+ KN++K+LDLE AVFQE+ENES +KKLIAD++A QPSIA..VQK VLKSFLH+++ ++Q.+..... .....
Fps11  236 KKDSVAEAKC KKIFNDLKIE QLYHE+EESI AKDLKAK+SQ VDESRGFKAD VLTAFLN+V+ ++S.+..... .....
 Fps4  285 RKDSEKEQ+C RAVFN+LNIQ DIYHK+EEET +SNLREK+AN IDESRGFKAE VLTLFLN+++ H+K.+..... .....
Ggps6  274 HLTFACPQGG ...AIRQFMH LWFGHHLADW SPVMKIA... .......... .......... .......... .....
Ggps7  275 HLSAACQHGH ...ATQHFIQ AWFDKKLAAV S......... .......... .......... .......... .....
Ggps3  347 QLLGF..DSD ...KVAPLLA ..LANYIAYR QN........ .......... .......... .......... .....
Ggps5  282 QLSGF..DPV ...KVAPLIA ..LANYIAYS PN........ .......... .......... .......... .....
Ggps4  345 QLEGF..DSR ...KAAPLIA ..LADYIAYR DN........ .......... .......... .......... .....
 Fps1  273 ..RHLRN+D+ ...DGAALAY ..ICELVAAR DH........ .......... .......... .......... .....
 Fps2  273 SLKQLAEQSL ...DTSALEA ..LAD+IIQR NK........ .......... .......... .......... .....
Ggps1  258 SIPSCPAEAM LAEMVRRYAD KIVPAQVAAR V......... .......... .......... .......... .....
Ggps2  258 NKSASKEELM SSADIIKKYS LDYAYNLAEK YYKNAIDSLN QVSSKSDIPG KALKYLAEFT I+RR+..... .....
Ggps8  293 LQSLPNPNGV RELLDALLAR LAHRAA.... .......... .......... .......... .......... .....
Ggps9  352 TGDEE.VKRY AVAYMESTGS FEYTRKVIKV LVDRARQMTE DIDDGRGKSG GIHKILDRIM LHQEENVAQK NGKKE
```

Figure 12. Multiple sequence alignment of genes encoding FPS and GGPS from various species; Fps1 (*B. stearothermophilus*), Fps2 (*E. coli*), Fps3 (*A. thaliana*), Fps4 (*K. lactis*), Fps5 (*L. albus* pFPS3), Fps6 (*L. albus* pFPS1), Fps7 (*H. sapiens*), Fps8 (*R. norvegicus*), Fps9 (*Z. mays*), Fps10 (*G. gallus*), Fps11 (*S. cerevisiae*), Ggps1 (*R. sphaeroides*), Ggps2 (*S. acidocaldarius*), Ggps3 (*A. thaliana*), Ggps4 (*C. annuum*), Ggps5 (*L. albus*), Ggps6 (*E. herbicola*), Ggps7 (*E. uredovora*), Ggps8 (*M. xanthus*), and Ggps9 (*N. crassa*). ♦ represents residues conserved in most sequences, + represents residues conserved in most FPS sequences, and * represents residues conserved in most GGPS sequences.

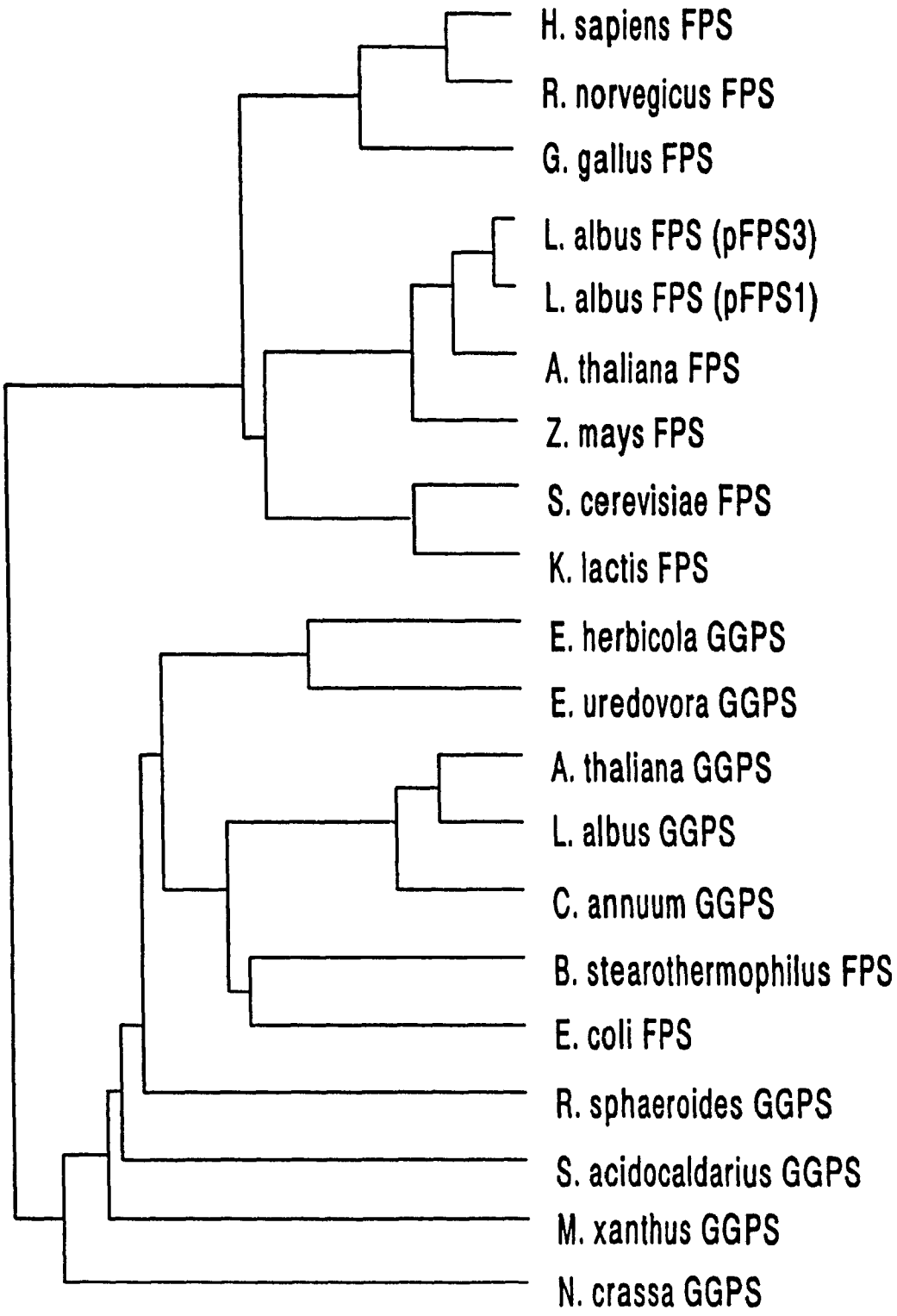
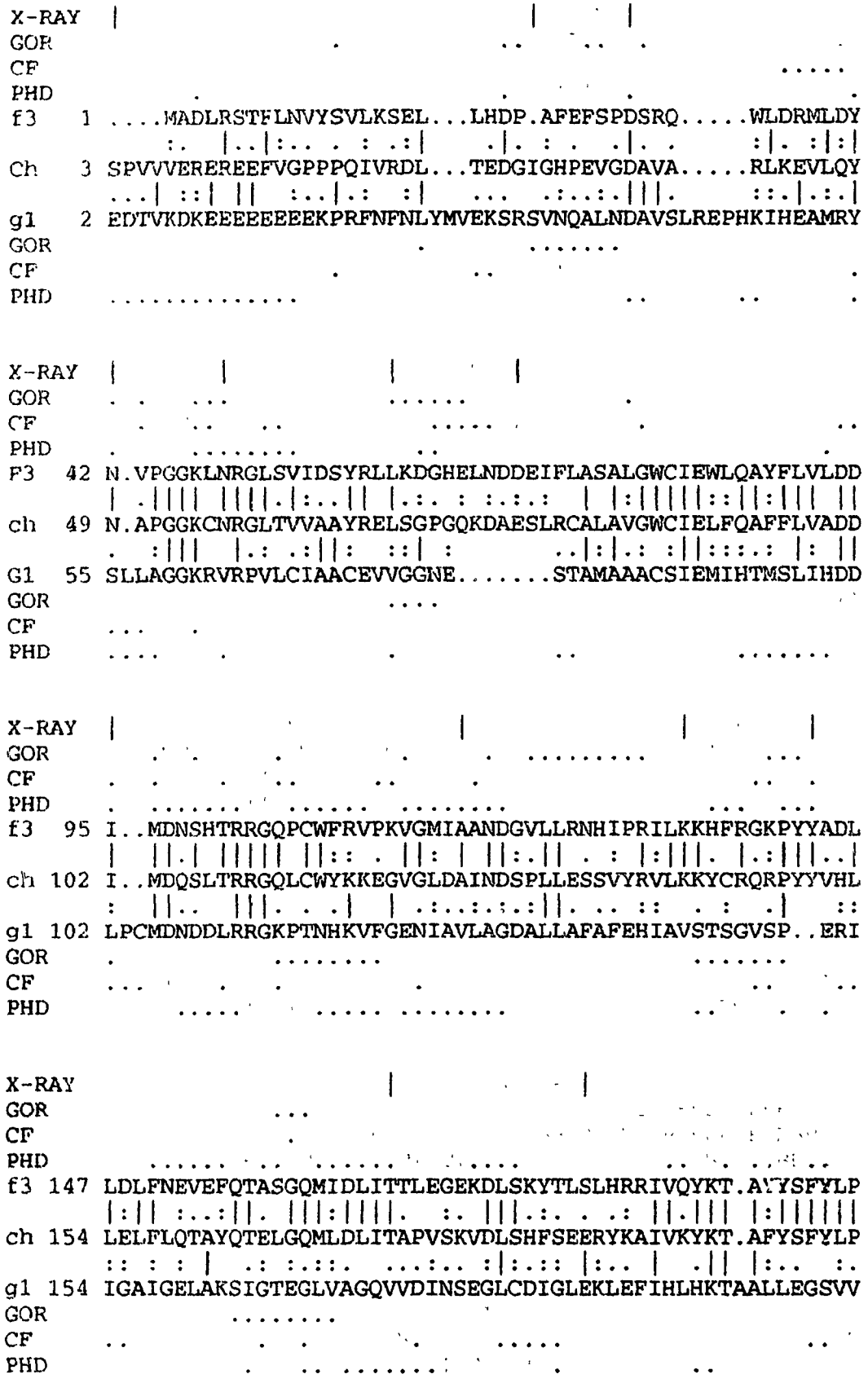


Figure 13. Dendrogram of multiple sequence alignment of FPS and GGPS sequences from various species.


X-RAY
GOR
CF
PHD
f3 1 ....MADLRSTFLNVYSVLKSEL...LHDP.AFEFSPDSRQ.....WLDRMLDY
Ch 3 SPVVVEREREEFVGPPPQIVRDL...TEDGIGHPEVGDAVA.....RLKEVLQY
g1 2 EDTVKDKEEEEEEEEKPRFNFNLYMVEKSRSVNQALNDAVSLREPHKIHEAMRY
X-RAY
GOR
CF
PHD
F3 42 N.VPGGKLNRGLSVIDSYRLLKDGHELNDDEIFLASALGWCIEWLQAYFLVLDD
ch 49 N.APGGKCNRGLTVVAAYRELSGPGQKDAESLRCALAVGWCIELFQAFFLVADD
G1 55 SLLAGGKRVRPVLCIAACEVVGGNE.....STAMAAACSIEMIHTMSLIHDD
X-RAY
GOR
CF
PHD
f3 95 I..MDNSHTRRGQPCWFRVPKVGMIAANDGVLLRNHIPRILKKHFRGKPYYADL
ch 102 I..MDQSLTRRGQLCWYKKEGVGLDAINDSPLLESSVYRVLKKYCRQRPYYVHL
g1 102 LPCMDNDDLRRGKPTNHKVFGENIAVLAGDALLAFAFEHIAVSTSGVSP..ERI
X-RAY
GOR
CF
PHD
f3 147 LDLFNEVEFQTASGQMIDLITTLEGEKDLSKYTLSLHRRIVQYKT.AYYSFYLP
ch 154 LELFLQTAYQTELGQMLDLITAPVSKVDLSHFSEERYKAIVKYKT.AFYSFYLP
g1 154 IGAIGELAKSIGTEGLVAGQVVDINSEGLCDIGLEKLEFIHLHKTAALLEGSVV
X-RAY
GOR
CF
PHD

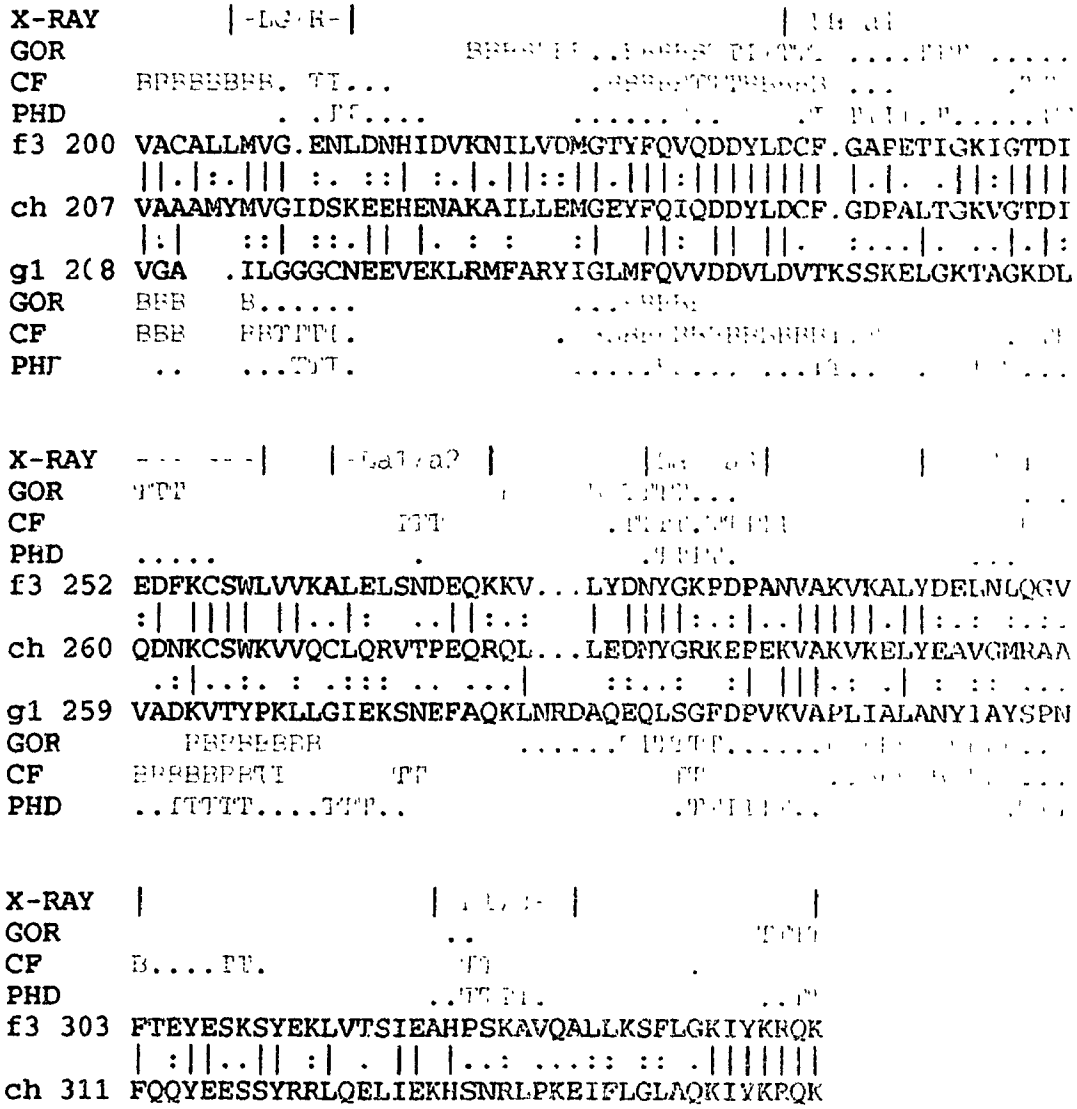


Figure 14. Sequence alignment of pFPS3 (f3), avian FPS (ch), and pGGPS1 (g1), with GOR, CF, and PHD secondary structure predictions and crystal structure. (H) represents a-helix-inducing residues, (B) represents b-strand-inducing residues, (T) represents turn-inducing residues, GOR (Garnier-Osguthorpe-Robson), CF (Chou-Fasman), PHD (protein-predict), X-RAY (secondary structure of avian FPS from crystal structure).

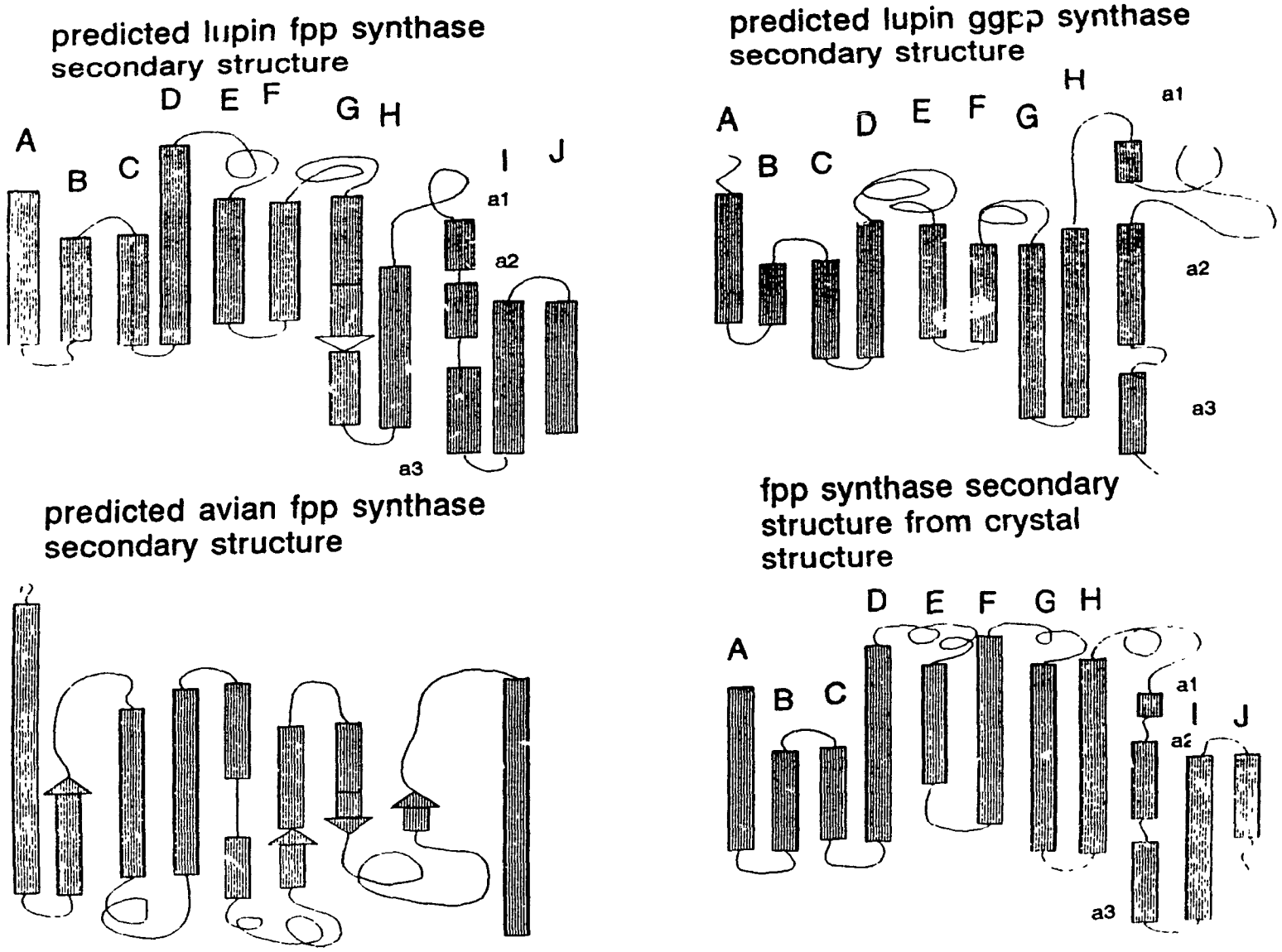


Figure 15. Graphical representation of the predicted secondary structures of pFPS3, pGGPS1, and avian liver FPS.

C.2.3 Secondary structure prediction

Three methods of secondary structure prediction were used in conjunction with sequence alignments in an effort to demonstrate that the recently published crystal structure for the avian FPS (Tarshis et. al., 1994) was an acceptable model for the lupin FPS and GGPS. The secondary structure prediction methods of Garnier-Osguthorpe-Robson (GOR), Chou-Fasman (CF), and Protein-predict (PHD) were used for both pFPS3 and pGGPS1 and predicted mainly α-helices and loops. The avian FPS crystal structure (Tarshis et. al., 1994) is also largely α-helical. Additional evidence for the structural models comes from a triple sequence alignment which shows a high degree of sequence similarity between these two sequences and the avian FPS. The predicted helices and loops are also well aligned in both the lupin FPS and GGPS and in the avian FPS (Fig. 14).

A graphical representation of the information in Figure 14 shows that when the three methods are combined, the overall predicted secondary structures for the lupin FPS and GGPS are similar both to each other and to the avian FPS crystal structure (Fig. 15). Panel 3 of Figure 15 shows a CF prediction of the avian FPS. Comparison of this predicted structure with that known from the crystal structure shown in panel 4 demonstrates that a single secondary structure prediction is not reliable alone. Therefore, the combination of three secondary structure prediction methods, sequence alignment, and alignment with the known crystal structure of avian FPS (Fig. 14) was designed to make structural predictions as accurate as possible.

The structures predicted for the lupin FPS and GGPS are close enough to that of

the avian FPS known structure that they were accepted as models for this study. Aside from slight differences in the lengths of some of the helices, the most obvious difference between the lupin FPS and GGPS is that the carboxy-terminal α-helices, I and J, present in the lupin and avian FPS, are not present in the GGPS.

C.3 Structure-function analysis

C.3.1 Production of mutant prenyltransferases

The observation that the GGPS structure was lacking the I and J helices was used as a starting point from which to examine the structure-function relationship between these two enzymes. Using PCR mutagenesis to introduce restriction sites, a series of truncated FPS clones and chimeric FPS-GGPS clones were made to determine whether these structural differences played a significant role in determining the product length specificity of these two enzymes. All clones, including the full-length FPS and GGPS, were inserted into the His-tag expression vector pTrc-His (Invitrogen) to facilitate purification.

Two truncated pFPS3 clones were constructed, one missing only helix J and the other missing helices I and J (Fig. 16). This was accomplished by using a 3' PCR primer with a stop codon just 5' of a restriction site. The stop codon truncated the coding sequence of the gene and the restriction site facilitated subcloning.

In order to make the chimeric constructs, three pGGPS1 constructs were made. The first of these is the full-length GGPS in the expression vector. The second is also a full-length GGPS, but with an XhoI restriction site introduced just 5' of the stop codon

to be used for the introduction of sections from the pFPS3 clone. The third GGPS construct was truncated to remove the three small, carboxy-terminal α-helices (α1, α2, and α3) and introduce a KpnI site at the 3' end of the sequence coding for the loop between helix H and α-1.

The pGGPS1 constructs with the introduced restriction sites were then used for the construction of chimeric clones (Fig.17). Two constructs were made from the full length GGPS with the XhoI site introduced. The first had both the I and J helices transferred from pFPS3 and the second had only helix I added. The third chimeric construct combined the truncated GGPS with the KpnI at the 5' end with the three small α-helices (α1 - α3) and helices I and J from pFPS3 (Table 5).

Table 5. Names and descriptions of clones

Clone number	Clone name	Clone description
1	C	control (no insert)
2	pF3IJ	pFPS3 minus J helix
3	pF3a3I	pFPS3 minus I and J helices
4	pG1F3Ha1	pGGPS1 minus helices α1-α3 with helices α1-α3, I, and J transferred from pFPS3
5	pG1F3a3Ia	pGGPS1 with helix I transferred from pFPS3
6	pG1F3a3Ib	pGGPS1 with helices I and J transferred from pFPS3
7	pGK201E	K201E site-directed mutant of pGGPS1
8	pGK201R	K201R site-directed mutant of pGGPS1
F3	pFPS3	full-length FPS3 clone
G1	pGGPS1	full-length GGPS1 clone

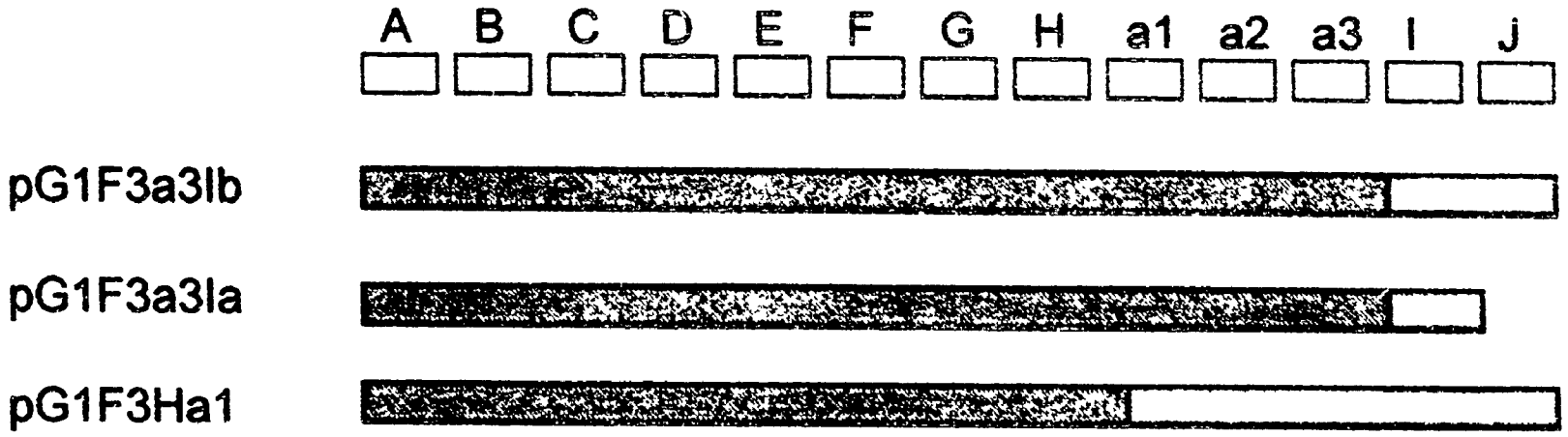


Figure 17. Chimeric constructs

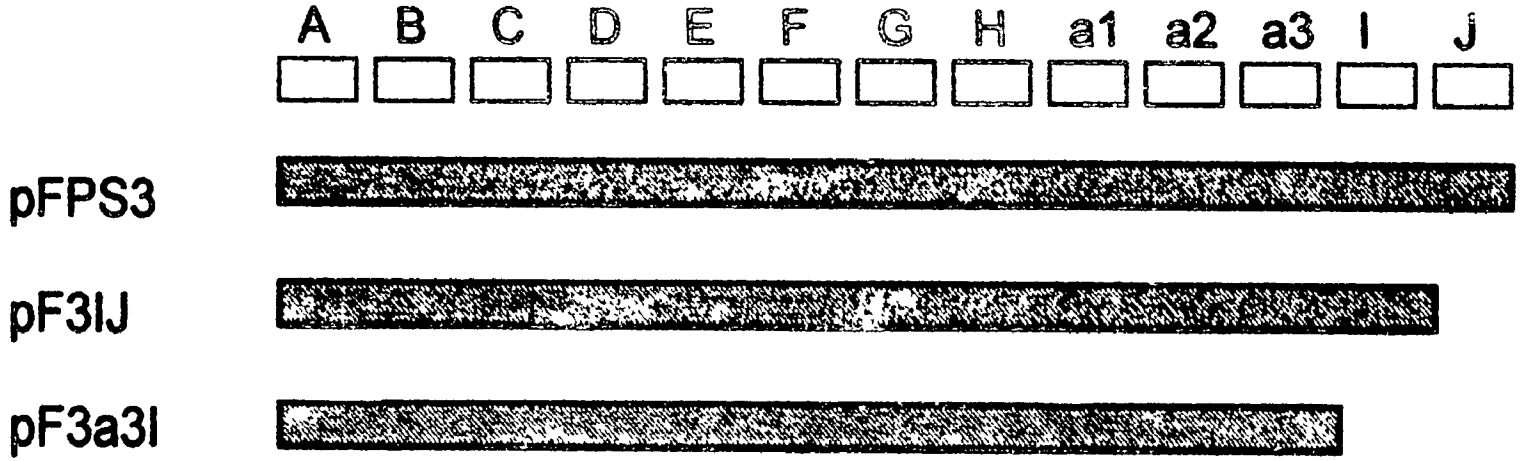


Figure 16. FPS constructs

C.3.2 Expression and purification of mutants

Large-scale expression and purification was carried out for the full-length pFPS3 clone. Fractions eluted from the Ni-NTA column were analyzed by SDS-PAGE (Fig. 18) and show a large band of protein migrating near the expected molecular weight of 39 KD. Maximum activity was also found in these fractions. The His-tagged FPS protein was then purified to near homogeneity by DEAE anion exchange chromatography. Fractions were run on SDS-PAGE as shown in Figure 19.

Small-scale preparations of 50 mL cultures for all clones were partially purified on the Ni-NTA (Qiagen) resin. SDS-PAGE indicated a large band of protein at the predicted size (approximately 39 KD) for the full-length pFPS3 (Fig. 20), though this band comigrated with a prominent band which was visible in the control *E. coli* lysate. An immunoblot (Fig. 21) of a parallel gel showed that all constructs were being expressed and that the protein product was present in the soluble fraction of the lysate.

The insoluble fractions of the cell lysates were solubilized in 6M guanidine-HCl and samples run on SDS-PAGE (Fig. 22). As with the soluble protein, a parallel gel was immunoblotted (Fig. 23). The blot indicated that a portion of the protein produced from the gene constructs was also present in the insoluble fraction. Figures 24 and 25 show an SDS-PAGE and an immunoblot similar to those in Figures 22 and 23, but with double the amount of protein. The bands in the western in Fig. 25 are no darker than those in Fig. 23 despite the additional protein. It is possible that the presence of guanidine in these samples interfered with recognition of the protein by the FPS antibody. Therefore, the presence of the overexpressed protein in inclusion bodies cannot be ruled out.

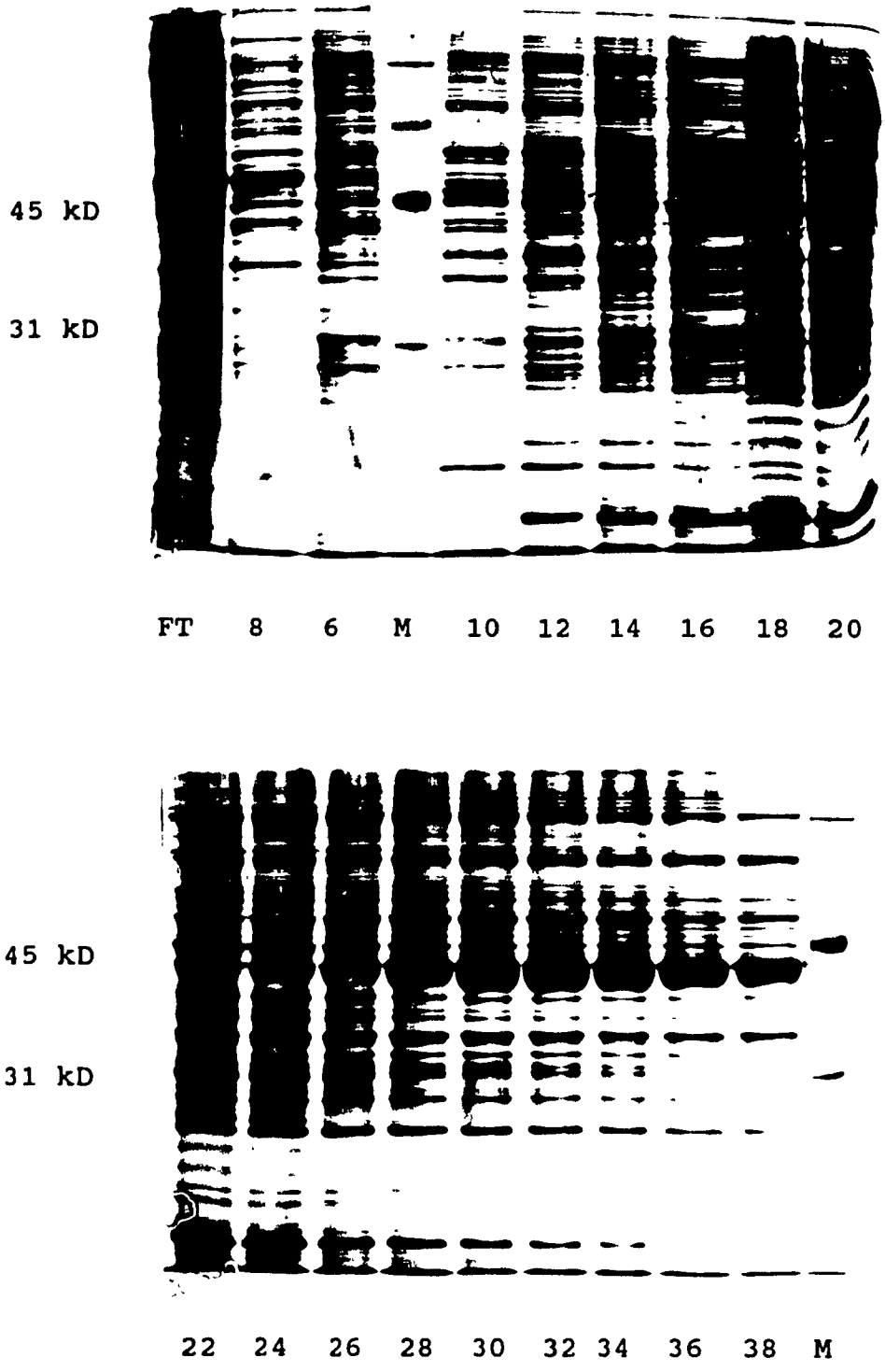


Figure 18. SDS-PAGE analysis of Ni-NTA column fractions from purification of pFPS3; FT (flow through), 8 (pH 8 wash), 6 (pH 6 wash), 10 - 38 (fractions 10 to 38 eluted from the Ni-NTA column), M (low range molecular weight markers).

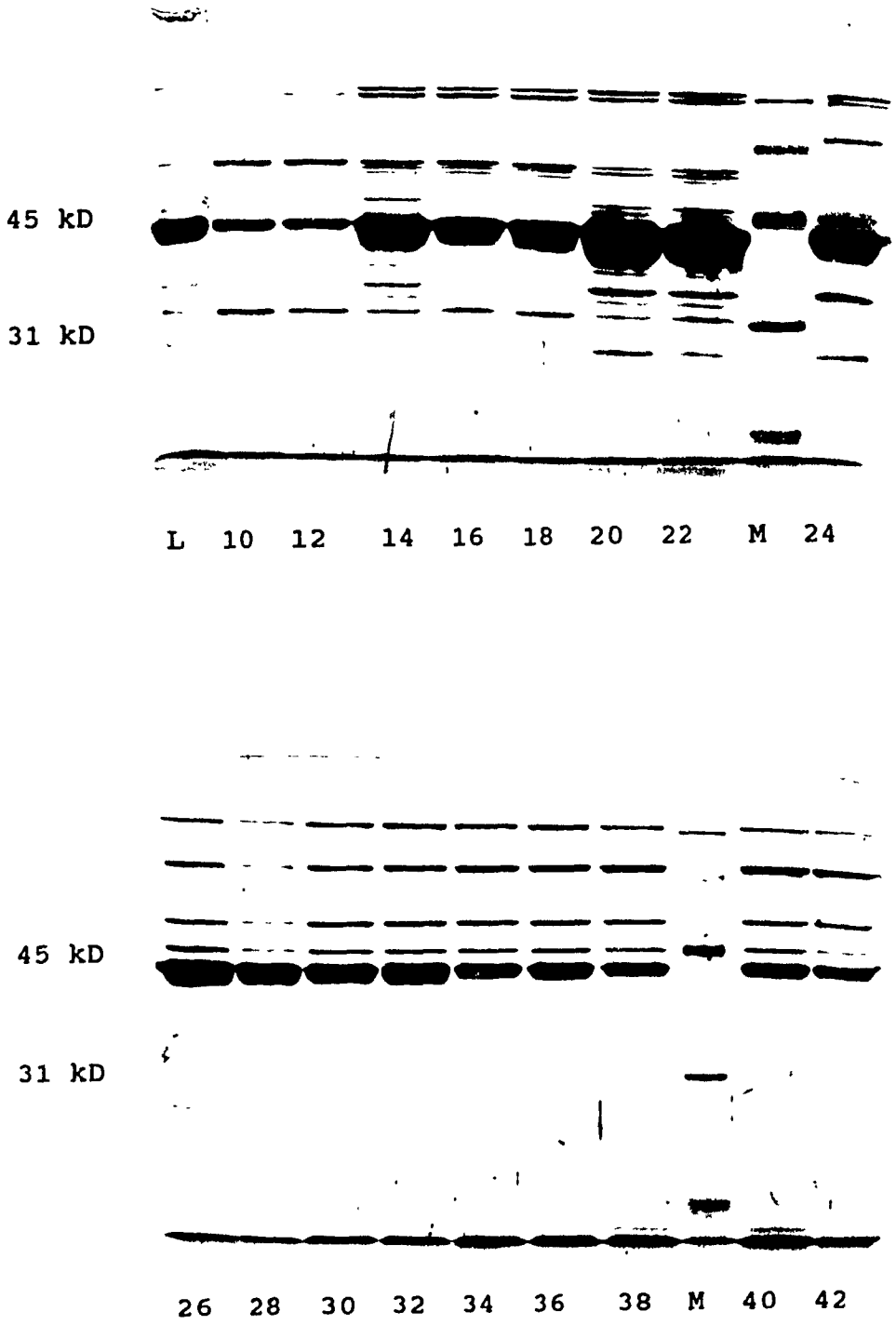


Figure 19. SDS-PAGE analysis of DEAE column fractions from purification of pFPS3; L (pooled, desalted, loaded Ni-NTA fractions), 10 - 42 (fractions 10 to 42 eluted from the DEAE column), M (low range molecular weight markers).

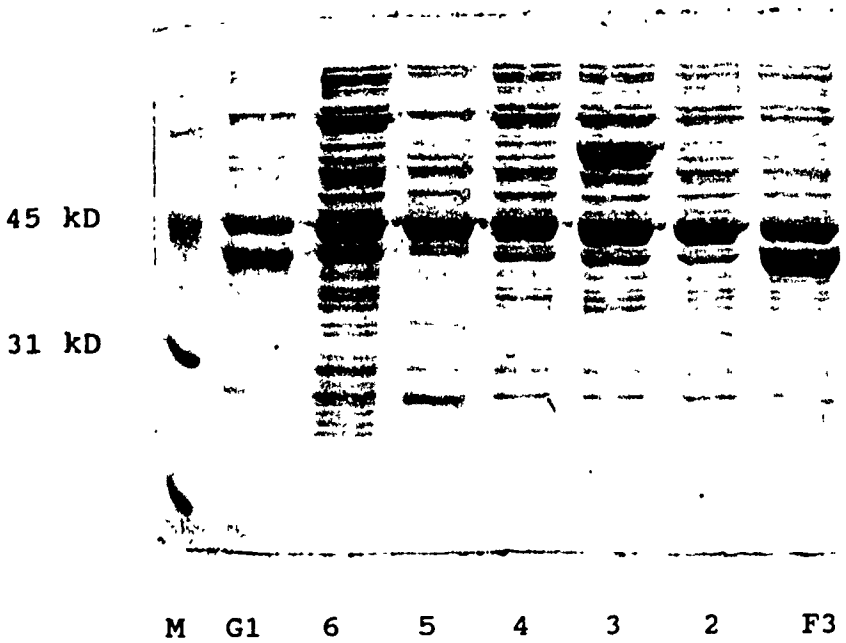


Figure 20. SDS-PAGE of Ni-NTA purified soluble lysate fractions; 2 (pF3IJ), 3 (pF3a3I), 4 (pG1F3Ha1), 5 (pG1F3a3Ia), 6 (pG1F3a3Ib), F3 (pFPS3), G1 (pGGPS1), M (low range molecular weight markers).

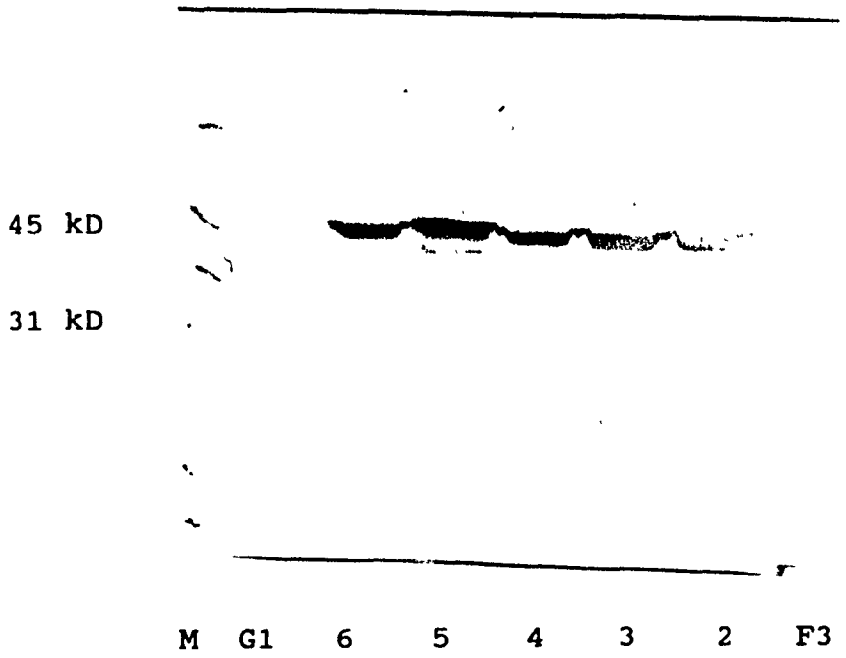


Figure 21. Immunoblot of Ni-NTA purified soluble lysate fractions; 2 (pF3IJ), 3 (pF3a3I), 4 (pG1F3Ha1), 5 (pG1F3a3Ia), 6 (pG1F3a3Ib), F3 (pFPS3), G1 (pGGPS1), M (low range molecular weight markers).

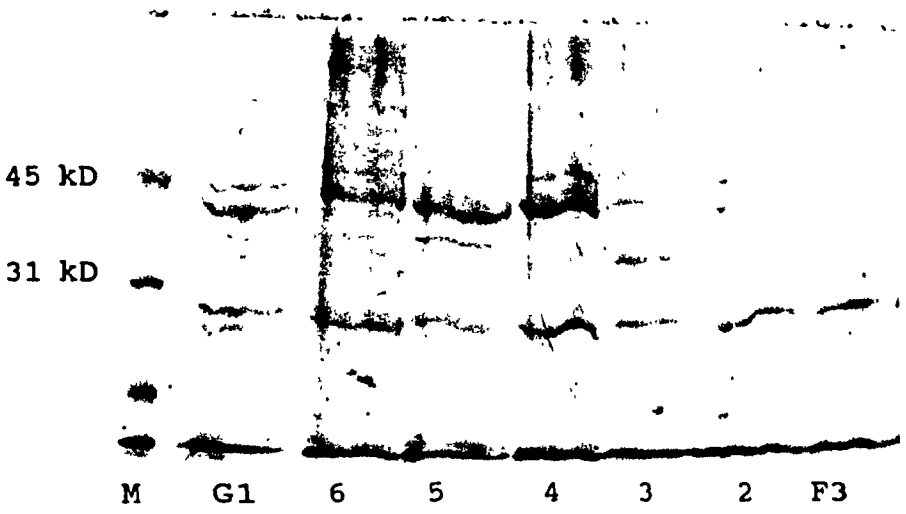


Figure 22. SDS-PAGE of solubilized insoluble lysate fractions; 2 (pF3IJ), 3 (pF3a3I), 4 (pG1F3Ha1), 5 (pG1F3a3Ia), 6 (pG1F3a3Ib), F3 (pFPS3), G1 (pGGPS1), M (low range molecular weight markers).

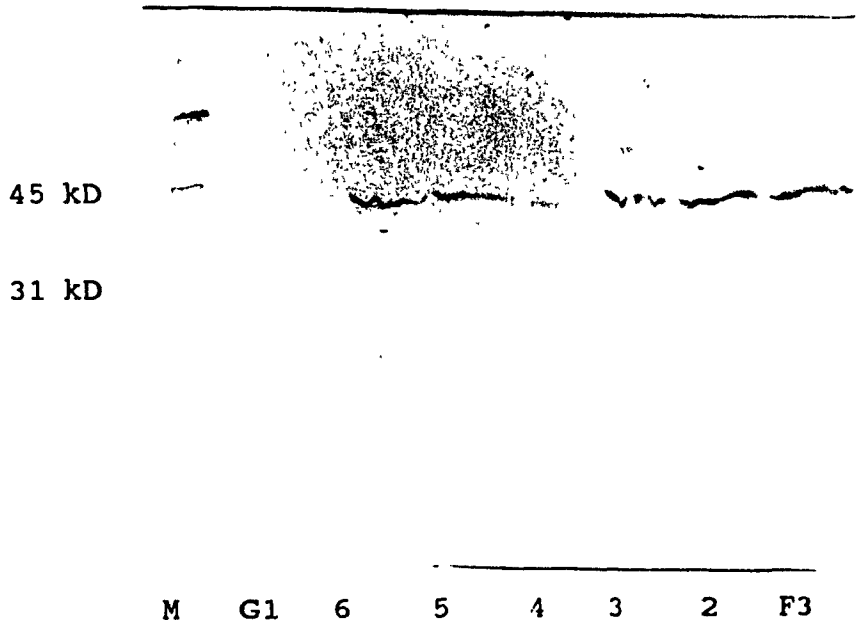


Figure 23. Immunoblot of solubilized insoluble fractions; 2 (pF3IJ), 3 (pF3a3I), 4 (pG1F3Ha1), 5 (pG1F3a3Ia), 6 (pG1F3a3Ib), F3 (pFPS3), G1 (pGGPS1), M (low range molecular weight markers).

Figure 24. SDS-PAGE of solubilized insoluble lysate fractions with protein doubled as compared to Figure 24; 2 (pF3IJ), 3 (pF3a3I), 4 (pG1F3Ha1), 5 (pG1F3a3Ia), 6 (pG1F3a3Ib), F3 (pFPS3), G1 (pGGPS1), M (low range molecular weight markers).

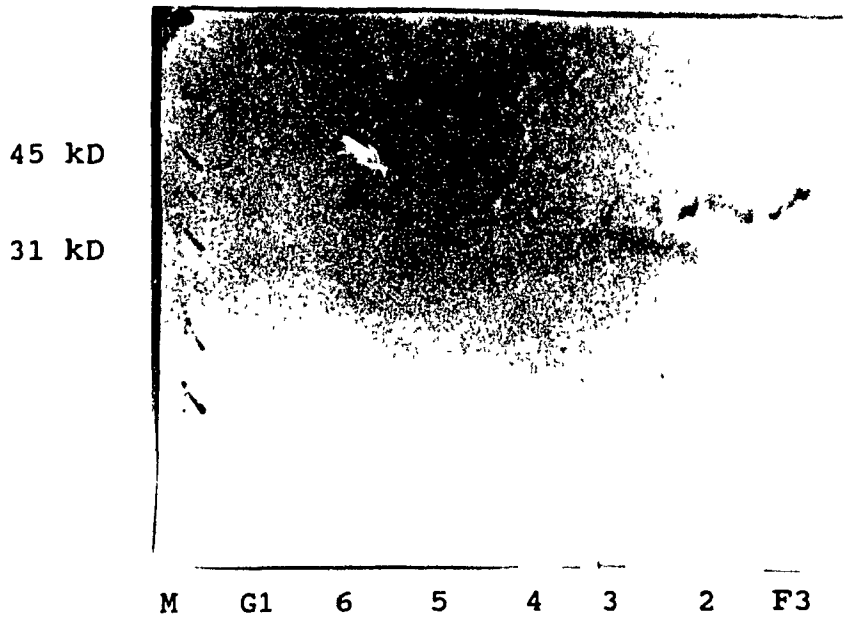


Figure 25. Immunoblot of solubilized insoluble lysate fractions with protein doubled as compared to Figure 25; 2 (pF3IJ), 3 (pF3a3I), 4 (pG1F3Ha1), 5 (pG1F3a3Ia), 6 (pG1F3a3Ib), F3 (pFPS3), G1 (pGGPS1), M (low range molecular weight markers).

C.3.3 Activity of mutants

Ni-NTA purified lysates of the 50 mL *E. coli* cultures of each gene construct were tested for FPS and GGPS activity with each of the three allylic substrates (DMAPP, GPP, and FPP). Enzymatic assays were acid hydrolysed, extracted, and used for scintillation counting. The comparative activities are shown in Figure 26. It was not possible to determine the specific activities due to the inability to express enough protein to purify to homogeneity. The absolute level of enzymatic activity of 6 of the 7 modified clones was approximately 1 % or less than that of the intact pFPS3 or pGGPS1 expressed in pTrc-His in *E. coli*. The other clone, pF3IJ, a modified pFPS3 clone with the terminal J helix removed, had approximately 2 % of the activity of pFPS3 when DMAPP was used as the allylic substrate (Fig. 26).

The most marked differences observed among the modified clones were changes in relative preference for the three allylic substrates. Substrate preference is illustrated in Figure 26, in which raw data have been normalized by dividing the allylic substrate activity of each clone by the IPP activity (labeled IPP only with no allylic substrate) for that clone. Whereas the intact clones of pFPS3 and pGGPS1 both have highest preference for GPP, the modified clones exhibit the highest activity with either DMAPP or FPP.

The activity profile of pF3IJ (construct #2) did not change greatly in relation to pFPS3. FPP was still not a favoured substrate. It was more active with either DMAPP or GPP, however, its preference was for DMAPP, in contrast to pFPS3 (Fig. 26). Product analysis showed that it was capable of producing FPP when either DMAPP or

GPP was used as the allylic substrate (Fig.27).

pF3a3I, pG1F3Ha1, and pG1F3a3Ib (constructs 5, 7, and 8, respectively) had no apparent allylic substrate preference and were relatively inactive as compared to pFPS3 and pGGPS1 (Fig.26). No products were visible with TLC for these constructs. Only pF3a3I showed any products, although they were very faint. The product observed from pF3a3I was FPP, as for the wild-type.

pG1F3a3Ia, pGK201E, and pGK201R showed similar allylic substrate preference patterns of FPP>GPP>DMAPP. This is markedly different from full-length clones, both of which had substrate preferences of GPP>DMAPP>FPP (Fig. 26). Product analysis by TLC of the activities of these clones showed products only for pGK201R and only with FPP from which it produced GGPP. It is possible that pGK201E produces FPP when GPP is the allylic substrate, but the TLC spot is too faint for confirmation (Fig. 28).

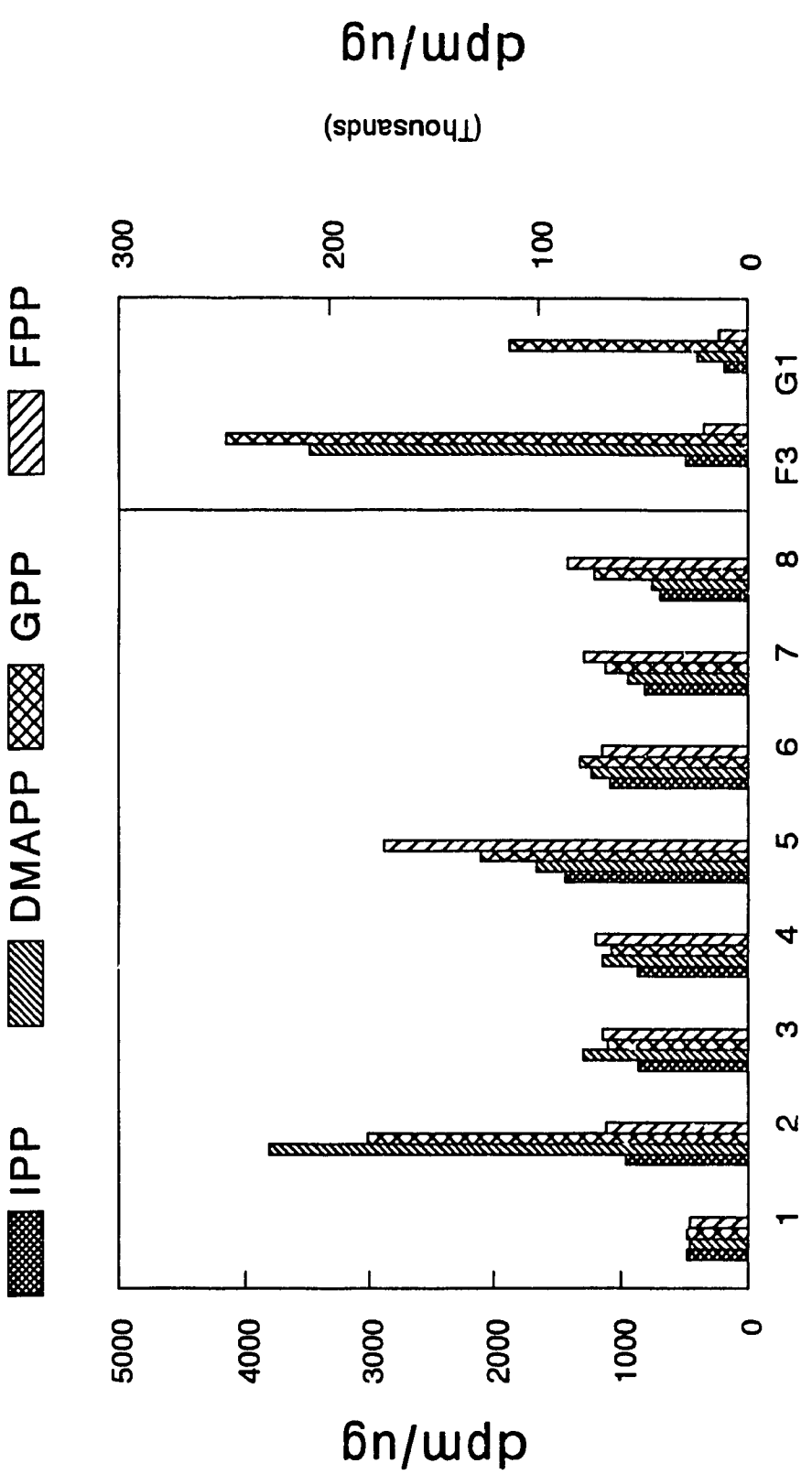


Figure 26. Comparative activities of constructs using partially purified protein from the Ni-NTA column (Qiagen). Numbering scheme as described in Table 5 on pg. 59.

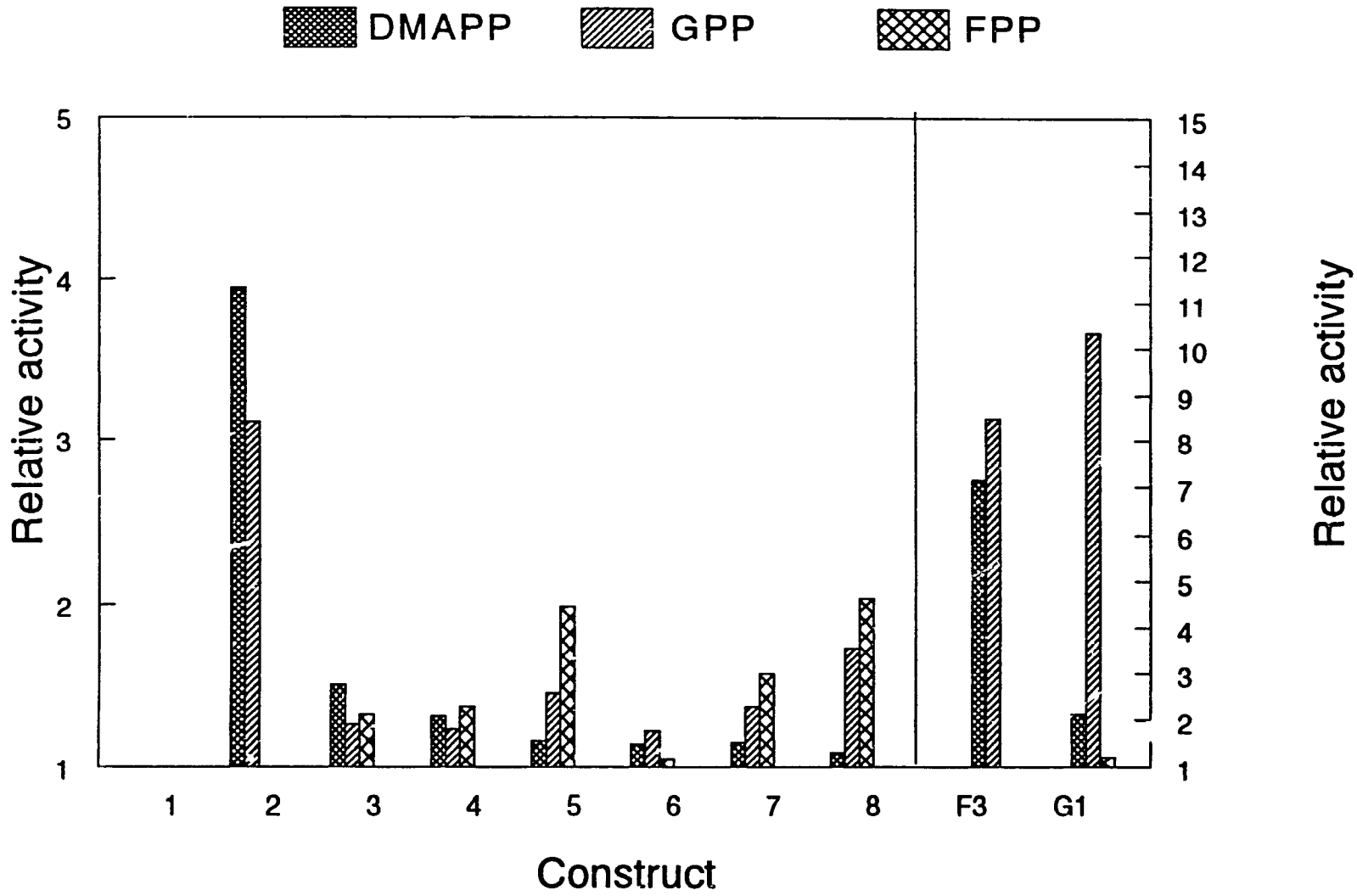


Figure 27. Relative activities of constructs using partially purified protein from the Ni-NTA column (Qiagen). Numbering scheme as described in Table 5 on pg. 59.

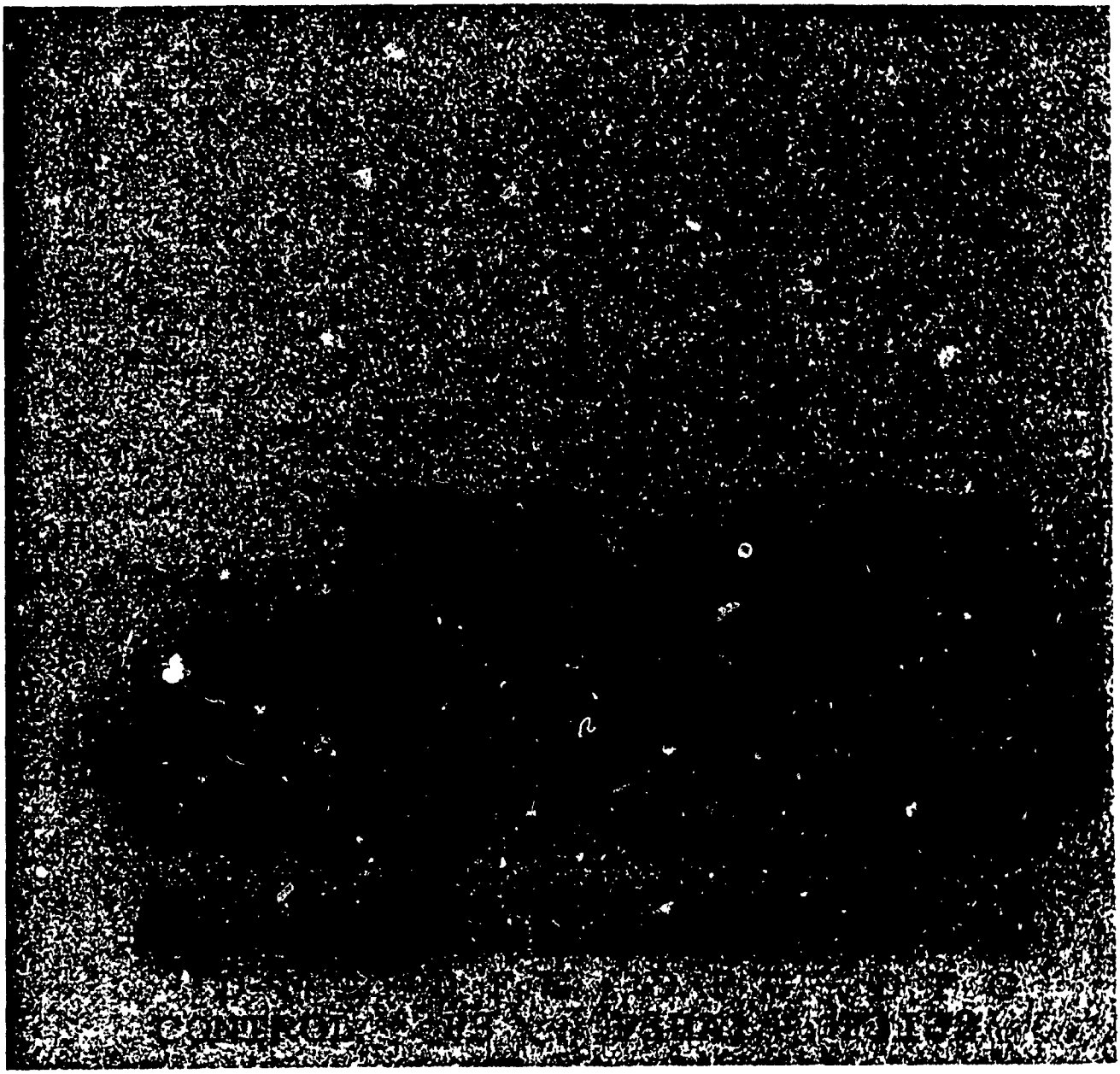


Figure 28. TLC product analysis of pFPS3 and truncated constructs pF3IJ and pF3a3I. (-) represents blank with IPP only, (D) represents DMAPP as the allylic substrate, (G) represents GPP as the allylic substrate, (F) represents FPP as the allylic substrate, hydrolyzed products are labeled on the left side as IOH (isopentenyl alcohol), GOH (geraniol), FOH (farnesol), and GGOH (geranylgeraniol).

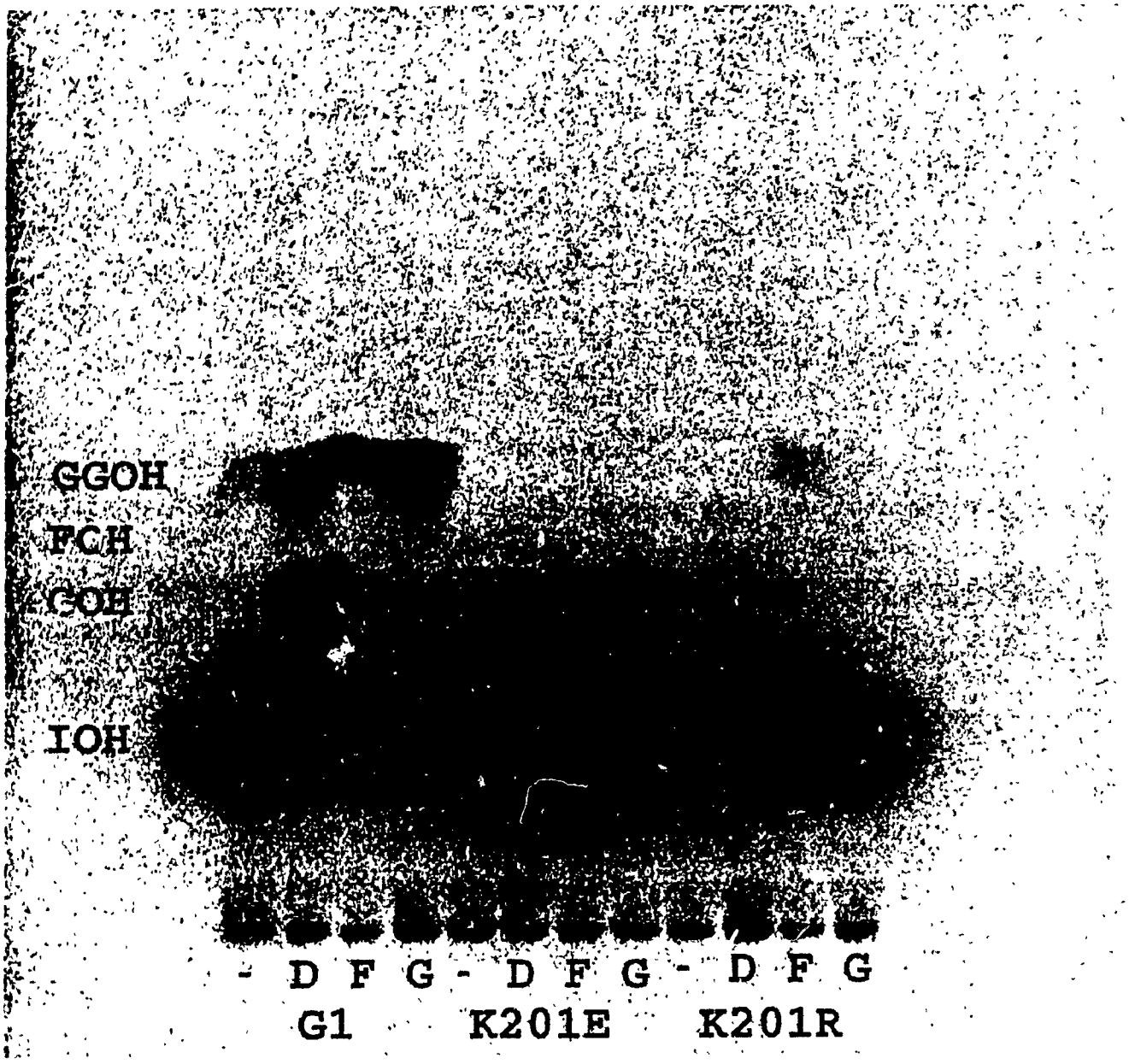


Figure 29. TLC product analysis of pGGPS1 and site-directed mutants pGK201E and pGK201R. (-) represents blank with IPP only, (D) represents DMAPP as the allylic substrate, (G) represents GPP as the allylic substrate, (F) represents FPP as the allylic substrate, hydrolyzed products are labeled on the left side as IOH (isopentenyl alcohol), GOH (geraniol), FOH (farnesol), and GGOH (geranylgeraniol).

D. DISCUSSION

D.1 Cloning of genes encoding FPS and GGPS from *Lupinus albus*

Two distinct FPS (pFPS1 and pFPS3) clones and one GGPS (pGGPS1) clone were isolated from a 10-day lupin seedling root cDNA library. The isolation of two FPS clones indicates that this gene is present as a gene family, a finding similar to some other species such as rat. In some Solanaceous plant species the genes encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR), an enzyme further upstream from FPS and GGPS in the isoprenoid biosynthetic pathway, exist as small gene families, with members having either constitutive or inducible gene expression. In tomato, at least one member of the HMGR gene family is induced by wounding and pathogenic agents and has been associated with isoprenoid phytoalexin biosynthesis (Weissenborn, et al., 1994). As FPS and GGPS are branch-point enzymes in the biosynthesis of isoprenoids, they are potential sites for regulation of the carbon flow in the various branches of the pathway. There are a variety of mechanisms by which regulation can be accomplished, including gene expression with transcriptional and/or translational control, the substrate affinities of the enzymes downstream of the branch points, and compartmentalization. The presence of two distinct, expressed FPS genes in lupin suggests they may be regulated at either the transcriptional level or by compartmentalization or both.

The presence of two FPS clones indicates that gene duplication occurred at some point in the past. The high level of identity at the nucleic acid level, 85%, as compared to 70% identity between the lupin FPS sequences and that of *A. thaliana* suggests that the gene duplication occurred after the divergence of these two species. In contrast,

there is only 39% nucleic acid sequence identity between the pGGPS1 clone and either of the FPS clones. If these genes arose from a common ancestral gene, it would appear that the gene duplication giving rise to these two enzymes predated the divergence of the phylogenetic lineages of these species.

Several functional domains also appear to have been conserved in FPS and GGPS. The multiple sequence alignment (Fig. 12) shows the five small regions of amino acids that are conserved amongst all the FPS and GGPS clones. At least one of the aspartate-rich domains, usually domain V, has been found to exist in almost all prenyltransferases and cyclases (Chappell, 1995). Since these aspartates have been shown to be involved in binding and catalysis, it is expected that they would be found in other enzymes catalyzing prenyl transfers. The amino acid sequence alignment also highlights those residues most conserved amongst FPS sequences and those most conserved amongst GGPS sequences. Both residues conserved among FPS and those conserved among GGPS are distributed throughout the length of the sequences, supporting evolution within each group from common ancestral FPS and GGPS enzymes.

The relative degree of similarity between FPS and GGPS sequences is shown in the dendrogram in Fig. 13, in which sequences cluster according to their phylogenetic origin. A notable exception to the clustering of the genes are the two bacterial FPS genes. These are more similar to the plant GGPS genes than to other FPS genes. Two of three cloned plant GGPS genes, those from *A. thaliana* and *C. annuum*, have plastid targeting signals. Many nuclear genes whose gene products are targeted to the plastid are thought to have had an evolutionary origin in the plastid genome. The clustering of

the plant GGPS sequences with bacterial sequences supports this notion.

D.2 Structural models for the lupin FPS and GGPS

Structural similarity of the FPS and GGPS genes is also suggested by comparison of their secondary structure predictions Chen et al. (1994) aligned six FPS, six GGPS, and one hexaprenyl pyrophosphate synthase (HPS) sequences and predicted their secondary structures. Their proposed structure, while not identical to that found in the crystal structure of avian FPS (Tarshis et al., 1994), is also almost entirely α-helical.

The secondary structure predictions for pFPS3 and pGGPS1 indicate that the lupin FPS and GGPS enzymes are likely to have entirely α-helical structures and that the avian FPS crystal structure is an acceptable model for these enzymes. The lupin FPS and GGPS have very similar predicted secondary structures; the major difference between them being that pGGPS1 is 26 amino acids shorter than pFPS3. The missing region in pGGPS1 corresponds to the two carboxy-terminal helices, I and J, found in the FPS sequences.

D.3 Structure-function relationship of lupin FPS and GGPS

The present work was designed to investigate the possibility that the presence or absence of the two carboxyl terminal helices of FPS could control the product length specificity of the enzyme. A series of truncated pFPS3 and chimeric pGGPS1-pFPS3 constructs were produced with the aim of making the structure of pFPS3 more like that of pGGPS1 and the structure of pGGPS1 more like that of pFPS3.

In all cases, modification of the cloned gene greatly reduced the activity in comparison to pFPS3 and pGGPS1. The level of activity per μg of partially purified protein from Ni-NTA column (Qiagen) (Fig. 26) indicates that the peak activity for most of the clones is reduced 50- to 100-fold as compared to pGGPS1 and approximately 200-fold as compared to pFPS3. The only exception is pF3IJ, which was missing only the J helix. While its activity is much lower than the full-length clones, it is 2- to 12-fold more active than the other constructs (Fig. 26). Product analysis of the pFPS3, pF3IJ, and pF3a3I clones (Fig. 28) confirmed that truncation of the FPS did not change the activity profile, as the products were FPP in all three cases when DMAPP or GPP was the allylic substrate. Substrate preference comparisons were possible with the partially purified proteins, however, comparison of absolute levels of activities between constructs and wild-type clones were less meaningful. The apparent absolute reductions of activity were probably a caused at least in part by reduced levels of expression, as the SDS-PAGE profiles show (Fig. 20). It is also likely that some of the actvity loss resulted from the perturbations in the tertiary structure caused by the trucations.

The most striking change in the activities for the constructs was seen for G1F3a3Ia with respect to relative substrate preference among the three allylic substrates (Fig. 27). Compared to pGGPS1, this construct had an approximately 5-fold reduction in preference for GPP accompanied by a 2-fold increase in preference for FPP. In this chimeric construct activity is greatly reduced, but the %GGPS activity is higher than in the full-length clone. Perhaps the addition of a section of foreign protein disturbs the structure in such a way that the allylic binding site looses affinity for the substrate or

allows water to enter the binding site. As a result, perhaps the longest substrate (FPP) would provide greater contact through its hydrophobic moiety to bind in a productive conformation. Similarly, it is possible that the structure of the allylic binding site was altered so as to be better able to accept larger substrates, which should bind with greater affinity due to their larger hydrocarbon moieties. If this were the case, then FPP could bind productively without occupying the homoallylic site and thus with less competition with IPP (King and Rilling, 1977). As a result, it would appear to be more active as a substrate than it usually is with the wild-type. The other two chimeric clones, pG1F3Ha1 and pG1F3a3Ib, were almost completely inactive (Fig. 27) and showed little substrate preference. While in both cases low levels of expression likely contribute to the reduced activity levels, changes in the overall topology of the protein and active site seem to have been too large for the enzyme to accommodate. None of the chimeric clones were active enough to distinguish products with TLC.

D.4 Site-directed mutants

The activity of the two site-directed mutants, K201E and K201R, was reduced approximately 100-fold as compared to pGGPS1 (Fig. 26). Their activity profiles, however, as seen by substrate preference were distinct. The relative activity of the K201E mutant with the substrate GPP was almost completely eliminated, while those for DMAPP and FPP were not greatly changed Fig. 27). Site-directed mutagenesis of FPS has recently been reported by Koyama et al. (1995) in which the conserved FQ residues of domain V were altered. In addition, a mutant FPS from yeast has been described

(Blanchard and Karst, 1993) in which the conserved lysine corresponding (domain IV) to position 201 in the lupin pGGPS1 clone was changed to a glutamate. Koyama et. al. (1995) observed drastic reductions in activity similar to that in both of the site-directed mutants of the lupin pGGPS1. They also found that the K_m for DMAPP of the F220A mutant was increased 25-fold. They suggested that the FQ motif maintains productive binding by the interaction of phenylalanine with the methyl group at the 3'-position *via* aromatic hydrophobicity and by the glutamine hydrogen bonding with the oxygen at the 1'-position of the allylic substrate. Site-directed mutagenesis of the aspartate residues of domain V by Marrero et al. (1992) showed that this domain binds to the pyrophosphate moiety of the allylic substrate. These proposed substrate binding mechanisms at domain V indicate the likely orientation of the substrate in the cleft. The pyrophosphate moiety of the allylic substrate is bound by the aspartates and arginines at the carboxy-terminal of helix H. The hydrocarbon moiety must then point down helix H into the cleft if the model of Koyama et al. (1995) for binding of the 3'-position methyl group and the 1'-position oxygen by the FQ motif is valid. This model would have the hydrocarbon moiety directed toward the conserved KT motif of domain IV, which is located close to the middle of helix G (Fig. 5). It is possible then that the KT motif of domain IV is involved in productive binding of the allylic substrate in a similar manner to the FQ motif of domain V. The threonine could bind the 7'-position methyl group by aromatic hydrophobicity as in the case of the phenylalanine with the 3'-position methyl group. The positively charged lysine could then interact with the π-electrons of the 6'-7' double bond. This could allow longer allylic substrates like GPP and FPP to be held securely

in a manner productive for catalysis.

Such a model might explain the results observed with the K201E mutation. The loss of the lysine at this position decreased the affinity of the enzyme for the GPP to a greater extent than its affinity for the larger allylic substrate, FPP. The longer substrate may be stabilized by interactions lower in the cleft, which could account for the change in the relative order of substrate preference observed for this mutant. Activity with DMAPP is nearly unaltered because it does not interact with domain IV. The K201R mutation is conservative with respect to charge and if the model were correct, it would not predict the large drop in the activity with the longer substrates observed for the K201E mutant. However, the K201R activity profile is similar to that of the K201E mutant, with a decreased affinity for GPP and that for FPP and DMAPP remaining relatively stable, as compared to pGGPS1. The enzyme appears to be very sensitive to even conservative changes in the conserved lysine of domain IV. This could be due to changes that cause the allylic substrate to bind in a non-productive manner. These findings parallel those of Blanchard and Karst (1993) who found that the FPS domain IV lysine to glutamate mutant lacked FPS activity but retained GPS activity.

D.5 Future work

The two approaches to engineering of FPS and GGPS had the effect of changing both the level of activity and the substrate specificity of these enzymes. While changes in product length specificity were not observed at the concentrations of substrate used, one of the chimeric enzymes (pG1F3a3Ia) and both of the site-directed mutants

(pGK201E and pGK201R) showed altered substrate preference profiles. Although activity was detectable for all constructs, products were only visible on TLC plates for pF3IJ, pF3a3I, and GK201R. Higher substrate concentrations or very long TLC exposure times could be used to probe substrate and product length specificity, especially for those constructs whose products were not visible on TLC. Increased expression levels in order to purify the overexpressed proteins in a manner similar to that used for pFPS3 is required for a more complete study of these constructs, including determination of specific activities and K_m constants for the various substrates. Pptimization of the *E. coli* system or possibly another expression system, such as Baculovirus insect cell culture system, would increase levels of expression of the protein in a soluble form and facilitate purification and analysis. Kinetic analysis of the wild-type, the truncated and chimeric constructs, and the site-directed mutants would clarify the apparent changes in substrate affinity discussed above.

The availability of the FPS, GGPS, truncated FPS, chimeric FPS-GGPS, and site-directed GGPS mutants provides an ideal experimental system for the study of the structure-function relationship of these enzymes. Possible avenues for future experimental work include:

(1) Fluoresence and circular dichroism spectroscopy in order to probe structural alterations caused by the changes introduced.

(2) X-ray crystalography in the presence and absence of substrates.

(3) Computer modeling followed by site-directed mutagenesis with the goal of determining the factors controling the product length specificity of these enzymes.

(4) Photoaffinity labeling using allylic analogues of increasing length in order to map the binding site for the allylic substrate.

(5) Use of substrate analogues in conjunction with the FPS and GGPS constructs and site-directed mutants for the purpose of mapping the binding sites for the substrates.

This work provides insight into substrate binding that would have application to a number of other classes of enzymes, notably the isoprenoid cyclases. The cyclases catalyze intramolecular prenyl transfer reactions and are responsible for much of the variety observed in the 23 000 known members of the isoprenoid family. Some of these compounds are of commercial importance and currently those of pharmaceutical and industrial importance are harvested from plants (Chappell, 1995). An understanding of structure-function relationship in these enzymes will provide the tools necessary for the design of chimeric isoprenoid biosynthetic genes resulting in the production of novel isoprenoids. These may have applications in industry, pharmaceuticals, and in phytoprotection, as the biosynthetic pathway is ubiquitous and some plants already use isoprenoid phytoanticipins and phytoalexins as a defense mechanism.

E. REFERENCES

Aitken, S.M., Attucci, S., and Gulick, P.J. (1995) A cDNA encoding geranylgeranyl pyrophosphate synthase from white lupin. Plant Phys. 108: 837-838

Anderson, M.S., Yarger, J.G., Burck, C.L., and Poulter, C.D. (1989) Farnesyl diphosphate synthetase - molecular cloning, sequence, expression of an essential gene from *Saccharomyces cerevisiae*. J. Biol. Chem. 264: 19176-19184

Ashby, M.N., Spear, D.H., and Edwards, P.A. (1990) Prenyltransferases from yeast to man. In: Attic A.D., ed. Molecular biology of atherosclerosis. Amsterdam: Elsevier Science Publishers. pp 27-34

Attucci, S., Aitken, S.M., and Gulick, P.J. (1995a) A cDNA encoding farnesyl pyrophosphate synthase in white lupin. Plant Physiol. 108: 835-836

Attucci, S., Aitken, S.M., Gulick, P.J., and Ibrahim, R.K. (1995b) Farnesyl pyrophosphate synthase from white lupin: molecular cloning, expression, and purification of the expressed protein Arch. Biochem. Biophys. 321: 493-501

Barnard, G.F. and Popjak, G. (1981) Human liver prenyltransferase and its characterization. Biochim. Biophys. Acta 661: 87-99

Barnard, G.F. and Popjak, G. (1980) Characterization of liver prenyltransferase and its inactivation by phenylglyoxal. Biochim. Biophys. Acta 617: 169-182

Bartley, G.E. and Scolnik, P.A. (1994) Nucleotide sequence of an arabidopsis cDNA fro geranylgeranyl pyrophosphate synthase. Plant Physiol. 104: 1469-1470

Blanchard, L. and Karst, F. (1993) Characterization of a lysine-to-glutamic acid mutation in a conservative sequence of farnesyl diphosphate synthase from *Saccharomyces cerevisiae*. Gene 125: 185-189

Brems, D.N., Bruenger, E., and Rilling, H.C. (1981) Isolation and characterization of a photoaffinity-labeled peptide from the catalytic site of prenyltransferase. Biochemistry 20: 3711-3718

Brems, D.N. and Rilling, H.C. (1979) Photoaffinity labeling of the catalytic site of prenyltransferase. Biochemistry 18: 860-864

Brems, D.N. and Rilling, H.C. (1977) On the mechanism of the prenyltransferase reaction. Metal ion dependent solvolysis of an allylic pyrophosphate. J. Am. Chem. Soc. 99: 8351-8352

Camara, B. (1985) Carotene synthesis in capsicum chromoplasts. Methods in Enzymology 147: 244-253

Carattoli, A. Romano, N., Ballario, P., Morelli, G., and Macino, G. (1991) The neurospora crassa carotenoid biosynthetic gene (albino 3) reveals highly conserved regions among prenyltransferases. J. Biol. Chem. 266: 5854-5859

Chappell, J. (1995) The biochemistry and molecular biology of isoprenoid metabolism. Plant Physiol. 107: 1-6

Chen, A., Kroon, P.A., and Poulter, C.D. 1994. Isoprenyl diphosphate synthases: protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. Protein Sci. 3: 600-607

Chen, A. and Poulter, C.D. (1993) Purification and Characterization of farnesyl diphosphate/geranylgeranyl diphosphate syntase - a thermostable bifunctional enzyme from methanobacterium thermoautotrophicum. J. Biol. Chem. 268: 1102-1107

Cornforth, J.W., Cornforth, R.H., Popjak, G., and Yengoyan, L. (1966) Studies on the biosynthesis of cholesterol. XX. Steric course of decarboxylation of 5-pyrophosphmevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesyl pyrophosphate. J. Biol. Chem. 241: 3970-3987

Cornforth, J.W., Cornforth, R.H., Donninger, C., and Popjak, G. (1966b) Proc. Roy. Soc. (Lond.). Ser. B. 163: 492

Delourme, D., Lacroute, F., and Karst, F. (1994) Cloning of an *Arabidopsis thaliana* cDNA coding for farnesyl diphosphate syntnase by functional complementation in yeast. Plant Mol. Biol. 26: 1867-1873

Dogbo, O. and Camara, B. (1987) Purification of isopentenyl pyrophosphate isomerase and geranylgeranyl pyrophosphate synthase ?from capsicum chromoplasts by affinity chromatography. Biochim. Biophys. Acta 920: 140-148

Eberhardt, N.L. and Rilling, H.C. (1975) Prenyltransferase from saccharomyces cerevisiae. J. Biol. Chem. 250: 863-866

Ericsson, J., Runquist, M., Thelin, A., Andersson, M., Chojnacki, T., and Dallner, G. (1993) Distribution of prenyltransferases in rat tissues - Evidence for a cytosolic all-trans-geranylgeranyl diphosphate synthase. J. Biol. Chem. 268: 832-838

Fujisaki, S., Hara, H., Nishimura, Y., Horiuchi, and Nishino, T. (1990) Cloning and nucleotide sequence of *ispA* gene responsible for farnesyl diphosphate synthase activity in *Escherichia coli*. J. Biochem. 108: 995-1000

Holloway, P.W. and Popjak, G. (1967) The purification of 3,3-dimethylallyl- and geranyl-transferase and of isopentenyl pyrophosphate isomerase from pig liver. Biochem. J. 104: 57-70

Ito, M., Kobayashi, M., Koyama, T., and Ogura, K. (1987) Stereochemical analysis of prenyltransferase reactions leading to (Z)- and (E)-polyprenyl chains. Biochemistry 26: 4745-4750

Joly, A. and Edwards, P.A. (1993) Effect of site-directed mutagenesis of conserved aspartate and arginine residues upon farnesyl diphosphate synthase activity. J. Biol. Chem. 268: 26983-26989

King, H.L. and Rilling, H.C. (1977) Avian liver prenyltransferase. The role of metal in substrate binding and the orientation of substrates during catalysis. Biochemistry 16: 3815-3819

Koike-Takeshita, A., Koyama, T., Obata, S., and Ogura, K. (1995) Molecular cloning and nucleotide sequences of the genes for two essential proteins constituting a novel enzyme system for heptaprenyl diphosphate synthesis. J. Biol. Chem. 270: 18396-18400

Koyama, T., Tajima, M., Nishino, T., and Ogura, K. (1995) Significance of Phe-220 and Gln-221 in the catalytic mechanism of farnesyl diphosphate synthase of *Bacillus stearothermophilus*. Biochem. Biophys. Res. Commun. 212: 681-686

Koyama, T., Saito, A., Ogura, K., Obata, S., and Takeshita, A. (1994a) Site-directed mutagenesis of farnesyl diphosphate synthase: effect of substitution on the three carboxy-terminal amino acids. Can. J. Chem. 72: 75-79

Koyama, T., Obata, S., Saito, K., Takeshita-Koike, A., and Ogura, K. (1994b) Structural and functional roles of the cysteine residues of *Bacillus stearothermophilus* farnesyl diphosphate synthase. Biochemistry 33: 12644-12648

Koyama, T., Obata, S., Osabe, M., Takeshita, A., Yokoyama, K., Uchida, M., Nishino, T., and Ogura, K. (1993) Thermostable farnesyl diphosphate synthase of *Bacillus stearothermophilus*: molecular cloning, sequence determination, overproduction, and purification. J. Biochem. 113: 355-363

Koyama, T., Ogura, K., and Seto, S. (1977) Construction of a chiral center by use of the stereospecificity of prenyltransferase. J. Am. Chem. Soc. 99: 1999-2000

Kuntz, M., Romer, S., Suire, C., Hugueney, P., Weil, J.H., Schantz, R., and Camara, B. (1991) identification of a cDNA for the plastid-located geranylgeranyl pyrophophate synthase from capsicum annuum: correlative increase in enzyme activity and transcript level druing fruit ripening. Plant J. 2: 25-34

Laskovics, F.M. and Poulter, C.D. (1981) Prenyltransferase: determination of the binding mechanism and individual kinetic constants for farnesylpyrophosphate synthetase by rapid quench and isotope partitioning experiments. Biochemistry 20: 1893-1901

Laskovics, F.M., Krafcik, J.M., and Poulter, C.D. (1979) Prenyltransferase - kinetic studies of the 1'-4 coupling reaction with avian liver enzyme. J. Biol. Chem. 254: 9458-9463

Li, C.P. and Larkins, B.A. (1995) Nucleotide sequence of a maize farnesyl pyrophosphate synthase. unpublished.

Lutke-Brinkhaus, F. and Rilling, H.C. (1988) Purification of geranylgeranyl diphosphate synthase from *Phycomyces blakesleanus*. Arch. Biochem. Biophys., 266: 607-611

Marrero, P., Poulter, C.D., and Edwards, P.A. (1992) Effects of site-directed mutagenesis of the highly conserved aspartate residues in domain II of farnesyl diphosphate synthase activity. J. Biol. Chem. 267: 21873-21878

Math, S.K., Hearst, J.E., and Poulter, C.D. (1992) The crtE gene in erwinia herbicola encodes geranylgeranyl diphosphate synthase. Proc. Natl. Acad. Sci. USA 89: 6761-6764

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

Nishino, T., Ogura, K., and Seto, S. (1972) Substrate specificity of farnesyl pyrophosphate synthetase. J. Am. Chem. Soc. 94: 6849-6853

Nishino, T., Ogura, K., and Seto, S. (1973) Comparative specificity of prenyltransferase of pig liver and pumkin with respect to artificial substrates. Biochim. Biophys. Acta 302: 33-37

Nishino, T., Ogura, K., and Seto, S. (1971) Reactivity of artificial substrates for prenyltransferase. Biochim. Biophys. Acta 235: 322-325

Ogura, K. Saito, A., and Seto, S. (1974) A novel substrate for prenyltransferase. Formation of a nonallylic *cis*-homofarnesyl pyrophosphate. J. Am. Chem. Soc. 96: 4037-4038

Ogura, K., Nishino, T., Koyama, T., and Seto, S. (1970) Enzyme condensation of 3-methyl-2-alkenyl pyrophosphates with isopentenyl pyrophosphates. J. Am. Chem. Soc. 92: 6036-6041

Parker, T.S., Popjak, G., Sutherland, K., and Wong, S.M (1978)

Inhibition of liver prenyltransferase by alkyl phosphonates and phosphonophosphates. Biochim. Biophys. Acta 530: 24-34

Popjak, G. (1971) Specificity of enzymes of sterol biosynthesis. Harvey Lect. 65: 127-156

Popjak, G., Holloway, P.W., Bond, R.P.M, and Roberts, M. (1969) Analogues of geranyl pyrophosphate as inhibitors of prenyltransferase. Biochem. J. 111: 333-343

Popjak, G., Holloway, P.W., and Baron, J.M (1969) Synthesis of 10,11-dihydrofarnesyl pyrophosphate from 6,7-dihydrogeranyl pyrophosphate by prenyltransferase. Biochem. J. 111: 325-332

Popjak, G., Rabinowitz, J.L., and Baron, J.M. (1969) Artificial substrates for prenyltransferase. Biochem. J. 113: 861-868

Poulter, C.D. and Rilling, H.C. (1981) Prenyl transferases and isomerase. In: Biosynthesis of Isoprenoid Compounds Vol. 1. Ed. John W. Porter and Sandra L. Spurgeon. John Wiley and Sons, N.Y. pp. 161- 224

Poulter, C.D., Mash, E.A., Argyle, J.C., Muscio, O.J., and Rilling, H.C. (1979) Farnesyl pyrophosphate synthetase. Mechanistic studies of the 1'-4 coupling reaction in the terpene biosynthetic pathway. J. Am. Chem. Soc. 101: 6761-6763

Poulter, C.D. and Rilling, H.C. (1978) The prenyl transfer reaction. Enzymatic and mechanistic studies of the 1'-4 coupling reaction in the terpene biosynthetic pathway. Accnts. Chem. Res. 11: 307-313

Poulter, C.D., Argyle, J.C., and Mash, E.A. (1978) Farnesyl pyrophosphate synthetase. Mechanistic studies of the 1'-4 coupling reaction with 2-fluorogeranyl pyrophosphate. J. Biol. Chem. 253: 7227-7233

Poulter, C.D. and Satterwhite, D.M. (1977) Mechanism of the prenyl-transfer reaction. Studies with (E)- and (Z)-3-trifluoromethyl-2-buten-1-yl pyrophosphate. Biochemistry 16: 5470-5477

Poulter, C.D. and Rilling, H.C. (1976) Prenyltransferase: the mechanism of the reaction. Biochemistry 15: 1079-1083

Reed, B.C., and Rilling, H.C. (1976) Substrate binding of avian liver prenyltransferase. Biochemistry 15: 3739-3745

Reed, B.C. and Rilling, H.C. (1975) Crystallization and partial characterization of prenyltransferase from avian liver. Biochemistry 14: 50-54

Rost, B., Casadio, R., Fariselli, P., and Sander, C. (unpublished)

Rost, B. and Sander, C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. Proteins, 19:55-72

Saito, A. and Rilling, H.C (1978) Prenyltransferase - product binding and reaction termination. J. Biol. Chem. 254: 8511-8515

Sagami, H., Morita, Y., and Ogura, K. (1994) Purification and properties of geranylgeranyl-diphosphate synthase from bovine brain. J. Biol. Chem. 269: 20561-20566

Sheares, B.T., White, S.S., Molowa, D.T., Chan, K., Ding, V.D.H., Kroon, P.A., Bostedor, R.G., and Karkas, J.D. (1989) Cloning, analysis, and bacterial expression of human farnesyl pyrophosphate synthetase and its regulation in Hep G2 cells. Biochemistry 28: 8129-8135

Shinka, T., Ogura, K., and Seto, S. (1975) Comparative specificity of geranylgeranyl pyrophosphate synthetase of *Micrococcus lysodeikticus* and pumpkin. J. Biochem. 78: 1177-1171

Song, L., Poulter, C.D. (1994) Yeast farnesyl-diphosphate synthase: site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II. Proc. Natl. Acad. Sci. USA 91: 3044-3048

Spear, D.H., Kutsuani, S.Y., Correll, C.C., and Edwards, P.A. (1992) Molecular cloning and promoter analysis of the rat liver farnesyl diphosphate synthase gene. J. Biol. Chem. 267: 14462-14469

Tachibana, A., Tanaka, T., Taniguchi, M., and Oi, S. (1993) Purification and characterization of geranylgeranyl diphosphate synthase from methanobacterium thermoformicicum SF-4. Biosci. Biotech. Biochem. 57: 1129-1133

Tarshis, L.C., Yan, M., Poulter, C.D., and Sacchettini, J.C. (1994) Crystal structure of recombinant farnesyl diphosphate synthase at 2,6A resolution. Biochemistry 33: 10871-10877

Teruya, J.H., Kutsunai, S.Y., Spear, D.H., Edwards, P.A., and Clarke, C.F. (1990) Testis-specific transcription initiation of rat farnesyl pyrophosphate synthetase mRNA. Mol. Cell. Biol. 10: 2315-2326

Weissenborn, D.L., Denbow, C.J., Laine, M., Lang, S.S., Yu, X., and Cramer, C.L. (1995) HMG-CoA reductase and terpenoid phytoalexins: molecular specialization within a complex pathway. Plant Physiol.

Wiedemann, M., Misawa, N., and Sandmann, G. (1993) Purification and enzymatic characterization of the geranylgeranyl pyrophosphate synthase from erwinia uredovora after expression in escherichia coli. Arch. Biochem. Biophys. 306: 152-157

Wilkin, D.J., Kutsunai, S.Y., and Edwards, P.A. (1990) Isolation and sequence of the human farnesyl pyrophosphate synthetase cDNA. J. Biol. Chem. 265: 4607-4614

Yeh, L.S. and Rilling, H.C. (1977) Purification and properties of pig liver prenyltransferase: interconvertible forms of the enzyme. Arch. Biochem. Biophys. 183: 718-725