

Manuscript for:
FEBS Letters

Send correspondence to:
Sophie Vandermoten
Gembloux Agricultural University
Functional and Evolutionary Entomology
Passage des Déportés, 2
B-5030 Gembloux, BELGIUM
T. +32(0)81 622 287
F. +32(0)81 622 312
email : entomologie@fsagx.ac.be

Characterization of a novel aphid prenyltransferase displaying dual geranyl/farnesyl diphosphate synthase activity

Sophie Vandermoten^{1*}, Benoit Charlotheaux², Sébastien Santini², Stephanie E. Sen³, Catherine Béliveau⁴,
Micheline Vandenbol⁵, Frédéric Francis¹, Robert Brasseur², Michel Cusson^{4†} and Éric Haubruge^{1†}

¹ Gembloux Agricultural University, Department of Functional and Evolutionary Entomology, Passage des Déportés 2, B-5030 Gembloux, Belgium

² Gembloux Agricultural University, Center of Numerical Molecular Biophysics, Passage des Déportés 2, B-5030 Gembloux, Belgium

³ Department of Chemistry, Indiana University-Purdue University at Indianapolis (IUPUI) 402 North Blackford Street, Indianapolis, IN, 46202, USA

⁴ Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 rue du PEPS, Québec, Québec G1V4C7, Canada

⁵ Gembloux Agricultural University, Department of Animal and Microbial Biology, Passage des Déportés 2, B-5030 Gembloux, Belgium

* Author for correspondence

† These authors co-supervised the work of SV with similar amounts of work conducted at the two affiliated institutions

Running Title: Aphid geranyl/farnesyl diphosphate synthase

Key words: aphid, isoprenyl diphosphate synthase, geranyl diphosphate synthase, farnesyl diphosphate synthase, bifunctional enzyme, homology modelling.

Abstract

We report on the cDNA cloning and characterization of a novel short-chain isoprenyl diphosphate synthase from the aphid *Myzus persicae*. Of the three IPPS cDNAs we cloned, two yielded prenyltransferase activity following expression in *E. coli*; these cDNAs encode identical proteins except for the presence, in one of them, of an N-terminal mitochondrial targeting peptide. Although the aphid enzyme was predicted to be a farnesyl diphosphate synthase by BLASTP analysis, rMpIPPS, when supplied isopentenyl diphosphate and dimethylallyl diphosphate as substrates, typically generated geranyl diphosphate (C10) as its main product, along with significant quantities of farnesyl diphosphate (C15). Analysis of an MpIPPS homology model pointed to substitutions that could confer GPP/FPP synthase activity to the aphid enzyme.

Introduction

Short-chain isoprenyl diphosphate synthases (scIPPSs) are a class of prenyltransferases that catalyze the condensation of isopentenyl diphosphate (IPP; C5) with an allylic co-substrate. These enzymes are distinguished by the specific chain length of the prenyl diphosphate product generated. Thus, geranyl diphosphate synthase (GPPS) catalyzes a single condensation of IPP with dimethylallyl diphosphate (DMAPP; C5) to form geranyl diphosphate (GPP; C10), whereas farnesyl diphosphate synthase (FPPS) and geranylgeranyl diphosphate synthase (GGPPS) catalyze one and two additional condensations with IPP, respectively, generating farnesyl diphosphate (FPP; C15) and geranylgeranyl diphosphate (GGPP; C20). These products are precursors to many biologically essential compounds classified in three groups: monoterpenes, sesquiterpenes and diterpenes, respectively [1] (Figure 1).

The deduced amino acid sequences of all scIPPSs display high similarity, and sequence alignments reveal seven conserved regions [2]. Regions II and VI contain the substrate-binding regions, with two characteristic aspartate-rich motifs (DDx(xx)xD; where x represents any amino acid residue). Region II, which includes the first aspartate-rich motif (FARM), is designated as the chain length determination region, whereas, the second aspartate-rich motif (SARM) in region VI is responsible for IPP binding [3].

In insects, FPPS is involved in the biosynthetic pathway of the sesquiterpenoid juvenile hormone, which has important morphogenetic and gonadotropic functions [4]. In many aphid species, FPPS is also predicted to play a key role in the biosynthetic pathway of the alarm pheromone E- β -farnesene [5].

Whereas FPPS and GGPPS occur nearly ubiquitously in plants, animals, fungi and bacteria, GPPS appears to have a more limited distribution in nature, having been identified

almost exclusively in essential oil (monoterpene)-producing plants. However, monoterpenes are also important natural products in some insects, serving as semiochemicals. Recently evidence for the existence of GPPS activity in insects was provided for scolytid [6] and chrysomelid beetles [7]. In aphids, the presence of a GPPS is also expected inasmuch as the cyclopentanoids nepetalactone and nepetalactol, which are derived from GPP, are essential components of the sex pheromone [8].

In spite of the presumed importance of scIPPSs in insect metabolism, a limited number of insect IPPS sequences have been reported to date, and characterization of the recombinant proteins has only recently begun [6,9-13].

In this paper, we report on the cloning of three scIPPS cDNAs from *M. persicae* (two long forms, MpIPPS1-L and MpIPPS2-L and a short form, MpIPPS1-S), on their heterologous expression and on the characterization of one of these isoforms, including the identification of its principal *in vitro* products, prediction of its 3D-structure, and analysis of the generated 3D models. Although MpIPPS1 isoforms were initially predicted to be FPPSs, based on their amino acid sequences, *in vitro* assays revealed a dual GPP/FPP synthase activity, generating both GPP and FPP.

Materials and Methods

Biological material

Myzus persicae individuals were reared on bean plants (*Vicia faba* L.) at $20 \pm 2^\circ\text{C}$, under a 16 h L:8 h D photoperiod, in a dedicated environmental chamber.

Cloning and sequencing of M. persicae IPPS cDNAs

Total RNA was isolated from whole-body *M. persicae* individuals, using the Trizol reagent (Invitrogen) according to manufacturer's instructions. We constructed a cDNA library as described [14]. Briefly, 3 μg of total RNA was reverse transcribed using an oligo dT primer with the following sequence: TTTTGTACAAGC(T)₁₆, followed by synthesis of the second cDNA strand and ligation of an adaptor (upper primer; 5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCCGCCGGGCAGGT-3' and lower primer 5'-ACCTGCCCGG-3'); the latter was used for amplification of the cDNA using an adaptor-specific primer (ASP; 5'-CTAATACGACTCACTATAGGGC-3') in conjunction with the oligo dT primer.

To amplify a core fragment, a set of degenerate primers were designed against highly conserved regions among IPPS proteins [11]. The primer set included the forward primers P1, 5'-TAYAAAYRYIVBIVRIGGIAARHWIAAYMGNGG-3', and P3, 5'-AYIMGIMGIGGIVMIMYITGYTGG-3', and the reverse primer P2, 5'-RTCYTGDATRTCIGTICCIWYYTTNCC-3'. Two PCR reactions were successively performed, using the primer pairs P1-P2 and P3-P2.

To amplify the 5'- and 3'-ends of the cDNA fragments, specific primers (SP) were designed against the partial sequences obtained and utilized in conjunction with ASP and oligo dT primers. For amplification of MpIPPS1 isoforms, the specific primer sequence used

for 5' RACE was SP-1, 5'-AAGCAGTCTTATATTTTACAATAGC-3', and those used for 3'-RACE were, SP-5, 5'-GACCCAGAAATCTTTCGACAAGC-3', and SP-6, 5'-GAAGATGGTAAATGTTCATGGT -3'. For amplification of the MpIPPS2-L homolog, 5'-RACE, was performed using SP-1, and for 3'-RACE the specific primer sequence was, SP-7, 5'-TATACACACATTCTGGAAGTGTTTTATGACG-3'.

Construction of expression vectors

After amplification by PCR, using specific primers to introduce *Eco*RI and *Xho*I restriction sites at the 5' and 3' ends, respectively, the coding region of each cDNA was cloned, in phase with a His-tag at the N-terminus, into the expression vector pET30b (Novagen) previously cut with the same restriction enzymes. Fidelity of the inserts was confirmed by sequencing.

Expression and purification of recombinant proteins

Escherichia coli strain BL21 (DE3) was transformed with each MpIPPS construct. Single colony transformants were cultured at 37°C for 3 h in 10 ml of LB medium containing 1% glucose and 30 µg/ml kanamycin, with shaking at 200 rpm. This starter culture was used to inoculate a flask containing 200 ml of the same medium, and cells were grown under the same conditions to an OD₆₀₀ of ~ 0.6. Induction was initiated by addition of 1 mM isopropyl-β-D-thiogalactoside (IPTG), after replacing the medium with 200 ml of LB medium containing 30 µg/ml kanamycin. The cells were incubated overnight at 15°C with shaking at 150 rpm, and then harvested by centrifugation (3000 g, 10 min, 4°C).

After resuspending the resulting bacterial pellet in 10 ml of 25 mM Tris-HCl buffer pH 8, containing 5 mM MgCl₂ and 2 mM β-mercaptoethanol (BME), the cells were first treated with lysozyme (1 mg) for 10 min at room temperature, and then sonicated for 3 min

using 1 sec bursts at 4°C (Branson sonicator, 80W, 20% power). The clarified sample was subjected to Co²⁺ metal affinity chromatography using TALON Resin (1 ml, Clontech), which had been equilibrated with binding buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM MgCl₂, 5 mM imidazole and 2 mM BME). The column was washed with 4 ml of binding buffer, followed by 4 ml of wash buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM MgCl₂, 25 mM imidazole, 2 mM BME). Proteins were eluted with 25 mM Tris-HCl pH 8 containing 150 mM NaCl, 5 mM MgCl₂, 150 mM imidazole and 2 mM BME. The protein fractions were finally placed in 50 mM Tris-HCl pH 7 containing 2 mM BME, using a PD-10 desalting column (Amersham Biosciences). The purity of each rMpIPPS was verified by SDS-PAGE and protein concentration was determined by the microassay procedure of the Bio-Rad Protein Assay, using BSA as standard. The procedure described above was also used to produce the previously characterized FPPS of *Drosophila melanogaster* (rDmFPPS) [13]. Thrombine cleavage of the His-tag was carried out as described [13].

Prenyltransferase product analysis

To identify the products formed *in vitro* by rMpIPPS and to compare them to those formed by rDmFPPS, the purified enzyme suspensions (0.5 µg) were assayed in 50 mM Tris-HCl, pH 7, containing 50 mM MgCl₂ and 2 mM BME, using 10 to 100 µM allylic substrate (DMAPP or GPP) and 4 µM [¹⁴C]IPP, in a total volume of 50 µl. The reactions were incubated for 40 min at 37°C. Samples were further incubated overnight at 37°C in the presence of 2 U shrimp alkaline phosphatase in 10x phosphatase buffer (Roche), 2% Triton X-100, 5% glycerol and 2.5 mg/ml BSA, to convert the diphosphates into the corresponding alcohols. Acetonitrile (final concentration of 50%) and authentic standards (geraniol, farnesol and geranylgeraniol; Sigma) were added to the samples prior to HPLC analysis. Reverse-phase HPLC was carried out on a Dionex quaternary pump equipped with a C-18 column (25

cm x 4.6 mm, 5 μ m; Supelco). Products were fractionated using an acetonitrile-water gradient (60 to 70% ACN over 10 min; 70 to 80% ACN over 20 min; held at 80% ACN for 15 min) at a flow rate of 1 ml/min, with standards detected at 210 nm. The amount of radioactivity in each 1-min fraction was assessed by liquid scintillation counting using a Beckman LS6000SC.

Kinetic studies

The K_m values for GPP, DMAPP and IPP were obtained by double reciprocal Lineweaver–Burk plots of the amount of total allylic diphosphate product formed at increasing concentrations of substrate. K_m values for GPP and DMAPP were determined in the range of 4–200 μ M, with a fixed IPP concentration of 4 μ M. The K_m value for IPP was determined for concentrations between 0.5 and 12 μ M, at 100 μ M GPP. Assays were conducted with 5.5 μ g of recombinant proteins in a final volume of 25 μ l, using 50 mM Tris-HCl pH 7 containing 50 mM $MgCl_2$, 2 mM BME. The reaction was incubated for 20 min at 37°C, and then quenched by the addition of 25 μ l of 2.5 M HCl, which served to hydrolyze the allylic diphosphate formed. After 10 min the reaction was partitioned between aqueous and organic phase by extraction with diethyl ether/pentane (1:1, v:v) and analyzed by liquid scintillation counting.

Molecular Modelling

The aphid IPPS displays 47% sequence identity to chicken; 3D models of the MpIPPS1-S dimer (residues 15-362) with bound DMAPP, GPP, or FPP were thus developed by homology modelling using MODELLER [15] and the appropriate PDB avian structures as template (1UBX, 1UBY, and 1UBW). Homology models were then minimized using GROMACS [16] and successfully validated with PROCHECK [17]. Models were generated

similarly for the WT sequence of the avian enzyme; the PDB files used as templates correspond to a double mutant.

All residues located at less than 5 Å from DMAPP, GPP or FPP in the avian and in the *M. persicae* 3D models were identified with PEX [18]. These interactions were determined for the dimers. These residues were then sorted out with respect to the portion(s) of the ligand involved in the interaction (OPP, C1-C5, C6-C10, and C11-C15).

Results

Molecular cloning of M. persicae IPPS

The degenerate PCR cloning approach that we employed permitted amplification of three separate cDNAs, referred to here as MpIPPS1-L, MpIPPS1-S and MpIPPS2-L. These encode proteins of 395 (45.3 kDa), 363 (41.6 kDa) and 396 (45.7kDa) amino acids, with calculated pI's of 7.5, 5.83 and 8.23, respectively.

The two MpIPPS1 isoforms are identical except for the presence of an N-terminal extension of 32 amino acids in MpIPPS-L, a feature also present in MpIPPS2-L (Figure 2). As described for *Agrotis ipsilon* and *D. melanogaster* FPPSs, the N-terminal extensions of aphid isoforms have the features of a mitochondrial targeting motif [19] and display the RxxS consensus sequence, which is currently reported as a mitochondrial targeting cleavage motif [20].

On the basis of BLASTp analyses, the proteins reported here were initially identified as FPPSs. A CLUSTAL alignment revealed that the amino acid sequences show ~50% similarity to FPPSs of other organisms. The predicted translated products have the seven conserved sequence regions identified by Koyama et al. [2], including the two aspartate-rich motifs (regions II and VI) (Figure 2).

Functional expression and product formation

Both the long and short isoforms of rMpIPPS1 displayed significant catalytic activity in the presence of IPP and either allylic substrate, but rMpIPPS2, with or without its N-terminal extension, failed to show any detectable activity under the same experimental conditions; as a consequence, the present report focuses on the characterization of rMpIPPS1. The specific activity of rMpIPPS1-S was approximately four times greater than that of

rMpIPPS1-L. This may result from the mitochondrial targeting peptide in MpIPPS-L affecting protein folding and catalysis; the mature mitochondrial enzyme is predicted to lose its N-terminal extension after gaining access into its subcellular target [19], suggesting that MpIPPS1-L is not the catalytically functional (i.e. mature) protein. For this reason, only rMpIPPS1-S was used in subsequent assays.

The investigation of the nature of the end-products formed *in vitro* was performed using rMpIPPS1-S and DmFPPS for comparison. Unexpectedly, the recombinant aphid enzyme generally yielded GPP as its major product, a tendency that became more pronounced with increasing concentrations of DMAPP, while DmFPPS catalyzed almost exclusively FPP formation (Figures 3A & 3B). The observed predominance of GPP is unlikely to be an artifact of our assay conditions, given that similar product profiles were obtained under long incubations and that cleavage of the His-tag altered neither the activity nor the product profile (data not shown). Under conditions where GPP was utilized as allylic substrate, only FPP was formed, with trace amounts of GGPP (Figure 3C).

Preliminary assessments of K_m values for rMpIPPS1-S ($K_{m/GPP} = 68 \mu\text{M}$, $K_{m/DMAPP} = 45 \mu\text{M}$ and $K_{m/IPP} = 15 \mu\text{M}$) indicate that the $K_{m/GPP}$ value reported here for this enzyme may be higher than that reported earlier for insect FPPSs [13,21,22], which suggests that the DMAPP substrate might compete with nascent polyprenyl chain and thus induce the abortion of the condensation reaction and the subsequent release of GPP.

Prediction and analysis of MpIPPS1-S structures

In order to identify the active-site residues that might be responsible for the end-product profile of the aphid enzyme, we compared the active site of the aphid and avian enzymes in dimeric 3D models generated by homology modelling.

Since DMAPP, GPP and FPP all contain the same diphosphate group and first five carbon atoms, we postulated that differences in the binding of allylic compounds should not be related to the substitution of residues located in the vicinity of these portions of the ligand. We thus focused our analysis on the residues that are close to the end of the allylic chain of GPP or FPP (from C6 to C15). This assumption appeared to be correct since only the Ser123/Ala117 substitution (avian/aphid numbering) is involved in interactions with the unconsidered portions of the ligands.

Comparison of aphid and avian active-sites

Seven positions (named “A” to “G” in Figure 5) of the aphid/avian sequence alignment correspond to a residue that is in direct contact with the GPP/FPP chain end but that is not conserved between the two species (Table 1 & Figure 4B). Six belong to the A chain and one is on the B chain (position E). In both enzymes, the 7 residues are all close to the FPP C11-C15 group. Most of them are also close to the C6-C10 group of GPP and FPP. Four substitutions are located in conserved region II (positions A to D). These residues are the ones whose side chains are closest to the end of the allylic chain in the aphid enzyme. The Phe113/Gln107 non-conservative substitution (position C) corresponds to the closest one in both enzymes. In the active-site of the aphid enzyme, Tyr106 from the A chain and Val141 from the B chain are further from the ligand than the corresponding Phe112 and Phe147 avian residues (positions B and E). In contrast, the aphid Leu110 residue is much closer to the FPP C11-C15 portion than the corresponding avian residue (Ala116 – position D).

Analysis of a multiple alignment from various organisms

In order to assess the significance of the substitutions reported above, we analyzed amino acid variation at the 7 positions referred to above in a multiple alignment of IPPSs

from various organisms (Figure 5). Located upstream of the first DDx(xx)xD motif, positions A to D are in a region more conserved than positions E to G. Positions F and G are the most variable with non-conservative substitutions seen even among related organisms. Residues at position E are also poorly conserved even among related organisms. Nevertheless, most of the residues at these positions are hydrophobic. The A position is always occupied by a bulky hydrophobic residue. Most of the residues at the B position are aromatic; a Tyr residue is observed in all aphid sequences as well as in those of some other organisms. Some variation is observed at the C position, where the aphid sequences display a Gln residue that contrasts with the aromatic residue generally observed in other sequences. It should be noted, however, that the aphid enzyme shares this feature with the *Choristoneura fumiferana* type-1 FPPS [11], the exact product(s) of which remain to be determined [12]. With the exception of *T. cruzi*, *E. coli*, *B. stearothersophilus* and *D. melanogaster*, the D position contains a hydrophobic residue with a side chain of variable size.

Discussion

Although prenyltransferases capable of producing both FPP and GGPP were reported earlier for two archeas [23] and more recently for the protozoan parasite *Toxoplasma gondii* [24], our study conducted on the aphid *M. persicae* led to the identification of the first prenyltransferase that can generate both GPP and FPP. In general FPPSs have been reported to yield about 3 times or more FPP than GPP [13,25-27]. Unexpectedly, rMpIPPS1-S yielded a GPP:FPP ratio varying between ~50:50 and ~70:30, depending on DMAPP concentration, while the dipteran rDmFPPS produced about 90% FPP. The latter observation suggests that rMpIPPS1 plays a predominant role in GPP formation, and that its function is not limited to FPP production, as originally hypothesized. Nevertheless, unlike other GPPSs known, the rMpIPPS1-S was shown here to be capable of using GPP as allylic co-substrate to generate significant amounts of FPP. Under our *in vitro* assay conditions, the GPP/FPP ratio was dependent on DMAPP concentration. At higher DMAPP concentration, the rMpIPPS1-S produced increasing amounts of GPP. This observation is in agreement with earlier reports indicating that the product profile of scIPPSs may be, to a certain degree, modulated by substrate concentration [28]. *In vivo*, the DMAPP concentration could be under the control of IPP isomerase, which catalyzes the isomerization of IPP to DMAPP, or could be regulated by substrate compartmentalization.

Analyses of the homology models we generated for MpIPPS1-S revealed that, with the exception of Ser123/Ala117, all the amino acid residues interacting with the diphosphate moiety are conserved between MpIPPS1-S and the avian FPPS. The amino acid residues interacting with the C6-C15 portion of the ligand are, however, less conserved, including seven substitutions among which one or several could account for the end-product distribution that we observed.

Some of the positions identified here (Phe112/Tyr106, Phe113/Gln107, Ala116/Leu110) have been shown to be involved in product chain-length selectivity in the avian enzyme, using site-directed mutagenesis [29,30].

Among these positions, the Phe113/Gln107 and the Ala116/Leu110 substitutions are the most likely to be associated with the specific product distribution observed for the aphid enzyme because of the proximity of the corresponding residues from the allylic chain, the nature of the substitution, and the variations observed in the sequence alignment of FPPSs sequence from various organisms.

The Ala116 to Leu110 substitution would reduce the size of the active-site cavity disfavoring the production of longer allylic compounds as a result of steric hindrance, according to the observations of Poulter and coworkers [29,30]. However, this explanation appears to be unlikely since several other FPPSs also have a Leu or a similar residue (Ile, Met) at this position (Figure 5). Thus, the most striking singularity in the residues lining the active-site of the aphid enzyme appears to be the Gln107 at the C position, a substitution that represents a slight decrease in steric hindrance relative to the aromatic residue typically found at this position, but a change in the nature of the residue's side chain (polar instead of hydrophobic). This modification could account for the observed aphid product distribution according to two complementary hypotheses. First, the presence of a Gln residue at this position would generate unfavorable interactions between the polar side-chain of the residue and the end of the GPP/FPP hydrophobic allylic chain of the product [31]. As a result shorter products would be favored. Second, chain elongation in the active-site would require the disruption of hydrogen bonds between the Gln107 amide group and the neighboring residues which would favor the production of products with shorter allylic chains.

Whether the Phe113/Gln107 substitution, alone or in combination with the other substitutions identified here, is responsible for the dual GPP/FPP synthase activity of the

enzyme will need to be assessed by site-directed mutagenesis coupled with appropriate enzyme assays. The possibility that other residues, outside the active site, may be involved in defining this enzyme's product selectivity also needs to be examined.

In summary, the existence of a GPP/FPP synthase in aphids is perhaps not too surprising, given that these insects require both monoterpenes and sesquiterpenes as precursors for sex pheromone, alarm pheromone and juvenile hormone biosynthesis. Whether another aphid scIPPS exists that is dedicated to FPP synthesis is not clear at this stage; MpIPPS2 is certainly a good candidate but we yet have to identify the conditions under which it is active so that its product selectivity may be characterized.

Acknowledgments

This work was supported by a PhD scholarship to SV, awarded by the “Fonds pour la formation à la Recherche dans l’Industrie et l’Agriculture” (Belgium), as well as by a grant to MC from the Canadian Biotechnology Strategy Fund. Bioinformatics analyses were supported by the Ministère de la Région Wallonne contract n°215140 (aBUSTEC) and by the Belgian Program on Interuniversity Attraction Poles initiated by the Federal Office for Scientific, Technical and Cultural Affairs (IAP. P6/19 PROFUSA). RB thanks the “Fonds National pour la Recherche Scientifique” (Belgium).

References

- [1] Wang, K., Ohnuma, S. (1999) Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. *Trends Biochem. Sci.* 24 445-451.
- [2] Koyama, T., Obata, S., Osabe, M., Takeshita, A., Yokoyama, K., Uchida, M., Nishino, T., Ogura, K. (1993) Thermostable farnesyl diphosphate synthase of *Bacillus stearothermophilus*: molecular cloning, sequence determination, overproduction, and purification. *J. Biochem* 113, 355-363.
- [3] Liang, P.H., Ko, T.Z., Wang, A.H.J. (2002) Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* 269, 3339-3354.
- [4] Wyatt, G.R., Davey, K.G. (1996) Cellular and molecular action of juvenile hormone II. Roles of juvenile hormone in adult insects. *Adv. Insect Physiol.* 26, 1-155.
- [5] Kislow, C., Edwards, L.J. (1972) Repellent odour in aphids. *Nature* 235, 108-109.
- [6] Gilg, A.B., Bearfield, J.C., Titigger, C., Welch, W.H., Blomquist, G.J. (2005) Isolation and functional expression of an animal geranyl diphosphate synthase, and its role in bark beetle pheromone biosynthesis. *Proc. Natl. Acad. Sci., USA* 102, 9760-9765.
- [7] Burse, A., Schmidt, A., Frick, S., Kuhn, J., Gershenzon, J., Boland, W. (2007) Iridoid biosynthesis in *Chrysomelina* larvae: Fat body produces early terpenoid precursors. *Insect Biochem. Mol. Biol.* 37, 255-265.
- [8] Dawson, G.W., Griffiths, D.C., Janes, N.F., Mudd, A., Pickett, J.A., Wadhams, L.J., Woodcock, C.M. (1987) Identification of an aphid sex pheromone. *Nature* 325, 614-616.

- [9] Castillo-Gracia, M., Couillaud, F. (1999) Molecular cloning and tissue expression of an insect farnesyl diphosphate synthase. *Eur. J. Biochem.* 262, 365-370.
- [10] Kikuchi, K., Hirai, M., Shiotsuki, T. (2001) Molecular cloning and tissue distribution of farnesyl pyrophosphate synthase from the silkworm. *J. Insect Biotech. Sericol.* 70, 167-172.
- [11] Cusson, M., Béliveau, C., Sen, S.E., Vandermoten, S., Rutledge, R.G., Stewart, D., Francis, F., Haubruge, E., Rehse, P., Huggins, D.J., Dowling, A.P.G., Grant, G.H. (2006) Characterization and tissue-specific expression of two lepidopteran farnesyl diphosphate synthase homologs: implications for the biosynthesis of ethyl-substituted juvenile hormones. *Proteins* 65, 742-758.
- [12] Sen, S.E., Cusson, M., Trobaugh, C., Béliveau, C., Richard, T., Graham, W., Mimms, A., Roberts, G. (2007) Purification, properties and heteromeric association of type-1 and type-2 lepidopteran farnesyl diphosphate synthases. *Insect Biochem. Mol. Biol.* 37, 819-828.
- [13] Sen, S.E., Trobaugh, C., Beliveau, C., Richard, T., Cusson, M. (2007) Cloning, expression and characterization of a dipteran farnesyl diphosphate synthase. *Insect Biochem. Mol. Biol.* 37, 1198-1206.
- [14] Matz, M.V. (2000) Amplification of representative cDNA samples from microscopic amounts of invertebrate tissue to search for new genes? in: *Green Fluorescent Protein: Applications and Protocols* (Hicks B.W., Eds), pp. 1-21. Humana Press Inc., Totowa, NJ.
- [15] Sali, A., Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.

- [16] Berendsen, H.J.C., van der Spoel, D., van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* 91, 43-56.
- [17] Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.
- [18] Thomas, A., Bouffieux, O., Geurickx, D., Brasseur, R. (2001) Pex, analytical tools for PDB files. I. GF-Pex: basic file to describe a protein. *Proteins* 43, 28-36.
- [19] Martín, D., Piulachs, M.D., Cunillera, N., Ferrer, A., Bellés, X. (2007) Mitochondrial targeting of farnesyl diphosphate synthase is a widespread phenomenon in eukaryotes. *Biochim. Biophys. Acta* 1773, 419-426.
- [20] von Heijne, G., Steppuhn, J., Herrmann, R.G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* 180, 535-545.
- [21] Koyama, T., Matsubara, M., Ogura, K. (1985) Isoprenoid enzyme systems of silkworm. II. Formation of the juvenile hormone skeletons by farnesyl pyrophosphate synthase II. *J. Biochem.* 98, 457-463.
- [22] Sen, S.E., Sperry, A.E. (2002) Partial purification of farnesyl diphosphate synthase from whole-body *Manduca sexta*. *Insect Biochem. Mol. Biol.* 32, 889-899.
- [23] Fujiwara, S., Yamanaka, A., Hirooka, K., Kobayashi, A., Imanaka, T., Fukusaki, E. (2004) Temperature-dependent modulation of farnesyl diphosphate/geranylgeranyl diphosphate synthase from hyperthermophilic archaea. *Biochem. Biophys. Res. Commun.* 325, 1066-1074.
- [24] Ling, Y., Li, Z., Miranda, K., Oldfield, E., Moreno, S.N.J. (2007) The farnesyl diphosphate/geranylgeranyl diphosphate synthase of *Toxoplasma gondii* is a

- bifunctional enzyme and a molecular target of bisphosphonates. *J. Biol. Chem.* 282, 30804-30816.
- [25] Blanchard, L., 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.
- [26] Ohnuma, S., Narita, K., Nakazawa, T., Ishida, C., Takeuchi, Y., Ohto, C., Nishino, T. (1996) A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. *J. Biol. Chem.* 271, 30748-30754.
- [27] Mekkiengkrai, D., Sando, T., Hirooka, K., Sakdapipanich, J., Tanaka, Y., Fukusaki, E., Kobayashi, A. (2004) Cloning and characterization of farnesyl diphosphate synthase from the rubber-producing mushroom *Lactarius chrysorrheus*. *Biosci. Biotechnol. Biochem.* 68, 2360-2368.
- [28] Ohnuma, S., Hirooka, K., Ohto, C., Nishino, T. (1997) Conversion from archaeal geranylgeranyl diphosphate synthase to farnesyl diphosphate synthase. Two amino acids before the first aspartate-rich motif solely determine eukaryotic farnesyl diphosphate synthase activity. *J. Biol. Chem.* 272, 5192-5198.
- [29] Tarshis, L.C., Proteau, P.J., Kellogg, B.A., Sacchettini, J.C., Poulter, C.D. (1996) Regulation of product chain length by isoprenyl diphosphate synthases. *Proc. Natl. Acad. Sci. USA* 93, 15018-15023.
- [30] Fernandez, S.M.S., Kellogg, B.A., Poulter, C.D. (2000) Farnesyl diphosphate synthase. Altering the catalytic site to select for geranyl diphosphate activity. *Biochemistry* 39, 15316-15321.

- [31] Lins, F., Brasseur, R. (1995) The hydrophobic effect in protein folding. *FASEB J.* 9, 535-540.

Figure 1. Abbreviated biosynthetic pathways for monoterpenes, sesquiterpenes and diterpenes, including the reactions catalyzed by short-chain isoprenyl diphosphate synthases (scIPPSs).

Figure 2. The cDNA coding region of *M. persicae* IPPS isoforms. The RxxS mitochondrial targeting cleavage motif present at the N terminus is denoted by a blue box. The conserved regions (I–VII) are denoted by black boxes. Positions of the Asp residues within the two aspartate-rich motifs (II–VI) common to all scIPPS are denoted by red boxes.

Figure 3. Radio-HPLC analysis of prenyltransferase products, following their conversions to the corresponding alcohols. (A) rDmFPPS incubated in the presence of [¹⁴C]IPP and DMAPP, (B) rMpIPPS-S incubated in the presence of [¹⁴C]IPP and DMAPP, (C) rMpIPPS-S incubated in the presence of [¹⁴C]IPP and GPP. Values are the mean and standard deviation (error bars) of determinations made from three separate protein preparations.

Figure 4: (A) Overall structure of the *M. persicae* homodimeric model (red) with the superimposed avian homodimeric model (green). Among the amino acid residues interacting with the C6-C15 region of the FPP allylic chain, those that show differences in identity at equivalent positions between the two proteins have their side chains displayed in blue. The FPP molecule is drawn as thin lines. (B) Detailed stereo-view of the FPP C6-C15 interaction site in the A subunit (*M. persicae* residues in red, chicken in green). Residues are labeled with respect to the *M. persicae* sequence. Note that Val-141 is on the B chain (i.e. subunit). Hydrogens are omitted.

Figure 5: Sequence alignment of aphid IPPSs (including MpIPPS) with FPPS from different organisms and *Ips pini* GPPS. Only the regions that include the seven positions with residues that are candidates for a role in the product selectivity of the aphid enzyme are presented.

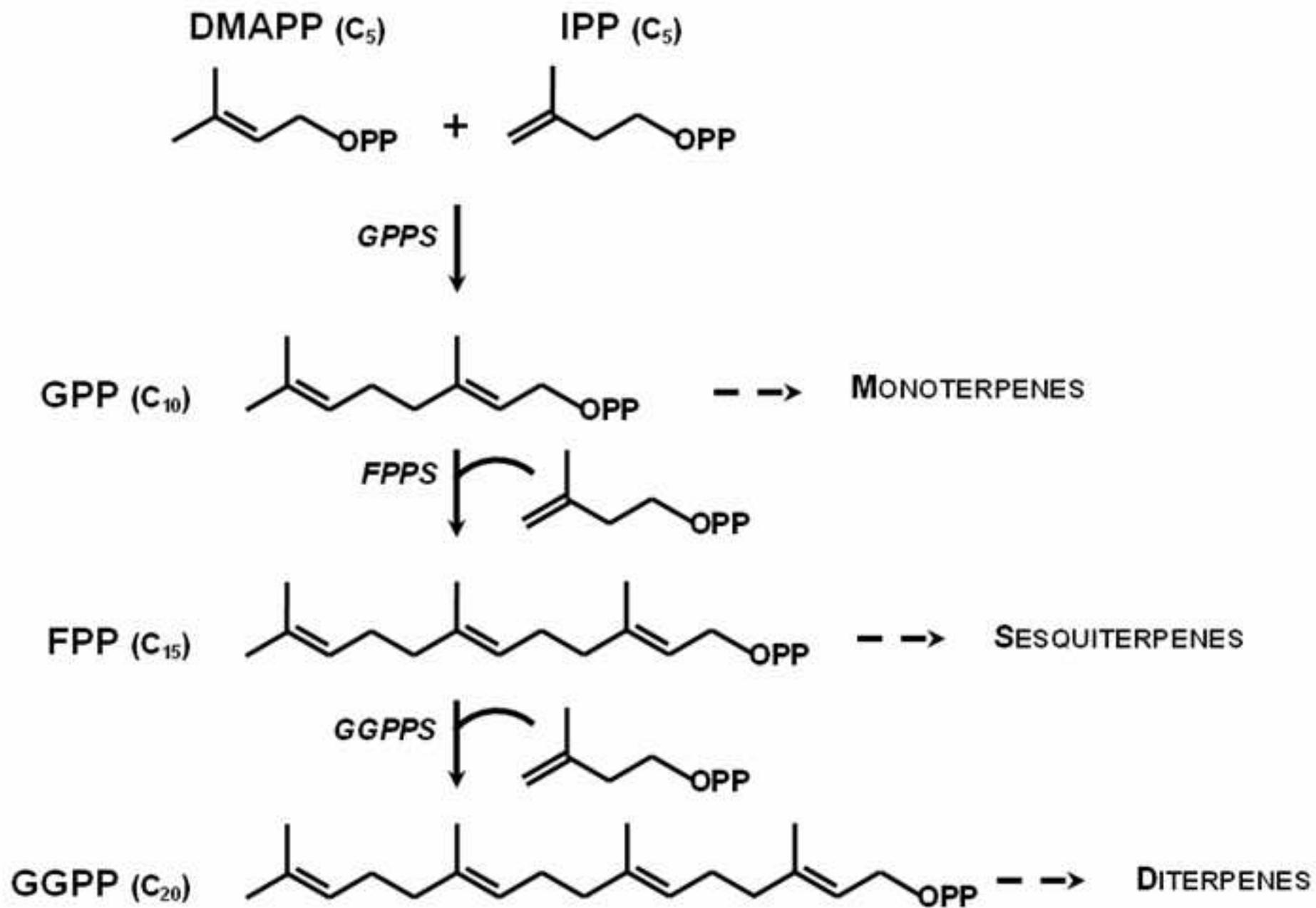
These residues are highlighted in black and designated “A” to “G” at the top of the alignment. Sequence sources: (1) ZP_01700053; (2) Q08291; (3) AAK71861; (4) P08524; (5) AAH10004; (6) NP_114028; (7) P08836 ; (8) NP_493027; (9) AAB07248; (10) AY968586; (11) EU447786; (12) AY968585; (13) AY968583; (14) AY968584; (15) AAD27853; (16) XP_308653; (17) XP_396224; (18) AY966009; (19) AY953507; (20) AY954920; (21) AY954919; (22) AAX55632.

Table 1: Distances (in Å) between the side chains of selected amino acid residues and GPP and FPP allylic ligands in the aphid homodimeric IPPS active site, as compared to distances determined for residues at equivalent positions in the avian structure. Residues at an identical position in the sequence alignment are presented on the same line. The chain (i.e., subunit) bearing the residue is indicated in the left-most column. The distances correspond to the distance between the mass center of the nearest atoms of the corresponding groups. Distances above the 5 Å cutoff are omitted. Results are presented only for one ligand because of the symmetry of the dimer.

Table 1:

		APHID			AVIAN			
		C6-C10		C11-C15	RESIDUE	C6-C10		C11-C15
RESIDUE		GPP	FPP	FPP		GPP	FPP	FPP
Chain A	LEU 103			2.5	PHE 109	3.4		3.3
	TYR 106	4.7	4.5	2.3	PHE 112	2.9	2.7	2.2
	GLN 107	2.1	2.1	1.5	PHE 113	2.1	1.7	1.8
	LEU 110	2.0	2.3	2.3	ALA 116	2.6	2.1	4.4
	THR 172	3.8	4.7	2.6	ALA 178	3.6		2.4
	ALA 176	2.7	3.2	3.1	GLU 182	2.0	2.9	2.3
Chain B	VAL 141	4.1	4.4	3.7	PHE 147	2.4	3.2	2.6

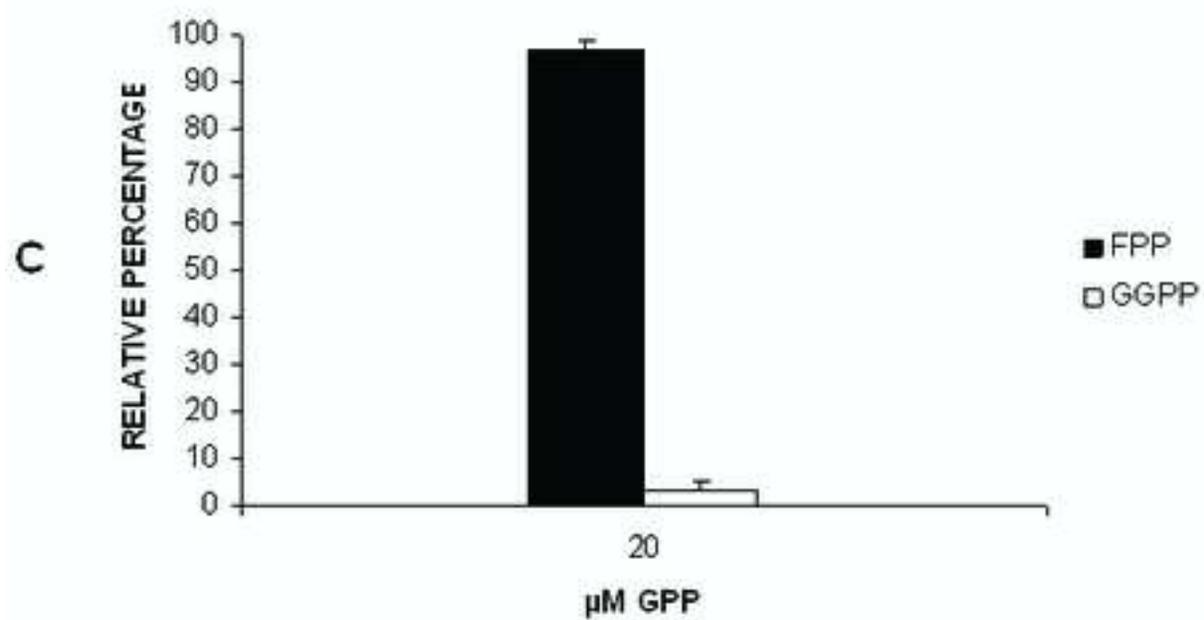
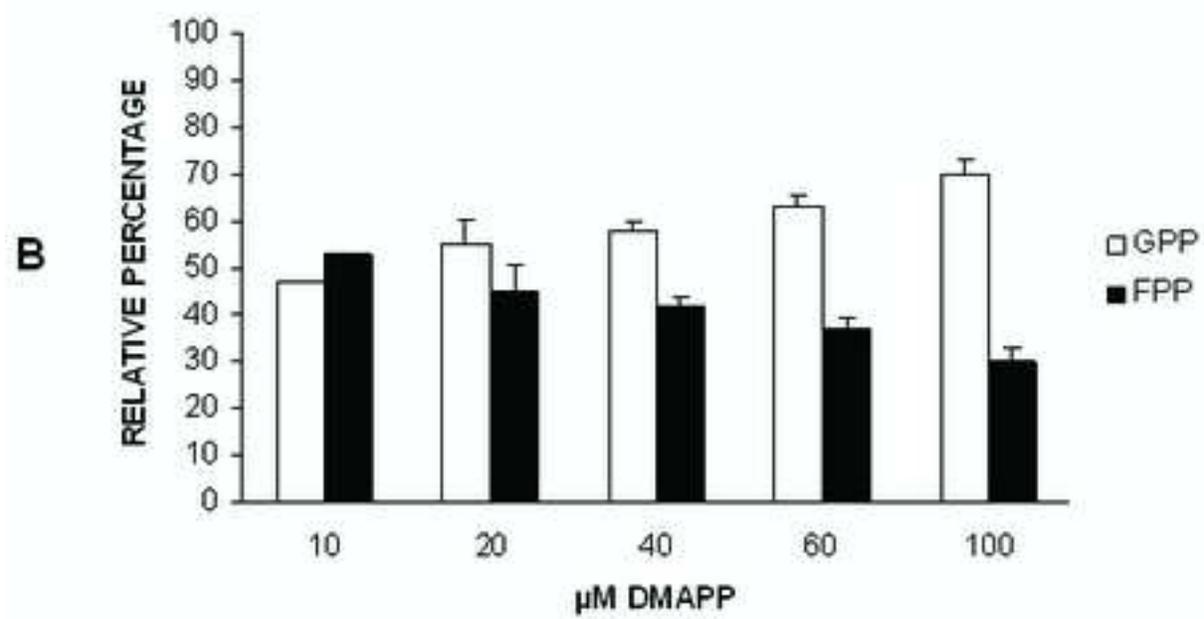
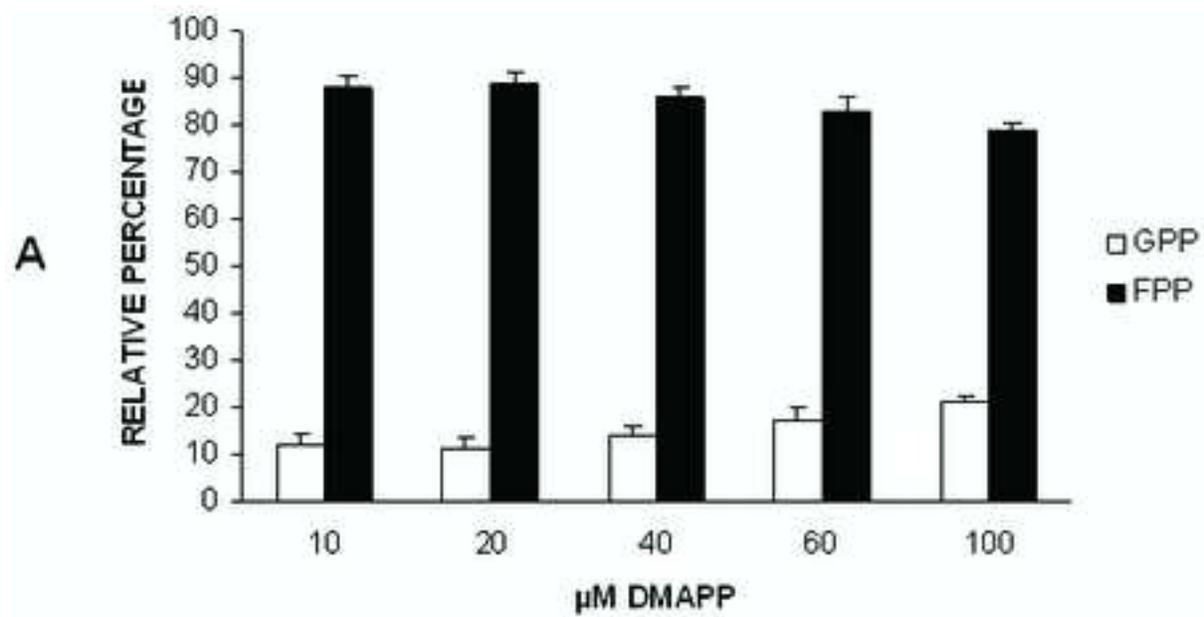
Figure(s)
[Click here to download high resolution image](#)



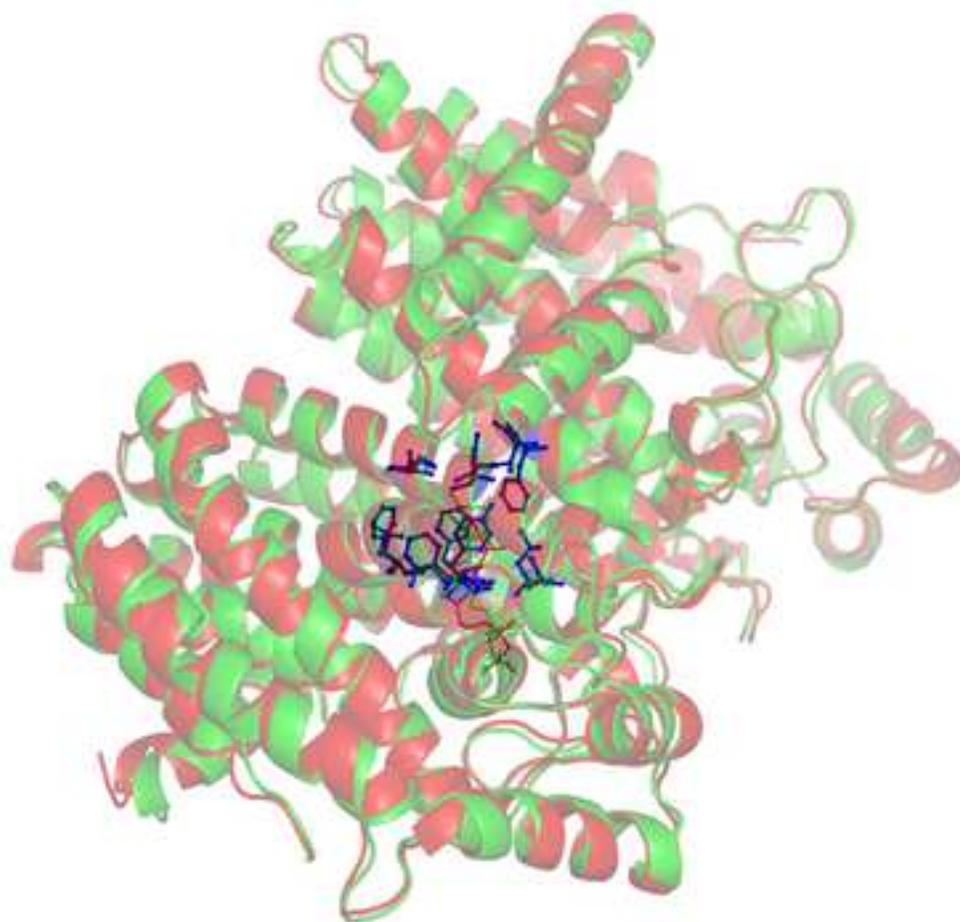
Figure(s)

[Click here to download high resolution image](#)

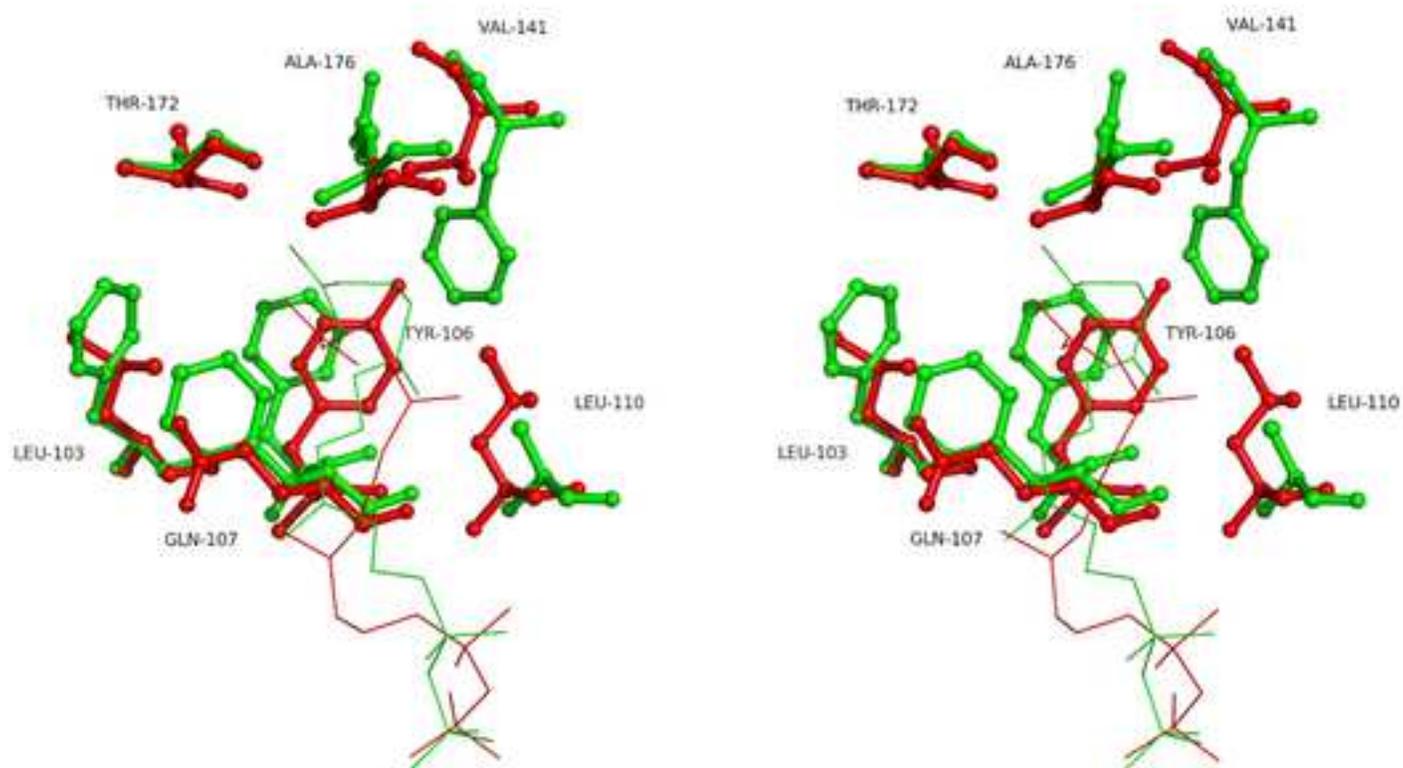
	<u>RxxS</u>							
MpIPPS1-L	MNKMLT-FTR	ALSRRSAFLL	CDSAVV-RNN	NE <u>RAMS</u> TVRA	PFVPPVITGT	AVSKDETRDF	MAVF PDVVRD	LTDTGRNLDV
MpIPPS1-S	-----	-----	-----	----MSTVRA	PFVPPVITGT	AVSKDETRDF	MAVF PDVVRD	LTDTGRNLDV
MpIPPS2-L	MYKLTTFTR	ALSRRTAFLS	CPAVAAGREN	HE <u>RSMS</u> VILP	PFVTSVITGA	LVS KDELGNE	MAVF PDIVKD	LTDTSLQLNK
	I				II			
MpIPPS1-L	PDVTKWLAKL	LQYNVEGGKK	NRGLALVLSY	KMLSSPSDQT	DENLKLSYIL	GWCVEILQAY	QLVLDDIMDN	AITRRGRPCW
MpIPPS1-S	PDVTKWLAKL	LQYNVEGGKK	NRGLALVLSY	KMLSSPSDQT	DENLKLSYIL	GWCVEILQAY	QLVLDDIMDN	AITRRGRPCW
MpIPPS2-L	PDVTKWLENL	LQYNVEGGKN	NRGLILVSSF	KMLSSPSYLT	DENLRLSYIL	GWCVEILQAY	QLVMDDLMDN	AITRRGRPCW
	III			IV				
MpIPPS1-L	YRHNDIGLMA	VNDGVLLLEQA	IYQLIKKYFK	DKPYYTHILE	LFFDVTMKTA	MGQCLDMLTA	NSFKSKKLEK	YTMENYTAIV
MpIPPS1-S	YRHNDIGLMA	VNDGVLLLEQA	IYQLIKKYFK	DKPYYTHILE	LFFDVTMKTA	MGQCLDMLTA	NSFKSKKLEK	YTMENYTAIV
MpIPPS2-L	YRHNDVGLMA	VNDGILLEQA	IYQLLKKYFK	DKPYYTHILE	LFYDVTMKSA	MGQCLDMLTA	KSFKSKKLEK	YTMENYKAIV
	V			VI				
MpIPPS1-L	KYKAYYSFF	LPVCLAMRMT	NINDPEIFRQ	AKTILLEMGH	FFQVQDDFLD	CYGDPDVMGK	IGTDIEDGKC	SWLAVVALQK
MpIPPS1-S	KYKAYYSFF	LPVCLAMRMT	NINDPEIFRQ	AKTILLEMGH	FFQVQDDFLD	CYGDPDVMGK	IGTDIEDGKC	SWLAVVALQK
MpIPPS2-L	KYKAYYSFV	LPVCLAMRMT	NVNDQEIFRQ	AKVILLDMGQ	FFQIQDDYLD	CYGNPEITGK	IGTDIEDGKC	SWLAVKALQK
	VII							
MpIPPS1-L	VNSEQKKIME	DNYGIDDPAN	VAIKDLYAQ	LKLADTFHLY	EEESYKLICT	HIQQLSRGLS	QDMFFKLEK	IYKRTL
MpIPPS1-S	VNSEQKKIME	DNYGIDDPAN	VAIKDLYAQ	LKLADTFHLY	EEESYKLICT	HIQQLSRGLS	QDMFFKLEK	IYKRTL
MpIPPS2-L	VTTEQKKIME	DNYGIDDQFN	VAVIKDLYQQ	LKLQNTFLLY	EEESYNLIRT	CIQNFSPGLS	QDMFFKLEK	IYKRTL



A



B



Figure(s)

[Click here to download high resolution image](#)

	A	BC	D		E		F	G		
<i>Gallus gallus</i> ¹	107	ELFQFFFLVADDI--MD	121		141	DAIN-DSFLLESSVYRV	156	176	QTAYQT--ELGQMLD	188
<i>Homo sapiens</i> ²	159	ELLQFFFLVADDI--MD	173		193	DAIN-DANLLEACIYRL	208	228	QSSYQT--EIGQTLD	240
<i>Rattus rattus</i> ³	93	ELLQFFFLVLDLDDI--MD	107		127	DAIN-DALLLEAAIYRL	142	162	QSSYQT--EIGQTLD	174
<i>Caenorhabditis elegans</i> ⁴	89	EILQSFYLIADDI--MD	103		123	SAIN-DAFIMDSFVEDI	138	158	KSKQKT--LIGQFLD	170
<i>Trypanosoma cruzi</i> ⁵	88	EFLQPHYLMEDDI--MD	102		122	QCAINDGIILKSWTQIM	138	158	KVDYAT--AVGQMYD	170
<i>Saccharomyces cerevisiae</i> ⁶	90	ELLQYFVLVADDM--MD	104		124	IAIN-DAFMLEAAIYKL	139	159	EVTFQT--ELGQLMD	171
<i>Arabidopsis thaliana</i> ⁷	83	EWLQYFVLVLDLDDI--MD	97		117	IAIN-DGILLRNHIHRI	132	152	EVEFQT--ACGQMID	164
<i>Escherichia coli</i> ⁸	74	ECIHPYSLIHDDL PAMD	90		110	AILA-GDALQTLAFSIL	125	147	SASGIAGMCGGQALD	161
<i>Bacillus stearothermophilus</i> ⁹	76	EMINTYSLIHDDL PSM D	92		112	AILA-GDGLLTYAPQLI	127	149	KAAGPEGMVAGQAAD	163
<i>Myzus persicae</i> IPPS1 ¹⁰	101	EILQYQLVLDLDDI--MD	115		135	MAVN-DGVLLEQAIYQL	150	170	DVTMKT--AMGQCLD	182
<i>Myzus persicae</i> IPPS2 ¹¹	135	EILQYQLVMDL--MD	149		169	MAVN-DGILLEQAIYQL	184	204	DVTMKS--AMGQCLD	216
<i>Acyrtosiphon pisum</i> ¹²	101	EILQYQLVLDLDDI--MD	115		135	MAVN-DGVLLEQAIYQL	150	170	DVTMKT--AMGQCLD	182
<i>Aphis fabae</i> ¹³	101	EILQYQLVLDLDDI--MD	115		135	MAVN-DGILLEQTIYQL	150	170	DVTMKT--SMGQCLD	182
<i>Megoura viciae</i> ¹⁴	101	EILQYQLVLDLDDI--MD	115		135	MAVN-DGILLEQTIYQL	150	170	DVAMKT--AMGQCLD	182
<i>Drosophila melanogaster</i> ¹⁵	161	EMLSFFFIISDDV--MD	174		194	TAIN-DALMIENAMYAI	209	229	EITYIT--TCGQSLD	241
<i>Anopheles gambiae</i> ¹⁶	186	EMLHSMFLIMDDV--MD	200		220	SGVN-DALMIEAAIAHL	235	255	EMKFIT--TIGQSLD	267
<i>Apis mellifera</i> ¹⁷	92	EIMQFHTMIDDI--LD	106		124	GAIN-DGLILETCIYKL	139	159	KVSYNTKAAHGECLD	173
<i>Dendroctonus jeffreyi</i> ¹⁸	168	EMVHAYVLILDDI--MD	182		202	TAVN-DAVMIENAVYLL	217	237	EGNLKT--TLGQSLD	249
<i>Ips pini</i> ¹⁹	174	EMLHTYFLIIDDI--ID	188		208	TAVY-DAVMENGVYLL	223	243	DMALKT--SLGQSLD	255
<i>Choristoneura fumiferana</i> (FPPS1) ²⁰	168	EMFHTHQLILNDI--ME	182		202	SSIN-DAILVQSAMYST	217	237	EMLLKC--STGQFLE	249
<i>Choristoneura fumiferana</i> (FPPS2) ²¹	137	EILQGFVLVLDLDDI--MD	151		171	AAVN-DSLSLMFSSIFYV	186	206	ESLMHT--SIGQHLD	218
<i>Ips pini</i> (GPPS) ²²	147	EILQYFCMLDDI--MD	161		181	NAVTDVCLMEMTFEL	196	216	NLLFLT--HMGQGYD	228