

**STRUCTURAL FEATURES CONFERRING DUAL GERANYL/FARNESYL DIPHOSPHATE
SYNTHASE ACTIVITY TO AN APHID PRENYLTRANSFERASE**

**Sophie Vandermoten^{1,6}, Sébastien Santini^{2,4,6}, Éric Haubruge^{1,7}, Fabien Heuze², Frédéric Francis¹,
Robert Brasseur², Michel Cusson^{3,7} and Benoit Charlotiaux^{2,5,6}**

¹ Gembloux Agricultural University, Department of Functional and Evolutionary Entomology, Gembloux, Belgium

² Gembloux Agricultural University, Center of Numerical Molecular Biophysics, Gembloux, Belgium

³ Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Québec, Canada

Present addresses:

⁴ Information Génomique et Structurale, UPR2589, Parc Scientifique de Luminy, Marseille, France

⁵ Center for Cancer Systems Biology (CCSB) and Department of Cancer Biology, Dana-Farber Cancer Institute, and
Department of Genetics, Harvard Medical School, Boston, USA.

Contribution:

⁶These authors contributed equally

⁷These authors co-supervised the work with similar amounts of work conducted at the two affiliated institutions

Send correspondence to: Sophie VANDERMOTEN, Gembloux Agricultural University, Department of
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

ABSTRACT

In addition to providing lipid chains for protein prenylation, short-chain isoprenyl diphosphate synthases (scIPPSs) play a pivotal role in the biosynthesis of numerous mevalonate pathway end-products, including insect juvenile hormone and terpenoid pheromones. For this reason, they are being considered as targets for pesticide development. Recently, we characterized an aphid scIPPS displaying dual geranyl diphosphate (GPP; C₁₀)/farnesyl diphosphate (FPP; C₁₅) synthase activity *in vitro*. To identify the mechanism(s) responsible for this dual activity, we assessed the product selectivity of aphid scIPPSs bearing mutations at Gln107 and/or Leu110, the fourth and first residue upstream from the “first aspartate-rich motif” (FARM), respectively. All but one resulted in significant changes in product chain-length selectivity, effectively increasing the production of either GPP (Q107E, L110W) or FPP (Q107F, Q107F–L110A); the other mutation (L110A) abolished activity. Although some of these effects could be attributed to changes in steric hindrance within the catalytic cavity, molecular dynamics simulations identified other contributing factors, including residue-ligand Van der Waals interactions and the formation of hydrogen bonds or salt bridges between Gln107 and other residues across the catalytic cavity, which constitutes a novel product chain-length determination mechanism for scIPPSs. Thus the aphid enzyme apparently evolved to maintain the capacity to produce both GPP and FPP through a balance between these mechanisms.

KEYWORDS

Aphid, isoprenyl diphosphate synthase, pheromone and juvenile hormone biosynthesis, homology modelling, molecular dynamics.

INTRODUCTION

Isoprenoids form a very large family of compounds that play essential roles in numerous biochemical pathways, *e.g.* as components of electron carriers, as cell membrane constituents, in subcellular targeting and regulation, as hormones, as well as serving in communication and defense in plants and insects. Because isoprenoids are often involved in life-dependent processes, some of the enzymes responsible for their biosynthesis have been used as targets for the development of drugs (*e.g.*, statins, bisphosphonates) and herbicides (Cromartie et al. 1999), and some are being considered as targets for insecticide design (Cusson et al. 2006; Palli and Cusson 2007; Vandermoten et al. 2008b; Vandermoten et al. 2009).

Isoprenoids generated by the mevalonate pathway are built through a common biosynthetic mechanism in which the C₅ dimethylallyl diphosphate (DMAPP) undergoes a 1'-4 condensation with its homoallylic isomer, isopentenyl diphosphate (IPP), to form geranyl diphosphate (GPP, C₁₀), which may then undergo additional condensations with IPP to yield longer products such as farnesyl diphosphate (FPP, C₁₅) and geranylgeranyl diphosphate (GGPP, C₂₀). Generally, each of these diphosphates is generated by a specific short-chain isoprenyl diphosphate synthase (scIPPS; also referred to as “prenyltransferase”) named according to its final product: geranyl diphosphate synthase (GPPS), farnesyl diphosphate synthase (FPPS), and geranylgeranyl diphosphate synthase (GGPPS) (Bellés et al. 2005).

Site-directed and random mutagenesis studies aimed at scIPPSs (Ohnuma et al. 1996a; Ohnuma et al. 1996b; Tarshis et al. 1996; Ohnuma et al. 1997; Narita et al. 1999; Fernandez et al. 2000; Kawasaki et al. 2003), combined with the elucidation of their 3D structures (Tarshis et al. 1994; Kavanagh et al. 2006a; Kavanagh et al. 2006b), have provided insights into the mechanism of product chain-length determination (Liang et al. 2002). Several studies demonstrated that the product chain length of wild-type (WT) FPPSs and GGPPSs is in large part determined by the nature of the amino-acid residues located at the fourth and fifth positions upstream from the “first aspartate-rich motif” (FARM), a well conserved motif among scIPPSs (Ohnuma et al. 1996a; Ohnuma et al. 1996b; Tarshis et al. 1996; Narita et al. 1999).

The side-chains of these amino-acid residues can either prevent or allow additional condensations to take place, with smaller side-chains typically permitting further product elongation.

Residues other than those at positions –4 and –5 relative to the FARM may also play a role in the chain length selectivity of these prenyltransferases. For example, the Ala116Trp (–1 relative to the FARM) and Asn144Trp (+23 relative to the FARM) avian FPPS mutants displayed a smaller binding pocket for the hydrocarbon chain of the allylic substrate, with altered product selectivity favoring synthesis of shorter products such as GPP (Fernandez et al. 2000; Vandermoten et al. 2008a). In yeast GGPPS, in which a Ser residue occupies both the fourth and fifth positions upstream from the FARM, other residues located on distinct α -helices and deeper within the catalytic cavity prevent chain elongation beyond C₂₀ (Chang et al. 2006).

Here we follow up on an earlier study in which we provided the evidence for an aphid prenyltransferase displaying dual GPP/FPP synthase activity (Vandermoten et al. 2008a), a finding that was corroborated by an independent study (Lewis et al. 2008). In an effort to determine whether the aphid enzyme was unique to the aphid species from which we originally cloned it, the green peach aphid, *Myzus persicae*, we undertook the cloning of its cDNA from the pea aphid, *Acyrtosiphon pisum*. Production of the recombinant *A. pisum* ortholog (ApIPPS) and analysis of its catalytic properties revealed the same dual GPP/FPP synthase activity found for the *M. persicae* enzyme (MpIPPS). To identify the mechanism responsible for this dual activity, both *in vitro* and *in silico* studies were performed. These analyses provide new insights into the product chain-length determination mechanism of scIPPSs.

MATERIAL AND METHODS

Cloning and sequencing of ApIPPS cDNA. Cloning of an ApIPPS cDNA was carried out using the same PCR strategy employed for the cloning of MpIPPS cDNAs (for primer sequences, see Supplemental Table 1). The complete coding region of ApIPPS was cloned into the pGEM-T Easy vector (Promega) and then into the expression vector pET-30b (Novagen) after amplification by PCR, using specific primers to introduce *Eco*RI and *Xho*I restriction sites at the 5' and 3' ends, respectively.

Site-directed mutagenesis of ApIPPS. Mutagenesis was performed on pGEM-T Easy vector (Promega) containing the ApIPPS coding region with *EcoRI* and *XhoI* restriction sites at the 5' and 3' ends, respectively. Site-specific mutations were introduced using the GeneTailor site-directed mutagenesis system (Invitrogen) (for primer sequences, see Supplemental Table 1). After mutations were introduced into target sites and confirmed by sequencing, the mutated coding regions contained between the *EcoRI* and *XhoI* restriction sites were ligated into the pET30b vector. Fidelity of each insert was confirmed by sequencing.

Overproduction and purification of recombinant proteins. Each construct for expression of a protein mutant was introduced into *E. coli* strain BL21 (DE3). The procedures used for expression and purification of ApIPPS mutant and WT enzymes were as described (Vandermoten et al. 2008a). Purity of the mutant enzymes was assessed by SDS-PAGE, and protein concentration was determined by the microassay procedure of the Bio-Rad Protein Assay, using BSA as standard.

Prenyltransferase assay and product analysis. Enzyme assays were conducted in the presence of DMAPP and [^{14}C]IPP as substrates, and the nature of products formed *in vitro* by each recombinant protein was determined by HPLC/LSC analysis of hydrolyzed products as described (Vandermoten et al. 2008a). Proportions of GPP and FPP produced were calculated on the basis that they incorporate 1 and 2 units of [^{14}C]IPP, respectively; in our previous report (Vandermoten et al. 2008a), the proportions presented were simply calculated from the total radioactivity associated with each product. Given that all assays were carried out at five different DMAPP concentrations and that increasing DMAPP concentrations had the same effect on product distribution with WT and mutant proteins (Supplemental Fig. 3), these assays were used as an indirect form of replication.

Molecular dynamics simulations. Homodimeric 3D models corresponding to the ApIPPS WT sequence (residues 15-362) were built using MODELLER (Sali and Blundell 1993) as previously described (Vandermoten et al. 2008a) to serve as initial structures for the molecular dynamics simulations. Models without ligand or with DMAPP, GPP or FPP were constructed using the appropriate PDB avian structures as templates (PDB ID codes: 1UBV, 1UBX, 1UBY, and 1UBW, respectively)

(Tarshis et al. 1994; Tarshis et al. 1996). All models were successfully validated with PROCHECK (Laskowski et al. 1993), which was not surprising given the high level of sequence similarity (48% of sequence identity and only two single residue indels in coil regions that are not involved in the formation of the catalytic site). The molecular dynamics simulation protocol was similar to the one used by (Santini et al. 2009). Briefly, DMAPP, GPP and FPP coordinates were submitted to the PRODRG server (Schuettelkopf and van Aalten 2004) to generate topologies adapted by hand to the G43a2 version of GROMOS96 (Schuler et al. 2001). Each homodimer complexed with 4 Mg^{2+} ions and ligands was included in a triclinic box (9 x 11 x 9 nm) and filled with approximately 26,000 simple-point-charge water molecules. The systems were neutralized with Na^+ ions at the most favorable places, then minimized by 2000 steps of steepest descent and 1000 steps of conjugate gradient using the GROMACS software (Berendsen et al. 1995). The electrostatic interactions were treated with the PME method (Essmann et al. 1995) and the Van der Waals interactions with the shift method. 1.2 nm cut-offs were applied.

The minimized systems were submitted to 50 ps heating runs at 300 K under position restraints on protein heavy atoms, Mg^{2+} ions and ligand atoms, and then to 4 ns runs without restraints. Temperature and pressure (1 bar) were maintained using the bath method (Berendsen et al. 1984). Bonds were maintained with the LINCS algorithm (Hess et al. 1997).

Analysis of the molecular dynamics trajectories. Root Mean Square Deviations (RMSD) were computed over the 4 ns of the runs using the initial structure as reference. In the light of these results, the first 500 ps of each run were considered as the equilibration phase of the simulations and were not taken into account for the rest of the analyses (Supplemental Figure 1). The Root Mean Square Fluctuations (RMSF) were computed similarly for each Ca atom.

Interaction energies along the trajectories were estimated for both electrostatic and Van der Waals terms between the groups of atoms under investigation. For each structure, the number of contacts between the residue at position 110 and the ligand was defined as the number of pairs of atoms that have their mass center at less than 0.5 nm from each other. This number was computed using the `g_mindist`

subroutine within the GROMACS package. The same routine was used to determine the minimum distance between residue 110 and the ligand. Hydrogen bonds were determined using the `g_hbond` subroutine, with default parameters. H-bonds between the residue at position 107 and T172, T75, Q179 or K209 were classified as “crossing” H-bonds after visual inspection of the 3D models and trajectories.

Since all simulations were done with dimers, the two monomers were regarded as repetitions, making the assumption that the molecular mechanisms in each pocket are independent of one another. For each variable measured, the calculations were performed on each monomer individually. Single average and standard deviation values were then calculated from 0.5 to 4 ns, combining the two monomers.

For molecular visualization, we used both PYMOL (deLano 2002) and VMD software (Humphrey et al. 1996). All 3D images were generated with PYMOL. Residues are numbered with respect to the WT sequence of ApIPPS.

RESULTS

Cloning of *A. pisum* IPPS. A full-length cDNA encoding a scIPPS was isolated from our *A. pisum* cDNA library (Genbank accession no. [AY968585](#)); the predicted protein (ApIPPS) showed a high level of sequence identity (97%) with MpIPPS1-L (Genbank accession no. [AAV33491](#)), including the presence of a Gln residue at the fourth position upstream from the FARM (Fig. 1A and Supplemental Figure 2). ApIPPS displayed the same dual GPP/FPP synthase activity reported for MpIPPS (Vandermoten et al. 2008a), with increasing concentrations of DMAPP leading to rising proportions of the GPP product (Fig. 1B). The splice variant that we cloned contained an N-terminal mitochondrial targeting peptide almost identical to that found in MpIPPS1-L, the presence of which was shown to impair enzymatic activity (Vandermoten et al. 2008a). Therefore, only the short form of ApIPPS (*e.g.*, the mature protein without the targeting peptide) was used in the present assays and analyses.

Design of mutated ApIPPS and analysis of *in vitro* end-products. Comparison of the ApIPPS and avian FPPS active sites led to the identification of two amino acid residues that could be responsible, either alone or in combination, for the dual activity of the aphid enzyme: Gln107 and Leu110

(Vandermoten et al. 2008a). We predicted the existence of two distinct regulatory mechanisms involving the residues at each of these two positions: (i) a mechanism involving hydrogen bonds for position 107 and (ii) a mechanism involving steric hindrance for position 110. Therefore, we constructed five ApIPPS mutants by means of site-directed mutagenesis: Q107F, Q107E, L110A, L110W, and Q107F-L110A. For each position, one substitution was chosen to decrease the putative constraint of the residue on product chain length (Q107F, L110A), while the other was selected to increase it (Q107E, L110W). The substitutions in the double mutant were chosen to mimic the chicken enzyme, thereby allowing us to determine whether the residues at these two positions could entirely account for the uncommon properties of the aphid enzyme.

With the exception of L110A, all mutant enzymes presented a product profile in agreement with the initial predictions based on the homology models (Vandermoten et al. 2008a). The average IPP conversion rates for all DMAPP concentrations combined were generally high for WT (95%) and mutant proteins (82% for Q107E; 97% for Q107F; and 97% for Q107F-L110A), with the exception of the L110W mutant, which displayed a conversion rate of 52% and the L110A mutant, which failed to show any detectable activity.

Significant changes in product selectivity were observed for each mutant, with individual proteins falling into one of two categories: (i) those that formed greater amounts of GPP (Q107E and L110W) and (ii) those that formed greater amounts of FPP (Q107F and Q107F-L110A) than the wild-type ApIPPS (Fig. 1C). In addition, the mutated enzymes generally displayed a DMAPP-concentration-dependent effect similar to that observed for the WT enzyme (Supplemental Figure 3). However, DMAPP concentration had a far smaller impact on product chain-length selectivity than the mutations.

Investigation of the ApIPPS-ligand interactions by molecular dynamics simulations. To gain insight into the molecular mechanisms involved in the regulation of the IPPS end-product chain length, we performed 4 ns molecular dynamics simulations on 3D models of all mutants, without ligand or with DMAPP, GPP, or FPP. To assess the validity of the simulations, we verified the stability of the system (protein-ligand) along the trajectories. We therefore examined the total energy of the system, the RMSD

of the protein C α atoms, and their RMSF. Altogether, these preliminary analyses demonstrated the validity of the simulations, which should represent a good quality sample of the conformational space (see Supplemental Figure 1 for a full report on these analyses). Furthermore, they indicated that both WT and mutant protein structures tend to stabilize after the first 500 ps and could thus be subjected to more detailed analyses during the remaining 3.5 ns of the simulations. We then turned our attention to the impact of substitutions at positions 110 and 107 on the product chain length selectivity of the enzyme.

Interactions between the residue at position 110 and the ligand. By substituting Trp or Ala for Leu at position 110 (Leu, Ala or Trp; referred to here as X110), we aimed to increase or decrease the steric hindrance within the active site and thus favor the production of shorter or longer products, respectively. To assess this effect, we first calculated the minimum distance between the X110 residue and the ligand. In the presence of DMAPP, greater distances were observed when an Ala residue was present at position 110 (*e.g.*, L110A and Q107F-L110A mutants), although differences with other proteins were not significant except for the L110W mutant (Fig. 2A). Sharper differences were observed in the presence of GPP, with the minimum distance to the Ala residue in the L110A mutant being ~0.24 nm greater than that calculated for the Leu residue in the WT, and Q107F and Q107E mutants; not surprisingly, this difference was even greater between the L110A and L110W mutants (Fig. 2A), reflecting the fact that GPP and the Trp residue almost made steric clashes. The double mutant displayed an intermediate profile and greater variability. No significant differences were observed for this variable among the FPP-liganded structures (Fig. 2A).

The residue at position 110 could also affect ligand stabilization within the active site, which may affect the observed product profile. To assess this possibility, we calculated the number of X110-ligand atom pairs in interaction (*i.e.* separated by < 0.5 nm), and the energy associated with the X110-ligand interaction. Very few atoms of X110 interacted with DMAPP, in either WT or mutant enzymes. However, the number of atom pairs in interaction increased in the presence of GPP and FPP, with a significantly higher number of interactions observed for the L110W mutant, but a near absence of GPP-Ala interacting atom pairs for the L110A mutant (Fig. 2B). With GPP as ligand, the number of interactions calculated for

the Q107F-L110A mutant was intermediate between those of the L110A and Q107F mutants. The main contribution to the energy of interaction between X110 and the ligand came from the Van der Waals energy term (Lenard-Jones energy), a finding that is consistent with the nature of the side chain of the X110 residue and of the allylic backbone of the ligands. The energy of interaction between the X110 residue and the ligand displayed trends similar to those observed for the number of interacting atom pairs, with the highest disparities among mutants observed in the presence of GPP (Fig. 2C). The lowest degree of GPP stabilization by the X110 residue was seen with the L110A and Q107F-L110A mutants. The replacement of Leu by Ala led to an average decrease of 0.9 kCal/mol in GPP stabilization relative to the WT enzyme. Contrastingly, the L110W mutation resulted in a slight increase in GPP stabilization.

The above results are consistent with observations made by visual inspection of the structures along the trajectories. In the WT enzyme bound to GPP, L110 interacted with the ligand and allowed for ligand elongation in spite of its bulkiness (Fig. 2D). Replacement of Leu by Ala, suppressed X110-GPP interactions and the GPP ligand became wobbly within the active site cavity, presumably due to a lack of stabilization (Fig. 2D). This conclusion is further supported by the observation that deformation and movement of the GPP ligand along the trajectory were greater for the L110A and Q107F-L110A mutants than for the other proteins (Supplemental Figure 4). On the other hand, a Trp residue at position 110 increased steric hindrance within the cavity, thereby impeding additional condensations (Fig. 2E). In fact, production of GPP by the L110W mutant was found to require structural rearrangements (displacement of the helix bearing the X110 residue in monomer 1) without which the Trp residue would have made steric clashes with the ligand (Fig. 2E). The active site of the L110W mutant could even accommodate an FPP molecule, following a change in the orientation of the side chain of the Trp residue and a kink in the ligand backbone (Fig. 2F).

Hydrogen bonds involving the residue at position 107. Q107 was predicted to be the most important residue in defining the unique product selectivity of aphid IPPS (Vandermoten et al. 2008a). The glutamine amide group would form hydrogen bonds (H-bonds) with other residues of the active site cavity, and reorganization of these H-bonds would be required to accommodate longer acyl chains, thus

restraining chain elongation. To test this hypothesis, we focused on the H-bonds formed by the residue at position 107 (X107) in the WT and mutant structures, without ligand or in the presence of DMAPP, GPP or FPP. We found that fourteen distinct residues could compete to make H-bonds with the side-chain of X107, at least transiently, in one of the simulated system (Fig. 3A). We classified these bonds as either “crossing” or “non-crossing”, depending on whether they involved a partner across the active site cavity or not, respectively (Fig. 3A). Four residues out of the fourteen were involved in crossing H-bonds: T172, T175, Q179 and/or K209. H-bonds with K209 were further sorted out with respect to the portion of the Lys residue (backbone or side chain) involved in the interaction. In the presence of FPP, no crossing H-bonds were observed (Fig. 3B), which is consistent with the fact that these bonds are not compatible with ligand elongation. Since the side-chain of a Phe has neither H-bond acceptor nor donor, no H-bonds were detected for the Q107F and Q107F-L110A mutants.

In the WT enzyme, Q107 made H-bonds approximately half of the time in the absence of ligand or in the presence of DMAPP (Fig. 3B). In the structure without ligand, the majority of these H-bonds crossed the active site cavity, and K209 was by far the most frequent partner, with interactions involving primarily its backbone (Fig. 3C). Presence of a GPP molecule in the active site virtually eliminated H-bond formation while the FPP ligand allowed these interactions to occur in 15% of the frames, although none crossed the cavity (Fig. 3B).

The replacement of Q107 by a Glu residue led to a ≥ 3 -fold increase in the total number of H-bonds formed (Fig. 3B). Crossing H-bonds were also more frequent than in the WT enzyme, but only in the absence of ligand or in the presence of DMAPP. Again, K209 was the preferred partner, but these interactions involved primarily the side chain of K209 (Fig. 3C).

Differences in the total number of H-bonds were observed between simulations involving the L110A mutant and the WT structures. Notably, crossing H-bonds, in the L110A structure, were formed in the presence of either DMAPP or GPP (Fig. 3B). In the absence of ligand, Q107 interacted with K209, as in the WT enzyme, but made additional H-bonds with Q179; these interactions were maintained

following introduction of a DMAPP molecule in the active site (Fig. 3C). In the presence of GPP, most of the crossing H-bonds were made with the backbone of K209 (Fig. 3C).

In the L110W structures with ligand, Q107 made more H-bonds than in the WT enzyme (Fig. 3B). This mutation further enhanced the formation of crossing H-bonds in the presence of DMAPP. With a GPP molecule in the active site, the frequency of crossing H-bonds rose to an average of 0.95 (Fig. 3B). This pronounced increase was mostly due to interactions with T175 and, to a lesser extent, T172 (Fig. 3C). In the absence of ligand, crossing H-bonds involved primarily the side chains of Q179 and K209. The same residues interacted with Q107 in the presence of DMAPP, although the interaction with K209 involved its backbone instead of its side-chain (Fig. 3C).

Amongst the H-bonds reported above, the high proportion of interactions between X107 and K209 (Fig. 3C) suggested that these bonds could play an important role in defining the product selectivity of the WT and mutant enzymes, particularly in the case of the Q107E mutant. To further assess this hypothesis, we computed the Coulomb energy between these residues along the trajectories. The X107-K209 electrostatic interactions were consistent with the H-bond profiles of the mutants. For example, the energy calculated for the Q107F and Q107F-L110A mutants was 0 kCal/mol, irrespective of the liganded state of the enzyme (Fig. 4A,B); this finding concurs with the fact that the Phe residue at position 107 is not polar. In the absence of ligand, the Coulomb energy varied between -2.5 and -5 kCal/mol for the WT, L110A and L110W structures, whereas it was much lower for the Q107E mutant (-17.8 kCal/mol, on average; Fig. 4A,B). This sharp decrease resulted primarily from the formation of salt bridges between the side chains of the Glu and Lys residues. In the presence of DMAPP, the Coulomb energy rose to > -2 kCal/mol for all structures examined except for the Q107E mutant, in which it remained relatively low (-8.4 kCal/mol; Fig. 4B). In the presence of GPP or FPP, the interactions between Q107 and K209 were negligible, as expected from the X107-associated H-bond profiles in these structures.

Again, visual inspection of the structures along the trajectories corroborated the above observations. In the WT enzyme, the H-bonds formed between Q107 and K209 (backbone and side chain) needed to be broken to accommodate GPP within the active site, as illustrated by superimposition of unliganded and

GPP-liganded structures (Fig. 4C). Substituting Glu for Gln at position 107 led to the formation of a salt bridge between E107 and the side chain of K209, the presence of which was not compatible with the formation of GPP (Fig. 4D). Interestingly, because the X107 residue is surrounded by two Tyr residues (Y106, Y213), the Phe residue in the Q107F mutant could be stabilized through stacking interactions between the three aromatic rings. In this position, F107 did not hinder product formation within the active site (Fig. 4E). In the L110W mutant, the steric constraint imposed by the Trp residue induced displacement of Q107, allowing the formation of specific H-bonds with T172 and T175 in this mutant enzyme (Fig. 4F).

DISCUSSION

Most organisms produce distinct scIPPSs to generate each of the universal C₁₀, C₁₅, and C₂₀ isoprenoid diphosphate precursors. These prenyltransferases tend to display characteristic residues within and near their allylic substrate binding site (FARM), including some that have been shown to play a role in product chain-length selectivity. For this reason, the FARM and the residues in its vicinity have been designated the “chain-length determination” (CLD) region of scIPPSs (Tarshis et al. 1996). The identification of an aphid prenyltransferase that displays dual GPP/FPP synthase activity prompted us to take a closer look at its CLD region and to explore the molecular adaptations responsible for its uncommon product chain-length selectivity. Here, on the basis of *in vitro* and *in silico* characterization of WT and site-directed ApIPPS mutants, we identified two residues, located at position -4 (Q107) and -1 (L110) relative to the FARM, that appear to play a key role in conferring dual activity to the aphid enzyme.

Mechanism conferring dual activity. Product chain-length selectivity of scIPPSs has been proposed to be essentially under the control of steric constraints within the active site cavity, with residues located upstream from the FARM (Ohnuma et al. 1996a; Ohnuma et al. 1996b; Tarshis et al. 1996; Narita et al. 1999) and on other α -helices (Chang et al. 2006) being responsible for this effect. As supported by our results on the product chain-length selectivity of the designed scIPPS mutants, which are

in agreement with our initial predictions and molecular dynamics analyses, the present work shows that additional factors can affect product chain-length selectivity. These factors include stabilization of the ligand and reorganization of the interactions among residues of the active site cavity upon ligand entry and elongation. The latter factor, however, appears to be the most critical one in the aphid enzyme, given the atypical nature of the Gln residue at position -4 relative to the FARM, a site that is generally occupied by an aromatic residue in FPPSs (Vandermoten et al. 2008a). The polar side-chain of Q107 can indeed make H-bonds with residues across the active site cavity, disfavoring the production of longer products. Thus, it is not surprising that the Q107E mutation, which converted these H-bonds to more stable salt bridges, altered the IPP conversion rate slightly (82%) and the product selectivity in favor of GPP production. At position 110, in addition to reducing the IPP conversion rate (52%), substituting Trp for Leu increased steric hindrance within the active site, converting the enzyme into a strict GPPS, an outcome similar to that reported for avian FPPS following replacement of Ala with Trp at the equivalent position (Fernandez et al. 2000). In the aphid enzyme, however, the impact of this Trp residue on steric hindrance was enhanced by its indirect effect on the frequency of crossing H-bonds formed by Q107.

Surprisingly, the L110A mutation, designed to favour FPP formation through a reduction in steric hindrance, suppressed enzymatic activity. Although a protein miss-folding problem could be suspected as the cause for this lack of activity, this scenario appears very unlikely given that (i) the replacement of Leu by Ala constitutes a minor modification of the protein, (ii) an Ala residue is observed at the equivalent position in some FPPSs (e.g., avian FPPS) and (iii) the double mutant bearing the exact same substitution in addition to a Q107F mutation presented an activity similar to the WT enzyme (97% and 95% IPP conversion rate, respectively). In addition, we did not observe any problem with the expression or purification of the mutant enzyme (see Supplemental Figure 5 for SDS-PAGE gel). Rather, the absence of activity appears to result from a lack of ligand stabilization within the active site cavity. Ala110 would indeed not stabilize the ligand enough in the active site to compensate for the need to disrupt Q107-associated crossing H-bonds for ligand elongation. The latter hypothesis is supported by the fact that enzymatic activity was restored in the Q107F-L110A double-mutant, in which Phe107 cannot make H-

bonds across the active site cavity. Moreover, the product profile of the double-mutant, which generated a lower proportion of FPP than the Q107F single mutant, further demonstrates that steric hindrance within the active site pocket does not, alone, explain prenyltransferase product chain-length selectivity, and that ligand stabilization is also a crucial parameter.

Functional significance. We do not know whether the product selectivity of the aphid prenyltransferase is modulated *in vivo* to favor the production of either GPP or FPP as a function of the desired end product (*e.g.*, mono- or sesquiterpene) in a given tissue. It has been shown, however, that under *in vitro* conditions where the GPP:FPP product ratio was ~1.6:1, rMpIPPS could generate significant quantities of both pinene (monoterpene) and germacrene A (sesquiterpene) in a “linked” assay where the appropriate terpene synthase was added to the assay buffer (Lewis et al. 2008). As suggested by our data, some degree of *in vivo* modulation of the GPP:FPP product ratio could be achieved by altering the local DMAPP concentration, either through regulation of the enzyme responsible of the conversion of IPP into DMAPP, IPP isomerase, or enzyme/substrate compartmentalization. The IPP isomerase of the aphid *A. pisum* has been cloned by our group (ApIPPI; Genbank accession no. [FJ824667](#)), and, like the GPP/FPP synthase studied here, likely contains a N-terminal mitochondrial targeting signal (unpublished data). Since both long and short transcripts (*e.g.*, with and without the sequence encoding the mitochondrial targeting signal) of the GPP/FPP synthase of *M. persicae* have been shown to be produced (Vandermoten et al. 2008a), the potential exists for the regulation of subcellular localization and concentration of each isomerase and synthase and their products.

Evolutionary origins. It is unclear why evolution would have favored the development of a dual GPP/FPP synthase in aphids, as opposed to distinct GPP and FPP synthases in other organisms. Although we cannot, at this stage, rule out the existence of such distinct enzymes in the pea aphid, the only clearly different prenyltransferase that we were able to identify in this species’ genome through Blast analysis (Genbank accession no. [XP_001947197](#)) has a sequence that strongly suggests it is a polyprenyl diphosphate synthase (*e.g.*, whose product is longer than C20), as opposed to an scIPPS (data not shown). These enzymes are typically involved in generating the isoprenyl side chains of coenzyme Q family

members, and some use FPP as their allylic diphosphate co-substrate (Tonhosolo et al. 2005), which could be provided by the ApIPPS characterized here.

Although all *cis*-prenyltransferases, the class of enzymes to which the aphid scIPPS belongs, are believed to be derived from a common ancestor (Chen and Poulter 1993), GPPSs appear to be rare in the animal kingdom, as plants have a near monopoly on monoterpenes. However, some insects (namely aphids and beetles) use monoterpenes as pheromones. Aphids thus require both GPP and FPP as precursors for sex pheromone, alarm pheromone and juvenile hormone biosynthesis (Kislow and Edwards 1972; Schooley and Baker 1985; Dawson et al. 1987), and ApIPPS could well be involved in all of these pathways. The first (and only to this day) animal GPPS to have been characterized was cloned from the bark beetle *Ips pini* (IpGPPS), in which distinct GPP and FPP synthases were found (Gilg et al. 2005). However, plant GPPSs, which can be either hetero- or homodimeric (Burke et al. 1999; Burke and Croteau 2002) and IpGPPS seem to have evolved independently from their respective FPP or GGPP synthase precursors (Vandermoten et al. 2009). The product chain-length determination mechanism of IpGPPS has not been identified and hypotheses about it cannot be readily proposed given that its CLD region is very similar to that of FPPSs (Gilg et al. 2005). Although type-1 lepidopteran FPPSs also contain a Gln residue at position –4 relative to the FARM, a phylogenetic analysis that included these and several other insect FPPSs and MpIPPS (Cusson et al. 2006) strongly suggests that the two enzymes have evolved independently (*e.g.*, within their respective taxon) from a conventional FPPS. In the case of the aphid enzyme, this would have occurred through mutations altering the nature of the residue at the fourth and, possibly, the first positions upstream from the FARM (some FPPSs already have a Leu residue at position –1). Given that these putative mutations would not have suppressed the formation of FPP, it is conceivable that they became fixed, without further divergence towards distinct GPP and FPP synthases.

As indicated above, lepidopteran type-1 FPPS also displays a Gln residue at the fourth position upstream from the FARM, but the recombinant enzyme showed neither GPP nor FPP synthase activity in the presence of DMAPP and IPP, unless assayed in the presence of the more conventional type-2 protein, in which case it enhanced the FPPS activity of the latter enzyme (Sen et al. 2007). In addition,

lepidopteran type-1 FPPS contains several other uncommon active-site substitutions relative to other IPPSs, some of which are expected to interact with the Gln residue in ways that are different from those identified here for the aphid enzyme. Finally, the production of monoterpenes by lepidopteran insects seems rare as it has been reported only once (Vandermoten et al. 2008b). However, many coleopteran insects share with aphids an ability to produce both mono- and sesquiterpenoid compounds, and could therefore have independently acquired an scIPPS with features similar to those described here for the aphid enzyme. Whether there exist IPPSs displaying similar features and product chain-length determination mechanisms in other insect taxa has yet to be established.

ACKNOWLEDGEMENT

We thank C. Béliveau and M. Vandenbol for helpful advice on site-directed mutagenesis. This work was supported by a PhD scholarship to S.V., awarded by the “Fonds pour la formation à la Recherche dans l’Industrie et l’Agriculture” (Belgium) and an NSERC Discovery Grant to MC. S.S. was supported by a grant from the FUSAGx (Gembloux - Belgium), and B.C. by the Belgian Program on Interuniversity Attraction Poles initiated by the Federal Office for Scientific, Technical and Cultural Affairs (IAP. P6/19 PROFUSA). R.B. thanks the “Fonds National pour la Recherche Scientifique” (Belgium).

ABBREVIATIONS USED

FARM, first aspartate-rich motif; scIPPS, short-chain isoprenyl diphosphate synthase; GPPS, geranyl diphosphate synthase; FPPS, farnesyl diphosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; MpIPPS, *Myzus persicae* isoprenyl diphosphate synthase; ApIPPS, *Acyrtosiphon pisum* isoprenyl diphosphate synthase; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl diphosphate; FPP farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; RMSD, root mean square deviation; RMSF, root mean square fluctuation; CLD, chain-length determination.

LIST OF SUPPLEMENTAL MATERIAL

Supplemental Table 1: Sequences of primers designed.

Supplemental Figure 1: Molecular dynamics global parameters.

Supplemental Figure 2: Alignment of MpIPPS (AAY33491) and ApIPPS (AY968585) amino acid sequences.

Supplemental Figure 3: Radio-HPLC analysis of the in vitro products

Supplemental Figure 4: Ligand motion within the active site cavity.

Supplemental Figure 5: SDS-PAGE of L110A mutant obtained during each step of the purification.

REFERENCES

- Bellés, X., Martín, D. and Piulachs, M.D., 2005. The mevalonate pathway and the synthesis of juvenile hormone in insects. *Annu. Rev. Entomol.* 50: 181-199.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R., 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81: 3684-3690.
- Berendsen, H.J.C., van der Spoel, D. and van Drunen, R., 1995. GROMACS: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* 91: 43-56.
- Burke, C.C. and Croteau, R., 2002. Geranyl diphosphate synthase from *Abies grandis*: cDNA isolation, functional expression, and characterization. *Arch. Biochem. Biophys.* 405: 130-136.
- Burke, C.C., Wildung, M.R. and Croteau, R., 1999. Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proc. Natl. Acad. Sci., USA* 96: 13062-13067.
- Chang, T.H., Guo, R.T., Ko, T.P., Wang, A.H. and Liang, P.H., 2006. Crystal structure of type-III geranylgeranyl pyrophosphate synthase from *Saccharomyces cerevisiae* and the mechanism of product chain length determination. *J. Biol. Chem.* 281: 14991-15000.
- Chen, A. and Poulter, C.D., 1993. Purification and characterization of farnesyl diphosphate/geranylgeranyl diphosphate synthase. A thermostable bifunctional enzyme from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* 268: 11002-11007.
- Cromartie, T.H., Fisher, K.J. and Grossman, J.N., 1999. The discovery of a novel site of action for herbicidal bisphosphonates. *Pest. Biochem. Physiol.* 63: 114-126.
- 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. and 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.

- Dawson, G.W., Griffiths, D.C., Janes, N.F., Mudd, A., Pickett, J.A., Wadhams, L.J. and Woodcock, C.M., 1987. Identification of an aphid sex pheromone. *Nature* 325: 614-616.
- deLano, W.L., 2002. The PyMOL Molecular Graphics System. on world wide web.
- Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G., 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103: 8577-8593.
- Fernandez, S.M.S., Kellogg, B.A. and Poulter, C.D., 2000. Farnesyl diphosphate synthase. Altering the catalytic site to select for geranyl diphosphate activity. *Biochemistry* 39: 15316-15321.
- Gilg, A.B., Bearfield, J.C., Titigger, C., Welch, W.H. and 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.
- Hess, B., Bekker, H., Berendsen, H. and Fraaije, J., 1997. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18: 1463-1472.
- Humphrey, W., Dalke, A. and Schulten, K., 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14: 33-8, 27-8.
- Kavanagh, K.L., Dunford, J.E., Bunkoczi, G., Russell, R.G. and Oppermann, U., 2006a. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J. Biol. Chem.* 281: 22004-22012.
- Kavanagh, K.L., Guo, K., Dunford, J.E., Wu, X., Knapp, S., Ebetino, F.H., Rogers, M.J., Russe, R.G.G. and Oppermann, U., 2006b. The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. *Proc. Natl. Acad. Sci., USA* 103: 7829-7834.
- Kawasaki, T., Hamano, Y., Kuzuyama, T., Itoh, N., Seto, H. and Dairi, T., 2003. Interconversion of the product specificity of type I eubacterial farnesyl diphosphate synthase and geranylgeranyl diphosphate synthase through one amino acid substitution. *J. Biochem. (Tokyo)* 133: 83-91.
- Kislow, C. and Edwards, L.J., 1972. Repellent odour in aphids. *Nature* 235: 108-109.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26: 283-291.
- Lewis, M.J., Prosser, I.M., Mohib, A. and Field, L.M., 2008. Cloning and characterisation of a prenyltransferase from the aphid *Myzus persicae* with potential involvement in alarm pheromone biosynthesis. *Ins. Mol. Bio.* 17: 437-443.
- Liang, P.H., Ko, T.Z. and Wang, A.H.J., 2002. Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* 269: 3339-3354.
- Narita, K., Ohnuma, S. and Nishino, T., 1999. Protein design of geranyl diphosphate synthase. Structural features that define the product specificities of prenyltransferases. *J. Biochem. (Tokyo)* 126: 566-571.
- Ohnuma, S., Hirooka, K., Ohto, C. and 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.
- Ohnuma, S., Nakazawa, T., Hemmi, H., Hallberg, A.M., Koyama, T., Ogura, K. and Nishino, T., 1996a. Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. *J. Biol. Chem.* 271: 10087-10095.
- Ohnuma, S., Narita, K., Nakazawa, T., Ishida, C., Takeuchi, Y., Ohto, C. and Nishino, T., 1996b. 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.
- Palli, S.R. and Cusson, M., 2007. Future insecticides targeting genes involved in the regulation of molting and metamorphosis. In: Ishaaya, I., Nauen, R. and Horowitz, A.R. (Eds), *Insecticide Design Using Advanced Technologies*. Springer-Verlag, Berlin, pp. 105-134.
- Sali, A. and Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234: 779-815.
- Santini, S., Crowet, J.M., Thomas, A., Paquot, M., Vandenbol, M., Thonart, P., Wathélet, J.P., Blecker, C., Lognay, G., Brasseur, R., Lins, L. and Charlotiaux, B., 2009. Study of *Thermomyces lanuginosa* lipase in the presence of tributylglycerol and water. *Biophys. J.* 96: 4814-4825.
- Schooley, D.A. and Baker, F.C., 1985. Juvenile hormone biosynthesis. In: Kerkut, G.A. and Gilbert, L.I. (Eds), *Comprehensive insect physiology, biochemistry and pharmacology*. Pergamon Press, Oxford, pp. 363-389.
- Schuettelkopf, A.W. and van Aalten, D.M.F., 2004. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica D60*: 1355-1363.
- Schuler, L.D., Daura, X. and Van Gunsteren, W.F., 2001. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* 22.
- Sen, S.E., Cusson, M., Trobaugh, C., Béliveau, C., Richard, T., Graham, W., Mimms, A. and 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.
- Tarshis, L.C., Proteau, P.J., Kellogg, B.A., Sacchettini, J.C. and Poulter, C.D., 1996. Regulation of product chain length by isoprenyl diphosphate synthases. *Proc. Natl. Acad. Sci. USA* 93: 15018-15023.
- Tarshis, L.C., Yan, M., Poulter, C.D. and Sacchettini, J.C., 1994. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. *Biochemistry* 33: 10871-10877.
- Tonhosolo, R., D'Alexandri, F.L., Genta, F.A., Wunderlich, G., Gozzo, F.C., Eberlin, M.N., Peres, V.J., Kimura, E.A. and Katzin, A.M., 2005. Identification, molecular cloning and functional characterization of an octaprenyl pyrophosphate synthase in intra-erythrocytic stages of *Plasmodium falciparum*. *Biochem. J.* 392: 117-126.
- Vandermoten, S., Charlotiaux, B., Santini, S., Sen, S.E., Béliveau, C., Vandenbol, M., Francis, F., Brasseur, R., Cusson, M. and Haubruge, E., 2008a. Characterization of a novel aphid

prenyltransferase displaying dual geranyl/farnesyl diphosphate synthase activity. FEBS Lett. 582: 1928-1934.

Vandermoten, S., Francis, F., Cusson, M. and Haubruge, E., 2008b. La biosynthèse des isoprénoïdes chez les pucerons : Une cible potentielle de nouveaux bio-insecticides? Biotechnologie, Agronomie, Société et Environnement 12: 451-460.

Vandermoten, S., Haubruge, E. and Cusson, M., 2009. New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. Cell. Mol. Life Sci. in press DOI: 10.1007/s00018-009-0100-9.

FIGURE LEGENDS

Fig. 1. A) Alignment of ApIPPS and MpIPPS amino acid residues 106 to 115, along with the corresponding region of chicken FPPS (GgFPPS), for comparison; arrows point to the two residues chosen for mutagenesis. B) Radio-HPLC analysis of *in vitro* products (GPP and FPP) of WT rApIPPS, following 40 min incubation at 37°C, in the presence of 4 μ M [14 C]IPP and increasing concentrations of DMAPP, and conversion of the products to their corresponding alcohols. C) Same as in B, but for WT and all mutant proteins assayed at 10 μ M DMAPP. The L110A mutant displayed no detectable activity.

Fig. 2. Interactions between residues at position 110 (X110) and ligands, as determined during 3.5 ns molecular dynamics simulations (from 500 to 4000 ps). A) Minimum distance (mean \pm SD; nm) between X110 and the ligand, computed for both subunits. B) Number of X110-ligand atom pairs (mean \pm SD) less than 5 nm apart, computed for both subunits. C) Evolution of the Lennard-Jones energy (kCal/mol) between X110 and GPP along the trajectories in one subunit. D-F) Representative snapshots of the active site cavity. The X110 residue and the ligand are shown as sticks, while the α -helix bearing the X110 residue is displayed as a ribbon. Hydrogen atoms are omitted. In all cases, dimeric structures were superimposed using the WT structure as reference.

Fig. 3. H-bonds involving the residue at position 107 (X107) and other residues within the active site, as determined during 3.5 ns molecular dynamics simulations (from 500 to 4000 ps). A) Stereo view of a snapshot of the WT enzyme showing the fourteen residues that could make crossing (colored) and non-

crossing (grey; light-grey in the case of the tyrosine residue shown behind the space-fill FPP molecule) H-bonds with X107 along the trajectories. Dashed lines denote pairs of residues that can make crossing H-bonds. L110 is also displayed (black). Hydrogen atoms are omitted. B) Average frequency of occurrence of crossing, non-crossing and total H-bonds between the side chain of X107 and any other residue, computed for both subunits. C) Frequency of occurrence of H-bonds formed between X107 and K209/side-chain, K209/backbone, Q179, T175 and T172 (average for two monomers).

Fig. 4. Specific interactions between the residue at position 107 (X107) and K209, as determined during 3.5 ns molecular dynamics simulations (from 500 to 4000 ps). A) Evolution of the Coulomb energy between X107 and K209 in one subunit during the simulation, without ligand. B) Mean coulomb energy (\pm SD) between X107 and K209 for both subunits. C-F) Representative snapshots of the active site cavity. For reasons of clarity, only selected residues from one subunit are shown. Hydrogen atoms are omitted and H-bonds are represented by dashed lines. In all cases, dimeric structures were superimposed using the WT structure with GPP as reference.