Insect Biochemistry and Molecular Biology 161 (2023) 104001

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Characterisation of geranylgeranyl diphosphate synthase from the sandfly *Lutzomyia longipalpis*

Charles Ducker [a], Stanley French [a], Monika Pathak [b], Harry Taylor [a], Adam Sainter [a], William Askem [a], Ingrid Dreveny [b], Antônio Euzébio Goulart Santana [c], John A. Pickett [d], Neil J. Oldham [a,*]

[a] School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, UK
[b] School of Pharmacy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK
[c] Center of Engineering and Agrarian Science, Federal University of Alagoas, Maceio, Brazil
[d] School of Chemistry, Cardiff University, Main Building, Park Pl, Cardiff, CF10 3AT, UK

A B S T R A C T

Leishmaniasis is a debilitating and often fatal neglected tropical disease. Males from sub-populations of the *Leishmania*-harbouring sandfly, *Lutzomyia longipalpis*, produce the diterpene sex and aggregation pheromone, sobralene, for which geranylgeranyl diphosphate (GGPP) is the likely isoprenoid precursor. We have identified a GGPP synthase (*lz*GGPPS) from *L. longipalpis*, which was recombinantly expressed in bacteria and purified for functional and kinetic analysis. *In vitro* enzymatic assays using LC-MS showed that *lz*GGPPS is an active enzyme, capable of converting substrates dimethylallyl diphosphate (DMAPP), (*E*)-geranyl diphosphate (GPP), (*E,E*)-farnesyl diphosphate (FPP) with co-substrate isopentenyl diphosphate (IPP) into (*E,E,E*)-GGPP, while (*Z,E*)-FPP was also accepted with low efficacy. Comparison of metal cofactors for *lz*GGPPS highlighted Mg^{2+} as most efficient, giving increased GGPP output when compared against other divalent metal ions tested. In line with previously characterised GGPPS enzymes, GGPP acted as an inhibitor of *lz*GGPPS activity. The molecular weight in solution of *lz*GGPPS was determined to be ~221 kDa by analytical SEC, suggesting a hexameric assembly, as seen in the human enzyme, and representing the first assessment of GGPPS quaternary structure in insects.

1. Introduction

The mevalonate (MEV) pathway is an essential metabolic pathway utilised by eukaryotes, archaea and certain bacteria, through which the 5-carbon compound isopentenyl diphosphate, also known as isopentenyl pyrophosphate, (IPP) is synthesized. IPP can then be reversibly converted to the allylic dimethylallyl diphosphate (DMAPP) through action of IPP isomerases. Sequential head to tail condensation reactions of DMAPP and IPP, mediated by isoprenyl diphosphate synthase (IDS) enzymes, yields longer isoprenoid chains such as geranyl diphosphate (GPP, C_{10}), farnesyl diphosphate (FPP, C_{15}) and geranylgeranyl diphosphate (GGPP, C_{20}, Fig. 1A). These are essential for diverse biochemical processes including protein prenylation and terpene/terpenoid biosynthesis (Buhaescu and Izzedine, 2007). IDS enzymes are divided into *cis* and *trans* superfamilies, defined by the stereochemistry of the allylic double bonds formed during catalysis (*Z* or *E* respectively) (Chen et al., 2020a). For completeness, it is noted that an alternative pathway to IPP and DMAPP, known as the methyl-D-erythritol phosphate (MEP) pathway, exists in most bacteria, the apicoplasts of some parasites and in plant plastids, but it is believed that insects, and animals in general, are incapable of performing these metabolic transformations (Banerjee and Sharkey, 2014).

The New World phlebotomine sandfly *Lutzomyia longipalpis* is a carrier of the *Leishmania* parasite, the causative pathogen of the neglected tropical disease leishmaniasis (Lainson et al., 1977). Male sandflies produce sex/aggregation pheromones, of which there are several defined geographical chemotypes in Brazil (Spiegel et al., 2016). These are variably comprised of three well-characterised pheromone components - the homosesquiterpenes 3-methyl-α-himachalene and 9-methylgermacrene-B, and the diterpene sobralene (Fig. 1A and B) (Hamilton et al., 1999a, 1999b; Palframan et al., 2018). Terpenes are produced by terpene synthases (TPS), which operate downstream of IDS activity. Studying the regulation of isoprenoid diphosphate production is a prerequisite to understanding subsequent secondary metabolism and terpene synthesis. The isoprenoid precursor for diterpenes is GGPP, which is produced by IDS enzymes known as geranylgeranyl diphosphate synthases (GGPPS). IDS enzymes are characterised by dual aspartate-rich motifs (generally DDxxD) which are important for metal

* Corresponding author.
E-mail address: neil.oldham@nottingham.ac.uk (N.J. Oldham).

https://doi.org/10.1016/j.ibmb.2023.104001
Received 23 May 2023; Received in revised form 2 August 2023; Accepted 20 August 2023
Available online 22 August 2023

cofactor coordination and diphosphate binding in the active site, and referred to as the first aspartate-rich motif (FARM) and second aspartate-rich motif (SARM). The amino acids at the −4 and −5 position from the FARM are generally considered as contributory factors in determining product chain length, and variations in this region categorise a GGPPS as type I (large residue at −4/−5; archaea), type II (two residue insertion in FARM; bacteria and plant) or type III (small residues at −4 and −5; fungi and animals)(Wang and Ohnuma, 1999).

GGPPS enzymes generally can accept DMAPP, GPP or FPP as allylic substrates (requiring three, two, and one equivalents of IPP to produce the final product respectively, see Fig. 1) but kinetic studies suggest that in the case of type III GGPPSs, FPP is the preferred substrate (Sagami et al., 1994). In plants, several gene paralogs are generally present, which can be ubiquitously expressed or restricted to various developmental stages and in specific tissues (Beck et al., 2013; Coman et al., 2014). *Arabidopsis thaliana* encodes 10 functional GGPPS proteins, with variable subcellular localisation ranging from cytosolic, mitochondrial and plastidial (Beck et al., 2013). However, essential plastidial GGPP production in *Arabidopsis* (preceding carotenoid biosynthesis) is derived from a single GGPPS, while two co-operative plastid-targeted enzymes carry out this function in tomato (*Solanum lycopersicum*), illustrating diversity in the regulation of isoprenoid chain synthesis in plants (Barja et al., 2021).

A relative paucity of information exists on the molecular characteristics of insect GGPPSs and their regulation. In contrast to plants, a single gene is usually present, indicated by the lethality imparted on *Drosophila melanogaster* following nonsense mutation, which highlights a lack of redundancy and the essential role of the enzyme (Lai et al., 1998). *GGPPS* expression in the spruce budworm moth *Choristoneura fumiferana* varied throughout development alongside other prenylation-linked genes, with increased expression directly preceding the pupal stage, followed by a reduction until adult emergence (Barbar et al., 2013). The bumblebee *Bombus terrestris* also showed age-dependent variation of *GGPPS* expression in the labial gland, where GGPP is converted to the diterpene pheromone component geranylcitronellol (Prchalová et al., 2016). *GGPPS* depletion indirectly influenced mono/sesquiterpene production in the aphids *Aphis gossypii* and *Megoura viciae* (Sun and Li, 2018; Song et al., 2022). Increased *GGPPS* expression in the green versus red morph pea aphid (*Acrythosiphon pisum*) correlated with higher carotenoid pigment levels in green morphs, while gene knockdown significantly reduced carotenoid levels in both without altering their chemical profile (Ding et al., 2019).

Structural analysis of human GGPPS found it to be hexameric (in the form of a trimer of dimers) with sequence alignments indicating this is also likely to be the case for insect members (Kavanagh et al., 2006; Barbar et al., 2013). This is in contrast to the dimeric form generally seen in farnesyl diphosphate synthases (FPPS) and fungal, bacterial, archaeal, and plant GGPPSs (Reed and Rilling, 1975; Chang et al., 2006; Kloer et al., 2006; Feng et al., 2020). One study also suggested that human GGPPS existed as an octamer in solution (Kuzuguchi et al., 1999). The human enzyme was found to co-crystallise with GGPP product bound in an inhibitory conformation in the active site, and evidence from

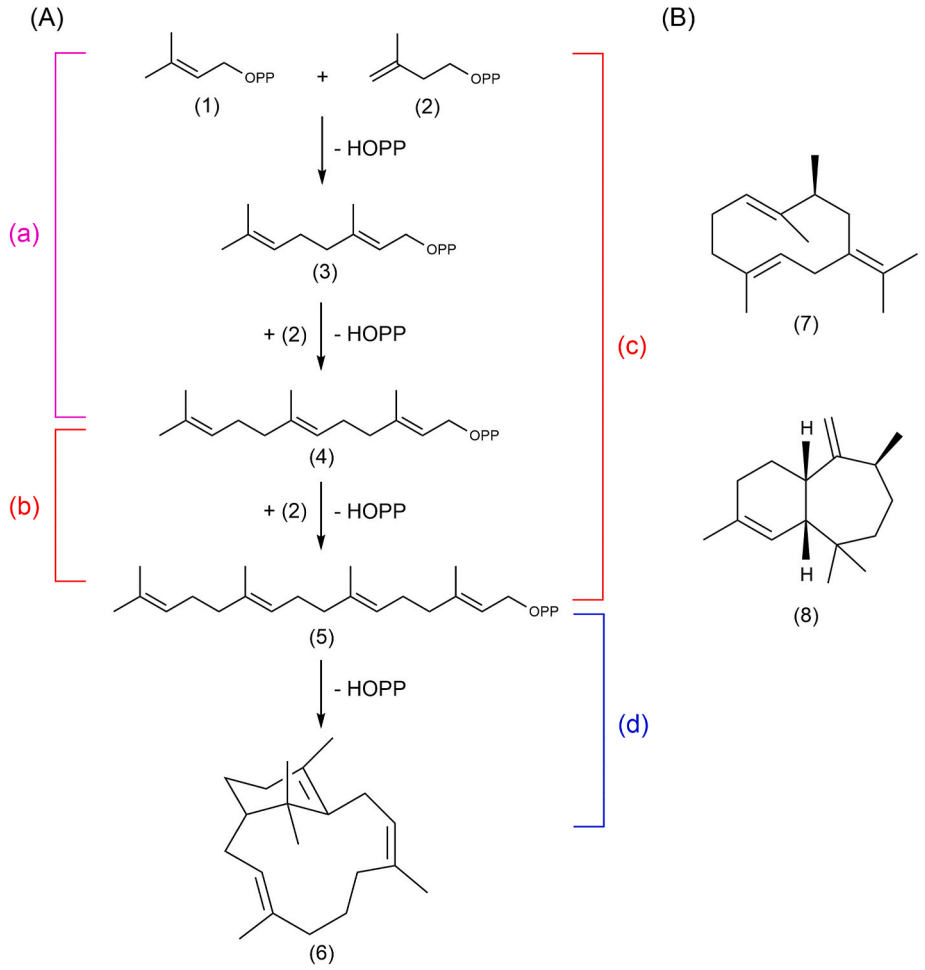

Fig. 1. Sex and aggregation pheromones in *L. longipalpis*. A. Schematic of the proposed IDS and TPS enzyme activity in *L. longipalpis* leading to diterpene pheromone production. DMAPP (1); IPP (2); (*E*)-GPP (3); (*E,E*)-FPP (4); (*E,E,E*)-GGPP (5); Sobralene (6), (a) FPPS activity; (b) GGPPS activity; (c) *De novo* GGPPS activity; (d) TPS activity. **B**. Structures of the homosesquiterpene pheromones. (*S*)-9-Methylgermacrene-B (7); (1*S*,3*S*,7*R*)-3-Methyl-α-himachalene (8).

C. fumiferana suggests this is also likely to extend to insects (Kavanagh et al., 2006; Barbar et al., 2013).

The enzymes responsible for isoprenoid production and downstream terpene synthesis in *L. longipalpis* have not been characterised to date. In this study, we have identified a *trans*-GGPPS enzyme from the sandfly genome, and following heterologous expression and purification of recombinant protein, have determined its functional activity and quaternary structure. Unravelling the complex pathways governing aggregation pheromone biosynthesis in *L. longipalpis* could allow for targeted monitoring of sandfly populations to control the spread of leishmaniasis.

2. Materials and methods

2.1. Phylogenetic analyses and protein structural modelling

Protein sequences corresponding to reported insect GGPPSs from *B. terrestris* (I1VX07), *D. melanogaster* (O61539), *A. pisum* (A0A8R2B6W9), *A. gossypii* (S5MIM6) and *C. fumiferana* (U3RD44) were retrieved from the UniProt database along with various other mammalian, fungal, bacterial and plant sequences and aligned with *L. longipalpis* (*lz*GGPPS, A0A7G3AJG6) using MUSCLE. The maximum likelihood tree was constructed in MEGA11 using LG + G + I method with partial deletion of gaps, and phylogeny was tested by bootstrap method (1000 replicates). Structural model prediction of monomeric *lz*GGPPS was carried out using I-TASSER, which was arranged into a hexamer in ChimeraX 1.4 using the structure of human GGPPS (PDB 2Q80) as a guide (Kavanagh et al., 2006; Yang and Zhang, 2015).

2.2. Protein expression and purification

The presence of protein localisation signals in *lz*GGPPS primary structure were predicted using TargetP-2.0 and iPSORT (Armenteros et al., 2019; Bannai et al., 2002). DNA encoding full-length *lz*GGPPS in expression vector pET100/D-TOPO (N-terminally His-tagged) was synthesized by GeneArt (Thermo Fisher Scientific, Loughborough, UK). The plasmid was transformed into BL21(DE3)pLysS *E. coli* competent cells for heterologous protein expression. Starter cultures of 10 mL LB (with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol) were incubated with shaking from a single colony overnight at 37 °C, and used to inoculate a 500 mL culture the following day. Upon culture turbidity reaching OD_{600} ~0.5, cells were induced with 0.75 mM IPTG and harvested after a further 3 h at 37 °C.

Cells were lysed in 25 mL wash buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 10% v/v glycerol, 20 mM imidazole, 1 mM DTT) supplemented with 1% Igepal CA-630 and cOmplete protease inhibitor EDTA-free (Roche, Welwyn Garden City, UK), rotated at 4 °C for 15 min, sonicated 10 × 10 s on ice (25 Hz, 60% amplitude) and centrifuged at 4 °C for 15 min at 14000 rpm. The soluble fraction was added to pre-washed 1 mL Nickel-NTA agarose (in wash buffer – 500 μL bed volume) and rotated at 4 °C for 2 h. Beads were washed with 3 × 10 mL wash buffer and protein eluted with 3 mL elution buffer (wash buffer containing 400 mM imidazole), before buffer exchange using PD-10 columns (Cytiva, Little Chalfont, UK) to protein storage buffer (for enzyme activity assays – 25 mM MOPSO pH 7.2, 100 mM NaCl, 10% v/v glycerol, 1 mM DTT; for analytical size-exclusion chromatography - same except 2.5% v/v glycerol plus 1 mM $MgCl_2$). Expression and purification of *lz*IDS11 was carried out with the same methodology. For *ps*TPS4, expression and purification conditions were identical except protein was eluted with 1 mL elution buffer, and concentrated and buffer exchanged using Vivaspin 500 (10 kDa MWCO; Merck, Feltham, UK) centrifugal concentrators.

2.3. SDS-PAGE and immunoblot

Proteins were resolved by SDS-PAGE on 5–20% gradient gels and transferred to Amersham Protran nitrocellulose membrane (Cytiva, Little Chalfont, UK). Membranes were reversibly stained with Ponceau S to ensure transfer was successful, and following blocking (5% w/v Marvel milk powder, 50 mM Tris-HCl pH 7.2, NaCl 150 mM), immunoblots were visualised using a mouse anti-hexahistidine tag primary antibody (1/1000; Bio-Rad AD1.1.10) and horseradish peroxidase-conjugated rabbit anti-mouse secondary antibody (1/5000; Agilent P0260; Agilent, Cheadle, UK).

2.4. Protein mass spectrometry

Protein (*lz*GGPPS - 7.5 μg) was submitted to SDS-PAGE, with bands excised, destained and dehydrated with acetonitrile. Bands were reduced (10 mM DTT), alkylated (55 mM iodoacetamide) and digested with trypsin (0.5 μg, Promega) at 37 °C for 18 h. Formic acid (1 μL) was added to protein digests to quench protease activity. The supernatant was removed from gel pieces, centrifuged at 3000×*g* for 2 min and retained for nanoLC-MS/MS analysis.

Samples (2.5 μL) were submitted to LC-MS/MS using an Ultimate 3000 RSLC nano HPLC system equipped with a Thermo PepMap 300 C18 column for separation (150 mm × 75 μm, 300 Å, 5 μm particle size) coupled to a Thermo LTQ FT Ultra Mass Spectrometer containing a nano-ESI source (Thermo Fisher, Loughborough, UK). The mobile phase consisted of a binary gradient of 95:5:0.1 water:acetonitrile:formic acid (A) and 5:95:0.1 water:acetonitrile:formic acid (B) with a flow rate of 400 μL/min (Table S1A). Protein sequencing was performed in data dependent acquisition (DDA) mode using collision induced dissociation (CID) fragmentation of the 3 most abundant ions in the precursor spectra. Thermo files were converted to mz5 format using MS Convert, part of the ProteoWizard package (Chambers et al., 2012). Database searching was performed using SearchGUI. X! Tandem was the search engine used, and post-processing was performed with Peptide Shaker (Barsnes and Vaudel, 2018); Peptide length 4–40, max missed cleavages 1, precursor tolerance 10 ppm with carbamidomethylation of C considered a fixed modification and oxidation of M and deamidation of N as variable modifications.

2.5. Isoprenyl diphosphate synthase (IDS) activity assays

IDS activity assays were carried out by combining *lz*GGPPS or *lz*IDS11 protein (5 μM) with allylic substrate [DMAPP or (*E*)-GPP or (*E,E* or *Z,E*)-FPP - 50 μM, Table S2] and IPP (150 or 100 or 50 μM for respective allylic substrate) in IDS buffer (25 mM MOPSO pH 7.2, 10% v/v glycerol, plus $MgCl_2$ or $ZnCl_2$ or Mn[II]Cl_2 as described in legends, 100 μL total volume). Assays were incubated at 30 °C for 1 h, before centrifugation at 5000 rpm for 2 min and addition of sodium diphosphate to 20 mM. For assays including (*E*)-3-methylpent-3-en-1-yl diphosphate (homo-IPP; HIPP), a 50:50 mix of IPP:HIPP was added to a final concentration as described above. HIPP was synthesized following a literature method (Hou et al., 2020). The spectroscopic data were essentially identical to those reported in the literature. ^{1}H NMR (400 MHz, D_2O) δ 5.31–5.21 (m, 1H), 3.86 (dd, $^3J_{P,H} = 6.8$ Hz, $^3J_{H,H} = 6.8$ Hz 2H), 2.21 (t, $^3J_{H,H} = 6.8$ Hz, 2H), 1.52 (s, 3H), 1.45 (d, $^3J_{H,H} = 6.7$ Hz, 3H); ^{13}C NMR (101 MHz, D_2O) δ 133.3, 121.2, 64.3 (d, $^2J_{P,C} = 4.9$ Hz, CH_2), 39.7 (d, $^3J_{P,C} = 12.0$ Hz, CH_2), 14.8, 12.6; ^{31}P NMR (162 MHz, D_2O) δ −7.14 (d, $^2J_{P,P} = 21.5$ Hz), −10.60 (d, $^2J_{P,P} = 21.8$ Hz).

Products were identified by reverse phase liquid chromatography-mass spectrometry (LC-MS) using a Waters Acquity Ultra Performance Liquid Chromatography (UPLC) system equipped with a BEH C18 column (1.7 μM, 2.1 mm × 50 mm; Waters Corporation, Wilmslow, UK) coupled to a Waters Synapt G1 Quadrupole Time-of-Flight MS operated in negative electrospray ionisation mode. Sample injection volume was 5 μL with a flow rate of 0.3 mL per minute and a column temperature of 20 °C, with the mobile phase consisting of a binary gradient of aqueous ammonium bicarbonate (5 mM, solvent A) and analytical-grade acetonitrile (solvent B; Table S1B/C give gradient parameters for short and

long separation methods). The capillary voltage was set at −2.3 kV, with desolvation gas flow rate and temperature at 600 L/h and 500 °C respectively. Products (and standards) were identified from their deprotonated ions (EIC *m/z* 313.1 for GPP, *m/z* 381.2 for FPP, *m/z* 395.2 for homo-FPP (HFPP),*m/z* 449.2 for GGPP, *m/z* 463.2 for homo-GGPP (HGGPP) and *m/z* 477.2 for bishomo-GGPP (HHGGPP)). For further confirmation of product by MS/MS, a fixed collision energy of 15 V was used to fragment the parent ion and release diphosphate (*m/z* 158.9) and phosphate ions (*m/z* 79.0). For confirmation of double bond stereochemistry, after incubation at 30 °C IDS assays were treated with 20 U shrimp alkaline phosphatase (New England Biolabs, M0371S, Ipswich, USA) for 4 h at 37 °C, with resulting alcohol products extracted using cyclohexane and submitted to GC-MS (see 2.7).

2.6. Diphosphate assay and Michaelis-Menten kinetic analysis

Assays for diphosphate release from enzyme activity were carried out using the EnzChek Pyrophosphate Assay Kit (Invitrogen E6645, Thermo Fisher, Loughborough, UK) with a modified protocol for 96-well plates described previously (Barja and Rodríguez-Concepción, 2020). Briefly, 5 μg *lz*GGPPS was combined with allylic substrate [DMAPP or (*E*)-GPP or (*E,E or Z,E*)-FPP – 50 μM] and IPP (150 or 100 or 50 μM) in reaction buffer (50 mM Tris-HCl pH 7.5, 1 mM $MgCl_2$, 0.2% Tween 20, 200 μL total volume) supplemented with 0.2 mM 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG), 0.2 U purine nucleoside phosphorylase (PNP) and 0.006 U inorganic diphosphatase. Following incubation for 1 h at 30 °C, absorbance was measured in triplicate at OD_{360} on a SPECTROstar Nano (BMG LABTECH, Aylesbury, UK), with background readings deducted and experimental samples quantified using a standard curve of sodium diphosphate (0–40 μM).

For kinetic analyses, enzyme activity was assessed by reading OD_{360} every minute for 30 min (30 °C) to obtain the initial rate. To determine apparent kinetic constants K_m and V_{max}, either (*E,E*)-FPP concentration was varied (8–240 μM) while holding IPP constant (100 μM), or IPP was varied (1–50 μM) with (*E,E*)-FPP constant (40 μM). Data were fitted by nonlinear regression with GraphPad Prism software using the Michaelis-Menten equation model. For inhibition studies, (*E,E,E*)-GGPP concentration was varied (0–100 μM) alongside constant concentrations of (*E,E*)-FPP (40 μM) and IPP (50 μM).

2.7. Terpene synthase assays

TPS activity assays were carried out by combining 5 μM *lz*GGPPS or *lz*IDS11 protein (or ~2 μM *ps*TPS4) with 50 μM (*E*)-GPP, (*E,E* or *Z,E*)-FPP or (*E,E,E*)-GGPP in TPS buffer (10 mM Tris-HCl pH 7.0, 10% v/v glycerol, 10 mM $MgCl_2$, 1 mM DTT, 200 μL total volume) overlaid with 200 μL cyclohexane. Assays were incubated at 30 °C for 90 min, before the aqueous and organic layers were mixed with a glass pipette, and separated by centrifugation at 3000 rpm for 2 min. The aqueous layer was discarded, and the organic layer evaporated to ~50 μL under a stream of nitrogen.

Products were identified by gas chromatography-mass spectrometry (GC-MS) using a Jeol AccuTOF GCx (Welwyn Garden City, UK) system equipped with a DB-5ms column (Agilent, Cheadle, UK). Sample injection volume was 5 μL, with helium carrier gas set at a constant flow rate of 1 mL/min and the inlet temperature set at 200 °C (Splitless mode). The oven started with a 4-min hold at 35 °C, followed by an increase of 10 °C/min to 300 °C and a 6 min hold. MS spectra were obtained in electron ionisation (EI+) mode within an *m/z* range of 40–500, following an 8-min solvent delay. Product identities were confirmed by comparing mass spectra and retention times to known standards/positive control assay products.

2.8. Analytical size-exclusion chromatography (SEC)

The native molecular weight (MW) in solution of *lz*GGPPS was assessed by analytical SEC using an ÄKTA pure fast protein liquid chromatography (FPLC) system (Cytiva, Little Chalfont, UK). Protein (0.25 mg) was applied using a 500 μL loading loop to a pre-equilibrated Superdex 200 Increase 10/300 GL column (Cytiva, Little Chalfont, UK) with a flow rate of 0.5 mL/min. Blue dextran 2000 (2000 kDa) was used to determine the column void volume (V_o), while two sets of protein standards were run separately for calibration curve construction [Thyroglobulin (669 kDa), Ferritin (440 kDa), Aldolase (158 kDa) and Ovalbumin (44 kDa)]; [Conalbumin (75 kDa), Carbonic anhydrase (29 kDa) and Aprotinin (6.5k Da)]. The partition coefficient (K_{av}) of *lz*GGPPS and protein standards was calculated using the equation $K_{av} = (V_e - V_o)/(V_t - V_o)$ where V_e = elution volume and V_t = total column volume. Protein standard K_{av} values were plotted against logMW using GraphPad Prism software and used to interpolate *lz*GGPPS apparent MW.

3. Results

3.1. The L. longipalpis genome contains a GGPPS homologue

We carried out blastp and tblastn searches to identify GGPPS homologues from the recently published *L. longipalpis* genome (Jacobina, NCBI accession PRJNA20279) and a transcriptomic study of the midgut of lab-raised female sandflies (Jacobina, NCBI accession PRJNA649559), looking for amino acid sequence similarities and particularly presence of the conserved aspartate rich motifs (Coutinho-Abreu et al., 2020; Labbé et al., 2023). Two strong candidates, which we named *lz*GGPPS1 (GITU01003271.1/A0A7G3AJG6) and *lz*GGPPS2 (A0A1B0CWK5) were identified (Yang and Wu, 2019). The former was obtained from both transcriptomic and genomic data and the latter from genomic data only. This study focuses on *lz*GGPPS1 (referred to from now on as *lz*GGPPS), as *lz*GGPPS2 was identical in sequence other than being truncated at the C-terminus (Fig. S1). Construction of a phylogenetic tree comprising various reported insect GGPPS proteins, in addition to examples from other eukaryotic organisms and bacteria, revealed that *lz*GGPPS clustered most closely to that of *D. melanogaster* (Fig. 2A). Comparative amino acid alignments of insect members showed high conservation, with the majority of variation present at the N and C-termini (Fig. 2B). Notably, the SARM was identical across insect proteins, with the final aspartate residue replaced with asparagine, as were the −4 and −5 positions upstream of the FARM (SS). Assessment of *lz*GGPPS for the presence of IPP binding motifs (IBMs) showed high conservation of residues expected to be important for IPP interaction and IDS activity (13/15 – Fig. S2)(Rebholz et al., 2023b). The N-terminus was interrogated for potential signalling/mitochondrial targeting peptides with none identified. We concluded that *lz*GGPPS was therefore likely to represent a *trans*-GGPPS enzyme from *L. longipalpis*.

3.2. lzGGPPS is an active enzyme capable of accepting various substrates

In order to study *lz*GGPPS functionality, recombinant protein was expressed in bacteria and purification was carried out using immobilised metal affinity chromatography (Fig. 3). Protein identity was confirmed with LC-MS/MS peptide mass fingerprinting and database searching (Table S3). Following incubation of *lz*GGPPS with various allylic substrates and IPP, assessment of product profile by LC-MS showed that the enzyme was active and was capable of synthesizing GGPP (Fig. 4A, S3). In agreement with other type III GGPPS enzymes, (*E,E*)-FPP appeared to be the preferred substrate, although *de novo* activity from DMAPP was also detected with an apparent (*E,E,E*) product configuration (Fig. 4B, S3E)(Sagami et al., 1994). To confirm product identity, MS/MS was utilised to fragment GGPP and release diphosphate and phosphate ions (Fig. S4). We saw no evidence of extension beyond GGPP to geranylfarnesyl diphosphate (data not shown). The stereochemistry of the GGPP produced from (*E,E*)-FPP and IPP was confirmed as (*E,E,E*), following conversion of the product to geranylgeraniol through

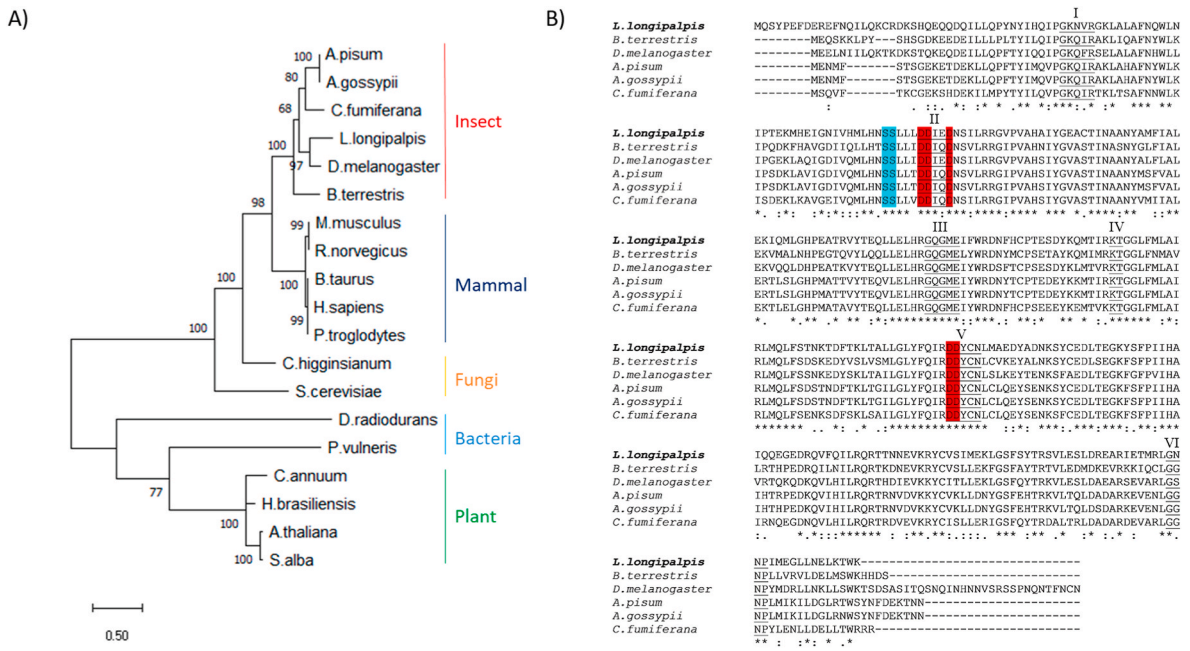

Fig. 2. Sequence analysis of geranylgeranyl diphosphate synthases (GGPPSs). **A**. Phylogenetic tree of selected insect, mammalian, fungi, bacterial and plant GGPPS enzymes inferred by maximum-likelihood analysis (unrooted). **B**. MUSCLE alignment of insect GGPPS enzymes. Conserved *trans*-IDS motifs are underlined (I-VI), with active site aspartate residues in the FARM (II) and SARM (V) highlighted in red (Ding et al., 2019). Residues upstream of the FARM (−4 and −5) important for determining product chain length are highlighted turquoise.

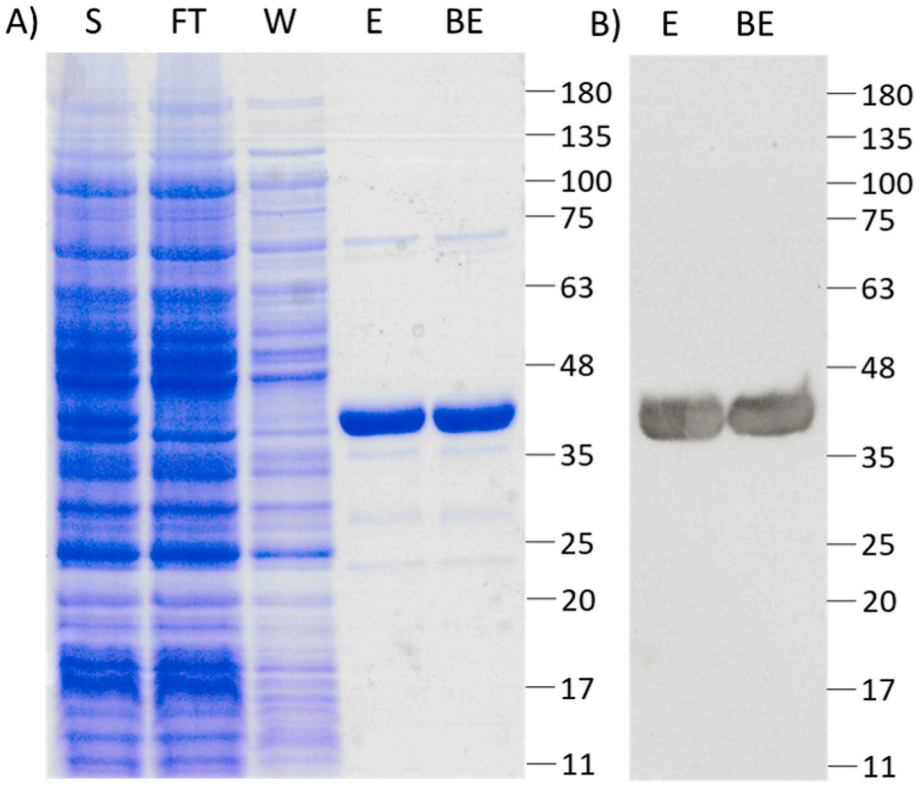

Fig. 3. Bacterial expression and purification of recombinant *lz*GGPPS. **A**. Protein resolved by SDS-PAGE (5–20% gradient) and visualised using Coomassie Blue R-250. **B**. Immunoblot of purified protein using αHis antibody. S - Soluble fraction; FT - Flow-through; W - Wash; E − Elution; BE - Buffer exchanged.

phosphatase treatment and assessment of retention time by GC compared to a dephosphorylated (*E,E,E*)-GGPP standard (Fig. S5). Enzyme activity was also assessed through an orthogonal plate assay screening for the release of inorganic diphosphate (PPi) following condensation of isoprenoid units (see 2.6). Pleasingly, the relative activity across allylic substrate precursors was broadly consistent with LC-MS data (Fig. 4C). Of note, we found IPP trilithium salt (Supelco 39784) to give high background in PPi readings while IPP triammonium

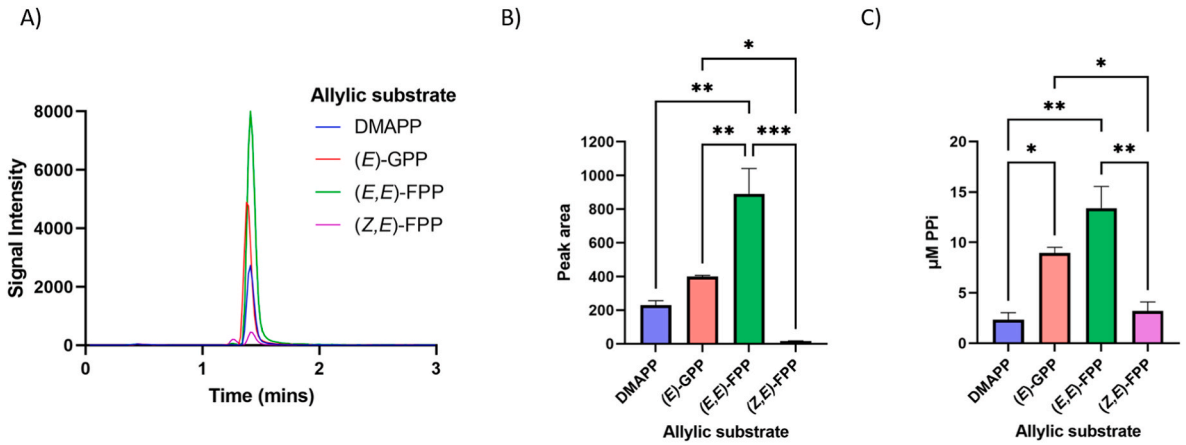

Fig. 4. Functional activity of recombinant *lz*GGPPS. **A**. Overlay of representative LC-MS resolved GGPP product peaks (EIC *m/z* 449.2) following incubation of *lz*GGPPS (5 μM) with various allylic substrates (50 μM) and IPP (50–150 μM) in the presence of 10 mM $MgCl_2$ for 1 h at 30 °C. **B**. GGPP product peak areas from LC-MS analysis displayed as averages of three independent experiments (integration of early eluting GGPP peak from (*Z,E*)-FPP is represented, ±SEM). **C**. Pyrophosphate release following incubation of *lz*GGPPS (5 μg) with various allylic substrates (50 μM) and IPP (50–150 μM) for 1 h at 30 °C displayed as averages of three independent experiments (±SEM). Statistical analyses were performed using one-way ANOVA with Tukey's multiple comparisons test; significance is reported in figures by $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

salt (Echelon I-0050) did not, meaning the latter was chosen for use in subsequent assays.

Interestingly, addition of EDTA (20 mM) to inactivate *lz*GGPPS resulted in release of a small amount of GGPP from the enzyme, as seen by LC-MS, which was also the case with incubation of active enzyme with allylic substrate only (DMAPP), but not IPP only (data not shown). In previous work, human GGPPS was shown to co-crystallise with sequestered GGPP bound in the active site, which may explain this observation. GGPP of bacterial origin co-purifies with *lz*GGPPS and is released following enzyme inactivation or competition for binding with allylic substrate precursors (Kavanagh et al., 2006). Levels of GGPP released in this way were always significantly lower (by LC-MS) than those seen with productive substrates - a result supported by the PPi kinetic assays. To rule out any IDS activity from contaminating *E. coli*-derived proteins, we carried out LC-MS assays with another IDS-like bacterially expressed protein preparation from *Lutzomyia* that we have found to be inactive. This enzyme, which we have named *lz*IDS11 (GITU01006826/A0A1B0CVV3), gave no diphosphate products with any combination of allylic substrate and IPP (data not shown) confirming the absence of activity carrying over from the expression system.

In addition to the more familiar (*E,E*)-isomer, (*Z,E*)-FPP was also tolerated as a substrate of *lz*GGPPS, albeit with significantly reduced output. This experiment yielded two distinct GGPP peaks (Fig. S3D). The early-eluting GGPP peak is likely to be (*E,Z,E*)-GGPP from (*Z,E*)-FPP extension, whereas the later peak is displaced (*E,E,E*)-GGPP from *E. coli* as described above. The homosesquiterpenes produced in *L. longipalpis* contain an additional methyl group, probably due to incorporation of homo-IPP (HIPP) into homo-FPP by FPPS enzymes, before cyclisation into the final terpene products (Hamilton et al., 1999a, 1999b). The diterpene, sobralene, is – by contrast – a regular diterpene. Despite this, we introduced HIPP into our IDS assays to assess whether *lz*GGPPS is capable of producing homo-isoprenoid chains when provided with the appropriate substrate. With an equimolar mix of HIPP and IPP, *lz*GGPPS was able to extend (*E,E*)-FPP into both homo-GGPP (HGPP) and GGPP, although the latter was the major product by a ratio of approximately 1:3, when taking into account the sequestered GGPP of bacterial origin (Fig. S6A). Similarly, incubating (*E*)-GPP with HIPP and IPP gave the full range of products possible, including bishomo-GGPP (HHGGPP), although regular GGPP was by far the most prevalent (GGPP:HGGPP:HHGGPP 20:3:1), suggesting HIPP is a non-optimal substrate for *lz*GGPPS (Fig. S6B).

The identity and prevalence of the divalent metal cation can impact insect IDS isoprenoid product specificity and activity (Frick et al., 2013; Rivera-Perez et al., 2015). A comparative analysis of the impact of various divalent metal ions on *lz*GGPPS enzymatic activity with DMAPP + IPP as substrates suggested that Mg^{2+} was the most effective cofactor (Fig. 5A). There was some activity with the inclusion of Zn^{2+} and Mn^{2+}, while Co^{2+} gave no product and led to formation of an insoluble precipitate. Interestingly, some enzyme activity was retained in the absence of metal addition, possibly due to *lz*GGPPS co-purifying with Mg^{2+} and/or GGPP from the bacterial expression system. With (*E*)-GPP as the allylic substrate, there was evidence of some minor FPP release, consistent with previously characterised GGPPS enzymes (Fig. 5B) (Sagami et al., 1994; Kavanagh et al., 2006; Barbar et al., 2013). However, increasing Mg^{2+} concentration appeared to inhibit FPP production. Further comparison of metal ions with (*E*)-GPP + IPP as substrates confirmed Mg^{2+} as the best *lz*GGPPS cofactor for promoting GGPP production (Fig. S7).

3.3. Pseudo terpenoid synthase activity in lzGGPPS

A recent study on the postman butterfly identified GGPPS-like enzymes exhibiting TPS activity for production of its anti-aphrodisiac pheromone, the monoterpene (*E*)-β-ocimene, (Darragh et al., 2021). We assessed whether *lz*GGPPS possesses similar activity by incubating the enzyme with substrate precursors commensurate with mono- (GPP), sesqui- (FPP) and diterpene (GGPP) synthesis. No mono- or diterpene activity could be detected by GC-MS analysis, which was also the case for (*Z,E*)-FPP (data not shown). However, when incubated with (*E,E*)-FPP, two product peaks were visible by GC-MS and identified as the sesquiterpene alcohols nerolidol (major peak) and farnesol (minor peak, Fig. S8A). This did not occur when the negative control (*lz*IDS11) was incubated with (*E,E*)-FPP, ruling out activity from contaminating bacterial phosphatases (Fig. S8B). Incubation of a characterised TPS from the striped flea beetle (*Phyllotreta striolata*, *ps*TPS4) with (*E,E*)-FPP produced (*E*)-nerolidol (as expected, Fig. S8C) (Beran et al., 2016). The spectrum of the product from *lz*GGPPS matched that produced by *ps*TPS4, and the retention time was identical, confirming that *lz*GGPPS produces (*E*)-nerolidol. Use of a dephosphorylated (*E,E*)-FPP standard confirmed the identity of the minor product as (*E,E*)-farnesol.

To estimate the relative TPS versus GGPPS activity of *lz*GGPPS, diphosphate release was screened following incubation of the enzyme with (*E,E*)-FPP in the presence or absence of IPP using a diphosphate

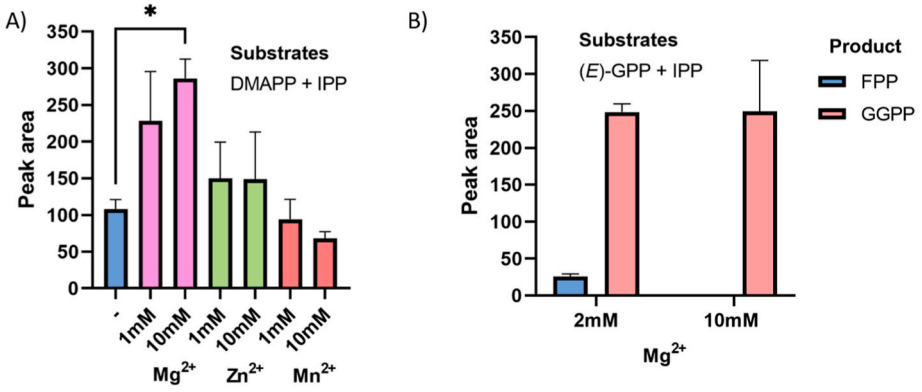

Fig. 5. Dependence of recombinant *lz*GGPPS activity on metal ion cofactors. A. LC-MS-derived GGPP product peak areas (EIC *m/z* 449.2) following incubation of *lz*GGPPS (5 μM) with DMAPP (50 μM) and IPP (150 μM) in the presence or absence of various divalent metal ions for 1 h at 30 °C displayed as averages of three independent experiments (±SEM). Statistical analyses were performed using one-way ANOVA with Dunnett's multiple comparisons test; significance is reported in figures by $^{*}P < 0.05$. b) FPP (*m/z* 381.2). **B.** GGPP (*m/z* 449.2) product peak areas following incubation of *lz*GGPPS (5 μM) with (*E*)-GPP (50 μM) and IPP (100 μM) in the presence 2 mM or 10 mM $MgCl_2$ for 1 h at 30 °C displayed as averages of three independent experiments (±SEM).

assay kit. PPi release (reflecting *lz*GGPPS activity) was negligible in the absence of IPP when compared against either *lz*GGPPS with IPP or *ps*TPS4 with (*E,E*)-FPP only (Fig. S9A). The diphosphate assay was also used to confirm that incubation of (*E,E*)-FPP with *lz*IDS11 yielded negligible background activity (Fig. S9B). We therefore conclude that *lz*GGPPS is predominantly an active GGPPS with some pseudo-TPS character caused by activation of the allylic substrate and quenching of the carbocation by water in the absence of IPP co-substrate to yield the isomeric sesquiterpene alcohols.

3.4. *lzGGPPS and product inhibition*

For a more detailed biochemical analysis of *lz*GGPPS, we carried out Michaelis-Menten assessments of the enzyme's kinetics. Activity of GGPPS enzymes has been shown to be stimulated by the addition of surfactants (Sagami et al., 1994; Kavanagh et al., 2006). In agreement with this, addition of 0.2% Tween 20 significantly increased PPi release by *lz*GGPPS (Fig. S9C). Under these conditions, when varying (*E,E*)-FPP concentration while holding IPP constant, *lz*GGPPS gave an apparent V_{max} of 0.48 μM PPi/min/μg protein and a K_m of 167.4 μM (Fig. 6A, Table 1). This represents a very high K_m compared with other reported values from GGPPS enzymes, which are generally in the range of 2–20 μM (Kavanagh et al., 2006; Chang et al., 2006; Barbar et al., 2013; Heider et al., 2014). With IPP as the variable substrate, V_{max} was 0.14 μM PPi/min/μg protein and K_m was 7.9 μM, which more closely aligns with previous studies of GGPPSs (Fig. 6B, Table 1). Introduction of GGPP into PPi assays significantly reduced enzyme velocity (Fig. 6C). This is in agreement with an established consensus that GGPP product acts as a negative feedback inhibitor, with phosphates bound in the allylic site of the catalytic cavity of type III enzymes (Fig. S10)(Sagami et al., 1994; Kavanagh et al., 2006; Barbar et al., 2013).

Table 1
Apparent kinetic parameters of *lz*GGPPS for substrates (*E,E*)-FPP and IPP.

Variable Substrate	Co-substrate	K_m (μM)	V_{max} (μM PPi/min/μg)	k_{cat} (min^{-1})
(*E,E*)-FPP	IPP	167.4 (±30.9)	0.48 (±0.05)	3.9
IPP	(*E,E*)-FPP	7.9 (±1.4)	0.14 (±0.01)	1.2

3.5. *lzGGPPS exhibits hexameric quaternary structure*

To determine the MW of *lz*GGPPS in solution, we employed analytical SEC to separate proteins according to hydrodynamic radius. Comparison of column retention to a range of protein standards gave an apparent MW of 221 kDa (Fig. 7A and B). Immunoblot of collected fractions confirmed that *lz*GGPPS was the source of the A_{280} peak (Fig. S11). Given that a monomer of recombinant *lz*GGPPS is ~41 kDa, this implies a hexameric native configuration is likely (~246 kDa), consistent with the arrangement of the human enzyme (Kavanagh et al., 2006)(Fig. 7C).

4. Discussion

In eukaryotes, synthesis of GGPP is essential for geranylgeranylation of target proteins, granting them a hydrophobic C-terminus and promoting interaction with cellular membranes (Wang and Casey, 2016). Geranylgeranylation of RAS proteins in *Bombyx mori* was found to be the sole form of prenylation, whereas in mammals farnesylation is more

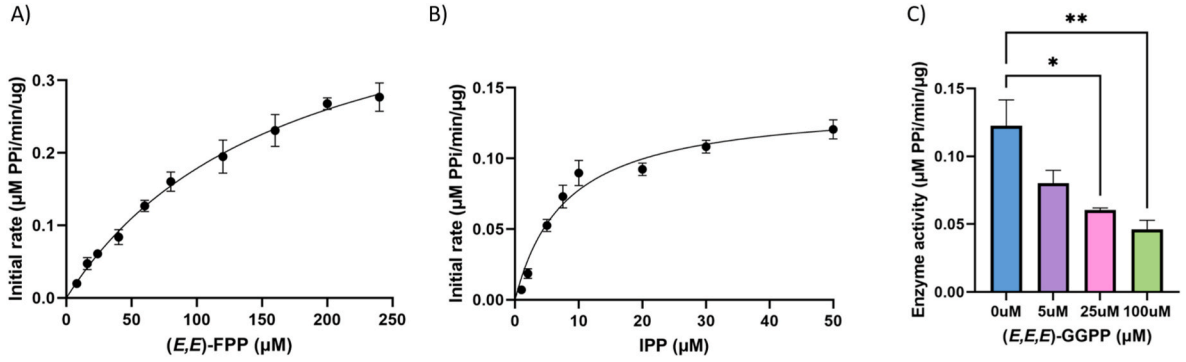

Fig. 6. Michaelis-Menten kinetic analysis of recombinant *lz*GGPPS activity. A. Plot measuring the initial rate of PPi release with varying concentrations of (*E,E*)-FPP and a constant IPP concentration (100 μM), displayed as averages of three independent experiments (±SEM). **B.** As for **A** with varying concentrations of IPP and a constant (*E,E*)-FPP concentration (40 μM). **C.** Initial rate of PPi release from *lz*GGPPS activity with varying concentrations of added GGPP and constant concentrations of (*E,E*)-FPP (40 μM) and IPP (50 μM), displayed as averages of three independent experiments (±SEM). Statistical analyses were performed using one-way ANOVA with Dunnett's multiple comparisons test; significance is reported in figures by $^{*}P < 0.05$; $^{**}P < 0.01$.

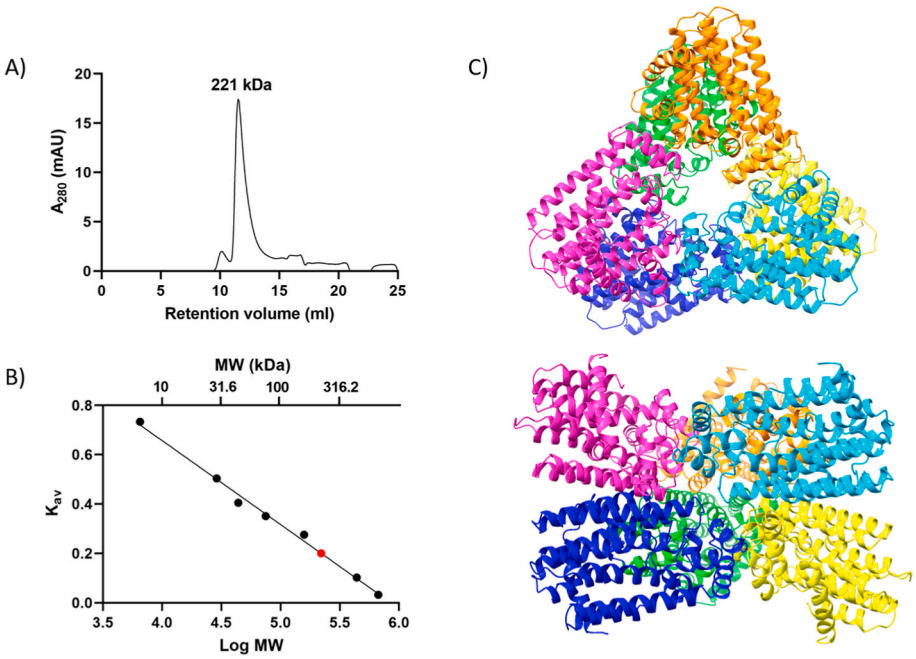

Fig. 7. Quaternary structure of recombinant *lz*GGPPS. **A**. Analytical SEC plot of A_{280} against retention volume following injection of *lz*GGPPS on Superdex 200 Increase 10/300 GL column (Cytiva). **B**. Plot of the partition coefficient (K_{av}) against log MW for a range of protein standards, with the relative position of *lz*GGPPS indicated in red. **C**. I-TASSER structural prediction of *lz*GGPPS arranged into a hexamer in line with structure of human GGPPS (Kavanagh et al., 2006).

prevalent (Moriya et al., 2010). Moreover, in some insects GGPP is required for production of diterpene/terpenoid semiochemicals, and for carotenoid synthesis (Hojo et al., 2011; Prchalová et al., 2016; Palframan et al., 2018; Ding et al., 2019). We have identified and biochemically characterised a GGPPS enzyme from New World phlebotomine sandflies, which is capable of producing the precursor for protein prenylation and/or sobralene production. This represents the first detailed study of an IDS from *L. longipalpis* and only the fourth functional study of an insect GGPPS; the others being in *C. fumiferana*, *A. gossypii* and *M. viciae* (Barbar et al., 2013; Zhang and Li, 2014; Song et al., 2022).

The catalytic behaviour of *lz*GGPPS appears to agree generally with other studies of type III enzymes, where allylic substrates are preferred in the order (*E,E*)-FPP > (*E*)-GPP > DMAPP. While some activity remained with the non-typical allylic substrate (*Z,E*)-FPP, this was significantly lower than that for (*E,E*)-FPP, as would be expected for a *trans*-IDS. We also found that *lz*GGPPS could accept HIPP as a substrate, albeit less efficiently than conventional IPP. Substrate promiscuity in IDS enzymes has been long established, particularly in FPPSs, which can accept a range of artificial isoprenoids at both allylic and non-allylic binding sites (see Couillaud et al., 2022 for a recent review). Experimental evidence from the microorganism *Sulfolobus acidocaldarius* GGPPS has found it to be similarly open to atypical substrates, including HIPP (Ohnuma et al., 1998; Nagaki et al., 2004). Lepidoptera have evolved FPPS enzymes that selectively extend homo (ethyl-branched)-DMAPP for juvenile hormone synthesis, which can also act as a non-optimal substrate for spruce budworm GGPPS (Sen et al., 2007; Barbar et al., 2013). The fact that some chemotypes of *L. longipalpis* produce the homosesquiterpene pheromone components 9-methylgermacrene-B (7) and 3-methyl-α-himachalene (8) with highly defined regiochemistry (the additional carbon is incorporated in the second isoprene unit only) raises the question of how this is achieved mechanistically. Is it, for example, controlled by the flux/location of HIPP vs IPP, or do one or more of the enzymes in the biosynthetic pathway select for a specific HFPP isomer? If the latter, it could be envisaged that either a specialist HFPPS exists, or that the homosesquiterpene synthases favour one form of HFPP, or that these two properties work in tandem. Sobralene (6) is a regular diterpene, possessing only C_5 isoprene units, the precursor of which would be the principal product of the *L. longipalpis* GGPPS described here. Despite this, our results show that the enzyme can incorporate HIPP, albeit at a reduced rate compared to IPP, if the former is made available. Thus, the interplay between substrate-availability and the presence of IDSs and TPSs with varying specificities is likely to be the key to different terpene product profiles in this organism.

Despite appearing to be the most efficient substrate of *lz*GGPPS, (*E,E*)-FPP gave a very high apparent K_m value following kinetic analysis. Currently, the only steady-state kinetic report of an insect GGPPS is for spruce budworm, with a K_m of 2.1 μM with respect to FPP, suggesting *lz*GGPPS is a significantly less efficient enzyme with a lower allylic substrate affinity (Barbar et al., 2013). Nonetheless, for IPP K_m readings are comparable between spruce budworm and sandfly, calculated at 4.0 μM and 7.9 μM respectively (Barbar et al., 2013). Despite use of different methods in assessment of enzyme velocity, V_{max} estimates were also broadly similar (for (*E,E*)-FPP 2456 pmol/h/μg protein for *C. fumiferana* versus 5760 pmol/h/μg for *L. longipalpis*; for IPP 1904 pmol/h/μg versus 1680 pmol/h/μg) (Barbar et al., 2013). Moreover, k_{cat} readings for *lz*GGPPS with regard to IPP (1.2 min^{-1} = 0.02 s^{-1}) reflect a ~10-fold reduction compared to human enzyme (0.204 s^{-1}) while being close to values reported for *S. cerevisiae* (0.025 s^{-1}) (Kavanagh et al., 2006; Chang et al., 2006). Comparison to closer evolutionary relatives of *L. longipalpis*, such as *D. melanogaster* (*quemao*) will inform whether these kinetic constants are similar across dipteran insects (Lai et al., 1998).

In contrast to plants and bacteria, insect TPSs have begun to be characterised only relatively recently. From the examples we know to date, they appear to have evolved from IDS ancestors following gene duplication (Gilg et al., 2009; Beran et al., 2016; Lancaster et al., 2019). Genome analysis of *Heliconius melponene* butterflies identified several predicted GGPPSs, with expression analysis revealing a male bias in two of these, and subsequent functional characterisation revealing TPS activity in both (Darragh et al., 2021). We found that *lz*GGPPS is able to dephosphorylate (*E,E*)-FPP in the absence of IPP and release (*E*)-nerolidol and (*E,E*)-farnesol as products. (*E*)-Nerolidol production from FPP by heterologously expressed insect TPSs has been described on more than one occasion (Beran et al., 2016; Darragh et al., 2021). It is a

common constituent in plant essential oil and can stimulate a defensive response against herbivorous insects and pathogens (Chan et al., 2016; Chen et al., 2020b). (*E,E*)-Farnesol is an important intermediate in juvenile hormone production in insects (Noriega, 2014). More broadly, production of terpenoid alcohols in insects by IDS-like TPS enzymes has also been reported in the tea green leafhopper *Empoasca onukii* (geraniol) and the stinkbugs *Murgantia histrionica*, *Halymorpha halys* and *Nezara viridula* (all sesquipiperitol)(Zhou et al., 2019; Lancaster et al., 2018; Rebholz et al., 2023a). However, sesquiterpenoid alcohol production by *lz*GGPPS occurs at low levels and is likely due to quenching of the allylic carbocation intermediate by water in the absence of nucleophilic IPP, giving either nerolidol or farnesol depending on the site of attack. This reaction may be even less favourable *in vivo* through tight control of FPP/IPP substrate flux from metabolic pathways. Furthermore, the use of recombinant proteins and substrate at concentrations far above those found *in cellulo* may also yield non-biologically relevant products.

Insect GGPPS enzymes have previously been predicted to adopt a hexameric arrangement based on conservation of three hydrophobic regions important for inter-dimer interactions in the human enzyme (Barbar et al., 2013). For the first time, we have confirmed this to be the case, with evidence pointing to a complex of 5–6 monomer units and ruling out the dimeric complex seen in non-animal GGPPSs. The apparent molecular weight in solution of previously-described human GGPPS also appeared to correspond to a complex numbering 5–6 units, with the crystal structure confirming a propeller-bladed trimer of dimers (Kavanagh et al., 2006). It is hence likely that this assembly is represented across all insect GGPPS enzymes.

In summary, we have elucidated the enzymatic function of a protein responsible for isoprenoid biosynthesis in *L. longipalpis*, the first to be characterised for this species complex. Further work is required to identify potential TPSs associated with downstream terpene synthesis, and how this is regulated in the sandfly. A greater understanding and biochemical reconstitution of these pathways could give a route to sustainable aggregation pheromone production for use in control pathways - a potentially viable route to limit the spread of leishmaniasis.

Data Access Statement

Protein mass spectrometry data created during this research are openly available from the University of Nottingham data repository at https://doi.org/10.17639/nott.7330.

Acknowledgements

The authors gratefully acknowledge the Biotechnology and Biological Sciences Research Council (BBSRC) for research grant BB/V003933/1. We also thank Robert Layfield and Daniel Scott for provision of lab space and equipment, Ben Pointer-Gleadhill for assistance with GC-MS and Cameron Baines for preparing the image of the *lz*GGPPS hexameric model.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmb.2023.104001.

References

Armenteros, J.J.A., Salvatore, M., Emanuelsson, O., Winther, O., Von Heijne, G., Elofsson, A., Nielsen, H., 2019. Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2, 1–14. https://doi.org/10.26508/lsa.201900429.

Banerjee, A., Sharkey, T.D., 2014. Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. Nat. Prod. Rep. 31, 1043–1055. https://doi.org/10.1039/c3np70124g.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., Miyano, S., 2002. Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18, 298–305. https://doi.org/10.1093/bioinformatics/18.2.298.

Barbar, A., Couture, M., Sen, S.E., Béliveau, C., Nisole, A., Bipfubusa, M., Cusson, M., 2013. Cloning, expression and characterization of an insect geranylgeranyl diphosphate synthase from Choristoneura fumiferana. Insect Biochem. Mol. Biol. 43, 947–958. https://doi.org/10.1016/j.ibmb.2013.07.004.

Barja, M.V., Ezquerro, M., Beretta, S., Diretto, G., Florez-Sarasa, I., Feixes, E., Fiore, A., Karlova, R., Fernie, A.R., Beekwilder, J., Rodríguez-Concepción, M., 2021. Several geranylgeranyl diphosphate synthase isoforms supply metabolic substrates for carotenoid biosynthesis in tomato. New Phytol. 231, 255–272. https://doi.org/10.1111/nph.17283.

Barja, M.V., Rodríguez-Concepción, M., 2020. A simple in vitro assay to measure the activity of geranylgeranyl diphosphate synthase and other short-chain prenyltransferases. Methods Mol. Biol. (2083), 27–38.

Barsnes, H., Vaudel, M., 2018. SearchGUI: a highly adaptable common interface for proteomics search and de Novo engines. J. Proteome Res. 17, 2552–2555. https://doi.org/10.1021/acs.jproteome.8b00175.

Beck, G., Coman, D., Herren, E., Ruiz-Sola, M.Á., Rodríguez-Concepción, M., Gruissem, W., Vranová, E., 2013. Characterization of the GGPP synthase gene family in Arabidopsis thaliana. Plant Mol. Biol. 82, 393–416. https://doi.org/10.1007/s11103-013-0070-z.

Beran, F., Rahfeld, P., Luck, K., Nagel, R., Vogel, H., Wielsch, N., Irmisch, S., Ramasamy, S., Gershenzon, J., Heckel, D.G., Köllner, T.G., 2016. Novel family of terpene synthases evolved from trans-isoprenyl diphosphate synthases in a flea beetle. Proc. Natl. Acad. Sci. U.S.A. 113, 2922–2927. https://doi.org/10.1073/pnas.1523468113.

Buhaescu, I., Izzedine, H., 2007. Mevalonate pathway: a review of clinical and therapeutical implications. Clin. Biochem. 40, 575–584. https://doi.org/10.1016/j.clinbiochem.2007.03.016.

Chambers, M.C., MacLean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L., Mallick, P., 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nat. Biotechnol. 30, 918–920. https://doi.org/10.1038/nbt.2377.

Chan, W.K., Tan, L.T.H., Chan, K.G., Lee, L.H., Goh, B.H., 2016. Nerolidol: a sesquiterpene alcohol with multi-faceted pharmacological and biological activities. Molecules. https://doi.org/10.3390/molecules21050529.

Chang, T.H., Guo, R.T., Ko, T.P., Wang, A.H.J., 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. https://doi.org/10.1074/jbc.M512886200.

Chen, C.C., Zhang, L., Yu, X., Ma, L., Ko, T.P., Guo, R.T., 2020a. Versatile cis-isoprenyl diphosphate synthase superfamily members in catalyzing carbon-carbon bond formation. ACS Catal. 10, 4717–4725. https://doi.org/10.1021/acscatal.0c00283.

Chen, S., Zhang, L., Cai, X., Li, X., Bian, L., Luo, Z., Li, Z., Chen, Z., Xin, Z., 2020b. (E)-Nerolidol is a volatile signal that induces defenses against insects and pathogens in tea plants. Hortic. Res. 7 https://doi.org/10.1038/s41438-020-0275-7.

Coman, D., Altenhoff, A., Zoller, S., Gruissem, W., Vranová, E., 2014. Distinct evolutionary strategies in the GGPPS family from plants. Front. Plant Sci. 5, 1–12. https://doi.org/10.3389/fpls.2014.00230.

Couillaud, J., Duquesne, K., Iacazio, G., 2022. Extension of the terpene chemical space: the very first biosynthetic steps. Chembiochem 23. https://doi.org/10.1002/cbic.202100642.

Coutinho-Abreu, I.V., Serafim, T.D., Meneses, C., Kamhawi, S., Oliveira, F., Valenzuela, J.G., 2020. Leishmania infection induces a limited differential gene expression in the sand fly midgut. BMC Genom. 21, 1–16. https://doi.org/10.1186/s12864-020-07025-8.

Darragh, K., Orteu, A., Black, D., Byers, K.J.R.P., Szczerbowski, D., Warren, I.A., Rastas, P., Pinharanda, A., Davey, J.W., Garza, S.F., Almeida, D.A., Merrill, R.M., McMillan, W.O., Schulz, S., Jiggins, C.D., 2021. A novel terpene synthase controls differences in anti-aphrodisiac pheromone production between closely related Heliconius butterflies. PLoS Biol. 19, 1–30. https://doi.org/10.1371/JOURNAL.PBIO.3001022.

Ding, B.Y., Niu, J., Shang, F., Yang, L., Chang, T.Y., Wang, J.J., 2019. Characterization of the geranylgeranyl diphosphate synthase gene in Acyrthosiphon pisum (Hemiptera: aphididae) and its association with carotenoid biosynthesis. Front. Physiol. 10, 1–10. https://doi.org/10.3389/fphys.2019.01398.

Feng, Y., Morgan, R.M.L., Fraser, P.D., Hellgardt, K., Nixon, P.J., 2020. Crystal structure of geranylgeranyl pyrophosphate synthase (CrtE) involved in cyanobacterial terpenoid biosynthesis. Front. Plant Sci. 11, 1–14. https://doi.org/10.3389/fpls.2020.00589.

Frick, S., Nagel, R., Schmidt, A., Bodemann, R.R., Rahfeld, P., Pauls, G., Brandt, W., Gershenzon, J., Boland, W., Burse, A., 2013. Metal ions control product specificity of isoprenyl diphosphate synthases in the insect terpenoid pathway. Proc. Natl. Acad. Sci. U.S.A. 110, 4194–4199. https://doi.org/10.1073/pnas.1221489110.

Gilg, A.B., Tittiger, C., Blomquist, G.J., 2009. Unique animal prenyltransferase with monoterpene synthase activity. Naturwissenschaften 96, 731–735. https://doi.org/10.1007/s00114-009-0521-1.

Hamilton, J.G.C., Hooper, A.M., Ibbotson, H.C., Kurosawa, S., Mori, K., Muto, S., Pickett, J.A., 1999a. 9-Methylgermacrene-B is confirmed as the sex pheromone of the sandfly Lutzomyia longipalpis from Lapinha, Brazil, and the absolute

stereochemistry defined as S. Chem. Commun. 2335–2336. https://doi.org/10.1039/a900242a.
Hamilton, J.G.C., Hooper, A.M., Mori, K., Pickett, J.A., Sano, S., 1999b. 3-Methyl-α-himachalene is confirmed, and the relative stereochemistry defined, by synthesis as the sex pheromone of the sandfly Lutzomyia longipalpis from Jacobina. Brazil. Chem. Commun. 355–356. https://doi.org/10.1039/a900242a.
Heider, S.A.E., Peters-Wendisch, P., Beekwilder, J., Wendisch, V.F., 2014. IdsA is the major geranylgeranyl pyrophosphate synthase involved in carotenogenesis in Corynebacterium glutamicum. FEBS J. 281, 4906–4920. https://doi.org/10.1111/febs.13033.
Hojo, M., Toga, K., Watanabe, D., Yamamoto, T., Maekawa, K., 2011. High-level expression of the Geranylgeranyl diphosphate synthase gene in the frontal gland of soldiers in Reticulitermes speratus (Isoptera: rhinotermitidae). Arch. Insect Biochem. Physiol. 77, 17–31. https://doi.org/10.1002/arch.20415.
Hou, A., Lauterbach, L., Dickschat, J.S., 2020. Enzymatic synthesis of methylated terpene analogues using the plasticity of bacterial terpene synthases. Chem. Eur J. 26, 2178–2182. https://doi.org/10.1002/chem.201905827.
Kavanagh, K.L., Dunford, J.E., Bunkoczi, G., Russell, R.G.G., Oppermann, U., 2006. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. J. Biol. Chem. 281, 22004–22012. https://doi.org/10.1074/jbc.M602603200.
Kloer, D.P., Welsch, R., Beyer, P., Schulz, G.E., 2006. Structure and reaction geometry of geranylgeranyl diphosphate synthase from Sinapis alba. Biochemistry 45, 15197–15204. https://doi.org/10.1021/bi061572k.
Kuzuguchi, T., Morita, Y., Sagami, I., Sagami, H., Ogura, K., 1999. Human geranylgeranyl diphosphate synthase: cDNA cloning and expression. J. Biol. Chem. 274, 5888–5894. https://doi.org/10.1074/jbc.274.9.5888.
Labbé, F., Abdeladhim, M., Abrudan, J., Araki, A.S., Araujo, R.N., Arensburger, P., Benoit, J.B., Brazil, R.P., Bruno, R.V., Rivas, G.B. da S., de Abreu, V.C., Charamis, J., Coutinho-Abreu, I.V., da Costa-Latgé, S.G., Darby, A., Dillon, V.M., Emrich, S.J., Fernandez-Medina, D., Gontijo, N.F., Flanley, C.M., Gatherer, D., Genta, F.A., Gesing, S., Giraldo-Calderón, G.I., Gomes, B., Aguiar, E.R.G.R., Hamilton, J.G.C., Hamarsheh, O., Hawksworth, M., Hendershot, J.M., Hickner, P.V., Imler, J.L., Ioannidis, P., Jennings, E.C., Kamhawi, S., Karageorgiou, C., Kennedy, R.C., Krueger, A., Latorre-Estivalis, J.M., Ligoxygakis, P., Meireles-Filho, A.C.A., Minx, P., Miranda, J.C., Montague, M.J., Nowling, R.J., Oliveira, F., Ortigão-Farias, J., Pavan, M.G., Pereira, M.H., Pitaluga, A.N., Olmo, R.P., Ramalho-Ortigao, M., Ribeiro, J.M.C., Rosendale, A.J., Sant'anna, M.R.V., Scherer, S.E., Secundino, N.F.C., Shoue, D.A., Moraes, C. da S., Gesto, J.S.M., Souza, N.A., Syed, Z., Tadros, S., Teles-De-freitas, R., Telleria, E.L., Tomlinson, C., Traub-Csekö, Y.M., Marques, J.T., Tu, Z., Unger, M.F., Valenzuela, J., Ferreira, F.V., de Oliveira, K.P.V., Vigoder, F.M., Vontas, J., Wang, L., Weedall, G.D., Zhioua, E., Richards, S., Warren, W.C., Waterhouse, R.M., Dillon, R.J., McDowell, M.A., 2023. Genomic analysis of two phlebotomine sand fly vectors of Leishmania from the New and Old World. PLoS Neglected Trop. Dis. https://doi.org/10.1371/journal.pntd.0010862.
Lai, C., McMahon, R., Young, C., Mackay, T.F.C., Langley, C.H., 1998. quemao, a Drosophila bristle locus, encodes geranylgeranyl pyrophosphate synthase. Genetics 149, 1051–1061. https://doi.org/10.1093/genetics/149.2.1051.
Lainson, R., Ward, R., Shaw, J., 1977. Experimental transmission of Leishmania chagasi, causative agent of neotropical visceral leishmaniasis, by the sandfly Lutzomyia longipalpis. Nature 266, 628–630.
Lancaster, J., Khrimian, A., Young, S., Lehner, B., Luck, K., Wallingford, A., Ghosh, S.K. B., Zerbe, P., Muchlinski, A., Marek, P.E., Sparks, M.E., Tokuhisa, J.G., Tittiger, C., Köllner, T.G., Weber, D.C., Gundersen-Rindal, D.E., Kuhar, T.P., Tholl, D., 2018. De novo formation of an aggregation pheromone precursor by an isoprenyl diphosphate synthase-related terpene synthase in the harlequin bug. Proc. Natl. Acad. Sci. U.S.A. 115, E8634–E8641. https://doi.org/10.1073/pnas.1800008115.
Lancaster, J., Lehner, B., Khrimian, A., Muchlinski, A., Luck, K., Köllner, T.G., Weber, D. C., Gundersen-Rindal, D.E., Tholl, D., 2019. An IDS-type sesquiterpene synthase produces the pheromone precursor (Z)-α-Bisabolene in Nezara viridula. J. Chem. Ecol. 45, 187–197. https://doi.org/10.1007/s10886-018-1019-0.
Moriya, K., Tsubota, T., Ishibashi, N., Yafune, A., Suzuki, T., Kobayashi, J., Shiotsuki, T., Utsumi, T., 2010. Bombyx mori Ras proteins BmRas1, BmRas2 and BmRas3 are neither farnesylated nor palmitoylated but are geranylgeranylated. Insect Mol. Biol. 19, 291–301. https://doi.org/10.1111/j.1365-2583.2009.00982.x.
Nagaki, M., Miki, Y., Nakada, M., Kawakami, J., Kitahara, H., Maki, Y., Gotoh, Y., Nishino, T., Koyama, T., 2004. Substrate specificities of several prenyl chain elongating enzymes with respect to 4-methyl-4-pentenyl diphosphate. Biosci. Biotechnol. Biochem. 68, 2070–2075. https://doi.org/10.1271/bbb.68.2070.
Noriega, F.G., 2014. Juvenile hormone biosynthesis in insects: what is new, what do we know, and what questions remain? Int. Sch. Res. Notices 1–16 (2014). https://doi.org/10.1155/2014/967361.
Ohnuma, S.I., Hemmi, H., Koyama, T., Ogura, K., Nishino, T., 1998. Recognition of allylic substrates in Sulfolobus acidocaldarius geranylgeranyl diphosphate synthase: analysis using mutated enzymes and artificial allylic substrates. J. Biochem. 123, 1036–1040. https://doi.org/10.1093/oxfordjournals.jbchem.a022040.
Palframan, M.J., Bandi, K.K., Hamilton, J.G.C., Pattenden, G., 2018. Sobralene, a new sex-aggregation pheromone and likely shunt metabolite of the taxadiene synthase cascade, produced by a member of the sand fly Lutzomyia longipalpis species complex. Tetrahedron Lett. 59, 1921–1923. https://doi.org/10.1016/j.tetlet.2018.03.088.
Prchalová, D., Buček, A., Brabcová, J., Žáček, P., Kindl, J., Valterová, I., Pichová, I., 2016. Regulation of isoprenoid pheromone biosynthesis in bumblebee males. Chembiochem 17, 260–267. https://doi.org/10.1002/cbic.201500415.
Rebholz, Z., Lancaster, J., Larose, H., Khrimian, A., Luck, K., Sparks, M.E., Gendreau, K. L., Shewade, L., Köllner, T.G., Weber, D.C., Gundersen-Rindal, D.E., O'Maille, P., Morozov, A.V., Tholl, D., 2023a. Ancient origin and conserved gene function in terpene pheromone and defense evolution of stink bugs and hemipteran insects. Insect Biochem. Mol. Biol. 152 https://doi.org/10.1016/j.ibmb.2022.103879.
Rebholz, Z., Shewade, L., Kaler, K., Larose, H., Schubot, F., Tholl, D., Morozov, A.V., O'Maille, P.E., 2023b. Emergence of terpene chemical communication in insects: evolutionary recruitment of isoprenoid metabolism. Protein Sci. 32, 1–20. https://doi.org/10.1002/pro.4634.
Reed, B.C., Rilling, H.C., 1975. Crystallization and partial characterization of prenyltransferase from avian liver. Biochemistry 14, 50–54.
Rivera-Perez, C., Nyati, P., Noriega, F.G., 2015. A corpora allata farnesyl diphosphate synthase in mosquitoes displaying a metal ion dependent substrate specificity. Insect Biochem. Mol. Biol. 64, 44–50. https://doi.org/10.1016/j.ibmb.2015.07.010.
Sagami, H., Morita, Y., Ogura, K., 1994. Purification and properties of geranylgeranyl-diphosphate synthase from bovine brain. J. Biol. Chem. 269, 20561–20566. https://doi.org/10.1016/s0021-9258(17)32030-6.
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. https://doi.org/10.1016/j.ibmb.2007.05.012.
Song, X., Qin, Y.G., Zhang, Y.H., Zhou, Y.B., Li, Z.X., 2022. Farnesyl/geranylgeranyl diphosphate synthases regulate the biosynthesis of alarm pheromone in a unique manner in the vetch aphid Megoura viciae. Insect Mol. Biol. 1–11. https://doi.org/10.1111/imb.12826.
Spiegel, C.N., Dias, D.B. do S., Araki, A.S., Hamilton, J.G.C., Brazil, R.P., Jones, T.M., 2016. The Lutzomyia longipalpis complex: a brief natural history of aggregation-sex pheromone communication. Parasites Vectors 9, 1–15. https://doi.org/10.1186/s13071-016-1866-x.
Sun, Z.J., Li, Z.X., 2018. The terpenoid backbone biosynthesis pathway directly affects the biosynthesis of alarm pheromone in the aphid. Insect Mol. Biol. 27, 824–834. https://doi.org/10.1111/imb.12521.
Wang, K., Ohnuma, S. ichi, 1999. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. Trends Biochem. Sci. 24, 445–451. https://doi.org/10.1016/S0968-0004(99)01464-4.
Wang, M., Casey, P.J., 2016. Protein prenylation: unique fats make their mark on biology. Nat. Rev. Mol. Cell Biol. 17, 110–122. https://doi.org/10.1038/nrm.2015.11.
Yang, J., Zhang, Y., 2015. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 43, W174–W181. https://doi.org/10.1093/nar/gkv342.
Yang, Z., Wu, Y., 2019. Improved annotation of Lutzomyia longipalpis genome using bioinformatics analysis. PeerJ (2019), 1–17. https://doi.org/10.7717/peerj.7862.
Zhang, H., Li, Z.-X., 2014. A type-III insect geranylgeranyl diphosphate synthase with a novel catalytic property. Protein Pept. Lett. 21, 615–623.
Zhou, Y., Liu, X., Yang, Z., 2019. Characterization of terpene synthase from tea green leafhopper being involved in formation of geraniol in tea (Camellia sinensis) leaves and potential effect of geraniol on insect-derived endobacteria. Biomolecules 9. https://doi.org/10.3390/biom9120808.