De novo formation of an aggregation pheromone precursor by an isoprenyl diposphosphate synthase-related terpene synthase in the harlequin bug

Jason Lancaster*, Ashot Khrimian*, Sharon Young†, Bryan Lehner*, Katrin Luck*, Anna Wallingford*, Saikat Kumar B. Ghosh†, Philipp Zerbe, Andrew Muchinski†, Paul E. Marek*, Michael E. Sparks*, James G. Tokuhisa*#, Claus Tittiger*, Tobias G. Köllner†, Donald C. Weber†, Dawn E. Gundersen-Rindal†, Thomas P. Kuhr†, and Dorothea Tholl*†#

*Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; †Invasive Insect Biocontrol and Behavior Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD 20705; ‡Department of Biochemistry and Molecular Biology, University of Nevada, Reno, NV 89557; ††Department of Biochemistry, Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany; ‡‡Department of Entomology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; †††Department of Plant Biology, University of California, Davis, CA 95616; and ‡Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

Edited by Rodney B. Croteau, Washington State University, Pullman, WA, and approved July 23, 2018 (received for review January 1, 2018)

Insects use a diverse array of specialized terpene metabolites as pheromones in intraspecific interactions. In contrast to plants and microbes, which employ enzymes called terpene synthases (TPSs) to synthesize terpene metabolites, limited information from few species is available about the enzymatic mechanisms underlying terpene pheromone biosynthesis in insects. Several stink bugs (Hemiptera: Pentatomidae), among them severe agricultural pests, release 15-carbon sesquiterpenes with a bisabolene skeleton as sex or aggregation pheromones. The harlequin bug, Murgantia histrionica, a specialist pest of crucifers, uses two stereoisomers of 10,11-epoxy-1-bisabol-3-ol as a male-released aggregation pheromone called murgantiol. We show that MhTPS (MhIDS-1), an enzyme unrelated to plant and microbial TPSs but with similarity to trans-iso-15-carboxylic acid synthases (IDTs) of the core terpene biosynthetic pathway, catalyzes the formation of (15S,6S,7R)-1,10-bisaboladien-1-ol (sesquipiperitol) as a terpene intermediate in murgantiol biosynthesis. Sesquipiperitol, a so-far-unknown compound in animals, also occurs in plants, indicating convergent evolution in the biosynthesis of this sesquiterpene. RNAi-mediated knockdown of MhTPS mRNA confirmed the role of MhTPS in murgantiol biosynthesis. MhTPS expression is highly specific to tissues lining the cuticle of the abdominal sternites of mature males. Phylogenetic analysis suggests that MhTPS is derived from a trans-ID5 progenitor and diverged from bona fide TPS genes including MhIDS-2, which functions as an (E,E)-farnesyl diphasphate (FPF) synthase. Structure-guided mutagenesis revealed several residues critical to MhTPS and MhFPFS activity. The emergence of an ID3-like protein with TPS activity in M. histrionica demonstrates that de novo terpene biosynthesis evolved in the Hemiptera in an adaptation for intraspecific communication.

Hemiptera | Pentatomidae | harlequin bug | aggregation pheromone | terpene synthase

Terpenes play important roles in mediating chemical interactions of microbes, plants, and animals (1–4). In particular, volatile terpene compounds function as long- and short-distance semiochemicals in organismal interactions (5–10). Insects are well known to release volatile terpenes as interspecific signals in chemical defense or as alarm, aggregation, and sex pheromones in intraspecific communication (11–16). Despite these important functions, little is yet known about the formation of terpene specialized metabolites in insects. In bacteria, fungi, and plants, volatile terpenes with 10-carbon (monoterpenes) and 15-carbon (sesquiterpenes) scaffolds are produced from the isoprenyl diphosphates, geranyl diphosphate (GPP) and farnesyl diphosphate (FPP), respectively, by enzymes called terpene synthases (TPSs) (17–23). As central intermediates in the core terpene metabolism, GPP and cis- or trans-FPP are assembled by trans- or cis-isoprenyl diphasphate synthases (IDTs) in condensations of the 5-carbon unit dimethylallyl diphasphate (DMAPP) with one or two molecules of its isomer isopentenyl diphasphate (IPP). Despite their low sequence similarity, plant- and microbial-type TPS and IDT proteins share a common functionally active a-helical domain and catalyze reactions initiated by diphasphate ionization and carbocation formation (24–26). These structural and mechanistic similarities have led to the hypothesis of a possible ancient re-creation of IDT enzymes from core metabolism to specialized TPS function. Although insects employ a core terpene metabolism that supports juvenile hormone biosynthesis and has been associated with pheromone production (27, 28), to date no plant- or microbial-type TPS genes have been reported from insect genomes. The absence of these genes has led to the general belief that terpenes evolved from a plant- or microbial-type TPS progenitor through duplication events in the insect genome. However, the biosynthetic origin and evolution of these terpenes is mostly unknown. We show that the harlequin bug, Murgantia histrionica, a stink bug pest (Hemiptera) of crucifer crops, produces a terpene aggregation pheromone by an enzyme that is unrelated to microbial and plant terpene synthases. M. histrionica terpene synthase activity is highly sex- and tissue-specific and creates a sesquiterpenic alcohol, so far unknown in animals, as a pheromone precursor. The enzyme evolved from ancestral isoprenyl diphasphate synthases and provides new evidence for de novo biosynthesis of terpenes in hemipteran insects. Knowledge of pheromone biosynthesis in stink bugs may lead to the development of new controls of these pests.

Significance

Many insects release volatile terpenes for chemical communication. However, the biosynthetic origin and evolution of these biochemicals are mostly unknown. We show that the harlequin bug, Murgantia histrionica, a stink bug pest (Hemiptera) of crucifer crops, produces a terpene aggregation pheromone by an enzyme that is unrelated to microbial and plant terpene synthases. M. histrionica terpene synthase activity is highly sex- and tissue-specific and creates a sesquiterpenic alcohol, so far unknown in animals, as a pheromone precursor. The enzyme evolved from ancestral isoprenyl diphasphate synthases and provides new evidence for de novo biosynthesis of terpenes in hemipteran insects. Knowledge of pheromone biosynthesis in stink bugs may lead to the development of new controls of these pests.
notion that insects are unable to synthesize terpene specialized metabolites de novo by the use of TPS enzymes and instead largely depend on the sequestration of terpene precursors from their host plants (15). However, studies on the biosynthesis of aggregation pheromones in Coleoptera (beetles) have shown that trans-IDS-like enzymes are able to convert GPP or FPP to terpene pheromones or their respective precursors. Gilg et al. (29, 30) demonstrated that the bark beetle Ips pini employs an IDS-like enzyme to produce the monoterpene myrcene from GPP as precursor of the aggregation pheromone ipsdienol. Recently, a similar finding has been reported from males of the flea beetle Phyllothea striolata, which synthesize the cyclic sesquiterpene aggregation pheromone (6R,7S)-bimachala-9,11-diene from (Z,E)-FPP by an IDS-type enzyme (31). From nine different IDS-type transcripts in the P. striolata transcriptome, two were found to encode bona fide trans- or cis-IDS enzymes, while four transcripts encode TPSs, suggesting an evolutionary origin of these enzymes from IDS progenitors. Whether insects other than beetles, especially those of earlier evolutionary origin, biosynthesize volatile terpenes de novo is unknown.

Here, we show that within the Hemiptera, stink bugs (Pentatomidae) use IDS-like proteins in pheromone biosynthesis. Several pentatomids release sesquiterpene aggregation/sex pheromones with a bisabolene carbon skeleton (32–39). Among these, the harlequin bug Murtanga histrionica, a crucifer specialist, produces a mixture of (3S,6S,7R,10S) and (3S,6S,7R,10R) stereoisomers of 10,11-epoxy-1-bisabolen-3-ol as a male-released aggregation pheromone dubbed murgantiol (40–42) (Fig. 1). We demonstrate that an enzyme with homology to IDS proteins (MhTPS) converts (E,E)-FPP 2 to sesquiperitol 3 as the presumed stereospecific alcohol precursor of murgantiol (Fig. 1), while a second trans-IDS protein (MhFPPS) catalyzes the formation of the MhTPS substrate (E,E)-FPP from IPP and DMAPP. MhTPS is transcribed at high levels in males, with a predominant localization in the subcuticular tissue of the abdominal sternites. A significant role of MhTPS in pheromone biosynthesis was confirmed by RNAi-mediated knockdown of MhTPS mRNA in M. histrionica males leading to reduced emission of murgantiol. Phylogenetic comparison of the M. histrionica enzymes with other insect IDS proteins suggests that in the Hemiptera proteins with TPS activity evolved from a trans-IDS progenitor. Sequence- and structure-guided mutagenesis identified several residues with critical function in MhTPS and MhFPPS activity. Together, our study suggests that, in comparison with

Fig. 1. Functional characterization of MhIDS-1 (MhTPS) from M. histrionica. Recombinant MhIDS-1 protein was expressed in E. coli and Sf9 cells and partially purified by affinity chromatography. Proteins were incubated with (E,E)-FPP 2 in the presence of Mg2+ and products were analyzed by GC-MS. (A) GC-MS chromatograms of enzyme products. Sf9 control cells express a housefly cytochrome P450 reductase. *, nonenzyme product; 3, (15S,6S,7R)-sesquiperitol; cu, γ-curcumene; far, (2E,6E)-farnesol; sesq, β-sesquiphellandrene; zi, α-zingiberene. (B) Mass spectra of enzymatic products with (15S,6S,7R)-sesquiperitol standard 3. (C) Formation of sesquiperitol 3 by M. histrionica TPS activity (boxed). A putative single or two-step pathway to murgantiol 1 is shown involving isomerization and epoxidation reactions. EV, empty vector; MdCPR, Musca domestica cytochrome P450 reductase.
microbial- and plant-type TPSs, insect TPS enzymes evolved more recently from IDS proteins and this event has occurred in multiple insect lineages including hemipteran insects.

**Results**

**Identification and Functional Characterization of IDS-Like Genes in M. histrionica.** We hypothesized that in the murgantiol biosynthetic pathway of *M. histrionica* an enzyme with terpene synthase activity synthesizes a bisabolene-type hydrocarbon or alcohol terpene precursor, which presumably undergoes further modification(s) including epoxidation to form the pheromone end products (Fig. 1). To identify genes involved in the formation of the murgantiol pheromone precursor, we queried a transcriptome dataset from different sexes and developmental stages of *M. histrionica* (43) with plant and microbial TPS sequences and *trans-IDS*-type sequences of *I. pini* and *F. pisi* (AAX55631.1) and GPPS/TPS (AAX55632.1). *Bombix mori* (silk moth) FPPPS1 (NP_001036889.1) and FPPPS2 (NP_001093301.1), and *Drosophila melanogaster* (fruit fly) FPPPS (NP_477380). When no genes with sequence similarity to plant or microbial TPS sequences could be identified, two *trans-IDS*-like sequences (MhIDS-1 and MhIDS-2) annotated as FPP synthase (FPPPS)-like genes were retrieved. According to the transcriptome results by Sparks et al. (43), MhIDS-1 (GEC01420512.1; MG662378.1) mRNA levels are low in mature females but show a ∼13-fold higher accumulation in mature males, while MhIDS-2 (GEC01414919.1; MG662379.1) transcript levels are equal in both sexes. cDNAs of both genes were amplified from RNA extracted from mature male bugs. MhIDS-1 encodes a 385-aa protein of 44.30 kDa while the MhIDS-2 protein contains 405 aa residues and has a size of 46.36 kDa (*SI Appendix*, Fig. S1). MhIDS-1 shares 22–24% sequence identity with insect FPPPSs from *I. pini*, *D. mela- nogaster*, and the black bean aphid *Aphis fabae* (AAY33488.2), whereas comparatively higher sequence identities with these proteins (40–45%) were found for MhIDS-2.

To determine the biochemical function of the detected IDS-like genes, cDNAs encoding full-length proteins were cloned in the bacterial protein expression vector pEXP-5, generating an N-terminal histidine-tag fusion. When tested for sesqui-terpene synthase activity with (*E*,*E*)-FPP 2 as a substrate, the partially purified recombinant MhIDS-1 protein produced a terpene alcohol as its major product (Fig. 1 A, *SI Appendix*, Fig. S4). Using GC-MS, we identified the alcohol product as sesquipiperitol 3, a sesqui-terpene alcohol with a bisabolane skeleton, which is found in different plant species (44–46). The identification of sesquipiperitol was performed by comparisons of mass spectra and retention indices and further verified by chemical correlations (Fig. 1 A and B and *SI Appendix, SI Results, SI Materials and Methods*, and Fig. S2). Sesquipiperitol was also the main product in assays upon cleavage of the N-terminal histidine tag (*SI Appendix, Fig. S3A*). In addition to sesquipiperitol, small and varying amounts of the sesqui-terpene olefins, γ-curcumene, zingiberene, and β-sesquiphellandrene were detected (Fig. L4). Hot sample injection contributed to the formation of the olefin products by thermal dehydration of sesquipiperitol, as could be shown in contrast to cool-on-column injection (*SI Appendix, Fig. S3B*). Sesquipiperitol was also produced, although at lower levels, from (*Z*,*E*)-FPP, but almost no enzymatic activity was found with (*Z*,*Z*)-FPP as the substrate (*SI Appendix, Fig. S3C*). The recombi- nant enzyme did further convert GPP to several monoterpens (*SI Appendix, Fig. S3C*). However, when incubated with IPP and DMAPP, no formation of terpene products was observed, indicating that MhIDS-1 was unable to synthesize prenyl diphosphates for subsequent conversion into terpene products (*SI Appendix, Fig. S3C*). Accordingly, formation of FPP by MhIDS-1 was not observed. Because of its TP activity and lack of IDS activity, we designate MhIDS-1 hereinafter as MhTPS. We further tested the activity of MhTPS produced in insect *Sp9* cells (Fig. L4). Recombinant MhTPS protein expressed without a histidine tag under these conditions generated the same enzymatic products upon incubation with (*E*,*E*)-FPP as those produced by the bacterially expressed enzyme (Fig. L4). An alignment of the MhTPS amino acid sequence with those of *I. pini* and *F. pisi* TPS proteins suggested the presence of a putative N-terminal targeting peptide, although the Rxx5 motif indicative of a mitochondrial targeting sequence is absent from the MhTPS protein (*SI Appendix*, Fig. S1). Truncation of MhTPS (M4–R56) resulted in the loss of enzymatic activity.

Kinetic analysis of MhTPS with (*E*,*E*)-FPP as the substrate revealed an apparent *Kₘ* value of 4.0 ± 0.7 μM and a *Vₘ₅₆* of 675.3 ± 53.7 pkat/mg. The *kₗ₅₆* value was 0.03 ± 0.003 s⁻¹ and *kₗ₅₆/Kₘ₆* was 7.5 ± 0.5 × 10⁻⁶ s⁻¹M⁻¹. *Kₘ₆*, *kₗ₅₆*, and *kₗ₅₆/Kₘ₆* values of MhTPS1 were similar to those of plant sesqui-TPS enzymes such as (*E*)-β-caryophyllene synthase from *Artemisia annua* (47).

Recombinant MhTPS1 did not produce zingiberenol 4 as a possible precursor of murgantiol. We tested whether changes in cofactor type and concentration or modifications of pH conditions would alter the enzymatic product profile and activity. No change in product specificity was observed when Mg²⁺ was substituted with Co²⁺, although this metal ion has been found to modify product specificity of a GPP/FPP synthase in the leaf beetle *Phaedon brassicae* (*P*IDS-1) (48). Activity increased by approximately twofold between 0.1 and 10 mM Mg²⁺, while the opposite was the case for Co²⁺ (*SI Appendix, Fig. S3D*). Activity was highest at pH 7 (*SI Appendix, Fig. S3E*) and no change in product outcome was found under lower or higher pH conditions.

In contrast to MhTPS1, partially purified recombinant MhIDS-2 protein did not show any TP activity when assayed with different isomers of FPP or GPP as substrates. Instead, MhIDS-2 produced (*E*,*E*)-FPP from IPP and DMAPP, indicating that this protein functions as a trans-IDS (*SI Appendix, Fig. S4*). The enzyme was unable to synthesize any other isomer of FPP, which suggests that (*E*,*E*)-FPP is the main isomeric form produced by *M. histrionica* (*SI Appendix, Fig. S4*). Removal of a putative transit peptide (M1–559) (*SI Appendix, Fig. S1*) led to a substantially higher production of (*E*,*E*)-FPP by the truncated MhIDS-2 protein (*SI Appendix, Fig. S4*). Because of its FPPS activity, we designate MhIDS-2 hereinafter as MhFPPS.

**Absolute Configuration of Sesquipiperitol.** To further support the role of sesquipiperitol as a precursor of murgantiol, we determined the stereospecific configuration of sesquipiperitol at C-6 and C-7, which we predicted to be the same as that of murgantiol (*SI Appendix, SI Results, SI Materials and Methods*, and Figs. S2 and S5). Oxidation of enzymatically produced sesquipiperitol to sesquipiperitene 5 concluded a relative 65,7R or 6R,7S configuration (*SI Appendix, Fig. S2*). Further conversion of sesquipiperitol to bisabolane 7 determined the configuration at C-7 to be (R) (*SI Appendix, Fig. S5A*). This result unambiguously confirmed a 65,7R configuration of sesquipiperitol 3, which is identical to the C-6, C-7 configuration of murgantiol 1. Chiral GC analysis and 2D NMR recordings determined an (S) configuration at C-1 (*SI Appendix, SI Results, SI Materials and Methods*, and Fig. S5 B and C).

**Sex- and Tissue-Specific Expression of MhTPS.** We compared transcript abundance of MhTPS and MhFPPS between different sexes, developmental stages, and tissues by RT-PCR and qRT-PCR analyses. In agreement with mRNA levels determined by transcriptome analysis (43), MhTPS transcript levels were significantly higher (∼37-fold) in mature males than in mature and immature females (Fig. 2 A and *SI Appendix, Fig. S6A*). MhTPS transcript abundance in mature males was also significantly higher than that of nymphs and immature males, which do not emit murgantiol (Fig. 2A). As expected from transcriptome data (43), MhFPPS transcript levels were comparatively higher than those of MhTPS.
in mature females (SI Appendix, Fig. S6A). Tissue-specific transcripts of \textit{MhTPS} localized to the tissue lining the cuticle of the abdominal sternites of mature males, while 200- to 2,000-fold lower transcript levels of \textit{MhTPS} were observed in the thorax, fat body, and midgut (Fig. 2B and SI Appendix, Fig. S6B).

**Sesquipiperitol Synthase Activity in \textit{M. histrionica} Protein Extracts.** To further confirm that the enzymatic activity for the formation of sesquipiperitol is present in protein lysates of \textit{M. histrionica}, we extracted protein from the cuticle-specific tissue of mature male bugs and incubated the lysate with (\textit{E,E})-FPP. GC-MS analysis of hexane extracts obtained from the aqueous phase of the assay demonstrated the specific formation of sesquipiperitol in the presence of (\textit{E,E})-FPP (Fig. 3A). By contrast, protein extracts from cuticle-associated tissue of mature females did not produce sesquipiperitol, confirming the male-specific biosynthesis of this compound (Fig. 3A). Sesquipiperitol was not synthesized in extracts of the male head and thorax, fat body, or gut tissue when incubated with (\textit{E,E})-FPP, which further supports the tissue specificity of this enzymatic reaction (Fig. 3B). We did not find further conversion of sesquipiperitol to the putative downstream intermediate zingiberenol or to murgantiol or other products, suggesting that possible downstream enzymatic activities were not supported or potentially inhibited under our selected extraction and assay conditions. Alternatively, compound conversion might have been hampered by the loss of tissue integrity and compartmentalization.

**Verification of in Vivo \textit{MhTPS} Function in Murgantiol Biosynthesis.** To verify the in vivo role of \textit{MhTPS} in murgantiol biosynthesis, we injected males 3–5 d posteclosion with \textit{MhTPS}-derived sequence-specific dsRNA. Transcript abundance was significantly reduced (∼35- to 40-fold) 13–15 d postinjection compared with males injected with \textit{lacZ} or Ringer’s solution, or non-injected males (\textit{P} < 0.0001) (Fig. 4A). Pheromone analysis of males at the same time postinjection showed significantly lower emission (approximately sevenfold) of murgantiol in \textit{MhTPS} dsRNA males than control males (Fig. 4B), confirming a substantial role of \textit{MhTPS} in murgantiol biosynthesis.

**Phylogenetic Analysis of \textit{M. histrionica} TPS and FPPS and Determination of Residues with Function in TPS and IDS Catalytic Activity.** A phylogenetic analysis based on maximum likelihood was performed to assess the evolutionary relationship of \textit{M. histrionica} TPS and FPPS with other insect IDS and TPS enzymes (Fig. 5). The dataset included \textit{trans-IDS} proteins from Coleoptera, Lepidoptera, and Hemiptera with known GPPS/FPPS or FPPS activities, the GPPS/TPS from \textit{I. pini}, and the recently characterized IDS and TPS enzymes from \textit{P. striolata}. To compare the relationship of the \textit{M. histrionica} enzymes with similar enzymes in the Pentatomidae, we retrieved IDS-like sequences from the brown marmorated stink bug \textit{Halyomorpha halys} based on publicly available transcriptome datasets of this insect (49). For \textit{H. halys} two IDS-like sequences were identified, of which the IDS-1 sequence clusters with \textit{M. histrionica} TPS (MG662378.1) (38.3% sequence identity), suggesting that \textit{H. halys} IDS-1 might be a functional TPS enzyme. \textit{M. histrionica} TPS and \textit{H. halys} IDS-1 form a clade separated from \textit{I. pini} GPPS/TPS, a cis-IDS (FPPPS) and the TPS clade of \textit{P. striolata} (Fig. 5). \textit{M. histrionica} FPPS (MG662379.1) and the more closely related \textit{H. halys} IDS-2 protein are positioned in a larger clade of insect proteins with bona fide \textit{trans}-FPPS or GPPS/FPPS activity from Coleoptera, Lepidoptera, and Hemiptera (Fig. 5). A broader phylogenetic analysis including insect GPPPSs and plant \textit{trans-IDS} proteins supports an evolutionary origin of the pentatomid TPSs together with the coleopteran TPSs from a \textit{trans-IDS} progenitor that gave rise to proteins with GPPS/FPPS or TPS activities (SI Appendix, Fig. S7).

![Fig. 2.](https://example.com/fig2.png) **Relative transcript abundance of \textit{MhTPS} in \textit{M. histrionica} determined by qRT-PCR.** (A) \textit{MhTPS} transcript abundance at different developmental stages and in different sexes. Young is 3-d adult; mature is 15-d adult. (\textit{n} = 3, ±SD). (B) \textit{MhTPS} transcript abundance in different tissues of adult males. Ab. Cuticle, tissue lining the abdominal cuticle; Ab. ST, Abdominal soft tissue minus midgut. (\textit{n} = 3, ±SD). Gene expression was normalized against 18S rRNA. Transcript abundance is shown relative to that in nymphs (\textit{A}) or the male head tissue (\textit{B}) (set to 1). Significance was determined using one-way ANOVA and means grouped by Tukey’s honestly significant difference test.
Sequence comparisons between insect IDS and IDS-derived TPS enzymes show that both types of proteins contain a conserved first and second aspartate-rich motif (FARM and SARM, respectively) (SI Appendix, Fig. S1). These motifs are required for the coordinated binding of Mg$^{2+}$ ions with the allylic substrate to initiate catalysis through carbocation formation and rearrangement (24–26). While residue substitutions occur in the SARM of TPSs from I. pini and P. striolata, no such substitutions are present in the TPSs of stink bugs (SI Appendix, Fig. S1). Notably, in comparison with bona fide trans-IDS proteins, the TPS proteins substitute aromatic with nonaromatic amino acid residues at positions 4 and 5 upstream of the FARM (SI Appendix, Figs. S1, S8). To determine whether these substitutions affect the orientation of (E,E)-FPP at the active site of the enzyme, we positioned (E,E)-FPP in the active site of crystallized Gallas gallus FPPS and an M. histrionica TPS homology model. This comparative positioning via ligand docking indicated that the amino acid changes upstream of the FARM (GgFPPS F112, F113 – MhTPS M118, S119) likely cause a different orientation of the prenyl side chain of the FPP substrate (Fig. 6A). According to the dynamic mutagenesis resulted in the loss of TPS activity of the recombinant MhTPS protein (Fig. 6B). By contrast, the reciprocal substitutions F95M and F96S in a truncated MhFPPS protein did not abolish IDS activity but converted the enzyme into a 20-carbon geranylgeranyl diphosphate (GGPP) synthase (SI Appendix, Fig. S8B). In comparison with (E,E)-FPP, docking of DMAPP in both MhTPS and GgFPPS models did not result in different positions of this allylic diphosphate (SI Appendix, Fig. S9A). However, DMAPP did not have substantial inhibitory effects on the MhTPS-catalyzed reaction of (E,E)-FPP to sesquiperipetal, indicating a limited affinity of the enzyme for this allylic diphosphate (SI Appendix, Fig. S9B).

To further define determinants of TPS or IDS catalytic specificity, we identified amino acids that are exclusively conserved in TPS or IDS sequences (SI Appendix, Fig. S1). Most of these residues were found on helices and loops outside of the predicted active side cavity of the MhTPS model, making their impact on catalysis less predictable. Among residues of further interest were Y172DAW340 in MhTPS because these amino acids replace a highly conserved KKKRK motif in MhFPPS and other IDS proteins (SI Appendix, Fig. S1, box M2) with W80 positioned at the bottom of the active-site cavity (SI Appendix, Fig. S9C). Reciprocal substitution of these residues in MhTPS and MhFPPS abolished both TPS and IDS activities, indicating their critical functions in both enzymes (SI Appendix, Fig. S8 A and B). Loss of activity was also observed when residues KC135KG137 in MhTPS (K135 is a conserved amino acid in insect TPSs) were substituted with the corresponding RC170PC172 sequence of full-length MhFPPS, which contains a Cys residue highly conserved in insect IDS proteins (SI Appendix, Fig. S1, box M3). The same situation was found for the reverse mutation in MhFPPS (SI Appendix, Figs. S1, S8 A and B, and S9D), again indicating critical functions of these residues in both proteins.

**Discussion**

*M. histrionica* TPS Functions as a Sesquiterpene Alcohol Synthase. We found that the TPS activity associated with the recombinant MhTPS enzyme and with crude lysates of male *M. histrionica* stereospecifically convert (E,E)-FPP to the (13,65,7R) isomer of sesquieryperitol 3. The synthesis of sesquieryperitol most likely proceeds by formation of a bisabolyl cation followed by a hydride shift and subsequent quenching of the carbocation by water (SI Appendix, Fig. S10). The 65,7R configuration of sesquieryperitol corroborates its function as an intermediate in the formation of

![Fig. 4. Effects of RNA interference on *M. histrionica* MhTPS expression and murgantol emission. (A) MhTPS transcript abundance in adult males 12 d postinjection normalized to 18S RNA. LacZ was used as a negative control (n = 3, ±SEM). (B) Amount of murgantol detected in headspace collections 10–12 d postinjection (n = 9, ±SEM). Bars in each figure with the same letter are not different according to a generalized log-linear model (α = 0.05). (A) \( \chi^2 = 63.13, P < 0.0001 \). (B) \( \chi^2 = 4883.3, P < 0.0001 \).](image)

![Fig. 5. Majority-rule phylogram inferred from maximum-likelihood analysis of FPPS and TPS enzymes of *M. histrionica* (bold) with related IDS proteins of *H. halys*, TPS and IDS proteins of *P. striolata*, GPPS/TPS of *I. pini*, and other insect trans-(GPPS)/FPPS proteins. The maximum likelihood method was based on the Le and Gascuel (59) (LG G+1) model. Bootstrap values (n = 1,000 replicates) are shown next to each node. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Proteins with known or putative TPS activity are highlighted by the gray box. The tree was rooted using a GPPS from *D. melanogaster*. Full species names are listed in SI Appendix, Fig. S7.](image)
were substituted with KKxR, a highly conserved gene and its corresponding enzyme.

The sesquiterpenes α-curcumene, zingiberene, and β-sesquiphellandrene were reported in animals. The occurrence of sesquipiperitol as an enzymatic product in the red-shouldered backpack stink bug, H. halys (33, 50), suggests that adaptations of terpene-specialized metabolism for intraspecific communication have occurred in multiple lineages throughout insect evolution. To date, IDS-type TPSs with similar functions have only been characterized in the coleopterans I. pini and P. striolata (29–31). Phylogenetic analysis indicates that pentatomid and coleopteran TPSs are derived from a trans-IDS progenitor that diverged early from true trans-IDS enzymes. The separate clustering of pentatomid TPSs from P. striolata and I. pini TPSs (Fig. 5 and SI Appendix, Fig. S7) suggests that proteins with TPS activity might have emerged independently in these and other insect taxa. However, phylogenetic comparisons with additional TPSs from other insect lineages will be necessary to further support this hypothesis.

In our approach to identify residues that discriminate between TPS or IDS catalytic activity, we found that MhTPS lost activity when nonaromatic amino acids at positions 4 and 5 upstream of the FAR motif were substituted with aromatic amino acids of equivalent position in insect IDSs. This effect was likely caused by a modified position of the FPP substrate. Molecular docking of (E,E)-FPP in the active-site cavity of an MhTPS homology model showed an extension of the FPP prenyl side chain into the cavity (Fig. 6A), which appears to be critical for facilitating a subsequent cyclization and formation of the terpene product. For IDS enzymes, substitutions of aromatic and nonaromatic amino acids upstream of the FARM determine the chain length of the enzymatic product: nonaromatic residues facilitate the synthesis of products with extended prenyl chains in long-chain trans-IDS enzymes (≥C20) (54). Indeed, substitution of the two Phe residues with Met and Ser upstream of the FARM in MhFPPS resulted in a change from FPPS to GGPPS activity (SI Appendix, Fig. S8B). This change in IDS product specificity correlates with the presence of nonaromatic residues in other insect GGPP synthases (55).

The diagram shows the functional analysis of amino acid residues upstream of the first aspartate-rich motif of MhTPS. (A) Position of (E,E)-FPP in the active site of G. gallus FPPS (green) and the M. histrionica TPS homology model (blue). Residues used for site-directed mutagenesis are illustrated. (B) TPS activity of the MhTPS wild type and protein variant with residue substitutions at positions 4 and 5 upstream of the FAR motif. TPS activity was also observed when residues S2–DAW were substituted with KKxR, a highly conserved motif in insect IDS proteins. The charged side chain of the introduced Arg residue facing toward the hydrocarbon tail of (E,E)-FPP in this MhTPS variant may prevent further cyclization (SI Appendix, Fig. S9C). The KKxR motif in insect IDSs is equivalent to the KRLR motif in the FPPS of Escherichia coli. The basic residues of this motif interact with the diphosphate moiety of IPP (56). Therefore, the loss of MhTPS activity by reversely substituting the KKxR motif with the SDAW sequence is most likely due to improper binding of the IPP substrate.

Reciprocal substitutions of the highly conserved Lys and Cys residues downstream of the FARM of TPS or IDS proteins, respectively, also caused a loss of activity in both mutants. Although, according to our model, these residues are positioned on a loop without immediate proximity to the docked substrate (SI Appendix, Fig. S9D), they may be involved in substrate binding and/or in the closure of the active site upon substrate binding, as has been shown, for example, for loop residues on the β-domain of TPS enzymes (57).

In summary, we were able to identify residues with critical function in TPS and IDS activity. Further structural analysis, ideally via protein crystallization and combinatorial mutations paired with the identification of epistatic residue networks, will be necessary to fully identify the residues controlling the transition from insect IDS to TPS enzyme activity. This evolutionary transition likely required a combination of residue substitutions to change substrate affinities and specificities for

**Fig. 6.** Functional analysis of amino acid residues upstream of the first aspartate-rich motif of MhTPS. (A) Position of (E,E)-FPP in the active site of G. gallus FPPS (green) and the M. histrionica TPS homology model (blue). Residues used for site-directed mutagenesis are illustrated. (B) TPS activity of the MhTPS wild type and protein variant with residue substitutions at positions 4 and 5 upstream of the FAR motif. MhTPS wt, MhTPS M1, MhTPS M118F/S119F, standard, drug, and TPS product. Relative abundances of the MhTPS variants were measured by liquid chromatography–tandem mass spectrometry and are converted to terpene trans abundances (TTA) (15, 29, 77, 80).

**Table 1.** Summary of protein variants, their substitutions, and the effect on TPS activity.

<table>
<thead>
<tr>
<th>Protein Variant</th>
<th>Substitutions</th>
<th>Effect on TPS Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MhTPS wt</td>
<td>none</td>
<td>wild type activity</td>
</tr>
<tr>
<td>MhTPS M1</td>
<td>M118F/S119F</td>
<td>loss of activity</td>
</tr>
<tr>
<td>MhTPS M118F/S119F</td>
<td>F12/S112</td>
<td>loss of activity</td>
</tr>
</tbody>
</table>

**Fig. 7.** Loss of MhTPS activity was also observed when residues S2–DAW were substituted with KKxR, a highly conserved motif in insect IDS proteins. The charged side chain of the introduced Arg residue facing toward the hydrocarbon tail of (E,E)-FPP in this MhTPS variant may prevent further cyclization (SI Appendix, Fig. S9C). The KKxR motif in insect IDSs is equivalent to the KRLR motif in the FPPS of Escherichia coli. The basic residues of this motif interact with the diphosphate moiety of IPP (56). Therefore, the loss of MhTPS activity by reversely substituting the KKxR motif with the SDAW sequence is most likely due to improper binding of the IPP substrate.

**Fig. 8.** Reciprocal substitutions of the highly conserved Lys and Cys residues downstream of the FARM of TPS or IDS proteins, respectively, also caused a loss of activity in both mutants. Although, according to our model, these residues are positioned on a loop without immediate proximity to the docked substrate (SI Appendix, Fig. S9D), they may be involved in substrate binding and/or in the closure of the active site upon substrate binding, as has been shown, for example, for loop residues on the β-domain of TPS enzymes (57).
DMAPP and IPP and gain a cyclization function following the binding of (E,E)-FPP as a single allylic substrate.

**Materials and Methods**

Two putative trans-IDS-like genes (MhIDS-1, MhIDS-2) identified in the transcriptome of *M. histrionica* were amplified as full-length cDNAs from RNA extracted from adult males and inserted in the prokaryotic expression vector pEXPS-N/TOPO, generating an N-terminal histidine tag. A truncated version of MhIDS-2 lacking a putative N-terminal transit peptide was cloned in the same vector, and a truncated version of MhIDS-1 was synthesized and cloned in the pET19b expression vector. For expression in insect cells, the full-length MhIDS-1 cDNA was cloned without an N-terminal His-tag into the BaculoDirect vector. Recombinant partially purified MhIDS-1 and MhIDS-2 proteins expressed in *E. coli* and lysates of *Syph* cells expressing MhIDS-1 were tested for TPS and IDS activities and the reaction products were analyzed by GC-MS and liquid chromatography tandem MS, respectively. Identification and determination of the stereospecific configuration of the main enzymatic product of MhIDS-1 (MhTPS), sesquiperitol, were performed by spectroscopic and chemical transformations in combination with chiral GC and NMR analysis. Kinetic properties of MhTPS were examined in assays using [1-3H]-(E,E)-FPP as substrate. Developmental, sex- and tissue-specific expression of the MhTPS transcript or MhTPS activity were determined by qRT-PCR and TPS activity assays of crude lyophilized male headspace. MhTPS involvement in pheromone production was verified by injecting young males of *M. histrionica* with dsRNA of MhTPS or laez (control) followed by qRT-PCR and pheromone analysis in the headspace. Phylogenetic analysis was performed using maximum likelihood. MhTPS and MhFPPS protein variants were generated by site-directed mutagenesis, expressed in *E. coli*, and analyzed for TPS and FPPS activity. Further details on organisms, experimental and analytical procedures, data analysis, phylogenetic analysis, protein homology modeling and substrate docking are provided in SI Appendix, SI Materials and Methods.

**Acknowledgments** We thank Megan V. Herlihy and Jeremy K. Turner for assisting with harlequin bug rearing and volatile collection and Filadello Cicchino for assisting with data analyses. This work was supported by the Department of Agriculture National Institute of Food and Agriculture Grant 2016-67013-24759 (to D.T., D.E.G.-R., A.K., T.P.K., C.T., and D.C.W.) and the Virginia Polytechnic Institute and State University Department of Biochemistry and Biological Sciences and the Translational Plant Sciences Program.