Revised submission to *FEBS Journal*

**Title**
A *cis*-prenyltransferase from *Methanosarcina acetivorans* catalyzes both head-to-tail and non-head-to-tail prenyl condensation

**Authors**
Takuya Ogawa, Koh-ichi Emi, Kazushi Koga, Tohru Yoshimura, and Hisashi Hemmi

**Affiliation**
Department of Applied Molecular Bioscience, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 460-8601, Japan

**Corresponding author**
Hisashi Hemmi
Address: Department of Applied Molecular Bioscience, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan
Phone: +81-52-789-4134
Fax: +81-52-789-4120
E-mail: hhemmi@agr.nagoya-u.ac.jp

**Running title**
Irregular side reaction of a *cis*-prenyltransferase
Abbreviations

DMAPP, dimethylallyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; GPP, geranyl diphosphate; HMBC, hetero-nuclear multiple-bond connectivity; HSQC, hetero-nuclear single quantum coherence; IPP, isopentenyl diphosphate; MOPS, 3-morpholinopropanesulfonic acid

Keywords

cis-prenyltransferase; archaea; glycosyl carrier lipid; isoprenoid; diterpenoid
35 Abstract

36 Cis-prenyltransferase usually consecutively catalyzes the head-to-tail condensation
37 reactions of isopentenyl diphosphate to allylic prenyl diphosphate in the production of
38 (E,Z-mixed) polyprenyl diphosphate, which is the precursor of glycosyl carrier lipids.
39 Some recently discovered homologs of the enzyme, however, catalyze the
40 non-head-to-tail condensation reactions between allylic prenyl diphosphates. In the
41 present study, we characterize a cis-prenyltransferase homolog from a methanogenic
42 archaeon, Methanosarcina acetivorans, to obtain information on the biosynthesis of the
43 glycosyl carrier lipids within it. This enzyme catalyzes both head-to-tail and
44 non-head-to-tail condensation reactions. The kinetic analysis shows that the main
45 reaction of the enzyme is consecutive head-to-tail prenyl condensation reactions
46 yielding polyprenyl diphosphates, while the chain lengths of the major products seem
47 shorter than expected for the precursor of glycosyl carrier lipids. On the other hand, a
48 subsidiary reaction of the enzyme, i.e., non-head-to-tail condensation between
49 dimethylallyl diphosphate and farnesyl diphosphate, gives a novel diterpenoid
50 compound, geranyllavandulyl diphosphate.
Introduction

Cis-prenyltransferases [1] catalyze consecutive head-to-tail condensation reactions of isopentenyl diphosphate (IPP) to allylic prenyl diphosphate to produce the precursor of glycosyl carrier lipids such as dolichol for eukaryotes and undecaprenol for bacteria, which are required for N-linked protein glycosylation and peptide glycan biosynthesis, respectively (Fig. 1). The intermediates and the final product of the reaction, i.e., (poly)prenyl diphosphates, are the oligomers/polymers of C5 prenyl moieties, and the prenyl moieties that are newly condensed by the action of the enzyme have a double bond with Z-configuration. The primer substrate of the enzyme is usually (all-E) prenyl diphosphate such as (E,E)-farnesyl diphosphate (FPP), in which case the intermediate of the condensation reaction should be (E,Z-mixed) prenyl diphosphate. The intermediate is used as an allylic substrate for the next condensation step with IPP to provide a longer intermediate, and the condensation is repeated until the carbon chain of the product finally reaches the required length, which can be anticipated from the chain lengths of the glycosyl carrier lipids, e.g., C_{85-105} for human and C_{55} for general bacteria.

Recently, a few homologs of cis-prenyltransferase were shown to catalyze the non-head-to-tail condensation between allylic prenyl diphosphates (Fig. 1). For example, lavandulyl diphosphate synthase from lavender uses two dimethylallyl diphosphate (DMAPP) molecules to yield a non-allylic product, lavandulyl diphosphate, which is a precursor of a monoterpene alcohol, lavandulol [2]. Cyclolavandulyl diphosphate synthase from Streptomyces sp. CL190 catalyzes both the same condensation between two DMAPPs and an additional cyclization of the intermediate lavandulyl diphosphate to give a non-standard allylic prenyl diphosphate, which is used for the modification of a secondary metabolite to form lavanducyanin [3]. Another non-standard allylic prenyl
diphosphate, isosesquilavandulyl diphosphate, is synthesized from geranyl diphosphate (GPP) and DMAPP by the action of a cis-prenyltransferase homolog from *Streptomyces* sp. strain CNH-189 [4]. That compound is used for the biosynthesis of meroterpenoid antibiotics, merochlorins. The substrate specificities of these enzymes are high: They do not catalyze a standard head-to-tail condensation between allylic prenyl diphosphate and IPP.

With the exception of a few previous characterization studies on the enzymes from *Sulfolobus acidocaldarius* [5], *Aeropyrum pernix* [6], and *Thermococcus kodakarensis* [7], which yield polyprenyl diphosphates with carbon chains up to C\textsubscript{55-65}, archaeal cis-prenyltransferases have not been sufficiently elucidated. The structures of glycosyl carrier lipids also remain unclear in most archaeal species. The detailed structures have been determined only for the lipids from *Halobacterium halobium* [8], *Halofex volcanii* [9, 10], *S. acidocaldarius* [11], and *Pyrococcus furiosus* [12] (Fig. 1). The chain lengths of the lipids are C\textsubscript{45-60}, which approximates their bacterial C\textsubscript{55} counterpart. The archaeal lipids, however, are classified as dolichols because their $\alpha$-terminal prenyl moiety is reduced. Interestingly, some of them also have reduced prenyl moieties at their $\omega$-termini [9-12]. On the other hand, older reports concerning glycosyl carrier lipids from methanogenic archaea have suggested that both dolichol and polyprenol are used in some species [13-15]. A methanogenic archaeon, *Methanosarcina acetivorans*, possesses three *cis*-prenyltransferase homologs, which is exceptional. This situation motivated us to characterize the homologs from *M. acetivorans* in the present study, and in the process we encountered a *cis*-prenyltransferase that catalyzes both head-to-tail and non-head-to-tail prenyl condensations.
Results

By searching databases such as KEGG (http://www.kegg.jp), we found three *cis*-prenyltransferase homologs, MA1831, MA3723 and MA4402, encoded in the genome sequence of *M. acetivorans*. Among them, MA3723 is the closest relative of the previously characterized archaeal *cis*-prenyltransferases from *S. acidocaldarius* and *T. kodakarensis*, while MA1831 is closest to the enzyme from *A. pernix*. A phylogenetic analysis of *cis*-prenyltransferase homologs led to the construction of a phylogenetic tree (Fig. 2) with three distinct (but not always monophyletic) clusters of the hypothetical proteins from Euryarchaeota. A cluster colored in yellow, to which MA3723 belongs, contains one homolog from each species of Euryarchaeota including *T. kodakarensis*, suggesting that it is composed of the orthologs of the essential enzyme required for glycosyl carrier lipid biosynthesis. Another cluster colored in light blue contains MA4402 and its supposed orthologs found in the genomes of some limited Euryarchaeota species. The third cluster forming an isolated, beige-colored branch is composed of the orthologs of MA1831, which exist only in *Methanosarcina* spp. The *cis*-prenyltransferases from *Saccharomyces cerevisiae*, *Escherichia coli*, and the two species of Crenarchaeota, *S. acidocaldarius* and *A. pernix*, independently take deeply branched positions.

We constructed the pET32a-based *E. coli* expression system of each homolog from *M. acetivorans* to obtain recombinant proteins as the fusion with an N-terminal Thioredoxin-His6-S-tag. MA1831 and MA4402 were efficiently expressed in *E. coli* and were purified using a polyhistidine tag, while the major portion of MA3723 seemingly formed an inclusion body even with the coexpression of chaperons, leaving only a small portion in the purified fraction (Fig. 3A-C). In a prenyltransferase assay
using $^{14}$C-labeled IPP and an allylic prenyl diphosphate [DMAPP, GPP, FPP or geranylgeranyl diphosphate (GGPP)], only MA1831 catalyzed the condensation of IPP with prenyl diphosphates longer than C$_5$. Although the purified MA1831 fraction contained some protein contaminants, the affinity-purified fraction from the mock cells, which contained similar contaminants and was used as the negative control, did not show prenyltransferase activity. The enzyme did not accept DMAPP as an allylic substrate, which resembles the substrate specificity of cis-prenyltransferases that are responsible for glycosyl carrier lipid biosynthesis [1]. MA3723 and MA4402 were completely inactive, even when they were mixed or coexpressed. We cannot exclude the possibility that the addition of the long (>160 a.a.) N-terminal tag inactivates MA3723 and MA4402, while such N-terminal tagging was utilized in our previous study on cis-prenyltransferase (and also for MA1831 in the present work) without causing inactivation [6]. Given these conditions, we started the characterization of the recombinant MA1831 in order to obtain information on the biosynthesis of the glycosyl carrier lipid in *M. acetivorans*.

Reversed phase TLC analyses of the products from the prenyltransferase reactions with MA1831 were performed after the acid phosphatase treatment of them. It should be noted that Triton X-100 was at first added in these reactions because a detergent is sometimes required as the activator of cis-prenyltransferases [16, 17]. The major products from these standard prenyl condensation reactions using the allylic substrate (GPP, FPP or GGPP) and IPP with the same concentration had a chain length of C$_{30}$ (Fig. 4A). The chain lengths of the products were, however, still shorter than those of the glycosyl carrier lipids from *M. acetivorans*. As mentioned in our previous paper [18], we detected undecaprenol and dihydroundecaprenol (though the position of
the double bond reduction was unclear) via LC-MS analysis of the lipid that was
extracted from \textit{M. acetivorans} and then treated with phosphatase (Fig. 5). These C_{55}
polyprenols are considered to be the glycosyl carrier lipids of the archaeon because the
lipid from a relative species, \textit{Methanosarcina barkeri}, also has been reported as
undecaprenol [15]. To establish the chain length of the final product of MA1831, we
raised the ratio of IPP against the allylic substrate to 10, and removed Triton X-100,
which generally activates \textit{cis}-prenyltransferases by enhancing turnover, but at the same
time causes the production of shorter products [19, 20]. As a result, the chain lengths of
the major products increased to C_{35-45} (Fig. 4B), which was still shorter than C_{55}.
Moreover, we tried the removal of the N-terminal tag from the recombinant MA1831 to
investigate its effect on product chain length. Enterokinase treatment at 22°C of the
fusion enzyme, however, resulted in inactivation of the prenyltransferase. Therefore, the
gene of MA1831 was subcloned into a pET48b(+) vector because the fusion protein
obtained from the system can be cleaved by human rhinovirus 3C protease at 4°C. By
the treatment with the protease (Fig. 3D), an N-terminal Thioredoxin-His$_6$-tag was
removed without a significant loss of MA1831 activity. The tag-free MA1831 was
purified and used for prenyltransferase assay under the same conditions, i.e., 10 times
larger amount of IPP against that of an allylic substrate, without Triton X-100, and so on.
The removal of the tag slightly elongated the chain lengths of major products to C_{35-50}
(Fig. 4C). The formation of a C_{55} product, i.e., undecaprenyl diphosphate, was observed
only when GPP was used as the allylic substrate. The longest products, however, were
shorter than C_{55} when FPP and GGPP were the allylic substrate.

Next, as a trial, we examined the activity of MA1831 to catalyze non-standard
reactions that included a non-head-to-tail prenyl condensation between allylic prenyl
diphosphates. First, $^{14}$C-labeled FPP was reacted with the same amount of an allylic
substrate using either DMAPP, GPP, FPP, or GGPP in the absence of Triton X-100.
Only the combination of FPP and DMAPP gave a new product with a chain length that
was estimated to be $C_{20}$ because its $R_f$ value approximated that of geranylgeraniol in the
reversed-phase TLC analysis (Fig. 6A). The chain length did not change when the ratio
of DMAPP against FPP was increased to 10 (Fig. 6B). The product from FPP and
DMAPP was also analyzed via normal-phase TLC. A new major product migrated
slightly farther than FPP, but was distinctly separated from prenyl alcohols and prenyl
monophosphates, suggesting that the product had a diphosphate group (Fig. 6C).
When $^{14}$C-labeled GGPP was used, faint spots of new products were observed when
DMAPP or GPP was used as the counter substrate (Fig. 6D). These results demonstrated
that MA1831 accepts DMAPP as the prenyl acceptor substrate alternative to IPP, when
FPP is the prenyl donor substrate. It should be noted that the activity of non-head-to-tail
prenyl condensation between FPP and DMAPP was also observed with the tag-free
MA1831 derived from the pET48b-based system (data not shown), while all the data
shown above were obtained using the fusion enzyme from the pET32a-based system.
To establish the compatibility of the non-standard reaction with a standard
head-to-tail condensation, kinetic analyses of the N-terminal tagged MA1831 were
performed in the absence of Triton X-100. To assay the standard reaction, we quantified
the radioactivity incorporated from $[^{14}$C]IPP into the mixture of hydrophobic products.
The $K_m$ for IPP was 7.21 µM when the concentration of FPP was maintained at 100 µM,
while the $V_{max}$ was 4.71 nmol IPP·min$^{-1}$·mg protein$^{-1}$. We also observed the inhibitory
effect of DMAPP against the standard reaction to roughly estimate the affinity of the
enzyme to DMAPP, which acts as a mixed inhibitor, as shown in Fig. 7, not just as a
competitor against IPP. DMAPP itself is not a preferred prenyl donor substrate for
MA1831 and thus unlikely to compete with FPP, but it is conceivable that the product of
the non-head-to-tail condensation between FPP and DMAPP competes with FPP. In the
interest of simplification, however, the $K_i$ of DMAPP was calculated using an equation
for competitive inhibition against IPP. In the presence of 150 µM DMAPP (and 100 µM
FPP), the $K_m$' for IPP was increased to 12.2 µM, while the $V_{max}$' was decreased to 2.92
nmol IPP·min$^{-1}$·mg protein$^{-1}$. The calculated $K_i$ of DMAPP was 217 µM, which was
much larger than the $K_m$ and $K_m$' for IPP. Moreover, the specific activity of MA1831
was measured using 5 µM $^{14}$C-labeled FPP. When 100 µM IPP was used as the counter
substrate, the specific activity was 4.77 nmol IPP·min$^{-1}$·mg protein$^{-1}$, which almost
approximated the $V_{max}$ of the standard reaction. This result suggested that the
concentration of FPP was sufficiently higher than its $K_m$, which seemed below 1 µM but
could not be determined accurately because the radioactivity of the products
incorporated from $^{14}$C-labeled IPP was too low when FPP was added at lower
concentrations where a Michaels and Menten curve would be given. When the counter
substrate was changed to 400 µM DMAPP, the concentration exceeding its $K_i$, the
specific activity of the non-standard reaction was 0.477 nmol DMAPP·min$^{-1}$·mg
protein$^{-1}$, which was 10-fold smaller than the standard reaction. These data clearly
showed that the non-standard reaction between FPP and DMAPP was a side reaction,
and that the main reaction of MA1831 was the standard head-to-tail condensation
between the allylic substrates and IPP.

Such substrate specificity, however, has never been reported for usual
$cis$-prenyltransferases, nor for their homologs that catalyze only non-head-to-tail prenyl
condensations. Thus, the structure of the irregular product of MA1831 obtained from
the subsidiary non-standard reaction between FPP and DMAPP attracted our interest. The LC-ESI-MS analysis of the product gave a negative ion with an \( m/z \) value that was equal to that of GGPP, while the retention time of the product was slightly different from that of GGPP (Fig. 8). This demonstrated that the product was an isomer of GGPP. Then we hydrolyzed the product with acid phosphatase and purified the resultant alcohol using HPLC. As outlined in Figs. 9-11 and Table 1, \(^1\text{H}\) and \(^{13}\text{C}\)-NMR, \(^1\text{H}-^1\text{H}\)-COSY, \(^1\text{H}-^{13}\text{C}\)-edited-hetero-nuclear single quantum coherence (HSQC), and \(^1\text{H}-^{13}\text{C}\)-hetero-nuclear multiple-bond connectivity (HMBC) analyses gave structural information that was sufficient to allow us to determine that the alcohol was geranyllavandulol. Therefore, we concluded that the irregular product of MA1831 is geranyllavandulyl diphosphate. The structure suggests that the reaction mechanism of the non-head-to-tail condensation between FPP and DMAPP that is catalyzed by MA1831 (Fig. 12) resembles that of lavandulyl diphosphate synthase.
Discussion

Among the three cis-prenyltransferase homologs from *M. acetivorans*, only MA1831 showed enzymatic activity. This result suggests that MA1831 is responsible for the biosynthesis of glycosyl carrier lipids in *M. acetivorans*. Nevertheless, the major products from the standard head-to-tail prenyl condensation reactions catalyzed by MA1831 were, under all reaction conditions we used for the *in vitro* assays, shorter than undecaprenyl diphosphate, which is the supposed precursor of C55 glycosyl carrier lipids in *M. acetivorans*. A slight amount of the C55 product was produced only when GPP was used as the allylic substrate against a 10-fold larger amount of IPP, but was not produced when the allylic substrate was either FPP or GGPP, which are general precursors for glycosyl carrier lipids. Furthermore, the probable orthologs of MA1831 are possessed only by some species of genus *Methanosarcina* among the archaea, and also by bacterial species of class Clostridia, which implies that the ancestral gene of the archaeal MA1831 orthologs might have emerged by horizontal gene transfer. Under these circumstances, we cannot exclude the possibility that one of the other cis-prenyltransferase homologs, i.e., MA3723 or MA4402, is responsible for the biosynthesis of glycosyl carrier lipids, whereas neither of them showed activity in the present study for unknown reasons. Recently, the heteromer formation of cis-prenyltransferase in animals and yeast was reported [21]. The heteromeric enzymes are composed of a subunit similar to the usual homodimeric cis-prenyltransferases and another subunit that is homologous to Nogo-B receptor, which is a transmembrane protein containing a cis-prenyltransferase-like domain. Such a heteromeric enzyme system also exists in some higher plants and is involved in dolichol and rubber biosynthesis [22, 23]. Our phylogenetic analysis revealed that the organisms possessing
the probable ortholog of MA3723 always have a probable MA4402 ortholog and that
the cluster containing MA4402 is somewhat related to *S. cerevisiae* NUS1, the Nogo-B
receptor-like subunit of dehydrodolichyl diphosphate synthase. Based on the notion that
MA4402 and MA3723 might also form a heteromer to exert enzyme activity, we mixed
the separately purified proteins and also constructed their co-expression system, in
which each protein was expressed in *E. coli* as a fusion with the N-terminal tag. Both of
these trials, however, did not yield active *cis*-prenyltransferase (data not shown).

MA1831 also catalyzes the non-head-to-tail prenyl condensation between FPP
and DMAPP, probably as a side reaction. To our knowledge, the product
geranyllavandulyl diphosphate is a novel compound, and the geranyllavandulyl
structure has never been observed in natural compounds. It should be mentioned,
however, that geranyllavandulyltoluquinol has been reported as the hypothetical
precursor of irregular diterpenoid toluquinols, i.e., neobalearone and epineobalearone,
from the brown alga *Custoseira amentacea var. stricta* [24, 25]. Geranyllavandulyl
diphosphate is, however, unlikely to be the precursor of these natural compounds
because it does not possess the allylic diphosphate structure needed for prenyl donor
substrates. Instead, the precursor could be geranylisolavandulyl diphosphate if the
prenyltransfer reaction to toluquinol occurs singly, unlike the biosynthesis of
sophoraflavanone G that requires two dimethylallyltransferases to yield a lavandulyl
group [26]. The fact that geranyllavandulyl diphosphate cannot be a prenyl donor
suggests that the non-head-to-tail prenyl condensation catalyzed by MA1831 probably
has no physiological meaning. Such irregular prenyl chains, however, can extend the
structural variety of natural isoprenoids by substituting their regular prenyl groups, and
thus will be valuable for the synthesis of novel compounds. The characteristic amino
acid sequence of MA1831 that causes the low specificity toward prenyl acceptor substrates, i.e., IPP for head-to-tail condensation and DMAPP for non-head-to-tail condensation, remains unclear. In the future, however, mutagenic and structural studies should uncover the specific traits of the cis-prenyltransferase homologs that catalyze non-head-to-tail prenyl condensations such as lavandulyl diphosphate synthase, cyclolavandulyl diphosphate synthase, isosesquilavandulyl diphosphate synthase, and MA1831. Moreover, it seems important to determine whether the previously reported cis-prenyltransferases also catalyze such non-standard reactions.
Materials and Methods

Materials

Precoated reversed-phase thin-layer chromatography (TLC) plates, RP18, and normal-phase TLC plates, Silica gel 60, were purchased from Merck Millipore, Germany. [1-^{14}C]IPP was purchased from GE Healthcare, UK. Non-labeled IPP, DMAPP and FPP were donated by Dr. Chikara Ohto, Toyota Motor Co. GPP was donated by Drs. Kyozo Ogura and Tanetoshi Koyama, Tohoku University. GGPP was purchased from Larodan fine chemicals, Sweden.

Database search and phylogenic analysis

A database search for the homologs of known cis-prenyltransferases was performed using a web service provided by KEGG. The multiple alignment of archaeal cis-prenyltransferase homologs was performed using the Clustal Omega program [27] provided at the EMBL-EBI website (http://www.ebi.ac.uk/). The phylogenetic tree was constructed via the neighbor joining method using the CLC Sequence Viewer, ver. 7.5 (CLC bio, Denmark).

Expression of recombinant cis-prenyltransferase homologs from M. acetivorans

The genes of the cis-prenyltransferase homologs from M. acetivorans, ma1831, ma3723 and ma4402, were amplified from genomic DNA using KOD DNA polymerase (TOYOBO, Japan), and the primer pairs are listed below (the recognition sequences of the restriction enzymes are underlined): for ma1831, 5’-gtccatggacatggatacacctaaatttaaaagac-3’ and 5’-atgctcgagttatccgcccaatgtaatatcttg-3’; for ma3723, 5’-atcatggacatgtaccgtggaataaggagc-3’ and
5’-gacctcagagacttcaagccactcctcg-3’; and, for ma4402, 5’-gtcataagactcttttttggc-3’ and 5’-agtctcagtcctttctctg-3’. The amplified DNA fragments were digested with NcoI and XhoI, and each of the digested fragments was ligated with an NcoI/XhoI-digested pET32a(+) expression vector (Novagen, USA) to construct pET32a-MA1831, pET32a-MA3723, and pET32a-MA4402, respectively. For the subcloning of ma1831 into pET48b(+), the gene was amplified from the pET32a-MA1831 plasmid using the primers 5’-ctctttcagggacccatggatacacctaaatttaaaagac-3’ and 5’-cggatcctggtacccttatccgcccaatgtaatat-3’. The amplified DNA fragment was ligated with an SmaI-digested pET48b(+) vector to construct pET48b-MA1831 using an In-Fusion advantage PCR cloning kit (TaKaRa, Japan) following the manufacturer’s instructions. For the construction of the coexpression system of MA3723 and MA4402, the ma3723 gene, with the gene region encoding the N-terminal Thioredoxin-His6-S-fusion tag and containing the upstream ribosome-binding site, was amplified from pET32a-MA3723 using the primer pair 5’-cggaaagtgactcagaaaggagatatacatgagcg-3’ and 5’-ggtggtggtctcgtgactcagctttctttaaatc-3’. The amplified DNA and XhoI-digested pET32a-MA4402 were ligated to construct pET32a-MA4402/3723, in which ma4402 and ma3723 compose an artificial operon, using an In-Fusion advantage PCR cloning kit. After construction of the plasmids, the sequences of the inserted genes were verified by DNA sequencing. The plasmids constructed as described above were introduced into adequate E. coli host strains such as BL21(DE3), BL21(DE3) transformed with a chaperon-expression plasmid pG-KJE8 (Clontech, Japan), and KRX (Promega, USA). BL21(DE3) was
transformed with pET32a-MA1831 or pET48b-MA1831 and cultured in 300 mL LB medium supplemented with 100 mg/L ampicillin at 37°C until formation of the log phase, which was followed by overnight cultivation at 25°C after the addition of 0.2 mM IPTG. BL21(DE3)/pG-KJE8 was transformed with pET32a-MA3723 and cultured in 300 mL LB medium containing 100 mg/L ampicillin, 20 mg/L chloramphenicol, 0.4 g/L L-arabinose, and 5 µg/L tetracycline. The culture was grown at 37°C until log phase, and additional cultivation was performed at 22°C overnight after the addition of 0.1 mM IPTG. KRX was transformed with pET32a-MA4402 and cultured in 300 mL LB medium supplemented with 100 mg/L ampicillin at 37°C until log phase, and then at 22°C overnight after the addition of 0.1% L-rhamnose. The cultivation of KRX transformed with pET32a-MA3723/4402 was performed under the same conditions used for KRX/pET32a-MA4402.

The cells harvested from the culture were disrupted by sonication using a UP200S ultrasonic homogenizer (Hielscher Ultrasonics, Germany) in a HisTrap binding buffer (20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, pH7.4). The homogenate was centrifuged at 24,000 g for 30 min, and the supernatant was recovered as a crude extract. The crude extract was loaded onto a 1 mL HisTrap crude FF column (GE Healthcare). The column was washed with 10 mL of the HisTrap binding buffer (imidazole concentration was raised to 60 mM as needed), and the recombinant proteins were eluted from the column with 5 mL of a HisTrap elution buffer (20 mM sodium phosphate, 0.5 M NaCl, 500 mM imidazole, pH7.4). For the cleavage of a Thioredoxin-His6-tag, which was fused at the N-terminal of MA1831 purified from BL21(DE3)/pET48b-MA1831, 300 µg of the purified fusion enzyme was reacted with 5 units of polyhistidine-tagged human rhinovirus 3C protease (Funakoshi, Japan)
overnight at 4°C, after buffer exchange into 50 mM Tris-HCl, pH8.0, containing 150 mM NaCl. The solution of the protease reaction was loaded on a 1 mL HisTrap crude FF column equilibrated with the HisTrap binding buffer, and the flow-through fraction was used as the solution of purified tag-free MA1831. The level of purification was confirmed via SDS-PAGE.

Prenyltransferase assay and radio-TLC analysis of the products

A reaction mixture typically contains, in a total volume of 200 µL, 0.5 or 5 nmol [14C]IPP (2.04 GBq/mm or 81.4 MBq/mm, respectively), 0.5 nmol allylic prenyl diphosphate (DMAPP, GPP, FPP or GGPP), 20 µmol 3-morpholinopropanesulfonic acid (MOPS)-NaOH, pH7.0, 1 µmol MgCl₂, 0 or 0.1% Triton X-100, and 30-1,000 pmol of the purified cis-prenyltransferase homolog. After incubation at 40°C for 1 hr, the reaction was stopped by chilling. After the addition of 200 µL water saturated with NaCl, the hydrophobic products were extracted from the mixture with 600 µL 1-butanol equilibrated against saturated saline. Exclusively for the kinetic analyses, the butanol layer was washed with the same volume of saturated saline equilibrated against 1-butanol. Radioactivity in an aliquot of the butanol layer was measured using an LSC-5100 liquid scintillation counter (Aloka, Japan). An unusual unit, nmol IPP·min⁻¹·mg protein⁻¹, was used to express the specific activity of cis-prenyltransferase because this assay method gives the amount of IPP incorporated into the mixture of hydrophobic products rather than the amount of the products. Kinetic parameters were obtained via Lineweaver-Burk linear regression analysis. The remainder of the butanol extract was provided for product analysis via the radio-TLC method described below.

The non-head-to-tail prenyl condensation catalyzed by MA1831 was assayed
using $^{14}$C-labeled FPP and GGPP, which were synthesized using FPP synthase from *Geobacillus stearothermophilus* and GGPP synthase from *S. acidocaldarius*, respectively, as described elsewhere [18]. The assay mixture typically contained, in a total volume of 200 µL, ~0.15 nmol of the $^{14}$C-labeled allylic prenyl diphosphate (FPP or GGPP), 0.15 nmol of non-labeled allylic prenyl diphosphate (DMAPP, GPP, FPP or GGPP), 20 µmol MOPS-NaOH, pH7.0, 1 µmol MgCl$_2$, and 30-400 pmol of purified MA1831 fused with the N-terminal tag. After 1 hr incubation at 40°C, the hydrophobic products and substrates were extracted with 1-butanol as described above to be used in radio-TLC analysis.

For reversed-phase TLC analysis, the butanol extract containing radiolabeled products/substrates was treated overnight with 2 mg of potato acid phosphatase (Sigma Aldrich, USA) in a 2:1 mixture of methanol and 1 M sodium acetate buffer, pH4.6, following a method developed by Fujii et al. [28]. The resultant prenyl alcohols were extracted with $n$-pentane and analyzed with a RP18 reversed-phase TLC plate developed with acetone/H$_2$O (9:1). For normal-phase TLC analysis, the butanol extract was evaporated without phosphatase treatment, and then the residue was dissolved to methanol and analyzed with a Silica gel 60 normal-phase TLC plate developed with methanol/chloroform/5 mM NH$_4$HCO$_3$ (4:6:0.9). The distribution of radioactivity on the plate was visualized with a Typhoon FLA7000 multifunctional scanner (GE Healthcare) and, if required, quantified using the attached software Image Quant TL to calculate the specific activities based on the consumption of the prenyl-donor substrates, either IPP or DMAPP.

LC-ESI-MS analysis of the product from a non-standard MA1831 reaction
The reaction mixture for the preparation of a non-head-to-tail condensation product of MA1831 contained, in a final volume of 300 µL, 20 nmol DMAPP, 20 nmol FPP, 30 µmol MOPS-NaOH, pH7.0, 0.3 µmol MgCl2, and 0.4 nmol purified MA1831 fused with the N-terminal tag. After incubation for 1 hr at 40°C, the reaction was stopped by chilling. After the addition of 300 µL water saturated with NH4HCO3, the hydrophobic products/substrates were extracted with 600 µL 1-butanol equilibrated against water. The butanol layer was evaporated under a N2 stream, and the residue was dissolved with methanol/acetonitrile (1:1) to be used in LC-ESI-MS analysis.

LC-ESI-MS analysis was performed with an Esquire 3000 ion trap system (Bruker Daltonics, USA) connected to an Agilent 1100 Series HPLC (Agilent Technologies, USA). The compounds eluted from a COSMOSIL packed column 5C18-AR-II (2.0ID × 150 mm, Nacalai, Japan) were detected via UV absorption at 210 nm and ESI-MS in the negative mode. The mobile phase used for the analysis was made up of a mixture of solution A (25 mM NH4HCO3) and B (acetonitrile) at a flow rate of 0.2 mL·min⁻¹. Sample elution was performed with the program as follows: the ratio of B was 10% for the initial 3 min; a linear gradient from 10 to 100% B for the next 15 min; and, 100% B for an additional 10 min. After each analysis, the column was equilibrated with 10% B. The MS parameters were the same as those used in our previous work [29]: sheath gas, N2 of 30 psi; dry gas, N2 of 7.0 L·min⁻¹, 320°C; scanning range, 50-1,000 m/z; scan speed, 13,000 m/z·sec⁻¹; ion charge control target, 20,000; maximum accumulation time, 100 ms; averages, 10; and, rolling averaging, 2.

NMR analysis of the MA1831 product

The reaction mixture contained, in a final volume of 20 mL, 8 µmol DMAPP, 2
μmol FPP, 1 mmol MOPS-NaOH, pH7.0, 20 μmol MgCl$_2$, and 42 nmol purified MA1831 fused with the N-terminal tag. After 3 hr incubation at 40°C, 20 mL of 1 M sodium acetate buffer, pH4.6, containing 40 mg potato acid phosphatase was added. The same reaction was performed in a total of 6 tubes at 37°C overnight. The hydrolyzed product was extracted from the mixture with the same volume, and then again with a half volume of n-pentane. The pentane layer was merged, and evaporated under a N$_2$ stream. The residue was dissolved with 2 mL methanol. Typically, 100 μL of the methanol solution was provided for HPLC separation performed with a Shimadzu HPLC system equipped with a COSMOSIL Packed Column 5C$_{18}$-AR-II (10ID × 150 mm, Nacalai, Japan). The mobile phase was methanol/H$_2$O (9:1), and the flow rate was 3 mL·min$^{-1}$. The column temperature was kept at 40°C. The UV absorption of carbon-carbon double bonds was monitored at 210 nm, and a peak fraction was recovered. This separation process was repeated ~20 times, and the recovered fractions were merged. The fraction was evaporated, and the residue was dissolved with 500 μL CD$_3$OD. This evaporation-dissolution process was repeated twice for complete substitution of the solvent with CD$_3$OD. $^1$H and $^{13}$C NMR, $^1$H-$^1$H-COSY, $^1$H-$^{13}$C-edited-HSQC, and $^1$H-$^{13}$C-HMBC analyses of the sample were performed with an AVANCE III HD 600 NMR spectrometer equipped with a cryoprobe (600 MHz, Bruker).

Structural investigation of glycosyl carrier lipids from *M. acetivorans*

Total lipids were extracted via a method established by Bligh and Dyer [30] from ~3 g of *M. acetivorans* cells cultivated as described elsewhere [18]. The cells were suspended in 15 mL of water, and 70 mL of methanol and 37.5 mL of chloroform were
added. The mixture was stirred for 2 hr, and then 37.5 mL each of chloroform and water were added. After being stirred briefly, the mixture was allowed to settle and separate into two phases. The aqueous layer was extracted again with 10 mL each of chloroform and water. To the combined organic layer, an equivalent volume of methanol and an 80% volume of water were added and stirred, and the mixture was separated into two phases. After the organic layer was evaporated to dryness at 55°C under a stream of nitrogen gas, the residual was dissolved with 4 mL of methanol/chloroform (1:1) and mixed with an excess volume of cold acetone to be placed in the dark at 4°C for 3 hr. The insoluble matter collected by centrifugation and decantation was dissolved with 0.5 mL of 1-butanol and subjected to a phosphatase treatment as described above. The hydrolysate was extracted with n-pentane, evaporated to dryness, and again subjected to acetone precipitation. After centrifugation to remove insoluble matter, the supernatant was evaporated and dissolved to methanol/2-propanol (1:1) for LC-ESI/MS analysis. The analysis was performed under the same conditions described above, with the exception of the mobile phase being replaced with methanol/2-propanol/100 mg·L⁻¹ sodium acetate (9:10:1).
Acknowledgements

This work was supported in part by JSPS KAKENHI Grant Number 25108712 and 26660060 for H.H. and 245479 for T.O.

Author contributions

TO and HH planned the experiments; TO, KE and KK performed the experiments; TO, KE, KK, TY and HH analyzed the data; and, HH wrote the paper.
References


Table 1. $^{13}$C and $^1$H NMR assignments for geranyl lavandulol

<table>
<thead>
<tr>
<th>position</th>
<th>$\delta_C$</th>
<th>$\delta_H$ (multiplicity, $J$ in Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>63.7</td>
<td>3.52 (dd, 10.8, 18.5), 3.53 (dd, 10.8, 19.8)</td>
</tr>
<tr>
<td>2</td>
<td>49.9</td>
<td>2.25 (m)</td>
</tr>
<tr>
<td>3</td>
<td>27.9</td>
<td>2.04 (m), 2.21 (m)</td>
</tr>
<tr>
<td>4</td>
<td>122.4</td>
<td>~5.12 (m)</td>
</tr>
<tr>
<td>5</td>
<td>135.3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>39.4</td>
<td>~2.00 (m)</td>
</tr>
<tr>
<td>7</td>
<td>~26</td>
<td>~2.10 (m)</td>
</tr>
<tr>
<td>8</td>
<td>124.0</td>
<td>~5.12 (m)</td>
</tr>
<tr>
<td>9</td>
<td>134.4</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>39.4</td>
<td>~2.00 (m)</td>
</tr>
<tr>
<td>11</td>
<td>~26</td>
<td>~2.10 (m)</td>
</tr>
<tr>
<td>12</td>
<td>124.0</td>
<td>~5.12 (m)</td>
</tr>
<tr>
<td>13</td>
<td>130.7</td>
<td></td>
</tr>
<tr>
<td>5-methyl</td>
<td>14.8</td>
<td>1.63 (s)</td>
</tr>
<tr>
<td>9-methyl</td>
<td>14.7</td>
<td>1.61 (s)</td>
</tr>
<tr>
<td>$\omega$-cis</td>
<td>16.3</td>
<td>1.62 (s)</td>
</tr>
<tr>
<td>$\omega$-trans</td>
<td>24.5</td>
<td>1.69 (s)</td>
</tr>
<tr>
<td>1’</td>
<td>145.6</td>
<td></td>
</tr>
<tr>
<td>2’</td>
<td>111.31</td>
<td>4.75 (m), 4.84 (m)</td>
</tr>
<tr>
<td>1’-methyl</td>
<td>18.65</td>
<td>1.70 (s)</td>
</tr>
<tr>
<td>OH</td>
<td></td>
<td>4.64 (broad)</td>
</tr>
</tbody>
</table>
Figure legends

Figure 1. Reactions catalyzed by cis-prenyltransferase and its homologs
Some natural compounds biosynthesized from the products of the enzymes are also exemplified.

Figure 2. Phylogenetic tree of cis-prenyltransferase homologs
The enzymes reported to have activity as listed below are colored in red: ApeUPS, cis-prenyltransferase from A. pernix; SacUPS, from S. acidocaldarius; TkoUPS, from T. kodakarensis; EcoUPS, undecaprenyl diphosphate synthase from E. coli; SceRER2 and SceSRT1, dehydrodolichyl diphosphate synthases from Saccharomyces cerevisiae; and, SceNUS1, a heteromer subunit of S. cerevisiae dehydrodolichyl diphosphate synthase. Only SceNUS1 did not show enzyme activity independently. All other proteins colored in blue are cis-prenyltransferase homologs from Euryarchaeota: AF0707 and AF1219 (nowadays, AF_RS03590 and AF_RS06170, respectively), from Archaeoglobus fulgidus; Hvo_2315 and Hvo_2318, from Haloferax volcanii; MM_0014, MM_0618 and MM_1083, from Methanosarcina mazei; MBUR_0041, from Methanococcales burtonii; MCP_1817 and MCP_2939, from Methanocella paludicola; MTHE_0589 and MTHE_1499, from Methanosaeta thermophila; MJ_1372, from Methanocaldococcus jannaschii; MTH232, from Methanothermobacter thermautotrophicus; and, PH1590, from Pyrococcus horikoshii. The homologs from M. acetivorans, MA1831, MA3723 and MA4402, are colored in green. Bootstrap values over 70% were indicated at each node.
Figure 3. SDS-PAGE analysis of the purified recombinant proteins from *E. coli* cells.

(A) MA1831 expressed in BL21(DE3)/pET32a-MA1831. (B) MA3723 expressed in BL21(DE3)/pG-KJE8/pET32a-MA3723. (C) MA4402 expressed in KRX/pET32a-MA4402. (A-C) Lanes 1 and 2, pellet and soluble fractions from the centrifugation of lysed *E. coli* cells, respectively; lanes 3-5, flow-through, wash and elution fractions from a HisTrap affinity column, respectively. The expected band of the recombinant protein in the elution fraction (and also that in the pellet fraction if necessary) is indicated with an asterisk. (D) The cleavage of fusion MA1831. Lane 1, MA1831 fused with the N-terminal Thioredoxin-His$_6$-tag, which was purified from in BL21(DE3)/pET48b-MA1831; lane 2, MA1831 solution after reaction with human rhinovirus 3C protease; lane 3, flow-through fraction from a HisTrap affinity column, containing tag-free MA1831 indicated with an asterisk.

Figure 4. Radio-TLC analysis of the products from standard reactions of MA1831.

Reversed-phase TLC analyses were performed for the hydrolyzed products from head-to-tail prenyl condensation reactions. The enzymes used were: for (A and B), MA1831 fused with an N-terminal Thioredoxin-His$_6$-S-tag, obtained from the pET32a-based expression system; and for (C), MA1831 purified after the removal of an N-terminal Thioredoxin-His$_6$-tag, obtained from the pET48b-based system. The allylic substrates shown below the figures were reacted with the same amount (A) or with a 10-fold larger amount (B and C) of $^{14}$C-labeled IPP, in the presence (A) and absence (B and C) of 0.1% Triton X-100. s.f., solvent front; ori., origin.
Figure 5. LC-ESI-MS analysis of glycosyl carrier lipids from *M. acetivorans*

The UV chromatograms of (A) authentic polyprenols mainly containing undecaprenol (C55, open triangle) and dodecaprenol (C60) and (B) the phosphatase-treated total lipid extracted from *M. acetivorans*. The peak co-eluted with authentic undecaprenol and an ion with *m/z* of 790.1 is indicated by an open triangle, and another peak co-eluted with an ion with *m/z* of 792.1 is indicated by a closed triangle. (C and D) Extracted ion chromatograms with indicated *m/z* ranges. The *m/z* values of 790.1 and 792.1 correspond to [M+Na]^+ of undecaprenol and dihydroundecaprenol, respectively.

Figure 6. Radio-TLC analysis of the products from non-standard reactions of MA1831

(A) Reversed-phase TLC analysis for the hydrolyzed products from non-head-to-tail prenyl condensation between 14C-labeled FPP and the same amount of a non-labeled allylic substrate shown in the figure. An arrowhead indicates the spot of a new product from a non-head-to-tail condensation between FPP and DMAPP. (B) Reversed-phase TLC analysis of the hydrolyzed products from the reaction between 14C-labeled FPP and DMAPP with various ratios shown in the figure. (C) Normal-phase TLC analysis of the products from the reaction between 14C-labeled FPP and DMAPP, without phosphatase treatment. An arrowhead indicates the spot of the new product. (D) Reversed-phase TLC analysis of the hydrolyzed products from the reaction between 14C-labeled GGPP and the same amount of a non-labeled allylic substrate shown in the figure. Arrowheads emphasize the faint spots of probable reaction products. s.f., solvent front; ori., origin
Figure 7. Lineweaver-Burk plots for the determination of $K_m$ for IPP and $V_{max}$
Head-to-tail condensation reactions against 100 µM FPP were performed in the absence and presence of 150 µM DMAPP.

Figure 8. LC-ESI-MS analysis of the product from MA1831-catalyzed non-head-to-tail prenyl condensation between FPP and DMAPP
(A) UV chromatograms at 210 nm of the hydrophobic substrates/products extracted with 1-butanol from the reaction mixture containing MA1831 (upper panel) or not (middle panel), and that of authentic GGPP (lower panel). An arrowhead indicates the peak of a new product. (B) Full scan mass spectrum of the new product. The analysis was performed in the negative-ion mode.

Figure 9. NMR analyses of the hydrolyzed product from MA1831-catalyzed non-head-to-tail prenyl condensation between FPP and DMAPP
(upper panel) $^1$H-NMR spectrum. (lower panel) $^{13}$C-NMR spectrum. The atom numbering is shown in Figure 11.

Figure 10. 2D-NMR analyses of the hydrolyzed product from MA1831-catalyzed non-head-to-tail prenyl condensation between FPP and DMAPP
(upper panel) $^1$H-$^1$H-COSY spectrum. (lower panel) $^1$H-$^{13}$C-edited-HSQC spectrum. The atom numbering is shown in Figure 11.

Figure 11. 2D-NMR analyses of the hydrolyzed product from MA1831-catalyzed non-head-to-tail prenyl condensation between FPP and DMAPP (continued)
(upper panel) $^1$H-$^{13}$C-HMBC spectrum. (lower panel) Selected COSY (red) and HMBC (blue) data of the hydrolyzed product geranyllavandulol, with the numbering of the atoms.

**Figure 12. Hypothetical reaction mechanism of a non-head-to-tail condensation between FPP and DMAPP catalyzed by MA1831**

The specificity of proton abstraction by a base from the $E$- or $Z$-methyl group of DMAPP is unclear now.