JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. 2017, 139, 14556-14567
pubs.acs.org/JACS

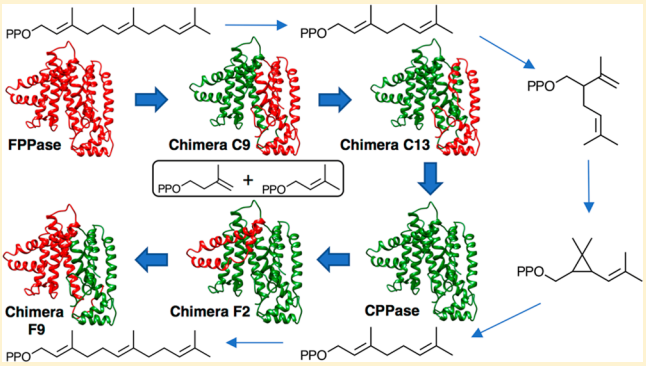
Structure–Function Studies of *Artemisia tridentata* Farnesyl Diphosphate Synthase and Chrysanthemyl Diphosphate Synthase by Site-Directed Mutagenesis and Morphogenesis

J. Scott Lee,[†] Jian-Jung Pan, Gurusankar Ramamoorthy, and C. Dale Poulter*

Department of Chemistry, University of Utah, 315 South 1400 East, RM 2020, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: The amino acid sequences of farnesyl diphosphate synthase (FPPase) and chrysanthemyl diphosphate synthase (CPPase) from *Artemisia tridentata* ssp. Spiciformis, minus their chloroplast targeting regions, are 71% identical and 90% similar. FPPase efficiently and selectively synthesizes the "regular" sesquiterpenoid farnesyl diphosphate (FPP) by coupling isopentenyl diphosphate (IPP) to dimethylallyl diphosphate (DMAPP) and then to geranyl diphosphate (GPP). In contrast, CPPase is an inefficient promiscuous enzyme, which synthesizes the "irregular" monoterpenes chrysanthemyl diphosphate (CPP), lavandulyl diphosphate (LPP), and trace quantities of maconelliyl diphosphate (MPP) from two molecules of DMAPP, and couples IPP to DMAPP to give GPP. *A. tridentata* FPPase and CPPase belong to the chain elongation protein family (PF00348), a subgroup of the terpenoid synthase superfamily (CL0613) whose members have a characteristic α terpene synthase α-helical fold. The active sites of *A. tridentata* FPPase and CPPase are located within a six-helix bundle containing amino acids 53 to 241. The two enzymes were metamorphosed into one another by sequentially replacing the loops and helices of the six-helix bundle from enzyme with those from the other. Chain elongation was the dominant activity during the N-terminal to C-terminal metamorphosis of FPPase to CPPase, with product selectivity gradually switching from FPP to GPP, until replacement of the final α-helix, whereupon cyclopropanation and branching activity competed with chain elongation. During the corresponding metamorphosis of CPPase to FPPase, cyclopropanation and branching activities were lost upon replacement of the first helix in the six-helix bundle. Mutations of active site residues in CPPase to the corresponding amino acids in FPPase enhanced chain-elongation activity, while similar mutations in the active site of FPPase failed to significantly promote formation of significant amounts of irregular monoterpenes. Our results indicate that CPPase, a promiscuous enzyme, is more plastic toward acquiring new activities, whereas FPPase is more resistant. Mutations of residues outside of the α terpene synthase fold are important for acquisition of FPPase activity for synthesis of CPP, LPP, and MPP.



INTRODUCTION

Isoprenoid compounds are readily distinguished from other natural products by the distinctive methylbutyl (isoprene) units embedded within their carbon skeletons. With over 80 000 members,[1] they constitute the largest and most structurally diverse group of compounds synthesized in nature.[1] Common metabolites of the isoprenoid biosynthetic pathway include sterols, carotenoids, ubiquinones, and mono-, sesqui-, and diterpenes.[1−3] The carbon skeletons of isoprenoid compounds are constructed from the fundamental C_5 isomeric building blocks isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are joined by chain elongation (1′−4), branching (1′−2), cyclopropanation (c1′−2−3), and cyclobutanation (c1′−2−3−2′) reactions (Figure 1).[2] The 1′−4 pattern is the most common attachment found between isoprene units and is typically called a "regular" or "head-to-tail" coupling.[4] The other three are termed "irregular" or "non-head-to-tail".[5] The 1′−4 pattern results from a reaction that joins C1 in DMAPP, or its higher isoprenologs, to C4 in IPP. Thus, beginning with DMAPP, a homologous series of linear C_{10}, C_{15} C_{20}, etc. allylic diphosphates is formed, where each member can be the substrate for construction of more complex skeletons, including those with c1′−2−3, 1′−2, and c1′−2−3−2′ substructures.[2]

Two structurally distinct families of enzymes catalyze chain elongation. While each contains enzymes selective for the synthesis of products with different chain lengths, one family synthesizes allylic isoprenoid diphosphates with *E*-double bonds, and the other, allylic diphosphates with *Z*-double bonds. *E*-polyprenyl diphosphate synthases typically produce short-chain products early in the isoprenoid pathway.[6] These enzymes belong to the chain elongation family (PF00348) within the terpenoid synthase superfamily (CL0613),[7] whose

Received: July 20, 2017
Published: September 19, 2017

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

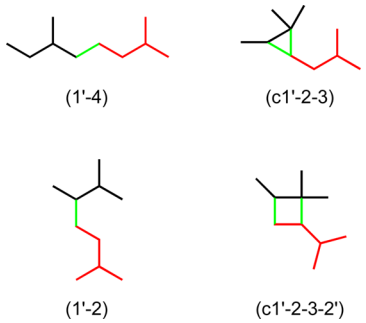


Figure 1. Four basic patterns for incorporation of isoprene units into metabolites by joining two isoprenoid diphosphates. Bonds between isoprene units are in green. The isoprene units are numbered according to the numbering schemes for IPP and DMAPP.

members also include protein families for cyclopropanation/rearrangement (PF00494) and cyclization (PF03936 and PF01397) (Figure 2). While all proteins in CL0613 contain a

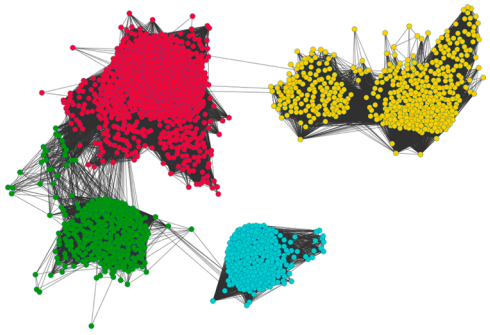


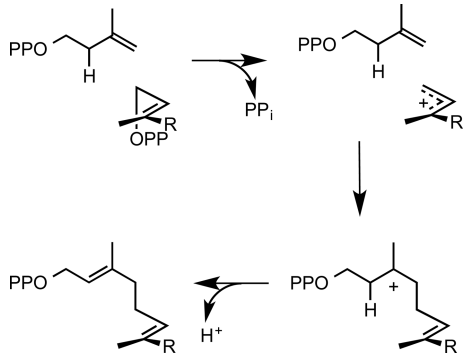
Figure 2. Sequence similarity network (Blast cutoff of e-7) for enzymes in the terpenoid synthase superfamily (CL0613); chain elongation (PF00348, red), cyclopropanation/rearrangement (PF00494, gold), cyclization (PF03936, green; PF01397, turquoise).

characteristic all α-helix α terpene synthase motif,[7−9] those belonging to PF03936 and PF01397 often have additional β and/or γ domains.[10] Enzymes in the *Z*-polyprenyl synthase family belong to PF01255 and synthesize longer-chain products.[11,12]

Chain elongation is the most common reaction in the isoprenoid pathway. The first step in *E*-selective chain elongation couples DMAPP and IPP to produce geranyl diphosphate (GPP). GPP can then couple with another molecule of IPP to give farnesyl diphosphate (FPP), and so on.[4] The isoprene units are joined by a dissociative electrophilic alkylation of the nucleophilic double bond in IPP by an electrophilic cation generated from the allylic diphosphate cosubstrate (DMAPP, GPP, FPP, etc.) (Scheme 1).[13,14] Extension of the allylic chain by sequential addition of molecules of IPP produces a homologous series of allylic isoprenoid diphosphates that in turn are substrates for branch point enzymes that feed into the numerous branches of the terpenoid pathway.[15]

The cyclopropanation reaction is also common, but this linkage is subsequently disassembled in most metabolites and is typically detected indirectly through rearranged carbon skeletons in downstream metabolites.[16] For example, the first pathway specific reaction in biosynthesis of sterols and carotenoids is a c1′−2−3 cyclopropanation, followed by a rearrangement to give structures with 1′−1 (head-to-head)

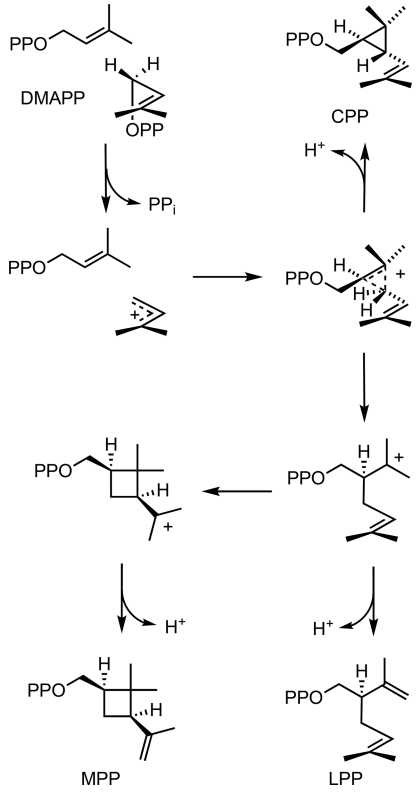
Scheme 1. Electrophilic Alkylation Mechanism for Chain Elongation[a]



[a]R = CH_3, C_6H_{11}, $C_{11}H_{19}$, etc.

linkages between the two component isoprenoid fragments.[17,18] The mechanism for cyclopropanation is also a dissociative electrophilic alkylation, where the allylic cation generated by one molecule of an allylic diphosphate attacks the allylic double bond in another molecule, as illustrated for the coupling of DMAPP to give chrysanthemyl diphosphate (CPP) in Scheme 2.[2] Metabolites with structures derived from 1′−2 branching, as

Scheme 2. An Integrated Electrophilic Alkylation-Rearrangement Mechanism for Formation of Cyclopropane, Branched, and Cyclobutane Skeletons



illustrated by lavandulyl diphosphate (LPP), are much less common,[19] while those with 1′−2−3−2′ cyclobutane structures, illustrated by maconelliyl diphosphate (MPP), have only been reported for mealy bug mating pheromones.[20,21] The 1′−2 and c1′−2−3−2′ structures are produced from the same

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

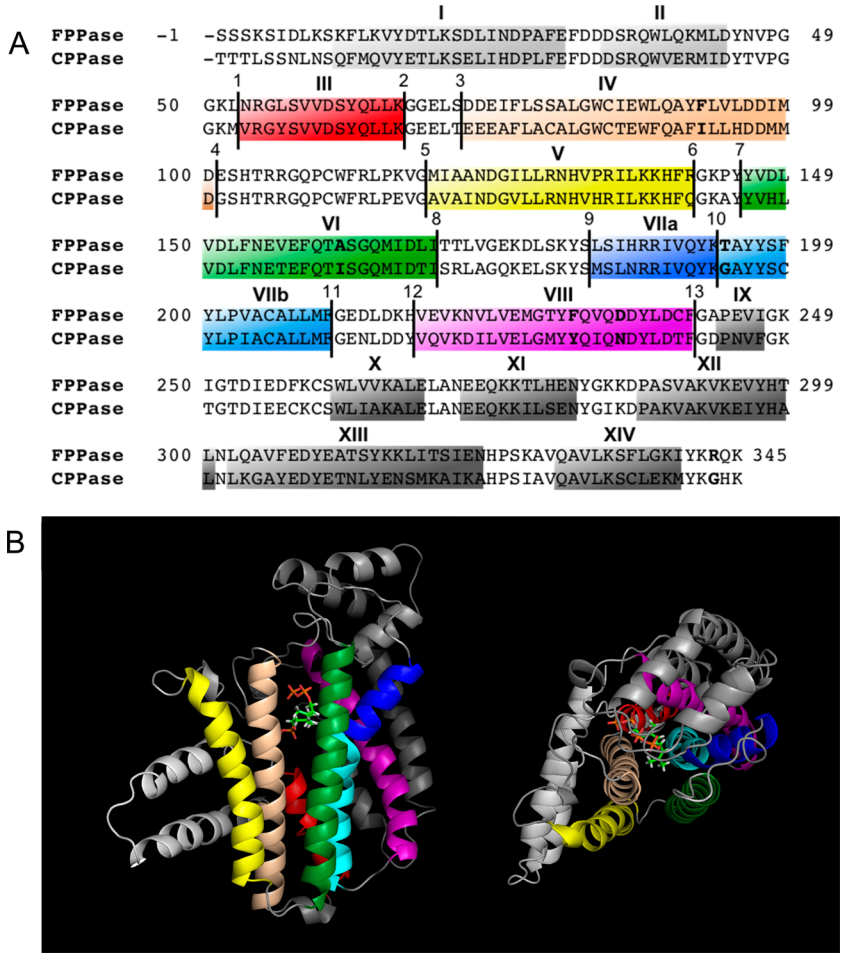


Figure 3. (A) Sequence alignments for *A. tridentata* FPPase and CPPase, where helices in the N-terminal sequence (light gray), six-helix bundle (colored), and C-terminal regions are highlighted. K193 and T194 lie on either side of a kink between helix VIIa (dark blue) and helix VIIb (cyan). The 13 junctions used for morphogenesis are located at the beginning and end of each helix and at the KT junction at the kink between VIIa and VIIb. Active site residues in CPPase that are different from those of FPPase are in bold. (B) Ribbon structures from the side and the top of a subunit in the *A. tridentata* FPPase homodimer.

protonated cyclopropane intermediates that give the c1′−2−3 skeletons.[2,22]

Enzymes that catalyze c1′−2−3 cyclopropanation typically couple two molecules of DMAPP to give chrysanthemyl diphosphate (CPP),[22] two molecules of FPP to give presqualene diphosphate,[23,24] or two molecules of geranylgeranyl diphosphate to give prephytoene diphosphate.[25] Presqualene and prephytoene diphosphates are intermediates in the sterol and carotenoid pathways, respectively. Enzymes that catalyze 1′−2 branching are found in PF00348 and the structurally unrelated *Z*-polyprenyl synthase family PF01255.[26] The only wild-type enzyme known to catalyze c1′−2−3−2′ cyclobutanation is CPPase, where cyclobutanation is a minor promiscuous activity.

The enzymes in CL0613 have distinctive aspartate/glutamate-rich sequences that, along with Mg^{2+} ions, bind the diphosphate moieties of their substrates and catalyze carbocation formation.[27,28] The structures of proteins in PF00348 and PF00494 suggest that they diverged from a common chain elongation ancestor early in the evolution of terpenoid biosynthesis.[10] This hypothesis is supported by the recent emergence of a new cyclopropanation enzyme, chrysanthemyl diphosphate synthase (CPPase) in chrysanthemums (*Chyrsanthemum cinerariifolium*)[22] and sagebrush (*Artemisia tridentata* ssp. Spiciformis)[29] from farnesyl diphosphate synthase (FPPase) in the same plants. The chain elongation and cyclopropanation enzymes are members of PF00348 and share structural features found in other chain elongation enzymes in that family.[6] *A. tridentata* contains two FPPases that are highly selective and efficient for chain elongation of DMAPP to FPP. In contrast, *A. tridentata* CPPase is an inefficient, promiscuous enzyme that converts two molecules of DMAPP to CPP (c1′−2−3, cyclopropanation), lavandulyl diphosphate (LPP, 1′−2, branching), and maconelliyl diphosphate (MPP, c1′−2−3−2′, cyclobutanation), as well as IPP and DMAPP to GPP (1′−4, chain elongation). These properties are consistent with those expected for a recently evolved enzyme that has gained a new function but has not yet optimized its ability to catalyze the new reaction.[30]

E-Polyprenyl chain elongases have five conserved regions within a six-helix bundle that contains characteristic aspartate-rich DDxx(xx)D diphosphate binding motifs.[31,32] Differences between *A. tridentata* FPPase and CPPase within this bundle are minimal (Figure 3).[6] The substantial similarity between FPPase and CPPase prompted us to construct chimeric *A. tridentata* proteins where FPPase was transformed into CPPase, and vice versa, by replacing conserved regions of one enzyme with those from the other.[6] The constructs retained the

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

ability to catalyze chain elongation while acquiring the ability to catalyze branching, cyclobutanation and cyclopropanation. We now present product and kinetic studies for chimeras where each helix and loop between the first and last helix in the six-helix bundle in one enzyme is replaced by the corresponding sequence in the other and for site-directed mutants where only active-site residues are replaced. Our results illustrate the importance of residues outside of the active site to influence the regioselectivity of the reactions.

RESULTS

Engineered Proteins. *1. Chimeras.* Chimeric enzymes were constructed from *A. tridentata* FPPase and CPPase by replacing the amino acids in one protein with those from the other at 13 different junctions, beginning at the N-terminus using megaprimer whole plasmid (MEGAWHOP) and structure–based combinatorial protein engineering (SCOPE) PCR protocols.[33,34] The first PCR reaction synthesized a megaprimer that codes for amino acid sequences from the N-terminus of the chimeric proteins, fused at the appropriate junction site, to a short segment that codes for amino acids in the C-terminal segment. The second PCR synthesized an expression plasmid for the chimeric protein (Figure S1).

Names for the chimeras are based on the enzyme from which the N-terminal sequence was derived, F for FPPase or C for CPPase, and the junction at which the sequence changes to the other enzyme. For instance, the "C6" chimera consists of CPPase sequence from the N-terminus to the sixth junction, with the remaining sequence from FPPase (see Figure 3). The precise locations of the helices were identified by comparisons of X-ray structures for FPPase and confirmed by the structure of chimera C3. Chimeras were constructed from FPPase and CPPase by fusing the proteins at 13 junctions to produce chimeric fusions at each helix/loop transition from amino acids 53 to 241, including the junction between amino acids 193 and 194 located on either side of a kink in helix VII. Two additional FPPase-CPPase chimeras constructed where the six-helix bundle consisting of sequences from amino acids 53 to 241 from one enzyme (core) were flanked with the corresponding N- and C-terminal sequences (scaffold) from the other. These core-scaffold chimeras, with cores from FPPase and CPPase, are named c-F-c and f-C-f, respectively.

2. Site-Directed Mutants. Six amino acids in the active site of FPPase (F92, A161, T194, F231, D235 and R343) differ from the corresponding residues in CPPase (I92, I161, G194, Y231, N235, and G343). In FPPase, F92 and A161 form part of the binding pocket for the hydrocarbon moieties in DMAPP and GPP, the polar amino acids are in contact with the diphosphate resides or the tightly bound Mg^{2+} ions, and F231 and D235 form part of the binding pocket for the hydrocarbon moiety in IPP. Substitution of A161 by I161 shortens the binding pocket for the hydrocarbon moiety of the allylic substrate in CPPase, decreasing the preference for binding GPP over DMAPP. Ten site-directed mutants of FPPase and CPPase were constructed where one, or a combination of, the active site residues in one enzyme were replaced with those from the other to give FPPase single mutants A161I, T194G, F231Y, D235N, double mutant F231Y/D235N (2mFPPase), triple mutant A161I/T194G/D235N (3mFPPase), quintuple mutant A161I/T194G/F231Y/D235N/R343G (5mFPPase) and sextuple mutant F92I/A161I/T194G/F231Y/D235N/R343G (6mFPPase); and CPPase single mutant N235D and triple mutant I161A/G194T/N235D (3mCPPase). All of the mutants contained an N-terminal His_6 tag to facilitate purification by Ni^{2+} affinity chromatography.

Product Studies. Products were determined for incubations with 0.5 mM IPP and 0.5 mM DMAPP (chain elongation conditions), which gave products from chain elongation, cyclopropanation, branching, and cyclobutanation reactions, and with 3 mM DMAPP (irregular conditions), which only gave products from cyclopropanation, branching, and cyclobutanation reactions. The structures and relative amounts of products were determined by hydrolysis of the diphosphates with alkaline phosphatase, extraction of the corresponding alcohols with methyl *tert*-butyl ether (MTBE), and analysis by GC on HP-5 or Chiral-HP GC columns and by GC/MS. Structures were established by comparing GC retention times and mass spectra with authentic samples (Figure 4).[6]

Under standard chain elongation conditions, FPPase gave *E,E*-FPP as the major product, accompanied by a smaller amounts of GPP and the *Z,E*-FPP (Table 1).[14] Since formation of FPP requires two molecules of IPP and equal concentrations

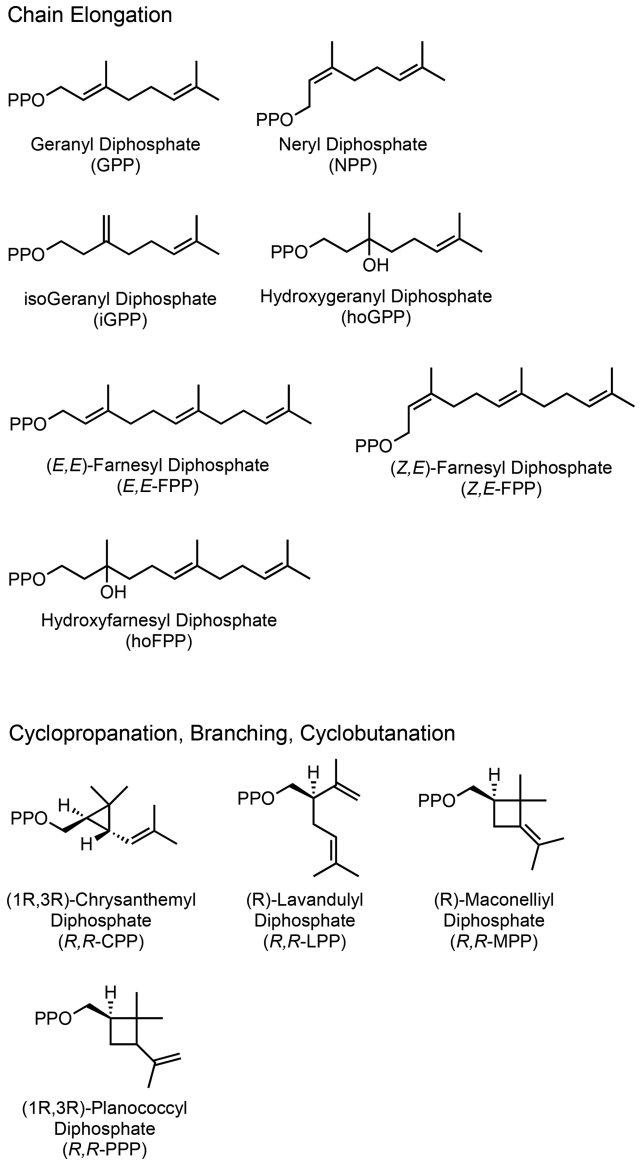


Figure 4. Products from FPPase, CPPase, and FPPase/CPPase chimeras and site-directed mutants.

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556–14567

Table 1. Relative Percentages of Products Formed by FPPase, CPPase, Chimeras and Site-Directed Mutants under Chain Elongation Incubation Conditions (0.5 mM DMAPP and IPP)[a]

	GPP	NPP	iGPP	hoGPP	*E,E*-FPP	*Z,E*-FPP	hoFPP	CPP	LPP
FPPase	8 ± 1	<1	nd[b]	nd	88 ± 1	4 ± 1	nd	nd	nd
C1	9 ± 2	<1	nd	nd	87 ± 2	4 ± 1	nd	nd	nd
C2	21 ± 1	<1	nd	nd	76 ± 1	3 ± 1	nd	nd	nd
C3	7 ± 1	<1	nd	nd	90 ± 1	2 ± 1	nd	nd	nd
C4	16 ± 1	<1	nd	nd	81 ± 1	3 ± 1	nd	nd	nd
C5	22 ± 3	1 ± 1	nd	nd	75 ± 3	2 ± 1	nd	nd	nd
C6	51 ± 6	2 ± 1	nd	nd	47 ± 6	<1	nd	nd	nd
C7	60 ± 2	2 ± 1	nd	nd	38 ± 2	<1	nd	nd	nd
C8	66 ± 2	2 ± 1	nd	nd	31 ± 2	nd	nd	nd	nd
C9	92 ± 1	3 ± 1	nd	nd	3 ± 2	nd	nd	nd	nd
C10	76 ± 2	2 ± 1	nd	nd	21 ± 2	nd	nd	nd	nd
C11	70 ± 4	2 ± 1	nd	nd	27 ± 4	nd	nd	nd	nd
C12	85 ± 1	3 ± 1	nd	nd	10 ± 1	nd	nd	nd	nd
C13	63 ± 1	2 ± 1	nd	nd	22 ± 1	nd	nd	3 ± 1	8 ± 1
A161I	21 ± 2	2 ± 1	nd	nd	70 ± 2	7 ± 1	nd	nd	nd
T194G	7 ± 1	<1	nd	nd	85 ± 2	8 ± 1	nd	nd	nd
F231Y	71 ± 2	<1	nd	nd	29 ± 1	<1	nd	nd	nd
D235N	9 ± 1	<1	nd	nd	84 ± 2	7 ± 1	nd	nd	nd
2m	71 ± 2	nd	nd	nd	29 ± 1	nd	nd	nd	nd
3m	62 ± 3	nd	nd	nd	29 ± 1	9 ± 1	nd	nd	nd
5m	69 ± 3	5 ± 1	nd	nd	24 ± 1	2 ± 1	nd	nd	nd
6m	93 ± 3	3 ± 1	nd	nd	4 ± 3	<1	nd	nd	nd
c-F-c	12 ± 3	nd	nd	nd	88 ± 3	nd	nd	nd	nd
CPPase	47 ± 1	3 ± 1	2 ± 1	nd	2 ± 1	nd	nd	35 ± 1	11 ± 1
F1	65 ± 1	2 ± 1	<1	nd	4 ± 1	nd	nd	18 ± 1	11 ± 1
F2	93 ± 1	2 ± 1	<1	nd	4 ± 1	nd	nd	nd	nd
F3	81 ± 6	2 ± 1	<1	nd	19 ± 6	<1	nd	nd	nd
F4	76 ± 2	2 ± 1	<1	nd	24 ± 2	<1	nd	nd	nd
F5	65 ± 2	1 ± 1	nd	nd	33 ± 2	1 ± 1	nd	nd	nd
F6	72 ± 3	1 ± 1	nd	nd	28 ± 3	<1	nd	nd	nd
F7	55 ± 6	1 ± 1	nd	nd	43 ± 4	1 ± 1	nd	nd	nd
F8	14 ± 2	<1	nd	nd	81 ± 2	3 ± 1	nd	nd	nd
F9	8 ± 2	nd	nd	nd	89 ± 2	3 ± 1	nd	nd	nd
F10	26 ± 1	<1	nd	nd	72 ± 1	2 ± 1	nd	nd	nd
F11	38 ± 4	<1	nd	nd	60 ± 4	2 ± 1	nd	nd	nd
F12	53 ± 4	1 ± 1	nd	nd	46 ± 4	<1	nd	nd	nd
F13	21 ± 5	<1	nd	nd	76 ± 5	3 ± 1	nd	nd	nd
N235D	63 ± 3	4 ± 1	7 ± 1	22 ± 1	nd	nd	nd	<1	4 ± 1
3m	70 ± 2	5 ± 1	6 ± 1	17 ± 1	2 ± 1	<1	<1	nd	nd
f-C-f	68 ± 2	2 ± 1	nd	nd	nd	nd	nd	1 ± 0.3	25 ± 1

[a]GPP, geranyl diphosphate; NPP, neryl diphosphate; iGPP, isogeranyl diphosphate; hoGPP, hydroxygeranyl diphosphate; *E,E*-FPP, *E,E*-farnesyl diphosphate; *Z,E*-FPP, *Z,E*-farnesyl diphosphate; hoFPP, hydroxyfarnesyl diphosphate; CPP, chrysanthemyl diphosphate; LPP, lavandulyl diphosphate. [b]nd, not detected.

of IPP and DMAPP were used in the incubations, at the point where IPP was exhausted, only ∼55% of the available DMAPP had been consumed to give an 11:1 ratio of *E,E*-FPP: GPP. The preponderance of *E,E*-FPP reflects substantial preference for utilization of GPP over DMAPP for chain elongation.[35]

As the N- to C-terminal metamorphorsis of FPPase into CPPase proceeded, FPP was the major product until the C6 chimera, where helix V was replaced and ∼ equal amounts of GPP and FPP were formed (Figure 5A). Small to trace amounts of NPP and *Z,E*-FPP were observed for the chimeras that produced substantial amounts of GPP and FPP, respectively. From the C6 chimera on, the amount of GPP was similar to or greater than FPP, reaching a maximum of 92% for C9. No irregular products were observed under chain elongation conditions until the metamorphosis had proceeded to C13, where all helices forming the active site bundle in FPPase had been replaced, at which point CPP and LPP constituted ∼11% (∼1:3 CPP:LPP) of the product mixture. Upon completion of the metamorphosis to CPPase, approximately equal amounts of chain elongation (mostly GPP) and irregular (∼3:1 CPP:LPP) products were seen.

The loss of activity for synthesis of irregular products was dramatic during the metamorphosis of CPPase to FPPase. Production of CPP and LPP plummeted upon replacement of helix III in F2 to give an enzyme that predominantly formed GPP. The remaining metamorphoses produced enzymes that mostly gave increasing amounts of FPP. Production of FPP increased steadily between F2 and F9, where it reached 90%, and decreased to 46% at F12 before rising again (Figure 5B).

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

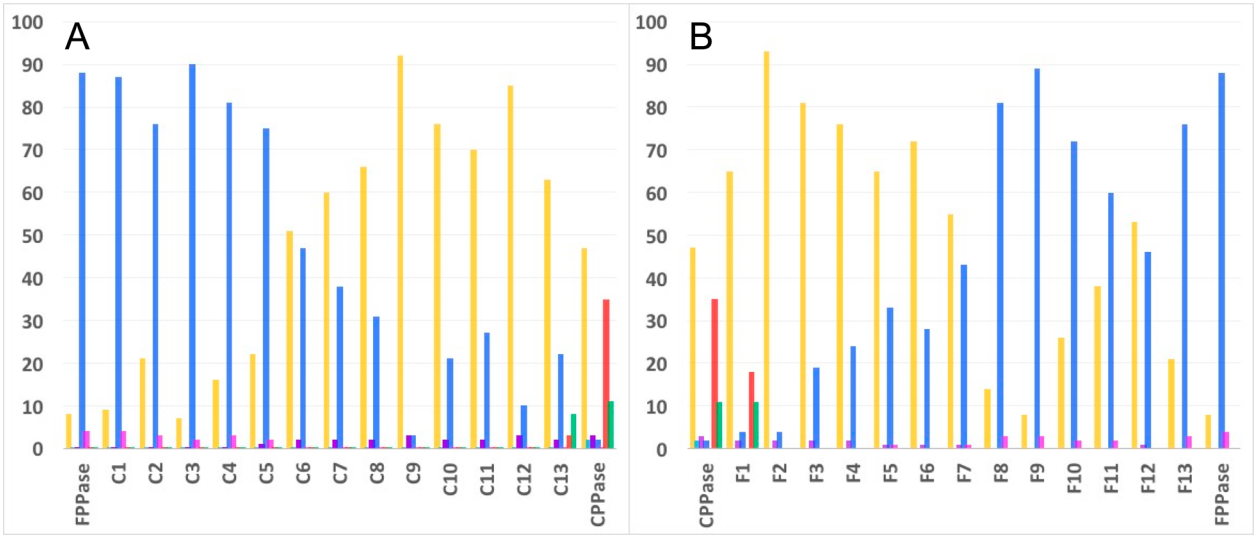


Figure 5. Products formed under chain elongation conditions during morphogenesis of FPPase to CPPase and of CPPase to FPPase. GPP, gold; *i*GPP, light blue; NPP, purple; *E,E*-FPP, dark blue; *Z,E*-FPP, pink; CPP, red; LPP, green. (A) FPPase to CPPase; (B) CPPase to FPPase.

The influence of the scaffolds surrounding the active-site helical bundles in FPPase and CPPase was evaluated with core-scaffold chimeras c-F-c and f-C-f. Synthesis of FPP was predominant for F13, C1, and the core-scaffold c-F-c chimera. In contrast, more diverse distributions were seen for C13, F1, and f-C-f. FPP, which constituted 23% of the product mixture for C13, was detected at 4% for F1 and not detected for the f-C-f chimera. Furthermore, the CPP:LPP ratios were ∼1:3 for C13, ∼2:1 for F1, and ∼1:25 for f-C-f. These core-scaffold chimeras highlight the ability of amino acid substitutions outside of the active site, and even outside the α terpenoid fold, to alter product distributions for the promiscuous CPPase while minimally effecting those of the more robust FPPase.

Single replacements of active site residues in FPPase with the corresponding amino acids in CPPase did not substantially alter the product distributions, with the exception of the F231Y mutation, which almost reversed the relative amounts of GPP and FPP. None of the FPPase mutants with multiple active site replacements gave irregular products in incubations with IPP and DMAPP. In sharp contrast, a single N235D mutation in CPPase substantially reduced the production of irregular products by reducing synthesis of CPP and FPP, and instead gave GPP as the major product, along with 30% of a 1:3 mixture of iGPP and hoGPP. Similar distributions of C_{10} products were seen for 3mCPPase, along with small amounts of both double bond isomers of FPP and hoFPP.

Only products with irregular carbon skeletons were formed when the enzymes were incubated with 3 mM DMAPP (irregular conditions). No products were seen for FPPase, C2, C9, F2–F7, F11–F13, and c-F-c. LPP (branching) and MPP (cyclobutanation) were the only products detected for chimeras C1, C3, C4, C8, C10, C11, and F10, as well as 5mFPPase and 6mFPPase (Table 2). The C5–C7 chimeras also produced trace amounts of planococcyl diphosphate (PPP), a double bond isomer of MPP. CPP was first observed as a minor product for C12 during the metamorphosis of FPPase to CPPase, and only became dominant for CPPase. C12, C13, F1, F8, F9, f-C-f, CPPase, the N235D CPPase mutant, and 3mCPPase synthesized compounds with all three of the irregular carbon skeletons.

Table 2. Relative Percentages of Products Formed by the Irregular Coupling of Two Molecules of DMAPP

	CPP	LPP	MPP	PPP
FPPase	nd	nd	nd	nd[a]
C1	nd	66 ± 8	34 ± 8	nd
C3	nd	63 ± 5	37 ± 5	nd
C4	nd	48 ± 1	52 ± 1	nd
C5	nd	47 ± 2	50 ± 1	3 ± 1
C6	nd	58 ± 1	39 ± 1	3 ± 1
C7	nd	57 ± 2	40 ± 2	3 ± 1
C8	nd	56 ± 4	44 ± 4	nd
C10	nd	57 ± 1	43 ± 1	nd
C11	nd	94 ± 1	6 ± 1	nd
C12	2 ± 0.3	95 ± 1	3 ± 1	nd
C13	27 ± 0.5	72 ± 1	1 ± 1	nd
5m	nd	63 ± 2	37 ± 2	nd
6m	nd	89 ± 2	11 ± 1	nd
CPPase	79 ± 1	20 ± 1	1 ± 1	nd
F1	64 ± 1	35 ± 1	1 ± 1	nd
F8	5 ± 1	86 ± 1	9 ± 1	nd
F9	4 ± 1	86 ± 1	10 ± 1	nd
F10	nd	83 ± 1	17 ± 1	nd
N235D	19 ± 1	79 ± 1	2 ± 1	nd
3m	2 ± 1	80 ± 1	17 ± 1	nd
f-C-f	16 ± 0.7	83 ± 1	1 ± 1	nd

[a]nd, not detected.

Kinetic Studies. Apparent Michaelis–Menten constants were determined from single-point measurements of initial rates for product formation. [β-^{32}P]IPP and [β-^{32}P]DMAPP were synthesized from the corresponding monophosphate esters by incubation with [γ-^{32}P]ATP and isopentenyl phosphate kinase.[36] After quenching, assay mixtures were spotted on silica TLC plates and the developed plates analyzed by phosphorimaging. Radioactivity only appeared in products from chain elongation for incubations with [β-^{32}P]IPP and DMAPP. For chain elongation, values for k_{cat} and K_M were obtained by fits to the equations for a sequential bisubstrate binding mechanism.[37] The equations for the coupling of two molecules of DMAPP have squared terms for the substrate in

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556–14567

Table 3. Kinetic Constants for Chain Elongation[a]

	k_{cat}^{DMAPP} (s^{-1})	k_{cat}^{GPP} (s^{-1})	K_M^{DMAPP} (μM)	K_M^{GPP} (μM)	K_M^{IPP} (μM)
FPPase	0.96 ± 0.09	2.6 ± 0.1	9 ± 2	1.1 ± 0.3	11 ± 3
C1	1.2 ± 0.1	1.5 ± 0.1	4 ± 1	3.5 ± 0.9	12 ± 5
C2	0.80 ± 0.07	0.81 ± 0.05	74 ± 6	38 ± 7	60 ± 20
C3	0.63 ± 0.06	0.22 ± 0.01	200 ± 20	36 ± 7	23 ± 7
C4	1.0 ± 0.1	0.26 ± 0.01	330 ± 50	26 ± 4	34 ± 9
C5	0.61 ± 0.04	0.23 ± 0.01	110 ± 20	16 ± 3	14 ± 3
C6	0.23 ± 0.02	0.42 ± 0.03	29 ± 3	25 ± 5	8 ± 2
C7	0.188 ± 0.008	0.147 ± 0.009	140 ± 20	240 ± 50	11 ± 2
C8	0.132 ± 0.006	(7.7 ± 0.3) × 10^{-3}	18 ± 2	680 ± 70	14 ± 2
C10	0.64 ± 0.03	0.010 ± 0.001	120 ± 20	4000 ± 800	59 ± 10
C11	(7.0 ± 0.6) × 10^{-2}	nd	2000 ± 300	nd	1300 ± 400
C12	0.110 ± 0.008	nd	2300 ± 300	nd	1400 ± 400
C13	(1.11 ± 0.03) × 10^{-2}	nd	350 ± 50	nd	1600 ± 400
c-F-c	(1.2 ± 0.3) × 10^{-2}	nd	290 ± 50	nd	80 ± 20
CPPase	(5.1 ± 0.2) × 10^{-3}	nd	510 ± 90	nd	1700 ± 200
F1	(1.25 ± 0.09) × 10^{-3}	(3.2 ± 0.1) × 10^{-4}	150 ± 30	6.1 ± 0.9	18 ± 2
F8	(2.1 ± 0.1) × 10^{-3}	3.2 × 10^{-3} (247 μM)[b]	1000 ± 300	nd	1000 ± 200
F9	(9.1 ± 0.4) × 10^{-3}	8.9 × 10^{-3} (247 μM)[b]	1000 ± 200	nd	2000 ± 300
F10	(6.7 ± 0.4) × 10^{-3}	(7.1 ± 0.4) × 10^{-3}	14 ± 4	7 ± 2	11 ± 2
F11	(2.7 ± 0.1) × 10^{-3}	1.7 × 10^{-3} (247 μM)[b]	1100 ± 300	nd	1600 ± 300
F12	(1.77 ± 0.06) × 10^{-3}	3.1 × 10^{-3} (82 μM)[b]	440 ± 80	nd	1200 ± 100
F13	(3.7 ± 0.2) × 10^{-3}	3.0 × 10^{-3} (247 μM)[b]	400 ± 70	nd	2800 ± 400
f-C-f	(8.7 ± 0.3) × 10^{-3}	9.7 × 10^{-3} (247 μM)[b]	530 ± 70	nd	3300 ± 300

[a][DMAPP] and [IPP] were optimized to give maximal values for k_{cat}. Apparent kinetic constants were determined at a fixed optimal [substrate 1] for varied [substrate 2]. nd, not determined. [b]Values for K_M^{GPP} were not determined because of low turnover, large differences in optimal concentrations of IPP and GPP and substrate inhibition at high [IPP]. k_{cat}^{GPP} is an apparent value at optimal [GPP] and 6 mM IPP. Where no values for K_M^{GPP} are listed, k_{cat} is an apparent value at the indicated [GPP].

the numerator and denominator as the result of DMAPP binding at both the donor and acceptor sites. Consequently, we could not obtain Michaelis constants for DMAPP binding at individual donor and acceptor sites by varying [DMAPP] and instead report K_{50}^{DMAPP}, the concentration of DMAPP that gives half-maximal velocity.

The kinetic constants for chain elongation are presented in Table 3. We were not able to measure values for chimeras C9 and F2−F7 because of a combination of low activities and poor levels of expression. The catalytic efficiency for chain elongation of GPP to FPP (CE^{GPP}, defined by k_{cat}^{GPP}/K_M^{GPP} = 2.4 × 10^6 s^{-1} M^{-1}) by sagebrush FPPase was similar to CE^{GPP} for the avian enzyme and approximately 100-fold slower than the diffusion controlled limit for substrate binding (Table 4).[35] CE^{DMAPP} for chain elongation of DMAPP to GPP was ∼20-fold lower. This difference favors synthesis of FPP and minimizes accumulation of GPP during chain elongation from DMAPP to FPP. The values for CE^{DMAPP} and CE^{GPP} dropped during morphogenesis of FPPase into CPPase and shifted from CE^{DMAPP} < CE^{GPP} to CE^{DMAPP} > CE^{GPP}, with a concomitant preference for synthesis of GPP. From FPPase to C3, CE^{DMAPP} decreased 31-fold, while CE^{GPP} decreased 393-fold. From C3 to C10, CE^{DMAPP} was similar, while CE^{GPP} continued to decline 10^3-fold, (10^6-fold relative to FPPase). Beginning at C11, CE^{DMAPP} dropped substantially, followed by smaller decreases to a CE^{DMAPP} for CPPase that was 10^4-fold less than for FPPase. Values for CE^{GPP} cannot be measured for these enzymes because of a combination of low k_{cat}s and high K_M^{DMAPP}s. During morphogenesis of CPPase to FPPase, activity for chain elongation remained low, except for a spike at F10 due to lower values for K_M^{DMAPP} and K_M^{GPP} (also observed at F1), for which we have no explanation.

Table 4. Catalytic Efficiencies for Chain Elongation

	$k_{cat}^{DMAPP}/K_M^{DMAPP}$ (s^{-1} M^{-1})	k_{cat}^{GPP}/K_M^{GPP} (s^{-1} M^{-1})
FPPase	1.0 × 10^5	2.4 × 10^6
C1	3.0 × 10^5	4.3 × 10^5
C2	1.1 × 10^4	2.1 × 10^4
C3	3.2 × 10^3	6.1 × 10^3
C4	3.0 × 10^3	1.0 × 10^4
C5	5.5 × 10^3	1.4 × 10^4
C6	7.9 × 10^3	1.7 × 10^4
C7	1.34 × 10^3	6.1 × 10^2
C8	7.3 × 10^3	11
C10	5.3 × 10^3	2.5
C11	35	–
C12	48	–
C13	32	–
f-C-f	41	–
CPPase	10	–
F1	8.3	52
F8	2.1	–
F9	9.1	–
F10	4.8 × 10^2	1 × 10^3
F11	2.5	–
F12	4	–
F13	9.6	–
c-F-c	16	–

The catalytic efficiencies for synthesis of CPP and LPP from two molecules of DMAPP ($k_{cat}^{DMAPP}/K_{50}^{DMAPP}$) and for synthesis of GPP from IPP and DMAPP ($k_{cat}^{DMAPP}/K_M^{DMAPP}$) by CPPase were ∼10 000-fold lower than for synthesis of GPP by FPPase (Tables 4 and 5). During morphogenesis of CPPase to FPPase, the rate for synthesis of the irregular products

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

Table 5. Kinetic Constants for Formation of Irregular Products

	k_{cat}^{DMAPP} (s^{-1})	K_{50}^{DMAPP} (μM)	k_{cat}/K_{50}^{DMAPP} (s^{-1} M^{-1})
C1	3.7×10^{-4}[b]	nd[a]	–
C3	6.9×10^{-4}[b]	nd	–
C4	5.0×10^{-4}[b]	nd	–
C5	1.1×10^{-3}[b]	nd	–
C6	6.6×10^{-4}[b]	nd	–
C7	9.5×10^{-4}[b]	nd	–
C8	4.0×10^{-4}[b]	nd	–
C10	9.6×10^{-4}[b]	nd	–
C11	8.6×10^{-4}[b]	nd	–
C12	5.9×10^{-4}[b]	nd	–
C13	$(1.1 \pm 0.05) \times 10^{-2}$	2100 ± 200	5.2
f-C-f	1.8×10^{-3}[b]	nd	–
CPPase	$(2.0 \pm 0.08) \times 10^{-2}$	4100 ± 400	4.8
F1	1.9×10^{-3}[b]	nd	–
F8	3.2×10^{-5}[b]	nd	–
F9	3.0×10^{-4}[b]	nd	–
F10	6.3×10^{-4}[b]	nd	–

[a]nd, not determined. [b]Measured at 6 mM DMAPP.

dropped 10-fold for the F1 chimera and 100- to 1000-fold for the remaining chimeras in the series. Irregular activity was not detected for FPPase. For the FPPase to CPPase series, a low level of irregular activity was first detected for C1 and remained low through C12, where k_{cat}^{DMAPP} was ~20−50-fold lower than that for wt CPPase. For these chimeras, catalytic efficiencies for the irregular couplings were substantially less than for chain elongation and the major irregular products from branching (LPP) and cyclobutanation (MPP) were typically not detected in competition with chain elongation. At C13, k_{cat}^{DMAPP} increased to a level similar to CPPase. However, LPP was the major irregular product and the catalytic efficiency for chain elongation was ~6-fold greater than that for irregular coupling. In comparison, CPP was the dominant irregular product for CPPase, and the catalytic efficiencies for irregular coupling and chain elongation were similar.

Structural Studies. Models for *A. tridentata* FPPase·IPP·DMAPP, C13·IPP·DMAPP, C13·IPP·DMAPP, CPPase·IPP·DMAPP, and CPPase·DMAPP·DMAPP were constructed from X-ray structures of *E. coli* FPPase liganded with IPP and an unreactive thiolo analogue of DMAPP[38,39] in a "closed" active conformation (Brookhaven PDB: 1RQI; 2.4 Å) and apo C3 (Brookhaven PDB: 4KK2; 2.2 Å). C3, which is the only protein in the *Artemisia* group whose crystal structure has been reported, was used as the base structure for the modeling, as described in Supporting Information.

Six active-site amino acids F92, A161, T194, F231, D235, and R343 in FPPase are replaced by I92, I161, G194, Y231, N235, and G343, respectively, in CPPase. The locations of substrates and amino acids at positions 193, 194, 231, 235, and 343 in the *A. tridentata* enzymes are shown in Figure 6 with IPP in the acceptor region of *A. tridentata* FPPase (part A) and IPP (part B) or DMAPP (part C) in the acceptor region of *A. tridentata* CPPase, including water molecules identified in the X-ray structure of *E. coli* FPPase.IPP.thioDMAPP. DMAPP is located in the donor regions of the active sites. In each case, the structures are oriented so carbon-2 in the substrate in the acceptor site is located directly behind carbon-3. This orientation highlights the differences among the conformations

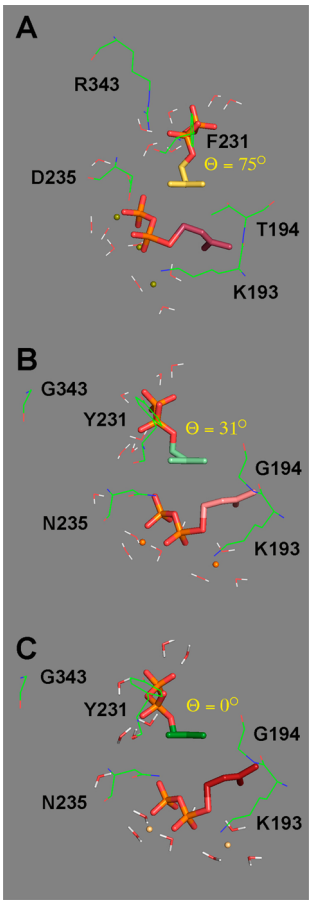


Figure 6. Active site structures for *A. tridentata* FPPase and CPPase. (A) FPPase·IPP·DMAPP: isopentenyl moiety (yellow), dimethylallyl moiety (raspberry), Mg^{2+} (olive). (B) CPPase·IPP·DMAPP: isopentenyl moiety (pale green), dimethylallyl moiety (salmon), Mg^{2+} (orange). (C) CPPase·DMAPP·DMAPP: dimethylallyl moiety (isopentenyl region, forest green), dimethylallyl moiety (firebrick red), Mg^{2+} (light orange).

of the hydrocarbon moieties of isopentenyl and dimethylallyl bound in the acceptor site. For IPP, the carbon atoms attached to carbon-3 are restricted to a common plane with carbon-1 located above the plane. In *A. tridentata* FPPase·IPP·DMAPP, the dihedral angle between carbon-1 and the methyl group (Θ) is 75°, which is similar to the dihedral angle seen in *E. coli* FPPase·IPP·DMAPP.[29] For *A. tridentata* CPPase·IPP·DMAPP, Θ = 31°. In contrast to IPP, all of the carbon atoms in the dimethylallyl moiety are restricted to a common plane (Θ = 0°). Corresponding active site structures for C13 are essentially identical to those for CPPase.

In FPPase, F231 is located perpendicular to the plane of C2, C3, C4, and the methyl group in IPP with its edge ~4 Å above the C3−C4 double bond in position to stabilize developing positive charge in the IPP moiety during the electrophilic alkylation reaction. One of the terminal R343 guanidinium nitrogens is located ~3.4 Å almost directly above the center of the F231 aromatic ring. The other terminal nitrogen is located 3.9 Å from one of the nonbonding P1 oxygen atoms in IPP. R343 seems to have two roles—to help immobilize the diphosphate group in IPP and to lock the aromatic ring of F231 over the double bond in IPP. D235 is part of the DDXXD motif that binds the diphosphate unit in DMAPP. The carboxylate group in D235 is coordinated with one of the

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

active site magnesium atoms that bind the diphosphate group in DMAPP and each of the carboxylate oxygens are about ~3.5 Å from the IPP methyl group. The side chain hydroxyl group in T194 is hydrogen bonded to the amide carbonyl of G228 and the T194 methyl group is 4.8 Å from the methylene carbon at C4 in IPP. The amino group in K193 helps bind the the diphosphate group in DMAPP in FPPase·IPP·DMAPP and CPPase·IPP·DMAPP. The nonconserved amino acids at positions 92 and 161 in FPPase and CPPase are located in a hydrophobic pocket that interacts with the terminal isoprene unit in GPP and do not make contact with DMAPP.

In CPPase, the plane of the aromatic ring in Y231 makes a ~50° angle with the plane of the C1−C4 carbons in IPP and its edge is 3.6 Å above the IPP methyl group. The diphosphate moiety in IPP is anchored to Y231 by a hydrogen bond between a nonbonding P1 oxygen and the hydroxyl group in the tyrosine side chain. The amide group in N235 is twisted relative to the position of the D235 carboxylate in FPPase and is located 3.1 Å from the IPP methyl group. The backbone chain of G194 is displaced toward IPP in CPPase relative to the T194 backbone, and is located 4.8 Å from the IPP C4 methylene carbon. Water molecules occupy similar locations in the active sites of FPPase and CPPase where they coordinate with the three magnesium atoms that bind to the diphosphate in DMAPP and form a shell around the diphosphate in IPP.

Figure 7 (part A) shows active site overlays for *A. tridentata* FPPase·IPP·DMAPP and *A. tridentata* CPPase·IPP·DMAPP

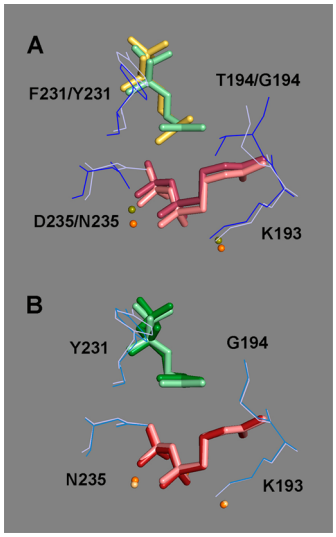


Figure 7. Overlays of active sites of *A. tridentata* FPPase and CPPase. (A) FPPase·IPP·DMAPP, IPP (yellow), DMAPP (raspberry), Mg^{2+} (olive), and amino acids (blue); CPPase·IPP·DMAPP, IPP (pale green), DMAPP (salmon), Mg^{2+} (orange), and amino acids (sky blue). (B) CPPase·IPP·DMAPP, IPP (pale green), DMAPP (salmon), Mg^{2+} (orange), and amino acids (sky blue); CPPase·DMAPP·DMAPP, DMAPP (acceptor, dark green; donor, firebrick red), Mg^{2+} (light orange), and amino acids (light blue).

and are arranged to superimpose the carbon-2-carbon-3 bonds of the substrates in the acceptor sites. The only substantial conformational changes seen between the active site amino acids in FPPase and CPPase relative to the substrates are at positions 194, 231, and 235. The side chains in F231 and Y231 are positioned to stabilize the tertiary carbocationic intermediate formed during the alkylation step (see Scheme 1). However, the Y231 aromatic ring in CPPase is displaced upward toward phosphorus-1 in IPP relative to the F231 phenyl group in FPPase. The IPP methyl group is located above the D235 carboxylate equidistant from both carboxylate oxygens. In contrast, the amide group in N235 side chain amide of CPPase is displaced upward to within hydrogen bonding distance of the non-bridging oxygen of P2 in DMAPP. These changes alter the topology of the binding pocket occupied by the methyl group in IPP and move C1 in IPP toward the plane of the four other carbon atoms in IPP. The T194 to G194 mutation allows the binding pocket to accommodate the methylene group as the dihedral angle in IPP decreases from 75° in FPPase to 31° in CPPase. As shown in Figure 7 (part B), CPPase can accommodate either IPP or DMAPP in the acceptor region of the active site by small changes in the conformation of the diphosphate moiety and the tilt of the aromatic ring in Y231.

DISCUSSION

The α terpenoid synthase fold, found in all members of CL0613, most likely evolved from a gene duplication-fusion event of a primordial 4-helix bundle protein.[10] Presumably an ancestral fusion protein, with hallmark DDXXD/DDXXD motifs on mirror image helices, evolved the ability to bind an allylic diphosphate and catalyze the electrophilic alkylation reactions required for chain elongation to synthesize the linear isoprenoid diphosphates (PF00348), which lie at the core of the isoprenoid biosynthetic pathway. Subsequent mutations provide a logical pathway for evolution of cyclopropanation/rearrangement enzymes with DTLED/DYKED motifs (PF00494), and cyclases with DDXXD/(N,D)D(L,I,V)X(S,T)-XXXE motifs (PF03936, PF01397) located at similar positions in the α-terpenoid synthase fold. Chain elongation is an essential activity, and enzymes belonging to PF00348 are found in all forms of life. The cyclopropanation-rearrangement enzymes in PF00494 are not universally distributed, but they are required for biosynthesis of sterols,[23] carotenoids,[18] and membrane-affiliated pentacyclic triterpenes[40] in Archaea, Bacteria, and Eukaryota.

Carbocations are remarkably indiscriminate in their cyclizations, rearrangements, and reactions with nucleophiles. As a group, the α terpenoid synthases, which catalyze electrophilic reactions through carbocationic intermediates, are highly promiscuous.[10] Many of the wild-type enzymes give multiple products and seemingly minor mutations in highly selective α terpenoid synthases often result in promiscuous behavior. This behavior has apparently been exploited during evolution of the isoprenoid pathway to produce the unusually large number of different carbon skeletons found among isoprenoid compounds.[1]

A. tridentata CPPase is an example where a parental FPPase, which is highly optimized to synthesize FPP from IPP and DMAPP, has evolved the capability to synthesize CPP, LPP, and much less efficiently, MPP, from DMAPP. In contrast to some highly selective terpenoid cyclases, where a single mutation can introduce significant promiscuous behavior,[10] the architecture of FPP synthase robustly supports chain elongation. Although trace amounts of irregular products were seen early during morphogenesis of FPPase to CPPase when DMAPP is the sole substrate, they were not detected in incubations with IPP and DMAPP, where chain elongation is a competitive reaction, until the process had reached the C13 chimera and all of the FPP helices surrounding the active site were replaced by the corresponding helices from CPP.

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

Likewise, no irregular products were formed in competition with chain elongation in the site-directed mutants. In contrast, formation of irregular products in competition with chain elongation ceased very early during morphogenesis of CPPase to FPPase.

Evolution of FPPase to CPPase involves changes in substrate binding to favor DMAPP over GPP in the donor site and permit DMAPP to compete with IPP for the acceptor site. The selectivity of the donor site in FPPase can be altered to favor binding DMAPP over GPP in a straightforward manner by reducing the size of the hydrophobic pocket.[41] In contrast, the ability to exclude DMAPP from the acceptor site is a highly conserved feature of chain elongation enzymes that is difficult to manipulate. A major factor that allows FPPase to discriminate between IPP and DMAPP appears to be related to differences in the conformational flexibility of the two molecules.[42] The hydrocarbon moiety in DMAPP is rigid, with all of its carbon atoms confined to a common plane, while rotation about the carbon-1-carbon-2 bond in IPP allows carbon-1 to move out of the plane. In the *A. tridentata* FPPase·IPP·DMAPP terinary complex, carbon-1 is displaced from the plane by ∼75° to give a conformation that cannot be duplicated by DMAPP. In contrast, the out of plane displacement for carbon-1 of IPP in *A. tridentata* CPPase·IPP·DMAPP complex is only 31°, and the bound conformations of IPP and DMAPP in the acceptor site of *A. tridentata* CPPase are similar (Figure 7B).

Six active site amino acids in FPPase, F92, A161, T194, F231, D235, and R343, are not conserved in CPPase. Site-directed FPPase mutants where one or more of these residues are replaced with the corresponding amino acids from CPPase catalyze chain elongation with no detectable amounts of irregular products in incubations with IPP and DMAPP, including the 6m mutant, where all six of the amino acids are replaced. Trace quantities of LPP and MPP are seen when 6m-FPPase is incubated with DMAPP, similar to the behavior observed in the C1−C12 chimeras.

Kinetic constants for FPPase, CPPase, and the mutant proteins reflect these differences in substrate binding at the donor and acceptor sites. For chain elongation, K_M^{GPP} increases >4000-fold during the morphogenesis of FPPase to CPPase, while K_M^{DMAPP} increases only ∼50 fold. A concomitant ∼170-fold increase in K_M^{IPP} results in a 10 000-fold decrease in catalytic efficiency (k_{cat}/K_M) for the chain elongation reaction. FPPase does not synthesize irregular products from DMAPP, even under forcing conditions. Very low levels of branching and cyclobutanation products are seen for incubations of the C1−C12 chimeras but are only detected when DMAPP is the sole substrate. At C13, where all of the loops and helices in the α terpenoid synthase fold in FPPase are replaced, the activity for synthesis of irregular products increases substantially and competes with chain elongation in incubations with IPP and DMAPP, with CPP and LPP constituting 11% of the products. The final morphogenesis of C13 to wild type *A. tridentata* CPPase results in small adjustments that increase the ratio of irregular to chain elongation products from ∼1:10 to ∼1:1 and the ratio of irregular products CPP to LPP from ∼1:3 to 3:1.

The amino acids at positions 194, located at the VIIa/VIIb kink in helix VII, and at positions 231 and 235 in helix VIII are in contact with the hydrocarbon moiety IPP in FPPase·IPP·DMAPP and CPPase·IPP·DMAPP. In FPPase, the IPP methyl group and the C4 methylene carbon are approximately equidistant (∼4 Å) from the edge of the aromatic ring of F231 and the methyl group is ∼3.5 Å above the plane of the D235 carboxylate (Figure 8A). Rotations away from the 75°

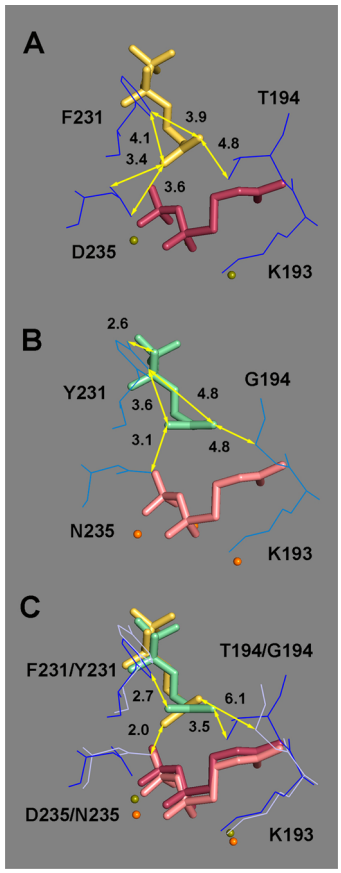


Figure 8. Interactions between C4 olefinic methylene group and the methyl group in IPP with amino acids at positions 231 and 235 in *A. tridentata* FPPase and CPPase. Same color-coding as Figure 6. (A) FPPase. (B) CPPase. (C) An overlay of FPPase and CPPase.

angle will generate unfavorable steric interactions with the edge of the aromatic ring. The positions of the Y231 aromatic ring and the N235 amide groups in CPPase·IPP·DMAPP are both "tilted" relative to their conformations in FPPase·IPP·DMAPP to reduce the twist in IPP to 31° (Figure 8B). A shift in the location of helix VIII in CPPase·IPP·DMAPP with respect to the rest of the six-helix bundle moves the binding pocket for the methyl group in IPP, defined by amino acids at positions 231 and 235, "upward" relative to their locations in FPPase. This motion relaxes the sterically induced twist about the C2−C3 bond in IPP from Θ from 75° in FPPase·IPP·DMAPP to 31° in CPPase·IPP·DMAPP (Figure 8C). The accompanying "downward" motion of the olefinic methylene group in IPP is accommodated by the space created by the mutation T194G.

In FPPase·IPP·DMAPP the distances between C1 of DMAPP and the C3−C4 olefinic carbons in IPP are 3.3 and 3.6 Å, respectively, and increase to 3.9 and 4.0 Å in CPPase·IPP·DMAPP. In CPPase·DMAPP·DMAPP, the distances between C1 of the donor DMAPP and the C2−C3 olefinic carbons in the acceptor DMAPP are 4.1 and 4.2 Å. These distances between electrophile and nucleophile in the FPPase and CPPase complexes are compatible with a dissociative electrophilic alkylation of the double bonds in IPP and DMAPP by the electrophilic dimethylallyl cation sandwiched between

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

the acceptor substrate and the inorganic pyrophosphate leaving group.

The conformations for FPPase·IPP·DMAPP, CPPase·IPP·DMAPP, and CPPase·DMAPP·DMAPP shown in Figure 8 indicate how the active site is remolded to accommodate DMAPP in the acceptor region. However, the minimum number of mutations needed to convert FPPase to a catalyst where CPP is a major product is not obvious. While substitutions of some active site residues are required for the remolding, it is apparent that amino acid substitutions outside of the active site are important. This is illustrated by replacing the scaffold around the six-helix bundle core of CPPase with the scaffold from FPPase, resulting in a ∼2-fold increase in LPP and a 35-fold decrease in CPP. The ability of seemingly small changes in the active sites of FPPase and CPPase, and even less apparent changes distal to the active site, to alter the course of the reactions they catalyze is a characteristic of many enzymes in the terpenoid synthase superfamily, especially those in the cyclase groups PF01397 and PF03936,[10] and is a consequence of the sensitivity of carbocations to their environments.

CONCLUSIONS

A. tridentata FPPase and CPPase are closely related proteins. FPPase catalyzes the chain elongation of DMAPP to FPP by sequential addition of two molecules of IPP, while CPPase catalyzes chain elongation of DMAPP to GPP along with condensation of two molecules of DMAPP to give irregular isoprenoid carbon skeletons with cyclopropane, branched, and cyclobutane structures. Stepwise replacement of the helices and loops surrounding the active site of FPPase by the corresponding regions from CPPase results in the gradual metamorphosis of the selective efficient chain elongation enzyme to a promiscuous relative that inefficiently catalyzes chain elongation and cyclopropanation/branching/cyclobutanation reactions. The initial replacements shift the preference for synthesis of FPP to GPP, accompanied by a moderate loss in catalytic efficiency. Formation of irregular products only becomes competitive with chain elongation upon replacement of helix VIII. Structural models suggest that the high selectivity for chain elongation by FPPase results because IPP binds to the acceptor region of the active site in a conformation that cannot be adopted by DMAPP. Synthesis of irregular products becomes competitive with chain elongation when the topology of the active site allows IPP and DMAPP to be bound in similar conformations.

A. tridentata CPPase is an example of an enzyme that has evolved recently from a highly specialized parent. The origins of FPPase date back to the very beginning of cellular life, and the enzyme has perfected its ability to catalyze chain-elongation. In contrast, *A. tridentata* CPPase has recently evolved from *A. tridentata* FPPase, presumably by gene duplication and random mutagenesis[43] but is still a promiscuous inefficient catalyst in comparison with FPPase. While the biosynthesis of irregular monoterpenes provides no obvious benefit to sagebrush, the insecticidal properties of pyrethrum synthesized from CPP by *C. cinerafolium* could provide continual evolutionary pressure for optimization of CPPase.[22] The approximately 80 000 occurring isoprenoid compounds found in nature, with numerous different carbon skeletons often formed from the same substrates, is a testament to the ability of enzyme evolution to harness the inherent diversity of carbocation chemistry.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b07608.

Experimental Section; Supporting Figures and Tables (PDF)

AUTHOR INFORMATION

Corresponding Author

*poulter@chemistry.utah.edu

ORCID

C. Dale Poulter: 0000-0001-7682-3095

Present Address

†Mucosal Inflammation Program, Department of Medicine, University of Colorado Anschutz Medical Campus, Aurora, Colorado 80045, United States.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institute of General Medical Sciences (GM 21328).

REFERENCES

(1) Buckingham, J.; Cooper, C. M.; Purchase, R. *Natural Products Reference Desk*; CRC Press, Taylor & Francis Group: Boca Raton, FL, 2016; p 235.

(2) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. *J. Am. Chem. Soc.* **2008**, *130*, 1966−1971.

(3) Tarshis, L. C.; Proteau, P. J.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15018−15023.

(4) Poulter, C. D. *Phytochem. Rev.* **2006**, *5* (1), 17−26.

(5) Christianson, D. W. *Curr. Opin. Chem. Biol.* **2008**, *12*, 141−150.

(6) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. *Science* **2007**, *316*, 73−76.

(7) EBML-EBI, http://pfam.xfam.org/family/polyprenyl_synt#tabview=tab2.

(8) Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871−10877.

(9) Cao, R.; Zhang, Y.; Mann, F. M.; Huang, C.; Mukkamala, D.; Hudock, M. P.; Mead, M. E.; Prisic, S.; Wang, K.; Lin, F.-Y.; Chang, T.-K.; Peters, R. J.; Oldfield, E. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 2417−2432.

(10) Christianson, D. W. *Chem. Rev.* **2017**, *117*, 11570−11648.

(11) Fujihashi, M.; Zhang, Y.-W.; Higuchi, Y.; Li, X.-Y.; Koyama, T.; Miki, K. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 4337−4342.

(12) Paterson-Jones, J. C.; Gilliland, M. G.; Van Staden, J. *J. Plant Physiol.* **1990**, *136*, 257−263.

(13) Poulter, C. D.; Rilling, H. C. *Acc. Chem. Res.* **1978**, *11*, 307−313.

(14) Thulasiram, H. V.; Poulter, C. D. *J. Am. Chem. Soc.* **2006**, *128*, 15819−15823.

(15) Resh, M. D. *Nat. Chem. Biol.* **2006**, *2*, 584−590.

(16) Poulter, C. D.; Marsh, L. L.; Hughes, J. M.; Argyle, J. C.; Satterwhite, D. M.; Goodfellow, R. J.; Moesinger, S. G. *J. Am. Chem. Soc.* **1977**, *99*, 3816−3823.

(17) Rilling, H. C.; Epstein, W. W. *J. Am. Chem. Soc.* **1969**, *91*, 1041−1042.

(18) Altman, L. J.; Ash, L.; Kowerski, R. C.; Epstein, W. W.; Larsen, B. R.; Rilling, H. C.; Muscio, F.; Gregonis, D. E. *J. Am. Chem. Soc.* **1972**, *94*, 3257−3259.

(19) Gunawardena, K.; Rivera, S. B.; Epstein, W. W. *Phytochemistry* **2002**, *59*, 197−203.

(20) Zhang, A.; Amalin, D.; Shirali, S.; Serrano, M. S.; Franqui, R. A.; Oliver, J. E.; Klun, J. A.; Aldrich, J. R.; Meyerdirk, D. E.; Lapointe, S. L. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9601−9606.

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567

(21) Bierl-Leonhardt, B. A.; Moreno, D. S.; Schwarz, M.; Fargerlund, J.; Plimmer, J. R. *Tetrahedron Lett.* **1981**, *22*, 389−392.
(22) Rivera, S. B.; Swedlund, B. D.; King, G. J.; Bell, R. N.; Hussey, C. E.; Shattuck-Eidens, D. M.; Wrobel, W. M.; Peiser, G. D.; Poulter, C. D. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 4373−4378.
(23) Epstein, W. W.; Rilling, H. C. *J. Biol. Chem.* **1970**, *245*, 4597−4605.
(24) Pandit, J.; Danley, D. E.; Schulte, G. K.; Mazzalupo, S.; Pauly, T. A.; Hayward, C. M.; Hamanaka, E. S.; Thompson, J. F.; Harwood, H. J. *J. Biol. Chem.* **2000**, *275*, 30610−30617.
(25) Pan, J. J.; Solbiati, J. O.; Ramamoorthy, G.; Hillerich, B. S.; Seidel, R. D.; Cronan, J. E.; Almo, S. C.; Poulter, C. D. *ACS Cent. Sci.* **2015**, *1*, 77−82.
(26) Iwata-Reuyl, D.; Math, S. K.; Desai, S. B.; Poulter, C. D. *Biochemistry* **2003**, *42*, 3359−3365.
(27) Demissie, Z. A.; Erland, L. A. E.; Rheault, M. R.; Mahmoud, S. S. *J. Biol. Chem.* **2013**, *288*, 6333−6341.
(28) Chayet, L.; Rojas, M. C.; Cori, O.; Bunton, C. A.; McKenzie, D. C. *Bioorg. Chem.* **1984**, *12*, 329−338.
(29) Hosfield, D. J.; Zhang, Y.; Dougan, D. R.; Broun, A.; Tari, L. W.; Swanson, R. V.; Finn, J. *J. Biol. Chem.* **2004**, *279*, 8526−8529.
(30) Hemmerlin, A.; Rivera, S. B.; Erickson, H. K.; Poulter, C. D. *J. Biol. Chem.* **2003**, *278*, 32132−32140.
(31) Bergthorsson, U.; Andersson, D. I.; Roth, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (43), 17004−17009.
(32) Chen, A.; Kroon, P. A.; Poulter, C. D. *Protein Sci.* **1994**, *3*, 600−607.
(33) Miyazaki, K.; Takenouchi, M. *Biotechniques* **2002**, *33*, 1036−1038.
(34) O'Maille, P. E.; Bakhtina, M.; Tsai, M.-D. *J. Mol. Biol.* **2002**, *321*, 677−691.
(35) Laskovics, F. M.; Poulter, C. D. *Biochemistry* **1981**, *20*, 1893−1901.
(36) Chen, M.; Poulter, C. D. *Biochemistry* **2010**, *49*, 207−217.
(37) Laskovics, F. M.; Krafcik, J. M.; Poulter, C. D. *J. Biol. Chem.* **1979**, *254*, 9458−9463.
(38) Phan, R. M.; Poulter, C. D. *J. Org. Chem.* **2001**, *66*, 6705−6710.
(39) Phan, R. M.; Poulter, C. D. *Org. Lett.* **2000**, *2*, 2287−2289.
(40) Prades, J.; Vögler, O.; Alemany, R.; Gomez-Florit, M.; Funari, S. S.; Ruiz-Gutiérrez, V.; Barceló, F. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808*, 752−760.
(41) Stanley Fernandez, S. M.; Kellogg, B. A.; Poulter, C. D. *Biochemistry* **2000**, *39*, 15316−15321.
(42) Rilling, H. *Pure Appl. Chem.* **1979**, *51*, 597−608.
(43) Näsvall, J.; Sun, L.; Roth, J. R.; Andersson, D. I. *Science* **2012**, *338*, 384−387.

DOI: 10.1021/jacs.7b07608
J. Am. Chem. Soc. 2017, 139, 14556−14567