



# HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2017 April 15.

Published in final edited form as:

ACS Chem Biol. 2016 April 15; 11(4): 889–899. doi:10.1021/acscchembio.5b00960.

## Structure and Function of Fusicoccadiene Synthase, a Hexameric Bifunctional Diterpene Synthase

Mengbin Chen<sup>†</sup>, Wayne K. W. Chou<sup>‡</sup>, Tomonobu Toyomasu<sup>¶</sup>, David E. Cane<sup>‡</sup>, and David W. Christianson<sup>†,§,\*</sup>

<sup>†</sup>Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania, 19104-6323, United States

<sup>‡</sup>Department of Chemistry, Brown University, Box H, Providence, Rhode Island, 02912, United States

<sup>¶</sup>Department of Bioresource Engineering, Faculty of Agriculture, Yamagata University, Wakaba-cho 1-23, Tsuruoka, Yamagata, Japan

<sup>§</sup>Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA 02138, United States

### Abstract

Fusicoccin A is a diterpene glucoside phytotoxin generated by the fungal pathogen *Phomopsis amygdali* that causes the plant disease constriction canker, first discovered in New Jersey peach orchards in the 1930's. Fusicoccin A is also an emerging new lead in cancer chemotherapy. The hydrocarbon precursor of fusicoccin A is the tricyclic diterpene fusicoccadiene, which is generated by a bifunctional terpenoid synthase. Here, we report X-ray crystal structures of the individual catalytic domains of fusicoccadiene synthase: the C-terminal domain is a chain elongation enzyme that generates geranylgeranyl diphosphate, and the N-terminal domain catalyzes the cyclization of geranylgeranyl diphosphate to form fusicoccadiene. Crystal structures of each domain complexed with bisphosphonate substrate analogues suggest that three metal ions and three positively charged amino acid side chains trigger substrate ionization in each active site. While *in vitro* incubations reveal that the cyclase domain can utilize farnesyl diphosphate and geranyl diphosphate as surrogate substrates, these shorter isoprenoid diphosphates are mainly converted into acyclic alcohol or hydrocarbon products. Gel filtration chromatography and analytical ultracentrifugation experiments indicate that full-length fusicoccadiene synthase adopts hexameric quaternary structure, and small-angle X-ray scattering data yield a well-defined molecular envelope illustrating a plausible model for hexamer assembly.

---

\*To whom correspondence should be addressed: Department of Chemistry, University of Pennsylvania, 2001 Roy and Diana Vagelos Laboratories, 231 South 34<sup>th</sup> Street, Philadelphia, PA, 19104. Tel: 215-898-5714; Fax: 215-573-2201; chris@sas.upenn.edu.

### Accession Codes

The atomic coordinates and crystallographic structure factors of unliganded PaFS<sub>CT</sub>, the PaFS<sub>CT</sub>-Co<sup>2+</sup><sub>3</sub>-pamidronate complex, the PaFS<sub>344</sub>-Mn<sup>2+</sup><sub>3</sub>-neridronate complex, and the PaFS<sub>344</sub>-Mg<sup>2+</sup><sub>3</sub>-pamidronate complex have been deposited in the Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)) with accession codes 5ERN, 5ERO, 5ER8, and 5ERM, respectively.

### Notes

The authors declare no competing financial interest.

### Supporting Information

Supporting Information Available: detailed methods, figures, and tables. This material is available free of charge via the Internet.

Terpenoids (also known as terpenes or isoprenoids) constitute the largest and most diverse group of natural products, with more than 75,000 discovered to date. They play vital roles in defense, regulation, and communication, and they have been utilized since times of antiquity as therapeutics, fragrances, and flavorings.<sup>1</sup> Strikingly, the chemical origins of such vast chemodiversity are rooted in just two simple C<sub>5</sub> precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Prenyl diphosphate synthases utilize one molecule of DMAPP and one or more molecules of IPP as substrates in head-to-tail chain elongation reactions to generate linear, achiral isoprenoid diphosphates such as C<sub>10</sub> geranyl diphosphate (GPP), C<sub>15</sub> farnesyl diphosphate (FPP), and C<sub>20</sub> geranylgeranyl diphosphate (GGPP).<sup>2</sup> In turn, terpenoid cyclases transform these linear isoprenoid diphosphates into myriad hydrocarbon products containing multiple rings and stereocenters.<sup>3–7</sup> Terpenoid cyclases vary in the number of products generated, with some enzymes generating a single cyclization product and other enzymes producing numerous cyclic products.<sup>8</sup>

Fusicoccin A is a C<sub>20</sub> diterpene glucoside containing an unusual tricyclic 5-8-5 ring system (Figure 1).<sup>9–11</sup> This phytotoxin is generated by the fungal pathogen *Phomopsis amygdali* and causes constriction canker, a plant disease first discovered in New Jersey peach orchards in the 1930's and subsequently identified in Georgia and Maryland orchards.<sup>12–14</sup> More recently, however, fusicoccin A has been explored in new approaches to cancer chemotherapy, since this novel diterpene exhibits promising antitumor properties.<sup>15</sup> Much of the biological activity of fusicoccin A derives from its ability to mediate protein-protein interactions involving the 14-3-3 protein.<sup>16–18</sup> Additionally, fusicoccin A blocks the proliferation and migration of glioblastoma cells by inhibiting focal adhesion kinase, which mediates the aggressive invasive phenotype of human gliomas.<sup>19</sup>

Fusicocadiene, the hydrocarbon precursor of fusicoccin A, is generated by a bifunctional diterpene synthase.<sup>20</sup> Unlike structurally characterized plant diterpene synthases that exhibit  $\alpha\beta\gamma$  domain architecture,<sup>21–23</sup> and bacterial diterpene synthases that adopt a single  $\alpha$  domain fold<sup>24,25</sup> ( $\alpha$ ,  $\beta$ , and  $\gamma$  domain nomenclature is outlined by Oldfield<sup>26</sup>), fusicocadiene synthase from *P. amygdali* (PaFS) is unique in that it is expected to adopt  $\alpha\alpha$  domain architecture: the C-terminal  $\alpha$  domain catalyzes the isoprenoid chain elongation reaction that generates GGPP, which is then cyclized by the N-terminal  $\alpha$  domain to form fusicocca-2,10(14)-diene (Figure 1). Divalent metal ions are essential for catalysis by PaFS, and the amino acid sequence contains signature metal binding motifs:<sup>27,28</sup> the C-terminal domain contains two aspartate-rich DDXX(D,N) motifs characteristic of a chain elongation enzyme, and the N-terminal domain contains one DDXXD motif and one NSE/DTE motif, characteristic of a cyclase. The overall sequence identity between the N- and C-terminal domains is 19%.

The purpose of the current study is to establish a structural context for understanding catalysis by PaFS, including an unusual transannular proton transfer proposed for the cyclization mechanism.<sup>29</sup> We report that PaFS is an allosteric hexamer that exhibits cooperativity in the cyclization of GGPP, and it can also utilize GPP and FPP as aberrant substrates. X-ray crystal structures of the individual catalytic domains allow us to construct a plausible model of the full-length PaFS hexamer based on small-angle X-ray scattering data.

## RESULTS AND DISCUSSION

PaFS was previously prepared and characterized as a GST fusion protein, which enabled the identification of chain elongation and cyclase domains.<sup>20</sup> Here, we report the preparation of a simpler PaFS construct bearing an N-terminal hexahistidine tag, which facilitates the preparation and purification of wild-type and mutant enzymes. Because we were unable to crystallize full-length PaFS, we hypothesized that potential flexibility in the 60-residue polypeptide linker might be hindering crystallization. Therefore, we prepared 3 individual domain constructs (details provided in the Supporting Information): (1) PaFS<sub>350</sub>, which contains residues 1-350 of the N-terminal cyclase domain, was used for activity measurements; (2) PaFS<sub>344</sub>, which contains residues 1-344 of the N-terminal cyclase domain, was used for crystallography (PaFS<sub>344</sub> and PaFS<sub>350</sub> exhibit comparable activity, but we were unable to crystallize PaFS<sub>350</sub>); and (3) PaFS<sub>CT</sub>, which contains residues 390-719 of the C-terminal GGPP synthase domain, was used for activity measurements and crystallography.

### Catalytic activity of full-length PaFS

Saturation kinetics are not observed for the generation of diterpene products from the C<sub>5</sub> substrates DMAPP and IPP by full-length PaFS. Interestingly, full-length PaFS exhibits approximately twice the activity at higher DMAPP concentrations compared with a reconstituted equimolar mixture of the full-length but partially inactivated mutants D92A PaFS and D474A PaFS, in which the cyclase and GGPP synthase domains, respectively, have been deactivated (Supplementary Information Figure S1). Covalent connection of the GGPP synthase and the GGPP cyclase domains affords a modest enhancement in the rate of formation of cyclic diterpenes, but this effect is not sufficiently large to suggest product channeling between active sites.

Full-length PaFS exhibits a sigmoidal dependence of reaction rate on substrate concentration for the direct cyclization of GGPP with a Hill coefficient of 2.2, with substrate inhibition being observed at higher GGPP concentrations (Figure 2a, Table 1). Thus, full-length PaFS is an allosteric enzyme for GGPP cyclization. When incubated with GGPP, full-length PaFS generates 64% fusicoccadiene, 9%  $\delta$ -araneosene, and one additional unidentified diterpene product.

In the absence of IPP, full-length PaFS can also solvolyze the shorter chain FPP and GPP as alternative substrates (Figure 2a) to yield predominantly acyclic products (Supplementary Information Figure S2): FPP is converted to farnesol (60.5%), nerolidol (14.0%), and farnesene (14.0%), while GPP is converted to a mixture of geraniol (59.5%) and linalool (35.0%). The binding of smaller C<sub>10</sub> or C<sub>15</sub> substrates in the active site of the N-terminal cyclase domain presumably allows binding of water in the vacant portion of the active site normally occupied by the olefinic tail of the C<sub>20</sub> GGPP substrate, thereby resulting primarily in formation of the derived acyclic allylic alcohols, with some competing deprotonation in the case of FPP. Finally, the cyclase-catalyzed solvolysis of FPP exhibits cooperativity, with a Hill coefficient of 3.0, while the reaction of GPP does not (Figure 2a).

### Catalytic activity of PaFS<sub>350</sub>

The N-terminal cyclase domain construct PaFS<sub>350</sub> is catalytically active, but the absence of the C-terminal domain presumably results in subtle structural changes in the N-terminal domain that impact substrate binding and increase the  $K_m$ , giving rise to moderate reductions in catalytic efficiency,  $k_{cat}/K_m$  (Table 1). The same trend is observed with PaFS<sub>350</sub> as with full-length PaFS, in that the catalytic efficiency ( $k_{cat}/K_m$ ) of the reaction with GGPP is lower than that with FPP. In contrast to the behavior of full-length PaFS, reactions catalyzed by PaFS<sub>350</sub> using substrates GPP, FPP, and GGPP exhibit classical Michaelis-Menten kinetics, with no evidence of cooperativity (Figure 2b). Therefore, the C-terminal domains are required for allosteric communication among N-terminal domain subunits of the full-length enzyme.

When incubated with GGPP, PaFS<sub>350</sub> generates a mixture of cyclic diterpenes consisting of 66% fusicoccadiene, 4%  $\delta$ -araneosene, and one unidentified diterpene product, comparable to the product array of full-length PaFS (Figure 2c). By contrast, when incubated with the corresponding monoterpene and sesquiterpene analogues, PaFS<sub>350</sub> converts GPP exclusively to a mixture of acyclic monoterpenes consisting predominantly of geraniol (64.5%) and linalool (28.5%), while FPP is converted to a mixture comprising mostly farnesene isomers (60.2%) and nerolidol (9.5%) (Supplementary Information Figure S2). Compared with full-length PaFS, the truncation mutant PaFS<sub>350</sub> essentially lost the ability to convert FPP into farnesol (1.0%), instead favoring deprotonation of the farnesyl cation to produce acyclic farnesene hydrocarbons. Subtle structural changes in the active site presumably caused by the domain truncation may lead to the observed alteration in product outcome. Alternatively, it is also possible that the C-terminal elongation domain in full-length PaFS makes some contribution to the catalytic solvolysis of FPP in the absence of IPP.

### Catalytic activity of PaFS<sub>CT</sub>

The measurement of steady-state kinetic parameters for the chain elongation reaction was complicated by the existence of an ambiguous mixture of oligomeric species detected for the C-terminal domain construct PaFS<sub>CT</sub>. Accordingly, we measured only the specific activity of PaFS<sub>CT</sub> (Table 1).

To measure the steady-state kinetic parameters in full-length PaFS for the chain elongation reaction catalyzed exclusively by the C-terminal domain, the N-terminal cyclization domain was inactivated by the D92A mutation, which abolished the binding of catalytically essential  $Mg^{2+}$  ions. As expected, the resultant D92A PaFS did not generate any cyclic products yet retained GGPP synthase activity. The chain elongation reaction in the D92A mutant using the 5-carbon substrates DMAPP and IPP was extremely sluggish, however, thereby preventing determination of the steady-state kinetic parameters. In contrast, GPP and FPP were processed with increasing efficiency when co-incubated with IPP to generate GGPP (Table 1). This result presumably reflects the imperative for GGPP biosynthesis by this processive enzyme, in that the product of one cycle of chain elongation generates a better substrate for the next cycle of chain elongation until the isoprenoid chain length reaches the maximum limit set by the depth of the active site cavity.

### Crystal structure of PaFS<sub>CT</sub>

Crystals of the unliganded and liganded GGPP synthase domain PaFS<sub>CT</sub> belong to two distinct space groups (Supporting Information Table S1). Regardless of the presence or absence of ligand, both crystal forms reveal a hexameric quaternary structure in the crystal lattice (Figure 3a). The hexamer can also be described as a trimer of isologous dimers with D<sub>3</sub> point group symmetry. Extensive surface area is buried between each monomer of the dimer (1,854 Å<sup>2</sup> per monomer), and between one dimer and the next (1,960 Å<sup>2</sup> per dimer), as calculated by PISA (<http://www.ebi.ac.uk/pdbe/pisa>). The active site of each monomer is oriented inward toward a central cavity approximately 22 Å in diameter. Helices B, E, and F of each monomer mediate dimer assembly, in similar fashion to that first observed in the avian FPP synthase dimer (Supplementary Information Figure S3).<sup>30</sup> With regard to dimer-dimer interactions in hexamer assembly, helices α2 and α3 of one dimer engage loops connecting helices D and E of one monomer and helices F and G of the other monomer in the neighboring dimer. While similar quaternary structure is observed for human GGPP synthase,<sup>31</sup> GGPP synthases from yeast and mustard crystallize simply as isologous dimers.<sup>32,33</sup>

The structure of the PaFS<sub>CT</sub> monomer clearly reveals the characteristic class I terpenoid synthase fold (Figure 3b). Like other class I terpenoid synthases,<sup>4</sup> PaFS<sub>CT</sub> undergoes a significant conformational change upon the binding of ligands. In unliganded PaFS<sub>CT</sub>, the α1 loop is disordered. This loop flanks the ~14 Å-deep active site cleft, the contour of which is defined mainly by nonpolar residues. In the liganded enzyme, the α1 loop becomes ordered and α-helices D, H, and α3 shift to fully enclose the active site. The substrate analogue pamidronate (Supplementary Information Figure S4) binds in the DMAPP binding site, and the bisphosphonate moiety of pamidronate mimics the diphosphate moiety of DMAPP (as inhibitors of human FPP synthase, bisphosphonates are formulated as drugs such as Fosamax to treat osteoporosis<sup>34</sup>). The bisphosphonate moiety of pamidronate coordinates to 3 Co<sup>2+</sup> ions, which are also coordinated by two aspartate-rich metal binding motifs, D<sup>474</sup>DFQD and D<sup>605</sup>DYQN on helices D and H, respectively (Figure 3c and Supplementary Information Figure S5a). Water molecules complete metal coordination polyhedra so that each metal ion is 6-coordinate with octahedral geometry. The bisphosphonate moiety is also stabilized by hydrogen bonds with R483, K619, and K629; an additional hydrogen bond is made with K561 in some monomers. Typically, diphosphate or bisphosphonate binding to terpenoid synthases is accommodated by 3 metal ions and 3 basic residues, and these interactions ensure complete active site closure.<sup>28</sup>

### Crystal structure of PaFS<sub>344</sub>

Crystallization of the GGPP cyclization domain PaFS<sub>344</sub> complexed with 3 Mn<sup>2+</sup> ions and the bisphosphonate substrate analogue neridronate (Supplementary Information Figure S4) was achieved by *in situ* proteolysis. Crystals diffracted X-rays to 2.5 Å resolution and were moderately twinned (22%). The initial electron density map was phased using the single-wavelength anomalous diffraction of the Mn<sup>2+</sup> ions. The refined structure of the PaFS<sub>344</sub>-Mn<sup>2+</sup><sub>3</sub>-neridronate complex (Supplementary Information Figures S5b and S6) then facilitated structure determination of the PaFS<sub>344</sub>-Mg<sup>2+</sup><sub>3</sub>-pamidronate complex (Figure 4 and Supplementary Information Figure S5c) from highly twinned crystals (41%) using

molecular replacement. Since  $Mg^{2+}$  is generally the metal ion required by terpenoid cyclases *in vivo*, this structure with its 3 bound  $Mg^{2+}$  ions is the more biologically relevant.

Two monomers oriented in antiparallel fashion occupy the asymmetric unit of PaFS<sub>344</sub> crystals, and these monomers interact through 583 Å<sup>2</sup> of contact surface area (calculated with PISA). This relatively minimal buried surface area suggests that PaFS<sub>344</sub> is a monomer in solution, consistent with the results of gel filtration chromatography. Like PaFS<sub>CT</sub>, the PaFS<sub>344</sub> monomer adopts the characteristic class I terpenoid synthase fold (Figure 4a). However, two extra segments in PaFS<sub>344</sub>, E2-T16 at the N-terminus and P328-L344 at the C-terminus, are fully ordered and distinguish PaFS<sub>344</sub> from PaFS<sub>CT</sub>. The active site of PaFS<sub>344</sub> is locked in a fully closed conformation through the binding of 3  $Mn^{2+}$  ions and neridronate (Supplementary Information Figures S5b and S6). The  $Mn^{2+}$  ions are also coordinated by the aspartate-rich metal binding motif (D<sup>92</sup>DVTD<sup>96</sup>) on helix D as well as the “NSE” motif (N<sup>232</sup>DIWS<sup>236</sup>WPKE<sup>240</sup>) on helix H. Water molecules complete metal coordination polyhedra so that each metal ion is 6-coordinate with octahedral or distorted octahedral geometry. Additionally, the bisphosphonate moiety receives hydrogen bonds from R188, K239, R325, and Y326. Thus, as observed in PaFS<sub>CT</sub>, 3 metal ions and 3 basic residues ensure the molecular recognition of the bisphosphonate moiety, and these interactions ensure complete active site closure.

The higher resolution structure (2.3 Å) of the PaFS<sub>344</sub>- $Mg^{2+}$ <sub>3</sub>-pamidronate complex determined from highly twinned crystals is generally identical to that of the PaFS<sub>344</sub>- $Mn^{2+}$ <sub>3</sub>-neridronate complex, with an r.m.s. deviation of 0.28 Å for 273 C $\alpha$  atoms. The  $Mg^{2+}$  coordination polyhedra in this structure (Figure 4b and Supplementary Information Figure S5c) are essentially identical to  $Mn^{2+}$  coordination polyhedra (Supplementary Information Figure S6), suggesting that these two metal ions may be interchangeable for function. Additionally, hydrogen bond interactions with the bisphosphonate groups of pamidronate and neridronate are essentially identical in both structures.

While the N- and C-terminal domains of PaFS exhibit 19% mutual amino acid sequence identity, it is interesting that PaFS<sub>344</sub> exhibits a higher sequence identity (24%) with the fungal sesquiterpene cyclase aristolochene synthase from *Penicillium roqueforti*. In fact, this is the highest sequence identity shared between PaFS<sub>344</sub> and any terpenoid synthase, the next highest being with aristolochene synthase from the orthologous *Aspergillus terreus*, with which it shares 22% sequence identity. A search of available terpenoid cyclase structures using Dali<sup>35</sup> indicates that PaFS<sub>344</sub> most closely resembles aristolochene synthase from *A. terreus*<sup>36,37</sup> ( $Z = 25.7$ , root-mean-square deviation (rmsd) = 2.5 Å for 257 C $\alpha$  atoms). Among diterpene cyclases, PaFS<sub>344</sub> most closely resembles the bacterial cyclase CotB2 ( $Z = 15.7$ , rmsd = 3.3 Å for 221 C $\alpha$  atoms),<sup>24</sup> although the amino acid sequence identity between these two proteins is only 14%. PaFS<sub>344</sub> exhibits comparable structural similarity with the  $\alpha$  domain of taxadiene synthase from the Pacific yew ( $Z = 15.6$ , rmsd = 3.3 Å for 236 C $\alpha$  atoms) as well as fungal *ent*-kaurene synthase ( $Z = 15.6$ , rmsd = 3.2 Å for 217 C $\alpha$  atoms).<sup>21,25</sup> Therefore, the diterpene cyclase domain of PaFS is more similar to sesquiterpene cyclases than to diterpene cyclases based on both sequence and structural comparisons.



A feature that distinguishes PaFS<sub>344</sub> from other terpenoid cyclases is the T58–F65 loop, which mediates a direct connection between the cyclase active site and the polypeptide linker segment D329–S389 that connects the C-terminal GGPP synthase domain and the N-terminal cyclase domain in full-length PaFS. Main chain atoms of the G60–P61 segment at the tip of this loop make hydrogen bonds with the side chains of N333 and Q336 in the D329–L344 fragment of the linker segment (D329–S389) present in the PaFS<sub>344</sub> construct (Figure 4c). Additionally, the main chain NH group of G60 hydrogen bonds with S64, which is immediately adjacent to F65, the ring face of which partially defines the active site contour. Given that the linker segment connecting the catalytic domains in full-length PaFS is about 60 residues long, it is possible that in addition to direct domain-domain interactions, the linker could transmit interdomain as well as intrasubunit conformational changes between active sites to enable cooperativity in catalysis. The functional importance of interdomain linkers in this regard has been reviewed by Ma and colleagues.<sup>38</sup>

Highly reactive carbocation intermediates in the cyclization cascade require not only a fully enclosed active site for protection from bulk solvent, but they also require a chemically inert active site capable of stabilizing these high-energy intermediates and their intervening transition states. To better understand how the cyclization domain of PaFS enables this chemistry, a model of the product fusicoccadiene docked in the active site of PaFS<sub>344</sub> was generated (Figures 5a,b). The ring faces of aromatic residues F65, F89, and W197 appear to be ideally positioned to stabilize proposed<sup>29</sup> carbocation centers through cation- $\pi$  interactions. Additionally, the structure of PaFS<sub>344</sub> reveals a characteristic break in helix G at the V192-G193 segment with the C-terminal end of helix G1 oriented toward the active site. The best fit of fusicoccadiene in the active site is such that sites of developing positive charge on the face of the macrocycle are oriented toward the backbone carbonyl of V192, consistent with possible carbocation stabilization by the helix G1 macrodipole.<sup>39</sup>

Given the general lack of acidic or basic residues in the PaFS<sub>344</sub> active site cavity, it is possible that inorganic pyrophosphate (PP<sub>i</sub>) serves as a general base to quench the final carbocation intermediate to yield fusicoccadiene (Figure 2c). A similar role for the PP<sub>i</sub> co-product was first considered in the aristolochene synthase mechanism<sup>40</sup> and subsequently in the FPP synthase mechanism.<sup>41</sup> As docked in the active site of PaFS<sub>344</sub>, the correct end of the 5-8-5 carbon skeleton is oriented toward the PP<sub>i</sub>-Mg<sup>2+</sup><sub>3</sub> cluster so as to enable a final PP<sub>i</sub>-mediated deprotonation to generate fusicoccadiene (Figures 5a,b). The PP<sub>i</sub> anion could similarly mediate the final deprotonation that generates the minor product  $\delta$ -araneosene.

### Hexameric quaternary structure of PaFS

Gel filtration chromatography initially indicated that full-length PaFS is an oligomer with either 5 or 6 subunits; analytical ultracentrifugation experiments confirmed that full-length PaFS is a 502-kD hexamer (Supplementary Information Figure S7). Therefore, small-angle X-ray scattering (SAXS) was used to view the molecular envelope and determine the molecular dimensions of PaFS in solution. The radius of gyration ( $R_g$ ) from Guinier analysis of low-angle scattering data is 53.7 Å, and the pair-distance distribution function  $P(r)$  yields an  $R_g$  value of 52.5 Å and a maximum dimension ( $D_{max}$ ) value of 155.4 Å (Supplementary

Information Figure S8). These dimensions are consistent with the assembly of the full-length protein as a hexamer.

Upon initial inspection of the *ab initio* molecular envelope of PaFS calculated from SAXS data (Figure 6), it was evident that the hexameric architecture of PaFS<sub>CT</sub> docked with 3 PaFS<sub>344</sub> dimers would not fit satisfactorily within this envelope. Indeed, our best attempt at manual fitting in this regard yielded  $\chi = 7.4$  ( $\chi$  is a measure of the goodness of fit of the protein model to the X-ray scattering data; a lower  $\chi$  value indicates a better fit). We subsequently hypothesized that the C-terminal domains observed in the crystal structure of PaFS<sub>CT</sub> could dissociate into 3 dimers. Indeed, dimeric quaternary structure is observed for GGPP synthases from yeast<sup>32</sup> and mustard;<sup>33</sup> the PaFS dimer (Supplementary Information Figure S3) is identical to these. Accordingly, to optimize the fit of PaFS domains into the molecular envelope, we utilized the program SASREF<sup>42</sup> to allow the 3 PaFS<sub>CT</sub> dimers of the crystallographic hexamer to separate by 5 Å and reorient by 15°. Subsequently, 3 PaFS<sub>344</sub> crystallographic dimers were docked at the periphery, all while maintaining overall three-fold symmetry. The resulting model yielded a satisfactory fit to the molecular envelope with  $\chi = 1.5$  (Figure 6). The C-terminus of PaFS<sub>344</sub> and the N-terminus of PaFS<sub>CT</sub> are approximately 25 Å apart, suggesting that the flexible 60-residue long linker connecting the C-terminal and N-terminal domains may reside between the PaFS<sub>CT</sub> and PaFS<sub>344</sub> dimers. This would facilitate the possible function of the linker in mediating interdomain communication required for allosteric catalysis.

### Implications for bifunctional catalysis

The catalytic efficiencies for the conversion of GGPP and FPP into hydrocarbon products by full-length PaFS in the absence of IPP are almost equal, which is curious in view of the fact that until now PaFS was believed to be solely a diterpene cyclase.<sup>20</sup> However, if the intracellular and/or local concentrations of GGPP exceed those of FPP, then aberrant processing of FPP would not be an issue. It is certainly possible that the coordinate expression and covalent assembly of the GGPP synthase and cyclase domains ensure a higher local concentration of GGPP so that the cyclization of GGPP by the N-terminal domain is favored. When full-length PaFS is incubated with DMAPP and IPP, diterpene products are exclusively generated. Thus, intermediates GPP and FPP do not dissociate from the C-terminal domain for aberrant reactions with the N-terminal domain.

Interestingly, when full-length PaFS is incubated solely with a 1:1 ratio of FPP and IPP, the ratio of sesquiterpene to diterpene products generated is 57:43, while with an FPP:IPP ratio of 1:8, the sesquiterpene:diterpene product ratio decreases slightly to 35:65, indicating that the amount of GGPP produced and cyclized is increased compared with the direct, anomalous cyclization of FPP. Even so, since FPP is not released from the C-terminal domain during GGPP biosynthesis from DMAPP and IPP, local concentrations of FPP should remain well below the  $K_m$  for FPP utilization by the N-terminal domain, thereby ensuring that the anomalous processing of FPP by the cyclase domain will be minimal *in vivo*.

While there does not appear to be strong channeling from the C-terminal domain to the N-terminal domain, the generation of diterpene products by full-length PaFS is roughly twice



as fast as that observed for an equimolar mixture of two PaFS mutants bearing inactivated chain elongation and cyclase domains, respectively (Supplementary Information Figure S1). This is consistent with the similarly modest enhancement of the rate of sesquiterpene product formation by a bifunctional fusion protein engineered between farnesyl diphosphate synthase and 5-*epi*-aristolochene synthase, perhaps reflecting a modestly advantageous proximity effect influencing the transient local concentration of the acyclic intermediate.<sup>43</sup> This modest enhancement could also reflect the possibility of an agglomeration or enzyme-clustering effect<sup>44</sup> resulting in part from the hexameric quaternary structure of full-length PaFS.

Notable, too, is the fact that PaFS<sub>350</sub> is most closely related to the fungal sesquiterpene cyclase aristolochene synthase (amino acid sequence identities of 24% and 22%, respectively, with the enzymes from *P. roqueforti* and *A. terreus*). In comparison, the amino acid sequence identity between the N-terminal and C-terminal domains of PaFS is 19%. Thus, the evolutionary origins of the  $\alpha\alpha$  domain architecture of PaFS are not clear, particularly in view of the catalytic versatility of the cyclase domain with both C<sub>15</sub> and C<sub>20</sub> substrates. While it is tempting to speculate that  $\alpha\alpha$  domain architecture arose from a gene duplication and fusion event involving a primordial diterpene synthase, it is also possible that the cyclization domain of PaFS evolved from an ancestral sesquiterpene synthase, developing the ability to utilize GGPP as well as FPP subsequent to fusion with a GGPP synthase.

Finally, it is instructive to compare PaFS with the bifunctional sesterterpene cyclase ophiobolin F synthase from *Aspergillus clavatus* (AcOS), which generates tricyclic products from C<sub>25</sub> geranylarnesyl diphosphate, each containing a common 5-8-5 tricyclic ring structure similar to that of fusicoccadiene.<sup>45</sup> The cyclization domains of AcOS and PaFS share 42% amino acid sequence identity, and there is significant conservation of the residues that contribute to the active site contour in each cyclase. Although most paired residues are isosteric or strictly conserved, two are smaller: L217 and A220 of AcOS correspond to W225 and V228 of PaFS, respectively. As a result, the active site of AcOS is slightly larger, so that it can accommodate the isoprenoid tail of a C<sub>25</sub> substrate as distinguished from a C<sub>20</sub> substrate. However, the active site contours are otherwise highly similar (Figure 5c), so as to accommodate and enforce the cyclization cascade leading to the formation of the 5-8-5 tricyclic product. Future experiments will probe the role of this contour in PaFS as a specialized template for the generation of tricyclic 5-8-5 hydrocarbon products.

## METHODS

The PaFS gene was cloned from a plasmid encoding full-length PaFS.<sup>20</sup> The full-length construct, plus the single-domain constructs PaFS<sub>CT</sub>, PaFS<sub>350</sub>, and PaFS<sub>344</sub>, were each prepared using a pET-28a vector (Novagen) with an N-terminal His-tag, expressed using *Escherichia coli* BL21(DE3)RIL-CodonPlus (Novagen), and purified as described in the Supporting Information.

Steady-state kinetics were measured for full-length PaFS and single-domain constructs with GPP, FPP, and GGPP, and product arrays were evaluated using gas chromatography-mass

spectrometry using previously described methods.<sup>46</sup> Sedimentation equilibrium experiments to determine the quaternary structure of full-length PaFS are outlined in the Supporting Information.

Crystals of unliganded PaFS<sub>CT</sub>, the PaFS<sub>CT</sub>-Co<sup>2+</sup><sub>3</sub>-pamidronate complex, the PaFS<sub>344</sub>-Mn<sup>2+</sup><sub>3</sub>-neridronate complex, and the PaFS<sub>344</sub>-Mg<sup>2+</sup><sub>3</sub>-pamidronate complex were grown by the sitting-drop method at 4 °C; following X-ray diffraction data collection at the synchrotron, initial electron density maps were phased by molecular replacement or by single wavelength anomalous diffraction of manganese in the PaFS<sub>344</sub>-Mn<sup>2+</sup><sub>3</sub>-neridronate complex as described in the Supporting Information. Each structure was satisfactorily refined using PHENIX.<sup>47</sup> Data collection and refinement statistics are recorded in Supplementary Information Table S1.

SAXS measurements were performed at the SIBYLS beamline at the Advanced Light Source using full-length N333A/Q336A PaFS. To reconstruct the *ab initio* molecular envelope of PaFS, the algorithm described by Svergun and colleagues was used as implemented in the program GASBOR.<sup>48</sup> Complete experimental details are outlined in the Supporting Information, and SAXS statistics are recorded in Supplementary Information Table S2.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Funding

We thank the NIH for grants GM56838 (D.W.C.) and GM30301 (D.E.C.) in support of this research. D.W.C. thanks the Radcliffe Institute for Advance Study for the Elizabeth S. and Richard M. Cashin Fellowship.

We are grateful to R. Marmorstein at the Perelman School of Medicine, University of Pennsylvania, for access to dynamic light scattering and analytical ultracentrifugation equipment. Additionally, we thank the National Synchrotron Light Source, Beamline X-29 (Brookhaven National Laboratory), and the Stanford Synchrotron Radiation Lightsource, Beamline 14-1, for access to X-ray crystallographic data collection facilities, and to the Advanced Light Source, SIBYLS Beamline 12.3.1, for access to SAXS data collection facilities. We thank V. Stojanoff at NSLS and S. Russi at SSRL for their assistance during data collection.

## ABBREVIATIONS

<b>AcOS</b>	ophiobolin F synthase from <i>Aspergillus clavatus</i>
<b>DMAPP</b>	dimethylallyl diphosphate
<b>EDTA</b>	ethylenediaminetetraacetic acid
<b>FPP</b>	farnesyl diphosphate
<b>GC-MS</b>	gas chromatography-mass spectrometry
<b>GPP</b>	geranyl diphosphate
<b>GGPP</b>	geranylgeranyl diphosphate

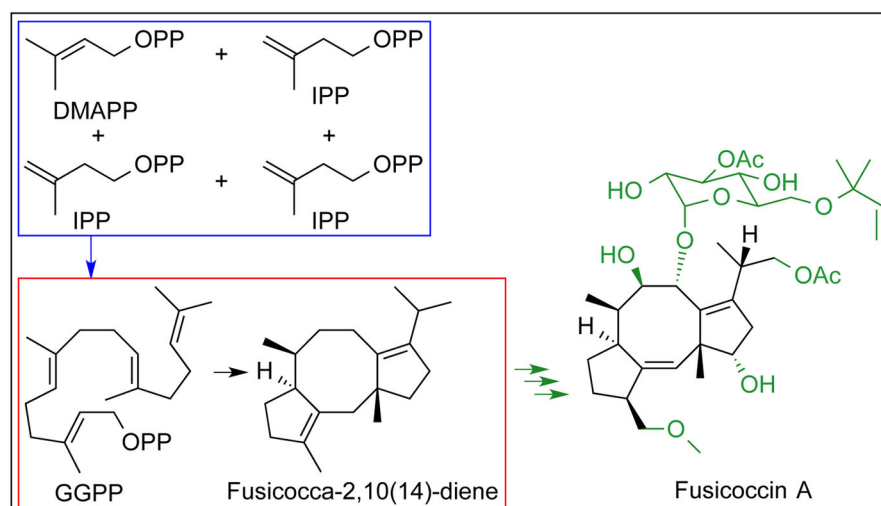
<b>IPP</b>	isopentenyl diphosphate
<b>PaFS</b>	fusicoccadiene synthase from <i>P. amygdali</i>
<b>PaFS<sub>344</sub></b>	fusicoccadiene synthase construct containing N-terminal residues 1-344
<b>PaFS<sub>350</sub></b>	fusicoccadiene synthase construct containing N-terminal residues 1-350
<b>PaFS<sub>CT</sub></b>	fusicoccadiene synthase construct containing C-terminal residues 390-719
<b>PP<sub>i</sub></b>	inorganic pyrophosphate
<b>rmsd</b>	root-mean-square deviation
<b>SAXS</b>	small-angle X-ray scattering
<b>TCEP</b>	tris(2-carboxyethyl)phosphine hydrochloride

## References

- Gershenson J, Dudareva N. The function of terpene natural products in the natural world. *Nat Chem Biol.* 2008; 3:408–414. [PubMed: 17576428]
- Poulter CD, Rilling HC. The prenyl transfer reaction. Enzymic and mechanistic studies of the 1'-4 coupling reaction in the terpene biosynthetic pathway. *Acc Chem Res.* 1978; 11:307–313.
- Davis, EM.; Croteau, R. *Biosynthesis: Aromatic polyketides, isoprenoids, alkaloids.* Vol. 209. Berlin: Springer-Verlag Berlin; 2000. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes; p. 53-95.
- Christianson DW. Structural biology and chemistry of the terpenoid cyclases. *Chem Rev.* 2006; 106:3412–3442. [PubMed: 16895335]
- Christianson DW. Unearthing the roots of the terpenome. *Curr Opin Chem Biol.* 2008; 12:141–150. [PubMed: 18249199]
- Cane DE. Isoprenoid biosynthesis. Stereochemistry of the cyclization of allylic pyrophosphates. *Acc Chem Res.* 1985; 18:220–226.
- Cane DE. Enzymatic formation of sesquiterpenes. *Chem Rev.* 1990; 90:1089–1103.
- Fischbach MA, Clardy J. One pathway, many products. *Nat Chem Biol.* 2007; 3:353–355. [PubMed: 17576415]
- Ballio A, Chain EB, de Leo P, Erlanger BF, Mauri M, Tonolo A. Fusicoccin: a new wilting toxin produced by *Fusicoccum amygdali* Del. *Nature.* 1964; 203:297–297.
- Ballio A, Brufani M, Casinovi CG, Cerrini S, Fedeli W, Pellicciari R, Santurbano B, Vaciano A. The structure of fusicoccin A. *Experientia.* 1968; 24:631–635. [PubMed: 5697762]
- Barrow KD, Barton DHR, Chain EB, Ohnsorge UFW, Thomas R. The constitution of fusicoccin. *Chem Commun.* 1968:1198–1200.
- Daines RH, Cohoon DF, Leone I, Brennan E. Control of *Fusicoccum* canker of peach by nutrition, defoliation, and protective fungicides. *Phytopathology.* 1958; 48:400–407.
- Weaver LO. The constriction disease of peach in Maryland. *Plant Dis Rep.* 1951; 35:144.
- Lalancette N, Polk DF. Estimating yield and economic loss from constriction canker of peach. *Plant Dis.* 2000; 84:941–946.
- de Vries-van Leeuwen IJ, Kortekaas-Thijssen C, Nzigou Mandouckou JA, Kas S, Evidente A, de Boer AH. Fusicoccin-A selectively induces apoptosis in tumor cells after interferon- $\alpha$  priming. *Cancer Lett.* 2010; 293:198–206. [PubMed: 20153922]
- Bunney TD, De Boer AH, Levin M. Fusicoccin signaling reveals 14-3-3 protein function as a novel step in left-right patterning during amphibian embryogenesis. *Development.* 2003; 130:4847–4858. [PubMed: 12930777]

17. de Boer AH, de Vries-van Leeuwen IJ. Fusicoccanes: diterpenes with surprising biological functions. *Trends in Plant Sci.* 2012; 17:360–368. [PubMed: 22465041]
18. Milroy L-G, Brunsveld L, Ottmann C. Stabilization and inhibition of protein-protein interactions: the 14-3-3 case study. *ACS Chem Biol.* 2013; 8:27–35. [PubMed: 23210482]
19. Bury M, Andolfi A, Rogister B, Cimmino A, Mégalizzi V, Mathieu V, Feron O, Evidente A, Kiss R. Fusicoccin A, a phytotoxic carbocyclic diterpene glucoside of fungal origin, reduces proliferation and invasion of glioblastoma cells by targeting multiple tyrosine kinases. *Transl Oncol.* 2013; 6:112–123. [PubMed: 23544164]
20. Toyomasu T, Tsukahara M, Kaneko A, Niida R, Mitsuhashi W, Dairi T, Kato N, Sassa T. Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. *Proc Natl Acad Sci U S A.* 2007; 104:3084–3088. [PubMed: 17360612]
21. Koksall M, Jin Y, Coates RM, Croteau R, Christianson DW. Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature.* 2011; 469:116–120. [PubMed: 21160477]
22. Köksal M, Hu H, Coates RM, Peters RJ, Christianson DW. Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. *Nat Chem Biol.* 2011; 7:431–433. [PubMed: 21602811]
23. Zhou K, Gao Y, Hoy JA, Mann FM, Honzatko RB, Peters RJ. Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis*. *J Biol Chem.* 2012; 287:6840–6850. [PubMed: 22219188]
24. Janke R, Gorner C, Hirte M, Bruck T, Loll B. The first structure of a bacterial diterpene cyclase: CotB2. *Acta Cryst.* 2014; D70:1528–1537.
25. Liu W, Feng X, Zheng Y, Huang C-H, Nakano C, Hoshino T, Bogue S, Ko T-P, Chen C-C, Cui Y, Li J, Wang I, Hsu S-TD, Oldfield E, Guo R-T. Structure, function and inhibition of *ent*-kaurene synthase from *Bradyrhizobium japonicum*. *Sci Rep.* 2014; 4:6214. [PubMed: 25269599]
26. Cao R, Zhang Y, Mann FM, Huang C, Mukkamala D, Hudock MP, Mead ME, Pristic S, Wang K, Lin F-Y, Chang T-K, Peters RJ, Oldfield E. Diterpene cyclases and the nature of the isoprene fold. *Proteins: Struct, Funct, Bioinf.* 2010; 78:2417–2432.
27. Chen A, Kroon PA, Poulter CD. Isoprenyl diphosphate synthases: Protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. *Protein Sci.* 1994; 3:600–607. [PubMed: 8003978]
28. Aaron JA, Christianson DW. Trinuclear metal clusters in catalysis by terpenoid synthases. *Pure Appl Chem.* 2010; 82:1585–1597. [PubMed: 21562622]
29. Toyomasu T, Tsukahara M, Kenmoku H, Anada M, Nitta H, Ohkanda J, Mitsuhashi W, Sassa T, Kato N. Transannular proton transfer in the cyclization of geranylgeranyl diphosphate to fusicoccadiene, a biosynthetic intermediate of fusicoccins. *Org Lett.* 2009; 11:3044–3047. [PubMed: 19530695]
30. Tarshis LC, Yan M, Poulter CD, Sacchettini JC. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. *Biochemistry.* 1994; 33:10871–10877. [PubMed: 8086404]
31. Kavanagh KL, Dunford JE, Bunkoczi G, Russell RG, Oppermann U. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J Biol Chem.* 2006; 281:22004–22012. [PubMed: 16698791]
32. Chang T-H, Guo R-T, Ko T-P, Wang AH-J, Liang P-H. Crystal structure of type-III geranylgeranyl pyrophosphate synthase from *Saccharomyces cerevisiae* and the mechanism of product chain length determination. *J Biol Chem.* 2006; 281:14991–15000. [PubMed: 16554305]
33. Kloer DP, Welsch R, Beyer P, Schulz GE. Structure and reaction geometry of geranylgeranyl diphosphate synthase from *Sinapis alba*. *Biochemistry.* 2006; 45:15197–15204. [PubMed: 17176041]
34. Ezra A, Golomb G. Administration routes and delivery systems of bisphosphonates for the treatment of bone resorption. *Adv Drug Delivery Rev.* 2000; 42:175–195.
35. Holm L, Rosenström P. Dali server: conservation mapping in 3D. *Nucl Acids Res.* 2010; 38:545–549.

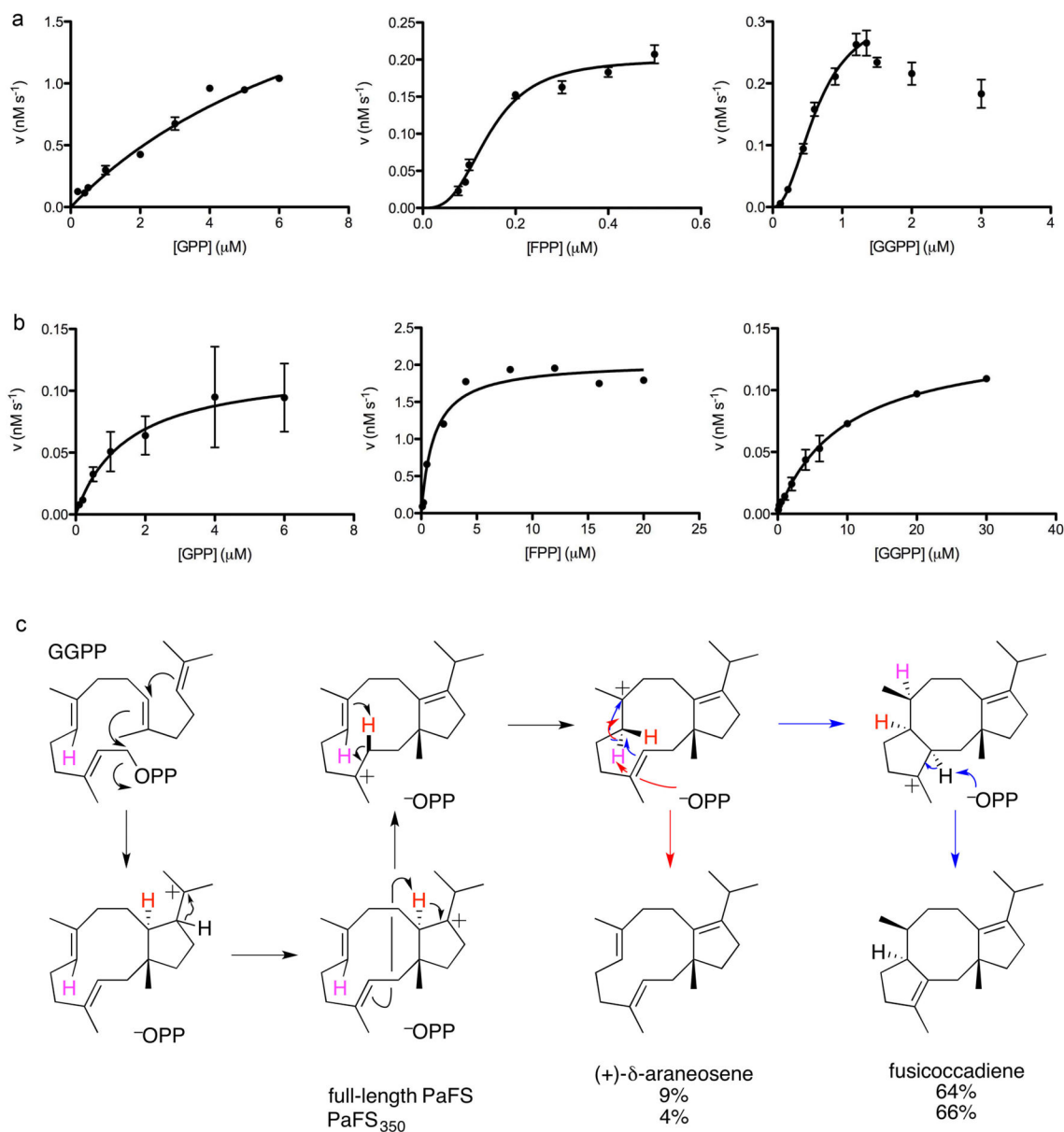
36. Shishova EY, Di Costanzo L, Cane DE, Christianson DW. X-ray crystal structure of aristolochene synthase from *Aspergillus terreus* and the evolution of templates for the cyclization of farnesyl diphosphate. *Biochemistry*. 2007; 46:1941–1951. [PubMed: 17261032]
37. Chen M, Al-lami N, Janvier M, D'Antonio EL, Faraldos JA, Cane DE, Allemann RK, Christianson DW. Mechanistic insights from the binding of substrate and carbocation intermediate analogues to aristolochene synthase. *Biochemistry*. 2013; 52:5441–5453. [PubMed: 23905850]
38. Ma B, Tsai CJ, Haliloglu T, Nussinov R. Dynamic allostery: Linkers are not merely flexible. *Structure*. 2011; 19:907–917. [PubMed: 21742258]
39. Baer P, Rabe P, Citron CA, de Oliveira Mann CC, Kaufmann N, Groll M, Dickschat JS. Hedycaryl synthase in complex with nerolidol reveals terpene cyclase mechanism. *Chem Bio Chem*. 2014; 15:213–216.
40. Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW. Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. *J Biol Chem*. 2000; 275:25533–25539. [PubMed: 10825154]
41. Hosfield DJ, Zhang Y, Dougan DR, Broun A, Tari LW, Swanson RV, Finn J. Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. *J Biol Chem*. 2004; 279:8526–8529. [PubMed: 14672944]
42. Petoukhov MV, Svergun DI. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J*. 2005; 89:1237–1250. [PubMed: 15923225]
43. Brodelius M, Lundgren A, Mercke P, Brodelius PE. Fusion of farnesyl diphosphate synthase and *epi*-aristolochene synthase, a sesquiterpene cyclase involved in capsidiol biosynthesis. *Eur J Biochem*. 2002; 269:3570–3577. [PubMed: 12135497]
44. Castellana M, Wilson MZ, Xu Y, Joshi P, Cristea IM, Rabinowitz JD, Gitai Z, Wingreen NS. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat Biotechnol*. 2014; 32:1011–1018. [PubMed: 25262299]
45. Chiba R, Minami A, Gomi K, Oikawa H. Identification of ophiobolin F synthase by a genome mining approach: a sesterterpene synthase from *Aspergillus clavatus*. *Org Lett*. 2013; 15:594–597. [PubMed: 23324037]
46. Li R, Chou WKW, Himmelberger JA, Litwin KM, Harris GG, Cane DE, Christianson DW. Reprogramming the chemodiversity of terpenoid cyclization by remolding the active site contour of *epi*-isozizaene synthase. *Biochemistry*. 2014; 53:1155–1168. [PubMed: 24517311]
47. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst*. 2010; D66:213–221.
48. Svergun DI, Petoukhov MV, Koch MHJ. Determination of domain structure of proteins from X-ray solution scattering. *Biophys J*. 2001; 80:2946–2953. [PubMed: 11371467]



**Figure 1. Fusicoccadiene synthase from *P. amygdali* (PaFS)**

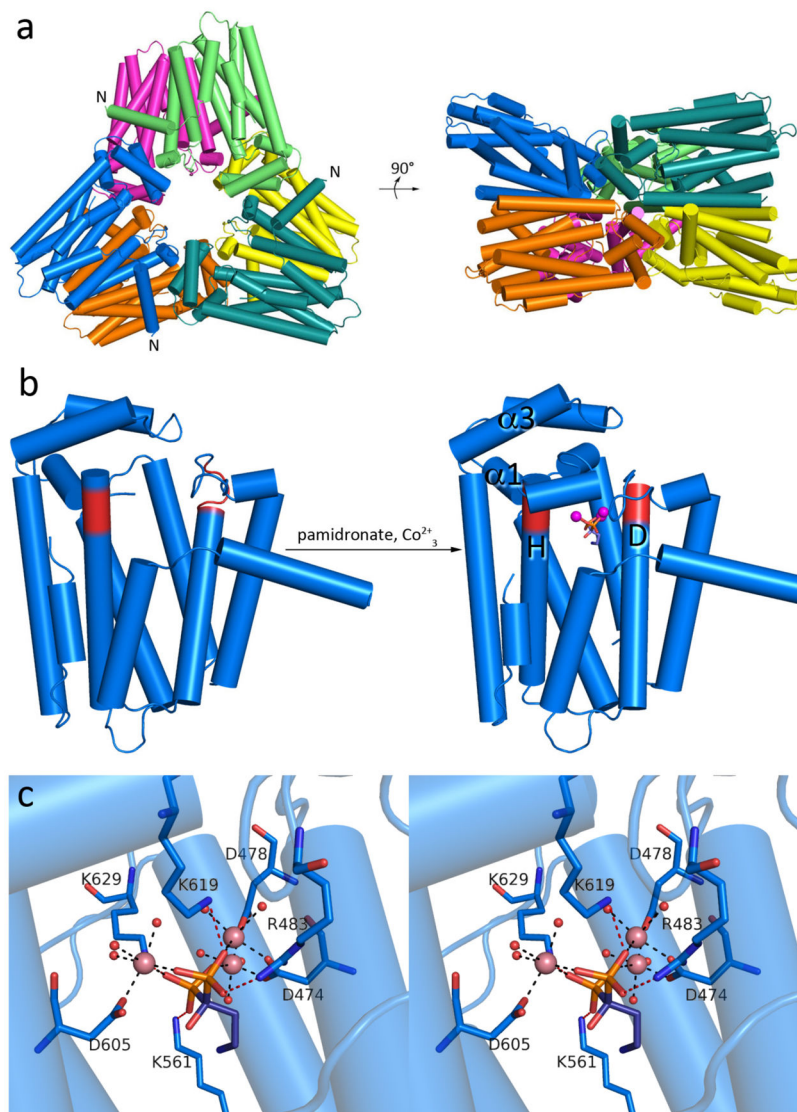
Geranylgeranyl diphosphate (GGPP) is generated from one molecule of dimethylallyl diphosphate (DMAPP) and three molecules of isopentenyl diphosphate (IPP) in the C-terminal α domain (blue box), and GGPP is cyclized in the N-terminal α domain to form fusicocca-2,10(14)-diene (red box). Further biosynthetic modifications (green) yield the phytotoxin fusicoccin A.





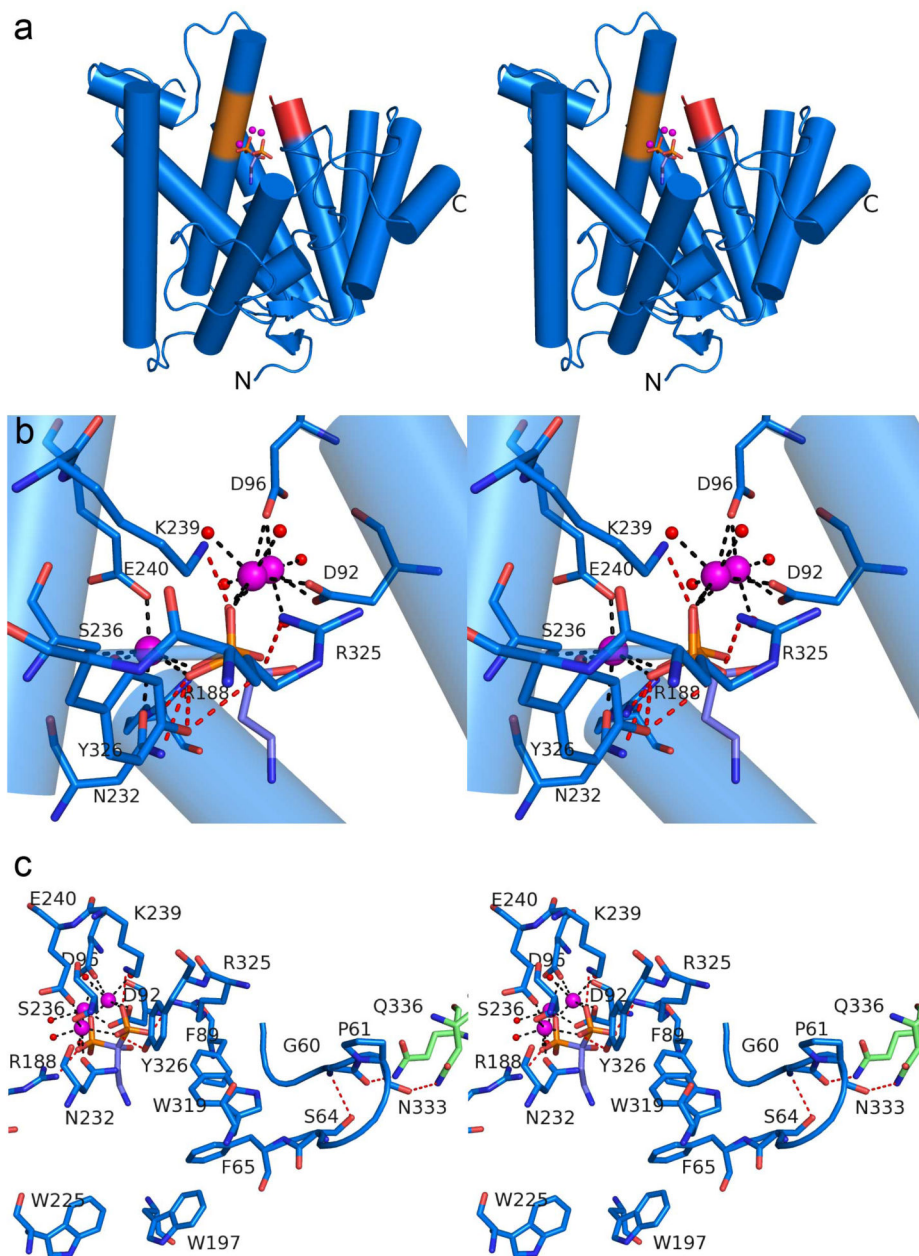
**Figure 2. Catalytic activity measurements**

(a) Generation of hydrocarbon products from substrates GPP, FPP, and GGPP in the absence of IPP (the concentrations of full-length PaFS in these experiments were 74.5 nM, 7.45 nM, and 14.9 nM, respectively). Catalysis with GPP exhibits Michaelis-Menten kinetics, whereas catalysis with FPP or GGPP exhibits cooperativity based on the sigmoidal dependence of catalytic activity on substrate concentration. With GGPP, substrate inhibition is evident at higher concentrations. (b) PaFS<sub>350</sub> (33.5 nM) exhibits Michaelis-Menten kinetics for the generation of hydrocarbon products from substrates GPP, FPP, and GGPP. (c) Proposed cyclization mechanism<sup>29</sup> for the generation of fusicocadiene (black and blue arrows) and δ-araneosene (black and red arrows) by full-length PaFS and PaFS<sub>350</sub> (a third, unidentified diterpene product is also generated).



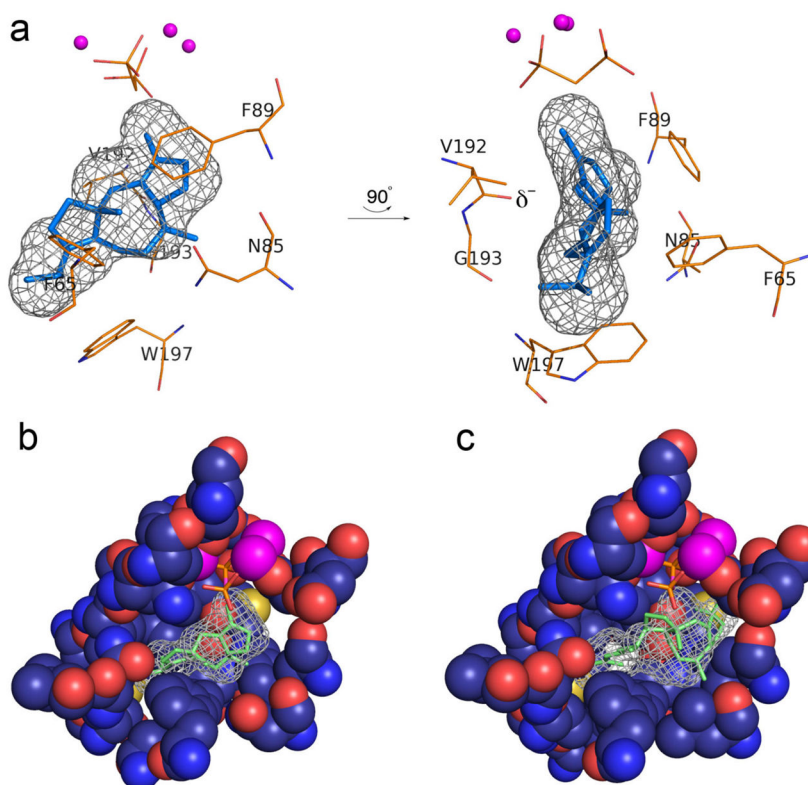
**Figure 3. PaFS C-terminal GGPP synthase domain**

a) PaFS<sub>CT</sub> adopts the  $\alpha$  fold of a class I terpenoid synthase and crystallizes as a hexamer, or a trimer of dimers. The N-termini of selected subunits are labeled and indicate the point of connection to the missing N-terminal domain. Two perpendicular orientations are shown. (b) The binding of 3 Co<sup>2+</sup> ions and pamidronate triggers complete closure of the active site of PaFS<sub>CT</sub>. (c) Stereoview of intermolecular interactions in the PaFS<sub>CT</sub>-Co<sup>2+</sup><sub>3</sub>-pamidronate complex. Metal coordination and hydrogen bond interactions are shown as black and red dashed lines, respectively (metal-ligand distances are recorded in Supplementary Information Table S3). Pamidronate binds in the DMAPP binding site, so interactions of the phosphonate groups mimic interactions with the diphosphate group of DMAPP that trigger ionization and formation of the allylic cation that initiates the chain elongation reaction.



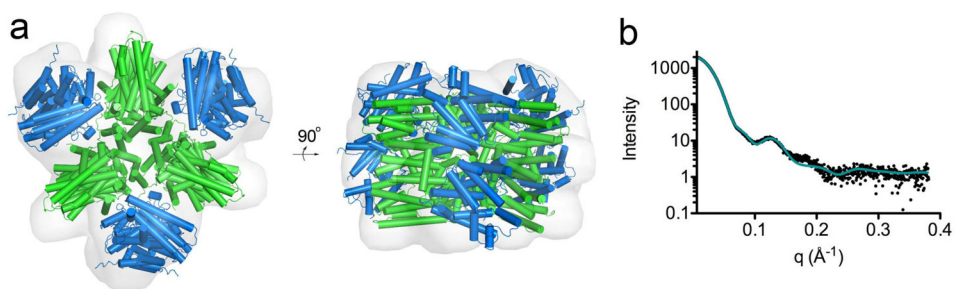
**Figure 4. PaFS N-terminal GGPP cyclase domain**

(a) Stereoview of the PaFS<sub>344</sub>-Mg<sup>2+</sup><sub>3</sub>-pamidronate complex reveals that the N-terminal domain adopts the α fold of a class I terpenoid synthase. (b) Stereoview of the PaFS<sub>344</sub>-Mg<sup>2+</sup><sub>3</sub>-pamidronate complex showing metal coordination and hydrogen bond interactions as black and red dashed lines, respectively (metal-ligand distances are recorded in Supplementary Information Table S3). (c) The G60-F65 hairpin segment may facilitate communication between cyclization domain active sites in full-length hexameric PaFS by mediating interactions between active site residue F65 and interdomain linker residues N333 and Q336 (green).



**Figure 5. Models of enzyme-product complexes**

(a) Model of product fusicoccadiene bound in the active site of PaFS<sub>344</sub>. The three-dimensional active site contour is represented by black meshwork, and fusicoccadiene was manually fit into this meshwork. The position of the Mg<sup>2+</sup><sub>3</sub>-PP<sub>i</sub> cluster is modeled after the Mg<sup>2+</sup><sub>3</sub>-bisphosphonate cluster in the PaFS<sub>344</sub>-Mg<sup>2+</sup><sub>3</sub>-pamidronate complex. Two perpendicular views are shown; the backbone carbonyl of V192 at the helix G break ( $\delta^-$ ) is oriented toward the location of a proposed carbocation intermediate in the PaFS mechanism. (b) Cut-away view of the active site pocket of PaFS<sub>344</sub>, showing how the surface contour (black meshwork) is defined by residues lining the pocket. (c) Ophiobolin F docked in the active site of ophiobolin F synthase modeled after the PaFS<sub>344</sub>-fusicoccadiene complex in (b). The extra active site volume resulting from the W225L and V228A substitutions readily accommodate the larger C<sub>25</sub> sesterterpene.



**Figure 6. Structure of the full-length PaFS hexamer in solution**

(a) Model of the hexamer of full-length N333A/Q336A PaFS fit into the three-fold averaged *ab initio* molecular envelope generated from SAXS data. PaFS<sub>344</sub> and PaFS<sub>CT</sub> dimers are marine and green, respectively. (b) Theoretical scattering profile (teal solid line) overlaid with experimental data (black) indicates an excellent fit of the molecular envelope in (a), with  $\chi = 1.5$  as calculated with SASREF.

**Table 1**

Catalytic activity of PaFS constructs

Substrate	Full-length PaFS			
	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\mu\text{M}$ )	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{M}^{-1}\text{s}^{-1}$ )	$h$
GPP <sup>a</sup>	0.036 ± 0.007	9.2 ± 2.6	3.9 × 10 <sup>3</sup>	1.0
FPP <sup>b</sup>	0.027 ± 0.001	0.14 ± 0.02	1.9 × 10 <sup>5</sup>	3.0
GGPP <sup>c</sup>	0.021 ± 0.002	0.63 ± 0.03	3.3 × 10 <sup>4</sup>	2.2
Substrate	Cyclization domain, PaFS <sub>350</sub>			
	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\mu\text{M}$ )	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{M}^{-1}\text{s}^{-1}$ )	$h$
GPP <sup>a</sup>	0.0036 ± 0.0005	1.5 ± 0.6	2.7 × 10 <sup>3</sup>	1.0
FPP <sup>d</sup>	0.069 ± 0.004	1.5 ± 0.3	4.6 × 10 <sup>4</sup>	1.0
GGPP <sup>c</sup>	0.0042 ± 0.0002	9.6 ± 1.1	4.4 × 10 <sup>2</sup>	1.0
Substrates for GGPP biosynthesis	Full-length D92A PaFS, with inactivated cyclase domain			GGPP synthase domain, PaFS <sub>CT</sub>
	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\mu\text{M}$ )	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{M}^{-1}\text{s}^{-1}$ )	specific activity ( $\text{nmol}\cdot\text{s}^{-1}\text{mg}^{-1}$ )
GPP + IPP	0.26 ± 0.03	2.1 ± 0.6	1.2 × 10 <sup>5</sup>	0.77 ± 0.03
FPP + IPP	1.6 ± 0.2	1.9 ± 0.5	8.4 × 10 <sup>5</sup>	1.96 ± 0.08

<sup>a</sup>Yields geraniol and linalool as major products.<sup>b</sup>Yields (*E,E*)-farnesol and (*E*)-nerolidol as major products.<sup>c</sup>Yields fusicoccadiene,  $\delta$ -araneosene, and one unidentified diterpene as major products.<sup>d</sup>Yields (*E*)- $\beta$ -farnesene and (*E*)-nerolidol as major products.