—— SUPPLEMENTARY INFORMATION ——

Taxadiene Synthase Structure and Evolution of Modular Architecture in Terpenoid Biosynthesis

Mustafa Köksal[1], Yinghua Jin[2,3], Robert M. Coates[2], Rodney Croteau[4], and David W. Christianson[1]

[1]*Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104-6323 USA*

[2]*Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801*

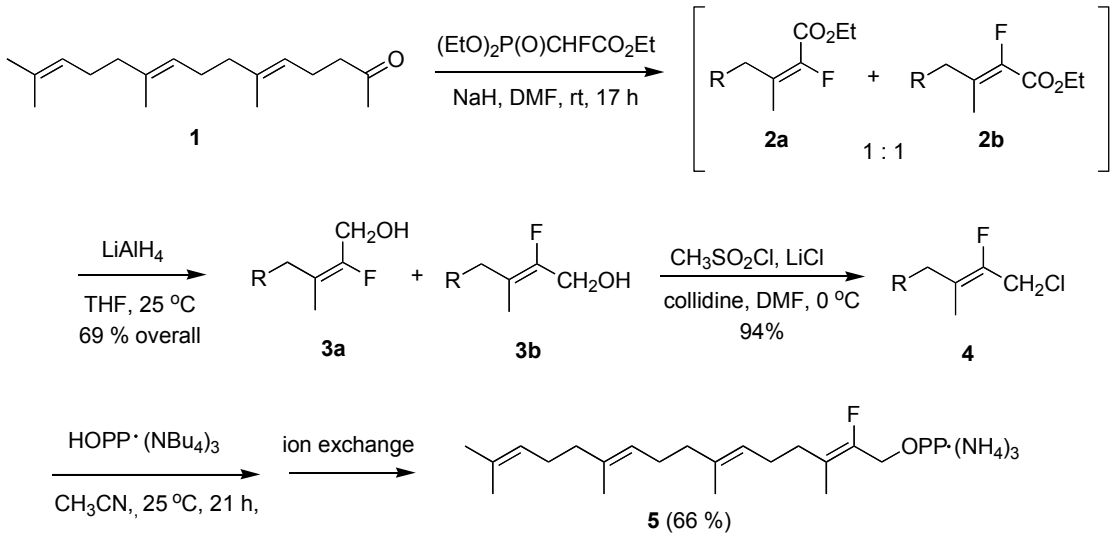
[3]*Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309*

[4]*Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340*

—— SUPPLEMENTARY METHODS ——

Synthesis of isoprenoid diphosphate ligands. Ligand structures are illustrated in Figure S1. 13-Aza-13,14-dihydrocopalyl diphosphate (ACP) was available from previous research[29]. The previously unknown 2-fluorogeranylgeranyl diphosphate (FGP, **5**) was synthesized in 4 steps from (*E*, *E*)-farnesylacetone (**1**) as outlined in Scheme S1 following procedures similar to those used in the preparation of 2-fluorogeraniol, 2-fluorofarnesol, and their diphosphate derivatives[22,41-43]. Full details are described below. Non-standard abbreviations are; -OPP for diphosphate.

Scheme S1



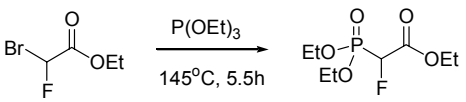
Materials were purchased from commercial suppliers and used without further purification, unless otherwise indicated. Ether and tetrahydrofuran (THF) were distilled from sodium/benzophenone immediately prior to use. CH_2Cl_2 was distilled from CaH_2 immediately

before use. CH_3CN and DMF were distilled from P_2O_5 and stored over 3 Å molecular sieves. A commercial sample of methanesulfonyl chloride (Aldrich) was distilled from P_2O_5 before use. All the reactions, except those performed in aqueous solvent, were conducted under dry nitrogen in flame-dried glassware. Unless otherwise specified, solvents were evaporated using a rotary evaporator after workup.

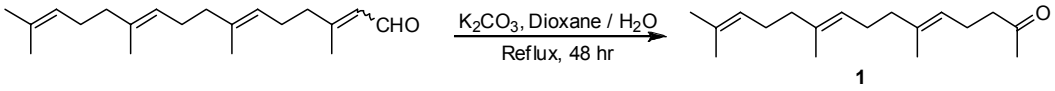
Flash column chromatography refers to the procedure of Still[44] and was performed by using a 100-150 times weight excess of silica gel 60 230-400 mesh ASTM from Merck. Fractions were analyzed by TLC using silica gel 60 F254 250 μm precoated-plates. Ion exchange column chromatography for diphosphate purification was carried out according to the procedure of Poulter[45]. Dowex AG 50W-X8 was obtained from BioRad (H^+ form) and converted to the ammonium form before ion exchange chromatography according to the method of Poulter[45].

IR spectra were recorded on a Perkin Elmer Spectrum BX spectrophotometer as neat liquids on NaCl plates. LR, HR ESI, EI and CI mass spectra were recorded on a Micromass ZAB-SE Spectrometer (Waters Corporation, Beverly, MA, USA) in the Mass Spectrometry Laboratory of the School of Chemical Sciences, University of Illinois, Urbana-Champaign. NMR spectra were taken on Unity 400, Unity 500, or Unity 600 spectrometers. $CHCl_3$ (7.26 ppm), C_6H_6 (7.15 ppm), H_2O (4.67 ppm), and CH_3OH (4.87 ppm) were used as internal references in ^1H NMR spectroscopy. ^{19}F NMR and ^{31}P NMR spectra were obtained with 85 % H_3PO_4 (0 ppm) and CCl_3F (0 ppm) as external references, respectively. ^1H NMR data are reported in the order: chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet), number of protons, coupling constants (*J* in Hz), assignment.

Unless otherwise specified, the purity of all compounds was ≥ 95 % based on NMR spectral integration and/or GC analysis.



Triethyl 2-Fluoro-2-phosphonoacetate. The procedure by Burton[46] was followed with some modification. A solution of triethyl phosphite (6.0 g, 36 mmol) and ethyl bromofluoroacetate (2.0 g, 11 mmol) was heated in a sealed tube for 5.5 h at 145 °C. Removal of low-boiling fractions by simple distillation gave the desired pure product (2.42 g, 91 %) as a colorless oil: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.29-1.36 (m, 9H, C*H*$_3$), δ 4.18-4.27 (m, 4H, C*H*$_2$), δ 4.31 (q, 2H, J = 7.1 Hz, C*H*$_2$OCO), δ 5.18 (dd, 1H, J_{HF} = 47.9 Hz, J_{HP} = 12.6 Hz, C*H*F); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 14.2 (s, *C*H$_3$), δ 16.5 (d, J_{CP} = 5.6 Hz, 2*C*H$_3$), δ 62.6 (s, *C*H$_2$), δ 64.3 (d, J_{CP} = 6.0 Hz, *C*H$_2$), δ 64.4 (d, J_{CP} = 5.8 Hz, *C*H$_2$), δ 84.5 (dd, J_{CF} = 196.3, 159.2 Hz, *C*HF), δ 165 (d, J_{CF} = 22.9 Hz, *C*O); ^{19}F NMR ($CDCl_3$, 376 MHz) δ –210.9 (dd, J_{PF} = 71.8 Hz); ^{31}P NMR ($CDCl_3$, 162 MHz) δ 11.2 (d, J_{PF} = 71.8 Hz). The NMR data are consistent with the reported values by Elkik and Imbeaux[47].

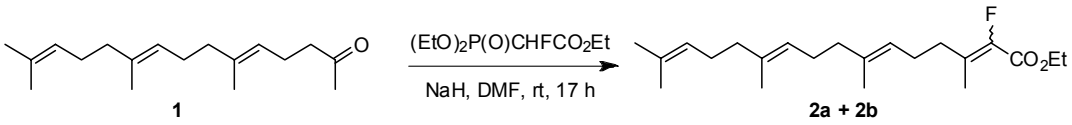


1

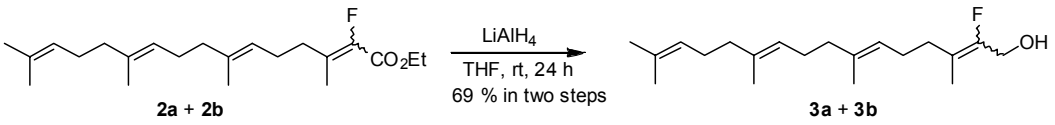
(*E*, *E*)-Farnesylacetone (1). Oxidative rearrangement of geranyllinalool with pyridinium chlorochromate[48-50] afforded a mixture of (*Z*, *E*, *E*)- and (*E*, *E*, *E*)-geranylgeranials. Retroaldol hydrolysis[51] of the mixture of α,β-unsaturated aldehydes gave farnesylacetone in 85 % yield as described below.

A suspension of the mixture of isomeric geranylgeranials (2.36 g, 8.20 mmol) and K_2CO_3 (2.95 g, 21.3 mmol) in dioxane (30 mL) and water (30 mL) was refluxed for 48 h (oil bath temperature 120 °C). It was then cooled to room temperature, and the orange mixture was diluted with water (80 mL). The organic layer was separated, and the aqueous layer was extracted with

hexane (4 x 50 mL). The combined hexane fractions were dried over Na_2SO_4 and concentrated to give the crude product (2.19 g) as a red oil. Flash column chromatography purification with 8 % ether in hexane as the eluent afforded farnesylacetone (**1**, 1.87 g, 85 %): TLC R_f 0.63 (20 % EtOAc in hexane); ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.60 (s, 3H, CH_3), 1.61 (s, 3H, CH_3), 1.62 (d, 3H, J = 0.5 Hz, CH_3), 1.69 (d, 3H, J = 1.1 Hz CH_3), 1.96-2.10 (m, 8H, CH_2), 2.14 (s, 3H, CH_3), 2.27 (q, 2H, J = 7.2 Hz, CH_2), 2.47 (t, 2H, J = 7.5 Hz, CH_2), 5.06-5.12 (m, 3H, vinyl H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 16.18, 16.19, 17.9, 22.6, 25.9, 16.7, 26.9, 30.1, 39.8, 39.9, 43.9, 122.7, 124.2, 124.5, 131.4, 135.2, 136.6, 209.0. The NMR data are consistent with the values reported by Kato *et al*.[52].



Ethyl 2-Fluoro-3,7,11,15-tetramethyl-hexadeca-2,6,10,14-tetraenoates (2a + 2b). Conditions for the following reaction were based on those described by Komatsu and Kitazume[53]. Farnesylacetone (800 mg, 3.05 mmol) was converted to esters **2a** and **2b** using triethyl 2-fluoro-2-phosphonoacetate (890 mg, 3.66 mmol) and NaH (122 mg, 3.05 mmol) in DMF at room temperature for 17 h. A 1:1 mixture (GC) of trans/cis isomers of esters (**2a** + **2b**, 1.05 g) was obtained as an orange oil. The mixture was used in the following step without further purification.

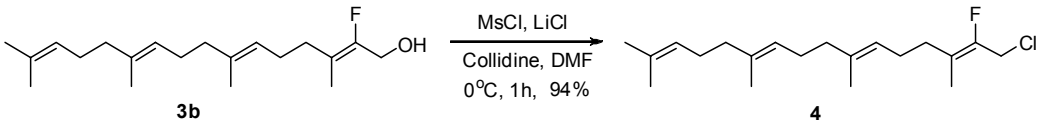


(2*E/Z*, 6*E*, 10*E*)-2-Fluoro-3,7,11,15-tetramethyl-hexadeca-2,6,10,14-tetraen-1-ols (3a + 3b). Reduction of the ester mixture (**2a + 2b**) to the corresponding alcohols was carried out following

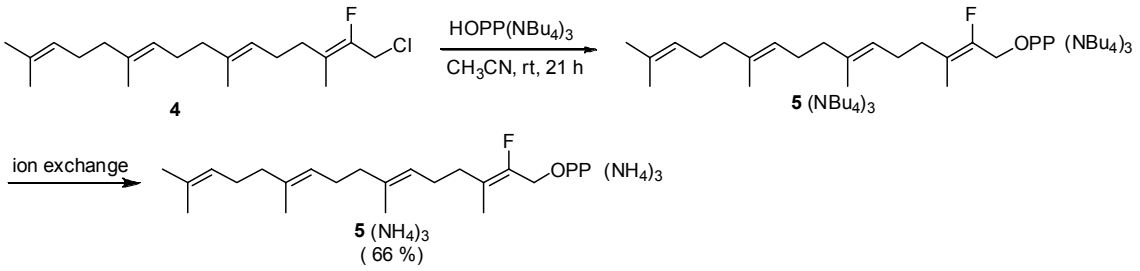
the procedure by Cane[54]. A solution of the ester (**2a** + **2b**, 1.00 g, 2.85 mmol) in THF (70 mL) was stirred and cooled at 0 ^{o}C as $LiAlH_4$ (130 mg, 3.42 mmol) was added. The suspension was allowed to warm to room temperature and stirred for 24 h. The reaction was quenched by the addition of saturated NH_4Cl (100 mL). The mixture was extracted with ether (3 x 100 mL). The combined ether solution was washed with water (2 x 100 mL) and brine (2 x 100 mL), dried over $NaSO_4$, and concentrated to give a 1:1 mixture of trans and cis isomers of alcohol **3** (802 mg) as a yellow oil. Purification by flash column chromatography (8 % EtOAc in hexane) on silica gel afforded the cis isomer followed by the trans isomer (total 705 mg, 69 % in two steps).

Cis isomer (**3a**): TLC R_f 0.18 (10 % EtOAc in hexane); IR (neat) υ_{max} 3350, 2966, 2920, 2858, 1448, 1382, 1149, 1014 cm^{-1}; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.61 (m, 9H, C*H*$_3$), 1.69 (d, 3H, *J* = 1.0 Hz C*H*$_3$), 1.70 (d, 3H, *J* = 3.3 Hz C*H*$_3$), 1.97-2.14 (m, 12H, C*H*$_2$), 4.21 (d, 2H, *J* = 22.8 Hz, C*H*$_2$OH), 5.08-5.12 (m, 3H, vinyl *H*); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 13.7, 13.8, 16.2 (d, *J* = 4.8 Hz), 17.9, 25.9, 26.6 (d, *J* = 3.0 Hz), 26.7, 26.9, 31.93, 31.97, 39.9 (d, *J* = 5.2 Hz), 58.0 (d, *J* = 32.1 Hz), 116.0 (d, *J* = 13.5 Hz), 123.2, 124.2, 124.6, 131.5, 135.4, 136.8, 153.8 (d, *J* = 244.0 Hz); ^{19}F NMR ($CDCl_3$, 470 MHz) δ -120.0 (t, *J* = 22.3 Hz); LR-MS (EI): m/z 308.2; HR-MS (EI): Calcd for $C_{20}H_{33}FO$, 308.2504. Found, 308.2507.

Trans isomer (**3b**): TLC R_f 0.12 (10% EtOAc in hexane); IR (neat) υ_{max} 3339, 2967, 2921, 2857, 1450, 1383, 1159, 1014 cm^{-1}; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.57 (s, 1H, O*H*), 1.61 (m, 9H, C*H*$_3$), 1.69 (d, 3H, *J* = 3.1 Hz, C*H*$_3$), 1.69 (s, 3H, C*H*$_3$)1.96-2.14 (m, 12H, C*H*$_2$), 4.25 (dd, 2H, *J* = 22.4, 6.2 Hz, C*H*$_2$OH), 5.08-5.16 (m, 3H, vinyl *H*); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 15.5, 15.6, 16.2 (d, *J* = 1.3 Hz), 17.9, 25.9, 26.1 (d, *J* = 1.7 Hz), 26.8, 27.0, 29.99, 30.05, 39.9 (d, *J* = 4.0 Hz), 58.2 (d, *J* = 32.2 Hz), 116.2 (d, *J* = 15.5 Hz), 123.4, 124.4, 124.6, 131.5, 135.2, 136.0, 153.0 (d, *J* = 242.1 Hz); ^{19}F NMR ($CDCl_3$, 470 MHz) δ -121.5 (t, *J* = 21.8 Hz); LR-MS (EI): m/z 308.3; HR-MS (EI): Calcd for $C_{20}H_{33}FO$, 308.2504. Found, 308.2508.



(2*Z*, 6*E*, 10*E*)- 1-Chloro-2-fluoro-3,7,11,15-tetramethyl-hexadeca-2,6,10,14-tetraene (4). Fluoro alcohol **3b** was converted to chloride **4** under Meyers' conditions[55]. Reaction of **3b** (100 mg, 1.32 mmol), LiCl (137 mg, 3.24 mmol), *s*-collidine (393 mg, 3.24 mmol), and MsCl (112 mg, 0.97 mmol) in DMF (2.5 mL) for 1 h at 0 ˚C provided chloride **4** as a yellow oil (99 mg, 94%). The product was converted to the diphosphate directly without purification. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.61 (m, 9H, CH_3), 1.69 (d, 3H, J = 1.2 Hz, CH_3), 1.71 (d, 3H, J = 2.9 Hz, CH_3), 1.96-2.19 (m, 12H, CH_2), 4.19 (d, 2H, J = 22.4 Hz, CH_2Cl), 5.08-5.14 (m, 3H, vinyl *H*); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 15.9, 16.0, 16.2 (d, J = 2.2 Hz), 17.9, 25.9, 26.2, 26.8, 27.0, 30.18, 30.24, 39.4 (d, J = 34.3 Hz), 39.9 (d, J = 5.2 Hz), 119.0 (d, J = 16.1 Hz), 123.4, 124.4, 124.6, 131.5, 135.2, 136.2, 149.9 (d, J = 240.2 Hz); ^{19}F NMR ($CDCl_3$, 376 MHz) δ -117.6 (t, J = 23.1 Hz).



(2*Z*, 6*E*, 10*E*)-2-Fluoro-3,7,11,15-tetramethyl-hexadeca-2,6,10,14-tetraen-1-yl Diphosphate, Trisammonium Salt (5). Diphosphorylation was carried out as previously, using Poulter's method[45]. The reaction of chloride **4** (99 mg, 0.30 mmol) with $HOPP(NBu_4)_3$ (592 mg, 0.61 mmol) and 3 Å molecular sieves (670 mg) in dry acetonitrile (5 mL) for 21 h at room temperature provided crude diphosphate **5** as a brown oil (567 mg). ^{31}P NMR analysis indicated it was a 1: 2.2 mixture of inorganic diphosphate and organic diphosphate. Ion exchange

chromatography on cation exchange resin (Aldrich, Dowex AG 50 W-X8 100-200 mesh) using the buffer (30 mL, 5:1 v/v 2M $(NH_4)_2CO_3$: 1-propanol) and lyophilization, followed by washing with MeOH (5 x 10 mL) to remove the inorganic diphosphate afforded a yellowish solid. Further purification by recrystallization from MeOH gave the pure ammonium salt **5** as a white powder (55 mg, 66 %). NMR data for tetrabutylammonium salt of diphosphate **5**: ^{1}H NMR (CD_3OD, 400 MHz) δ 0.96 (t, 36H, *J* = 7.4 Hz, C*H*$_3$), 1.36 (sextet, 24H, *J* = 7.4 Hz, C*H*$_2$), 1.54-1.64 (m, 36H, 12C*H*$_2$, 4C*H*$_3$), 1.69 (d, *J* = 3.0 Hz, CH_3), 1.89-1.95 (m, 4H, C*H*$_2$), 1.97-2.05 (m, 8H, C*H*$_2$), 3.16-3.20 (m, 24H, C*H*$_2$), 4.62 (dd, 2H, *J* = 23.6, 5.3 Hz, CH_2OPP), 5.01-5.08 (m, 3H, vinyl *H*); ^{13}C NMR (CD_3OD, 100 MHz) δ 14.1, 16.0, 16.8, 17.9, 20.8, 24.6, 24.9, 26.1, 27.1, 27.7, 32.5, 40.9, 41.2, 57.0, 59.6, 63.8 (d, *J* = 5.3 Hz) , 110.9, 118.4, 123.2, 123.3, 132.8, 136.2, 139.9, 156.3; ^{31}P NMR (CD_3OD, 162 MHz) δ 4.24 (s, inorganic monophosphate), -5.90 (s, inorganic diphosphate), -7.05 (d, *J* = 19.1 Hz), –8.74 (d, *J* = 19.6 Hz), -10.0 (s, organic monophosphate); ^{19}F NMR (CD_3OD, 376 MHz) δ -119.6 (t, *J* = 24.0 Hz). NMR data for trisammonium salt of diphosphate **5**: ^{1}H NMR (CD_3OD, 400 MHz) δ 1.58 (s, 6H, C*H*$_3$), 1.59 (d, 3H, *J* = 1.0 Hz, C*H*$_3$), 1.65 (d, 3H, *J* = 1.0 Hz, C*H*$_3$), 1.72 (d, 3H, *J* = 2.9 Hz C*H*$_3$), 1.93-2.09 (m, 12H, C*H*$_2$), 4.61 (dd, 2H, *J* = 23.2, 5.4 Hz, C*H*$_2$OPP), 5.05-5.12 (m, 3H, vinyl *H*); ^{31}P NMR (CD_3OD, 162 MHz) δ -8.49 (d, *J* =19.0 Hz), -8.49 (d, *J* =19.8 Hz); ^{19}F NMR (CD_3OD, 376 MHz) δ -119.6 (t, *J* = 24.0 Hz); Low resolution ESI: m/z 467.4; HR-MS (ESI): Calcd for $C_{20}H_{34}FP_2O_7$, 467.1764. Found, 467.1782.

References for Supplementary Methods

41. Karp, F. *et al*. Inhibition of monoterpene cyclases by inert analogues of geranyl diphosphate and linalyl diphosphate. *Arch Biochem Biophys* **468**, 140-146 (2007).

42. Faraldos, J. A., Zhao, Y., O'Maille, P. E., Noel, J. P. &, Coates, R. M. Interception of the enzymatic conversion of farnesyl diphosphate to 5-epi-aristolochene by using a fluoro substrate

analogue: 1-fluorogermacrene A from (2*E*,6*Z*)-6-fluorofarnesyl diphosphate. *Chem Bio Chem* **8**, 1826-1833 (2007).

43. Shishova, E. Y. *et al*. X-ray crystallographic studies of substrate binding to aristolochene synthase suggest a metal ion binding sequence for catalysis. *J Biol Chem* **283** 15431-15439 (2008).

44. Still, W. C., Kahn, M. & Mitra, A. Rapid chromatographic technique for preparative separations with moderate resolution. *J Org Chem* **43**, 2923-2925.

45. Woodside, A. B. & Huang, Z., Poulter, C. D. (1993) Trisammonium geranyl diphosphate. *Org Synth Coll Vol VIII*, 616-620 (1978).

46. Thenappan, A. & Burton, D. J. Reduction-olefination of esters: a new and efficient synthesis of α-fluoro α,β-unsaturated esters. *J Org Chem* **55**, 4639-4642 (1990).

47. Elkik, E. & Imbeaux, M. A convenient synthesis of ethyl (diethoxyphosphoryl) fluoroacetate from ethyl fluoroacetate. *Synthesis* 861-862 (1989).

48. Babler, J. H. & Coghlan, M. J. A facile method for the bishomologation of ketones to α,β-unsaturated aldehydes: application to the synthesis of the cyclohexanoid components of the boll weevil sex attractant. *Synth Commun* **6**, 469-474 (1976).

49. Sundararaman, P. & Herz, W. Oxidative rearrangements of tertiary and some secondary allylic alcohols with chromium(VI) reagents. A new method for 1,3-functional group transposition and forming mixed aldol products. *J Org Chem* **42**, 813-819 (1977).

50. Piancatelli, G., Scettri, A. & D'Auria, M. Pyridinium chlorochromate: A versatile oxidant in organic synthesis. *Synthesis* 245-258 (1982).

51. Coates, R. M. & Freidinger, R. M. Total synthesis of (±)-sesquicarene. *Tetrahedron* **26**, 3487-3493 (1970).

52. Kato, T., Suzuki, M., Kobayashi, T. & Moore, B. P. Synthesis and pheromone activities of optically active neocembrenes and their geometrical isomers, (*E*,*Z*,*E*)- and (E,E,Z)-neocembrenes. *J Org Chem* **45**, 1126-1130 (1980).

53. Komatsu, Y. & Kitazume, T. Preparation of 2,6-difluoromanoalogues derivatives. *J Fluorine Chem* **102**, 61-67 (2000).

54. Cane, D. E., Yang, G., Xue, Q. & Shim, J. H. Trichodiene synthase. Substrate specificity and inhibition. *Biochemistry* **34**, 2471-2479 (1995).

55. Collington, E. W & Meyers, A. I. Facile and specific conversion of allylic alcohols to allylic chlorides without rearrangement. *J Org Chem* **36**, 3044-3045 (1971).

—— SUPPLEMENTARY TABLES ——

Table S1. Data Collection and Refinement Statistics

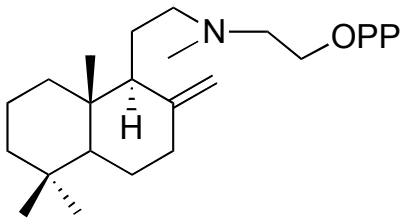
	TXS-ACP complex	TXS-FGP complex
Data Collection		
Space group	$P2_12_12_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	55.46, 72.41, 206.93	54.05, 201.98, 81.43
α, β, γ (º)	90, 90, 90	90, 91.60, 90
Wavelength (Å)	0.945	1.008
Resolution (Å)	42.1-1.82	43.4–2.25
R_{merge}*	0.120 (0.530)	0.130 (0.454)
$I/\sigma I$	14.1 (3.1)	5.7 (2.2)
Completeness (%)	99.6 (98.7)	85.4 (86.9)
Redundancy	6.6 (6.0)	2 (1.9)
Refinement		
Resolution (Å)	42.1-1.82	43.4–2.25
No. reflections	72,900	66,937
R_{work}/R_{free}	0.167 / 0.205	0.187 / 0.250
No. atoms#		
Protein	6103	12049
Ligand/ion	73	66
Water	953	1206
Average B factors (Å^2)		
Main chain	21	23
Side chain	26	24
Ligand/ion	36	27
Water	35	23
R.m.s. deviations		
Bonds lengths (Å)	0.006	0.005
Bond angles (º)	0.9	0.9

*Number in parentheses refers to the outer resolution shell of data.
#Per asymmetric unit.

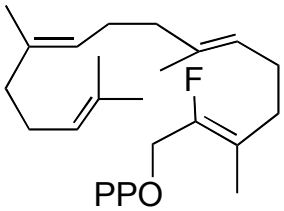
Table S2. Volume Analysis of Substrates, Products, and Active Sites of Terpene Synthases

Enzyme	Abbrev.	Origin	Terpene	PDB code	Closed con-formation	No. of carbons	Substrate V_{vdW} (A^3)	Product V_{vdW} (A^3)	Cavity V_{vdW} (A^3)
Isoprene synthase	ISPS	plant	hemi-	3N0G	model	5	94	94	173
Limonene synthase	LS	plant	mono-	2ONG	X-ray	10	167	155	174
Bornyl diphosphate synthase	BPPS	plant	mono-	1N20	X-ray	10	167	152	224
5-epi-aristolochene synthase	5EAS	plant	sesqui-	5EAT	X-ray	15	240	224	363
epi-isozizane synthase	EIZS	bacteria	sesqui-	3KB9	X-ray	15	240	217	261
trichodiene synthase	TDS	fungi	sesqui-	1JFG	X-ray	15	240	224	281
aristolochene synthase	ARS	fungi	sesqui-	2OA6	X-ray	15	240	224	279
taxadiene synthase	TXS	plant	di-	3P5R	model	20	306	290	579
squalene-hopene cyclase	SHC	bacteria	tri-	1SQC	X-ray	30	475	434	946
oxidosqualene cyclase	OSC	human	tri-	1W6K	X-ray	30	482	450	616

—— SUPPLEMENTARY FIGURES ——



13-aza-13,14-dihydrocopalyl diphosphate (ACP)



2-fluorogeranylgeranyl diphosphate (FGP)

Figure S1. Structures of unreactive isoprenoid ligands. 13-aza-13,14-dihydrocopalyl diphosphate (ACP) and 2-fluorogeranylgeranyl diphosphate (FGP) were cocrystallized with TXS (OPP = diphosphate).

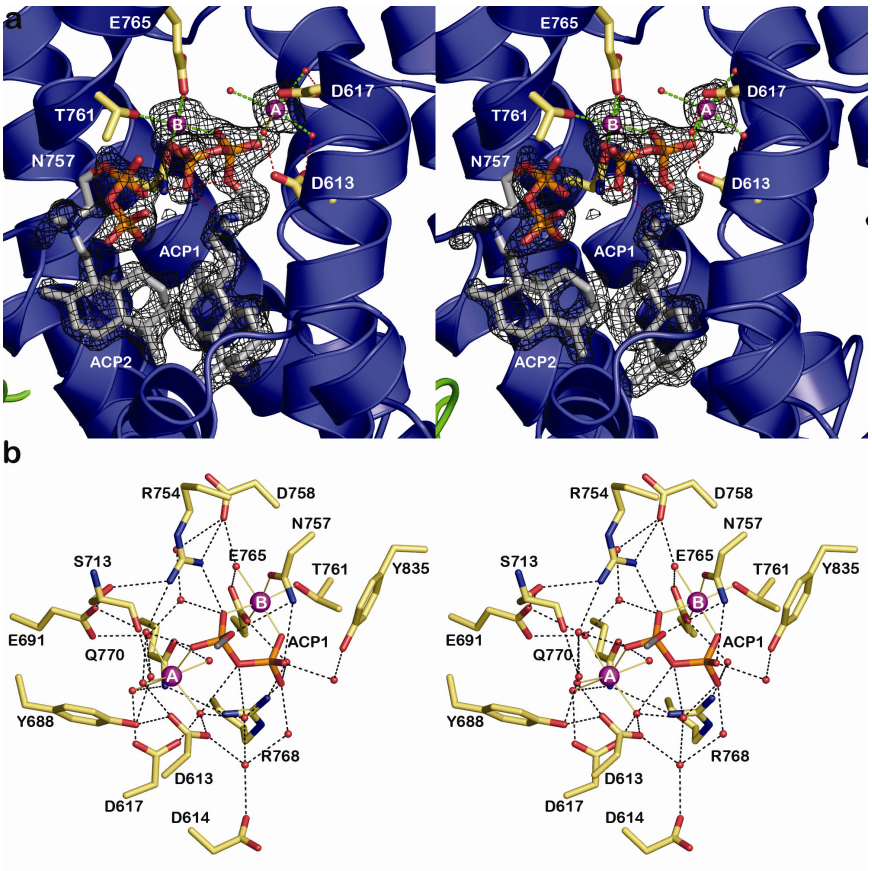


Figure S2. ACP binding in taxadiene synthase. (a) Simulated annealing $|F_o|$-$|F_c|$ omit map in which two molecules of ACP and 2 Mg^{2+} ions are omitted from the structure factor calculation (contoured at 3.0σ). Atoms are color coded as follows: carbon = yellow (TXS) or gray (ACP), nitrogen = blue, oxygen = red, phosphorus = orange; magnesium ions and water molecules appear as purple or red spheres, respectively. Metal coordination and hydrogen bond interactions are indicated by green and black dashed lines, respectively. The first molecule of ACP (ACP1) binds with its isoprenoid moiety nested within the hydrophobic active site and its diphosphate group engaged in metal coordination interactions. However, only two metal ions are bound in this complex: Mg^{2+}_A ion is not directly coordinated by D613 and D617 (the aspartate-rich motif),

but two of the water molecules coordinated to Mg^{2+}_{A} donate hydrogen bonds to D613 and D617; Mg^{2+}_{B} is chelated by the side chains of N757, T761, and E765 (the "NSE/DTE" motif). It is interesting to note that a terminal diphosphate oxygen accepts a hydrogen bond from the 13-aza moiety of ACP1, which is presumed to be protonated and positively charged. This demonstrates that the diphosphate group can engage in basic interactions with molecules bound in the active site. The second molecule of ACP (ACP2) binds in the outer active site in the vicinity of the disordered J-K loop. It is characterized by weaker electron density, which may reflect some degree of disorder. Its diphosphate group makes no metal coordination interactions and makes a hydrogen bond interaction with only R580 (not shown for clarity). (b) Intermolecular interactions in the TXS-ACP1 complex; metal coordination and hydrogen bond interactions are indicated by solid yellow and black dashed lines, respectively. The isoprenoid moiety of ACP is truncated to one carbon (gray) for clarity.

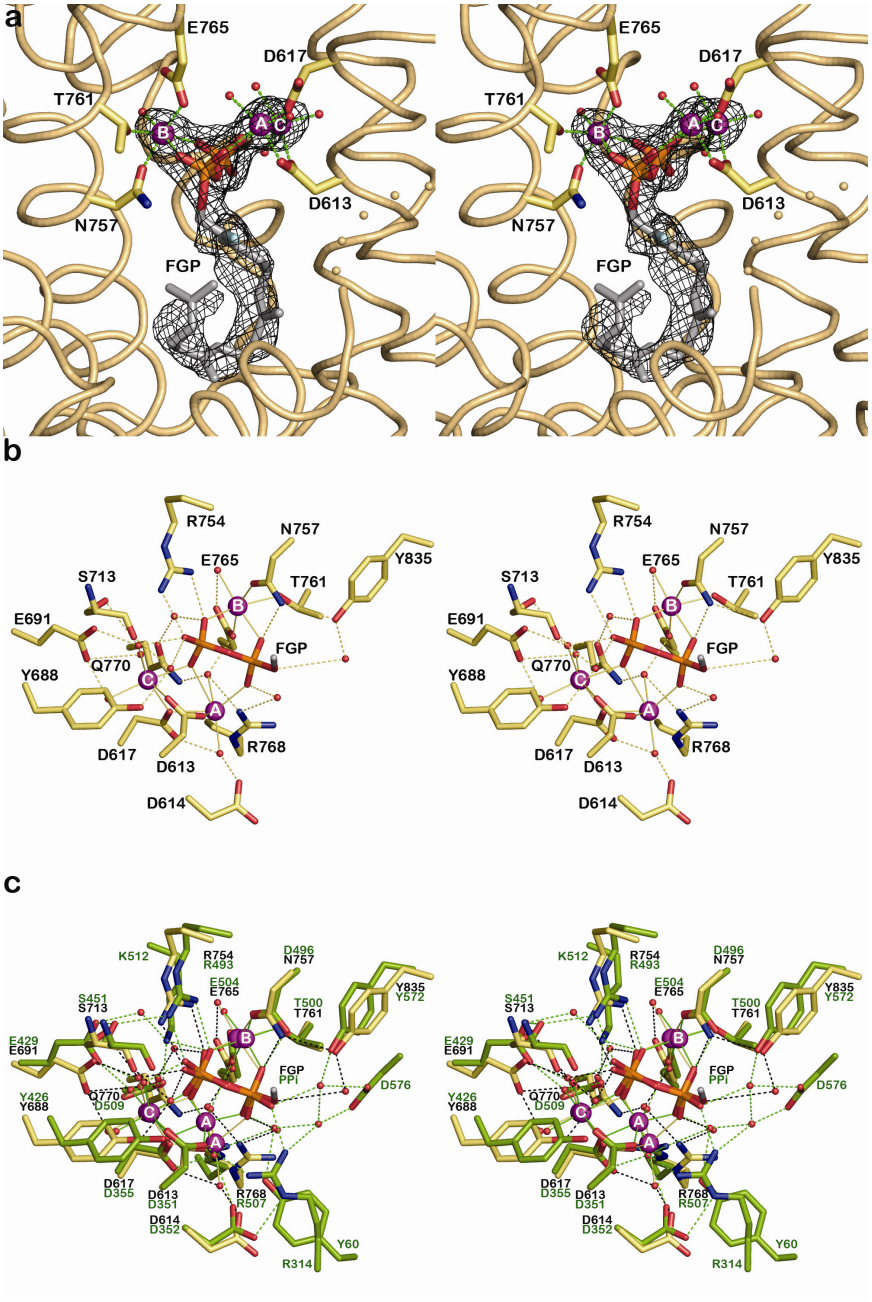


Figure S3. Comparison of the substrate binding in taxadiene synthase with bornyl diphosphate synthase. (a) Stereoview of Figure 3a, simulated annealing |Fo|-|Fc| omit map showing the binding of FGP to TXS (contoured at 3.0σ). (b) Stereoview of Figure 3b, showing

diphosphate molecular recognition in the TXS-FGP complex. (c) Least-squares superposition of the TXS and bornyl diphosphate synthase active sites showing conserved features of diphosphate molecular recognition. Metal coordination and hydrogen bond interactions are indicated by green and black dashed lines, respectively. Atoms are color coded as follows: carbon = yellow (TXS) or green (bornyl diphosphate synthase), nitrogen = blue, oxygen = red, phosphorus = orange; magnesium ions and water molecules appear as purple or red spheres, respectively. Residue labels are black for TXS and green for bornyl diphosphate synthase.

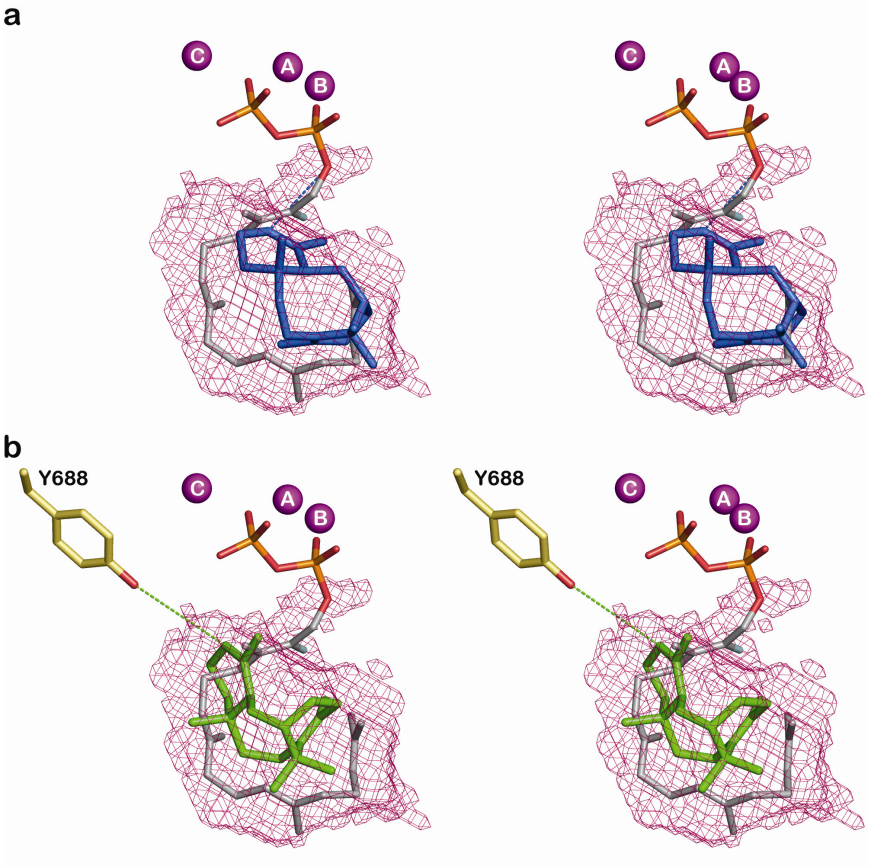


Figure S4. Possible orientations of taxadiene in the active site cavity of taxadiene synthase. (a) Stereoview of Figure 4c, showing one orientation of taxadiene (blue) fits in the active site cavity such the H5β atom of the preceding taxen-4-yl cation would be oriented toward the diphosphate leaving group, suggesting that the PP_i anion could serve as the stereospecific base that terminates the cyclization cascade. Three Mg^{2+} ions and FGP are shown for reference; all protein atoms are omitted for clarity. A blue dashed line indicates the possible trajectory of proton transfer. (b) An alternate fit of taxadiene (green) in the active site cavity orients H5β atom of the preceding taxen-4-yl cation toward the Y688, suggesting that the phenolic hydroxyl group

could also serve as the stereospecific base that terminates the cyclization cascade. A green dashed line indicates the possible trajectory of proton transfer.