

Biosynthesis of Lycosantalol, a *cis*-Prenyl Derived Diterpenoid

Jiachen Zi,^{†,||} Yuki Matsuba,^{‡,||} Young J. Hong,[§] Alana J. Jackson,[†] Dean J. Tantillo,[§] Eran Pichersky,[‡] and Reuben J. Peters^{*,†}

[†]Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011, United States

[‡]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109, United States

[§]Department of Chemistry, University of California, Davis, California 95616, United States

S Supporting Information

ABSTRACT: Terpenoid natural products are generally derived from isoprenyl diphosphate precursors with trans double-bond configuration, and no diterpenoid derived from the cisoid precursor (*Z,Z,Z*)-nerylneryl diphosphate (**1**) has yet been identified. Here further investigation of a terpenoid biosynthetic gene cluster from tomato is reported, which resulted in identification of a biosynthetic pathway from **1**, in a pathway featuring a number of interesting transformations. Compound **1** is first cyclized to a tricyclene core ring structure analogous to that found in α -santalene, with the resulting diterpene termed here lycosantalene (**2**). Quantum chemical calculations indicate a role for the diphosphate anion coproduct in this cyclization reaction. Subsequently, the internal *cis* double bond of the neryl side chain in **2** is then further transformed to an α -hydroxy ketone moiety via an epoxide intermediate (**3**). Oxygen labeling studies indicate **3** undergoes oxidative conversion to lycosantalol (**4**). Thus, in addition to elucidating the cisoid origins of **4**, this work has further provided mechanistic insight into the interesting transformations required for its production.

Until recently, terpenoid biosynthesis was thought to be exclusively derived from transoid linear precursors, with C₁₀ monoterpenes derived from geranyl diphosphate, C₁₅ sesquiterpenes from (*E,E*)-farnesyl diphosphate (*e,e*-FPP), and C₂₀ diterpenes from (*E,E,E*)-geranylgeranyl diphosphate (GGPP). These precursors are formed by short-chain *trans*-isoprenyl diphosphate synthases,¹ with the separate family of *cis*-prenyl transferases (CPTs) thought to be confined to the production of longer chain length isoprenoids.² However, it has now been reported that CPT family members do participate in terpenoid biosynthesis, as tomato produces the linear *cis*-neryl diphosphate (NPP) as a monoterpene precursor³ and (*Z,Z*)-farnesyl diphosphate (*z,z*-FPP) as a sesquiterpene precursor,⁴ and lavender produces the irregular lavandulyl diphosphate as a monoterpene precursor.⁵

The CPTs producing NPP and *z,z*-FPP are orthologues found in domesticated and wild tomato plants; *Solanum lycopersicum* and *Solanum habrochaites*, respectively. This *CPT1* gene is found in a region that also contains the separate, subsequently acting terpene synthases (TPSs).^{3,4} In addition, this region contains a gene encoding another CPT, *CPT2*, along with those encoding several additional TPSs as well as

uncharacterized CYPs, alcohol acyltransferases, and an alcohol oxidase.⁶ Thus, it appears that tomatoes contain a terpenoid biosynthetic gene cluster, which is an unusual but not entirely uncommon occurrence in plant genomes.⁷

Previous characterization of the *S. lycopersicum* CPT2 (*SICPT2*) suggested that this produces (*Z,Z,Z*)-nerylneryl diphosphate (NNPP, **1**),⁸ with subsequent production of an unidentified olefin by a nearby TPS (*SITPS21*).⁶ In addition, there are two CYPs in the cluster, suggesting that these also might function in biosynthesis of the final terpenoid natural product, although one or the other of these appears to be nonfunctional in specific species of *Solanum* (i.e., encode a pseudogene). For example, in *S. lycopersicum* only one CYP appears to be functional, *CYP71D51*, which is present as the only gene in between *SICPT2* and *SITPS21*. While the NPP derived monoterpenes are readily detected from *S. lycopersicum*, the unknown olefin derived from the activity of *SICPT2* and *SITPS21* is not,⁶ suggesting that this might be further transformed by at least the mono-oxygenase activity of *CYP71D51*.

Characterization of the hypothesized *cis*-prenyl derived diterpenoid product of the *SITPS21-CYP71D51-SICPT2* subcluster from *S. lycopersicum* was undertaken using a previously described modular metabolic engineering system.⁹ To verify the production of **1** by *SICPT2*, we expressed *SICPT2* in *Escherichia coli* along with a previously described plasmid that overexpresses key enzymes from the endogenous isoprenoid precursor metabolic pathway, increasing flux into terpenoid production.¹⁰ This enabled isolation of the dephosphorylated derivative (*Z,Z,Z*)-nerylnerol in amounts (ca. 2 mg) sufficient for comprehensive NMR analysis, with comparison of the resulting series of NOE correlations between H₃-17/H-14, H-14/H₃-18, H₃-18/H-10, H-10/H₃-19, H₃-19/H-6, H-6/H₃-20, and H₃-20/H-2, to those observed with the dephosphorylated derivative of GGPP, (*E,E,E*)-geranylgeraniol (Figures S1–S11 and Table S1, Supporting Information (SI)), confirming the *cis* configuration of the internal double-bonds and hence the production of **1** by *SICPT2* (Scheme S1, SI).

Coexpression of *SITPS21* along with *SICPT2* led to production of the previously reported unknown olefin,⁶ which also was isolated in amounts (ca. 3 mg) sufficient for comprehensive NMR analysis (Figures S2 and S12–S17 and Table S2, SI). This compound was found to contain a

Received: August 23, 2014

Published: November 18, 2014

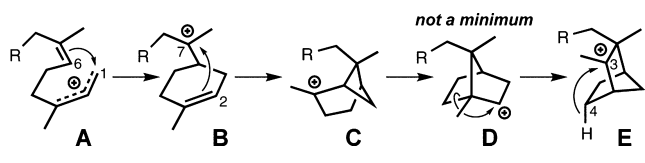


tricyclo[2.2.1.0^{2,6}]heptane ring structure analogous to that found in tricyclene, albeit here with a neryl side-chain. On the basis of the match between this core ring structure and that found in the sesquiterpene α -santalene, the term lycosantalene (**2**) is proposed for the diterpene characterized here (Scheme S2, SI). Given the production of small amounts of 7*R*-(+)- α -santalene by SITPS21 from *z,z*-FPP,⁶ it seems likely that lycosantalene also exhibits the corresponding absolute stereochemistry. Consistent with this hypothesis, **2** also exhibits positive optical rotation: $[\alpha]_D^{23}$, +7.9 (*c* 0.1, CHCl₃).

Interestingly, (+)- α -santalene is the major product of the *z,z*-FPP specific TPS from *S. habrocaites*,⁴ which might suggest some role for the cisoid configuration in formation of the observed ring structure. Use of a *cis*-isoprenyl diphosphate precursor does enable direct formation of the cyclohexanyl (terpinyl-type) carbocation intermediate, which otherwise requires isomerization of the allylic diphosphate from C1 to C3 to enable rotation around the C2–C3 bond (Scheme S3, SI). However, it should be noted that α -santalene also is the major product of *e,e*-FPP specific TPSs from other plants.¹¹

Formation of the strained tricyclo[2.2.1.0^{2,6}]heptane ring structure is of some mechanistic interest and was explored here by quantum chemical calculations (QCC; see SI for details),^{12–18} the application of which to terpene cyclization has recently been experimentally validated.¹⁹ The reaction proceeds through a series of intermediates, with the initial allylic carbocation formed by diphosphate ionization (A) undergoing C1–C6 cyclization to a terpinyl-type carbocation (B), which undergoes C2–C7 cyclization to a bicyclic pinyl-type carbocation (C) that rearranges to a bornyl-type secondary carbocation (D), a transition state rather than intermediate, and is further rearranged to a camphyl-type carbocation (E) that is predicted to undergo concerted C3–C4 ring closure/C4 deprotonation to form the tricyclene core ring structure (Scheme 1). Notably, some support for the suggested stereoselective loss of the pro-*R*-hydrogen in tricyclene formation has been previously reported.²⁰

Scheme 1. Carbocation Series for Tricyclene Ring Forming Reaction Indicated by QCC (R = H, Prenyl, or Neryl for Mono-, Sesqui-, or Diterpene Cyclization, Respectively)



QCC analysis in the absence of the pyrophosphate (PPi) coproduct indicated that the conversion of cation B to cation E is concerted (but asynchronous), while inclusion of PPi led to a stepwise mechanism (in which cation D is still not a minimum), due to stabilization of particular carbocations by the anionic PPi. However, such stabilization required substantial shifts in the relative orientation of the PPi and various carbocationic intermediates (particularly B and C relative to the initial A and terminal E; see SI for details). Given the tight binding of the diphosphate/PPi observed in TPS cocrystal structures,²¹ such reorientation seems unlikely to occur in the restricted context of the active site. Intriguingly, inclusion of PPi in the QCC analysis does seem to bias the reaction toward cyclopropane formation as the C3–C2–C4 angle in E is significantly more acute (79° versus 95°),

shortening the distance between C3–C4 from 2.28 to 2.01 Å (Figure 1 and Schemes S4 and S5, SI). Thus, beyond orienting

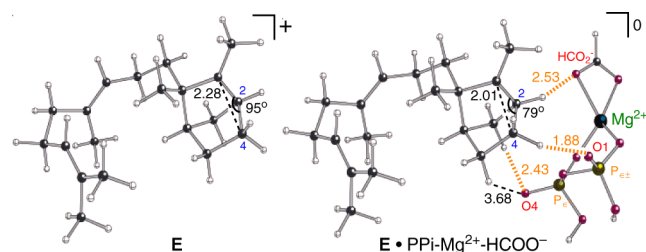
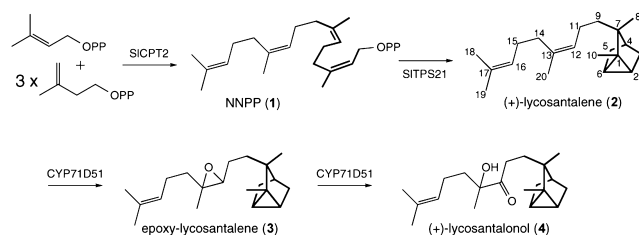


Figure 1. QCC structure (left, B3LYP/6-31+G(d,p); right, B3LYP/6-31G(d); distances in Å) for cation E in the absence (left) or presence (right) of the PPi coproduct.

the substrate to enable primary and secondary cyclization, simply restricting the conformational freedom of the olefinic carbocation intermediates relative to the PPi coproduct may help explain the ability of certain TPSs to form tricyclene structures. Such restriction may be enabled by increased hydrocarbon chain length (i.e., size of the C7 substituent, due to mass effects and/or increased interaction with the TPS), as SITPS21 quite specifically produces **2** (C₂₀; R = neryl), as any other product being formed must be present as less than 3% of **2**. By contrast, the previously characterized α -santalene (C₁₅; R = prenyl) synthases are less specific (i.e., produce significant amount of other sesquiterpenes), and the only molecularly characterized monoterpene synthase that produces tricyclene (C₁₀; R = H) only yields this as a minor product.²²

To characterize the ability of CYP71D51 to react with **2** we obtained a synthetic gene, codon optimized and N-terminally modified as previously shown to enable functional expression of plant microsomal CYP in *E. coli*,^{23–30} and coexpressed this (*sCYP71D51*) along with SICPT2 and SITPS21, as well as a plant CYP reductase to provide the requisite high-energy electrons. In addition to **2**, two new products were observed in these cultures, with masses suggestive of successive oxygenation reactions, i.e., product **3** had an apparent molecular ion of *m/z* = 288 versus 272 for **2**, while that for **4** was 304 (Scheme 2 and

Scheme 2. Biosynthesis of **4** from **1** via **2** and **3** Catalyzed by SICPT2, SITPS21, and CYP71D51



Figures S18 and S24, SI), consistent with the calculated molecular masses for a diterpene olefin (272 Da), oxy derivative (288 Da), and dioxy derivative (304 Da). Compounds **3** and **4** were isolated in amounts (ca. 0.7 and 1.5 mg, respectively) sufficient for comprehensive NMR analysis, which indicated that **3** is the $\Delta^{12,13}$ epoxide derivative of **2** (Figures S2 and S19–23 and Table S3, SI), while **4** contains a C13,12 α -hydroxy ketone (Figures S25–S30 and Table S4, SI) and also exhibits positive optical rotation; $[\alpha]_D^{23}$, +13.8 (*c* 0.2, CHCl₃). Purified **3** fed to CYP71D51 *in vitro* led to the production of **4**

(Figure S31, SI), indicating the intermediacy of **3** in biosynthesis of **4**, for which the term lycosantalanol is proposed here.

Production of the α -hydroxy ketone moiety in **4** by CYP71D51 was probed by labeling experiments using $^{18}\text{O}_2$. In vitro reactions run under $^{18}\text{O}_2$ with **2** as substrate led to production of doubly labeled **4** (Figure S32, SI), demonstrating that both oxygens are inserted by CYP71D51. Formation of the α -hydroxy ketone in **4** from the epoxide found in **3** can be envisioned as proceeding via the addition of oxygen to either C12 or C13. In vitro reactions run under $^{18}\text{O}_2$ with **3** as substrate led to production of singly labeled **4**, MS fragmentation (Figure S33, SI) indicates **4** arises from the addition of oxygen to C12 of **3**, suggesting formation of either a epoxide-hemiketal or 12-hydroperoxy-13-ol intermediate that goes on to form the observed α -hydroxy ketone.

The studies reported here illuminate the biosynthesis of (+)-lycosantalanol (**4**), whose production is encoded by a *SITPS21-CYP71D51-SICPT2* subcluster of a larger terpenoid biosynthetic gene cluster in the genome of *S. lycopersicum* (tomato). Notably, **4** is derived from a *cis*-prenyl precursor, specifically the NNPP (**1**) produced by SICPT2. The use of **1** enables direct C1–C6 cyclization in the formation of (+)-lycosantalene (**2**) catalyzed by SITPS21. QCC analysis of this reaction indicates an important role for the steric restrictions imposed on the pyrophosphate coproduct by SITPS21 in production of the strained tricyclic core ring structure of **2**. CYP71D51 catalyzed formation of the α -hydroxy ketone found in **4** was shown to proceed via epoxy-lycosantalene (**3**), with subsequent addition of oxygen to C12. Thus, this work has further provided mechanistic insight into the interesting transformations required for the production of **4** from the cisoid precursor **1** (Scheme 2).

■ ASSOCIATED CONTENT

Supporting Information

Experimental methods, sequence of the synthetic gene for CYP71D51, detailed QCC results, supplemental schemes and figures, references, and additional details of QCC analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*rjpeters@iastate.edu

Author Contributions

^{||}These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from the NIH (GM076324 to R.J.P.) and NSF (IOS-1025636 to E.P. and CHE-0957416 and CHE-030089 [supercomputing resources via XSEDE] to D.J.T.).

■ REFERENCES

- (1) Bohlmann, J.; Gershenzon, J. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10402.
- (2) Kellogg, B. A.; Poulter, C. D. *Curr. Opin. Chem. Biol.* **1997**, *1*, 570.
- (3) Schillmiller, A. L.; Chauvinhold, I.; Larson, M.; Xu, R.; Charbonneau, A. L.; Schmidt, A.; Wilkerson, C.; Last, R. L.; Pichersky, E. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10865.

- (4) Sallaud, C.; Rontein, D.; Onillon, S.; Jabes, F.; Duffe, P.; Giacalone, C.; Thoraval, S.; Escoffier, C.; Herbette, G.; Leonhardt, N.; Causse, M.; Tissier, A. *Plant Cell* **2009**, *21*, 301.

- (5) Demissie, Z. A.; Erland, L. A.; Rheault, M. R.; Mahmoud, S. S. *J. Biol. Chem.* **2013**, *288*, 6333.

- (6) Matsuba, Y.; Nguyen, T. T.; Wiegert, K.; Falara, V.; Gonzales-Vigil, E.; Leong, B.; Schafer, P.; Kudrna, D.; Wing, R. A.; Bolger, A. M.; Usadel, B.; Tissier, A.; Fernie, A. R.; Barry, C. S.; Pichersky, E. *Plant Cell* **2013**, *25*, 2022.

- (7) Osbourn, A. *Plant Physiol.* **2010**, *154*, 531.

- (8) Akhtar, T. A.; Matsuba, Y.; Chauvinhold, I.; Yu, G.; Lees, H. A.; Klein, S. E.; Pichersky, E. *Plant J.* **2013**, *73*, 640.

- (9) Cyr, A.; Wilderman, P. R.; Determan, M.; Peters, R. J. *J. Am. Chem. Soc.* **2007**, *129*, 6684.

- (10) Morrone, D.; Lowry, L.; Determan, M. K.; Hershey, D. M.; Xu, M.; Peters, R. J. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1893.

- (11) Jones, C. G.; Moniodis, J.; Zulak, K. G.; Scaffidi, A.; Plummer, J. A.; Ghisalberti, E. L.; Barbour, E. L.; Bohlmann, J. *J. Biol. Chem.* **2011**, *286*, 17445.

- (12) Tantillo, D. J. *Nat. Prod. Rep.* **2011**, *28*, 1035.

- (13) Tantillo, D. J. *Nat. Prod. Rep.* **2013**, *30*, 1079.

- (14) Hong, Y. J.; Tantillo, D. J. *Org. Biomol. Chem.* **2010**, *8*, 4589.

- (15) Hong, Y. J.; Tantillo, D. J. *Chem. Commun.* **2012**, *48*, 1571.

- (16) Weitman, M.; Major, D. T. *J. Am. Chem. Soc.* **2010**, *132*, 6349.

- (17) Major, D. T.; Weitman, M. *J. Am. Chem. Soc.* **2012**, *134*, 19454.

- (18) This is Part 9 in our series on diterpene-forming carbocation rearrangements. Part 8: Hong, Y. J.; Tantillo, D. J. *Helv. Chim. Acta* **2014**, *97*, 1475–1480. Part 7: Hong, Y. J.; Tantillo, D. J. *Nature Chem.* **2014**, *6*, 104–111.

- (19) Zu, L.; Xu, M.; Lodewyk, M. W.; Cane, D. E.; Peters, R. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2012**, *134*, 11369.

- (20) Pyun, H. J.; Wagschal, K. C.; Jung, D. I.; Coates, R. M.; Croteau, R. *Arch. Biochem. Biophys.* **1994**, *308*, 488.

- (21) Gao, Y.; Honzatko, R. B.; Peters, R. J. *Nat. Prod. Rep.* **2012**, *29*, 1153.

- (22) Falara, V.; Akhtar, T. A.; Nguyen, T. T.; Spyropoulou, E. A.; Bleeker, P. M.; Chauvinhold, I.; Matsuba, Y.; Bonini, M. E.; Schillmiller, A. L.; Last, R. L.; Schuurink, R. C.; Pichersky, E. *Plant Physiol.* **2011**, *157*, 770.

- (23) Zi, J.; Peters, R. J. *Org. Biomol. Chem.* **2013**, *11*, 7650.

- (24) Wang, Q.; Hillwig, M. L.; Okada, K.; Yamazaki, K.; Wu, Y.; Swaminathan, S.; Yamane, H.; Peters, R. J. *J. Biol. Chem.* **2012**, *287*, 6159.

- (25) Morrone, D.; Chen, X.; Coates, R. M.; Peters, R. J. *Biochem. J.* **2010**, *431*, 337.

- (26) Swaminathan, S.; Morrone, D.; Wang, Q.; Fulton, D. B.; Peters, R. J. *Plant Cell* **2009**, *21*, 3315.

- (27) Wang, Q.; Hillwig, M. L.; Peters, R. J. *Plant J.* **2011**, *65*, 87.

- (28) Wang, Q.; Hillwig, M. L.; Wu, Y.; Peters, R. J. *Plant Physiol.* **2012**, *158*, 1418.

- (29) Wu, Y.; Hillwig, M. L.; Wang, Q.; Peters, R. J. *FEBS Lett.* **2011**, *585*, 3446.

- (30) Wu, Y.; Wang, Q.; Hillwig, M. L.; Peters, R. J. *Biochem. J.* **2013**, *454*, 209.