ORGANIC LETTERS

2000
Vol. 2, No. 15
2257−2259

Arabidopsis thaliana LUP1 Converts Oxidosqualene to Multiple Triterpene Alcohols and a Triterpene Diol

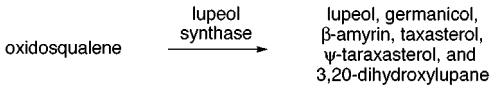
Michael J. R. Segura, Michelle M. Meyer, and Seiichi P. T. Matsuda*

Department of Chemistry and Department of Biochemistry and Cell Biology, Rice University, 6100 South Main Street, Houston, Texas 77005

matsuda@rice.edu

Received May 4, 2000

ABSTRACT



The *Arabidopsis thaliana LUP1* gene encodes an enzyme that converts oxidosqualene to pentacyclic triterpenes. Lupeol and *β*-amyrin were previously reported as LUP1 products. Further investigation described here uncovered the additional products germanicol, taraxasterol, *ψ*-taraxasterol, and 3,20-dihydroxylupane. These results suggest that the 80 known $C_{30}H_{50}O$ compounds that are structurally consistent with being oxidosqualene cyclase products may be derived from fewer than 80 enzymes and that some $C_{30}H_{52}O_2$ compounds may be direct cyclization products of oxidosqualene.

The structural diversity of triterpenoids[1] is generated by multiple pathways for cyclization, oxidation, and derivatization of either squalene or oxidosqualene. Oxidosqualene cyclases convert oxidosqualene to cyclic triterpenoids and thereby produce triterpenoid skeletal diversity. Over 80 $C_{30}H_{50}O$ triterpenes are structurally consistent with being oxidosqualene cyclase products.[2] However, the biological origins of most triterpenoid skeletons remain unstudied because the cyclases have been recalcitrant to purification from natural sources. Most triterpenoids are isolated from plants, which usually contain multiple oxidosqualene cyclase activities. Although several plant oxidosqualene cyclases have been purified,[3] it is often difficult to obtain large enough amounts of these membrane proteins for extensive analyses. Preparations containing single oxidosqualene cyclases are now available from recombinant expression systems, which can provide enough activity for de novo characterization of minor products. We describe here preparative-scale expression and incubation experiments to identify minor triterpene products of *Arabidopsis thaliana* lupeol synthase.[4]

The *Arabidopsis* lupeol synthase was expressed in the yeast lanosterol synthase mutant SMY8,[5] which cannot cyclize oxidosqualene and consequently lacks background activity. A homogenate of this strain converted oxidosqualene to products that were chromatographically separated into two fractions: a triterpene alcohol mixture that was inseparable by conventional silica gel chromatography and a more polar component. A preparative scale reaction (400 mg racemic oxidosqualene) was performed to provide enough material to characterize the minor products. The triterpene alcohols were acetylated, and a portion of the acetates was fractionated by argentation HPLC.[6] Three pure triterpene acetates were obtained, and comparing their ^{13}C NMR spectra with published values indicated that the products (Scheme 1) were lupeol (**1**), *β*-amyrin (**2**), and taraxasterol (**4**). Germanicol (**3**) and *ψ*-taraxasterol (**5**) were identified from a mixture of the corresponding acetates. Additional NMR experiments

(1) Connolly, J. D.; Hill, R. A. *Nat. Prod. Rep.* **1999**, *16*, 221−240, and references therein.

(2) Matsuda, S. P. T. In *Biochemical Principles and Mechanisms of Biosynthesis and Biodegradation of Polymers*; Steinbüchel, A., Ed.; Wiley-VCH: Weinheim, 1998; pp 300−307.

(3) (a) Abe, I.; Sankawa, U.; Ebizuka, Y. *Chem. Pharm. Bull.* **1992**, *40*, 1755−1760. (b) Abe, I.; Ebizuka, Y.; Seo, S.; Sankawa, U. *FEBS Lett.* **1989**, *249*, 100−104. (c) Abe, I.; Sankawa, U.; Ebizuka, Y. *Chem. Pharm. Bull.* **1989**, *37*, 536−538. (d) Abe, I.; Ebizuka, Y.; Sankawa, U. *Chem. Pharm. Bull.* **1988**, *36*, 5031−5034.

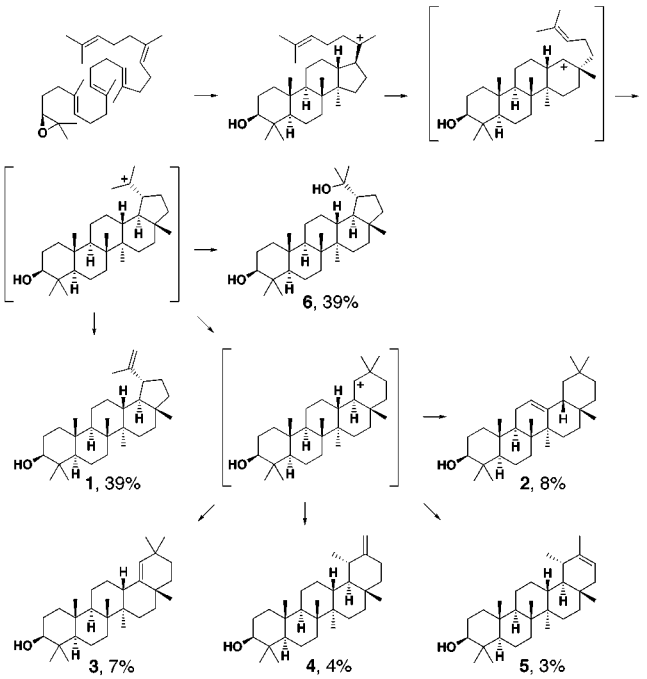
(4) Herrera, J. B. R.; Bartel, B.; Wilson, W. K.; Matsuda, S. P. T. *Phytochemistry* **1998**, *49*, 1905−1911.

(5) Corey, E. J.; Matsuda, S. P. T.; Baker, C. H.; Ting, A. Y.; Cheng, H. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 327−331.

(6) Ruan, B.; Shey, J.; Gerst, N.; Wilson, W. K.; Schroepfer, G. J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11603−11608.

10.1021/ol006016b CCC: $19.00 © 2000 American Chemical Society
Published on Web 07/06/2000

Scheme 1. LUP1 Converts Oxidosqualene to at Least Five Different $C_{30}H_{50}O$ Triterpene Alcohols and a $C_{30}H_{52}O_2$ Diol[8]



(^{1}H, ^{13}C, DEPT, COSYDEC, HSQC, and HMBC) confirmed the structures, and some assignments were revised.[7] The polar fraction was shown to be composed of 3,20-dihydroxylupane (**6**)[9] by similar NMR experiments. Although none of these compounds have been described from *Arabidopsis*, the presence of an enzyme that forms these compounds suggests they are probably *Arabidopsis* metabolites. Lupeol and β-amyrin have been shown to be enzymatic oxidosqualene cyclase products in crude homogenates.[10] β-Amyrin synthases[11] from *Panax ginseng* and lupeol synthases[12] from *Olea europaea* and *Taraxacum officinale* are known. These enzymes form single products and are phylogenetically distinct from LUP1.[12] The present study provides the first demonstration of direct enzymatic conversion of oxidosqualene to germanicol, taraxasterol, ψ-taraxasterol, and 3,20-dihydroxylupane.

All six compounds can be derived from the lupeyl cation (Scheme 1). Deprotonation from C-29 provides lupeol, and quenching with water forms 3,20-dihydroxylupane. Expanding the E ring provides the germanicyl cation, which can deprotonate directly from C-18 to form germanicol or undergo hydride shifts and deprotonation at C-12 to form β-amyrin. A methyl shift to the taraxasteryl cation followed by deprotonation from C-30 or C-21 yields taraxasterol and ψ-taraxasterol, respectively. Such broad product diversity is unprecedented in oxidosqualene cyclases. Although traces of parkeol (~1%) have been isolated from some incubations with cycloartenol synthases,[13] no byproducts have been reported from lanosterol synthases. β-Amyrin and lupeol synthases[11,12] are known that generate single products. Some biological role for LUP1 is implied by *Arabidopsis* maintaining the *LUP1* gene without accumulating mutations that abolish encoded enzymatic activity. If generating lupeol were the primary role of LUP1, it should evolve to attain greater accuracy as other lupeol synthases have done. LUP1 product heterogeneity could be useful for several biological functions. Pure compounds form crystals more readily than mixtures do, and LUP1 may provide triterpenes for a role (such as epicuticular wax) in which an amorphous solid is physiologically preferable. Alternatively, product heterogeneity could generate a diverse array of defense compounds relatively efficiently.

Although enzymes that convert a single substrate to multiple products are unusual, examples are known in terpene

(7) Lupeol and β-amyrin were characterized in the previous study.[5] **Germanicol**: selected ^{1}H NMR signals (500 MHz, $CDCl_3$, 25 °C) δ 0.733 (3H, s, H-27), 0.844 (3H, s, H-24), 0.849 (3H, s, H-23), 0.904 (3H, s, H-25), 0.937 (3H, s, H-29), 0.944 (3H, s, H-30), 1.017 (3H, s, H-28), 1.078 (3H, s, H-26), 2.046 (3H, s, acetate), 4.485 (1H, m, H-3α), 4.862 (1H, m, H-19); ^{13}C NMR (125 MHz, $CDCl_3$, 25 °C) δ 14.55 (C-27), 16.08 (C-26), 16.52 (C-24), 16.76 (C-25), 18.14 (C-6), 21.12 (C-11), 21.32 (acetate) 23.69 (C-2), 25.25 (C-28), 26.18 (C-12), 27.51 (C-15), 27.91 (C-23), 29.18 (C-30), 31.36 (C-29), 32.35 (C-20), 33.33 (C-21), 34.34 (C-17), 34.52 (C-7), 37.14 (C-10), 37.36 (C-22), 37.69 (C-16), 37.82 (C-4), 38.39 (C-13), 38.62 (C-1), 40.76 (C-8), 43.32 (C-14), 51.13 (C-9), 55.58 (C-5), 80.96 (C-3), 129.77 (C-19), 142.67 (C-18), 171.02 (acetate). **ψ-Taraxasterol**: selected ^{1}H NMR signals (500 MHz, $CDCl_3$, 25 °C) δ 0.733 (3H, s, H-28), 0.844 (3H, s, H-24), 0.854 (3H, s, H-23), 0.879 (3H, s, H-25), 0.948 (3H, s, H-27), 0.989 (3H, d, H-29), 1.045 (3H, s, H-26), 1.635 (3H, s, H-30), 2.046 (3H, s, acetate) 4.485 (1H, m, H-3α), 5.261 (1H, d, H-21); ^{13}C NMR (125 MHz, $CDCl_3$, 25 °C) δ 14.69 (C-27), 16.04 (C-26), 16.35 (C-25), 16.52 (C-24), 17.70 (C-28), 18.18 (C-6), 21.32 (acetate), 21.61 (C-11), 21.62 (C-30), 22.53 (C-29), 23.69 (C-2), 27.02 (C-15), 27.59 (C-12), 27.94 (C-23), 34.16 (C-7), 34.38 (C-17), 36.32 (C-19), 36.69 (C-16), 37.01 (C-10), 37.79 (C-4), 38.44 (C-1), 39.20 (C-13), 41.08 (C-8), 42.17 (C-22), 42.33 (C-14), 48.68 (C-18), 50.33 (C-9), 55.38 (C-5), 80.99 (C-3), 118.88 (C-21), 139.86 (C-20), 171.02 (acetate). **Taraxasterol**: selected ^{1}H NMR signals (500 MHz, $CDCl_3$, 25 °C) δ 0.843 (3H, s, H-24), 0.850 (3H, s, H-23), 0.855 (3H, s, H-28), 0.877 (3H, s, H-25), 0.927 (3H, s, H-27), 1.021 (3H, s, H-26), 1.021 (3H, d, H-29), 2.046 (3H, s, acetate), 4.485 (1H, m, H-3α), 4.609 (2H, m, H-30); ^{13}C NMR (125 MHz, $CDCl_3$, 25 °C) δ 14.72 (C-27), 15.88 (C-26), 16.33 (C-25), 16.49 (C-24), 18.18 (C-6), 19.47 (C-28), 21.32 (acetate), 21.46 (C-11), 23.69 (C-2), 25.49 (C-29), 25.62 (C-21), 26.15 (C-12), 26.64 (C-15), 27.94 (C-23), 33.99 (C-7), 34.53 (C-17), 37.05 (C-10), 37.80 (C-4), 38.29 (C-16), 38.44 (C-1), 38.86 (C-22), 39.16 (C-13), 39.38 (C-19), 40.92 (C-8), 42.04 (C-14), 48.65 (C-18), 50.40 (C-9), 55.45 (C-5), 80.98 (C-3), 107.11 (C-30), 154.67 (C-20), 171.02 (acetate).

(8) Relative amounts of $C_{30}H_{50}O$ triterpenes were determined by GC-FID quantitation after acetylation (<1% error). The diol content was determined by integrating characteristic signals for lupeol and 3,20-dihydroxylupane in the crude product ^{1}H NMR spectrum (~5% error).

(9) **3*β*, 20-Dihydroxylupane**: selected ^{1}H NMR signals (500 MHz, $CDCl_3$, 25 °C) δ 0.764 (3H, s, H-24), 0.810 (3H, s, H-28), 0.841 (3H, s, H-25), 0.957 (3H, s, H-27), 0.973 (3H, s, H-23), 1.059 (3H, s, H-26), 1.122 (3H, s, H-29/30), 1.225 (3H, s, H-29/30), 3.198 (1H, m, H-3α); ^{13}C NMR (125 MHz, $CDCl_3$, 25 °C) δ 14.83 (C-27), 15.38 (C-24), 16.15 (C-25), 16.15 (C-26), 18.33 (C-6), 19.21 (C-28), 21.38 (C-11), 24.76 (C-29/30), 27.38 (C-2), 27.56 (C-15), 27.98 (C-23), 28.74 (C-21), 29.05 (C-12), 31.54 (C-29/30), 34.54 (C-7), 35.55 (C-16), 37.07 (C-10), 37.44 (C-13), 38.68 (C-1), 38.84 (C-4), 40.20 (C-22), 41.34 (C-8), 43.51 (C-14), 44.64 (C-17), 48.29 (C-18), 49.93 (C-19), 50.27 (C-9), 55.17 (C-5), 73.50 (C-20), 79.00 (C-3).

(10) (a) Corey, E. J.; Ortiz de Montellano, P. R. *J. Am. Chem. Soc.* **1967**, *89*, 3362−3363. (b) Barton, D. H. R.; Jarman, T. R.; Watson, K. C.; Widdowson, D. A.; Boar, R. B.; Damps, K. *J. Chem. Soc., Perkin Trans. 1* **1975**, 1134−1138.

(11) (a) Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Eur. J. Biochem.* **1998**, *256*, 238−244. (b) Kushiro, T.; Shibuya, M.; Ebizuka, Y. In *Towards Natural Medicine Research in the 21st Century*, *Excerpta Medica International Congress Series 1157*; Ageta, H., Aimi, N., Ebizuka, Y., Fujita, T., Honda, G., Eds.; Elsevier Science: Amsterdam, 1998; pp 421−428.

(12) Shibuya, M.; Zhang, H.; Endo, A.; Shishikura, K.; Kushiro, T.; Ebizuka, Y. *Eur. J. Biochem.* **1999**, *266*, 302−307.

(13) (a) Hart, E. A.; Hua, L.; Darr, L. B.; Wilson, W. K.; Pang, J.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **1999**, *121*, 9887−9888. (b) Godzina, S. M.; Lovato, M. A.; Meyer, M. M.; Foster, K. A.; Wilson, W. K.; Gu, W.; de Hostos, E. L.; Matsuda, S. P. T. *Lipids* **2000**, *36*, 249−255.

biosynthesis. The major product of *Alicyclobacillus acidocaldarius* squalene-hopene cyclase[14] is the triterpene diploptene, but significant amounts of diplopterol (22%) and five minor triterpene hydrocarbons (<1% each) are also present.[15] Monoterpene and sesquiterpene cyclases can be even more prolific; γ-humulene synthase converts farnesyl diphosphate to 52 sesquiterpene hydrocarbons.[16] The discovery of an oxidosqualene cyclase that makes multiple products has important implications for triterpene biosynthesis. It is now unnecessary to postulate a dedicated enzymatic activity for all triterpene skeletons, some of which could be minor products of other enzymes.

The biosynthesis of 3,20-dihydroxylupane has not been studied in any of the six different plants[17] from which it has been isolated, and the *Arabidopsis* route is not necessarily universal. Hydrating free lupeol is in principle a plausible alternative. Like *Arabidopsis*, all of these plants are members of Asteridae or Rosidae, which form a monophyletic group.[18] The close relationship of plants that contain 3,20-dihydroxylupane suggests that they may share the biosynthetic route shown in Scheme 1.

Most triterpenes have an oxygen atom at C-3 (these are probably oxidosqualene cyclase products that retain the C-3 oxygen from the starting epoxide). Most triterpenes also contain other oxygen atoms, some of which are introduced by mechanisms related to that shown in Scheme 1. Squalene-hopene cyclases[19] consistently form water-quenched species as minor products, and dammarenediol is formed similarly by quenching the dammarenyl cation.[20] The mechanism is also prevalent in sesquiterpene biosynthesis. Patchouli alcohol,[21] *epi*-cedrol,[22] and *epi*-cubenol[23] arise by quenching an intermediate carbocation formed by cyclizing farnesyl diphosphate. Figure 1 shows the naturally occurring $C_{30}H_{52}O_2$ triterpene diols (3,18-neohopanediol (**8**),[24] 3,14-taraxeranediol (**9**),[25] 3,20-taraxastanediol (**10**),[26] 3,22-stictanediol (**11**),[27] 3,21-gammaceranediol (**12**),[28] litsomentol (**13**),[29] and 13,17,21-polypodatriendiol (**14**)[30]). All of these diols have an alcohol where a carbocation plausibly exists during cyclization or rearrangement, and they could result from water quenching an intermediate carbocation. The experiments with LUP1 described here indicate that direct formation of $C_{30}H_{52}O_2$ triterpene diols from oxidosqualene in this manner must be considered a plausible alternative to indirect routes to these compounds.

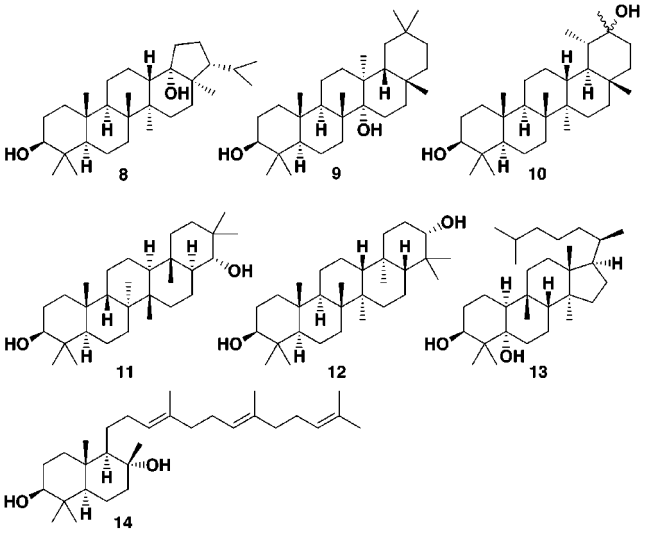


Figure 1. Naturally occurring $C_{30}H_{52}O_2$ diols that could be formed by quenching an intermediate carbocation with water.

Acknowledgment. We thank Jennifer B. R. Herrera for useful discussions and for acetylating the triterpene alcohol fraction and William K. Wilson for valuable advice regarding NMR. M.J.R.S. was a Robert A. Welch Scholar and was supported by NIH training grant T32 GM08362. M.M.M was an American Society of Pharmacognosy Undergraduate Fellow. The National Institutes of Health (AI41598) and the Robert A. Welch Foundation (C-1323) funded this research.

Supporting Information Available: 500 MHz ^{1}H and 125 MHz ^{13}C NMR spectra for 3,20-dihydroxylupane, taraxasteryl acetate, and a mixture of germanicyl acetate and ψ-taraxasteryl acetate. This material is available free of charge at http://pubs.acs.org.

OL006016B

(14) Ochs, D.; Kaletta, C.; Entian, K.-D.; Beck-Sickinger, A.; Poralla, K. *J. Bacteriol.* **1992**, *174*, 298−302.
(15) Pale-Grosdemange, C.; Feil, C.; Rohmer, M.; Poralla, K. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2237−2240.
(16) Steele, C. L.; Crock, J.; Bohlmann, J.; Croteau, R. *J. Biol. Chem.* **1998**, *273*, 2078−2089.
(17) (a) Tsichritzis, F.; Jakupovic, J. *Phytochemistry* **1990**, *29*, 3173−3187. (b) Ulubelen, A.; Topcu, G.; Lotter, H.; Wagner, H.; Eris, C. *Phytochemistry* **1994**, *36*, 413−415. (c) Gearien, J. E.; Klein, M. *J. Pharm. Sci.* **1975**, *64*, 104−108. (d) Hui, W.-H.; Li, M.-M. *Phytochemistry* **1977**, *16*, 111−112. (e) Dantanarayana, A. P.; Kumar, N. S.; Sultanbawa, M. U. S.; Balasubramaniam, S. *J. Chem. Soc., Perkin Trans. 1* **1981**, 2717−2723. (f) Yuruker, A.; Orjala, J.; Sticher, O.; Rali, T. *Phytochemistry* **1998**, *48*, 863−866.
(18) Pearson, L. C. *The Diversity and Evolution of Plants*; CRC Press: Boca Raton, 1995.
(19) Kannenberg, E. L.; Poralla, K. *Naturwissenschaften* **1999**, *86*, 168−176.
(20) Kushiro, T.; Ohno, Y.; Shibuya, M.; Ebizuka, Y. *Biol. Pharm. Bull.* **1997**, *20*, 292−294.
(21) Munck, S. L.; Croteau, R. *Arch. Biochem. Biophys.* **1990**, *282*, 58−64.
(22) (a) Hua, L.; Matsuda, S. P. T. *Arch. Biochem. Biophys.* **1999**, *369*, 208−212. (b) Mercke, P.; Crock, J.; Croteau, R.; Brodelius, P. E. *Arch. Biochem. Biophys.* **1999**, *369*, 213−222.
(23) Cane, D. E.; Ke, N. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 105−107.
(24) Achari, B.; Pal, A.; Pakrashi, S. C. *Tetrahedron Lett.* **1975**, 4275−4278.
(25) Hui, W.-H.; Li, M.-M. *J. Chem. Soc., Perkin Trans. 1* **1976**, 23−30.
(26) Anjaneyulu, V.; Prasad, K. H.; Ravi, K.; Connolly, J. D. *Phytochemistry* **1985**, *24*, 2359−2367.
(27) Chin, W. J.; Corbett, R. E.; Heng, C. K.; Wilkins, A. L. *J. Chem. Soc., Perkin Trans. 1* **1973**, 1437−1446.
(28) Tanaka, R.; Matsunaga, S. *Phytochemistry* **1992**, *31*, 3535−3539.
(29) Govindachari, T. R.; Viswanathan, N.; Mohamed, P. A. *Tetrahedron* **1971**, *27*, 4991−5009.
(30) Boar, R. B.; Couchman, L. A.; Jaques, A. J.; Perkins, M. J. *J. Am. Chem. Soc.* **1984**, *106*, 2476−2477.