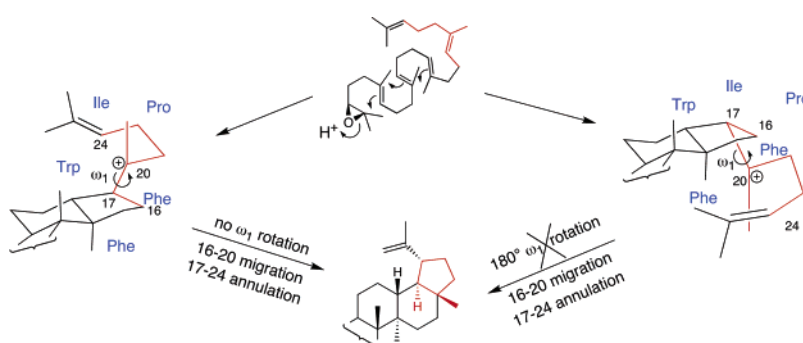# Structure and Reactivity of the Dammarenyl Cation: Configurational Transmission in Triterpene Synthesis

Quanbo Xiong,[†] Flavio Rocco,[‡] William K. Wilson,[†] Ran Xu,[§] Maurizio Ceruti,*,[‡] and Seiichi P. T. Matsuda*,[†,§]

Department of Biochemistry and Cell Biology and Department of Chemistry, Rice University, 6100 Main Street, Houston, Texas 77005, and Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via Pietro Giuria 9, 10125 Torino, Italy

matsuda@rice.edu; maurizio.ceruti@unito.it

The dammarenyl cation (**13**) is the last common intermediate in the cyclization of oxidosqualene to a diverse array of secondary triterpene metabolites in plants. We studied the structure and reactivity of **13** to understand the factors governing the regio- and stereospecificity of triterpene synthesis. First, we demonstrated that **13** has a 17β side chain in *Arabidopsis thaliana* lupeol synthase (LUP1) by incubating the substrate analogue (18*E*)-22,23-dihydro-20-oxaoxidosqualene (**21**) with LUP1 from a recombinant yeast strain devoid of other cyclases and showing that the sole product of **21** was 3β-hydroxy-22,23,24,25,26,27-hexanor-17β-dammaran-20-one. Quantum mechanical calculations were carried out on gas-phase models to show that the 20-oxa substitution has negligible effect on substrate binding and on the activation energies of reactions leading to either C17 epimer of **13**. Further molecular modeling indicated that, because of limited rotational freedom in the cyclase active site cavity, the C17 configuration of the tetracyclic intermediate **13** can be deduced from the angular methyl configuration of the pentacyclic or 6−6−6−6 tetracyclic product. This rule of configurational transmission aided in elucidating the mechanistic pathway accessed by individual cyclases. Grouping of cyclases according to mechanistic and taxonomic criteria suggested that the transition between pathways involving 17α and 17β intermediates occurred rarely in evolutionary history. Two other mechanistic changes were also rare, whereas variations on cation rearrangements evolved readily. This perspective furnished insights into the phylogenetic relationships of triterpene synthases.

Introduction

Triterpenoids comprise a complex family of natural products that have a wide range of structural, regulatory, ecochemical, and medicinal properties.[1] Nature synthesizes over 100 different triterpene skeletons from squalene or 2,3-oxidosqualene by variations on three fundamental pathways (Scheme 1).[1a,2,3] In many prokaryotes, squalene (**1**) is cyclized via the 17α-deoxydammarenyl cation (**3**)

[†] Department of Biochemistry and Cell Biology, Rice University.
[‡] Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino.
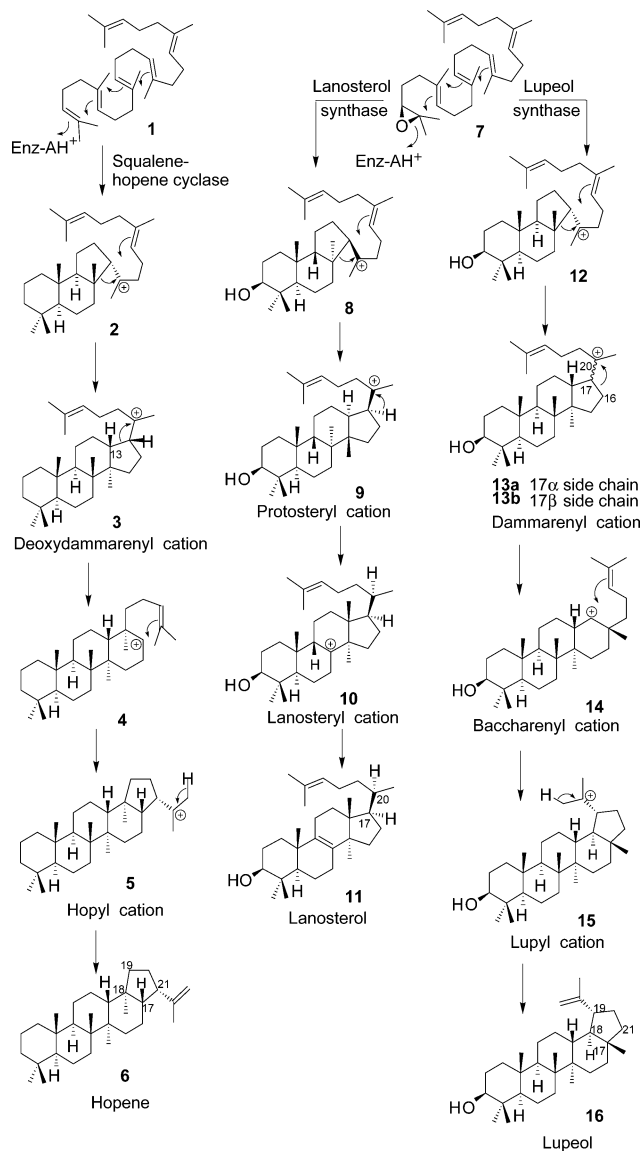[§] Department of Chemistry, Rice University.

(1) (a) Nes, W. R.; McKean, M. L. *Biochemistry of steroids and other isopentenoids*; University Park Press: Baltimore, MD, 1977. (b) Connolly, J. D.; Hill, R. A. *Nat. Prod. Rep.* **2001**, *18*, 560−578. (c) Harrewijn, P.; van Oosten, A. M.; Piron, P. G. M. *Natural terpenoids as messengers: a multidisciplinary study of their production, biological functions and practical applications*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001. (d) Hostettmann, K.; Marston, A. *Saponins*; Cambridge University Press: New York, 1995; Chapter 7. (e) Waller, G. R.; Yamasaki, K. *Saponins Used in Traditional and Modern Medicine*; Plenum Press: New York, 1996.

SCHEME 1. Three Fundamental Pathways of (Oxido)squalene Cyclization

through application of basic rules governing the reactivity of cations. Nevertheless, much greater complexity than originally envisioned has been revealed by recent studies[6] on molecular modeling,[7] protein crystallography,[8,9] biomimetic synthesis,[10] mechanism,[3,11] natural product isolation,[3,12] evolutionary origins,[13] mutant cyclases,[14] inhibitors,[15] substrate analogues,[16] and characterization of cloned cyclases.[17] Molecular modeling and experiment have cast doubt on the seminal hypothesis that cyclization is a concerted process.[2] These doubts were reinforced by a recent X-ray structure of squalene-hopene cyclase (SHC) indicating that squalene is not fully folded in the active site.[8] Another longstanding misconception was that

(6) Reviews: (a) Cyclization: Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189−2206. (b) Abe, I.; Prestwich, G. D. In *Isoprenoids Including Carotenoids and Steroids*; Cane, D. E., Ed.; Vol. 2 in Comprehensive Natural Products Chemistry; Barton, D., Nakanishi, K., Meth-Cohn, O., Eds.; Pergamon/Elsevier: Amsterdam, The Netherlands, 1999; pp 267−298. (c) Hoshino, T.; Sato, T. *Chem. Commun.* **2002**, 291−301. (d) Dev, S. In *Handbook of Triterpenoids*; Dev, S., Ed.; CRC Press: Boca Raton, FL, 1989; Vol. 1, Chapter 1, pp 7−53. (e) Ruzicka, L. *Pure Appl. Chem.* **1963**, *6*, 493−523. (f) Ourisson, G.; Rohmer, M.; Poralla, K. *Annu. Rev. Microbiol.* **1987**, *41*, 301−333. (g) Kannenberg, E. L.; Poralla, K. *Naturwissenschaften* **1999**, *86*, 168−176.

(7) (a) Jenson, C.; Jorgensen W. L. *J. Am. Chem. Soc.* **1997**, *119*, 10846−10854. (b) Rajamani, R.; Gao, J. *J. Am. Chem. Soc.* **2003**, *125*, 12768−12781. (c) Gao, D.; Pan, Y.-K. *J. Am. Chem. Soc.* **1998**, *120*, 4045−4046. (d) Hess, B. A., Jr. *J. Am. Chem. Soc.* **2002**, *124*, 10286−10287. (e) Hess, B. A., Jr. *Org Lett.* **2003**, *5*, 165−167. (f) Hess, B. A., Jr. *Collect. Czech. Chem. Commun.* **2003**, *68*, 202−210. (g) Hess, B. A., Jr. *Collect. Czech. Chem. Commun.* **2004**, *69*, 261−266. (h) Hess, B. A., Jr.; Smentek, L. *Org. Lett.* **2004**, *6*, 1717−1720. (i) Vrcek, V.; Saunders, M.; Kronja, O. *J. Org. Chem.* **2003**, *68*, 1859−1866. (j) Nishizawa, M.; Yadav, A.; Imagawa, H.; Sugihara, T. *Tetrahedron Lett.* **2003**, *44*, 3867−3870. (k) Schulz-Gasch, T.; Stahl, M. *J. Comput. Chem.* **2003**, *24*, 741−753. (l) Hess, B. A., Jr. *Eur. J. Org. Chem.* **2004**, 2739−2242 (Erratum 2747). (m) Vrcek, V.; Vrcek, I. V.; Kronja, O. *Croat. Chem. Acta* **2001**, *74*, 801−813 (n) Fouillet, C. C. J.; Mareda, J. *J. Mol. Struct.* (*THEOCHEM*) **2002**, *589−590*, 7−26. (o) Hess, B. A., Jr.; Smentek, L. *Mol. Phys.* **2004**, *102*, 1201−1206. (p) Nishizawa, M.; Yadav, A.; Iwamoto, Y.; Imagawa, H. *Tetrahedron* **2004**, *60*, 9223−9234.

(8) Reinert, D. J.; Balliano, G.; Schulz, G. E. *Chem. Biol.* **2004**, *11*, 121−126.

(9) (a) Thoma, R.; Schulz-Gasch, T.; D'Arcy, B.; Benz, J.; Aebi, J.; Dehmlow, H.; Hennig, M.; Stihle, M.; Ruf, A. *Nature* **2004**, *432*, 118−122. (b) Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J. Mol. Biol.* **1999**, *286*, 175−187. (c) Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811−1815 (d) Lenhart, A.; Weihofen, W. A.; Pleschke, A. E. W.; Schulz, G. E. *Chem. Biol.* **2002**, *9*, 639−645. (e) Zhu, X.; Heine, A.; Monnat, F.; Houk, K. N.; Janda, K. D.; Wilson, I. A. *J. Mol. Biol.* **2003**, *329*, 69−83.

(10) Bartlett, P. A. In *Asymmetric Synthesis*; Morrison, J. D., Ed.; Academic Press: New York, 1984; Vol. 3, pp 341−409. (b) Johnson, W. S.; Telfer, S. J.; Cheng, S.; Schubert, U. *J. Am. Chem. Soc.* **1987**, *109*, 2517−2518. (c) Johnson, W. S.; Bartlett, W. R.; Czeskis, B. A.; Gautier, A.; Lee, C. H.; Lemoine, R.; Leopold, E. J.; Luedtke, G. R.; Bancroft, K. J. *J. Org. Chem.* **1999**, *64*, 9587−9595. (d) van Tamelen, E. E. *Acc. Chem. Res.* **1968**, *1*, 111−120. (e) Demuth, M.; Heinemann, C. *J. Am. Chem. Soc.* **1999**, *121*, 4894−4895. (f) Zhou, S.-Z.; Sey, M.; De Clercq, P. J.; Milanesio, M.; Viterbo, D. *Angew. Chem., Int. Ed.* **2000**, *39*, 2861−2863. (g) Berckmoes, K.; De Clercq, P. J.; Viterbo, D. *J. Am. Chem. Soc.* **1995**, *117*, 5857−5858. (h) Huang, A. X.; Xiong, Z.; Corey, E. J. *J. Am. Chem. Soc.* **1999**, *121*, 9999−10003. (i) Corey, E. J.; Wood, H. B., Jr. *J. Am. Chem. Soc.* **1996**, *118*, 11982−11983. (j) Nishizawa, M.; Takenaka, H.; Hirotsu, K.; Higuchi, T.; Hayashi, Y. *J. Am. Chem. Soc.* **1984**, *106*, 4290−4291.

(11) (a) Corey, E. J.; Staas, D. D. *J. Am. Chem. Soc.* **1998**, *120*, 3526−3527. (b) Johnson, W. S.; Lindell, S. D.; Steele, J. *J. Am. Chem. Soc.* **1987**, *109*, 5852−5853. (c) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1289−1296. (d) Corey, E. J.; Virgil, S. C. *J. Am. Chem. Soc.* **1991**, *113*, 4025−4026. (e) Corey, E. J.; Virgil, S. C.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Singh, V.; Sarshar, S. *J. Am. Chem. Soc.* **1995**, *117*, 11819−11820.

(12) (a) Connolly, J. D.; Hill, R. A. *Nat. Prod. Rep.* **2003**, *20*, 640−659. (b) Rosa-Putra, S.; Nalin, R.; Domenach, A.-M.; Rohmer, M. *Eur. J. Biochem.* **2001**, *268*, 4300−4306. (c) Pale-Grosdemange, C.; Feil, C.; Rohmer, M.; Poralla, K. *Angew. Chem., Int. Ed.* **1998**, *37*, 2237−2240.

to hop-22(29)-ene (hopene, **6**) and related pentacyclic triterpenes. In eukaryotes, triterpene precursors of membrane sterols are produced by cyclization of (3*S*)-2,3-oxidosqualene (**7**) to the protosteryl cation (**9**), which undergoes a series of 1,2-hydride and methyl shifts, culminating in deprotonation to lanosterol (**11**) or cycloartenol. Secondary metabolites in plants account for the greatest diversity of triterpene skeletons, almost all of which arise from the dammarenyl cation (**13a** or **13b**). This key intermediate can have many fates,[3] such as rearrangement to the baccharenyl (**14**) and lupyl (**15**) cations to furnish lupeol (**16**).

The classic scheme for understanding triterpene biosynthesis[1a,4,5] accounts for known triterpene skeletons[3]

(2) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812−2833.

(3) Xu, R.; Fazio, G. C.; Matsuda, S. P. T. *Phytochemistry* **2004**, *65*, 261−291.

(4) Eschenmoser, A.; Ruzicka, L.; Jeger, O.; Arigoni, D. *Helv. Chim. Acta* **1955**, *38*, 1890−1904.

(5) Stork, G.; Burgstahler, A. W. *J. Am. Chem. Soc.* **1955**, 77, 5068−5077.

the protosteryl cation (**9**) has a 17α side chain[18] (sterol numbering[19]). Corey and co-workers[11d,e] showed that the side chain is 17β by demonstrating that lanosterol synthase converts oxidosqualene analogue **17** to **18** (Scheme 2).

In pentacyclic triterpenes such as lupeol, the C17 configuration of the precursor dammarenyl cation **13** is unknown. This intermediate is commonly drawn with a 17β side chain (**13b**), albeit without experimental evidence or theoretical justification. The C17 stereochem-
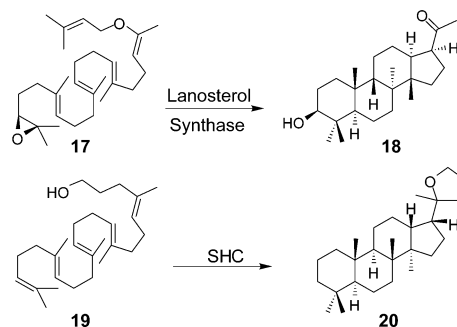
(13) (a) Rohmer, M.; Bouvier, P.; Ourisson, G. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 847−851. (b) Ourisson, G.; Nakatani, Y. *Chem. Biol.* **1994**, *1*, 11−23. (c) Harlampidis, K.; Trojanowska, M.; Osbourn, A. E. *Adv. Biochem. Eng. Biotechnol.* **2002**, *75*, 31−49. (d) Pearson, A.; Budin, M.; Brocks, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15352−15357. (e) Qi, X.; Bakht, S.; Leggett, M.; Maxwell, C.; Melton, R.; Osbourn, A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8233−8238. (f) Brocks, J. J.; Logan, G. A.; Buick, R.; Summons, R. E. *Science* **1999**, *285*, 1033−1036. (g) Bode, H. B.; Zeggel, B.; Silakowski, B.; Wenzel, S. C.; Reichenbach, H.; Müller, R. *Mol. Microbiol.* **2003**, *47*, 471−481. (h) Nes, W. R. *Lipids* **1974**, *9*, 596−612. (i) Volkman, J. K. *Org. Geochem.* **2005**, *36*, 139−159.

(14) (a) Reviews: ref 6c and: Segura, M. J. R.; Jackson, B. E.; Matsuda, S. P. T. *Nat. Prod. Rep.* **2003**, *20*, 304−317. (b) Sato, T.; Abe, T.; Hoshino, T. *Chem. Commun.* **1998**, 2617−2618. (c) Sato, T.; Hoshino, T. *Biosci. Biotechnol. Biochem.* **1999**, *63*, 2189−2198. (d) Wu, T.-K.; Griffin, J. H. *Biochemistry* **2002**, *41*, 8238−8244. (e) Kushiro, T.; Shibuya, M.; Masuda, K.; Ebizuka, Y. *J. Am. Chem. Soc.* **2000**, *122*, 6816−6824. (f) Hoshino, T.; Abe, T.; Kouda, M. *Chem. Commun.* **2000**, 441−442. (g) Hoshino, T.; Shimizu, K.; Sato, T. *Angew. Chem.*, *Int. Ed.* **2004**, *43*, 6700−6703. (h) Wu, T.-K.; Chang, C.-H. *ChemBioChem* **2004**, *5*, 1712−1715. (i) Dang, T.; Prestwich, G. D. *Chem. Biol.* **2000**, *7*, 643−649.

(15) (a) Cattel, L.; Ceruti, M. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 353−373. (b) Viola, F.; Ceruti, M.; Cattel, L.; Milla, P.; Poralla, K.; Balliano, G. *Lipids* **2000**, *35*, 297−303. (c) Mercer, E. I. *Prog. Lipid Res.* **1993**, *32*, 357−416.

(16) (a) Abe, I.; Rohmer, M. *J. Chem. Soc.*, *Perkin Trans. 1* **1994**, 783−791. (b) Stach, D.; Zheng, Y.-F.; Perez, A. L.; Oehlschlager, A. C.; Abe, I.; Prestwich, G. D.; Hartman, P. G. *J. Med. Chem.* **1997**, *40*, 201−209. (c) Abe, I.; Dang, T.; Zheng, Y. F.; Madden, B. A.; Fei, C.; Poralla, K.; Prestwich, G. D. *J. Am. Chem. Soc.* **1997**, *119*, 11333−11334. (d) Abe, I.; Tanaka, H.; Noguchi, H. *J. Am. Chem. Soc.* **2002**, *124*, 14514−14515. (e) Tanaka, H.; Noguchi, H.; Abe, I. *Org. Lett.* **2004**, *6*, 803−804. (f) Tanaka, H.; Noguchi, H.; Abe, I. *Tetrahedron Lett.* **2004**, *45*, 3093−3096. (g) Abe, I.; Sakano, Y.; Sodeyama, M.; Tanaka, H.; Noguchi, H.; Shibuya, M.; Ebizuka, Y. *J. Am. Chem. Soc.* **2004**, *126*, 6880−6881 (h) Abe, I.; Sakano, Y.; Tanaka, H.; Lou, W.; Noguchi, H.; Shibuya, M.; Ebizuka, Y. *J. Am. Chem. Soc.* **2004**, *126*, 3426−3427. (i) Xiong, Q.; Zhu, X.; Wilson, W. K.; Ganesan, A.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **2003**, *125*, 9002−9003. (j) Hoshino, T.; Nakano, S.; Kondo, T.; Sato, T.; Miyoshi, A. *Org. Biomol. Chem.* **2004**, *2*, 1456−1470. (k) Nakano, S.; Ohashi, S.; Hoshino, T. *Org. Biomol. Chem.* **2004**, *2*, 2012−2022. (l) Hoshino, T.; Ohashi, S. *Org. Lett.* **2002**, *4*, 2553−2556. (m) Abe, I.; Tanaka, H.; Noguchi, H. *Baiosaiensu to Indasutori* **2004**, *62*, 587−590; *Chem. Abstr. 2004*, *141*, 312963.

(17) (a) Herrera, J. B.; Bartel, B.; Wilson, W. K.; Matsuda, S. P. T. *Phytochemistry* **1998**, *49*, 1905−1911. (b) Segura, M. J.; Meyer, M. M.; Matsuda, S. P. T. *Org. Lett.* **2000**, *2*, 2257−2259. (c) Husselstein-Muller, T.; Schaller, H.; Benveniste, P. *Plant Mol. Biol.* **2001**, *45*, 75−92. (c) Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Eur. J. Biochem.* **1998**, *256*, 238−244. (d) Shibuya, M.; Adachi, S.; Ebizuka, Y. *Tetrahedron* **2004**, *60*, 6995−7003. (e) Zhang, H.; Shibuya, M.; Yokota, S.; Ebizuka, Y. *Biol. Pharm. Bull.* **2003**, *26*, 642−650. (f) Corey, E. J.; Matsuda, S. P. T.; Bartel, B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11628−11632. (g) Kelly, R.; Miller, S. M.; Lai, M. H.; Kirsch, D. R. *Gene* **1990**, *87*, 177−183. (h) Buntel, C. J.; Griffin, J. H. *J. Am. Chem. Soc.* **1992**, *114*, 9711−9713. (i) Roessner, C. A.; Min, C.; Hardin, S. H.; Harris-Haller, L. W.; McCollum, J. C.; Scott, A. I. *Gene* **1993**, *127*, 149−150. (j) Ochs, D.; Kaletta, C.; Entian, K.-D.; Beck-Sickinger, A.; Poralla, K. *J. Bacteriol.* **1992**, *174*, 298−302. (k) Fazio, G. C.; Xu, R.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **2004**, *126*, 5678−5679. (l) Hayashi, H.; Huang, P.; Takada, S.; Obinata, M.; Inoue, K.; Shibuya, M.; Ebizuka, Y. *Biol. Pharm. Bull.* **2004**, *27*, 1086−1092.

(18) (a) Cornforth, J. W. *Angew. Chem.*, *Int. Ed.* **1968**, *7*, 903−911. (b) Nes, W. R.; Varkey, T. E.; Krevitz, K. *J. Am. Chem. Soc.* **1977**, *99*, 260−262.

(19) Sterol numbering is generally used for cationic intermediates and double bond positions, but "20-oxa" always refers to squalene numbering. Standard carbon atom numbering is used for pentacyclic triterpenes, squalene, and derivatives thereof.

SCHEME 2

istry is apparently lost during D-ring expansion via the baccharenyl cation (**14**) and is not obvious from the product structure.[20]
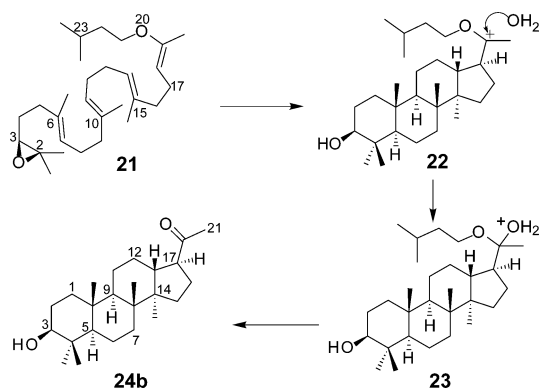
Nature can clearly access both 17α and 17β intermediates. Most known dammarenediols and other unrearranged tetracyclic dammarane derivatives in plants have a 17β side chain configuration inherited from intermediate **13b**,[3] as do tetracycles produced by pentacyclic triterpene synthase mutants.[14e] However, a tetracyclic triterpene from an edible plant has 17α stereochemistry,[21a] which is evidently maintained from **13a**. In addition, many pentacyclic triterpenes in angiosperms, gymnosperms, and ferns appear to be derived from **13a**.[3] Also accessing the 17α-(deoxy)dammarenyl cation (**3** or **13a**) are products of SHC, as signified by several observations,[2,12c,16a,c,e,f] including the conversion by SHC of the truncated hydroxysqualene analogue **19** to the tetrahydrofuran derivative **20**.[14b] The opposite configuration of the protosteryl and deoxydammarenyl cations at C8, C9, C13, C14, and C17 suggests that the chirality in ring C might be conveyed to the new stereocenter at C17. This potential relay of asymmetry and evidence that **13a** is accessed by many plant cyclases, juxtaposed against results favoring **13b** as the mechanistic intermediate, prompted us to study the C17 configuration of dammarenyl cations in triterpene synthesis.

We first determined experimentally the mechanistic pathway used by LUP1, a representative angiosperm cyclase. We showed that LUP1 cyclizes the 20-oxaoxidosqualene analogue **21** to an enzymatic product of 17β configuration and thus accesses **13b**. Molecular modeling allayed concerns about possible distortion of the reaction pathway by the 20-oxa substitution and led to a general rule of configurational transmission consistent with our experimental finding. Application of this rule to individual triterpenes helped to reveal the mechanistic pathway used for cyclization. These results led to a classification of cyclases according to mechanistic criteria and provided insights into evolutionary relationships among triterpene synthases.

(20) **15** could arise from **13b** by D-ring expansion/annulation or from **13a** by 180° side chain rotation about the C17−C20 bond, D-ring expansion via a baccharenyl cation (**14**) with a vertical[29a] p orbital, and annulation onto the β face of the C17 cation.

(21) (a) Isolation of (23*E*)-17α-dammara-20,23-diene-3β,25-diol from *Borassus flabellifer*: Révész, L.; Hiestand, P.; La Vecchia, L.; Naef, R.; Naegeli, H.-U.; Oberer, L.; Roth, H.-J. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1521−1526. See also: (b) Scholz, D.; Baumann, K.; Grassberger, M.; Wolff-Winiski, B.; Rihs, G.; Walter, H.; Meingassner, J. G. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2983−2986.

SCHEME 3

FIGURE 1. Partial ^{1}H NMR spectra (500 MHz, CDCl$_3$) of the crude product from reaction of LUP1 homogenate with **21** (A) or **24a** (B).

Results and Discussion

Experimental Evidence for the 17β-Dammarenyl Cation Intermediate in Lupeol Synthesis. Our investigation of the C17 configuration of the dammarenyl cation required modifying the corresponding experiment with the protosteryl cation. That work[11d,e] involved in vitro incubation of 20-oxaoxidosqualene (**17**) with lysed yeast, whose sole oxidosqualene cyclase is lanosterol synthase. An analogous experiment with lysed plant cells would be difficult to interpret because plants contain cycloartenol synthase (CAS, accessing **9**) as well as various cyclases accessing the dammarenyl cation. We avoided the complication of multiple cyclases by using heterologous expression technology to insert a recombinant *Arabidopsis thaliana* lupeol synthase gene (*LUP1*) into a yeast lanosterol synthase mutant (SMY8) lacking any cyclases.[17a,b] Expression of *LUP1* from the Gal promoter in SMY8 yeast provided high enzymatic activity free of other cyclases.[17a,b]

In vitro incubation of lysed yeast containing recombinant LUP1 with substrate analogue **21**[22] (Scheme 3) gave a polar TLC spot that was absent in a control incubation lacking **21**. This enzymatic product was isolated chromatographically and identified as 3β-hydroxy-22,23,24,-25,26,27-hexanordammaran-20-one (**24b**) by ^{1}H, ^{13}C, and 2D NMR and by GC-MS of the acetate derivative (see the Supporting Information). NOESY interactions between H-17 and the 14α-methyl and between other proton pairs clearly demonstrated the 17β-acetyl stereochemistry. The spectral data were compatible with previous reports of **24b**, which has been isolated from *Euphorbia supina*[23] and prepared by dammar resin oxidation,[21a,24] by dipterocarpol degradation,[25] and as an intermediate in a synthesis of *dl*-dammarenediol.[10c] GC analysis of the acetylated ether-soluble fraction indicated 61% conversion of **21** to **24b**, assuming that only the (*S*)-enantiomer underwent enzymatic cyclization.

Although our reaction and isolation conditions were much milder than the acidic or basic conditions required to epimerize 20-ketosteroids,[11d,26] we rigorously excluded the possibility that **24b** might have been formed from
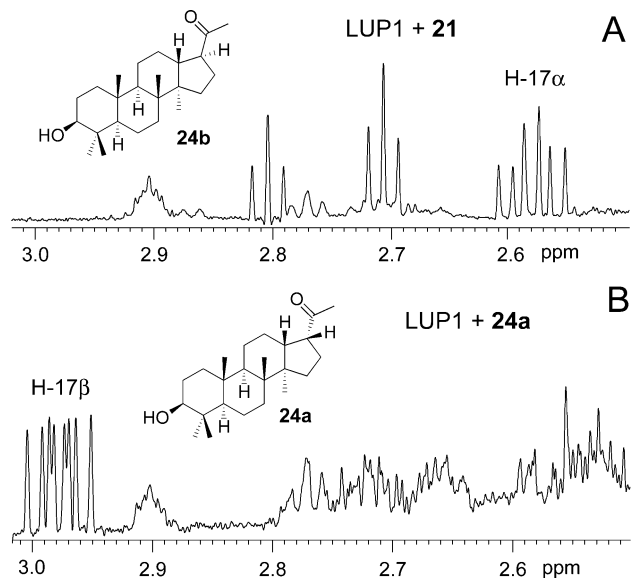
its 17α epimer **24a** during workup. First, we prepared a standard of epimer **24a** by deprotection of its TBDMS ether using 20% aqueous HF in CH$_3$CN. Only slight epimerization occurred at C17 under these conditions, whereas heating with 20 mM *p*-toluenesulfonic acid was necessary for nearly complete conversion of **24a** to **24b**. These observations were compatible with mPW1PW91/6-311+G(2d,p)//B3LYP/6-31G(d) calculations predicting that an equilibrium mixture should contain roughly 0.5% of the 17α-acetyl epimer (see the Supporting Information). If the enzymatic reaction initially produced **24a**, followed by epimerization to **24b**, the isolated product should contain at least ~0.5% **24a**. The 500 MHz ^{1}H NMR spectrum, which resolved distinctive signals of the epimers (e.g. H-17 and H-21), indicated the absence of **24a** at a detection limit of 0.1% of **24b** (Figures 1 and S1). Confirming this result, GC-MS of the acetylated enzymatic reaction product showed no additional component with *m/z* 402, and **24a** was absent in GC analyses (see the Supporting Information). In a further experiment, no detectable C17 epimerization occurred from incubation of yeast homogenate with **24a** (Figure 1); this experiment was done under the same incubation and workup conditions used in the enzymatic conversion of **21** to **24b**, except that no substrate was added. Taken together, these results demonstrate that **24b** is not an isomerized product of its 17α-epimer **24a**.
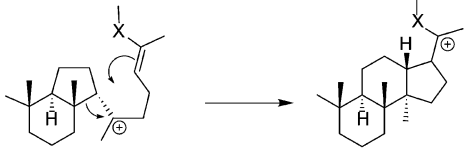
Compound **24b** is rationalized as the hydrolytic product of hemiketal **23**, which is evidently derived from the tetracyclic cation **22** by quenching with water. The presence of water in this region of the active site cavity is supported by the formation of lupane-3β,20-diol as a byproduct of lupeol synthesis.[14e,17b] The absence of a Δ^{24} double bond in **22** precludes direct loss of the side chain as a C$_5$ cation, as occurs with the intermediate from **17**.[11c] The lack of ring-expansion products of **22** is attributable to the stabilization of the C20 cation by electron donation from the adjacent oxygen atom. The good yield of **24b**, the absence of byproducts, and the behavior of 20-oxa substrate analogues with other cyclases[2,11a,b] indicate that

(22) Ceruti, M.; Viola, F.; Dosio, F.; Cattel, L.; Bouvier-Nave, P.; Ugliengo, P. *J. Chem. Soc., Perkin Trans. 1* **1988**, 461−469.
(23) Tanaka, R.; Matsuda, M.; Matsunaga, S. *Phytochemistry* **1987**, *26*, 3365−3366.
(24) Mills, J. S. *J. Chem. Soc.* **1956**, 2196−2202.
(25) Crabbé, P.; Ourisson, G.; Takahashi, T. *Tetrahedron* **1958**, *3*, 279−302.

TABLE 1. Relative Energies in the Formation of 17α (26a, 28a) and 17β (26b, 28b) Dammarenyl Cation Models[a]

	25a	X=C		26a	X=C	17α-SC
	25b	X=C		26b	X=C	17β-SC
	27a	X=O		28a	X=O	17α-SC
	27b	X=O		28b	X=O	17β-SC

	reactant	TS	product
relative enthalpy (kcal/mol)			
25a → 26a (17α)	0.0	7.1	−2.2
25b → 26b (17β)	0.0	8.1	−8.2
27a → 28a (17α, 20-oxa)	0.0	7.3	−15.5
27b → 28b (17β, 20-oxa)	0.0	7.7	−21.6
relative free energy (kcal/mol)			
25a → 26a (17α)	0.0	10.6	2.5
25b → 26b (17β)	0.0	10.9	−4.4
27a → 28a (17α, 20-oxa)	0.0	10.3	−10.6
27b → 28b (17β, 20-oxa)	0.0	10.0	−17.6

[a] Enthalpies and free energies (relative to the reactant) are from mPW1PW91/6-311+G(2d,p)//B3LYP/6-31G(d) calculations. Conformers **25a** and **25b** had very similar enthalpies and free energies, as did **27a** and **27b**: see Table S3, which also includes ZPE corrections and energies from other DFT methods.

24b is not the product of a minor aberrant pathway. The C17 configuration of the dammarenyl cation is frozen in **24b**. Thus, the C17 stereochemistry of **24b** supports the formation of the 17β-dammarenyl intermediate **13b** in lupeol biosynthesis.

Molecular Modeling of Dammarenyl Cation Formation. The possibility that the 20-oxa substitution of **21** might perturb the energetics of cyclization was explored by modeling D-ring formation of the 17α- and 17β-dammarenyl cations with and without the 20-oxa moiety (Table 1). The 20-oxa substitution stabilized the product cation and had a strong exothermic effect on the net reaction energy. This exothermic effect was enhanced in the 17β-dammarenyl cation models, which did not share the unfavorable steric interactions between the C14α and C20 methyl groups of the 17α epimers. By contrast, the C17 configuration and the 20-oxa substitution had little effect on activation energies.

Greater detail was provided by QST3 transition-state[27] and path[28] calculations (Figure 2) showing that the transition state occurs during C-ring expansion, well before ring D has closed. The stabilizing effect of the 20-oxa substitution comes into play only as the annulation process nears completion and thus does not influence the activation energy. The transition states and activation energies closely resemble computational results for protosteryl cation formation by Hess,[7d] who first proposed bridged transition states in C-ring expansion/D-ring formation as an alternative to anti-Markovnikov intermediates.[2,7b,j,p]

(26) (a) Fujimoto, H.; Tanaka, O. *Chem. Pharm. Bull.* **1970**, *18*, 1440−1445. (b) Brunke, E. J. *Tetrahedron* **1979**, *35*, 781−788.

(27) This method is based on a quadratic synchronous transit (QST) approach to finding the transition state: Peng, C.; Schlegel, H. B. *Isr. J. Chem.* **1993**, *33*, 449−454.

(28) Ayala, P. Y.; Schlegel, H. B. *J. Chem. Phys.* **1997**, *107*, 375−384.
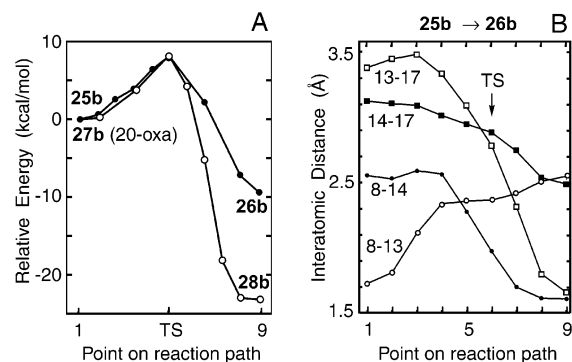
FIGURE 2. Changes in energy (A) and geometry (B) from path calculations modeling 17β-dammarenyl cation formation with and without the 20-oxygen. Energies in panel A are from mPW1PW91/6-311+G(2d,p)//B3LYP/6-31G* calculations. Additional geometry changes and commentary are given in Table S4.

The above energy calculations, together with additional studies,[29] were used to address the concern that interpretation of experiments with 20-oxa substrates should "take into account the steric and electronic perturbations on the cyclization mechanism engendered by the CH₂ to O substitution".[6b] We present evidence that the 20-oxa substitution does not significantly affect enzyme−substrate interactions before or during cyclization or otherwise cause substrate misfolding in the active site cavity. We also show that the energetics of tetracyclization is not perturbed and that no inherent preference is given to a particular tetracyclic C17 epimer. Many of the following arguments apply to protosteryl as well as dammarenyl cation formation.

(1) The strongest evidence against perturbation of enzyme−substrate interactions by the 20-oxa substitution is the high conversion (61%) of the 20-oxa substrate analogue **21** to tetracycle **24** in our in vitro incubation. Consistent with this observation is the weak inhibition by **21** of lanosterol and β-amyrin synthases.[22,30] Computation[31] and intuition suggest that a residue interacting strongly with the vinyl ether would also bind to the 2,3-epoxy group and thus interfere with natural substrate binding. Most polar residues in cyclase active site cavities are encumbered by hydrogen bonds[8,9a−d] and tend to be unavailable for substrate binding. Other modeling results[29b] indicated that the 20-oxa substitution has little effect on the weak CH/π interactions between substrate and aromatic residues. This evidence against distorting effects of the 20-oxa substitution also holds *during* cyclization because the 20-oxa region of the substrate remains uncharged until tetracyclization.

(29) (a) Matsuda, S. P. T.; Wilson, W. K.; Xiong, Q. Manuscript to be submitted for publication (on mechanistic insights into triterpene synthesis from quantum mechanical calculations). (b) Matsuda, S. P. T.; Wilson, W. K.; Xiong, Q. Manuscript to be submitted for publication (on enzymatic stabilization in triterpene synthesis).

(30) The I_{50} values were 80 μM (rat liver microsomes, lanosterol synthase) and 300 μM (β-amyrin synthase from *Pisum sativum*). Values for K_i of **21** and K_m of oxidosqualene were 40 and 30 μM (respectively) for lanosterol synthase and both 250 μM for β-amyrin synthase.[22] Compound **21** is a reversible inhibitor, whereas **17** appears to be an irreversible inhibitor. After tetracyclization of **17**, the terminal prenyl moiety can be lost as an allylic cation that is apparently capable of denaturing the cyclase, as judged by affinity labeling results.[11c]

(31) In a study modeling the interactions of the 2,3-epoxy and 20-oxa functionalities with serine and tyrosine, the epoxide and vinyl ether models showed very similar DFT binding energies.[29b]

SCHEME 4 [a]

C16 migration
Lupeol synthesis

C13 migration
Hopene synthesis
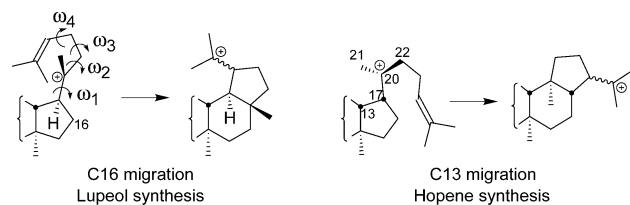
[a] Dihedral Angle Definitions: ω_1, H17−C17−C20−C21; ω_2, C17−C20−C22−C23; ω_3, C20−C22−C23−C24; ω_4, C22−C23−C24−C25.

(2) The 20-oxa substitution does not interfere energetically with the usual folding of the substrate as judged by the similarity of B3LYP/6-31G* geometries for oxidosqualene, **21**, and **17** (all folded for lupeol formation)[29a] and by the small energy penalty for geometries of **21** or **17** with atoms C18 and C22 held in their native relative position in oxidosqualene.[29a]

(3) The similarity of activation energies in Table 1 indicates that the 20-oxa moiety does not give preference to a particular C17 epimer or perturb the energetics of cyclization prior to D-ring closure. This conclusion assumes that product formation is under kinetic control, i.e., that D-ring formation is irreversible. The irreversibility of the 20-oxa pathways is clearly demonstrated by the energetics in Table 1 and supported by the lack of byproducts accompanying **24b**. Interestingly, our bare substrate calculations and those of others[7b,32] suggest some reversibility in 17α-dammarenyl cation formation.[33]

Saturation of the 22−23 double bond also has very limited effect on substrate folding and cyclization of **21** as judged by the ability of 22,23-dihydro-**7** to undergo tetracyclization to products of the dammarenyl cation.[16h] Any misfolding of the terminal isoprene unit of **21** should be inconsequential because this moiety does not participate in cyclization or stabilization of the C20 cation of **21**. The foregoing experimental and theoretical evidence establishes the 17β-dammarenyl cation stereochemistry only for the single enzyme LUP1 of *A. thaliana*. This limited result is generalized in the following studies.

Molecular Modeling of Dammarenyl Cation Reactivity. We studied dammarenyl cation reactivity to understand the process by which the C17 chirality is lost and the C20 stereocenter is created. Our study paralleled that described above for C-ring expansion/D-ring formation, which is analogous to D-ring expansion/E-ring formation. We modeled the reactivity of the dammarenyl cations using the full side chain, in contrast to the model compounds **25−28**, which were truncated at C23 to avoid distracting conformational complexity arising from the lack of a terminal double bond in the 20-oxa substrate.

Side chain flexibility of the four single bonds, corresponding to dihedral angles ω_1, ω_2, ω_3, and ω_4 of Scheme 4, can produce dozens of conformers. Some conformers are folded, and others, such as the initial tetracyclic

species formed in the SHC active site,[8] are in a partially extended conformation. Suitably folded conformers can form ring E directly, whereas extended conformers can either rearrange to 6−6−6−5 tetracyclic products such as euphol, undergo ring expansion/rearrangement to 6−6−6−6 tetracycles, or fold after tetracyclization to form pentacycles.

Only 8 low-energy conformers can form pentacyclic products directly. Dihedrals ω_2 and ω_4 prefer values of roughly +100° and −100°, the ω_2 value determining whether D-ring expansion proceeds by migration of C13 or C16 and the ω_4 value determining whether the resulting isopropyl substituent has an α or β configuration. For given values of ω_2 and ω_4, only one ω_3 rotamer (+*gauche* or −*gauche*) produces a conformer in which C17 is sufficiently close to C24 for facile E ring closure.

Two ω_1 rotamers leading to D-ring expansion are possible; the C20 methyl can be either anti or syn to H17. We propose that only the anti ω_1 rotamer should be considered in enzymatic mechanisms. The anti ω_1 rotamer results directly from D-ring formation as a consequence of the 18*E* double bond geometry of oxidosqualene. Formation of the syn rotamer would require a ca. 180° rotation of the side chain about the C17−C20 bond. Such rotation inside the active site cavity of oxidosqualene cyclases seems sterically improbable,[16l] as the C20 methyl and C22 hydrogens would sweep out a large volume. The crystal structure[8] of 2-azasqualene in SHC indicates the usual snug fit of substrate in the active site as judged by the large number of enzyme−substrate atomic distances of <3.3 Å (Table S5). At the estimated[8] location of the C20 methyl after tetracyclization, residues F601, F605, I261, and W169 form a constricting girdle, and the protruding P263 would interfere with rotation at C22 and C23. A crystal structure of lanosterol in lanosterol synthase[9a] shows similar obstacles, notably F696 and H232. These barriers restrict side chain rotation of the protosteryl cation **9** and thereby enforce the 20*R* stereochemistry for lanosterol.[11d,e]

However, this picture is complicated by the high energy of the partially cyclized substrate (favoring enzyme mobility that might permit rotation), the likelihood of an extended side chain conformation (as observed with SHC[8]), potential stabilization of the C20 cation by aromatic residues, and our ignorance of plant cyclase geometries. Thus, a more extensive argument against 180° rotation is required. As indicated in Table 2, the stereochemistry of the product D−E ring junction is a function of ω_1 rotation, 17α/17β side chain stereochemistry, C13 vs C16 migration, and C17 epimerization. Each of the 8 potential regio- and stereoisomeric products can be formed in two ways. For hopenes derived from the 17α-dammarenyl cation and lupeol skeletons derived from the 17β-dammarenyl cation, trans D−E ring junctions can be formed without ω_1 rotation or C17 epimerization as well as from a rotation−epimerization combination. Cis ring junctions require either C17 epimerization or ω_1 rotation. We consider two cases separately, according to whether the side chain is folded initially.

If the side chain is already folded after tetracyclization, as shown in Scheme 4, then D-ring expansion/E-ring formation can occur directly by an energetically favorable pathway with a modest energy of activation.[29a] Rotation about C17−C20 of one lupeol-type fold to another (mak-

(32) Unfavorable energetics in hopene synthesis were also predicted by Rajamani and Gao,[7b] who suggested that an interplay of thermodynamic and kinetic factors may determine the product distribution.

(33) Our small models did not include stabilization of the C20 cation by the terminal double bond of the substrate. Inclusion of this stabilization, which does not exist for the 24,25-dihydro-20-oxa analogue **21** (or for extended side chain conformers, as observed in SHC[8]), would reduce the net energy differences between the 20-oxa and 20-carbo species by a few kcal/mol.[29a]

TABLE 2. Mechanistic Origins of the D−E Ring Junction Stereochemistry of 6,6,6,6,5 Triterpenes[a]

ω_1 rotation	side chain configuration	C13/C16 migration	C17 epimerization	product structure[b]
0	17α	C13	−	**hopene-βα**
0	17α	C13	+	hopene-αα
0	17α	C16	−	lupeol-αβ
0	17α	C16	+	lupeol-αα
0	17β	C13	−	hopene-αβ
0	17β	C13	+	hopene-ββ
0	17β	C16	−	**lupeol-βα**
0	17β	C16	+	lupeol-ββ
180	17α	C13	−	hopene-ββ
180	17α	C13	+	hopene-αβ
180	17α	C16	−	lupeol-ββ
180	17α	C16	+	**lupeol-βα**
180	17β	C13	−	hopene-αα
180	17β	C13	+	**hopene-βα**
180	17β	C16	−	lupeol-αα
180	17β	C16	+	lupeol-αβ

[a] The ω_1 rotations are very roughly 180° (anti to syn) or 0°. [b] Skeletal type (hopene or lupeol) and stereochemistry of the C17 and C18 substituents (lupeol numbering). The parent lupeol and hopene products are shown in boldface. The lupeol-ββ entries are likely precursors of ursenes.[3]

TABLE 3. Geometries and Relative Energies of Various Side Chain Conformers of the Tetracyclic Hopene Precursor[a]

energy, kcal/mol	ω_1, deg	ω_2, deg	ω_3, deg	ω_4, deg	C13−C17, Å	C16−C17, Å
0.0[b]	144	81	−81	122	1.66	1.55
3.7[c]	164	−1	177	115	1.74	1.54
3.3	26	112	179	115	1.57	1.63
4.4	30	180[d]	−179	115	1.57	1.63
4.9	90[d]	150[d]	176	115	1.58	1.57
4.9	−90[d]	−150[d]	172	113	1.58	1.58

[a] Relative mPW1PW91/6-311+G(2d,p)//B3LYP/6-31G* electron energies without ZPE corrections for conformers of model **31a**. Dihedral angles $\omega_1 - \omega_4$ are defined in Scheme 4. See Table S6 for additional data. [b] Conformer folded for direct E-ring formation. [c] Conformer whose side chain dihedrals resemble those of the X-ray structure of 2-azasqualene in SHC. [d] Torsion angle frozen during geometry optimization.

ing the C20 methyl syn to H17) would entail a convoluted process involving additional rotations about ω_2 and ω_3, and the same considerations apply to hopene-type folding. Although a hopene-type fold could rotate intact to a lupeol-type fold without perturbing ω_2 or ω_3, this would require a cavity so large that the enzyme could not reliably preorganize the substrate for regio- and stereospecific cyclization. Pathways entailing unfolding and refolding are thermodynamically improbable because most rotations of the folded conformer are energetically unfavorable, including rotations of ω_1 from anti to syn (Table 3).

If the side chain is not initially folded, as in SHC,[8] we also argue that 180° rotation of ω_1 does not occur enzymatically. We present several lines of reasoning, none of which is conclusive by itself, but together are compelling. (1) Formation of lupeol or hopene structures via ω_1 rotation would also entail C17 epimerization (Table 2), a combination of two energetically unfavorable processes. (2) If 180° rotation about C17−C20 can occur, then so can the smaller rotation to a cation with its 2p orbital oriented vertically[29a] to facilitate a 1,2-hydride shift from C17 to C20, leading to tetracyclic products such

as euphol. The availability of such alternative pathways leading to tetracyclic products is indicated by the conversion of 22,23-dihydro-2,3-oxidosqualene to euph-7-en-3β-ol and bacchar-12-en-3β-ol by a recombinant β-amyrin synthase.[16h] Thus, a cyclase that permits 180° ω_1 rotation should be quite inaccurate. (3) The active site in the vicinity of C17 and C20 (after tetracyclization) must be constricting in at least one dimension to preorganize (oxido)squalene appropriately for regio- and stereospecific C- and D-ring formation. Even such a modest constriction would probably interfere sterically with the C20 substituents during 180° ω_1 rotation. (4) Additional arguments stemming from steric interactions between the side chain and the 14α-methyl suggest that 180° ω_1 rotation and C16 migration are energetically unfavorable for the 17α-dammarenyl cation (see the Supporting Information).

Taken together, the above lines of reasoning indicate that 180° rotation about C17−C20 of the dammarenyl cation is highly unlikely in enzymatic cyclizations. Our logic is not based on enzymatic stabilization of the C20 cation in a specific position or orientation; such arguments are fundamentally unsound.[29b] The principle of least motion[34] was also avoided because the unsolvated active site cavity permits the tetracyclic cation to continue exploring the potential energy surface until deprotonation can occur. Sinnott[34b] has suggested that least motion effects are linked to solvent disruption. Thus, least motion arguments may be more applicable to biomimetic cyclizations.

The foregoing analysis, indicating a single ω_1 rotamer, leaves 4 folded conformers each for the 17α- and 17β-dammarenyl cations. Geometrical parameters and relative energies of these 8 conformers are summarized in Table 4. Conformers with the side chain folded toward C13 show elongation of the C16−C17 bond due to hyperconjugation with the nominally vacant 2p orbital of the C20 cation. Similar elongation of the C13−C17 bond is observed for side chain conformers folded toward C16. The ω_1 dihedral deviates from anti by ca. 30° for optimal orbital overlap for hyperconjugative cation stabilization, and this geometry reduces steric interactions between the side chain and the C16 or C12 hydrogens. Thus, folding of the side chain facilitates and guides D-ring expansion.

Lupeol, its 19β epimer, hopene, and its 21β epimer (from models **29a**, **29b**, **31a**, **31b**) show favorable enthalpies for E-ring formation. All four structural types are known in nature.[3] According to our mechanistic analysis, the other four conformers (**33a**, **33b**, **35a**, **35b**) would lead to products with a 13,18-syn-CDE ring fusion (**34a**, **34b**, **36a**, **36b**), in which ring D would likely adopt a twist/boat conformation. Nature does not appear to use these conformers, either to produce syn-CDE pentacycles with a trans-DE ring junction[35] or to form 6−6−6−6 tetracycles.[3]

We optimized likely transition-state structures leading to the hopyl and lupyl cations (**5** and **15**). Their geometries, shown in Figure 3, indicate that the transition state for D-ring expansion/E-ring formation, like that of

(34) (a) Hine, J. *Adv. Phys. Org. Chem.* **1977**, *15*, 1−61. (b) Sinnott, M. L. *Adv. Phys. Org. Chem.* **1988**, *24*, 113−204 (see especially p 161).
(35) In hancolupenol and hancokinol formation, the protosteryl cation undergoes D-ring expansion by C16 migration but circumvents the syn-CDE ring fusion through E-ring annulation from the 17α face.[3]

TABLE 4. Calculated Geometries and Enthalpies of Folded Dammarenyl Cation Models Differing in Side Chain Conformation[a,b]

configuration (Me/H/^{i}Pr)[c]	C13 migration				C16 migration			
	17α side chain		17β side chain		17α side chain		17β side chain	
	αβα	αββ	βαα	βαβ	αβα	αββ	βαα	βαβ
	31a → 32a	31b → 32b	35a → 36a	35b → 36b	33a → 34a	33b → 34b	29a → 30a	29b → 30b
C13−C17	1.66	1.67	1.64	1.64	1.56	1.56	1.54	1.54
C16−C17	1.55	1.55	1.56	1.56	1.62	1.62	1.64	1.64
C17−C24	3.33	3.33	3.36	3.41	3.38	3.38	3.43	3.48
C20−C24	3.24	3.14	3.13	3.22	3.24	3.13	3.15	3.21
ω_1	144	151	−163	−160	−144	−144	155	155
ω_2	81	89	−90	−88	−111	−94	94	93
ω_3	−81	−67	66	78	82	67	−70	−78
ω_4	122	−93	96	−121	−123	93	−93	128
ΔH reactant vs **31a** or **29a**	0.0	3.8	−1.5	−1.4	5.6	5.7	0.0	0.3
ΔH product vs **32a** or **30a**	0.0	−5.9	0.5	6.9	23.9	17.7	0.0	6.4
ΔH(product − reactant)	−3.3	−9.7	2.1	8.3	8.6	2.3	−9.8	−3.7

[a] Interatomic distances (Å) and dihedral angles (deg) are from B3LYP/6-31G* reactant geometries (sterol atom numbering). Enthalpies (ΔH, kcal/mol) are from mPW1PW91/6-311+G(2d,p)//B3LYP/6-31G* calculations. Energies from other DFT methods are given in Table S7. [b] The C22−C23 bonds were 1.61−1.62 Å, and the H17−C17 bonds were ca. 1.09 Å. During optimization of **33a**, ω_1 was frozen at −144° to prevent minimization to a −107° angle (hyperconjugation with the C17−H17 bond). [c] Configuration of the angular methyl and hydrogen at the DE ring junction and of the isopropyl cation substituent, respectively.

C-ring expansion/D-ring formation,[7d] occurs during ring expansion and before closure of the new ring. These transition-state structures are cyclopropane/carbonium ions and represent lower energy pathways than the 6−6−6 and 6−6−6−6 secondary carbocations that have often been invoked as intermediates in (oxido)squalene cyclization.[29a]

Configurational Transmission. Our postulate that 180° rotation about the C17−C20 bond does not occur after enzymatic tetracyclization has an important corollary, namely that the configuration of the C20 methyl following D-ring expansion is determined by the C17 stereochemistry of the dammarenyl cation. In the 17β-dammarenyl cation, the C20 methyl is situated on the β face and will become a β substituent after D-ring expansion, regardless of whether C13 or C16 migration occurs (Figure 4). Analogously, the 17α-dammarenyl cation will produce an angular methyl of α configuration. These conclusions are consistent with (1) our experimental results on the cyclization of 20-oxa substrate **21** by lupeol synthase, (2) the accepted C17 stereochemistry for the deoxydammarenyl intermediate **3** in hopene biosynthesis, and (3) a comprehensive survey of the mechanistic pathways leading to triterpene skeletons.[3] Thus, although the C17 stereocenter may be destroyed during D-ring expansion, the configurational information is transmitted to the C20 stereocenter (corresponding to

C17 of lupeol or C18 of hopene). By analogous reasoning, the postulate forbidding 180° rotation about the C17−C20 bond in enzymatic reactions also applies to the protosteryl cation and its (unknown) 17α epimer. Obviously, the postulate does not apply to nonenzymatic reactions.[36]

In our model of concerted D-ring expansion/E-ring formation, the C17−H17 bond remains vertical relative to the plane of the ABCD ring system, thus retaining the original C17 configuration and thereby generating a trans DE ring junction. However, the existence of ursenes and other pentacyclic triterpenes derived from DE-cis intermediates[3] points to the operation of alternative mechanism(s) of E-ring formation, probably involving C17 epimerization via a C17 cation with the C17−H17 bond in the ABCD plane. It might be supposed that 17β-dammarenyl cations undergoing D-ring expansion will always do so by migration of C16 to C20 and that C13 migration will occur only with 17α-dammarenyl cations. Although this appears to be the case in nature[3] and can be rationalized by steric interference disfavoring the unobserved migrations (and a CDE-syn ring fusion in the case of E-ring formation), it is unlikely to be an ironclad

(36) Molecular mechanics calculations suggest that barriers to ω_1 rotation are <4 kcal/mol (Figure S11), but solvent effects could raise these barriers sufficiently to give selectivity in low-temperature reactions.
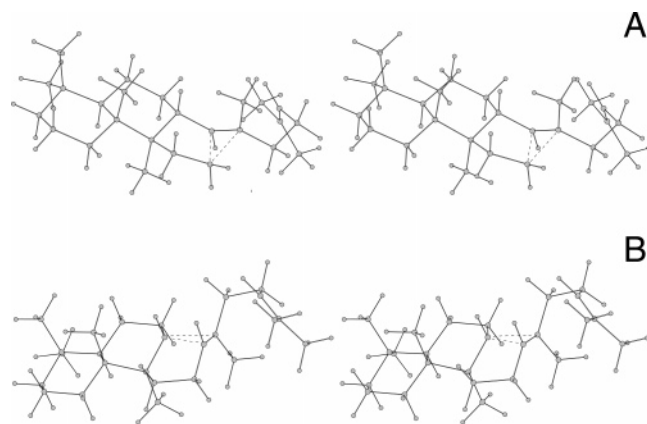
Interatomic distances in Å

TS to lupyl cation (A)		TS to hopyl cation (B)	
C16-C17	2.02	C13-C17	2.19
C16-C20	1.86	C13-C20	1.96
C17-C24	3.16	C17-C24	2.96
C20-C24	3.04	C20-C24	3.06

FIGURE 3. Stereoscopic views of transition state structures for E-ring formation leading to the lupyl (A) or hopyl (B) cation. Geometry optimization was done at the B3LYP/6-31G* level with QST3, and vibrational frequency calculations confirmed the structures to be saddle points.
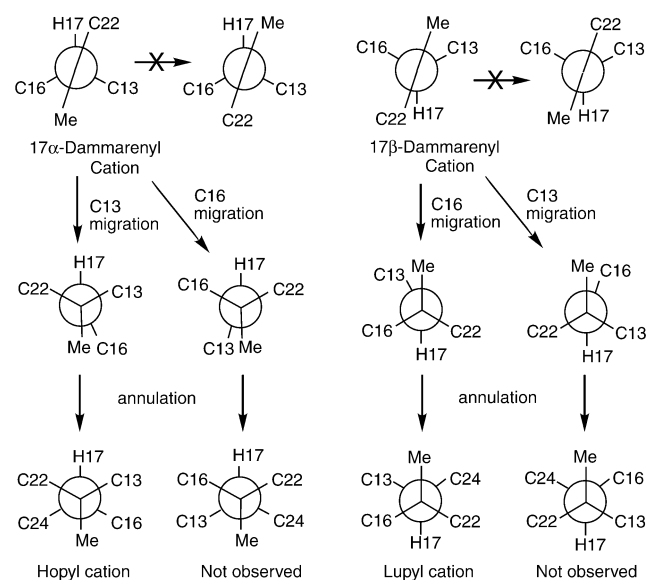

FIGURE 4. Newman projections for D-ring expansion of 17α- and 17β-dammarenyl cations **13a** and **13b**, followed by annulation to form ring E. The view is from C20 to C17 (sterol atom numbering). The unobserved products (with syn-CDE ring fusion) are shown as idealized projections.

rule, as both C13 and C16 migration occur for the protosteryl cation.[3] Because of the reversed C13 configuration in the protosteryl cation, C13 migration may be favored in pentacyclization to avoid a CDE-syn ring fusion, and a 17α-protosteryl cation may prefer C16 migration in gentle biomimetic reactions. In enzymatic cyclization, the reaction path is determined more by substrate folding and the geometry of the active site cavity than by modest differences in gas-phase energies of cationic intermediates.[29b]

We have studied models of many stereoisomers of the lupyl cation that differ in configuration at C17, C18, or

SCHEME 5. Euphol and Tirucallol Could Each Arise from Either 13a or 13b Depending on Whether the Side Chain Rotates Clockwise (cw) or Counterclockwise (ccw) about the C17−C20 Bond As Viewed from C17[a]

38a	20R	Euphol
38b	20S	Tirucallol
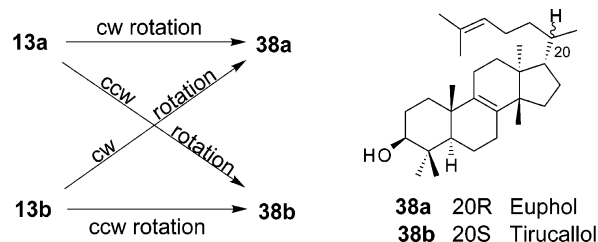
[a] These rotations are nominally 90° or −90°.

C19 (lupeol numbering). CDE-syn isomers (modeled in Table 4) and the isolupyl cation (leading to nepehinol, **37b**) are substantially higher in enthalpy than the lupyl
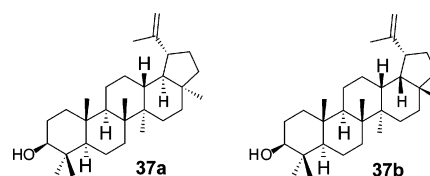
cation. Interestingly, nature does not appear to make **37a**, whose low energy derives from its anti CDE ring fusion with a cis DE ring junction.[37] The rule of configuration transmission predicts that enzymatic formation of **37a** would proceed via the 17α-dammarenyl cation, which in nature produces hopene rather than lupeol skeletons.

The C17 stereochemistry of the dammarenyl cation cannot be deduced by the rule of configurational transmission for lanosterol, euphol (**38a**), tirucallol (**38b**), and other rearranged tetracycles. Rearrangement requires a nominal[38] +90° or −90° rotation of ω_1 to give a C20 cation oriented[29a] to permit a 1,2-hydride shift of H17 to C20. In the absence of enzymatic or solvation effects, the +90° and −90° rotations differ little in energy (Table 3). Rearranged triterpenes of 20R or 20S configuration could be formed from either **13a** or **13b**, depending on whether the 17,20-hydride shift occurred after a clockwise or counterclockwise rotation about the C17−C20 bond (Scheme 5).[1a,3] These cryptic mechanisms can be elucidated by using 20-oxa substrate analogues to determine the C17 configuration of (deoxy)dammarenyl cation intermediates. We argue below on phylogenetic grounds that these cyclases in higher plants access **13b**. Eventually, the mechanistic pathways will be revealed through correlations between sequence and active site geometries from crystal structures.

Unlike euphol and tirucallol formation, side chain deprotonation of dammarenyl cations should leave the C17 stereochemistry intact. However, careful workers have pointed out that the stereochemistry of a (deoxy)-

(37) Lupeol was about 1 kcal/mol higher in energy than **37a** (see the Supporting Information).

(38) The values of +90° or −90° for rotation are relative to the ω_1 value of 180° near the end of D-ring formation. As hyperconjugation with C13−C17 or C16−C17 is established, ω_1 changes by about 20−30° (see trends in Table S4). From this reference point, the rotation needed for optimal H17 migration is either 60−70° or 110−120°, similar to the values used by others.[11d,e]

dammarenyl cation might be reversed by an intervening mechanistic pathway entailing a C17−C20 hydride shift, 180° side chain rotation, and a C20−C17 hydride shift.[12c] Our arguments against 180° rotation about C17−C20 now exclude this possibility. Thus, the detection of 17β-dammara-20(21),24-diene as a trace (0.1%) byproduct of hopene synthesis[12c] demonstrates formation of some 17β-deoxydammarenyl intermediate and shows that stereocontrol by I261 in SHC is not quite absolute. Other hopene byproducts involve 17−20 hydride shifts;[12c] as in lanosterol synthesis,[11d] the side chain rotation is clockwise as viewed from C17.

The similarity in predicted activation energies for 17α- and 17β-dammarenyl cation formation (Table 1) indicates that, in the absence of enzymatic or solvation effects, the C17 stereochemistry is not dictated by the carbon configurations in rings B and C. In enzymatic cyclizations, the C17 configuration is evidently controlled through substrate folding involving dihedrals ω_A, ω_B, ω_C, and ω_D. Dihedral ω_A (C8−C13−C14−C15) must be ca. −90° for C-ring expansion to give a 14α-methyl substituent, and dihedrals ω_B and ω_C must be roughly −90° and +70°, respectively, to situate C13 and C17 close enough for D-ring formation. The C17 stereochemistry is determined by ω_D, values of ca. −120° leading to the 17α product and +90° leading to the 17β isomer. Substrates held in one of these 2 folded conformations can react by an energetically favorable pathway to give a single C20 cationic intermediate.
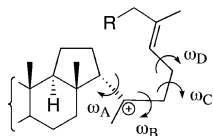

Other conformers might be thought to produce tricycles such as thalianol[17k] rather than a tetracycle. However, ω_D in SHC[8] is initially 155° because C19 of squalene is constricted between F601 below and P263 above, thus forcing the C19 methyl into the plane of the sinusoidally folded substrate. During formation of ring C, contraction of the substrate toward the fixed[8] C10 methyl draws the C19 methyl past F601. As the C18−C19 double bond is positioned for cation-π addition, the C19 methyl must point down because an upward orientation is blocked by the bulky I261 above C19. This stereocontrol is disrupted in I261A and I261G mutants, which produce mixtures of 17α and 17β tetracycles.[14f]

Some relay of asymmetry does occur during tetracyclization of folded (oxido)squalene.[4,5] Because of substrate folding and the absence of intermediates in ABC ring formation,[29a] the trans geometry of the substrate double bonds leads to opposite substituent orientations at C5 and C10 and at C8 and C9, and this stereochemical control can be achieved under suitable biomimetic conditions.[10] The C13 and C14 substituent orientations are likewise opposite in tetracyclic intermediates (**3**, **9**, **13a**, **13b**) because ω_A (like ω_1) does not undergo 180° rotation in the active site cavity before ring expansion. All known enzymatic annulations create the C8 and C14 methyls anti, whereas biomimetic reactions may produce some syn isomer.[10h] By contrast, enzymes have more capacity to generate diversity by folding the substrate into ener-

getically unfavorable conformers. The thermodynamic preference for opposite orientations of C9 and C10 (anti ABC ring fusion) is more readily overcome enzymatically than biomimetically.[10i,j]

Chemotaxonomic and Phylogenetic Insights. Most 3-hydroxypentacyclic triterpenes in angiosperms have a β-methyl group at the DE ring fusion or are derived from such a cation.[3] Because of configurational transmission, these compounds must arise from the 17β-dammarenyl cation. Similar reasoning indicates that this mechanistic route is also followed in the biosynthesis of some pentacyclic triterpenes in ferns and other lower plants, whether originating from squalene (e.g. **39**) or oxidosqualene. In addition, some 6−6−6−6 tetracycles (e.g. **40**) and 3-deoxy triterpenes (e.g. **41**) in angiosperms access the 17β-dammarenyl cation.[3] In contrast, pentacyclic triterpenes in prokaryotes arise from the 17α-(deoxy)dammarenyl cation,[2,12c,16a,b] and a similar pathway provides a hexacyclic C_{35} homologue of hopene.[16d,m,39] Some pentacyclic triterpenes in ferns (e.g. **42**), mosses, gymnosperms (e.g. **43**), and angiosperms (e.g. **44**) likewise have an α-methyl at the DE-ring fusion[3] and therefore follow the same mechanistic route. Product structure also reveals other mechanistic features, e.g. whether intermediates have boat or chair configurations in ring B and whether squalene or oxidosqualene was the substrate.
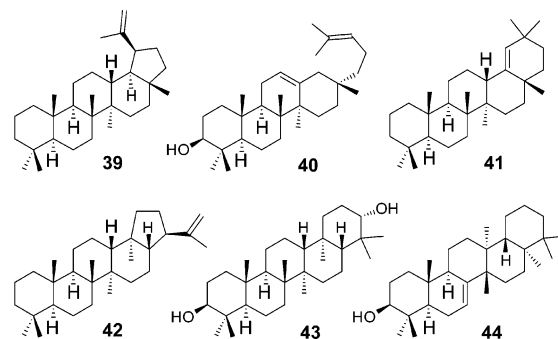
These mechanistic features and simple taxonomic considerations illuminate the relative ease with which specific catalytic capabilities can evolve. The ring B chair/boat, the 17α/17β, and squalene/oxidosqualene transitions occurred rarely in evolutionary history. These 3 binary divisions were used to organize cyclases into 8 possible categories, 5 of which make triterpene structures known in nature (Figure 5). Within a category, evolutionary changes involving E-ring expansion, the site of deprotonation, and rearrangements by 1,2-hydride and methyl shifts are relatively common. Somewhat less common are transitions between pentacyclization and tetracyclization and between five-membered and six-membered D-ring formation.

Transitions between mechanistic pathways can be interpreted as barriers to evolutionary change.[40] The hierarchy of mechanistic barriers in Figure 5 derives from both mutation probabilities and reproductive fitness. The importance of mutation probability in the evolution of

(39) See Supporting Information.

(40) Crossing evolutionary barriers is easier when cyclization produces few rings; e.g. 17α/17β transitions are more facile for enzymes making tetracycles rather than pentacycles (see the Supporting Information).
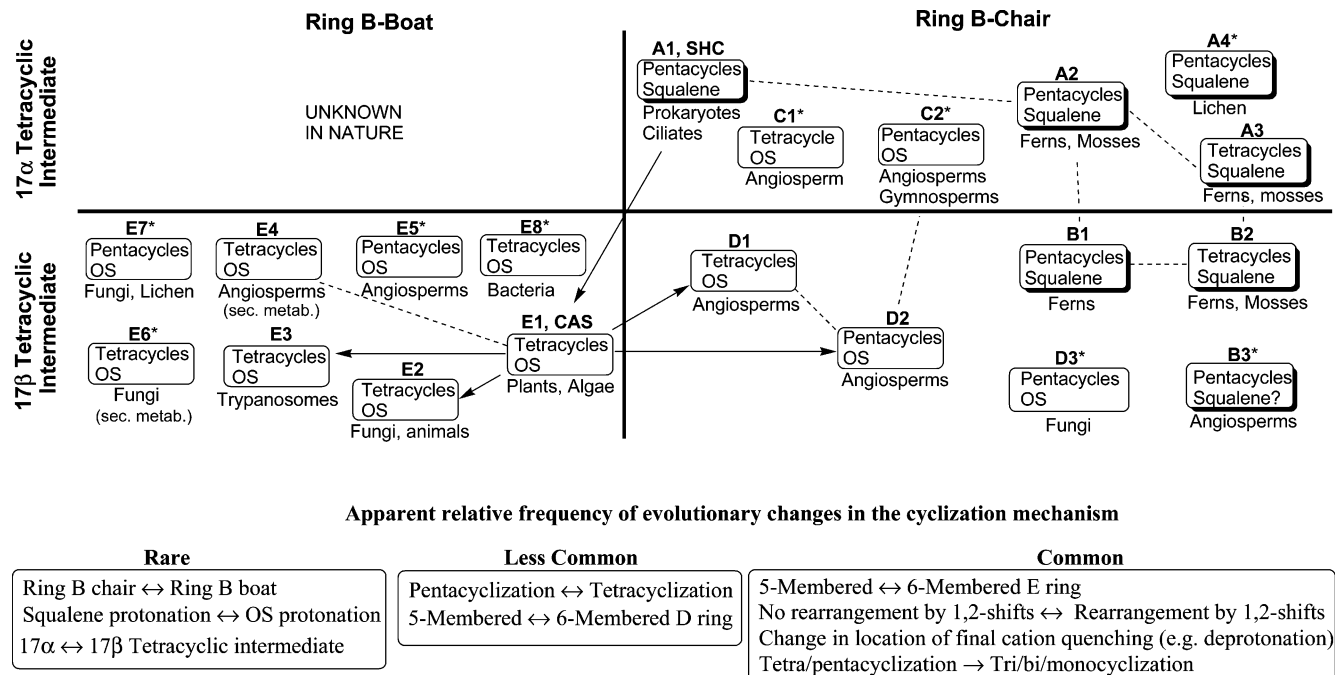
Ring B-Boat **Ring B-Chair**

Apparent relative frequency of evolutionary changes in the cyclization mechanism

Rare	Less Common	Common
Ring B chair ↔ Ring B boat Squalene protonation ↔ OS protonation 17α ↔ 17β Tetracyclic intermediate	Pentacyclization ↔ Tetracyclization 5-Membered ↔ 6-Membered D ring	5-Membered ↔ 6-Membered E ring No rearrangement by 1,2-shifts ↔ Rearrangement by 1,2-shifts Change in location of final cation quenching (e.g. deprotonation) Tetra/pentacyclization → Tri/bi/monocyclization

FIGURE 5. Evolutionary relationships of triterpene synthases. Cyclases are classified into eight possible categories based on mechanistic criteria corresponding to the three rare evolutionary changes. Five categories (A–E) are known in nature. Categories of cyclases with squalene as substrate (A, B) are shown as shadowed boxes; cyclases in other categories (C–E) use oxidosqualene (OS). Cyclases are further divided into groups according to taxonomic and structural features (i.e. whether the triterpene product is tetracyclic or pentacyclic). Minor groups (not widely distributed in the specified taxa) are denoted by an asterisk. Cycloartenol is regarded as tetracyclic. Solid arrows represent evolutionary pathways suggested by phylogenetic reconstructions from sequence data, and dashed lines represent tentative pathways. The stereochemistry of C17 and ring B refers to that of the unrearranged tetracyclic intermediate (**3**, **9**, **13a**, or **13b**). Cyclase groups producing mono-, bi-, and tricyclic products or unusual cyclization patterns are not shown. See also Figure S12.

secondary triterpene metabolites is supported by artificial evolution results. Common mechanistic transitions, e.g. changes in the location of cation deprotonation or extent of rearrangement, can be effected with point mutations that alter hydrogen-bonding networks,[6c,14a,e] and good product accuracy can be obtained with double mutants.[14a] In contrast, the 3 principal mechanistic barriers have not been traversed accurately and efficiently by mutagenesis, chimeric methods, or directed evolution, although limited transitions for 17α/17β[14f] and squalene/oxidosqualene[6c,14g,i] have been achieved with simple mutations.

In the evolutionary transition from a squalene cyclase to an oxidosqualene cyclase,[41] the ability to cyclize squalene must be lost to prevent the cyclization of this oxidosqualene precursor. This transition can be achieved by simple disruptive mutations,[6c,14g,i] which are inefficient or inaccurate. In the multistep evolution of lanosterol synthase from SHC, this transition was accomplished by decreasing enzymatic stabilization of the aspartate anion and relying on the 3-hydroxyl of oxidosqualene intermediates to help stabilize the anion.[9a,29b] However, this hydrogen bond keeps the substrate bound to aspartate during cyclization, whereas the substrate contracts from both ends in SHC.[8] Because folding for D-ring formation occurs *during* cyclization, this change in contraction mandates a redesigned folding mechanism. These modi-

fications presumably required an improbable combination of numerous mutations, consistent with the rarity of this evolutionary transition.

Cyclases of primary metabolism seldom cross mechanistic barriers because high accuracy is required to prevent the accumulation of byproducts.[42] Early in eukaryotic evolution, CAS arose from SHC. This remarkable transformation (presumably comprising intermediate steps) incorporated the 3 rare mechanistic transitions: 17α/17β, squalene/oxidosqualene, and ring-B boat/chair. Since this ancient[13f] feat, evolutionary changes for eukaryotic cyclases that make primary metabolites have been mechanistically minor, the modifications being limited to the site at which the C8/C9 lanosteryl cation is deprotonated. Whereas CAS removes a C19 proton from the C9 cation, parkeol synthases deprotonate the C9 cation at C11, and lanosterol synthases remove the C9 proton from the C8 cation.[29b] Notably, all eukaryotic cyclases of primary metabolism employ similar mechanisms, i.e., tetracyclization of oxidosqualene to the 17β-protosteryl cation, followed by rearrangement to the C8/C9 lanosteryl cation. These enzymes produce only 3 of

(41) SHC and other prokaryotic squalene cyclases can accept oxidosqualene as substrate, but oxidosqualene is usually absent in prokaryotes. In eukaryotes, the abundance of squalene relative to oxidosqualene may explain how some cyclases in ferns and mosses can make 3-deoxyhopene skeletons.

(42) Tolerance to byproducts of primary metabolism, which can disrupt membrane function, signaling, and steroid hormone activity, varies among organisms. For example, yeast are more tolerant than humans to aberrant metabolites and intermediates of sterol synthesis: Ruan, B.; Lai, P. S.; Yeh, C. W.; Wilson, W. K.; Pang, J.; Xu, R.; Matsuda, S. P. T.; Schroepfer, G. J., Jr. *Steroids* **2002**, *67*, 1109–1119. Control over sterol homeostasis may also be facilitated by the prevailing segregation of primary and secondary triterpene metabolism in separate mechanistic categories.

100 known triterpene skeletons[3] that can be formed directly from oxidosqualene.

Enzyme accuracy is less important for triterpene secondary metabolites, which are found mainly in plants. Relaxing the requirement for accuracy effectively lowers the activation energy[43] for evolutionary change because inaccurate mutants may also survive to reproduce. Selective pressure can then lead to rapid evolution of cyclases, whose course is channeled in part by the different heights of mechanistic barriers in triterpene synthesis. The numerous cyclases of secondary metabolism in angiosperms (groups D1 and D2) originated from CAS via an early ring-B boat/chair transition. These enzymes evolved to produce the majority of known triterpene skeletal types with only rare crossing of the principal mechanistic barriers of Figure 5.[44] Instead, many alternatives were explored among the easier evolutionary steps, resulting in the overwhelming predominance of 17β/ring-B-chair/oxidosqualene cyclases in angiosperms. The near absence of 17α-dammarenyl cyclases in higher plants reflects the rarity of the 17β to 17α evolutionary step.[45] We thus propose that most angiosperm cyclases making 20S tetracyclic products such as **38b** access the 17β dammarenyl cation but reverse the direction of ω_1 rotation from the clockwise motion that gives 20R products (Scheme 5).

A few angiosperm triterpenes are derived from the protosteryl cation (groups E4 and E5). Cucurbitadienol synthase (group E4), the only characterized cyclase in these groups, is closely related to the CAS family[17d] and differs mechanistically only in the extent of rearrangement and site of deprotonation. Arborinols from group E5 may represent a link between SHC and CAS on the basis of geochemical evidence showing that their degradation products occur in Permian and Triassic sediments (probably predating angiosperms).[46] The 3-deoxy pentacycles from squalene cyclization in angiosperms (group B3) could have arisen from group D2 by an oxidosqualene/squalene transition or have been conserved from ancestors in lower vascular plants. Geochemical evidence indicates that pteridosperms (seed ferns) contained these triterpene skeletons.[47]

Unlike higher plants, ferns and mosses make numerous secondary metabolites that lack the 3β-hydroxyl group and appear to be products of squalene cyclization (groups A2, A3, B1, B2).[48] These metabolites show less diversity than those of angiosperms but nevertheless achieved nearly all the easier mechanistic transitions of Figure 5 and apparently also crossed the 17α/17β barrier. The corresponding cyclases have not been characterized, and their ancestry is not obvious. Evolution from CAS would entail a reversal of at least two rare evolutionary steps. We suggest that squalene cyclases in lower vascular plants may have originated from an early lateral gene transfer from prokaryotes. Consistent with this position, a fragmentary sequence (GenBank AF159795) from the fern *Pteridium aquilinum* encodes a protein ca. 50% identical to squalene-hopene cyclases from blue-green algae, but only ca. 30% identical to eukaryotic oxidosqualene cyclases. Lateral gene transfer probably also accounts for many anomalous cyclases in fungi, protists, and bacteria. This is a likely origin of several minor groups (A4, D3, E6, E7, part of E8), although a different evolutionary pathway was proposed for lichen triterpenes.[49]

Conclusions. The dammarenyl cation represents the first branch point from which numerous mechanistic pathways of triterpene synthesis diverge. Molecular modeling and conformational analysis indicate that the formation and reactivity of this tetracyclic intermediate is not governed by any relay of asymmetry from rings ABC. Instead, cyclases control the regio- and stereoselectivity of reaction through substrate folding, some of which occurs initially and some during cyclization.

We established from experimental and theoretical evidence that this folding gives a 17β-dammarenyl cation in LUP1-catalyzed biosynthesis. We then generalized this result by showing that, although the C17 configuration may be lost during biomimetic D-ring expansion, the configurational information is relayed to C20 in enzymatic reactions as a consequence of limited rotational freedom for cationic intermediates in the active site cavity. D-ring expansion of 17β-dammarenyl and protosteryl cations places the ring-DE angular methyl on the β face, and the opposite configuration is produced by the 17α isomers. This rule of configurational transmission applies to pentacyclic triterpenes and 6−6−6−6 (but not 6−6−6−5) tetracycles. The angular methyl configuration is taken from the mechanistic intermediate preceding any 1,2 shifts or other rearrangements. Thus, the structure of most triterpenes reveals which of the three fundamental cyclization pathways of Scheme 1 was accessed.

This knowledge provided the foundation for classifying (oxido)squalene cyclases according to the mechanism

(43) Qualitatively, the evolutionary barriers between mechanism types are analogous to energy barriers between chemical reactants and products. Energy minima correspond to mutants that optimize reproductive fitness, adaptability, and survival. This metaphor can be extended to other topographical features of potential energy surfaces, e.g., lateral gene transfer is like a tunnel through the barrier.

(44) (a) Evolution from simpler to more complex cyclization mechanisms has been proposed,[17e] but angiosperm cyclase mechanisms probably evolved non-progressively as a maze. (b) For a given clade, the mechanistic category for secondary metabolism may have been set by accident of ancestry, but this initial choice was largely immutable, at least under the evolutionary conditions of recent eras.

(45) The rare example a tetracyclic 17α-dammarane derivative isolated in small amounts from an angiosperm[21] (group C1) may represent a minor product of an inaccurate enzyme. Alternatively, an accurate cyclase might have originated by a rare event of lateral gene transfer or evolution from group D1 or C2.

(46) (a) Hauke, V.; Adam, P.; Trendel, J.-M.; Albrecht, P.; Schwark, L.; Vliex, M.; Hagemann, H.; Püttmann, W. *Org. Geochem.* **1995**, *23*, 91−93. (b) Borrego, A. G.; Blanco, C. G.; Püttmann, W. *Org. Geochem.* **1997**, *26*, 219−228. (c) Ourisson, G.; Albrecht, P.; Rohmer, M. *Trends Biochem. Sci.* **1982**, *7*, 236−239. Arborinols of apparent microbial origin also occur in recent sediments: (d) Jaffé, R.; Hausmann, K. B. *Org. Geochem.* **1995**, *22*, 231−235. (e) Hanisch, S.; Ariztegui, D.; Püttmann, W. *Org. Geochem.* **2003**, *34*, 1223−1235.

(47) Paull, R.; Michaelsen, B. H.; McKirdy, D. M. *Org. Geochem.* **1998**, *29*, 1331−1343.

(48) The alternative of oxidosqualene cyclization followed by removal of the C3 oxygen has been disproven for a representative fern: Barton, D. H. R.; Mellows, G.; Widdowson, D. A. *J. Chem. Soc. D* **1971**, 110−116.

(49) Stictane triterpene synthases in lichen (group E7) were suggested to originate from squalene cyclases (group A4) by a squalene/oxidosqualene transition, followed by a ring B chair/boat transition: Wilkins, A. L. *Bibl. Lichenol.* **1993**, *53*, 277−288. A 17α/17β transition would also be needed. The intermediary 3-oxygenated triterpenes are known in lichen, although a separate group is not given for these compounds. The possibility that pentacyclic triterpene synthases in vascular plants originated via mosses from cyclases in lichen has been considered: Kondratyuk, S. Y.; Galloway, D. J. *Bibl. Lichenol.* **1995**, *57*, 327−345.

used for cyclization. During evolution, some mechanistic patterns (e.g. cation rearrangements) evolved rapidly to create the vast diversity of triterpene skeletons. Other mechanistic changes occurred rarely, notably transitions between 17α and 17β stereochemistry of tetracyclic intermediates, between ring-B chair and ring-B boat intermediates, and between squalene and oxidosqualene protonation. These three formidable barriers divide cyclases into 8 possible mechanistic categories, of which 5 are known in nature. This perspective furnishes clues to evolutionary relationships from product structure alone. Considering the difficulty of expressing cyclases heterologously and the relative ease of chromatographic/spectroscopic analysis of biological material, these structural insights will likely remain valuable in the foreseeable future. The mechanistic paradigm also provides a framework for understanding the large amount of sequence information that is becoming available for triterpene synthases. This information, combined with insights from geochemical analyses, cyclase crystal structures, and evolutionary studies of related enzymes[13e,50] should illuminate the complex patterns of phylogenetic relationships among triterpene synthases.

Experimental Section

Components of synthetic complete medium were obtained from Fisher Scientific (Pittsburgh, PA). Heme (in the form of hemin), ergosterol, and epicoprostanol (internal GC standard) were from Sigma/Aldrich (St. Louis, MO). Chemical synthesis of (18*E*)-22,23-dihydro-20-oxaoxidosqualene (**21**) has been described.[22] The 3β-(*tert*-butyldimethylsilyloxy)-22,23,24,25,26,27-hexanor-17α-dammaran-20-one was obtained from Novartis Pharma AG and was evidently prepared analogously to the known[21a] 3β-*tert*-butyldiphenylsilyl ether. ¹H NMR and ¹³C NMR spectra were acquired from dilute CDCl₃ solutions at 25 °C. GC/MS and GC conditions are given in the Supporting Information.

Yeast Culture. SMY8 yeast (*MAT***a** *erg7::HIS3 hem1::TRP1 ura3-52 trp1-Δ63 leu2-3,112 his3-Δ200 ade2* Gal⁺) were transformed with *Arabidopsis thaliana* lupeol synthase (*LUP1*) as reported previously.[17a,b] The resulting recombinant yeast strain was cultured in synthetic complete medium lacking uracil and containing 2% glucose, ergosterol (20 μg/mL, supplied from a stock of 2 mg/mL in 1:1 ethanol/Tween 80), and hemin (13 μg/mL, supplied from a stock of 1.3 mg/mL hemin chloride in 1:1 ethanol/10 mM NaOH). The cells were grown in 0.25 L of this medium and then induced in a 2-L Erlenmeyer flask containing 1 L of an identical medium but with 2% galactose instead of glucose. This culture was shaken at 250 rpm for 24 h at 30 °C, followed by centrifugation at 800 × g for 10 min and collection of the cell pellet (14 g).

In Vitro Incubation of LUP1 with 21. The yeast cell pellet (described above) was suspended in sodium phosphate buffer (20 mL, pH 6.2) and lysed in a French Press at 1900 psi. A portion (10 mL) of the resulting 40% slurry was incubated with 1 mL of racemic **21** solution (1 mg/mL in 2.5% Tween-80) for 12 h at 23 °C. The reaction was terminated by addition of 20 mL of ethanol. After removal of denatured protein and cell debris by centrifugation at 15 000 × g for 10 min, the supernatant was concentrated by rotary evaporation in vacuo. The residue was partitioned with 3 × 20 mL of ethyl ether, and the combined ether extract was used for chromatographic isolation of the enzymatic product (98%) and quantitative GC analysis (2%).

TLC (silica gel; ethyl ether/hexane 2:3) showed, in addition to ergosterol (R_f 0.24), lupeol, and other mobile triterpenes, a

spot at R_f 0.21 that was absent in the control incubation lacking substrate **21**. The combined ether extract was dissolved in 0.5 mL of hexane and loaded onto a 0.63-g silica gel plug in a Pasteur pipet. After elution with ethyl ether/hexane (1:4), a total of 10 fractions (3 mL/fraction) were collected. Fractions 5−7 were combined and loaded on a 20 × 20 cm² analytical silica gel TLC plate and developed with ethyl ether/hexane 1:2. The UV-inactive band below ergosterol afforded **24b** (ca. 200 μg), which was identified by NMR, GC, and GC/MS (see the Supporting Information).

3β-Hydroxy-22,23,24,25,26,27-hexanor-17α-dammaran-20-one (24a). A solution of the TBDMS ether of **24a** (1 mg) in 20% aqueous HF in CH₃CN (2 mL) was stirred at room temperature for 1 h. The reaction was quenched with 5% NaHCO₃ and washed with 0.1 M phosphate buffer (pH 7). The organic phase was passed through a silica gel plug and eluted with ether. Evaporation of the eluate gave **24a**, which was characterized by ¹H and ¹³C NMR, GC, GC/MS, and TLC (see the Supporting Information). The TLC mobilities (silica gel, ethyl ether/hexane 2:3) of **24a** and **24b** were identical. When the TBDMS ether of **24a** (1 mg) was stirred in 20 mM *p*-toluenesulfonic acid in acetone (0.5 mL) at 55 °C for 30 h, complete conversion to **24b** was observed.

In Vitro Incubation with 24a. Compound **24a** (150 μg) was incubated with LUP1 yeast homogenate (10 mL) under the same conditions used in the enzymatic conversion of **21** to **24b**, except that no substrate was added. NMR analysis of the crude sterol fraction (Figure 1) showed **24a** and no **24b** (detection limit 0.1%).

Molecular Modeling. Quantum mechanical calculations were done with Gaussian 98 or Gaussian 03,[51] as described in the Supporting Information. Geometry optimizations were done with B3LYP/6-31G*. Because B3LYP appeared to underestimate the exothermicity of annulations, we used mPW1PW91 for single-point energies.

(51) (a) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.9 and A.11.2; Gaussian, Inc.: Pittsburgh, PA, 1998. (b) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03*, Revision B.02 and C.01; Gaussian, Inc.: Wallingford CT, 2004.

(50) Rezen, T.; Debeljak, N.; Kordis, D.; Rozman, D. *J. Mol. Evol.* **2004**, *59*, 51−58.

Supporting Information Available: NMR spectra of **24a** and **24b**, GC of the crude enzymatic reaction mixture, molecular modeling results, atomic coordinates, and other supplementary data. This material is available free of charge via the Internet at http://pubs.acs.org.

JO050147E