Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

An *Arabidopsis* oxidosqualene cyclase catalyzes iridal skeleton formation via Grob fragmentation

Quanbo Xiong, William K. Wilson, and Seiichi P. T. Matsuda*

Department of Chemistry and Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77005, USA Email: matsuda@rice.edu

Table of Contents

Materials and Methods

Components of synthetic complete medium and TLC plates (E. Merck brand) were obtained from Fisher Scientific (Pittsburgh, PA, USA). Heme (in the form of hemin) and ergosterol were from Sigma (St. Louis, MO, USA). Bis(trimethylsilyl)trifluoroacetamide (BSTFA) was purchased from Aldrich Chemical Co. (Milwaukee, WI, USA).

GC-MS was performed on an Agilent 6890 GC equipped with an Agilent 5973 MSD containing an electron-impact ion source. Samples (5 µL) were introduced by split injection (40:1) at 280 °C on a Restek 35% phenyl/65% methylpolysiloxane column (30 m × 0.25 mm × 0.1 µm), with the oven temperature held isothermally at 270 °C. The

helium carrier gas flow rate was 20 cm/s, and mass spectra (*m/z* 50 to 550) were obtained at 70 eV after a 3-min solvent delay.

For GC-MS derivatization to trimethylsilyl (TMS) ethers, a 10% aliquot of the NSL was dissolved in 100 μL of BSTFA-pyridine (1:1) and heated at 40 °C for 2 h. The reaction mixtures were used directly for GC-MS analysis without evaporating the reagents.

NMR spectra were measured in dilute solution at 25 °C on a Bruker Avance 500 MHz spectrometer and referenced to tetramethylsilane at 0 ppm (^{1}H) or $CDCl_3$ at 77.0 ppm (^{13}C).

Saponification was done as follows. To the yeast cell pellet (1 g) was added 2 mg of butylated hydroxytoluene (BHT) in 400 μL ethanol and 10% KOH in 80% ethanol (5 mL). The mixture was purged under a nitrogen stream for 2 min and then tightly capped and heated at 70 °C for 2 h. After cooling to room temperature, the reaction mixture was diluted with 5 mL of water and extracted with 3 × 15 mL hexanes. The combined hexane extracts were washed with 3 × 5 mL of water and 5 mL of brine and evaporated in vacuo to a residue comprising the non-saponifiable lipids (NSL).

Cloning of At5g42600 and plasmid construction

The first strand of cDNA of the open reading frame *PEN5/At5g42600*, 2286bp was transcribed from mRNA obtained from 2-day old *Arabidopsis thaliana* seedlings using a RETROscript RT-PCR kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. The following primers were designed (enzyme restriction sites are capitalized).

SalIF (5'- atatGTCGACatgtggagactgcgaattggagct -3')
XhoIR (5'- atcagggaCTCGAGcaagatgtt -3')
XhoIF (5'- aacatcttgCTCGAGtccctgat-3')
NotIR (5'- gGCGGCCGCtcaagaaacaagcagacgcaga-3')

SalIF contains the start ATG and 5' sequence with a Sal I site for subcloning, and XhoIR spans a Xho I site in the coding sequence. XhoIF is the reverse complement of

XhoIR, and NotIR is the 3' oligo that contains the stop site and a Not I site for subcloning. PCR was performed with a TripleMaster PCR system (Eppendorf, Westbury, NY) with a program for 40 cycles: 94 °C, 15 s for denaturation, 55 °C, 15 s for annealing and 72 °C, 30 s for extension. After treating with Ex-Taq polymerase (Panvera, Madison, WI, USA) at 72 °C for 10 min, the front region amplicons were purified with a Qiaquick Kit (Qiagen, Chatsworth, CA, USA) and ligated into the pGEM-T (Promega, Madison, WI, USA) vector with a Quick Ligation Kit (NEB, Beverly, MA, USA), and *E. coli* strain DH5α was transformed with the ligation mixtures. The rear region was combined to the front region on T-vector after digestion with Not I/Xho I. The whole length insert was sequenced using T7/SP6 and internal primers to ensure that no splicing or PCR errors had been introduced. It was excised with Sal I/Not I and subcloned into the yeast expression vector pRS426GAL.[1] The resulting plasmid was named pXQ11.2 and was transformed into the yeast lanosterol synthase mutant SMY8[2] and the squalene epoxidase/lanosterol synthase double mutant RXY6[3] using the lithium acetate method,[4] to generate RXY6[pXQ11.2] and SMY8[pXQ11.2], respectively.

In vivo production and characterization of the $C_{30}H_{52}O$ alcohol 2 and byproducts

Small-scale cultures: A 100-mL culture of SMY8[pXQ11.2] was grown in synthetic complete medium lacking uracil with 2% galactose as a carbon source and supplemented with ergosterol (20 μg/mL), hemin (3 μg/mL), and Tween 80 (5 mg/mL). A control culture was grown under the same conditions except that glucose was used instead of galactose. The cultures were incubated to saturation with shaking at 250 rpm for 24 h at 30 °C. Centrifugation at 500 × g for 10 min yielded ca. 1 g of yeast cells, which were saponified to give the NSL.

The ^{1}H NMR spectrum of the NSL from the galactose-induced sample showed, in comparison with the control sample, a characteristic methyl doublet at 0.764 ppm (J = 6.8 Hz), a methyl singlet at 0.891 ppm, and the absence of a doublet-doublet at 3.2 ppm. The peak height of the methyl doublet was about 3% of the H-18 signal of ergosterol at 0.632 ppm.

Analytical TLC on silica gel of the galactose-induced NSL sample eluted with CH_2Cl_2/hexane (3:1) showed, above ergosterol (Rf = 0.27), an extra spot at Rf = 0.33, which was not observed in the control sample.

The total ion chromatogram (TIC) of the control sample gave the usual peaks of oxidosqualene, dioxidosqualene and ergosterol at 4.9, 6.5 and 8.3 min, respectively. In contrast, the galactose-induced sample showed an extra peak at 5.2 min and significantly decreased intensity for the oxidosqualene peak. Unlike typical triterpene alcohol TMS derivatives with M^+ at *m/z* = 498, the peak at 5.2 min had an unusual M^+ of *m/z* = 500, corresponding to a likely molecular formula $C_{33}H_{60}OSi$ (the formula weight of a TMS-triterpene alcohol + 2H); the MS pattern (Figure S1) was expected for the TMS ether of a partially cyclized triterpene. This mass spectrum of **2** is entirely different from the mass spectrum of **1** (Figure S2); the isolation of **1** is described in a later section (see below). The reconstructed ion chromatogram (RIC) at *m/z* = 500 showed no other peaks (Figure S3), and the total abundance of several tiny peaks with ion *m/z* = 498 (Figure S4) amounted to no more than 1% of the above M^+ = 500 product peak, suggesting it as a dominant enzyme product.

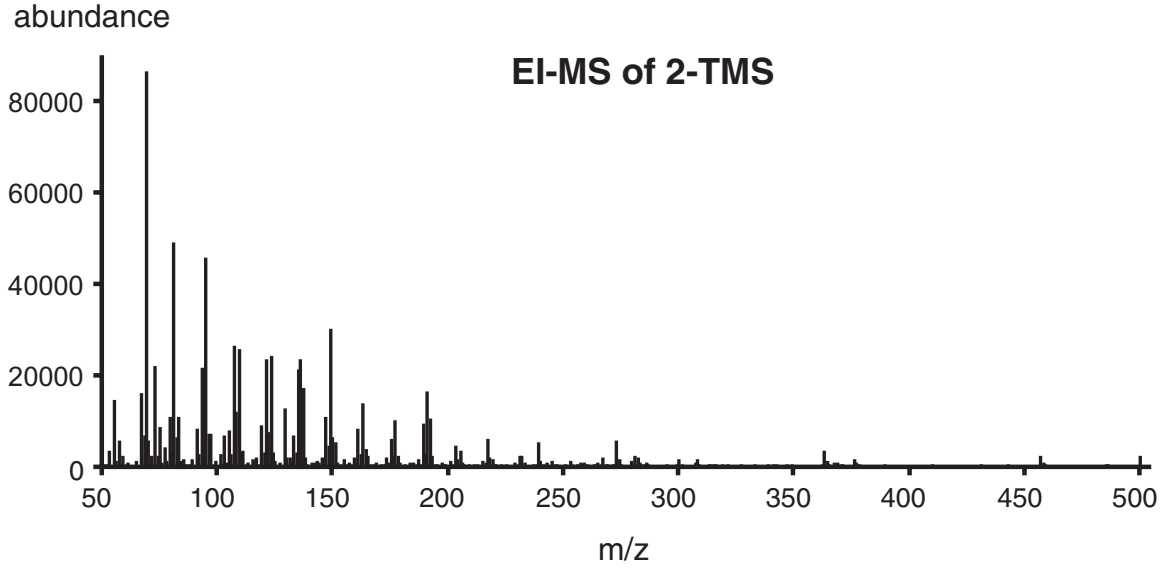


Figure S1. Electron-impact mass spectrum of the TMS ether of compound **2**.

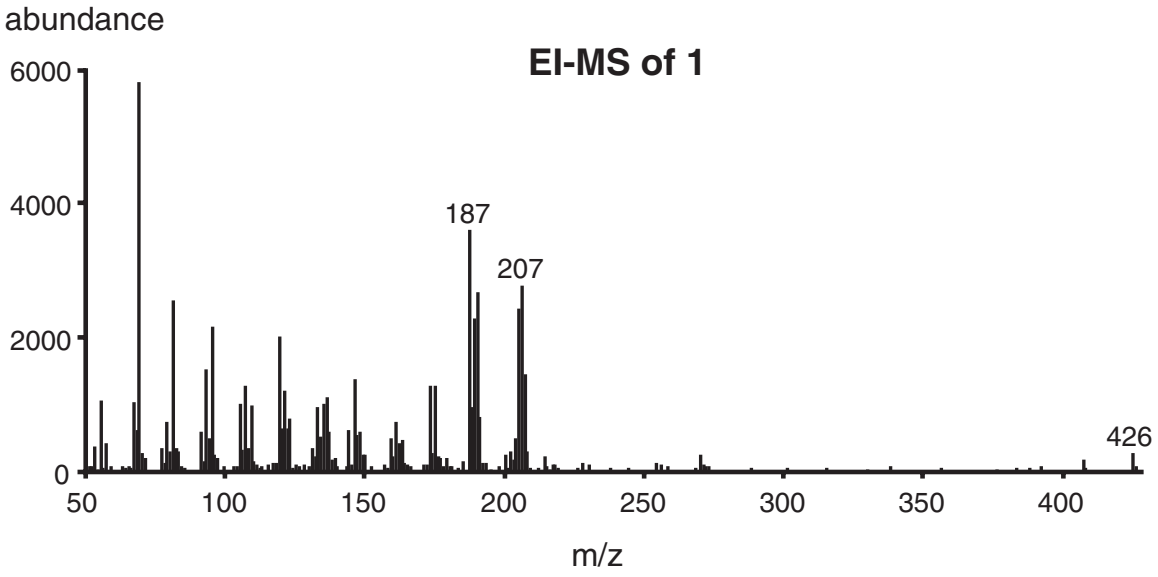


Figure S2. Electron impact mass spectrum of compound **1**.

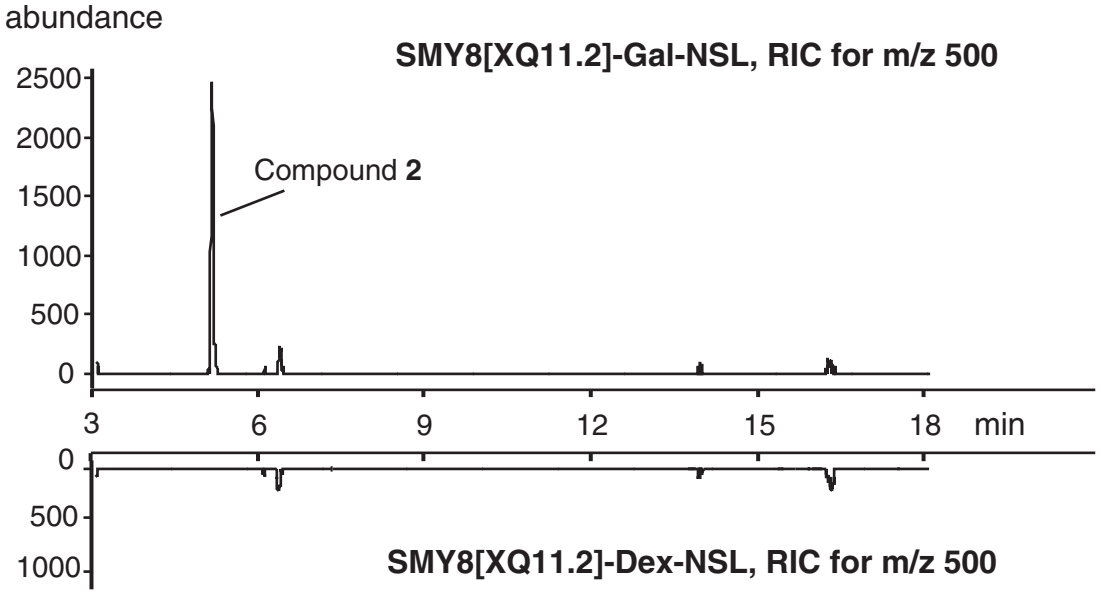


Figure S3. Reconstructed ion chromatogram (*m/z* = 500) for the GC-MS of crude NSL of SMY8[XQ11.2]. Upper panel: galactose-induced cells. Lower panel: cells cultured in dextrose-containing medium without galactose induction (control). The results showed an alcohol as major product peak at 5.2 min with M^+ at *m/z* = 428.

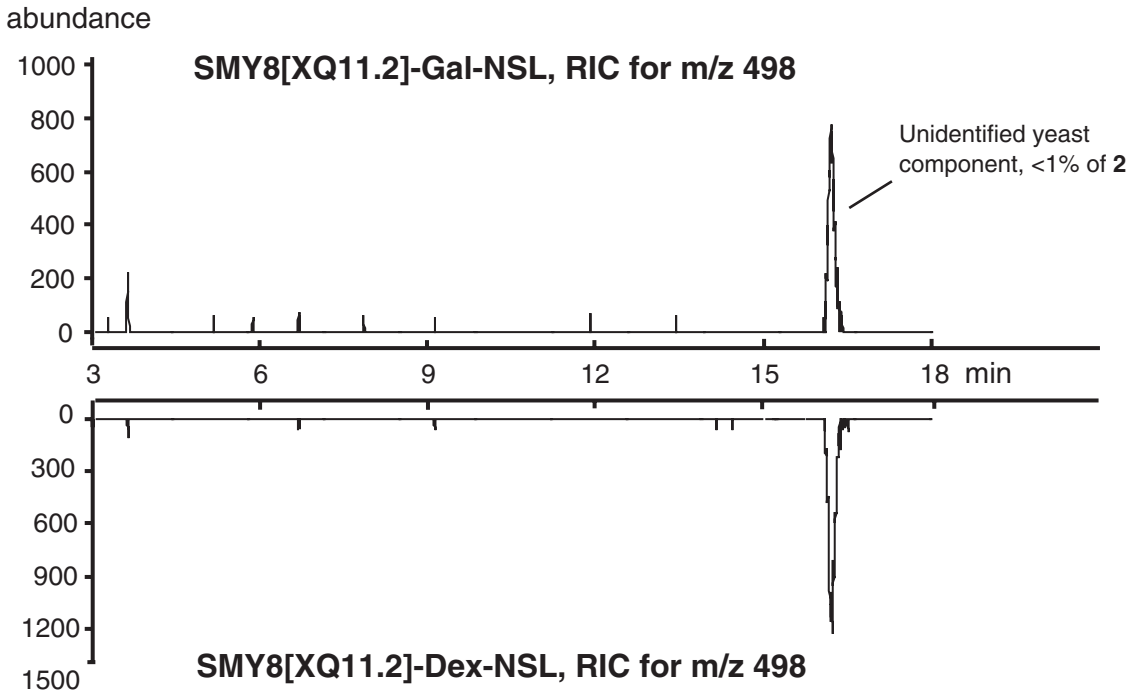
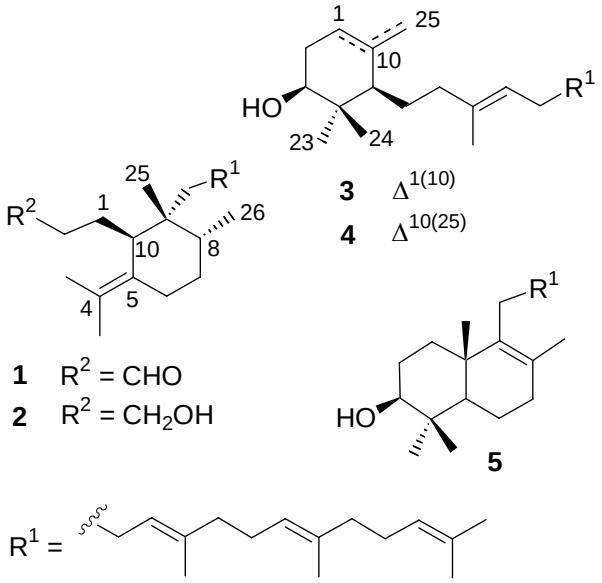


Figure S4. Reconstructed ion chromatogram (*m/z* = 498) for the GC-MS of crude NSL of SMY8[XQ11.2]. Upper trace: galactose-induced cells; Lower trace: cells cultured in dextrose-containing medium without galactose induction (control). These data indicate only trace amounts of oxidosqualene cyclized products (< 1%) with *m/z* = 498 (M^+ for TMS derivatives of $C_{30}H_{50}O$ triterpene alcohols).

Large-scale culture; isolation of **2**: Cell pellets (21 g) from a 3-L scale galactose-induced SMY8[pXQ11.2] culture were saponified to give 48 mg of NSL. This material was loaded on a chromatography column containing 7 g silica gel (230-400 mesh) and eluted with CH_2Cl_2/hexane (5:1). A colorless oil (ca 1 mg) was isolated; this corresponded to the TLC spot at Rf = 0.33 described above. This material was used for NMR structure determination described below.

Other fractions from the column chromatography were thoroughly analyzed by ^{1}H NMR and GC-MS. Three very minor products were observed in fractions preceding those containing **2**. Two of these minor products, obtained as mixtures, were identified as camelliol C (**3**)[5] and achilleol A (**4**)[6] by comparison of GC-MS and ^{1}H NMR data with those of standard samples and with literature values. The distinctive ^{1}H NMR signals

were δ = 0.831 (s, H-24), 0.970 (s, H-23) and 5.237 (m, H-1) for **3**; and δ = 0.715 (s, H-24), 1.031 (s, H-23), 4.607 (br s, H-25*Z*), 4.874 (br s, H-25*E*) for **4**. Structures are given in Scheme S1.



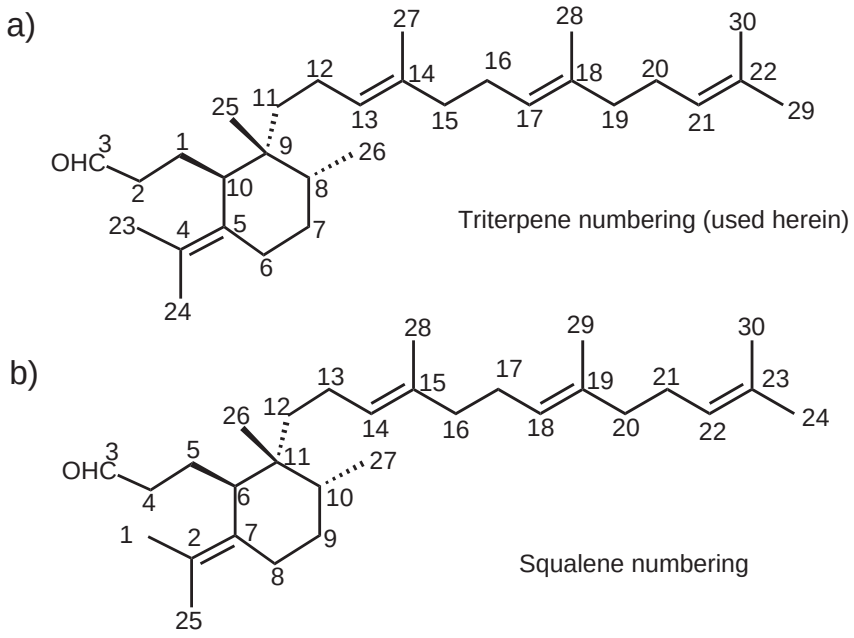
Scheme S1. Product structures from cyclization of oxidosqualene by AthMRN1.

A third minor product (**5**), obtained in a mixture with **3** and **4**, showed methyl singlets at 0.804, 0.946, and 1.007 ppm and a doublet-doublet at ca. 3.2 ppm. The three upfield methyl singlets suggested a bicyclic ring system with a double bond at Δ7, Δ8, or Δ8(26). The latter possibility was excluded by the lack of broad olefinic singlets in the δ 4.5-5.0 region. Like **3** and **4**, compound **5** showed a molecular ion at *m/z* = 498 as the TMS ether by GC-MS.

Strong Gaussian apodization provided coupling constants for the "singlets" of **5** to give the following assignments: δ = 1.007 (d, 0.3 Hz, H23), 0.804 (d, 0.3 Hz, H24), 0.946 (d, 0.7 Hz, H25). Model compounds for Δ7 and Δ8 structures showed the following chemical shifts for H23, H24, and H-25: 4,4-dimethyl lathosterol δ = 0.996, 0.899, 0.861 ppm; 9β-lanosta-7,24-dien-3β-ol, 1.017, 0.87, 0.984 ppm; lanosterol δ = 1.001, 0.811, 0.983 ppm. Thus, a Δ7 structure with either 9α or 9β stereochemistry was excluded by the poor

agreement for the H24 signal. NMR comparisons for a Δ8 structure (lanosterol) were favorable, and **5** was tentatively assigned as the Δ8 bicyclic alcohol.

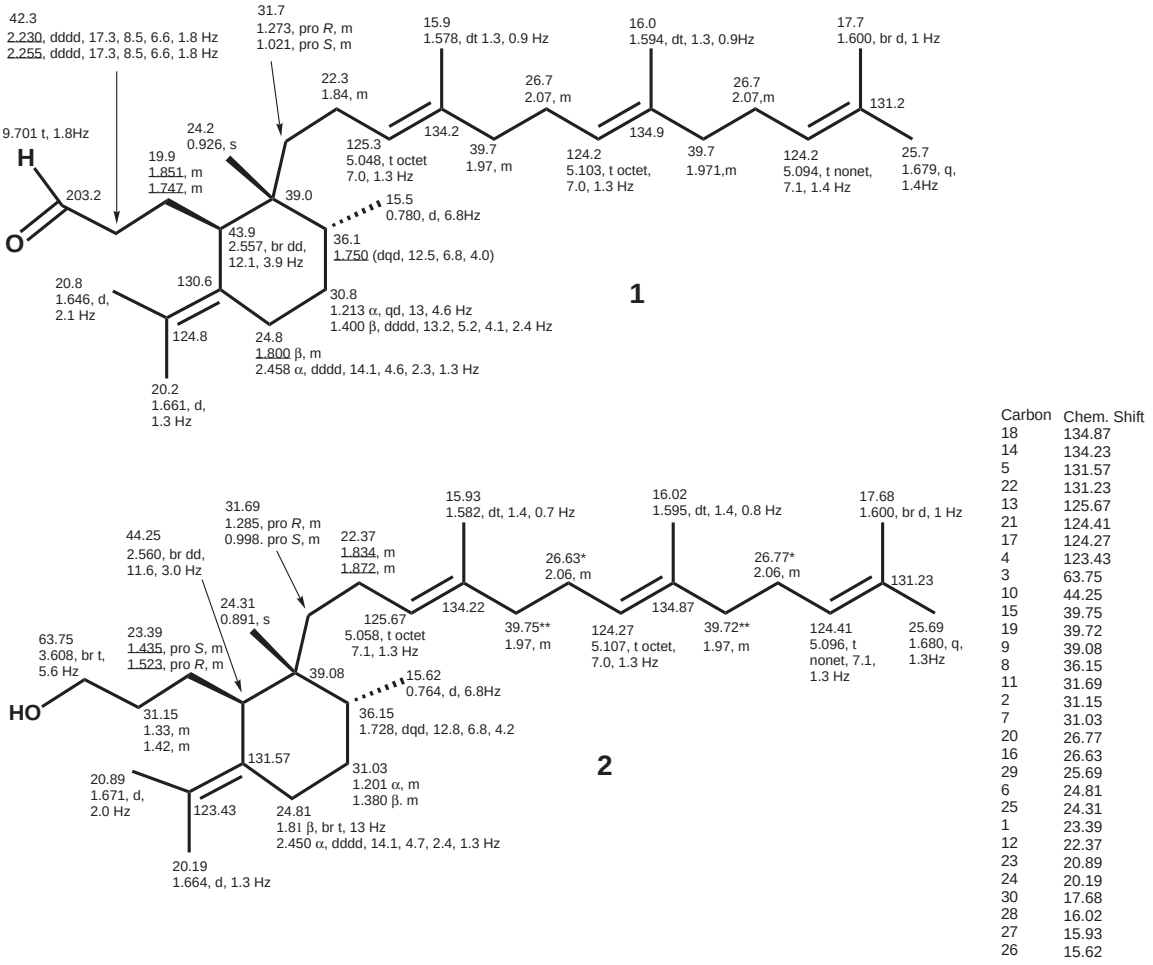
We have used triterpene numbering, which differs from the squalene numbering used by F. Marner and coworkers. These numbering systems are compared in Scheme S2.



Scheme S2. Numbering systems for marneral: a) triterpene numbering used herein; b) squalene numbering used by F. Marner and coworkers.

Structure determination of 2

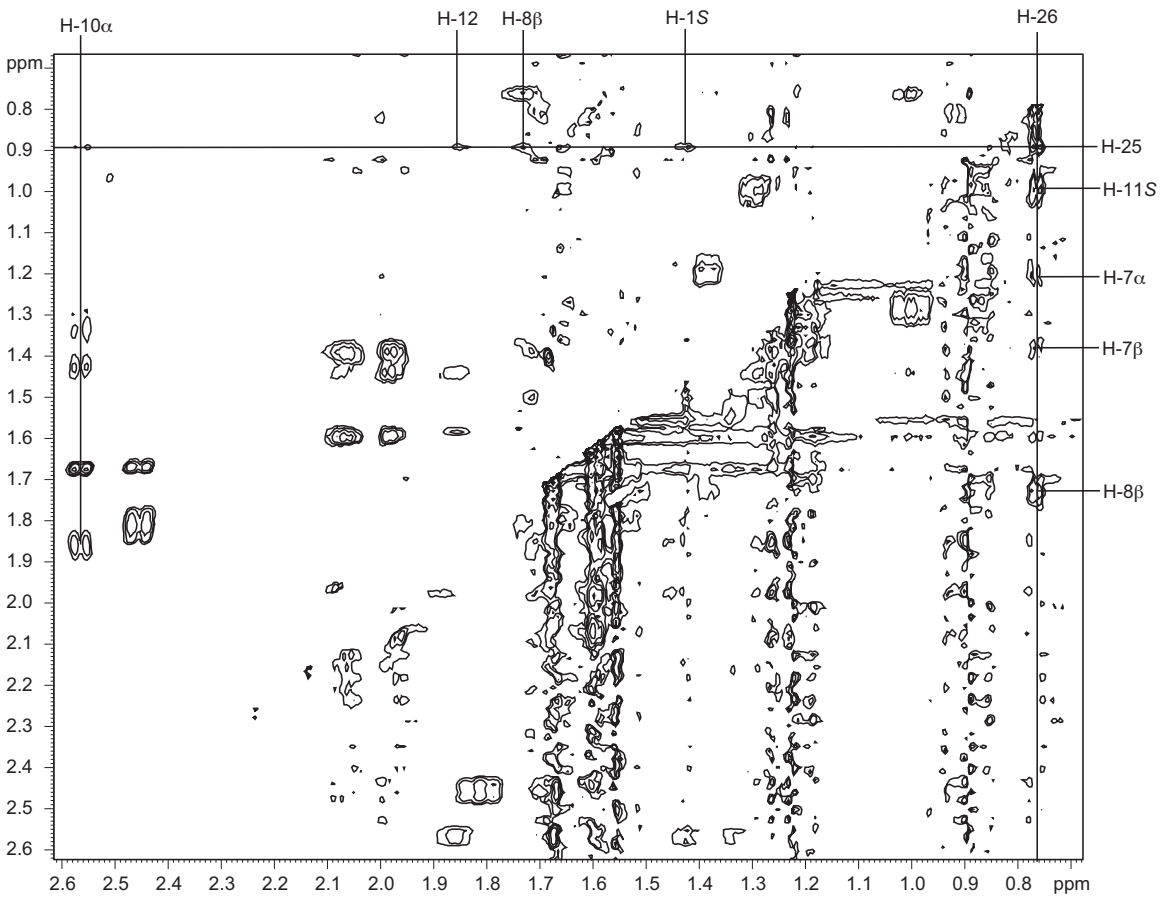
The ^{1}H, ^{13}C and DEPT NMR spectra of **2** showed two aliphatic and six allylic methyls, suggesting a monocyclic system from oxidosqualene. However, **2** did not show the usual 3α proton of triterpenes at ca. 3.2 ppm (dd, J = 11, 4 Hz, 1 H) but instead had a signal at 3.608 ppm (t, J = 5.6 Hz, 2 H) corresponding to CH_2OH; a similar resonance is seen in several reported *seco*-A ring triterpenes.[7] The carbon skeleton of **2** (Figure S5) was established from COSYDEC, HSQC, and HMBC spectra (revealing atom connectivities for carbon atoms 1-11) and from chemical shift comparisons with data for thalianol[3] and camelliol C[5] (confirming the side chain structure). These results provided most of the NMR signal assignments for the iridal skeletons of **1** and **2** shown in Figure S5.

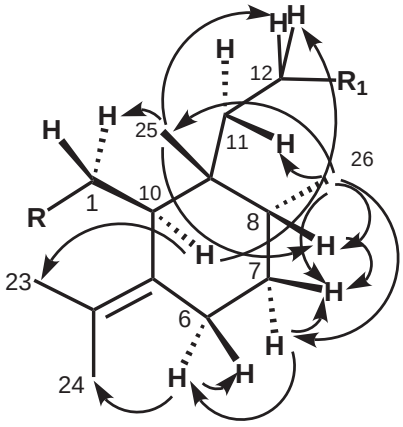


Carbon	Chem. Shift
18	134.87
14	134.23
5	131.57
22	131.23
13	125.67
21	124.41
17	124.27
4	123.43
3	63.75
10	44.25
15	39.75
19	39.72
9	39.08
8	36.15
11	31.69
2	31.15
7	31.03
20	26.77
16	26.63
29	25.69
6	24.81
25	24.31
1	23.39
12	22.37
23	20.89
24	20.19
30	17.68
28	16.02
27	15.93
26	15.62

Figure S5. ^{1}H and ^{13}C NMR assignments of **1** and **2**. The ^{13}C NMR chemical shifts of **1** were deduced from HSQC and HMBC spectra and have an accuracy of about ±0.1 ppm. ^{1}H chemical shifts of **1** and **2** have an accuracy of about ±0.001, 0.003, or 0.01 ppm (values given to 3 decimal places, underlined values, or values to 2 decimal places). ^{13}C shifts marked by asterisks may be interchanged.

The relative 8α-methyl configuration of **2** was determined from the NOESY spectrum (Figure S6) as described in the main text. The absolute stereochemistry at C9 and C10 is derived from the C5 and C10 stereochemistry of triterpene skeletons formed by enzymatic oxidosqualene cyclization. Thus, the absolute configuration of **2** at C-8 is *R*. This is the same configuration as in 8-deoxy-16-hydroxyiridal (10-deoxy-17-hydroxyiridal by Marner's numbering) found in *Iris sibirica*;[8] this stereochemistry is in accordance with the origin of all the naturally occurring iridals.





2

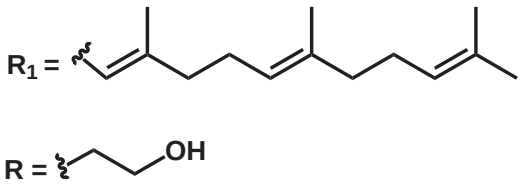


Figure S6. NOESY spectrum of **2** (above) and summary of NOESY correlations (below). Only the positive contours of the phase-sensitive NOESY spectrum are shown.

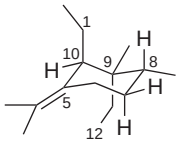
Assignment of the C-8 configuration of **2**, which was based primarily on NMR results, was confirmed by quantum mechanical modeling. This confirmation was done because the ^{1}H and ^{13}C NMR data of compound **2** were virtually identical with those reported for its C-8 epimer (triterpene numbering) from a biomimetic synthesis.[9] This high similarity was surprising since γ effects[10] of the 8-methyl on chemical shifts of C-6 and C-10 should be markedly different for the two C-8 epimers, unless by accident the epimeric methyls are both equatorial or both axial because of unusual conformational effects.

The quantum mechanical calculations were performed with Gaussian 03 software (Linux and Windows version B.02)[11] using a $C_{15}H_{28}$ model structure of **2** that included the ring system, the isopropylidene group, and truncated side chains, as shown in Table S1. We optimized model structures with B3LYP/6-31G*, determined the Boltzmann distribution of conformers from B3PW91/6-311G(2d,p) energies, and calculated NMR shieldings using the GIAO method.

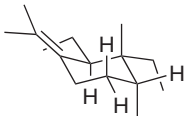
Table S1 indicates that the chair-1 conformer (A) for the 8α-methyl epimer and the chair-1 conformer (D/E) for 8β-methyl epimer have substantially lower energy than other conformers. The Boltzmann distribution shows the 8α-methyl epimer to exist as almost 100% conformer A and the 8β-methyl epimer to be 88% conformer D/E (rotamers about the C8-C9-C11-C12 bond) and 12% G. Conformers B, C, and F had negligible population. The predominantly chair conformations support the arguments in the main text for assignment of the 8α-methyl configuration to **2**.

^{13}C and ^{1}H NMR shieldings for the $C_{15}H_{28}$ model were predicted at the B3PW91/6-311G(2d,p) level using Gaussian 03. (Shieldings for the two ethyl side chains were ignored.) A single conformer (A) was used for the 8α-epimer. For the 8β-epimer, shieldings were averaged according to the Boltzmann distribution of conformers D, E, and G. The shieldings, which were converted to chemical shifts using empirical adjustments,[12] showed rms deviations of 0.07 and 0.78 ppm (^{1}H and ^{13}C data for the 8α-epimer) versus 0.24 and 2.42 ppm (^{1}H and ^{13}C data for the 8β-epimer). The much lower deviations between predicted and observed values for the 8α-methyl isomer were in the expected range for such calculations[12] and validate our configurational assignment from 2D NMR spectra.

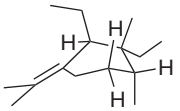
Table S1. Relative energies (kcal mol^{-1}) and Boltzmann distribution of conformers for $C_{15}H_{28}$ models corresponding to 8α- and 8β-methyl epimers of compound **2**.[a]



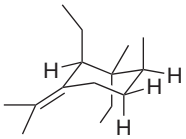
A



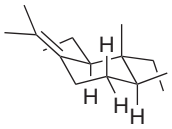
B



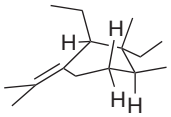
C



D/E



F



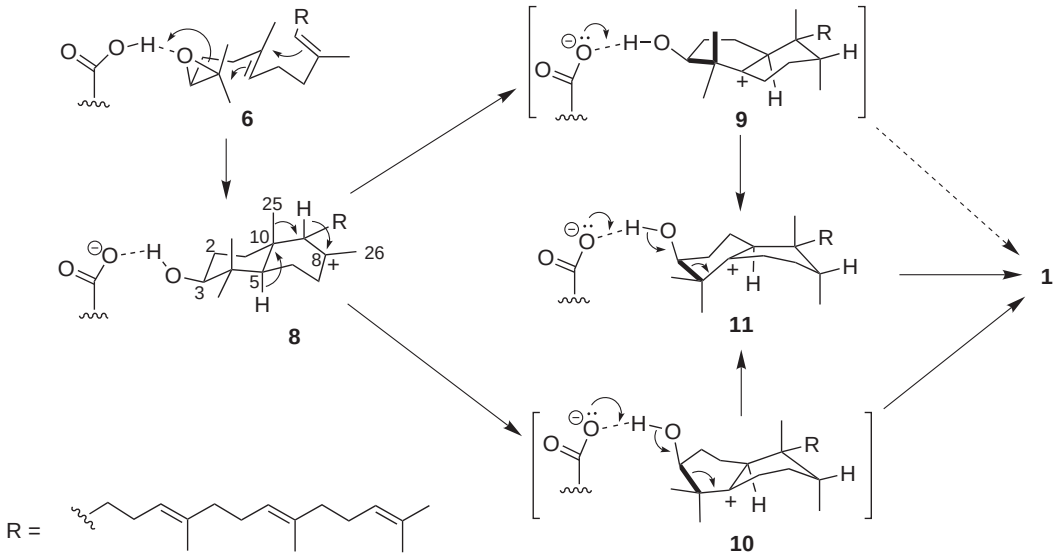
G

	8α-methyl epimer			8β-methyl epimer			
	Chair-1 *anti* (A)	Chair-2 (B)	Boat (C)	Chair-1, *anti* (D)	Chair-1, *gauche* (E)	Chair-2 (F)	Boat (G)
B3LYP/6-31G*	0	8.2	6.8	0	0.3	3.1	1.1
B3PW91/6-311G**	0	7.8	6.5	0	0.4	3.0	1.0
B3PW91/6-311G(2d,p)	0	7.9	6.5	0	0.4	3.0	0.9
Boltzmann distrib.[b]	100	0	0	57	31	0	12

[a] Conformers D/E are side-chain rotamers of the same chair conformation, differing according to whether the C8-C9-C11-C12 rotamer is *anti* or *gauche*. Chair-1 (*anti*) of the 8β-methyl isomer was 3 kcal/mol higher than that of the 8α-methyl epimer. [b] Calculated percentage distribution based on B3PW91/6-311G(2d,p) energies.

Conformational analysis of ring-A inversion prior to Grob fragmentation

After the C5 hydrogen in **8** migrates to C10, ring A inverts to the other chair form, which is now energetically more favorable. This involves reversing the axial/equatorial orientation of the substituents in ring A to achieve compatibility with the axial 10α-hydrogen. Chair-chair inversion normally proceeds stepwise, half of the ring a time. For the bicyclic C5 cation, we identified two intermediate conformers en route to the inverted chair conformer (Scheme S3). Inversion of axial/equatorial orientations for the C1 substituent, with partial inversion at C2, leads to the twist conformer **9**. Similar inversion of C3 and C4 substituents gives boat conformer **10** (Table S2).



Scheme S3. Suggested mechanistic pathway from oxidosqualene (**6**) to marneral (**1**). Bonds elongated by hyperconjugation with the C5 cation are shown thickened.

Table S2. Changes in axial and equatorial substituent orientations as ring A inverts from the C1 side (via twist **9**) or the C4 side (via boat **10**).[a]

Substituent	Original Chair	Twist (**9**)	Boat (**10**)	Inverted Chair(**11**)
10α	ax	ax	ax	ax
1α	ax	eq	ax	eq
1β	eq	ax	eq	ax
2α	eq	eq	eq	ax
2β	ax	eq	ax	eq
3α	ax	ax	eq	eq
3β	eq	eq	ax	ax
4α	eq	eq	ax	ax
4β	ax	ax	eq	eq

[a] Axial (ax) and equatorial (eq) orientations are unconventionally defined here: axial bonds are approximately orthogonal to the least squares plane of the ring, and equatorial bonds are tilted slightly or moderately from the plane of the ring. The original chair orientations are in magenta, and the inverted chair orientations are in blue. Owing to the *sp2* center at C5, proper nomenclature is pseudo-chair, pseudo-boat, quasi-axial, etc.

DFT calculations suggest almost barrierless Grob fragmentation

In this section, we show that **9**, **10**, and **11** can undergo essentially barrierless Grob fragmentation to **1**. The gas-phase density functional theory (DFT) calculations are quite appropriate for molecuclar modeling in the unsolvated active site cavity of cyclases. Calculations were done with Gaussian 03 software.[11] Geometry optimizations were performed with B3LYP/6-31G*, and single-point energies were calculated with mPW1PW91/6-311+G(2d,p) and SCF=tight. All models were considered as closed-shell systems (restricted calculations) with the frozen-core approximation.

In order to model the deprotonation of the 3-hydroxyl in the Grob fragmentation, we added to the model the protonating aspartate (as acetate) and two water molecules hydrogen-bonded to acetate. The distance between acetate and ring B was fixed (see below) to prevent possible oxonium ion formation during geometry optimization. For each of **9**, **10**, and **11**, we placed the acetate somewhat above the substrate ring system, as is found in a lanosterol synthase crystal structure.[13] Water molecules were added on either side to reduce the gas-phase basicity of the acetate. These waters were positioned away from the 3-hydroxyl to avoid their eventual interaction with the incipient 3-aldehyde. The positioning of the acetate-$2H_2O$ moiety was somewhat arbitrary and differed among the three conformers; this precluded energy comparisons between conformer models with acetate-$2H_2O$. Initial minimizations were done by constraining various dihedral and valence angles in order to achieve natural hydrogen bonding distances for acetate. Then these angular constraints were lifted, and the position of C7, C9, and the acetate methyl were frozen. For a given conformer, these three atoms remained frozen in the same positions for calculations at various C3-C4 bond distances. Figure S7 shows stereoviews of **9**, **10**, and **11** optimized with acetate-$2H_2O$.

In the boat and chair conformers (**10** and **11**), the empty *2p* orbital of the C5 cation is oriented roughly parallel to the C3-C4 bond. This orientation elongates the C3-C4 bond through hyperconjugation.[14] The C3-C4 bond is activated for cleavage by its elongation and by its interaction with the vacant *2p* orbital at C5 (the beginnings of a π bond). In the twist conformer (**9**), the *2p* orbital is partially aligned with both the C3-C4 bond and the C4-C24 bond, as reflected by their modest elongation (1.576 and 1.578 Å). The initial orientation of the *2p* orbital is dependent on the ring-B conformation, which is subject to

steric effects of the active site cavity. As the C3-C4 bond length increases during fragmentation, the alignment of this bond with the *2p* orbital improves.

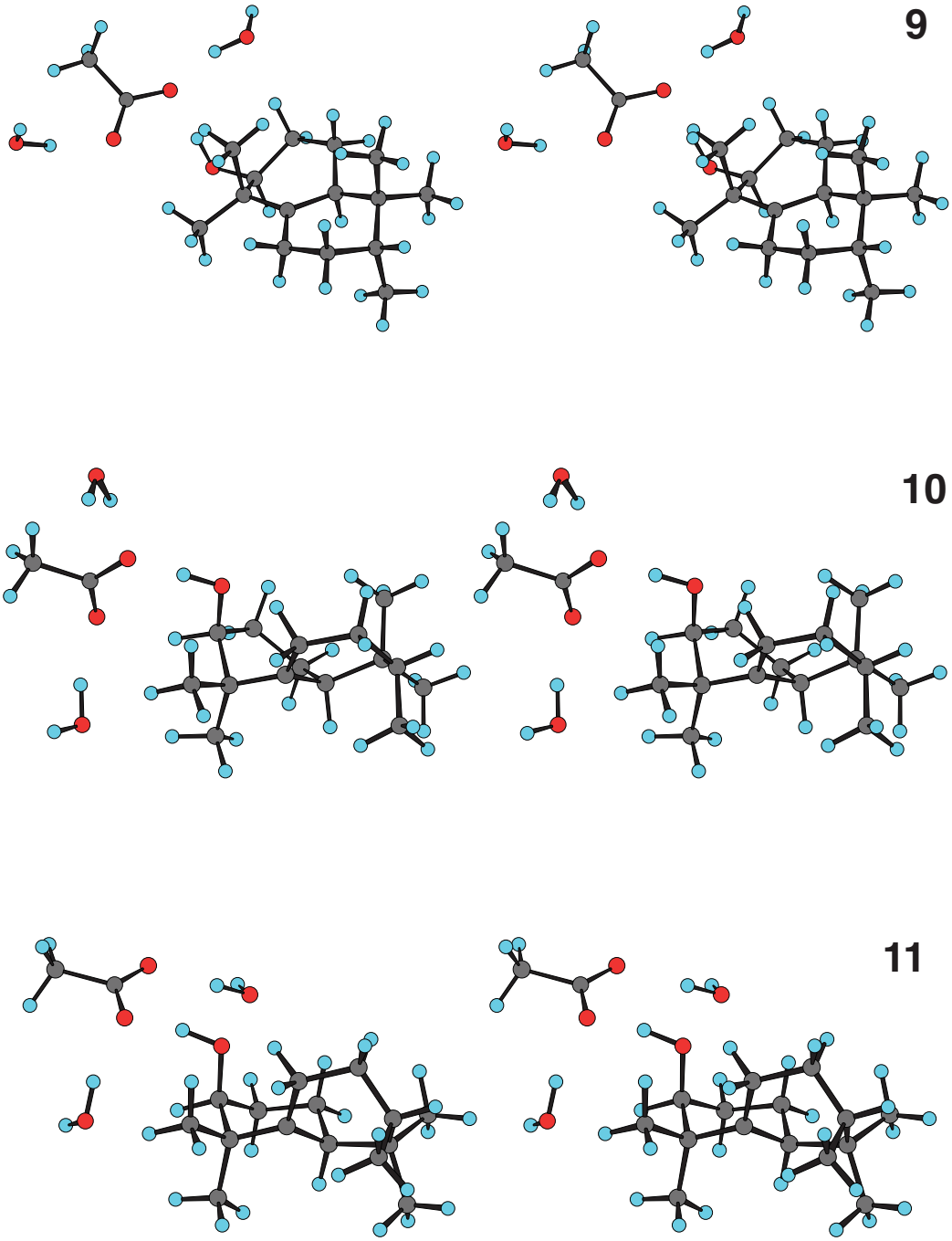


Figure S7. Stereoviews of twist (**9**), boat (**10**), and inverted chair (**11**) conformers of the C5 cation from B3LYP/6-31G* geometries. The models include acetate and two water molecules. The C3-C4 bond length was fixed at 1.58 Å in **9** and 1.60 Å in **10** and **11**.

We studied the Grob fragmentation of conformers **9**, **10**, and **11** by carrying out a relaxed potential energy surface (PES) scan along the C3-C4 bond distance. This was done by fixing the C3-C4 bond at particular values. We obtained B3LYP/6-31G* geometry-optimized structures for a C3-C4 bond length at 0.1 Å intervals from 1.6-2.0 Å for **9**, 1.6-2.5 Å for **10**, and 1.58-3.0 Å for **11**. For each structure, a single-point energy calculation was done with mPW1PW91/6-311+G(2d,p); this method avoids the underestimation of cyclization energies observed with B3LYP/6-31G*. (Fragmentation energies, which correspond to reverse annulation, are overestimated by B3LYP/6-31G*.)

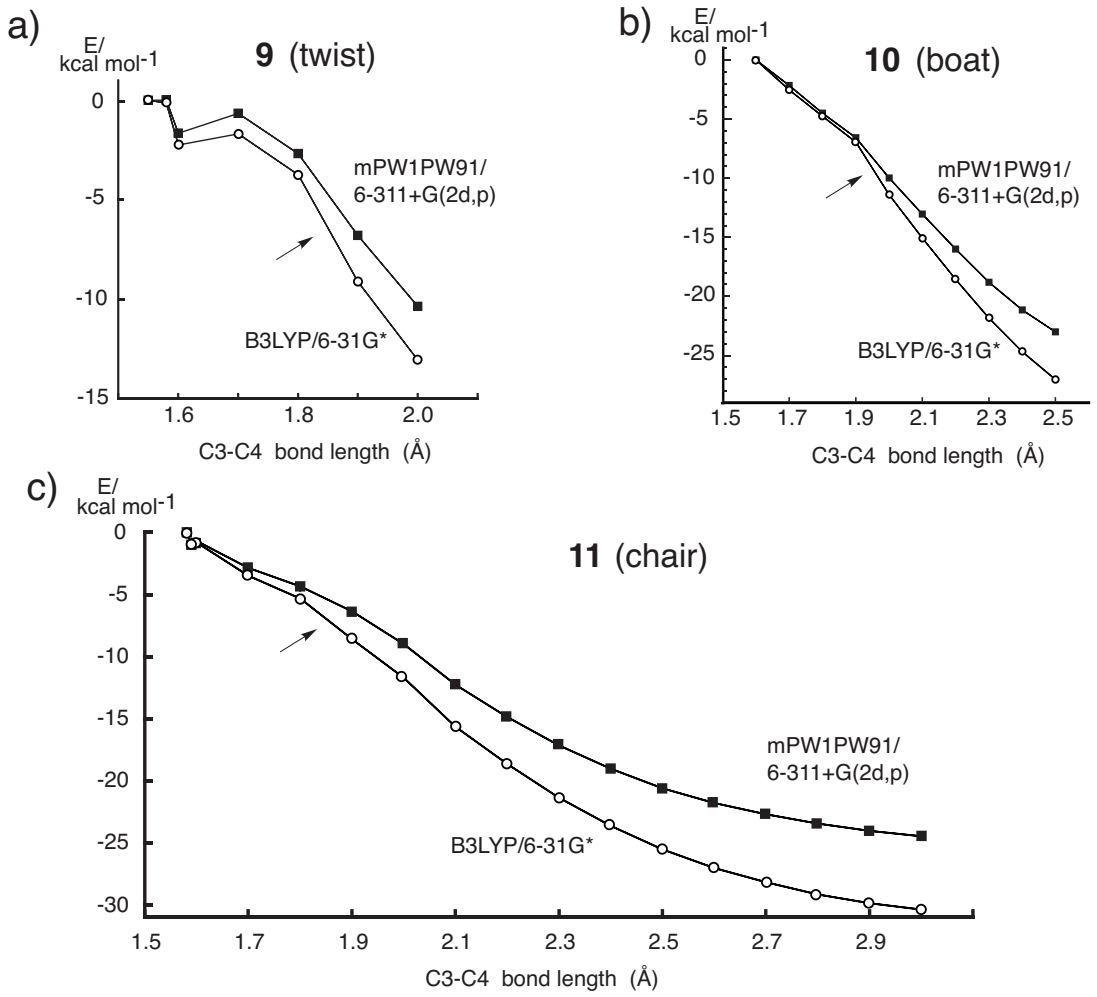


Figure S8. Energetics of Grob fragmentation of **9**, **10**, and **11**, as described by PES scans in which the C3-C4 bond length is varied. Arrows mark the point of proton transfer from the 3-hydroxyl to acetate in these model calculations. Absent enzymatic effects, **9** and **10** are ca. 3 kcal/mol higher in energy than **11**.

The PES scans, shown in Figure S8, suggest the existence of a barrierless reaction path corresponding to Grob fragmentation. The slight barriers at a C3-C4 bond length of about 1.6 Å are insignificant relative to the energy of the cationic intermediate at room temperature. The active-site residues undoubtedly modify the substrate conformation, and these effects will be more important than minor energy barriers found in nonenzymatic models. The apparent reaction energy of 25-30 kcal mol^{-1} is a rough estimate that is markedly influenced by the limited stabilization of acetate in our model.

The use of one-dimensional PES scans to estimate the minimum energy path (MEP) of a reaction has limitations, especially when the reaction is not described well by a single bond length or angle. Cramer et al.[15] have studied a reaction showing large differences between the MEP and a single-parameter cross section of the PES. Such differences can usually be detected by monitoring other geometric parameters for discontinuities. We monitored the geometries of each point on the PES scans for any abrupt changes, such as in conformational changes of ring A or proton transfer from hydroxyl to acetate (data not shown). The PES scan for **9** showed a modest geometry discontinuity in the region of 1.6-1.7 Å, probably associated with bringing the C3-C4 bond into better alignment with the cationic *2p* orbital. At about 1.90 Å, the PES scans for **9**, **10**, and **11** showed a discontinuity in the OH distances for the hydroxyl proton. This discontinuity corresponds to proton transfer from the 3-hydroxyl to acetate, a process that should have a small activation energy, especially if the aspartate has some mobility. Because no other discontinuities were found for **9** and **10**, we inferred that their fragmentation to **1** is essentially barrierless.

The methodology we used to model the Grob fragmentation of **9** and **10** is not ideal. A desirable approach would be to model these conformers enclosed in the enzyme active site; however, no crystal structure exists for MRN1. Nevertheless, the predominant nonsteric enzyme-substrate interaction is proton transfer from the substrate 3-hydroxyl to the aspartate of MRN1, and we have modeled this interaction. We have used PES scans rather than path calculations to study the fragmentation. The PES scans may not correspond to the lowest energy pathway but they do indicate a pathway that is strongly exothermic and shows essentially no barriers. Use of PES scans is most appropriate for reactions that can be described by changes in the length of a single bond or angle, and

fragmentation of **9** and **10** is well described by the cleavage of the C3-C4 bond. We note that IRC and path calculations are also fallible, as these methods do not necessarily find the global MEP. Furthermore, molecules at room temperature are not constrained to the MEP and follow moderately higher energy pathways, which are abundant in plateau regions of the PES. A molecular dynamics study would be informative but is computationally too expensive to justify, especially in the absence of crystallographic data for MRN1.

Evidence of 1 as the direct cyclization product

SMY8[pXQ11.2] and RXY6[pXQ11.2], each in galactose-containing medium (100 mL), were cultured to saturation. Cell pellets were divided into the following groups, each equivalent to 30 ml culture, and processed as indicated in Table S3.

Table S3. Summary of small-scale incubations to detect the direct enzymatic product **1.**[a]

Group	Yeast strain	24 h incubation	Incubated substrate	Processing method[b]
A	SMY8[pXQ11.2]	-	-	Saponification
B	SMY8[pXQ11.2]	-	-	EtOH precipitation
C	SMY8[pXQ11.2]	+	none	EtOH precipitation
D	RXY6[pXQ11.2]	+	none	EtOH precipitation
E	RXY6[pXQ11.2]	+	**6**	EtOH precipitation
F	RXY6[pXQ11.2]	+	**1**	EtOH precipitation

[a] Group D represents a control for *in vitro* reactions E and F. [b] Triterpene products were isolated by saponification or by ethanol precipitation, as described below.

Cell pellets of group A were saponified as described above to get the NSL. The cells of groups B-F were suspended in 0.1 M sodium phosphate buffer (1 mL, pH 6.2) in 1.5-mL Eppendorf tubes. To tube E was added 100 μg racemic oxidosqualene[16] in 4% Tween 80 (50 μL). The cells were then broken by vortexing with glass beads (0.5 mm

diameter, 0.5 mL) for 10 min. Tube F contained **1** (50 μg) in 4% Tween 80 (50 μL). To tubes B-D was also added 4% Tween 80 (50 μL) for consistency. Cell homogenate in tube B was extracted with ethanol/hexane immediately whereas tubes C-F were incubated at ambient temperature overnight before extraction. In each tube, 20 μg of epicoprostanol (Sigma, 99% purity) was added as an internal standard for quantitation.

To the cell homogenate of tube B and reaction mixtures of tubes C-F were added two volumes of ethanol. The precipitated protein and cell debris were removed by centrifugation. The solvent was evaporated under a N_2 stream until there was no smell of ethanol. The remaining aqueous phase was partitioned with hexane (3 × 2 mL). The hexane phase was washed with H_2O (2 × 1 ml), brine (1 ml), and dried under N_2 to give the hexane extract.

Efforts to quantitate **1** by GC-MS on these samples were not successful due to its thermal instability in the ion source. ^{1}H NMR was used to measure the ratio of **1** and **2**, based on the doublet methyl peak intensity of **2** at 0.764 ppm and **1** at 0.780 ppm. The presence of these two compounds was also well supported by other distinguishing signals, such as the C$\underline{H}_2$OH of **2** at 3.608 ppm and the aldehyde proton of **1** at 9.701 ppm.

As shown in Figure S9, the predominant product in the NSL of SMY8[pXQ11.2] (A) was alcohol **2** (12:1 ratio with aldehyde **1**). By contrast, immediate ethanol precipitation/hexane extraction of the SMY8[pXQ11.2] cell homogenate (B) led to predominantly to **1** rather than **2** (10:3 ratio). The high ratio of **2** to **1** in (A) was not due to the decomposition of **1**, but to the conversion of **1** to **2** because the combined amount of **1** and **2** in (A) was more than twice the amount of (B). The overall lower amount of **2** and **1** in (B) was likely due to inefficient extraction. However, additional incubation of the SMY8[pXQ11.2] homogenate for 24 h (C) resulted in the decomposition of **1**, but no apparent increase of **2** was observed. In the case of the RXY6[pXQ11.2] strain, incubation with oxidosqualene led exclusively to production of **1** (E) (73% conversion, assuming only the *(S)*-enantiomer participated in the enzymatic reaction), without detectable **2** (1% detection limit). Unlike results for the cell homogenate of SMY8[pXQ11.2], **1** was almost intact after 24 h incubation in RXY6[pXQ11.2] (F), presumably due to the different cell components of SMY8[pXQ11.2] and RXY6[pXQ11.2].

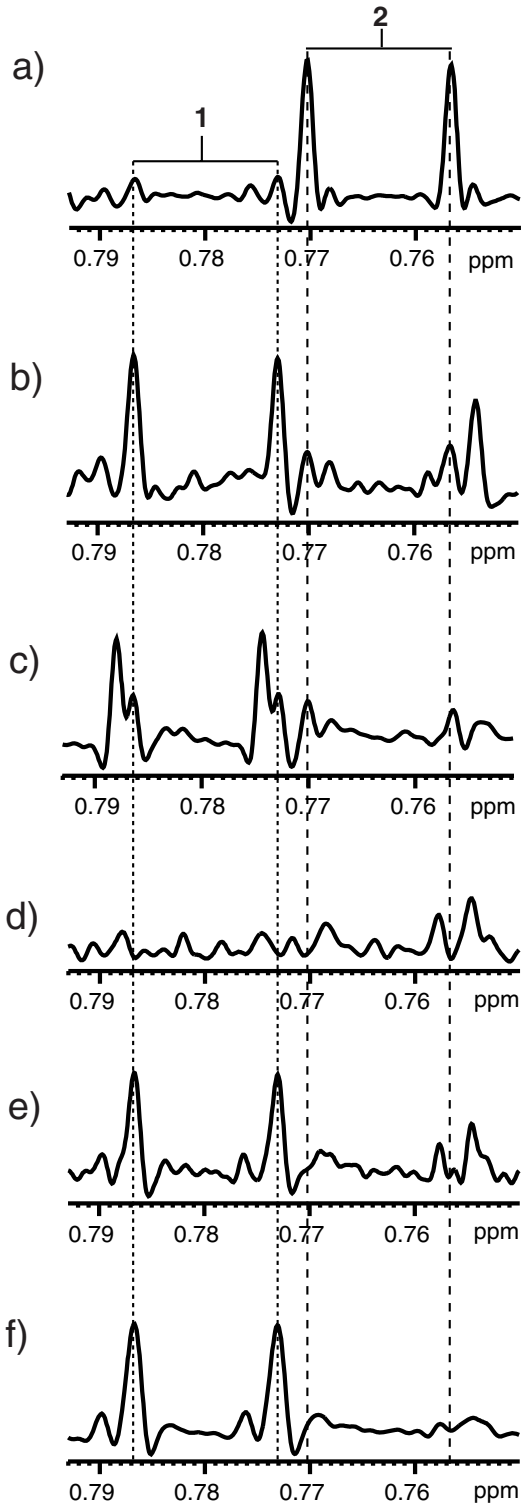


Figure S9. ^{1}H NMR analysis of the enzymatic production of compound **1** versus **2**. Spectra a) through f) correspond to groups A-F of Table S3.

These data demonstrated that the direct enzymatic product was aldehyde **1**, not alcohol **2**. Although reduction of **1** to **2** could have happened partially *in vivo* by a yeast alcohol dehydrogenase or possibly during extraction and work up, most reduction evidently occurred during the saponification process, where the reaction mechanism remains to be studied.

Sequence alignments (Figure S10) and phylogenetic analysis (Figure S11)

The marneral synthase protein sequence was aligned with related protein sequences using ClustalW and the Gonnet protein weight matrix with gap penalty 10.0 and gap length penalty 0.20. PAUP 4.0b5 was used to generate a 50% majority-rule consensus tree. The bootstrap method was performed for 100 replicates using parsimony as the optimality criterion and with full heuristic search. All characters were weighted equally and gaps were treated as "missing". Starting trees were obtained via stepwise addition. The addition sequence was simple, with one tree held at each step during stepwise addition. The branch-swapping algorithm was used with tree-bisection-reconnection, and the steepest descent option was not invoked. The "MulTrees" option was in effect, topological constraints were not enforced, and trees were unrooted. Cyclases include squalene-hopene synthases from *Alicyclobacillus acidocaldarius* (AacSHC),[17] *Methylococcus capsulatus* (McaSHC);[18] lanosterol synthase from *Homo sapiens* (HsaERG7),[19] *Saccharomyces cerevisiae* (SceERG7);[20] cycloartenol synthase from *Dictyostelium discoideum* (DdiCAS1),[21] *Abies magnifica* (AmaCAS1), [22] *Costus speciosus* (CspCAS1),[23] *Oryza sativa* (Osa CAS1),[24] *Avena strigosa* (AstCAS1),[25] *Avena prostrata* (AprCAS1),[25] Avena sativa (AsaCAS1),[26] *Betula platyphylla* (BplCAS1 and BplCAS2),[27] *Pisum sativum* (PsaCAS1),[28] *Glycyrrhiza glabra* (GglCAS1),[29] *Luffa cylindrica* (LcyCAS1),[30] *Cucurbita pepo* (CpeCAS1),[31] *Arabidopsis thaliana* (AthCAS1),[32] *Panax ginseng* (PgiCAS1);[33] lupeol synthases from *A. thaliana* (AthLUP1, AthLUP2),[34,35,36] *Olea europaea* (OeuLUP2),[37] *Taraxacum officinale* (TofLUP2),[37] *Glycyrrhiza glabra* (GglLUP1),[29] *Betula platyphylla* (BplLUP1),[27]; β-amyrin synthases from *Panax ginseng* (PgiBAM1, PgiBAM2),[33,38] *Betula platyphylla* (BplBAM1),[27] *Medicago truncatula* (MtrBAM1),[39] *Pisum sativum* (PsaBAM1),[40] *Glycyrrhiza glabra* (GglBAM1),[41] *Avena prostrata* (AprBAM1),[25]

Avena sativa (AsaBAM1);[26] mixed-amyrin synthases from *Pisum sativum* (PsaAAM1),[40] *Lotus japonicus* (LjaAAM2);[39b] mixed triterpene synthase from *Costus speciosus* (CspMTS1);[23] isomultiflorenol synthase from *Luffa cylindrica* (LcyIMS1);[42] cucurbitadienol synthase from *Cucurbita pepo* (CpeCUB1),[43] thalianol synthase (PEN4) from *A. thaliana*[3], and other members of the *A. thaliana* LUP and PEN clades.[36] AthLUP5 and AthPEN6 have been shown to access dammarenyl type intermediates,[44] as do AthPEN2 and AthPEN3.[45]

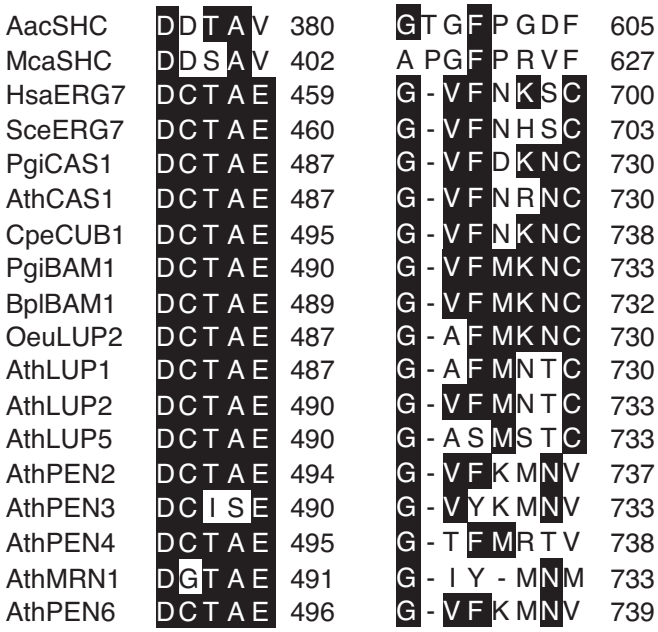


Figure S10. Amino acid sequence alignment of cyclases showing the distinctive G488 of Ath*MRN1* in the DCTAE motif; other oxidosqualene cyclases have cysteine at this position. *MRN1* has a residue deletion between positions 730 and 731 that is not present in other cyclases; this deletion occurs in a region that may be important for C and D ring formation.[13,46] Cyclases are defined and cited in the preceding text.

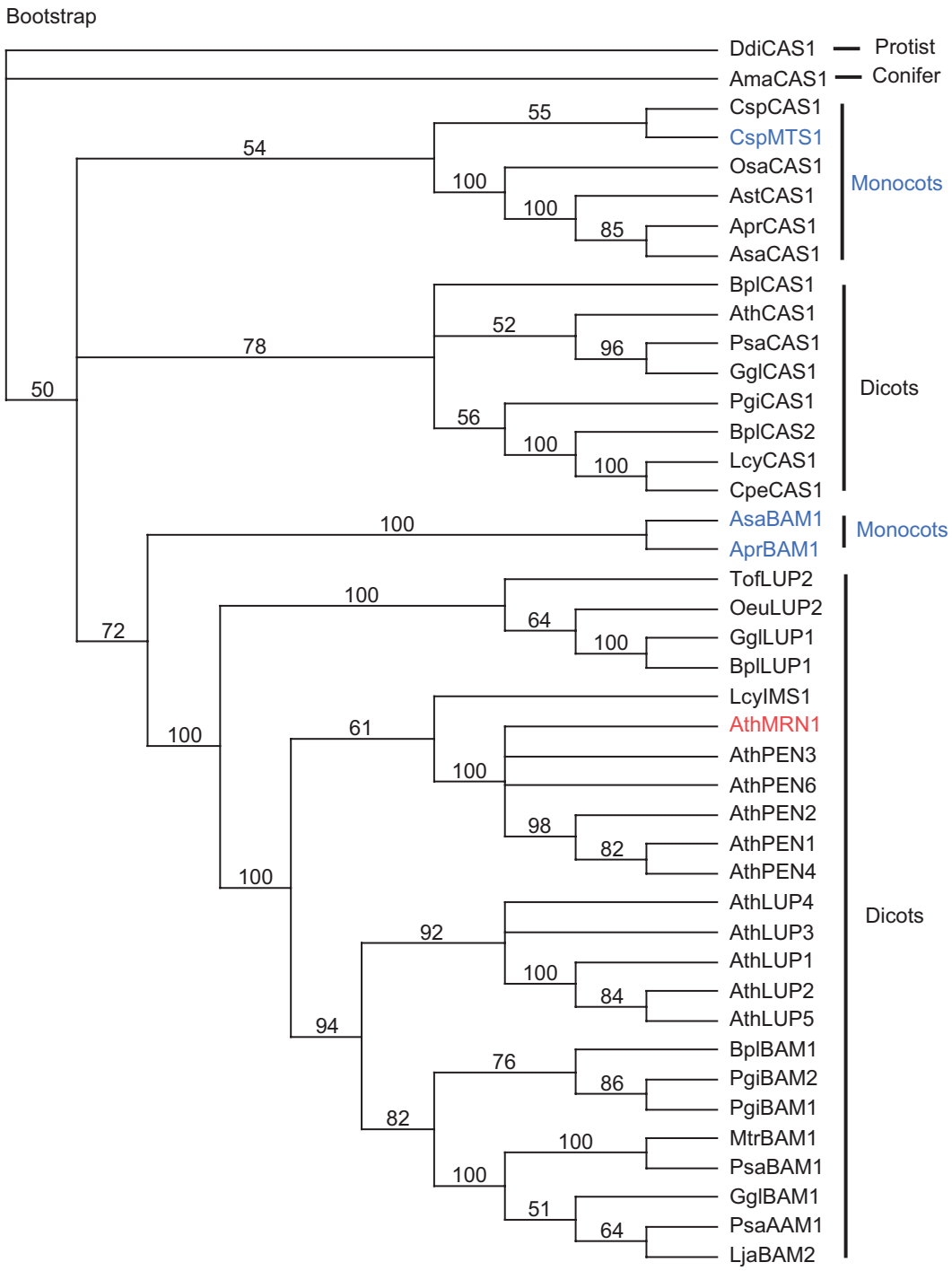


Figure S11. Phylogenetic analysis of plants oxidosqualene cyclases. MRN1 (PEN5, red) is closely related to the *Arabidopsis* PEN1-4 and PEN6 proteins and is distant from the known nonsteroidal triterpene synthases from monocots (blue). Cyclases are defined and cited in the preceding text.

References for Supporting Information

[1] a) H. Liu, J. Krizek, A. Bretscher, *Genetics* **1992**, *132*, 665-673; b) J. B. Herrera, B. Bartel, W. K. Wilson, S. P. T. Matsuda, *Phytochemistry* **1998**, *49*, 1905-1911.

[2] E. J. Corey, S. P. T. Matsuda, C. H. Baker, A. Y. Ting, H. Cheng, *Biochem. Biophys. Res. Commun.* **1996**, *219*, 327-331.

[3] G. C. Fazio, R. Xu, S. P. T. Matsuda, *J. Am. Chem. Soc.* **2004**, *126*, 5678-5679.

[4] R. H. Schiestl, R. D. Gietz, *Curr. Gent.* **1989**, *16*, 339-46.

[5] T. Akihisa, K. Arai, Y. Kimura, K. Koike, W. C. M. C. Kokke, T. Shibata, T. Nikaido, *J. Nat. Prod.* **1999**, *62*, 265-268.

[6] A. F. Barrero, E. J. Alvarez-Manzaneda Roldan, R. Alvarez-Manzaneda Roldan, *Tetrahedron Lett.* **1989**, *30*, 3351-3352.

[7] T. Akihisa, K. Koike, Y. Kimura, N. Sashida, T. Matsumoto, M. Ukiya, T. Nikaido, *Lipids* **1999**, *34*, 1151-1157.

[8] F. J. Marner, I. Longerich, *Liebigs Ann. Chem*. **1992**, 269-272.

[9] F. J. Marner, T. Kasel, *J. Nat. Prod.* **1995**, *58*, 319-323.

[10] (a) D. M. Grant, B. V. Cheney, *J. Am. Chem. Soc.* **1967**, *89*, 5315-5318; (b) K. Pihlaja, E. Kleinpeter, *Carbon-13 NMR Chemical Shifts in Structural and Stereochemical Analysis*, VCH, Weinheim, **1994**, pp. 58-71.

[11] Gaussian 03, Revisions B.05 and C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004**.**

[12] L.-W. Guo, W. K. Wilson, C. H. L. Shackleton, unpublished results.

[13] R. Thoma, T. Schulz-Gasch, B. D'Arcy, J. Benz, J. Aebi, H. Dehmlow, M. Hennig, M. Stihle, A. Ruf, *Nature* **2004**, *432*, 118-122.

[14] (a) P. Buzek, P. v. R. Schleyer, S. Sieber, *Chemie in unserer Zeit* **1992**, *26*, 116-128. (b) T. Laube, *Acc. Chem. Res.* **1995**, *28*, 399-405.

[15] C. J. Cramer, S. E. Denmark, P. C. Miller, R. L. Dorow, K. A. Swiss, S. R. Wilson, *J. Am. Chem. Soc.* **1994**, *116*, 2437-2447.

[16] R. G. Nadeau, R. P. Hanzlik, *Method. Enzymol*. **1969**, *15*, 346-351.

[17] D. Ochs, C. Kaletta, K.-D. Entian, A. Beck-Sickinger, K. Poralla, *J. Bacteriol.* **1992**, *174*, 298-302.

[18] A. Tippelt, L. Jahnke, K. Poralla, *Biochim. Biophys. Acta* **1998**, *1391*, 223-232.

[19] a) C. H. Baker, S. P. T. Matsuda, D. R. Liu, E. J. Corey, *Biochem. Biophys. Res. Commun.* **1995**, *213*, 154-160; b) C. K. Sung, M. Shibuya, U. Sankawa, Y. Ebizuka, *Biol. Pharm. Bull.* **1995**, *18*, 1459-1461.

[20] E. J. Corey, S. P. T. Matsuda, B. Bartel, *Proc. Natl. Acad. Sci.* USA **1994**, *91*, 2211-2215.

[21] S. M. Godzina, M. A. Lovato, M. M. Meyer, K. A. Foster, W. K. Wilson, W. Gu, E. L. de Hostos, S. P. T. Matsuda, *Lipids* **2000**, *36*, 249-255.

[22] J. B. R. Herrera, S. P. T. Matsuda, unpublished results.

[23] N. Kawano, K. Ichinose, Y. Ebizuka, *Biol Pharm Bull* **2002**, *25*, 477-482.

[24] L. B. Darr, S. M. Godzina, S. P. T. Matsuda, *Plant Physiol.* **1999**, *121*, 686.

[25] X. Qi, S. Bakht, M. Leggett, C. Maxwell, R. Melton, A. Osbourn, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 8233-8238.

[26] K. Haralampidis, G. Bryan, X. Qi, K. Papadopoulou, S. Bakht, R. Melton, A. E. Osbourn, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13431-13436.

[27] H. Zhang, M. Shibuya, S. Yokota, Y. Ebizuka, *Biol. Pharm. Bull.* **2003**, *26*, 642-650.

[28 M. Morita, M. Shibuya, M.-S. Lee, U. Sankawa, Y. Ebizuka, *Biol. Pharm. Bull.* **1997**, *20*, 770-775.

[29] H. Hayashi, N. Hiraoka, Y. Ikeshiro, T. Kushiro, M. Morita, M. Shibuya, Y. Ebizuka, *Biol. Pharm. Bull.* **2000**, *23*, 231-234.

[30] H. Hayashi, N. Hiraoka, Y. Ikeshiro, K. Yazaki, S. Tanaka, T. Kushiro, M. Shibuya, Y. Ebizuka, *Plant Physiol.* **1999**, *121*, 1384.

[31] H. Hayashi, P. Huang, S. Takada, M. Obinata, K. Inoue, M. Shibuya, Y. Ebizuka, *Biol. Pharm. Bull.* **2004**, *27*, 1086-1092.

[32] E. J. Corey, S. P. T. Matsuda, B. Bartel, *Proc. Natl. Acad. Sci.* USA **1993**, *90*, 11628-11632.

[33] T. Kushiro, M. Shibuya, Y. Ebizuka, *Eur. J. Biochem.* **1998**, *256*, 238-224.

[34] J. B. R. Herrera, B. Bartel, W. K. Wilson, S. P. T. Matsuda, *Phytochemistry* **1998**, *49*, 1905-1911.

[35] T. Kushiro, M. Shibuya, K. Masuda, Y. Ebizuka, *Tetrahedron Lett.* **2000**, *41*, 7705-7710.

[36] T. Husselstein-Muller, H. Schaller, P. Benveniste, *Plant Mol. Biol.* **2001**, *45*, 75-92.

[37] a) T. Kushiro, M. Shibuya, Y. Ebizuka, *Tetrahedron Lett.* **1999**, *40*, 5553-5556. b) M. Shibuya, H. Zhang, A. Endo, K. Shishikura, T. Kushiro, Y. Ebizuka, *Eur. J. Biochem.* **1999**, *266*, 302-307.

[38] T. Kushiro, M. Shibuya, Y. Ebizuka, in *Towards Natural Medicine Research in the 21st Century, Excerpta Medica International Congress Series 1157* (Eds.: H. Ageta, N. Aimi, Y. Ebizuka, T. Fujita, G. Honda), Elsevier Science, Amsterdam, **1998**, pp. 421-428.

[39] a) H. Suzuki, L. Achnine, R. Xu, S. P. T. Matsuda, R. A. Dixon, *Plant Journal* **2002**, *32*, 1033-1048; b) I. Iturbe-Ormaetxe, K. Haralampidis, K. Papadopoulou, A. E. Osbourn, *Plant Mol. Biol.* **2003**, *51*, 731-743.

[40] M. Morita, M. Shibuya, T. Kushiro, K. Masuda, Y. Ebizuka, *Eur. J. Biochem.* **2000**, *267*, 3453-3460.

[41] H. Hayashi, P. Y. Huang, A. Kirakosyan, K. Inoue, N. Hiraoka, Y. Ikeshiro, T. Kushiro, M. Shibuya, Y. Ebizuka, *Biol. Pharm. Bull.* **2001**, *24*, 912-916.

[42] H. Hayashi, P. Huang, K. Inoue, N. Hiraoka, Y. Ikeshiro, K. Yazaki, S. Tanaka, T. Kushiro, M. Shibuya, Y. Ebizuka, *Eur. J. Biochem.* **2001**, *268*, 6311-6317.

[43] M. Shibuya, S. Adachi, Y. Ebizuka, *Tetrahedron* **2004**, *60*, 6995-7003.

[44] Y. Ebizuka, Y. Katsube, T. Tsutsumi, T. Kushiro, M. Shibuya, *Pure Appl. Chem.* **2003**, *75*, 369-374.

[45] Q. Xiong, S. P. T. Matsuda, unpublished results.

[46] T. Hoshino, K. Shimizu, T. Sato, *Angew. Chem.* **2004**, *116*, 6868-6871; *Angew. Chem. Int. Ed.* **2004**, *43*, 6700-6703.

Atomic coordinates for molecular modeling

Coordinates are given for **9**, **10**, and **11** with the C3-C4 bond frozen at 1.60 Å. Coordinates for the same structures but optimized with acetate-2H_2O are also given.

All atomic coordinates herein are from B3LYP/6-31G* geometry optimizations. Frozen bonds are described by atom numbering from the figures in the text, not by atom numbering from the coordinate files. For economy of space, coordinates are given in condensed format. These data are easily converted to tabular form by global find-and-replace routines available in most word processors. First, replace the paragraph mark with nothing; spaces might also need to be deleted; then replace "|" with the paragraph mark. If desired, commas can be replaced by spaces or the tab mark.

Compound **9** (twist), C3-C4 bond frozen at 1.60 Å


```
1|1|UNPC-UNK|FOpt|RB3LYP|6-31G(d)|C15H27O1(1+)|PCUSER|03-Jun-2005|0||#
 B3LYP/6-31G* OPT GEOM=MODREDUNDANT||iridal C5 cation, Twist, PES160 w
ithout enzyme groups||1,1|C,-0.2474688718,1.803516109,-0.7060118861|C,
-1.69237326,2.0587925118,-0.2158576048|C,-2.3320406139,0.8761447531,0.
5144845872|C,-1.2884995766,0.0269264997,1.3804234805|C,-0.1739522975,-
0.4371252026,0.5232741731|C,0.4452428911,-1.7532186927,0.7852641379|C,
1.8368868189,-2.0208218316,0.1956315452|C,1.9857338404,-1.4501429986,-
1.2219368216|C,1.6709248288,0.0798049294,-1.2289172722|C,0.2087939181,
0.3274164883,-0.6851339581|O,-3.4252752891,1.2622698738,1.3113562231|H
,-3.3110356037,2.1811717367,1.6011619244|C,2.6870884765,0.8264372265,-
0.3312293721|C,1.1904498696,-2.283014338,-2.2448902635|C,-0.6004448553
,0.9606021244,2.447037465|C,-2.0464099004,-1.0908626779,2.1246220745|C
,1.7990944829,0.6317664414,-2.6644542408|H,0.4553348732,2.3800510824,-
0.0982882639|H,-0.1355558734,2.1752359161,-1.7253813848|H,-1.681642071
4,2.9234955974,0.4606840709|H,-2.3442822818,2.3376713024,-1.0490282693
|H,-2.7436358674,0.1663835934,-0.2148142899|H,0.3868320964,-1.99208850
33,1.8528747727|H,-0.2955412444,-2.4490718568,0.3363364249|H,2.0079277
319,-3.1019936945,0.192674088|H,2.5904250317,-1.5963146705,0.866109164
9|H,3.0446161976,-1.5393869973,-1.4973456571|H,-0.4663228412,-0.220209
2653,-1.3931222118|H,2.5687390025,0.5960828491,0.7339916762|H,3.707914
9644,0.5508699641,-0.6168338762|H,2.6066484571,1.9111361299,-0.4450203
872|H,0.1023541092,-2.2280606786,-2.1004306914|H,1.3964597031,-1.97440
89321,-3.2725051342|H,1.4689904357,-3.3388541672,-2.1626569308|H,-0.01
1840583,1.7719474042,2.0144880872|H,-1.4185767494,1.3909653814,3.03416
55209|H,0.037010773,0.383281823,3.1214722758|H,-2.4359554446,-1.845162
9098,1.4318899395|H,-1.4321721011,-1.592832038,2.8767506926|H,-2.89978
06125,-0.6399910605,2.6337683482|H,1.0020527456,0.2767625601,-3.325604
4029|H,1.7909692109,1.7251272536,-2.6842709526|H,2.7541137492,0.312705
1363,-3.095750584||Version=IA32W-G03RevC.02|HF=-662.8079268|RMSD=5.138
e-009|RMSF=8.403e-004|Dipole=0.0882424,0.2113513,0.4153893|PG=C01 [X(C
15H27O1)]||@
```


Compound **10** (boat), C3-C4 bond frozen at 1.60 Å


```
1\1\GINC-DFTC\FOpt\RB3LYP\6-31G(d)\C15H27O1(1+)\BILLW\05-Jun-2005\0\\#
 B3LYP/6-31G* OPT GEOM=MODREDUNDANT\\iridal C5 cation, Boat, PES160 wi
thout enzyme groups, O3-C5 not frozen, PES287 c\\1,1\C,0.5167413277,-1
.652771945,-1.1880496167\C,1.949352534,-1.6869640223,-0.5962212919\C,2
```

.4129729316,-0.3988136612,0.0958449089\C,1.7265930595,0.9271694108,-0.
4791782068\C,0.2482838978,0.7688441794,-0.4182443448\C,-0.6319709795,1
.9102849605,-0.1108707907\C,-1.5404239571,1.5227594774,1.1082282801\C,
-2.4033397857,0.301844233,0.7729895547\C,-1.5254109578,-0.9414298978,0
.4264908708\C,-0.4125295803,-0.5146807362,-0.7100473404\O,2.0777745367
,-0.3647135793,1.4804140219\H,2.5922536935,-1.0430194415,1.9468963852\
C,-0.8372996386,-1.4875817076,1.6896963439\C,-3.4842218259,0.649970866
8,-0.2670838491\C,2.2881758728,2.1551319091,0.2638362247\C,2.057960558
6,1.0833762433,-2.0096428611\C,-2.3935046849,-2.0669724231,-0.16891502
07\H,0.0282367944,-2.6123996282,-1.0148113329\H,0.5917672946,-1.570314
6946,-2.2772171291\H,2.0387495056,-2.4976264704,0.1355737959\H,2.65092
18198,-1.9340041281,-1.3996683893\H,3.495995915,-0.2713254595,-0.03609
44765\H,-0.0842099145,2.8296936953,0.0864078906\H,-1.3030931125,2.0835
447077,-0.9636482973\H,-2.1604503691,2.396225277,1.3326323357\H,-0.909
3864671,1.3459696424,1.983985458\H,-2.9307804563,0.0305510423,1.698633
8052\H,-1.0922846236,-0.2609405094,-1.5420803257\H,-0.224592786,-0.747
6278592,2.2055305891\H,-1.6060038289,-1.8518193809,2.3805440343\H,-0.1
904877023,-2.3378619898,1.452621415\H,-3.0813468109,0.8078931643,-1.27
67988321\H,-4.2379278061,-0.1368300654,-0.3445518025\H,-4.0038196343,1
.5691059144,0.0228901039\H,1.9783929796,2.1734156749,1.3102582956\H,3.
3810388285,2.1031107297,0.2445005518\H,2.002670561,3.0946481573,-0.217
8563879\H,1.7331955065,0.2400783006,-2.6203430231\H,1.6146065974,2.000
0882604,-2.4082331365\H,3.1462595039,1.1718738354,-2.0903955896\H,-2.7
996086274,-1.8198012792,-1.1540501721\H,-1.8431848485,-3.0075068632,-0
.254420937\H,-3.2373825454,-2.2584118064,0.5033845774\\Version=x86-Lin
ux-G03RevB.02\State=1-A\HF=-662.8087427\RMSD=7.562e-09\RMSF=2.275e-03\
Dipole=-0.4348092,-0.52337,0.1537199\PG=C01 [X(C15H27O1)]\\@

Compound **11** (chair), C3-C4 bond frozen at 1.60 Å

1\1\GINC-DFTB\FOpt\RB3LYP\6-31G(d)\C15H27O1(1+)\BILLW\06-Jun-2005\0\\#
 B3LYP/6-31G* OPT GEOM=MODREDUNDANT\\iridal C5 cation, chair, PES160 w
ithout enzyme groups, O3-C5 not frozen, B3LYP 2\\1,1\C,-0.6333517665,1
.0915362967,-1.5397749818\C,-1.8927079253,0.3021764524,-1.8816323428\C
,-2.5167532351,-0.2949443017,-0.6208927574\C,-1.5001890581,-1.24957936
,0.1634981831\C,-0.3128776157,-0.3710324067,0.4330835845\C,0.035455880
1,0.0069704018,1.804588666\C,1.481385567,0.4915602194,2.079712483\C,2.
40218125,0.3047168671,0.8613519138\C,1.7530093029,0.9664722083,-0.3954
510274\C,0.4096276865,0.2238403506,-0.7137738878\O,-2.8506214655,0.725
8268119,0.3132883873\H,-3.6800794759,1.152374585,0.0446309096\C,1.5412
036541,2.4746770648,-0.1213461852\C,2.7759064169,-1.1788284633,0.69044
72975\C,-2.205412492,-1.7946784246,1.4198961433\C,-1.1236165812,-2.463
8511335,-0.7392136077\C,2.6763221692,0.8429102555,-1.6254340516\H,-0.8
981245756,1.971737736,-0.9511864494\H,-0.1180101651,1.4302371547,-2.44
1997973\H,-2.6250371019,0.9763407017,-2.344663094\H,-1.6905343483,-0.4
820679193,-2.6189943846\H,-3.3917420958,-0.914097679,-0.8602514874\H,-
0.6973739463,0.834708951,1.9564021219\H,-0.3028481967,-0.7519032843,2.
514282002\H,1.8672881509,-0.0805073225,2.9294303344\H,1.4684397718,1.5
365969274,2.3969756609\H,3.3301839576,0.8538213943,1.0635895245\H,0.63
2267091,-0.6148631446,-1.3882851206\H,0.7690103726,2.6811486032,0.6271
355239\H,2.475410036,2.916082578,0.241465534\H,1.2659279165,3.01309305
14,-1.0326922799\H,1.9059597028,-1.8253059688,0.5028079442\H,3.4779912
886,-1.3358705465,-0.132363863\H,3.2535243244,-1.550227292,1.602846278
6\H,-2.5419650866,-1.0015099766,2.088679408\H,-3.0898936252,-2.3596987
618,1.109635732\H,-1.5612757846,-2.4869529152,1.9723830869\H,-0.663179
7102,-2.1893640158,-1.6885569073\H,-0.4423611898,-3.1336909209,-0.2056
491493\H,-2.0367470396,-3.0263612284,-0.9553338015\H,2.8628847998,-0.1

973339652,-1.9100081854\H,2.2491492952,1.3536452876,-2.4957193808\H,3.
645007841,1.3116773146,-1.4212256593\\Version=x86-Linux-G03RevB.02\HF=
-662.8148151\RMSD=4.372e-09\RMSF=4.157e-04\Dipole=0.9535733,0.2229017,
0.3137852\PG=C01 [X(C15H27O1)]\\@

Compound **9** (twist) with acetate-2H_2O, C3-C4 bond frozen at 1.60 Å. Also, the position of C7, C9, and the acetate methyl were fixed.

1|1|UNPC-UNK|FOpt|RB3LYP|6-31G(d)|C17H34O5|PCUSER|03-Jun-2005|0||# B3L
YP/6-31G* OPT GEOM=MODREDUNDANT||iridal C5 cation intermediate, Twist
PES160 from PES158 HF coord, restart again||0,1|C,1.3921360225,-1.8573
299771,0.2059649846|C,0.124120028,-1.9693499035,1.0840629727|C,-0.3784
839511,-0.6476138485,1.6811829004|C,-0.0930940063,0.6159989384,0.74212
20263|C,1.2928479968,0.7008619706,0.3288610067|C,1.9174109832,2.038470
9833,0.1045160061|C,3.1726559934,2.08085801,-0.77167|C,4.1501539974,0.
9421290162,-0.4433629974|C,3.4446470014,-0.4416589891,-0.611271|C,2.16
36870108,-0.5202769934,0.3060689919|O,-1.7032359537,-0.7133078547,2.09
93279079|H,-2.298140965,-1.0259188965,1.3616219373|C,3.0464350143,-0.6
343049908,-2.0945770017|C,4.8262309917,1.1664120212,0.9228700046|C,-0.
6658250027,0.4122279389,-0.745239981|C,-0.7337690162,1.8718849308,1.34
64900315|C,4.4281180138,-1.5705539831,-0.2344239987|H,1.1073610167,-2.
0145990086,-0.8377220091|H,2.0875200231,-2.6622939723,0.4563949984|H,-
0.6705199785,-2.3915289202,0.4625579912|H,0.2856550157,-2.6638848716,1
.9154330017|H,0.2014540514,-0.4194628136,2.5894718901|H,1.1651099705,2
.7562729741,-0.2324629853|H,2.1626489864,2.3704459818,1.1314350058|H,3
.6624149746,3.0519020128,-0.6348519968|H,2.8756609873,2.0358030078,-1.
8243809997|H,4.9493559999,0.9692880237,-1.1971349946|H,2.5358270051,-0
.4584259876,1.3550529936|H,2.2309090113,0.0261710043,-2.4072380043|H,3
.9064890151,-0.4341869857,-2.744260999|H,2.722705021,-1.6599929928,-2.
2909280027|H,4.12988499,1.0924080146,1.7676450026|H,5.6301439975,0.447
2450281,1.1012100063|H,5.2712809829,2.1673260251,0.9593000064|H,-0.435
7699913,-0.5519680609,-1.1952439757|H,-1.7751350024,0.4820809446,-0.59
03019974|H,-0.380678006,1.2303449398,-1.4145839815|H,-0.2336350125,2.1
579889362,2.2798470279|H,-0.7226440312,2.7275309311,0.6657250317|H,-1.
7745920118,1.6372539187,1.5710720392|H,4.6396930112,-1.6004819827,0.83
98410018|H,4.0467230202,-2.5531749851,-0.5256290004|H,5.3804160144,-1.
4250909773,-0.7581509961|C,-5.6884699986,-0.4472830179,-0.059321|C,-4.
1596629996,-0.381441998,-0.1076530165|O,-3.521496979,-1.4398329739,0.2
027899823|O,-3.615867998,0.7074239879,-0.4481440219|H,-6.0045119843,-0
.7661880128,0.9396550053|H,-6.0385979967,-1.2102700249,-0.7642719913|H
,-6.1329740126,0.5199509771,-0.3020400022|H,-4.6895349756,2.1596320008
,-0.7744210382|O,-5.3542099644,2.8319620083,-1.0473290468|H,-5.5211709
616,2.5977580048,-1.9721390464|O,-1.7214549848,-2.5344520009,-1.584101
0002|H,-2.0126509849,-3.4574000025,-1.6248599611|H,-2.4053929869,-2.10
3062991,-1.0137160133||Version=IA32W-G03RevC.02|State=1-A|HF=-1044.336
284|RMSD=7.784e-009|RMSF=1.038e-003|Dipole=5.2224016,-0.1438939,-0.260
8253|PG=C01 [X(C17H34O5)]||@

Compound **10** (boat) with acetate-2H_2O, C3-C4 bond frozen at 1.60 Å. Also, the position of C7, C9, and the acetate methyl were fixed.

1\1\GINC-APSARA\FOpt\RB3LYP\6-31G(d)\C17H34O5\BILLW\28-May-2005\0\\# B
3LYP/6-31G* OPT GEOM=ALLCHECK\\iridal C5 cation intermediate, Aboat PE
S 160 from 170 coord\\0,1\C,-1.7675493856,1.5898917735,-1.3791625168\C
,-0.3332452728,1.1803946175,-1.768110293\C,0.5646076385,0.6123026678,-

0.6389913482\C,-0.0531749682,0.8086801834,0.8238076809\C,-1.2977102671
,0.0812624128,0.5976124438\C,-1.3118721146,-1.3464410119,0.8980641387\
C,-2.6134382401,-2.1522724819,0.7168605707\C,-3.8413344168,-1.25866079
26,0.4735691111\C,-3.5309392401,-0.2476464819,-0.6727554293\C,-2.38229
26974,0.7164095431,-0.2199769842\O,0.8804293805,-0.7143715673,-0.90638
88954\H,1.8859019804,-0.871275924,-0.7107686994\C,-3.1333796576,-1.041
1194305,-1.9428780508\C,-4.3225200795,-0.6253636762,1.7927833502\C,0.9
213221485,0.1947049956,1.8510677775\C,-0.1710490991,2.3165602907,1.158
2247847\C,-4.7833111646,0.5867380028,-1.0090818413\H,-2.4171701489,1.5
133564787,-2.2540176288\H,-1.8003550402,2.6413605981,-1.0904353376\H,-
0.3671623994,0.3973885768,-2.530727533\H,0.1666553181,2.0413319949,-2.
2250578401\H,1.5110732871,1.1737633173,-0.5338972849\H,-0.4962943402,-
1.6454424195,0.1667516799\H,-0.8308048387,-1.5368884149,1.8619918208\H
,-2.7744086273,-2.7607015866,1.6137077927\H,-2.4948540291,-2.860943964
8,-0.1074155761\H,-4.6549194728,-1.9020348024,0.1111621967\H,-2.832925
9027,1.438740156,0.4826366295\H,-2.1431689885,-1.5013694753,-1.8707665
656\H,-3.8614344679,-1.8392761479,-2.1295576298\H,-3.1282936006,-0.401
2510287,-2.8304525794\H,-3.5568428476,-0.0010497256,2.2698287331\H,-5.
2142506215,-0.0069577528,1.6554061015\H,-4.5803995089,-1.4145871443,2.
5077748961\H,1.2662450514,-0.8008107457,1.5698125641\H,1.7964889755,0.
8451259545,1.921819171\H,0.4513132808,0.1476943487,2.8410129972\H,-0.7
512233941,2.9037757689,0.4496617081\H,-0.6386658117,2.4221882069,2.144
3055081\H,0.8353355913,2.7419864626,1.2251679974\H,-5.1143873191,1.202
0234614,-0.1660677375\H,-4.6001117106,1.2600978864,-1.8534900925\H,-5.
6141574228,-0.0703834846,-1.2916941513\C,5.5774757599,-0.1532544819,-0
.1916144293\C,4.0618493728,-0.0145613414,-0.342748802\O,3.5681319687,1
.1419624873,-0.3823036674\O,3.3763548354,-1.0910451107,-0.4205171801\H
,6.0734427644,0.444493838,-0.9631909889\H,5.8909277033,-1.196408462,-0
.2659523936\H,5.8853864188,0.257272572,0.7778129854\H,4.2856930294,-2.
69625657,-0.5209155093\O,4.8721862766,-3.4803003104,-0.4296700278\H,4.
9215610013,-3.5967168087,0.5305071597\H,3.402433229,2.4167738562,0.821
3171041\O,3.085833342,3.0049747691,1.5464281968\H,3.7352565385,2.87947
01033,2.253322117\\Version=x86-Linux-G03RevB.05\HF=-1044.3433116\RMSD=
5.003e-09\RMSF=3.021e-03\Dipole=4.28262,-0.7544828,1.2555246\PG=C01 [X
(C17H34O5)]\\@

Compound **11** (chair) with acetate-2H_2O, C3-C4 bond frozen at 1.60 Å. Also, the position of C7, C9, and the acetate methyl were fixed.

1\1\GINC-DFTC\FOpt\RB3LYP\6-31G(d)\C17H34O5\BILLW\25-May-2005\0\\# B3L
YP/6-31G* OPT GEOM=MODREDUNDANT\\iridal C5 cation, Achair, PES160 from
 PES158 coord, restart from dftb\\0,1\C,1.0203870274,1.3963049478,-0.9
784280127\C,-0.2925179879,0.85049198,-1.5230230118\C,-0.8684170167,-0.
2350279951,-0.6047290041\C,0.1675880095,-1.4392220605,-0.4133909886\C,
1.3054670108,-0.6866460621,0.1197160055\C,1.3808710089,-0.4779590529,1
.571731004\C,2.5456030083,0.3921559464,2.052553\C,3.7969940093,0.14343
39354,1.1905720039\C,3.5306230133,0.6315149256,-0.273656\C,2.081380013
9,0.2067129299,-0.7637830005\O,-1.2467460052,0.2998600201,0.6109969922
\H,-2.2557350081,0.0549090429,0.7642319945\C,3.7439380216,2.1612849243
,-0.3213000114\C,4.2552970023,-1.3222090658,1.3124540155\C,-0.46157199
78,-2.4647710502,0.5536540175\C,0.3775040004,-2.139471075,-1.776402982
5\C,4.5399770136,0.0091429131,-1.2633799931\H,0.8361060349,1.921920957
7,-0.0379130167\H,1.4788830449,2.1038889321,-1.6736640171\H,-1.0238049
684,1.6697639971,-1.5376140187\H,-0.1622689948,0.4726429703,-2.5440960
091\H,-1.7173370284,-0.7568729796,-1.0716050002\H,0.3757580099,-0.0247
890486,1.774132999\H,1.3282120051,-1.4444420494,2.0867600103\H,2.75497

10069,0.1715049529,3.1048550022\H,2.264997013,1.4504079471,2.011608992
9\H,4.6095340115,0.7651849347,1.5883480005\H,2.1763010112,-0.239753078
7,-1.7561369969\H,3.0824520227,2.7085249326,0.356601983\H,4.7762270221
,2.3906319216,-0.0341130109\H,3.5929400257,2.5613399174,-1.3289870148\
H,3.5651069998,-2.0282400662,0.8331780196\H,5.2398520023,-1.4783460737
,0.8629300183\H,4.3271139992,-1.6061330583,2.3689870178\H,-0.803510993
7,-2.0055930424,1.4812380151\H,-1.3319050016,-2.9157900462,0.070322020
5\H,0.2573619962,-3.2591550545,0.7883580215\H,0.8071740056,-1.50469608
64,-2.552283989\H,1.0297189919,-3.0104130804,-1.6484669756\H,-0.593834
0037,-2.4977140685,-2.1296669778\H,4.4253850084,-1.076261087,-1.349167
985\H,4.4045670181,0.436827906,-2.2640139967\H,5.5710570139,0.21669991
03,-0.9539239922\C,-5.9719459984,0.0762661258,0.667629\C,-4.5121349926
,0.5269170924,0.6235789943\O,-4.2491209651,1.7318170855,0.485584985\O,
-3.6457680165,-0.4253469267,0.7567440001\H,-6.6433659823,0.9131791391,
0.4651869948\H,-6.1962790082,-0.339975862,1.6564930033\H,-6.1473820194
,-0.7245328758,-0.0602749939\H,-3.5859440505,-1.8233409365,-0.33227598
92\O,-3.3662370664,-2.4611289472,-1.0571079846\H,-3.9708440618,-2.2147
439387,-1.7725909856\H,-2.690458946,2.5960290466,0.0200449764\O,-2.101
3659314,3.2257310295,-0.4582150285\H,-2.7413819184,3.7763540417,-0.933
2770319\\Version=x86-Linux-G03RevB.02\State=1-A\HF=-1044.3508104\RMSD=
3.540e-09\RMSF=3.344e-03\Dipole=3.5580852,-0.9054208,-0.7498417\PG=C01
 [X(C17H34O5)]\\@