Enzymes

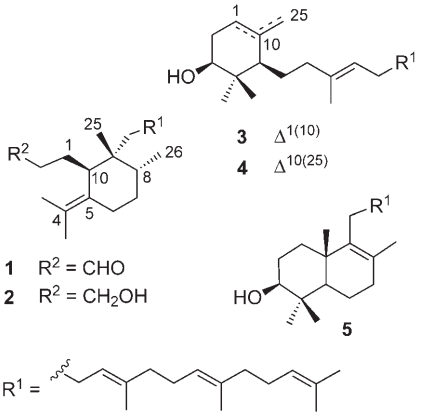
DOI: 10.1002/ange.200503420

An *Arabidopsis* Oxidosqualene Cyclase Catalyzes Iridal Skeleton Formation by Grob Fragmentation**

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

Oxidosqualene cyclases generate over 100 different triterpene skeletons from a single substrate.[1] This diversity is achieved by variations in three mechanistic motifs: cation–π annulation, cationic migration (including 1,2-shifts and ring expansions), and deprotonation. We encountered a fourth mode while characterizing the *Arabidopsis thaliana* triterpene synthase gene *At5g42600*. The encoded enzyme annulates oxidosqualene to a bicyclic intermediate, which undergoes 1,2-shifts to a C5 cation. Ring A is then cleaved to form the 3,4-*seco* aldehyde **1** by a Grob fragmentation,[2] which has not previously been documented in triterpene biosynthesis. Aldehyde **1** has never been isolated but was hypothesized to be the first carbocyclic precursor[3] in the biosynthesis of iridals. These unusual triterpenoids have been found only in Iridaceae,[3,4] a monocot family that is phylogenetically distant from *Arabidopsis*. Iridals possess diverse biological activities[5] and are the precursors of irones, which give the prized violet scent to orris oil.[3b] Herein, we describe the cloning and heterologous expression of the cyclase gene, the structure elucidation of **1**, the fragmentation mechanism, and phylogenetic implications.

The *At5g42600* gene was amplified by the reverse-transcription polymerase chain reaction (RT-PCR) from *A. thaliana* mRNA and subcloned into the pRS426GAL vector.[6a,b] The resultant plasmid pXQ11.2 was expressed in the yeast hosts SMY8 and RXY6; SMY8 lacks lanosterol synthase,[6c] and RXY6 additionally lacks squalene epoxidase.[6d] Triterpene products were isolated from the non-saponifiable lipids (NSL) of SMY8[pXQ11.2] and identified by GC mass-spectrometric and NMR spectroscopic analysis. We observed only traces of $C_{30}H_{50}O$ triterpene alcohols, namely, **3**,[7a] **4**,[7b] and **5**[7c] (Scheme 1). The predominant product was a novel 3,4-*seco* triterpene alcohol **2** bearing a typical iridal skeleton, and its $C_{30}H_{52}O$ formula suggested a postcyclization reduction. Oxidation of **2** with Dess–Martin periodinane[8] gave the putative precursor **1**, which provided spectral data that enabled us to detect this elusive aldehyde in the NSL of SMY8[pXQ11.2] (Figure 1a). Interestingly, direct hexane extraction of cell homogenates without saponification showed **1** rather than **2** as the major enzymatic product from both in vivo SMY8[pXQ11.2] cultures and an in vitro reaction of RXY6[XQ11.2] with oxidosqualene (Figure 1b,c). These results indicated that **1** was the direct enzymatic product and that **2** arose from further metabolism and/or during saponification.



Scheme 1. MRN1 products: marneral (**1**), marnerol (**2**), camelliol C (**3**), achilleol A (**4**), and tentative structure **5**. Triterpene numbering is used.

The 8α-methyl configuration in **1** and **2** was determined from NOESY spectra and ^{1}H NMR coupling constants.

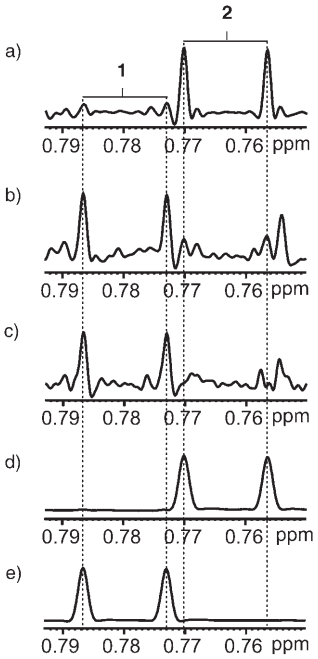


Figure 1. ^{1}H NMR spectra ($CDCl_3$, 500 MHz, 8α-methyl signals) showing the enzymatic formation of **1** versus **2**. a) Hexane extract of the NSL of SMY8[pXQ11.2] cell pellets. b) Hexane extract of SMY8-[pXQ11.2] cell homogenate. c) Hexane extract of RXY6[pXQ11.2] cell homogenates incubated with oxidosqualene. d) Purified sample of **2**. e) Purified sample of **1** from Dess–Martin oxidation. Ratios of **1**/**2** in a), b), and c) were 1:12, 10:3, and 99:1, respectively.

[*] Dr. Q. Xiong, W. K. Wilson, Prof. Dr. S. P. T. Matsuda
Department of Chemistry and
Department of Biochemistry and Cell Biology
Rice University, Houston, TX 77005 (USA)
Fax: (+1) 713-348-5154
E-mail: matsuda@rice.edu

[**] This work was supported by The National Science Foundation (MCB-0209769).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

Conformational heterogeneity of the cyclohexane ring was taken into account by calculating the Boltzmann distribution from B3PW91/6-311G(2d,p)//B3LYP/6-31G* energies (see the Supporting Information). The results showed the same relative stereochemistry observed in natural 10-deoxy-17-hydroxyiridal,[3a] and thus validated the hypothesis of Marner and Longerich[3] that cyclization of oxidosqualene into iridals proceeds via a B-ring boat intermediate to **1**. In recognition of the pioneering work of Marner on iridals, we named **1** marneral and **2** marnerol. Marneral synthase (MRN1) provides the first definitive example of ring cleavage accompanying oxidosqualene cyclization.

A likely mechanism that leads to **1** is shown in Scheme 2. Oxidosqualene (**6**) is protonated by aspartate 487 and cyclized to the bicyclic carbocation **8**. A series of 1,2-shifts moves H9 to C8, C25 to C9, and H5 to C10 as the positive charge migrates to C5. The last migration forces a conformational reorganization of ring A, which can proceed via twist **9** (inversion of C1) or boat **10** (inversion of C3 and C4) en route to the fully inverted chair **11**. Conformers **10** and **11**, which place the C3–C4 bond in hyperconjugation with the C5 cation, readily form **1** by barrierless[9] Grob fragmentation,[2] with transfer of the 3-hydroxy proton to the hydrogen-bonded aspartate. Conformer **9** may also undergo fragmentation but is hyperconjugatively poised for 4β-methyl migration as well. Other enzymes that generate a C5 cation make Δ5 products by deprotonation at C6 or 3-keto products by 4β-methyl and 3α-hydride shifts. MRN1 avoids these alternatives and accurately selects only the fragmentation pathway. This selectivity appears to be a function of the B-ring conformation and the aspartate mobility.

Sequence alignments for MRN1 revealed distinctive residues that may contribute to its unusual catalytic mechanism (Figure 2). The crystal structure of human lanosterol synthase[14] shows that Cys456 is hydrogen bonded to the protonating aspartate residue (Asp455). Unlike other oxidosqualene cyclases, MRN1 contains glycine in place of cysteine. This modification may facilitate the deprotonation of the axial hydroxy group in **10** and **11** by increasing the mobility and basicity of the aspartate residue in MRN1. MRN1 also differs from other oxidosqualene cyclases by a deletion between positions 730 and 731. Residues in this region are involved in substrate folding to form rings C and D.[1a,c,14] A single-residue deletion in the corresponding region of squalene-hopene cyclase led to incomplete cyclization, with some stereochemical inversion.[15] The deletion in MRN1 may have a similar effect. More detailed understanding of the enzyme mechanism awaits a crystal structure of a marneral synthase.

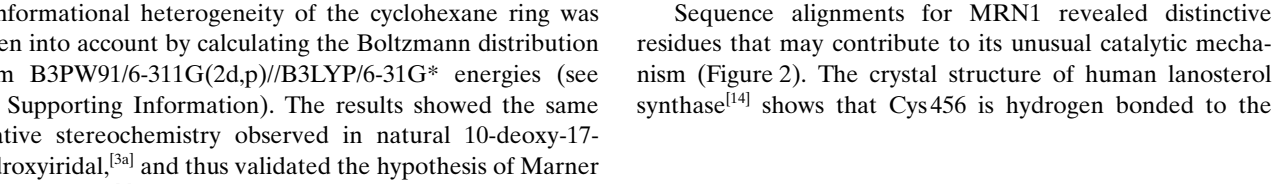
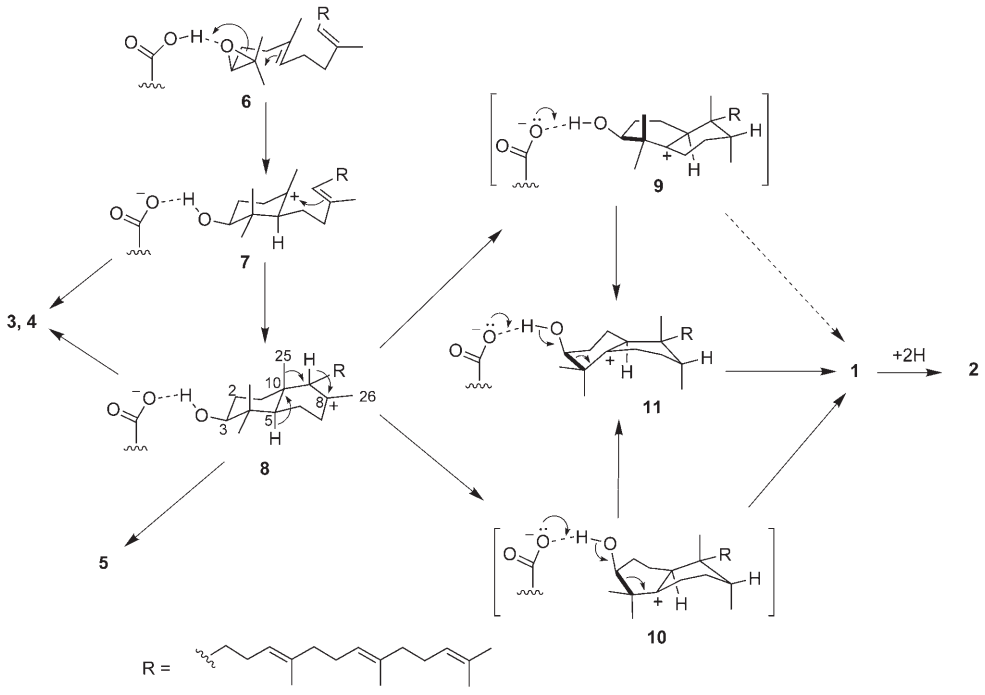


Figure 2. Partial amino acid sequence alignments of MRN1 with squalene-hopene cyclase from *Alicyclobacillus acidocaldarius* (AacSHC),[11] lanosterol synthase from *Homo sapiens* (HsaERG7),[12] and cycloartenol synthase from *A. thaliana* (AthCAS1).[13] More extensive alignments are given in the Supporting Information.



Scheme 2. Suggested mechanistic pathway from oxidosqualene (**6**) to marneral (**1**). Bonds elongated by hyperconjugation with the C5 cation are shown thickened. By-products **3** and **4** could arise from either **8**[10] or **7**.

No monocot marneral synthase has been characterized, but it seems clear that the *Arabidopsis* and *Iris* marneral synthases evolved independently from one another. A phylogenetic tree (see the Supporting Information) shows distant relationships between monocot and eudicot cyclases[16] and indicates that MRN1 arose within the eudicot PEN clade[17] from enzymes that generate the all-chair dammarenyl cation. MRN1 has diverged relatively little from the other PEN genes (70–80% identical) but has developed the ability to cleave ring A and achieved a rare evolutionary transition from a B-ring chair to B-ring boat (**7**) mechanism.[18] This case is a striking example of how enzymes with close phylogenetic affinity can have fundamental differences in mechanism and product structure.

Grob fragmentation is a useful synthetic reaction but its role in natural-products biosynthesis is not widely recognized.[19] Several other 3,4-*seco* triterpenoids with a C4–C5

tetrasubstituted olefin apparently arise via a B-ring boat intermediate by Grob fragmentation. These include lepidolide[20a] and (24*E*)-3,4-secocucurbita-4,24-diene-3,26-dioic acid[20b] from *Russula lepida*. A similar biosynthetic pathway can be ascribed to camelliol B[7b] and the sesquiterpene coumarins secodriol and secodrial from *Achillea ochroleuca*.[20c] An analogous mechanism via a B-ring chair may pertain to crystallopicrin from *Cortinarius* species,[20d] helianol, isohelianol[20e] and sasanquol from *Camellia sasanqua*,[20f] and graminol A from *Triticum aestivum*.[20g] However, most 3,4-*seco* triterpenoids[21a–d] with a C4–C5 saturated bond arise from a discrete postcyclization step, as do some diterpenoids, such as geayine[21e] and 3,4-secosonderianol.[21f]

In summary, our cloning and characterization of an *Arabidopsis* cyclase uncovered the elusive carbocyclic precursor of iridal triterpenes and provided the first experimental proof of Grob fragmentation in triterpene synthesis. Molecular modeling and sequence comparison suggested how cyclases can modify the reaction kinetics to select either the Grob fragmentation or other termination pathways. The independent development of iridal biosynthesis in monocots and the eudicot PEN clade underscores the complications in using chemotaxonomy to understand phylogenetic relationships and shows that nonsteroidal cyclases can readily evolve to generate diverse triterpene skeletons. Our results exemplify the power of genome mining to reveal novel structures and catalytic activities that conventional analysis of the native organism would not detect.

Experimental Section

cDNA of the *At5g42600* (*PEN5*, 2286 bp) was transcribed from mRNA obtained from 2-day-old *A. thaliana* seedlings. A culture of SMY8[pXQ11.2] (100 mL) was induced with galactose in a synthetic complete medium lacking uracil and grown to saturation with shaking at 30 °C. The cell pellet was saponified with ethanolic KOH at 70 °C for 2 h, followed by extraction with hexane. Analysis of the organic extracts by NMR spectroscopy (500 MHz, $CDCl_3$, 25 °C) and GC-MS (35 % phenyl polysiloxane column, 30 m) showed **2** as the only triterpene alcohol. Chromatographic purification (silica gel, CH_2Cl_2/hexane 5:1) of the NSL (48 mg) from a 3-L culture gave, in addition to fractions containing traces of **3**–**5** (< 1 % of total enzymatic products), alcohol **2** (1 mg) as a colorless oil: ^{1}H NMR (500 MHz, $CDCl_3$): $\delta =$ 0.764 (d, $J = 6.8$ Hz, 3H), 0.891 (s, 3H), 1.58–1.68 (6 br s, 18H), 1.728 (dqd, $J = 12.8$, 6.8, 4.2 Hz, 1H), 2.560 (dd, $J = 11.6$, 2.6 Hz, 1H), 3.608 (t, $J = 5.6$ Hz, 2H), 5.05–5.11 ppm (m, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): $\delta = 15.6$, 15.9, 16.0, 17.7, 20.2, 20.9, 22.4, 23.4, 24.3, 24.8, 25.7, 26.6, 26.8, 31.0, 31.2, 31.7, 36.2, 39.1, 39.7, 39.8, 44.3, 63.8, 123.4, 124.3, 124.4, 125.7, 131.2, 131.6, 134.2, 134.9 ppm; MS (EI, trimethylsilyl ether) *m*/*z*: 500 [M^+], 69 (base); the carbon skeleton of **2** was elucidated by HSQC, HMBC, and COSYDEC and by comparison with the NMR data for the side chain of thalianol[6d] and camelliol C.[7a] The C8 configuration was deduced as follows: In the NOESY spectrum, H26 correlated with H25 and H11*S*, and H25 correlated with H8 and H10α; the ^{1}H NMR spectrum showed couplings of 12.8 and 4.2 Hz between H8 and the C7 protons; these data are fully compatible with the 8α-methyl chair conformer with the equatorial C8 and C9 methyl groups but are incompatible with any conformer of the 8β-methyl isomer. Conformational analysis and the NOESY spectrum are given in the Supporting Information.

Oxidation of **2** to **1**: Dess–Martin periodinane reagent[8] (100 mg) was added to alcohol **2** (500 μg) in CH_2Cl_2 (2 mL). The reaction mixture was stirred for 24 h at room temperature. After removal of the solvent, the residue was partitioned between hexane/H_2O. The hexane fraction was purified by chromatography on silica gel (CH_2Cl_2/hexane 1:1) to give aldehyde **1** in 80 % yield as a colorless oil: ^{1}H NMR (500 MHz, $CDCl_3$): $\delta = 0.780$ (d, 6.8 Hz, 3H), 0.920 (s, 3H), 1.57–1.68 (6 br s, 18H), 1.750 (1H, dqd, $J = 12.5$, 6.8, 4.0 Hz, 1H), 2.556 (dd, $J = 12.0$, 3.8 Hz, 1H), 5.04–5.11 (m, 3H), 9.710 ppm (t, $J =$ 1.8 Hz, 1H); ^{13}C NMR (125 MHz, $CDCl_3$): $\delta = 15.5$, 15.9, 16.0, 17.7, 19.0, 20.2, 20.8, 22.3, 24.2, 24.8, 25.7, 26.7, 26.7, 30.7, 31.6, 36.0, 39.0, 39.7, 39.7, 42.3, 43.9, 124.2, 124.2, 124.8, 125.3, 130.6, 131.2, 134.2, 134.9, 203.2 ppm; MS (EI) *m*/*z*: 426 [M^+], 207, 187, 69 (base).

The samples for the NMR data in Figure 1a,b were obtained from 30-mL cultures of SMY8[pXQ11.2] by a) extraction with hexane of the NSL from the cell pellet or b) by precipitation of homogenized cells with ethanol, partial evaporation of the supernatant, and extraction with hexane. The samples for Figure 1c were obtained by incubation of a cell homogenate from a 30-mL RXY6[pXQ11.2] culture with 2,3-oxidosqualene for 16 h at ambient temperature, followed by the workup used in (b).[22]

Received: September 28, 2005

Keywords: cyclization · fragmentation · iridals · oxidosqualene · terpenoids

[1] a) K. U. Wendt, G. E. Schulz, E. J. Corey, D. R. Liu, *Angew. Chem.* **2000**, *112*, 2930–2952; *Angew. Chem. Int. Ed.* **2000**, *39*, 2812–2833; b) R. Xu, G. C. Fazio, S. P. T. Matsuda, *Phytochemistry* **2004**, *65*, 261–291; c) K. U. Wendt, *Angew. Chem.* **2005**, *117*, 4032–4037; *Angew. Chem. Int. Ed.* **2005**, *44*, 3966–3971.

[2] a) C. A. Grob, P. W. Schiess, *Angew. Chem.* **1967**, *79*, 1–14; *Angew. Chem. Int. Ed. Engl.* **1967**, *6*, 1–15; b) C. A. Grob, *Angew. Chem.* **1969**, *81*, 543–554; *Angew. Chem. Int. Ed. Engl.* **1969**, *8*, 535–546; c) P. Weyerstahl, H. Marschall in *Comprehensive Organic Synthesis, Vol. 6* (Eds.: B. M. Trost, I. Fleming), Pergamon, Oxford, **1991**, pp. 1041–1070; d) T. Constantieux, J. Rodriguez in *Science of Synthesis: Houben–Weyl Methods of Molecular Transformations, Vol. 26* (Ed.: J. Cossy), Georg Thieme, New York, **2005**, pp. 413–462.

[3] a) F.-J. Marner, I. Longerich, *Liebigs Ann. Chem.* **1992**, 269–272; b) F.-J. Marner, *Curr. Org. Chem.* **1997**, *1*, 153–186.

[4] H. Ito, S. Onoue, Y. Miyake, T. Yoshida, *J. Nat. Prod.* **1999**, *62*, 89–93.

[5] a) T. Yoshida, H. Ito, *Curr. Top. Phytochem.* **2000**, *4*, 135–145; b) J.-P. Bonfils, F. Pinguet, S. Culine, Y. Sauvaire, *Planta Med.* **2001**, *67*, 79–81; c) L. Shao, N. E. Lewin, P. S. Lorenzo, Z. Hu, I. J. Enyedy, S. H. Garfield, J. C. Stone, F.-J. Marner, P. M. Blumberg, S. Wang, *J. Med. Chem.* **2001**, *44*, 3872–3880; d) K. Takahashi, S. Suzuki, Y. Hano, T. Nomura, *Biol. Pharm. Bull.* **2002**, *25*, 432–436; e) F. Benoit-Vical, C. Imbert, J. P. Bonfils, Y. Sauvaire, *Phytochemistry* **2003**, *62*, 747–751.

[6] a) H. Liu, J. Krizek, A. Bretscher, *Genetics* **1992**, *132*, 665–673; b) J. B. R. Herrera, B. Bartel, W. K. Wilson, S. P. T. Matsuda, *Phytochemistry* **1998**, *49*, 1905–1911; c) E. J. Corey, S. P. T. Matsuda, C. H. Baker, A. Y. Ting, H. Cheng, *Biochem. Biophys. Res. Commun.* **1996**, *219*, 327–331; d) G. C. Fazio, R. Xu, S. P. T. Matsuda, *J. Am. Chem. Soc.* **2004**, *126*, 5678–5679.

[7] a) T. Akihisa, K. Arai, Y. Kimura, K. Koike, W. C. M. C. Kokke, T. Shibata, T. Nikaido, *J. Nat. Prod.* **1999**, *62*, 265–268; b) A. F. Barrero, E. J. Alvarez-Manzaneda Roldan, R. Alvarez-Manzaneda Roldan, *Tetrahedron Lett.* **1989**, *30*, 3351–3352; c) for the tentative structural evidence suggested by ^{1}H NMR spectroscopic and GC mass-spectrometric analysis, see the Supporting Information.

[8] D. B. Dess, J. C. Martin, *J. Org. Chem.* **1983**, *48*, 4155–4156.

[9] A cross-section of the potential energy surface was obtained by varying the C3–C4 bond lengths in **10** and **11**, as described in the

Supporting Information; the energy decreased with essentially no barrier as the C3–C4 bond was broken and the 3-hydroxy proton was transferred to aspartate (modeled as acetate stabilized by two hydrogen-bonded water molecules).

[10] Molecular modeling suggested that monocycles **3** and **4** might arise from a bicyclic cation by a mechanism that we now identify as Grob fragmentation; see the Supporting Information for Q. Xiong, X. Zhu, W. K. Wilson, A. Ganesan, S. P. T. Matsuda, *J. Am. Chem. Soc.* **2003**, *125*, 9002 – 9003.

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

[12] 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.

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

[14] 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.

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

[16] These relationships are also supported by the following reports: a) K. Haralampidis, G. Bryan, X. Qi, K. Papadopoulou, S. Bakht, R. Melton, A. Osbourn, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13431 – 134316; b) X. Qi, S. Bakht, M. Leggett, C. Maxwell, R. Melton, A. Osbourn, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8233 – 8238.

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

[18] This mechanistic transition is rare, at least among tetra- and pentacyclic triterpenes: Q. Xiong, F. Rocco, W. K. Wilson, R. Xu, M. Ceruti, S. P. T. Matsuda, *J. Org. Chem.* **2005**, *70*, 5362 – 5375.

[19] The lability of **1** and the weak expression of *MRN1* in *A. thaliana* suggest that **1** and other potential fragmentation products may easily be overlooked in surveys of natural products; for microarray results of MRN1, see: https://www.genevestigator.ethz.ch/ and P. Zimmermann, M. Hirsch-Hoffmann, L. Hennig, W. Gruissem, *Plant Physiol.* **2004**, *136*, 2621 – 2632.

[20] a) J. W. Tan, Z. J. Dong, Z. H. Ding, J. K. Liu, *Z. Naturforsch. C* **2002**, *57*, 963 – 965; b) J. W. Tan, Z. J. Dong, J. K. Liu, *Helv. Chim. Acta* **2000**, *83*, 3191 – 3197; c) H. Greger, O. Hofer, W. Robien, *Phytochemistry* **1983**, *22*, 1997 – 2003; d) W. Steglich in *Biologically Active Molecules* (Ed.: U. P. Schlunegger), Springer, Berlin, **1989**, pp. 1 – 8; e) T. Akihisa, Y. Kimura, K. Koike, T. Shibata, Z. Y. Yoshida, T. Nikaido, T. Tamura, *J. Nat. Prod.* **1998**, *61*, 409 – 412; f) T. Akihisa, K. Yasukawa, Y. Kimura, S. Yamanouchi, T. Tamura, *Phytochemistry* **1998**, *48*, 301 – 305; g) T. Akihisa, K. Koike, Y. Kimura, N. Sashida, T. Matsumoto, M. Ukiya, T. Nikaido, *Lipids* **1999**, *34*, 1151 – 1157.

[21] a) W. J. Baas, *Phytochemistry* **1985**, *24*, 1875 – 1889; b) J. D. Connolly, R. A. Hill, *Nat. Prod. Rep.* **2002**, *19*, 494 – 513; c) J. D. Connolly, R. A. Hill, *Nat. Prod. Rep.* **2003**, *20*, 640 – 659; d) L. M. Liao, P. C. Vieira, E. Rodrigues-Filho, J. B. Fernandes, M. F. G. F. Silva, *Z. Naturforsch. C* **2002**, *57*, 403 – 406; e) G. Palazzino, E. Federici, P. Rasoanaivo, C. Galeffi, F. Delle Monache, *Gazz. Chim. Ital.* **1997**, *127*, 311 – 314; f) A. A. Craveiro, E. R. Silveiro, *Phytochemistry* **1982**, *21*, 2571 – 2574.

[22] The Supporting Information contains additional details on cloning, transformation, expression, saponification, isolation, structure elucidation of **1**–**5**, NMR spectral data, GC-MS data, molecular modeling, sequence alignments, and phylogenetic analysis.

