Article

Conformational Energetics of Cationic Backbone Rearrangements in Triterpenoid Biosynthesis Provide an Insight into Enzymatic Control of Product

La#szlo# Ku#rti, Rong-Jie Chein, and E. J. Corey

J. Am. Chem. Soc., **2008**, 130 (28), 9031-9036 • DOI: 10.1021/ja800980h • Publication Date (Web): 18 June 2008

J|A|C|S
ARTICLES

Published on Web 06/18/2008

Conformational Energetics of Cationic Backbone Rearrangements in Triterpenoid Biosynthesis Provide an Insight into Enzymatic Control of Product

László Kürti, Rong-Jie Chein, and E. J. Corey*

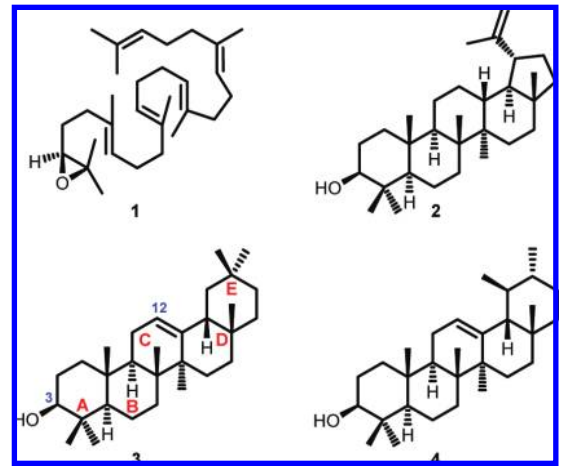
Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received February 7, 2008; E-mail: corey@chemistry.harvard.edu

Abstract: 2,3-(*S*)-Oxidosqualene ($C_{30}H_{50}O$) serves as a versatile starting point for the remarkable biosynthesis of many isomeric naturally occurring triterpenoids of formula $C_{30}H_{50}O$. These biosyntheses all involve polycyclization via cationic intermediates. The fully cyclized primary products then are converted to various structures by cationic rearrangements involving the polycyclic backbone. The energetics of these rearrangements has been examined by B3LYP 6−31 G* DFT calculations and by *ab initio* Hartree−Fock calculations at the 6−31G* or 3−21G(*) level. The results have led to the conclusion that the biosynthesis of friedelin, the most drastically rearranged of the pentacyclic triterpenes, involves a complex nonstop process, with no stable intermediates between 2,3-(*S*)-oxidosqualene and friedelin. It is proposed that this single-reaction biosynthesis consists of pentacyclization to the lupanyl cation followed directly by a sequence of 10 suprafacial 1,2-shifts of carbon and hydrogen, driven by the large exergonicity of the pentacyclization and electrostatic acceleration of the rearrangement steps.

The primary tetracyclic and pentacyclic triterpenes ($C_{30}H_{50}O$) comprise a set of natural products that are formed from the precursor 2,3-(*S*)-oxidosqualene **1** ($C_{30}H_{50}O$) by cation-olefin-type cyclization.[1] Although there is also scant information on the enzymes that produce them and the way in which they function to control product formation, it is clear that one important role of the enzyme is to organize **1** into a folded conformation that corresponds to the stereochemistry of the initial cyclization product.[2] A huge number of natural products arise from these isomeric primary cyclization products by further oxidation. Such oxidative reactions take place at almost any of the carbons and result in a tremendous structural diversity. The structures of the $C_{30}H_{50}O$ parents bear a relationship to one another that can be depicted mechanistically by skeletal interconversion of carbocations through 1,2-suprafacial shifts[3] of carbon and hydrogen. Thus, protonation of lupeol (**2**, original member of the lupane family) followed by suprafacial 1,2-rearrangement can produce two of the most common $C_{30}H_{50}O$ parent members, β-amyrin (**3**, oleane family) and α-amyrin (**4**, ursane class).

The crux of the present study is the analysis of the intermediates involved in the biosynthesis of the $C_{30}H_{50}O$ triterpenoid friedelin (**5**), the structure of which was first determined by one of us with Ursprung in February 1955.[4,5] A remarkable result in the structural investigation on friedelin was the finding that reduction of the ketonic function to hydroxyl and subsequent acid treatment produces a mixture of $\Delta^{12,13}$ and $\Delta^{13,18}$ oleanes (i.e., 3-deoxy-β-amyrins, **6**). This chemical linkage of the friedelane and oleane structures, which was proposed to occur via a sequence of 1,2-shifts of H and CH_3 along the pentacyclic backbone, suggested that friedelin **5** might arise biosynthetically from β-amyrin **3** by protonation of the 12,13-double bond at C(12) and suprafacial shifts of CH_3 and H along the backbone



(1) For a comprehensive survey see: Dev, S.; Gupta, A. S.; Patwardhan, S. A. *CRC Handbook of Terpenoids; Triterpenoids, Volume I and II*; CRC Press: Boca Raton, FL, 1986.
(2) For recent reviews see: (a) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812–2833. (b) Xu, R.; Fazio, G. C.; Matsuda, S. P. T. *Phytochemistry* **2004**, *65*, 261–291. (c) Yoder, R. A.; Johnston, J. N. *Chem. Rev.* **2005**, *105*, 4730–4756. (d) Eschenmoser, A.; Arigoni, D. *Helv. Chim. Acta* **2005**, *88*, 3011–3050.
(3) (a) Corey, E. J.; Ursprung, J. J. *J. Am. Chem. Soc.* **1956**, *78*, 183–188. (b) Corey, E. J.; Ursprung, J. J. *Chem. Ind.* **1954**, 1387–1388.
(4) (a) Corey, E. J. Abstracts 14th National Organic Symposium, June 1955, pp 81−82. (b) Corey, E. J.; Ursprung, J. J. *J. Am. Chem. Soc.* **1955**, *77*, 3667–3668. (c) Corey, E. J.; Ursprung, J. J. *J. Am. Chem. Soc.* **1956**, *78*, 5041–5051.
(5) Friedelin was one of the earliest pentacyclic triterpenoids to be studied; see: Chevreul, M *Ann. Chim.* **1807**, *62*, 323–325.

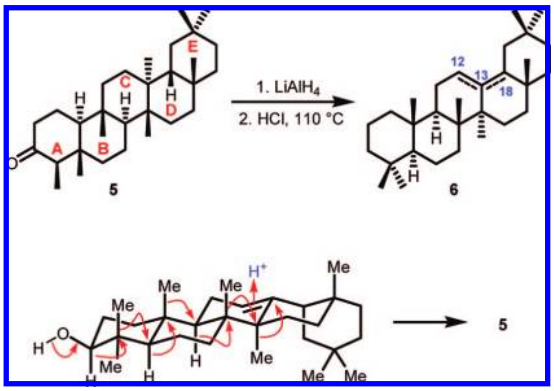


Figure 1. Representation of the rearrangement of β-amyrin (**3**) to friedelin (**5**) by way of carbocations and suprafacial 1,2-shifts of CH_3 and H.

leading, after proton loss from C(3), to the isomeric ketone **5**, as summarized in Figure 1.

Although it may seem contradictory that this biosynthetic pathway occurs in the opposite direction to the energetically downhill (i.e., exergonic) transformation of reduced **5** to **6**, overall a biosynthetic pathway from β-amyrin could also be exergonic, since it is accompanied by the loss of a C−C π-bond (ca. 65 kcal/mol π-strength) plus an RO-H σ-bond (ca. 105 kcal/mol) (total of 170 kcal/mol) and the gain of a C═O π-bond (ca. 90 kcal/mol π-strength) plus a C−H σ-bond (ca. 95 kcal/mol); (total of 185 kcal/mol). Thus, changes in the bonding energies could provide an overall driving force for the conversion of β-amyrin **3** to **5** of about 15 kcal/mol and make such a biosynthetic pathway energetically favorable, even in the face of an increase in eclipsing and steric strains for **5** (which are

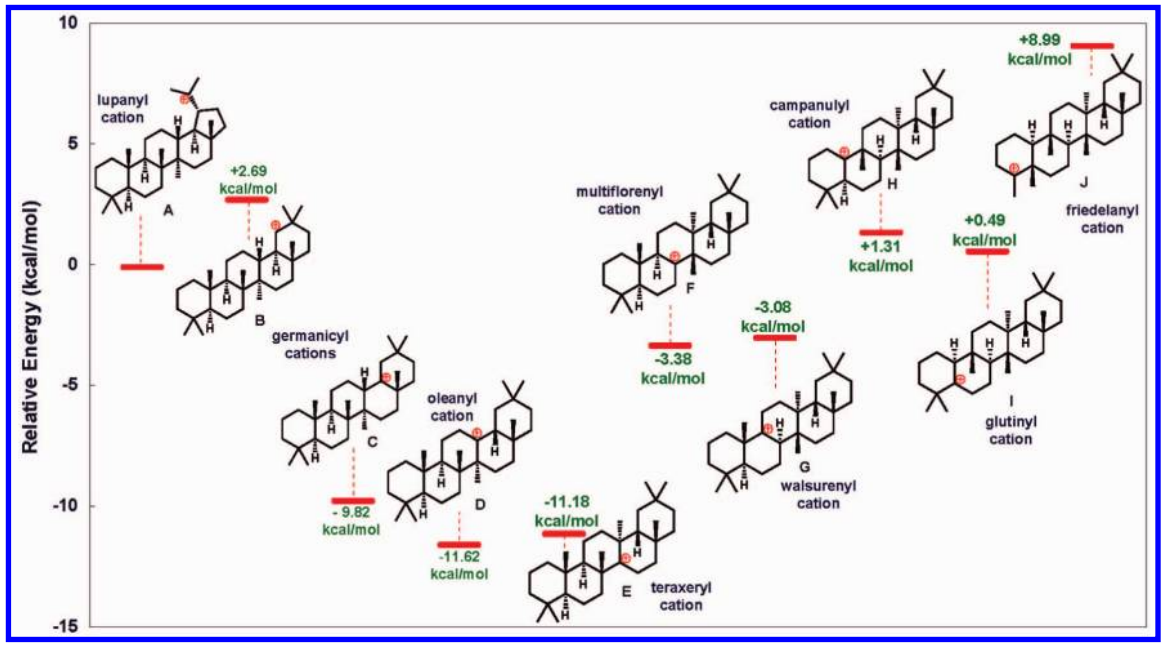


Figure 2. Equilibrium geometry calculations for triterpene cations using DFT/6−31G* method.

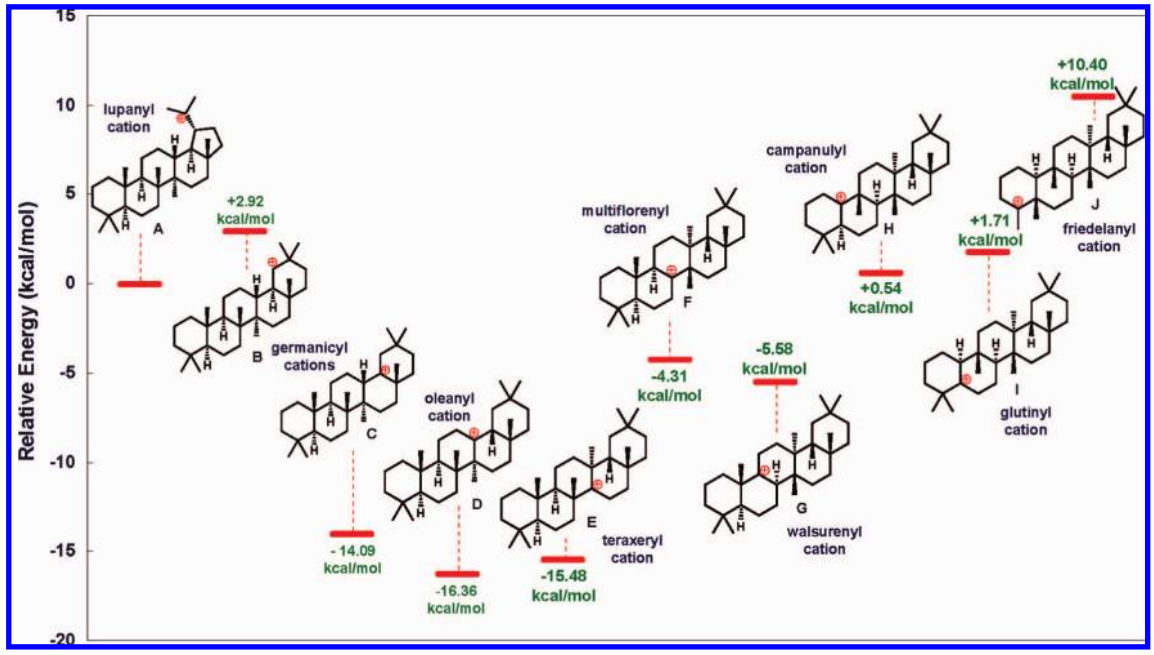


Figure 3. Hartree−Fock/6−31G* single-point energy calculations for triterpene cations.

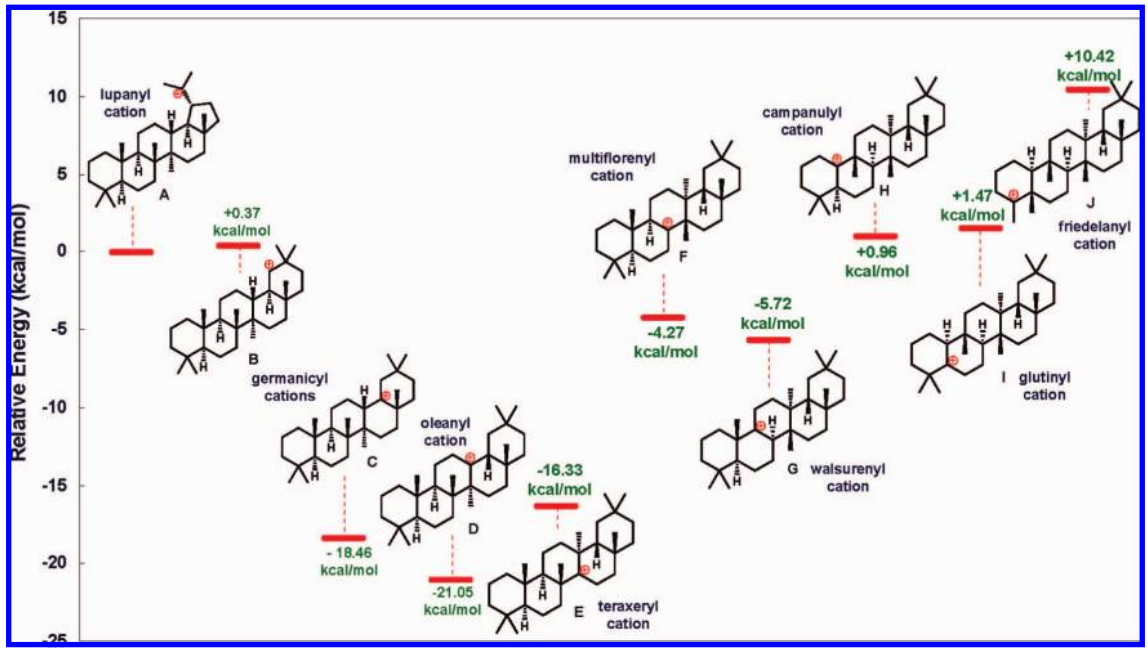


Figure 4. Hartree−Fock/3−21G(*) single-point energy calculations for triterpene cations.

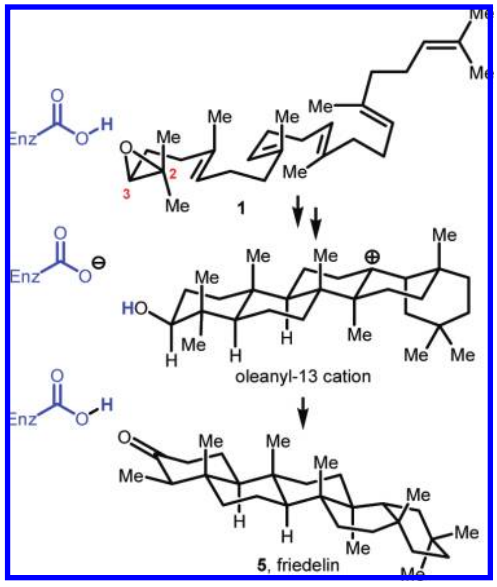


Figure 5. Proposal for the nonstop conversion of 2,3-(*S*)-oxidosqualene to friedelin.

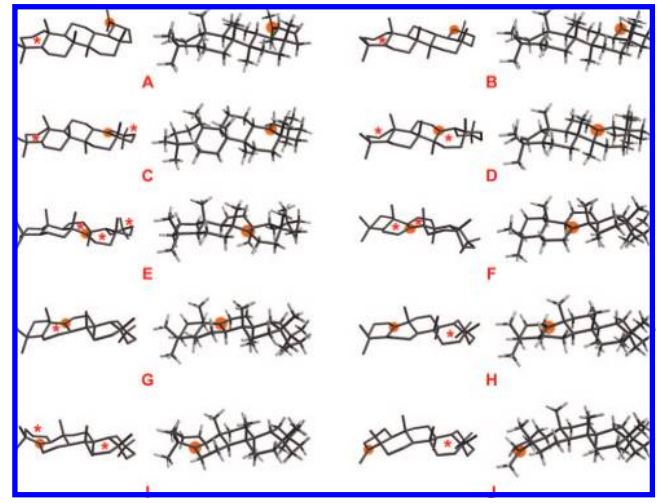


Figure 6. Conformations of cations **A**−**J**. The equilibrium geometries were obtained by using DFT/6−31G* method. The carbocations in the tube structures are highlighted with an orange filled circle. Nonchair ring conformations in these intermediate structures are indicated by an asterisk placed in the center (or above) of the nonchair ring.

not insignificant since the D ring has a twist boat form). Nonetheless, understanding just how friedelin synthase catalyzes the energetically uphill rearrangements of carbon and hydrogen still remains a problem.

We have examined the conformational strain in the carbocations which are intermediates along the pathway connecting the lupanyl cation with the friedelanyl cation using both DFT and Hartree−Fock *ab initio* computational methodology.[6] In addition, we have performed calculations in which C^+ in each cation is replaced by an isoelectronic boron atom. The results of the DFT calculations on the cations using a 6−31G* basis set are shown in Figure 2 for the equilibrium geometry. The values shown differ quantitatively, but not qualitatively from Hartree−Fock calculations using the 6−31G* or the smaller 3−21G(*) basis set (see Figures 3 and 4). The qualitatively similar results from the three calculations lend confidence in the significance of the relative energies.

The data shown in Figure 2 indicate that the rearrangement of the lupanyl cation **A** to the germanicyl cation **B** and further to the oleanyl cation **D** is driven by relief of conformational strain amounting to about 10 kcal/mol. Thus, the biosynthesis of β-amyrin (**3**) from the lupanyl cation should be thermodynamically favorable, requiring only that a proton acceptor be positioned on β-amyrin synthase at a location proximate to a hydrogen at carbon-12.

On the other hand, the backbone rearrangement of the oleanyl cation **D** to the friedelanyl cation **J** is thermodynamically unfavorable by about 20 kcal/mol. This raises and important question: how does friedelin synthase overcome the intervening energy barriers, which are huge compared to that for proton loss from a tertiary carbocation, and force the rearrangement

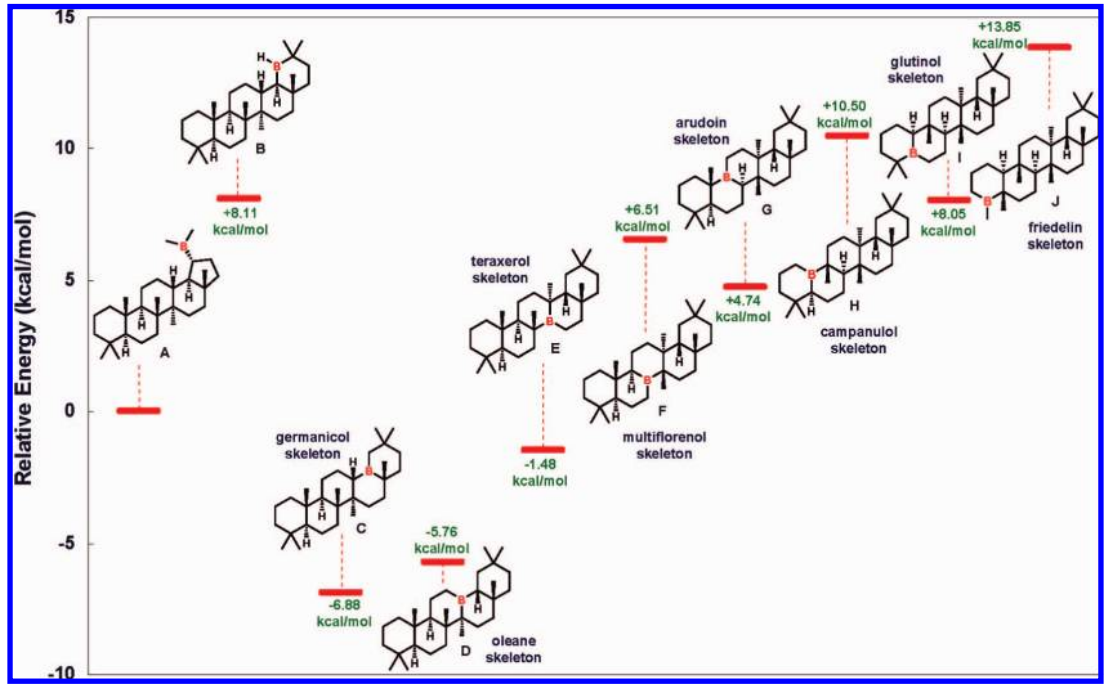


Figure 7. Hartree−Fock/3−21G(*) single-point energy calculations for isoelectronic borane analogs of triterpene cations.

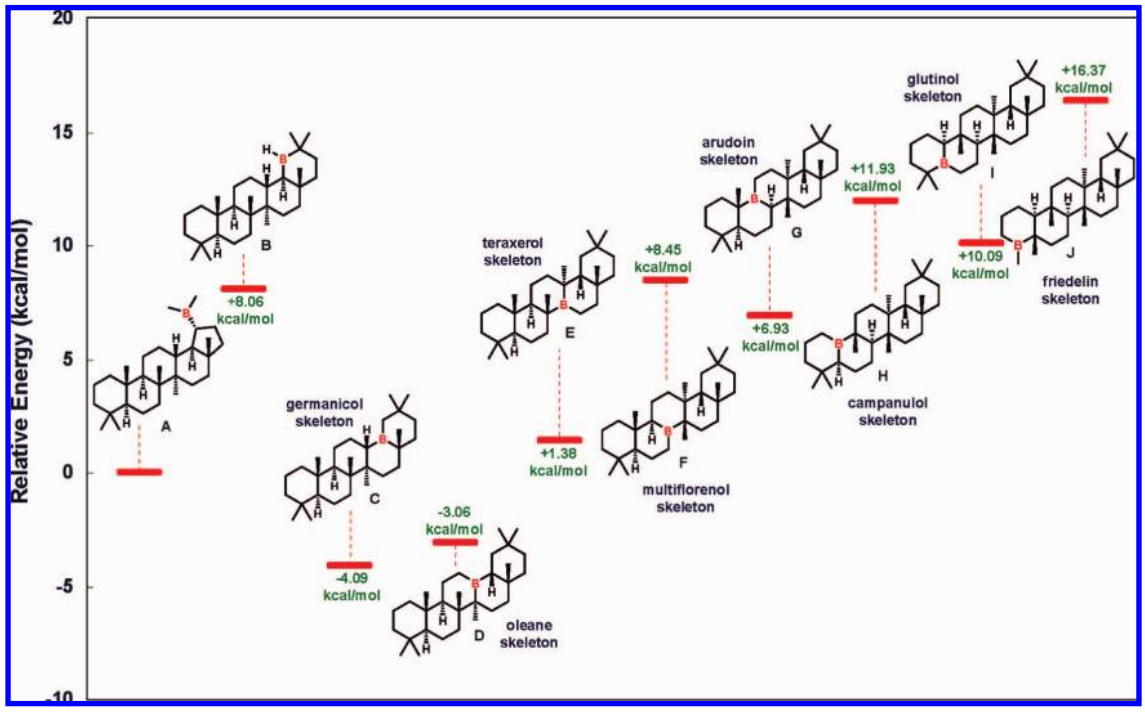


Figure 8. Hartree−Fock/6−31G* single-point energy calculations for isoelectronic borane analogs of triterpene cations.

of the cation **D** to the friedelanyl cation **J**? In this connection it is relevant that the overall pentacyclization process is very exergonic to form cation **A**. A simple calculation of bond energy changes for the 2,3-(*S*)-oxidosqualene (**1**) → lupanyl cation cyclization reveals that this change could be exergonic by as much as 50 kcal/mol.[7] This figure must be reduced by the energy required to separate charges during the cyclization. In the biosynthesis of lanosterol from 2,3-(*S*)-oxidosqualene, the cyclization is initiated by the transfer of a proton from aspartic acid COOH (D456 in the lanosterol synthase of *S. cerevisiae*) to the oxirane oxygen.[2,8] If friedelin biosynthesis is initiated in the same way, the cyclization occurs as shown in Figure 5.

During this exergonic cyclization of **1** into the oleanyl cation - aspartate ion pair, positive and negative charges are separated. In addition the conformation of the enzyme might change to one of higher energy. Electrostatic calculations reveal that the electrostatic energy of the separated charges could be on the order of 18 kcal/mol (assuming 9 Å separation distance and a dielectric constant of 2 for intervening triterpene scaffold). This electrostatic factor would reduce the exothermicity of the

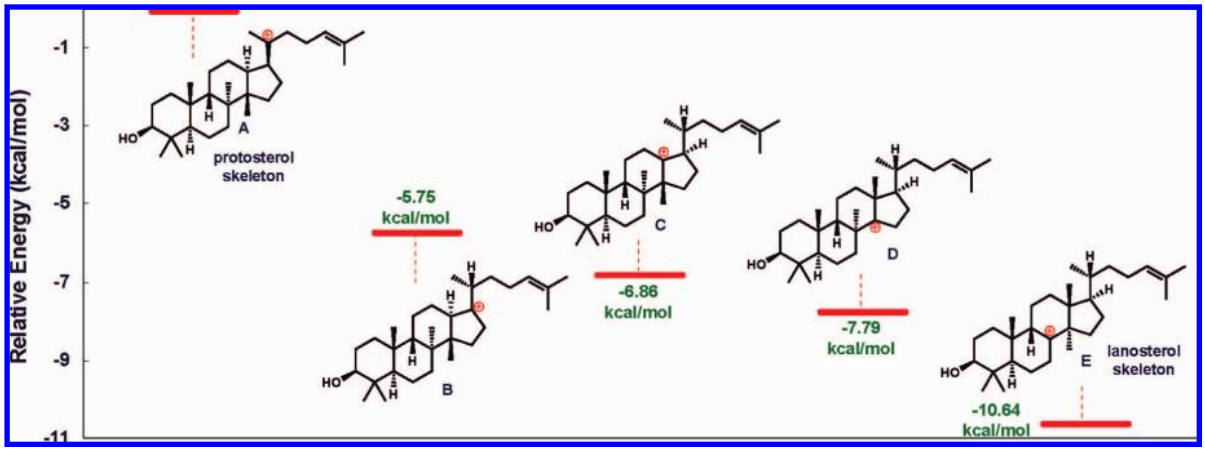


Figure 9. Equilibrium geometry calculations for cationic intermediates in lanosterol biosynthesis using DFT/6−31G* method.

transformation of **1** to the initial pentacyclic cation to ca. 30 kcal/mol. Forcing the conformation of the enzyme into one of higher energy as a result of pentacycle formation would reduce this energy still further.

Even though the exothermicity of cyclization might be reduced somewhat by induced conformational strain in the enzyme itself, the overall cyclization process is still likely to be exothermic. The electrostatic attraction between the anion, Enz-COO^- in Figure 5, and the triterpene carbocation plus any conformational strain in the enzyme might well be sufficient to guide the backbone rearrangement toward the friedelin skeleton over the energy landscape pictured in Figure 2. As the rearrangement of the oleanyl cation to the friedelanyl cation proceeds, the force between the charges increases (as the inverse square of the distance separating them) and the electrostatic driving force intensifies.[9]

These considerations lead us to propose the following hypothesis: friedelin may be biosynthesized from 2,3-(*S*)-oxidosqualene *nonstop by a single enzyme* in a sequence involving cationic pentacyclization *and* subsequent multistep rearrangement of carbon and hydrogen via carbocationic intermediates. Basically, the process is driven by the strong overall exergonic nature of the transformation **1** → **5**. We also offer the additional hypothesis that friedelin synthase produces friedelin rather than the other known pentacyclic triterpenes, because the catalytic site lacks a properly positioned proton acceptor to divert any of the carbocation intermediates to the other triterpenes (e.g., lupeol or β-amyrin).

This argument then leads to three other hypotheses. First, each of the pentacyclic triterpenes, lupeol, germanicol, α-amyrin, β-amyrin, teraxerol, multiflorenol, glutinol, etc., may be produced directly from **1** without passing through stabilized intermediates (i.e., nonstop cyclization and rearrangement) by a single protein, with the structure of the product being controlled by the location of the proton acceptor on the enzyme.[10] Second, isomerizing interconversions within the plant of the type lupeol → β-amyrin, β-amyrin → multiflorenol, or β-amyrin → friedelin are unlikely to occur; that is, once produced, enzymic products such as β-amyrin are not further rearranged. Finally, there could be high homology between the different pentacyclic triterpene synthases, and they could belong on a common phylogenetic tree with a common ancestor.

The conformations of cations **A**−**J**,that were obtained by conducting DFT 6−31G* equilibrium geometry calculations, are shown in Figure 6. Nonchair ring conformations in these intermediate structures are indicated by an asterisk placed in the center (or above) of the nonchair ring.

We have also carried out energy calculations for the boranes which are isoelectronic with the triterpene cations **A**−**J** of Figures 2, 3, and 4 using Hartree−Fock methodology at the 6−31G* and 3−21G(*) levels with the results summarized in Figures 7 and 8. Several interesting points emerge. First, the topology of the energy landscape is quite similar for the carbocations and their borane analogs. Second, the HF/3−21G(*) and HF/6−31G* calculations resulted in qualitatively similar relative energies for the borane series **A**−**J**. Third, just as the tertiary lupanyl cation **A** is energetically more stable than the secondary germanicyl cation **B**, the borane analog of the lupanyl cation is more stable than that of the germanicyl cation.

The biosynthesis of lanosterol and cholesterol involves the cyclization of **1** to the protosterol cation which then undergoes backbone rearrangement of H and CH_3 to form a C(8) cation with the lanosterol skeleton.[2,11] We have also modeled this rearrangement of the various cations by DFT calculations with a 6−31G* basis set, with results shown in Figure 9. In contrast to the biosynthesis of the friedelin skeleton from the oleane skeleton, it is evident that the backbone rear-

(6) (a) Kong, J.; et al. *J. Comput. Chem.* **2000**, *21*, 1532–1548. (b) Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*; Wavefunction, Inc.: Irvine, CA, 2003.

(7) Bonds broken in **1**: Enz-COOH, C(2)-O, 5 π(C═C) for a total of 460 kcal/mol; bonds made in the lupanyl cation: C(3)−O−H, 5 s(C−C) for a total of 510 kcal/mol; difference 50 kcal/mol, exergonic. For a similar result on the endothermicity of cyclization, see: Matsuda, S. P. T.; Wilson, W. K.; Xiong, Q. *Org. Biomol. Chem.* **2006**, *4*, 530–543.

(8) (a) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1277–1288. (b) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1289–1296.

(9) The thermodynamically endergonic rearrangement of β-amyrin to teraxerol and on to multiflorenol, which cannot be effected by acid catalysis, has been realized chemically, but only by oxidatively driven rearrangement; see: Agata, I.; Corey, E. J.; Hortmann, A. G.; Klein, J.; Proskow, S.; Ursprung, J. J. *J. Org. Chem.* **1965**, *30*, 1698–1710.

(10) Perfect product control may not be attained by some of the triterpene synthases, perhaps not surprisingly given the high reactivity of the intermediate cations. The Matsuda group have shown using high sensitivity GC-MS analysis that the baruol synthase of Arabidopsis thaliana makes 22 minor products (0.02−3% each) in addition to baruol (90%); see: Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Kolesnikova, M. D.; Onak, C. S.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **2007**, *129*, 11213–11222.

rangements involved in lanosterol biosynthesis are exergonic. The finding that a chemically generated protosterol cation (**A** in Figure 9) rearranges within a few minutes in CH_2Cl_2 solution at −95 °C to the lanosterol skeleton[11b] is totally consistent with the energy profile summarized in Figure 9. Given this energy profile, it is not surprising that there are only a few known natural products with the protosterol skeleton.

In summary, the results described here provide a new and, we hope, useful view of the fascinating rearrangements that lead to the biosynthesis of many polycyclic triterpenoids and catalytic requirements on the enzymes that direct them. We close by noting that there have been a number of other publications dealing with computational aspects of cation-olefin cyclizations related to triterpene/sterol biosynthesis.[12]

Acknowledgment. L.K. is a Fellow of the Damon Runyon Cancer Research Foundation (DRG: 1961-07). R.-J. C. is a recipient of a Taiwan Merit Scholarship.

Supporting Information Available: Tabulated energies and coordinates obtained by B3LYP/6−31G* DFT and *ab initio* calculations for the structures described herein. This material is available free of charge via the Internet at http://pubs.acs.org.

JA800980H

(11) (a) Corey, E. J.; Virgil, S. C. *J. Am. Chem. Soc.* **1990**, *112*, 6429–6431. (b) Corey, E. J.; Virgil, S. C. *J. Am. Chem. Soc.* **1991**, *113*, 4025–4026.

(12) (a) Jenson, C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1997**, *119*, 10846–10854. (b) Gao, D.; Pan, Y.-K.; Byun, K.; Gao, J. *J. Am. Chem. Soc.* **1998**, *120*, 4045–4046. (c) Rajamani, R.; Gao, J. *J. Am. Chem. Soc.* **2003**, *125*, 12768–12781. (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. *Eur. J. Org. Chem.* **2004**, *223*, 9–2242. (g) Hess, B. A., Jr.; Smentek, L. *Org. Lett.* **2004**, *6*, 1717–1720.