Reaction Mechanisms

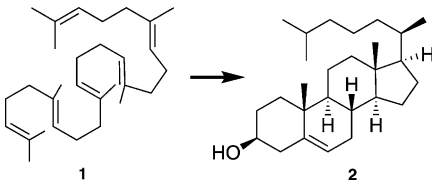
DOI: 10.1002/ange.201302886

The Concerted Nature of the Cyclization of Squalene Oxide to the Protosterol Cation

B. Andes Hess Jr. and Lidia Smentek*

Dedicated to Professor Rudolf Zahradník on the occasion of his 85th birthday

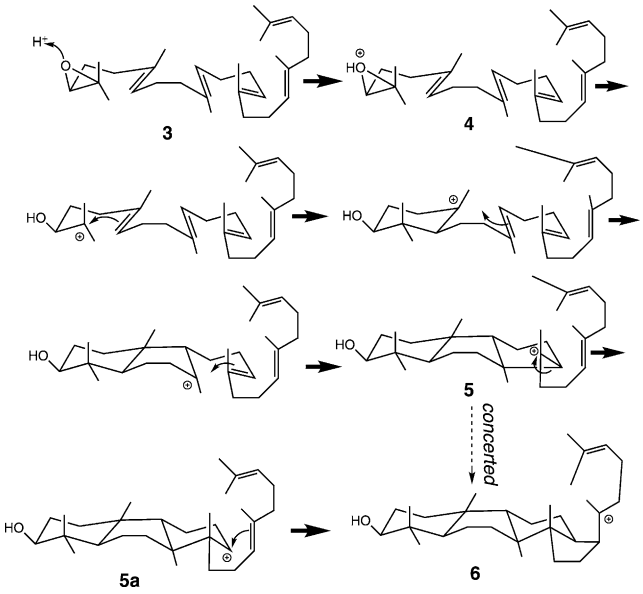
In 1926 Heilbron first suggested that squalene (**1**) was involved in the metabolic process that produces cholesterol (**2**) in animals (Scheme 1).[1]



Scheme 1. Conversion of squalene to cholesterol.

In 1952 this was confirmed by Langdon and Bloch when they reported that feeding mice isotopically labeled squalene oxide gave rise to labeled cholesterol.[2] This biochemical conversion of an acyclic hydrocarbon, squalene, to cholesterol that contains four rings and eight chiral centers remains one of the most astounding biological reactions known. In 1953 Woodward and Bloch showed how squalene could fold to give the basic steroidal ring system.[3] In 1966 Corey[4] and van Tamelen[5] independently reported that squalene is first epoxidized with the cyclization initiated by protonation of the epoxide (**3**). Since then the primary focus has been on the cyclizations, with proposals that range from a fully concerted formation of the four rings to a stepwise mechanism in which each ring is formed to give a distinct carbocation intermediate.[6]

In Scheme 2 the overall cyclization after protonation of the epoxide is broken down into steps, which may or may not be concerted. Until 2002, it was generally agreed that the formation of the C ring involves first a five-membered C ring (**5**), which with an anti-Markovnikov ring expansion gives the six-membered C ring (**5a**) and subsequent cyclization of the five-membered D ring yields the protosterol cation (**6**). We showed that there is an alternate explanation for this cyclization that avoids the anti-Markovnikov ring expansion, namely a concerted ring expansion of the initially formed five-membered C ring and formation of the D ring (**5**→**6**).[7] Subsequently Tantillo has found a number of cases in terpene



Scheme 2. Cyclization of squalene oxide (**3**) to the protosterol cation (**6**).

biosynthesis where secondary carbocations are similarly avoided.[8] We have also studied various steps of the cyclization of the A–C rings from both a stereochemical point of view as well as concerted vs. stepwise cyclizations.[9]

In 2004 Schulz reported the X-ray structure of 2-azasqualene encapsulated in hopene cyclase, in which the azasqualene appears to be "held" by the enzyme in a conformation that would allow at least the first three rings to cyclize with the correct stereochemistry on the pathway to the formation of the squalenes.[10] We were able to use his results to locate a transition structure that links the initially protonated squalene to the tricyclic, tertiary carbocation intermediate in a concerted, but highly asynchronous, reaction.[11] These results strongly supported the concerted nature of the formation of these three rings.

The cyclization of squalene oxide to lanosterol differs in one key aspect: In the case of hopene cyclization, the B ring is formed in a chair conformation, while in the lanosterol cyclization the B ring is formed in a boat conformation. The question that will be addressed here is: Does a conformation of oxidosqualene exist that on protonation and ring-opening would allow the concerted formation of the A–C rings with the B ring formed in the less-stable boat conformation as well as the proper stereochemistry of all the formed rings? For this

[*] Prof. Dr. B. A. Hess Jr., Prof. Dr. L. Smentek
Department of Chemistry, Vanderbilt University
Nashville, TN 37235 (USA)
E-mail: b.andes.hess@vanderbilt.edu

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201302886.

cyclization to be concerted (6-6-5, chair–boat formation), a conformation of the squalene oxide must exist that once protonated has the electrophilic C–C double bonds positioned to interact with the developing positively charged carbon atoms. Our earlier studies on the individual cyclizations have shown that for a concerted cyclizations to occur, the conformation about the key C–C bond must not exist as a minimum if a positive charge is developed in the vicinity of the achimerically assisting double bond that supplies the electrons for the newly forming C–C σ bond.[9c,d]

A search[12] for a transition structure[13] that would involve rings A and B being formed (chair–boat) in a concerted fashion using BB1K/6-31G* was carried out on a truncated, protonated squalene. The epoxide, which would open under acid catalysis to give a carbocation, was not included because of problems that arose from the presence of the conjugate base of the acid in the computations (see below for a further discussion of this point). The search yielded transition structure **7** (Figure 1). An intrinsic reaction coordinate (IRC) calculation on **7** showed (Figure 2) that it links

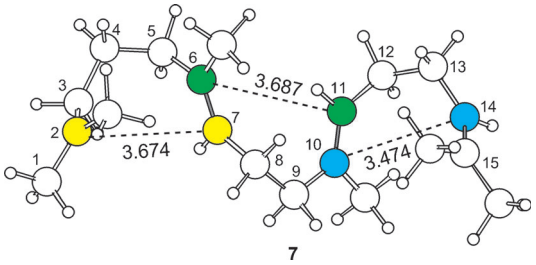


Figure 1. A truncated, protonated squalene transition structure. The two yellow carbon atoms form the chair A ring, the two green carbon atoms the boat B ring, and the two blue carbon atoms form the five-membered C ring. Distances in Ångstroms.

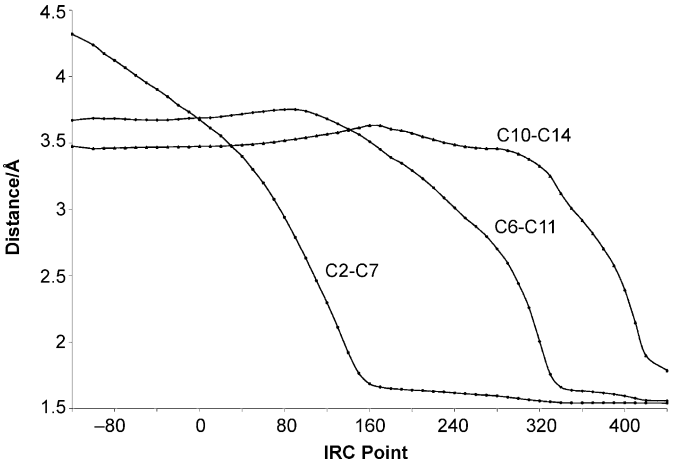


Figure 3. IRC of the conversion of **8** to **9** via transition structure **7**. The change in bond distances of the three ring-forming bonds is plotted against the IRC point in the concerted formation of the A–C rings. Distances are in Ångstroms.

a reactant conformer (**8**) and the tricyclic product **9**, which confirms that conformer **8** can undergo a reaction in which all three rings (A–C) are formed in an asynchronous, concerted fashion. This is easily seen in Figure 3, since rings B and C do not begin to significantly form until the previous ring is almost completely closed. The asynchronicity can be further seen by the structures of points 180 and 335 (Figure 4) in which the A and B rings, respectively, are essentially formed.

During the course of this concerted reaction 47 kcal mol^{-1} of energy is released, but as seen from Figure 5 the amount released with the formation of each ring differs. The

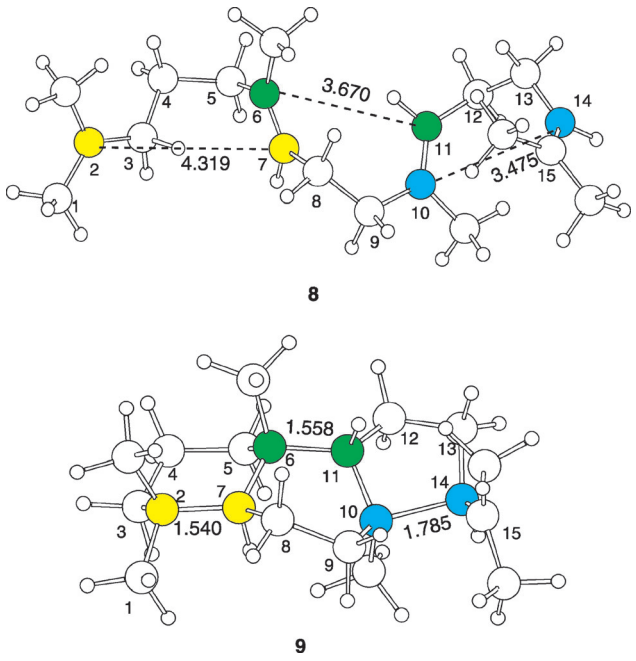


Figure 2. Reactant (**8**) and product (**9**) linked by transition structure **7** as shown by IRC calculations. Distances in Ångstroms.

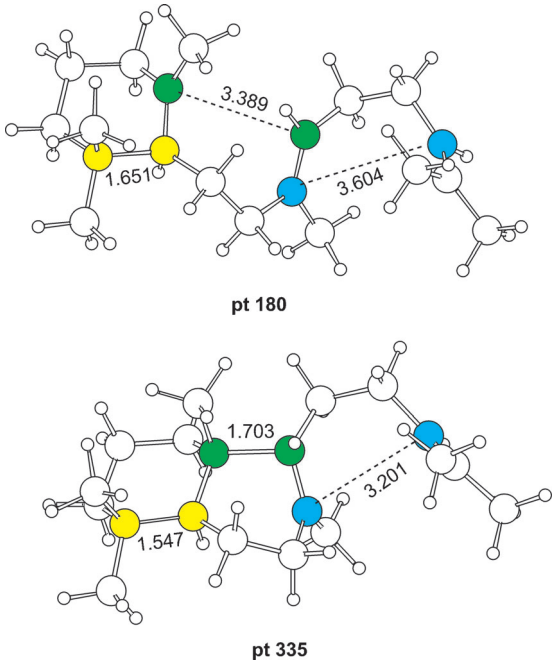


Figure 4. Structures of IRC points 180 and 335. Distances in Ångstroms.

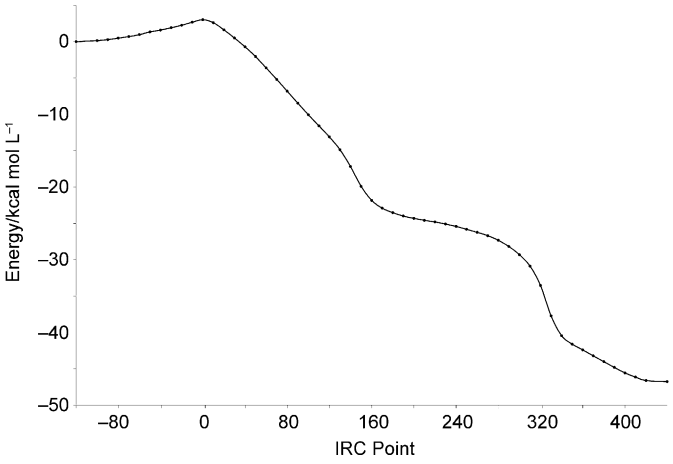
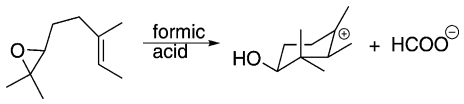


Figure 5. The energy of each IRC point relative to the energy of reactant **8**. The transition structure is at point 0.

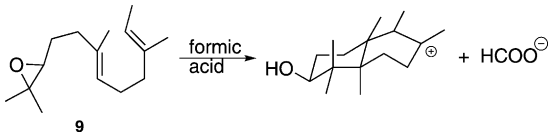
formation of ring A releases about 20 kcal mol^{-1} of energy in contrast to rings B and C, which combined release about 27 kcal mol^{-1} of energy. This difference is presumably due to ring B being formed in a boat conformation, which releases less energy than when ring B is formed in a chair conformation. With zero point energy corrections the overall release of 39 kcal mol^{-1} for the ABC concerted ring formation compares with 48 kcal mol^{-1} for that of the conversion of squalene to the hopenes.[14] Again the difference arises primarily from the boat versus the chair formed in ring B.

While the above results support the concerted ring closure of the A–C rings, the question remains whether or not the ring opening of the epoxide on protonation is also part of an overall concerted reaction of the formation of the A–C rings from squalene oxide. In 1998 Pan and Gao found that protonation of a model system of squalene oxide with formate did indeed proceed in concert with the ring opening of the epoxide ring and formation of the A ring of squalene oxide (Scheme 3).[15] This model reaction was repeated with BB1K/



Scheme 3. Gao and Pan's concerted epoxide and ring A formation model reaction.

6-31G* and additional carbons to confirm whether or not the epoxidation and formation of rings A and B would also be concerted (Scheme 4). A transition structure was located for this reaction.



Scheme 4. Model system expanded to include carbon atoms that would form both the A and B rings on acid catalysis of the epoxide.

However, while the epoxide did undergo acid-catalyzed, concerted ring opening and ring closure of ring A, the IRC failed before ring B was fully closed owing to a very flat region on the potential energy surface. This was likely due to the fact that once ring A had closed, the formate ion began to migrate toward the positive charge that had been formed on closure of ring A, which negated the necessity of ring B to close. This phenomenon was also observed by Pan and Gao.[15] Presumably in the cavity of the enzyme the conjugate base, which is formed on protonation of the epoxide, would not be able to "migrate" to the positive charge in ring A. We therefore conclude that for the protosterol cyclization shown in Scheme 2, all steps through the formation of ring C may indeed take place in an asynchronous, concerted reaction, without the formation of intermediate carbocations as minima. What is required for this pathway to be followed by squalene oxide is that the enzyme must "cradle" squalene oxide in a conformation similar to the one we have located. Unlike in the case of the squalene to hopene cyclization, there are no experimental data available that would show that such a conformation of squalene oxide exists within the enzyme. We do note though that the reacting conformer (**10**, Figure 6) found previously[11] for the concerted cyclization of

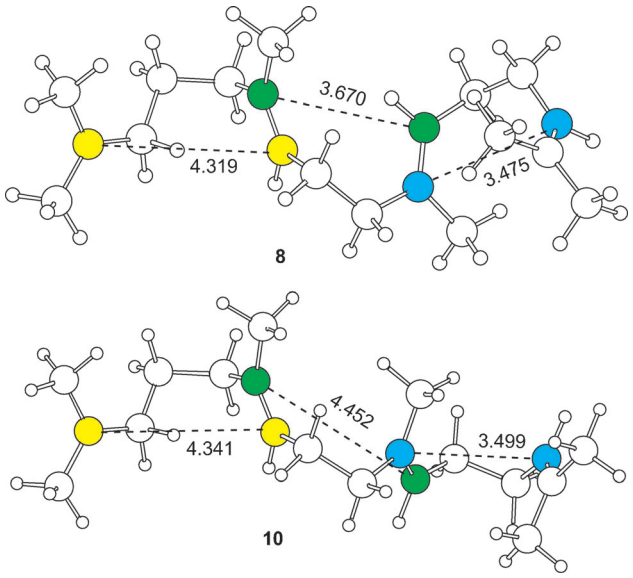


Figure 6. Model reactants for the concerted prostosterol (**8**) and hopene (**10**) cyclizations obtained with BB1K/6-31G* computations. Distances in Ångstroms.

rings A–C in squalene on the hopene pathway has some resemblance to the conformer of the reactant for the squalene oxide cyclization (**8**, Figure 6).

In particular it is seen that the conformations about all C–C bonds from C1 to C9 are the same. The first difference occurs about the C9–C10 bond, which of course must be the case, since the ring B is formed in a chair conformation in hopene and a boat in protosterol.

These results presented so far have no bearing on whether the next ring (the five-membered D ring) is formed in concert

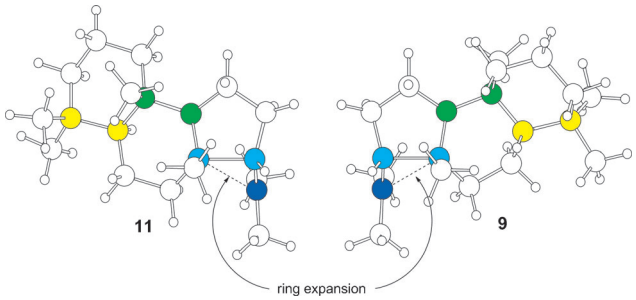


Figure 7. C_{20} model tricyclic products for the concerted hopene (**11**) and prostosterol (**9**) A–C cyclizations obtained with BB1K/6-31G* computations. The dark blue atoms on ring expansion of the C ring would become part of the six-membered C ring.

with rings A–C or tricyclic **9** exists as a distinct carbocation intermediate. However, as shown in Figure 7, the C_{20} model products of the hopene cyclization (**11**) and the protosterol cyclization (**9**) have almost identical C rings (5-membered). In both cases in the enzyme-catalyzed cyclizations the next step would be the ring expansion of ring C and a fourth cyclization to yield the 5-membered D ring.

Since we found in the case of the hopene cyclization that with the C_{25} model (this would have allowed D-ring formation), its cyclization stopped at the five-membered C ring;[11] we conclude that this will likely also be the case in the protosterol cyclization. That is for the conversion of the squalene oxide to the protosterol carbocation we propose that the A–C rings are formed in concert, along with the ring-opening of the protonated epoxide, with a pause at the intermediate carbocation **9**. In both cases in the next step ring C would expand to a six-membered ring concomitantly with the formation of the six-membered D ring. We have computed both the energies of activation and reaction for the hopene and lanosterol concerted pathways of the C_{25} models (see Figure 8) with (BB1K/6-31G*).[16] The activation and reaction energy for this conversion was calculated to be 7.2 and -10.1 kcal mol^{-1} for the hopene pathway and 5.5 and -6.3 kcal mol^{-1} for the lanosterol pathway (B ring with a boat conformation). While these activation energies are not large, they are not insignificant. This may in part account for these

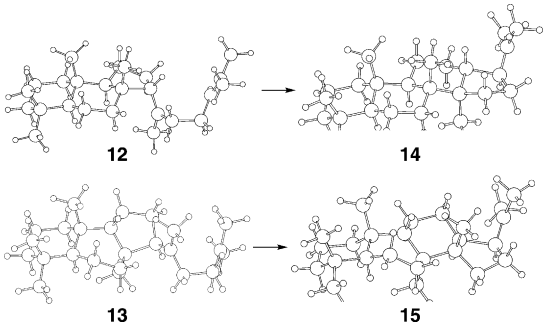


Figure 8. BB1K/6-31G* reactants and products for the C_{25} model system for the conversion of the 5-membered C ring to the 6-membered C ring concomitant with formation of the 5-membered D ring for the hopene and lanosterol pathways (**12**→**14** and **13**→**15**, respectively).

reactions "stopping" or "pausing" at the 6-6-5 tricycle. On the other hand given that there is a significant amount of energy released in the formation of the tricycle within the cavity of the enzyme, it is likely that these intermediate carbocations do not have a significant lifetime and proceed quickly through the final C-ring expansion and D-ring formation. The overall release of energy for the formation of the A–D rings in protosterol is calculated to be 46 kcal mol^{-1} and for the hopene pathway 58 kcal mol^{-1}, which are in reasonable agreement with Matsuda's computed values (*m*PW1PW91/6-311 + G(2d,p)) of 44 kcal for the conversion of protonated squalene oxide to the protosterol carbocation and 54 kcal mol^{-1} for the formation of the A–D rings in the hopene pathway (with a chair conformation of the B ring).[16] Our calculations do not take into account the ring opening of the epoxide; however, BB1K/6-31G* calculations on a model protonated epoxide and its conversion to a ring opened tertiary carbocation showed the process to be slightly endothermic (+0.8 kcal mol^{-1}). Hence our computed energies of the overall processes described here are reasonable estimates of the overall energetics of these two cyclizations.

The X-ray structure of oxidosqualene cyclase (OSC) has been reported,[17] but unfortunately unlike in the case of squalene hopene cyclase no crystal structure has been obtained with an encapsulated, "deactivated" substrate. Hence at first sight our results presented here appear to be not quite as compelling as those we reported for the cyclization of squalene,[11] where the X-ray structure of the azasqualene encapsulated in squalene hopene cyclase was very similar in structure to the reactant conformer of protonated squalene, which was shown to proceed with the concerted formation of the A–C rings. However, we have previously carried out detailed conformational studies with models of the ring closures of both the B and C rings,[9a,13] in which it was shown that the conformations required for the formation of both the B and C rings do not exist when a positive charge exists in the A and B rings respectively; these conformations simply collapse to give the B and C rings, respectively. Hence, since the conformer that we have located (**8**) is the one required to form the B ring as a boat conformer as well as yield the proper stereochemistry of all three rings, this is indeed compelling evidence that the protonated epoxide proceeds with the concerted ring-opening of the epoxide and formation of rings A–C. As a consequence it seems unlikely that the monocyclic and bicyclic intermediate carbocations need to be invoked as a number recent reports do for the mechanism of the enzymatic cyclization of squalene oxide as well as for squalene.[18]

Received: April 8, 2013
Revised: July 8, 2013
Published online: September 3, 2013

Keywords: carbocations · density functional calculations · reaction mechanisms · squalene · terpenoids

[1] I. M. Heilbron, E. D. Kamm, W. M. Owens, *J. Chem. Soc.* **1926**, 1630–1644.

[2] R. G. Langdon, K. Bloch, *J. Am. Chem. Soc.* **1952**, *74*, 1869 – 1870.

[3] R. B. Woodward, K. Bloch, *J. Am. Chem. Soc.* **1953**, *75*, 2023 – 2024.

[4] E. J. Corey, W. E. Russey, P. R. Ortiz de Montellano, *J. Am. Chem. Soc.* **1966**, *88*, 4750 – 4751.

[5] E. E. van Tamelen, J. D. Willett, R. B. Clayton, *J. Am. Chem. Soc.* **1967**, *89*, 3371 – 3373.

[6] For reviews, see: a) D. U. Cane, *Chem. Rev.* **1990**, *90*, 1089 – 1103; b) 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; c) I. Abe, M. Rohmerk, G. D. Prestwich, *Chem. Rev.* **1993**, *93*, 2189 – 2206; d) J. D. Connolly, R. A. Hill, *Nat. Prod. Rep.* **2001**, *18*, 560 – 578; e) R. F. Xu, G. C. Fazio, S. P. T. Matsuda, *Phytochemistry* **2004**, *65*, 261 – 291; f) R. A. Yoder, J. N. Johnston, *Chem. Rev.* **2005**, *105*, 4730 – 4756; g) D. W. Christianson, *Chem. Rev.* **2006**, *106*, 3412 – 3442; h) I. Abe, *Nat. Prod. Rep.* **2007**, *24*, 1311 – 1331; i) J. Degenhardt, T. G. Köllner, J. Ger *s*henzon, *Phytochemistry* **2009**, 70, 1621 – 1637; j) D. J. Tantillo, *Nat. Prod. Rep.* **2011**, *28*, 1035 – 1053.

[7] B. A. Hess, Jr., *J. Am. Chem. Soc.* **2002**, *124*, 10286 – 10287.

[8] a) Y. J. Hong, D. J. Tantillo, *Org. Lett.* **2006**, *8*, 4601 – 4604; b) M. W. Lodewyk, P. Gutta, D. J. Tantillo, *J. Org. Chem.* **2008**, *73*, 6570 – 6579; c) Y. J. Hong, D. J. Tantillo, *J. Am. Chem. Soc.* **2009**, *131*, 7999 – 8015; d) Y. J. Hong, D. J. Tantillo, *J. Am. Chem. Soc.* **2010**, *132*, 5375 – 5386.

[9] a) B. A. Hess, Jr., *Org. Lett.* **2003**, *5*, 165 – 167; b) B. A. Hess, Jr., *Collect. Czech. Chem. Commun.* **2003**, *68*, 202 – 210; c) B. A. Hess, Jr., *Eur. J. Org. Chem.* **2004**, 2239 – 2242; d) B. A. Hess, Jr., L. Smentek, *Mol. Phys.* **2004**, *102*, 1201 – 1206.

[10] D. J. Reinert, G. Balliano, G. E. Schulz, *Chem. Biol.* **2004**, *11*, 121 – 126.

[11] L. Smentek, B. A. Hess, Jr., *J. Am. Chem. Soc.* **2010**, *132*, 17111 – 17117.

[12] a) All calculations were performed with Gaussian 09 (Revision A.02), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, 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, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, **2009**; b) Geometries were optimized without symmetry constraints using BB1K/6-31G* (M. Weitman, D. T. Major, *J. Am. Chem. Soc.* **2010**, *132*, 6349 – 6360; Y. Zhao, B. J. Lynch, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 2715 – 2719; C. Lee, W. Yang, R. Parr, *Phys. Rev. B* **1988**, *37*, 785 – 789; P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213 – 222). All structures were characterized by frequency calculations, and reported energies include zero-point energy corrections (unscaled). Intrinsic reaction coordinate (IRC) calculations (C. Gonzalez, H. B. Schlegel, *J. Phys. Chem.* **1990**, *94*, 5523 – 5527; K. Fukui, *Acc. Chem. Res.* **1981**, *14*, 363 – 368) were also used to verify the identity of all transition structures.

[13] The initial guess for this transition structure made use of the results of our earlier calculations on a model system of the part of squalene that contained only those carbons necessary for the A(chair)–B(boat)ring closures (see Ref. [11]). The conformation of the remaining carbons that would form the C ring was chosen to give the proper stereochemistry of the C rings in the product.

[14] BB1K/6-31G* calculations were carried out for the previous B3LYP/6-31G study of the concerted ring closure of the ABC rings in the cyclization of squalene to the hopenes (Ref. [11]), the results of which are given in the Supporting Information.

[15] D. Gao, Y.-K. Pan, K. Byun, J. Gao, *J. Am. Chem. Soc.* **1998**, *120*, 4045 – 4046.

[16] Our earlier calculation on this C-ring expansion/D-ring formation was carried out with B3LYP/6-31G*, which both Matsuda (S. P. T. Matsuda, W. K. Wilson, Q. Xiong, *Org. Biomol. Chem.* **2006**, *4*, 530 – 543) and Major (M. Weitman, D. T. Major, *J. Am. Chem. Soc.* **2010**, *132*, 6349 – 6360) have noted does not give reliable energies for cyclizations of hydrocarbons. Major has found that the BB1K method on the other hand does.

[17] 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.

[18] a) T.-K. Wu, T.-T. Wang, C.-H. Chang, Y.-T. Liu, W.-S. Shie, *Org. Lett.* **2008**, *10*, 4959 – 4962; b) T.-K. Wu, C.-H. Chang, Y.-T. Liu, T.-T. Wang, *Chem. Rec.* **2008**, *8*, 302 – 325; c) T.-K. Wu, Y.-C. Chang, Y.-T. Liu, C-.H. Chang, H.-Y. Wen, W.-H. Li, W.-S. Shie, *Org. Biomol. Chem.* **2011**, *9*, 1092 – 1097; d) B. Tian, L. Eriksson, *J. Phys. Chem. B* **2012**, *116*, 13857 – 13862; e) G. Siedenburg, D. Jendrossek, *Appl. Environ. Microbiol.* **2011**, *77*, 3905 – 3915.

