J|A|C|S ARTICLES

Published on Web 11/16/2010

Compelling Computational Evidence for the Concerted Cyclization of the ABC Rings of Hopene from Protonated Squalene

Lidia Smentek and B. Andes Hess, Jr.*

Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37235, United States

Received May 7, 2010; E-mail: b.andes.hess@vanderbilt.edu

Abstract: The long-standing question of what is the nature of the cyclization of squalene to form tetracyclic and pentacyclic triterpenes has been addressed computationally. Using the DFT method with an intrinsic reaction coordinate calculation, we find that the first three rings of protonated squalene were formed without the intermediacy of mono- or bicyclic carbocations. The cyclization, calculated in the gas phase, proceeds in a highly asynchronous, concerted reaction to yield the tricyclic, tertiary carbocation with a 5-membered C ring. The fourth double bond of squalene is not properly oriented for the ring expansion of the C ring in concert with the formation of the 5-membered ring.

Introduction

The ability of Nature to synthesize complex organic molecules is perhaps best exemplified by two closely related biosyntheses, the conversion of squalene (**1**) to lanosterol (**2**), the precursor to all steroids, and its conversion to hopene (**3**) (Figure 1). The link between squalene and steroids was first suggested in 1926[1,2] when Heilbron stated "That squalene is intimately connected with metabolic processes is self-evident, and in this connection we desire to direct attention to a possible relationship between this hydrocarbon, stigmasterol ($C_{30}H_{50}O$) and cholesterol ($C_{27}H_{46}O$)." This connection was confirmed 25 years later by the report that feeding mice with isotopically labeled squalene leads to their production of labeled cholesterol.[3] Earlier this year isolation of the animal steroid progesterone from the vascular plant, *Juglans regia*, has provided striking evidence of the close connection between the common ancestry of steroids and triterpenes such as the hopenes isolated from the plant world.[4]

In 1934 Robinson was the first to address the question of the mechanism of the cyclization of squalene to yield cholesterol,[5] though his original suggestion of how the squalene molecule folded to provide the four rings was shown 19 years later by Woodward and Bloch to be incorrect.[6] Over the next 10 years a mechanism for the cyclization emerged that involved carbocations,[7] though generally the cyclizations of the four rings were thought to be concerted. In 1966 Corey[8] and van Tamelen[9] provided convincing evidence that squalene was epoxidized and then protonated to initiate the cyclizations that lead to lanosterol. In the case of the formation of hopene, the enzyme is thought to protonate squalene directly to yield **4a** (Figure 2). In 1967 van Tamelen suggested the possibility of "a stepwise annelation process involving conformationally rigid carbonium ions" rather than a concerted reaction and the possibility of a tricyclic intermediate (**4e**) in which the C ring exists as a five-membered ring.[10] The question then arose how the six-membered C ring is formed, since it would have to involve an anti-Markovnikov ring expansion to a secondary carbocation (**4d**) followed by ring closure of the D ring.

Over the next 20 years much experimental work on the enzymatic cyclization of squalene was carried out in order to learn more about how the enzyme actually is able to chaperone such a complex biosynthetic reaction yielding highly stereospecific products. It included mutated enzymes as well as chemically modified squalenes. This work has been extensively reviewed.[11−17] The apparent renewed interest in the biosynthesis of terpenoids is due in part to the potential of metabolic engineering to increase yields of desired terpenoids as well producing new, "synthetic" terpenoids of commercial value.[18]

In 2002, based on density functional calculations (DFT), we proposed that the formation of the 6−6−6 ABC secondary carbocation (**4e** in Figure 2) can nicely be avoided by a concerted reaction in which the 6−6−5 ABC carbocation (**4c**

(1) Heilbron, I. M.; Kamm, E. D.; Owens, W. M. *J. Chem. Soc.* **1926**, 1630.
(2) Channon, H. J. *Biochem. J.* **1926**, *20*, 400.
(3) Langdon, R. G.; Bloch, K. *J. Am. Chem. Soc.* **1952**, *74*, 1869.
(4) Pauli, G. F.; Friesen, J. B.; Gödecke, T.; Farnsworth, N. R.; Glodny, B. *J. Nat. Prod.* **2010**, *73*, 338.
(5) Robinson, R. *J. Soc. Chem. Ind.* **1934**, *53*, 1062.
(6) Woodward, R. B.; Bloch, K. *J. Am. Chem. Soc.* **1953**, *75*, 2023.
(7) Eschenmoser, A.; Ruzika, L.; Jeger, O.; Arigoni, D. *Helv. Chim. Acta* **1955**, *38*, 1890.
(8) Corey, E. J.; Russey, W. E.; Ortis de Montellano, P. R. *J. Am. Chem. Soc.* **1966**, *88*, 4750.
(9) van Tamelen, E. E.; Willett, J. D.; Clayton, R. B.; Lord, K. E. *J. Am. Chem. Soc.* **1966**, *88*, 4752.
(10) van Tamelen, E. E.; Willett, J. D.; Clayton, R. B. *J. Am. Chem. Soc.* **1967**, *89*, 3371.
(11) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812.
(12) Connolly, J. D.; Hill, R. A. *Nat. Prod. Rep.* **2001**, *18*, 560.
(13) Xu, R. F.; Fazio, G. C.; Matsuda, S. P. T. *Phytochemistry* **2004**, *65*, 261.
(14) Yoder, R. A.; Johnston, J. N. *Chem. Rev.* **2005**, *105*, 4730.
(15) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412.
(16) Abe, I. *Nat. Prod. Rep.* **2007**, *24*, 1311.
(17) Degenhardt, J.; Köllner, T. G.; Gerhenzon, J. *Phytochemistry* **2009**, *70*, 1621.
(18) Roberts, S. C. *Nat. Chem. Biol.* **2007**, *3*, 387.

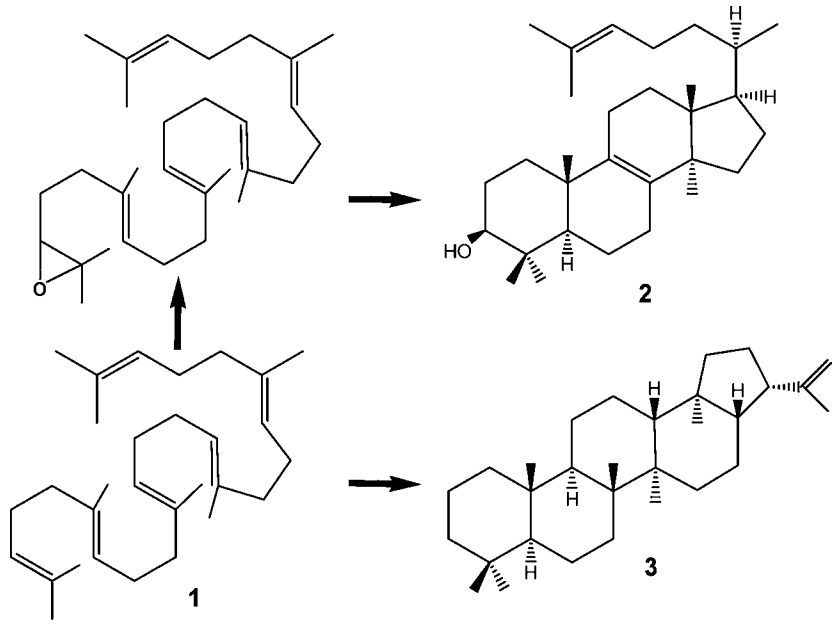

Figure 1. Conversion of squalene oxide and squalene to lanosterol and hopene.

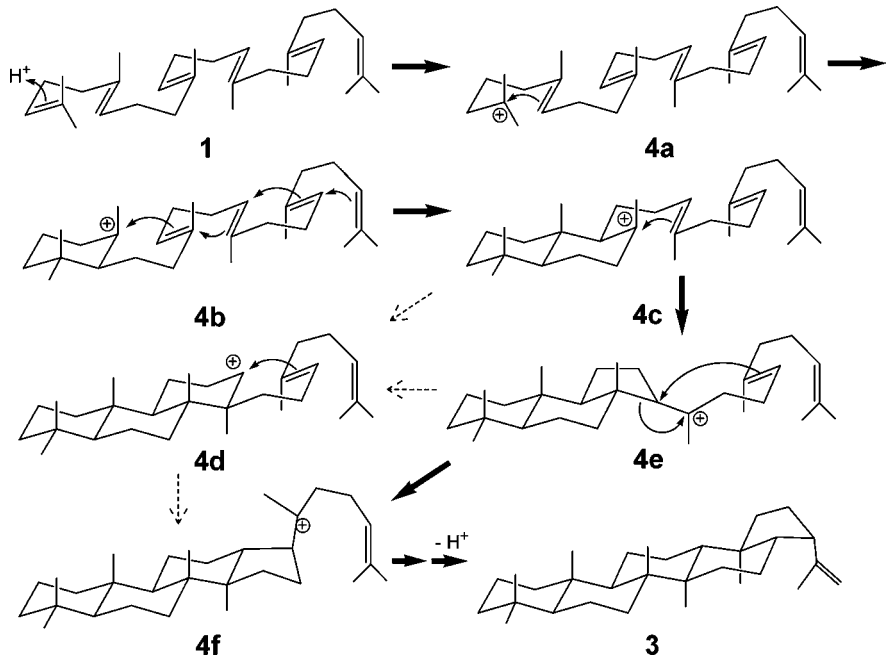

Figure 2. Pathway for the conversion of squalene (**1**) to hopene (**3**). Note that **4c** might undergo an anti-Markovnikov ring closure (dashed arrow) to give the secondary carbocation **4d** or undergo the Markovnikov ring closure to **4e**, which subsequently can undergo a concerted ring expansion of the C ring and formation of the D ring to yield **4f** or an anti-Markovnikov ring expansion to form **4d**. A second ring expansion/formation of **4f** gives **3**.

in Figure 2) undergoes a concerted but highly asynchronous reaction. A transition structure was located which links the 6−6−5 ABC tertiary carbocation (from either squalene or squalene oxide, e.g., **4e** in Figure 2) with the 6−6−6−5 ABCD tertiary carbocation (**4f** in Figure 2).[19] While a truncated model system was used in these calculations, a more complete model of squalene gave the analogous transition structures for the squalene and squalene oxide cyclizations (structures **5** and **6** in Figure 3, respectively).[20] Since this novel mechanism was suggested, which avoids an anti-Markovnikov cyclization, a number of other examples have been proposed in terpene chemistry in which anti-Markovnikov cyclizations and rearrangements are avoided via highly asynchronous, concerted reactions.[21−24]

While much has been learned in the past 50 years about the mechanism of the cyclizations of squalene, still a fundamental question remains, namely, whether the overall process is concerted or not. Are the rings formed via discrete carbocation intermediates,

(19) Hess, B. A., Jr. *J. Am. Chem. Soc.* **2002**, *124*, 10286.
(20) Hess, Jr., B. A.; Smentek, L. Unpublished results.
(21) Hong, Y. J.; Tantillo, D. J. *Org. Lett.* **2006**, *8*, 4601.
(22) Lodewyk, M. W.; Gutta, P.; Tantillo, D. J. *J. Org. Chem.* **2008**, *73*, 6570.
(23) Hong, Y. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2009**, *131*, 7999.
(24) Hong, T. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2010**, *132*, 5375.

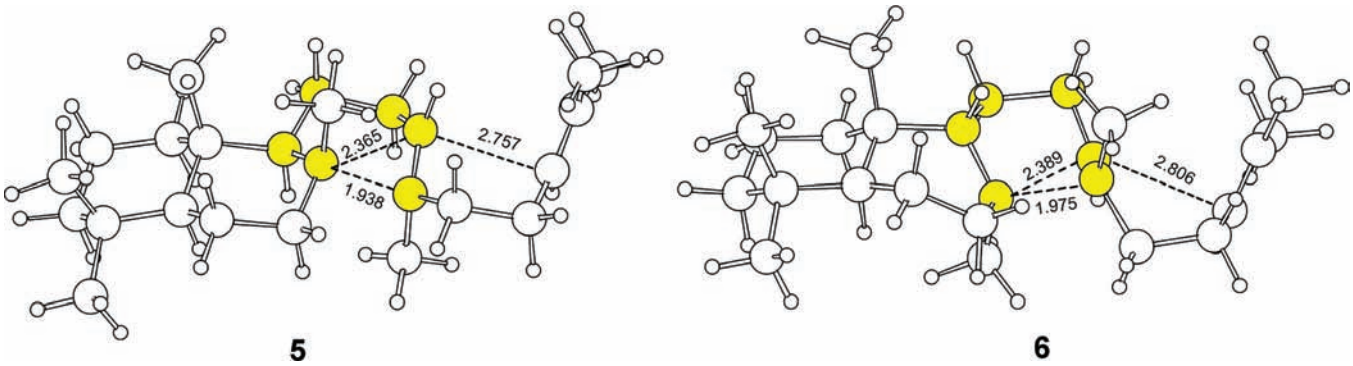

Figure 3. DFT transition structures for the C−D ring formation for hopene (**5**) and lanosterol (**6**). Note that **5** has a chair−chair AB ring system and **6** a chair−boat AB ring system. The hydroxyl group was not included in these calculations, since it is quite distant from the reacting part of the molecule. Distances between selected carbon pairs are given in Å.

or is the squalene cradled within the enzymes in such a way that a concerted, highly asynchronous reaction forms the rings?[25−27] Over the years there has been much speculation and a myriad of proposals about the answer to this question with as yet no definitive answers. However, there appears to be some agreement that in both squalene and squalene epoxide cyclizations the 6−6−5 ABC carbocation is formed at least fleetingly as an intermediate in the overall cyclization process.[11,13,28−30] One of the more convincing pieces of experimental evidence for the intermediacy of the 6−6−5 tricylic carbocation (**4e**) was published by Hoshino in which he used a mutant SHC as well as a truncated squalene with a hydroxyl group at the C-22 position in which the cyclizations were "stopped" at 6−6−5 tricyclics.[31] He concluded that the "trapping" of the 6−6−5 tricylic carbocation (**4e**) in both experiments provides strong evidence for the intermediacy of **4e** in the enzymatically catalyzed cyclization of squalene. Subsequently, Hoshino and Abe reported additional trapping experiments using diols of squalene,[32] which also supported the intermediacy of **4e**. Corey provided similar evidence for the intermediacy of the 6−6−5 carbocation in the conversion of squalene oxide to lanosterol.[33]

In 2004 Reinert, Balliano, and Schulz published a remarkable experiment, which contributed a significant piece of circumstantial evidence in support of a concerted pathway in the cyclization of squalene by squalene−hopene cyclase.[34] They were able to obtain an X-ray structure of 2-azasqualene (**7**, the nitrogen deactivated the cyclization process) encapsulated in the enzyme. On the basis of the X-ray analysis of this system they suggested that the squalene molecule was held in such a conformation that the interaction of the developing positive charges with the ring-forming C−C double bond would likely lead to barrierless ring closures of the A−D rings. They suggested that the reaction would pause at the 6−6−6−5 tetracyclic, tertiary carbocation (**4f** in Figure 2) before undergoing the concerted ring expansion of ring D and formation of the five-membered E ring. Their results provide a conformer of squalene that may be used as a starting point for computationally addressing the nature of the cyclization of squalene.

In contrast to Schulz's suggestion of concerted cyclizations, Hoshino and Sato, based on an extensive study of the cyclization of squalene in a series of site-directed mutation experiments, concluded that in addition to the tricyclic cation, the bicyclic and monocylic cations were also likely to be intermediates in the conversion of squalene to hopene.[35] By using various mutant SHC's for the cyclization of squalene, they were able to isolate mono-, bi-, and tricyclic triterpene products. They argued that these mutants changed the ability of the enzyme to stabilize the various cationic interemediates in the cyclization, which led to the formation of these mono-, bi-, and tricyclic products. Additional studies reported by Hoshino[36] in which unnatural amino acids were incorporated into the catalytic sites of SHC were also used to support the distinct intemediacy of carbocations along the course of the formation of the A−C rings.

An earlier computational study has addressed the question of the overall concertedness of the cyclization of squalene to yield hopene, published in 2003 by Gao.[37] He carried out a QM/MM calculation (AM1 was used for the QM part of the method) on the cyclization and based on the result concluded that the bicyclic A−B 6−6 bicyclic is formed as an intermediate in a concerted reaction from protonated squalene, which subsequently undergoes concerted cyclizations to form the tetra- and pentacyclic systems. However, in a recently published experimental paper[38] based on kinetic studies of the cyclization of farnesyl acetate both in solution and on zeolites, including secondary isotope effects, the authors concluded that the

(25) A referee suggested that there might be some confusion over how such cyclizations have been described in the past. While the term "cascade" has been widely used to describe the multiple cyclizations that must take place to form the tetra- and pentacyclic systems, it is our understanding that this term implies neither a concerted nor a stepwise mechanism but rather simply a series of cyclizations. The stepwise mechanism in this case implies one or more intermediate carbocations in the pathway being studied. In a concerted reaction the cyclizations might be synchronous, that is, all occurring at the same time, or asynchronous, that is, one cyclization proceeding to some degree before the next one begins to take place. See refs 26 and 27.

(26) Dewar, M. J. S.; Gie, C. *Acc. Chem. Res.* **1992**, *25*, 537.
(27) Tantillo, D. J. *J. Phys. Org. Chem.* **2008**, *21*, 561.
(28) Xiong, Q.; Rocco, F.; Wilson, W. K.; Xu, R.; Ceruti, M.; Matsuda, S. P. T. *J. Org. Chem.* **2005**, *70*, 5362.
(29) Matsuda, S. P. T.; Wilson, W. K.; Xiong, Q. *Org. Biomol. Chem.* **2006**, *4*, 530.
(30) Wu, T.-K.; Li, W.-H.; Chang, C.-H.; Wen, H.-Y.; Liu, Y.-T.; Chang, Y.-C. *Eur. J. Org. Chem.* **2009**, 5731.
(31) Hoshino, T.; Kouda, M.; Abe, T.; Ohashi, S. *Biosci. Biotechnol. Biochem.* **1999**, *63*, 2038.
(32) Abe, T.; Hoshino, T. *Org. Biomol. Chem.* **2005**, *3*, 3127.
(33) Corey, E. J.; Virgil, S. C.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Singh, V.; Sarshar, S. *J. Am. Chem. Soc.* **1995**, *117*, 11819.
(34) Reinert, D. J.; Balliano, G.; Schulz, G. E. *Chem. Biol.* **2004**, *11*, 121.
(35) Hoshino, H.; Sato, T. *Chem. Commun.* **2002**, 291.
(36) Morikubo, N.; Fukuda, Y.; Ohtake, K.; Shinya, N.; Kiga, D.; Sakamoto, K.; Asanuma, M.; Hirota, H.; Yokoyama, S.; Hoshino, T. *J. Am. Chem. Soc.* **2006**, *128*, 13184.
(37) Rajamnai, R.; Gao, J. *J. Am. Chem. Soc.* **2003**, *125*, 12768.
(38) Raptis, C.; Lykakis, I. N.; Tsangarakis, C.; Stratakis, M. *Chem.—Eur. J.* **2009**, *15*, 11918.

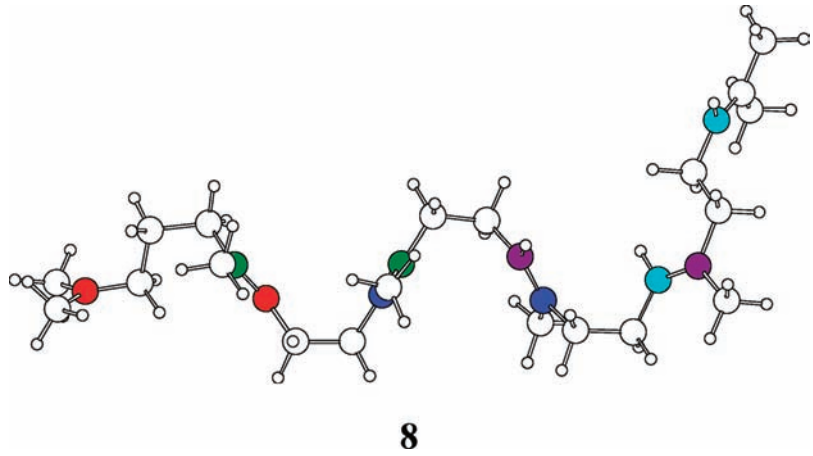

Figure 4. DFT structure of the conformer of protonated squalene found obtained using the X-ray structure[34] of 2-azaqualene as a starting point. The two atoms of the same color denote which carbons form the five rings in hopene.

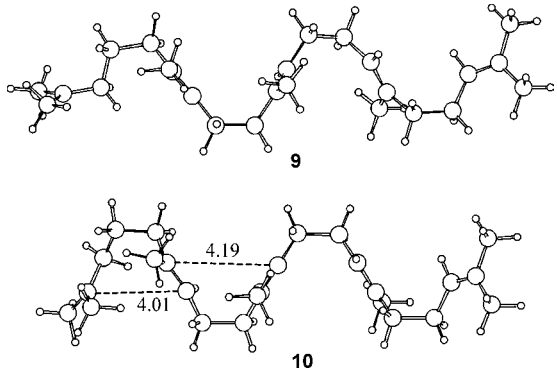

Figure 5. DFT structures of the C-25 carbocation conformer of truncated squalene (**9**) and starting transition structure (**10**) for the IRC calculation. The two dashed lines connect the two atoms that will form the A and B rings in the cyclization. The two distances given are in Å.

formation of the bicyclic decalin product (chair,chair) is a concerted process along with protonation of the farnesyl acetate, which is in disagreement with Gao's computational results as well as with Hoshino's results mentioned above.

We previously carried out several studies on small model systems on the cyclizations of the individual rings A, B, and C in an attempt to delineate the mechanism of their cyclizations to give the 6−6−5 ring system. In 2004 it was shown that a 10-carbon open-chain carbocation model system cyclized by the double bond undergoing an electrophilic attack by the carbocation to form the A ring in a chair conformation.[39] Subsequently, we reported a 15-carbon model system with 10 ring-forming carbons plus 5 methyl groups to obtain evidence for the concerted cyclization of the A and B rings of squalene (chair, chair).[40] Further support of AB concerted ring formation was provided in 2008[41] when we reported the generation of a DFT conformer of protonated squalene whose structure is remarkably similar to the X-ray structure of 2-azasqualene encapsulated in SHC.[34] A similar conformational study of a model system of the B−C ring closure to the five-membered ring suggested that this ring closure is likely to be concerted in the lanosterol biosynthesis.[42]

While, as mentioned above, there is experimental evidence for the interemediacy of the ABC 6−6−5 carbocation (**4e**), no unequivocal computational or experimental evidence of the nature of the cyclization of these three rings has been obtained. Gao's results suggest at least one intermediate cation intermediate. While our previous results are suggestive that the A−B−C cyclizations might be concerted, they do not consider these cyclizations in a unified manner, which is the main endeavor of the research reported here. Herein we present results using the intrinsic reaction coordinate method (IRC)[43,44] with DFT calculations to study the concertedness or lack thereof of the A−D ring formation in the cyclization of squalene. Of course, all of these calculations are done in the gas phase in the absence of an enzyme. However, if it can be shown that a conformer of squalene exists in the gas phase that once protonated will react to some degree in a concerted fashion; then one can conclude that after the protonation step the main role of the enzyme in these cyclizations might be to hold the squalene chain in the proper conformation, especially if this gas-phase conformation is similar to that found by Schulz for 2-azasqualene encapsulated in squalene hopene cyclase (SHC).[34]

Methods

All calculations were performed with GAUSSIAN 09W.[45] Geometries were optimized, and all IRC calculations[43,44] were performed with the 6-31G*[46] basis set and the B3LYP functional.[47,48] Single-point energy calculations were carried out on all stationary points with mPW1PW91/6-31G*//B3LYP/6-31G* as suggested by Matsuda.[29] Since the main conclusions of this paper are based on the IRC method, one should be aware that pathways calculated with this method might not always agree with those determined experimentally.[49]

Results

In order to begin the study of the pathway (concerted or not) of A−D ring formation in the biosynthesis of hopene from squalenes, one must choose as a starting point a conformation

(39) Hess, B. A., Jr. *Collect. Czech. Chem. Commun.* **2004**, *69*, 261.
(40) Hess, B. A., Jr.; Smentek, L. *Org. Lett.* **2004**, *6*, 1717.
(41) Hess, B. A., Jr.; Smentek, L. *Collect. Czech. Chem. Commun.* **2008**, *73*, 786.
(42) Hess, B. A., Jr. *Org. Lett.* **2003**, *5*, 165.
(43) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363.
(44) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523.
(45) Frisch, M. J. et al. *Gaussian 09*, revision A.02; Gaussian, Inc.; Wallingford, CT, 2009.
(46) Hariharan, P. C.; Pople, J. A. *Theor Chim. Acta* **1973**, *28*, 213.
(47) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
(48) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(49) Ammal, S. C.; Yamataka, H.; Aida, M.; Dupuis, M. *Science* **2003**, *299*, 1555.

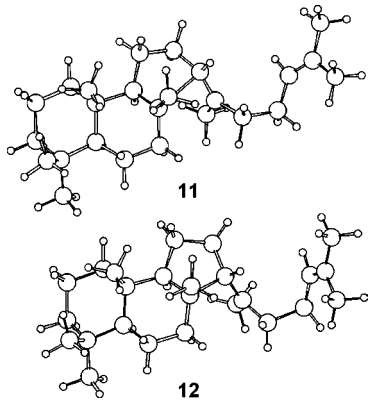

Figure 6. Structures of the two minima located for the ABC carbocation.

of squalene from the almost countless number of its possible conformers. The X-ray structure of 2-azasqualene encapsulated in SHC[34] in fact provided a very important starting "guess" for this conformation. We earlier found a computed DFT conformer of protonated squalene (**8**) very close in structure to that of 2-azasqualene mentioned above (Figure 4). Since we were interested in the formation of only the first four rings, a model system was chosen that includes the first 25 carbon atoms of squalene (see Figure 5). Replacement of the first 15 carbons atoms (those that would involve formation of the A and B rings) of this 25 carbon unit with the previously found concerted transition structure for the AB ring formation (see above) provided an initial guess for the geometry of a transition structure that would be key in initiating the cascade of ring closures. The search for a transition structure that utilized this guess very quickly gave the transition structure (**10**) shown in Figure 5. An IRC calculation with **10** as the starting transition structure provided in one direction structure **9**. This was encouraging since the structure of **9** had been found to be very similar to the first 25 carbons of the enzyme-encapsulated 2-azasqualene.[34] The opposite direction (product formation) led to a very lengthy IRC pathway, which turned out to give a highly asynchronous, concerted cyclization of the six-membered chair conformers of the A and B rings, and finally, a minimum (the only one) was reached in which the five-membered C ring had been formed. The initial minimum found, **11** in Figure 6, clearly shows the cyclized five-membered C ring.

The newly formed C−C bond of the C ring is rather long (2.44 Å),[50] which is an artifact of the B3LYP functional as shown by Matsuda (see ref 28 and Table 7 in ref 28). He found that with other functionals this minimum has a much shorter C−C bond length of the newly formed bond in the 5-membered ring. Interestingly, we were able to locate a second minimum (**12** in Figure 6) only 0.7 kcal/mol lower in energy than **11**. The above-mentioned long C−C bond in **11** is 1.73 Å and in **12** is still rather long, due likely to hyperconjugation as suggested by Matsuda for the analogous elongated C−C bond in the dammarenyl cation.[28] Close examination of the structures of **11** and **12** indicates that they are essentially identical with the exception of the length of the bond forming the C ring.

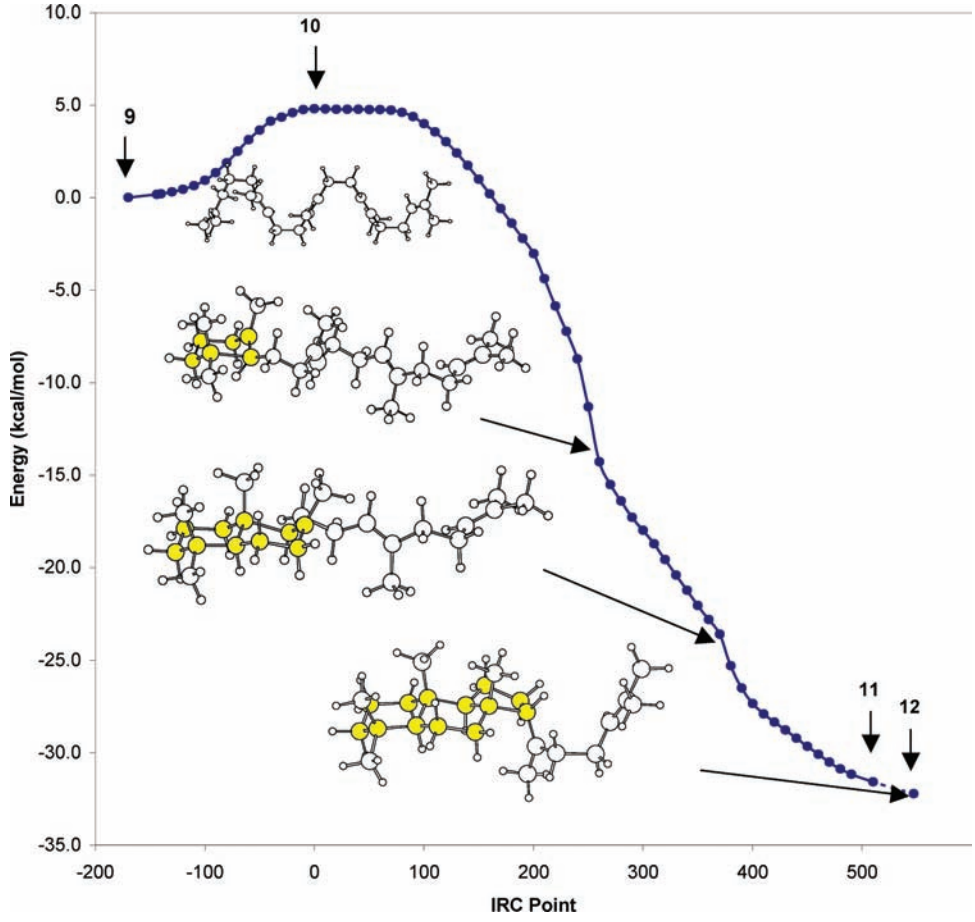

Figure 7. DFT IRC pathway of the formation of the A, B, and C rings for the hopene cyclization. This complete pathway was obtained in a single IRC calculation. Energies relative to structure **9** are plotted against every 10th IRC point. Negative points link the **9** with transition structure **10** and for positive points **10** with the ABC minimum **11**.

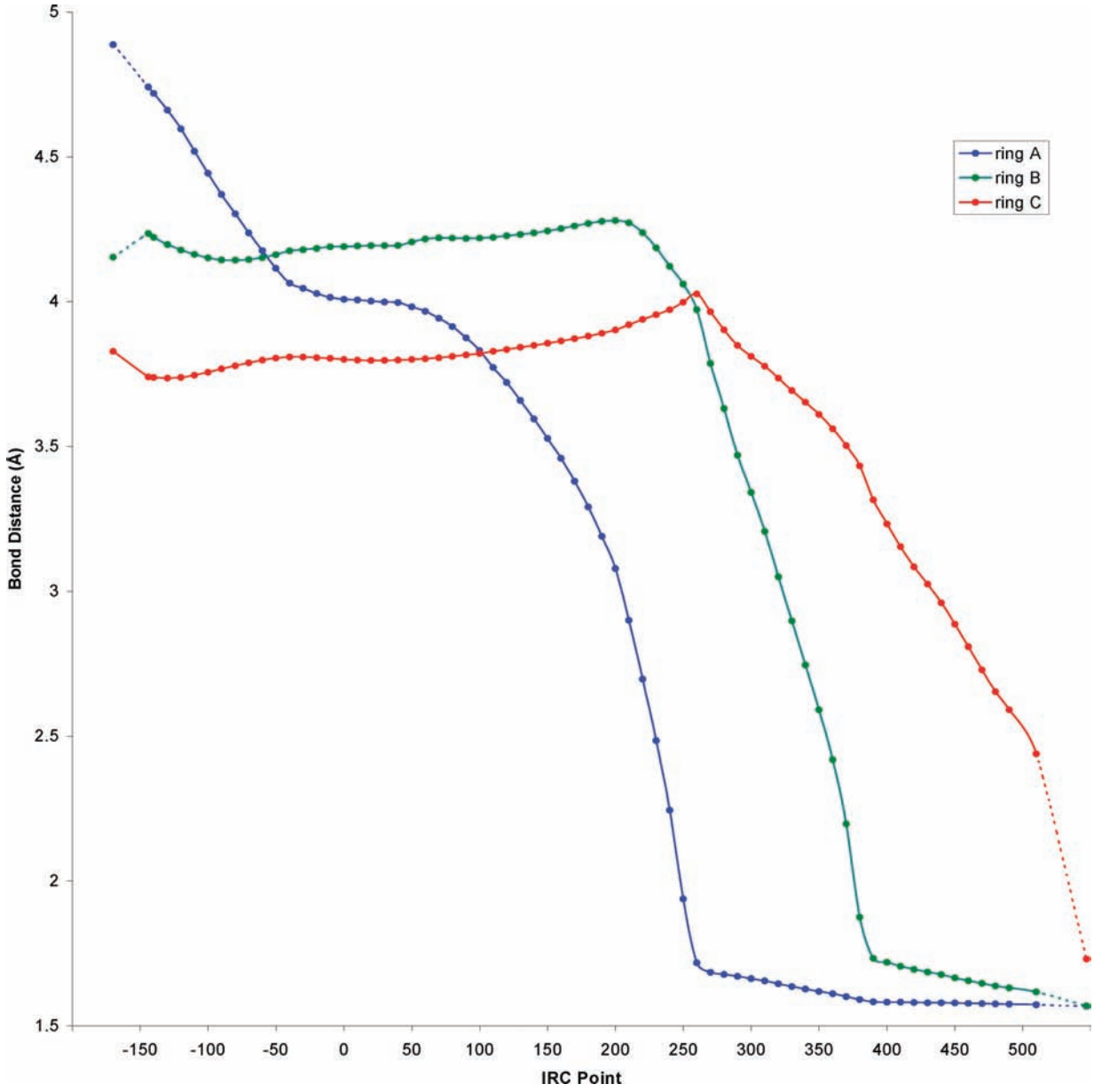

Figure 8. Change in ring-forming C−C bond distances vs IRC point. The three points at the far left are the corresponding distances in **9**, and the three points at the far right the corresponding bond distances in **12**. The transition structure is at point 0.

Due to the flatness of the potential surface in this region, no transition structure could be located connecting **11** with **12**.

In Figure 7 relative energies of points along the IRC pathway are plotted against every 10th point. In the plot are also presented the transition structure (**10**) as well the two structures along the pathway in which the A and A−B rings have been formed and finally the tricyclic product **12**.

Discussion

The absence of any minimum on the IRC pathway shown in Figure 7 indicates that formation of these three rings, while highly asynchronous, is truly of a concerted nature in the gas phase. The three "steps" corresponding to the formation of the three rings are clearly seen in Figure 7, which supports the asynchronicity of these cyclizations. From this figure one can also see that the amount of energy released during the cyclization of each ring decreases for A, B, and C, from 14, 12, to 7 kcal/mol, respectively. These energies do not include zero-point energy corrections. However, they are in qualitative agreement with energies for the cyclization of the A−C rings in the hopene cyclization reported by Matsuda.[28,29] Quantitative agreement is not to be expected here; since Matsuda has previously shown that while the B3LYP functional when used with the 6-31G* basis set typically gives trustworthy geometries, there is significant error in cyclization energies.[28,29]

Further insight into the asynchronicity of the A−C cyclizations may be obtained by following three key bond distances of the new sp^3-sp^3 C−C bonds being formed during the course of the overall concerted reaction (see Figure 8). The high degree of asynchronicity of the three ring closures is indeed clearly seen from this plot, since each bond is well on its way to forming before the next bond begins to form. Note that the

(50) This was also found to be the case for our earlier study of the model system for the ring expansion, ring closure of the C and D rings. See ref 19.

A-ring bond undergoes a significant change between **9** and the transition structure. The activation energy for this step (**9** to **10**) is on the order of 4 kcal/mol. It is likely that the enzyme produces the initial carbocation with the squalene having a conformation closer to that of the transition structure than conformer **9** does, but obviously this conclusion is speculative.

The concertedness of A−C ring formation contradicts earlier conclusions based on both previous experimental and computational results. However, we would argue that the experimental results discussed in the Introduction would also support the highly ascynchronous concerted pathway we found with the IRC calculation presented above, since the IRC mechanism does indeed indicate that there is significant positive charge built up (as a result of the asynchronicity) along the pathway at the carbons that then interact with a double bond to form the next ring in the cyclization equence. Anything that would destabilize the build up of these positive charges (such as a mutant SHC) in the concerted mechanism would also likely give rise to the mono- and bicyclic products isolated by Hoshino and Sato.[31] We propose that their elegant site-directed mutation experiments are not able to differentiate between the intermediacy of multiple carbocations on the cyclization pathway and a concerted mechanism that is highly asynchronous. We would again argue that Hoshino's results with unnatural amino acids[36] would also support a highly asynchronous concerted reaction.

Given the concertedness of the formation of the A−C rings, the question must be addressed why does the cyclization process apparently stop or pause with the formation of the 5-membered C ring? Formation of the D ring requires not only a ring closure involving the next double bond in the squalene chain but also the ring expansion of the C ring from a 5-membered to 6-membered ring. In the case of the formation of rings A−C its concerted nature depends primarily on the juxtaposition of a double bond with a tertiary carbocationic center. In the case of the D ring the eventual ring closure requires build up of a positive charge on what formally would be a secondary carbocationic center before there can be significant interaction between this center and the carbon−carbon double bond. However, as noted by Schulz, this 6−6−5 carbocation when encapsulated in the enzyme appears to lack the stabilization afforded by the enzyme for the 6−6 carbocation and the 6−6−6−5 carbocation: "A putative fourth cation at C15 after the formation of the 6−6−5 tricycle must be short lived because it has no stabilizing partner. In contrast, the fifth cation (the 6−6−6−5 cation) at C19 generated with the 6−6−6−5 tetracycle is well accommodated...".[34] The enzyme by raising the energy of the 6−6−5 cation would "lower" the activation energy for the concerted ring expansion/ring formation (6−6−5 to 6−6−6−5) possibly allowing this step also in the presence of the enzyme to continue the overall concerted cyclization of squalene. Of course, our gas-phase DFT calculations do not account for any enzymatic effects. Matsuda calculated this step to have an activation energy of 7 kcal/mol, which is a significant barrier in the absence of the enzyme.[28] He also calculated the energy of reaction for this combined step to be −0.3 kcal/mol with B3LYP/6-31G* and −4.7 kcal/mol with mPW1PW91/6-311+G(2d,p), the latter method giving much more reliable energies as shown by him.

Another factor must also be considered. The originally encapsulated squalene during the course of the formation of the A−C rings undergoes significant geometric changes within the enzyme. It is possible that before the concerted C-ring expansion and D-ring cyclization can take place a reorganization of the encapsulating enzyme must occur in order to favorably chaperone this next rather complicated transformation. In Figure 9 there is a plot of the three C−C bond distances from the latter part of the IRC, the two bonds that form the 5- or 6-membered C ring, and the bond that would form the 5-membered D ring. It is seen from the plot that there is no tendency for the reacting carbocation to undergo the C-ring expansion and D-ring cyclization. For this to occur there would have to be a crossing of the two lower curves (ring expansion) and a marked shortening of the D-ring-forming C−C bond, neither of which is apparent from the plot. Hence, at least within our model, there is simply no tendency for the D-ring formation to be in concert with the A−C ring formation.

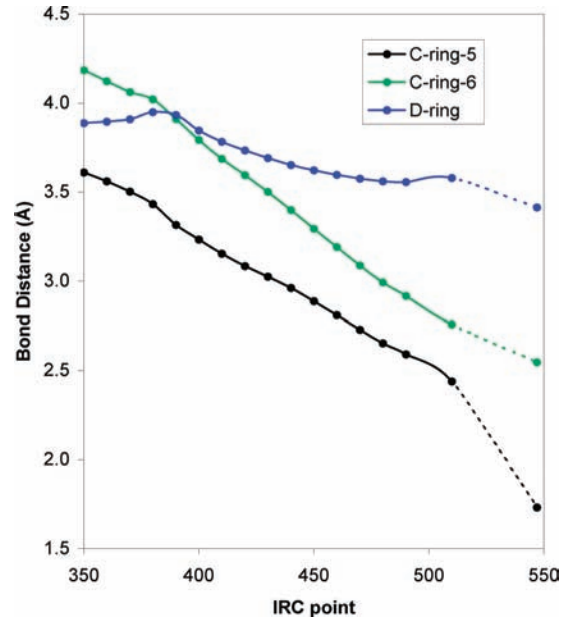

Figure 9. Change in bond distances crucial for the formation of the 6-membered C ring and the 5-membered D ring during the latter part of the IRC pathway. The dotted lines at the right connect minima **11** and **12**.

Conclusions

Beginning with a computationally found conformer of squalene that is very similar to the structure found by Schulz[34] of 2-azasqualene encapsulated in SHC, an IRC pathway was located that avoids the intermediacy of either the mono- or the bicyclic carbocation intermediates previously proposed by others based on both experimental and compuational results. The calculations described here give very strong support to a concerted, highly asynchronous cyclization of rings A−C in the enzymatically controlled conversion of squalene to hopene with formation of the intermediate 6−6−5 tricyclic carbocation prior to formation of the D and E rings. These results parallel Matsuda's finding in the cyclization of squalene oxide to lanosterol, lupeol, and hopen-3β-ol, where he was unable to locate minima corresponding to the formation of the A and B rings.[29] It is argued that experimental results previously used in support of a stepwise mechanism with intermediate cations do no rule out a highly asynchronous, concerted pathway for the cyclization of the A−C rings of hopene.

Supporting Information Available: Coordinates and energies for all structures computed and complete ref 45. This material is available free of charge via the Internet at http://pubs.acs.org.

JA1039133