The cyclization mechanism of squalene in hopene biosynthesis: the terminal methyl groups are critical to the correct folding of this substrate both for the formation of the five-membered E-ring and for the initiation of the polycyclization reaction†

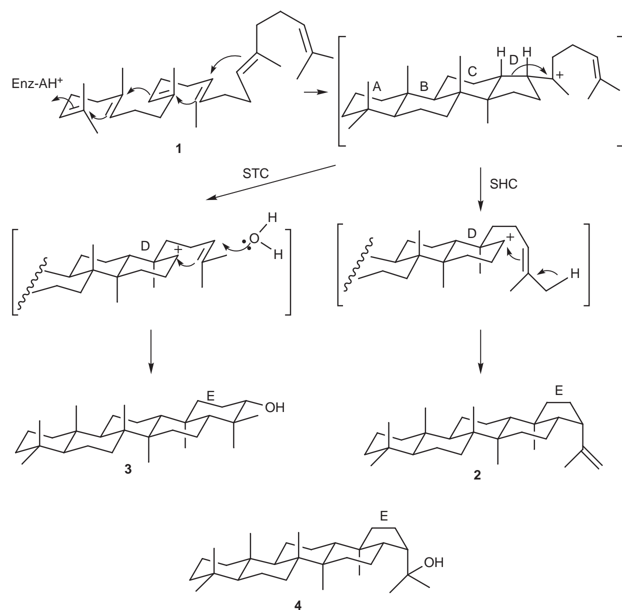
Tsutomu Hoshino,*[ab] and Tomohiro Kondo[b]

[a] *Department of Applied Biological Chemistry, Faculty of Agriculture, and* [b] *Graduate School of Science and Technology, Niigata University, Ikarashi, Niigata 950-2181, Japan. E-mail: hoshitsu@agr.niigata-u.ac.jp*

Received (in Cambridge, UK) 18th February 1999, Accepted 10th March 1999

Incubations of C(23)-norsqualenes 5 and 6, lacking one of the two terminal methyl groups, with squalene-hopene cyclase gave unprecedented products 7 and 8 having a tetrahymanol skeleton together with a neohopane skeleton 12, strongly suggesting that the two geminal methyls of squalene 1 are critical to the formation of the five-membered E-ring in hopene biosynthesis and also are required to initiate the cyclization reactions of 1 into the pentacyclic triterpenes 2 and 3.

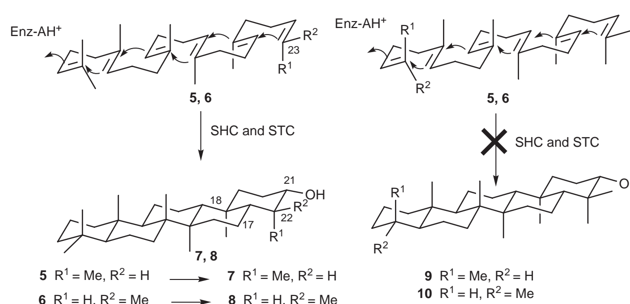
The biocyclization of squalene **1** into hop-22(29)-ene **2** catalyzed by squalene-hopene cyclase [EC 5.4.99.-] (SHC) or into tetrahymanol **3**, by squalene-tetrahymanol cyclase [EC 5.4.99.-] (STC), is one of the most intricate biochemical reactions (Scheme 1).[1] Compound **4**, hopan-22-ol, is a minor by-product. The cyclization proceeds *via* precise enzymatic control to form the fused 6/6/6/6/5- or 6/6/6/6/6-ring system and nine new stereocenters. The polyene cyclization reaction of **1** is analogous to that of 2,3-oxidosqualene into lanosterol catalyzed by lanosterol synthase.[1] Recently, it has been proposed that a ring expansion process from the five- to the six-membered ring is involved in the D-ring formation of hopene biosynthesis, prior to the further cyclization (Scheme 1).[2a,b] Formation of the five-membered intermediate D-ring is consistent with the Markovnikov rule. The expansion process of the D-ring may also be responsible for tetrahymanol biosynthesis.[2c] Such a ring expansion reaction was also demonstrated for the oxidosqualene cyclization.[3] Computational energy calculations of the five- and six-membered intermediates have supported the feasibility of such a ring expansion process.[4] The structure of **2** differs from that of **3** only in the terminal E-ring, *i.e.* five- and six-membered rings for **2** and **3**, respectively. To gain insight into the different cyclization mechanism between SHC and STC, the substrate analogues **5** and **6**, lacking one of the two terminal methyls in squalene backbone **1**, were separately incubated with SHC or STC, resulting in the formation of **7**, **8** and **12** as SHC enzyme products, and **7** and **8** as STC products; SHC, despite responsibility for the formation of the five-membered E-ring, did produce a tetrahymanol skeleton having a six-membered E-ring. We show herein that the terminal methyl groups play a critical role for the determination of either five- or six-membered E-ring formation.



Scheme 1

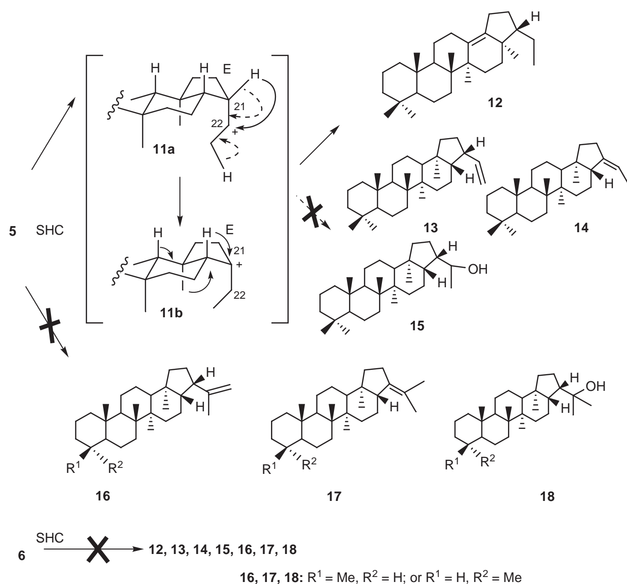
The analogues **5** and **6** were synthesized as follows: (±)-2,3-oxidosqualene was treated with HIO_4 to give the C_{27}-aldehyde, which was then subjected to a Wittig reaction with $EtPPh_3Br$ in the presence of Bu^nLi in THF. The C(23)-norsqualenes **5** and **6** thus obtained were separated with a SiO_2 column (10% $AgNO_3$) with hexane; **6** followed by **5.** The 7 methyl signals of **5** in $CDCl_3$ were as follows: δ_H 1.66 [3H, s, (*E*)-CH_3C=] and 1.58 [18H, (*Z*)-CH_3C=], while those of **6** were: δ_H 1.66 [3H, s, (*E*)-CH_3C=], 1.61[3H, brd, *J* 5.8, (*E*)-CH_3C=], 1.58 [12H, s, (*Z*)-CH_3C=] and 1.57 [3H, s, (*Z*)-CH_3C=].[5]

Separate incubations of **5** and **6** with cell-free homogenates from the cloned *E. coli* harboring SHC under catalytic optimal conditions[6] afforded highly polar products **7** and **8**, respectively (Scheme 2). From the incubation mixtures of **5**, another product **12** was found together with some minor products, both of which had lower polarities than **7** on TLC. Separation of each product was performed using a SiO_2 column with hexane–EtOAc, but the minor products, present in negligible amounts, were inseparable even on $AgNO_3$–SiO_2 TLC. Isolation yields of **7** and **12** from **5** were 17 and 12%, respectively, while that of **8** from **6** was 16%; neither **12** nor the minor products were detected from **6** (Scheme 3), using the same quantities of the substrate and the cell-free extracts as for **5**.



Scheme 2

† Details of orbital interactions and 2D NMR analyses for **7**, **8** and **12** are available from the RSC website, see: http://www.rsc.org/suppdata/cc/1999/731/



Scheme 3

Independent incubations of **5** and **6** with cell-free extracts from *Tetrahymena pyriformis* STC[7] also gave **7** and **8** (1 : 2.5 ratio), respectively. No other products were detected. The products with STC were indistinguishable from those with SHC *via* GC-MS.

The structures of highly polar compounds **7**[5] ($C_{29}H_{50}O$, EI-HRMS: *m/z* 414.3834; requires 414.3862) and **8**[5] ($C_{29}H_{50}O$, *m/z* 414.3854) were determined *via* NMR analysis. The signals at δ_C 74.1 and 76.7 in **7** and **8**, respectively, proved the involvement of a hydroxy group. Detailed analyses using 2D NMR revealed that **7** and **8** had a pentacyclic tetrahymanol skeleton possessing a chair conformation of the E-ring.[5] The hydroxy groups of both **7** and **8** were in the same equatorial orientation at the 21-position, but the arrangements of the methyl at the 22-position were different; axial for **7** by taking account of the coupling constants of H-21 [δ_H 3.69 (*ddd, J* = 11.2, 5.2, 5.2 Hz)] and equatorial for **8** owing to the *ddd* splitting of H-21 [δ_H 3.05 (*J* = 11.0, 11.0, 4.8 Hz)]. Product **12**[5] ($C_{29}H_{48}$, EI-HRMS: *m/z* 396.3722; requires 396.3756) did not have a hydroxy group. One of the 7 methyl groups appeared as a triplet (*J* = 7.6 Hz) and the other 6 methyls as singlets in the ^{1}H NMR spectrum, showing the presence of one ethyl group in **12**, and detailed analysis revealed a neohopene skeleton,[2a] but with a carbon skeleton of C_{29}.

Formation of the tetrahymanol skeleton by SHC has never been reported before, and also is quite interesting from the evolutionary aspect of squalene cyclases. The E-ring formation proceeded with complete stereoselectivity; the (23*Z*)-methyl group of **5** was axial, while the corresponding *E*-methyl of **6** was equatorial during the E-ring formation. The methyl orientations of each product from SHC agreed with those from STC (Scheme 2), and were consistent with the previous report that the (23*E*)-methyl of natural **1** was arranged in an equatorial orientation during the E-ring formation.[7a] The hydroxy groups of both **7** and **8** were introduced in the same equatorial disposition as a result of nucleophilic attack of a water molecule in an equatorial direction on the C-21 cation. The cyclization of the six-membered E-ring formation would be a concerted reaction under stereoelectronic control and explained in terms of HOMO–LUMO interactions.[8]

It is noticeable that **12** was produced only from **5**, and not from **6** (Scheme 3). This fact implicates that the (23*Z*)-methyl of **1** is more important than the (23*E*)-methyl for the formation of the hopanyl cation **11** with the 5-membered E-ring. However, the two terminal methyls would be necessary for the complete building of the five-membered E-ring, since the six-membered species **7** and **8** were produced in significant amounts when one of the two methyls was absent. Why was the tetrahymanol skeleton produced by the incubations of the norsqualene **5** and **6** with SHC? One plausible answer may be that the two geminal methyls strongly bind to SHC to acquire the desired conformation, shown in Scheme 1, during the formation of the five-membered E-ring, and the substrate affinity would become looser when one of the terminal methyls is absent, which would have led to the formation of a tetrahymanol skeleton under stereoelectronic control. The binding force of the *Z*-methyl to the SHC would be stronger than the corresponding *E*-methyl, because **12** was produced only from **5**, and not from **6**. The absence of **13** from **5**, although the hopanyl cation **11** has been produced, suggests that the migration of the hydride (1, 2-shift) must be fast compared with the two deprotonation reactions for the formations of **13** and **14**. This would be due to a greater stability of the tertiary C21-carbocation intermediate **11b** compared to the secondary C22-cation **11a**. The hopanol analogue **15** was also not found. At the present time, we cannot propose which terminal methyl of either the *Z*- or *E*-isomers of the natural **1** is responsible for the proton elimination when the double bond in **2** is introduced. Compounds **9**, **10**, **16**, **17** and **18** were also not detected in any reaction mixture from either SHC or STC, all of which are the presumed enzymic products based on the idea that the polycyclization could be initiated from the methyl-deficient part (Scheme 2 and 3). This finding strongly suggested that the two geminal methyls are indispensable for the initiation of polycyclizations by both STC and SHC, which is in contrast to the report that 2,3-*trans*-1′-norsqualene 2,3-oxide, lacking one methyl on the epoxide ring, was cyclized by the lanosterol synthase.[1,9]

This work was partly supported by a Grant-in-Aid to T.H. (No.0966011) from the Ministry of Education, Science and Culture, Japan.We are also indebted to Dr M. Takahashi, Tsukuba University, for the gift of *Tetrahymena pyriformis* (GL strain).

Notes and references

1 I. Abe, M. Rohmer and G. D. Prestwich, *Chem. Rev.*, 1993, **93**, 2189.

2 (*a*) C. Pale-Grosdemange, C. Feil, M. Rohmer and K. Poralla, *Angew. Chem. Int. Ed.*, 1998, **37**, 2237; (*b*) T. Sato, T. Abe and T. Hoshino, *Chem. Commun.*, 1998, 2617; (*c*) I. Abe and M. Rohmer, *J. Chem. Soc., Chem. Commun.*, 1991, 902.

3 E. J. Corey, S. C. Virgil, H. Cheng, C. H. Baker, S. P. T. Matsuda, V. Singh and S. Sarshar, *J. Am. Chem. Soc.*, 1995, **117**, 11 819; E. J. Corey and H. Cheng, *Tetrahedron Lett.,* 1996, **37**, 2709; T. Hoshino and Y. Sakai, *Chem. Commun.,* 1998, 1591.

4 C. Jenson and W. L. Jorgensen *J. Am. Chem. Soc.*, 1997, **119**, 10 846.

5 Analyses of NMR data (^{1}H-^{1}H COSY 45, HOHAHA, NOESY, DEPT, HMQC and HMBC) unequivocally supported the structures of **5**, **6**, **7**, **8** and **12**.

6 T. Sato, Y. Kanai and T. Hoshino, *Biosci. Biotechnol. Biochem.,* 1998, **62**, 407.

7 (*a*) P. Bouvier, Y. Berger, M. Rohmer and G. Ourisson, *Eur. J. Biochem.*, 1980, **112**, 549; (*b*) M. Renoux and M. Rohmer, *Eur. J. Biochem.*, 1986, **155**, 125.

8 Given that a water molecule attacks in the axial direction, a more hindered interaction would occur between the LUMO at C17–C18 and the HOMO at C21–C22 due to constrained overlapping (carbon numbering shown in Scheme 2).

9 R. B. Clayton, E. E. van Tamelene and R. G. Nadeau, *J. Am. Chem. Soc.*, 1968, **90**, 820.

Communication 9/01351B