Preprint Title	Unraveling the Role of Prenyl Side-Chain Interactions in Stabilizing the Secondary Carbocation in the Biosynthesis of Variexenol B
Authors	Moe Nakano, Rintaro Gemma and Hajime Sato
Publication Date	04 Juli 2023
Article Type	Letter
Supporting Information File 1	SI_variexenol.pdf; 2.7 MB
ORCID® iDs	Hajime Sato - https://orcid.org/0000-0001-5185-096X

Unraveling the Role of Prenyl Side-Chain Interactions in Stabilizing the Secondary Carbocation in the Biosynthesis of Variexenol B

Moe Nakano[1], Rintaro Gemma[1], and Hajime Sato[1,2,*]

Address: [1] Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, 4-4-37 Takeda, Kofu, Yamanashi 400-8510, Japan and [2] PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332–0012.

Email: hsato@yamanashi.ac.jp

* Hajime Sato

Abstract

Terpene cyclization reactions involve a number of carbocation intermediates. In some cases, these carbocations are stabilized by through-space interactions with π orbitals. Several terpene/terpenoids, such as sativene, satalene, bergamotene, ophiobolin and mangicol, possess prenyl side-chains that do not participate in the cyclization reaction. The role of these prenyl side-chains has been partially investigated, but remains elusive in the cyclization cascade. In this study, we focus on variexenol B that are synthesized from iso-GGPP, as recently reported by Dickschat and co-workers, and investigate the possibility of through-space interactions with prenyl side-chains using DFT calculations. Our calculations show that (i) unstable secondary cation is stabilized by the cation-π interaction from prenyl side-chains, thereby lowering the activation

energy, (ii) the four-membered ring formation is completed through bridging from the exomethylene, and (iii) the annulation from the exomethylene proceeds in a barrierless manner.

Keywords

DFT, carbocation, cation-π interaction, terpene, biosynthesis

Introduction

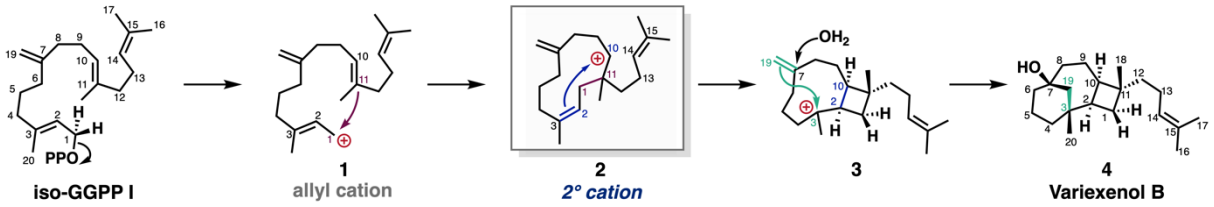
Terpene/terpenoids are most aboundant natural product in nature. There have been reported over 80,000 terpenoids compounds to date.[1-3] One of the most intriguing point is that all diversified strutures are synthesized from common starting material, isoprenoids. Reactions that generate complex cyclic structures and multiple stereogenic centers from linear achiral precursors offer many valuable insights from a fundamental organic chemistry perspective.

Terpene cyclization cascade generally involves the multi-step domino type reaction. Therefore, it is challenging to reveal the detailed reaction mechanism solely by the experimental method. To address this issue, computational chemistry including DFT[4-8], QM/MM[9-15] and QM/MM MD[16] calculations have been used for the biosynthetic studies of terpene/terpenoids.[17]

Terpene-forming reactions, which involve various types of carbocation species stabilized by hyperconjugative interactions, through-space interactions, and C-H π interactions, have been intensively investigated by Tantillo and co-workers, who have contributed greatly to revealing the intriguing nature of carbocations. [6,18,19]

We have also elucidated various new insights of carbocation chemistry, such as the C–H π interaction between carbocation intermediate and Phe residue of terpene

cyclase in sesterfisherol biosynthesis [20], and the intricated rearrangement reaction mechanism promoted by the equilibrium state of homoallyl cation and cyclopropylcarbinyl cation in trichobrasilenol biosynthesis [21], by combinined methods of computational and experimental chemistry.



Scheme 1: Proposed Biosynthetic Pathway for Variexenol B.

Recently, Dickschat et al. reported the synthesis of novel diterpene compounds, variexenol B, using a substrate analogue called iso-GGPP (Scheme 1).[22] This biosynthetic pathway has two interesting aspects. First, this cyclization cascade involves prenyl side-chain that do not participate in the cyclization cascade. This type of terpene compounds have been reported, such as santalene, bergamotene, mangicol, etc. The idea that the reaction mechanism changes due to differences in prenyl side-chains has been studied by Tantillo and co-workers.[23-25] They reported that the carbocation intermediates traversed in the biosynthesis of pinene/camphene and ylangene/sativene change based on the presence or absence of prenyl side-chains. In their study, it was argued that the extent of hyperconjugation determines whether the reaction proceeds in a stepwise or concerted manner.

The second interesting aspect of variexenol B biosynthesis is that it has exomethylene. Terpene with exomethylene as a starting material is rare. Several terpene cyclizations with exomethylene have been known, such as with caryolene and crotinsulidane

diterpenoids, and the reaction mechanisms have been analyzed.[26-29] It would be interesting to see how exomethylene reacts in the cyclization of variexenol B.

In this study, we investigated the biosythetic pathways using DFT calculation to validate the above-mentioned aspects.

Results and Discussion

Through our computational analysis, the detailed structures of the intermediates and transition states were elucidated. Interestingly, we have found the interaction between secondary carbocation and side prenyl side-chain. Figure 1 shows the computed biosynthetic pathway and energy diagram without cation-π interaction, while Figure 2 shows the computed biosynthetic pathway including cation-π interaction from the prenyl side-chain.

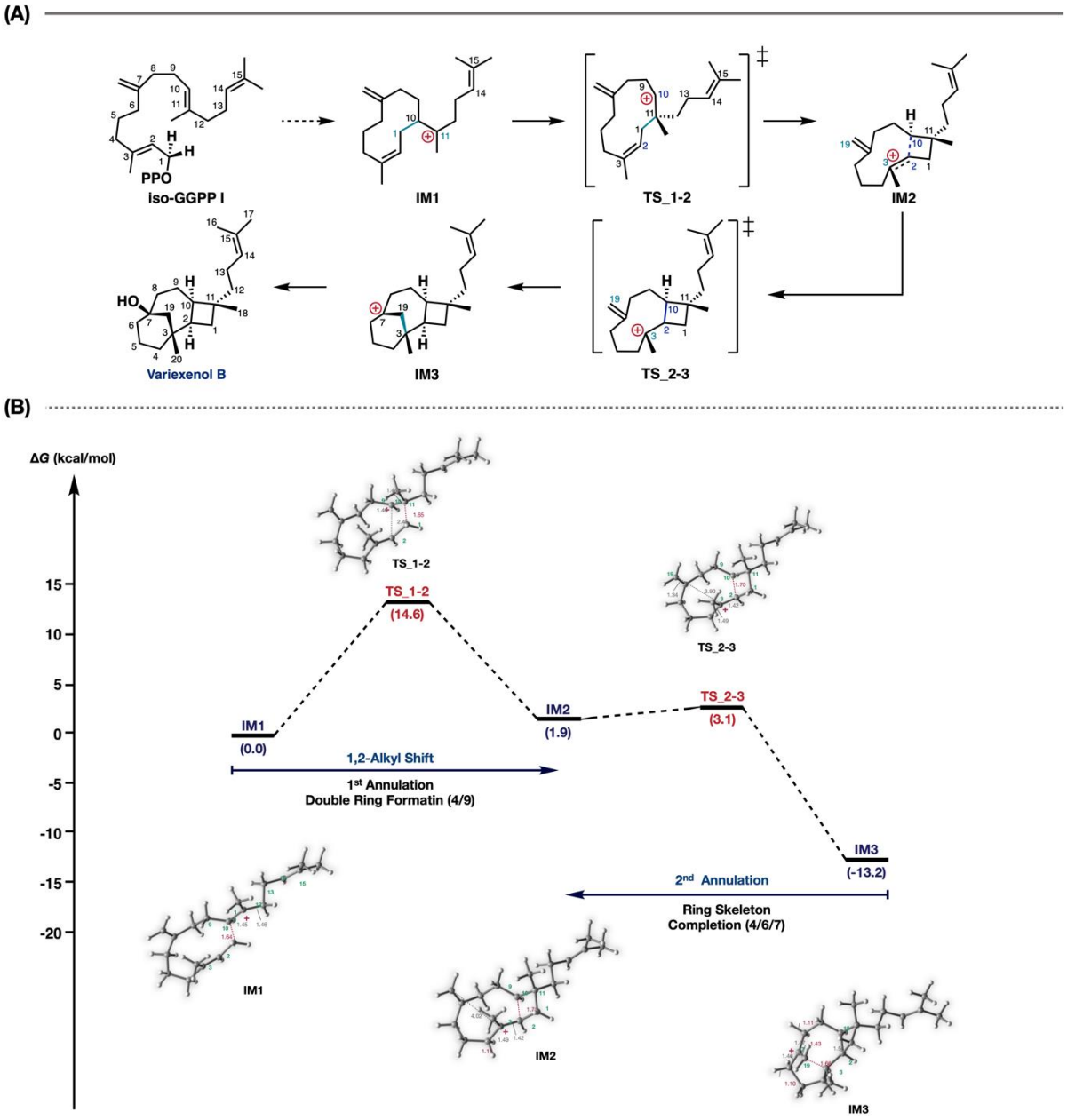


FIgure 1: (A) Results of DFT Evaluation of the Whole Pathway of Variexenol B Without Cation-π Interaction. (B) Energy Diagram of Variexenol B without cation-π interaction. IM stand for intermediate and TS stands for transition state. Potential energies (kcal/mol, Gibbs free energies calculated at the mPW1PW91/6-31+G(d,p)//M06-2X/6-31+G(d,p) level) relative to **IM1** are shown in parentheses.

Our research began with the application of DFT calculations to the putative biosynthetic pathway of variexenol B (Figure 1). It was revealed that variexenol B biosynthetic pathway undergoes a two-step reaction process. Contrary to putative biosynthetic pathway, the formation of the C1–C11 and C2–C10 bonds was found to be concerted, due to the formation of a secondary carbocation at the C10 position. Then, the tertiary carbocation formed at the C3 position undergoes virtually barrierless cyclization from the exomethylene to yield **IM3**.

We next investigated the effect of the prenyl side-chain in variexenol B biosynthesis. Although several terpene compounds with prenyl side-chains have been reported, it remains unclear whether these prenyl side-chains are located inside or outside the active site during the cyclization process. Therefore, we searched for conformations in which the side chain is closer to the carbocation center and performed calculations.

It was found that the structure with the prenyl side-chain containing the C14=C15 double bond positioned inwards was more advantageous than the pathway shown in Figure 1. Calculations based on the specified structure are shown in Figure 2. In this pathway, the C14=C15 double bond interacts with the secondary carbocation at C10, reducing the activation energy of the first step by approximately 4.7 kcal/mol. Moreover, due to the stabilization of the secondary carbocation-like intermediate **IM2**, the reaction proceeds stepwise rather than concertedly. It was found that the final cyclization reaction from the exomethylene proceeds without a barrier, similar to the previous pathway.

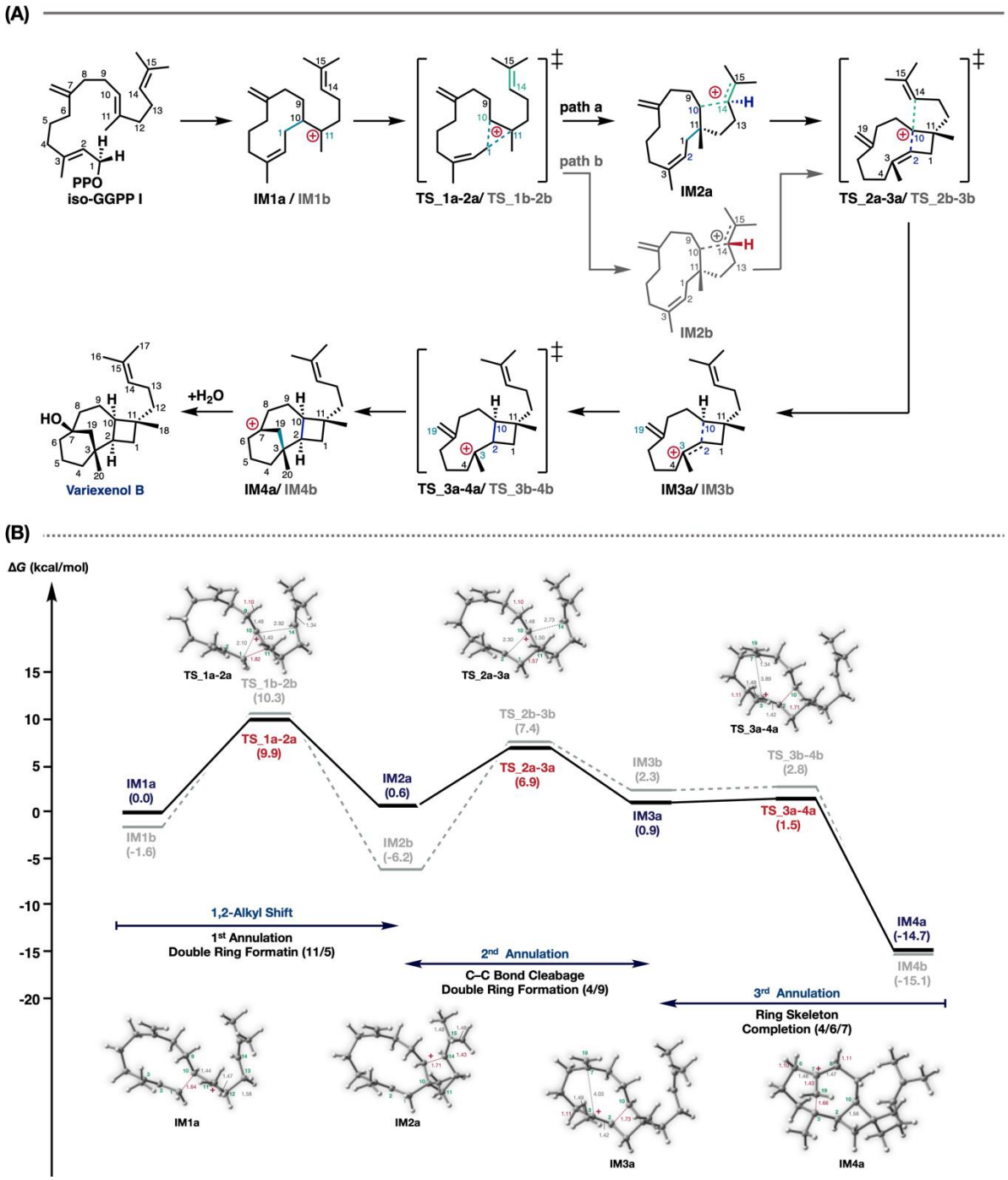


FIgure 2: (A) Results of DFT evaluation of the whole pathway of variexenol B including cation-π interaction from the prenyl side-chain. Path a has an α-hydrogen at the C14 position in **IM2**, while path b has the opposite orientation. (B) Energy Diagram of Variexenol B with consideration of cation-π interaction. Potential energies (kcal/mol, Gibbs free energies calculated at the mPW1PW91/6-31+G(d,p)//M06-2X/6-31+G(d,p) level) relative to **IM1** are shown in parentheses.

Regarding the orientation of the prenyl side-chain, two pathways can be considered depending on whether the hydrogen at C14 is pointing; α-hydrogen (path a) or β-hydrogen (path b). Both pathways follow similar reaction mechanisms. However, when comparing path a and path b, the most striking energy difference is in the step from **IM2a/b** to **IM3a/b** (Figure 2B). The energy barrier of this step is 6.3 kcal/mol for path a, whereas 13.6 kcal/mol for path b, with a difference of 7.3 kcal/mol. Although the stabilization of the intermediate **IM2b** is greater in path b, the activation energy suggests that path a is more favorable.

Generally, the activation energies for terpene cyclization reactions are often below 10 kcal/mol. However, in the case of complex rearrangement reactions involving secondary carbocations, which we recently discovered, reactions with activation energies around 16 kcal/mol have been reported.[21] In the pathway shown in Figure 1, the highest energy barrier was 14.6 kcal/mol. Conversely, in Figure 2, path a had an energy barrier of 9.9 kcal/mol and path b 13.6 kcal/mol. From these results, it can be concluded that although all three pathways have the potential to advance the reaction, the most energetically favorable pathway is path a, as shown in this study.

To our knowledge, there have not been reported the interaction from the prenyl side-chain towards the carbocation center. Systems with secondary carbocations on rings bearing prenyl side-chains are commonly observed in steroid biosynthesis. These type of cyclization reactions have been vigorously studied by Hess [30-35] and Wu [36,37]. In these systems, the secondary carbocation and the double bond of the neighboring prenyl side-chain interacts and promptly induce C–C bond formation. There have been no reports of results where, as in our case, the cation is stabilized without bond formation.

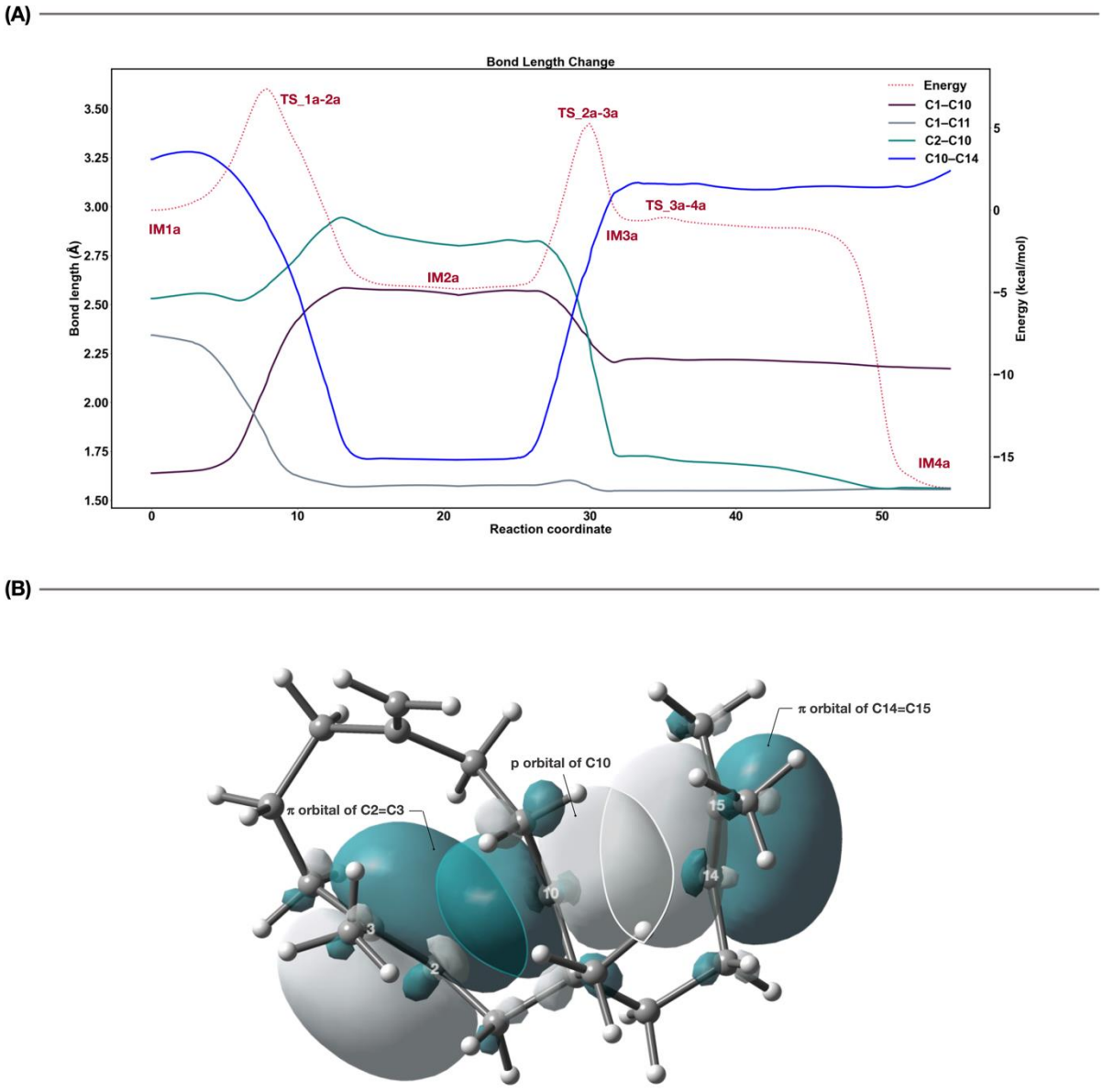


Figure 3: (A) A Representative Example of the Evolution of Key Bond Lengths in the Conversion of path a. (B) Key Representative Orbitals of **TS_2a–3a** computed by DFT calcualtions.

To investigate the details of carbocations and hyperconjucations in variexenol B biosynthetic pathway, we carried out the bond length change analysis on the bonds that contribute most to the reaction from **IM1a** to **IM4a** (figure 2).

In the process from **IM1a** to **TS_1a–2a**, the C1–C10 bond ruptures as C1 shifts towards C11. Subsequently, in the transition towards **IM2a**, a complete formation of the C1–

C11 bond occurs. At this point, the vacant orbital of the carbocation at C10 interacts with the π orbital of the C14=C15 double bond.

The distance between C10 and C14 is 1.71 Å, which is hardly to recognize as a single C–C bond, since the distance is greatly elongated. moreover, the bond length of C14=15 is 1.43Å, which is close to the double bond length. Judging from the bond length alone, it is not impossible to conclude that the C10–C14 bond is formed, but considering the rational mechanism of organic reactions, bond cleavage does not occur immediately after the bond is formed.

On the other hand, the C10–C14 bond length of **IM2b** is 1.64Å, which is the bond length when hyperconjugated, which is commonly observed in terpene-forming reactions. This relatively short bond length appears to contribute the stability of **IM2b**. This energy difference between **IM2a** and **IM2b** appears to be due to small differences in conformation caused by the stereochemistry of H14.

In **TS_2a–3a**, the C10 secondary carbocation is stabilized and sandwiched between the two π orbitals of C2=C3 and C14=C15. The status of this orbital interaction is depicted in Figure 3B. This interaction forms the C2–C10 bond and the reaction proceeds to **IM3a**.

Regarding the 4-membered ring formation, the C2–C10 bond in **IM3** is 1.73 Å, which is hard to recognize as a single bond. However, the C2–C10 bond is 1.56 Å and the 4-membered ring bond is completed (Figure 3B) because the hyperconjugation effect is eliminated by the removal of the C3 carbocation by annulation from exomethylene. Based on the key bond analysis, we have successfully elucidated the details of variexenol B biosynthesis.

Conclusion

In conlcusion, we have investigated the detailed reaction mechanism of the variexenol biosynthesis. We have revealed three new insights (i) the possibility of stabilization of the secondary carbocation by prenyl side-chain of the intermediate, (ii) four membered ring formation is completed by the bridging reaction, and (iii) annulation from the exomethylene is the barrierless process.

To date, when constructing the computational model, we have sometimes truncated the prenyl side-chains that do not participate in the cyclization cascade in order to reduce the computational cost.[34,39] However, as demonstrated in this study, the possibility of cation-π interactions lowering the activation energy of annulation requires caution when constructing computational models in the future.

Furthermore, future research is expected to determine whether there is space in the enzyme active site for these prenyl side-chains to fold and approach the reaction center, as seen in X-ray crystallographic analysis.

Experimental

All calculations were carried out using the Gaussian 16 package.[39] Structure optimizations were done with the M06-2X [40] density functional theory method and the 6-31+G(d,p) basis set without any symmetry restrictions. M06-2X was selected because of its accuracy in calculating terpene-forming reaction and its proven track record of being used in previous studies of reaction mechanism analysis.[41,42] Vibrational frequency calculations at the same level of theory with optimization were performed to verify that each local minimum has no imaginary frequency and that each **TS** has only a single imaginary frequency. Conformational search was done with conflex program.[43-45] Intrinsic reaction coordinate (IRC) calculations [46-49] for all

TSs were performed with GRRM11[50] based on Gaussian 16. Single-point energies were calculated at the mPW1PW91/6-31+G(d,p) level based on the optimized structure by the M06-2X method. The utility of relative Gibbs free energy energies (G_{rel}) based on single-point energy at the mPW1PW91 level has been previously validated for a wide variety of terpene-forming reactions.[21,38,51]

Supporting Information

Supporting information features IRC plot, 3D representations of all computed structures, cartesian coordinates, energies, and imaginary frequencies.

Acknowledgements

This work was supported by a JSPS KAKENHI Grant-in-Aid for Early-Career Scientists (No. 22K14791(H.S.)), a JSPS KAKENHI Grant-in-Aid for Transformative Research Areas (No. 22H05125 (H.S.)), a MEXT grant for Leading Initiative for Excellent Young Researchers (No. JPMXS0320200422 (H.S.)), JST PRESTO (No. JPMJPR21D5 (H.S.)), the Uehara Memorial Foundation (No. 202110117 (H.S.)), the Terumo Life Science Foundation (No. 21-III4030 (H.S.)), Inamori foundation, and Astellas Foundation for Research on Metabolic Disorders (H.S.). We would like to express our deepest gratitude to Dr. Makoto Obata for his insightful feedback and continuous mentorship that significantly contributed to the success of this research.

References

1. Rudolf, J. D.; Alsup, T. A.; Xu, B.; Li, Z. *Nat. Prod. Rep.* **2021**, *38*, 905–980..

2. Christianson, D. W. *Chem. Rev.* **2017**, *117*, 11570–11648.

3. Dickschat, J. S. *Nat. Prod. Rep.* **2016**, *33*, 87–110.

4. Tantillo, D. J. *Nat. Prod. Rep.* **2011**, *28*, 1035–1053.

5. Tantillo, D. J. *Nat. Prod. Rep.* **2013**, *30*, 1079–1086.

6. Tantillo, D. J. *Chem. Soc. Rev.* **2010**, *39*, 2847–2854.

7. Tantillo, D. J. *Angew. Chem. Int. Ed.* **2017**, *56*, 10040–10045.

8. Hong, Y. J.; Tantillo, D. J. *Chem. Soc. Rev.* **2014**, *43*, 5042–5050.

9. Weitman, M.; Major, D. T. *J. Am. Chem. Soc.* **2010**, *132*, 6349–6360.

10. Major, D. T.; Weitman, M. *J. Am. Chem. Soc.* **2012**, *134*, 19454–19462.

11. Raz, K.; Levi, S.; Gupta, P. K.; Major, D. T. *Curr. Opin. Biotech.* **2020**, *65*, 248–258.

12. Das, S.; Shimshi, M.; Raz, K.; Eliaz, N. N.; Mhashal, A. R.; Ansbacher, T.; Major, D. T. *J. Chem. Theory Comput.* **2019**, *15*, 5116–5134.

13. Ansbacher, T.; Freud, Y.; Major, D. T. *Biochemistry* **2018**, *57*, 3773–3779.

14. Dixit, M.; Weitman, M.; Gao, J.; Major, D. T. *ACS Catal.* **2018**, *8*, 1371–1375.

15. Dixit, M.; Weitman, M.; Gao, J.; Major, D. T. *ACS Catal.* **2017**, *7*, 812–818.

16. Lou, T.; Li, A.; Xu, H.; Pan, J.; Xing, B.; Wu, R.; Dickschat, J. S.; Yang, D.; Ma, M. *J. Am. Chem. Soc.* **2023**, *145*, 8474–8485.

17. Sato, H.; Saito, K.; Yamazaki, M. *Front. Plant Sci.* **2019**, *10*, 802.

18. Hong, Y. J.; Tantillo, D. J. *Chem. Sci.* **2013**, *4*, 2512–2518.

19. Hong, Y. J.; Tantillo, D. J. *Org. Lett.* **2015**, *17*, 5388–5391.

20. Sato, H.; Narita, K.; Minami, A.; Yamazaki, M.; Wang, C.; Suemune, H.; Nagano, S.; Tomita, T.; Oikawa, H.; Uchiyama, M. *Sci. Rep.* **2018**, *8*, 2473.

21. Sato, H.; Hashishin, T.; Kanazawa, J.; Miyamoto, K.; Uchiyama, M. *J. Am. Chem. Soc.* **2020**, *142*, 19830–19834.

22. Li, H.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2022**, *61*, e202211054.

23. Hong, Y. J.; Tantillo, D. J. *Chem. Commun.* **2012**, *48*, 1571–1573.

24. Hong, Y. J.; Tantillo, D. J. *Org. Biomol. Chem.* **2010**, *8*, 4589–4600.

25. Zi, J.; Matsuba, Y.; Hong, Y. J.; Jackson, A. J.; Tantillo, D. J.; Pichersky, E.; Peters, R. J. *J. Am. Chem. Soc.* **2014**, *136*, 16951–16953.

26. Nguyen, Q. N. N.; Tantillo, D. J. *Beilstein J. Org. Chem.* **2013**, *9*, 323–331.

27. Ortega, D. E.; Nguyen, Q. N. N.; Tantillo, D. J.; Toro-Labbé, A. *J. Comput. Chem.* **2016**, *37*, 1068–1081.

28. Campos, R. B.; Tantillo, D. J. *J. Org. Chem.* **2018**, *83*, 1073–1076.

29. Hong, Y. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2010**, *132*, 5375–5386.

30. Hess, B. A. *European J. Org. Chem.* **2004**, *2004*, 2239–2242.

31. Smentek, L.; Hess, B. A., Jr *J. Am. Chem. Soc.* **2010**, *132*, 17111–17117.

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

33. Hess, B. A., Jr *Org. Lett.* **2003**, *5*, 165–167.

34. Hess, B. A., Jr; Smentek, L. *Org. Lett.* **2004**, *6* (11), 1717–1720.

35. Hess, B. A., Jr; Smentek, L. *Angew. Chem. Int. Ed.* **2013**, *52* (42), 11029–11033.

36. Chen, N.; Wang, S.; Smentek, L.; Hess, B. A., Jr; Wu, R. *Angew. Chem. Int. Ed..* **2015**, *54* (30), 8693–8696.

37. Zhang, F.; Wang, Y.; Yue, J.; Zhang, R.; Hu, Y.-E.; Huang, R.; Ji, A.-J.; Hess, B. A., Jr; Liu, Z.; Duan, L.; Wu, R. *Sci Adv* **2023**, *9*, eadh1418.

38. Sato, H.; Li, B. X.; Takagi, T.; Wang, C.; Miyamoto, K.; others, *JACS Au* **2021**, *1*, 1231–1239.

39. Gaussian 16, Revision C.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2016.

40. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

41. Dixit, M.; Weitman, M.; Gao, J.; Major, D. T. *ACS Catal.* **2017**, *7*, 812–818.

42. Zev, S.; Gupta, P. K.; Pahima, E.; Major, D. T. *J. Chem. Theory Comput.* **2022**, *18*, 167–178.

43. Goto, H.; Osawa, E. *J. Am. Chem. Soc.* **1989**, *111*, 8950–8951.

44. Goto, H.; Osawa, E. *J. Chem. Soc., Perkin Trans. 2* **1993**, 187–198.

45. CONFLEX 9, H. Goto, S. Obata, N. Nakayama and K. Ohta, CONFLEX Corporation, Tokyo, Japan, **2021**.

46. Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363–368.

47. Page, M.; Doubleday, C.; Jr., J. W. M. *J. Chem. Phys.* **1990**, *93*, 5634–5642.

48. Ishida, K.; Morokuma, K.; Komornicki, A. *J. Chem. Phys.* **1977**, *66*, 2153–2156.

49. Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523–5527.

50. Maeda, S.; Ohno, K.; Morokuma, K. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683–3701.

51. Matsuyama, T.; Togashi, K.; Nakano, M.; Sato, H.; Uchiyama, M. *JACS Au* **2023**, *3*, 1596–1603.