Predicting Rearrangement-Competent Terpenoid Oxidation Levels

Christina H. McCulley, Dean J. Tantillo*

Department of Chemistry, University of California – Davis, Davis, CA USA

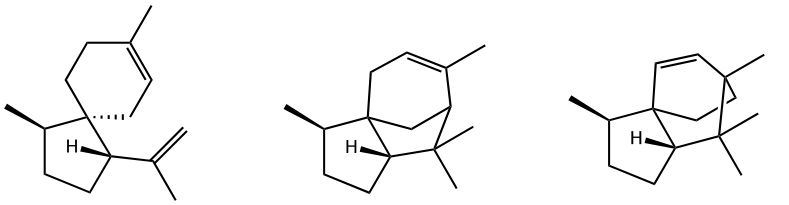
Abstract

Results of density functional calculations on rearrangements of potential biosynthetic precursors to the sesquiterpenoid illisimonin A reveal that only some possible precursors, those with certain specific oxidation patterns, are rearrangement-competent.

Introduction

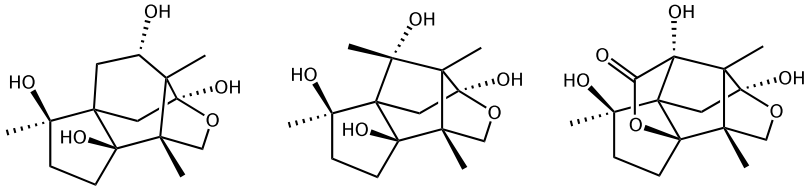
The diversity of Nature's library of terpenoid natural products arises, for the most part, from two sources. First, acyclic, achiral metabolic intermediates—oligoisoprenyl diphosphates—are converted into terpenes, which are often polycyclic and rich with stereogenic centers (Figure 1a), by terpene synthase-mediated carbocation cyclization/rearrangement reactions.[1] Hundreds of possible hydrocarbon skeletons are known for sesquiterpenes, for example. Second, heteroatoms are added—most often oxygen atoms—by enzymes such as cytochrome P450s, to produce terpenoids (Figure 1b).[2] In some cases, oxidation is accompanied by further rearrangement.[3] An example of a skeletal rearrangement that may occur after oxidation is discussed here.

(a) representative terpenes



1 **2** **3**

(b) representative terpenoids

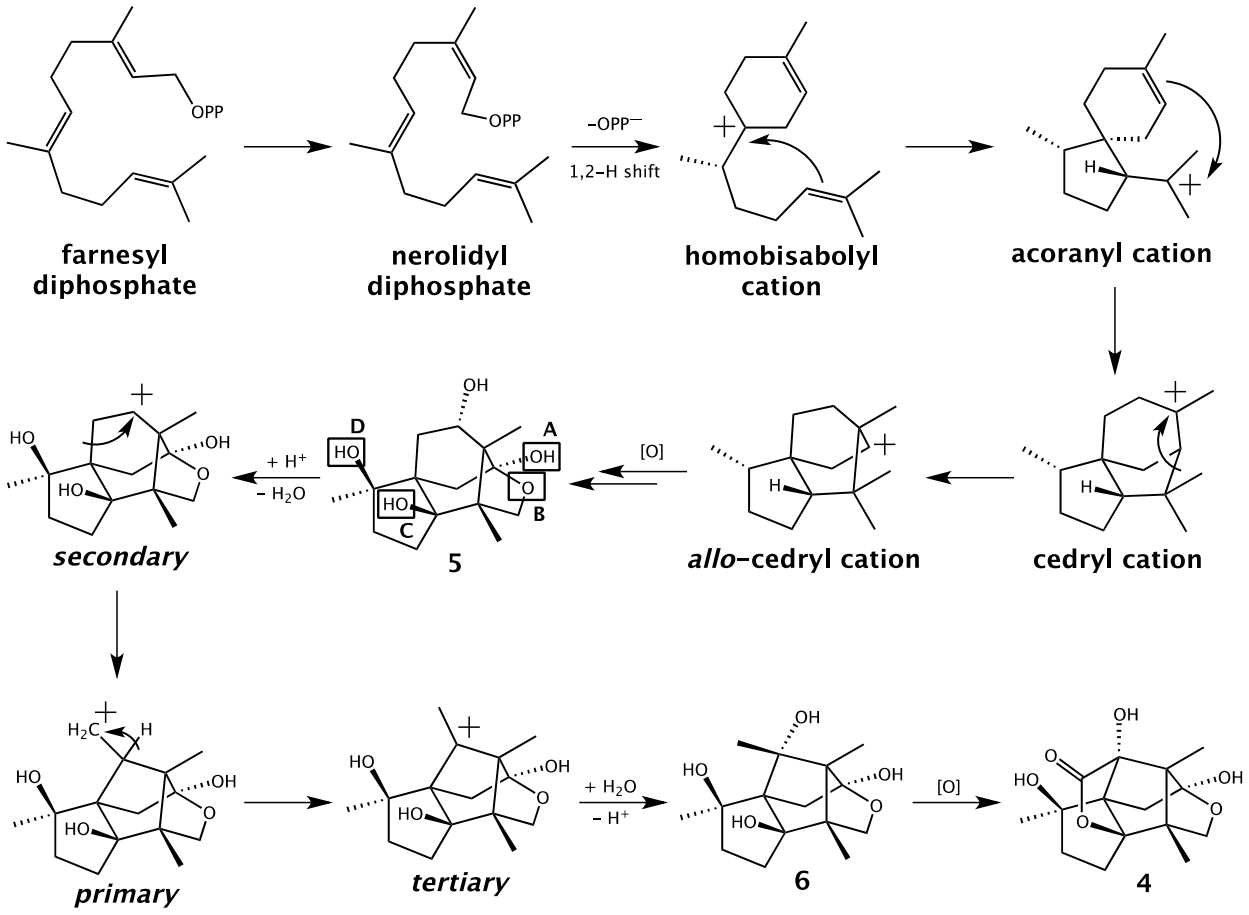


5 **6** **4**

Figure 1. Representative terpene and terpenoid structures, here sesquiterpenes and sesquiterpenoids: acorene (**1**), cedrene (**2**) and *allo*-cedrane (**3**, also known as khusiene).[5]

The authors of the paper describing the isolation of illisimonin A (**4**, Figure 1) proposed that a skeletal rearrangement occurs from a highly oxidized form of *allo*-cedrane, **5→6** (Scheme 1).[4] We report the results of density functional theory (DFT) computations used to probe the feasibility of rearrangement at this oxidation level and others.[6] These results indicate that rearrangement from a less (but still) oxidized structure is accompanied by a lower overall barrier. We hope that our approach will find utility in predicting rearrangement-competent oxidation levels for other terpenoids, which could facilitate both biosynthetic experiments and total synthesis efforts.[7]

Scheme 1. Rearrangements proposed in ref. 4 (stereochemistry and nerolidyl diphosphate added).



Methods

Computations were performed at the mPW1PW91/6-31+G(d,p) level of theory, a level of theory used previously to model many carbocation rearrangements,[6,8] using *Gaussian09*.[9] All minima and transition state structures (TSSs) were confirmed by frequency analysis, and intrinsic reaction coordinate (IRC) calculations were performed on unconstrained transition state structures to confirm the minima to which they are connected.[10] All energies shown are free energies at room temperature, except for those in IRC plots, which are electronic energies. Three-dimensional molecular images were generated with *CylView*.[11]

Results and Discussion

The previously proposed rearrangement

The secondary→primary→tertiary carbocation rearrangement mechanism in Scheme 1 was the focus of our attention, as it is exceedingly unlikely that a primary carbocation is a minimum.[12] For the originally proposed system, two epimers of the primary carbocation are possible. Both were found to be transition state structures (TSSs), not minima (Figure 2). In one case (Figure 2, top), the core H–C–C substructure (top left in each ball-and-stick image) is partly bridged, while in the other (Figure 2, bottom) an intramolecular CH–O interaction[13] provides an alternative to bridging. These modes of delocalization are representative of the products connected to each TSS: bridging leads to a 1,2-hydride shift to form the tertiary carbocation originally proposed as an intermediate (Figure 2, top), while the CH–O interaction presages an intramolecular proton transfer to form a C=C double bond. In neither case, however, is a simple secondary carbocation (Scheme 1) found to connect to the TSS in the reactant direction.[6,14] Instead, oxygen-bridged structures are found – in one case (Figure 2, top) bridging to the site of the putative primary carbocation and in the other case (Figure 2, bottom) bridging to the site of the putative secondary carbocation connected to the TSS by a 1,2-alkyl shift (see the Supporting Information for IRC plots).[15] Thus, for neither configuration is the primary carbocation a minimum, nor is it connected to the previously proposed secondary and tertiary carbocations, i.e., the two proposed steps are not simply merged into a concerted process as has been observed for many other carbocation rearrangements.[6,12,14] In any case, the barriers for rearrangement associated with the TSSs in Figure 2 are predicted to be too high to be relevant to biology without considerable intervention (on the order of 10 kcal/mol barrier lowering),[16] so other oxidation levels were examined.

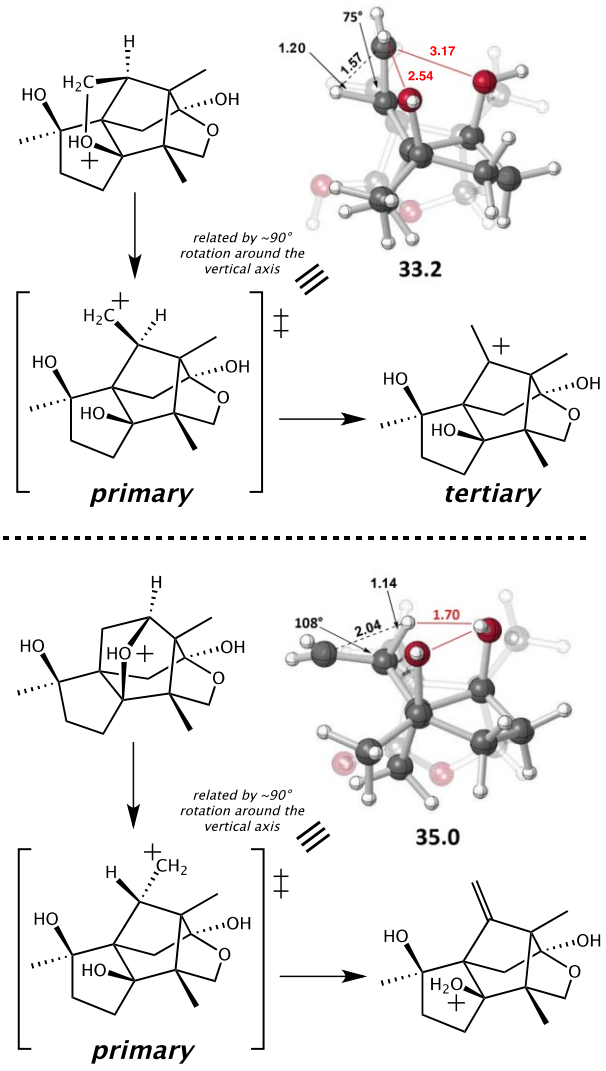


Figure 2. Rearrangements found for the fully substituted system (oxygens at positions A, B, C and D are shown), i.e., cations derived from **5**. Ball-and-stick images of the computed transition state structures (mPW1PW91/6-31+G(d,p)) resembling primary carbocations for the two epimers. Selected distances are shown in Å. Predicted rearrangement barriers ($\Delta G^{\ddagger}$) in kcal/mol are listed below each transition state structure.

The hydrocarbon case

Carbon skeletons of terpenoids are usually constructed at the hydrocarbon stage, i.e., through the action of synthases/cyclases that produce terpenes,[1,3] thus we looked for an energetically viable pathway that involved non-oxygenated carbocations preceding **5** in Scheme 1. Again, we began by locating primary carbocation-like structures for two epimers at the

stereogenic center next to the site of the putative primary carbocation. In both cases, primary carbocation-like TSSs were again found (Figure 3, first column) and again, these were associated with high rearrangement barriers (here, approximately 40 kcal/mol). Note that these TSSs lack the possibility of intramolecular CH–O or O–C^{+} interactions. One of these TSSs is connected to a cedryl cation (an epimer of that shown in Scheme 1) and the sans-oxygen version of the originally proposed tertiary carbocation (Figure 4). The other TSS is connected to a prezizyl cation[17] and the sans-oxygen version of the originally proposed tertiary carbocation (Figure 5).[18]

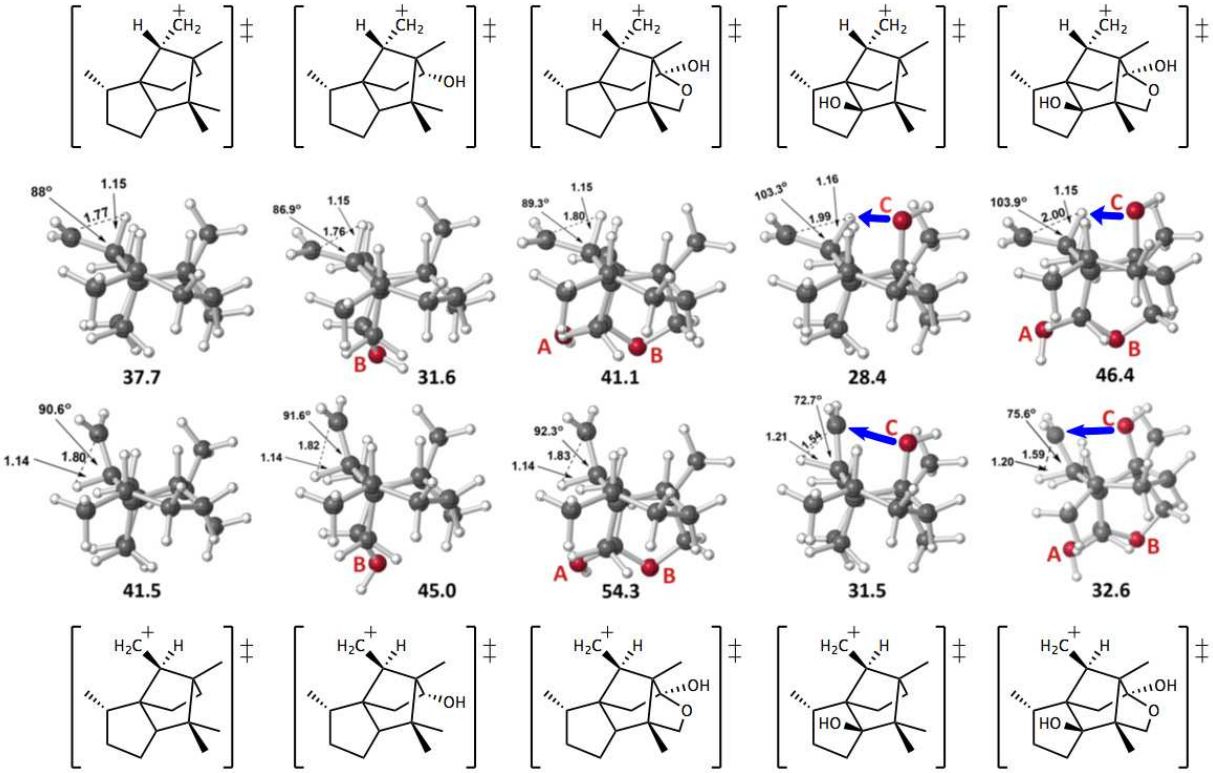


Figure 3. Representative transition state structures with various carbons oxidized (see SI for others), oriented to show $^{+}$C–C–H substructures (top left of each ball-and-stick image) side-on to facilitate comparisons. Predicted activation barriers ($\Delta G^{\ddagger}$) are shown below each in kcal/mol. Selected distances are shown in Å. Key CH–O and O–C^{+} interactions are highlighted in blue.

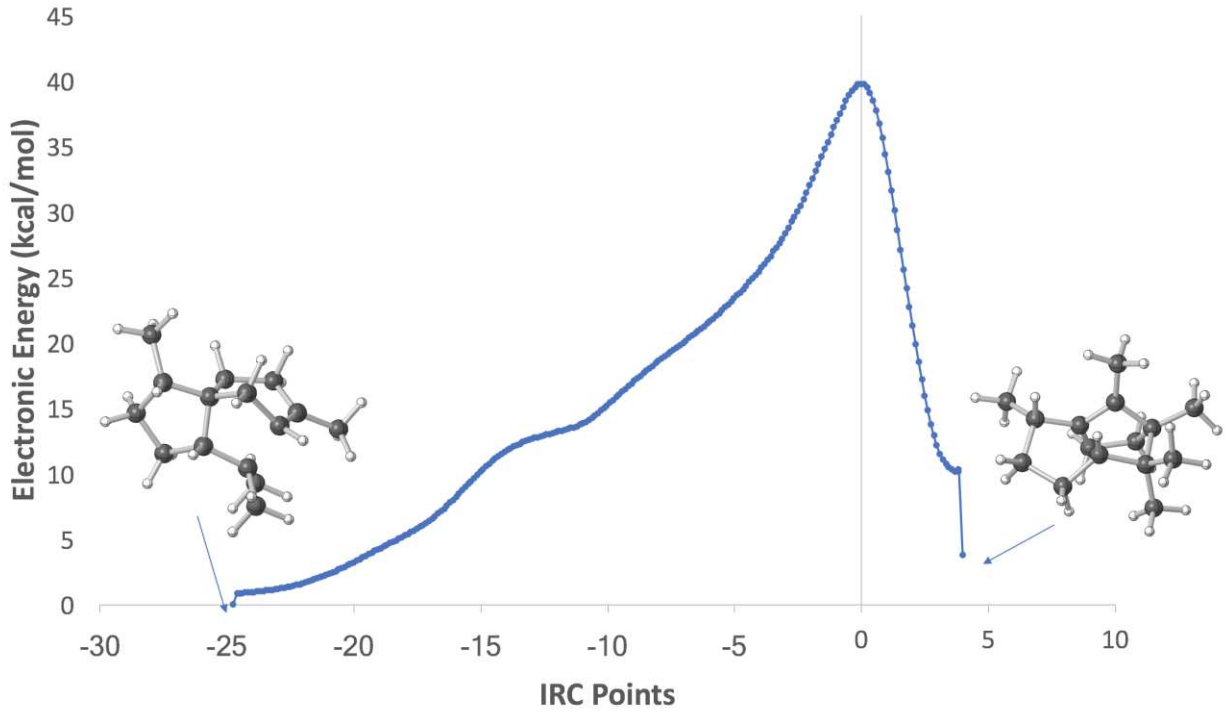


Figure 4. IRC plot for rearrangement of cedryl cation epimer.

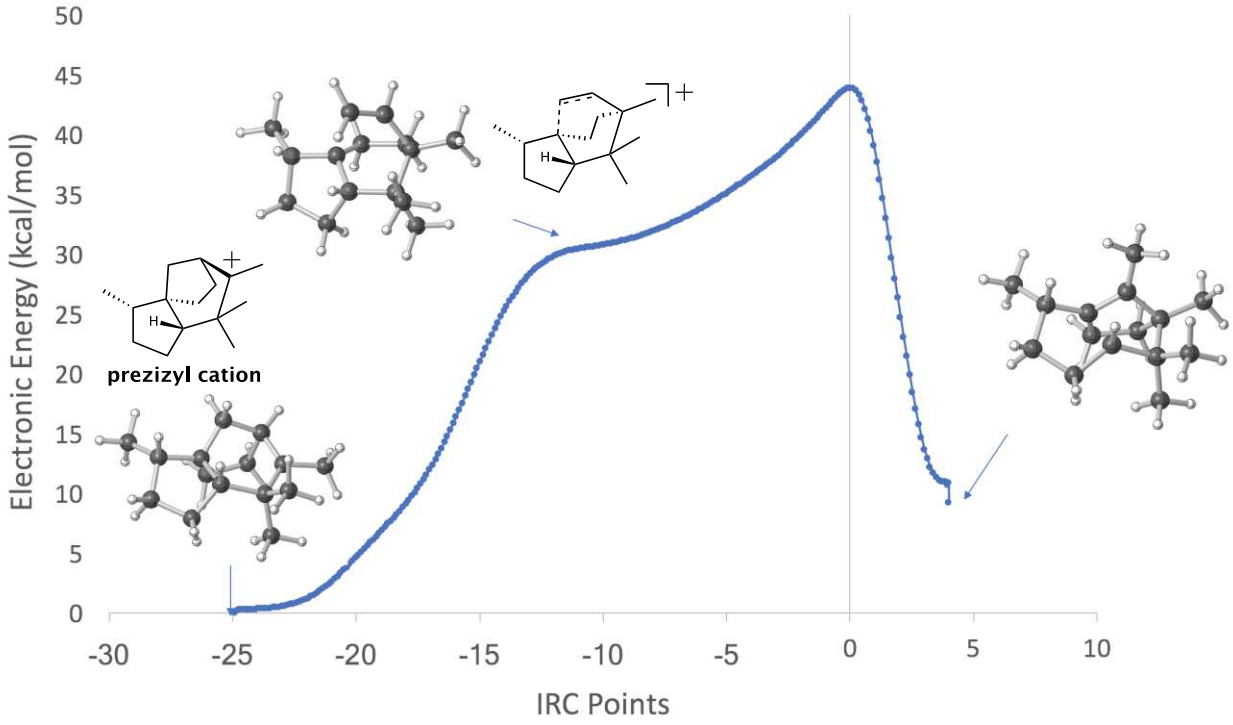


Figure 5. IRC plot for rearrangement of prezizyl cation.

Varying the oxidation level

Carbocations with a variety of alternative oxidation levels were explored in hopes that one or more would allow for an energetically viable rearrangement of the carbon skeleton. Oxygens were added at positions A, B and C (as labeled in Scheme 1), in various combinations. All TSSs for the rearrangements examined (Figure 3) contain a hyperconjugated primary carbocation center (some on the verge of H-bridging, e.g., two bottom right structures in Figure 3). However, these TSSs connect minima with a variety of polycyclic skeletons (see SI for details). For systems without an oxygen at position C, much higher barriers are predicted for migration of hydrogen to the bottom face (as drawn in Figure 3) of the carbocation center,[19] and the difference in predicted barriers increases as more oxygens are added. While it is difficult to pin down the origins of these barrier differences, since different types of reactant minima are connected to very similar TSSs, it is observed that in the preferred TSSs, the C–H bond α to the primary carbocation center is slightly longer and its hydrogen is slightly closer to the CH_2 group, i.e., it is closer to bridging. Having an OH group at position C (Figure 3, column 4) leads to the lowest barrier, ~28 kcal/mol, likely due in part to the presence of a favorable internal CH–O interaction between the hyperconjugated H and the OH group (~1.7 Å; note that the epimeric TSS also has a favorable interaction, in that case between the carbocation center and an oxygen lone pair; see SI for additional details on these interactions). This TSS leads to an epimer (at the spiro center) of an acorenyl cation as reactant and an alkene resulting from intramolecular proton transfer (analogous to the reaction shown at the bottom of Figure 2) as product. Structures with an OH at position C and additional OH groups did not lead to lower barriers, likely the result of differences in strain and additional avenues for selective reactant stabilization (e.g., Figure 2).

Theozymes

In hopes that noncovalent interactions with an enzyme active site might lower the rearrangement barrier into a reasonable range (~20 kcal/mol),[16] we undertook theozyme calculations.[20] In these, structures were reoptimized in the presence of various combinations of benzene, indole, phenol and formamide – mimics of sidechains of phenylalanine, tryptophan, tyrosine and asparagine/glutamine (or backbone amides), respectively (see SI for details), on the

system from Figure 3 with the lowest barrier (Figure 6, bottom). In many cases, alkylation of the residue models by the carbocations occurred, so constraints were added during optimization to mimic the restrictions that could be imposed by an enzyme active site.[21] The system for which the lowest barrier was found is shown at the top of Figure 6: formamide/phenol with the primary carbocation center constrained to be 2.92 Å from one *ortho* carbon of the phenol (corresponding to a reasonable carbocation–π distance).[22] The predicted rearrangement barrier for this system is only 12.6 kcal/mol (and the reaction is predicted to be exergonic by ~9 kcal/mol) and the reactant in this case is a hydroxyl-bearing *allo*-cedryl cation (Figure 6). Carbocation-π interactions[22] between the phenol and both the primary carbocation center and the hyperconjugated hydrogen, as well as a CH–O interaction[13] between one CH at the primary carbocation center and the amide oxygen (~2 Å) conspire to provide selective stabilization to the TSS, for which these groups bear the largest positive charge (NBO charges on H: +0.43 in TSS, +0.27 in reactant; see SI for additional details). While this barrier is derived from a theozyme with constraints, and while no enzyme for this reaction has actually been isolated, the arrangement shown is not unusual.[23] Theozymes may also lower the rearrangement barrier into the biologically reasonable range for other systems in Figure 3, but the system described here has the advantage of starting with the lowest inherent rearrangement barrier (by at least 3 kcal/mol; Figure 3). *Our proof-of-principle results show that rearrangement is at least possible and lead us to propose that one of the compounds shown in Figure 7 is likely the biosynthetic precursor to ilisimonin A.* Related oxidized *allo*-cedranes have been described previously.[4b,24]

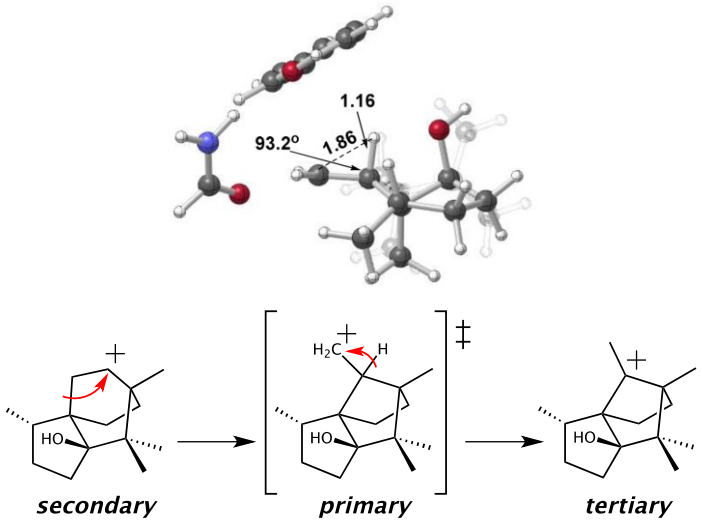


Figure 6. Transition state structure for the migration shown in the presence of a formamide/phenol theozyme. Selected distances are shown in Å.

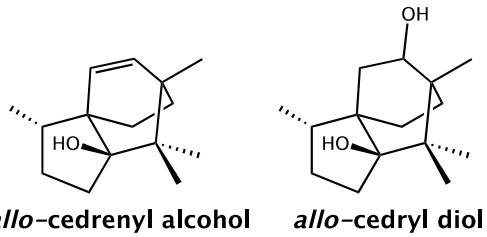


Figure 7. Proposed, rearrangement-competent, biosynthetic precursors to illisimonin A.

Conclusions

Our computational study has shed light on several aspects of illisimonin A biosynthesis: (1) Primary carbocations are almost certainly not involved as minima. (2) Rearrangement at the oxidation level of **5** would require an unreasonably large degree of enzymatic intervention. (3) Rearrangement at the hydrocarbon level would as well. (4) Some structures with intermediate oxidation levels have lower rearrangement barriers. (5) Specifically oriented noncovalent interactions can lower some barrier for these structures into the biologically reasonable range. We hope that an illisimonin A-producing enzyme will be isolated and characterized so that these predictions can be tested. We also hope that the approach described herein will find utility in predicting rearrangement-capable oxidation levels for structures involved in the biosynthesis (or synthesis) of other natural products.

ASSOCIATED CONTENT

Supporting Information

Additional details on computations, coordinates and energies for computed structures (PDF)

AUTHOR INFORMATION

Corresponding Author

**djtantillo@ucdavis.edu*

ORCID

Dean J. Tantillo: 0000-0002-2992-8844

Notes

The authors declare no competing financial interest.

Acknowledgements

We are grateful to the US National Science Foundation (CHE-1565933 and CHE-030089 for computer time via the XSEDE program) for support. CHM acknowledges a fellowship from UC Davis' Chemical Biology Program (supported in part by an NIH T32 Training Grant). We are grateful to Innes Tsui for help with NBO calculations.

References

1. Cane, D. E. *Chem. Rev.* **1990**, *90*, 1089-1103: "Enzymic Formation of Sesquiterpenes".
2. (a) Pateraki, I.; Heskes, A. M.; Hamberger, B. *Adv. Biochem. Eng. Biotechnol.* **2015**, *148*, 107-139: "Cytochromes P450 for Terpene Functionalisation and Metabolic Engineering". (b) Janocha, S.; Schmitz, D.; Bernhardt, R. *Adv. Biochem. Eng. Biotechnol.* **2015**, *148*, 215-250: "Terpene Hydroxylation with Microbial Cytochrome P450 Monooxygenases". (c) Banerjjee, A.; Hamberger, B. *Phytochem. Rev.* **2018**, *17*, 81-111: "P450s Controlling Metabolic Bifurcations in Plant Terpene Specialized Metabolism".
3. Rudolf, J. D.; Chang, C.-Y. *Nat. Prod. Rep.*, in press, DOI: 10.1039/C9NP00051H: "Terpene Synthases in Disguise: Enzymology, Structure, and Opportunities of Non-canonical Terpene Synthases".

4. (a) Ma, S.-G.; Li, M.; Lin, M.-B.; Li, L.; Liu, Y.-B.; Qu, J.; Li, Y.; Wang, X.-J.; Wang, R.-B.; Xu, S.; Hou, Q.; Yu, S.-S. *Org. Lett.* **2017**, *19*, 6160-6163: "Illisimonin A, a Caged Sesquiterpenoid with a Tricyclo[5.2.1.$0^{1,6}$]decane Skeleton from the Fruits of *Illicium simonsii*. (b) An elegant total synthesis that served to reassign the absolute configuration (and demonstrated the reverse of the **5**-to-**6**-type rearrangement for a related non-biosynthetic analog): Burns, A. S.; Rychnovsky, S. D. *J. Am. Chem. Soc.* **2019**, *141*, 13295-13300: "Total Synthesis and Structure Revision of (−)-Illisimonin A, a Neuroprotective Sesquiterpenoid from the Fruits of *Illicium simonsii*".

5. Tomita, B.; Isono, T.; Hirose, Y. *Tetrahedron Lett.* **1970**, 1371-1372: "Terpenoids. XXVIII. Acorane Type Sesquiterpenoids from Juniperus rigida and Hypothesis for the Formation of New Tricarbocyclic Sesquiterpenoids".

6. (a) Tantillo, D. J. *Nat. Prod. Rep.* **2011**, *28*, 1035-1053: "Biosynthesis via Carbocations: Theoretical Studies on Terpene Formation". (b) Tantillo, D. J. *Angew. Chem. Int. Ed.* **2017**, *56*, 10040-10045: "Importance of Inherent Substrate Reactivity in Enzyme Promoted Carbocation Cyclization/Rearrangements". (c) Sato, H.; Saito, K.; Yamazaki, M. *Front. Plant Sci.* **2019**, *10*, 802: "Acceleration of Mechanistic Investigation of Plant Secondary Metabolism Based on Computational Chemistry". (d) Major, D. T.; Freud, Y.; Weitman, M. *Curr. Opin. Chem. Biol.* **2014,** *21*, 25-33: "Catalytic Control in Terpenoid Cyclases: Multiscale Modeling of Thermodynamic, Kinetic, and Dynamic Effects".

7. (a) Chen, K.; Baran, P. S. *Nature* **2009**, *459*, 824-828: "Total Synthesis of Eudesmane Terpenes by Site-Selective C–H Oxidations". (b) Hung, K.; Condakes, M. L.; Novaes, L. F. T.; Harwood, S. J.; Morikawa, T.; Yang, Z.; Maimone, T. J. *J. Am. Chem. Soc.* **2019**, *141*, 3083-3099 : "Development of a Terpene Feedstock-Based Oxidative Synthetic Approach to the Illicium Sesquiterpenes".

8. (a) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664-675: "Exchange Functionals with Improved Long-rang Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW Models". (b) Matsuda, S.P.T.; Wilson, W.K.; Xiong, Q. *Org. Biomol. Chem.* **2006**, 4, 530-543: "Mechanistic insights into triterpene synthesis

from quantum mechanical calculations. Detection of systematic errors in B3LYP cyclization energies".

9. Gaussian 09, Revision B.01, M. J. Frisch *et al.,* Inc., Wallingford CT, **2009**.
10. (a) Fukui, K. *Acc. Chem. Res.* **1981**, 14, 363-368: "The path of chemical reactions - the IRC approach". (b) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H. B.; Ding, L.; Morokuma, K. *Chem. Rev.* **2015**, 115, 5678-5796: "The ONIOM Method and Its Applications". (c) Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. *Int. J. Quantum Chem.* **2015**, 115, 258-269: "Intrinsic reaction coordinate: Calculation, bifurcation, and automated search".
11. CYLview, 1.0b; Legault, C.Y., Universite de Sherbrooke, **2009** (http://www.cylview.org).
12. (a) Hong, Y. J.; Giner, J.-L.; Tantillo, D. J. *J. Org. Chem.* **2013**, *78*, 935-941: "Triple-Shifts and Thioether Assistance in Rearrangements Associated with an Unusual Biomethylation of the Sterol Side Chain". (b) Castineira Reis, M; Silva Lopez, C.; Faza, O. N.; Tantillo, D. J. *Chem. Sci.* **2019**, *10*, 2159-2170: "Pushing the Limits of Concertedness. A Waltz of Wandering Carbocations".
13. Johnston, R. C.; Cheong, P. H.-Y. *Org. Biomol. Chem.* **2013**, *11*, 5057-5064: "C-H•••O Non-Classical Hydrogen Bonding in the Stereomechanics of Organic Transformations: Theory and Recognition".
14. (a) Tantillo, D. J. *Chem. Soc. Rev.* **2010**, *39*, 2847-2854: "The Carbocation Continuum in Terpene Biosynthesis - Where are the Secondary Cations?". (b) Tantillo, D. J. *J. Phys. Org. Chem.* **2008**, *21*, 561-570: "Recent Excursions to the Borderlands between the Realms of Concerted and Stepwise: Carbocation Cascades in Natural Products Biosynthesis". (c) Hess, B. A., Jr. *J. Am. Chem. Soc.* **2002**, *124*, 10286-10287: "Concomitant C-Ring Expansion and D-Ring Formation in Lanosterol Biosynthesis from Squalene without Violation of Markovnikov's Rule".
15. (a) Chan, H. S. S.; Nguyen, Q. N. N.; Paton, R. S.; Burton, J. W. *J. Am. Chem. Soc.* **2019**, *141*, 15951-15962: "Synthesis, Characterization, and Reactivity of Complex Tricyclic Oxonium Ions, Proposed Intermediates in Natural Product Biosynthesis". (b) Gunbas, G.; Sheppard, W. L.; Fettinger, J. C.; Olmstead, M. M.; Mascal, M. *J. Am. Chem. Soc.* **2013**, *135*, 8173-

8176: “Extreme Oxatriquinanes: Structural Characterization of α-Oxyoxonium Species with Extraordinarily Long Carbon–Oxygen Bonds”.

16. Wang, S. C.; Tantillo, D. J. *J. Org. Chem.* **2008**, *73*, 1516-1523: "Theoretical Studies on Synthetic and Biosynthetic Oxidopyrylium–Alkene Cycloadditions. Pericyclic Pathways to Intricarene"

17. Hong, Y. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2009**, *131*, 7999-8015: "Consequences of Conformational Preorganization in Sesquiterpene Biosynthesis. Theoretical Studies on the Formation of the Bisabolene, Curcumene, Acoradiene, Zizaene, Cedrene, Duprezianene, and Sesquithuriferol Sesquiterpenes.

18. As expected for concerted reactions with multiple asynchronous events, reaction coordinates for both processes involve distinct shoulders, often regions with structures resembling previously proposed intermediates.[6,12,14]

19. This difference is related to the difference between hyperconjomers: Rauk, A.; Sorensen, T. S.; Schleyer, P. v. R. *J. Chem. Soc., Perkin Trans. 2* **2001**, 869-874: “Tertiary Cyclohexyl Cations. Definitive Evidence for the Existence of Isomeric Structures (Hyperconjomers)”.

20. (a) Tantillo, D. J.; Chen, J.; Houk, K. N. *Curr. Op. Chem. Biol.* **1998**, *2*, 743-750: "Theozymes and Compuzymes: Theoretical Models for Biological Catalysis". (b) Zhang, X.; DeChancie, J.; Gunaydin, H.; Chowdry, A. B.; Clemente, F. R.; Smith, A. J. T.; Handel, T. M.; Houk, K. N. *J. Org. Chem.* **2008**, *73*, 889-899: “Quantum Mechanical Design of Enzyme Active Sites”.

21. Hare, S. R.; Pemberton, R. P.; Tantillo, D. J. *J. Am. Chem. Soc.* **2017**, *139*, 7485-7493: "Navigating Past a Fork in the Road—Carbocation–π Interactions Can Manipulate Dynamic Behavior of Reactions Facing Post-Transition State Bifurcations"

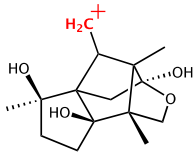
22. (a) Hong, Y. J.; Tantillo, D. J. *Chem. Sci.* **2013**, *4*, 2512-2518: "C-H•••π Interactions as Modulators of Carbocation Structure - Implications for Terpene Biosynthesis". (b) Hong, Y. J.; Tantillo, D. J. *Org. Lett.* **2015**, *17*, 5388-5391: "Tension between Internal and External Modes of Stabilization in Carbocations Relevant to Terpene Biosynthesis - Modulating Minima Depth via C–H---π Interactions". (c) Sato, H.; Narita, K.; Minami, A.; Yamazaki, M.; Wang, C.; Suemune, H.; Nagano, S.; Tomita, T.; Oikawa, H.; Uchiyama, M. *Scientific Rep.*

2018, *8*, 2473: “Theoretical Study of Sesterfisherol Biosynthesis: Computational Prediction of Key Amino Acid Residue in Terpene Synthase”.

23. (a) Christianson, D. W.. *Chem. Rev.* **2006,** *106*, 3412-3442: “Structural Biology and Chemistry of the Terpenoid Cyclases”. (b) Christianson, D. W.. *Chem.l Rev.* **2017,** *117*, 11570-11648: “Structural and Chemical Biology of Terpenoid Cyclases”.

24. (a) Tomita, B.; Hirose, Y. *Phytochemistry* **1973**, *12*, 1409-1414: “*Allo*-Cedrol: A New Tricarbocyclic Sesquiterpene Alcohol”. (b) Liu, H.; Chen, M.; Lang, Y.; Wang, X.; Zhuang, P. *Nat. Prod. Res.*, in press, DOI: 10.1080/14786419.2018.1538222: “Sesquiterpenes from the Fruits of *Illicium Simonsii* maxim”. (c) Schmidt, T. J.; Muller, E.; Fronczek, F. R. *J. Nat. Prod.* **2001**, *64*, 411-414: “New *allo*-Cedrane Type Sesquiterpene Hemiketals and Further Sesquiterpene Lactones from Fruits of *Illicium floridanum*”. (d) Zhang, G.; Zhuang, P.; Wang, X.; Yu, S.; Ma, S.; Qu, J.; Li, Y.; Liu, Y.; Zhang, Y.; Yu, D. *Planta Med.* **2013**, *79*, 1056-1062: “Sesquiterpenes from the Roots of *Illicium ijadifengpi*”

TOC graphic



transition state structure rather than intermediate