J|A|C|S ARTICLES

Published on Web 04/15/2010

Challenges Posed to Bornyl Diphosphate Synthase: Diverging Reaction Mechanisms in Monoterpenes

Michal Weitman and Dan Thomas Major*

Department of Chemistry and the Lise Meitner-Minerva Center of Computational Quantum Chemistry, Bar-Ilan University, Ramat-Gan 52900, Israel

Received December 5, 2009; E-mail: majort@mail.biu.ac.il

Abstract: The simplest form of terpenoid chemistry is found for the monoterpenes, which give plants fragrance, flavor, and medicinal properties. Monoterpene synthases employ geranyl diphosphate as a substrate to generate an assortment of cyclic products. In the current study we present a detailed analysis of the multiple gas-phase reaction pathways in the synthesis of bornyl cation from geranyl diphosphate. Additionally, the fate of the proposed bornyl cation intermediate in the bornyl diphosphate synthase reaction is investigated by molecular dynamics simulations. We employ accurate density functional theory (DFT) methods after careful validation against high-level ab initio data for a set of model carbocations. The gas-phase results for the monoterpene reactions indicate a diverging reaction mechanism with multiple products in the absence of enzymatic control. This complex potential energy surface includes several possible bifurcation points due to the presence of secondary cations. Additionally, the suggested bornyl cation intermediate in the bornyl diphosphate synthase reaction is studied by molecular dynamics simulations employing a hybrid quantum mechanics (DFT)−molecular mechanics potential energy function. The simulations suggest that the bornyl cation is a transient species as in the gas phase and that electrostatic steering directs the formation of the final product, bornyl diphosphate.

Introduction

Most enzymes catalyze reactions at rates that approach the encounter rate between the enzyme and substrate.[1] Typical turnover numbers, reflected in catalytic rate constant k_{cat}, range from 50 to 5000 molecules per second.[2] Such turnover numbers are indeed impressive and display a remarkable uniformity across a broad range of chemical reactions catalyzed by enzymes. The catalytic effect of enzymes, which constitute a rate enhancement of up to 20 orders of magnitude, is largely ascribed to transition-state (TS) stabilization, due to a preorganized polar environment in the enzyme active site.[3,4]

Interestingly, in the generation of many natural molecules, synthetic control and precision, rather than rate enhancement, seem to be the key enzymatic tools for generating structural diversity from a limited pool of simple metabolites.[5] For example, terpenoids (isoprenoids) form a ubiquitous family of structurally and stereochemically diverse natural compounds,[6−11] with only modest rate enhancements.[12−14] Thus far more than 55 000 isoprenoids have been identified, which all employ the five-carbon precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP).[15] IPP can be elongated into sequentially longer isoprenoid diphosphate precursors such as geranyl diphosphate (GPP; C10), farnesyl diphosphate (FPP; C15), and geranylgeranyl diphosphate (GGPP; C20). These may subsequently undergo highly specific cyclization reactions to generate a variety of natural products in the presence of the diphosphate moiety and Mg^{2+} or Mn^{2+} ions. A most remarkable feature of these enzymes is that they employ one of only three acyclic precursors, GPP, FPP, or GGPP, to produce a variety of compounds with unique architecture and stereochemistry. Cyclic terpenoid componds are formed through electrophilic cyclizations involving highly reactive carbocation intermediates which must be sequestered within the active site to avoid abortive side reations.[16] The chemistry of terpenes has been studied extensively by experimental methods.[8,16,17] However, due to the highly reactive nature of carbocations, the existence of many intermediate structures is merely postulated. Computational studies have emerged that address the nature of the

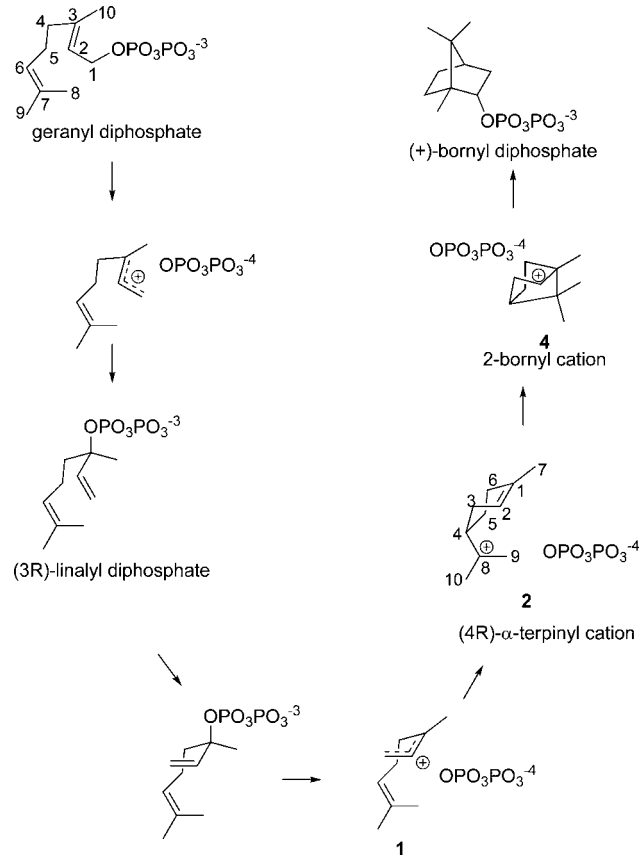
(1) Alberty, R. A.; Hammes, G. G. *J. Phys. Chem.* **1958**, *62*, 154–159.
(2) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90–93.
(3) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210–3235.
(4) Zhang, X.; Houk, K. N. *Acc. Chem. Res.* **2005**, *38*, 379–385.
(5) Clardy, J.; Walsh, C. *Nature* **2004**, *432*, 829–837.
(6) Christianson, D. W. *Curr. Opin. Chem. Biol.* **2008**, *12*, 141–150.
(7) Christianson, D. W. *Science* **2007**, *316*, 60–61.
(8) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412–3442.
(9) Lesburg, C. A.; Caruthers, J. M.; Paschall, C. M.; Christianson, D. W. *Curr. Opin. Struct. Biol.* **1998**, *8*, 695–703.
(10) Yoshikuni, Y.; Ferrin, T. E.; Keasling, J. D. *Nature* **2006**, *440*, 1078–1082.
(11) Degenhardt, J.; Kollner, T. G.; Gershenzon, J. *Phytochemistry* **2009**, *70*, 1621–1637.
(12) Winstein, S.; Valkanas, G.; Wilcox, C. F. J. *J. Am. Chem. Soc.* **1972**, *94*, 2286–2290.
(13) Williams, D. C.; McGarvey, D. J.; Katahira, E. J.; Croteau, R. *Biochemistry* **1998**, *37*, 12213–12220.
(14) Cane, D. E.; Chiu, H. T.; Liang, P. H.; Anderson, K. S. *Biochemistry* **1997**, *36*, 8332–8339.
(15) Sacchettini, J. C.; Poulter, C. D. *Science* **1997**, *277*, 1788–1789.
(16) Cane, D. E. *Chem. Rev.* **1990**, *90*, 1089–1103.
(17) Gershenzon, J.; Dudareva, N. *Nat. Chem. Biol.* **2007**, *3*, 408–414.

intermediates in the carbocation reaction cascades, both in the gas phase[18−36] and in enzymes.[37,38]

The simplest form of terpenoid chemistry is found for the monoterpenes, which bestow plants with fragrance, flavor, and medicinal properties.[11,39,40] These volatile organic compounds have also been suggested to be important players in organic aerosol formation, thereby influencing radiative forcing and global warming.[41] Monoterpene synthases employ GPP as a substrate to generate an assortment of cyclic monoterpene products. An extensively studied monoterpenoid cyclase is (+)-bornyl diphosphate synthase (BPPS).[8] The crystal structure of BPPS was solved in the presence of several ligands mimicking intermediate states of the proposed reaction, including the product (+)-bornyl diphosphate (BPP).[42] These structures shed considerable light on key steps in the reaction mechanism. The currently accepted reaction mechanism in BPPS is thought to involve the following steps (Scheme 1):[8] (1) binding of GPP, (2) metal-activated ionization to yield an allylic transoid carbocation,[43] (3) formation of (3*R*)-linalyl diphosphate (LPP), (4) rotation around the C2−C3 bond and ionization to generate a cisoid allylic cation in the reactive conformation, (5) cyclization to form a (4*R*)-α-terpinyl cation, (6) further cyclization to yield (+)-2-bornyl cation, and (7) reincorporation of the pyrophosphate by the (+)-2-bornyl carbocation, to yield the final product BPP in ca. 75% yield.[44] This mechanism has received accumulating support from experiments[42,44−51] with isotopically labeled substrates[45,46,48,51] as well as substrate and intermediate analogues.[42] Kinetic studies have suggested that the rate-limiting chemical step is the cleavage of the carbon−oxygen bond, yielding the reactive allylic carbocation.[46]

Following the formation of the terpinyl cation (Scheme 1), many side reactions, yielding other monoterpene products, could

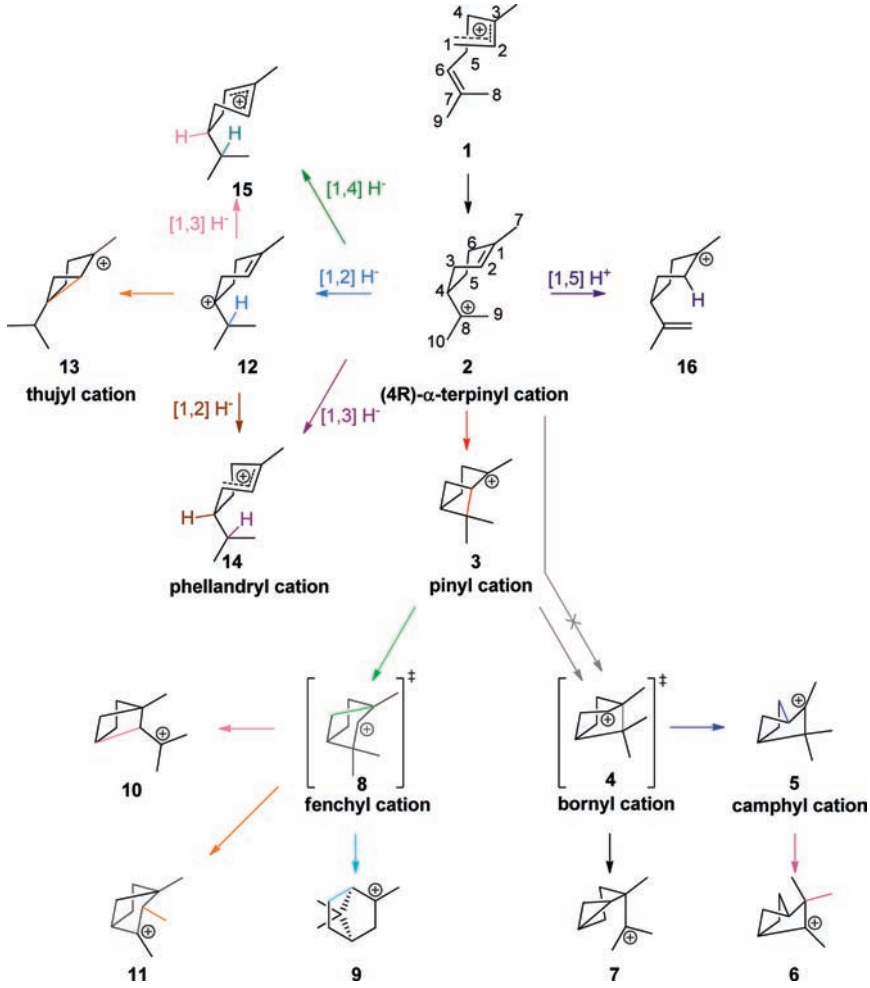
Scheme 1. Cyclization of Geranyl Diphosphate To Produce (+)-Bornyl Diphosphate by the Enzyme Bornyl Diphosphate Synthase



be envisaged (Scheme 2). For instance, migrations, hydride transfers, proton transfers, or deprotonations could yield additional products such as limonene, terpinolene, camphene, α-pinene, *β*-pinene, *β*-phellandrene, α-terpinene, *γ*-terpinene, α-thijuene, sabinene, and 3-carene. Indeed, in addition to the main product BPP, BPPS also produces significant amounts of (+)-α-pinene, (+)-camphene, and (±)-limonene.[44] The principle governing the product distribution in terpene enzymes is thought to be substrate folding in the active site and the positioning of active-site bases.[8] In an initial attempt to answer these questions, we map out an extensive reaction scheme in the gas phase for the monoterpenes. A detailed analysis of the inherent reaction profile is a prerequisite for understanding the enzymatic reaction of BPPS and other monoterpenes. The current study employs density functional theory (DFT) calculations. Initially, the

(18) Raghavachari, K.; Haddon, R. C.; Schleyer, P. v. R.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1983**, *105*, 5915–5917.
(19) Sieber, S.; Schleyer, P. v. R.; Vancik, H.; Mesic, M.; Sunko, D. E. *Angew. Chem., Int. Ed.* **1993**, *32*, 1604–1606.
(20) Schleyer, P. v. R.; Sieber, S. *Angew. Chem., Int. Ed.* **1993**, *32*, 1606–1608.
(21) Schreiner, P. R.; Severance, D. L.; Jorgensen, W. L.; Schleyer, P. v. R.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1995**, *266*, 3–2664.
(22) Schreiner, P. R.; Schleyer, P. v. R.; Schaefer, H. F., III. *J. Org. Chem.* **1997**, *62*, 4216–4228.
(23) Gutta, P.; Tantillo, D. J. *J. Am. Chem. Soc.* **2005**, *128*, 6172–6179.
(24) Xiong, Q.; Rocco, F.; Wilson, W. K.; Xu, R.; Ceruti, M.; Matsuda, S. P. T. *J. Org. Chem.* **2005**, *70*, 5362–5375.
(25) Matsuda, S. P. T.; Wilson, W. K.; Xiong, Q. *Org. Biomol. Chem.* **2006**, *4*, 530–543.
(26) Hong, Y. J.; Tantillo, D. J. *Org. Lett.* **2006**, *8*, 4601–4604.
(27) Bojin, M. D.; Tantillo, D. J. *J. Phys. Chem. A* **2006**, *110*, 4810–4816.
(28) Gutta, P.; Tantillo, D. J. *Org. Lett.* **2007**, *9*, 1069–1071.
(29) Hong, Y. J.; Tantillo, D. J. *J. Org. Chem.* **2007**, *72*, 8877–8881.
(30) Lodewyk, M. W.; Gutta, P.; Tantillo, D. J. *J. Org. Chem.* **2008**, *73*, 6570–6579.
(31) Tantillo, D. J. *J. Phys. Org. Chem.* **2008**, *21*, 561–570.
(32) Wang, S. C.; Tantillo, D. J. *Org. Lett.* **2008**, *10*, 4827–4830.
(33) Hong, Y. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2009**, *131*, 7999–8015.
(34) Hong, Y. J.; Tantillo, D. J. *Nature Chem.* **2009**, *1*, 384–389.
(35) Hess, A. B. J. *J. Am. Chem. Soc.* **2002**, *124*, 10286–10287.
(36) Werstiuk, N. H. *J. Chem. Theory Comput.* **2007**, *3*, 2258–2267.
(37) Rajamani, R.; Gao, J. *J. Am. Chem. Soc.* **2003**, *125*, 12768–12781.
(38) Allemann, R. K.; Young, N. J.; Ma, S.; Truhlar, D. G.; Gao, J. *J. Am. Chem. Soc.* **2007**, *129*, 13008–13013.
(39) McCaskill, D.; Croteau, R. *Adv. Biochem. Eng. Biotechnol.* **1997**, *55*, 107–146.
(40) Davis, E. M.; Croteau, R. *Top. Curr. Chem.* **2000**, *209*, 53–95.
(41) Kiendler-Scharr, A.; Wildt, J.; Dal Maso, M.; Hohaus, T.; Kleist, E.; Mentel, T. F.; Tillmann, R.; Uerlings, R.; Schurr, U.; Wahner, A. *Nature* **2009**, *461*, 381–384.
(42) Whittington, D. A.; Wise, M. L.; Urbansky, M.; Coates, R. M.; Croteau, R. B.; Christianson, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15375–15380.
(43) Chayet, L.; Rojas, M. C.; Cori, O.; Bunton, C. A.; McKenzie, D. C. *Bioorg. Chem.* **1984**, *12*, 329–338.
(44) Wise, M. L.; Savage, T. J.; Katahira, E.; Croteau, R. *J. Biol. Chem.* **1998**, *273*, 14891–14899.
(45) Cane, D. E.; Saito, A.; Croteau, R.; Shaskus, J.; Felton, M. *J. Am. Chem. Soc.* **1982**, *104*, 5831–5833.
(46) Croteau, R. B.; Shaskus, J. J.; Renstrøm, B.; Felton, N. M.; Cane, D. E.; Saito, A.; Chang, C. *Biochemistry* **1985**, *24*, 7077–7085.
(47) Croteau, R.; Felton, M. *Arch. Biochem. Biophys.* **1981**, *207*, 460–464.
(48) Croteau, R.; Felton, N. M.; Wheeler, C. J. *J. Biol. Chem.* **1985**, *260*, 5956–5962.
(49) Croteau, R.; Satterwhite, D. M.; Wheeler, C. J.; Felton, N. M. *J. Biol. Chem.* **1989**, *264*, 2075–2080.
(50) Wise, M. L.; Pyun, H.-J.; Helms, G.; Assink, B.; Coates, R. M.; Croteau, R. B. *Tetrahedron* **2001**, *57*, 5327–5334.
(51) Croteau, R.; Satterwhite, D. M.; Cane, D. E.; Chang, C. C. *J. Biol. Chem.* **1986**, *261*, 13438–13445.

Scheme 2. Possible Carbocation Pathways for the Bornyl Diphosphate Synthase Catalysis of Geranyl Diphosphate to (+)-Bornyl Diphosphate



performance of several DFT methods on a carefully selected database of small carbocations is compared with high-level ab initio methods. Subsequently, the optimal DFT level is employed in the study of the monoterpene carbocation reaction cascade. A key question arising from the gas-phase study is the plausibility of a stable secondary bornyl cation in BPPS. To address this issue we perform activated and free energy molecular dynamics (MD) simulations on a hybrid quantum mechanics/molecular mechanics (QM/MM) potential energy surface (PES), where the quantum region is treated with the optimal DFT method found in the gas-phase study. Although QM/MM simulations have been employed in the past in the study of terpenes, the current work is the first to go beyond the semiempirical QM level, which is essential in the treatment of highly reactive carbocations.

Computational Details

Gas-Phase Modeling. Model calculations on small carbocation compounds (Figure 1) were performed at high-level ab initio and various DFT levels in order to assess the most reliable DFT method for the current study of monoterpenes. The model structures are cations of ethane, propane, and butane. These structures are not meant to represent an exhaustive set of all possible carbocations for these compounds. Rather, the intent is to choose some key structures that contain important classical and nonclassical carbocation structures. The target ab initio levels employed herein were the MP4 and CCSD(T) methods with the 6-311+G(2d,p) basis set and frozen cores, optimized at the MP2/6-311+G(2d,p) level.[52] The DFT methods considered include both pure and hybrid functionals. In particular, the PBEPBE,[53] PBEPBEh,[54] B3LYP,[55−57] MPW1PW91,[58] B98,[59] BMK,[60] MPWB1K,[61] and BB1K[62] density functionals were tested in conjunction with the 6-31+G(d,p) basis set, which is a practical basis set for use in larger systems.[52] On the basis of the model calculations, BB1K/6-31+G(d,p) was chosen as the most appropriate level for the study of the monoterpenes. The optimization at all DFT levels and MP2 employed analytical derivatives. All stationary points were characterized with frequency calculations. All BB1K/6-31+G(d,p) frequencies were scaled by 0.96 when computing zero-point energy.[62] To verify that the TSs correspond to the expected reactant and product wells, we performed intrinsic reaction coordinate (IRC) calculations for all TSs.[63−66] All gas-phase calculations were performed with the Gaussian 03 program.[67]

(52) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.
(53) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
(54) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
(55) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(56) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(57) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
(58) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664–675.
(59) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
(60) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405–3416.
(61) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908–6918.
(62) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2715–2719.

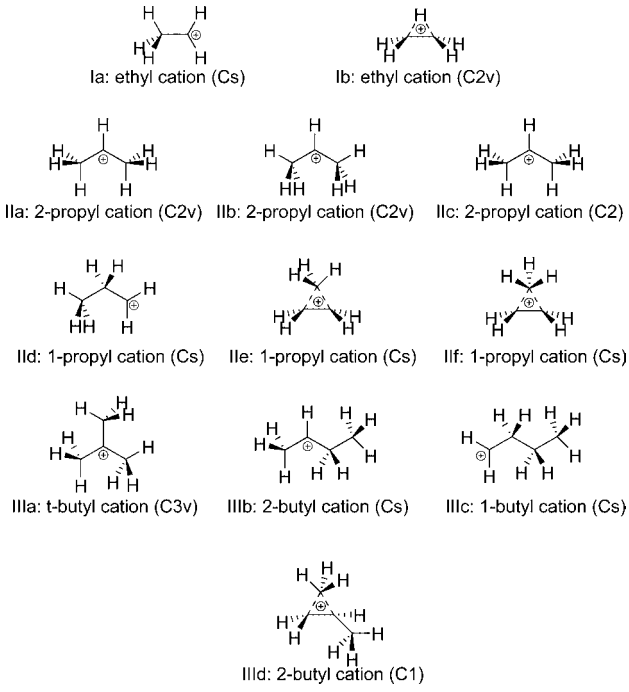


Figure 1. Model carbocation structures.

Enzyme Modeling. Overview. The question of the fate of the bornyl cation was addressed by two different sets of QM/MM MD simulations: (1) We performed activated MD where the bornyl cation was used as a starting point for a series of unbiased MD simulations. In these simulations, 10 initial starting structures were taken from a single 1 ns MD simulation, using AM1 for the QM region for the carbocation, thus avoiding rearrangement as described below. Specifically, one structure was extracted every 100 ps from the 1 ns simulation, and each of these structures was used as a starting point for the QM(BB1K)/MM MD simulation. Importantly, 10 independent MD simulations, using these uncorrelated starting points, were run with QM(BB1K)/MM. (2) To quantify the results of the activated QM/MM dynamics, we performed free energy simulations that involve multiple trajectories along a predefined reaction coordinate.

Model of Solvated Enzyme−Coenzyme−Substrate Complex. The X-ray crystallography structure of dimeric (+)-bornyl diphosphate synthase from *Salvia officinalis* has been published by Whittington et al.[42] The enzyme is a homodimer comprised of two α helical domain units containing 598 amino acids each. Each monomer possesses an independent active-site pocket that is sequestered from water, although a single specific water molecule is present in the binding site. The enzyme was solved at 2.4 Å resolution, with the diphosphate and 2-azabornane occupying the active site (PDB code 1N23). The 2-azabornane in the active site was manually modified to the 2-bornyl cation.

Hydrogen atoms were added to the enzyme using the HBUILD module of the CHARMM program,[68] while the hydrogen atoms of the substrate were added manually. The protonation states of all ionizable residues were assigned on the basis of physiological pH. His residues were modeled as neutral or protonated moieties with hydrogen positioned at either Nδ or Nϵ or both, depending on their hydrogen-bonding pattern with surrounding amino acid residues or water molecules. This enzyme structure was employed as a starting point for extensive MD simulations.

Potential Energy Surface. The PES in the current study is described by a hybrid QM/MM Hamiltonian.[69] In this treatment, the reactive fragment wherein the chemistry occurs is treated via QM to allow for bond breaking and forming, while the structural and electrostatic effects of the enzyme and solvent are included via MM. The QM region contains the carbocation, diphosphate, and Mg^{2+} ions in one of the active sites and is described by the BB1K functional,[62] in conjunction with a hybrid basis set. On H, C, and O atoms we employed the 6-31G(d) basis set,[52] while on P and Mg we employed the CEP effective core potential basis set.[70] The performance of the 6-31G(d) basis set in comparison with the 6-31+G(d,p) basis set used in the gas-phase study was estimated for the bornyl and camphyl cations. The camphyl cation is more stable than the bornyl cation by 20.32 and 20.35 kcal/mol using BB1K and the 6-31G(d) and 6-31+G(d,p) basis sets, respectively. The geometry difference was also found to be negligible (Table S173, Supporting Information). The MM part is described by the CHARMM22 force field.[71] Water molecules are treated by the TIP3P model.[72] The electrostatic part of the QM/MM interactions is treated with electrostatic embedding (the MM partial charges are included in the Fock/Kohn−Sham operators),[69] wherein the MM partial charges are allowed to polarize the QM region. Additionally, van der Waals (vdW) interactions between QM and MM atoms are included.

Stochastic Boundary Molecular Dynamics. The current MD study employed stochastic boundary conditions for the enzymatic reaction due to the size of BPPS and the high cost of the QM/MM simulations.[73] The enzyme was soaked in a sphere with a radius of 24 Å, and the Langevin stochastic boundary ranged from 20 to 24 Å. The temperature of the constant particle−volume−temperature (NVT) simulations was 298 K. The simulations employed the Leap-Frog integration scheme with a time step of 1 fs.[74] TIP3P water and protein hydrogens were constrained using the SHAKE algorithm.[74] For the MM group atoms, the nonbonded interactions were set to zero at distances beyond 14 Å. The electrostatic forces were shifted to zero from a distance of 12 Å, while the vdW interaction energy was switched to zero at 12 Å. No cutoff was employed in computing the QM/MM interactions. The system was initially heated slowly to 298 K during the course of 25 ps and thereafter equilibrated for 75 ps. Subsequently, the system was further simulated for 1 ns. During these simulations, the QM region was composed of the carbocation only and treated by the AM1 Hamiltonian.[75] We note that, at the semiempirical level, no carbocation rearrangement occurs due to very high barriers.[75] Subsequently, for the activated dynamics, the QM region was enlarged to include the diphosphate and Mg^{2+} ions, and the QM method was switched to DFT as described above. Ten equally separated snapshots were extracted from the 1 ns QM(AM1)/MM simulation and used as starting points for 10 independent QM(BB1K)/MM MD simulations. The initial velocities were taken directly from the QM(AM1)/MM simulation restart files. These 10 MD simulations lasted 5 ps each, for a combined simulation time of 50 ps.

(63) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154–2161.
(64) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1991**, *95*, 5853–5860.
(65) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Phys.* **2004**, *120*, 9918–9924.
(66) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Theory Comput.* **2005**, *1*, 61–69.
(67) Frisch, M. J.; et al. *Gaussian 03*, revision D.02; Gaussian, Inc.: Pittsburgh, PA, 2003.
(68) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(69) Gao, J. *Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials*; VCH: New York, 1995; Vol. 7.
(70) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026–6033.
(71) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(72) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(73) Brooks, C. L., III; Brünger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843–865.
(74) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, U.K., 1987.
(75) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

Table 1. Benchmark Calculations for Model Carbocation Systems

	Ia	Ib	IIa	IIb	IIc	IId	IIe	IIf	IIIa	IIIb	IIIc	IIId
MP4SDTQ[a]	6.67	0.00	0.48	3.82	0.00	19.83	6.83	6.73	0.00	14.61	34.49	10.98
CCSDT[a]	6.12	0.00	0.43	3.67	0.00	19.48	7.46	7.35	0.00	14.36	34.04	11.50
B3LYP[b]	3.48	0.00	0.09	3.56	0.00	21.00	10.39	10.60	0.00	13.96	35.70	12.43
MPW1PW91[b]	4.79	0.00	0.15	3.63	0.00	21.12	8.08	8.59	0.00	14.19	36.07	12.04
PBEPBE[b]	4.28	0.00	0.01	3.85	0.00	21.37	8.59	9.03	0.00	13.94	36.28	12.05
PBEPBEh[b]	5.09	0.00	0.16	3.70	0.00	21.19	7.76	8.34	0.00	14.28	36.20	11.93
B98[b]	3.65	0.00	0.10	3.48	0.00	20.81	9.04	9.30	0.00	13.98	35.51	12.05
BMK[b]	4.34	0.00	0.21	3.44	0.00	20.78	5.68	6.15	0.00	14.46	35.72	10.22
MPWB1K[b]	6.03	0.00	0.30	3.72	0.00	21.07	6.64	7.25	0.00	14.45	36.04	11.22
BB1K[b]	5.96	0.00	0.28	3.71	0.00	21.06	6.70	7.30	0.00	14.35	35.98	11.22

[a] Optimized at the MP2/6-311++G(2d,p) level. [b] Optimized with 6-31+G(d,p) basis set.

Free Energy Simulations. The classical potential of mean force (PMF) as a function of the reaction coordinate was defined as

$$W^{\mathrm{cl}}(\zeta) = -RT \ln \rho(\zeta) + C \qquad (1)$$

where ρ is the unbiased probability density along the reaction coordinate ζ, R is the gas-phase constant, T is the temperature, and C is an arbitrary constant.[76] The reaction coordinate was defined as a linear combination of the forming C2−O bornyl carbocation−diphosphate bond and the geometric camphyl→bornyl rearrangement coordinate, $\zeta = R_{C2-O} + R_{C1-C6} - R_{C2-C6}$. To facilitate uniform sampling along the reaction coordinate, a biasing potential is added to the potential energy of the system, $U^{\mathrm{bias}}(\zeta) = U^{\mathrm{umbr}}(\zeta) + U^{\mathrm{harm}}(\zeta)$. The umbrella potential, $U^{\mathrm{umbr}}(\zeta)$, is defined as $-W^{\mathrm{cl}}(\zeta)$, which essentially removes the reaction barrier, enabling uniform sampling. Furthermore, the reaction coordinate is divided into regions (windows) centered around ζ_0, and a harmonic potential, $U^{\mathrm{harm}}(\zeta) = k(\zeta - \zeta_0)^2/2$, is added. This allows the enzyme/substrate/solvent system to relax along the reaction coordinate. Practically, since $W^{\mathrm{cl}}(\zeta)$ is required for $U^{\mathrm{umbr}}(\zeta)$ and the optimal force constant, k, in $U^{\mathrm{harm}}(\zeta)$ is not known initially, the PMF is obtained by employing adaptive umbrella sampling MD simulations,[77] wherein the biasing potential is refined in an iterative manner. This simulation provides a biased PMF. The weighed histogram analysis method (WHAM) is used to combine the data from the different windows as well as to remove the effect of the biasing potential, yielding the unbiased PMF.[78] In the PMF simulations, the starting system was taken from the end of the 1 ns production run which used the QM(AM1)/MM. The PES was then changed to QM(BB1K)/MM and the reaction coordinate divided into 15 windows. Each window was briefly equilibrated for 1 ps and sampled for 5 ps each. Thus, the combined PMF simulation time was 75 ps.

All QM(AM1)/MM simulations used the CHARMM program with the SQUANTM module,[68] while all QM(BB1K)/MM simulations employed CHARMM interfaced with the GAMESS-UK program.[79]

Results

Model Gas-Phase Reactions. A key question in computational enzymology is the accuracy of the potential energy function. Previously, we have successfully employed modified reaction-specific semiempirical methods that combine high accuracy with good efficiency.[80−84] During the initial stages of this study, we attempted to develop semiempirical Hamiltonians tailored for carbocation reactions, but due to the inherent limitations of semiempirical formalism and implementation this was unsuccessful. Thus, a series of model calculations was performed in order to assess the performance of various DFT methods for carbocations. The model compounds chosen were various ethyl, propyl, and butyl carbocations, which have been studied extensively in the past by Tantillo et al.[27] and Schleyer and co-workers.[85−87] The models are shown in Figure 1, and the accompanying results are presented in Table 1.

The ability of these functionals to treat nonclassical bridging hydrogen was tested by comparing the ethyl carbocations **Ia** and **Ib**. At the MP4 and CCSD(T) levels the nonclassical carbocation **Ib** is more stable by 6.67 and 6.12 kcal/mol, respectively. All the functionals tested show reasonable performance and are within approximately 3 kcal/mol of the target values. Noteworthy, at the MPWB1K and BB1K level, the values are 6.03 and 5.96 kcal/mol, in good agreement with the high-level ab initio results.

Furthermore, the ability of the functionals to accurately differentiate between primary and various secondary classical carbocations, as well as between classical and nonclassical carbocations, was tested. Initially, three different conformers of the 2-propyl cation were investigated (**IIa**−**c**). Two of these possess C_{2v} symmetry with staggered (**IIa**) and eclipsed (**IIb**) conformations, while one has a staggered C_2 (**IIc**) symmetry, which is obtained by a slight rotation of the methyl groups in opposite directions. At all levels of theory the relative stability is **IIc** > **IIa** > **IIb**. **IIa** is less stable than **IIc** by 0.48 and 0.43 kcal/mol at the MP4 and CCSD(T) levels, respectively, while **IIb** is less stable than **IIc** by 3.82 and 3.67 kcal/mol, respectively. All DFT methods are accurate to within ca. 0.5 kcal/mol of these values. The primary cation, **IId**, is less stable by 19.83 and 19.48 kcal/mol at the MP4 and CCSD(T) levels, respectively. All DFT methods slightly overestimate this difference, by ca. 1.0−1.5 kcal/mol. The bridging, nonclassical carbocations, **IIe** and **IIf**, are less stable than **IIc** by 6.83 and 6.73 kcal/mol at the MP4 level, respectively, and by 7.46 and 7.35 kcal/mol at the CCSD(T) level. This difference between

(76) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690–7697.
(77) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
(78) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(79) Guest, M. F.; Bush, I. J.; van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. *Mol. Phys.* **2005**, *103*, 719–747.
(80) Major, D. T.; York, D. M.; Gao, J. *J. Am. Chem. Soc.* **2005**, *127*, 16374–16375.
(81) Major, D. T.; Nam, K.; Gao, J. *J. Am. Chem. Soc.* **2006**, *128*, 8114–8115.
(82) Major, D. T.; Gao, J. *J. Am. Chem. Soc.* **2006**, *128*, 16345–16357.
(83) Rubinstein, A.; Major, D. T. *J. Am. Chem. Soc.* **2009**, *131*, 8513–8521.
(84) Major, D. T.; Heroux, A.; Orville, A. M.; Valley, M. P.; Fitzpatrick, P. F.; Gao, J. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20734–20739.
(85) Schleyer, P. v. R.; Maerker, C.; Buzek, P.; Sieber, S. *Stable Carbocation Chemistry*; John Wiley and Sons: New York, 1997.
(86) Raghavachari, K.; Whiteside, R. A.; Pople, J. A.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1981**, *103*, 5649–5657.
(87) Carneiro, J. W. d. M.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1990**, *112*, 4064–4066.

the classical and nonclassical structures is well reproduced by the MPWB1K and BB1K methods, which have both been developed for chemical kinetics. The MPWB1K and BB1K methods are within 0.8 kcal/mol of the CCSD(T) values, respectively.

The final set of test cases involves the butyl carbocations and contains primary (**IIIc**), secondary (**IIIb**), and tertiary carbocations (**IIIa**) as well as a bridged carbocation (**IIId**). At the ab initio levels the relative stability is **IIIa** > **IIId** > **IIIb** > **IIIc**. At the DFT levels there is comparably good performance for all methods.

The geometries predicted by all DFT methods are in good agreement with the geometries obtained by MP2/6-311+G(2d,p) (Tables S171 and S172, Supporting Information). The average unsigned bond distance errors are 0.005 Å (B3LYP), 0.003 Å (MPW1PW91), 0.012 Å (PBEPBE), 0.004 Å (PBEPBEh), 0.006 Å (B98), 0.007 Å (BMK), 0.007 Å (MPWB1K), and 0.006 Å (BB1K). The unsigned average bond angle errors are 0.791 (B3LYP), 0.689 Å (MPW1PW91), 0.980 Å (PBEPBE), 0.718 Å (PBEPBEh), 0.722 Å (B98), 0.488 Å (BMK), 0.566 Å (MPWB1K), and 0.580 Å (BB1K).

The BB1K functional was also tested with the 6-31G(d) basis set for the model compounds. The difference between BB1K with 6-31G(d) and 6-31+G(d,p) is negligible in most cases. The largest difference for the two basis sets was between compounds **Ia** and **Ib** (classical and bridged ethyl cation), where the values were 5.20 and 5.96 kcal/mol in favor of **Ib** using 6-31G(d) and 6-31+G(d,p), respectively. This difference is due to the added polarization functions on hydrogens in the latter basis set. For all the other model compounds, the difference ranged between 0.00 and 0.20 kcal/mol for the two basis sets.

Of the DFT methods tested, BB1K and MPWB1K perform best, while B3LYP performs slightly worse than the other functional for energetics, while PBEPBE has the largest errors for geometries. Previous studies have noted that B3LYP systematically underestimates reaction barrier heights and that this functional is inaccurate for interactions that are subject to medium-range correlation energy and for alkane isomerization energies.[88,89]

Monoterpene Reaction Mechanism in the Gas Phase. On the basis of the model calculations of the previous section, we employed the BB1K/6-31+G(d,p) level for the monoterpene reaction. Indeed, the BB1K functional has shown excellent general performance.[88,89]

In order to clarify the mechanism of the bornyl cation formation, we investigated the reaction sequence **1** → **4** (Tables 2 and 3). While dealing with this mechanism, the paths leading to additional monoterpene products such as camphene, fenchone, α- and *β*-phellandrene, α- and *β*-pinene, sabinene, and *β*-thujene were also investigated. We stress that not all mechanistic possibilities were exhausted.

The carbocation reaction cascade commences with the allyl cation **1** (Scheme 2). This reactive intermediate in its extended form was used as a reference point with its relative potential energy, including zero-point correction, set to 0.00 kcal/mol, and all subsequent values for minimum structures are relative to this species. The prefolded allyl cation cyclizes spontaneously to form (4*R*)-α-terpinyl cation, **2**. The *β* form may also be formed, but herein we focus on the α form which is relevant to (+)-BPPS. This intermediate may exist in four principal conformations categorized by cyclohexene ring puckering (C6endo and C6exo, where endo is defined as being on the same side of the cyclohexene ring as the C4 hydrogen in **2**) and chair conformation (C4 substituent axial and equatorial). The relative energies are as follow: C4ax−C6exo, −15.36 kcal/mol; C4ax−C6endo, −15.05 kcal/mol; C4eq−C6exo, −13.98 kcal/mol; and C4eq−C6endo, −17.96 kcal/mol. We note that only the axial form is capable of undergoing further cyclization through the C8-position. Moreover, after further cyclization, only a single principal conformer exists due to the rigidity of the molecules.

Subsequent Markovnikov addition of the isopropyl cation moiety to the *π* bond leads directly to the formation of the pinyl

Table 2. Relative Potential Energies (kcal/mol) for Gas-Phase Monoterpenes in Scheme 2

compound	conformer	Δ*E*[a]
1		0.00
2	C4-ax, C6-exo	–15.36
	C4-ax, C6-endo	–15.05
	C4-eq, C6-exo	–13.98
	C4-eq, C6-endo	–17.96
3		–20.18
4[b]		–11.98
5		–31.50
6		–31.50
7		–9.08
8[b]		–12.02
9		–31.59
10		–9.73
11		–35.42
12	C6-exo	–18.72
	C6-endo	–18.72
13	C6-exo	–26.19
	C6-endo	–26.19
14	C4-ax, C6-exo	–27.99
	C4-ax, C6-endo	–29.70
15	C4-ax	–19.20
16	C4-ax, C6-exo	–13.04
	C4-ax, C6-endo	–10.92

[a] Computed with BB1K/6-31+G(d,p). Values include zero-point energy. [b] These structures were found to be TSs.

Table 3. Relative Potential Energies (kcal/mol) for Gas-Phase Monoterpene Reaction Transition States

reactant	product	conformer	Δ*E*[a]
2	**3**	C4-ax, C6-exo	1.94
		C4-ax, C6-endo	1.63
2	**12**	C4-ax, C6-exo	2.54
		C4-ax, C6-endo	4.27
		C4-eq, C6-exo	2.60
		C4-eq, C6-endo	4.20
2	**14**	C4-ax, C6-exo	2.11
		C4-ax, C6-endo	2.32
2	**15**	C4-ax, C6-exo	10.66
2	**16**	C4-ax, C6-exo	11.91
		C4-ax, C6-endo	11.61
4	**7**		26.27
5	**6**		8.45
8	**10**		14.83
8	**11**		2.41
12	**14**	C6-exo	3.96
		C6-endo	3.16
12	**15**	C6-exo	16.45
		C6-endo	15.64
12	**13**	C6-exo	3.47
		C6-endo	3.47

[a] Computed with BB1K/6-31+G(d,p). Values include zero-point energy.

(88) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
(89) Riley, K. E.; Op't Holt, B. T.; Merz, K. M. J. *J. Chem. Theory Comput.* **2007**, *3*, 407–433.

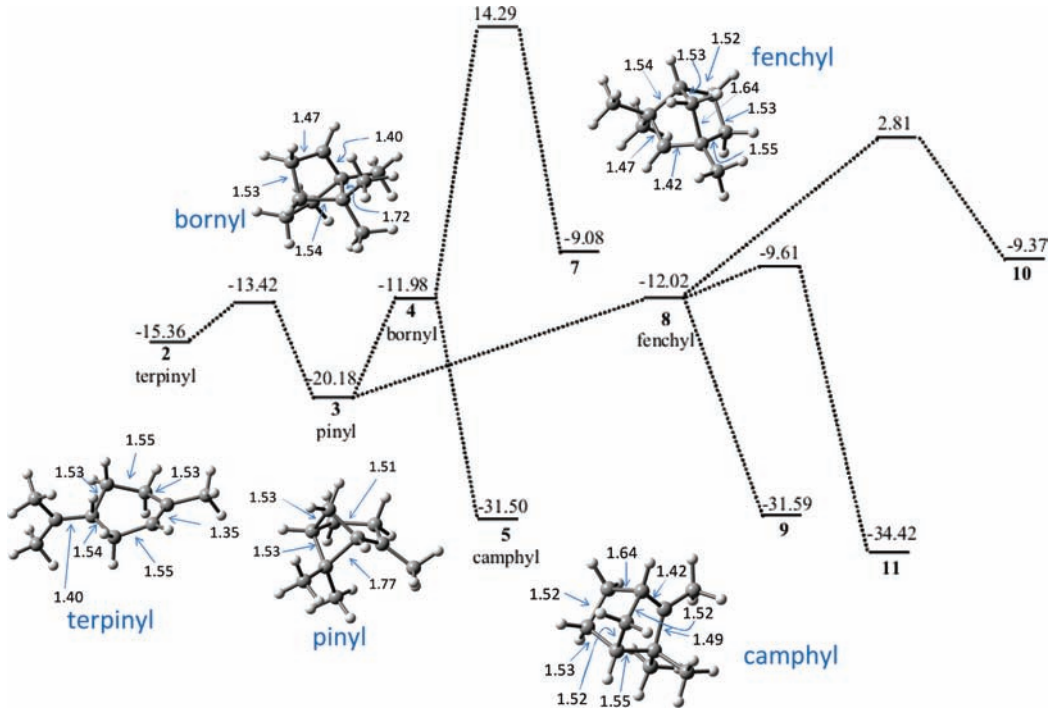


Figure 2. Potential energy diagram (including zero-point energy) for key steps in the gas-phase reaction of monoterpenes.

cation **3** through a C2−C8 bond formation, with a relative energy of −20.18 kcal/mol. The barrier for this step is 1.94 kcal/mol for C6exo and 1.63 kcal/mol for C6endo. The pinyl cation has a slight nonclassical nature, bearing a C2−C8 bond distance of 1.775 Å and a C1−C8 distance of 2.252 Å (Figures 2 and S1, Supporting Information). Additionally, the C2 hybridization is nearly sp_2. A following 1,2 isopropyl migration leads to the formation of the secondary cation bornyl, **4**, with a relative energy −11.98 kcal/mol. This secondary cation could not be located as a minimum but rather corresponds to a TS. This finding is in agreement with the findings of Schleyer and co-workers.[19] Similarly, other terpene secondary cations have been found to be TSs by Tantillo et al.[33] We could not locate a TS directly connecting **2** to **4**. Subsequently, the bornyl cation may rearrange spontaneously to give the camphyl cation, **5**. The relative energy of the camphyl cation is −31.50 kcal/mol. **5** may be characterized as a nonclassical cation, having a C1−C6 distance of 2.083 Å, a C2−C6 bond length of 1.643 Å, and an almost complete C2 sp_2 hybridization (Figures 2 and S1). An IRC calculation confirmed that bornyl is a TS connecting pinyl and camphyl. The lower stability of the pinyl cation relative to camphyl may be attributed to ring strain of the four-membered ring in **3**. A 1,2 methyl shift with an energy barrier of 8.45 kcal/mol yields **6**, a tertiary cation that is isoenergetic to the camphyl cation. Indeed, compounds **5** and **6** are enantiomers. Alternatively, the bornyl cation may rearrange through a C2 electrophilic addition to C4, releasing a tertiary cation as a leaving group, yielding compound **7**, with a relative energy of −9.08 kcal/mol. This latter process has a barrier of 26.27 kcal/mol. Interestingly, this suggests that the bornyl cation is a possible bifurcation point (Figure 2), although this is difficult to prove.[90] IRC calculations indicate that the TS connecting bornyl, **4**, to **7** is along the reaction path **3** → **4** → **7** (Figures 2 and 3). The bornyl structure obtained from the IRC calculation,

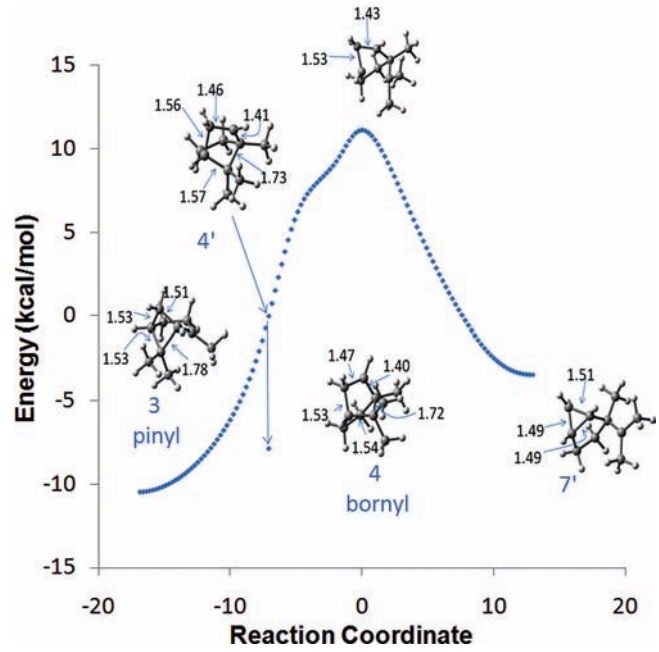


Figure 3. Intrinsic reaction coordinate plot for the transformation between **3** and **7** in the gas phase.

4′ (a prime indicates that the structure was obtained from the IRC path and is slightly different from the one obtained via optimization), is 7.8 kcal/mol above the bornyl structure, **4**, obtained from direct TS optimization (Figure 3).

For the pinyl cation one may envision a 1,2 rearrangement of the C3 bridge yielding the fenchyl cation, **8**, through a C1−C3 bond formation. This secondary cation, with a relative energy of −12.02 kcal/mol, is found to be a transient species in the gas phase, corresponding to a TS between pinyl cation, **3**, and cation **9**. This was confirmed by IRC calculations. The geometry of **8** indicates a partially nonclassical nature: the C1−C3 bond

(90) Ess, D. H.; Wheeler, S. E.; Iafe, R. G.; Xu, R.; Celebi-Olcum, N.; Houk, K. N. *Angew. Chem., Int. Ed.* **2008**, *47*, 7592–7601.

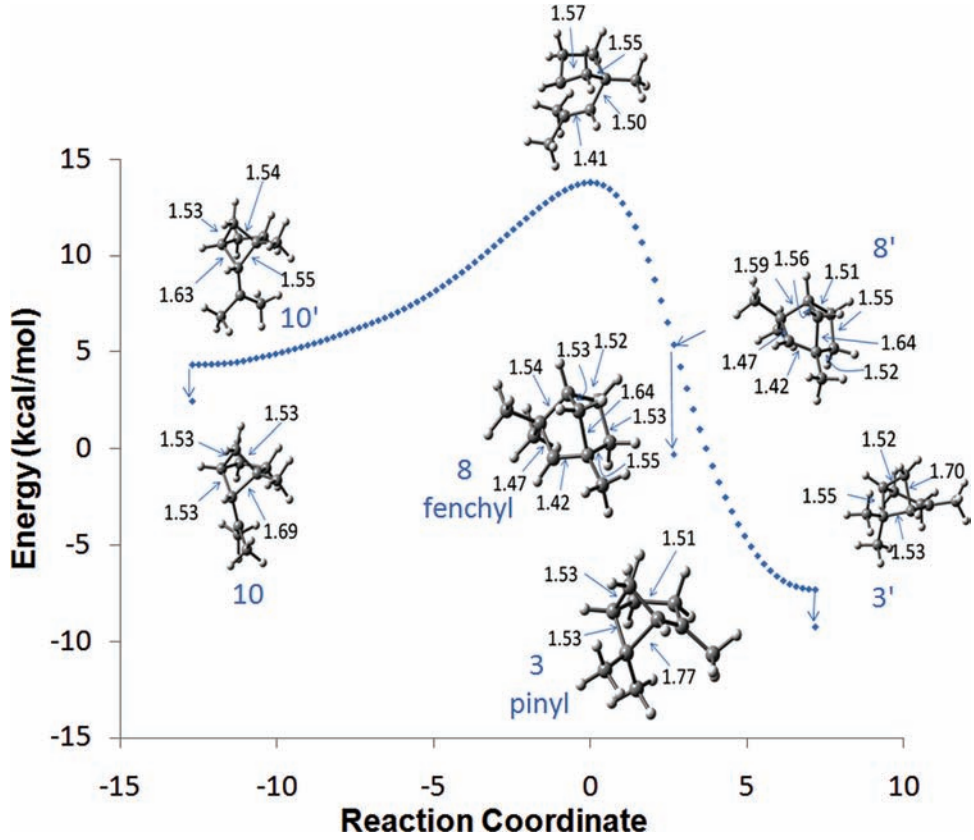


Figure 4. Intrinsic reaction coordinate plot for the transformation between **3** and **10** in the gas phase.

distance is 1.637 Å, while the C2−C3 distance is 2.094 Å (Figures 2 and S1). Additional pathways leading to fenchyl are possible,[91,92] although these were not pursued here. Cation **9** has a relative energy of −31.59 kcal/mol and distances C1−C6 of 2.096 Å and C2−C6 of 1.647 Å. The fenchyl cation may further rearrange to form two additional products, **10** and **11**. **10** (−9.73 kcal/mol) is generated by migration of C4 to form a bond between C2 and C4 with a barrier of 14.83 kcal/mol, while **11** (−35.42 kcal/mol) is formed by a 1,2 methyl migration between C8 and C2 with a barrier of 2.41 kcal/mol. **10** is characterized by a strained cyclobutane ring with a C1−C2 distance of 1.693 Å. In **11**, the C4−C5 bond distance is 1.642 Å, while the C5−C8 distance is 2.102 Å, indicative of the bridged nature of the C5 methylene. Thus, fenchyl presents an additional potential bifurcation point on the monoterpene PES serving as a branching point when approached from both **10** and **11** (Figures 2, 4, and 5). The connections between the two TSs linking fenchyl and **10** and **11**, respectively, were confirmed by IRC calculations. In Figure 4 we show the IRC path between pinyl, **3**, and **10**. Along this path, fenchyl, **8**, is located as a point connecting the TS and pinyl, thereby suggesting it is a possible bifurcation point. The fenchyl structure obtained from the IRC calculation, **8′**, is 6.8 kcal/mol above the fenchyl structure, **8**, obtained from direct TS optimization. Interestingly, the pinyl cation obtained from the IRC calculation, **3′**, is a conformer of the pinyl, **3**, obtained by direct optimization. In Figure 5 we display the IRC path between **9** and **11**. Here fenchyl is observed as a point connecting the TS and **11**, suggesting fenchyl may serve as an additional potential bifurcation point. The fenchyl structure obtained from the IRC calculation, **8′**, is 6.1 kcal/mol below the fenchyl structure, **8**, obtained from direct TS optimization. Nonetheless, we cannot rule out that there exists a dyotropic mechanism connecting **9** and **11**.

Returning to the terpinyl cation, **2**, several additional pathways are possible. The thujyl cation **13** is obtained through a sequence of steps beginning with a 1,2 hydride transfer from (4*R*)-α-terpinyl cation, **2**, to **12** with barriers ca. 2.6 and 4.2 kcal/mol for the C6exo and C6endo conformers, respectively. The relative energy of **12**, which may exist in two principal pre-enantiomeric (C4) endo/exo (C6) conformations, is −18.72 kcal/mol. Subsequently the latter closes to yield the fused thujyl cyclopropane, **13** (−26.19 kcal/mol), with a barrier of 3.47 kcal/mol. Interestingly, the thujyl cation may exist in two enantiomeric forms with the C3 methylene group on opposite sides of the ring pseudo plane. The cyclopropane ring has bond lengths C2−C4 1.69 Å, C2−C3 1.54 Å, and C3−C4 1.43 Å, indicative of considerable ring strain. The former two bonds are somewhat long, while the latter bond is short, which is indicative of a stabilizing interaction of the vacant p-orbital on the C1 carbocation center and valence cyclopropane orbitals.[52] This interaction is likely the source of the unusual stability of **13**, in spite of the cyclopropyl moiety. In comparison, **7** is considerably less stable, and indeed in this case the carbocation is not α to the cyclopropyl moiety. Interestingly, a 1,2 methyl migration in **7** yields a cyclopropylcarbinyl cation which is considerably more stable (results not shown).

The allylic cation phellandryl, **14**, can be formed directly from the terpinyl cation, **2**, by 1,3 hydride transfer yielding −27.99 and −29.70 kcal/mol for the exo and endo conformers,

(91) Haseltine, R.; Huang, E.; Ranganayakulu, K.; Sorensen, T. S.; Wong, N. *Can. J. Chem.* **1975**, *53*, 1876–1890.
(92) Huang, E.; Ranganayakulu, K.; Sorensen, T. S. *J. Am. Chem. Soc.* **1972**, *94*, 1779–1780.

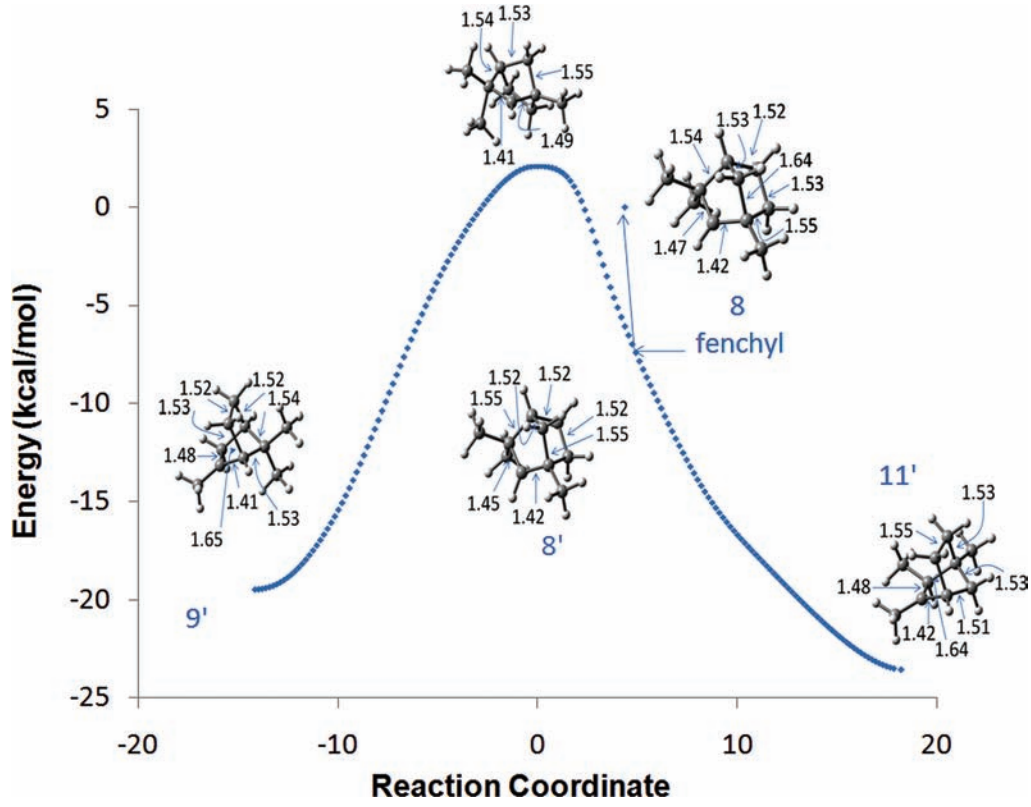


Figure 5. Intrinsic reaction coordinate plot for the transformation between **9** and **11** in the gas phase.

respectively. The associated barriers are 2.11 and 2.32 kcal/mol, respectively. **14** may also be formed from **12** via a 1,2 hydride transfer with a barrier of 3.16 and 3.96 kcal/mol for the endo and exo conformers, respectively. The allylic cation, **15**, can be formed directly from terpinyl cation, **2**, by 1,4 hydride transfer with energy −19.20 kcal/mol and a barrier of 10.66 kcal/mol. Conversely, **15** may be formed from **12** via a 1,3 hydride transfer with a barrier of 16.45 for the exo and 15.64 kcal/mol for the endo conformer.

The terpinyl cation, **2**, may undergo a 1,5 proton transfer in the C4-axial conformation to form a tertiary cation **16**, with relative energy of −13.04 and −10.92 kcal/mol for the exo and endo forms, respectively. Similar intramolecular proton transfers have been studied in several other terpenes, such as in the biosynthesis of trichodiene[26] and aristolochene.[38] It is worth mentioning that the C6 exo/endo ring puckering greatly influences the conformers' relative stability. The main difference between the two conformers seems to be steric congestion between the C3 and C6 methylene group in the endo form. The associated barriers are 11.91 and 11.61 kcal/mol for the exo and endo forms.

Fate of the Bornyl Cation in BPPS. The previous section dealt exclusively with the gas-phase reactions of monoterpenes. Although important insights may be obtained from such calculations, it is necessary to include the structural, electrostatic, and dynamic effects of the enzyme to understand the enzyme mechanism. In an attempt to address a key question in BPPS, namely the fate of the bornyl cation, we constructed a model of the enzyme based on a crystal structure that served as a starting point for extensive MD simulations, as detailed in the Computational Details section. The interactions in the enzyme/substrate/water system are treated via a hybrid QM/MM potential, while the dynamics of the system is treated by Newtonian MD simulations.

The question of the fate of the bornyl cation in BPPS was addressed by two different sets of MD simulations: (1) With activated MD, the bornyl cation in BPPS was used as a starting point for a series of unbiased MD simulations. In these simulations, 10 initial starting structures were taken from a single 1 ns MD simulation using QM(AM1)/MM (AM1 used due to efficiency and to avoid rearrangement of bornyl). Specifically, one structure was extracted every 100 ps from the 1 ns simulation and used as a starting point for the QM(BB1K)/MM MD simulation. Ten independent MD simulations, using these uncorrelated starting points, were then run with QM(BB1K)/MM to observe the fate of the bornyl cation. As will be described below, these simulations suggest that BPP is the main product, while camphyl is expected to be a minor product. (2) To quantify the relative stability of bornyl, BPP, and camphyl, we then performed free energy simulations using QM(BB1K)/MM. As will be discussed below, these simulations suggest that camphyl is a potential side product in this reaction.

The fate of the bornyl cation in BPPS was investigated by QM(BB1K)/MM MD simulations. The QM region comprises the carbocation, the diphosphate moiety, and three magnesium ions, while the remainder of the enzyme is treated by MM (Figure 6). A structural comparison between the average enzyme structures and the gas-phase geometries is presented in Table S173 (Supporting Information). A series of activated dynamics simulations initiated with the bornyl cation bound to BPPS were run. From a set of 10 uncorrelated MD simulations lasting 5 ps each, we observe that the lifetime of the 2-bornyl carbocation is less than 0.3 ps (Table 4). In nine of the simulated experiments, the carbocation recombines with the diphosphate moiety to form BPP within the first 300 fs of the MD simulations (Table 4, Figure 6). The initial C2−O3 distance is ca. 2.5−2.6 Å, while the bonded distance is 1.460 ± 0.047 Å. In one of the 10 simulations, the 2-bornyl

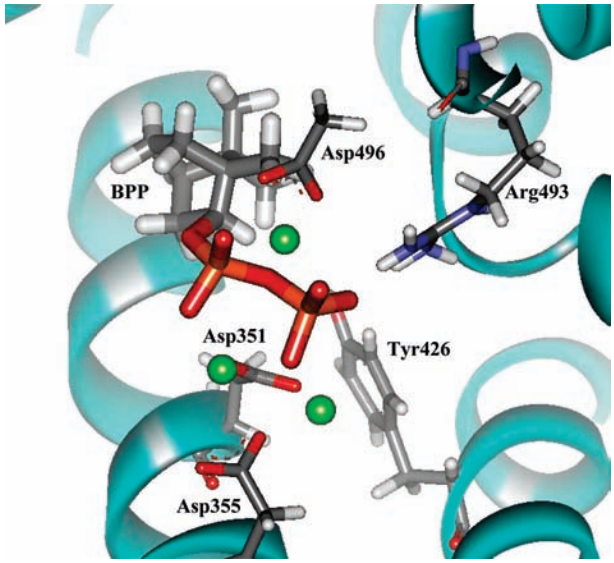


Figure 6. Molecular dynamics snapshot of (+)-bornyl diphosphate in bornyl diphosphate synthase.

Table 4. Bornyl Cation Lifetime, τ, from 10 Independent QM(BB1K)/MM Activated Molecular Dynamics Simulations[a]

simulation	τ (fs)	product
1	280	BPP[b]
2	111	BPP[b]
3	106	BPP[b]
4	144	BPP[b]
5	231	CAM[c]
6	258	BPP[b]
7	240	BPP[b]
8	106	BPP[b]
9	90	BPP[b]
10	279	BPP[b]

[a] End of bornyl cation lifetime was defined by the formation of a bond (distance <1.6 Å) including the C2 cation and the diphosphate moiety O3 atom (forming bornyl diphosphate) or C2 and C6 (forming camphyl). [b] Bornyl diphosphate. [c] Camphyl cation. See text for description of the nature of the cation.

carbocation rearranged to form the camphyl cation within ca. 200 fs (Table 4). BPP and the camphyl cation remained stable for the remainder of the 5 ps simulation. The average C1−C8 distance in bornyl is 1.591 ± 0.040 Å, while the average C2−C8 distance is 2.321 ± 0.042 Å. The average C2−C6 distance in camphyl is 1.608 ± 0.037 Å, while the average C1−C6 distance is 2.247 ± 0.065 Å.

Once formed, the camphyl cation may be subject to a nucleophilic attack by the diphosphate at the C2 position coupled to rearrangement to bornyl, en route to forming BPP. To quantify the likelihood of such a nucleophilic attack at the C2 position and concomitant phosphate−bornyl recombination, we computed the free energy profile for this step using PMF simulations. Inspection of the free energy profile in combination with the PMF MD trajectories reveals that the transformation between camphyl and bornyl diphosphate is a gradual process: initially a camphyl → bornyl rearrangement takes place, followed by bornyl−diphosphate recombination (Figure 6). As may be seen from the free energy profile in Figure 7, camphyl is more stable than bornyl by 5.3 kcal/mol. The formation of the C2−phosphate bond to yield BPP is highly exothermic, with

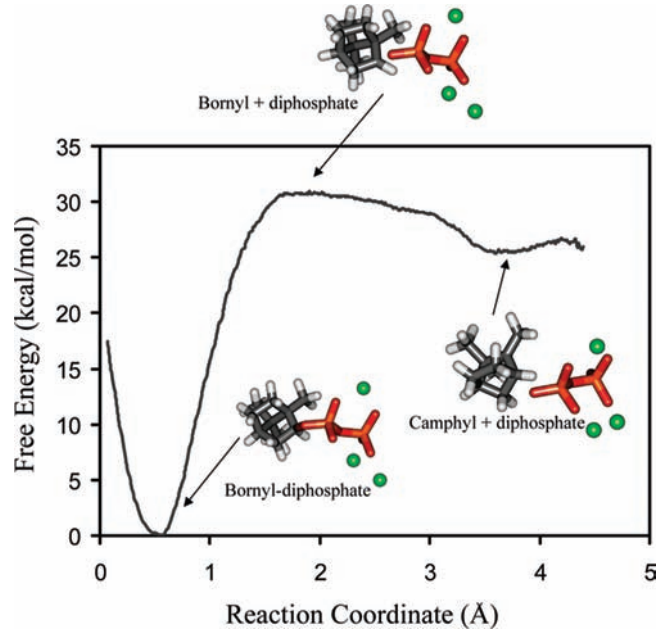


Figure 7. Computed potential of mean force for the interconversion of camphyl cation and (+)-bornyl diphosphate in bornyl diphosphate synthase.

a free energy of 30.9 kcal/mol relative to the unbound bornyl and diphosphate.

Discussion

A main finding in this study, based on the gas-phase results, is the multitude of possible side reactions in the carbocation cyclization reactions leading to (+)-bornyl diphosphate (BPP). Seemingly, the (4*R*)-α-terpinyl cation is an early branching point in monoterpene synthesis where numerous pathways with low energy barriers are possible.

Interestingly, we find that in the gas phase the (4*R*)-α-terpinyl cation may lose its C4 hydrogen through a low-barrier 1,2 hydride transfer (Scheme 2, Table 3). Subsequently, two enantiomers of the thujyl cation may be formed, both originating from (4*R*)-α-terpinyl. Clearly, both thujyl enantiomers could be formed by the (4*S*)-isomer as well. Hong and Tantillo have suggested that such a loss of stereochemistry in the sesquiterpene bisabolyl cation may explain how a terpene might bind only a single isomer but yield products that might indicate the binding of both isomers.[33] A similar scrambling occurs when camphyl, **5**, rearranges to its enantiomer, **6**, which might misleadingly suggest that the reaction pathway started from the (4*S*)-enantiomer, as **6** may be obtained directly from the (4*S*)-α-terpinyl isomer. Likewise, **6** might rearrange to yield **5**.

Additionally, we find that key cationic intermediates in the biosynthesis of several monoterpenes such as bornyl and fenchyl are TSs in the gas phase. These secondary cations may therefore not be viable intermediates in monoterpene synthases. These findings are in agreement with findings of Hong and Tantillo for other terpenes as well.[33] Nonetheless, one might envision that the enzyme electrostatic environment is able to stabilize such species. This question will be addressed below.

We also find several potential bifurcation points, which have recently been observed for more complex terpenes.[34] Bornyl (**4**) constitutes one possible bifurcation point, connected to **7** on the one hand and to pinyl (**3**) and camphyl (**5**) on the other hand (Figures 2 and 3). Fenchyl comprises an additional dual

possible bifurcation point that may be reached from **10** or **11** prior to branching into pinyl (**3**) or **9** (Figures 2, 4, and 5). This complex gas-phase PES and the unstable nature of the bornyl cation raise the question of the challenge posed to the enzyme BPPS. Thus far, the intimate details of the enzymatic reaction remain elusive. To understand the complex chemistry of terpene synthases, it is essential to go beyond gas-phase studies of terpene chemistry. Indeed, it is necessary to study these reactions within a flexible enzyme environment.

A key question in the BPPS reaction mechanism is the nature of the bornyl carbocation in the enzyme. The 2-norbornyl cation has been studied extensively in the gas phase and in solution by Schleyer et al. and others, and it was found that the classical and nonclassical forms of the carbocation are TSs and not stable species.[18−22,36] Indeed, it was stated over two decades ago that *"classical" and partially bridged 2-norbornyl cations are not expected to exist at least as isolated (gas phase) entities*.[18] In analogy, in our gas-phase studies we conclude that the secondary 2-bornyl cation is a transient species that immediately rearranges to camphyl or pinyl cations. The energy difference between camphyl and bornyl cations is ca. 20 kcal/mol. However, even the camphyl cation has a slight nonclassical nature, with a resemblance to the original 2-bornyl cation. On the other hand, the enzymatic synthesis of BPP is postulated to proceed via the 2-bornyl cation. In the enzyme one might envision that the 2-bornyl cation is stabilized due to electrostatic interactions in the active site. In the BPPS crystal structure with the 2-bornyl cation analogue 2-azabornane,[42] the ammonium moiety is in direct ionic interaction with the diphosphate, suggesting that the secondary cation might be stabilized. Additionally, the ammonium interacts with an active-site water molecule. However, Hong and Tantillo have postulated that secondary cations might be avoided in the biosynthesis of terpenes, on the basis of extensive gas-phase studies of terpenes.[33] Thus, it is of great interest to investigate whether the 2-bornyl carbocation might be a viable, stable intermediate in BPPS. To gain some insight into this, we performed activated MD simulations of the bornyl cation in BPPS as detailed in the Computational Details and Results sections. Specifically, we performed QM/MM NVT simulations of BPPS where the carbocation, diphosphate, and Mg^{2+} ions were treated at the BB1K level with a hybrid basis set (see Computational Details), while the remaining enzyme was treated classically. From a set of 10 uncorrelated MD simulations of 5 ps each, we observe that the lifetime of the 2-bornyl carbocation is less than 0.3 ps (Table 4). In nine of the simulated experiments, the carbocation recombined with the diphosphate moiety to form BPP within the first 300 fs of the MD simulations (Table 4, Figure 6). On the basis of visual inspection of the relative orientation of the substrate in the active site, we suggest that the diphosphate oxygen responsible for the recombination is likely the one initially bonding to C3 in (3*R*)-linalyl diphosphate, in agreement with experiment.[8] We find that the diphosphate steers the substrate at a subpicosecond rate via electrostatic forces. The average velocity of the substrate in the active site is ca. 1500 m/s, assuming that the diphosphate is held firmly in place by an extensive network of interactions. Such an electrostatic steering mechanism seems essential in order to avoid rapid rearrangement to camphyl cation or other related side products. Nonetheless, in one of the 10 simulations, the 2-bornyl carbocation rearranged to form the camphyl cation within ca. 200 fs (Table 4). This is in agreement with the work of Croteau and co-workers, who have demonstrated that (+)-camphene is produced in significant amounts by BPPS from common sage (*Salvia officinalis*).[44] Croteau and co-workers also observed formation of (+)-α-pinene and (±)-limonene. To estimate the relative stability of the carbocation precursors to these additional side-products, further simulations will have to be performed. Based on the activated MD simulations, we conclude that the 2-bornyl carbocation is not a stable intermediate in the BPPS reaction. Once formed, it is expected to undergo rapid recombination to yield the product BPP or rearrangement to form camphyl.

Analysis of the structural features of the bornyl cation in the enzyme provides additional insight into BPPS function. The average C1−C8 distance in bornyl is 1.591 ± 0.040 Å, while the average C2−C8 distance is 2.321 ± 0.042 Å. In comparison, these relative distances in the gas phase are 1.724 and 2.196 Å, respectively (Table S173, Supporting Information). The average C2−C6 distance in camphyl is 1.608 ± 0.037 Å, while the average C1−C6 distance is 2.247 ± 0.065 Å. In comparison, these relative distances in the gas phase are 1.643 and 2.091 Å, respectively. Thus, the carbocations display a more classical nature in the enzyme, due to the electrostatic environment in BPPS which favors charge localization.

Once formed, the camphyl cation may be subject to a nucleophilic attack by the diphosphate at the C2 position coupled to rearrangement to bornyl, en route to forming BPP. An attack at the C1 cationic position of camphyl cation seems less likely due to an unfavorable positioning in the active site, and the lack of rotational freedom due to the above-mentioned active-site water molecule. To quantify the likelihood of such a nucleophilic attack at the C2 position and concomitant phosphate−bornyl recombination, we computed the free energy profile for this step. Inspection of the free energy profile in combination with the MD trajectories reveals that the transformation between camphyl and bornyl diphosphate is a gradual process: initially a camphyl → bornyl rearrangement takes place, followed by bornyl−diphosphate recombination (Figure 6). As may be seen from the free energy profile in Figure 7, there is a slight barrier of 5.3 kcal/mol for the camphyl → bornyl rearrangement. This barrier is the free energy difference between camphyl and bornyl in the enzyme environment. In comparison, in the gas phase the free energy difference is ca. 20 kcal/mol (with added enthalpy and entropy contributions to the values in Table 2). We therefore conclude that the enzyme environment stabilizes the bornyl cation relative to camphyl by ca. 15 kcal/mol. Nonetheless, the bornyl cation is a TS in the enzyme as well, in agreement with the suggestion of Hong and Tantillo.[33] Interestingly, the formation of the C2−phosphate bond is highly exothermic and irreversible. Thus, once formed, the bornyl diphosphate is unlikely to rearrange back to yield the camphyl cation, although abrupt deprotonation of camphyl cation to yield camphene is possible. We note that additional mechanistic possibilities such as a concerted ring-closure of (4*R*)-α-terpinyl cation to 2-bornyl cation and capture by the diphosphate cofactor may also be envisaged for this process. Moreover, additional side reactions are possible as well.

The electrostatic steering mechanism suggested here is in agreement with the proposal by Peters and co-workers, who suggested that the ionized pyrophosphate group exerts a significant electrostatic effect in terpene synthases.[93] Indeed, it has been suggested that the electrostatic effect directs the product outcome in a so-called substrate-assisted catalytic specificity

(93) Xu, M.; Wilderman, P. R.; Peters, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7397–7401.

mechanism. According to this explanation, terpene synthases drive the reactions so that the intermediates have their carbocation oriented toward the pyrophosphate charges. Herein we show the first atomic-level description of such an electrostatic steering effect, where the terpene synthase guides the intermediate away from side products and toward the ultimate product.

Conclusion

This work describes a theoretical study of the carbocation-based cyclization starting from the geranyl allylic cation yielding (+)-bornyl diphosphate (BPP).

This study focused on the multitude of possible reaction paths for bornyl diphosphate in the gas phase. We find that the secondary cations, bornyl and fenchyl, are transition states. Additionally, these secondary cations are potential bifurcation points,[90] adding to the complexity of the potential energy surface. Indeed, the challenges posed to the enzyme bornyl diphosphate synthase (BPPS) are immense. To address one aspect of the BPPS catalytic strategy, we studied the nature of the 2-bornyl cation and the role of the diphosphate moiety. Indeed, we find that the BPPS environment stabilizes the 2-bornyl cation considerably relative to camphyl when compared to the gas phase, although bornyl is likely a transient species in the enzyme as well. The diphosphate moiety plays a crucial role in steering the intermediate 2-bornyl cation toward the BPP product formation. We speculate that this diphosphate−cation electrostatic interaction is a general tool for terpenes in modulating reaction products and specificity, in addition to the initial substrate folding in the active site. We note that the above conclusions indicate that the enzyme modifies the potential energy surface significantly due to a highly charged environment. Moreover, the enzyme tends to disfavor the slightly nonclassical carbocation nature of the bornyl and camphyl cations. Together, these findings suggest that mechanistic conclusions regarding enzymatic cyclization of terpenes based on gas-phase studies alone might be premature.

Considerable questions remain: How does the initial fold facilitate formation of α-terpinyl cation? Is the ring-closure of α-terpinyl cation anti-Markovnikov, yielding the 2-bornyl secondary cation? How does BPPS avoid the multitude of side reactions? Is the topology of the suggested bifurcation points preserved in the enzyme? These questions are currently under investigation in our laboratory, along with related questions on other terpenes.

Acknowledgment. This work has been supported by a start-up grant from Bar-Ilan University and the Alon fellowship from the Council for Higher Education—Planning and Budgeting Committee of the Israel Science Foundation.

Supporting Information Available: Figures showing monoterpene carbocations; IRC figures; tables of Cartesian coordinates for model compounds, gas-phase structures, and absolute energies; complete refs 67 and 71. This material is available free of charge via the Internet at http://pubs.acs.org.

JA910134X