## Supporting Information for

# Electrostatically Guided Dynamics - the Root of Fidelity in a Promiscuous Terpene Synthase? 

Dan Thomas Major* and Michal Weitman

Department of Chemistry and the Lise Meitner-Minerva Center of Computational Quantum Chemistry, Bar-Ilan University, Ramat-Gan 52900, Israel
majort@biu.ac.il

Figure S1. Two dimensional slice of the multidimensional classical potential of mean force ( $\mathrm{kcal} / \mathrm{mol}$ ) for the formation of linalyl cation from (3R)-linalyl diphosphate in BPPS. Continuous contour levels correspond to $5 \mathrm{kcal} / \mathrm{mol}$.


Figure S2. One dimensional cut of the multidimensional classical potential of mean force ( $\mathrm{kcal} / \mathrm{mol}$ ) for the formation of (4R)- $\alpha$-terpinyl cation from linalyl cation in BPPS.


Figure S3. Classical Potential of Mean Force ( $\mathrm{kcal} / \mathrm{mol}$ ) as a function of the C2-C7 and C2-O3 distances from $\mathrm{QM}(\mathrm{M} 06-2 \mathrm{X}) / \mathrm{MM}$ simulations.


Careful inspection of the free energy surface in Fig. 3 indicates that the branching ratio into BPP and camphyl cation is dependent on the detailed positioning of the bornyl cation in the active site. Specifically, using the C2-O3 distance as a measure of the proximity of the bornyl cation to the diphosphate, we suggest that at greater distances from O3 (ca. $3.0 \AA$ ) the steepest descent path on the PMF leads towards the camphyl cation. At intermediate C2-O3 distances (ca. $2.8 \AA$ ) a mixed distribution might be expected as the steepest descent path does not seem to clearly favor either direction. At short C2-O3 distances (<. $2.6 \AA$ ) we expect the formation of BPP to dominate, as the free energy surface gradient leads towards the formation of BPP. Interestingly, our simulations suggest that the position of the pinyl cation relative to the diphosphate moiety fluctuates somewhat with $\mathrm{C} 2-\mathrm{O} 3$ distances ranging between 2.5-3.0 $\AA$. Indeed, the free energy surface of the pinyl cation and the dividing surface region between pinyl and bornyl cations is rather flat with respect to the C2O3 distance (Fig. S3). Thus, the enzyme is unlikely to be able to accurately control the position distribution of pinyl or bornyl carbocations relative to the diphosphate moiety. Such lack of precise control is not surprising in the absence of directional interactions such as hydrogen bonds between the carbocation and the enzyme matrix.

A key question is the exact nature of the bornyl cation and what determines its fate. In our previous combined gas-phase and initial enzymatic study of the BPPS catalyzed reaction we suggested that the bornyl cation is not a stable intermediate but rather a short-lived species. To probe the reactivity of bornyl cation we ran an initial set of activated dynamics trajectories, where the starting point was a trajectory pool of the bornyl cation collected during the course of our free energy simulations. In total
we ran 75 trajectories commencing with the bornyl cation. The trajectories are presented in Fig. 7 and in all cases, camphyl cation or BPP is formed within a few hundred fs, suggesting that bornyl cation is a short-lived species in BPPS. Moreover, once a product is formed it remains stable for the remainder of the simulation. Inspection of the data in Fig. 7B suggests that formation of camphyl cation is favored at longer C2-O3 distances. At shorter C2-O3 distances there is a clear preference for BPP. At intermediate $\mathrm{C} 2-\mathrm{O} 3$ initial distances, we obtain a mixture of BPP and camphyl carbocation. Based on these simulations, we conclude that either BPP or camphyl cation may be formed, if the trajectories are initiated from the bornyl carbocation ensemble, while ignoring the phase-space history of bornyl cation.

Figure S4. Snapshot from QM(M06-2X)/MM molecular dynamics simulation of the terpinyl cation in BPPS.


Figure S5. Snapshot from QM(M06-2X)/MM molecular dynamics simulation of the pinyl cation in BPPS.


Figure S6. Snapshot from QM(M06-2X)/MM molecular dynamics simulation of the pinyl $\rightarrow$ bornyl cation transition state in BPPS.


Figure S7. Snapshot from QM(M06-2X)/MM molecular dynamics simulation of the bornyl cation in BPPS.


Figure S8. Snapshot from $\mathrm{QM}(\mathrm{M} 06-2 \mathrm{X}) / \mathrm{MM}$ molecular dynamics simulation of bornyl diphosphate in BPPS.


Figure S9. Snapshot from $\mathrm{QM}(\mathrm{M} 06-2 \mathrm{X}) / \mathrm{MM}$ molecular dynamics simulation of bornyl diphosphate in BPPS.


Figure S10. Lifetime distribution (fs) for the bornyl cation in BPPS during QM(M062 X )/MM activated dynamics MD simulations. The lifetime is defined as the difference between the time of bornyl cation formation $(\mathrm{R}(\mathrm{C} 3-\mathrm{C} 7)<1.7 \AA)$ and the formation of either BPP $(\mathrm{R}(\mathrm{C} 2-\mathrm{O} 3)<2.0 \AA)$ or the camphyl cation ( $\mathrm{R}(\mathrm{C} 3-\mathrm{C} 4)-\mathrm{R}(\mathrm{C} 2-$ C4) $>0.0 \AA$ ). The bornyl cation was fully formed prior to formation of either BPP or the camphyl cation in all simulations.


Table S1. Relative M06-2X/6-31+G(d,p) potential energies, including zero-point contributions ( $\mathrm{kcal} / \mathrm{mol}$ ), for selected monoterpene cations in the gas-phase.

| Compound | $\boldsymbol{\Delta E}$ |
| :--- | ---: |
| Linalyl | 0.00 |
| Terpinyl | -15.16 |
| Pinyl | -20.61 |
| Bornyl | -11.46 |
| Camphyl | -30.84 |

Table S2. Ensemble averaged distances ( $\AA$ ) between potential carbocation deprotonation sites and diphosphate or water oxygens in BPPS. The molecular dynamics simulations employed $\mathrm{QM}(\mathrm{M} 06-2 \mathrm{X}) / \mathrm{MM} .{ }^{a}$

| Carbocation | Atom | O 3 | O 7 | OW |
| :--- | :--- | :--- | :--- | :--- |
| (4R)- $\alpha$-Terpinyl | C 7 | $4.79 \pm 0.23$ | $5.84 \pm 0.31$ | $3.88 \pm 0.29$ |
|  | C 8 | $5.70 \pm 0.30$ | $6.00 \pm 0.35$ | $3.60 \pm 0.34$ |
|  | C 9 | $5.53 \pm 0.17$ | $6.26 \pm 0.40$ | $4.40 \pm 0.47$ |
| (+)-Pinyl | C 3 | $2.64 \pm 0.13$ | $5.00 \pm 0.16$ | $4.21 \pm 0.18$ |
|  | C 4 | $2.99 \pm 0.14$ | $6.44 \pm 0.15$ | $5.59 \pm 0.16$ |
|  | C 10 | $3.19 \pm 0.18$ | $4.44 \pm 0.24$ | $4.27 \pm 0.30$ |
| (+)-Camphyl | C 4 | $3.54 \pm 0.79$ | $6.05 \pm 0.49$ | $5.28 \pm 0.45$ |
|  | C 10 | $4.45 \pm 0.60$ | $4.28 \pm 0.76$ | $3.60 \pm 0.60$ |

[^0]Table S2 displays key ensemble averaged distances between deprotonation sites in (4R)- $\alpha$-terpinyl, $(+)$-pinyl, and (+)-camphyl cations and active site bases. In the case of (4R)- $\alpha$-terpinyl the active site water molecule is the most likely quenching agent, either via deprotonation or hydroxylation (Fig. S5). Deprotonation at either C8 or C9 would yield limonene. Deprotonation of the C4 position in (+)-pinyl cation would yield $\alpha-(+)$-pinene, while deprotonation at C 10 would result in $\beta-(+)$-pinene (Fig. S6). In both cases, the most likely base is the O3 of the diphosphate moiety or the active site water. However, in the pinyl cation orientations we observed in our simulations, there is no active site base available for anti-periplanar deprotonation, and further reorientation would be necessary prior to deprotonation. This supports the
small amount of pinene formed in BPPS. (+)-camphyl may undergo deprotonation at C6 by the diphosphate O3, which is directly coupled to a C1-C6 bond formation, to produce (+)- $\beta$-camphene (Fig. S10). Alternatively, the active site water molecule may deprotonate the camphyl cation at C 10 to yield $(+)-\alpha$-camphene. Our current simulations support such a deprotonation process.

Table S3. Relative M06-2X/6-31+G(d,p) potential energies, including zero-point contributions ( $\mathrm{kcal} / \mathrm{mol}$ ), for selected carbocation models in the gas-phase.

|  | M06-2X | BB1K $^{a}$ | MP4SDTQ $^{a}$ |
| :--- | :--- | :--- | :--- |
| Ia | 5.52 | 5.96 | 6.67 |
| Ib | 0.00 | 0.00 | 0.00 |
| IIa | 0.35 | 0.28 | 0.48 |
| IIb | 3.43 | 3.71 | 3.82 |
| IIc | 0.00 | 0.00 | 0.00 |
| IId | 20.23 | 21.06 | 19.83 |
| IIe | 6.06 | 6.70 | 6.83 |
| IIf | 5.74 | 7.30 | 6.73 |
| IIIa | 0.00 | 0.00 | 0.00 |
| IIIb | 14.45 | 14.35 | 14.71 |
| IIIc | 35.02 | 35.98 | 34.49 |
| IIId | 10.76 | 11.22 | 10.98 |





3a: t-butyl cation ( $C_{3 v}$ )


3d: 2-butyl cation ( $C_{1}$ )
${ }^{a}$ Ref. 1.

Table S4. Cartesian coordinates ( $\AA$ ) linalyl cation optimized at the M062X/6$31+G(d, p)$ level.

|  | X | Y | Z |
| :---: | :---: | :---: | :---: |
| C | -2.70937 | 0.86646 | -0.81408 |
| H | -3.11999 | -0.0825 | -0.48279 |
| H | -2.66434 | 1.03046 | -1.88663 |
| C | -2.28377 | 1.82012 | 0.04589 |
| H | -1.88863 | 2.75557 | -0.34186 |
| C | -2.3668 | 1.68831 | 1.46799 |
| C | -2.89615 | 0.48647 | 2.14859 |
| H | -2.87277 | -0.42669 | 1.5567 |
| H | -2.38282 | 0.32994 | 3.10156 |
| H | -3.94457 | 0.70817 | 2.40418 |
| C | -1.85092 | 2.76863 | 2.29219 |
| H | -1.93695 | 3.74802 | 1.81682 |
| H | -2.27939 | 2.77801 | 3.29626 |
| C | -0.27675 | 2.4588 | 2.45325 |
| H | -0.15939 | 1.4671 | 2.89595 |
| H | 0.17166 | 2.44061 | 1.45678 |
| C | 0.33911 | 3.53492 | 3.291 |
| H | 0.68481 | 4.41315 | 2.74938 |
| C | 0.48217 | 3.4999 | 4.62623 |
| C | 1.1596 | 4.62689 | 5.35484 |
| H | 0.47783 | 5.06944 | 6.08946 |
| H | 2.02605 | 4.25251 | 5.91054 |


| H | 1.49666 | 5.41076 | 4.67447 |
| :--- | ---: | ---: | ---: |
| C | 0.00708 | 2.37565 | 5.50439 |
| H | -0.50374 | 1.57592 | 4.96383 |
| H | 0.85252 | 1.9345 | 6.04314 |
| H | -0.67989 | 2.76377 | 6.26441 |

Energy (a.u.) =-390.806391

NImag $=0$

Table S5. Cartesian coordinates ( $\AA$ ) of terpinyl cation optimized at the M062X/6$31+G(d, p)$ level.

|  | X |  | Z |  |
| :--- | ---: | ---: | ---: | :---: |
| C | -2.03454 | 5.27827 | 3.78273 |  |
| H | -2.34151 | 4.72727 | 4.67983 |  |
| H | -2.18946 | 6.33889 | 3.99707 |  |
| C | -2.85124 | 4.8364 | 2.59359 |  |
| H | -3.6327 | 5.48951 | 2.21813 |  |
| C | -2.60804 | 3.64472 | 2.04214 |  |
| C | -3.36205 | 3.06959 | 0.88148 |  |
| H | -4.13956 | 3.75311 | 0.53586 |  |
| H | -2.68752 | 2.85774 | 0.04455 |  |
| H | -3.83542 | 2.1221 | 1.16132 |  |
| C | -1.50858 | 2.81053 | 2.66188 |  |
| H | -1.8398 | 2.40714 | 3.62865 |  |
| H | -1.26383 | 1.94485 | 2.03728 |  |


| C | -0.20743 | 3.60143 | 2.84828 |
| :--- | ---: | ---: | ---: |
| H | 0.50561 | 2.98062 | 3.41119 |
| H | 0.28897 | 3.84452 | 1.90704 |
| C | -0.54454 | 5.07878 | 3.50528 |
| H | -0.16461 | 5.8153 | 2.79711 |
| C | 0.39586 | 4.73543 | 4.52119 |
| C | -0.01828 | 4.06922 | 5.7819 |
| H | -0.88339 | 3.41626 | 5.65518 |
| H | 0.8031 | 3.5324 | 6.25745 |
| H | -0.33447 | 4.87856 | 6.4585 |
| H | 1.82501 | 5.09816 | 4.36081 |
| H | 2.07022 | 5.43731 | 3.35365 |
| H | 1.98963 | 5.94119 | 5.05035 |

Energy (a.u.) $=-390.830548$

NImag $=0$

Table S6. Cartesian coordinates ( $\AA$ ) of pinyl cation optimized at the M062X/6$31+G(d, p)$ level.

|  | X |  | Z |  |
| :--- | ---: | ---: | ---: | :---: |
| C | -2.46052 | 3.61079 | 4.62474 |  |
| H | -2.89686 | 2.60937 | 4.5758 |  |
| H | -2.84485 | 4.11883 | 5.50729 |  |
| C | -2.65298 | 4.45059 | 3.36753 |  |


| H | -3.35916 | 5.27689 | 3.31289 |
| :---: | :---: | :---: | :---: |
| C | -2.13458 | 3.9403 | 2.17704 |
| C | -2.33025 | 4.6272 | 0.87804 |
| H | -2.80845 | 5.60232 | 0.98189 |
| H | -1.38481 | 4.71832 | 0.3338 |
| H | -2.97767 | 3.98651 | 0.26331 |
| C | -1.36672 | 2.66022 | 2.2418 |
| H | -2.13153 | 1.87022 | 2.27863 |
| H | -0.79282 | 2.49846 | 1.32635 |
| C | -0.48327 | 2.57864 | 3.52636 |
| H | -0.59085 | 1.59823 | 3.99463 |
| H | 0.56922 | 2.69677 | 3.2568 |
| C | -0.92327 | 3.67062 | 4.49257 |
| H | -0.33801 | 3.67089 | 5.41652 |
| C | -0.95701 | 5.02666 | 3.76406 |
| C | -1.25513 | 6.24448 | 4.62535 |
| H | -0.31851 | 6.51916 | 5.1251 |
| H | -1.5623 | 7.09542 | 4.01237 |
| H | -2.00683 | 6.06917 | 5.39446 |
| C | 0.10891 | 5.37023 | 2.7409 |
| H | 0.37111 | 4.56344 | 2.05574 |
| H | -0.16695 | 6.25552 | 2.16107 |
| H | 1.01344 | 5.6222 | 3.3074 |

Energy (a.u.) $=-390.839234$

NImag $=0$

Table S7. Cartesian coordinates ( $\AA$ ) bornyl cation optimized at the M062X/6$31+G(d, p)$ level.

|  | X | Y | Z |
| :---: | :---: | :---: | :---: |
| C | -1.43031 | 0.66942 | -1.15186 |
| H | -2.33636 | 0.15738 | -1.51271 |
| H | -1.45314 | 1.66489 | -1.60796 |
| C | -0.20903 | -0.05461 | -1.57826 |
| H | 0.2273 | 0.07249 | -2.57241 |
| C | 0.44509 | -0.73956 | -0.53345 |
| C | 1.8231 | -1.31045 | -0.76965 |
| H | 2.46224 | -0.61914 | -1.32555 |
| H | 2.30671 | -1.53959 | 0.18324 |
| H | 1.7555 | -2.2455 | -1.33345 |
| C | -0.56381 | -1.64566 | 0.25633 |
| H | -0.98942 | -2.37871 | -0.43254 |
| H | 0.01858 | -2.20191 | 0.99428 |
| C | -1.62759 | -0.72326 | 0.89291 |
| H | -2.64233 | -0.99465 | 0.59539 |
| H | -1.59001 | -0.76338 | 1.98309 |
| C | -1.24032 | 0.6655 | 0.3749 |
| H | -1.73239 | 1.49448 | 0.88776 |
| C | 0.31054 | 0.68806 | 0.45918 |
| C | 0.99757 | 1.93592 | -0.09502 |
| H | 0.57456 | 2.31929 | -1.026 |
| H | 0.92299 | 2.73195 | 0.65189 |


| H | 2.06223 | 1.74578 | -0.25815 |
| :--- | ---: | ---: | ---: |
| C | 0.9392 | 0.35404 | 1.80339 |
| H | 0.62503 | -0.59713 | 2.23022 |
| H | 2.03004 | 0.37141 | 1.7411 |
| H | 0.63887 | 1.14844 | 2.49876 |

Energy (a.u.) $=-390.824652$

NImag $=1$

Table S8. Cartesian coordinates $(\AA)$ of camphyl cation optimized at the M062X/6$31+G(d, p)$ level.

|  | X | Y | Z |
| :--- | :--- | :--- | :---: |
| C | 1.25582 | -0.44709 | -1.32746 |
| H | 2.33285 | -0.58146 | -1.44711 |
| H | 0.77142 | -0.66294 | -2.28114 |
| C | 0.90417 | 0.95102 | -0.81189 |
| H | 1.21322 | 1.86292 | -1.3174 |
| C | -0.39115 | 0.87061 | -0.22085 |
| C | -1.27598 | 2.00987 | 0.07941 |
| H | -2.08435 | 1.96204 | -0.66732 |
| H | -1.7511 | 1.91847 | 1.05952 |
| H | -0.77765 | 2.97369 | -0.02957 |
| C | 1.51528 | 0.85544 | 0.73891 |
| H | 2.47613 | 1.36173 | 0.61596 |
| H | 0.96061 | 1.42105 | 1.49498 |


| C | 1.58061 | -0.6555 | 1.00718 |
| :--- | ---: | ---: | ---: |
| H | 2.60171 | -1.02551 | 0.89602 |
| H | 1.23762 | -0.91467 | 2.00912 |
| C | 0.68821 | -1.20836 | -0.11877 |
| H | 0.65648 | -2.29583 | -0.19979 |
| C | -0.71758 | -0.56202 | 0.06439 |
| C | -1.69304 | -1.00845 | -1.06403 |
| H | -1.33012 | -0.79564 | -2.07106 |
| H | -1.83035 | -2.0891 | -0.97095 |
| H | -2.66698 | -0.52977 | -0.93778 |
| C | -1.38575 | -0.80262 | 1.42155 |
| H | -0.85053 | -0.3248 | 2.24626 |
| H | -2.41368 | -0.43098 | 1.42459 |
| H | -1.42238 | -1.87762 | 1.6194 |

Energy (a.u.) $=-390.855540$

NImag $=0$

## Methods.

Potential Energy Surface. The potential energy surface (PES) in the current study is described by a hybrid QM/MM Hamiltonian. ${ }^{2,3}$ In this treatment, the reactive fragment wherein the chemistry occurs is treated via QM to allow for bond breaking and forming, while the environmental effects of the enzyme and solvent are included via MM. The QM region contains the carbocation, diphosphate, and $\mathrm{Mg}^{2+}$ ions in one of the active sites and is described by the M06-2X functional, ${ }^{4}$ in conjunction with a hybrid basis set. ${ }^{1}$ On H and C atoms we employed the $6-31 \mathrm{G}(\mathrm{d})$ basis set, ${ }^{5}$ on O atoms we employed $6-31+\mathrm{G}(\mathrm{d})$, while on P and Mg we employed a minimal STO-3G Gaussian basis set. ${ }^{5}$ In simulations of deprotonation of camphyl cation, the active site water was also treated as QM. The MM part is described by the CHARMM22 force field, ${ }^{6}$ while water molecules are treated by the TIP3P model, with van der Waals parameters modified as implemented in the standard CHARMM topology files. ${ }^{7}$ The electrostatic part of the QM-MM interactions are treated with electrostatic embedding, ${ }^{3}$ wherein the MM partial charges are introduced into the Kohn-Sham equations, allowing polarization of the QM region. Additionally, van der Waals interactions between QM and MM atoms are included. No non-bonded interaction cutoffs were employed in computing the $\mathrm{QM} / \mathrm{MM}$ interactions.

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. ${ }^{8}$ The enzyme is a homodimer comprised of two a helical domain units containing 598 amino acids each. Each monomer possesses an independent active site pocket, which is sequestered from water, although a single specific water molecule is present in the binding site. The
enzyme was solved at $2.4 \AA$ 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, ${ }^{9,10}$ while the hydrogen atoms of the substrate were added manually. The protonation states of all ionizable residues were assigned based on physiological pH . His residues were modeled as neutral or protonated moieties with hydrogen positioned either at Nd or Ne 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.

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. ${ }^{11}$ The enzyme was soaked in a sphere with a radius of $20 \AA$, and the Langevin stochastic boundary ranged from 16-20 $\AA$. This size system gives behavior similar to that of our previous $24 \AA$ BPPS setup. Counterions were added beyond the stochastic boundary region to neutralize the system. 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. ${ }^{12}$ TIP3P water and protein hydrogens were constrained using the SHAKE algorithm. ${ }^{12}$ For the MM group atoms, the non-bonded interactions were set to zero at distances beyond $14 \AA$. The electrostatic forces were shifted to zero from a distance of $12 \AA$, while the vdW interaction energy was switched to zero at $12 \AA$. No cutoff was employed in computing the $\mathrm{QM} / \mathrm{MM}$ interactions. The system was initially heated slowly to 298 K during the course of 25 psec , and thereafter equilibrated for 75 psec .

Subsequently, the system was further simulated for 1 nsec . During these simulations, the QM region was composed of the carbocation only and treated by the AM1 Hamiltonian. ${ }^{13}$

Multidimensional Adaptive Umbrella Sampling. The reaction coordinates were defined by means of geometric variables relevant to different stages of the reaction mechanism (See Scheme 1 and Fig. S5-S10). We performed three separate sets of multidimensional umbrella sampling studies using the RXNCOR module in CHARMM: (i) Starting from (3R)-linalyl diphosphate and ending with terpinyl cation formation. (ii) Starting from the terpinyl cation and ending with BPP or camphyl cation. (iii) Deprotonation of camphyl cation to yield camphene. In set (i) we defined three reaction coordinate variables: $\quad \zeta_{1}=R(\mathrm{Cl}-\mathrm{O} 3), \quad \zeta_{2}=R(\mathrm{C} 3-\mathrm{O} 3)$, and $\zeta_{3}=R(\mathrm{Cl}-\mathrm{C} 6)$. This set accounts for heterolytic cleavage of geranyl diphosphate and (3R)-linalyl diphosphate to yield the linalyl-diphosphate ion-pair, as well as formation of the terpinyl cation. In set (ii) we defined four reaction coordinates: $\zeta_{1}=R(\mathrm{C} 2-\mathrm{C} 7), \quad \zeta_{2}=R(\mathrm{C} 3-\mathrm{C} 7), \quad \zeta_{3}=R(\mathrm{C} 3-\mathrm{C} 4)-R(\mathrm{C} 2-\mathrm{C} 4), \quad$ and $\zeta_{4}=R(\mathrm{C} 2-\mathrm{O} 3)$. The former two coordinates in set (ii) represent formation of pinyl and bornyl cations, respectively, from terpinyl cation, whereas the latter two define bornyl $\rightarrow$ camphyl migration and BPP formation, respectively. In set (iii) we defined two reaction coordinates, describing proton transfer from camphyl cation to the active site water molecule, and from water to the diphosphate moiety. Specifically, $\zeta_{1}=R(\mathrm{Cl0}-H)-R\left(H-O_{W}\right)$ and $\zeta_{2}=R\left(O_{W}-H_{W}\right)-R\left(H_{W}-O 7\right)$.

To facilitate uniform sampling along the reaction coordinates, biasing potentials may be added to the potential energy of the system along a selected reaction coordinate,
$\zeta, U^{\text {bias }}(\zeta)=U^{u m b r}(\zeta)+U^{\text {harm }}(\zeta)$. The umbrella potential, $U^{u m b r}(\zeta)$, is defined as $-W^{c l}(\zeta)$, which essentially removes the reaction barriers, enabling uniform sampling. Furthermore, the reaction coordinate is divided into regions (windows) centered around $\zeta_{0}$, and a harmonic potential, $U^{\text {harm }}(\zeta)=k\left(\zeta-\zeta_{0}\right)^{2} / 2$, is added. Practically, since neither $W^{c l}(\zeta)$, which is required for $U^{u m b r}(\zeta)$, nor the optimal force constant, $k$, in $U^{\text {harm }}(\zeta)$, are known initially, the PMF is obtained by employing adaptive umbrella sampling molecular dynamics simulations. ${ }^{14}$ The umbrella simulations provide a biased PMF. A multidimensional version of 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. ${ }^{15,16}$ The reaction coordinates where evenly spaced with $0.10-0.25 \AA$ separation between centers of windows. Each window was briefly equilibrated for 2 ps and sampled for $5-10 \mathrm{ps}$ each. The combined QM(M06-2X)/MM PMF simulation time was over 600 ps .

Nuclear Quantum Effects Simulations. To account for zero-point vibrational energy and tunneling effects in the deprotonation of camphyl cation, we used Monte Carlo path-integral simulations. ${ }^{17}$ In the path-integral simulation approach, classical particles are replaced by a ring of quasi-particles (beads) to describe quantum delocalization. Previously, we have implemented the primitive approximation (PA) path-integral method within CHARMM. ${ }^{18-21}$ Recently, we also implement a higher order path-integral method based on work of S. A. Chin (CH) together with the staging algorithm. ${ }^{17}$ Using the CH approach, we included nuclear quantum corrections with the same hybrid quantum mechanics/molecular mechanics ( $\mathrm{QM} / \mathrm{MM}$ ) potential
energy surface used for all the simulations in this work. In practice we employed Eq. S1 to compute the nuclear quantum corrections at a temperature of 298 K as implemented in a development version of CHARMM. ${ }^{17}$

$$
\begin{equation*}
e^{-\tau(T+V)}=e^{-v_{1} \tau W_{i}} e^{-t_{2} \tau T} e^{-v_{2} \tau W_{j}} e^{-t_{1} \tau T} e^{-v_{1} \tau W_{k}} e^{-2 t_{0} \tau T} \tag{S1}
\end{equation*}
$$

Here $\tau=\beta / P, P$ is the number of beads, $V$ is the potential energy of the system, $W_{i}$ is an effective potential, $T$ is the kinetic energy, and $v_{1}, v_{2}, t_{0}, t_{1}$, and $t_{2}$.are parameters. ${ }^{17}$

In all simulations the bead sampling was performed by simultaneously moving all beads at each path-integral step using the staging algorithm. The number of classical configurations employed was ca. 300 while 10 Monte Carlo path-integral steps were performed at each classical configuration.

## References:

(1) Weitman, M.; Major, D. T. J. Am. Chem. Soc. 2010, 132, 6349.
(2) Warshel, A.; Levitt, M. J. Mol. Biol. 1976, 103, 227.
(3) Gao, J. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials; VCH: New York, 1995; Vol. 7.
(4) Zhao, Y.; Truhlar, D. Theor. Chem. Acc. 2008, 120, 215.
(5) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. Ab Initio Molecular Orbital Theory; John Wiley \& Sons: New York, 1986.
(6) MacKerell, J., A. D. et al. J. Phys. Chem. B 1998, 102, 3586.
(7) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J. Chem. Phys. 1983, 79, 926.
(8) 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.
(9) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J. Comp. Chem. 1983, 4, 187.
(10) Brooks, B. R.; III, C. L. B.; Jr., A. D. M.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. J. Comp. Chem. 2009, 30, 1545.
(11) Brooks III, C. L.; Brünger, A.; Karplus, M. Biopolymers 1985, 24, 843.
(12) Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Oxford University Press: Oxford, U.K., 1987.
(13) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. J. Am. Chem. Soc. 1985, 107, 3902.
(14) Torrie, G. M.; Valleau, J. P. J. Comp. Phys. 1977, 23, 187.
(15) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. J. Comp. Chem. 1992, 13, 1011.
(16) Doron, D.; Kohen, A.; Major, D. T. J. Chem. Theory Comput. 2012, 8, 2484.
(17) Azuri, A.; Engel, H.; Doron, D.; Major, D. T. J. Chem. Theory Comp. 2011, 7, 1273.
(18) Gao, J. L.; Wong, K. Y.; Major, D. T. Journal of Computational Chemistry 2008, 29, 514.
(19) Major, D. T.; Gao, J. L. Journal of Molecular Graphics and Modelling 2005, 24, 121.
(20) Major, D. T.; Gao, J. L. Journal of Chemical Theory and Computation 2007, 3, 949.
(21) Major, D. T.; Garcia-Viloca, M.; Gao, J. L. Journal of Chemical Theory and Computation 2006, 2, 236.


[^0]:    ${ }^{a}$ The distances were averaged over 4,000-10,000 trajectory frames.

