BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

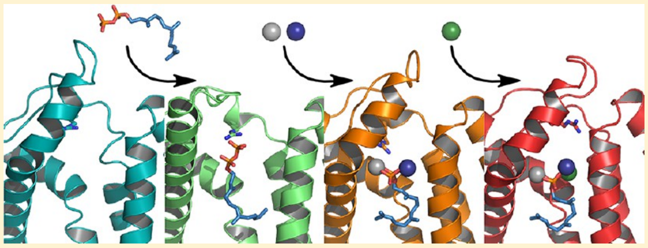
Conformational Change and Ligand Binding in the Aristolochene Synthase Catalytic Cycle

Marc W. van der Kamp,*,[†] Jitnapa Sirirak,[†] Jolanta Żurek,[†,‡] Rudolf K. Allemann,[‡] and Adrian J. Mulholland*,[†]

[†]Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.

[‡]School of Chemistry and Cardiff Catalysis Institute, Cardiff University, Park Place, Cardiff CF10 3AT, Wales, U.K.

Ⓢ *Supporting Information*



ABSTRACT: Terpene synthases are potentially useful biocatalysts for the synthesis of valuable compounds, such as anticancer drugs and antibiotics. The design of altered activities requires better knowledge of their mechanisms, for example, an understanding of the complex conformational changes that are part of their catalytic cycle, how they are coordinated, and what drives them. Crystallographic studies of the sesquiterpene synthase artistolochene synthase have led to a proposed sequence of ligand binding and conformational change but have provided only indirect insight. Here, we have performed extensive molecular dynamics simulations of multiple enzyme−ligand complexes (over 2 μs in total). The simulations provide clear evidence of what drives the conformational changes required for reaction. They support a picture in which the substrate farnesyl diphosphate binds first, followed by three magnesium ions in sequence, and, after reaction, the release of aristolochene and two magnesium ions followed by the final magnesium ion and diphosphate. Binding of farnesyl diphosphate leads to an increased level of sampling of open conformations, allowing the first two magnesium ions to bind. The closed enzyme conformation is maintained with a diphosphate moiety and two magnesium ions bound. The open-to-closed transition reduces flexibility around the active site entrance, partly through a lid closing over it. The simulations with all three magnesium ions and farnesyl diphosphate bound provide, for the first time, a realistic model of the Michaelis complex involved in reaction, which is inaccessible to experimental structural studies. These insights could help with the design of altered activities in a range of terpene synthases.

Terpene synthases make up a class of enzymes involved in the biosynthesis of the largest and most structurally diverse class of natural products, the terpenoids.[1] The enzymes convert linear substrates such as geranyl diphosphate, farnesyl diphosphate (FDP), and geranylgeranyl diphosphate into complex cyclic hydrocarbons often with very high regio- and stereospecificity. Many resulting natural terpenoids have practical applications, including artemisinin (antimalarial), camphor (e.g., moth repellant and antimicrobial), gossypol (male contraceptive and antimalarial), steroids, and paclitaxel (anticancer). Sesquiterpene synthases catalyze the conversion of FDP into more than 300 different known natural sesquiterpene products that share the formula $C_{15}H_{24}$.[2] In addition, novel terpenoid compounds can be formed by using non-natural substrate analogues (see, e.g., ref 3) and by applying enzyme engineering to adjust product specificity.[4] The unique product specificity of each natural enzyme relies on specific choreographing of the conformational and chemical changes in the catalytic cycle. Mutagenesis studies have indeed indicated that control of the folding and dynamics of the substrate in the enzyme active site must be crucial for product specificity,[5,6] but it is often not understood how such control is achieved. In addition, the involvement of sesquiterpene synthases in catalysis is more active and multifaceted than previously thought; they go beyond simply providing a template for the substrate to fold up correctly and react.[2] To fulfill the potential of (sesqui)terpene synthases as biocatalysts for valuable and potentially novel terpenoid compounds, a better understanding of the way in which these enzymes influence substrate binding and reaction to control product specificity is needed. Such understanding may in turn provide routes for re-engineering enzymes to modify the terpenoids produced.

Structures of several sesquiterpene synthases from X-ray crystallography,[7−12] including the fungal aristolochene synthases from *Penicillium roqueforti* (PR-AS)[13] and *Aspergillus terreus* (AT-AS),[14] reveal that they all share a common α-helical structure (despite a lack of significant sequence similarity), the class I terpene synthase fold. The aristolochene synthases PR-AS and AT-AS in particular have been studied extensively.[5,13−25] Their sequences are very similar (61% identical and 85% similar), and their three-dimensional structures are closely related, with nearly identical active sites.[14] Details of the conversion of farnesyl diphosphate (FDP) into products by

Received: July 8, 2013
Revised: October 3, 2013
Published: October 9, 2013

dx.doi.org/10.1021/bi400898k | *Biochemistry* 2013, 52, 8094−8105

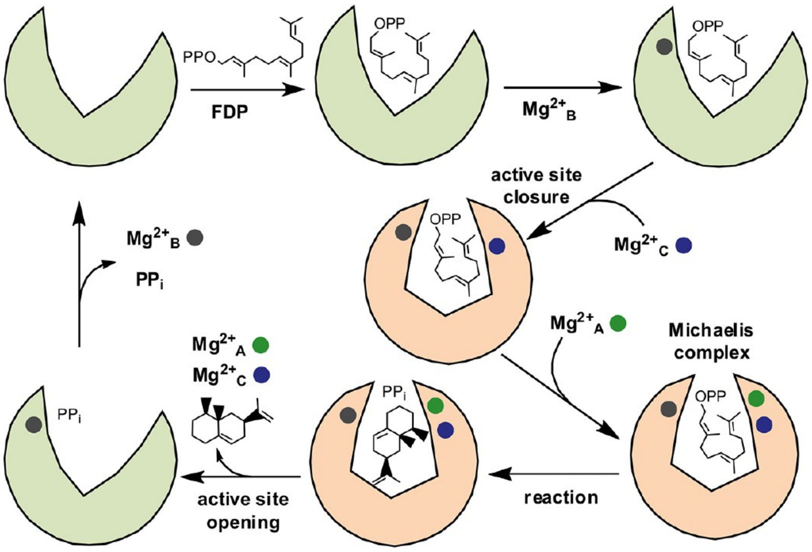


Figure 1. Ligand binding cycle proposed by Shishova et al.[25]

aristolochene synthases are shared with other sesquiterpene synthases, such as a common mechanism for formation of the active enzyme−substrate complex, FDP ionization, and cyclization,[7,9,10,16,25,26] and cation stabilization by the diphosphate moiety and aromatic amino acid residues.[7,18,20,23,27] The cyclization reaction is initiated through the expulsion of diphosphate (PP_i) from the substrate through its coordination to three magnesium ions, which are bound by two binding sites on opposite sides of the active site. Closure of the active site [proposed to be triggered by the binding of FDP and two Mg^{2+} ions on either side of the diphosphate moiety (see below)] is most likely a prerequisite for catalysis in all sesquiterpene synthases;[28] without such closure, the reaction, if it takes place at all, would have a much reduced specificity.

X-ray crystallography studies of AT-AS first revealed the closed conformation of a class I sesquiterpene synthase, in complex with diphosphate and three Mg^{2+} ions.[14] Subsequent crystallography studies with FDP and fluorinated analogues led to a hypothesis for the sequence of ligand binding in the formation of the active closed complex.[25] In particular, the open structures with Mg^{2+}B and either monofluorinated FDP (PDB entry 3BNY, chain D) or PP_i (PDB entry 3BNX, chain D) bound and the largely closed structure with Mg^{2+}B, Mg^{2+}C, and PP_i bound (PDB entry 3CKE, chain D, or 3CKEd) suggested that binding of Mg^{2+}C in addition to Mg^{2+}B and FDP triggers active site closure. Together with the previously determined closed (PDB entry 2OA6, chain D, or 2OA6d) and open (PDB entry 2E4O) crystal structures,[14] these structures suggest the ligand binding−unbinding sequence depicted in Figure 1. According to this proposal, first, FDP binds with its isoprenoid tail sequestered into the hydrophobic active site cavity, possibly together with Mg^{2+}B. Thereafter, binding of Mg^{2+}C triggers the conformational change from the open to the closed form. Subsequently, Mg^{2+}A binds to generate the catalytically active Michaelis complex and the chemical reaction commences. As the active site reopens, the hydrocarbon product, Mg^{2+}A, and Mg^{2+}C are released first, followed by the release of diphosphate and Mg^{2+}B. The crystallography studies did not, however, identify a productively bound substrate (i.e., the Michaelis complex), and the precise details of the sequence and mechanism of ligand binding and the related conformational change remain unknown.

In this work, we perform a series of molecular dynamics (MD) simulations on AT-AS with all combinations of ligands absent or present (FDP or PP_i and Mg^{2+} ions), starting from the closed conformation. At least three simulations of 30 ns were performed for each system, and certain runs were further extended to 100 or 200 ns (Table 1), for a grand total of 2.35 μs of simulation time. Analysis of the resulting trajectories provides detailed insight into ligand binding and conformational change as well as the identification of a likely Michaelis complex in aristolochene synthase. This potentially reactive complex, which has remained elusive so far, offers insights into the first reaction step, expulsion of diphosphate and cyclization. Such insights are important to fully understand the cascade that leads to the efficient formation of complex (sesqui)terpenoids and may provide routes for re-engineering enzymes to modify the terpenoids produced.

Table 1. Protein−Ligand Systems Simulated

apo[a]	MgB
FDP[b]	MgC
FDP and MgA	PP_i and MgA
FDP and MgB	PP_i and MgB
FDP and MgC	PP_i and MgC
FDP, MgA, and MgB	PP_i, MgA, and MgB
FDP, MgA, and MgC	PP_i, MgA, and MgC
FDP, MgB, and MgC	PP_i, MgB, and MgC
FDP, MgA, MgB, and MgC[c]	PP_i, MgA, MgB, and MgC

[a] Protein without any ligands. One of the three simulations was extended to 200 ns. [b] One of the three simulations was extended to 100 ns; three additional 30 ns simulations were run with the farnesyl chain in a different starting position. [c] Two additional simulations (100 and 30 ns in length) were run with Glu94 interacting with Mg^{2+}A and Mg^{2+}C (interaction with Mg^{2+}C was established spontaneously during equilibration; interaction with Mg^{2+}A was established after production simulation for 6.5 and 0.06 ns, respectively). Nine additional simulations of 30 ns of this system were run with the farnesyl chain in three alternative binding modes (three each), and with modified Mg^{2+} parameters.[34]

METHODS

Starting Structures and Parameters. All models were based on the X-ray crystal structure of aristolochene synthase from *A. terreus* (AT-AS) in the closed conformation, with PP_i and three Mg^{2+} ions bound,[14] chain D of PDB entry 2OA6 (2OA6d). All Asp, Glu, Lys, and Arg residues were modeled in their ionized forms. All His residues were modeled as neutral, with His43, -138, and -274 protonated on $N\delta1$ and His26, -77, -214, and -235 on $N\varepsilon2$, in line with the optimal hydrogen bonding network[29] calculated with the WHAT-IF web interface (http://swift.cmbi.ru.nl). Depending on the ligands included (see Table 1), PP_i and Mg^{2+} ions were deleted or PP_i was changed to FPP. The diphosphate moiety was modeled as fully deprotonated. The CHARMM27 force field[30] with CMAP corrections[31] was used. Parameters in line with this force field were used for PP_i and FPP (topology and parameters are included as Supporting Information). The TIP3P model adapted for use with CHARMM force fields[30] was used for water. All systems were solvated using VMD,[32] to form a large rectangular box extending at least 14 Å from the protein. Water molecules distant from the protein and each other were replaced with Na^+ ions to yield overall neutral systems. The farnesyl chain was built such that it is in line with cyclization to form germacrene A. In analogy with the cocrystal structure of AT-AS with Mg^{2+}B and 2F-FDP (PDB entry 3BNY, chain D), the farnesyl chain was attached to the diphosphate oxygen that points into the binding pocket closest to Mg^{2+}A (O3 in 2OA6d). In contrast to the conformation in PDB entry 3BNY (and a model of FDP bound to PR-AS[13]), the farnesyl chain itself was modeled such that the end of the chain is on the same side of the pocket as Mg^{2+}B (mode I). Preliminary simulations with the chain modeled the other way around, i.e., with the end of the chain on the side of Mg^{2+}A and Mg^{2+}C (mode II), lead to more linear farnesyl conformations that are not in line with cyclization (see also below). When the farnesyl chain is modeled in mode I, it bends at C5 and again at C8, closely aligned to the Phe87 and Tyr67 side chains (mutations of the equivalent Phe112 and Tyr92 in PR-AS to Ala lead to predominantly linear products[18,23]). To explore other possible FDP binding modes in the fully formed complex (i.e., with all three Mg^{2+} ions bound), we also simulated two alternative farnesyl conformations, mode II (see above) and mode III. In the modeled starting position of mode III, the farnesyl chain is aligned similarly into the binding pocket as in mode I but is connected to the diphosphate oxygen farthest from Mg^{2+}A. This is similar to the structure observed in cocrystal structures of other terpene synthases with fluorinated substrate analogues and all three Mg^{2+} ions bound (e.g., epi-aristolochene synthase[33] PDB entries 3LZ9, 3M00, 3M01, and 3M02). These additional simulations were performed with updated Mg^{2+} parameters.[34]

Simulation Details. All simulations were performed with AMBER11,[35] which faithfully represents CHARMM force fields.[36] After initial equilibration of the solvent (1000 steps of minimization and 50 ps MD at 300 K), systems were relaxed with 1000 steps of energy minimization, with positional restraints of 5 kcal mol^{-1} $Å^{-2}$ applied to all $C\alpha$ atoms and Mg^{2+} ions. With the same restraints in place, systems were equilibrated at 300 K and 1 atm in 50 ps of MD in the *NVT* ensemble followed by 100 ps of MD in the *NPT* ensemble. In another 50 ps of *NPT* MD, positional restraints were removed gradually. Thereafter, production MD was run at 300 K and 1 atm. Throughout, a 2 fs time step with SHAKE applied to bonds involving hydrogen and a direct space cutoff of 10 Å for nonbonded interactions with PME for long-range electrostatics were used. Langevin dynamics was used for temperature control (collision frequency of 5 ps^{-1} for equilibration and 2 ps^{-1} for production), and pressure was controlled by coupling to an external bath (AMBER11 default settings) for *NPT* conditions. The majority of production *NPT* MD runs were performed using the single-precision/double-precision model[37] on single M2050 GPU cards, reaching 11 ns/day for the systems with ~45K atoms. No significant difference from equivalent runs on multiple CPUs was observed. Three runs with different starting velocities were performed for each system, for ≥30 ns (see Table 1).

Analysis. Analysis of trajectories was performed with the AmberTools1.5 program ptraj or cpptraj unless specified otherwise. To assess the conformations sampled in the simulations with respect to the open and closed crystal structures, specific $C\alpha-C\alpha$ distances (d) were measured throughout the trajectories and normalized on the basis of the distance in 2OA6d (closed X-ray structure, d_C) and 2E4Oc (open X-ray structure, d_O) as follows:

$$d_{\mathrm{norm}} = \frac{1}{d_O - d_C}d - \frac{1}{d_O - d_C} + 1$$

so that a d_{norm} of 1 refers to closed and a d_{norm} of 2 to open. Hydrogen bonds were defined by a donor−acceptor distance of <3.6 Å and a donor−H−acceptor angle of >135°. Secondary structure assignments equivalent to DSSP[38] were performed with WORDOM.[39] Calculation of electrostatic potentials (in the presence of 0.05 M NaCl) was performed using the Adaptive Poisson−Boltzmann Solver (APBS).[40]

RESULTS

Conformational Change. All simulations started from ATAS in the closed conformation.[14] The combination of a bound diphosphate ligand and Mg^{2+} ions on either side is likely to be necessary to maintain a closed conformation,[2,25] so a change to a (more) open conformation is expected for systems that do not contain these. The conformations sampled typically did not, however, show very large deviations from the starting structure. This is consistent with the moderate $C\alpha$ root-mean-square deviation (rmsd) between open and closed crystal structures (1.94 Å between 2E4Oc and 2OA6d). When conformational change (e.g., opening of the active site) occurred, it typically happened within the first 20 ns; for all systems, the overall conformation showed little change in the last 10 ns of the simulation (the $C\alpha$ rmsd in the last 10 ns is never larger than 1.5 Å from the average structure in this period). In one of the FDP simulations, conformations sampled had a significantly higher $C\alpha$ rmsd compared to the starting structure over the final 10 ns of the simulation (3.4 Å). This is only in part due to a shift toward the crystallographic open active site conformation (the average $C\alpha$ rmsd against the open crystal structure, 2EO4c, is 2.9 Å). Visual inspection of the trajectory indicates that the large conformational change in this simulation is due to a wide opening of the active site [beyond what is seen in the open crystal structure (see further below)].

To quantify the opening of the active site (independent of further conformational change), two distances were used (see Figure S1 of the Supporting Information): (1) Asp91 $C\alpha-$Glu227 $C\alpha$, indicative of the opening of the active site entrance, and (2) Phe87 Cα–Ser216 Cα, indicative of the active site cavity itself, where the reactions take place. A further distance was used to monitor longer-range conformational change that may be needed to allow opening of the active site: Val124 Cα–Asp253 Cα [Val124 and Asp253 are distant from the active site entrance and lie on opposite sides, at the top of helices E and α1, respectively. They show significant changes in position between the open and closed crystal structures, and these changes are closely related to significant shifts in nearby helices (D, D1, and A–C1 loop and G1, α1, and I, respectively)].[14] A number of trends in the change from closed to open (or lack thereof) can be identified (Figure 2). First, a degree of long-distance conformational change is required for active site opening; when the distances used to monitor the active site entrance and cavity are similar to or larger than that of the open crystal structure (normalized distances of ≥ 2), the distance related to long-range conformational change is at least between those of the open and closed structures (normalized distance of >1.5). The distances used to monitor the active site cavity and long-range change are correlated (e.g., $R = 0.83$ for average distances in the final 2 ns over all simulations). Second, when the active site opens, the flexibility of the active site region is much larger than when it stays closed: closure leads to ordering of the active site. Third, in general, systems that change from closed to open stay open. There is one notable exception to this trend: in the simulations of the apoenzyme, partial closure can occur (see Figure 2B). By averaging the normalized distances of the active site and its entrance over 25–30 ns of the simulation, we obtained an overall measurement of the "openness" of conformations (Figure 2A). The simulations indicate that the enzyme in solution can adopt conformations with a wider active site opening than observed in the "open" crystal structures. Notably, such "more open" conformations are predominantly sampled with certain ligand combinations, such as FDP, FDP with MgB, and PP_i with MgB. Conversely, other ligand combinations show sampling of both the open and the closed structure (e.g., MgB, MgC, FDP with MgC, and FDP with MgA and MgC). One simulation, FDP run 2, shows a wide opening of the active site cavity as well as the active site entrance [normalized distance of ~3 (Figure 2C)]. This is accompanied by the largest long-range opening by far.

Crystallographic studies suggest a change in the ordering and/or flexibility at the top of helix H and the H-α1 loop when going from open to closed: the position of residues 231–239 could not be defined in open structures, whereas in the closed structure, they form a well-defined cap of the active site.[14] In addition, residues 227–230 are not α-helical in open crystal structures, whereas residues 227–232 form an extension of helix H in the closed structure. Because of the lack of well-defined electron density for the top of helix H (Cα B factors for residues 227–232 are 56–85 Å^2, whereas they are <40 Å^2 for residues 210–224) and the H–α1 loop (residues 231–239 are missing) in the open structures, their secondary structure and position cannot be determined by crystallography alone. Our simulations inform on the conformation of this region in relation to the ligand binding cycle. In almost all simulations, including those where active site opening takes place, residues 227–231 are α-helical for >95% and residue 232 is α-helical for >80% (measured from 20 to 30 ns). This is also the case for the extended simulations (apo to 200 ns and FDP to 100 ns), indicating that the amount of helicity in open conformations is unlikely to change significantly on the nanosecond to microsecond time scale. In addition, residue 233 is often

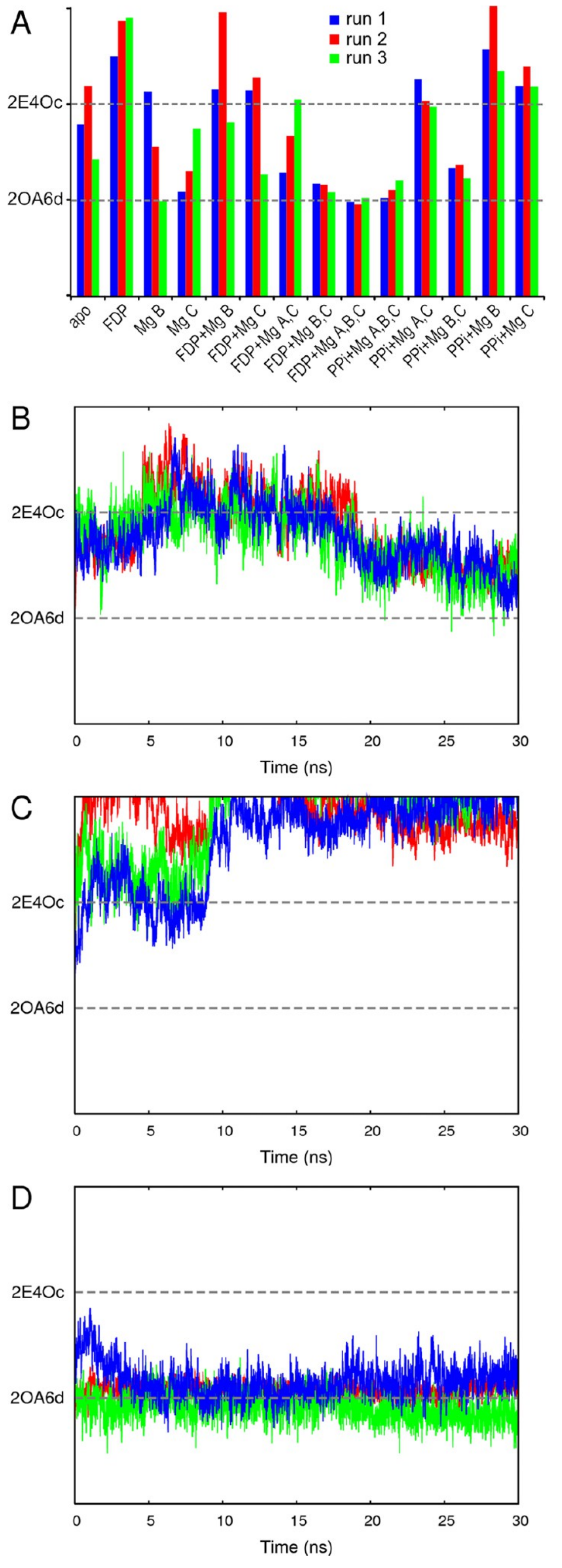


Figure 2. Comparison of simulations to closed (PDB entry 2OA6, chain D) and open (PDB entry 2E4O, chain C) crystal structures. (A) Average of normalized distances 1 and 2 (see Methods and Figure S1 of the Supporting Information) over the final 5 ns of each simulation. Systems with Mg^{2+}A but without Mg^{2+}C are omitted. (B) Normalized distances 1 (red, $d\text{C}\alpha_{\text{Asp91}}-\text{C}\alpha_{\text{Glu227}}$), 2 (green, $d\text{C}\alpha_{\text{Phe87}}-\text{C}\alpha_{\text{Ser216}}$), and 3 (blue, $d\text{C}\alpha_{\text{Val124}}-\text{C}\alpha_{\text{Asp253}}$) during 30 ns of apo run 3. (C) *Idem*, for FDP run 2. (D) *Idem*, for run 1 of FDP with MgA, MgB, and MgC.

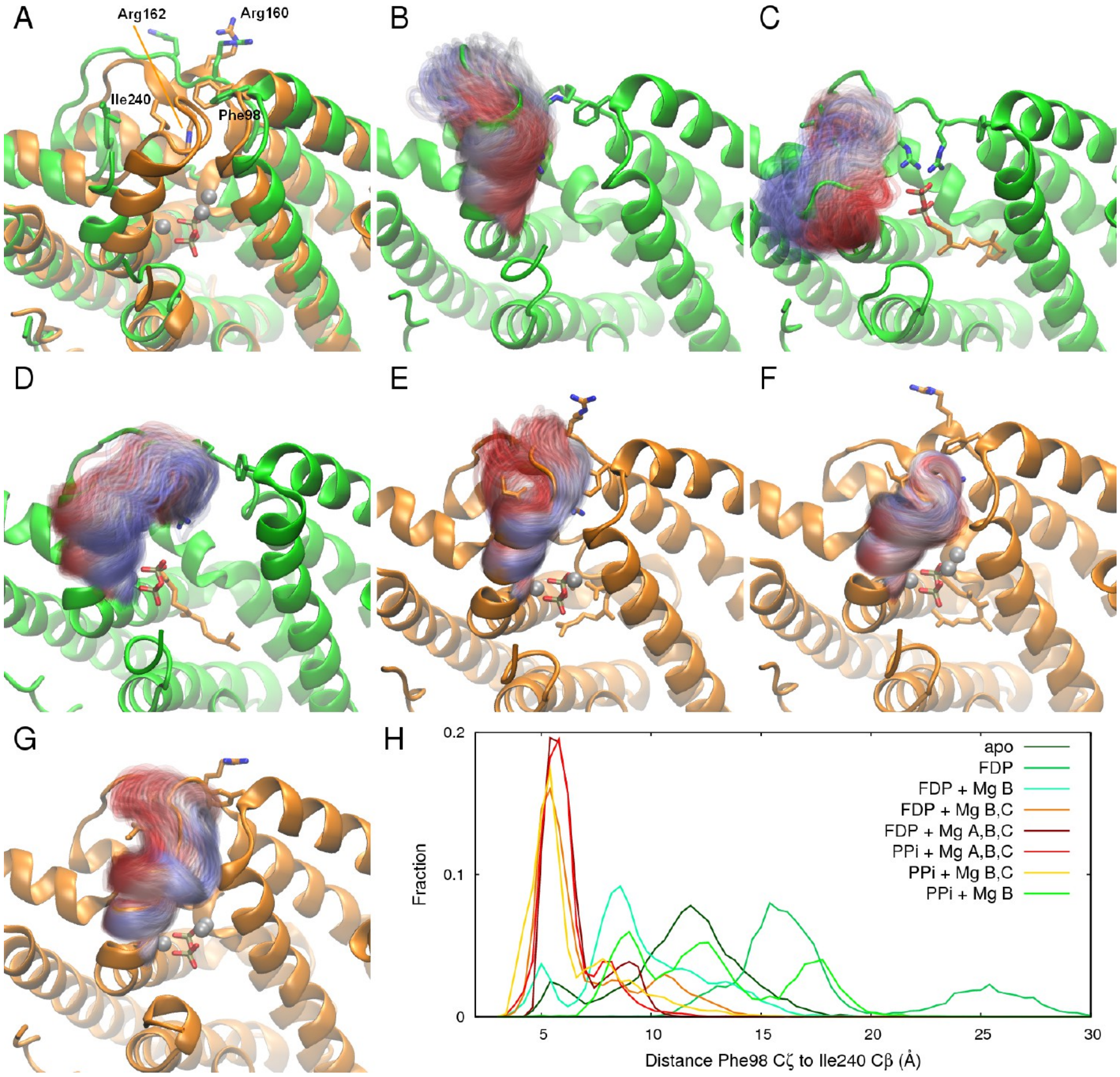


Figure 3. Active site entrance: lid opening and flexibility. (A) Open (2E4Oc, green) and closed (2OA6d, orange) crystal structures with annotated residues Phe98 and Ile240 (which interact when the lid is closed) and Arg160 and Arg162 (on the F−G1 loop near the active site entrance). (B) Representative snapshot from apo run 1, with residues 227−242 shown in transparent cartoon every 100 ps between 5 ns (blue) and 30 ns (red). (C) *Idem* from FDP run 3. (D) *Idem* from run 2 of FDP with MgB. (E) *Idem* from run 2 of FDP with MgB and MgC. (F) *Idem* from run 1 of FDP with MgA, MgB, and MgC. (G) *Idem* from run 1 of PP$_i$ with MgA, MgB, and MgC. (H) histogram of Phe98 Cζ−Ile240 Cβ distances from 15 to 30 ns in all three simulations of selected systems, in 0.4 Å bins.

included in the helix (between 27 and 100% depending on the simulation). The few cases in which fraying of the C-terminal side of the helix occurs involve systems that are unlikely to be catalytically relevant (e.g., PP$_i$ with MgA and MgC) or without the diphosphate moiety present (MgB or MgC). These data indicate that helix H extends to at least residue 232 in both closed and open conformations. There is, however, significant flexibility in the top of the helix when open conformations are adopted (Figure 3).

The H−α1 loop (which follows helix H) forms a lid over the active site entrance in the closed crystal structure (Figure 3A). This lid shows significant flexibility, sometimes also in simulations where the active site stays closed (Figure 3G), consistent with the high *B* factors in the closed crystal structure. The entrance, however, remains shielded from bulk solvent in all simulations with the closed active site. Hydrophobic interactions between Phe98 and Ile240 appear to be important to keep the lid "closed" over the active site entrance (Figure 3). In simulations in which the active site opens up, the loop becomes very flexible (with respect to the active site entrance and the rest of the protein); this increased flexibility is consistent with the lack of electron density observed in crystal structures of the open conformation.[14,25] On the time scale of our simulations, restructuring of the F−G1 loop that would result in Arg162 pointing outward into solvent, as found in the open crystal structures and only possible with an open lid, is not observed (Figure 3).

Interactions and Binding of Mg^{2+} Ions and the Diphosphate Moiety. Mg^{2+}B is bound by the NSE/DTE motif in terpene synthases.[28] In the closed crystal structure of AT-AS, Mg^{2+}B interacts with Asn219 Oδ1, Ser223 Oγ, and Glu227 Oδ2 (Figure 4A). In the simulations of the equivalent system (PP$_i$ with MgA, MgB, and MgC), these interactions are largely maintained, although some loss of interactions with Asn219 and Ser223 can occur. When FDP is present and the active site stays closed (FDP with MgA, MgB, and MgC and

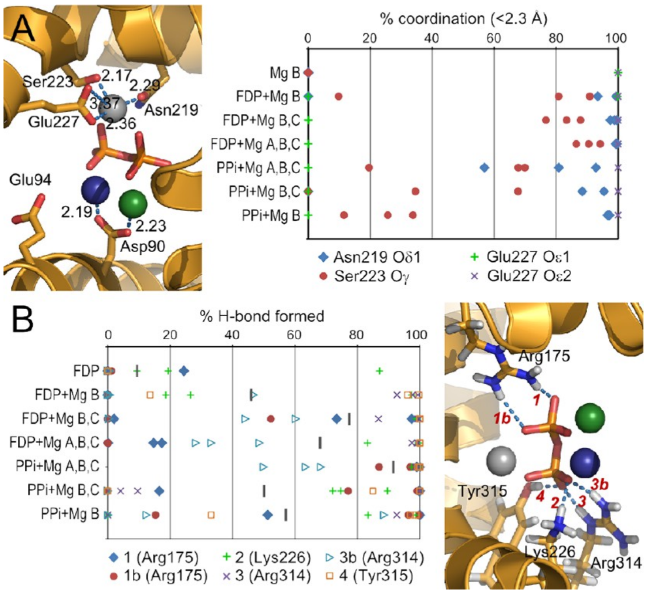


Figure 4. Interactions of $Mg^{2+}B$ and diphosphate with protein side chains. (A) Coordination of $Mg^{2+}B$ to the NSE motif, expressed as the percentage of simulation time from 15 to 30 ns that the distance to $Mg^{2+}B$ is <2.3 Å. (B) Hydrogen bonds to the diphosphate oxygens expressed as the percentage of simulation time from 15 to 30 ns that hydrogen bonds are formed. The average occupation of hydrogen bonds 1, 2, 3, 3b, and 4 (those present in the closed crystal structure) over all three simulations is indicated by a gray bar. In the molecular images, Mg^{2+} ions are colored as in Figure 1 (green for $Mg^{2+}A$, gray for $Mg^{2+}B$, and blue for $Mg^{2+}C$).

FDP with MgB and MgC), all interactions present in the crystal structure are well maintained (Figure 4A). When only $Mg^{2+}B$ is present, the NSE motif is disrupted; interactions with Asn219 and Ser223 are lost, and $Mg^{2+}B$ is held by both Glu227 carboxylate oxygens.

$Mg^{2+}A$ and $Mg^{2+}C$ are bound by the "aspartate-rich" DDXXD/E motif in terpene synthases;[28] for AT-AS, this is D^{90}DLLE (located on helix D1). In the closed crystal structure (2OA6d), Asp90 (the first Asp in the motif) is the only residue interacting directly with $Mg^{2+}A$ and $Mg^{2+}C$ (Oδ1 with $Mg^{2+}A$ and Oδ2 with $Mg^{2+}C$). In the largely closed crystal structure without $Mg^{2+}A$ (3CKEd), Asp90 is also the only residue interacting with $Mg^{2+}C$ (now with both carboxylate oxygens). The simulations are in agreement with the crystal structures: Asp90 Oδ1−$Mg^{2+}A$ and Asp90 Oδ2−$Mg^{2+}C$ distances are always <2.3 Å in all simulations of systems that contain both these Mg ions, and Asp90 Oδ1−$Mg^{2+}C$ and Asp90 Oδ2−$Mg^{2+}C$ distances are <2.3 Å in simulations of PP_i with MgB and MgC (equivalent to the situation in 3CKEd). For FDP with MgC, FDP with MgB and MgC, and PP_i with MgC, either one or both Asp90 carboxylate oxygens interact with $Mg^{2+}C$. In systems that have $Mg^{2+}A$ present at the start but not $Mg^{2+}C$, $Mg^{2+}A$ moves to the $Mg^{2+}C$ position. Several of these simulations do not show a full restoration of interactions between $Mg^{2+}A$ (now shifted toward position C) and the diphosphate and are therefore not discussed in detail. In the majority of simulations, only Asp90 interacts directly with $Mg^{2+}A$ and $Mg^{2+}C$ (as in the crystal structures). Notably, in run 3 of FDP with MgA, MgB, and MgC, Glu94 (the final residue in the aspartate-rich motif) moved to a position where it interacted with $Mg^{2+}A$ and $Mg^{2+}C$ simultaneously. This may be relevant to the conformation in solution; the side chain position of this conserved residue in the closed crystal structure (2OA6d) is likely influenced by the presence of a glycerol molecule directly next to it. Furthermore, additional simulations of FDP with MgA, MgB, and MgC with the Glu94−$Mg^{2+}A$ and −$Mg^{2+}C$ interaction in place indicated that this interaction may assist in the formation of a reactive Michaelis complex (see below).

Hydrogen bonds between the diphosphate moiety and protein side chains were also monitored (Figure 4B; hydrogen bond 1b is not fully formed in the crystal structure, where dN−O = 3.31 Å). All hydrogen bonds depicted in Figure 4B are formed for the majority of the simulation time for PP_i with MgA, MgB, and MgC (the system equivalent to the crystal structure). In FDP with MgA, MgB, and MgC, FDP with MgB and MgC, and PP_i with MgB and MgC, however, several hydrogen bonds are less stable. FDP with MgA, MgB, and MgC and FDP with MgB and MgC predominantly lose the hydrogen bonds with Arg175 (1 and 1b in Figure 4B), whereas hydrogen bond interactions with Lys226, Arg314, and Tyr315 are maintained for the majority of the simulation time. A reduced level of hydrogen bonding in PP_i with MgB and MgC is consistent with the partially closed crystal structure with the same ligands present (3CKEd), where only two hydrogen bonds (3 and 4 in Figure 4B) have donor−acceptor distances of ≤3 Å. Partial loss of the hydrogen bond donated by Tyr315 is observed in PP_i with MgB and MgC. Overall, the presence of these hydrogen bond interactions during simulation follows a clear trend (Figure 4B): the level is lowest in the FDP system, increases upon addition of $Mg^{2+}B$ and $Mg^{2+}C$, and then peaks in the system of PP_i with MgA, MgB, and MgC, followed by a decrease when $Mg^{2+}A$ and $Mg^{2+}C$ are omitted.

In the simulations with only FDP present, interactions with residues that lie more deeply in the active site entrance, Arg314 and Tyr315, are lost (although hydrogen bonds to different diphosphate oxygens are formed later in the simulation in some cases). The original interactions with Arg175 and Lys226 are also significantly weakened or lost, but they can be replaced with hydrogen bonds to other diphosphate oxygens (especially for Lys226). Further, new interactions are formed with positively charged residues around the edge of the active site entrance. It appears that the diphosphate can be "picked up", either by residues Arg162 and Arg160 [stable interactions are formed in three of six simulations (see Figure 3C)] or by Lys54 (the other three simulations).

Michaelis Complex Formation. From the results presented above, it is clear that closure of the active site requires the presence of $Mg^{2+}B$ and $Mg^{2+}C$ in conjunction with the diphosphate moiety. Together, they link both sides of the active site entrance and thereby allow the active site to close around the farnesyl chain. Experiments with fluorinated FDP have indicated that expulsion of PP_i and the initial cyclization of the farnesyl chain (by formation of the C1−C10 bond, yielding a germacryl cation) is a concerted process[3] (in the sense that no stable intermediate is formed; C−O bond dissociation likely leads to a short-lived allyl cation that rapidly cyclizes). For this concerted reaction to take place, C1 and C10 must be in reasonable proximity in the Michaelis complex. The C1−C10 distance was monitored in all simulations with FDP present. As expected, the simulations that result in opening of the active site (FDP, FDP with MgB, and FDP with MgC) show a wide range of C1−C10 distances from ∼3.5 to ∼11.5 Å (Figure 5E). Once the enzyme is "locked" into a closed conformation, the maximum C1−C10 distance sampled is reduced to ∼7.2 Å.

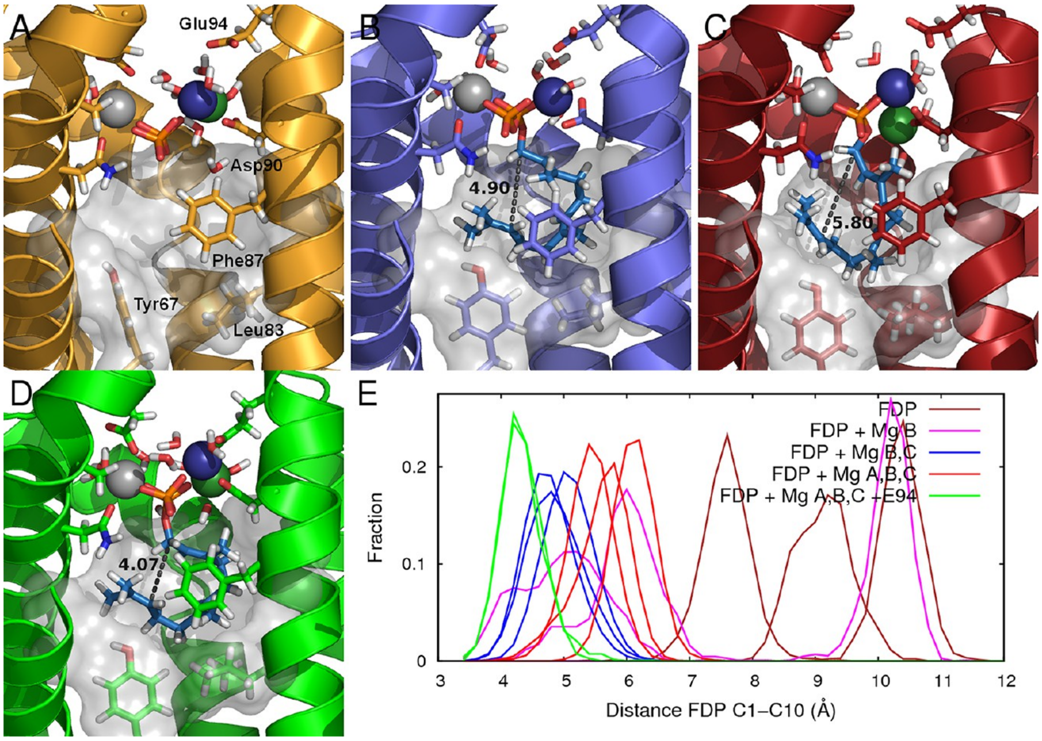


Figure 5. Active site cavity and Michaelis complex formation. (A) The closed crystal structure (2OA4d), with hydrogens added. Crystallographic water molecules interacting with PP_i and/or the Mg^{2+} ions are shown, as are selected protein side chains and the surface of the bottom part of the active site cavity (transparent). Mg^{2+} ions are colored as in Figure 1 (green for Mg^{2+}A, gray for Mg^{2+}B, and blue for Mg^{2+}C). (B–D) Representative conformations obtained after energy minimization with FDP and Mg^{2+}B and Mg^{2+}C bound (B), with FDP and Mg^{2+}A, Mg^{2+}B, and Mg^{2+}C bound (C), and with FDP and Mg^{2+}A, Mg^{2+}B, and Mg^{2+}C bound and Glu94 interacting with Mg^{2+}A and Mg^{2+}C (D). The farnesyl chain of FDP is colored light blue, and the C1–C10 distance is labeled; other depictions are as in panel A. (E) Histogram of FDP C1–FDP C10 distances from 15 to 30 ns of individual simulations, collected in 0.2 Å bins.

Surprisingly, simulations of FDP with MgB and MgC show a shorter average C1–C10 distance than simulations of FDP with MgA, MgB, and MgC. The two additional simulations of FDP with MgA, MgB, and MgC that feature a direct interaction of Glu94 with Mg^{2+}A and Mg^{2+}C, however, feature short C1–C10 distances (average of 4.44 Å from 15 to 30 ns). C1–C10 distances between 4.2 and 4.4 Å are most frequently sampled (Figure 5E), and occasionally, distances of <3.5 Å are reached. This situation remained unchanged upon extension of one of these simulations to 100 ns.

Representative structures for the three sets of simulations with FDP in the closed active site were obtained by minimization of selected conformations (Figure 5B–D). In all three structures, the active site cavity has a very similar shape. Compared to the crystal structure, the diphosphate moiety is positioned somewhat higher in the active site in FDP with MgB and MgC and FDP with MgA, MgB, and MgC, allowing C1 and C10 to be farther apart. In the simulations of FDP with MgA, MgB, and MgC with Glu94 interacting with Mg^{2+}A and Mg^{2+}C, the diphosphate moiety is in a position similar to that in the closed crystal structure, and the farnesyl chain is lined up tightly against the wall of the active site cavity. In the equivalent simulations without Glu94 interacting with Mg^{2+}A and Mg^{2+}C, the farnesyl chain is not optimally aligned with the active site cavity. The representative structure with Glu94 interacting with Mg^{2+}A and Mg^{2+}C indicates a putative Michaelis complex: C1 and C10 have approached each other within 4.07 Å, and bond formation (probably concurrent with expulsion of PP_i)[3] would result in the experimentally observed stereochemistry around C10 (*S*).

Alternative Farnesyl Binding Modes. In the simulations described above, the farnesyl moiety was modeled as being linked to the diphosphate oxygen closest to Mg^{2+}A, and the end of the farnesyl chain near Mg^{2+}B [binding mode I (see Figures 5 and 6A)]. It is possible that the farnesyl binding mode in a reactive Michaelis complex is different, and therefore, two alternative modes were investigated. In mode II, the farnesyl moiety is linked to the same diphosphate oxygen, but the end of the moiety lies on the side of Mg^{2+}A and Mg^{2+}C (see Figure 6B,C). In mode III, the farnesyl moiety was linked to the other oxygen pointing into the active site [farthest from Mg^{2+}A (see Figure 6D)]. This connection is also observed in crystal structures from other terpene synthases (see Discussion). Three additional 30 ns simulations were performed for all three modes, using the improved force field parameters for Mg^{2+}.[34] In all simulations, the overall enzyme conformation remained closed, and Mg^{2+}– and diphosphate–protein interactions were mostly similar as in the previous simulations with all three Mg^{2+} ions. Notably, stable interactions between Glu94 and Mg^{2+}A, Mg^{2+}C, or both were formed in all simulations.

In all three simulations of mode I, the farnesyl chain remains folded in the active site cavity with C1 and C10 in the proximity, in line for cyclization (Figure 6A,E). In the run that shows the largest average C1–C10 distance, the farnesyl moiety has adopted a conformation that is incompatible with germacryl cation formation: cyclization would lead to the methyl groups attached to C3 and C7 being on opposite sides of the ring. The average C1–C10 distance in the other two runs (from 15 to 30 ns) is 4.75 Å (similar although somewhat longer than in the previous simulations of FDP with MgA, MgB, MgC, and

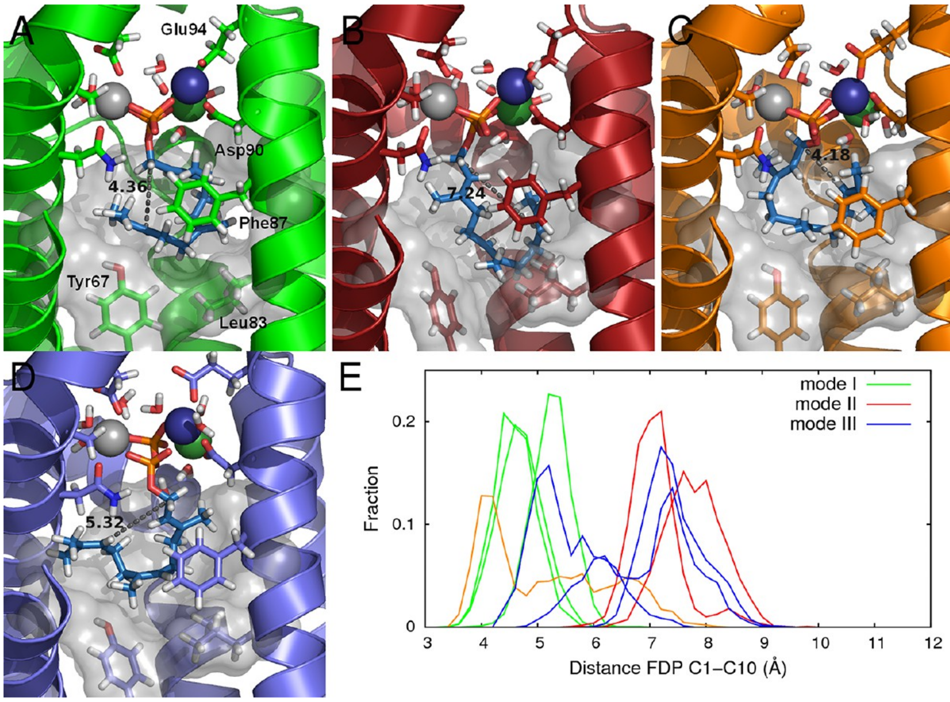


Figure 6. Alternative FDP binding modes. (A–D) Conformations obtained after energy minimization from additional simulations of FDP with $Mg^{2+}A$, $Mg^{2+}B$, and $Mg^{2+}C$ bound (green for $Mg^{2+}A$, gray for $Mg^{2+}B$, and dark blue for $Mg^{2+}C$, as in Figure 1) and new Mg^{2+} parameters,[34] with different FDP binding modes. The same protein residues are shown as in Figure 5. FDP binding mode I, identical to the majority of simulations (A); FDP binding mode II, with the farnesyl chain looped round in the opposite direction (B and C); and FDP binding mode III, with the farnesyl chain attached to the alternative diphosphate oxygen (D). In panel C, $Mg^{2+}A$ has lost direct interactions with the diphosphate oxygens. (E) Histogram of FDP C1–FDP C10 distances from 15 to 30 ns of individual simulations, collected in 0.2 Å bins. The orange line corresponds to the mode II simulation depicted in panel C.

Glu94). In all three simulations of mode II, the farnesyl chain ends up in a more linear conformation after 30 ns of MD (average C1–C10 distance of 7.10 Å in the final 2 ns). In one of the three simulations, shorter distances are sampled initially, but $Mg^{2+}A$ loses direct interaction with the diphosphate in the first 0.3 ns (Figure 6C and 6E, orange line); a water molecule nestles between them. In mode III, three populations of C1–C10 distances can be distinguished, with approximate average distances of 5.2, 6.2, and 7.3 Å (Figure 6E). Conformations with the largest distance are mostly sampled when $Mg^{2+}B$ has lost interactions with Ser223 and Asn219 from the NSE binding motif and has moved up higher in the entrance tunnel. As a consequence, the diphosphate group has moved up as well, allowing partial linearization of the farnesyl chain. A few different conformations are responsible for C1–C10 distances of <7 Å. One main conformation of the farnesyl moiety displays the shortest C1–C10 distances (Figure 6D). In this conformation, the farnesyl moiety bends at C5, and the end of the chain (C12 and C13) nestles between helices H and J.

■ DISCUSSION

Ligand Binding and Conformational Change. The primary motivation for this study was to test the hypothesis for the sequence of Mg^{2+} and FDP binding suggested by Shishova et al.[25] (Figure 1) and to obtain a detailed picture of the conformational changes involved. The nature of binding of Mg^{2+} to enzymes is widely important, as there are many examples in which Mg^{2+} ions are used by enzymes to bind phosphate moieties that are crucial for their function. Shishova et al. proposed the following sequence on the basis of their crystallographic studies[14,25] (Figure 1). (1) FDP binds (with the farnesyl chain in the hydrophobic active site cavity), and $Mg^{2+}B$ binds. (2) Binding of $Mg^{2+}C$ triggers a conformational change in the protein from its open form to its closed form. (3) $Mg^{2+}A$ binds to generate the catalytically active Michaelis complex. (4) After reaction, $Mg^{2+}A$ is released while the enzyme remains closed. (5) Dissociation of $Mg^{2+}C$ triggers reopening of the enzyme, followed by release of the hydrocarbon product (aristolochene). (6) Diphosphate and $Mg^{2+}B$ are released. The existence of several structures of FDP (or fluorinated analogues) bound to the enzyme in the absence of Mg^{2+} ions implies that FDP binding occurs before binding of Mg^{2+} ions. Our simulations with only $Mg^{2+}B$ or $Mg^{2+}C$ present indicate that the enzyme active site may not (fully) open up, making it difficult for FDP to enter and adopt the correct orientation. Furthermore, when FDP binds first, a distinctly open conformation is obtained that appears to be poised for Mg^{2+} binding; the Mg^{2+} binding sites can be accessed from solvent and are electrostatically very favorable for cation binding (Figure S2 of the Supporting Information). In accordance with the hypothesis, the enzyme remains in the open form when both FDP and $Mg^{2+}B$ are bound, and the presence of the $Mg^{2+}B$ ion helps orient the diphosphate moiety: the frequency of interactions between the diphosphate and residues Arg314 and Tyr315 increases. Our simulations further suggest that the diphosphate moiety of FDP (in the absence of Mg^{2+} ions) may first be recognized by positively charged side chains that line the entrance to the active site (Lys54, Arg160, and Arg162). This in turn allows interactions with side chains present in the active site channel (Lys226 and Arg175) to be formed; binding of $Mg^{2+}B$ then allows the stable formation of interactions with the side chains that lie more deeply in the active site (Arg314 and Tyr315).

Our simulations strongly indicate that both Mg^{2+}B and Mg^{2+}C need to bind alongside FDP for the full open-to-closed transition to occur; all simulations with these two ions and a diphosphate moiety present remain largely in the closed conformation, whereas all the other systems show a transition to the open conformation. The existence of crystal structures with Mg^{2+}B bound along with a diphosphate moiety (and the absence of equivalent structures with Mg^{2+}C), the largely preorganized Mg^{2+}B binding motif, and the observation from our simulations that diphosphate−protein interactions are strengthened more upon Mg^{2+}B binding than upon Mg^{2+}C binding together strongly indicate that Mg^{2+}B will bind before Mg^{2+}C. In addition, the conformations sampled in our simulations indicate that the accessibility of the Mg^{2+}B binding site would be very limited with FDP and Mg^{2+}C bound, whereas the Mg^{2+}C binding site is accessible. The simulations suggest the following sequence of conformational change. When Mg^{2+}C binds to the enzyme in addition to FDP and Mg^{2+}B, the Mg^{2+}B binding motif and Mg^{2+}A and Mg^{2+}C binding motif (on helix H and helix D1, respectively) are brought closer together. The result is a shift of helix H toward helix D1 that narrows the active site entrance and thereby allows hydrophobic residues Phe98 (on the D1−D loop) and Ile240 (on the H−α1 loop) to interact with each other. This interaction ensures that the lid closes over the active site, shielding it largely from bulk solvent. Solvent is still present in the active site entrance (and, indeed, interacts with the Mg^{2+} ions), and the Mg^{2+}A binding site is reasonably accessible.

Although it was recently shown that the Mg^{2+} parameters implemented in the CHARMM27 force field used here somewhat overestimate Mg^{2+}−phosphate attractions,[34] our simulations predict Mg^{2+}−protein conformations consistent with the data obtained from crystallography. For example, Asp90 ligates to Mg^{2+}C with both its carboxylate oxygens as in the structure of 3CKEd[25] (PP_i with Mg^{2+}B and Mg^{2+}C) and with one carboxylate oxygen each to Mg^{2+}A and Mg^{2+}C as in 2OA6d[14] (PP_i with Mg^{2+}B and Mg^{2+}C). Our simulations clearly indicate that binding a Mg^{2+} ion at position A is not supported without the presence of Mg^{2+} at position C; when starting at position A, it moves quickly toward position C. Mg^{2+}A must therefore bind after Mg^{2+}C binds and thus will be the last ligand to bind before reaction can commence. Similarly, it is likely to be the first to dissociate after completion of the reaction. Crystal structure 3CKEd and our simulations show that the enzyme remains in the closed conformation with PP_i and Mg^{2+}B and Mg^{2+}C bound. When Mg^{2+}C (or Mg^{2+}B) is released, the enzyme active site opens up. Evidence from crystallography[25] (3BNXd with PP_i and Mg^{2+}B bound) and the fact that Mg^{2+}B has a preorganized binding site indicate that Mg^{2+}C is most likely to dissociate first. The opening of the active site will then allow the hydrocarbon product (predominantly aristolochene) to be released. Our simulations indicate that interactions between PP_i or Mg^{2+}B and the protein are weakened both in the absence and presence of Mg^{2+}C, so the release of PP_i and Mg^{2+}B is likely to follow soon after Mg^{2+}C dissociates and the product is released. The enzyme thereby returns to its apo state, ready for a new catalytic cycle to start.

Michaelis Complex and Consequences for the Reaction Pathway. In the starting structures with FDP, the farnesyl chain was modeled into the active site cavity in line with mutagenesis studies conducted with PR-AS. Alternative positions of the farnesyl chain were tested, but these did not lead to conformations compatible with cyclization [the distance between C1 and C10 increased to >6 Å (results not shown)]. Mutation of Phe112 [equivalent to Phe87 in AT-AS (see Figure 4)] to Ala leads to predominantly linear farnesene products [36% (*E*)-*β*-farnesene and 53.5% (*E*,*E*)-*α*-farnesene].[23] Mutation of Tyr92 to residues with decreasing bulk also led to the accumulation of linear farnesene products, with relative amounts linearly dependent on residue size; the Tyr92 to Ala mutation led to ∼80% of linear farnesene products.[5] In addition, mutation of Leu108 (Leu83 in AT-AS) to smaller residues also leads to predominantly linear farnesene products (79% for Leu108Ala and 76% for Leu108Ser).[21] These three residues are therefore clearly important for enforcing a conformation of the farnesyl chain in line with cyclization. The putative Michaelis complex obtained from our simulations (Figure 5) helps to explain the role of these residues; all three are located such that the farnesyl chain cannot adopt a more linear conformation and has to “bend” near C7 and C8. [These residues may play a similar role in alternative binding modes (see below and Figure 6).] Furthermore, the putative Michaelis complex is consistent with the formation of a germacryl cation via an S_N2-like reaction directly from FDP; it has been shown that the inversion of configuration at C1 that is necessary for such a reaction takes place.[16] The C10−C11 double bond is oriented such that C1−C10 bond formation results in an *S* configuration around C10, which is again consistent with formation of a germacryl cation that after proton abstraction would form (−)-germacrene A (side product and putative intermediate) and subsequently (+)-aristolochene. The putative Michaelis complex revealed by our simulations therefore appears to be a reasonable model. In this model, the farnesyl chain is connected to the diphosphate oxygen closest to Mg^{2+}A (O3 in PDB entry 2OA6). This is different from the binding mode observed in the cocrystal structures of epi-aristolochene synthase with 2-fluorofarnesyl diphosphates (PDB entries 3LZ9, 3M00, 3M01, and 3M02),[33] limonene synthase with 2-fluorogeranyl or 2-fluorolinalyl diphosphate (PDB entries 2ONG and 2ONH),[41] bornyl diphosphate synthase with 3-aza-2,3-dihydrogeranyl or bornyl diphosphate (PDB entries 1N20 and 1N24),[42] and taxadiene synthase with 2-fluorogeranylgeranyl diphosphate (PDB entry 3P5R),[43] where the substrate or intermediate analogue is attached to the other oxygen (equivalent to O6 in PDB entry 2OA6). These crystal structure data suggest that FDP is more likely to bind with the farnesyl chain connected to O6. On the other hand, thermodynamically favorable binding modes of substrate analogues (as captured in crystal structures) may not always correspond to the productive binding modes of substrates.[44] The additional simulations of two alternative binding modes (mode II with the farnesyl chain connected to O3 but with the farnesyl chain looped the other way around and mode III with the farnesyl chain connected to O6) indicate that these binding modes in AT-AS sample conformations that are not “cyclization ready”; i.e., C1 and C10 are generally not in close proximity to each other. These findings may be influenced by (in)accuracies of the force field parameters and the limited sampling in the simulations performed here; catalytically relevant conformations may have been missed. In addition, it is possible that a linear farnesyl cation is formed in a conformation in which C1 and C10 are relatively far apart, and thereafter, ring closure takes place (in a concerted but highly asynchronous process).

For a definitive assignment, therefore, further work is needed to investigate the reactivity of different binding modes, e.g., using quantum mechanics/molecular mechanics (QM/MM) methods,[45] similar to the elegant QM/MM dynamics studies recently performed on bornyl phosphate synthase.[46]

Interestingly, our simulations indicate that to form a Michaelis complex with short C1−C10 distances (in line with cyclization), interaction of Glu94 with the Mg^{2+} ions may be required. [Notably, in simulations with the improved Mg^{2+} parameters, interactions between Glu94 and Mg^{2+} ions are always formed, independent of the farnesyl binding mode (see Figure 6).] Glu94 is part of the conserved metal-binding DDXXD/E motif, and mutation of the equivalent Glu119 in PR-AS to Asp or Gln leads to a 4.4- or 9-fold reduction in k_{cat}, respectively, stronger than the effect of the Asp115Glu mutation (2.7 times reduced k_{cat}) but weaker than the effect of the Asp115Asn mutation (which leads to an inactive enzyme).[22] It is therefore possible that an interaction between Glu94 and Mg^{2+}A, Mg^{2+}C, or both is required for efficient reaction, but it is less important than the presence of a carboxylate group at position 90. In the crystal structure of the closed enzyme (with PP_i and Mg^{2+}A, Mg^{2+}B, and Mg^{2+}C bound), the Glu94 side chain is not directed to Mg^{2+}A or Mg^{2+}C, possibly because of the nearby glycerol molecule.[14]

Once the germacryl cation is formed by PP_i expulsion and C1−C10 bond formation, a carbocationic reaction cascade will result in the final products, predominantly (+)-aristolochene in this case. To form germacrene A from the germacryl cation, deprotonation at C12 or C13 is required. There are no protein side chains in the vicinity that could act as a base for this deprotonation. On the basis of the putative Michaelis complex indicated by our simulations (Figure 5D), however, a relatively small shift of the germacryl cation formed from this complex would be needed to bring C12 or C13 near the PP_i oxygens. An oxygen attached to the phosphorus that was not linked to the farnesyl chain points into the active site, without interacting with Mg^{2+} ions or water; its pK_a can therefore be expected to be relatively high. This would create the protonated $HOPP_i$ species, which is consistent with the proposal that PP_i conducts the required (re)protonation of germacrene A as well as the deprotonation necessary to form the final product.[2,14] Recent studies provide support for this proposal, by ruling out alternatives[22,24] and indicating that the nature of the diphosphate binding residues is important for progression from germacrene A to aristolochene.[47] In analogy, QM/MM dynamics simulations of chemical reactions in bornyl diphosphate synthase suggested that the side products camphene and pinene may be formed through deprotonation by the diphosphate moiety (via an active site water molecule) of camphyl and pinyl carbocations.[46]

■ CONCLUSIONS

Our molecular dynamics simulations, with all combinations of ligands, strongly support the overall ligand binding cycle as proposed by Shishova et al.[25] Further details of the transitions in the catalytic cycle are revealed. Whereas the apoenzyme can sample open to near-closed conformations, the initial binding of FDP and one Mg^{2+} appears to shift the equilibrium toward distinctly open conformations, allowing further ligands to bind. Only in the presence of the diphosphate moiety, Mg^{2+}B, and Mg^{2+}C can a closed conformation (necessary for catalysis) be maintained. Once the enzyme is locked in the closed conformation, the farnesyl chain is forced into a conformation consistent with cyclization. When Glu94 and Asp90 coordinate to Mg^{2+}A and Mg^{2+}C, a putative Michaelis complex, poised for cyclization and in line with the current evidence for the mechanism of aristolochene formation, was obtained. This detailed model of a Michaelis complex in the active, closed form of a sesquiterpene synthase could assist in engineering of this class of enzymes to help produce a range of terpenoid compounds with high stereospecificity. The sequence of ligand binding and conformational change and further structural details revealed by this study are likely to be shared among most or all terpene synthases with the class I α-helical fold; crystallographic studies of terpene synthases other than aristolochene synthase confirm this (see ref 2 and references cited therein).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Topology and additional parameters used for FDP and PP_i, depiction of distances used for following the closed to open transition (Figure S1), and snapshots of representative structures with their electrostatic potential mapped onto the solvent accessible surface (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: Marc.vanderKamp@bristol.ac.uk. Phone: +44 117 928 9097.
*E-mail: Adrian.Mulholland@bristol.ac.uk. Phone: +44 117 928 9097.

Funding

A.J.M. is an EPSRC Leadership Fellow (Grant EP/G007705/1) and (with M.W.v.d.K.) thanks EPSRC for support. J.S. thanks the Royal Thai Government for funding.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was conducted using the computational facilities of the Advanced Computing Research Centre, University of Bristol.

■ ABBREVIATIONS

FDP, farnesyl diphosphate; PP_i, inorganic diphosphate; AT-AS, *A. terreus* aristolochene synthase; PR-AS, *P. roqueforti* aristolochene synthase; MD, molecular dynamics; PDB, Protein Data Bank.

■ REFERENCES

(1) Christianson, D. W. (2008) Unearthing the roots of the terpenome. *Curr. Opin. Chem. Biol. 12*, 141−150.

(2) Miller, D. J., and Allemann, R. K. (2012) Sesquiterpene synthases: Passive catalysts or active players? *Nat. Prod. Rep. 29*, 60−71.

(3) Yu, F. L., Miller, D. J., and Allemann, R. K. (2007) Probing the reaction mechanism of aristolochene synthase with 12,13-difluorofarnesyl diphosphate. *Chem. Commun.*, 4155−4157.

(4) Li, J.-X., Fang, X., Zhao, Q., Ruan, J.-X., Yang, C.-Q., Wang, L.-J., Miller, D. J., Faraldos, J. A., Allemann, R. K., Chen, X.-Y., and Zhang, P. (2013) Rational engineering of plasticity residues of sesquiterpene synthases from *Artemisia annua*: Product specificity and catalytic efficiency. *Biochem. J. 451*, 417−426.

(5) Deligeorgopoulou, A., and Allemann, R. K. (2003) Evidence for differential folding of farnesyl pyrophosphate in the active site of aristolochene synthase: A single-point mutation converts aristolochene synthase into an (E)-β-farnesene synthase. *Biochemistry 42*, 7741–7747.

(6) Greenhagen, B. T., O'Maille, P. E., Noel, J. P., and Chappell, J. (2006) Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *Proc. Natl. Acad. Sci. U.S.A. 103*, 9826–9831.

(7) Aaron, J. A., Lin, X., Cane, D. E., and Christianson, D. W. (2010) Structure of epi-isozizaene synthase from *Streptomyces coelicolor* A3(2), a platform for new terpenoid cyclization templates. *Biochemistry 49*, 1787–1797.

(8) Gennadios, H. A., Gonzalez, V., Di Costanzo, L., Li, A. A., Yu, F. L., Miller, D. J., Allemann, R. K., and Christianson, D. W. (2009) Crystal structure of (+)-δ-cadinene synthase from *Gossypium arboreum* and evolutionary divergence of metal binding motifs for catalysis. *Biochemistry 48*, 6175–6183.

(9) Lesburg, C. A., Zhai, G. Z., Cane, D. E., and Christianson, D. W. (1997) Crystal structure of pentalenene synthase: Mechanistic insights on terpenoid cyclization reactions in biology. *Science 277*, 1820–1824.

(10) Rynkiewicz, M. J., Cane, D. E., and Christianson, D. W. (2001) Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. *Proc. Natl. Acad. Sci. U.S.A. 98*, 13543–13548.

(11) Starks, C. M., Back, K. W., Chappell, J., and Noel, J. P. (1997) Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science 277*, 1815–1820.

(12) McAndrew, R. P., Peralta-Yahya, P. P., DeGiovanni, A., Pereira, J. H., Hadi, M. Z., Keesling, J. D., and Adams, P. D. (2011) Structure of a three-domain sesquiterpene synthase: A prospective target for advanced biofuels production. *Structure 19*, 1876–1884.

(13) Caruthers, J. M., Kang, I., Rynkiewicz, M. J., Cane, D. E., and Christianson, D. W. (2000) Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. *J. Biol. Chem. 275*, 25533–25539.

(14) Shishova, E. Y., Di Costanzo, L., Cane, D. E., and Christianson, D. W. (2007) X-ray crystal structure of aristolochene synthase from *Aspergillus terreus* and evolution of templates for the cyclization of farnesyl diphosphate. *Biochemistry 46*, 1941–1951.

(15) Calvert, M. J., Taylor, S. E., and Allemann, R. K. (2002) Tyrosine 92 of aristolochene synthase directs cyclisation of farnesyl pyrophosphate. *Chem. Commun.*, 2384–2385.

(16) Cane, D. E., Prabhakaran, P. C., Salaski, E. J., Harrison, P. H. M., Noguchi, H., and Rawlings, B. J. (1989) Aristolochene biosynthesis and enzymatic cyclization of farnesyl pyrophosphate. *J. Am. Chem. Soc. 111*, 8914–8916.

(17) Cane, D. E., Prabhakaran, P. C., Oliver, J. S., and Mcilwaine, D. B. (1990) Aristolochene biosynthesis: Stereochemistry of the deprotonation steps in the enzymatic cyclization of farnesyl pyrophosphate. *J. Am. Chem. Soc. 112*, 3209–3210.

(18) Deligeorgopoulou, A., Taylor, S. E., Forcat, S., and Allemann, R. K. (2003) Stabilisation of eudesmane cation by tryptophan 334 during aristolochene synthase catalysis. *Chem. Commun.*, 2162–2163.

(19) Faraldos, J. A., and Allemann, R. K. (2011) Inhibition of (+)-aristolochene synthase with iminium salts resembling eudesmane cation. *Org. Lett. 13*, 1202–1205.

(20) Faraldos, J. A., Antonczak, A. K., Gonzalez, V., Fullerton, R., Tippmann, E. M., and Allemann, R. K. (2011) Probing eudesmane cation-π interactions in catalysis by aristolochene synthase with non-canonical amino acids. *J. Am. Chem. Soc. 133*, 13906–13909.

(21) Faraldos, J. A., Gonzalez, V., Senske, M., and Allemann, R. K. (2011) Templating effects in aristolochene synthase catalysis: Elimination versus cyclisation. *Org. Biomol. Chem. 9*, 6920–6923.

(22) Felicetti, B., and Cane, D. E. (2004) Aristolochene synthase: Mechanistic analysis of active site residues by site-directed mutagenesis. *J. Am. Chem. Soc. 126*, 7212–7221.

(23) Forcat, S., and Allemann, R. K. (2006) Stabilisation of transition states prior to and following eudesmane cation in aristolochene synthase. *Org. Biomol. Chem. 4*, 2563–2567.

(24) Miller, D. J., Gao, J. L., Truhlar, D. G., Young, N. J., Gonzaleza, V., and Allemann, R. K. (2008) Stereochemistry of eudesmane cation formation during catalysis by aristolochene synthase from *Penicillium roqueforti*. *Org. Biomol. Chem. 6*, 2346–2354.

(25) Shishova, E. Y., Yu, F. L., Miller, D. J., Faraldos, J. A., Zhao, Y. X., Coates, R. M., Allemann, R. K., Cane, D. E., and Christianson, D. W. (2008) X-ray crystallographic studies of substrate binding to aristolochene synthase suggest a metal ion binding sequence for catalysis. *J. Biol. Chem. 283*, 15431–15439.

(26) Vedula, L. S., Zhao, Y. X., Coates, R. M., Koyama, T., Cane, D. E., and Christianson, D. W. (2007) Exploring biosynthetic diversity with trichodiene synthase. *Arch. Biochem. Biophys. 466*, 260–266.

(27) Tantillo, D. J. (2011) Biosynthesis via carbocations: Theoretical studies on terpene formation. *Nat. Prod. Rep.*, 28.

(28) Christianson, D. W. (2006) Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev. 106*, 3412–3442.

(29) Hooft, R. W. W., Sander, C., and Vriend, G. (1996) Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins: Struct., Funct., Genet. 26*, 363–376.

(30) MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B 102*, 3586–3616.

(31) MacKerell, A. D., Feig, M., and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem. 25*, 1400–1415.

(32) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33.

(33) Noel, J. P., Dellas, N., Faraldos, J. A., Zhao, M., Hess, B. A., Smentek, L., Coates, R. M., and O'Maille, P. E. (2010) Structural elucidation of cisoid and transoid cyclization pathways of a sesquiterpene synthase using 2-fluorofarnesyl diphosphates. *ACS Chem. Biol. 5*, 377–392.

(34) Allnér, O., Nilsson, L., and Villa, A. (2012) Magnesium ion–water coordination and exchange in biomolecular simulations. *J. Chem. Theory Comput. 8*, 1493–1502.

(35) Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem. 26*, 1668–1688.

(36) Crowley, M. F., Williamson, M. J., and Walker, R. C. (2009) CHAMBER: Comprehensive support for CHARMM force fields within the AMBER software. *Int. J. Quantum Chem. 109*, 3767–3772.

(37) Götz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., and Walker, R. C. (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput. 8*, 1542–1555.

(38) Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577–2637.

(39) Seeber, M., Felline, A., Raimondi, F., Muff, S., Friedman, R., Rao, F., Caflisch, A., and Fanelli, F. (2011) Wordom: A user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. *J. Comput. Chem. 32*, 1183–1194.

(40) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037–10041.

(41) Hyatt, D. C., Youn, B., Zhao, Y., Santhamma, B., Coates, R. M., Croteau, R. B., and Kang, C. (2007) Structure of limonene synthase, a

simple model for terpenoid cyclase catalysis. *Proc. Natl. Acad. Sci. U.S.A. 104*, 5360−5365.

(42) Whittington, D. A., Wise, M. L., Urbansky, M., Coates, R. M., Croteau, R. B., and Christianson, D. W. (2002) Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc. Natl. Acad. Sci. U.S.A. 99*, 15375−15380.

(43) Köksal, M., Jin, Y., Coates, R. M., Croteau, R., and Christianson, D. W. (2011) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature 469*, 116−120.

(44) Vedula, L. S., Rynkiewicz, M. J., Pyun, H. J., Coates, R. M., Cane, D. E., and Christianson, D. W. (2005) Molecular recognition of the substrate diphosphate group governs product diversity in trichodiene synthase mutants. *Biochemistry 44*, 6153−6163.

(45) Van der Kamp, M. W., and Mulholland, A. J. (2013) Combined Quantum Mechanics/Molecular Mechanics (QM/MM) methods in computational enzymology. *Biochemistry 52*, 2708−2728.

(46) Major, D. T., and Weitman, M. (2012) Electrostatically guided dynamics: The root of fidelity in a promiscuous terpene synthase? *J. Am. Chem. Soc. 134*, 19454−19462.

(47) Faraldos, J. A., Gonzalez, V., and Allemann, R. K. (2012) The role of aristolochene synthase in diphosphate activation. *Chem. Commun. 48*, 3230−3232.