Supporting information for Mechanistic Molecular Modeling of Copalyl Diphosphate Synthase from *Arabidopsis thaliana* (AtCPS)

To model the diphosphate attachment for the full ligand, geometric constraints were added to model the bonding between the carbon tail and the diphosphate. The diphosphate was treated as a separate ligand. A chlorine atom was used to get the correct position of the O7 end of the diphosphate, with angle and distance constraints set between the carbon and diphosphate. Details on these constraints can be seen in the table and figure below:

Table S1. Modeling diphosphate-terpene bond with Rosetta constraints

Atoms constrained	Type	Constraint Value
Cl-PPi_O7	Distance	0.0 ± 0.1 Å
C52-PPi_O7	Distance	1.8 ± 0.1 Å
C50-C52-PPi_O7	Angle	110.6 ± 10.0 degrees
C52-PPi_O7-Ppi_P2	Angle	124.8 ± 10.0 degrees

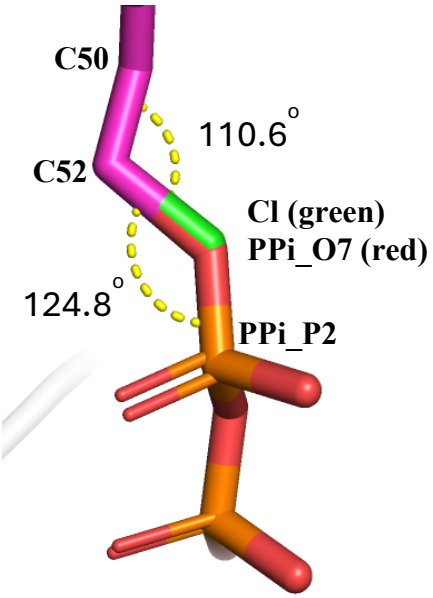

Table S1 (left): detailing the constraints and allowed deviation to model the reattachment of the diphosphate to the terpene ligand. Figure S1 (right) shows what the result of these constraints looks like in a docking model.

Important details for modeling the wild type mechanism with water bound (Figure 1 in main text) are shown below. The full constraint block for the wild type docking model can be found in an attached .zip file

Table S2. Water modeled to act as a base constraints:

Atoms constrained	Type	Constraint Value
O (water) – NE2 (H263)	Distance	2.8 ± 0.3 Å
O (water) – NE2 – CE1 (H263)	Angle	127.0 ± 12.0°
O (water) – OD1 (N322)	Distance	3.0 ± 0.5 Å

O (water) – OD1 – CG (N322)	**Angle**	**120.0 ± 12.0°**
O (water) – C12 (methyl carbon)	**Distance**	**2.5 ± 0.5 Å**
O (water) – C12 – C6	**Angle**	**110.0 ± 11.0°**
O (water) – N (K508 backbone)	**Distance**	**3.0 ± 0.3 Å**
O (water) – N – CA (K508 backbone)	**Angle**	**127.1 ± 15.0°**

Table S3. H263 modeled to act as a base constraints:

Atoms constrained	Type	Constraint Value
O (water) – NE2 (H263)	**Distance**	**2.8 ± 0.3 Å**
O (water) – NE2 – CE1 (H263)	**Angle**	**127.0 ± 12.0°**
O (water) – OD1 (N322)	**Distance**	**3.0 ± 0.5 Å**
O (water) – OD1 – CG (N322)	**Angle**	**120.0 ± 12.0°**
NE2 (H263) – C12 (methyl carbon)	**Distance**	**2.5 ± 0.5 Å**
NE2 (H263) – C12 – C6	**Angle**	**110.0 ± 11.0°**

ND1, the other nitrogen of the histidine side chain, was also modeled as a base maintaining the hydrogen bond between NE2 and the water. After filtering between poses with ND1 acting as a base and water acting as a base, only water acting as a base remained viable.

Table S4. General constraints used for all docking modeling D397 protonating the terminal π bond of GGPP to initiate the cyclization reaction and a protein-protein constraint where N425 hydrogen bonds to D379, activating it to react with the terpene precursor.

Atoms constrained	Type	Constraint Value
OOC (D379) – C18 (terpene)	**Distance**	**2.5 ± 0.5 Å**
OOC (D397) – C18 – C24 (terpene)	**Angle**	**109.0 ± 10.0°**

ND2 (N425) – OOC (D397)	**Distance**	**2.6 ± 0.3 Å**
ND2 – CG (N425) – OOC (D397)	**Angle**	**$147.2 \pm 10.0^{\circ}$**
CG – OOC (D397) – ND2 (N425)	**Angle**	**$120.3 \pm 10.0^{\circ}$**

Table S5. PPi and protein constraints for full ligand docking.

Atoms constrained	Type	Constraint Value
NZ (K463) – O3_Ppi	**Distance**	**2.8 ± 0.5 Å**
N (G209) – O2_PPi	**Distance**	**2.8 ± 0.3 Å**

Table S6. Water addition constraints:
Two waters were considered to add to the carbocation. Constraints between the two experiments are shown in the below table, and it is noted where they differ by the constraint between the oxygen of the water to the carbocation (C12). H_2O-1 had the same constraints to N322 and H263 shown in Table S2.

Atoms constrained	Type	Constraint Value
O (water 1 or 2) – C6	**Distance**	**3.0 ± 0.3 Å**
O (water 1 or 2) – C6 – C13 (β-hydroxy product)	**Angle**	**$100.0 \pm 15.0^{\circ}$**
H - O (water 1 or 2) – C6 – C13 (β-hydroxy product)	**Dihedral**	**$-21.2. \pm 15.0^{\circ}$**
O (water 1 or 2) – C6 – C13 – C9 (β-hydroxy product)	**Dihedral**	**$-163.8 \pm 15.0^{\circ}$**
O (water 1 or 2) – C6 – C13 (α-hydroxy product)	**Distance**	**3.0 ± 0.5 Å**
O (water 1 or 2) – C6 – C13 (α-hydroxy product)	**Angle**	**$134.4 \pm 30.0^{\circ}$**

O (water 1 or 2) – C6 – C13 – C9 (α-hydroxy product)	**Dihedral**	**50. 2 ± 20.0˚**

Figure S2. QM intermediate models of water addition.

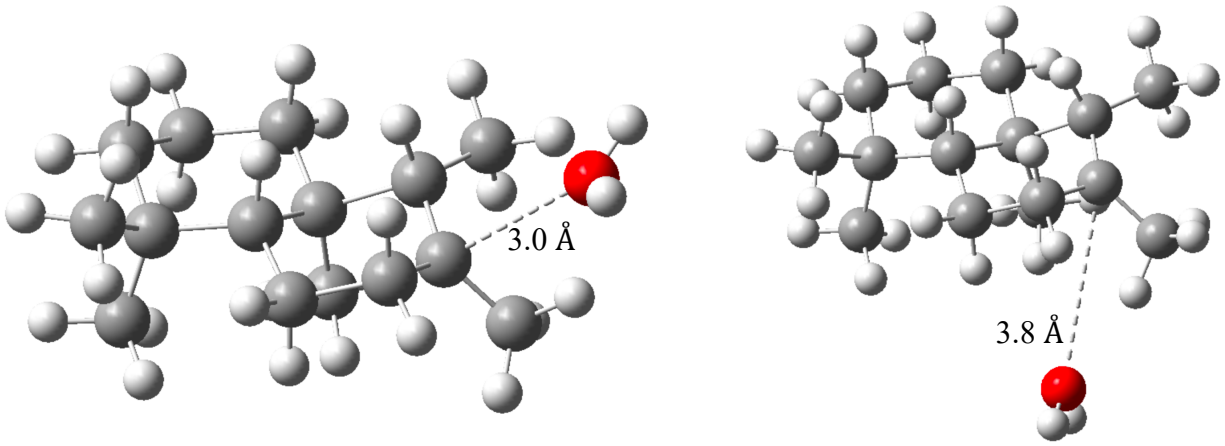

Figure S2: QM intermediate models with an implicit water modeled above and below a truncated model of intermediate A (labda-13E-en-8-yl$^+$) for production of 8-beta-hydroxy-ent-CPP (left) and 8-alpha-hydroxy-ent-CPP. Due to 8-alpha-hydroxy-ent-CPP model not being representative of the necessary conformation needed for water to add, constraints for 8-alpha-hydroxy-ent-CPP formation were made looser in the docking studies (see Table S6).

Figure S3. RMSD results for water addition compared to wild type.

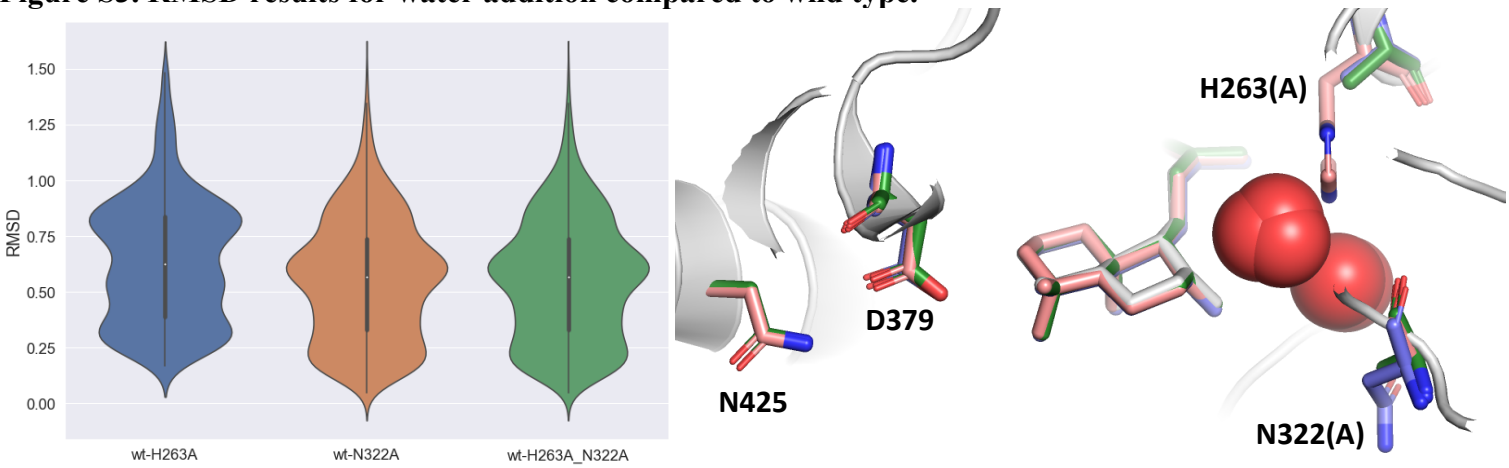

Figure S3: (Left) Violin plot showing carbon pair-wise RMSD between the docked terpene ligand between the wt-H263A mutant (blue), wt-N322A mutant (orange), and wt-H263A_N322A double mutant (green). (Right) An overlay of the lowest RMSD structures between poses docked in the wild type system (white), H263A mutant (blue), N322A mutant (tan), H263A_N322A double mutant (green). Waters for each mutant system are also shown as red spheres.

Figure S4. View of active site cavity to investigate alternative bases

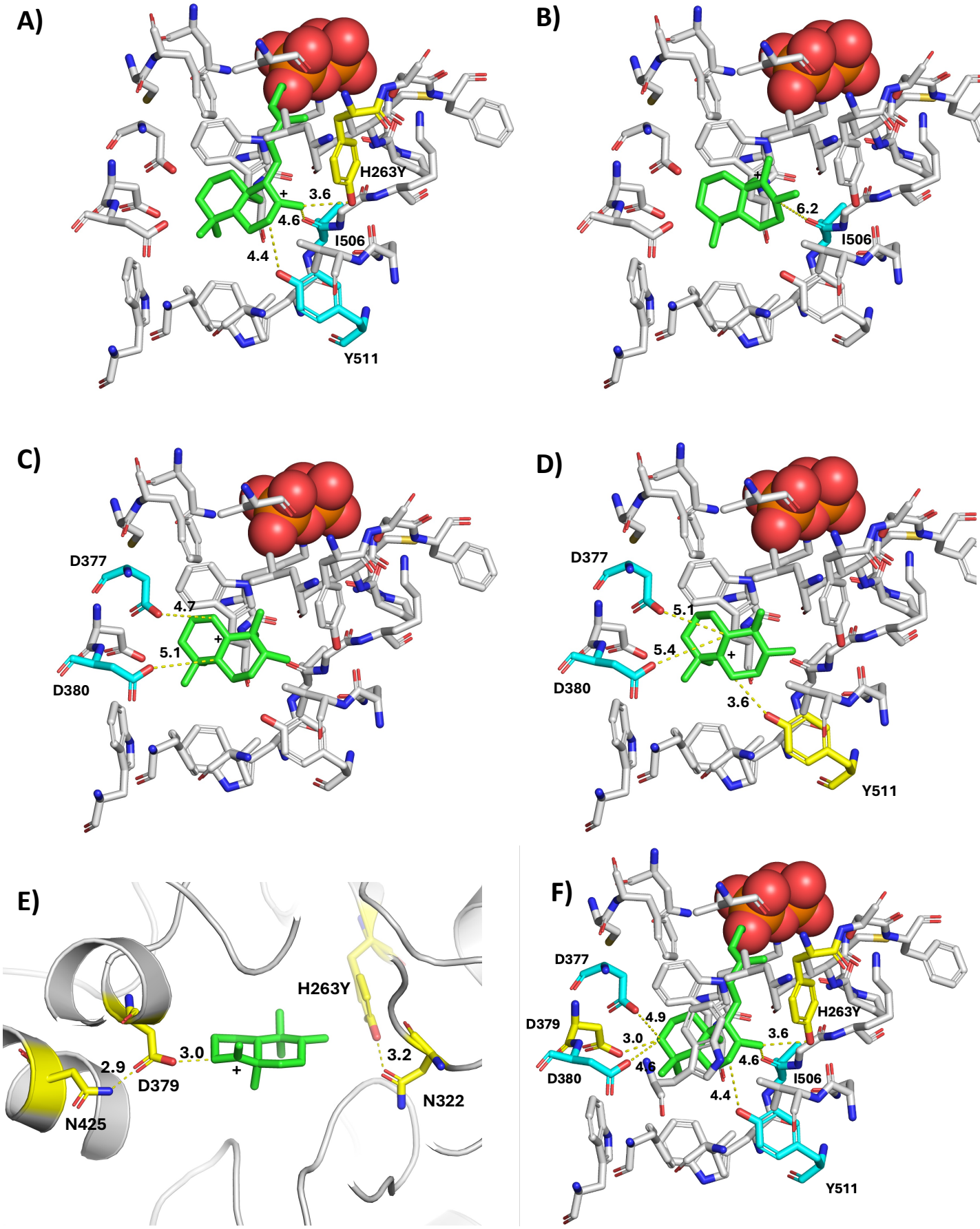

Figure S4: Showing representative binding pose of intermediates showcasing the location of the carbocation and nearby potential bases in the active site. For figures S4.A-S4.D, residues within 3-4 angstroms to a potential site of deprotonation are highlighted in yellow, while residues beyond 4 angstroms are shown in cyan. Not all residues beyond 4 angstroms are highlighted. A) Representative binding pose of intermediate A, untruncated. B) Representative binding pose of intermediate B1, truncated C) Representative binding pose of Intermediate C, truncated D) Representative binding pose of Intermediate D, truncated E) Representative binding pose of intermediate E, truncated, bound and being deprotonated by D379 to C3, the same carbon that initiates cyclization of GGPP. F) Identical to figure S4.A but showing the omitted residue W269 which was removed for clarity. All distances shown are in units of angstroms.

Figure S5. H263Y mutant acting as a base.

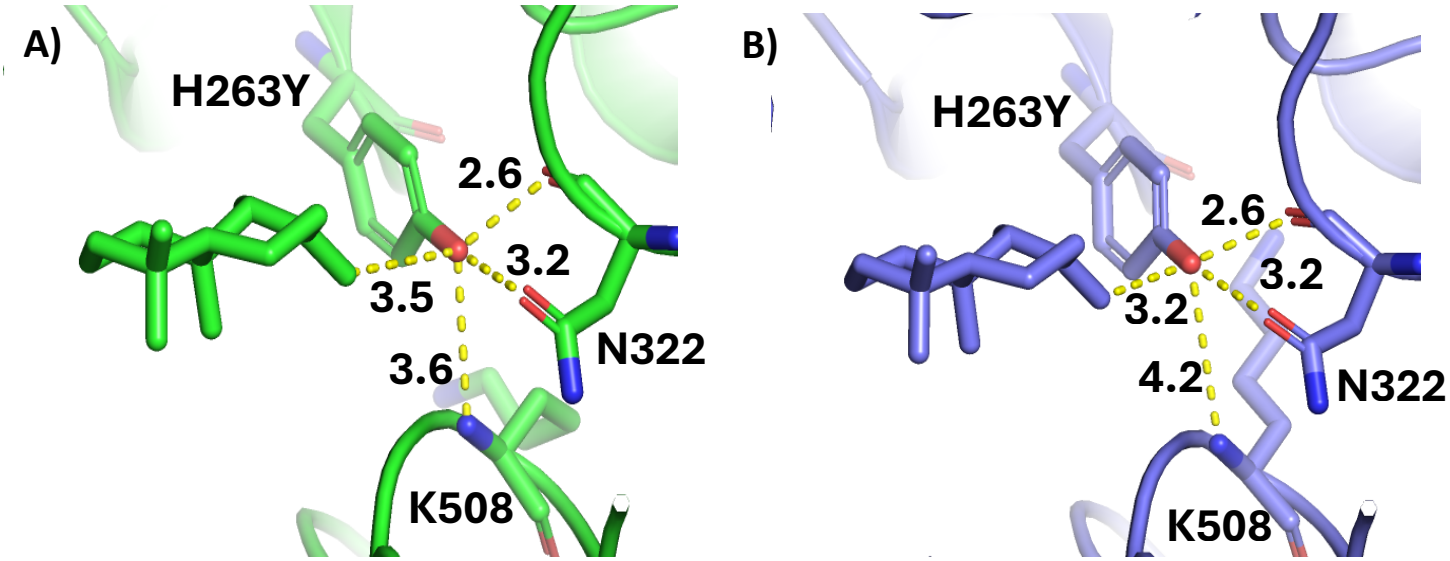

Figure S5: A) Representative binding pose of H263Y docking with no explicit constraints between H263Y and the ligand. B) Representative binding pose of H263Y constrained to the terminal methyl group to simulate it acting as a base. H263Y All distance values provided are in units of angstroms.

Figure S6. Investigating Y511 deprotonating intermediate A

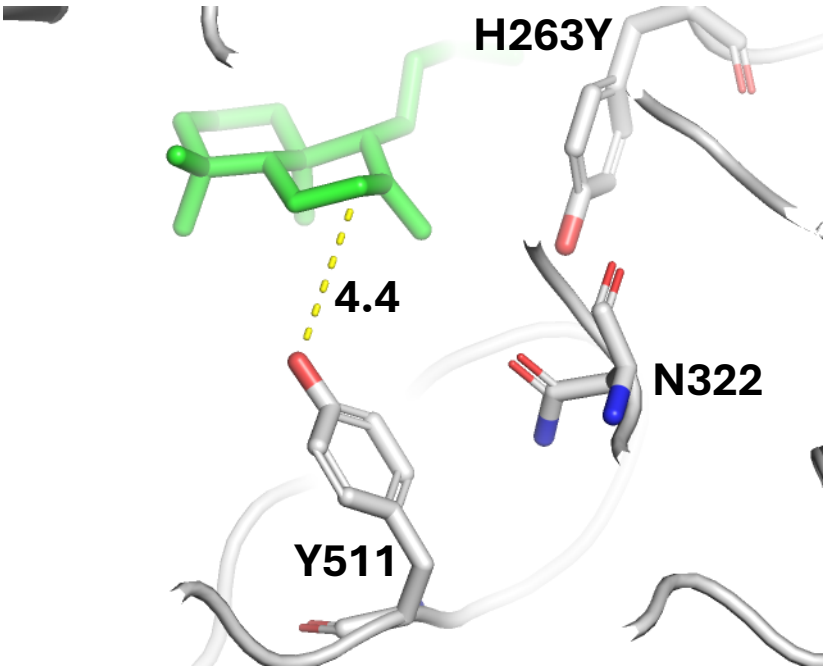

Figure S6: Representative binding pose of H263Y showing Y511 to the ligand, identifying that it is too far from a point of deprotonation and is at a poor angle to deprotonate the ent-labda-13E-en-8-yl^{+} intermediate (green).