

Resolving the Hydride Transfer Pathway in Oxidative Conversion of Proline to Pyrrole

Atanu Acharya, Dongqi Yi, Anna Pavlova, Vinayak Agarwal,* and James C. Gumbart*



Cite This: *Biochemistry* 2022, 61, 206–215



Read Online

ACCESS |



Metrics & More

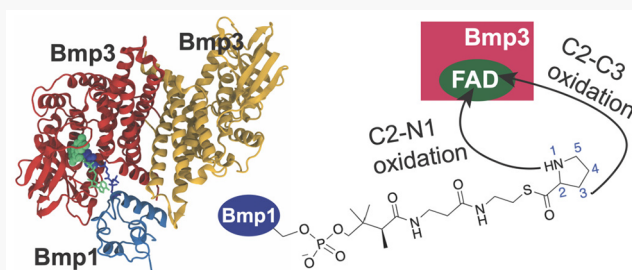


Article Recommendations



Supporting Information

ABSTRACT: Thiotemplated pyrrole is a prevailing intermediate in the synthesis of numerous natural products in which the pyrrole is tethered to a carrier protein (CP). Biosynthesis of the pyrrole requires oxidation of an L-proline side chain. Herein, we investigate the biocatalytic mechanism of proline-to-pyrrole synthesis by molecular dynamics simulations, quantum mechanics/molecular mechanics simulations, and electronic structure calculations using the recently reported (Thapa, H. R., et al. *Biochemistry* 2019, 58, 918) structure of a type II nonribosomal protein synthetase (NRPS) Bmp3-Bmp1 (Oxidase-CP) complex. The substrate (L-proline) is attached to the Bmp1 (CP), and the catalytic site is located inside the flavin-dependent oxidase (Bmp3). We show that the FAD isoalloxazine ring is stabilized in the catalytic site of Bmp3 by strong hydrogen bonding with Asn123, Ile125, Ser126, and Thr158. After the initial deprotonation followed by an enamine–imine tautomerization, oxidation of the C2–C3 or C2–N1 bond, through a hydride transfer (from either C3 or N1), is required for the pyrrole synthesis. Computational results indicate that the hydride transfer is more likely to occur from C3 than N1. Additionally, we demonstrate the elasticity in the oxidase active site through enzymatic synthesis of proline derivatives.



The pyrrole heterocycle is widespread in biological chemistry. In primary metabolism, pyrroles serve to bind metal ions in porphyrins. The metal binding capacity of pyrroles is exploited by bacterial secondary metabolites as well.¹ Pyrroles intercalate DNA and serve as a core synthon around which many pharmaceutical molecules are constructed.² Pyrroles in several secondary metabolite biosynthetic pathways are constructed via a type II nonribosomal peptide synthetase (NRPS) adenylation–oxidation cascade that involves the four-electron oxidation of the pyrrolidine side chain of L-proline (Figure 1a).³ Prior to oxidation, the L-proline carboxylate is activated by adenylation, catalyzed by an adenytransferase, which then transfers L-proline to a carrier protein (CP) phosphopantetheinyl thiol.^{4–7} The phosphopantetheinyl thiol is derived from coenzyme A (CoA) and is appended post-translationally to a serine side chain hydroxyl of the apo-CP to create the holo-CP.⁸ An oxidase then catalyzes the conversion of the pyrrolidine to pyrrole in a sequence of two two-electron oxidations with electrons funneled to molecular oxygen via a flavin cofactor.^{9,10} The thiotemplated pyrrole thus developed can be delivered to type I polyketide synthases for chain elongation,¹¹ to type I NRPSs for ligation to other amino acids,¹² and to other enzymes such as halogenases¹³ or methyltransferases¹⁴ for further modification.

The biosynthesis of thiotemplated pyrroles using the type II NRPS scheme mentioned above relies on intramolecular protein–protein interactions, of the CP with the adenytransferase and of the CP with the oxidase, as well as catalysis

in the respective adenytransferase and oxidase active sites. Biochemically validated atomistic details for both sets of protein–protein interactions are now available.^{5,10} Catalysis in the adenytransferase active site follows the well-accepted two-step process involving the ATP-mediated activation of the amino acid carboxylate followed by thioesterification of proline to the CP phosphopantetheinyl thiol. Conversion of this thiotemplated proline-derived pyrrolidine to pyrrole via a pyrroline intermediate⁹ in the oxidase active site remains less resolved. The structure and overall mechanism of pyrroli(di)ne oxidases resemble those of the mitochondrial fatty acid acyl-CoA dehydrogenases (Figure 1b). Catalysis in fatty acid dehydrogenases proceeds by abstraction of a proton from the fatty acyl C_α by a catalytic glutamate residue followed by hydride transfer from C_β to the flavin cofactor leading to the C_α–C_β oxidation.^{15,16} Pyrrolidine and pyrroline thioesters are unstable, which has prevented a structural description of the substrate- or intermediate-bound states of the pyrroli(di)ne oxidases.^{10,17,18} Furthermore, a lack of insight into substrate and intermediate binding in the oxidase active site has

Received: November 12, 2021

Revised: January 11, 2022

Published: January 24, 2022



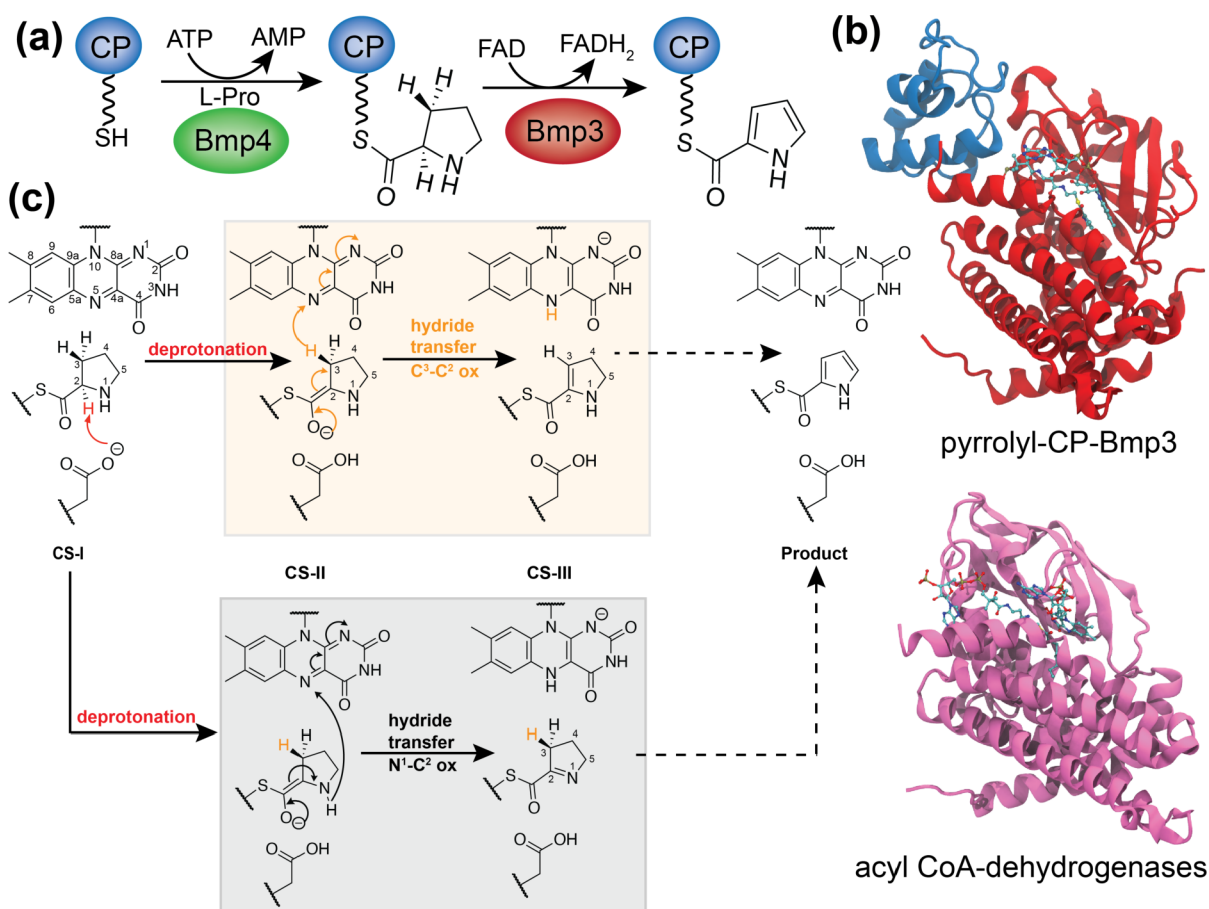


Figure 1. Chemical steps involved in pyrrole biosynthesis. (a) Adenylation–oxidation cascade illustrated as a Chemdraw scheme formally showing the product of the adenylation. (b) Overall structural comparison of pyrrolyl-CP-Bmp3 (Protein Data Bank entry 6CXT) and pig liver mitochondrial acyl CoA-dehydrogenases (Protein Data Bank entry 3MDE). The flavin cofactor and the phosphopantetheinylated substrate/product are shown in ball-and-stick representation. (c) Key steps and plausible mechanisms in the Bmp3-catalyzed pyrrole biosynthesis. The catalytic state (CS) and product state are defined, as well. Atom numbering schemes of pyrrolidine and isoalloxazine rings are shown in the CS-I state. The hydride transfer steps from C3 and from N1 are highlighted in orange and gray, respectively.

precluded deciphering of the site of hydride transfer from the pyrroli(di)ne to the flavin cofactor. Two mechanistic routes are conceivable, which converge to identical outcomes with interspersed imine–enamine tautomerizations (Figure 1c).¹⁹ The first is that the enzyme acts as an amine oxidase, which would involve hydride transfer from ^{prolyl}N1 to ^{flavin}N5, as has been proposed previously.^{17,20} The second possibility is that the pyrroli(di)ne oxidase acts as a C α –C β desaturase with the hydride being transferred from ^{prolyl}C3 to ^{flavin}N5. At present, these two routes are indistinguishable.

Herein, we investigate the first two steps of pyrrole biosynthesis using molecular dynamics (MD), quantum mechanics/molecular mechanics (QM/MM) simulations, and electronic structure calculations. In particular, we assess the feasibility of the Glu242 residue to initiate the biosynthesis. More importantly, we provide mechanistic insights into the hydride transfer from the substrate to the FAD bound to the Bmp3 enzyme. We investigate the energy barriers for the unique hydride transfer process from the C3 atom of the substrate instead of from the N1 atom. Furthermore, by enzymatic synthesis of proline derivatives, we demonstrate elasticity in the oxidase active site, which supports the computational description of substrate binding.

METHODS

System Preparation and MD Simulations. We used the tetrameric complex structure (Protein Data Bank entry 6CY8) of the holo Bmp3-Bmp1 system at 2.7 Å resolution.¹⁰ To test the stability and the functional relevance of the Bmp3-Bmp1 complex, we prepared a system containing the substrate, a single bound Bmp1 (chain A), and a dimer of Bmp3 (chains B and C), where only chain B contains the catalytic FAD molecule (Figure 3a). Therefore, we call chains B and C active and inactive Bmp3, respectively. Bmp3 structures have eight missing residues (residues 24–31), which were modeled using SWISS model.²¹ The Bmp1 structure is missing five N-terminal residues and one C-terminal residue, which were added using the VMD PSFgen plugin.²² We solvated this complex in an explicit TIP3P water box²³ with at least 12 Å of solvent padding around the protein. The system was neutralized by adding 28 Na⁺ ions. We used the CHARMM36m force field²⁴ for protein residues. Initial force-field parameters for the FAD molecule and the substrate were taken from CGenFF,²⁵ which assigns them by analogy to existing CHARMM parameters. Charges on the isoalloxazine ring of FAD and the pyrrolidine ring of the substrate (in the CS-I state) were further improved using the Force Field Toolkit (ffTK), as implemented in VMD.^{26,27} Parameters for

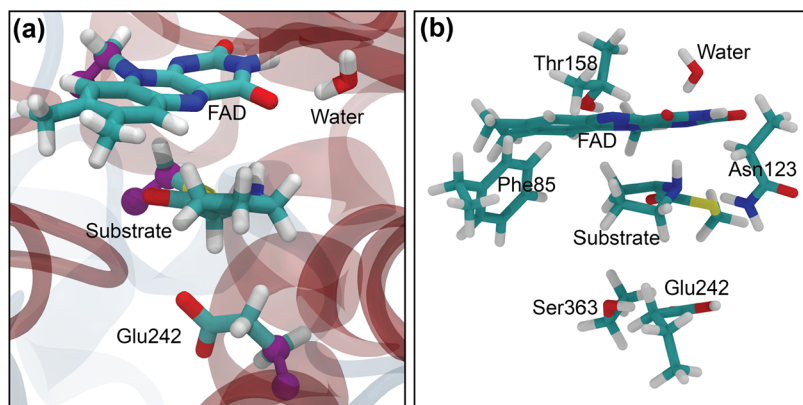


Figure 2. Computational models used in this study. (a) QM/MM scheme adopted in the simulations of the CS-I state. QM-MM bonds at the boundary are colored purple. (b) Cluster model adopted for the transition state calculations.

the dihedral angles, where the pyrrolidine rings connect with the phosphopantetheine (Ppant) appendage, were also improved using fTK (see the [Supporting Information](#) for details). MD simulations were performed in the *NPT* ensemble using periodic boundary conditions at a constant temperature of 310 K and a constant pressure of 1 atm. Long-range electrostatics were computed using the particle-mesh Ewald (PME) method.²⁸ We used a 12 Å cutoff with a switching function between 10 and 12 Å. All MD simulations were performed using a 2 fs time step.

We first minimized the modeled missing pieces followed by minimization of the entire system. We equilibrated the solvation box for 0.5 ns with the proteins restrained. After that, we equilibrated the system for an additional 0.5 ns with a 2 kcal mol⁻¹ Å⁻² restraint on protein backbone atoms, followed by a 50 ns equilibration without any restraints. The final structure from this equilibration was used to initiate all QM/MM simulations. Additionally, we ran two independent production simulations for 500 ns, from which we collected trajectory data every 20 ps for subsequent analysis.

QM/MM Simulations. The protein system was extracted from the 50 ns equilibration in the CS-I state. Because the active-site charge distribution changes significantly during the reaction, each system was resolvated and neutralized before QM/MM simulations. The QM region included all residues directly involved in the reaction. The water molecule next to FAD is also included in the QM region to investigate its role in each step of the reaction. The resultant QM region included FAD, the substrate, Glu242, and one water molecule (Figure 2a). The rest of the system was treated as point charges.

The link-atom scheme was used to satisfy the valency of the terminal QM atom, where a hydrogen atom was added as the link atom. The redistributed charge and dipole (RCD) method was used to modify charges and dipoles in the QM/MM boundary region to avoid overpolarization. Additionally, we used the switch function to modify point charges and enforced net-zero electrostatic interactions between charges at the cutoff distance. The net charge of the system was also set to zero. The QM part of the QM/MM calculation was performed at the BP86/def2-SVP level of theory using ORCA.²⁹ The same level of theory was successfully used to model enzymatic hydride transfer from flavin mononucleotide (FMN) using the QM/MM methodology described previously.³⁰ Becke–Johnson damping (D3BJ) was used for dispersion corrections to the DFT.³¹ QM/MM simulations were performed using the

NAMD/ORCA interface³² and a 0.5 fs time step. After minimization, we increased the temperature of the system to 310 K followed by equilibration for 100 fs with backbone restraints. Finally, we ran one production simulation for the CS-II state using the *NPT* ensemble at 310 K and 1 atm. The same procedure was followed for QM/MM simulations of the CS-I and CS-II states. The difference between the two states was included in the QM part of the QM/MM simulation.

QM Calculations in a Cluster Model. The cluster model was constructed using the QM/MM equilibrated structure of the CS-II system. In addition to residues already included in the QM region, we added Ser363, Asn123, Thr158, and Phe85 residues, which are located close to the active site (Figure 2b). Phe85 and Thr158 can play important roles because the hydrogen bond of FAD with Thr158 and hydrophobicity near the active site are required to tune the catalysis.³³ In total, the cluster model included five residues, the substrate, FAD, and one water molecule. All C_α-atom positions were frozen during subsequent QM optimizations to preserve the overall structure of the residues. The CS-II and CS-III states and the TS were optimized at the CAM-B3LYP/6-31+G(d) level of theory with D3BJ dispersion corrections³¹ using Gaussian16.³⁴ While the B3LYP functional has been used before to describe electronic states of FAD,³⁵ the use of additional long-range corrections is usually recommended for molecules with fused rings.³⁶ Therefore, we used the CAM-B3LYP functional, along with an implicit dispersion correction, to correctly describe nonbonded interactions within the cluster model. The average dielectric constant value inside a protein is usually in the range of 6–7.³⁷ While some studies have estimated a slightly lower dielectric constant (3–4) inside a protein,^{38,39} it is unlikely that implicit solvents with a dielectric constant in the range of 3–7 will show discernible differences in molecular properties. Thus, we mimic the protein environment by using the CPCM implicit solvent model⁴⁰ with a low dielectric constant ($\epsilon = 6.2528$).

RESULTS AND DISCUSSION

Model Description and the Substrate-Bound State.

Prior genetic and enzymological description of a thiotemplated pyrrole biosynthetic pathway from marine bacteria had identified the L-proline-derived pyrrolidine/pyrrolyl CP to be localized to the N-terminus of the Bmp1 polypeptide [henceforth termed Bmp1(CP)] and the pyrrolidine oxidase as the Bmp3 polypeptide.⁷ Subsequent crystallographic

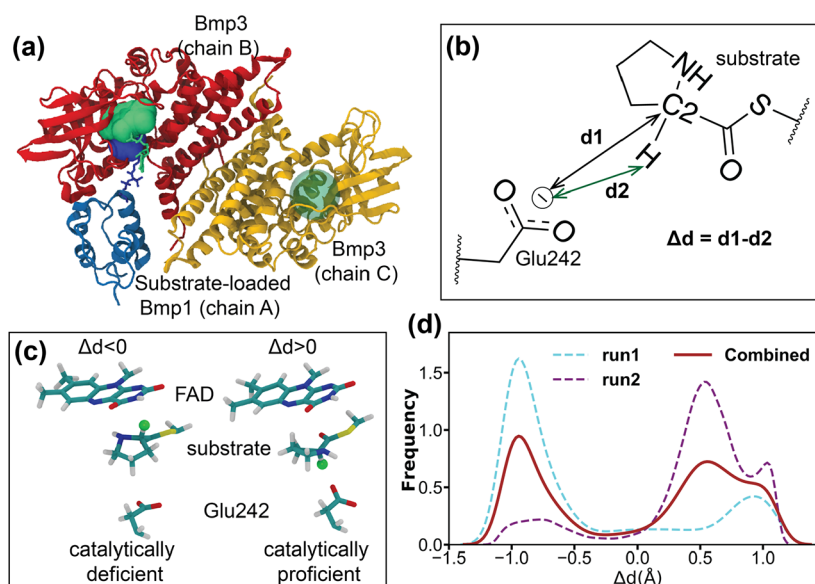


Figure 3. Distance and conformations of the substrate in the CS-I state. (a) Model Bmp1(CP)/Bmp3 system with one bound substrate (blue) and FAD (green). Active and inactive Bmp3 monomers are colored red and yellow, respectively. Substrate-bound Bmp1 is shown as a blue cartoon. The active site of chain C is lacking an FAD cofactor in the location shown as a green transparent sphere. (b) Definition of relevant distances. The center of the two oxygen atoms of Glu242 was used for the distance calculations. We defined $d_1 = d(\text{C2-Glu242})$, $d_2 = d(\text{C2-H-Glu242})$, and $\Delta d = d_1 - d_2$. (c) Representative structures of two conformations. C2-H atoms are shown as green spheres. (d) Distribution of the substrate conformations obtained from two independent MD simulations. Data from the two 500 ns simulations were pooled for this analysis.

description of the Bmp1(CP)/Bmp3 complex identified a tetrameric Bmp3 core similar to that of acyl-CoA dehydrogenases.¹⁰ Each Bmp3 monomer binds one Bmp1(CP), thus organizing a 4:4 Bmp1(CP)/Bmp3 stoichiometric complex. Each Bmp1(CP) contacts two Bmp3 monomers in this tetrameric complex. One of these Bmp3 monomers is the one to whose active site the phosphopantetheinylated molecular cargo is delivered by Bmp1(CP). The second Bmp3 monomer makes contact with only the Bmp1(CP) without participating in catalysis in our computational model. For computational analysis of catalysis occurring in the Bmp3 active site, our model system comprises a Bmp1(CP) peptide in complex with two Bmp3 peptides with the flavin cofactor preserved in the catalytically relevant Bmp3 monomer only. The model henceforth used for computation is illustrated in Figure 3a.

The structure of the pyrrolidine oxidase Bmp3 in complex with Bmp1(CP) was described in the product-bound state.¹⁰ The substrate-bound state, termed catalytic state I [CS-I (Figure 1)], was inaccessible. Other crystal structures of pyrrolidine oxidases were similarly not available with bound thioesterated substrates.^{18,41} Hence, we first sought to recover the physiological substrate-bound CS-I state for Bmp3. To describe the CS-I state, we equilibrated the system in which the thioesterified pyrrolidine heterocycle was modeled in the Bmp3 active site adjacent to the isoalloxazine moiety of FAD, followed by extended MD simulations. Two independent 500 ns simulations were run, as multiple replicas can more efficiently explore conformational space than a single longer run.⁴² The initial position of the substrate in the CS-I state was set as found in the structure of the pyrrole in the product-bound state. We analyzed two replicas of 500 ns MD simulations of the CS-I model. For the occupancy analysis, we pooled data from both replicas.

In this CS-I model, we observed that the FAD isoalloxazine ring is stabilized by hydrogen bonding with Asn123, Ile125,

Ser126, and Thr158 (Figure S1a–c). The percentage of all hydrogen bonding interactions among FAD, Bmp3, and water is provided in Table S1, and binding interactions of the FAD are illustrated in Figure S1. The pyrrolidine ring forms a hydrogen bond with the Asn123 side chain (Figure S2a), while the carbonyl group forms hydrogen bonds with the Ser363 backbone (Figure S2b,c). The substrate forms hydrogen bonds with Bmp3 predominantly through the Ppant moiety (Figure S2c–i), which attaches the pyrrolidine ring with the Bmp1-(CP). In summary, the FAD isoalloxazine ring and the substrate form several hydrogen bonds with the Bmp3 protein to stabilize the catalytic site.

Similarity to acyl-CoA dehydrogenases and prior mutational analysis suggest that the Glu242 side chain carboxylate acts as the catalytic base for abstraction of a proton from the pyrrolidine side chain to initiate catalysis (Figure 1). We¹⁰ and others¹⁸ had postulated that Glu242 abstracts the pyrrolidine C2 proton (pyrrolidine C2-H). Therefore, the distances between the Glu242 side chain carboxylate and pyrrolidine C2 and the pyrrolidine C2-H atoms are critical for catalysis. We calculated these distances along the MD trajectories (Figure S3).

We observed that the pyrrolidine C2 and pyrrolidine C2-H atoms stay close to the Glu242 side chain in MD simulations (Figure S3b). Because rotation is allowed around the pyrrolidine C2-carbonyl carbon bond, the pyrrolidine ring can adopt two dominant conformations. To quantify the population distribution, we analyzed the differences between pyrrolidine C2–Glu242 and pyrrolidine C2-H–Glu242 distances as $\Delta d = d_1 - d_2$ (Figure 3b). One conformation brings the pyrrolidine C2-H proximal to the Glu242 carboxylate, while the other conformation places the pyrrolidine C2-H distal to Glu242 (Figure 3c). Because the two conformations can interconvert by rotation around a single bond, the rotational energy barrier should be small. This rotation does not change the pyrrolidine C2–Glu242 distance. If the C2-H atom is pointed toward Glu242, then $\Delta d > 0$. In

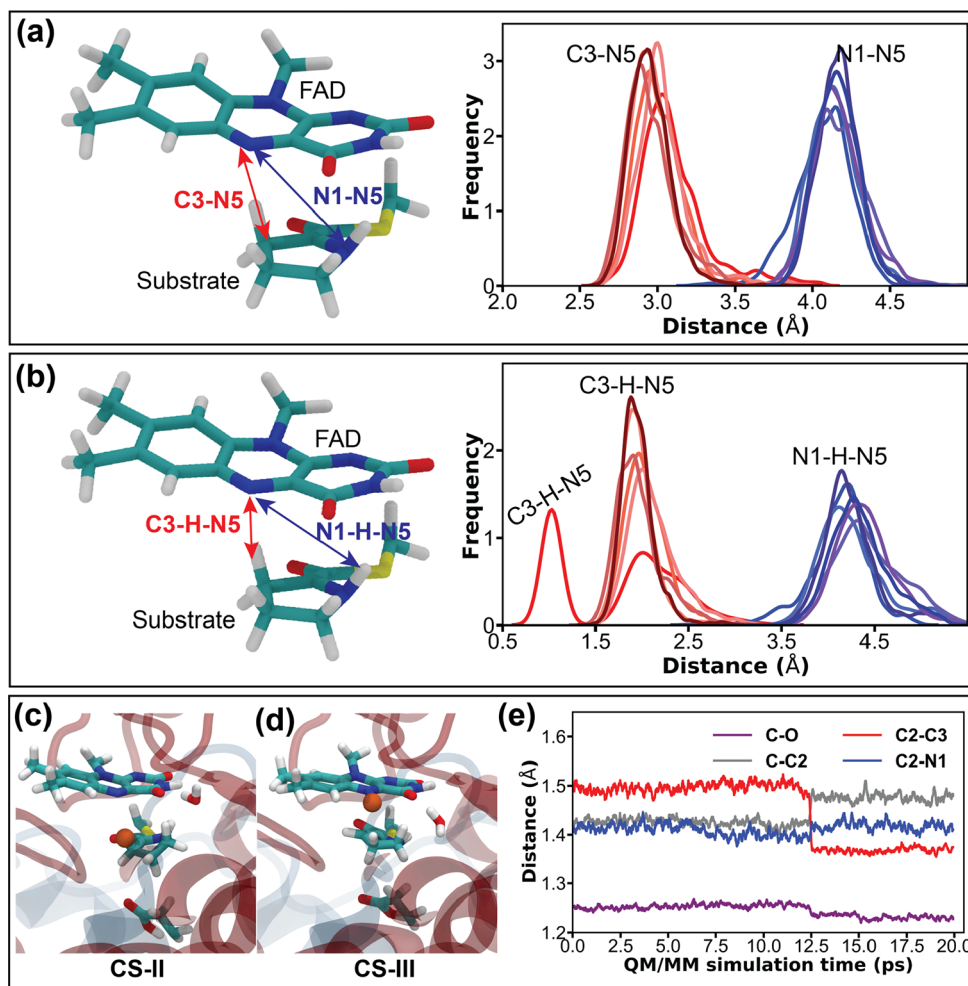


Figure 4. Distance distributions of the substrate in CS-II calculated from six independent QM/MM simulations. (a) Relative populations of distances between key pyrrolidine atoms (N1 and C3) and the ^{flavin}N5 atom. (b) Relative populations of distances from the ^{flavin}N5 atom to pyrrolidine C3-H and pyrrolidine N1-H. Representative structures of the system (c) before and (d) after the hydride transfer. The relevant hydrogen atom is shown as an orange sphere. (e) Analysis of key bond distances calculated along the QM/MM simulation in which C2–C3 oxidation is explicitly observed.

contrast, $\Delta d < 0$ when the C2-H atom is pointed away from the Glu242 side chain.

The combined data from two independent simulations show that the relative abundance of two rotamers are comparable (Figure 3d). The catalytically proficient conformation is represented by $\Delta d > 0$, because the deprotonated carboxylate of the Glu242 side chain is geometrically close to the pyrrolidine C2-H atom. We observed that the catalytically proficient conformation was favored in the second simulation, while the catalytically deficient conformation was favored in the first simulation (Figure 3c,d). In the $\Delta d > 0$ conformations, we observe a peak at $\Delta d \approx 0.5$ Å, which corresponds to a rotamer with the C2-H bond oriented parallel to the Glu242 carboxylate moiety (Figure S5). Nevertheless, all $\Delta d > 0$ conformations are catalytically efficient because the C2-H atom is close to Glu242 in these conformations with peaks at $\Delta d \approx 0.5$ and 1.0 Å.

We performed additional QM/MM simulations on the CS-I state to obtain more accurate dynamics of the catalytic site. Details of the QM/MM simulations are described in the Methods. From the QM/MM simulation, we observed that the most probable pyrrolidine C2–Glu242 distance is larger than the pyrrolidine C2-H–Glu242 distance, while both substrate atoms

remain close to the Glu242 side chain (also see Figure S6a,b). Note that we do not observe the catalytically deficient $\Delta d < 0$ state from QM/MM simulation of the CS-I state, possibly because of the shorter time scale (15 ps) of the QM/MM simulation. Within that time scale, QM/MM simulation sampled only conformations conducive to the transfer of a proton from the substrate to the Glu242 side chain (Figure S6a,b). While the substrate can bind in the active site in either conformation, catalysis can proceed in only one conformation and not in the other. On the basis of these results, we posit that the $\Delta d > 0$ state illustrated in Figure 3c is the catalytically proficient substrate-bound state for pyrrolidine oxidase Bmp3. Mutations of Glu242 in Bmp3 and the corresponding residue Glu244 in pyrrolidine oxidase AnaB render the enzymes inactive.^{10,17} Taken together, these results indicate the feasibility of the Glu242 protonation by the transfer of a proton from the C2 atom of the substrate. More importantly, the active site is open enough for free rotation around the C–C2 bond, allowing the pyrrolidine ring to fully rotate.

Hydride Transfer in the CS-II Intermediate. After thiotemplated substrate binding and abstraction of the pyrrolidine C2-H by the Glu242 side chain carboxylate, catalytic state II [CS-II (Figure 1)] is defined by the protonated Glu242 side

chain (Figure 1c). Pyrrolidine oxidation can proceed by tautomerization followed by the transfer of a hydride from pyrrolidine C3 to flavin N5, or via the transfer of a hydride from pyrrolidine N1 to flavin N5 (Figure 1c). Both of these routes have been proposed in the literature, but evidence supporting either of the two routes has been lacking as the enolate intermediate was not accessible experimentally. To address this knowledge gap, we performed six independent 20–25 ps QM/MM simulations for the CS-II intermediate state using the same QM/MM scheme adopted for the CS-I simulations. In the CS-I state, all carbon and nitrogen atoms of the pyrrolidine ring are sp^3 hybridized. In contrast, the pyrrolidine C2 and pyrrolidine C3 atoms (for C2–C3 oxidation) of the substrate are in sp^2 hybridization states (Figure 1). In the case of plausible C2–N1 oxidation, C2 and N1 will become sp^2 hybridized. Both mechanisms have been proposed before in the flavin-catalyzed biosynthesis processes.^{9,17,20} We investigated which mechanism is more likely to occur using a total of 145 ps of QM/MM simulations along with transition state calculations.

We analyzed the pyrrolidine C3–flavin N5 and pyrrolidine N1–flavin N5 distances to identify the closer hydride donor between pyrrolidine N1 and pyrrolidine C3 (Figure 4a). We observed that the most probable pyrrolidine C3–flavin N5 distance was shorter (~ 3.0 Å) than the pyrrolidine N1–flavin N5 distance (~ 4.2 Å) (Figure 4a). Additionally, we calculated the distance from the flavin N5 atom to pyrrolidine C3–H and pyrrolidine N1–H as illustrated in Figure 4b. The most probable pyrrolidine C3–H–flavin N5 distances were shorter than the pyrrolidine N1–H–flavin N5 distances (1.0–2.0 and 4.0–4.5 Å, respectively) as shown in Figure 4b. Therefore, pyrrolidine C3 and pyrrolidine C3–H atoms are closer to the FAD cofactor than pyrrolidine N1 and pyrrolidine N1–H, respectively. These results allow us to posit that the enolate intermediate tautomerizes such that the transfer of a hydride to the flavin cofactor occurs from pyrrolidine C3 and not from pyrrolidine N1. This then sets up the CS-III state in which the pyrrolidine C2–C3 bond should be unsaturated (Figure 1).

In support of the aforementioned assertion, in one of the QM/MM simulations, we captured the hydride transfer event from CS-II pyrrolidine C3 to flavin N5 (Figure 4c,d). Therefore, the pyrrolidine C3–H–flavin N5 distance distribution has two peaks for one of the simulations as illustrated in Figure 4b. We further queried the structural perturbation accompanying the hydride transfer in this simulation. We observed that the pyrrolidine C2–C3 bond length decreased from 1.50 to 1.37 Å, while the carbonyl C–pyrrolidine C2 bond length increased from 1.42 to 1.50 Å with the hydride transfer from the C3 atom (Figure 4e).

We optimized a model substrate in CS-II and CS-III in a low-dielectric solvent ($\epsilon = 6.2528$) using the conductor-like polarizable continuum model (CPCM) implicit solvent model to obtain ideal pyrrolidine C2–C3 and pyrrolidine C–C2 bond lengths. The optimization was performed at the CAM-B3LYP/6-31+G(d) level of theory. The optimized pyrrolidine C2–C3 and carbonyl C–pyrrolidine C2 bond lengths are 1.51 and 1.35 Å, respectively, in CS-II and 1.34 and 1.48 Å, respectively, in CS-III (Figure S8). The changes in these bond lengths match the probable bond lengths obtained from the QM/MM simulation of the states before and after the hydride transfer event. In contrast, pyrrolidine C2–N1 and carbonyl C–O distances remain similar before and after the hydride transfer event. Therefore, we observed only one peak in the most probable pyrrolidine C2–N1 distance (Figure S7c), indicating that pyrrolidine N1 did not participate in this hydride transfer event.

Taken together, these results indicate that pyrrolidine C2–C3 oxidation during the hydride transfer is more likely to occur than pyrrolidine C2–N1 oxidation.

In a larger context, diverse flavin-dependent dehydrogenases and oxidases frequently adopt a specific relative orientation of the isoalloxazine moiety and the oxidation site on the substrate.⁴³ The recurrent relative orientation is quantified⁴³ by a flavin N5–pyrrolidine C3 distance of 3.0–3.8 Å and a flavin N10–flavin N5–pyrrolidine C3 angle of 98–117° (except for cholesterol oxidase). We calculated these parameters from the six independent QM/MM simulations of CS-II. The results are summarized in Table 1. We found that both distance and angle

Table 1. Average Values of Geometric Parameters Calculated from Six Independent QM/MM Simulations of CS-II

QM/MM run	flavin N5–pyrrolidine C3 (Å)	flavin N10–flavin N5–pyrrolidine C3 (deg)
1	3.10	104.6
2	2.97	105.2
3	3.00	106.2
4	3.04	106.8
5	2.95	104.4
6	2.95	107.2

parameters are within the ranges observed in other flavin-dependent dehydrogenases and oxidases. Furthermore, the flavin N5–pyrrolidine C3 distances calculated from the QM/MM simulations are at the lower end of the range, providing additional support for the hydride transfer mechanism from pyrrolidine C3.

Experimentally Querying Active-Site Flexibility. Data presented above point toward the Bmp3 active site accommodating a large degree of flexibility in substrate binding with catalytically proficient and catalytically non-proficient modes of substrate binding being energetically equivalent. To further explore the substrate tolerance of the Bmp3 active site, we enzymatically generated the pyrrolidine C4–Me diastereomers and explored if they could be processed in the Bmp3 active site. Using previously described enzymatic synthesis schemes,^{44–46} L-leucine oxidative cyclization by the LdoA/NosE/NosF enzymes furnished the (2S,4S)-4-methylproline [1 (Figure 5)] and the GriE/GriF/ProC enzymes furnished the (2S,4R)-4-methylproline (2) diastereomer. The 4-methylproline isomers were purified from preparative scale enzyme reactions and confirmed by liquid chromatography/mass spectrometry after derivatization of the carboxylate (Figure S13a,b). Molecules 1 and 2 were used as substrates in a coupled adenylation–pyrrolidine oxidation reaction in which the adenylation Bmp4 transferred the proline derivatives to the phosphopantetheinyl thiol of holo-Bmp1 (CP) *in vitro* followed by oxidation by wild type Bmp3.⁷

As monitored by the formation of the characteristic cycloacylpantetheine fragmentation ejection ion,⁴⁷ full conversion of holo-Bmp1(CP) to 4-methylpyrrole-Bmp1(CP) was achieved in this reaction. While the activity of the adenylation Bmp4 would indeed be compromised as compared to activation of the native substrate L-proline, this result demonstrates elasticity in the Bmp3 active site, such that it can successfully incorporate diverse substrates.

Transition State and the Free Energy Barrier of Hydride Transfer. Although QM/MM simulations indicate the feasibility of the C2–C3 oxidation, they do not provide any

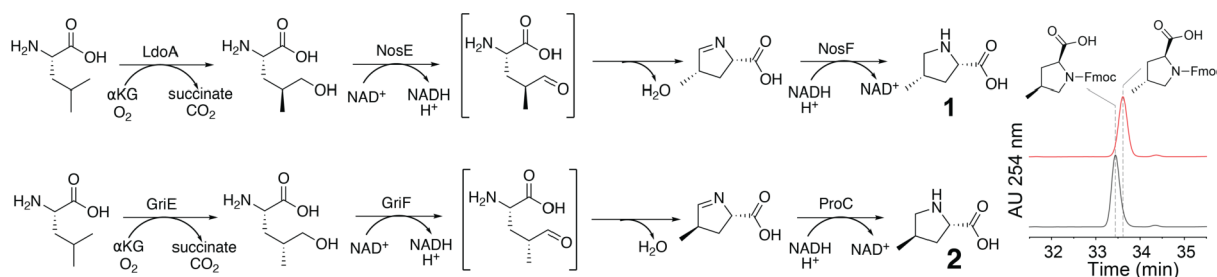


Figure 5. Oxidative conversion of L-leucine to 4-methylproline diastereomers **1** and **2** by the LdoA/NosE/NosF and GriE/GriF/ProC enzymes. After derivatization of the amino acids with Fmoc-Cl, the diastereomers could be chromatographically separated as illustrated by ultraviolet absorbance chromatograms recorded at 254 nm.

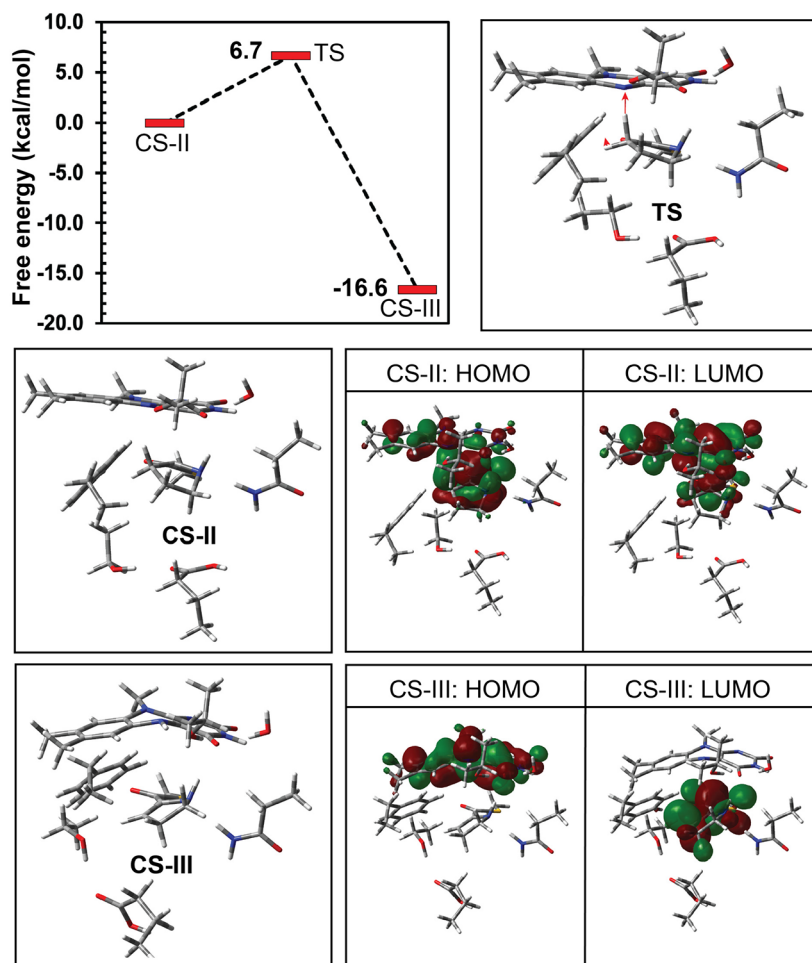


Figure 6. Transition state of hydride transfer. (Top left) Energy barrier for the plausible hydride transfer mechanism from the CS-II intermediate state. (Top right) Geometries of the TS between CS-II and CS-III. Optimized geometries and frontier molecular orbitals of the CS-II (middle panel) and CS-III (bottom panel) states.

information about the energy required for the hydride transfer event. To obtain additional insight into the energy required for hydride transfer, we performed transition state calculations using a cluster model. We discovered that the CS-II and CS-III intermediates are separated by 16.6 kcal/mol, with CS-III being lower in energy. Additionally, we observed that while the FAD isoalloxazine ring is planar in the CS-II intermediate, upon hydride transfer, the isoalloxazine ring bends out of plane around the short axis of the ring, leading to a reduction in the level of conjugation in the CS-III intermediate (Figure 6). The bending of the isoalloxazine ring has been observed in prior studies³⁵ when FAD has been reduced to FADH[−].

Furthermore, we observed that the HOMO and LUMO in the CS-II state are located on the FAD and the substrate (Figure 6), with a slightly larger contribution from the substrate. The overlap of the FAD and the substrate orbitals originates from the π – π^* overlap between them. However, the loss of planarity of the isoalloxazine ring in the CS-III intermediate leads to the abrogation of the π – π^* overlap. Hence, the HOMO and LUMO of the CS-III state are located only on FAD and the substrate, respectively. The localization of the HOMO on FAD in the CS-III intermediate confirms that the negative charge of the hydride is located on the FAD molecule in the CS-III intermediate.

The transition state (TS) is located 6.7 kcal/mol above the CS-II intermediate state. The TS is associated with a single imaginary frequency centered on the C3-H atoms, which captures the transfer of a hydride from C3 to N5 as indicated by the red arrow on the TS (Figure 6). Therefore, we conclude that the free energy barrier for the hydride transfer is 6.7 kcal/mol followed by structural rearrangements, which include slight bending of the isoalloxazine ring. In line with findings mentioned above, here, we additionally detect that the direct hydride transfer from N1 is geometrically constrained, and a few structural rearrangements must occur before the transfer. For example, the substrate must rotate around the C–C2 bond (see Figure 1 and Figure S8 for definitions) to come closer to flavin N5. A similar rotation around the C–C2 bond was observed in the CS-I state (Figure 3c,d). In contrast to the CS-I state, the C–C2 bond is a double bond in the CS-II state. Therefore, the rotational energy barrier is expected to be much higher in this case. We computed a potential energy surface (PES) along the O–C–C2–N1 dihedral angle (Figure S9) in a CPCM solvent model using a low dielectric constant ($\epsilon = 6.2528$) to mimic the protein environment. The geometry of the molecule was optimized along the dihedral angle at 10° intervals starting from the optimized CS-II state geometry of the substrate. Consistent with other calculations using the cluster model, the relaxed PES scan was performed at the CAM-B3LYP/6-31+G(d) level of theory. We found that the aforementioned rotation has a 33.6 kcal/mol energy barrier, which is much higher than the free energy barrier for FAD reduction by transfer of a hydride from the C3 atom. Although a complete rotation may not be required for C2–N1 oxidation, even an $\sim 70^\circ$ rotation is ~ 18 kcal/mol higher in energy than the initial state (Figure S9). To estimate the free energy of activation for this rotation, we also performed a TS calculation at the same level of theory. The TS is confirmed by a single imaginary frequency, which represents the rotation around the dihedral angle. We observed that the free energy barrier for this rotation is 33.5 kcal/mol while the two end states differ by only 0.5 kcal/mol. The free energy barrier of the rotation is ~ 5 -fold higher than that of the hydride transfer from C3, which does not require additional structural rearrangements. Therefore, we conclude that the C2–C3 oxidation of the substrate is more likely to occur than C2–N1 oxidation during the transfer of a hydride to the FAD.

SUMMARY AND CONCLUSIONS

We have simulated the initial state in the pyrrole biosynthesis in which the substrate protein is bound to an active and an inactive Bmp3 enzyme. We demonstrated that the tetrameric Bmp1/Bmp3 complex is not functionally important for the hydride transfer step (CS-II to CS-III). However, the tetramer may be required for subsequent steps in the catalytic cycle. One Bmp1/Bmp3 complex can perform the chemical reaction in the presence of a second inactive Bmp3. Furthermore, we demonstrated that the active site of Bmp3 can stabilize the FAD and proline substrate. We showed that Glu242 can initiate the biosynthesis process by capturing a proton from the substrate using MD and QM/MM simulations. Additionally, our results show a significant population of the substrate conformations that favors proton uptake by Glu242. Our simulations show that the C3 atom remains closer to the FAD hydride acceptor site than the N1 atom, and one QM/MM simulation captured the transfer of a hydride from C3 to FAD within 12 ps without any applied bias. We also determined that

transfer of a hydride to FAD is coupled with C2–C3 oxidation of the substrate and must overcome a very small (6.7 kcal/mol) energy barrier. In contrast, the alternative mechanism, via C2–N1 oxidation, requires rotation around the C=C bond, which has a 33.5 kcal/mol activation free energy barrier. Our findings provide mechanistic insights for a reaction that is well represented in biological chemistry but has evaded experimental characterization due to instability of the substrate and intermediate species.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biochem.1c00741>.

Details of modeling steps and additional computational and experimental details (PDF)

Details of the system (MOV)

Hydride transfer movie (MOV)

Accession Codes

Bmp3, UniProtKB entry F2K077; Bmp1, UniProtKB entry F2K074; Bmp4, UniProtKB entry F2K078.

AUTHOR INFORMATION

Corresponding Authors

Vinayak Agarwal – School of Chemistry and Biochemistry and School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; orcid.org/0000-0002-2517-589X; Email: vagarwal@gatech.edu

James C. Gumbart – School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; orcid.org/0000-0002-1510-7842; Email: gumbart@physics.gatech.edu

Authors

Atanu Acharya – School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; orcid.org/0000-0002-6960-7789

Dongqi Yi – School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

Anna Pavlova – School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.biochem.1c00741>

Funding

This work was supported by the National Science Foundation (Grant CHE-2004030 to V.A. and J.C.G.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Computational resources were provided by the Extreme Science and Engineering Discovery Environment under Grant TG-MCB130173 and the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology. This work also used the Hive cluster, which is supported by National Science Foundation Grant 1828187 and is managed by PACE.

REFERENCES

- (1) Hu, D. X.; Withall, D. M.; Challis, G. L.; Thomson, R. J. Structure, chemical synthesis, and biosynthesis of prodiginine natural products. *Chem. Rev.* **2016**, *116*, 7818–7853.
- (2) Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
- (3) Jaremko, M. J.; Davis, T. D.; Corpuz, J. C.; Burkart, M. D. Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein–protein interactions. *Nat. Prod. Rep.* **2020**, *37*, 355–379.
- (4) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. Conversion of L-proline to pyrrolyl-2-carboxyl-S-PCP during undecylprodigiosin and pyoluteorin biosynthesis. *Chem. Biol.* **2002**, *9*, 171–184.
- (5) Corpuz, J. C.; Podust, L. M.; Davis, T. D.; Jaremko, M. J.; Burkart, M. D. Dynamic visualization of type II peptidyl carrier protein recognition in pyoluteorin biosynthesis. *RSC Chem. Biol.* **2020**, *1*, 8–12.
- (6) Garneau-Tsodikova, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. Protein assembly line components in prodigiosin biosynthesis: characterization of PigA, G, H, I, J. *J. Am. Chem. Soc.* **2006**, *128*, 12600–12601.
- (7) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. *Nat. Chem. Biol.* **2014**, *10*, 640–647.
- (8) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Nat. Prod. Rep.* **2014**, *31*, 61–108.
- (9) Garneau, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. Characterization of the formation of the pyrrole moiety during clorobiocin and coumermycin A1 biosynthesis. *Biochemistry* **2005**, *44*, 2770–2780.
- (10) Thapa, H. R.; Robbins, J. M.; Moore, B. S.; Agarwal, V. Insights into thiotemplated pyrrole biosynthesis gained from the crystal structure of flavin-dependent oxidase in complex with carrier protein. *Biochemistry* **2019**, *58*, 918–929.
- (11) Yi, D.; Acharya, A.; Gumbart, J. C.; Gutekunst, W. R.; Agarwal, V. Gatekeeping Ketosynthases Dictate Initiation of Assembly Line Biosynthesis of Pyrrolic Polyketides. *J. Am. Chem. Soc.* **2021**, *143*, 7617–7622.
- (12) Höfer, I.; Crüsemann, M.; Radzom, M.; Geers, B.; Flachshaar, D.; Cai, X.; Zeeck, A.; Piel, J. Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chem. & Biol.* **2011**, *18*, 381–391.
- (13) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. Dichlorination of a pyrrolyl-S-carrier protein by FADH₂-dependent halogenase PltA during pyoluteorin biosynthesis. *Proc. Nat. Acad. Sci.* **2005**, *102*, 13843–13848.
- (14) Westrich, L.; Heide, L.; Li, S.-M. CloN6, a novel methyltransferase catalysing the methylation of the pyrrole-2-carboxyl moiety of clorobiocin. *ChemBioChem.* **2003**, *4*, 768–773.
- (15) Kim, J.-J. P.; Miura, R. Acyl-CoA dehydrogenases and acyl-CoA oxidases: Structural basis for mechanistic similarities and differences. *Eur. J. Biochem.* **2004**, *271*, 483–493.
- (16) Ghisla, S.; Thorpe, C. Acyl-CoA dehydrogenases. *Eur. J. Biochem.* **2004**, *271*, 494–508.
- (17) Mann, S.; Lombard, B.; Loew, D.; Méjean, A.; Ploux, O. Insights into the reaction mechanism of the prolyl–acyl carrier protein oxidase involved in anatoxin-a and homoanatoxin-a biosynthesis. *Biochemistry* **2011**, *50*, 7184–7197.
- (18) Lee, C.-C.; Ko, T.-P.; Chen, C.-T.; Chan, Y.-T.; Lo, S.-Y.; Chang, J.-Y.; Chen, Y.-W.; Chung, T.-F.; Hsieh, H.-J.; Hsiao, C.-D.; Wang, A. H.-J. Crystal Structure of PigA: A Prolyl Thioester-Oxidizing Enzyme in Prodigiosin Biosynthesis. *ChemBioChem.* **2019**, *20*, 193–202.
- (19) Walsh, C. T.; Wenciewicz, T. A. Flavoenzymes: versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* **2013**, *30*, 175–200.
- (20) Walsh, C. T.; Garneau-Tsodikova, S.; Howard-Jones, A. R. Biological formation of pyrroles: nature's logic and enzymatic machinery. *Nat. Prod. Rep.* **2006**, *23*, 517–531.
- (21) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303.
- (22) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (24) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71.
- (25) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
- (26) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid parameterization of small molecules using the force field toolkit. *J. Comput. Chem.* **2013**, *34*, 2757–2770.
- (27) Pavlova, A.; Gumbart, J. C. Parametrization of macrolide antibiotics using the force field toolkit. *J. Comput. Chem.* **2015**, *36*, 2052–2063.
- (28) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (29) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1327.
- (30) Willistein, M.; Bechtel, D. F.; Müller, C. S.; Demmer, U.; Heimann, L.; Kayastha, K.; Schünemann, V.; Pierik, A. J.; Ullmann, G. M.; Ermler, U.; Boll, M. Low potential enzymatic hydride transfer via highly cooperative and inversely functionalized flavin cofactors. *Nat. Commun.* **2019**, *10*, 2074.
- (31) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (32) Melo, M. C.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D.; Rocha, G. B.; Ribeiro, J. V.; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z. NAMD goes quantum: an integrative suite for hybrid simulations. *Nat. Methods* **2018**, *15*, 351.
- (33) Mancini-Samuelson, G. J.; Kieweg, V.; Sabaj, K. M.; Ghisla, S.; Stankovich, M. T. Redox properties of human medium-chain acyl-CoA dehydrogenase, modulation by charged active-site amino acid residues. *Biochemistry* **1998**, *37*, 14605–14612.
- (34) Frisch, M. J.; et al. *Gaussian 16*, rev. B.01; Gaussian, Inc.: Wallingford, CT, 2016.
- (35) Kabir, M. P.; Orozco-Gonzalez, Y.; Gozem, S. Electronic spectra of flavin in different redox and protonation states: a computational perspective on the effect of the electrostatic environment. *Phys. Chem. Chem. Phys.* **2019**, *21*, 16526–16537.
- (36) Acharya, A.; Chaudhuri, S.; Batista, V. S. Can TDDFT Describe Excited Electronic States of Naphthol Photoacids? A Closer Look with EOM-CCSD. *J. Chem. Theory Comput.* **2018**, *14*, 867–876.
- (37) Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the dielectric constant of proteins: smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *J. Chem. Theory Comput.* **2013**, *9*, 2126–2136.
- (38) Amin, M.; Küpper, J. Variations in Proteins Dielectric Constants. *ChemistryOpen* **2020**, *9*, 691.
- (39) Schutz, C. N.; Warshel, A. What are the dielectric constants of proteins and how to validate electrostatic models? *Proteins: Struct., Funct., Bioinf.* **2001**, *44*, 400–417.

- (40) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681.
- (41) Moncoq, K.; Regad, L.; Mann, S.; Méjean, A.; Ploux, O. Structure of the prolyl-acyl carrier protein oxidase involved in the biosynthesis of the cyanotoxin anatoxin-a. *Acta Crystallogr.* **2013**, *69*, 2340–2352.
- (42) Knapp, B.; Ospina, L.; Deane, C. M. Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. *J. Chem. Theory Comput.* **2018**, *14*, 6127–6138.
- (43) Fraaije, M. W.; Mattevi, A. Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem. Sci.* **2000**, *25*, 126–132.
- (44) Luesch, H.; Hoffmann, D.; Hevel, J. M.; Becker, J. E.; Golakoti, T.; Moore, R. E. Biosynthesis of 4-Methylproline in Cyanobacteria: cloning of nosE and nosF genes and biochemical characterization of the encoded dehydrogenase and reductase activities. *J. Org. Chem.* **2003**, *68*, 83–91.
- (45) Lukat, P.; Katsuyama, Y.; Wenzel, S.; Binz, T.; König, C.; Blankenfeldt, W.; Brönstrup, M.; Müller, R. Biosynthesis of methylproline containing griselimycins, natural products with anti-tuberculosis activity. *Chem. Sci.* **2017**, *8*, 7521–7527.
- (46) Zwick, C. R., III; Renata, H. Remote C–H hydroxylation by an α -ketoglutarate-dependent dioxygenase enables efficient chemo-enzymatic synthesis of manzacidin C and proline analogs. *J. Am. Chem. Soc.* **2018**, *140*, 1165–1169.
- (47) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* **2006**, *45*, 12756–12766.

Recommended by ACS

l-Threonine Transaldolase Activity Is Enabled by a Persistent Catalytic Intermediate

Prasanth Kumar, Andrew R. Buller, *et al.*

DECEMBER 18, 2020
ACS CHEMICAL BIOLOGY

READ 

Mutation of β Gln114 to Ala Alters the Stabilities of Allosteric States in Tryptophan Synthase Catalysis

Rittik K. Ghosh, Michael F. Dunn, *et al.*

OCTOBER 01, 2021
BIOCHEMISTRY

READ 

Cryptic Oxidative Transamination of Hydroxynaphthoquinone in Natural Product Biosynthesis

Tomohiro Noguchi, Tomohisa Kuzuyama, *et al.*

MARCH 16, 2022
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

READ 

Crystal Structures of Wild-Type and F448A Mutant *Citrobacter freundii* Tyrosine Phenol-Lyase Complexed with a Substrate and Inhibitors: Implications for the...

Robert S. Phillips and Steven Craig

SEPTEMBER 27, 2018
BIOCHEMISTRY

READ 

Get More Suggestions >

Supporting Information:

Resolving the hydride transfer pathway in oxidative conversion of proline to pyrrole

Atanu Acharya,[†] Dongqi Yi,[‡] Anna Pavlova,[†] Vinayak Agarwal,^{*,‡,¶} and James
C. Gumbart^{*,†,‡}

[†]*School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, United States*

[‡]*School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA
30332, United States*

[¶]*School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, United
States*

E-mail: vagarwal@gatech.edu; gumbart@physics.gatech.edu

Hydrogen bond analysis

The isoalloxazine ring forms hydrogen bonds with backbone and side chains of Thr158 (Figure S1a), Ser126 backbone and side chain (Figure S1b), and Asn123 backbone (Figure S1c). The Ser132, Tyr156 and Lys201 residues interact with the negatively charged phosphate and the adenine moieties of the FAD (Figure S1c,d). The isoalloxazine ring participate in 3-5 hydrogen bonds with Bmp3 residues (Figure S1e). Overall, the FAD molecule participates in 5-8 hydrogen bonds with Bmp3 residues (Figure S3a).

Furthermore, the FAD isoalloxazine C² carbonyl oxygen atom (see Figure 1 for the numbering scheme) is hydrogen bonded to a crystallographic water molecule. This water molecule mediates the hydrogen bonds between the FAD and Asn123 backbone, which is confirmed by similar occupancies of FAD-water and Asn123-water hydrogen bonds (18% and 22%). The water molecule also forms hydrogen bonds with the Thr158 backbone and Glu242 side chain. The FAD and Thr158 hydrogen bond was shown modulate the local environment, enabling catalytic proton transfer to glutamate.^{S1} In addition, the FAD isoalloxazine ring can form direct hydrogen bonds with the Asn123 backbone with 88% occupancy. The Asn123Ala mutant demonstrated reduction in catalytic activity (~20%) of Bmp3, indicating the importance of Asn123 residue in constructing the active site.^{S2} Furthermore, this water molecule does not remain near the FAD isoalloxazine ring in two 500-ns simulations. Rather, it forms hydrogen bonds with Thr158 and Glu242; both of these residues are located at the catalytic site. The percentage of all hydrogen bonding interactions between FAD, Bmp3, and water is provided in Table S1. Interestingly, the crystallographic water molecule is never replaced by any other water when it hops in the catalytic site between FAD, Asn123, Glu242, and Thr158. Overall, the presence of the water molecule may not be required for FAD binding, although it could be important for fine-tuning local pKa values for efficient catalysis. Experimental studies emphasized the role of the bound FAD in reducing the pKa of the C_α of thioesters^{S1,S3} enabling the deprotonation of the pyrrolidine C² atom.

Root mean square fluctuations (RMSF) analysis of Bmp1(CP) shows small fluctuations

Table S1: Relative populations of relevant hydrogen bonds. As substrate we included the pyrrolidine ring and the attached carbonyl group in our calculation. The Ppant moiety attaches the substrate to the Bmp1 carrier protein at the Ser35 residue. The distance and angle cut-off values of 3.5 Å and 30°, respectively, were used for the hydrogen bond analysis. Data from two independent 500-ns MD simulations of the CS-I state were used for this analysis.

FAD-Bmp3 hydrogen bonds			
Donor/Acceptor	Donor/Acceptor	Chain (Main/Side)	Population (%)
FAD	Thr158	Main	82
	Ile125	Main	74
	Ser126	Main	90
	Asn123	Main	88
	Thr158	Side	51
	Tyr156	Side	21
	Ser126	Side	82
	Ser132	Main	15
	Ser132	Side	22
	Leu361	Main	10
	Lys201	Side	17
Water-Bmp3 hydrogen bonds			
Donor/Acceptor	Donor/Acceptor	Chain (Main/Side)	Population
Water	Asn123	Main	22
	Thr158	Main	20
	Glu242	Side	11
Substrate-Bmp3 hydrogen bonds			
Donor/Acceptor	Donor/Acceptor	Chain (Main/Side)	Population
Substrate	Asn123	Side	9
	Ser363	Main	15
Ppant-Bmp3 hydrogen bonds			
Donor/Acceptor	Donor/Acceptor	Chain (Main/Side)	Population
Ppant	Ser132	Main	54
	Ser132	Side	10
	Ser135	Main	38
	Ser135	Side	42
	Ile134	Main	22
	Tyr178	Side	15
	Arg233	Side	14
	Arg243	Side	45

except at the N and C termini (Figure S4A). Moreover, the RMSF values near the substrate, e.g., residues 25 to 44 for Bmp1(CP), are even lower in both simulations, which indicates the stability of the Bmp1(CP)-Bmp3 binding interface in the model system. The RMSF values of Bmp3 residues of chains B and C are also similar for two simulations (Figure S4b). However, chain C of Bmp3 (catalytically deficient apo form) shows larger fluctuations for residues 129 to 135 compared to chain B (holo) of Bmp3 (Figure S4b). The Bmp3 residues 129 to 135 form a loop (Figure S1d), which is located near the FAD binding sites of each chain. Since chain C is lacking the FAD molecule in our model, a larger fluctuation is observed in chain C. However, this loop of chain C is far from the catalytic site we modeled in chain B, making it unlikely that such fluctuations will significantly impact it.

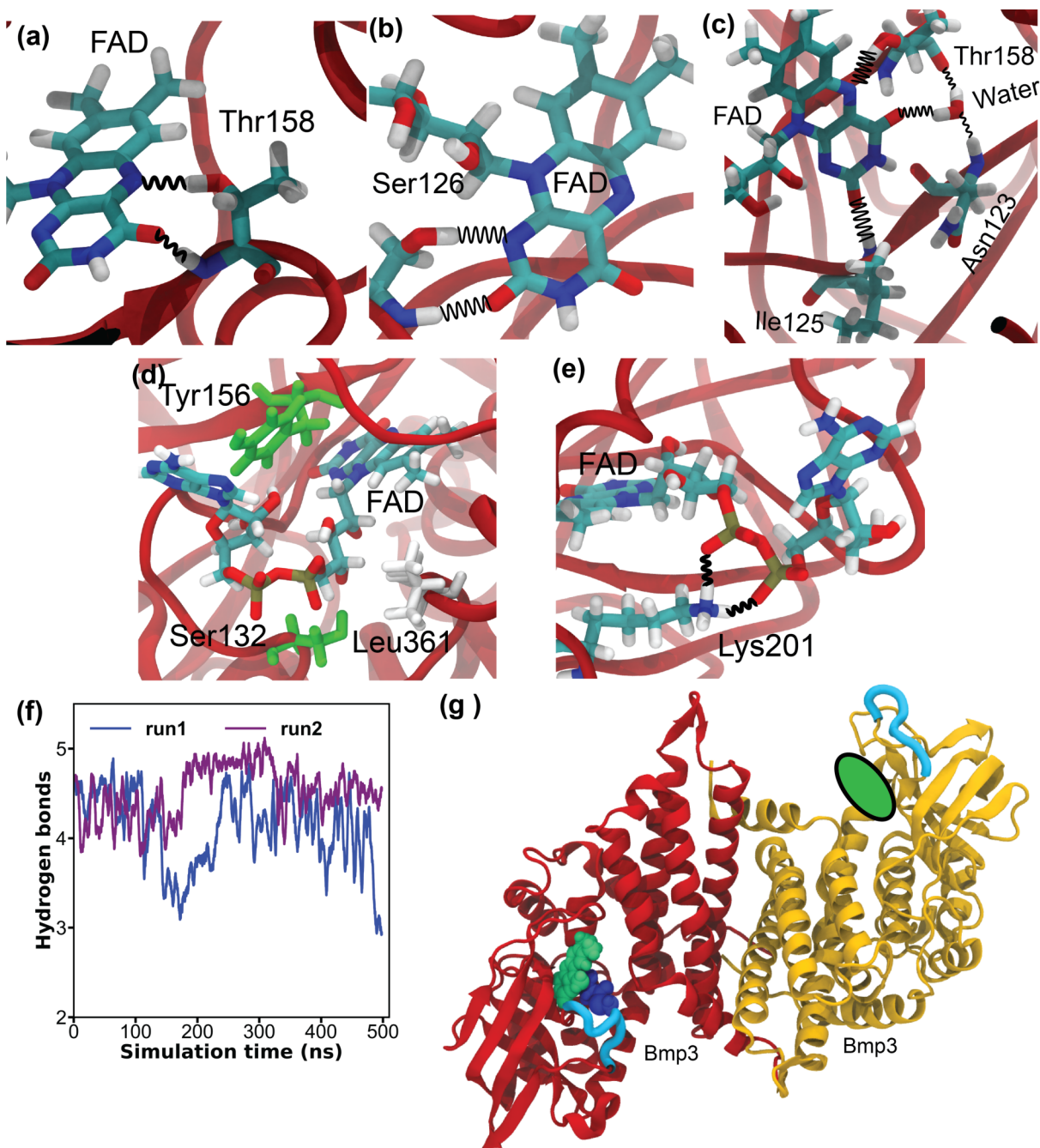


Figure S1: Hydrogen bonds between FAD isoalloxazine ring and (a-c) nearby amino acids and (d, e) adenine and phosphate moieties. (f) Number of hydrogen bonds formed between FAD isoalloxazine ring and Bmp3. The data is presented as moving average of consecutive 200 snapshots for clarity. (g) Bmp3 dimer used in this study. Only the chain B (red) contains substrate and FAD. The Bmp3 loop that assists in the FAD and substrate binding is shown in cyan. Since the yellow Bmp3 does not contain FAD, the cyan loop fluctuates more compared to the same loop in the red Bmp3. The missing FAD's location in the yellow Bmp3 is highlighted as a green oval.

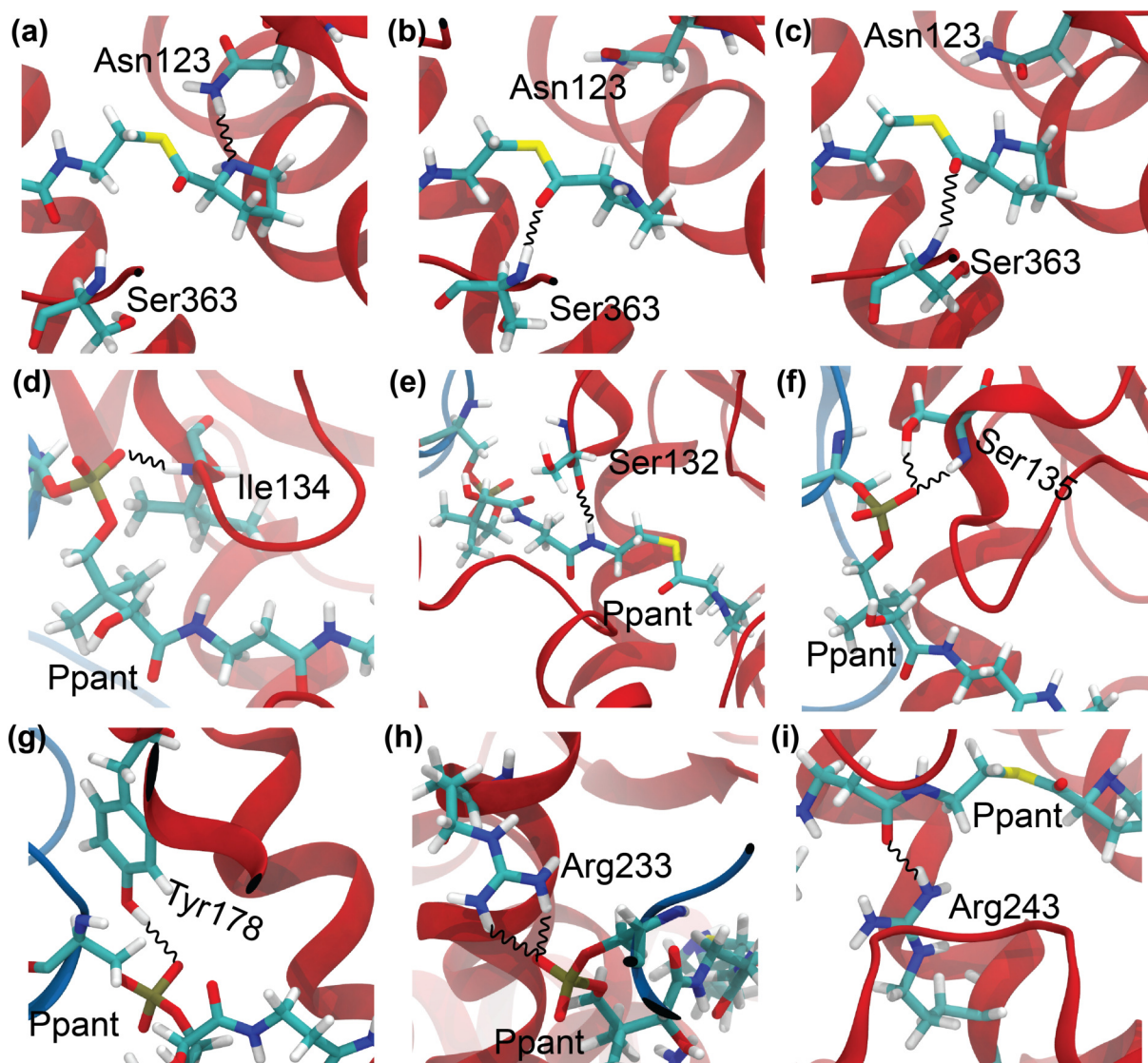


Figure S2: Relevant hydrogen bonds/salt bridges between the substrate and nearby amino acids.

CS-I MD

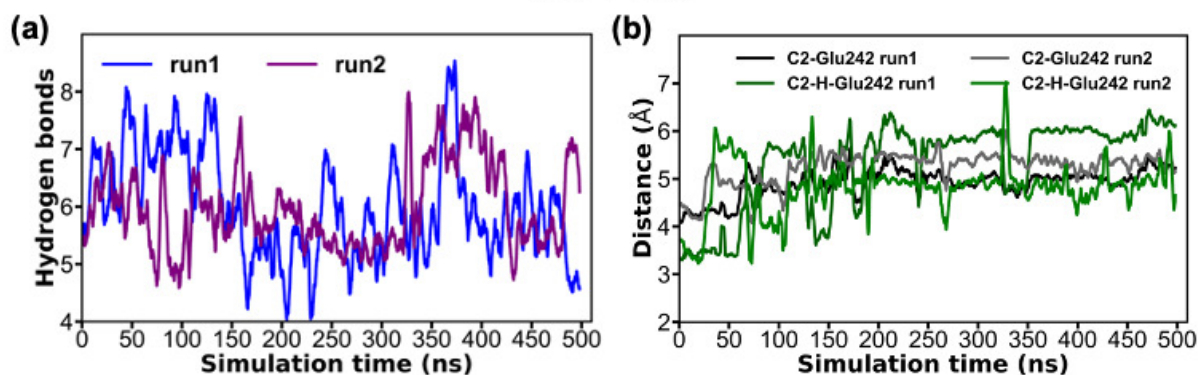


Figure S3: (a) Number of hydrogen bonds between FAD and Bmp3. Overall, the entire FAD molecule participates in 5-8 hydrogen bonds with Bmp3 residues. (b) Distance of the carboxylate oxygen atoms of Glu242 from the C2 or the attached hydrogen atom (C2-H). Data from two simulations were pooled for this analysis.

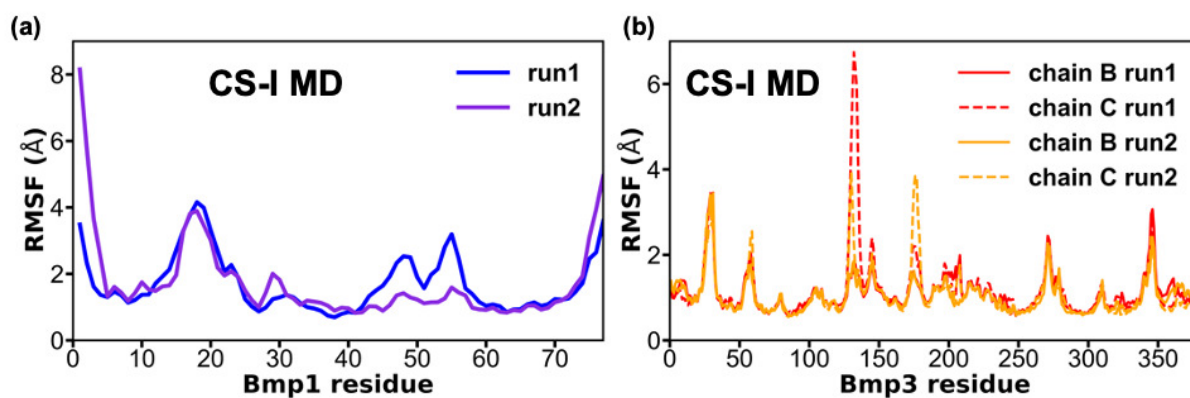


Figure S4: RMSF value of Bmp1 (a) and Bmp3 (b) from two independent MD simulations. Only chain B of Bmp3 includes a bound FAD molecule.

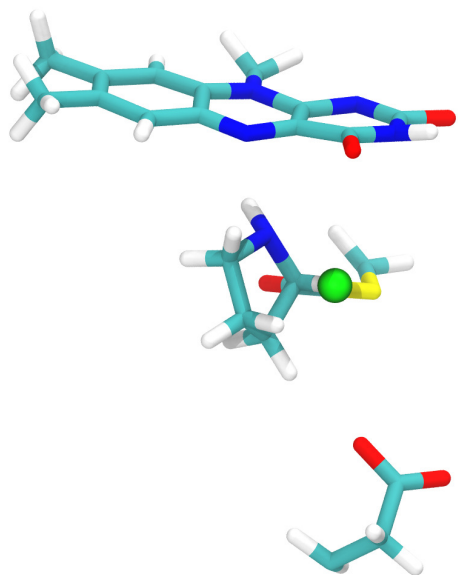


Figure S5: Representative conformation of CS-I with $\Delta d \approx 0.5 \text{ \AA}$.

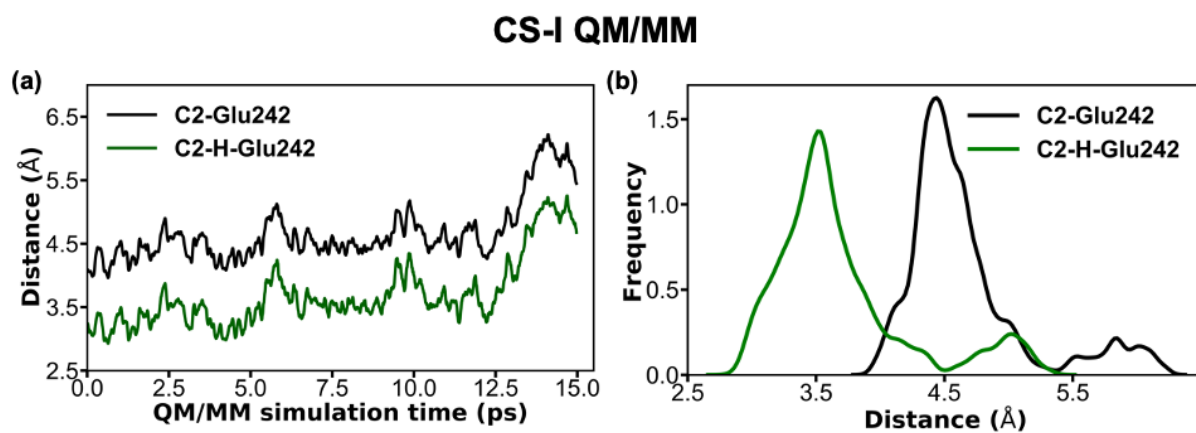


Figure S6: (a) Distance of the carboxylate oxygen atoms of Glu242 from the C2 and the attached hydrogen atom (C2-H) obtained from the QM/MM simulation of CS-I. (b) Relative populations of C2-Glu242 and C2-H-Glu242 distances are shown. Data from two simulations were pooled for this analysis.

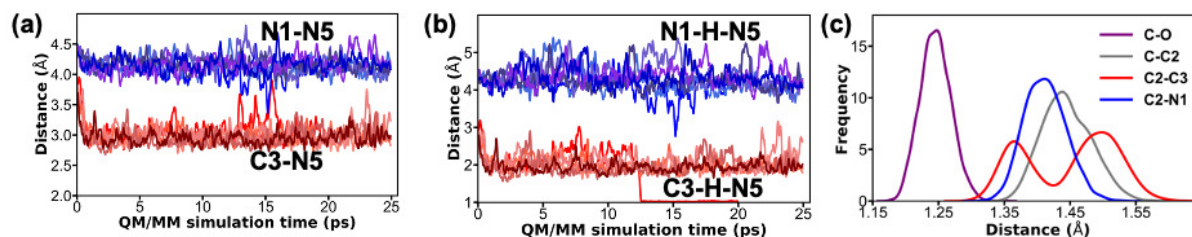


Figure S7: QM/MM simulation of the CS-II state. (a) C3–N5 (red hue) and N1–N5 (blue hue) distances at the CS-II intermediate state. (b) C3–H–N5 (red hue) and N1–H–N5 (blue hue) distances at the same state. (c) Analysis of key bond distances along the QM/MM simulation, where C2–C3 oxidation is explicitly observed.

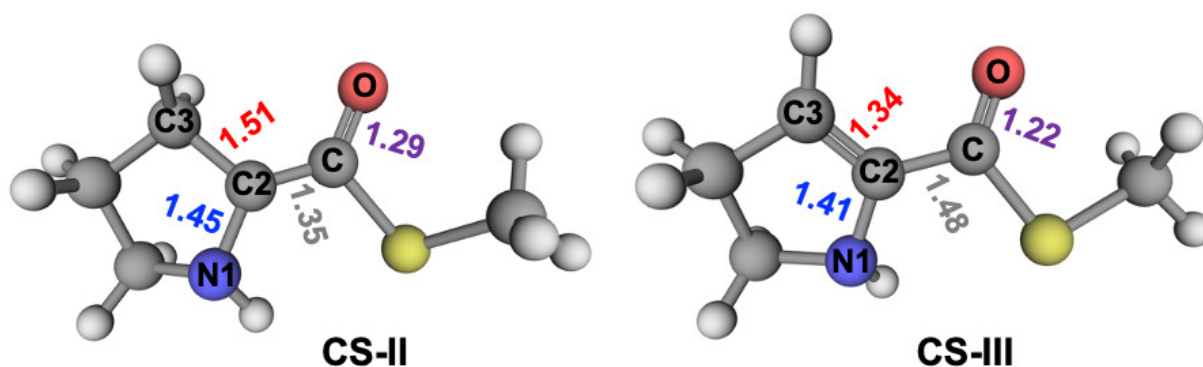


Figure S8: Optimized structures of the chemically active part of the substrate in the CS-II and CS-III states. Key geometric parameters are annotated.

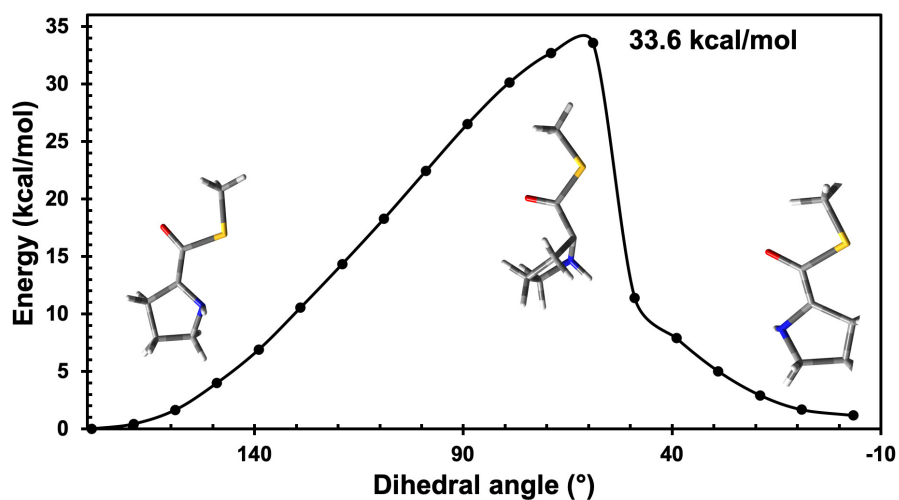


Figure S9: Potential energy curve along the O–C–C2–N1 dihedral angles in the CS-II intermediate state. In the cluster model, the value of this angle is 179°.

Optimization of force field parameters

Guassian16^{S4} was used for all quantum mechanical (QM) calculations and the Force Field Toolkit (ffTK)^{S5} was used for all parameter and partial atomic charge optimization. MP2/6-31G* level of theory was used for all geometry optimizations, Hessian calculations, and dihedral scans, as recommended for CHARMM-compatible parameters.^{S6} The initial parameters for FAD were obtained from the CGenFF website (<https://cgenff.umaryland.edu/>), which assigns parameters and partial atomic charges by analogy and provides a penalty to each.^{S7,S8} Only the charges and parameters with penalties higher than 10 were re-optimized with ffTK. Based on the locations of the atoms and bonded terms with high penalties, we constructed two reduced models of FAD for optimization (Figure S10). Model I was used for the optimization of the force-field parameters and the partial atomic charges within the isoalloxazine ring in FAD. These charges were optimized by fitting them to the minimized interactions energies and distances of a water molecules at HF/6-31G* level of theory, in accordance with CGenFF guidelines.^{S6} The QM energies were scaled by 1.16, as recommended for neutral compounds.^{S6} Good agreement with the target interaction energies (within 0.4 kcal/mol) was achieved (Table S2). For two of the atoms (N1 and N10), an interacting water molecule could not be placed without sterical clashes, and natural population analysis (NPA) was used to set the partial charges of these atoms instead (Figure S10A).^{S9-S11} Based on the results from NPA, the charges of N1 and N10 were set to -0.727 and -0.457, respectively.

All of the bond and angle parameters in need of optimization were located on the isoalloxazine ring, making model I sufficient for their optimization. The optimization was performed based on the QM Hessian calculation in internal coordinates, as implemented in ffTK.^{S5} Model I was also used to also optimize the dihedral force-field parameters within the isoalloxazine ring. Relaxed QM dihedral scans ranging from -60 to 60° at intervals of 10° were used to generate the target PES. The optimization of dihedral force-field constants was done using the simulated annealing protocol by Guvench et al., which is implemented in ffTK.^{S5,S12} Only structures with energies <10 kcal/mol in comparison to the optimized

structure were included in the fit.^{S12} Model II was used to optimize the dihedral parameters involved in connecting the isoalloxazine ring to the nucleotide. Good agreement with the target PES was achieved for both model I and model II (Figure S11).

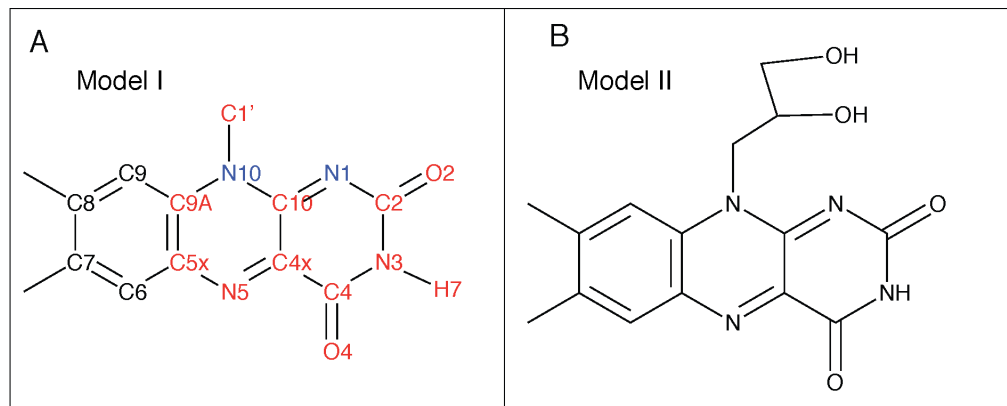


Figure S10: (a) Model I that was used for parametrization of FAD. The charges of the atoms in red were optimized using the standard CGenFF protocol by fitting to QM interactions with a water molecule. For the atoms in blue the charges were taken from NPA. All other charges were taken from the CGenFF website. (b) Model II that was used for parametrization of dihedrals connecting the isoalloxazine ring to the FAD molecule.

Table S2: Atom names and the corresponding QM and MM interaction energies with a water molecule for the atoms with fitted charges in FAD. Error in interaction energies and the fitted charges are also displayed. All energies are in kcal/mol, and QM energies are scaled by 1.16.

Atom name	QM energy (scaled)	MM energy	Error	Charge
O2	-6.36	-6.20	0.16	-0.510
O4	-5.55	-5.94	-0.39	0.461
C4	-1.71	-1.36	0.35	0.239
C4X	-2.75	-2.81	-0.06	0.597
C5X	-1.10	-1.09	0.01	0.043
C9A	-1.66	-1.67	-0.01	0.138
C10	-3.14	-2.95	0.19	0.538
H7	-6.23	-6.37	-0.14	0.312
C1'	-3.01	-2.98	0.031	0.085
N3	-0.74	-0.45	0.285	-0.365
N5	-1.62	-1.51	0.11	-0.642

For the substrate we constructed a model for which the initial parameters were also obtained from the CGenFF website (Figure S12A). The CHARMM36 force field has been carefully developed for the five- and six-membered carbohydrate rings with the aim to repro-

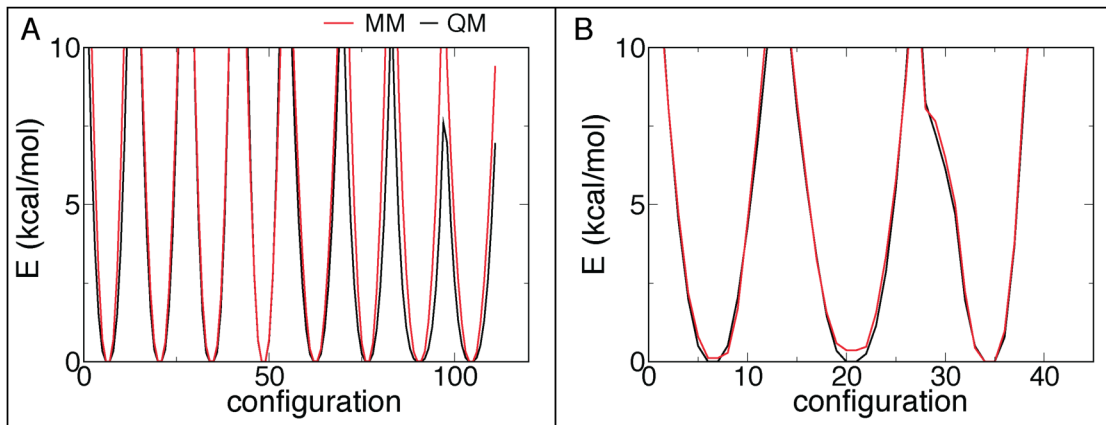


Figure S11: Comparison of QM (black) and MM-optimized (red) energies of dihedral scans performed in order to optimize the dihedral parameters. (a) Dihedral scans for model I. (b) Dihedral scans for model II.

duce both the QM target data, and the preferred ring puckering.^{S13,S14} Therefore, we chose to keep the bonded parameters for the five-membered ring of the substrate, despite penalties higher than 10 in some cases. The partial atomic charges with penalties higher than 10 and the dihedrals around the C2-C1 bond were re-optimized (Figure S12). We used the same approaches as with FAD, except the dihedral scan around the C2-C1 bond was performed in the range of -90 to 90° at intervals of 15° . The energies of interactions with water molecules were reproduced within 0.4 kcal/mol (Table S3) and good agreement between the QM and MM dihedral scans was achieved (Figure S12).

Table S3: Atom names and the corresponding QM and MM interaction energies with a water molecule for the atoms with fitted charges in the substrate. Errors in interaction energies, and the fitted charges are also displayed. All energies are in kcal/mol, and QM energies are scaled by 1.16.

Atom name	QM energy (scaled)	MM energy	Error	Charge
C0	-1.095	-0.491	0.604	-0.241
S	-2.695	-2.812	-0.117	-0.118
C1	-2.711	-2.836	-0.125	0.700
O1	-5.053	-5.126	-0.073	-0.516
C2	-1.999	-1.585	0.414	0.081
N1	-6.249	-6.276	-0.027	-0.787
H1	-3.798	-3.716	0.082	0.300

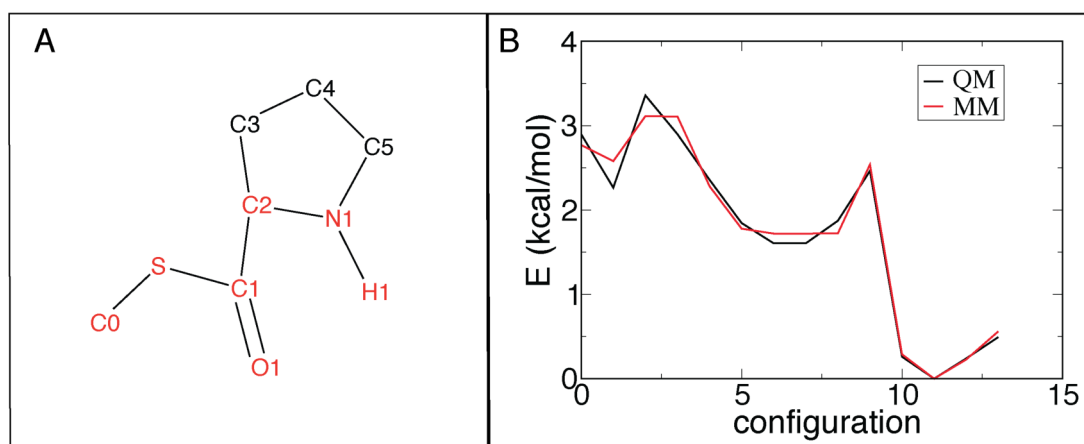


Figure S12: (a) Model of the substrate used for charge optimization and dihedral scans. The names of atoms with optimized charges are shown in red. (b) Comparison of QM (black) and MM (red) energies of the dihedral scan around C2-C1 bond.

Construction of expression plasmids

The construction of pET24-Sfp, pET28-Bmp4 (UNP F2K078) and pCDFDuet-Bmp3/Bmp1(CP) (UNP F2K077 and F2K074, respectively) vectors has been described previously.^{S2,S15} GriE, GriF, ProC, LdoA, NosE, and NosF gene fragments were synthesized by Twist Bioscience, and amplified by Phusion DNA polymerase. ProC fragment was cloned into pET28b(+) vector, while GriE, GriF, LdoA, and NosF was cloned into a previously described pET28MBP vector1 using NEBuilder HiFi DNA Assembly master mix. NosE was cloned into a pET24MBP vector which carries a N-terminal MBP and a C-terminal His6 tag. All vectors constructed were confirmed by Sanger sequencing and transformed into *Escherichia coli* BL21Gold(DE3) for expression of desired proteins.

Protein expression and purification

Purification of Sfp, Bmp4, Bmp3/apo-Bmp1(CP) followed the same protocol as described before.^{S2,S15} For the overexpression of other proteins, *E. coli* strains carrying expression plasmids were grown in 1 L terrific broth medium supplemented with appropriate antibiotics at 30 °C until OD₆₀₀ reached around 0.4. The culture temperature was reduced to 18 °C. When

OD₆₀₀ reached 0.7-0.8, protein expression was induced by the addition of 0.2 mM IPTG. The induced culture was grown at 18 °C for an additional 18 h. All subsequent steps of protein purification were conducted at 4 °C. Cells were harvested by centrifugation, resuspended in binding buffer (20 mM Tris-HCl (pH=8.0), 500 mM NaCl), and lysed by sonication. The lysate was clarified by centrifugation at 18,000 rpm for 45 min and loaded on 5 mL HisTrap HP column using AKTApurifier plus FPLC system. The column was washed extensively with wash buffer containing 20 mM Tris-HCl (pH=8.0), 30 mM imidazole, and 500 mM NaCl. Proteins were eluted with a linear gradient to 100% elution buffer (20 mM Tris-HCl (pH=8.0), 250 mM imidazole, 500 mM NaCl) over 10 column volumes (CVs). Purity of eluent fractions was checked by SDS-PAGE. Fractions containing proteins of interest were pooled. Combined protein solution was either concentrated with Amicon Ultra centrifugal filters with appropriate molecular weight cut-offs, and then desalted into storage buffer (binding buffer supplemented with 10% glycerol) using PD-10 columns or dialyzed in 2 L storage buffer overnight followed by concentration. Purified proteins were stored as small aliquots at -80 °C and fresh aliquots were used each time for enzyme assays.

Purification of amino acids

The general procedure for amino acid purification was similar to what has been previously described.^{S16} Amberchrom 50WX8 resin was slurry-packed with 1 M NH₄OH, washed first with 1 CV of 1 M NH₄OH followed by 2 CVs of H₂O. The resins were then charged with 1 M HCl and washed with H₂O until pH = 6–7 was achieved. The reaction mixtures from enzyme assays were loaded onto the column, washed with H₂O until pH= 6-7, and eluted with 1 M NH₄OH. The presence of amino acids in eluent fractions was checked by Thin Layer Chromatography (TLC) stained with basic aqueous KMnO₄ and heat as developing agents. Amino acids containing fractions were pooled, concentrated, and lyophilized.

Preparation of 4-methylproline

(2S, 4R)-4-methylproline. GriE reaction was first set up to produce (2S, 4R)-5-hydroxyleucine as previously described.^{S16} The reaction composed of 7.5 mM L-leucine, 22.5 mM α -ketoglutarate, 0.5 mM $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$, 7.5 mM L-ascorbate, 50 mM potassium phosphate buffer (pH=8.0) and 10 μM GriE in a total volume of 2 mL. The reaction was shaken overnight at 25 °C, 100 rpm under air. The 5-hydroxyleucine was purified using Amberchrom 50WX8 cation exchange column, and used as substrate for GriF/ProC reaction^{S17} which contained 4 mM (2S, 4R)-5-hydroxyleucine, 1 mM ZnSO_4 , 7.5 mM NAD^+ , 50 mM potassium phosphate buffer (pH=8.0), 10 μM GriF and 10 μM ProC in a total volume of 1.5 mL. The GriF/ProC assay was incubated at 30 °C overnight. The 4-methylproline product was purified with Amberchrom 50WX8 resin.

(2S, 4S)-4-methylproline. (2S, 4S)-5-hydroxyleucine was produced using LdoA assay as previously described.^{S18} LdoA assay was composed of 2 mM L-leucine, 6 mM α -ketoglutarate, 0.5 mM $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$, 2 mM L-ascorbate, 50 mM sodium acetate buffer (pH=5.0) and 20 μM LdoA in a total volume of 2.5 mL. The reaction was shaken overnight at 20 °C, 100 rpm under air. The (2S, 4S)-5-hydroxyleucine was then purified using Amberchrom 50WX8 column and used as substrates for NosE/NosF reaction.^{S19} The composition of NosE/NosF assay was similar to GriF/ProC reaction, except that 10 μM NosE/NosF was used as enzymes instead of GriF/ProC. NosE/NosF reaction was allowed to proceed at 37 °C overnight, and the 4-methylproline product was purified using Amberchrom 50WX8 column.

Derivatization of amino acids

To monitor the production of amino acids in enzyme assays, 50 μL solution that contained amino acids (diluted to less than 1 mM) was added with 50 μL MeCN, followed by the addition of 200 μL sodium borate buffer (200 mM, pH=10.0) and 200 μL Fmoc-Cl (1.25 mM prepared in MeCN). After incubation at room temperature for 20 min, the reaction was quenched with 200 μL 1-adamantanamine hydrochloride (12.5 mM dissolved in 1:1

H₂O/MeCN). After 2 min, the sample was centrifuged for 30 min to remove precipitates and the cleared solution was analyzed on a Luna 5 µm C8(2) 100 Å LC column (250×4.6 mm) on an Agilent 1260 Infinity HPLC system coupled to a Bruker amaZon ion-trap mass spectrometer. Water (solvent A) and MeCN (solvent B) supplemented with 0.1% (v/v) formic acid were used as the mobile phase. A flow rate of 0.5 mL.min⁻¹ was used with the following gradient: 0-3 min: linear gradient to 5% from 3% B, 3-9 min: linear gradient to 35% B, 9-34 min: linear gradient to 80% B, 34-37 min: linear gradient to 100% B, 37-40 min: 100% B, 40-41 min: linear gradient to 5% B, 41-42 min: 5% B, 42-43 min: linear gradient to 100% B, 43-44 min: 100% B, 44-45 min: linear gradient to 5% B.

Bmp1/Bmp3/Bmp4 assay

The Bmp1/Bmp3/Bmp4 assay containing 1 mM CoA-SH, 2.5 mM L-proline or (2S)-4-methylproline, 5 mM ATP, 10 mM MgCl₂, 2 mM TCEP, 30 mM HEPES-Na (pH= 7.9), 1 µM Sfp, 4 µM Bmp4 and 100 µM Bmp3/apo-Bmp1(CP) was conducted in a total volume of 100 µL and the reaction was allowed to proceed at 30 °C for 12 h. The reaction was quenched with equal volume of MeCN, and precipitates were removed by centrifugation. The supernatant was analyzed by Aeris 3.6 µm WIDEPORE XB-C18 LC column (250×4.6 mm) on an Agilent 1290 Infinity II UHPLC system coupled to a Bruker impact II Q-ToF mass spectrometer operating at room temperature. HPLC solvents (A: water and B: MeCN) were supplemented with 0.1% formic acid. A flow rate of 0.5 mL.min⁻¹ was used with the following gradient: 0-5 min: 5% B, 5-30 min: linear gradient to 75% B, 30-31 min: linear gradient to 95% B, 31-35 min: 95% B, 35-36 min: linear gradient to 5% B, 36-38 min: 5% B, 38-39 min: linear gradient to 95% B, 39-42 min: 95% B, 42-43 min: linear gradient to 5% B.

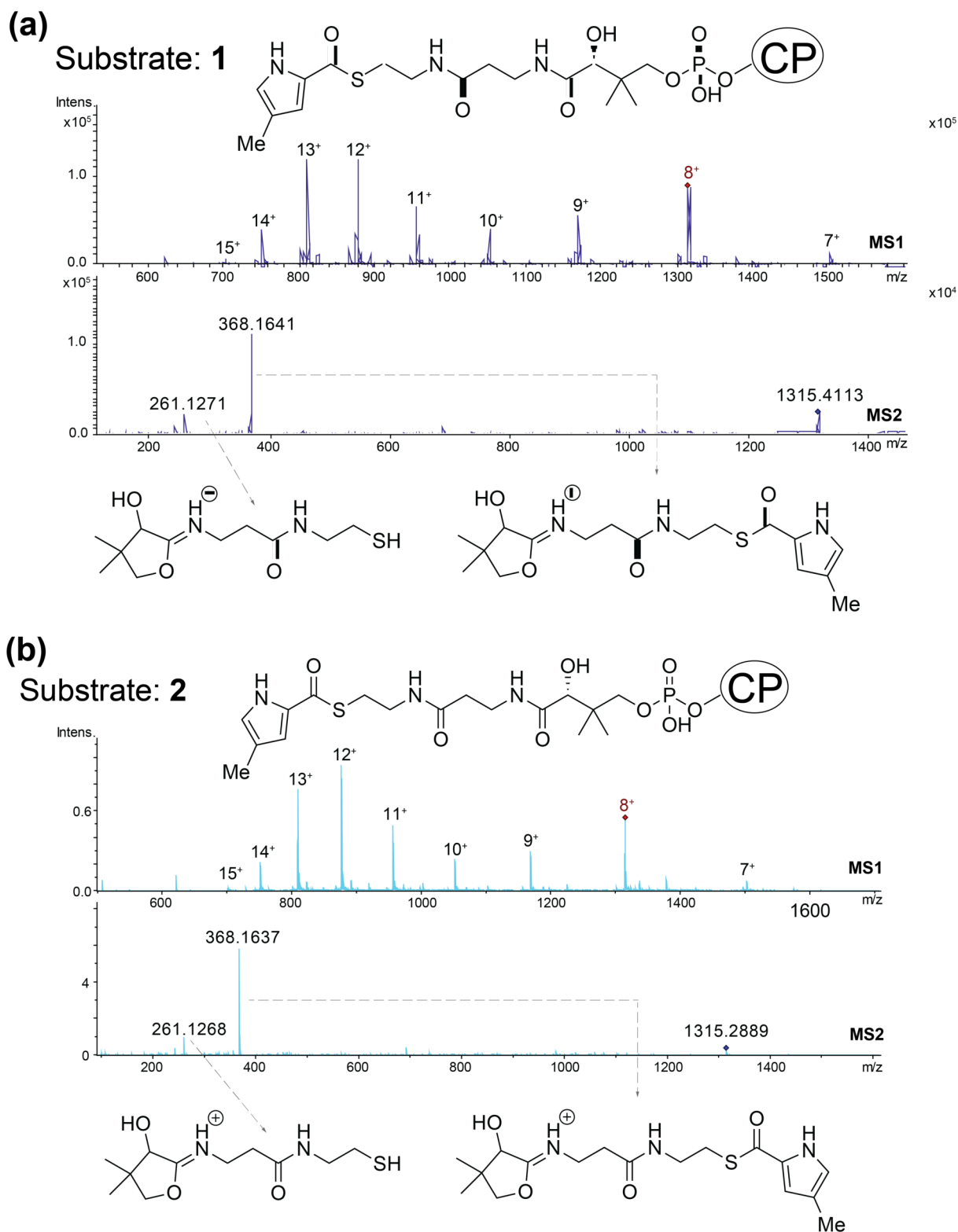


Figure S13: Mass spectrometric confirmation of the conversion of (a) **1** and (b) **2** to 4-methylpyrrolyl-Bmp1(CP) by Bmp4 and Bmp3 enzymes. The top panels illustrate the multiply charged MS1 ions detected for the 4-methylpyrrolyl-Bmp1(CP) product in each reaction, and the bottom panels illustrate the acyl-cyclopentetheine MS2 ejection ions demonstrating the formation of the 4-methylpyrrolyl product.

References

- (S1) Ghisla, S.; Thorpe, C. Acyl-CoA dehydrogenases. *Eur. J. Biochem* **2004**, *271*, 494–508.
- (S2) Thapa, H. R.; Robbins, J. M.; Moore, B. S.; Agarwal, V. Insights into thiotemplated pyrrole biosynthesis gained from the crystal structure of flavin-dependent oxidase in complex with carrier protein. *Biochemistry* **2019**, *58*, 918–929.
- (S3) Mansoorabadi, S. O.; Thibodeaux, C. J.; Liu, H. The Diverse Roles of Flavin Coenzymes Nature’s Most Versatile Thespians. *J. Org. Chem.* **2007**, *72*, 6329–6342.
- (S4) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian~16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.
- (S5) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid parameterization of small molecules using the force field toolkit. *J. Comp. Chem.* **2013**, *34*, 2757–2770.

- (S6) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
- (S7) Vanommeslaeghe, K.; Raman, E. P.; MacKerell Jr., A. D. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.
- (S8) Vanommeslaeghe, K.; MacKerell Jr., A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.
- (S9) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899–926.
- (S10) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural population analysis. *J. Chem. Phys.* **1985**, *83*, 735–746.
- (S11) Xu, Z.; Luo, H. H.; Tieleman, D. P. Modifying the OPLS-AA force field to improve hydration free energies for several amino acid side chains using new atomic charges and an off-plane charge model for aromatic residues. *J. Comput. Chem.* **2007**, *28*, 689–697.
- (S12) Guvench, O.; MacKerell Jr., A. D. Automated conformational energy fitting for force-field development. *J. Mol. Model.* **2008**, *14*, 667–679.
- (S13) Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W., 2nd; Mackerell, A. D., Jr CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J. Chem. Theory Comput.* **2011**, *7*, 3162–3180.

- (S14) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell Jr., A. D. Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543–2564.
- (S15) Agarwal, V.; Diethelm, S.; Ray, L.; Garg, N.; Awakawa, T.; Dorrestein, P. C.; Moore, B. S. Chemoenzymatic synthesis of acyl coenzyme A substrates enables in situ labeling of small molecules and proteins. *Org. Lett.* **2015**, *17*, 4452–4455.
- (S16) Zwick III, C. R.; Renata, H. Remote C–H hydroxylation by an α -ketoglutarate-dependent dioxygenase enables efficient chemoenzymatic synthesis of manzacidin C and proline analogs. *J. Am. Chem. Soc.* **2018**, *140*, 1165–1169.
- (S17) Lukat, P.; Katsuyama, Y.; Wenzel, S.; Binz, T.; König, C.; Blankenfeldt, W.; Brönstrup, M.; Müller, R. Biosynthesis of methyl-proline containing griselimycins, natural products with anti-tuberculosis activity. *Chem. Sci.* **2017**, *8*, 7521–7527.
- (S18) Hibi, M.; Kawashima, T.; Sokolov, P. M.; Smirnov, S. V.; Kodera, T.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. L-Leucine 5-hydroxylase of *Nostoc punctiforme* is a novel type of Fe (II)/ α -ketoglutarate-dependent dioxygenase that is useful as a biocatalyst. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2467–2472.
- (S19) Luesch, H.; Hoffmann, D.; Hevel, J. M.; Becker, J. E.; Golakoti, T.; Moore, R. E. Biosynthesis of 4-Methylproline in Cyanobacteria: cloning of nosE and nosF genes and biochemical characterization of the encoded dehydrogenase and reductase activities. *J. Org. Chem.* **2003**, *68*, 83–91.