ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Computational Insights into Amide Bond Formation Catalyzed by the Condensation Domain of Nonribosomal Peptide Synthetases

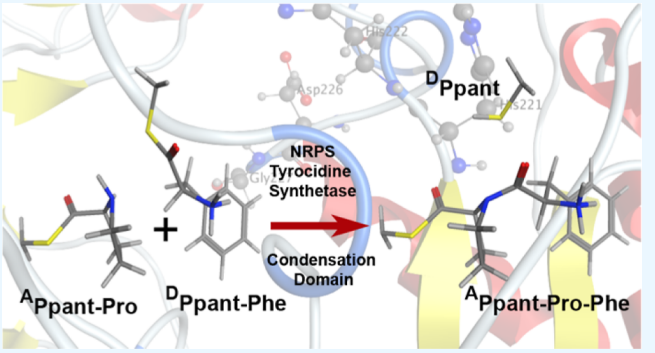
Basel Mansour and James W. Gauld*

Cite This: *ACS Omega* 2024, 9, 28556−28563

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Nonribosomal peptide synthetases (NRPSs) are important enzymes that synthesize an array of nongenetically encoded peptides. The latter have diverse physicochemical properties and roles. NRPSs are modular enzymes in which, for example, the condensation (C-) domain catalyzes the formation of amide bonds. The NRPS tyrocidine synthetase from *Brevibacillus brevis* is responsible for synthesizing the cyclic-peptide antibiotic tyrocidine. The first step is formation of an amide bond between a proline and phenylalanine which is catalyzed by a C-domain. In this study, a multiscale computational approach (molecular dynamics and QM/MM) has been used to investigate substrate binding and catalytic mechanism of the C-domain of tyrocidine synthetase. Overall, the mechanism is found to proceed through three exergonic steps in which an active site Histidine, His222, acts as a base and acid. First, His222 acts as a base to facilitate nucleophilic attack of the prolyl nitrogen at the phenylalanyl's carbonyl carbon. This is also the rate-limiting step with a free energy barrier of 38.8 kJ mol^{-1}. The second step is collapse of the resulting tetrahedral intermediate with cleavage of the S−C bond between the phenylalanyl and its Ppant arm, along with formation of the above amide bond. Meanwhile, the now protonated His222 imidazole has rotated toward the newly formed thiolate of the Ppant arm. In the final step, His222 acts as an acid, protonating the thiolate and regenerating a neutral His222. The overall mechanism is found to be exergonic with the final product complex being 46.3 kJ mol^{-1} lower in energy than the initial reactant complex.



INTRODUCTION

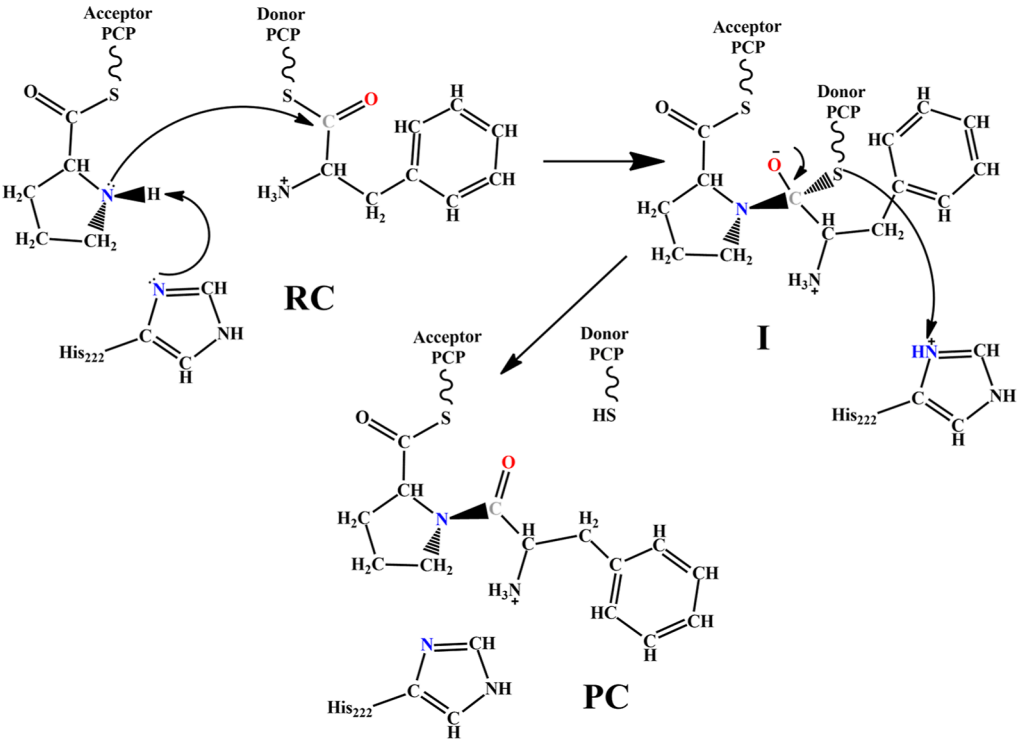
Bacteria and fungi are known to exploit numerous nonribosomal peptides (NRPs), possessing a diverse array of physicochemical properties, to perform a broad range of vital physiological functions.[1,2] For example, these functions include defense against pathogens, role as siderophores, or aid adaptation and responses to challenging environments.[3,4] As a result, some NRPs hold promise or are currently being used, for medicinal and industrial applications including as antibiotics, anticancer agents, insecticides, or immune system modulators.[5−7] Thus, there is considerable interest in elucidating how such diverse NRPs are synthesized.

Within the above organisms, the formation of NRPs is catalyzed by nonribosomal peptide synthetases (NRPSs).[8,9] The majority of NRPSs are organized in a modular structure,[8] where each module contains three main catalytically active domains responsible for specific chemical reactions.[10] Various combinations of domains and thus modules ultimately determine the NRP synthesized. For example, the initial (adenylation) domain selects and activates a specific amino acid. The resulting adenylated derivative is then transferred to a peptide carrier protein (PCP-)/thiolation domain[11] where the aminoacyl is transferred and covalently bound to a phosphopantetheine prosthetic (Ppant) arm.[12−14] This is then transported to the condensation (C-) domain which, crucially, catalyzes the formation of an amide bond between the aminoacyl group and the peptide already attached to the PCP domain of the preceding module, thereby extending the peptide.[10,11,13] That is, the C-domain catalyzes a key elongation step in the overall formation of NRPs.

Three different NRPS domains are capable of catalyzing the formation of amide bonds: the condensation, cyclization, and epimerization domains.[15,16] More specifically, the condensation and epimerization domains feature a highly conserved HHxxxDG motif,[17] with the second histidine (His222 in PDB ID: 2JGP) playing a proposed role in their catalytic mechanism, akin to other members of the noncanonical C-domain superfamily, for instance, chloramphenicol acetyltransferase, to which they belong. It is noted that the latter bacterial enzyme is key for bestowing their resistance to the

Received: March 15, 2024
Revised: June 10, 2024
Accepted: June 12, 2024
Published: June 22, 2024

Scheme 1. Proposed[17] Mechanism for the Amide Bond Formation between a Proline and Phenylalanine as Catalyzed by the Condensation Domain of the NRPS Tyrocidine Synthetase (RC = Reactant Complex; I = Intermediate; PC = Product Complex)



antibiotic chloramphenicol.[18,19] In contrast, the cyclization domain possesses a conserved DxxxxD motif.[20,21] The exact mechanism(s) by which these domains initiate and catalyze peptide bond formation remains unclear or unknown.

In particular, in the active site of the condensation domain, peptide bond formation necessitates the presence of two amino acid-loaded Ppant arms: one acting as the acceptor and the other as the donor.[16,22] This family is characterized by a pseudodimeric structure consisting of approximately 450 residues, forming a V-shape with an active site in between, covered by a latch that regulates the introduction of the Ppant arms.[21,23] Several studies have aimed to elucidate the role of key residues and the catalytic mechanism of the condensation domain. In particular, mutation of His222 to alanine eliminated its ability to catalyze amide bond formation.[10,21,24,25] It was proposed[10,17,21] that this histidine acts as a general base, accepting a proton from the α-amino of the aminoacyl group attached to the acceptor PCP. This is illustrated in Scheme 1 which uses the required amide bond formation between a proline and phenylalanine, the first step in the synthesis of the antibiotic tyrocidine, as an example. The acceptor's nitrogen (^{A}N) is then able to nucleophilically attack the carbonyl carbon ($^{D}C_{carb}$) of the aminoacyl carbon group attached to the donor PCP to form a tetrahedral intermediate. The final product is then formed via cleavage of the carbon-sulfur bond between $^{D}C_{carb}$ and the thiol sulfur of the donor Ppant arm (see Scheme 1).[10,17,21] However, another predominantly experimental study[19] in part considered the pK_a of the active site His222 in tyrocidine synthetase. Specifically, using the pK_a predictor tool H++ they suggested that it may in fact be protonated at physiological pH, thus raising questions about its ability to act as a base. Meanwhile, other experimental studies have suggested that the role of this critical second histidyl may instead be to stabilize transition states and tetrahedral intermediates.[26] Notably, they obtained an X-ray crystal structure of the NRPS surfactin A synthetase with a ligand in its active site (PDB ID: 2VSQ).[27] Based on its putative hydrogen bond network, they concluded that the second histidyl is likely not protonated at physiological pH and, furthermore, that its primary role is in substrate positioning.[28]

A greater understanding of the catalytic mechanism of the condensation domain of NRPSs holds the potential to also provide insights into, given their shared motifs, the NRPS epimerization domains and to, for example, aid the development of new peptides[29] and antibiotics.[30] Thus, in this present study, we have used a multiscale computational approach to examine the active site and catalytic mechanism of the condensation domain of tyrocidine synthetase from *Brevibacillus brevis*.[20] This NRPS catalyzes formation of the antibiotic tyrocidine, of which the first step is the C-domain catalyzed formation of an amide bond between a proline and phenylalanine (see Scheme 1). More specifically, docking, molecular dynamics (MD) simulations, and quantum mechanics/molecular mechanics (QM/MM) methods have been used to investigate the properties and roles of key active site residues as well as the overall catalytic mechanism.

COMPUTATIONAL METHODS

There is presently no available X-ray crystal structure of the condensation (C-) domain of tyrocidine synthetase from *B. brevis* with bound substrates. Thus, a suitable model was constructed from relevant known structures. More specifically, a suitable X-ray crystal structure of an unbound tyrocidine synthetase C-domain was obtained from the Protein Databank (PDB ID: 2JGP).[31] Concomitantly, a suitable X-ray crystal

https://doi.org/10.1021/acsomega.4c02531
ACS Omega 2024, 9, 28556−28563

structure of the PCP domain with an attached Ppant arm was also obtained from the Protein Databank (PDB ID: 7KVW).[32] A phenylalanine or proline was attached to the PCP arm to form the donor (^{D}Subs.) and acceptor (^{A}Subs.) ligands required for the catalyzed amide bond formation. It is noted that the Ppant arm is attached to the PCP protein via a phosphate group. In the chosen X-ray structure of the C-domain, there are two SO_4^{2-} groups: one near Arg45, mimicking the putative position of the phosphate group of ^{D}Subs., with the other hydrogen bonded to the active site His222 residue and positioned near Asp226, and Gly227.[31] To facilitate docking and construction of the fully bound reactive complex, the Molecular Operating Environment (MOE) program was used including for the initial determining of the protein's charges and hydrogens.[33] The protein surface was then calculated to aid identification of the potential required channel(s) for substrate binding (i.e., the two aminoacyl-loaded PCP complexes) that traversed the active site. Such a channel was identified that also passed through the sulfate group bound near Arg45. Therefore, ^{D}Subs. was manually docked in the channel near Arg45 with ^{A}Subs. docked on the other side and positioned as suggested by the observed hydrogen bond with His222.

MD Simulations. The Gromacs 2020.2 package was used for all MD simulations.[34] As the primary focus in this study is on the catalytic mechanisms of the condensation (C-) domain, which is at a distance from the PCP domains, only their aminoacyl-loaded Ppant arms were included in MD simulations and thus, the subsequent Hybrid Quantum Mechanics/MM calculation analysis. As noted above, the aminoacyl-loaded acceptor and donor arms are labeled as ^{A}Subs. and ^{D}Subs., respectively. For nonstandard residues, forcefield parameters were generated using acpype,[35] while the Amber99SB-ILDN forcefield was used for all other residues.[36] The substrate−ligand complexes were solvated to a distance of 10 Å from any protein atom while Cl^- and Na^+ ions were added to obtain a neutral system. Thus, the final solvated complex consisted of the condensation domain and the two model substrates (^{A}Subs and ^{D}Subs), which were 8381 atoms, and Na^+/Cl^- to neutralize the system's total charge and water molecules (which together was an additional 32937 atoms) for a total of 41318 atoms (Figure S1). The entire solvated neutralized system was placed inside a box of size 7.8 nm × 7.7 nm × 6.8 nm.

The solvated protein and protein−ligand complexes were then energy minimized, using the above forcefield until the root-mean-square gradient fell below 0.01 kcal/mol Å. This was followed by 200 ps equilibrations in both the *NVT* and *NPT* ensembles with a default time step of 2 fs. It is noted that the default Berendsen thermostat was used and set to 300 K. The default cutoff range for the nonbonded interactions, 10 Å, was used. For each system, a production 100 ns MD simulation was then run with a time step of 2 fs. More specifically, this protocol was applied to two systems: (a) enzyme only and (b) enzyme with both ^{A}Subs. and ^{D}Subs. bound in the active site. It is noted that 3 MD simulations were run on the fully bound enzyme-truncated substrate system for a period of 100 ns each (Figure S2). In all 3 simulations, the systems were similar and consistent for at least the period 0−30 ns, due to the truncated PCP substrates, and with the aminoacyl moieties of the substrates (^{A}Subs and ^{D}Subs) positioned adjacent to each other. Indeed, as part of the analysis, the rmsd of only substrates themselves, i.e., ^{A}Subs + ^{D}Subs, was also considered. These rmsd values were found to be generally low and consistent over the duration of the first 30 ns of the MD simulations. Hence, all MD analyses herein are based on this 30 ns MD simulation duration.

For the rmsd calculations of the entire enzyme, the backbone C, N, and O atoms of all residues were used. For the bound active site, all atoms of the two substrates, ^{A}Subs, and ^{D}Subs, were used, as well as all those of the His221, His222, Gly226, and Asp227 residues. If only the active site rmsd was considered, then only those of the His221, His222, Gly226, and Asp227 residues were used. Cluster analyses of the structures obtained were done using the gmx tool implemented in GROMACS, using the GROMOS method, with the cutoff set to 0.075 nm.

QM/MM Calculations. All ONIOM(QM/MM)[37,38] calculations were performed using the Gaussian16 suite of programs.[39] For the ONIOM(QM/MM)[37,38] studies, after cluster analysis of the above appropriate MD simulation a representative structure of the most populated conformation of the enzyme−ligands complex was chosen. More specifically, for the most populated cluster obtained based on rmsd of the active site residues (His221, His222, Asp226, Gly227), we selected those frames in which concomitantly the mechanistically relevant His222-N···HN-^{A}Subs and ^{A}Subs-N···C_{carb}-^{D}Subs were less than 3.0 and 3.5 Å, respectively. In addition, noting that in the MD simulations the structures were most consistent over the last 15−30 ns, a frame meeting the above distances within this period was selected. The high-level Quantum Mechanics (QM)-region contained the substrates, mechanistically important residues, and selected active site water molecules. The surrounding low-level MM region contained the rest of the protein as well as all waters up to 20 Å from the QM region. To help ensure structural integrity, all molecules 15−20 Å from the QM region were held fixed, all others being free to move.

For obtaining optimized structures, the QM region was described using the hybrid density functional theory (DFT) method B3LYP,[40,41] with inclusion of the empirical D3 dispersion correction developed by Grimme,[42,43] in combination with the 6−31G(d,p) basis set. Such methods have been widely used in the study of similar enzymatic reactions,[44−47] while Grimme's correction provides improved descriptions of noncovalent and dispersion interactions and the accuracy of kinetic barriers.[48] The MM region was described using the Amber force field[49] as implemented in Gaussian16.[39] Harmonic vibrational frequency calculations were also performed at this level of theory to determine the nature of the stationary points and to calculate Gibbs free energy corrections (ΔG_{corr}). It is noted that for free energy corrections, the default temperature of 298.15 K was used. Relative energies were obtained by performing single-point energy calculations on the above ONIOM(B3LYP-D3/6-31G(d,p):Amber) optimized structures at the ONIOM-(B3LYP-D3/6-311+G(2df,p):Amber) level of theory, and including the corresponding ΔG_{corr}. It is noted that all ONIOM calculations were conducted within a mechanical embedding (ME) formalism. This methodology has previously been successfully used to investigate a range of enzymatic mechanisms.[50−52] The % occurrence of key hydrogen bonds was calculated using distance and the angle cutoffs between the key atomic centers. For instance, the distance between the *X* and *Y* centers in the **X**(−H)···**Y** hydrogen bond was selected as

https://doi.org/10.1021/acsomega.4c02531
ACS Omega 2024, 9, 28556−28563

3.0 Å while concomitantly the $X\cdots H\cdots Y$ angle was taken into account and had to lie in the range 140−180°.

RESULTS AND DISCUSSION

As noted above, it has been suggested that His222 may be either neutral or protonated. Thus, using the default Protonate3D algorithm in MOE, we examined the pK_a of His222 and the effects of substrate binding on its value. In the unbound system, the pK_a of His222 was predicted to be 3.9. Upon substrate binding, the pK_a of His222 is further significantly decreased. That is, in the C-domain, His222 prefers to be neutral, and this is also the case upon substrate binding. As a result, herein, we have only considered the catalytic mechanism when His222 is neutral.

Docking and MD Simulations. As noted above, both the prolyl-loaded acceptor (^{A}Subs.) and phenylalanyl-loaded donor (^{D}Subs.) Ppant arms, the substrates of the first step in tyrocidine synthesis, were docked within the active site. In particular, as shown in Figure 1, the ligands were predicted to

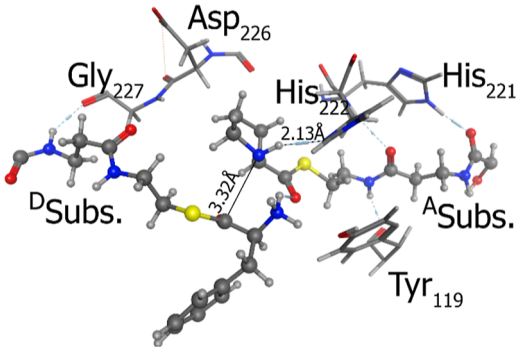


Figure 1. Illustration of the positioning, obtained by Docking, of the prolyl-loaded acceptor (^{A}Subs.) and phenylalanyl-loaded donor (^{D}Subs.) substrates within the active site of the NRPS tyrocidine synthetase.

preferably bind such that the prolyl −NH− ^{A}Subs. group was positioned between and near both the side chain imidazole of His222, and the carbonyl carbon of ^{A}Subs. at which nucleophilic attack is to occur. In particular, the ^{D}Subs.−NH···N−His222 and ^{D}Subs.−N(H)···(O)C−^{A}Subs. distances are 2.13 and 3.32 Å, respectively. In the subsequent MD simulation, these general positions were maintained with average ^{D}Subs.−NH···N−His222 and ^{D}Subs.−N(H)···(O)-C−^{A}Subs. distances of 2.2 and 3.3 Å, respectively.

It is noted that a plot of the rmsd of the backbone of the enzyme-ligand obtained over the course of the production MD simulation showed that its backbone is conformationally stable (Figure S3). Importantly, the rmsd of the key active site residues His221, His222, Asp226, and Gly227, over the course of the MD simulation, was plotted and is shown in Figure 2A. As can be seen, the rmsd varies from approximately 0.03 to 0.1 nm over the course of the MD simulation. Analysis of the contribution of each residue indicated that this variation is in large part due to conformational changes in the side chain of Asp226. In particular, its side chain carboxylate is hydrogen bonded to the backbone −NH− groups of Phe371 and Val372. In addition, it is also hydrogen bonded to a side chain $=NH_2^+$ of Arg359 which is involved in binding and positioning the substrate. That is, while Asp226 is not directly involved in substrate binding or catalysis in the C-domain of tyrocidine synthetase from *B. brevis*, it appears to have an important role as a second-shell residue that aids active site structure and indirectly, substrate binding.

The residue with the second greatest measured variability was His222. As noted above, its side chain imidazole (Im) is positioned directly adjacent to the ^{A}Subs. substrate and, more specifically, is proposed to deprotonate the ^{A}Subs. prolyl −NH− group (Figure 2B). Thus, to help gain greater insights into the potential role of His222 and the hydrogen bond network in which it is involved, the number and variation in hydrogen bonding of the imidazole His222 (His222-Im) with other residues, the substrate's −NH− group, as well as active site waters was examined. The average distances and % occurrence of key hydrogen bond interactions are given in Table 1.

Table 1. Average Distances, Standard Deviation (Std. Dev.), and % Occurrence of Key Hydrogen Bonds Involving the Imidazole Sidechain of His222 with Surrounding Residues and the ASubs. Substrate

hydrogen bond	average distance (Å)	std. dev. (Å)	% occurrence (%)
His222-Nε$H\cdots$Ocarb-Ala102	2.1	0.53	47
His222-N$\delta\cdots$HN-^{A}Subs.	2.2	0.03	69
His222-N$\delta\cdots$HO-Tyr119	3.9	0.96	4

As can be seen in Table 1, the Nε−H group of His222-Im is hydrogen bonded to the backbone carbonyl of Ala102, 47%, with an average distance of 2.1 Å. It is noted that this distance

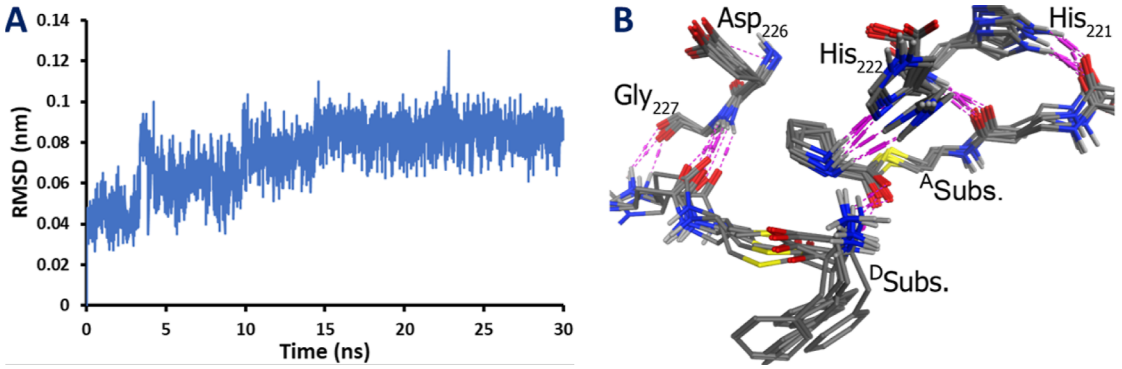


Figure 2. (A) Plot of the rmsd of the active site residues His221, His222, Asp226, and Gly227 obtained over the course of the MD simulation and (B) snapshots taken from cluster analysis of the MD simulation, of the QM region residues and substrates.

https://doi.org/10.1021/acsomega.4c02531
ACS Omega 2024, 9, 28556−28563

may be slightly larger due to an observed transitory (~700 ps) rotation of the His222-Im ring. Meanwhile, the Nδ center of His222-Im is quite consistently hydrogen bonded to the −NH− group of the ^{A}Subs. Substrate's prolyl residue, as evidenced by its % occurrence of 69% and concomitant lowest standard deviation (Table 1). Over the 30 ns duration of the MD simulation, the average distance between the nitrogen centers of the His222-N···HN-^{A}Subs hydrogen bond is 3.2 Å. It is noted that the His222-Im Nδ center also, but markedly less frequently (4%), forms hydrogen bonds with the phenolic hydroxyl of Tyr119. An interchangeable water was also observed to at times be situated near the Nδ center of His222-Im; however, it did not form hydrogen bonds with either the His222-Im Nδ center of the acceptor substrate's key prolyl −NH− group. Thus, an active site water is not directly involved in deprotonating the substrate. The conformational variability of Hs222-Im as well as key active site residues, ^{A}Subs. prolyl group and active site water are further illustrated in Figure 3. It clearly shows the consistent hydrogen bond network involving the backbone of Ala102, His222-Im, the substrate's −NH− group, and the high positional variability of the water.

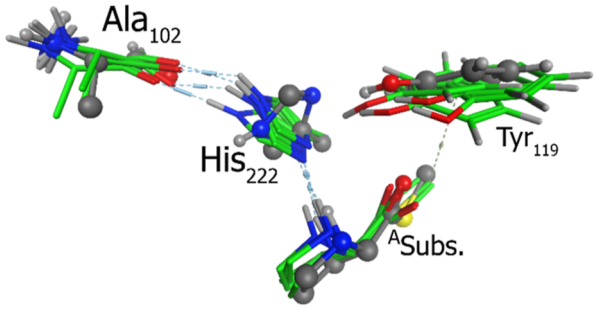


Figure 3. Overlaid snapshots from the most populated cluster obtained from rmsd analysis of His222 in the MD simulation.

The results of the docking and MD simulations support His222 as being the mechanistic base that deprotonates the prolyl−NH− group of the ^{A}Subs. A suitable starting structure was then selected after cluster analysis from the MD simulation to study the catalytic mechanism of the condensation domain of tyrocidine synthetase from *B. brevis*[20] using QM/MM (see Computational Methods).

QM/MM Investigation of the Enzymatic Mechanism of Tyrocidine Synthetase. The QM/MM optimized (see Computational Methods) structures obtained are shown in Figure 4. Selected key distances and angles are also provided (see also Tables S1−S7). The relative free energy surface obtained at the ONIOM(B3LYP-D3/6-311+G-(2df,p):Amber)-ME//ONIOM(B3LYP-D3/6-31G-(d,p):Amber)-ME + ΔG_{corr} level of theory (see Computational Methods), for the condensation domain of tyrocidine synthetase's catalyzed formation of an amide (peptide) bond between the prolyl and phenylalanyl moieties of ^{A}Subs. and ^{D}Subs. is provided in Figure 5.

In the optimized structure of the reactant complex **RC**, the imidazole of His222 forms a moderately short hydrogen bond with the prolyl −NH− of ^{A}Subs. with a length of 1.94 Å (Figure 4). It is noted that the strength of the interaction is underscored by the fact that the prolyl N−H distance in **RC** of 1.04 Å is 0.02 Å longer than for the optimized structure of just ^{A}Subs. obtained at the same level of DFT theory but in a dielectric medium (ε) of 4.0 and 78.35 (i.e., protein and water). Concomitantly, the lone pair on the prolyl nitrogen is directed toward the mechanistically important carbonyl carbon of ^{D}Subs. with a ^{A}Subs-N···O-^{D}Subs. distance of only 2.18 Å. Meanwhile, the $-NH_3^+$ terminus of ^{D}Subs indirectly hydrogen bonds, via water, with DSubs key carbonyl oxygen. It is noted that this interaction is observed in each of the subsequent intermediates, transition structures, and products. Thus, it may help facilitate the formation of an oxyanion during the mechanism, and its subsequent stabilization during the first step of the overall mechanism. Thus, clearly, in **RC**, the substrates are well positioned to react and, furthermore, His222 is well positioned to deprotonate the ^{A}Subs. −NH− group while a water molecule also seems well positioned to help stabilize the negative charge on the carbonyl oxygen of ^{D}Subs. It is noted that we have also considered electrostatic effects of the active site environment on this initial step, through single points at the *E*[ONIOM(B3LYP-D3/6-31G-(d,p):Amber)-EE//ONIOM(B3LYP-D3/6-31G-(d,p):Amber)-ME] − *E*[ONIOM(B3LYP-D3/6-31G-(d,p):Amber)-ME] level of theory, and then correcting the relative free energy previously obtained (see Computational Methods). It is noted that the activation barrier for this initial nucleophilic attack increased by 6.7 to 45.5 kJ mol^{-1} and is still calculated to be the overall rate-limiting step for this mechanism (Figure S4).

Indeed, the first step of the overall mechanism is the nucleophilic attack of the ^{A}Subs. prolyl −NH− nitrogen at the carbonyl carbon of the phenylalanyl moiety of ^{D}Subs. This reaction proceeds via tetrahedral transition structure **TS1**, at a cost of 38.8 kJ mol^{-1}, to give the tetrahedral intermediate **IC1**. The latter lies 22.2 kJ mol^{-1} lower in free energy than **RC**, and thus this step is exergonic and is also the rate-limiting step of the overall mechanism. In **TS1**, the prolyl −NH− proton has transferred onto the His222 Imidazole as evidenced by a His222-N−H bond length of just 1.07 Å (Figure 4). Meanwhile, the ^{A}Subs-N···C_{carb}-^{D}Subs. bond has partially formed as indicated by its now markedly shortened distance of 1.54 Å. Concomitantly, the C_{carb}−O_{carb} bond has lengthened to 1.28 Å, due to its increased oxyanion character that is partially stabilized by a strengthened hydrogen bond (1.72 Å) with the active site water. Meanwhile, the C_{carb}−S bond, which must cleave during the mechanism, has lengthened significantly by 0.23 Å to 2.08 Å. The key changes in **IC1** are that the ^{A}Subs-N···C_{carb}-^{D}Subs. bond has shortened further to just 1.54 Å and importantly, the His222 imidazole has rotated slightly such that its newly formed −NH− group points between the N and S centers of the formed intermediate.

The next step was found to be cleavage of the C_{carb}−S bond which occurs with concomitant formation of the amide bond between the initial prolyl and phenylalanyl moieties. This step proceeds via **TS2** at a cost of just 7.5 kJ mol^{-1} relative to IC1, to give the slightly lower energy intermediate **IC2** which has a relative free energy of −26.6 kJ mol^{-1} with respect to the initial **RC**. Thus, this second step is also exergonic but by just 4.4 kJ mol^{-1} (see Figure 5). Structurally, as shown in Figure 4, the C_{carb}−S bond has now broken while the newly formed C_{carb}−N amide bond length has shortened significantly to just 1.39 Å. The cleaved thiolate group originates from ^{D}Subs. (the thiolate sulfur of the donor Ppant arm) and now forms a moderately strong hydrogen bond of length 2.16 Å with the previously

https://doi.org/10.1021/acsomega.4c02531
ACS Omega 2024, 9, 28556−28563

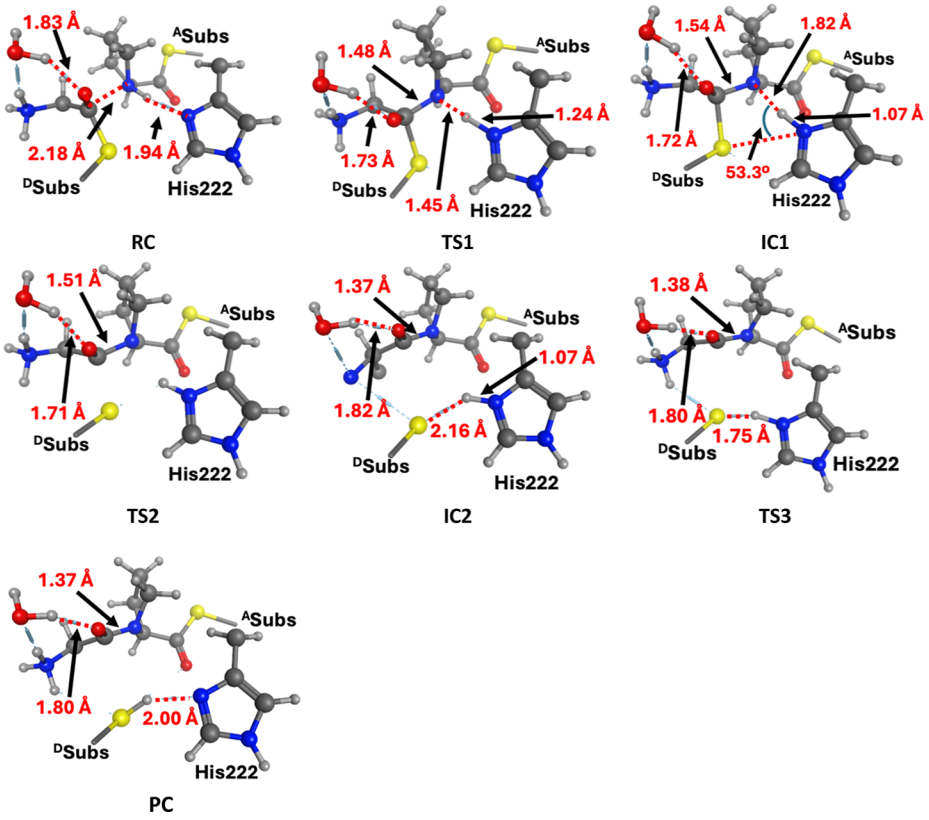


Figure 4. Optimized geometries (see Computational Methods) of the reactant and product complexes, intermediates, and transition structures obtained for tyrocidine synthetase's catalyzed amide bond formation between the prolyl and phenylalanyl of ^{A}Subs. and ^{D}Subs. For clarity, only selected bond lengths (Å) and angles (°), and key components of the QM region are shown. Full structures are provided in Tables S1−S7.

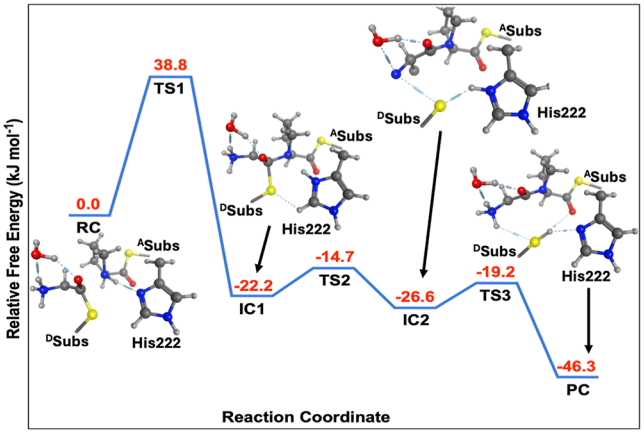


Figure 5. Relative free energy (kJ mol^{-1}) surface obtained (see Computational Methods) for tyrocidine synthetase's catalyzed amide bond formation between the prolyl and phenylalanyl moieties of ^{A}Subs. and ^{D}Subs.

formed protonated His222 side chain imidazole. The latter has undergone further rotation such that it is in fact no longer directed toward the former prolyl nitrogen of ^{A}Subs. (see Figure 4).

The third and final step of the overall mechanism is a proton transfer from His222 to the newly formed thiolate, which itself originated from the Ppant arm component of ^{D}Subs. That is, it is essentially regeneration of the initial neutral active site His222 imidazole with concomitant formation of the neutral donor Ppant arm complex. This proceeds via **TS3** with a barrier of just 7.4 kJ mol^{-1}, the lowest barrier of all 3 steps, to give the final product complex (**PC**) which is lower in free energy than **RC** by 46.3 kJ mol^{-1}. Thus, the overall catalytic mechanism is exergonic with each of its individual 3 steps also being exergonic (Figure 5). As shown in Figure 4, the former thiolate originating from the Ppant arm component of ^{D}Subs is now neutralized with an S−H bond length of 1.36 Å. It is also noted, however, that the newly formed C_{carb}−N amide bond has shortened further from that observed in **IC2** (1.38 Å) to 1.37 Å.

CONCLUSIONS

The NRPS tyrocidine synthetase from *B. brevis* is responsible for synthesizing the antibiotic tyrocidine. Its condensation (C)-domain catalyzes the first step in this overall process: amide (peptide) bond formation between the prolyl and phenylalanyl moieties of its two aminoacyl-Ppant substrates. In this study, a multiscale computational approach of docking, MD, and QM/MM methods, has been used to investigate substrate binding and the catalytic mechanism of the C-domain. For the latter method, optimized structures, harmonic vibrational frequencies, and free energy corrections (ΔG_{corr}) were obtained at the ONIOM(B3LYP-D3/6-31G-(d,p):Amber)-ME level of theory while relative free energies were obtained at the ONIOM(B3LYP-D3/6-311+G-(2df,p):Amber)-ME level of theory using the above optimized structures and corrections.

In the fully bound active site the prolyl substrate's (^{A}Subs) −HN− group directly hydrogen bonds to the side chain imidazole of His222. Concomitantly, the lone pair electrons of N-^{A}Subs are directed toward the carbonyl group of the

https://doi.org/10.1021/acsomega.4c02531
ACS Omega 2024, 9, 28556−28563

phenylalanyl substrate (^{D}Subs). The first step in the overall mechanism is shown to be an exergonic nucleophilic attack of the N-^{A}Subs center at the C_{carb}-^{D}Subs center, with concomitant transfer of the HN-^{A}Subs proton onto His222. This is the rate-limiting step of the overall mechanism with a free energy barrier of 38.8 kJ mol^{-1}. The resulting tetrahedral intermediate (**IC1**) is stabilized by an active site water.

The second step is the collapse of the tetrahedral intermediate (**IC1**) with cleavage of the S−C bond in ^{D}Subs, i.e., between the phenylalanyl moiety and its Ppant arm. This step, which is also found to be exergonic, occurs with a free energy barrier of just 7.5 kJ mol^{-1}. In the resulting intermediate, **IC2**, an amide bond has now been formed between the two substrate's prolyl and phenylalanyl moieties. Meanwhile, the now protonated imidazole of His222 has rotated such that it is now directed toward the newly formed thiolate sulfur of the Ppant arm (^{D}Ppant-S^{-}) to which the phenylalanine was formerly attached.

The final step is proton transfer from His222, i.e., His222 now acts as an acid, onto the ^{D}Ppant-S^{-} center, regenerating the initial neutral active site His222 residue. It is noted that the PCP domain is also regenerated. This process which occurs with a free energy barrier of 7.4 kJ mol^{-1} at the present level of theory is also found to be exergonic. As a result, the overall mechanism is found to be exergonic with the final product complex (**PC**) calculated to lie 46.3 kJ mol^{-1} lower in energy than the initial reactant complex, **RC**. The present results represent an important key step toward a larger-scale computational investigation of the effects, interplay, and roles of the multiple catalytic domains within NRPS modules.

As noted previously, the NRPS condensation domain shares its active site motif, H**H**xxxDG, with NRPS epimerization domains as well as several other important enzymes. Thus, these new insights into substrate binding and the catalytic mechanism of the C-domain of tyrocidine synthetase from *B. brevis*, potentially have wider importance and applicability.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c02531.

Plot of the rmsd of the peptide backbone for the fully bound enzyme−substrates complex and Gaussian 16 archive entries from all QM/MM single-point calculations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

James W. Gauld − *Department of Chemistry and Biochemistry, University of Windsor, Windsor, Ontario N9B 3P4, Canada;* orcid.org/0000-0002-2956-9781; Email: gauld@uwindsor.ca

Author

Basel Mansour − *Department of Chemistry and Biochemistry, University of Windsor, Windsor, Ontario N9B 3P4, Canada*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c02531

Author Contributions

Basel Mansour: calculations, analyses, and writing, James W. Gauld: mentoring, analyses, writing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Natural Science and Engineering Research Council of Canada (NSERC) for financial support. We thank Compute Canada and Digital Alliance for computational resources.

■ REFERENCES

(1) Matilla, M. A.; Krell, T. Plant Growth Promotion and Biocontrol Mediated by Plant-Associated Bacteria. In *Plant Microbiome: Stress Response*; Egamberdieva, D., Ahmad, P., Eds.; Springer Singapore: Singapore, 2018; pp 45−80.

(2) Giordano, D.; Coppola, D.; Russo, R.; Denaro, R.; Giuliano, L.; Lauro, F. M.; di Prisco, G.; Verde, C. Chapter Four—Marine Microbial Secondary Metabolites: Pathways, Evolution and Physiological Roles. In *Advances in Microbial Physiology*; Poole, R. K., Ed.; Academic Press, 2015; Vol. *66*, pp 357−428.

(3) Brader, G.; Compant, S.; Mitter, B.; Trognitz, F.; Sessitsch, A. Metabolic potential of endophytic bacteria. *Curr. Opin. Biotechnol.* **2014**, *27*, 30−37.

(4) Yu, A. O.; Leveau, J. H. J.; Marco, M. L. Abundance, diversity and plant-specific adaptations of plant-associated lactic acid bacteria. *Environ. Microbiol. Rep.* **2020**, *12*, 16−29.

(5) Katz, L.; Baltz, R. H. Natural product discovery: past, present, and future. *J. Indust. Microbiol. Biotechnol.* **2016**, *43*, 155−176.

(6) Santos-Aberturas, J.; Vior, N. M. Beyond Soil-Dwelling Actinobacteria: Fantastic Antibiotics and Where to Find Them. *Antibiotics* **2022**, *11*, 195.

(7) Adnani, N.; Rajski, S. R.; Bugni, T. S. Symbiosis-inspired approaches to antibiotic discovery. *Nat. Prod. Rep.* **2017**, *34*, 784−814.

(8) Payne, J. A.; Schoppet, M.; Hansen, M. H.; Cryle, M. J. Diversity of nature's assembly lines - recent discoveries in non-ribosomal peptide synthesis. *Mol. Biosyst.* **2017**, *13*, 9−22.

(9) Izoré, T.; Cryle, M. J. The many faces and important roles of protein-protein interactions during non-ribosomal peptide synthesis. *Nat. Prod. Rep.* **2018**, *35*, 1120−1139.

(10) Bloudoff, K.; Rodionov, D.; Schmeing, T. M. Crystal Structures of the First Condensation Domain of CDA Synthetase Suggest Conformational Changes during the Synthetic Cycle of Nonribosomal Peptide Synthetases. *J. Mol. Biol.* **2013**, *425*, 3137−3150.

(11) Keating, T. A.; Suo, Z.; Ehmann, D. E.; Walsh, C. T. Selectivity of the Yersiniabactin Synthetase Adenylation Domain in the Two-Step Process of Amino Acid Activation and Transfer to a Holo-Carrier Protein Domain. *Biochemistry* **2000**, *39*, 2297−2306.

(12) Keating, T. A.; Walsh, C. T. Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr. Opin. Chem. Biol.* **1999**, *3*, 598−606.

(13) Marahiel, M. A. Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. *J. Pept. Sci.* **2009**, *15*, 799−807.

(14) Mootz, H. D.; Schwarzer, D.; Marahiel, M. A. Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases. *ChemBioChem* **2002**, *3*, 490−504.

(15) Watanabe, K.; Oguri, H.; Oikawa, H. Diversification of echinomycin molecular structure by way of chemoenzymatic synthesis and heterologous expression of the engineered echinomycin biosynthetic pathway. *Curr. Opin. Chem. Biol.* **2009**, *13*, 189−196.

(16) Dekimpe, S.; Masschelein, J. Beyond peptide bond formation: the versatile role of condensation domains in natural product biosynthesis. *Nat. Prod. Rep.* **2021**, *38*, 1910−1937.

(17) Miller, B. R.; Gulick, A. M. Structural Biology of Nonribosomal Peptide Synthetases. *Methods Mol. Biol.* **2016**, *1401*, 3−29.

(18) Shaw, W. V. Chloramphenicol acetyltransferase: enzymology and molecular biology. *CRC Crit. Rev. Biochem.* **1983**, *14*, 1−46.

(19) Leslie, A. G.; Moody, P. C.; Shaw, W. V. Structure of chloramphenicol acetyltransferase at 1.75-A resolution. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4133.
(20) Bloudoff, K.; Fage, C. D.; Marahiel, M. A.; Schmeing, T. M. Structural and mutational analysis of the nonribosomal peptide synthetase heterocyclization domain provides insight into catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 95−100.
(21) Bloudoff, K.; Schmeing, T. M. Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. *Biochim. Biophys. Acta, Proteins Proteomics* **2017**, *1865* (11), 1587−1604.
(22) Hartwig, S.; Dovengerds, C.; Herrmann, C.; Hovemann, B. T. Drosophila Ebony: a novel type of nonribosomal peptide synthetase related enzyme with unusually fast peptide bond formation kinetics. *FEBS J.* **2014**, *281*, 5147−5158.
(23) Wheadon, M. J.; Townsend, C. A. Evolutionary and functional analysis of an NRPS condensation domain integrates β-lactam, D-amino acid, and dehydroamino acid synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2026017118.
(24) Muangkaew, P.; De Roo, V.; Zhou, L.; Girard, L.; Cesa-Luna, C.; Höfte, M.; De Mot, R.; Madder, A.; Geudens, N.; Martins, J. C. Stereomeric Lipopeptides from a Single Non-Ribosomal Peptide Synthetase as an Additional Source of Structural and Functional Diversification in Pseudomonas Lipopeptide Biosynthesis. *Int. J. Mol. Sci.* **2023**, *24*, 14302.
(25) Reeves, C. D. The Enzymology of Combinatorial Biosynthesis. *Crit. Rev. Biotechnol.* **2003**, *23*, 95−147.
(26) Bergendahl, V.; Linne, U.; Marahiel, M. A. Mutational analysis of the C-domain in nonribosomal peptide synthesis. *Eur. J. Biochem.* **2002**, *269*, 620−629.
(27) Tanovic, A.; Samel, S. A.; Essen, L.-O.; Marahiel, M. A. Crystal Structure of the Termination Module of a Nonribosomal Peptide Synthetase. *Science* **2008**, *321*, 659−663.
(28) Patel, K. D.; d'Andrea, F. B.; Gaudelli, N. M.; Buller, A. R.; Townsend, C. A.; Gulick, A. M. Structure of a bound peptide phosphonate reveals the mechanism of nocardicin bifunctional thioesterase epimerase-hydrolase half-reactions. *Nat. Commun.* **2019**, *10*, 3868.
(29) Niquille, D. L.; Folger, I. B.; Basler, S.; Hilvert, D. Biosynthetic Functionalization of Nonribosomal Peptides. *J. Am. Chem. Soc.* **2021**, *143*, 2736−2740.
(30) Felnagle, E. A.; Jackson, E. E.; Chan, Y. A.; Podevels, A. M.; Berti, A. D.; McMahon, M. D.; Thomas, M. G. Nonribosomal Peptide Synthetases Involved in the Production of Medically Relevant Natural Products. *Mol. Pharmaceutics* **2008**, *5*, 191−211.
(31) Samel, S. A.; Schoenafinger, G.; Knappe, T. A.; Marahiel, M. A.; Essen, L.-O. Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase. *Structure* **2007**, *15*, 781−792.
(32) Izoré, T.; Candace Ho, Y. T.; Kaczmarski, J. A.; Gavriilidou, A.; Chow, K. H.; Steer, D. L.; Goode, R. J. A.; Schittenhelm, R. B.; Tailhades, J.; Tosin, M.; Challis, G. L.; Krenske, E. H.; Ziemert, N.; Jackson, C. J.; Cryle, M. J. Structures of a non-ribosomal peptide synthetase condensation domain suggest the basis of substrate selectivity. *Nat. Commun.* **2021**, *12*, 2511.
(33) *Molecular Operating Environment (MOE)*, Chemical Computing Group ULC: 910-1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2023.
(34) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(35) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **2012**, *5*, 367.
(36) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712−725.
(37) Chung, L. W.; Hirao, H.; Li, X.; Morokuma, K. The ONIOM method: its foundation and applications to metalloenzymes and photobiology. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 327−350.
(38) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comput. Chem.* **2003**, *24*, 760−769.
(39) 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; 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.; 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*: Wallingford, CT, 2016.
(40) Tirado-Rives, J.; Jorgensen, W. L. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. *J. Chem. Theory Comput.* **2008**, *4*, 297−306.
(41) Govindarajan, M.; Ganasan, K.; Periandy, S.; Mohan, S. DFT (LSDA, B3LYP and B3PW91) comparative vibrational spectroscopic analysis of α-acetonaphthone. *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **2010**, *76*, 12−21.
(42) Grimme, S.; Steinmetz, M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031−16042.
(43) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112.
(44) Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J. Application of Quantum Mechanics/Molecular Mechanics Methods In The Study Of Enzymatic Reaction Mechanisms. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017**, *7*, No. e1281.
(45) Yagi, K.; Ito, S.; Sugita, Y. Exploring the Minimum-Energy Pathways and Free-Energy Profiles of Enzymatic Reactions with QM/MM calculations. *J. Phys. Chem. B* **2021**, *125*, 4701−4713.
(46) Kellie, J. L.; Wetmore, S. D. Selecting DFT methods for use in optimizations of enzyme active sites: applications to ONIOM treatments of DNA glycosylases. *Can. J. Chem.* **2013**, *91*, 559−572.
(47) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. *J. Chem. Theory Comput.* **2012**, *8*, 4637−4645.
(48) Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211−228.
(49) Yoda, T.; Sugita, Y.; Okamoto, Y. Comparisons of force fields for proteins by generalized-ensemble simulations. *Chem. Phys. Lett.* **2004**, *386*, 460−467.
(50) Wang, X.; Hirao, H. ONIOM (DFT:MM) Study of the Catalytic Mechanism of myo-Inositol Monophosphatase: Essential Role of Water in Enzyme Catalysis in the Two-Metal Mechanism. *J. Phys. Chem. B* **2013**, *117*, 833−842.
(51) Sharma, H.; Raju, B.; Narendra, G.; Motiwale, M.; Sharma, B.; Verma, H.; Silakari, O. QM/MM Studies on Enzyme Catalysis and Insight into Designing of New Inhibitors by ONIOM Approach: Recent Update. *ChemistrySelect* **2023**, *8*, No. e202203319.
(52) Yildiz, I. Computational Mechanistic Study of l-Aspartate Oxidase by ONIOM Method. *ACS Omega* **2023**, *8*, 19963−19968.

https://doi.org/10.1021/acsomega.4c02531
ACS Omega 2024, 9, 28556−28563