INTERACTIONS BETWEEN SMALL MOLECULE LIGANDS AND TARGET ENZYMES

EDITED BY: Sung-Kun (Sean) Kim, Dong-Woo Lee and Ki Duk Park

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INTERACTIONS BETWEEN SMALL MOLECULE LIGANDS AND TARGET ENZYMES

Topic Editors:
Sung-Kun (Sean) Kim, Northeastern State University, United States
Dong-Woo Lee, Yonsei University, South Korea
Ki Duk Park, Korea Institute of Science and Technology, South Korea

Table of Contents

Check for updates

Editorial: Interactions Between Small Molecule Ligands and Target Enzymes

Sung-Kun Kim[1]*, Ki Duk Park[2]* and Dong-Woo Lee[3]*

[1]Department of Natural Sciences, Northeastern State University, Broken Arrow, OK, United States, [2]Brain Science Institute, Korea Institute of Science and Technology, Seoul, South Korea, [3]Department of Biotechnology, Yonsei University, Seoul, South Korea

Editorial on the Research Topic

Interactions Between Small Molecule Ligands and Target Enzymes

The detail of the interactions between enzymes and ligands is a prerequisite for understanding cellular processes and developing novel therapeutic agents and chemical drugs. Although enzyme-ligand interactions occur via the complex formation driven by thermodynamics, the rate of complex formation and breakdown should also be considered on the reaction coordinate using kinetic parameters and thermodynamic stability. To gain insight into the molecular basis of enzyme-ligand interactions at the atomistic level, three-dimensional structural analysis of enzyme-ligand complexes is necessary. Nevertheless, it is noted that the protein-ligand interactions initiate an important conformational change and the binding states vary over the reaction period, representing a major challenge for the development of refined and optimized poses from the initial docking state. Accordingly, computational approaches have been conducted to advance the understanding of enzyme-ligand interactions. Recent advances in molecular docking and dynamic algorithms have provided more reliable and accurate illustrations of the enzyme-ligand complex, leading to the development of new regulatory ligands for target enzymes or protein engineering for better interactions.

In this research topic, we introduce the research articles and reviews tackling the interactions between enzyme and ligand using a wide variety of state-of-the-art techniques. We believe that the current, updated ideas and/or strategies to identify and develop new ligands for target enzymes can enhance the current understanding of structural flexibility or mutational effects on the ligand and its target, thus making an essential contribution to the fields of drug design and biotechnology.

The computational approaches are getting more critical to detailing the information of interactions between ligands and enzymes. The accuracy is substantially improved. Molecular dynamics simulation is the spearhead of the computational approaches to provide structural information. Adams et al. explained how the molecular dynamics simulations provide insight into antibiotic interactions using the enzyme L,L-diaminopimelate aminotransferase with the potential inhibitors. This study supports the great potential scope of the application of molecular dynamics to biological molecules to infer the function, form, and effect of interactions in situations where the crystal structure may not be known. Another example of a computational approach is an article published by Shaker et al. The authors attempted to design potent inhibitory lead molecules against *Pseudomonas aeruginosa* using an *in silico* ligand-based drug design. Subsequently, the researchers conducted a structure-based virtual screening with a library of potential inhibitors for anthranilate-CoA ligase, which is involved in the quorum signaling system of *P. aeruginosa*. The molecular dynamic simulation further validated resulting compounds in identifying improved inhibitors. Such a computational approach might provide

reasonably acceptable drug candidates against the resistant phenotypes of opportunistic infectious pathogens.

In this research topic series, two review articles described drug discovery. Hwang et al. introduces the structure and mechanism of modular PKS and NRPS, and includes examples of their repurposes. In addition, the authors discussed the tools of the design-build-test-learning cycle of synthetic biology and the future perspective of repurposing strategies. Vellanki et al. reviewed the recent effort in the discovery of novel non-immunosuppressive antifungal drugs against fungal pathogens. This review highlights the interactions of immunosuppressive drugs FK506 and rapamycin with immunophilin FKBP12 in targeting calcineurin and TOR kinase in human and fungal pathogens. Such a detailed inspection of ligand-bound protein structures will provide a useful guide to facilitate the development of new small molecules with antifungal activity.

Some research articles presented current issues we are facing. For example, extended-spectrum β-lactamases (ESBLs) are a serious threat to human health due to their enhanced hydrolytic activity against third-generation cephalosporins. Cao et al. described the structural mechanism underlying the hydrolysis of third-generation cephalosporins using non-catalytic-region (NCR)-ESBLs. Also, melanin, an important molecules in humans, plays a crucial role in protecting the skin from the harmful effects of ultraviolet radiation, but the abnormal accumulation of melanin causes skin pigmentary disorders. Kim et al. proposed a molecular mechanism for how 7,3′,3′-trihydroxyisoflavone (7,3′,4′-THIF) attenuates α-MSH-induced melanogenesis.

Interactions between ligands and proteins reveals many important biological and biochemical aspects. For example, protein-ligand binding structures enabled us to identify the functional role of a thermostable primordial protein. La et al. described the functional characterization of a gene encoding a putative keratin-degrading β-aspartyl peptidase. Determination

of its co-crystal structure with the substrate analog together with *in vivo* complementation experiments suggest that the viability of hyperthermophiles under stressful conditions may rely on the activity of BAP within cellular protein repair systems. Additionally, Kwon et al. describes the development of a fluorescent sensing system for monitoring enzyme activity in living cells. Specifically, the authors designed a DmpR-based genetic circuit relying on para-nitrophenol and alanine as regulators interacting with tyrosine phenol-lyase, allowing for straightforward control of its output signal without additional genetic modification. The developed genetic circuit system can be a useful tool for the accurate and rapid analysis of targeted enzyme activity in living cells.

The articles published here discussed a broad range of state-of-the-art computational and experimental studies, or a combination thereof, to understand the interactions between small ligands and their target biomolecules. We believe these articles help to emphasize the impact of the investigated processes in biomedicine and applications of enzyme-ligand interactions in biotechnology and other related science fields.

AUTHOR CONTRIBUTIONS

Check for updates

Comparative Molecular Dynamics Simulations Provide Insight Into Antibiotic Interactions: A Case Study Using the Enzyme L,L-Diaminopimelate Aminotransferase (DapL)

Lily E. Adams[1], Patrick Rynkiewicz[1], Gregory A. Babbitt[1], Jamie S. Mortensen[2], Rachel A. North[3], Renwick C. J. Dobson[3,4] and André O. Hudson[1]*

[1] Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, NY, United States, [2] Department of Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, United States, [3] Biomolecular Interaction Centre and School of Biological Sciences, University of Canterbury, Christchurch, New Zealand, [4] Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Melbourne, VIC, Australia

The L,L-diaminopimelate aminotransferase (DapL) pathway, a recently discovered variant of the lysine biosynthetic pathway, is an attractive pipeline to identify targets for the development of novel antibiotic compounds. DapL is a homodimer that catalyzes the conversion of tetrahydrodipicolinate to L,L-diaminopimelate in a single transamination reaction. The penultimate and ultimate products of the lysine biosynthesis pathway, *meso*-diaminopimelate and lysine, are key components of the Gram-negative and Gram-positive bacterial peptidoglycan cell wall. Humans are not able to synthesize lysine, and DapL has been identified in 13% of bacteria whose genomes have been sequenced and annotated to date, thus it is an attractive target for the development of narrow spectrum antibiotics through the prevention of both lysine biosynthesis and peptidoglycan crosslinking. To address the common lack of structural information when conducting compound screening experiments and provide support for the use of modeled structures, our analyses utilized inferred structures from related homologous enzymes. Using a comprehensive and comparative molecular dynamics (MD) software package—DROIDS (Detecting Relative Outlier Impacts in Dynamic Simulations) 2.0, we investigated the binding dynamics of four previously identified antagonistic ligands of DapL from *Verrucomicrobium spinosum*, a non-pathogenic relative of *Chlamydia trachomatis*. Here, we present putative docking positions of the four ligands and provide confirmatory comparative molecular dynamics simulations supporting the conformations. The simulations performed in this study can be applied to evaluate putative targets to predict compound effectiveness prior to *in vivo* and *in vitro* experimentation. Moreover, this approach has the potential to streamline the process of antibiotic development.

Keywords: L,L-diaminopimelate aminotransferase, molecular dynamics, ligand binding, antibiotics, L-lysine, diaminopimelate, peptidoglycan

INTRODUCTION

Novel targets and approaches to facilitate the development/discovery of antibiotics are highly sought after to combat the rise of single and multidrug resistant bacterial infections (Ventola, 2015). Bacterial enzymes that are involved in the anabolism of the nine nutritionally essential proteogenic amino acids (histidine, isoleucine, leucine, methionine, phenylalanine, threonine, tryptophan, valine, and lysine) are among the attractive targets. Humans cannot synthesize these amino acids, and they are instead obtained through the diet. Presumably, this means that targeting these bacterial amino acid synthesis pathways with antibiotics should be non-toxic for humans. The lysine biosynthesis pathway in particular is a well-known antibiotic target (Hutton et al., 2007; Triassi et al., 2014) because of its involvement in bacterial cell wall synthesis. In Gram-negative bacteria, *meso*-diaminopimelate (*meso*-DAP) is used as the crosslinker at the third position of a peptide stem that link individual peptidoglycan units to form the cell wall (**Figure 1E**); in Gram-positive bacteria, lysine serves this purpose (Hutton et al., 2007). Without the crosslinking of peptidoglycan, the cell wall loses stability and is unable to adequately protect and maintain separation of the internal and external environment, resulting in rupture. This critical role in peptidoglycan crosslinking and protein synthesis, as well as its lack of presence in animals (specifically humans), supports targeting the lysine biosynthetic pathways in the development of novel antibiotics.

Lysine biosynthesis occurs via one of two pathways: the alpha-aminoadipic (AAA) pathway is commonly found in yeast, fungi, and some archaea while the diaminopimelate (DAP) pathway is employed by photosynthetic organisms and most prokaryotes (Nishida et al., 1999; Velasco et al., 2002). DAP synthesis comprises three main steps: (1) conversion of aspartate to tetrahydrodipicolinate (THDP), (2) conversion of THDP to *meso*-diaminopimelate (*meso*-DAP), and (3) conversion of *meso*-DAP to lysine via an L,L-DAP intermediate (**Figure 1**). There are four known variants of the DAP pathway, distinguished by differences in the synthesis of *meso*-DAP: the two acyl pathways, the *meso*-diaminopimelate dehydrogenase (Ddh) pathway, and the L,L-diaminopimelate aminotransferase (DapL) pathway (Hudson et al., 2006). Notably, the DapL pathway variant utilizes DapL in a single transamination reaction to convert THDP directly to LL-DAP (Hudson et al., 2005, 2006).

DapL is a homodimer comprised of two subunits that are each ∼50 kDa in size (Hudson et al., 2006), functionally active as a dimer with two active sites on a hinge. Each subunit is composed of a major and minor arm, where the major arm composes the dimer interface for each subunit. The DapL enzyme transfers an amino group plus a proton and electron pair from a donor molecule (in this case glutamate) to an acceptor molecule (THDP). Three key regions of the enzyme are involved in each aminotransferase reaction—one loop on the minor arm of the enzyme and two loops from the major arm in the dimer interface. Thus, each active site requires coordination between both subunits for optimal function (Velick and Vavra, 1962; Liepman and Olsen, 2004).

A number of organisms of extreme biotechnological and clinical relevance employ the DapL pathway, including phyla *Chlamydophila, Treponema, Leptospira,* and *Bacteroides* (Berry et al., 1987; Burstain et al., 1991; Finegold and Sussman, 2002; Wexler, 2007; Lindsay et al., 2010; Ansdell, 2012). Additionally, the DapL pathway is home to numerous commensal and environmental organisms with major biotechnological applications (Adams, 2019). Understanding essential metabolic pathways at a molecular level in these species is critical for future treatment development. For example, *Chlamydia trachomatis* is the causative agent of one of the most reported sexually transmitted diseases in the United States and the leading cause of infection-mediated blindness worldwide (Mishori et al., 2012). It is an obligate intracellular pathogen with high rates of lateral gene transfer between serotypes and largely uncharacterized genetic properties (Mabey, 2008). *Verrucomicrobium spinosum*, another organism that employs the DapL pathway, is the closest free-living relative to *C. trachomatis* (Griffiths and Gupta, 2007), an organism for which a model system would be advantageous to support research independent of cell culture and biological containment. *V. spinosum*, in place, has high potential for use as a safe experimental model in research and development of antimicrobial compounds targeting *C. trachomatis* and related pathogens.

Because of its presence in medically and biotechnologically relevant organisms, DapL is an attractive target for antibiotic, herbicide, and algaecide development. A previous comprehensive screening analysis identified compounds with antibiotic potential that inhibit DapL from *Arabidopsis thaliana, Leptospira interrogans, Chlamydomonas reinhardtii,* and *Verrucomicrobium spinosum* (McKinnie et al., 2014). Four of these compounds (rhodanine, barbiturate, hydrazide, and thiobarbiturate), all of which are derived from classes with different structural elements, specifically inhibit the activity of *V. spinosum* DapL (*Vs*DapL). However, DapL crystal structures from *L. interrogans, C. reinhardtii,* and *V. spinosum* are either not published or do not exist, and the binding conformation of the effective compounds are not experimentally determined. Unfortunately, this scenario reflects a common situation in research settings where inhibitory compounds are screened against potential targets with only structural information inferred from a related species, resulting in unknown docking positions.

Informatics resources have been utilized in recent years to explore structure-guided drug design and structure-activity relationships (SAR), even in cases without experimentally determined structural information and in cases before *in vitro* experimentation. This method often involves the use of molecular docking to identify putative binding sites (Abdolmaleki et al., 2017), molecular dynamics to supplement and refine such docking (Iqbal and Shah, 2018), and/or subsequent SAR studies to predict the biological activity of the compound based on similar structures (Fan et al., 2010). However, most previous studies are limited in the scope of the molecular dynamics simulations performed, the size of the simulations, or size of the molecule analyzed. Adding in the often modeled structures further confounds results and requires post-processing and analysis.

FIGURE 1 | The Diaminopimelate pathways for lysine biosynthesis. **(A)** The synthesis of THDP from aspartate, **(B)** The acyl pathways, **(C)** the Ddh pathway, **(D)** the DapL pathway **(E)** Peptidoglycan crosslinking between disaccharide layers, GlcNAc-MurNAc, as mediated by meso-diaminopimelate/lysine and **(F)** The reversible transamination reaction catalyzed by DapL. DapC, acyl-diaminopimelate aminotransferase; DapE, acyl-diaminopimelate deacylase; DapD, 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate-N-acyltransferase; DapF, acyl-diaminopimelate deacylase, DapL, L,L-diaminopimelate aminotransferase; Ddh, diaminopimelate dehydrogenase; LysA, diaminopimelate decarboxylase.

Here, a comprehensive, comparative molecular dynamics (MD) simulation package, DROIDS (Detecting Relative Outlier Impacts in Dynamic Simulations 2.0) (Babbitt et al., 2018), was used in conjunction with SWISS-MODEL (Pettersen et al., 2004; Biasini et al., 2014) and AutoDock Vina (Trott and Olson, 2010) to investigate the binding dynamics of the identified putative inhibitory lead compounds and *Vs*DapL inferred from homology modeling. DROIDS calculates the local modes of interaction

between the residues, simulating inter- and intramolecular interactions, for two macromolecule sets over a time course for a defined number of replicates (Lewars, 2003; Jensen, 2007). It then compares the dynamics of the two macromolecules and presents an analysis of the differences in dynamic movement in the context of atomic fluctuations from the mean position. Here, we present putative docking positions of the four effective compounds tested and characterized via SAR and *in vitro*

analyses in previous work (Fan et al., 2010) and provide investigative MD simulation data supporting the structural inference. The methods and results presented here not only address the efficacy of these tools in a common scenario of investigative antibiotic development but also can be applied and modified to both supplement and provide a rational guide in laboratory method development.

METHODS

Multiple Sequence Alignment

Multiple sequence alignment was constructed using the Molecular Evolutionary Genetics Analysis (MEGA) (Kumar et al., 2016) tool with the DapL protein sequences from *V. spinosum* (NCBI Acc: WP_009961032.1), *A. thaliana* (UniProt: Q93ZN9), *C. trachomatis* (UniProt: G4NMX8), and *C. reinhardtii* (UniProt: A8IW39). Sequences were aligned via MUSCLE algorithm (Edgar, 2004). Conserved active site loops and residues were identified from the multiple sequence alignment, referencing those identified to interact with the natural ligand in the *A. thaliana* crystal structure and identified based on sequence homology between all four protein sequences.

Homology Modeling of *Vs*DapL

Three dimensional homology model was constructed via SWISS-MODEL (Biasini et al., 2014; Waterhouse et al., 2018) after alignment to the template L,L-diaminopimelate aminotransferase from *C. reinhardtii* (PDB 3QGU) with 53.3% sequence identity. The template was chosen as the crystal structure with the best sequence identity to the enzyme based on a basic local alignment search tool (*BLAST*) search with the HHBlits algorithm (Remmert et al., 2011) against the SWISS-MODEL template library (Guex et al., 2009; Benkert et al., 2011; Biasini et al., 2014; Bertoni et al., 2017; Waterhouse et al., 2018). The model was built based on the alignment between reference and query protein sequence. Conserved coordinates were directly copied while insertions and deletions were modeled from the SWISS-MODEL fragment library.
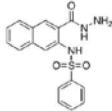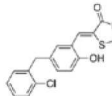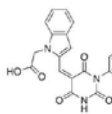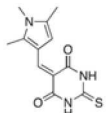
Structure Stabilization Testing

Structure stability and conformation testing was performed using the initial DROIDS pipeline development and prior to version 1 release. The solvated structures were each subjected to 1,000 cycles of minimization using a steepest-descent algorithm, followed by an 80-picosecond heating step where the temperature was gradually raised from 0 to 300K and then an 80-picosecond simulation allowing for the potential energy of the molecule to reach an equilibrium, followed by a 120-nanosecond sampling simulation. Atomic coordinates were calculated every 0.002 picoseconds and recorded every 0.2 picoseconds. Temperature and pressure were held constant for the equilibration and subsequent runs.

Compound Structure Generation and Docking

Lead compounds (McKinnie et al., 2014) (**Table 1**) were exported in the simplified molecular-input line entry system format from

TABLE 1 | Compounds identified by McKinnie et al. (2014) and the corresponding structures.

Compound	ChemBridge ID #	Chemical formula	Structure
Hydrazide	5925714	$C_{17} H_{15} N_3 O_3 S$	
Rhodanine	6523070	$C_{19} H_{16} Cl N O_2 S_2$	
Barbiturate	6072466	$C_{22} H_{17} N_3 O_5$	
Thiobarbiturate	6088649	$C_{12} H_{13} N_3 O_2 S$	

the ChemBridge database and the highest quality, lowest energy three dimensional coordinates were generated with the CORINA server (Sadowski et al., 1994; Schwab, 2010).

Three dimensional compounds were docked onto the DapL active site with the UCSF Chimera interface for Autodock Vina (Pettersen et al., 2004; Trott and Olson, 2010). The highest ranked conformation, with the best binding affinity reported, was chosen for each respective MD calculation set in DROIDS. For each run, the protein-ligand complex, the protein, and the ligand in its proper conformation identified by Vina were exported in PDB format. The resulting three files each were used.

Molecular Dynamics Preparation

All chemical artifacts in the structures were removed manually prior to running DROIDS. A structure-based sequence alignment was constructed in UCSF Chimera for both chains on each set of structures to be compared (in this case an apo-DapL structure and DapL docked with a ligand) and saved in clustal format. Amber16 and the AmberTools packages were used in ligand and complex preparation to generate atomic coordinate files for MD calculations. The program pdb4amber was used to remove all extraneous water molecules and add hydrogen atoms prior to renumbering residues for downstream analyses. The antechamber package was used to estimate and parameterize the force fields surrounding the ligand using the general Amber force field ("GAFF") simple harmonic function with AM1-BCC charge methods (Jakalian et al., 2000; Wang et al., 2004). Ligand coordinate files were generated from the antechamber ligand preparation. TeLEap was used to generate protein coordinate and topology files using the ff14SB protein force field parameters for both the apo-DapL and DapL-ligand complex (Maier et al., 2015). Atomic coordinates for both structures were utilized in subsequent MD calculations.

FIGURE 2 | Multiple protein sequence alignment of the *Vs*DapL to reference sequences. *Conserved active site residues; − − −, residues in loops lining the active site, and + denotes contribution from the opposing chain.
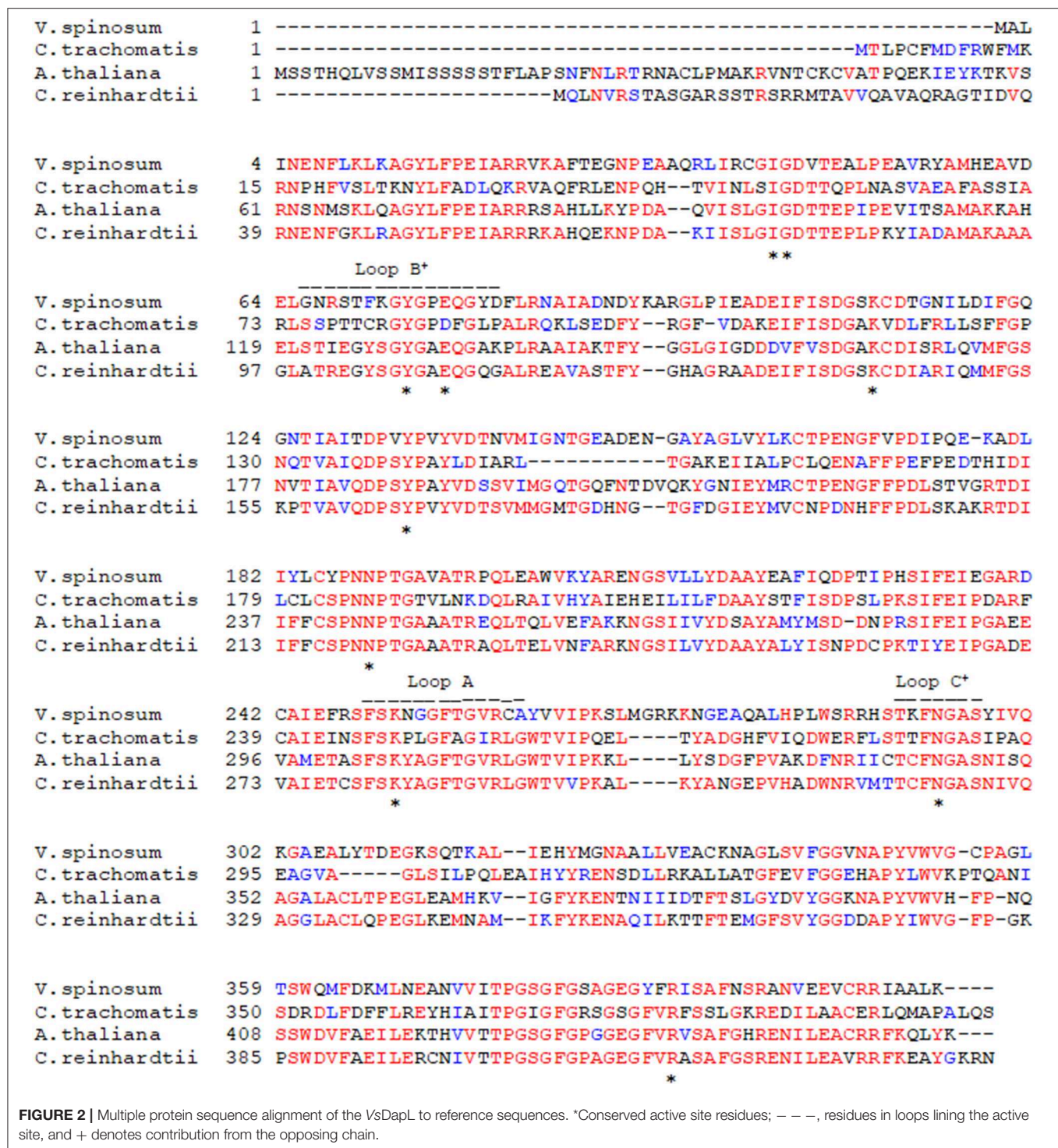
Molecular Dynamics Sampling

The solvated structures were each subject to 1,000 cycles of minimization using a steepest-descent algorithm, followed by a 100-picosecond heating step to 300 K, followed by a 10-nanosecond equilibration step to allow for the potential energy of the molecule to reach an equilibrium before 300 production sampling runs of 0.5 nanoseconds. Starting coordinates were randomly generated by the DROIDS software at the onset of each sampling run. Temperature and pressure were held constant for the equilibration and subsequent runs.
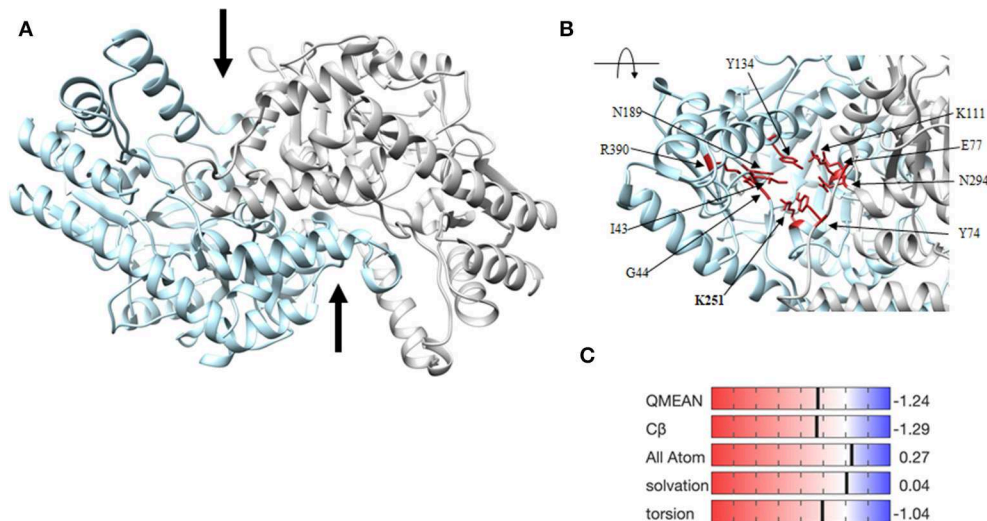
FIGURE 3 | Homology model of the *Vs*DapL enzyme. **(A)** Three dimensional ribbon model depicting the predicted structure, a homodimer (gray and blue). The two black arrows denote the predicted active sites. **(B)** Ribbon model of the rotated, centered *Vs*DapL active site showing active site residues (red) identified from multiple sequence alignment. **(C)** Z-score data when the model is compared to high quality X-ray crystal structures of similar size.
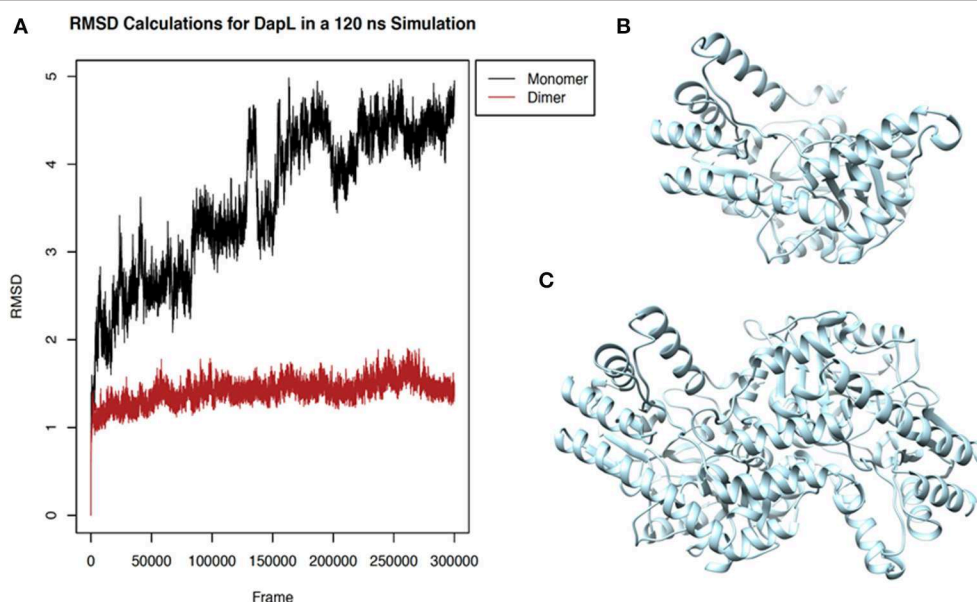
FIGURE 4 | Stability of *Vs*DapL comparing monomer and dimer. **(A)** Root Mean Square Deviation (RMSD) values for the fluctuation of the alpha-carbons of each amino acid from the mean position of the backbone calculated for the molecular dynamics sampling runs of the **(B)** *Vs*DapL monomer and **(C)** *Vs* DapL homodimer.

Molecular Dynamics Post-processing

The CPPTRAJ package provided in Ambertools was used to generate trajectories for each residue and calculate Root Mean Square Deviation (RMSD) at an amino acid resolution. The calculated atom fluctuation profiles were used to calculate signed symmetric Kullback–Leibler (KL) divergences in local atom fluctuation distributions from each amino acid on the polypeptide backbone, and *P*-values from a Benjamini–Hochberg corrected Kolmogorov-Smirnov (KS) test. All post-processing

analyses were performed in R and presented in graphical representation. UCSF Chimera was used for structure depiction.

RESULTS

Key Amino Acids of Known DapL Active Sites Are Conserved in *V. spinosum* DapL

To identify key active site amino acid residues in the DapL ortholog from *V. spinosum* (*Vs*DapL), we constructed a

multiple sequence alignment using the sequences of DapL orthologs where structures have already been experimentally determined [*V. spinosum* (PDB: WP_09961032.1), *A. thaliana* (UniProt: Q93ZN9), C. trachomatis (UniProt: G4NMX8), and C. reinhardtii (UniProt: A8IW39)]. The key residues in the active site were highly conserved across all organisms. Loops that line the active site in *V. spinosum* were predicted to reside between F249 and A261 (Loop A), as well as those from the opposing chain between residues G66 and D81 (Loop B), and T291 and S297 (Loop C). In addition, key conserved active site residues were identified as I43, G44, Y74, E77, K111, Y134, N189, K251, N294, R390 (**Figures 2**, **3B**) (McKinnie et al., 2014).

The three-dimensional homology model predicts a homodimer with two v-shaped clefts that correspond to the putative active sites (**Figure 3A**) and the active sites agree with the alignment findings above. The model quality, assessed with z-scores for the QMEAN function (−1.24), all atom pairwise energy (0.27), CBeta interaction energy (−1.29), solvation energy (0.04), and torsion angle energy (−1.04), indicates an acceptable structure prediction (**Figure 3C**).

DapL Is More Stable *in silico* as a Homodimer

To determine structural stability, RMSD calculations of the fluctuation of the alpha-carbons of each amino acid from their mean positions were used to compare the homodimer and monomer forms of *Vs*DapL. As a monomer, the early sampling of *Vs*DapL coordinates generate a RMSD value of 2 angstroms. Over time, however, the RMSD increases and nears 5 angstroms. On the contrary, the *Vs*DapL homodimer stabilizes just above a RMSD of 1 angstrom for the duration of the sampling time (**Figure 4**). The higher and consistently rising RMSD values resulting from the monomer dynamics indicate more instability in the system, while the constant lower RMSD values from the dimer indicate a much more stable structure. Overall, the dimer

TABLE 2 | The binding affinity for and residues that interact with each identified compound used in this study.

Compound	Binding affinity (Kd) (kcal/mol)	Interacting residues
Hydrazide	−7.6	43*, 74*, 111*, 110, 134*, 189*, 217, 250, 251**, 259, 294*
Rhodanine	−8.3	12, 15, 16, 135, 138, 156, 161, 361, 379
Barbiturate	−7.6	17, 18, 21, 77*, 78, 287, 290
Thiobarbiturate	−6.4	91, 95, 221, 231, 247, 316, 317

*Denotes residues that are associated with the active site. **Denotes catalytic residue.
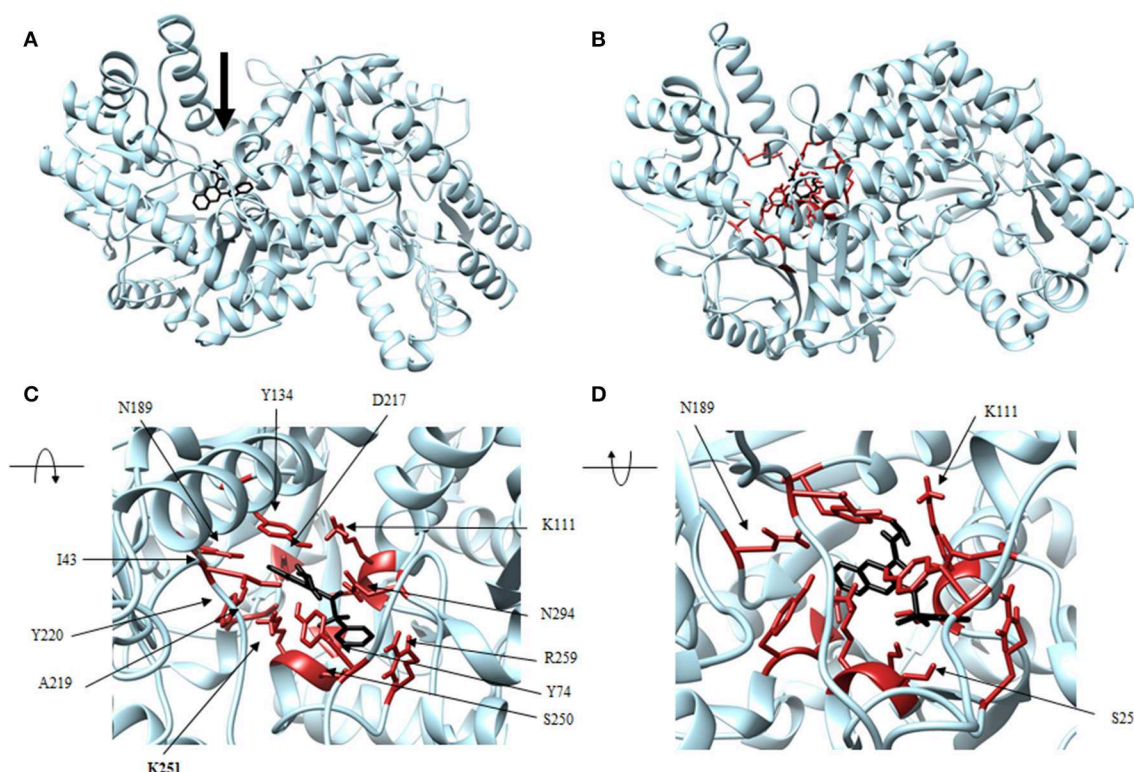
FIGURE 5 | Structure of hydrazide (black) associated with *Vs*DapL (blue). **(A)** Ribbon conformation showing binding pocket (black arrow). **(B)** Ribbon conformation showing interacting residues (red). **(C)** Rotated, centered binding pocket showing interacting residues (red, labeled). **(D)** Rotated, centered binding pocket, denoted by arrow, with two labeled interacting residues (red) for orientation.
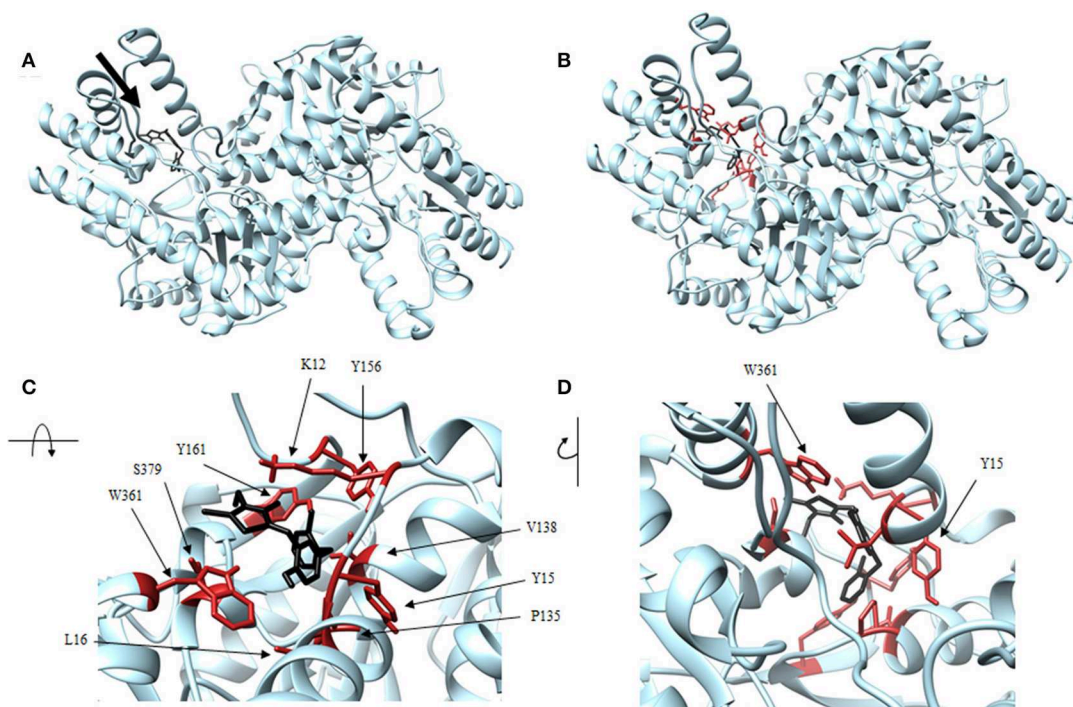
FIGURE 6 | Structure of rhodanine (black) associated with *Vs*DapL (blue). **(A)** Ribbon conformation showing binding pocket (black arrow). **(B)** Ribbon conformation showing interacting residues (red). **(C)** Rotated, centered binding pocket showing interacting residues (red, labeled). **(D)** Rotated, centered binding pocket, denoted by arrow, with two labeled interacting residues (red) for orientation.
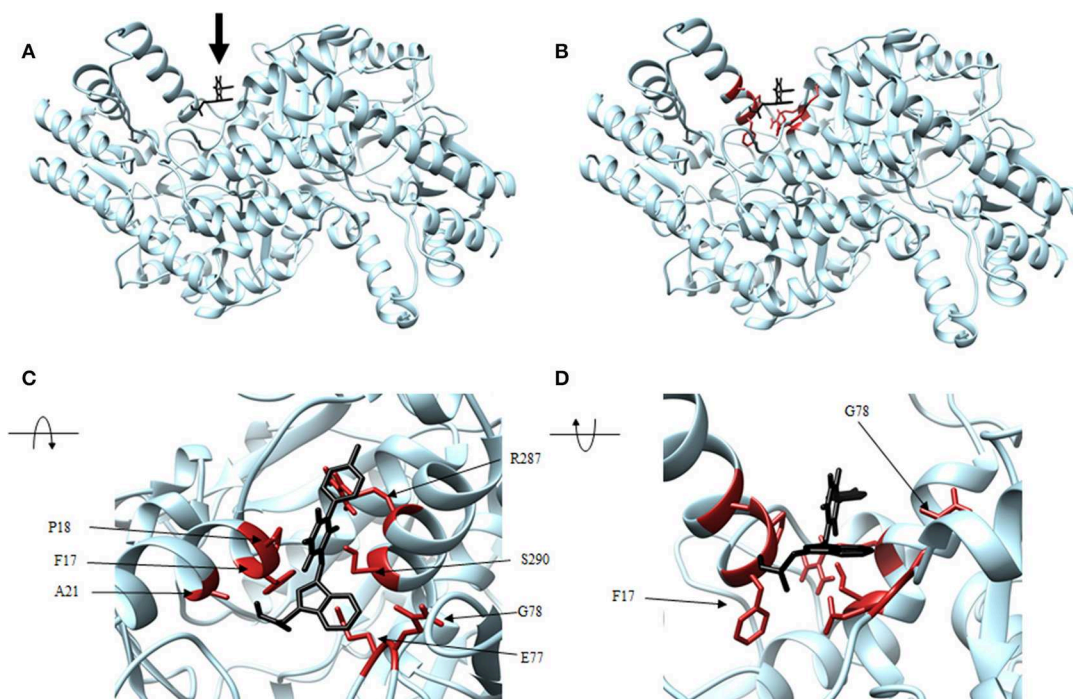
FIGURE 7 | Structure of barbiturate (black) associated with *Vs*DapL (blue). **(A)** Ribbon conformation showing binding pocket (black arrow). **(B)** Ribbon conformation showing interacting residues (red). **(C)** Rotated, centered binding pocket showing interacting residues (red, labeled). **(D)** Rotated, centered binding pocket, denoted by arrow, with two labeled interacting residues (red) for orientation.
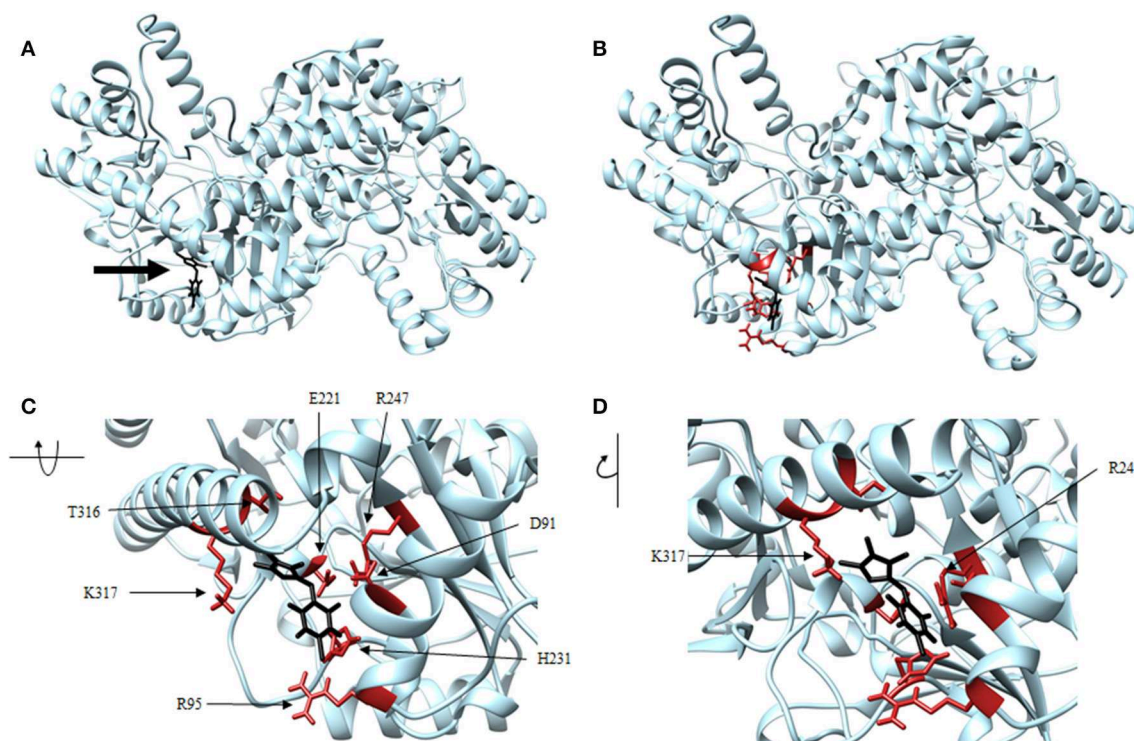
FIGURE 8 | Structure of thiobarbiturate (black) associated with *Vs*DapL (blue). **(A)** Ribbon conformation showing binding pocket (black arrow). **(B)** Ribbon conformation showing interacting residues (red). **(C)** Rotated, centered binding pocket showing interacting residues (red, labeled). **(D)** Rotated, centered binding pocket, denoted by arrow, with two labeled interacting residues (red) for orientation.

settled into a lowest-energy conformation and maintained that general pattern for the duration of the simulation, as opposed to a constantly fluctuating set of conformations in the monomer. This is also consistent with previous findings, where other DapL orthologs are homodimers in solution.

Rhodanine, Barbiturate, Hydrazide, and Thiobarbiturate Association With *Vs*DapL Is Supported by Molecular Dynamics Simulations

Four antagonist lead compounds (rhodanine, barbiturate, hydrazide, and thiobarbiturate) were docked into the *Vs*DapL homology model to represent modeling the molecular dynamics of macromolecular complexes without actual structural data. All antagonistic compounds successfully associated with the *Vs*DapL active site with a potential energy lower than the enzyme in its native state and with all potential energies at the same order of magnitude, but with varying degrees of affinity (−6.5 to −7.6 kcal/mol, data not shown) as indicated by the RMSD values when compared to the best identified conformation. The best docking position in the structure also varied for each compound (**Table 2**). The hydrazide molecule bound close to the active site (**Figure 5**), the rhodanine molecule docked in the minor arm of one subunit (**Figure 6**), the barbiturate molecule docked on the outer edge of the structure near an active site (**Figure 7**), and

the thiobarbiturate molecule docked on the outer surface of the protein (**Figure 8**).

In addition, identified contacts support that the four small molecule antagonists bound to different locations on DapL. Hydrazide bound to the majority of residues in the active site, including the catalytic lysine (K251). Rhodanine and thiobarbiturate did not bind to any residues that compose active site, and barbiturate bound near some active site residues but not directly in the active site (**Table 2**).

Molecular dynamics of each *Vs*DapL-antagonist macromolecular complex was compared to the unbound *Vs*DapL molecule in terms of differences in fluctuations for 500 replicate simulations with random initialization coordinates. The differences in fluctuation were then color-mapped to *Vs*DapL at an amino-acid resolution (**Figure 9**). Differences in fluctuation were found between the unbound molecule and bound complex for all four cases. The dynamics of most, if not all, *Vs*DapL residues that interacted with the compound in the pocket were stabilized upon association. In only two instances (**Figures 9B,C**) were the dynamics of a single residue amplified upon compound association. However, both residues are near loops and on the exterior of the protein, so some instability would be expected even upon settling into a bound conformation. Additionally, the color-mapping of the true *Vs*DapL active sites show stabilization upon compound association. This color mapping is supported by plots of the average fluctuation of each residue from the mean position of both bound and unbound structures, the signed KL
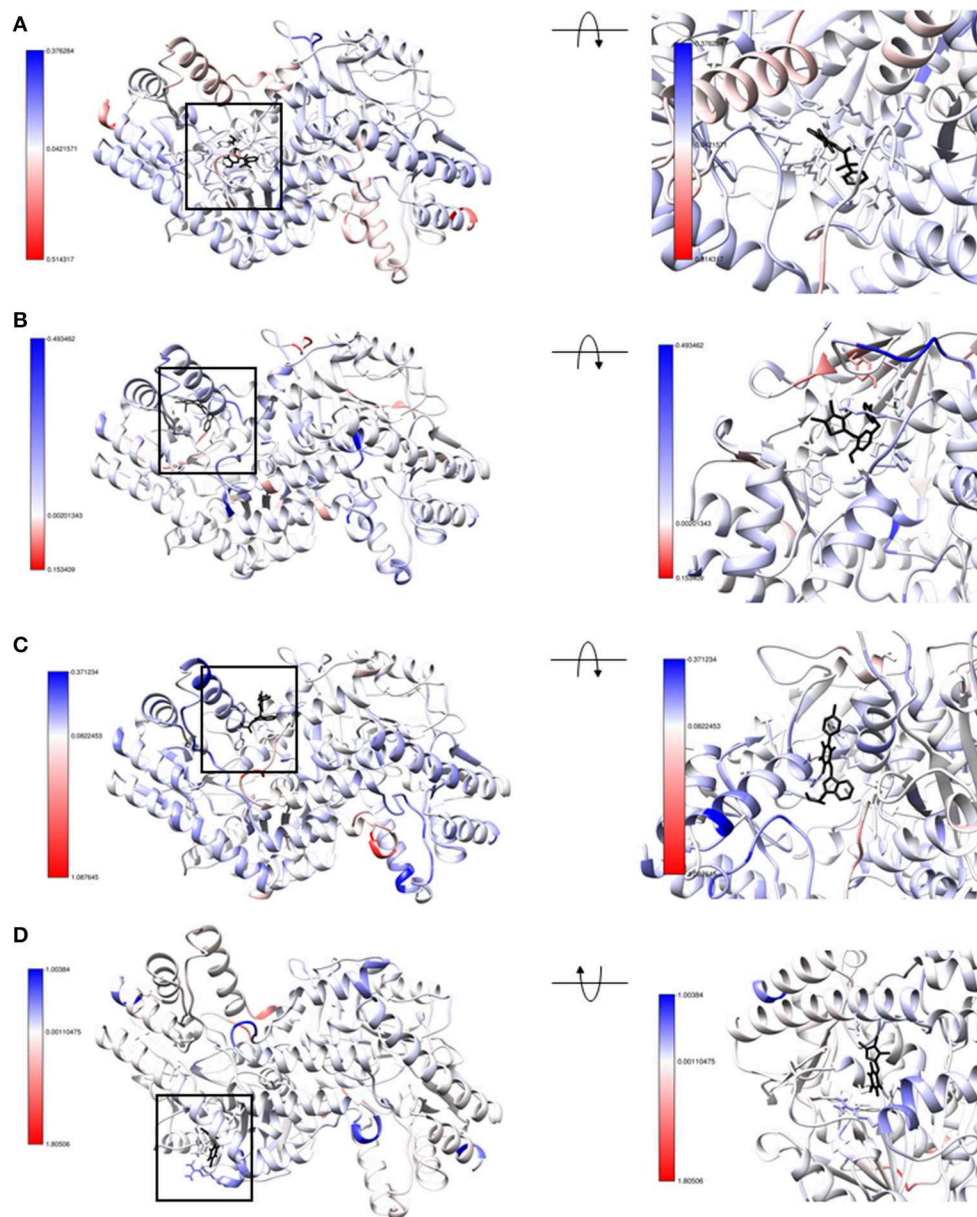
FIGURE 9 | Structure of **(A)** hydrazide **(B)** rhodanine **(C)** barbiturate and **(D)** thiobarbiturate associated with *Vs*DapL colored by signed symmetric KL divergences in local atom fluctuation distributions of each amino acid on the polypeptide backbone. The small molecule inhibitors are colored in black. The KL divergences shown are color-mapped to temperature on the protein with red indicating softer regions with amplified fluctuation and blue indicating stabilized regions with dampened fluctuation. The magnitude of divergence is shown in the color key beside the structure. (left) front-facing view with a black box to direct attention to the region (right) centered and rotated to show the interacting residues.

divergence of the fluctuations, and significance based on a KS test (**Supplemental Figure 1**). Overall, the local stabilization upon compound association indicates the binding positions identified with the docking software are accurate.

DISCUSSION

The *Vs*DapL homodimer is more stable than the monomer, and DapL is naturally crystallized as a homodimer. Based on the sample size and volume of data collected, combined with experimentally determined crystal structures of multiple DapL enzymes, we must emphasize that all DapL computational analyses (docking, molecular dynamics, etc.) should involve the homodimer. Conclusions drawn from calculations on one subunit are not biologically accurate. Because DROIDS is equipped to perform dynamics calculations on large, multi-subunit proteins and map the statistics at an amino acid resolution, it provided

us with a unique opportunity for DapL and other dimer dynamic analyses.

All antagonistic compounds in this study have the potential to interact with DapL *in vivo*, a conclusion supported by previous work (Fan et al., 2010; McKinnie et al., 2014). In our analyses, compounds were only docked into one active site because of the conformational change associated with the DapL mechanism of action. However, because all DapL structures characterized to date are indeed homodimers, we can speculate that the inhibitory lead compounds have the capacity to associate in at least two sites. In fact, a test docking of two rhodanine molecules onto DapL indicates better binding affinity for both together than the single molecule. A kinetic analysis of the docking pattern of these molecules would provide additional support for this hypothesis. However, docking protocols are not as robust as molecular dynamics simulations, nor do they flawlessly represent biological conditions.

Additionally, molecular dynamics simulations require precise and refined calculations that take into account extremely complex physical characteristics, and as a caveat they must be cautiously analyzed for robustness and consistency. For example, molecules must be in their lowest potential energy wells for consistent and meaningful calculations. This task becomes exponentially more difficult as the size of the molecule increases. If the atomic coordinates are calculated for longer duration, or sampled for a different number of replicates, it is possible that the refined, amino-acid resolution results would differ. However, the trends reported from our simulations, of stabilization in both the binding pocket and more loosely at the true active site of the enzyme, provide strong support for the putative binding positions of the compounds. Because the inhibition potential of the four small molecules was previously determined, we were able to draw conclusions based on the consistency in the results of the molecular dynamics simulations. Indeed, all simulations supported the previously drawn experimental conclusions. If the biological effect of these compounds was not known, SAR analyses could be performed with the top compounds identified in the docking and molecular dynamics screening. This would provide additional insights as to the potential biological effect of the compounds and the effect of modifications to these compounds to further guide future work. The efficacy of this study provides rationale for using any iteration of this work (modeling, docking, or molecular dynamics) as a springboard to suggest future experimental work in the rational design of experimental antibiotic compounds.

In conclusion, this and previous studies have shown that DapL is a potential target for the development of narrow spectrum antibiotics, and *Vs*DapL has high potential for the development of a model system for DapL from related pathogenic organisms. Additionally, this study supports the great potential scope of the application of molecular dynamics to biological molecules to infer function, form, and effect of interactions in situations where the crystal structure may not be known. The success of this exploratory study in identifying the putative binding locations of four inhibitory compounds only touches the surface of the capabilities of molecular dynamics software. If applied correctly, molecular dynamics software can provide direction for major questions in the biological sciences across disciplines.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

This work was conceived by AH and LA. The experiments were conducted by LA, PR, and JM. AH, LA, GB, RN, and RD wrote the manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00046/full#supplementary-material

Supplemental Figure 1 | Average atom fluctuation profiles (top plot), signed symmetric KL divergences in local atom fluctuation distributions of each amino acid on the polypeptide backbone (middle plot), and *P*-values from a Benjamini–Hochberg corrected KS test indicating significant differences in dynamics is also shown (bottom plot) for **(A)** hydrazide **(B)** rhodanine **(C)** barbiturate and **(D)** thiobarbiturate.

REFERENCES

Abdolmaleki, A., Ghasemi, J. B., and Ghasemi, F. (2017). Computer aided drug design for multi-target drug design: SAR /QSAR, molecular docking and pharmacophore methods. *Curr. Drug Targets* 18, 556–575. doi: 10.2174/1389450117666160101120822

Adams, L. E. (2019). *Genetics and Molecular Dynamics of L, L-Diaminopimelate Aminotransferase (DapL): an Enzyme Involved in Lysine and Peptidoglycan Biosynthesis.* Available online at: https://scholarworks.rit.edu/theses/10052 (accessed November 05, 2019).

Ansdell, V. (2012). "70 - Leptospirosis," in *Netter's Infectious Diseases*, eds E. C. Jong, and D. L. Stevens (Philadelphia, PA: W.B. Saunders), 425–429. doi: 10.1016/B978-1-4377-0126-5.00070-7

Babbitt, G. A., Mortensen, J. S., Coppola, E. E., Adams, L. E., and Liao, J. K. (2018). DROIDS 1.20: a GUI-based pipeline for GPU-accelerated comparative protein dynamics. *Biophys. J.* 114, 1009–1017. doi: 10.1016/j.bpj.2018.01.020

Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27, 343–350. doi: 10.1093/bioinformatics/btq662

Berry, C. D., Hooton, T. M., Collier, A. C., and Lukehart, S. A. (1987). Neurologic relapse after benzathine penicillin therapy for secondary syphilis in a patient with HIV infection. *N. Engl. J. Med.* 316, 1587–1589. doi: 10.1056/NEJM198706183162507

Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., and Schwede, T. (2017). Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* 7:10480. doi: 10.1038/s41598-017-09654-8

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42, W252–W258. doi: 10.1093/nar/gku340

Burstain, J. M., Grimprel, E., Lukehart, S. A., Norgard, M. V., and Radolf, J. D. (1991). Sensitive detection of *Treponema pallidum* by using the polymerase chain reaction. *J. Clin. Microbiol.* 29, 62–69.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

Fan, C., Clay, M. D., Deyholos, M. K., and Vederas, J. C. (2010). Exploration of inhibitors for diaminopimelate aminotransferase. *Bioorg. Med. Chem.* 18, 2141–2151. doi: 10.1016/j.bmc.2010.02.001

Finegold, S., and Sussman, M. (2002). "Anaerobic infections a clinical overview," in *Molecular Medical Microbiology*, (London, UK: Elsevier), 1867–1874.

Griffiths, E., and Gupta, R. S. (2007). Phylogeny and shared conserved inserts in proteins provide evidence that Verrucomicrobia are the closest known free-living relatives of chlamydiae. *Microbiology* 153, 2648–2654. doi: 10.1099/mic.0.2007/009118-0

Guex, N., Peitsch, M. C., and Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-Pdbviewer: a historical perspective. *Electrophoresis* 30(Suppl. 1), S162–S173. doi: 10.1002/elps.200900140

Hudson, A. O., Bless, C., Macedo, P., Chatterjee, S. P., Singh, B. K., Gilvarg, C., et al. (2005). Biosynthesis of lysine in plants: evidence for a variant of the known bacterial pathways. *Biochim. Biophys. Acta* 1721, 27–36. doi: 10.1016/j.bbagen.2004.09.008

Hudson, A. O., Singh, B. K., Leustek, T., and Gilvarg, C. (2006). An ll-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol.* 140, 292–301. doi: 10.1104/pp.105.072629

Hutton, C. A., Perugini, M. A., and Gerrard, J. A. (2007). Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Mol. Biosyst.* 3, 458–465. doi: 10.1039/B705624A

Iqbal, J., and Shah, S. J. A. (2018). Molecular dynamic simulations reveal structural insights into substrate and inhibitor binding modes and functionality of ecto-nucleoside triphosphate diphosphohydrolases. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-20971-4

Jakalian, A., Bush, B. L., Jack, D. B., and Bayly, C. I. (2000). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. method. *J. Comput. Chem.* 21, 132–146. doi: 10.1002/jcc.10128

Jensen, F. (2007). *Introduction to Computational Chemistry, 2nd Edn*, Chichester; Hoboken, NJ: John Wiley & Sons.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lewars, E. (2003). *Computational chemistry - introduction to the theory and applications of molecular and quantum mechanics. Comput. Chem.* 482. doi: 10.1007/978-90-481-3862-3

Liepman, A. H., and Olsen, L. J. (2004). Genomic analysis of aminotransferases in *Arabidopsis Thaliana. Crit. Rev. Plant Sci.* 23, 73–89. doi: 10.1080/07352680490273419

Lindsay, K. W., Bone, I., and Fuller, G. (2010). "Section V - multifocal neurological disease and its management," in *Neurology and Neurosurgery Illustrated, 5th Edn*, eds K. W. Lindsay, I. Bone, and G. Fuller (London, UK: Content Repository Only), 489–563

Mabey, D. (2008). "Trachoma: recent developments," *in Hot Topics in Infection and Immunity in Children IV Advances in Experimental Medicine and Biology*, eds A. Finn, and A. J. Pollard (New York, NY: Springer), 98–107.

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. doi: 10.1021/acs.jctc.5b00255

McKinnie, S. M. K., Rodriguez-Lopez, E. M., Vederas, J. C., Crowther, J. M., Suzuki, H., Dobson, R. C. J., et al. (2014). Differential response of orthologous l,l-diaminopimelate aminotransferases (DapL) to enzyme inhibitory antibiotic lead compounds. *Bioorg. Med. Chem.* 22, 523–530. doi: 10.1016/j.bmc.2013.10.055

Mishori, R., McClaskey, E. L., and WinklerPrins, V. (2012). *Chlamydia Trachomatis* infections: screening, diagnosis, and management. *Am. Fam. Physician* 86, 1127–1132.

Nishida, H., Nishiyama, M., Kobashi, N., Kosuge, T., Hoshino, T., and Yamane, H. (1999). A prokaryotic gene cluster involved in synthesis of lysine through the amino adipate pathway: a key to the evolution of amino acid biosynthesis. *Genome Res.* 9, 1175–1183.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2011). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* 9, 173–175. doi: 10.1038/nmeth.1818

Sadowski, J., Gasteiger, J., and Klebe, G. (1994). Comparison of automatic three-dimensional model builders using 639 X-ray structures. *J. Chem. Inf. Comput. Sci.* 34, 1000–1008. doi: 10.1021/ci00020a039

Schwab, C. H. (2010). Conformations and 3D pharmacophore searching. *Drug Discov. Today Technol.* 7, e245–e253. doi: 10.1016/j.ddtec.2010.10.003

Triassi, A. J., Wheatley, M. S., Savka, M. A., Gan, H. M., Dobson, R. C. J., and Hudson, A. O. (2014). L, L-diaminopimelate aminotransferase (DapL): a putative target for the development of narrow-spectrum antibacterial compounds. *Front. Microbiol.* 5:509. doi: 10.3389/fmicb.2014.00509

Trott, O., and Olson, A. J. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* 31, 455–461. doi: 10.1002/jcc.21334

Velasco, A. M., Leguina, J. I., and Lazcano, A. (2002). Molecular evolution of the lysine biosynthetic pathways. *J. Mol. Evol.* 55, 445–449. doi: 10.1007/s00239-002-2340-2

Velick, S. F., and Vavra, J. (1962). A kinetic and equilibrium analysis of the glutamic oxaloacetate transaminase mechanism. *J. Biol. Chem.* 237, 2109–2122.

Ventola, C. L. (2015). The Antibiotic resistance crisis. 40, 277–283.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174. doi: 10.1002/jcc.20035

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Wexler, H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* 20, 593–621. doi: 10.1128/CMR.00008-07

Repurposing Modular Polyketide Synthases and Non-ribosomal Peptide Synthetases for Novel Chemical Biosynthesis

Soonkyu Hwang[1], Namil Lee[1], Suhyung Cho[1], Bernhard Palsson[2,3,4] and Byung-Kwan Cho[1,5]*

[1] Systems and Synthetic Biology Laboratory, Department of Biological Sciences and KI for the BioCentury, Korea Advanced Institute of Science and Technology, Daejeon, South Korea, [2] Department of Bioengineering, University of California, San Diego, La Jolla, CA, United States, [3] Department of Pediatrics, University of California, San Diego, La Jolla, CA, United States, [4] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, [5] Intelligent Synthetic Biology Center, Daejeon, South Korea

In nature, various enzymes govern diverse biochemical reactions through their specific three-dimensional structures, which have been harnessed to produce many useful bioactive compounds including clinical agents and commodity chemicals. Polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) are particularly unique multifunctional enzymes that display modular organization. Individual modules incorporate their own specific substrates and collaborate to assemble complex polyketides or non-ribosomal polypeptides in a linear fashion. Due to the modular properties of PKSs and NRPSs, they have been attractive rational engineering targets for novel chemical production through the predictable modification of each moiety of the complex chemical through engineering of the cognate module. Thus, individual reactions of each module could be separated as a retro-biosynthetic biopart and repurposed to new biosynthetic pathways for the production of biofuels or commodity chemicals. Despite these potentials, repurposing attempts have often failed owing to impaired catalytic activity or the production of unintended products due to incompatible protein–protein interactions between the modules and structural perturbation of the enzyme. Recent advances in the structural, computational, and synthetic tools provide more opportunities for successful repurposing. In this review, we focused on the representative strategies and examples for the repurposing of modular PKSs and NRPSs, along with their advantages and current limitations. Thereafter, synthetic biology tools and perspectives were suggested for potential further advancement, including the rational and large-scale high-throughput approaches. Ultimately, the potential diverse reactions from modular PKSs and NRPSs would be leveraged to expand the reservoir of useful chemicals.

Keywords: polyketide synthase, non-ribosomal peptide synthetase, domain, module, repurposing

INTRODUCTION

Enzymes are biosynthetic protein machineries that recognize specific substrates through unique three-dimensional structures, and catalyze the conversion of these substrates into new biomolecules (Agarwal, 2006). Harnessing their diverse biochemical reactions has led to the production of many bioactive-compounds as clinical agents and commodity chemicals (Tibrewal and Tang, 2014). In addition, engineering such enzymes and repurposing their reactions into new pathways enhances the biocatalytic properties and the diversity of the natural products, respectively (Tibrewal and Tang, 2014). Biosynthesis of organic compounds has several advantages compared to classical chemical synthesis methods (Wenda et al., 2011). First, enzymes are not environmentally harmful, they act as non-toxic catalysts. The reaction conditions for the production of diverse chemicals are generally moderate in terms of temperature, pressure, and pH, while the classical chemical synthesis often requires extreme conditions. High selectivity of enzymes yields high purities with specific stereochemistry of the product and reduce undesired by-products and toxic intermediates. In nature, interestingly, the mechanisms underlying a large number of enzyme reactions have not been discovered yet. For example, the recent genome mining efforts on bacteria, fungi, and plants have revealed the richness of secondary metabolite biosynthetic gene clusters (smBGCs) including many unidentified smBGCs (Rutledge and Challis, 2015). Thus, novel non-natural chemicals and pathways have been constructed by the reprogramming of the multi-enzyme complex encoded by smBGCs (Bernhardt and O'Connor, 2009; Chevrette et al., 2019).

Type I modular polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) are prominent engineering targets due to their modular properties of enzyme assembly (Ladner and Williams, 2016). Type I modular PKS comprises several modules, each responsible for the incorporation and modification of one acyl-CoA substrate to synthesize the polyketide product, such as erythromycin (antibiotic), rapamycin (immunosuppressant), amphotericin B (antifungal), and other potential clinical agents (Staunton and Weissman, 2001). Likewise, NRPS comprises numerous modules, with each of them responsible for the incorporation and modification of one amino acid substrate to extend the polypeptide product, such as daptomycin (antibiotic), actinomycin D (antitumor), cyclosporine A (immunosuppressant), and other potential clinical agents (Walsh, 2016; Sussmuth and Mainz, 2017). As the organization and the order of these modules are co-linearly correlated with each unit of the final polyketide and polypeptide products, the target modules of rational engineering could be predicted for the production of novel derivatives having the modified unit at a specific position (Alanjary et al., 2019). Using this modular biosynthetic logic of type I modular PKS and NRPS, each reaction of a module was capable of being separated from the original assembly reactions and repurposed to construct de novo biosynthetic pathway for other chemicals (Pang et al., 2019). This approach is favorable in terms of (i) the potential diversity

of available synthetic parts governing unique chemical reactions, (ii) enabling retro-biosynthesis by combinatorial assembly of domains and modules, and (iii) the relative ease of engineering, owing to avoidance of the structural perturbation compared to the engineering within the multi-enzyme complex.

In this review, we briefly introduced the structure and mechanism of the modular PKS and NRPS, and thereafter focused on their repurposing examples, along with their advantages and limitations. Finally, tools in the design-build-test-learn cycle of synthetic biology and the future perspectives of the repurposing strategies were discussed.

MODULAR PKS AND NRPS ARCHITECTURE AND MECHANISM

Polyketide synthases are categorized into three types, namely, types I, II, and III, according to their organization and catalytic mechanisms (Yu et al., 2012). Among them, type I modular PKS has a hierarchical organization in which, the entire enzyme complex is composed of several subunits, each subunit is composed of several modules, and a module is composed of several domains (Bayly and Yadav, 2017) (**Figure 1A**). A minimal elongation module includes three domains; (i) an acyltransferase (AT) domain for loading the chain extender unit (typically malonyl- or methylmalonyl-CoA), (ii) an acyl carrier protein (ACP) for tethering and shuttling the extender unit or the polyketide intermediate, and (iii) a ketosynthase (KS) domain for catalyzing the condensation reaction between the extender unit of the downstream ACP domain and the polyketide intermediate attached at the KS active site which is translocated from the ACP domain of upstream module. Addition of other domains to this minimal elongation module modifies a polyketide backbone.

In the N- to C-terminus of a whole enzyme complex, loading, elongation, and termination modules are localized to catalyze the serial polyketide production (Barajas et al., 2017). Loading module (LM) initiates the chain formation from a broad range of priming starter units by acylation to the AT domain and transacylation to the ACP domain. According to the configuration of the domains, LMs are divided into subtype A, B, C, and D (Kornfuehrer and Eustáquio, 2019). In addition to AT-ACP didomain, type A LMs involve the condensation-incompetent KS domain that decarboxylates malonyl- or methylmalonyl-CoA to yield acetyl- or propionyl starter units, respectively. Type B LM consist of only the AT-ACP didomain and has a much broader range of substrates. Type C LM has a CoA-ligase domain rather than AT domain to incorporate carboxylic acid substrate in an ATP-dependent manner. Lastly, type D LM has a GCN5 N-acetyltransferase-like domain rather than AT domain that recently repurposed to catalyze decarboxylation (Skiba et al., 2020). A common characteristic of LMs is the absence of the condensation domain, which results a more flexible substrate specificity of the AT domain than those of downstream modules. Elongation modules consist of KS, AT, ACP, and other additional domains to serially catalyze the growth of the polyketide chain (Barajas et al., 2017; Zargar et al., 2018; Kornfuehrer and Eustáquio, 2019)
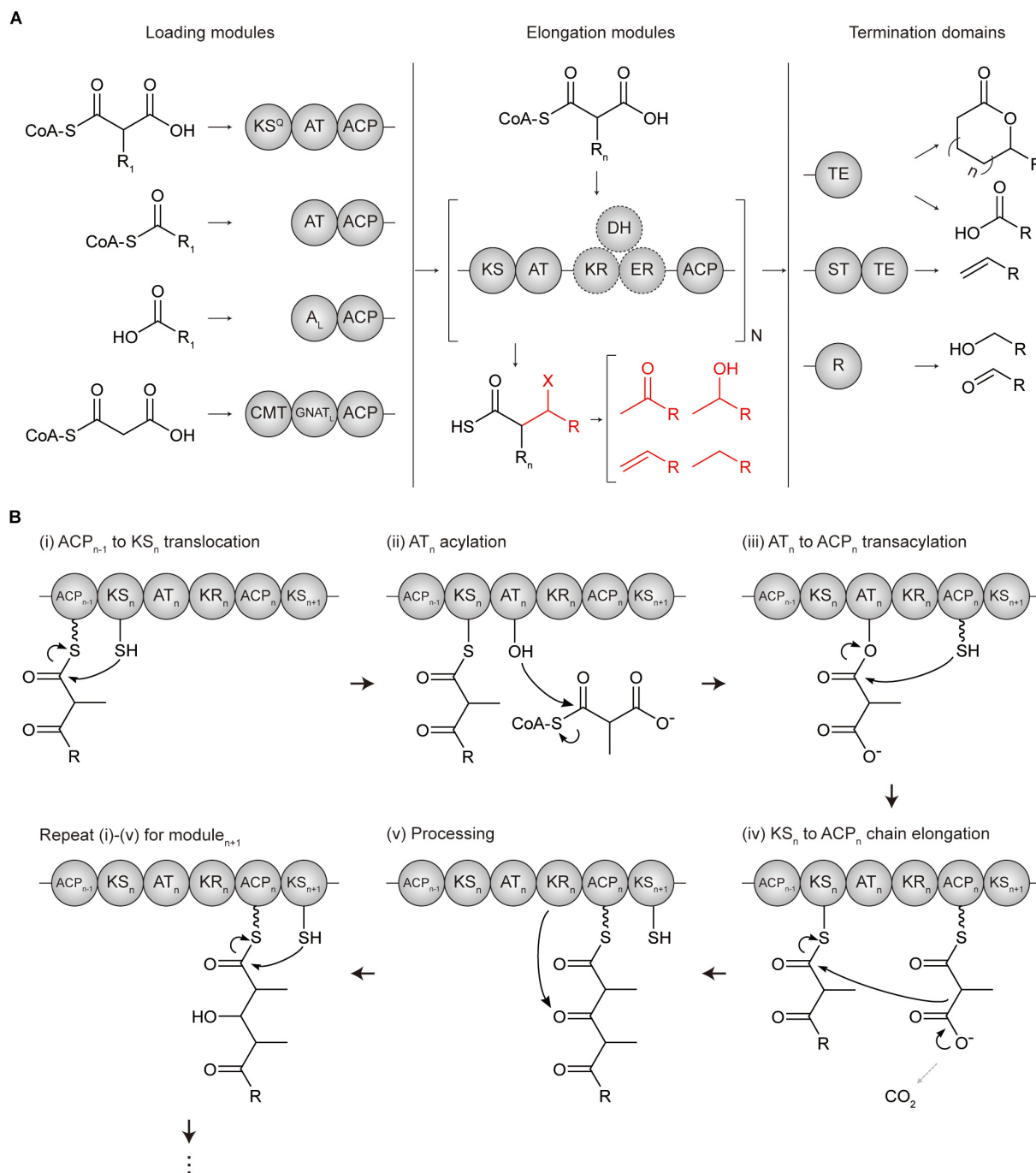
FIGURE 1 | Domain architectures and mechanisms of polyketide chain extension in modular PKS. **(A)** Overall flow scheme of polyketide biosynthesis with different domain architectures of modules. Four types of loading modules load the different substrates according to involved domains (chemical examples were indicated). Next, the extender unit is selected and condensed to the growing chain one by one per elongation module for N cycles. Optional reductive domains (dashed circles) reduce the β-carbon group resulting in different X groups (indicated in red). Finally, the growing polyketide chain is cleaved by three different types of offloading domains in termination modules producing different products, including linear carboxylic acids, macrocyclic acids, olefins, aldehydes, and primary alcohols. **(B)** Mechanism of polyketide chain extension for the elongation module$_n$. (i) ACP$_{n-1}$ to KS$_n$ translocation; the active site cysteine moiety of KS$_n$ receives the growing polyketide chain of ACP$_{n-1}$. (ii) AT$_n$ acylation; the cognate acyl unit is incorporated into the active site serine moiety of ATn to form the acyl-O-AT intermediate. (iii) AT$_n$ to ACP$_n$ transacylation; the acyl group of AT$_n$ is transacylated to the ACP$_n$. (iv) KS$_n$ to ACP$_n$ chain elongation; KS$_n$ catalyzes a decarboxylative Claisen condensation between the growing polyketide chain and the acyl extender unit of ACP$_n$ for the chain extension. (v) Processing; the extender units of ACP$_n$ are modified by a reductive loop or other additional domains. ACP, acyl carrier protein; A$_L$, CoA ligase-type domain; AT, acyltransferase; CMT, C-methyltransferase; GNAT$_L$, GCN5 N-acetyltransferase-like domain; DH, dehydratase; ER, enoylreductase; KR, ketoreductase; KS, ketosynthase; KS^Q, condensation-incompetent ketosynthase; R, reductive domain; ST, sulfotransferase; TE, thioesterase.

(**Figure 1B**). Termination module contains a thioesterase (TE) domain or reductive (R) domain as an offloading domain following the ACP domain to release the ACP-bound polyketide intermediate and complete polyketide biosynthesis. TE domain performs; (i) hydrolysis to yield the linear polyketide, or (ii) cyclization to yield the macrocyclic polyketide, or (iii) sulfonate transfer and decarboxylation with sulfotransferase (ST) domain to yield the terminal olefin. Alternatively, R domain catalyzes NADPH dependent two-electron reduction, yielding the aldehyde product.

Non-ribosomal peptide synthetases are categorized into two types according to their organizations and catalytic mechanisms, which are type I modular NRPS and type II standalone NRPS (Jaremko et al., 2019). Type I modular NRPS a resemblance to type I modular PKS, with a hierarchical organization of the enzyme complex-subunit-module-domain (Winn et al., 2016) (**Figure 2A**). A minimal elongation module includes three domains which are, an adenylation (A) domain for loading the amino acid extender unit (both proteinogenic and non-proteinogenic), an thiolation (T or PCP) domain for tethering and shuttling the extender unit transferred from the upstream A domain or the growing polypeptide transferred from T domain of the upstream module, and a condensation (C) domain for catalyzing the peptide bond formation between the amino acid extender unit of the downstream T domain and the growing polypeptide of the T domain of the upstream module. Other *cis*-acting and *trans*-acting domains to this minimal elongation module expand the extender amino acid unit.

Similar to PKS, loading, elongation, and termination module are localized to catalyze serial polypeptide production (Sussmuth and Mainz, 2017). NRPS LMs usually lack a C domain and are more variable than PKS. Many NRPS LMs carry the N-terminal modifications, which function as protection against degradation, modulating polarity, and providing specific properties such as membrane insertion. Several additional domains including formylating domain, CoA ligase domain, and other tailoring domains could be involved in LMs for acylation and formylation, among others. NRPS elongation module basically consists of three core domains C, A, and T corresponding to KS, AT, and ACP of PKS which play a role in chain elongation, substrate incorporation, and chain carrier function, respectively (Hur et al., 2012; Winn et al., 2016; Bloudoff and Schmeing, 2017; Sussmuth and Mainz, 2017) (**Figure 2B**). However, their overall structures and mechanisms are quite different, and there are also more various and distinct processing domains for NRPS. Examples include; epimerase (E) domain catalyzing the absolute configuration at the C_α atom, methyltransferse (MT) domain modifying the degree of C- or N-methylation, formylation (F) domain, cyclization (Cy), redox-active domain (Ox and Red), among others. These steps would be repeated between module$_n$ and module$_{n+1}$ till the termination module. Termination module of NRPS also contains an offloading domain such as a TE domain or reductive (R) domain after T domain to release the T-bound polypeptide intermediate for the completion of NRPS biosynthesis. The TE domain hydrolyzes or cyclize the intermediate to form a linear or cyclic polypeptide. Additionally, terminal C_T domain in fungal

NRPSs disconnects the oligopeptide by macrocyclization, and R* domain mediates Dieckmann-type cyclization of PK-NRP hybrids to obtain tetramate moieties (Sussmuth and Mainz, 2017). Diversity of NRPS is largely present in the type II NRPSs that consist of standalone or minimal domains, which have been reviewed in detail in other papers (Jaremko et al., 2019; McErlean et al., 2019).

Due to the collinear and modular biosynthetic logic, the structural diversity of the products of modular PKS and NRPS could be largely attributed to a few variables (Kornfuehrer and Eustáquio, 2019). The modules of PKS and NRPS are suitable devices for retro-biosynthesis, which starts from the target product and proceeds backward to precursors by stepwise combination of the independent module reactions (Pang et al., 2019).

MODULAR PKS AND NRPS ENGINEERING

As various strategies have been employed for the engineering of modular PKS and NRPS (Brown et al., 2018; Klaus and Grininger, 2018), we provided several landmark examples (**Figure 3**).

Substrate Exchange

Early attempts for the engineering of PKSs and NRPSs were conducted via precursor directed biosynthesis (PDB) and mutasynthesis (Alanjary et al., 2019) (**Table 1**). PDB is a strategy of creating a new product by providing an external substrate instead of the native substrate of the recognition domain (AT and A domain). This strategy is based on the native promiscuity of the recognition domains that are able to incorporate more than one kind of substrate. The advantages of this strategy are that the native enzyme could be free of any genetic engineering, and understanding of the structures and mechanisms of the enzymes is not required. This is because the promiscuity is easily confirmed by adding the target substrates and then measuring the kinetic profiles or the product formation. Despite this, native substrate and alternative substrates compete for the enzyme reaction that results in the formation of a mixture of major and minor products, hampering the yield and purity of the desired product. One of the successful examples of PDB was the incorporation of a series of 21 substrates consisting of monocyclic, polycyclic, branched aliphatic acids, benzoic acids, and heterocyclic acids to AT domain in loading module of rapamycin PKS, suggested by the previous report of relaxed substrate specificity of the domain (Lowden et al., 2004). Among the novel rapamycin analogs, one monocyclic aliphatic acid product has immunosuppressant activity comparable to the native product. In another example, native substrate L-Pro was altered by non-proteinogenic amino acids for pyreudione NRPS to generate various derivatives from pyreudione E to pyreudione K (Klapper et al., 2018). Although there were more examples for successful PDB, most of their products showed lower yield compared to the native products (Ward et al., 2007; Moran et al., 2009; Crawford et al., 2011; Xie et al., 2014).
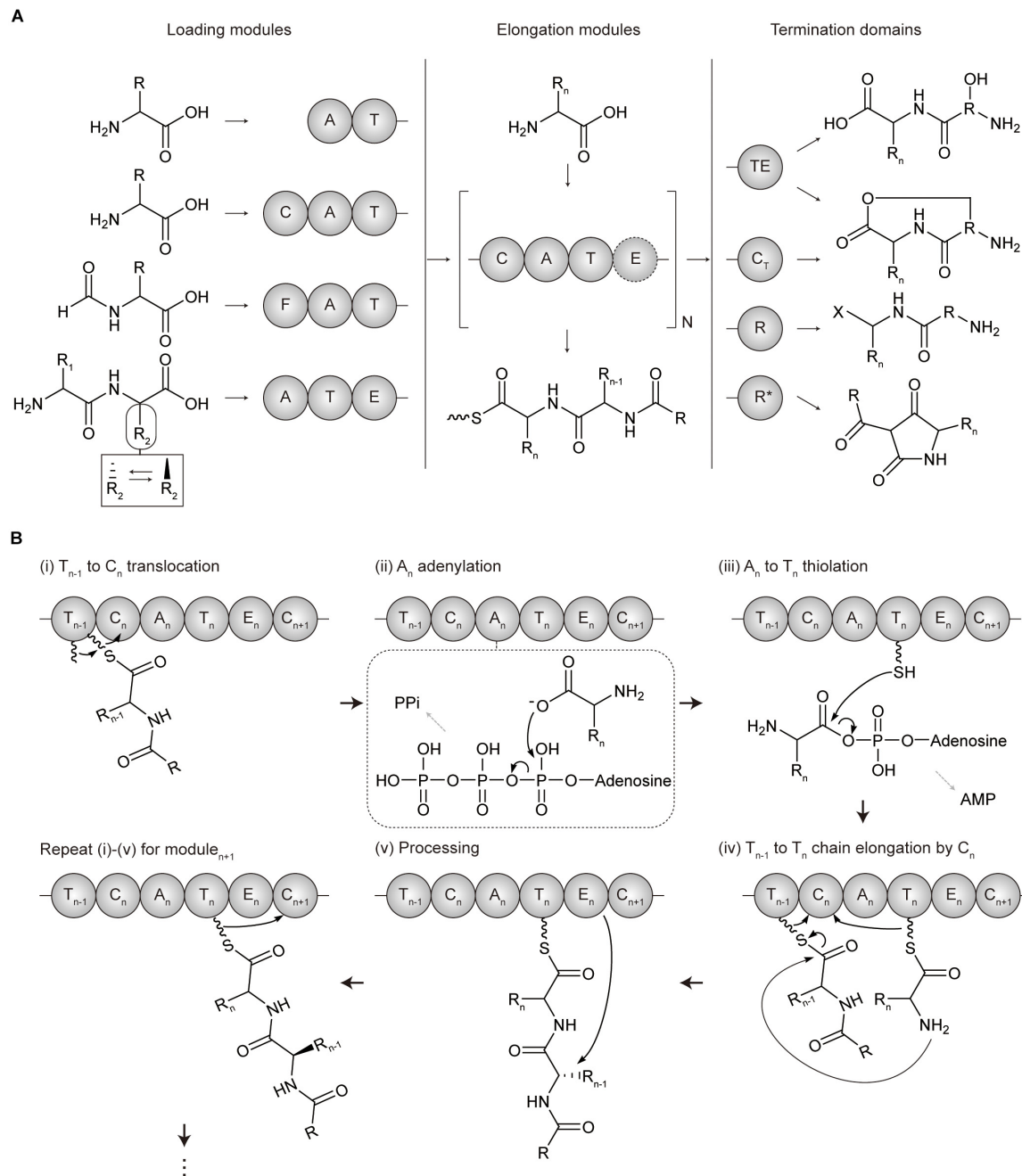
FIGURE 2 | Domain architectures and mechanisms of non-ribosomal peptide chain extension in modular NRPS. **(A)** Overall flow scheme of non-ribosomal peptide biosynthesis with different domain architectures of modules. Four representative types of loading modules load the different substrates according to the involved domains (chemical examples were indicated). Next, the extender unit is selected and condensed to the growing chain one by one per elongation module for N cycles. An example of optional processing domain is indicated by the dashed circles. Finally, the growing non-ribosomal peptide chain is cleaved by four representative types of offloading domains in termination modules, producing different products including linear peptides, macrocyclic peptides, aldehydes, and tetramate moieties. The terminal X group of the product from terminal R domain includes hydroxyl group (-OH), aldehyde group (-CHO), and other aldehyde derivatives (Barajas et al., 2015; Dan et al., 2019). **(B)** Mechanism of polyketide chain extension for the elongation module$_n$. (i) T_{n-1} to C_n translocation; the growing non-ribosomal peptide chain linked to Ppant arm of T_{n-1} domain translocates to the solvent channel of C_n domain donor site. (ii) A_n adenylation; the extender amino acid unit is activated by ATP to form aminoacyl-AMP in A_n domain. (iii) A_n to T_n thiolation; the aminoacyl-AMP intermediate of A_n is transferred to the Ppant arm of T_n domain to form aminoacyl thioester intermediate. (iv) T_{n-1} to T_n condensation at C_n; the aminoacyl thioester intermediate of T_n domain is translocated to the solvent channel of C_n domain acceptor site, and the peptide bond formation between the growing peptide of T_{n-1} domain and the amino acid extender unit of T_n domain elongates by adding one amino acid to the growing peptide. (v) Processing; the extender units of ACP$_n$ are modified by an epimerase (E) domain or other additional domains. A, adenylation domain; C, condensation domain; C_T, terminal condensation domain; F, formylating domain; R, reductive domain; R*, R-like domain; T, thiolation domain; TE, thioesterase.
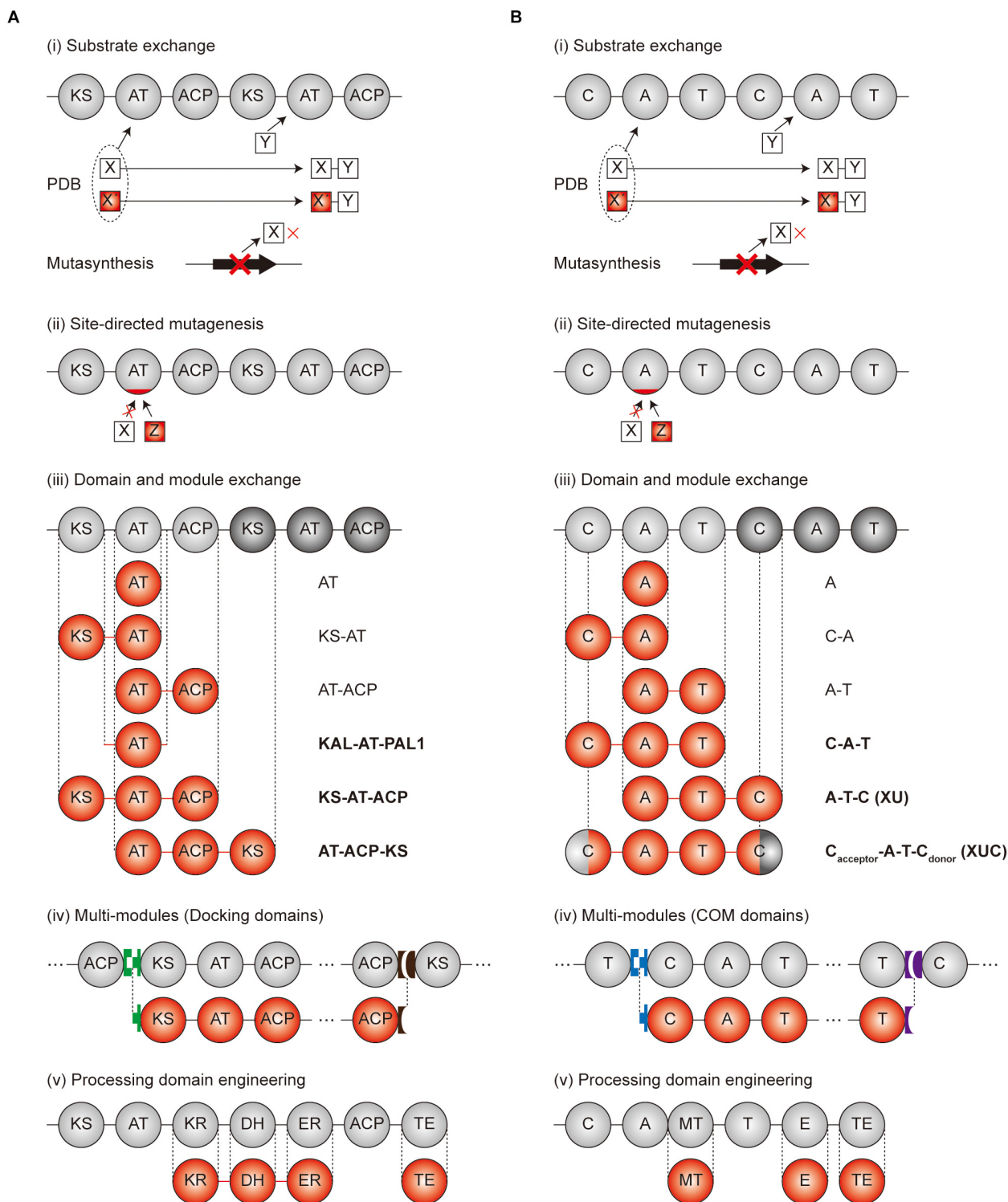
FIGURE 3 | Engineering scheme for modular PKS and NRPS. (A) Engineering strategies of PKSs. **(B)** Engineering strategies of NRPSs. Gray circles and red circles indicate the original and modified domains, respectively. Green, brown, blue, and purple blocks, shaped as lock-and-key models, are the docking domains for **(A)** and COM domains for **(B)**, respectively. Linkers were indicated as the lines between the domains. In case of (iii) domain and module exchange, the exchangeable units are indicated at the right of the domains. The units indicated as bold characters are currently the best exchangeable units. PDB, precursor-directed biosynthesis.

TABLE 1 | Selected examples of the substrate exchange strategy.

Category	Target	Engineering strategy (substrate change/genetic manipulation)	Product	References
Precursor directed biosynthesis	6-deoxyerythronolide B PKS	Native precursor propyl-CoA change to diketide precursor (SNPCs)/DEBS LM deletion	15-fluoroethyl-6 deoxyerythronolide B	Ward et al., 2007
Precursor directed biosynthesis	Rapamycin PKS	Native precursor 4,5-dihydroxycyclohex-1-enecarboxylic acid change to 21 analogs	Monocyclic aliphatic acids	Lowden et al., 2004
Precursor directed biosynthesis	Pyreudione NRPS	Native precursor L-Pro change to prolin derivatives	Pyreudione E to K	Klapper et al., 2018
Precursor directed biosynthesis	Iturin A NRPS	Native precursor L-Tyr change to 3-fluoro-L-Tyr	Fluorinated iturin A	Moran et al., 2009
Mutasynthesis	Geldanamycin PKS	Native precursor 3-amino-5-hydroxybenzoic acid change to 18 analogs/3-amino-5-hydroxybenzoic acid (AHBA) biosynthetic gene deletion	Geldanamycin variants	Eichner et al., 2009
Mutasynthesis	DEBS1-soraphen hybrid PKS	Native precursor benzoate change to *p*-fluoro, *p*-hydroxy, *m*-hydroxybenzoate/*badA* deletion with *encN* insertion	Triketide lactone variants	Garcia-Bernardo et al., 2004
Mutasynthesis	FK506 PKS	Native precursor allylmalonyl CoA change to 4-methylpentanoic acid/*tcsB* deletion	36-methyl-FK506	Mo et al., 2011
Mutasynthesis	Balhimycin NRPS	Native precursor β-OH-Tyr change to 3-Fht/*bhp* deletion	Fluorobalhimycin	Weist et al., 2002
Mutasynthesis	Balhimycin NRPS	Native precursor dihydroxyphenylglycine change to hydroxylated or methoxylated phenylglycines/*dpgA* deletion	Hydroxylated or methoxylated balhimycin	Weist et al., 2004
Mutasynthesis	Salinosporamide NRPS	Native precursor 5′-CIDA change to 5′-FDA/*salL* deletion	Fluorosalinosporamide	Eustaquio and Moore, 2008

To overcome the competition between native and altered substrates, mutasynthesis approaches were considered to increase the yield and purity of the desired product by feeding the alternative substrates together with the deletion of the biosynthetic genes of the native substrates in the expression host. For example, the mutasynthesis of novel FK506 derivatives was reported by the deletion of *tcsB*, resulting in modification of the FK506 C21 moiety (Mo et al., 2011). Utilizing the native promiscuity of AT domain of module 4 in FK506 PKS, the feeding of *trans*-2-hexenoic acid, 4-methylpentanoic acid, and 4-fluorocrotonic acid generated 36,37-dihydro-37-methyl-FK506, 36- methyl-FK506, and 36-fluoro-FK520, respectively. Interestingly, 36-methyl-FK506 not only has immunosuppressant function but has also improved neurite outgrowth activity. Other examples obtained various derivatives through incorporation of the non-natural units by mutasynthesis (Weist et al., 2002, 2004; Garcia-Bernardo et al., 2004; Eustaquio and Moore, 2008; Eichner et al., 2009). However, the requirement of genetic engineering and the dependency of native promiscuity are limitations of mutasynthesis. Despite these limitations, mutasynthesis is still useful when applied together with domain and module engineering in a synergistic manner. For example, non-natural precursors generated by semisynthetic or click chemistry could be fed, accompanied by the deletion of the native precursor biosynthetic genes and the mutations of PKS or NRPS enzymes to attach the pharmacophore-containing moiety or dye for novel chemical production (Alanjary et al., 2019).

Mutagenesis of Substrate Recognition Domains

The AT domain of PKS or the A domain of the NRPS have been major mutation targets for enzyme engineering (**Table 2**). Two directions of mutagenesis have been applied; (i) reducing or increasing the native promiscuity of the domain and (ii) creating *de novo* specificity of the domain (Musiol-Kroll and Wohlleben, 2018; Stanisic and Kries, 2019). Compared to domain and module exchange (see subsection "Domain and Module Exchange"), the modification of several residues in the AT or A domain could minimize the structural perturbation of entire enzyme assemblies, as well as minimally affecting the protein–protein interactions between adjacent domains and modules (Brown et al., 2018). In addition, this approach fundamentally changes the protein-substrate interaction enabling the incorporation of non-natural synthetic substrates.

The substrate specificity of the AT domain is determined by approximately 100 residues toward the C-terminus from the active site serine (Musiol-Kroll and Wohlleben, 2018). The most common substrates of elongation AT domains are malonyl- or methylmalonyl CoA (Chan et al., 2009; Barajas et al., 2017),

TABLE 2 | Selected examples of the substrate recognition domain mutagenesis.

Target	Engineering strategy (substrate change (mutation))	Rationale for engineering	Product	References
AT domain in module 4 of DEBS PKS	Methylmalonyl-CoA (YASH) to malonyl-CoA (HAFH)	Domain sequence alignment and structure-based selection	6-desmethyl-6-dEB	Reeves et al., 2001
AT domain in module 3 of epothilone PKS	Both methylmalonyl, malonyl-CoA (HASH) to methylmalonyl-CoA (YASH) or malonyl-CoA (HAFH)	Domain sequence alignment	Triketide lactones derivatives	Petkovic et al., 2008
AT domain in module 2 of DEBS PKS	Methylmalonyl-CoA (YASH) to non-natural propargyl, ethyl, allylmalonyl (VASH)	Domain sequence alignment and structure-based selection	Triketide lactones derivatives	Vogeli et al., 2018
AT domain in module 6 of DEBS PKS	Methylmalonyl-CoA to diverse bulky extender units (Q,Y,S to A,G,R)	Structure-based selection	Propargyl, heptenyl, benzyl-SNAC incorporation	Li et al., 2018
AT domain in module 5,6 of pikromycin PKS	Methylmalonyl-CoA to propargyl, ethyl, allyl, butylmalonyl CoA (Y755V, Y753V)	Domain sequence alignment and Homology modeling	10-dML analogs, narbonolide analogs	Kalkreuter et al., 2019
AT domain in module 1 of avermectin PKS	40 carboxylic acids to (2S)-methylbutyric acid (V222L or V222A)	Homology modeling	(2S)-methylbutyric acid (isolated module)	Wang et al., 2015
AT domain in module 6 of DEBS PKS	Methylmalonyl-CoA (YASH) to non-natural alkynyl-modified extender unit (RASH)	Domain sequence alignment and saturation mutagenesis	10-dML analogs	Koryakina et al., 2017
A domain of module 2 of surfactin NRPS	L-Asp to L-Asn (V299I, H322E, I330V)	Domain sequence alignment and Homology modeling	[Asn5]surfactin	Eppelmann et al., 2002
A domain of module 10 of CDA NRPS	L-Glu, Me-Glu to L-Gln, Me-Gln (K278G, Q236G)	Domain sequence alignment	CDA4a-10mQ	Thirlway et al., 2012
A domain of module 3 of Fusaricidins NRPS	L-Tyr, L-Val, L-Ile, L-(allo)-Ile, or L-Phe to only L-Phe (three mutants of three sites)	Domain sequence alignment	Fusaricidin analog (LI-F07)	Han et al., 2012
A domain of module 1 of Anabaenopeptin NRPS	L-Arg, L-Tyr to 4-azido-Phe (E204G,S243E or S243H)	Structure-based selection	Novel anabaenopeptin analog, clickable	Kaljunen et al., 2015
A domain of module 1 of gramicidin S NRPS	L-Phe to L-Tyr, O-propargyl-L-Tyr (W239S)	Structure-based selection and saturation mutagenesis	Propargylated DKP, clickable	Kries et al., 2014
A domain of module 1 of syringomycin NRPS	A domain of L-Ser-specifying EntF exchange to X-specifying SyrE. None to L-Ser (3.2 amino acid change per clone)	Domain sequence alignment, directed evolution (2 rounds), and iron growth selection	Enterobactin derivatives	Fischbach et al., 2007
A domain of AdmK of andrimid NRPS/PKS hybrid	L-Val to L-Ile, L-Leu, L-Ala, L-Phe (three sites saturation mutagenesis)	Domain sequence alignment, directed evolution, and high-throughput LC-MS/MS	Andrimid derivatives	Evans et al., 2011

(Continued)

TABLE 2 | Continued

Target	Engineering strategy (substrate change (mutation))	Rationale for engineering	Product	References
A domain of module 1 of tyrocidine NRPS	L-Phe to L-Thr (A301C, C331I, I3330V, W239M)	Domain sequence alignment, directed evolution, and PPi exchange assay	Tyrocidine derivatives	Villiers and Hollfelder, 2011
A domain of module 1 of bacillibactin NRPS	2,3-dihydroxybenzoic acid to 3-hydroxybenzoic acid and 2-aminobenzoic acid (four sites mutations)	Domain sequence alignment, directed evolution, and yeast surface display with FACS	3-hydroxybenzoic acid and 2-aminobenzoic acid (isolated module)	Zhang et al., 2013

while at least 20 malonyl-CoA analogs have been found to be incorporated (Yuzawa et al., 2017b). The substrates of the loading AT domain are more variable (Barajas et al., 2017). Sequence alignment analysis of prototypical 6-deoxyerythronolide B synthase (DEBS) and epothilone PKS suggested the specificity code is 'HAFH' for malonyl-CoA, 'YASH' for methylmalonyl-CoA, and 'HASH' for the both (Reeves et al., 2001; Petkovic et al., 2008). Changing the code results in altered specificity to corresponding substrates, although there may be the mixed substrate specificity for both substrates. The specific residues for changing the promiscuous specificity toward non-natural extender units were determined based on the domain sequence alignment, X-ray crystal structure, and homology modeling (Reeves et al., 2001; Petkovic et al., 2008; Li et al., 2018; Vogeli et al., 2018; Kalkreuter et al., 2019).

The specificity conferring code of the A domain in NRPS is more variable than PKS. Early studies about X-ray structure and sequence alignment suggested the eight specificity codes, referred to as Stachelhaus codes, which were later updated to broader range of residues by computational modeling (Conti et al., 1997; Stachelhaus et al., 1999; Challis et al., 2000; Rottig et al., 2011). Based on this, the A domain in the second module of surfactin NRPS was mutated at three of the eight specificity conferring residues (V299I, H322Q, and I330V), resulting in the successful L-Asp to L-Asn specificity change (Eppelmann et al., 2002). Moreover, the corresponding code of A domain in the third module of fusaricidins NRPS was determined by aligning its residues to those of the L-Phe specific A domain of gramicidin S and tyrocidine S NRPS. These mutations resulted in increased specificity toward L-Phe to produce the fusaricidin analog (LI-F07) (Han et al., 2012). Other examples are presented in **Table 2**.

The limitations of the site-directed mutagenesis approach are the effects of the other residues, outside the specificity conferring code. For instance, the specialized protein–protein interactions for the different PKSs and NRPSs such as KS-AT and C-A interfaces impeded the establishment of the universal code and the accurate prediction of specificity. This is also consistent with the frequent results of the unexpected mixed substrate specificity by AT or A domain mutagenesis. To overcome this, the directed evolution method that identifies the desired clone among the random mutagenesis libraries of 10^4 to 10^6 clones by iterative selection cycles and high-throughput screening has

been reported. The saturation mutagenesis of three sites in the AT domain of andrimid PKS generated a library of approximately 14,000 mutants, which were analyzed via the highly sensitive LC-MS/MS screening method (Evans et al., 2011). Consequently, mixed derivatives of adrimid containing L-Ile, L-Leu, L-Ala, and L-Phe, modified from L-Val, respectively, were obtained with improved antimicrobial activity. Together with other examples, the directed evolution approach was expected to become the universal strategy for the specificity change regardless of the kind of the enzymes (Fischbach et al., 2007; Villiers and Hollfelder, 2011; Zhang et al., 2013). However, there are still several requirements for feasible directed evolutions such as high-throughput assays, iterative cycles of mutagenesis for stability and selectivity, and effective screening. Alternatively, reducing the library size by the prediction of the residues for specificity change based on structural information and evolutionary evidence would be developed.

Domain and Module Exchange

Parallel to the site-directed mutagenesis approach, the domain or module exchange approaches have been applied to change the substrate specificity of PKS and NRPS (Sussmuth and Mainz, 2017; Kornfuehrer and Eustáquio, 2019). As the exchanging domains and modules are generally well-studied, the substrate specificity change is predictable and experimental design is more convenient. However, the correct splicing site should be precisely determined to preserve the protein–protein interactions between the acceptor and donor units, to minimize the perturbation of sophisticated conformational changes during chain elongation, and to maintain the overall structure of PKS and NRPS.

Domain or Module Exchange of PKS

In the case of PKS, successful examples were observed for the AT domain of the well-studied DEBS PKS exchange to other AT domains in different modules of DEBS PKS, or the modules from other PKSs (Oliynyk et al., 1996; Stassi et al., 1998; Hans et al., 2003; Petkovic et al., 2003) (**Table 3**). This AT exchange strategy is normally referred as 'AT domain swapping.' The splicing sites were determined as the rough boundaries of AT domains inferred by the sequence alignment between similar modules, and the AT domains were cloned by providing synonymous mutations to the boundaries for the introduction of restriction enzyme sites (Patel et al., 2004). Nevertheless, the

TABLE 3 | Selected examples of domain or module exchange of modular PKS.

Exchange unit	Target	Engineering strategy (substrate change/donor unit/fusion point)	Rationale for engineering	Product	References
AT-PAL1	AT domain in module 1 of DEBS PKS	Methylmalonyl to malonyl CoA specific/rapamycin AT domain in module 2 exchange/N,C-term RE site splicing	Domain sequence alignment	Two novel triketide lactones	Oliynyk et al., 1996
AT-PAL1-PAL2	AT domain in module 6 of DEBS PKS	Methylmalonyl to methylmalonyl or malonyl CoA specific/RAPS AT2, DEBS AT4, DEBS AT5 exchange/interdomain region RE site splicing	Domain sequence alignment	Triketide lactones derivatives	Hans et al., 2003
AT-PAL1	AT domain in module 4 of DEBS PKS	Methylmalonyl to malonyl CoA specific/rapamycin AT domain in module 2 exchange/N,C-term RE site splicing	Domain sequence alignment	6-desmethyl erythromycin D	Petkovic et al., 2003
AT-PAL1	AT domain in module 4 of DEBS PKS	Methylmalonyl to ethylmalonyl CoA specific/niddamycin AT domain in module 5 exchange/N,C-term RE site splicing	Domain sequence alignment	6-desmethyl-6-ethylerythromycin A	Stassi et al., 1998
AT-PAL1-PAL2	AT domain in module 1-5,7 of geldanamycin PKS	Methylmalonyl or methoxymalonyl to malonyl CoA specific/rapamycin AT domain in module 2,14 exchange/N,C-term RE site splicing	Domain sequence alignment	Geldanamycin derivatives	Patel et al., 2004
AT-ACP	LM (AT-ACP) of DEBS PKS	Methylmalonyl or malonyl CoA to 40 carboxylic acids specific/AVES LM (AT-ACP) exchange/ACP C-term region splicing	Domain sequence alignment	Novel antibiotic erythromycins	Marsden et al., 1998
KS^Q-AT-ACP	LM (KS^Q-AT-ACP) of tylactone PKS	Malonyl to Methylmalonyl CoA specific/Platenolide LM (KS^Q-AT-ACP) exchange/ACP-KS RE site splicing, synthetic linker connect	Domain sequence alignment	16-methyl platenollde I	Kuhstoss et al., 1996
KAL-AT-PAL1	AT domain in module 6 of DEBS PKS, AT domain in module 1 of β-lipomycin PKS	Methylmalonyl to methylmalonyl or malonyl CoA specific/epothilone AT domain in module 4 exchange, other various AT domains exchange/KAL-AT-PAL1	Domain sequence alignment and structure-based selection	3-hydroxycarboxylic Acid, short-chain ketones	Yuzawa et al., 2017b
KS-AT-KR-ACP	Insertion between module 1,2 (KS-AT-KR-ACP) of DEBS PKS	Methylmalonyl, malonyl CoA specific added/rapamycin module 2,5 (KS-AT-KR-ACP, AT-KR-ACP-KS) insertion/KS N-term, C-term RE site splicing	Domain sequence alignment	Novel octaketide macrolactones	Rowe et al., 2001
KS-AT-ACP, multi-modules	Modules of DEBS PKS	Same specificity/M1-M3, M1-M6, M1-RifM5, M2 to RifM5/Native RE site, conserved region of interpolypeptide linker	Domain sequence alignment	Triketide lactones derivatives	Gokhale et al., 1999

(Continued)

TABLE 3 | Continued

Exchange unit	Target	Engineering strategy (substrate change/donor unit/fusion point)	Rationale for engineering	Product	References
KS-AT-ACP	Modules of DEBS, soraphen, epothilone, geldanamycin, rifamycin, rapamycin, pikromycin, leptomycin PKS	Two module combinatorial biosynthesis of 14 module of 8 PKS/Conserved region of KS N-term and ACP C-term RE spicing	Domain sequence alignment	Triketide lactones derivatives	Menzella et al., 2005
KS-AT-ACP	Modules of DEBS 1,2,3 PKS	Methylmalonyl to propionyl, methylmalonyl, malonyl CoA specific/3 module combinatorial biosynthesis/conserved region of interpolypeptide linker	Domain sequence alignment	Triketide lactones derivatives	Klaus et al., 2016
AT-DH-KR-ACP-KS	Module 2 of neoaureothin PKS	Deletion of methylmalonyl CoA specific module 2/KS-AT linker conserved region and docking domain of ACP RE splicing	Domain sequence alignment	Homoaureothin	Sugimoto et al., 2015
KS-AT-ACP, KS-AT	Module 6 of DEBS PKS	Methylmalonyl-CoA to methylmalonyl or malonyl CoA specific/module 2,3,5 of DEBS PKS exchange/docking domain exchange to SYNZIP	Domain sequence alignment and structure-based selection	Triketide lactones derivatives	Klaus et al., 2019

swapping of only the AT domain was unfavorable in terms of impairment of catalytic activity owing to the non-native interactions between the swapped domains and neighboring domains such as KS, ACP, and interdomain linkers. The X-ray structural studies and mutagenesis studies of KS-AT and AT-ACP didomains revealed crucial residues in the interdomain linkers and domain interfaces for the specificity or catalytic activity, suggesting that the KS or ACP domain should be exchanged together with the AT domain (Kuhstoss et al., 1996; Marsden et al., 1998; Kim et al., 2004; Chen et al., 2007; Yuzawa et al., 2012; Miyanaga et al., 2016; Murphy et al., 2016). Recently, sequence alignments and structural studies related to module 6 of DEBS PKS and module 1 of β-lipomycin PKS provided ideal splicing sites located at interdomain linkers adjacent to AT domain despite of the protein–protein interactions with KS and ACP domains (Yuzawa et al., 2017b). The splicing sites were located at the conserved GTNAHVILE region of KS-AT linker (KAL) and conserved LPTY(A/P)FQ (H/R)xRYWL region of post AT linker 2 (PAL2) to minimize the non-native adjacent sequences, resulting in the universal KAL-AT-PAL1 unit for AT domain swapping. However, swapping by the KAL-AT-PAL1 unit is still remained to be validated for more various PKSs, because it may alter the protein–protein interactions between the domains and disrupt gatekeeping from downstream processing, which may result in incompatibility (Barajas et al., 2017).

In addition to determination of AT domain splicing sites, the exchange of a whole module (KS-AT-ACP) could be the alternative strategy to change the specificity. In this case, the interactions between upstream ACP and downstream KS domain should be considered, including covalent interdomain linkers and non-covalent interactions between domains. By splicing at the appropriate site of the interdomain linkers, 14 KS-AT-ACP modules from 8 PKSs was successfully isolated as a functional unit, and connected to generate a total of 154 combinatorial bimodular PKSs (Gokhale et al., 1999; Menzella et al., 2005). On the other hand, recent studies suggested that the evolutionary functional module is AT-ACP-KS rather than conventional KS-AT-ACP (Sugimoto et al., 2015; Keatinge-Clay, 2017). The sequence alignment of four aminopolyol PKSs supported this unit that the higher evolutionary correlation of the sequences between processing domains and downstream KS domain compared to the upstream KS domain (Zhang et al., 2017). Also, the first half of post-AT linker sequence showed higher correlation to the AT domain than the KS-AT linker, which refers AT-(processing domains)-ACP-KS is a more evolutionarily conserved unit. Moreover, the multi-modules of PKSs with the interpolypeptide non-covalent docking domain (DD) at both ends were able to be exchanged. The DD pairs were the compatible parts for exchanging the subunits that, using heterologous DD pairs led to the successful production of diketides and triketides (Menzella et al., 2005). However, several examples showed severe impairment of catalytic activity (Klaus et al., 2016). This is because the chain translocation step from the ACP domain of upstream module to the KS domain of downstream module was the bottleneck in addition to the chain elongation step (Khosla et al., 2007; Klaus et al.,

2016), emphasizing the importance of ACP-KS interaction, even when the DDs are compatible. Besides exchanging the module, incorporating the growing polypeptide with similar chain length to the native module was more successful than those of different chain length. Overall, the modular exchange strategies for PKS are diverse but the protein–protein interactions particularly for ACP-KS domains are important.

Domain or Module Exchange of NRPS

Similar to the exchange of domains and modules in PKSs, many engineering attempts for NRPSs have been reported. The swapping of only an A domain successfully resulted in the alteration of the specificity by the determination of proper splicing sites but, they were often hindered by non-native interactions between the swapping domains and neighboring domains such as C, T, and interdomain linkers (Stachelhaus et al., 1995; Doekel and Marahiel, 2000) (**Table 4**). The exchange of C-A didomain was more successful than the single A domain exchange in surfactin and pyoverdine NRPS (Tanovic et al., 2008; Calcott et al., 2014) but, failed in another report (Ackerley and Lamont, 2004). A-T didomain had a short interdomain linker with a conserved LPxP motif to interact with a key L-Tyr residue in the C-terminus of the A domain, and to assist the movement of the T domain during the catalytic cycle (Beer et al., 2014; Miller et al., 2014). Based on these characteristics, the A-T didomain in module 2 of actinomycin NRPS was successfully exchanged to A-T didomain in module 5 of actinomycin NRPS, but it still had low a yield (Schauwecker et al., 2000).

Considering both interactions within the C-A and A-T didomains, C-A-T module exchange occurred in several NRPSs. The most representative examples were, the daptomycin and A54145 NRPS modules of C-A-T or C-A-T-E spliced at the interdomain linker of the T and C domain that were changed to produce the novel daptomycin and A54145 derivatives (Yakimov et al., 2000; Nguyen et al., 2006, 2010; Butz et al., 2008). But the protein–protein interactions between T and C domain, consisting of a variable interdomain linker, were disturbed resulting in impaired activity. Different from the transacylation of PKS which leads to two separated steps of upstream ACP to KS, and then KS to downstream ACP, the C domain governs the peptide bond formation by one step and has strong stereoselectivity and side-chain selectivity for both donor and acceptor peptides (Clugston et al., 2003; Samel et al., 2007; Bloudoff and Schmeing, 2017; Li et al., 2017). Due to this strong correlation between all adjacent domains and linkers, finding the optimal splicing sites for the module exchange has remained a bottleneck. Recently, a suggested exchangeable unit A-T-C (XU) that the C-A linker was thought to be a better splicing site than other linkers in view of the N-terminal conserved region, the absence of secondary structures, and the fewer interactions between other domains (Bozhuyuk et al., 2018). Nonetheless, there was a requirement for the exchange of A-T-C unit that the C domain has specificity filters for both upstream and its own module. Therefore, downstream A-T-C unit should also be exchanged together with the target A-T-C unit. To avoid the limitations of XU, the XUC exchange unit which is $C_{acceptor}$-A-T-C_{donor} with the splicing site located at the intradomain conserved linker was

newly suggested (Bozhuyuk et al., 2019). Through dissecting the C domain specificity for upstream and downstream module, it was theoretically ideal that the exchange unit could be fused in a combinatorial manner. However, experimental results showed that the fusion between units from different genera caused a decrease in the yield. As the overall structure disruption by the C domain was thought to be the main reason, the solution to complement this limitation would be to research and discover more exchangeable XUC units.

There were several successful multi-module exchanges produced by using communication (COM) domain pairs such as DD domain pairs of PKS (Hahn and Stachelhaus, 2004; Chiocchini et al., 2006). In addition, the specificity of COM domain pairs has the ability to be altered via modification at their key residues resulting in the non-native COM domain (Hahn and Stachelhaus, 2006). Some examples, however, displayed a low product yield due to the disruption of non-covalent protein–protein interactions between the T and C domain. Overall, the modular exchange strategies for NRPS are diverse, similar to PKS, but the protein–protein interactions between domains, particularly those involving the C domain, are commonly important considerations.

Currently, the favorable strategies of exchanging module specificity for the production of novel chemical derivatives seems to be; (i) KAL-AT-PAL1 unit, AT-ACP-KS unit, using heterologous DD pairs for PKS, and (ii) XU, XUC, using heterologous COM domain pairs for NRPS, respectively. Although numerous studies for altering specificity of PKS and NRPS have been reported, more engineering trials should be accumulated for various PKSs and NRPSs to optimize the strategies employed.

Processing and Offloading Domain Engineering

In addition to the module specificity change, the incorporated extender unit could be further modified by various processing domains. Most of the processing target residues were the α-substituent and β-keto group of acyl-ACP intermediate of PKS, and R group of amino acid intermediate of NRPS (Barajas et al., 2017; Sussmuth and Mainz, 2017) (**Table 5**).

Processing and Offloading Domain Engineering of PKS

In PKS, ketoreductase (KR), dehydratase (DH), and enoyl-reductase (ER) domains are the most abundant processing domains located between the AT and ACP domain as a reductive loop, which governs the degree of β-carbonyl reduction. The KR domain performs the NADPH-mediated reduction of β-keto groups to β-hydroxyl groups and determines the stereochemistry of α-substituent and β-hydroxyl group. As the full deletion of the KR domain showed loss of specificity due to impaired protein folding and stability, the inactivation of the KR domain by its key residue mutations was more effective (Reid et al., 2003). In another approach the stereochemistry was altered by KR domain swapping, resulting in the effective stereocontrol of the product, which was generally difficult through modern synthetic methods (Annaval et al., 2015; Eng et al., 2016). Notably, substrate

TABLE 4 | Selected examples of domain or module exchange of modular NRPS.

Exchange unit	Target	Engineering strategy (substrate change/donor unit/fusion point)	Rationale for engineering	Product	References
A	A domain in module 7 of surfactin NRPS	L-Leu to L-Phe, L-Orn, L-Cys, L-Val specific/gramicidin S A domain in module 1,4,5 and ACV A domain in module 2,3 exchange/N,C-term RE site splicing	Domain sequence alignment	Five different surfactin variants	Stachelhaus et al., 1995
A	A domain in module 3 of hormaomycin NRPS	(β-Me)Phe to L-Thr, (3-Ncp)Ala, L-Val specific/Hormaomycin A subdomain in module 2,4,6 exchange/N,C-term RE site splicing	Homology modeling	Altered substrates (A domain assay)	Crüsemann et al., 2013
A, T-C-A	A domain in module 8,9 of tyrocidine NRPS	L-Orn and L-Leu to L-Ile, L-Phe specific/A domain in module 1 of bacitracin A NRPS/tyrocidine A (or T-C-A) domain in module 3/RE site splicing and fusion to TCALeuTTe.	Domain sequence alignment	Dipeptides (isolated module)	Doekel and Marahiel, 2000
A, C-A	A domain in module 10 of pyoverdine NRPS	L-Thr to L-Thr, L-Lys, L-Ser specific/Pyoverdine A domain in module 8, pyoverdine A domain in 3,4,5 modules, (A or C-A) of NRPS of other species exchange/deletion and genome integration by *attB*.	Domain sequence alignment	Pyoverdine derivatives	Calcott et al., 2014
A-T	A domain in module 2 of actinomycin NRPS	L-Val to N-methyl valine specific. Actinomycin (A-T) domain in module 5 exchange/C-A linker region, post T site RE splicing	Domain sequence alignment	Acyl-threonine–MeVal (isolated module)	Schauwecker et al., 2000
C-A-T, C-A-T-E	Modules 8-13 of daptomycin NRPS	L-Ala, L-Ser, 3 mGlu, L-Kyn to (L-Ser, L-Lys, L-Asn), (L-Ala, L-Asn), L-Glu, L-Trp, L-Ile specific/Modules 8-11 of A54145 NRPS/T-C, T-E, E-C linker RE splicing	Domain sequence alignment	Daptomycin derivatives	Nguyen et al., 2006
C-A-T, C-A-T-E	Modules 2,3,8,11,12,13 of A54145 NRPS	L-Glu, L-Asn, L-Lys, L-Asn, 3 mGlu to L-Asn, L-Asp, (L-Ala, L-Ser, L-Asn), (L-Ala, L-Ser), L-Glu specific/Many modules of A54145 NRPS/T-C, T-E, E-C linker RE splicing	Domain sequence alignment	A54145 derivatives	Nguyen et al., 2010
C-A-T-C	Modules 1,5 of surfactin NRPS	L-Glu, L-Asp to L-Gln, L-Asp specific/Module 1,5 of lichenysin A NRPS/C domain (HHXXXDG) active-site splicing	Domain sequence alignment	Recombinant lipopeptides	Yakimov et al., 2000
C-A + T-E	Insertion between module 4 and 5 of balhimycin NRPS	L-Hpg specific added/C$_5$A$_5$T$_4$E$_4$ of balhimycin NRPS/A-T linker region splicing	Domain sequence alignment	Balhimycin derivatives	Butz et al., 2008
A-T-C	Modules of xenotetrapeptide NRPS	Various changes/recombination of XtpS, GxpS, KolS, AmbS, GarS, GrsB, BicA/C-A linker region (WNATE) splicing.	Domain sequence alignment and structure-based selection	Xenotetrapeptides	Bozhuyuk et al., 2018
C$_{acceptor}$-A-T-C$_{donor}$	Modules of xenotetrapeptide NRPS	Various changes/recombination of XtpS, GxpS, KolS, AmbS, GarS, GrsB, BicA, SrfA, GrsAB, TycC, XeyS, Pax, others/C$_{acceptor}$-C$_{donor}$ linker splicing.	Domain sequence alignment and structure-based selection	Xenotetrapeptides	Bozhuyuk et al., 2019
C-A-T-E, multi-modules	Modules 4,5,6 of surfactin NRPS	L-Val, L-Asp, D-Leu to none/multimodule skipping by docking domain change/T/E-COM$_D$ or COM$_A$-C linker splicing	Domain sequence alignment	Surfactin derivatives	Chiocchini et al., 2006

specificity of the KR domain has been shown to be related to its substrate size, as it tends to be less active on smaller non-native substrates (Zheng et al., 2013). Next to the KR domain, DH domain promotes the dehydration of the β-hydroxyl group to form the double bond between α- and β-carbons, and even between β- and γ-carbons in some cases. Inactivation of DH domain by single mutation at the conserved active site motif resulted the retention of the β-hydroxyl group without specificity loss of the AT domain (Zhou et al., 2008). For the insertion of a non-native DH domain, it was important to consider the

TABLE 5 | Selected examples of processing and offloading domain engineering.

Target	Engineering strategy [engineering (+donor)/target sites]	Rationale for engineering	Product	References
KR domain in module 6 of DEBS PKS	KR inactivation/Y159F, S146A, K163E mutation	Domain sequence alignment and Homology modeling	3-Keto derivative of 6-deoxyerythronolide B	Reid et al., 2003
KR domain in module 2 of DEBS PKS	A1-type KR change to eight A2-type KR domains (amphotericin KR1, KR11, concanamycin KR4, KR10, elaiophylin KR4, oligomycin KR5, pimaricin KR7, candicidin KR13) and six B2-type KRs (DEBS KR1, lankamycin KR1, pikromycin KR1, lasalocid KR7, ECO-02301 KR19, stambomycin KR21)/AT-KR linker and KR-ACP linker RE splicing	Domain sequence alignment	Epimerized triketide lactones	Annaval et al., 2015
KR domain in module 1 of lipomycin PKS	A2-type KR change to A1-type amphotericin KR2 (or +DE2), and B1-type concanamycin KR2 (or +DE2)/post AT linker before DE and KR-ACP linker RE splicing	Domain sequence alignment and structure-based selection	*syn* form hydroxyacids	Eng et al., 2016
KR domain in module 2 of DEBS PKS	A1-type KR change to A2-type amphotericin KR2 and amphotericin KR11 with cognate DE/post AT linker before DE and KR-ACP linker RE splicing	Domain sequence alignment and structure-based selection	2D,2L-triketide lactone, ketolactones	Zheng et al., 2013
DH domain in module 18 and KR domain in module 21 of FR-008 PKS	KR and DH inactivation/Y1526F for KR21, H3084Y for DH18.	Domain sequence alignment	FR-008-V, -III, and -VI	Zhou et al., 2008
ER domain in module 4 of DEBS PKS	ER stereochemistry altered/Y52V for ER4 domain.	Domain sequence alignment and homology modeling	Triketide lactone with S to R methyl branch configuration switched	Kwan et al., 2008
KR domain in module 1 of borrelidin PKS	BorKR1 to Reductive loop of SpnB of spinosyn PKS and SpnBDH1 (in *cis*-double bond) to BorDH2 (in *trans*-double bond)/AT-DH linker region and ER-ACP linker region splicing	Domain sequence alignment, structure-based selection and Ppant ejection	Adipic acid	Hagen et al., 2016
TE domain in module 9 of tautomycetin PKS	Linear release change to macrocyclization by using TE domain in module 6 of pikromycin/TMC TE linker downstream splicing	Domain sequence alignment	Cyclized tautomycetin analog	Tripathi et al., 2016
E domain in module 4 of tyrocidine NRPS	L-Phe-D-Phe-L-Pro change to D-Phe-L-Pro. TycB2-3-AT.CATE(E_{TycA} or E_{TycB}) + TycB1-CAT/Te/T-E linker splicing	Domain sequence alignment	D-Phe-L-Pro (isolated module)	Stein et al., 2006
MT domain in module 2 of bassianolide NRPS, module 2 of beauvericin NRPS	*N*-methyl-L-Leu, *N*-methyl-L-Phe change to L-Leu, L-Phe. MT deletion/Intact domain deletion by overlap extension (SOE) PCR	Domain sequence information	*N*-desmethylbassianolide, *N*-desmethylbeauvericin B	Xu et al., 2019
MT domain insertion to module 6 of echinomycin NRPS	L-Ser change to *O* or *N*-methylated Ser. *O*-methylating MT domain of module 4 of kutznerides NRPS and *N*-methylating MT domain of module 3 of thiocoraline NRPS/intact domain insertion by RE splicing	Domain sequence alignment	*O* or *N*-methylated Ser (isolated module)	Lundy et al., 2018
Tyrocidine derivative with D-Phe4 connected to PEGA resin	D-Phe4 change to various amino acid libraries. TycC TE isolated for macrocyclization/intact domain splicing	Domain sequence alignment, combinatorial solid-phase chemistry	Cyclized tyrocidine analogs	Kohli et al., 2002

stereoselectivity of DH domain to the hydroxylated product by the upstream KR domain, which means that DH domain activity is strongly dependent to cognate KR domain (Barajas et al., 2017). Additionally, DH domain was also sensitive to the length of substrates that DH domains acting on natively long chain showed impaired activity to the shorter chain substrate (Faille et al., 2017). Lastly, ER domain promotes the reduction of the unsaturated α–β double bond formed by DH domain to generate the single bond, and determines the stereochemistry of the α-carbon. As the substrate of ER domain is the product of DH domain, the specificity of ER domain is dependent to both cognate KR and DH domain. Inactivation of ER domain by single mutation at the conserved residue switched the methyl branch configuration of triketide lactone product (Kwan et al., 2008). Moreover, ER domain affected the downstream ACP arrangement by the protein–protein interactions resulting in the different stereocontrol (Zhang et al., 2018). Three domains (KR, DH, ER) could be simultaneously introduced for the exchange of the KR domain in module 1 of borrelidin PKS (Hagen et al., 2016). An offloading TE domain was abundantly fused to other modules using the splicing site at the conserved linker between the offloading domain and ACP domain to produce truncated products. Moreover, they were regio-, stereospecific, which may alter the release and cyclization nature via swapping (Pinto et al., 2012; Barajas et al., 2015; Tripathi et al., 2016). Further engineering of other processing domains, such as the methyltransferase (MT) domain, should be conducted to utilize their potential to produce novel chemicals (Yuzawa et al., 2017c).

Processing and Offloading Domain Engineering of NRPS

Among a large variety of processing domains for non-ribosomal peptides, in this review we focused on the *cis*-acting processing domains located in the module.

Epimerase (E) domain governs *in situ* epimerization of the α-carbon of the T domain tethered L-amino acid during peptide elongation to generate D-amino acid. One example was the exchange of E domain in module 4 of tyrocidine NRPS ($E_{Tyc}B$) to E domain in loading module of tyrocidine NRPS ($E_{Tyc}A$), that showed the epimerization of L-Phe to D-Phe (Stein et al., 2006). Methyltransferse (MT) domain is another example of the engineering that provides *N*-methylation of the amino acid. This domain is usually integrated inside of A domain such as MT domain of cyclosporine A NRPS (Velkov et al., 2011). A recent engineering example of the MT domain in module 2 of bassianolide synthetase showed that the deletion of this domain generated the *N*-desmethylbassianolide without affecting the enzyme assembly lines (Xu et al., 2019). In another study, *O*-methylating MT domain was inserted between the A subdomains, resulting in the incorporation of *O* or *N*-methylated serine (Lundy et al., 2018). An offloading domain of NRPS was abundantly fused to other modules, similar to PKS, using the splicing site at the conserved linker between itself and the T domain to produce truncated products. Particularly, TE_{TycC} domain of tyrocidine NRPS was fused to other modules to generate various cyclized analogs by using its macrocyclization property (Kohli et al., 2001, 2002; Trauger et al., 2001).

Cyclization (Cy), Oxidase (Ox), and Formylation (F) domain were also expected to be the favorable targets for engineering. However, there have been no examples due to the upstream substrate specificity as well as other several mechanistic problems (Miller and Walsh, 2001; Schneider et al., 2003; Schoenafinger et al., 2006; Sussmuth and Mainz, 2017). Besides, other in-*trans* acting tailoring enzymes have the potential to be utilized for the production of novel derivatives but, studies are lacking. Therefore, further studies and engineering about various NRPS processing domains, including both in-*cis* and in-*trans* acting, should be quantitatively increased.

To sum up, the inactivation and swapping of the individual processing domain based on homology modeling and sequence alignment was successful, however, specialized protein–protein interactions hampered the full engineering of these domains, this required the case-by-case optimization of the swapping region.

REPURPOSING MODULAR PKS AND NRPS TO CONSTRUCT *DE NOVO* BIOSYNTHETIC PATHWAYS

As discussed above, numerous studies have repurposed modular PKSs and NRPSs via modifications and swapping of domains and modules in the original enzyme assembly lines, mainly for the production of novel derivatives. Otherwise, domains and modules were isolated and fused to elucidate the protein–protein interactions and substrate specificities of domains with the truncated product as a proof. Although there were several combinatorial examples of domain or module fusion to form the *de novo* enzyme assembly lines, their products were generally not the purpose of the study. Recently, technical advances of the separation and fusion of domains and modules from original enzyme assembly lines, owing to numerous engineering studies, opened the way for retro-biosynthesis. Retro-biosynthesis is a *de novo* pathway design that assembles reactions in a stepwise fashion in the reverse direction of synthesis, for the desired product (Birmingham et al., 2014) (**Figure 4**). In this approach, the final product formation is the sole selection criterion for each reaction step that reduced the screening effort for the high-throughput combinatorial experiments. By leveraging the potential diversity of the reactions from modular PKSs and NRPSs, the potential range of chemicals produced would be countless. Furthermore, this new pathway may have advantages in terms of reaction thermodynamics and economics. Despite these potentials, recently there have only been a few examples of repurposing PKSs and NRPSs due to experimental difficulties. In this section, we highlight recent efforts of repurposing PKS and NRPS domains and modules for the production of non-natural commodity chemicals or specialty chemicals by *de novo* pathway construction.

Redesigning Modular PKSs for Retro-Biosynthesis

Type I modular PKSs have favorable properties that functions and order of their modular catalytic domains determine the final
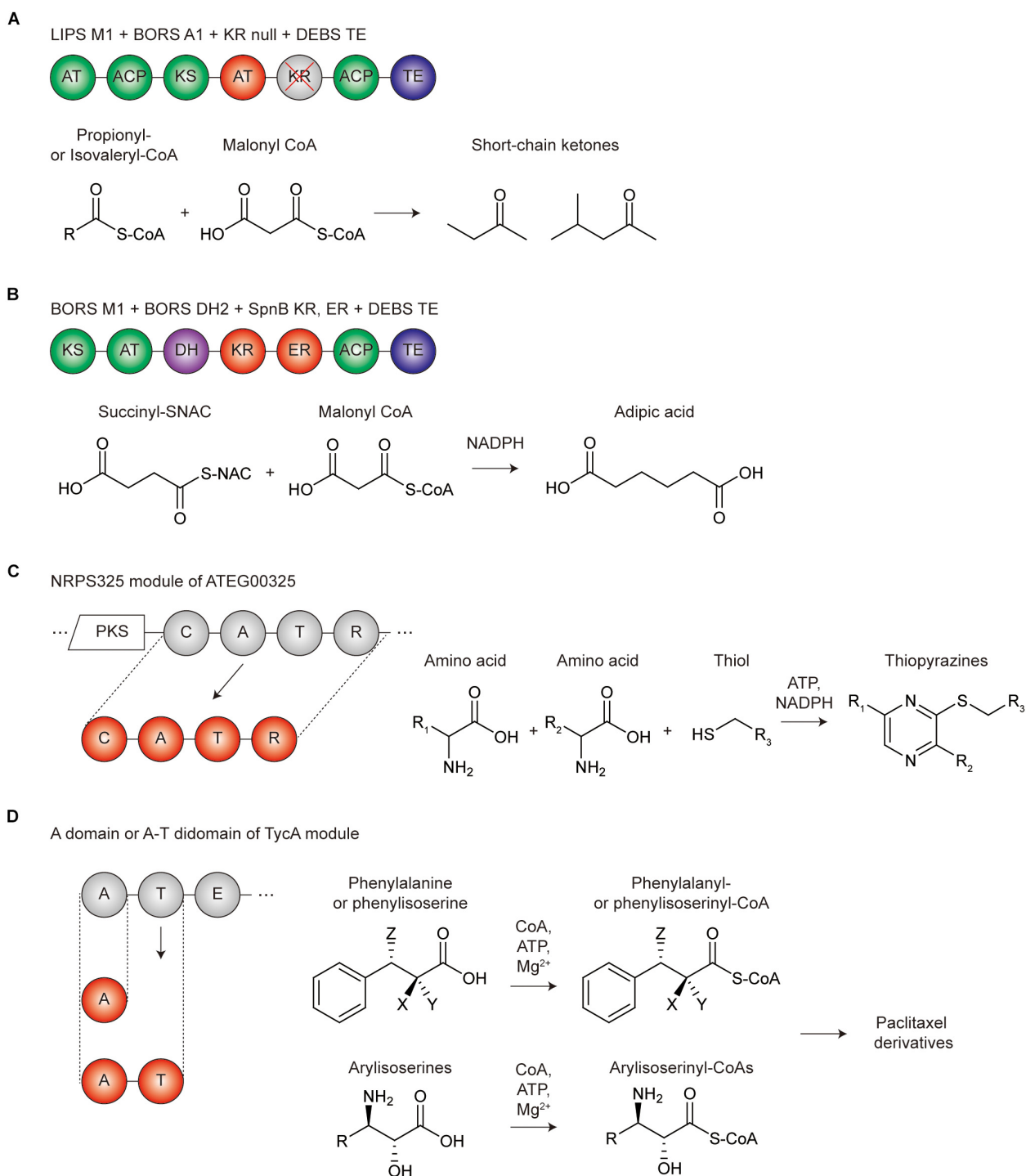
FIGURE 4 | Representative repurposing examples of modular PKS and NRPS for *de novo* biosynthetic pathways. **(A)** Repurposing the PKS domains and modules for the production of short-chain ketones. Green circles are the domains in module 1 of β-lipomycin PKS (LIPS M1), red circles are the AT domains in module 1 of borrelidin PKS (BORS A1), gray circles with the red crossed line are the inactivated KR domain (KR null), and the blue circles are the TE domain of DEBS PKS. **(B)** Repurposing the PKS domains and modules for the production of adipic acid. Green circles are the domains in module 1 of borrelidin PKS (BORS M1), red circles are the KR and ER domain in SpnB module of spinosyn PKS (SpnB KR, ER), and the blue circles are the TE domain of DEBS PKS. **(C)** Repurposing the NRPS module for the production of thiopyrazines. NRPS325 module of ATEG00325 PKS-NRPS hybrid megasynthetase was isolated (red circles) to promote the reaction for the thiopyrazine production itself. **(D)** Repurposing the NRPS domain for the production of paclitaxel derivatives. The A or A-T didomain in TycA module of tyrocidine A PKS was isolated (red circles) to be repurposed for the production of phenylalanyl-, phenylisoserinyl-, arylisoserinyl-CoAs, which are the precursors of the paclitaxel derivatives; X, NH_2 or H; Y, H or OH; Z, NH_2 or H.

TABLE 6 | Selected examples of modular PKS and NRPS repurposing for *de novo* biosynthetic pathways.

Category	Product	Host	Assembly line	References
Bio-based fuel	3-hydroxy acids	*In vitro*	LipPks1 from lipomycin PKS + TE from erythromycin PKS	Yuzawa et al., 2017b
Bio-based fuel	3-hydroxy acids	*Streptomyces venezuelae*	LipPks1 from lipomycin PKS + TE from erythromycin PKS	Yuzawa et al., 2017a
Bio-based fuel	Short chain ketones	*In vitro, Escherichia coli*	AT swapped LipPks1 from lipomycin PKS + TE from erythromycin PKS + KR null	Yuzawa et al., 2017b
Bio-based fuel	Short chain ketones	*Streptomyces albus*	AT swapped LipPks1 from lipomycin PKS + TE from erythromycin PKS + KR null + truncated N-terminal linker of KAL	Yuzawa et al., 2018
Industrial chemicals	Adipic acids	*In vitro*	Reductive loop swapped BolLM-BorMod1 from borrelidin PKS + TE from erythromycin PKS	Hagen et al., 2016
Industrial chemicals	Triketide lactones	*Escherichia coli*	154 designed bimodular PKSs based on 14 modules form eight PKSs	Menzella et al., 2005
Industrial chemicals	Triketide lactones	*Escherichia coli*	54 designed trimodular PKSs	Menzella et al., 2007
Industrial chemicals	Triketide lactones	*Streptomyces coelicolor*	DEBS1 + TE from erythromycin PKS	Regentin et al., 2004
Industrial chemicals	Triketide lactones	*Streptomyces venezuelae*	DEBS1 + TE from erythromycin PKS	Yan et al., 2012
Industrial chemicals	Triketide lactones	*Saccharomyces cerevisiae*	Mod2 of DEBS1 + TE from erythromycin PKS	Mutka et al., 2006
Potentially useful chemical	Thiopyrazine	*In vitro*	Isolated NRPS module (NRPS325) of isoflavipucine PKS-NRPS megasynthetase (ATEG00325)	Qiao et al., 2011
Anticancer	Paclitaxel derivatives	*In vitro*	Isolated NRPS domain (A or A-T domain) of tyrocidine synthetase TycA module + semisynthesis	Muchiri and Walker, 2012, 2017

product in a predictable manner. Most recently, an *in silico* toolkit called ClusterCAD, a computational platform for designing novel multi-modular type I PKS, has been developed and applied to several PKS repurposing studies (Eng et al., 2018). However, previous studies indicated that constructing multi-modular PKSs to produce novel chemical is still challenging (Weissman, 2016; Pang et al., 2019). The relationship between PKS module structure and acyl chain passage from one module to the next is not well understood thus, when testing the multi-modular assembly line, the acyl chain extension was frequently stalled at the middle of the synthesis process. Hence, presently, only recombination of one or two PKS modules to produce simple structured chemicals has been successfully conducted. To date, representative target molecules produced by using PKSs are divided into two categories: (i) fuels and (ii) industrial chemicals (**Table 6**).

Petroleum-based fuels are mixtures of highly reduced carbons with varying chain length as in the case of gasoline, which is mixture of C5 to C8 hydrocarbons. To date, bio-based hydrocarbon production has majorly relied on utilizing enzymes involved in isoprenoid and fatty acid biosynthesis (Beller et al., 2015). However, the alkane and alkene produced from these systems were generated as a mixture (Beller et al., 2010; Sukovich et al., 2010; Howard et al., 2013). As a way to overcome this limitation, theoretically, PKS modules can be engineered to produce hydrocarbons with specific carbon length. Recently, several cases have been reported that resulted in the over-production of pentadecaheptaene (PDH) by expressing iterative PKS of enediyne biosynthesis SgcE and its cognate TE SgcE10

in *Escherichia coli* or *Streptomyces* (Van Lanen et al., 2008; Zhang et al., 2008). Furthermore, through optimizing the enzyme ratio of SgcE and SgcE10, PDH production was enhanced, followed by the additional chemical reduction to overproduce the pentadecane (PD) (Liu et al., 2015). Presently, there is no example of producing hydrocarbons using only refactored modular PKS, but chemical reduction steps such as hydrogenation are required after the precursor production. Conversely, short-chain ketones have been successfully produced by using only repurposed PKS. Due to the high octane number, short-chain ketones have the potential to be utilized as gasoline replacements or gasoline oxygenates (Yuzawa et al., 2018). In addition, short-chain ketones are widely used as flavors and fragrances. PKS for short-chain ketone production was designed by repurposing the first PKS module from the lipomycin PKS and linking TE from erythromycin PKS (LipPks1 + TE). Originally, LipPks1 + TE was constructed for production of various 3-hydroxy acids depending on the starter acyl-CoAs (Yuzawa et al., 2017a,b). Engineering the same PKS by changing active site serine residue of KR domain to alanine deactivated the KR function and resulted in the production of various short-chain ketones (Yuzawa et al., 2017b).

A representative industrial chemical produced by re-designed PKS is adipic acid, a monomer component used to prepare the polymer Nylon 6,6. Conventionally, bio-based production of adipic acid has been conducted by producing *cis,cis*-muconic acid, followed by chemical hydrogenation (Niu et al., 2002), or constructing β-oxidation reversal and ω-oxidation pathways, which synthesized the dicarboxylic acid from glucose or glycerol (Clomburg et al., 2015). For PKS based C6 adipic

acid synthesis, a novel condensation strategy was proposed, which condenses C4 succinyl-CoA as the starter unit and C2 malonyl-CoA as the extender unit. Loading AT domain of the borrelidin PKS (BolLM) originally incorporates a *trans*-1,2-cyclo-pentanedicarboxylic acid CoA (CPDA-CoA) as the starter unit, but it has promiscuity to successfully recognize succinyl-CoA (Hagen et al., 2014, 2016). Furthermore, the first extension module of the borrelidin PKS (BorMod1) naturally incorporates malonyl-CoA. Thus, BolLM-BorMod1 linked with the TE domain from erythromycin produced 3-hydroxyadipic acid (Hagen et al., 2014). To convert 3-hydroxyadipic acid to adipic acid, further β-carbonyl processing governed by the DH and ER domains' is required, however, BorMod1 only contains a KR domain and lacks the DH and ER domains. Therefore, the reductive loop of BorMod1 was replaced with reductive loop from other PKSs containing full tridomains, resulted in successful synthesis of adipic acid (Hagen et al., 2016). Nevertheless, further studies are still needed for *in vivo* adipic acid production using PKS, as the starter unit, succinyl-CoA, is essential for the growth of the producing host and the promiscuity of the PKS modules will decrease the titer.

As fossil resources are depleted, the bio-based eco-friendly production of transport fuels and commodity chemicals, previously been produced from these fossil resources, is becoming more common. We believe that repurposing PKS will play a large role in this field. Additionally, further mechanistic understanding of PKS domains is expected to enable the design and implementation of PKS capable of producing even chemicals that do not occur naturally.

Redesigning Modular NRPSs for Retro-Biosynthesis

Non-ribosomal peptides are the most widely spread and structurally diverse secondary metabolites. NRPS modules are much more versatile than PKS modules in terms of the number of available substrate chemicals, including non-proteinogenic amino acids, thus the potential of producing novel chemicals by repurposing the module seems to be much higher. Nonetheless, current examples of NRPS repurposing are limited, compared to PKS, this may be due to the structural and mechanistic complexity of NRPS. In the case of PKS repurposing examples, the *de novo* pathway constructions for the useful chemical production were done by the combination of only unrelated PKS modules themselves. On the other hand, NRPS repurposing examples used a single NRPS domain or module reaction alone or combined with other non-NRPS enzymes. Thus, selected examples were discussed that showed the potential of further NRPS module repurposing.

The first example is a case that confirmed unexpected chemical production when a single NRPS module was isolated and expressed from the PKS-NRPS hybrid assembly line. Previously, novel PKS-NRPS assembly line was designed to produce tryptophan-containing preaspyridone analog by replacing the NRPS domain of PKS-NRPS megasynthetase (ApdA) of *Aspergillus nidulans*, which originally synthesizes tetramic acid preaspyridone, with the NRPS domain of

cyclopiazonic acid synthetase (CpaS) of *Aspergillus flavus* (Liu and Walsh, 2009). Based on these results, the same group dissected and tested the function of a single NRPS module (NRPS325) of PKS-NRPS megasynthetase (ATEG00325) from *Aspergillus terreus*, which produces isoflavipucine. Unexpectedly, the NRPS module can synthesize thiopyrazine *in vitro*, which is largely different from the original role of the module in the parent enzyme (Qiao et al., 2011). This showed that in a multi-modular assembly line, a single NRPS module is capable of producing potentially useful chemicals and can be utilized as a part to construct *de novo* biosynthetic pathways for non-natural compound production.

Another example was the repurposing of a tyrocidine NRPS domain (A domain or A-T didomain) to the paclitaxel biosynthesis pathway to produce various paclitaxel derivatives (Muchiri and Walker, 2012, 2017). Antimitotic, anticancer paclitaxel was originally produced from *Taxus brevifolia* or the semisynthetic method, which suffered from low yields and environmentally harmful reactions, respectively. Moreover, the synthesis of the most essential precursor, amino phenylpropanoyl CoA substrates, was a major limitation as, it required protection at their amino groups before synthetic thioesterification. The authors hypothesized that A domain in tyrocidine NRPS module TycA could be used as the potential chemoselective carboxylate CoA ligase that originally had phenylalanine specificity but showed potential promiscuity. As a result, α-, β-phenylalanyl, and (2*R*,3*S*)-phenylisoserinyl CoA were successfully generated by the A domain (Muchiri and Walker, 2012). By employing 16 substituted phenylisoserines, the A-T didomain of TycA converted them to their corresponding isoserinyl CoAs, resulting in the production of docetaxel, milataxel, and various other analogs (Muchiri and Walker, 2017). These examples showed that a single NRPS domain could be integrated to other biosynthetic or semisynthetic pathways and utilized as a part for the construction of *de novo* pathways for non-natural compound production.

Owing to the natural diversity of the amino acid substrates other than acyl CoA substrates, the substrate promiscuity of the NRPS modules seems to be broader than PKS modules. Although these aspects of NRPS hindered the understanding of their reactions, further advances would increase the potential of NRPS to be repurposed for the construction of *de novo* biosynthesis pathways of useful chemicals, by fusing the NRPS modules.

THE ROADMAP FOR REPURPOSING MODULAR PKS AND NRPS

As shown above, modular PKS and NRPS have been studied with the aim of (i) engineering for the production of novel chemicals, and (ii) repurposing the domain and module reactions toward *de novo* biosynthetic pathways for the production of useful chemicals. To overcome previous limitations of these approaches for the expansion of the diversity of their products and optimization of their productivity, the systematic strategy such as the design-build-test-learn cycle should be applied. The
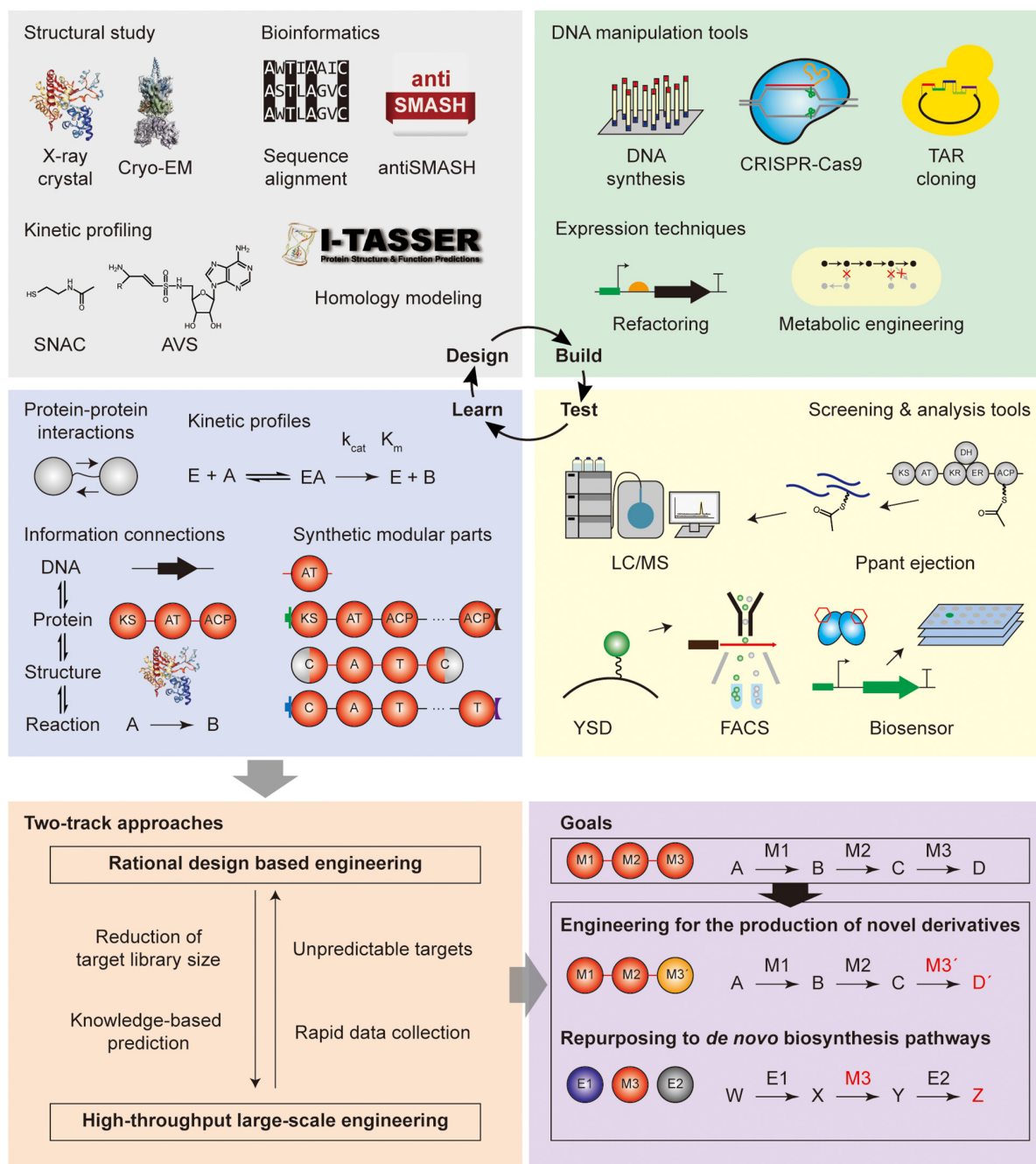
FIGURE 5 | Roadmap for repurposing modular PKS and NRPS. Design-build-test-learn cycle with the tools for each step was illustrated.

brief explanations, applications, and perspectives of the tools for each step are discussed below.

Design Tools

Design tools include structural, kinetic, mechanistic, and sequence-based techniques to obtain fundamental information for the experimental design of the repurposing. Specifically, the structural and kinetic knowledge is collected and sorted to the database, and bioinformatics tools are used for domain/module

selection and boundary identification based on the database. First, structural biology tools were used for the modular PKS and NRPS, which are macromolecules with numerous protein–protein interactions and dynamic conformational change during biosynthesis. Thus, the structures and interactions of the domains and one or two modules have been reported for high-resolution and dynamic scale. X-ray crystallography was the most commonly used technique that provided high-resolution images at the angstrom level (Conti et al., 1997;

Keatinge-Clay and Stroud, 2006; Tang et al., 2006; Buchholz et al., 2009). However, it would require the protein to crystallize which can be difficult for large structures. Additionally, the dynamic conformations and protein–protein interactions could not be inferred. To complement these limitations, conventional electron microscopy was utilized but it required additional staining, and suffered from relatively low resolution due to the deformation of flexible particles (Hoppert et al., 2001). Also, NMR spectroscopy was predominantly practiced in the structural determination of ACP and docking domains, and their conformation dynamics (Broadhurst et al., 2003; Richter et al., 2008; Lim et al., 2012). However, the size limitation of this technique hindered to be utilized for the whole module structure. Other alternative techniques also reported were the integrated utilization of X-ray crystallography and electron microscopy (Tarry et al., 2017), and Small-angle X-ray scattering (SAXS) or spectroscopic methods for the study of time-resolved conformations (Edwards et al., 2014; Alfermann et al., 2017). Recently, cryo-EM technique was applied to obtain the high-resolution and dynamic structural information of module 5 of pikromycin PKS and pentaketide-bound form of the module, but the homogeneous sample should be prepared carefully (Dutta et al., 2014; Whicher et al., 2014). By harnessing the rapid technical developments, the dynamic structures of the larger enzyme assemblies including multi-module level could be further elucidated.

In addition to the structural efforts, kinetic profiling of the modular PKS and NRPS has been vigorously utilized to measure the substrate specificity of AT and A domain, and to understand the protein–protein interactions between various domains or modules. This information is essential for the determination of the target and method for engineering. In case of the substrate specificity and affinity assays, integrated UV assay with NADPH consumption for acyltransferase activity in PKS (Lowry et al., 2013), and pyrophosphate exchange assay, pyrophosphate release assay, and hydroxylamine quenching assay for adenylation activity in NRPS have been performed (Villiers and Hollfelder, 2009; Wilson and Aldrich, 2010; Katano et al., 2013; Stanisic and Kries, 2019). The substrate specificity of AT, A, KS, and C domains in the elongation module toward growing polyketide chains or polypeptide tethered to phosphopantetheine (Ppant) arm of ACP and T domain was studied by using the mimicking molecule such as acyl- or aminoacyl-*N*-acetylcysteamine thioesters (acyl- or aminoacyl-SNACs) (Ehmann et al., 2000; Jensen et al., 2012). The studies of the protein–protein interactions commonly exploited the mechanism-based crosslinkers, inhibitors, and probes. Most of them covalently attached to the Ppant arm or other active sites of ACP or T domain. There were many examples of PKSs including phosphopantetheine analogs, photo-crosslinking benzophenone, and azide-alkyne click chemistry linkers for the understanding of KS-AT, AT-ACP, and ACP-KS interactions (Worthington et al., 2006; Ye and Williams, 2014; Ye et al., 2014; Ladner and Williams, 2016). The interactions were quantified by means of radioisotopic transfer assay, thermodynamics heat using calorimetry, and fluorescence, or MS-based techniques. For NRPS, mechanism-based inhibitors such as 5′-*O*-sulfamoyladenosine (AMS) and adenosine vinylsulfonamide (AVS), azide-alkyne click chemistry

linkers, and biotin probes were mainly used for the studies of NRPS A-T interactions (Finking et al., 2003; Sundlov et al., 2012; Ishikawa and Kakeya, 2014; Stanisic and Kries, 2019). These crosslinkers were versatile tools for the pretreatment of structural crystallization (Mitchell et al., 2012). As various tools have been used for the kinetic profiling, it might be better to compare between them for the reproducibility, or to establish the universal standard or tool.

Based on the information obtained from the structural and kinetic studies, computational approaches and bioinformatics would be required to design the detailed processes of the engineering of PKS and NRPS enzymes, or pathways (Medema and Fischbach, 2015; Alanjary et al., 2019). For example, the protein or DNA sequence boundaries of the domain, module, and linker for the swapping could be determined from the database of genome, protein, and smBGC (Marchler-Bauer et al., 2015; Blin et al., 2019; Sayers et al., 2020). In the case of the absence of the structural and kinetic profiling data, comparative analysis method was mainly used including sequence alignment tools, database search tools, and evolutionary analysis tools (Navarro-Munoz et al., 2020). Homology modeling tools such as I-TASSER, MODELLER, and SWISS-MODEL were frequently used to predict the structure of PKS or NRPS by comparative analysis with the reported similar structures for the engineering design (Roy et al., 2010; Waterhouse et al., 2018; Bitencourt-Ferreira and de Azevedo, 2019). Lastly, the integration of the enzyme reactions with their substrates, products, protein sequence, and gene sequence was applied to the automated pathway design pipeline tool of the combinations of reactions for the final products (Carbonell et al., 2014; Eng et al., 2018). Along with the structural and kinetic profiling tools, bioinformatic tools should be integrated and continuously updated to provide the universal, reproducible, and versatile strategy to involve the full pipeline of database search, comparative analysis, homology modeling, and experimental design.

Build Tools

Build tools are technical tools for preparing DNA parts, proteins, and hosts to test the experimental design from above. In recent decades synthetic biology tools have been rapidly developed, enabling large-scale, combinatorial, and high-throughput engineering (Winn et al., 2016; Lee et al., 2019). Despite these advanced tools, they are not frequently used for PKS and NRPS engineering as yet. In case of the genetic manipulation tools, the DNA sequence fragments of the domain and module for the engineering were obtained from the native source, mostly by traditional PCR amplification and restriction enzyme digestion. The rapid and inexpensive cost of DNA synthesis and the highly efficient, accurate DNA digestion tools such as CRISPR-Cas9 would replace the conventional restriction enzyme-based cloning method (Lee et al., 2015; Hughes and Ellington, 2017). Moreover, homologous recombination based DNA assembly tools such as Gibson assembly, linear-linear homologous recombination (LLHR), and yeast TAR cloning would be more efficient for the large size and numbers of DNA fragments than traditional ligation methods (Fu et al., 2012; Jiang et al., 2015; Lee et al., 2015). Site-directed mutagenesis for the PKS

and NRPS enzymes would be also easier by exploiting CRISPR-Cas9 system and the base-editors, even for *in vivo* genome editing (Komor et al., 2016). Directed evolution of synthetic libraries for the modular PKS and NRPS are feasible with these genetic manipulation tools. Furthermore, *in vitro* expression techniques such as protein purification tools and precursor synthesis tools as well as *in vivo* expression techniques in the heterologous host, such as promoter refactoring and metabolic pathway engineering, would simultaneously be developed with genetic manipulation tools (Chen et al., 2007; Weber et al., 2015; Winn et al., 2016).

Test Tools

Test tools were used to quantitatively measure the novel derivatives from engineering, or the products from novel pathways, to analyze the chemical structure of the products and to screen the desired clone from the libraries. The most widely applied tools for the quantitative measurement of the products were high-throughput mass spectrometry techniques that were also used for the screening and chemical analysis (Krug and Muller, 2014; Boiteau et al., 2018). By integration with other techniques such as liquid chromatography, the separation and evaluation of exact molecular weight, amount, and chemical moieties for the product could be interrogated. In many cases, the screening step is the bottleneck, which requires rapid and accurate selection of the desired clone among the extensive libraries. Leveraging the biological activity of the product for screening was the traditional screening assay logic that involves an inhibition zone assay for antibacterial activity, colorimetric assay for pigment product, and growth assay for the auxotrophic strain of the product such as siderophore (Balouiri et al., 2016; Cleto and Lu, 2017; Wehrs et al., 2019). As these methods were limited to specific products and the small size of the libraries, the universal and high-throughput screening strategies such as yeast surface display combined with FACS (Zhang et al., 2013), Ppant ejection strategy detecting the intermediate of the rate-limiting step by protease (Meluzzi et al., 2008), and biosensor development for the PKS and NRPS product such as the macrolide biosensor MphR (Kasey et al., 2018), along with mass spectrometry, would be a more favorable method.

Design-Build-Test-Learn (DBTL) Cycle for Modular PKS and NRPS Repurposing

Learning from the design-build-test steps of modular PKS and NRPS repurposing would be; (i) the updated information of the protein–protein interactions and the kinetic profiles to the substrate of the enzymes, (ii) the appropriate splicing or engineering sites for the construction of synthetic parts, (iii) the rationale for the selection of compatible sets, (iv) the integration of the information of cognate reactions, sequences, substrate, and products, and (v) the reversible connections between the

combination of reactions and the final product for retro-biosynthesis (Bayly and Yadav, 2017). Based on the current understanding, iterative DBTL cycles would ultimately achieve the goal of novel chemical productions as well as novel pathway generation (**Figure 5**).

CONCLUSION

We reviewed selected examples of modular PKS and NRPS repurposing for the generations of novel chemicals and pathways, followed by the roadmap in view of the synthetic biological DBTL cycle. The most important lesson from the repurposing examples was the requirement of careful considerations for dissecting the complex protein–protein interactions, despite their functionality in modular fashion. Therefore, the rational engineering design should continuously be improved by the characterization of the modular PKS and NRPS from the large amount of structural, kinetic, and genetic studies, to predict and understand the results. In parallel, the massive and rapid approaches, such as directed evolution and combinatorial strategy along with high-throughput screening, should be widely adopted to find the unpredictable factors from rational engineering. Technical advances for build tools could overcome the previous limitations for the identification of the exchangeable unit and the case-by-case optimization. Ultimately, the two approaches would complement each other by the reducing library size by rational design and the reflection of the updated information learned from the non-rational large-scale strategy. An automated pipeline of the modular PKS and NRPS repurposing for retro-biosynthesis is expected to greatly expand the reservoir of the bio-active compounds.

AUTHOR CONTRIBUTIONS

B-KC and SH conceived the study. SH, NL, and B-KC drafted the manuscript. SH, NL, SC, BP, and B-KC contributed to the final version of the manuscript.

FUNDING

REFERENCES

Ackerley, D. F., and Lamont, I. L. (2004). Characterization and genetic manipulation of peptide synthetases in *Pseudomonas aeruginosa* PAO1 in

order to generate novel pyoverdines. *Chem. Biol.* 11, 971–980. doi: 10.1016/j.chembiol.2004.04.014

Agarwal, P. K. (2006). Enzymes: an integrated view of structure, dynamics and function. *Microb. Cell Fact.* 5:2. doi: 10.1186/1475-2859-5-2

Alanjary, M., Cano-Prieto, C., Gross, H., and Medema, M. H. (2019). Computer-aided re-engineering of nonribosomal peptide and polyketide biosynthetic assembly lines. *Nat. Prod. Rep.* 36, 1249–1261. doi: 10.1039/c9np00021f

Alfermann, J., Sun, X., Mayerthaler, F., Morrell, T. E., Dehling, E., Volkmann, G., et al. (2017). FRET monitoring of a nonribosomal peptide synthetase. *Nat. Chem. Biol.* 13, 1009–1015. doi: 10.1038/nchembio.2435

Annaval, T., Paris, C., Leadlay, P. F., Jacob, C., and Weissman, K. J. (2015). Evaluating ketoreductase exchanges as a means of rationally altering polyketide stereochemistry. *Chembiochem* 16, 1357–1364. doi: 10.1002/cbic.201500113

Balouiri, M., Sadiki, M., and Ibnsouda, S. K. (2016). Methods for *in vitro* evaluating antimicrobial activity: a review. *J. Pharm. Anal.* 6, 71–79. doi: 10.1016/j.jpha.2015.11.005

Barajas, J. F., Blake-Hedges, J. M., Bailey, C. B., Curran, S., and Keasling, J. D. (2017). Engineered polyketides: synergy between protein and host level engineering. *Synth. Syst. Biotechnol.* 2, 147–166. doi: 10.1016/j.synbio.2017.08.005

Barajas, J. F., Phelan, R. M., Schaub, A. J., Kliewer, J. T., Kelly, P. J., Jackson, D. R., et al. (2015). Comprehensive structural and biochemical analysis of the terminal myxalamid reductase domain for the engineered production of primary alcohols. *Chem. Biol.* 22, 1018–1029. doi: 10.1016/j.chembiol.2015.06.022

Bayly, C. L., and Yadav, V. G. (2017). Towards precision engineering of canonical polyketide synthase domains: recent advances and future prospects. *Molecules* 22:235. doi: 10.3390/molecules22020235

Beer, R., Herbst, K., Ignatiadis, N., Kats, I., Adlung, L., Meyer, H., et al. (2014). Creating functional engineered variants of the single-module non-ribosomal peptide synthetase IndC by T domain exchange. *Mol. Biosyst.* 10, 1709–1718. doi: 10.1039/c3mb70594c

Beller, H. R., Goh, E. B., and Keasling, J. D. (2010). Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Appl. Environ. Microbiol.* 76, 1212–1223. doi: 10.1128/AEM.02312-09

Beller, H. R., Lee, T. S., and Katz, L. (2015). Natural products as biofuels and bio-based chemicals: fatty acids and isoprenoids. *Nat. Prod. Rep.* 32, 1508–1526. doi: 10.1039/c5np00068h

Bernhardt, P., and O'Connor, S. E. (2009). Opportunities for enzyme engineering in natural product biosynthesis. *Curr. Opin. Chem. Biol.* 13, 35–42. doi: 10.1016/j.cbpa.2009.01.005

Birmingham, W. R., Starbird, C. A., Panosian, T. D., Nannemann, D. P., Iverson, T. M., and Bachmann, B. O. (2014). Bioretrosynthetic construction of a didanosine biosynthetic pathway. *Nat. Chem. Biol.* 10, 392–399. doi: 10.1038/nchembio.1494

Bitencourt-Ferreira, G., and de Azevedo, W. F. Jr. (2019). Homology modeling of protein targets with MODELLER. *Methods Mol. Biol.* 2053, 231–249. doi: 10.1007/978-1-4939-9752-7_15

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310

Bloudoff, K., and Schmeing, T. M. (2017). Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. *Biochim. Biophys. Acta* 1865(11 Pt B), 1587–1604. doi: 10.1016/j.bbapap.2017.05.010

Boiteau, R. M., Hoyt, D. W., Nicora, C. D., Kinmonth-Schultz, H. A., Ward, J. K., and Bingol, K. (2018). Structure elucidation of unknown metabolites in metabolomics by combined NMR and MS/MS prediction. *Metabolites* 8:8. doi: 10.3390/metabo8010008

Bozhuyuk, K. A. J., Fleischhacker, F., Linck, A., Wesche, F., Tietze, A., Niesert, C. P., et al. (2018). *De novo* design and engineering of non-ribosomal peptide synthetases. *Nat. Chem.* 10, 275–281. doi: 10.1038/nchem.2890

Bozhuyuk, K. A. J., Linck, A., Tietze, A., Kranz, J., Wesche, F., Nowak, S., et al. (2019). Modification and *de novo* design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nat. Chem.* 11, 653–661. doi: 10.1038/s41557-019-0276-z

Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., and Weissman, K. J. (2003). The structure of docking domains in modular polyketide synthases. *Chem. Biol.* 10, 723–731. doi: 10.1016/s1074-5521(03)00156-x

Brown, A. S., Calcott, M. J., Owen, J. G., and Ackerley, D. F. (2018). Structural, functional and evolutionary perspectives on effective re-engineering of non-ribosomal peptide synthetase assembly lines. *Nat. Prod. Rep.* 35, 1210–1228. doi: 10.1039/c8np00036k

Buchholz, T. J., Geders, T. W., Bartley, F. E., Reynolds, K. A., Smith, J. L., and Sherman, D. H. (2009). Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem. Biol.* 4, 41–52. doi: 10.1021/cb8002607

Butz, D., Schmiederer, T., Hadatsch, B., Wohlleben, W., Weber, T., and Sussmuth, R. D. (2008). Module extension of a non-ribosomal peptide synthetase of the glycopeptide antibiotic balhimycin produced by *Amycolatopsis balhimycina*. *Chembiochem* 9, 1195–1200. doi: 10.1002/cbic.200800068

Calcott, M. J., Owen, J. G., Lamont, I. L., and Ackerley, D. F. (2014). Biosynthesis of novel pyoverdines by domain substitution in a nonribosomal peptide synthetase of *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* 80, 5723–5731. doi: 10.1128/AEM.01453-14

Carbonell, P., Parutto, P., Baudier, C., Junot, C., and Faulon, J. L. (2014). Retropath: automated pipeline for embedded metabolic circuits. *ACS Synth. Biol.* 3, 565–577. doi: 10.1021/sb4001273

Challis, G. L., Ravel, J., and Townsend, C. A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* 7, 211–224. doi: 10.1016/s1074-5521(00)00091-0

Chan, Y. A., Podevels, A. M., Kevany, B. M., and Thomas, M. G. (2009). Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep.* 26, 90–114. doi: 10.1039/b801658p

Chen, A. Y., Cane, D. E., and Khosla, C. (2007). Structure-based dissociation of a type I polyketide synthase module. *Chem. Biol.* 14, 784–792. doi: 10.1016/j.chembiol.2007.05.015

Chevrette, M. G., Gutierrez-Garcia, K., Selem-Mojica, N., Aguilar-Martinez, C., Yanez-Olvera, A., Ramos-Aboites, H. E., et al. (2019). Evolutionary dynamics of natural product biosynthesis in bacteria. *Nat. Prod. Rep.* doi: 10.1039/c9np00048h [Epub ahead of print].

Chiocchini, C., Linne, U., and Stachelhaus, T. (2006). *In vivo* biocombinatorial synthesis of lipopeptides by COM domain-mediated reprogramming of the surfactin biosynthetic complex. *Chem. Biol.* 13, 899–908. doi: 10.1016/j.chembiol.2006.06.015

Cleto, S., and Lu, T. K. (2017). An engineered synthetic pathway for discovering nonnatural nonribosomal peptides in *Escherichia coli*. *mBio* 8:e01474-17. doi: 10.1128/mBio.01474-17

Clomburg, J. M., Blankschien, M. D., Vick, J. E., Chou, A., Kim, S., and Gonzalez, R. (2015). Integrated engineering of beta-oxidation reversal and omega-oxidation pathways for the synthesis of medium chain omega-functionalized carboxylic acids. *Metab. Eng.* 28, 202–212. doi: 10.1016/j.ymben.2015.01.007

Clugston, S. L., Sieber, S. A., Marahiel, M. A., and Walsh, C. T. (2003). Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a DCL catalyst. *Biochemistry* 42, 12095–12104. doi: 10.1021/bi035090+

Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* 16, 4174–4183. doi: 10.1093/emboj/16.14.4174

Crawford, J. M., Portmann, C., Kontnik, R., Walsh, C. T., and Clardy, J. (2011). NRPS substrate promiscuity diversifies the xenematides. *Org. Lett.* 13, 5144–5147. doi: 10.1021/ol2020237

Crüsemann, M., Kohlhaas, C., and Piel, J. (2013). Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains. *Chem. Sci.* 4, 1041–1045. doi: 10.1039/C2SC21722H

Dan, Q., Newmister, S. A., Klas, K. R., Fraley, A. E., McAfoos, T. J., Somoza, A. D., et al. (2019). Fungal indole alkaloid biogenesis through evolution of a bifunctional reductase/Diels-Alderase. *Nat. Chem.* 11, 972–980. doi: 10.1038/s41557-019-0326-6

Doekel, S., and Marahiel, M. A. (2000). Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* 7, 373–384. doi: 10.1016/s1074-5521(00)00118-6

Dutta, S., Whicher, J. R., Hansen, D. A., Hale, W. A., Chemler, J. A., Congdon, G. R., et al. (2014). Structure of a modular polyketide synthase. *Nature* 510, 512–517. doi: 10.1038/nature13423

Edwards, A. L., Matsui, T., Weiss, T. M., and Khosla, C. (2014). Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J. Mol. Biol.* 426, 2229–2245. doi: 10.1016/j.jmb.2014.03.015

Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000). Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. *Chem. Biol.* 7, 765–772. doi: 10.1016/s1074-5521(00)00022-3

Eichner, S., Floss, H. G., Sasse, F., and Kirschning, A. (2009). New, highly active nonbenzoquinone geldanamycin derivatives by using mutasynthesis. *Chembiochem* 10, 1801–1805. doi: 10.1002/cbic.200900246

Eng, C. H., Backman, T. W. H., Bailey, C. B., Magnan, C., Garcia Martin, H., Katz, L., et al. (2018). ClusterCAD: a computational platform for type I modular polyketide synthase design. *Nucleic Acids Res.* 46, D509–D515. doi: 10.1093/nar/gkx893

Eng, C. H., Yuzawa, S., Wang, G., Baidoo, E. E., Katz, L., and Keasling, J. D. (2016). Alteration of polyketide stereochemistry from *anti* to *syn* by a ketoreductase domain exchange in a type I modular polyketide synthase subunit. *Biochemistry* 55, 1677–1680. doi: 10.1021/acs.biochem.6b00129

Eppelmann, K., Stachelhaus, T., and Marahiel, M. A. (2002). Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* 41, 9718–9726. doi: 10.1021/bi0259406

Eustaquio, A. S., and Moore, B. S. (2008). Mutasynthesis of fluorosalinosporamide, a potent and reversible inhibitor of the proteasome. *Angew. Chem. Int. Ed. Engl.* 47, 3936–3938. doi: 10.1002/anie.200800177

Evans, B. S., Chen, Y., Metcalf, W. W., Zhao, H., and Kelleher, N. L. (2011). Directed evolution of the nonribosomal peptide synthetase AdmK generates new andrimid derivatives *in vivo*. *Chem. Biol.* 18, 601–607. doi: 10.1016/j.chembiol.2011.03.008

Faille, A., Gavalda, S., Slama, N., Lherbet, C., Maveyraud, L., Guillet, V., et al. (2017). Insights into substrate modification by dehydratases from type I polyketide synthases. *J. Mol. Biol.* 429, 1554–1569. doi: 10.1016/j.jmb.2017.03.026

Finking, R., Neumuller, A., Solsbacher, J., Konz, D., Kretzschmar, G., Schweitzer, M., et al. (2003). Aminoacyl adenylate substrate analogues for the inhibition of adenylation domains of nonribosomal peptide synthetases. *Chembiochem* 4, 903–906. doi: 10.1002/cbic.200300666

Fischbach, M. A., Lai, J. R., Roche, E. D., Walsh, C. T., and Liu, D. R. (2007). Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11951–11956. doi: 10.1073/pnas.0705348104

Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P. M., et al. (2012). Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* 30, 440–446. doi: 10.1038/nbt.2183

Garcia-Bernardo, J., Xiang, L., Hong, H., Moore, B. S., and Leadlay, P. F. (2004). Engineered biosynthesis of phenyl-substituted polyketides. *Chembiochem* 5, 1129–1131. doi: 10.1002/cbic.200400007

Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science* 284, 482–485. doi: 10.1126/science.284.5413.482

Hagen, A., Poust, S., de Rond, T., Yuzawa, S., Katz, L., Adams, P. D., et al. (2014). *In vitro* analysis of carboxyacyl substrate tolerance in the loading and first extension modules of borrelidin polyketide synthase. *Biochemistry* 53, 5975–5977. doi: 10.1021/bi500951c

Hagen, A., Poust, S., Rond, T., Fortman, J. L., Katz, L., Petzold, C. J., et al. (2016). Engineering a polyketide synthase for *in vitro* production of adipic acid. *ACS Synth. Biol.* 5, 21–27. doi: 10.1021/acssynbio.5b00153

Hahn, M., and Stachelhaus, T. (2004). Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15585–15590. doi: 10.1073/pnas.0404932101

Hahn, M., and Stachelhaus, T. (2006). Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides. *Proc. Natl. Acad. Sci. U.S.A.* 103, 275–280. doi: 10.1073/pnas.0508409103

Han, J. W., Kim, E. Y., Lee, J. M., Kim, Y. S., Bang, E., and Kim, B. S. (2012). Site-directed modification of the adenylation domain of the fusaricidin nonribosomal peptide synthetase for enhanced production of fusaricidin analogs. *Biotechnol. Lett.* 34, 1327–1334. doi: 10.1007/s10529-012-0913-8

Hans, M., Hornung, A., Dziarnowski, A., Cane, D. E., and Khosla, C. (2003). Mechanistic analysis of acyl transferase domain exchange in polyketide synthase modules. *J. Am. Chem. Soc.* 125, 5366–5374. doi: 10.1021/ja029539i

Hoppert, M., Gentzsch, C., and Schorgendorfer, K. (2001). Structure and localization of cyclosporin synthetase, the key enzyme of cyclosporin biosynthesis in *Tolypocladium inflatum*. *Arch. Microbiol.* 176, 285–293. doi: 10.1007/s002030100324

Howard, T. P., Middelhaufe, S., Moore, K., Edner, C., Kolak, D. M., Taylor, G. N., et al. (2013). Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7636–7641. doi: 10.1073/pnas.1215966110

Hughes, R. A., and Ellington, A. D. (2017). Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* 9:a023812. doi: 10.1101/cshperspect.a023812

Hur, G. H., Vickery, C. R., and Burkart, M. D. (2012). Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat. Prod. Rep.* 29, 1074–1098. doi: 10.1039/c2np20025b

Ishikawa, F., and Kakeya, H. (2014). Specific enrichment of nonribosomal peptide synthetase module by an affinity probe for adenylation domains. *Bioorg. Med. Chem. Lett.* 24, 865–869. doi: 10.1016/j.bmcl.2013.12.082

Jaremko, M. J., Davis, T. D., Corpuz, J. C., and Burkart, M. D. (2019). Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein-protein interactions. *Nat. Prod. Rep.* 37, 355–379. doi: 10.1039/c9np00047j

Jensen, K., Niederkruger, H., Zimmermann, K., Vagstad, A. L., Moldenhauer, J., Brendel, N., et al. (2012). Polyketide proofreading by an acyltransferase-like enzyme. *Chem. Biol.* 19, 329–339. doi: 10.1016/j.chembiol.2012.01.005

Jiang, W., Zhao, X., Gabrieli, T., Lou, C., Ebenstein, Y., and Zhu, T. F. (2015). Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* 6:8101. doi: 10.1038/ncomms9101

Kaljunen, H., Schiefelbein, S. H. H., Stummer, D., Kozak, S., Meijers, R., Christiansen, G., et al. (2015). Structural elucidation of the bispecificity of A domains as a basis for activating non-natural amino acids. *Angew. Chem. Int. Ed. Engl.* 54, 8833–8836. doi: 10.1002/anie.201503275

Kalkreuter, E., CroweTipton, J. M., Lowell, A. N., Sherman, D. H., and Williams, G. J. (2019). Engineering the substrate specificity of a modular polyketide synthase for installation of consecutive non-natural extender units. *J. Am. Chem. Soc.* 141, 1961–1969. doi: 10.1021/jacs.8b10521

Kasey, C. M., Zerrad, M., Li, Y., Cropp, T. A., and Williams, G. J. (2018). Development of transcription factor-based designer macrolide biosensors for metabolic engineering and synthetic biology. *ACS Synth. Biol.* 7, 227–239. doi: 10.1021/acssynbio.7b00287

Katano, H., Watanabe, H., Takakuwa, M., Maruyama, C., and Hamano, Y. (2013). Colorimetric determination of pyrophosphate anion and its application to adenylation enzyme assay. *Anal. Sci.* 29, 1095–1098. doi: 10.2116/analsci.29.1095

Keatinge-Clay, A. T. (2017). Polyketide synthase modules redefined. *Angew. Chem. Int. Ed. Engl.* 56, 4658–4660. doi: 10.1002/anie.201701281

Keatinge-Clay, A. T., and Stroud, R. M. (2006). The structure of a ketoreductase determines the organization of the β-carbon processing enzymes of modular polyketide synthases. *Structure* 14, 737–748. doi: 10.1016/j.str.2006.01.009

Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., and Cane, D. E. (2007). Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem.* 76, 195–221. doi: 10.1146/annurev.biochem.76.053105.093515

Kim, C. Y., Alekseyev, V. Y., Chen, A. Y., Tang, Y., Cane, D. E., and Khosla, C. (2004). Reconstituting modular activity from separated domains of 6-deoxyerythronolide B synthase. *Biochemistry* 43, 13892–13898. doi: 10.1021/bi048418n

Klapper, M., Braga, D., Lackner, G., Herbst, R., and Stallforth, P. (2018). Bacterial alkaloid biosynthesis: structural diversity via a minimalistic nonribosomal peptide synthetase. *Cell Chem. Biol.* 25, doi: 10.1016/j.chembiol.2018.02.013

Klaus, M., D'Souza, A. D., Nivina, A., Khosla, C., and Grininger, M. (2019). Engineering of chimeric polyketide synthases using SYNZIP docking domains. *ACS Chem. Biol.* 14, 426–433. doi: 10.1021/acschembio.8b01060

Klaus, M., and Grininger, M. (2018). Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* 35, 1070–1081. doi: 10.1039/c8np00030a

Klaus, M., Ostrowski, M. P., Austerjost, J., Robbins, T., Lowry, B., Cane, D. E., et al. (2016). Protein-protein interactions, not substrate recognition, dominate

the turnover of chimeric assembly line polyketide synthases. *J. Biol. Chem.* 291, 16404–16415. doi: 10.1074/jbc.M116.730531

Kohli, R. M., Trauger, J. W., Schwarzer, D., Marahiel, M. A., and Walsh, C. T. (2001). Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry* 40, 7099–7108. doi: 10.1021/bi010036j

Kohli, R. M., Walsh, C. T., and Burkart, M. D. (2002). Biomimetic synthesis and optimization of cyclic peptide antibiotics. *Nature* 418, 658–661. doi: 10.1038/nature00907

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946

Kornfuehrer, T., and Eustáquio, A. S. (2019). Diversification of polyketide structures via synthase engineering. *MedChemComm* 10, 1256–1272. doi: 10.1039/C9MD00141G

Koryakina, I., Kasey, C., McArthur, J. B., Lowell, A. N., Chemler, J. A., Li, S., et al. (2017). Inversion of extender unit selectivity in the erythromycin polyketide synthase by acyltransferase domain engineering. *ACS Chem. Biol.* 12, 114–123. doi: 10.1021/acschembio.6b00732

Kries, H., Wachtel, R., Pabst, A., Wanner, B., Niquille, D., and Hilvert, D. (2014). Reprogramming nonribosomal peptide synthetases for "clickable" amino acids. *Angew. Chem. Int. Ed. Engl.* 53, 10105–10108. doi: 10.1002/anie.201405281

Krug, D., and Muller, R. (2014). Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* 31, 768–783. doi: 10.1039/c3np70127a

Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W., and Rao, R. N. (1996). Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* 183, 231–236. doi: 10.1016/s0378-1119(96)00565-3

Kwan, D. H., Sun, Y., Schulz, F., Hong, H., Popovic, B., Sim-Stark, J. C., et al. (2008). Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases. *Chem. Biol.* 15, 1231–1240. doi: 10.1016/j.chembiol.2008.09.012

Ladner, C. C., and Williams, G. J. (2016). Harnessing natural product assembly lines: structure, promiscuity, and engineering. *J. Ind. Microbiol. Biotechnol.* 43, 371–387. doi: 10.1007/s10295-015-1704-8

Lee, N., Hwang, S., Lee, Y., Cho, S., Palsson, B., and Cho, B. K. (2019). Synthetic biology tools for novel secondary metabolite discovery in *Streptomyces*. *J. Microbiol. Biotechnol.* 29, 667–686. doi: 10.4014/jmb.1904.04015

Lee, N. C., Larionov, V., and Kouprina, N. (2015). Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res.* 43:e55. doi: 10.1093/nar/gkv112

Li, R., Oliver, R. A., and Townsend, C. A. (2017). Identification and characterization of the sulfazecin monobactam biosynthetic gene cluster. *Cell Chem. Biol.* 24, 24–34. doi: 10.1016/j.chembiol.2016.11.010

Li, Y., Zhang, W., Zhang, H., Tian, W., Wu, L., Wang, S., et al. (2018). Structural basis of a broadly selective acyltransferase from the polyketide synthase of splenocin. *Angew. Chem. Int. Ed. Engl.* 57, 5823–5827. doi: 10.1002/anie.201802805

Lim, J., Sun, H., Fan, J. S., Hameed, I. F., Lescar, J., Liang, Z. X., et al. (2012). Rigidifying acyl carrier protein domain in iterative type I PKS CalE8 does not affect its function. *Biophys. J.* 103, 1037–1044. doi: 10.1016/j.bpj.2012.08.006

Liu, Q., Wu, K., Cheng, Y., Lu, L., Xiao, E., Zhang, Y., et al. (2015). Engineering an iterative polyketide pathway in *Escherichia coli* results in single-form alkene and alkane overproduction. *Metab. Eng.* 28, 82–90. doi: 10.1016/j.ymben.2014.12.004

Liu, X., and Walsh, C. T. (2009). Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. *Biochemistry* 48, 8746–8757. doi: 10.1021/bi901123r

Lowden, P. A., Bohm, G. A., Metcalfe, S., Staunton, J., and Leadlay, P. F. (2004). New rapamycin derivatives by precursor-directed biosynthesis. *Chembiochem* 5, 535–538. doi: 10.1002/cbic.200300758

Lowry, B., Robbins, T., Weng, C. H., O'Brien, R. V., Cane, D. E., and Khosla, C. (2013). *In vitro* reconstitution and analysis of the 6-deoxyerythronolide B synthase. *J. Am. Chem. Soc.* 135, 16809–16812. doi: 10.1021/ja409048k

Lundy, T. A., Mori, S., and Garneau-Tsodikova, S. (2018). Engineering bifunctional enzymes capable of adenylating and selectively methylating the side chain

or core of amino acids. *ACS Synth. Biol.* 7, 399–404. doi: 10.1021/acssynbio.7b00426

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221

Marsden, A. F. A., Wilkinson, B., Cortés, J., Dunster, N. J., Staunton, J., and Leadlay, P. F. (1998). Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* 279, 199–202. doi: 10.1126/science.279.5348.199

McErlean, M., Overbay, J., and Van Lanen, S. (2019). Refining and expanding nonribosomal peptide synthetase function and mechanism. *J. Ind. Microbiol. Biotechnol.* 46, 493–513. doi: 10.1007/s10295-018-02130-w

Medema, M. H., and Fischbach, M. A. (2015). Computational approaches to natural product discovery. *Nat. Chem. Biol.* 11, 639–648. doi: 10.1038/nchembio.1884

Meluzzi, D., Zheng, W. H., Hensler, M., Nizet, V., and Dorrestein, P. C. (2008). Top-down mass spectrometry on low-resolution instruments: characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorg. Med. Chem. Lett.* 18, 3107–3111. doi: 10.1016/j.bmcl.2007.10.104

Menzella, H. G., Carney, J. R., and Santi, D. V. (2007). Rational design and assembly of synthetic trimodular polyketide synthases. *Chem. Biol.* 14, 143–151. doi: 10.1016/j.chembiol.2006.12.002

Menzella, H. G., Reid, R., Carney, J. R., Chandran, S. S., Reisinger, S. J., Patel, K. G., et al. (2005). Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* 23, 1171–1176. doi: 10.1038/nbt1128

Miller, B. R., Sundlov, J. A., Drake, E. J., Makin, T. A., and Gulick, A. M. (2014). Analysis of the linker region joining the adenylation and carrier protein domains of the modular nonribosomal peptide synthetases. *Proteins* 82, 2691–2702. doi: 10.1002/prot.24635

Miller, D. A., and Walsh, C. T. (2001). Yersiniabactin synthetase: probing the recognition of carrier protein domains by the catalytic heterocyclization domains, Cy1 and Cy2, in the chain-initiating HWMP2 subunit. *Biochemistry* 40, 5313–5321. doi: 10.1021/bi002905v

Mitchell, C. A., Shi, C., Aldrich, C. C., and Gulick, A. M. (2012). Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains. *Biochemistry* 51, 3252–3263. doi: 10.1021/bi300112e

Miyanaga, A., Iwasawa, S., Shinohara, Y., Kudo, F., and Eguchi, T. (2016). Structure-based analysis of the molecular interactions between acyltransferase and acyl carrier protein in vicenistatin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 113, 1802–1807. doi: 10.1073/pnas.1520042113

Mo, S., Kim, D. H., Lee, J. H., Park, J. W., Basnet, D. B., Ban, Y. H., et al. (2011). Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J. Am. Chem. Soc.* 133, 976–985. doi: 10.1021/ja108399b

Moran, S., Rai, D. K., Clark, B. R., and Murphy, C. D. (2009). Precursor-directed biosynthesis of fluorinated iturin A in *Bacillus* spp. *Org. Biomol. Chem.* 7, 644–646. doi: 10.1039/b816345f

Muchiri, R., and Walker, K. D. (2012). Taxol biosynthesis: tyrocidine synthetase A catalyzes the production of phenylisoserinyl CoA and other amino phenylpropanoyl thioesters. *Chem. Biol.* 19, 679–685. doi: 10.1016/j.chembiol.2012.05.007

Muchiri, R., and Walker, K. D. (2017). Paclitaxel biosynthesis: adenylation and thiolation domains of an NRPS TycA PheAT module produce various arylisoserine CoA thioesters. *Biochemistry* 56, 1415–1425. doi: 10.1021/acs.biochem.6b01188

Murphy, A. C., Hong, H., Vance, S., Broadhurst, R. W., and Leadlay, P. F. (2016). Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chem. Commun.* 52, 8373–8376. doi: 10.1039/c6cc03501a

Musiol-Kroll, E. M., and Wohlleben, W. (2018). Acyltransferases as tools for polyketide synthase engineering. *Antibiotics* 7:62. doi: 10.3390/antibiotics7030062

Mutka, S. C., Bondi, S. M., Carney, J. R., Da Silva, N. A., and Kealey, J. T. (2006). Metabolic pathway engineering for complex polyketide biosynthesis in

Saccharomyces cerevisiae. FEMS Yeast Res. 6, 40–47. doi: 10.1111/j.1567-1356. 2005.00001.x

Navarro-Munoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., et al. (2020). A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* 16, 60–68. doi: 10.1038/s41589-019-0400-9

Nguyen, K. T., He, X., Alexander, D. C., Li, C., Gu, J. Q., Mascio, C., et al. (2010). Genetically engineered lipopeptide antibiotics related to A54145 and daptomycin with improved properties. *Antimicrob. Agents Chemother.* 54, 1404–1413. doi: 10.1128/AAC.01307-09

Nguyen, K. T., Ritz, D., Gu, J. Q., Alexander, D., Chu, M., Miao, V., et al. (2006). Combinatorial biosynthesis of novel antibiotics related to daptomycin. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17462–17467. doi: 10.1073/pnas.0608589103

Niu, W., Draths, K. M., and Frost, J. W. (2002). Benzene-free synthesis of adipic acid. *Biotechnol. Prog.* 18, 201–211. doi: 10.1021/bp010179x

Oliynyk, M., Brown, M. J., Cortes, J., Staunton, J., and Leadlay, P. F. (1996). A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.* 3, 833–839. doi: 10.1016/s1074-5521(96)90069-1

Pang, B., Valencia, L. E., Wang, J., Wan, Y., Lal, R., Zargar, A., et al. (2019). Technical advances to accelerate modular type I polyketide synthase engineering towards a retro-biosynthetic platform. *Biotechnol. Bioprocess Eng.* 24, 413–423.

Patel, K., Piagentini, M., Rascher, A., Tian, Z. Q., Buchanan, G. O., Regentin, R., et al. (2004). Engineered biosynthesis of geldanamycin analogs for Hsp90 inhibition. *Chem. Biol.* 11, 1625–1633. doi: 10.1016/j.chembiol.2004.09.012

Petkovic, H., Lill, R. E., Sheridan, R. M., Wilkinson, B., McCormick, E. L., McArthur, H. A., et al. (2003). A novel erythromycin, 6-desmethyl erythromycin D, made by substituting an acyltransferase domain of the erythromycin polyketide synthase. *J. Antibiot.* 56, 543–551. doi: 10.7164/antibiotics.56.543

Petkovic, H., Sandmann, A., Challis, I. R., Hecht, H. J., Silakowski, B., Low, L., et al. (2008). Substrate specificity of the acyl transferase domains of EpoC from the epothilone polyketide synthase. *Org. Biomol. Chem.* 6, 500–506. doi: 10.1039/b714804f

Pinto, A., Wang, M., Horsman, M., and Boddy, C. N. (2012). 6-Deoxyerythronolide B synthase thioesterase-catalyzed macrocyclization is highly stereoselective. *Org. Lett.* 14, 2278–2281. doi: 10.1021/ol300707j

Qiao, K., Zhou, H., Xu, W., Zhang, W., Garg, N., and Tang, Y. (2011). A fungal nonribosomal peptide synthetase module that can synthesize thiopyrazines. *Org. Lett.* 13, 1758–1761. doi: 10.1021/ol200288w

Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R., and McDaniel, R. (2001). Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* 40, 15464–15470. doi: 10.1021/bi015864r

Regentin, R., Kennedy, J., Wu, N., Carney, J. R., Licari, P., Galazzo, J., et al. (2004). Precursor-directed biosynthesis of novel triketide lactones. *Biotechnol. Prog.* 20, 122–127. doi: 10.1021/bp0341949

Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., et al. (2003). A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry* 42, 72–79. doi: 10.1021/bi0268706

Richter, C. D., Nietlispach, D., Broadhurst, R. W., and Weissman, K. J. (2008). Multienzyme docking in hybrid megasynthetases. *Nat. Chem. Biol.* 4, 75–81. doi: 10.1038/nchembio.2007.61

Rottig, M., Medema, M. H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011). NRPSpredictor2–a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* 39, W362–W367. doi: 10.1093/nar/gkr323

Rowe, C. J., Bohm, I. U., Thomas, I. P., Wilkinson, B., Rudd, B. A., Foster, G., et al. (2001). Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chem. Biol.* 8, 475–485. doi: 10.1016/s1074-5521(01)00024-2

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5

Rutledge, P. J., and Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* 13, 509–523. doi: 10.1038/nrmicro3496

Samel, S. A., Schoenafinger, G., Knappe, T. A., Marahiel, M. A., and Essen, L. O. (2007). Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase. *Structure* 15, 781–792. doi: 10.1016/j.str.2007.05.008

Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., and Karsch-Mizrachi, I. (2020). GenBank. *Nucleic Acids Res.* 48, D84–D86. doi: 10.1093/nar/gkz956

Schauwecker, F., Pfennig, F., Grammel, N., and Keller, U. (2000). Construction and *in vitro* analysis of a new bi-modular polypeptide synthetase for synthesis of N-methylated acyl peptides. *Chem. Biol.* 7, 287–297. doi: 10.1016/s1074-5521(00)00103-4

Schneider, T. L., Shen, B., and Walsh, C. T. (2003). Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation. *Biochemistry* 42, 9722–9730. doi: 10.1021/bi034792w

Schoenafinger, G., Schracke, N., Linne, U., and Marahiel, M. A. (2006). Formylation domain: an essential modifying enzyme for the nonribosomal biosynthesis of linear gramicidin. *J. Am. Chem. Soc.* 128, 7406–7407. doi: 10.1021/ja0611240

Skiba, M. A., Tran, C. L., Dan, Q., Sikkema, A. P., Klaver, Z., Gerwick, W. H., et al. (2020). Repurposing the GNAT Fold in the Initiation of Polyketide Biosynthesis. *Structure* 28, 63–74.e4. doi: 10.1016/j.str.2019.11.004

Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 6, 493–505. doi: 10.1016/S1074-5521(99)80082-9

Stachelhaus, T., Schneider, A., and Marahiel, M. A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* 269, 69–72. doi: 10.1126/science.7604280

Stanisic, A., and Kries, H. (2019). Adenylation domains in nonribosomal peptide engineering. *Chembiochem* 20, 1347–1356. doi: 10.1002/cbic.201800750

Stassi, D. L., Kakavas, S. J., Reynolds, K. A., Gunawardana, G., Swanson, S., Zeidner, D., et al. (1998). Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7305–7309. doi: 10.1073/pnas.95.13.7305

Staunton, J., and Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 18, 380–416. doi: 10.1039/a909079g

Stein, D. B., Linne, U., Hahn, M., and Marahiel, M. A. (2006). Impact of epimerization domains on the intermodular transfer of enzyme-bound intermediates in nonribosomal peptide synthesis. *Chembiochem* 7, 1807–1814. doi: 10.1002/cbic.200600192

Sugimoto, Y., Ishida, K., Traitcheva, N., Busch, B., Dahse, H. M., and Hertweck, C. (2015). Freedom and constraint in engineered noncolinear polyketide assembly lines. *Chem. Biol.* 22, 229–240. doi: 10.1016/j.chembiol.2014.12.014

Sukovich, D. J., Seffernick, J. L., Richman, J. E., Gralnick, J. A., and Wackett, L. P. (2010). Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Appl. Environ. Microbiol.* 76, 3850–3862. doi: 10.1128/AEM.00436-10

Sundlov, J. A., Shi, C., Wilson, D. J., Aldrich, C. C., and Gulick, A. M. (2012). Structural and functional investigation of the intermolecular interaction between NRPS adenylation and carrier protein domains. *Chem. Biol.* 19, 188–198. doi: 10.1016/j.chembiol.2011.11.013

Sussmuth, R. D., and Mainz, A. (2017). Nonribosomal peptide synthesis-Principles and prospects. *Angew. Chem. Int. Ed. Engl.* 56, 3770–3821. doi: 10.1002/anie.201609079

Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006). The 2.7-Angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11124–11129. doi: 10.1073/pnas.0601924103

Tanovic, A., Samel, S. A., Essen, L. O., and Marahiel, M. A. (2008). Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* 321, 659–663. doi: 10.1126/science.1159850

Tarry, M. J., Haque, A. S., Bui, K. H., and Schmeing, T. M. (2017). X-ray crystallography and electron microscopy of cross- and multi-module nonribosomal peptide synthetase proteins reveal a flexible architecture. *Structure* 25 78, 783–793.e4. doi: 10.1016/j.str.2017.03.014

Thirlway, J., Lewis, R., Nunns, L., Al Nakeeb, M., Styles, M., Struck, A.-W., et al. (2012). Introduction of a non-natural amino acid into a nonribosomal peptide antibiotic by modification of adenylation domain specificity. *Angew. Chem. Int. Ed. Engl.* 51, 7181–7184. doi: 10.1002/anie.201202043

Tibrewal, N., and Tang, Y. (2014). Biocatalysts for natural product biosynthesis. *Annu. Rev. Chem. Biomol. Eng.* 5, 347–366. doi: 10.1146/annurev-chembioeng-060713-040008

Trauger, J. W., Kohli, R. M., and Walsh, C. T. (2001). Cyclization of backbone-substituted peptides catalyzed by the thioesterase domain from the tyrocidine nonribosomal peptide synthetase. *Biochemistry* 40, 7092–7098. doi: 10.1021/bi010035r

Tripathi, A., Choi, S. S., Sherman, D. H., and Kim, E. S. (2016). Thioesterase domain swapping of a linear polyketide tautomycetin with a macrocyclic polyketide pikromycin in *Streptomyces* sp. CK4412. *J. Ind. Microbiol. Biotechnol.* 43, 1189–1193. doi: 10.1007/s10295-016-1790-2

Van Lanen, S. G., Lin, S., and Shen, B. (2008). Biosynthesis of the enediyne antitumor antibiotic C-1027 involves a new branching point in chorismate metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 105, 494–499. doi: 10.1073/pnas.0708750105

Velkov, T., Horne, J., Scanlon, M. J., Capuano, B., Yuriev, E., and Lawen, A. (2011). Characterization of the *N*-methyltransferase activities of the multifunctional polypeptide cyclosporin synthetase. *Chem. Biol.* 18, 464–475. doi: 10.1016/j.chembiol.2011.01.017

Villiers, B., and Hollfelder, F. (2011). Directed evolution of a gatekeeper domain in nonribosomal peptide synthesis. *Chem. Biol.* 18, 1290–1299. doi: 10.1016/j.chembiol.2011.06.014

Villiers, B. R., and Hollfelder, F. (2009). Mapping the limits of substrate specificity of the adenylation domain of TycA. *Chembiochem* 10, 671–682. doi: 10.1002/cbic.200800553

Vogeli, B., Geyer, K., Gerlinger, P. D., Benkstein, S., Cortina, N. S., and Erb, T. J. (2018). Combining promiscuous acyl-CoA oxidase and enoyl-CoA carboxylase/reductases for atypical polyketide extender unit biosynthesis. *Cell Chem. Biol.* 25, 833–839.e4. doi: 10.1016/j.chembiol.2018.04.009

Walsh, C. T. (2016). Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. *Nat. Prod. Rep.* 33, 127–135. doi: 10.1039/c5np00035a

Wang, F., Wang, Y., Ji, J., Zhou, Z., Yu, J., Zhu, H., et al. (2015). Structural and functional analysis of the loading acyltransferase from avermectin modular polyketide synthase. *ACS Chem. Biol.* 10, 1017–1025. doi: 10.1021/cb500873k

Ward, S. L., Desai, R. P., Hu, Z., Gramajo, H., and Katz, L. (2007). Precursor-directed biosynthesis of 6-deoxyerythronolide B analogues is improved by removal of the initial catalytic sites of the polyketide synthase. *J. Ind. Microbiol. Biotechnol.* 34, 9–15. doi: 10.1007/s10295-006-0156-6

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Weber, T., Charusanti, P., Musiol-Kroll, E. M., Jiang, X., Tong, Y., Kim, H. U., et al. (2015). Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. *Trends Biotechnol.* 33, 15–26. doi: 10.1016/j.tibtech.2014.10.009

Wehrs, M., Gladden, J. M., Liu, Y., Platz, L., Prahl, J.-P., Moon, J., et al. (2019). Sustainable bioproduction of the blue pigment indigoidine: expanding the range of heterologous products in R. toruloides to include non-ribosomal peptides. *Green Chem.* 21, 3394–3406. doi: 10.1039/C9GC00920E

Weissman, K. J. (2016). Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.* 33, 203–230. doi: 10.1039/c5np00109a

Weist, S., Bister, B., Puk, O., Bischoff, D., Pelzer, S., Nicholson, G. J., et al. (2002). Fluorobalhimycin–a new chapter in glycopeptide antibiotic research. *Angew. Chem. Int. Ed. Engl.* 41, 3383–3385. doi: 10.1002/1521-3773(20020916)41:18<3383::AID-ANIE3383>3.0.CO;2-R

Weist, S., Kittel, C., Bischoff, D., Bister, B., Pfeifer, V., Nicholson, G. J., et al. (2004). Mutasynthesis of glycopeptide antibiotics: variations of vancomycin's AB-ring amino acid 3,5-dihydroxyphenylglycine. *J. Am. Chem. Soc.* 126, 5942–5943. doi: 10.1021/ja0499389

Wenda, S., Illner, S., Mell, A., and Kragl, U. (2011). Industrial biotechnology—the future of green chemistry? *Green Chem.* 13, 3007–3047. doi: 10.1039/C1GC15579B

Whicher, J. R., Dutta, S., Hansen, D. A., Hale, W. A., Chemler, J. A., Dosey, A. M., et al. (2014). Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* 510, 560–564. doi: 10.1038/nature13409

Wilson, D. J., and Aldrich, C. C. (2010). A continuous kinetic assay for adenylation enzyme activity and inhibition. *Anal. Biochem.* 404, 56–63. doi: 10.1016/j.ab.2010.04.033

Winn, M., Fyans, J. K., Zhuo, Y., and Micklefield, J. (2016). Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* 33, 317–347. doi: 10.1039/c5np00099h

Worthington, A. S., Rivera, H., Torpey, J. W., Alexander, M. D., and Burkart, M. D. (2006). Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis. *ACS Chem. Biol.* 1, 687–691. doi: 10.1021/cb6003965

Xie, Y., Cai, Q., Ren, H., Wang, L., Xu, H., Hong, B., et al. (2014). NRPS substrate promiscuity leads to more potent antitubercular sansanmycin analogues. *J. Nat. Prod.* 77, 1744–1748. doi: 10.1021/np5001494

Xu, F., Butler, R., May, K., Rexhepaj, M., Yu, D., Zi, J., et al. (2019). Modified substrate specificity of a methyltransferase domain by protein insertion into an adenylation domain of the bassianolide synthetase. *J. Biol. Eng.* 13:65. doi: 10.1186/s13036-019-0195-y

Yakimov, M. M., Giuliano, L., Timmis, K. N., and Golyshin, P. N. (2000). Recombinant acylheptapeptide lichenysin: high level of production by *Bacillus subtilis* cells. *J. Mol. Microbiol. Biotechnol.* 2, 217–224.

Yan, J., Hazzard, C., Bonnett, S. A., and Reynolds, K. A. (2012). Functional modular dissection of DEBS1-TE changes triketide lactone ratios and provides insight into Acyl group loading, hydrolysis, and ACP transfer. *Biochemistry* 51, 9333–9341. doi: 10.1021/bi300830q

Ye, Z., Musiol, E. M., Weber, T., and Williams, G. J. (2014). Reprogramming acyl carrier protein interactions of an Acyl-CoA promiscuous trans-acyltransferase. *Chem. Biol.* 21, 636–646. doi: 10.1016/j.chembiol.2014.02.019

Ye, Z., and Williams, G. J. (2014). Mapping a ketosynthase:acyl carrier protein binding interface via unnatural amino acid-mediated photo-cross-linking. *Biochemistry* 53, 7494–7502. doi: 10.1021/bi500936u

Yu, D., Xu, F., Zeng, J., and Zhan, J. (2012). Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* 64, 285–295. doi: 10.1002/iub.1005

Yuzawa, S., Bailey, C. B., Fujii, T., Jocic, R., Barajas, J. F., Benites, V. T., et al. (2017a). Heterologous gene expression of N-terminally truncated variants of LipPks1 suggests a functionally critical structural motif in the N-terminus of modular polyketide synthase. *ACS Chem. Biol.* 12, 2725–2729. doi: 10.1021/acschembio.7b00714

Yuzawa, S., Deng, K., Wang, G., Baidoo, E. E., Northen, T. R., Adams, P. D., et al. (2017b). Comprehensive *in vitro* analysis of acyltransferase domain Exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synth. Biol.* 6, 139–147. doi: 10.1021/acssynbio.6b00176

Yuzawa, S., Kapur, S., Cane, D. E., and Khosla, C. (2012). Role of a conserved arginine residue in linkers between the ketosynthase and acyltransferase domains of multimodular polyketide synthases. *Biochemistry* 51, 3708–3710. doi: 10.1021/bi300399u

Yuzawa, S., Keasling, J. D., and Katz, L. (2017c). Bio-based production of fuels and industrial chemicals by repurposing antibiotic-producing type I modular polyketide synthases: opportunities and challenges. *J. Antibiot.* 70, 378–385. doi: 10.1038/ja.2016.136

Yuzawa, S., Mirsiaghi, M., Jocic, R., Fujii, T., Masson, F., Benites, V. T., et al. (2018). Short-chain ketone production by engineered polyketide synthases in *Streptomyces albus. Nat. Commun.* 9:4569. doi: 10.1038/s41467-018-07040-0

Zargar, A., Barajas, J. F., Lal, R., and Keasling, J. D. (2018). Polyketide synthases as a platform for chemical product design. *AIChE J.* 64, 4201–4207. doi: 10.1002/aic.16351

Zhang, J., Van Lanen, S. G., Ju, J., Liu, W., Dorrestein, P. C., Li, W., et al. (2008). A phosphopantetheinylating polyketide synthase producing a linear polyene to initiate enediyne antitumor antibiotic biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1460–1465. doi: 10.1073/pnas.0711625105

Zhang, K., Nelson, K. M., Bhuripanyo, K., Grimes, K. D., Zhao, B., Aldrich, C. C., et al. (2013). Engineering the substrate specificity of the DhbE adenylation domain by yeast cell surface display. *Chem. Biol.* 20, 92–101. doi: 10.1016/j.chembiol.2012.10.020

Zhang, L., Hashimoto, T., Qin, B., Hashimoto, J., Kozone, I., Kawahara, T., et al. (2017). Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolyol polyketides. *Angew. Chem. Int. Ed. Engl.* 56, 1740–1745. doi: 10.1002/anie.201611371

Zhang, L., Ji, J., Yuan, M., Feng, Y., Wang, L., Deng, Z., et al. (2018). Stereospecificity of enoylreductase domains from modular polyketide synthases. *ACS Chem. Biol.* 13, 871–875. doi: 10.1021/acschembio.7b00982

Zheng, J., Piasecki, S. K., and Keatinge-Clay, A. T. (2013). Structural studies of an A2-type modular polyketide synthase ketoreductase reveal features controlling α-substituent stereochemistry. *ACS Chem. Biol.* 8, 1964–1971. doi: 10.1021/cb400161g

Zhou, Y., Li, J., Zhu, J., Chen, S., Bai, L., Zhou, X., et al. (2008). Incomplete β-Ketone processing as a mechanism for polyene structural variation in the FR-008/candicidin complex. *Chem. Biol.* 15, 629–638. doi: 10.1016/j.chembiol.2008.05.007

Check for updates

Rational Drug Design for *Pseudomonas aeruginosa* PqsA Enzyme: An *in silico* Guided Study to Block Biofilm Formation

Bilal Shaker[1], Sajjad Ahmad[2], Thi Duc Thai[1], Seong-il Eyun[3] and Dokyun Na[1]*

[1] 84 Heukseok-ro, Dongjak-gu, Department of Biomedical Engineering, Chung-Ang University, Seoul, South Korea, [2] National Centre for Bioinformatics, Quaid-i-Azam University, Islamabad, Pakistan, [3] 84 Heukseok-ro, Dongjak-gu, Department of Life Science, Chung-Ang University, Seoul, South Korea

Pseudomonas aeruginosa is an opportunistic gram-negative bacterium implicated in acute and chronic nosocomial infections and a leading cause of patient mortality. Such infections occur owing to biofilm formation that confers multidrug resistance and enhanced pathogenesis to the bacterium. In this study, we used a rational drug design strategy to inhibit the quorum signaling system of *P. aeruginosa* by designing potent inhibitory lead molecules against anthranilate-CoA ligase enzyme encoded by the *pqsA* gene. This enzyme produces autoinducers for cell-to-cell communication, which result in biofilm formation, and thus plays a pivotal role in the virulence of *P. aeruginosa*. A library of potential drug molecules was prepared by performing ligand-based screening using an available set of enzyme inhibitors. Subsequently, structure-based virtual screening was performed to identify compounds showing the best binding conformation with the target enzyme and forming a stable complex. The two hit compounds interact with the binding site of the enzyme through multiple short-range hydrophilic and hydrophobic interactions. Molecular dynamic simulation and MM-PBSA/GBSA results to calculate the affinity and stability of the hit compounds with the PqsA enzyme further confirmed their strong interactions. The hit compounds might be useful in tackling the resistant phenotypes of this pathogen.

Keywords: multidrug resistance, biofilm, quorum sensing, drug discovery, structure-based virtual screening

INTRODUCTION

Pseudomonas aeruginosa is a gram-negative opportunistic bacterial pathogen that poses a significant threat to patients in hospital environments (Page and Heim, 2009). The World Health Organization (WHO) classified *P. aeruginosa* as the highest priority pathogen and declared an urgent demand for new antibiotics (Shrivastava et al., 2018). *P. aeruginosa* is extremely resistant to antibiotics due to intrinsic, evolved, and acquired mechanisms, such as decreased cell permeability, enzymatic drug inactivation via horizontal gene transfer, and biofilm development (Bonomo and Szabo, 2006; Zavascki et al., 2010). Moreover, the β-lactamase gene present in its genome makes cephalosporins and penicillin ineffective (Strateva and Yordanov, 2009). It has been estimated that

biofilm-forming *P. aeruginosa* is responsible for 65% of patient mortality and antibiotic resistance (Strateva and Yordanov, 2009).

Biofilm-embedded bacteria are 1,000 times more resistant to antibiotics than planktonic bacteria, and the biofilm also enables the bacteria to avoid a host's immune system (Gaddy and Actis, 2009). In a biofilm, bacterial cells are embedded inside the matrix of extracellular polymeric substances (EPSs) composed of proteins, exopolysaccharides, DNA, macromolecules, and lipids (Flemming and Wingender, 2001). All of these render the antibiotics impermeable and ineffective. The presence of persistent cells in biofilms also significantly contributes to their multidrug resistance property (Lewis, 2007). Overall, the complex morphology of biofilms is a significant barrier in designing potent therapeutic agents that can eradicate biofilm-associated infections successfully.

Quorum sensing (QS), a cell-to-cell communication system, reportedly plays a pivotal role in establishing persistent infections (Hentzer et al., 2003; Alhede et al., 2009; Van Gennip et al., 2009; Chiang et al., 2013). QS relies on the processing, secreting, and sensing of small diffusible quorum sensing signal molecules (QSSMs), known as autoinducers (Waters and Bassler, 2005). When a bacterial community reaches a certain threshold (density), which is reflected by AI concentration in the surrounding environment, bacterial gene transcription in the community becomes synchronized, which enables the community to behave collectively. A wide range of activities is controlled by AIs, including virulence factor secretion, swimming motility, secondary metabolite production, biofilm maturation, and antibiotic resistance (Ng and Bassler, 2009). A promising approach to control the growth of this pathogen is to interfere with QS-mediated signaling, which disrupts bacterial communication and attenuates virulence so that the bacteria can be eliminated from the host (Mühlen and Dersch, 2016; Wagner et al., 2016; Kamal et al., 2017). Therefore, QS inhibitors can be very useful and effective as a prophylactic to control antibiotic-resistant bacteria.

Pseudomonas aeruginosa has three major QS systems (*rhl, las,* and *pqs*) that mediate cell-to-cell communication and control the synthesis and secretion of virulence factors (Williams and Cámara, 2009; Jimenez et al., 2012; **Figure 1**). For instance, the *las* system positively regulates both the *rhl* and *pqs* systems by initiating the expression of AI receptors (RhlR and PqsR). RhlR and PqsR are also transcriptional activators when their respective AIs are bound. The *las* and *rhl* systems use two different AIs (acyl-homoserine-lactones, AHLs) (Alhede et al., 2009). The *pqs* system employs two signal molecules: 2-heptyl-3-hydroxy-4(1H)-quinoline or *Pseudomonas* quinolone signal (PQS) (Pesci et al., 1999) and its biosynthetic precursor 2-heptyl-4-hydroxyquinoline (HHQ) (Xiao et al., 2006). When PQS or HHQ molecules are bound to the PqsR transcriptional activator, PqsR induces the expression of various virulence genes, their biosynthetic genes, and biofilm formation-related genes (Calfee et al., 2001; Cao et al., 2001; Xiao et al., 2006). Although both PQS and HHQ bind to and activate PqsR, PQS is 100-fold more potent than HHQ (Xiao et al., 2006). Because the enzyme (PqsA) is responsible for the synthesis of the PQS signal

molecule (Grandclément et al., 2016), inhibition of the PqsA enzyme can disrupt biosynthesis of PQS signal molecule and consequently PqsR-dependent gene regulation and ultimately biofilm formation. The PQS act as a linker of the *las* and *rhl* quorum sensing systems in *Pseudomonas aeruginosa* and it has been shown that PQS system showed activity even in the *lasR* mutant strains, which represents that PQS could be a more significant modulator in the quorum sensing (McKnight et al., 2000). Also, PQS system is a broad regulatory system in *P. aeruginosa* influencing iron acquisition, outer membrane vesicle production, cytotoxicity through oxidative stress, and immune responses in the host cell (van Kessel, 2019).

To date, several SAR studies have been conducted to identify the potential inhibitors for QS system in *P. aeruginosa* and also provide important scaffolds for future QS inhibitor development (Ji et al., 2016; Fong et al., 2017). According to reported studies, sulfur-containing compounds possess excellent QS inhibiting properties and can demolish the bacteria virulence.

In this study, we predicted putative QS inhibitor based on *in silico* drug discovery technology. We prepared a library of 521 compounds containing dialogs of reported inhibitors. Further, their binding affinities were predicted through molecular docking to screen the most potent binder. Subsequently, Molecular Dynamic (MD) simulation was performed to validate the screened putative QS inhibitors affinity for the docked site and understand the dynamic behavior of the enzyme. Finally binding free energy for hit drug compounds was calculated with the MM-PBSA/GBSA approach (Miller et al., 2012), which leads to a possible mechanistic calculation useful for future experimental analysis. Our *in silico* results support the prediction that these predicted compounds can efficiently suppress the PqsA QS system and block virulence.

MATERIALS AND METHODS

Preparation of Enzyme Structure

To understand the interactions of the enzyme with ligands, the 3D structure of the target enzyme is required. Although the PqsA enzyme structure has not been experimentally proven, we subjected the protein sequence of *P. aeruginosa* PqsA (UniprotKB ID: Q9l4 × 3) to homology modeling using MODELLER9.24 (Eswar et al., 2007). In structural biology, comparative modeling or homology modeling is the most promising technology for significantly narrowing down the gap between experimentally determined structures and known protein sequences (Xiang, 2006). Two templates with maximum sequence identity and query coverage were used for better homology structure prediction (Template 1: PDB ID = 5OE3, query coverage = 77%, sequence identity = 100% and Template 2: PDB ID = 1ULT, query coverage = 96%, sequence identity = 25%). Multiple sequence alignment (MSA) of PqsA and the two templates was performed (**Supplementary Figure S1**). To date, only N terminal domain of the enzyme is determined due to the difficulty in getting the full-length protein structure. In this study, by using a close homology template we predicted the remaining structure of the PqsA in addition to the available N-terminal domain.

FIGURE 1 | Schematic diagram of rhl, pqs, and las QS systems in *P. aeruginosa.* The three major QS systems use C4-HSL, 3-oxo-C12-HSL, and HHQ/PQS for cell-to-cell communication. In addition, the three QS systems can interact with each other for facilitating a complex regulation of the physiology of *P. aeruginosa.*
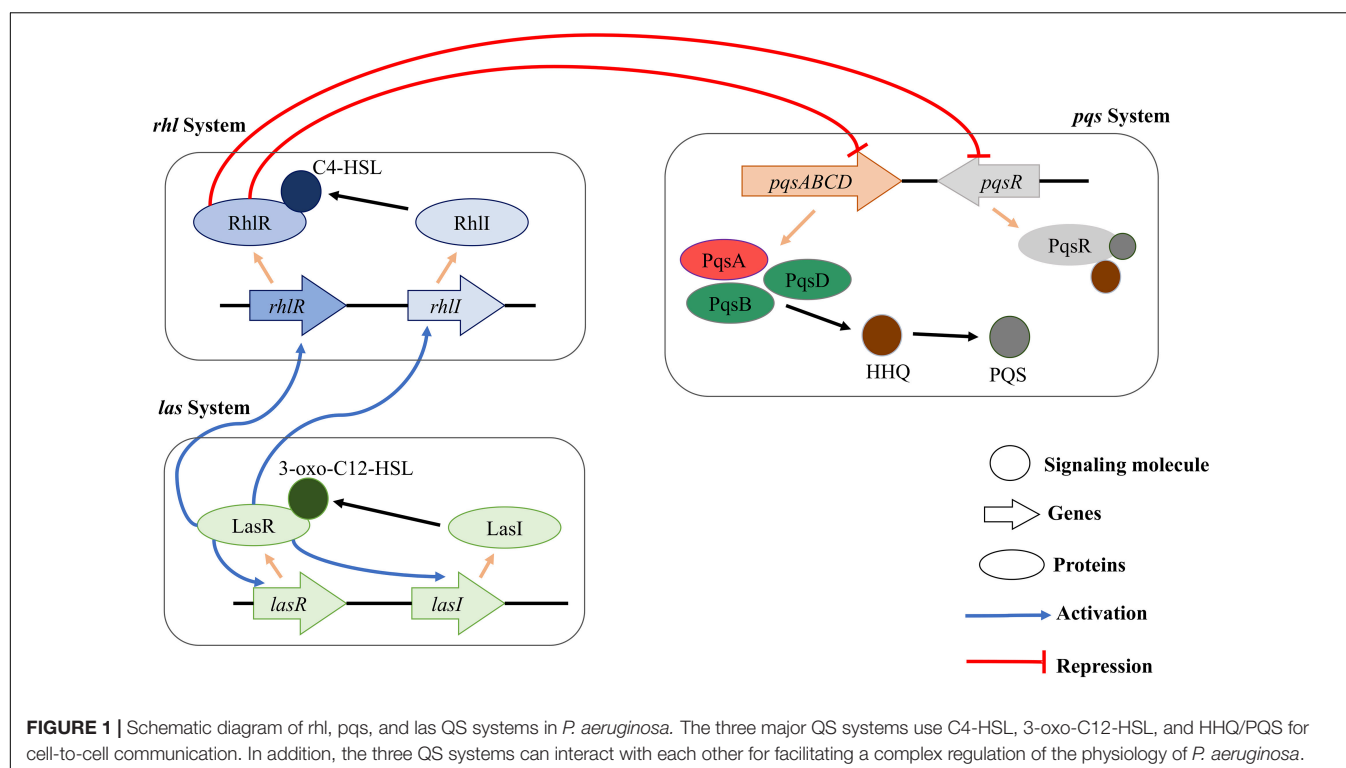
It is significant to highlight the dynamic aspect of full-length protein upon ligand or drug molecules binding (Itoh et al., 2015; Flydal et al., 2019). However, in this study the C-terminal region did not affect the binding pocket of the N-terminal domain, we used only the N-terminal domain structure for putative inhibitor screening. Different online servers were used for structure assessment, including the Swiss-Model structure assessment tool (Waterhouse et al., 2018), Ramachandran plot (Gopalakrishnan et al., 2007), and PDBsum (De Beer et al., 2014). University of California San Francisco (UCSF) Chimera (version 1.14) (Pettersen et al., 2004) was used for structural visualization and energy minimization of a total 1,500 steps divided into the first 750 steps of the steepest descent and the last 750 steps of the conjugate gradient (Ahmad et al., 2017). During minimization, standard residues were treated with AMBER ff14SB force-field.

QS Inhibitor Preparation

The QS inhibitors of PqsA were selected from those reported across various studies in the literature (Ji et al., 2016; Fong et al., 2017) and used as positive controls in this study. Moreover, 521 structural analogs to the reported inhibitors were generated from online repositories, such as ChemSpider (Ayers, 2012), ChEMBL (Davies et al., 2015), and PubChem (Kim et al., 2019). These web-servers provide a search module where users can find the available structures having a similar chemical composition (similar scaffolds) of the query structure. The similarity index was set to >85%, compounds having more than 85% structural similarity were selected for analog library. Discovery Studio (DS) and UCSF Chimera were used for ligand preprocessing, including protonation; ionization; and addition of

explicit counter ions, hydrogen atoms, or atomic partial charges. Energy minimization was performed using forcefield AMBER ff14SB for small molecules. The refined dataset was further utilized for computational experiments.

Here, the previously reported QS inhibitors of PqsA were used for conducting a docking study to reveal their binding affinity toward the target enzyme. ChemDraw (Mendelsohn, 2004), a chemical drawing tool, was used to draw 2D structures of the compounds and to convert the compounds into 3D structures. The compounds were eventually subjected to ligand preprocessing.

Ligand Binding Domain Analysis

To date, only the functional N-terminal domain of PqsA has been analyzed (Witzgall et al., 2017). For a brief binding pocket analysis, Witzgall et al. compared the N-terminal binding domain of PqsA from *P. aeruginosa* with the other CoA ligase-like anthraniloyl-CoA transacylase AuaE from *Stigmatella aurantiaca* (Sandmann et al., 2007), the PqsA ortholog HmqA from *Burkholderia pseudomallei* and *B. ambifaria* (Diggle et al., 2006; Vial et al., 2008), and an anthranilate-CoA from *Azoarcus evansii* (Schühle et al., 2001). According to the reported study, PqsA possesses highly conserved residues when compared with other aryl-CoA ligases, such as Gly307, Gly302, Ala278, Gly279, His308, and Tyr211.

Molecular Docking

Molecular docking was performed on the analog library through the PyRx interface (Dallakyan and Olson, 2015) of AutoDock Vina (Trott and Olson, 2010). During docking, the following
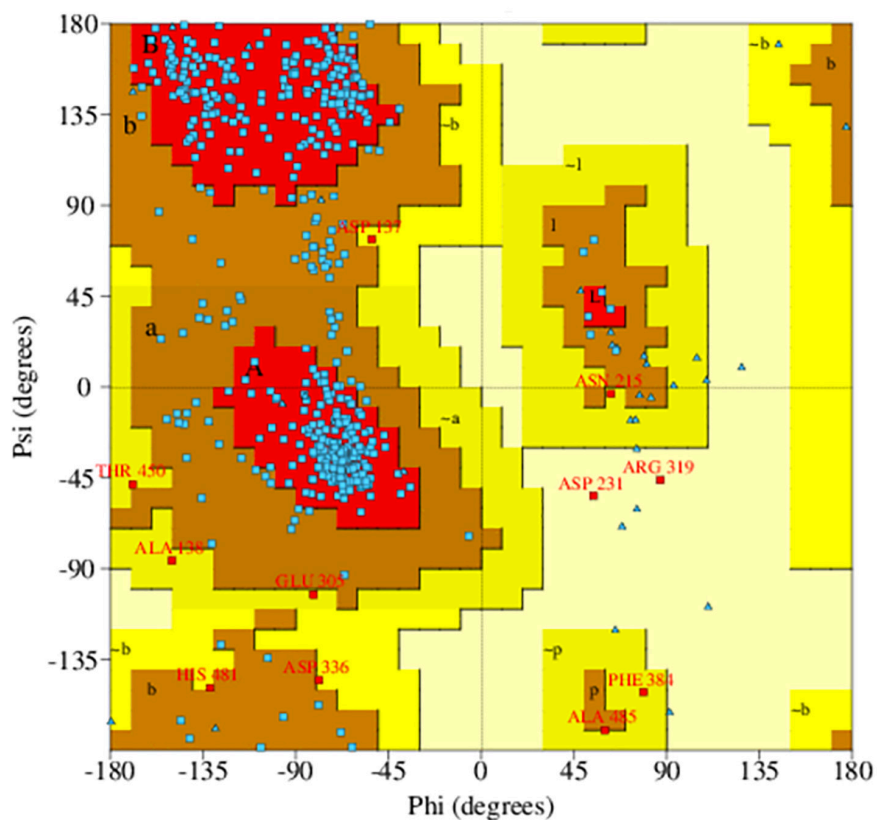
FIGURE 2 | Ramachandran plot of predicted PqsA structure. The blue dots indicate torsion angles distributed along the core secondary structure regions of the enzyme (shown in red). Yellow and pale-yellow areas contain favored and generously allowed torsion angles, respectively.
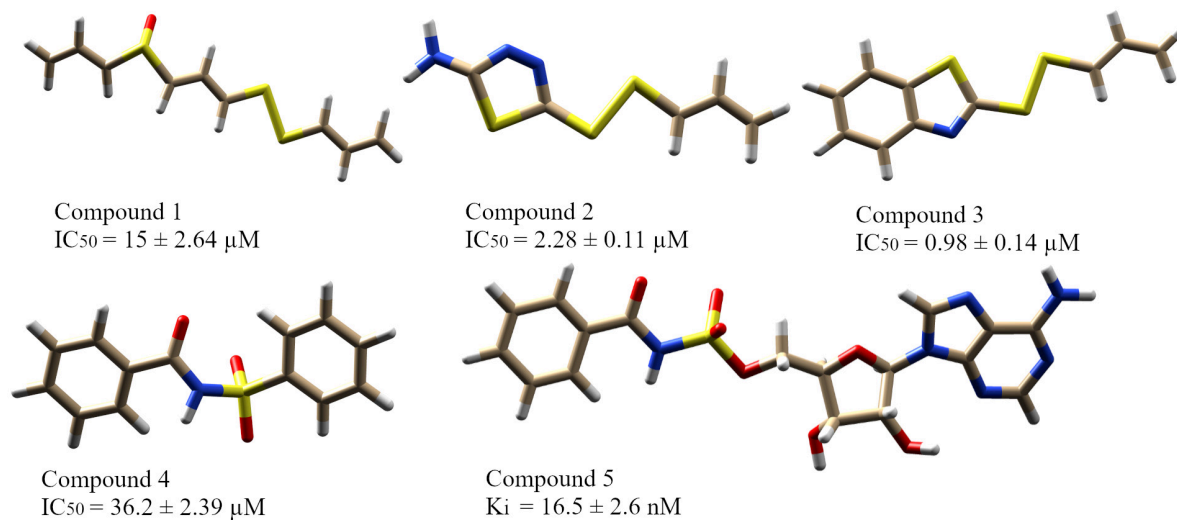
FIGURE 3 | The five reported PqsA inhibitors. The compounds are displayed in a stick representation along with their respective inhibitory constant values.

parameters were applied: exhaustiveness was set to 300, residues of the N-terminal domain involved in the binding pocket were established as the binding site, and the maximum number of poses was set to 300. The PyRx tool generates binding affinity values in the negative (a larger negative value implies a stronger binding affinity. The top 300 inhibitors showing a high binding affinity were retrieved for further analysis. The docking results of the top 10 inhibitors are shown in **Supplementary Table S1**. The docking protocol used herein

was first tested by docking a known co-crystallized compound at the N-terminal structure of PqsA (PDB ID = 5OE3) with 300 iterations. We found the same binding mode of the known inhibitors determined in crystallization studies. Docking reproducibility results by AutoDock Vina are shown in **Supplementary Figure S2**.

A brief binding analysis, including interaction, binding angles, binding poses, binding residues, and bond lengths, was visually conducted using DS, UCSF Chimera, and
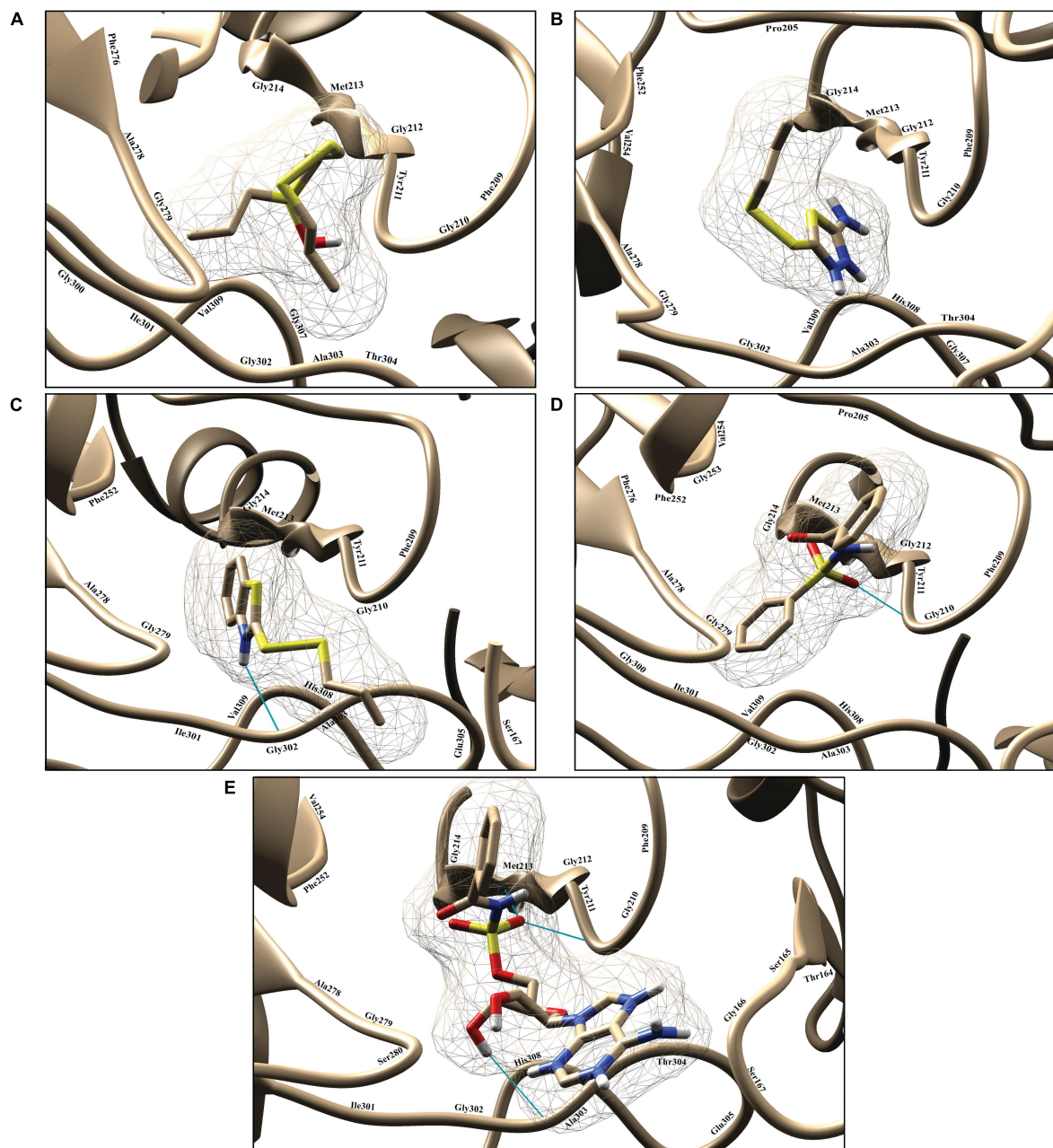
FIGURE 4 | Illustration of PqsA binding with five known inhibitors. The five known inhibitors are rendered in the active pocket: **(A–E)** correspond to compound 1–5. Inhibitors are displayed in stick representation, and backbone hydrogen bonds are shown as cyan-colored lines.

LigPlot (Wallace et al., 1995). Ligands with a strong binding affinity [6-amino-9-((2R,3R,4S,5R)-5-(((N-(2-aminobenzoyl)sulfamoyl)oxy)methyl)-3,4-dihydroxytetrahydrofuran-2-yl)-3H-purine-1,7,9-triium—named compound 1066—and 6-amino-9-((2R,3R,4S,5R)-3,4-dihydroxy-5-(((N-(3-hydroxy-2-naphthoyl)sulfamoyl)oxy)methyl)tetrahydrofuran-2-yl)-3H-purine-1,7,9-triium—named compound 1084] to the active pocket were selected as more potent QS inhibitors of PqsA from *P. aeruginosa*.

Bioavailability

The physicochemical properties of the molecule, including Lipinski's rule of five (molecular weight <500 Dalton, H-bond donor <5, H-bond acceptor <10, and cLog*P* < 5), toxicity prediction (tumorigenic, mutagenic, and irritant), and drug-likeness, were predicted using SwissADME (Daina et al., 2017) and PreADME (Lee et al., 2004).

Molecular Dynamic (MD) Simulations

The dynamic behavior of top docked complexes was studied using molecular dynamic (MD) simulations. The top two docked complexes were used in MD simulations carried out through the Sander Module of AMBER (Assisted model building with energy refinement) (Case et al., 2016). Primary coordinates of docked complexes were performed for MD simulations, including energy minimization (for the complete system, water and heavy atoms of the system), heating (at 300 K for the 20 picoseconds), equilibration (for 100 picoseconds with a time step of 2 nanoseconds), pressure (for 50 picoseconds), and production (for 100 ns). A general amber force field (GAFF) (Dickson et al., 2012) was used for putative inhibitor while ff03.rl (Case et al., 2014) used for the enzyme. Sodium ions were added randomly to neutralize the system. A simulation production run of 50 ns was accomplished to evaluate the

dynamics of the complex and check the docked conformation stability of the ligand. Langevin dynamics (Izaguirre et al., 2001) were used for temperature and pressure control, whereas the SHAKE algorithm (Kräutler et al., 2001) was applied for correct bond length. The production run was performed in constant volume and temperature (NVT) ensemble (Nosé, 1984) using the Berendsen algorithm (Lemak and Balabaev, 1994). MD simulation trajectories of each nanosecond were recorded and visualized and analyzed with Visual Molecular Dynamic (VMD) (Humphrey et al., 1996).

Binding Free Energies of Complexes

The binding free energy for top docked complexes was calculated using two methods: Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) and Molecular Mechanics-Generalized Born Surface Area (MMGBSA) incorporated with MMPBSA.py module of AMBER18 (Miller et al., 2012). In total, 50 frames from the trajectories were processed and the net energy of the system was calculated through the following equation.

$$\Delta G_{Binding} = \Delta G_{Complex} - \Delta G_{Receptor} - \Delta G_{Inhibitor}$$

Each of the terms in the equation involves the calculation of several energy components, including van der Waals energy, electrostatic energy, and internal energy summed from molecular mechanics and polar contribution toward solvation energy. The analysis also takes into account the contribution from non-polar terms toward solvation energy and inhibitor entropy.

RESULTS AND DISCUSSION

Overall Structure Assessment of PqsA

In rational drug discovery, the most fundamental step is to obtain a 3D structure of the target protein. The 3D structure

FIGURE 5 | 2D interaction maps of the top two predicted compounds. The compounds 1066 **(A)** and 1084 **(B)** are shown.
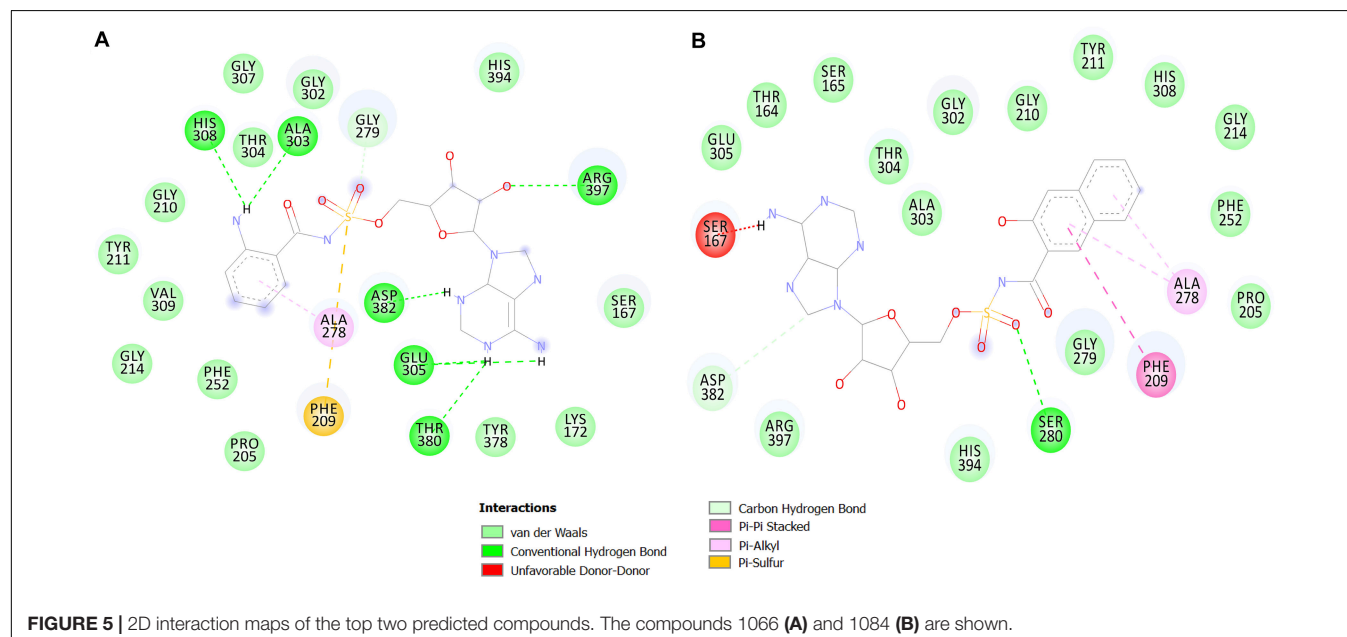
is used to understand the structural details and molecular function and to discover potent inhibitors of the target enzyme (Kopec et al., 2005). Because the full 3D structure of the PqsA from *P. aeruginosa* was unavailable, we performed comparative modeling using 5OE3 (PqsA N-terminal domain) from *P. aeruginosa* PAO1 and 1ULT from *Thermus thermophilus* using MODELLER9.24 (Eswar et al., 2007).

The predicted PqsA 3D structure is composed of a large N-terminal domain and a small C-terminal domain, and the domains are connected by a small flexible hinge (Gulick, 2009;

Supplementary Figure S3). The N-terminal domain of PqsA can be subdivided into three subdomains: two β-sheets connected by an internal 2-fold symmetry surrounded by α-helices and a distorted β-sheet followed by a flexible hinge region that links the N-terminal and C-terminal domains.

Ramachandran plot analysis was performed to assess the quality of the predicted structure. Most residues (77.5%) of the structure were grouped in the most favored region, whereas only 0.5% of the residues lie in the outlier region, which indicates a good structure quality (**Figure 2**). The secondary
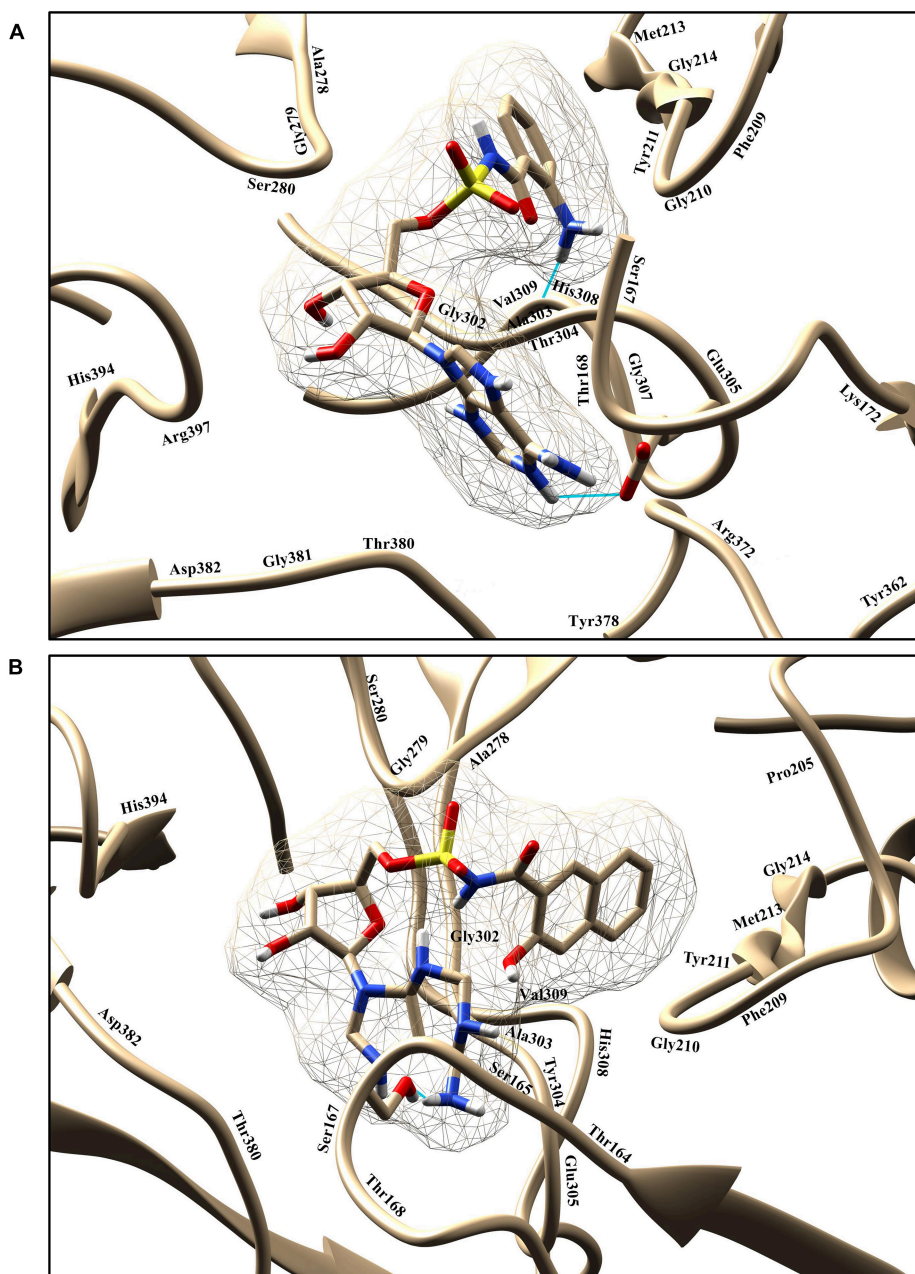
FIGURE 6 | Graphical representation of binding poses and residues that interact with the compounds 1066 **(A)** and 1084 **(B)**. The hydrogen bond is represented in cyan color.

structure contained the following: 17.6% strand, 20.3% alpha helix, 2.1% 3–10 helix, and 60% gamma turns, beta turns, helix-helix interactions, beta bulges, beta hairpins, and beta alpha beta motifs. The 2D structure of the protein was also predicted using MODELLER software, and it is illustrated in **Supplementary Figure S4**.

Reported QS Inhibitors of PqsA

In this study, five recently reported PqsA inhibitors were retrieved from the literature (Ji et al., 2016; Fong et al., 2017) and used as reference compounds to discover more potent inhibitors (**Figure 3**). The inhibiting potency of these compounds has been experimentally measured (Ji et al., 2016; Fong et al., 2017). The IC_{50} values of compounds 1–4 were 15 ± 2.64, 2.28 ± 0.11, 0.98 ± 0.14, and $36.2 \pm 2.39\ \mu M$, respectively. For compound 5, its constant K_i used to represent inhibitory potency was 16.5 ± 2.6 nM. Structure comparison of known compounds determined that compound 4 and 5 have a significant structure similarity, the 3D structures of the five compounds are shown in **Figure 3**.

A library of 521 analogous compounds to the reported PqsA inhibitors was generated to screen for potent inhibitory compounds of PqsA by utilizing similarity search tools from online compound databases (PubChem, ChemSpider, and ChEMBL) based on 85% structural similarity to the known selected inhibitors.

After a brief docking analysis, including binding affinity, binding poses, interacting residues, and enzyme–inhibitor interactions of all the compounds (five reported PqsA inhibitors and the analog library), compounds with a stronger binding affinity than that of controls were selected as potent PqsA inhibitory compounds.

Molecular Docking

Protein-ligand docking plays a pivotal role in predicting the accurate orientation of a ligand with its target protein (Morris and Lim-Wilby, 2008). For better understanding, molecular docking was divided into two phases: (1) enzyme docking with the known PqsA inhibitors and (2) enzyme docking with the analog library. Compounds with a strong binding

FIGURE 7 | Superimposition of the reported and predicted PqsA inhibitors. Inhibitors are displayed in stick representation, **(A)** superimposition of the top two predicted PqsA binders, compound 1066 is in brown while compound 1084 in cyan color and **(B)** Superimposition of the five reported PqsA inhibitors within the binding pocket. Yellow, blue, red, chartreuse and navy-blue colors represent compounds 1–5, respectively..
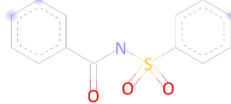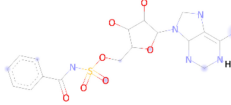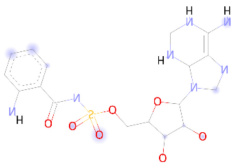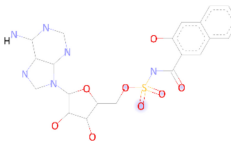
affinity and correct binding poses were selected. Binding affinity denotes the sum of total torsional energy, internal energy, and intermolecular energy subtra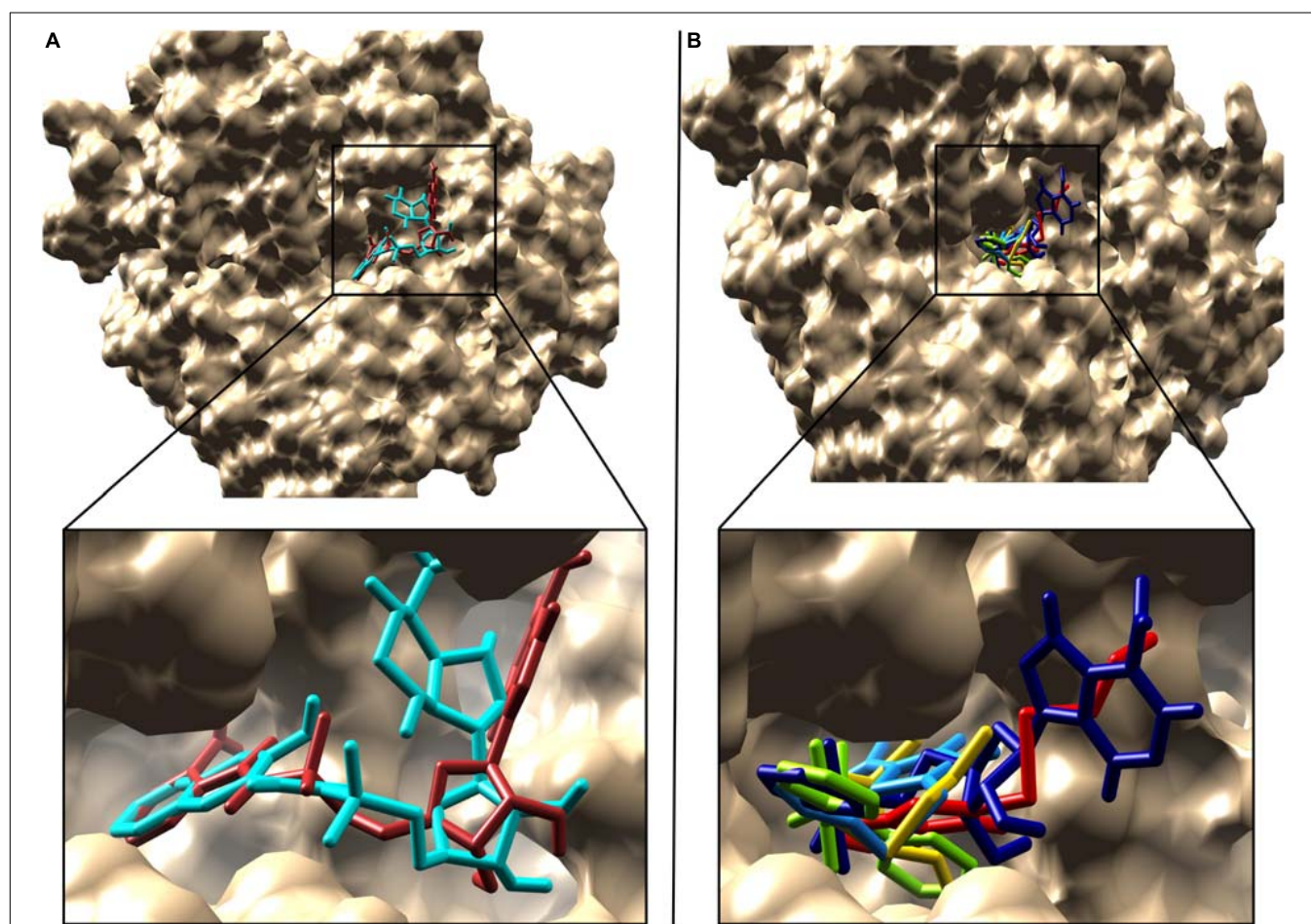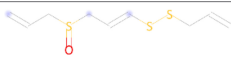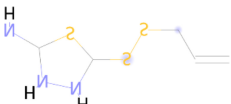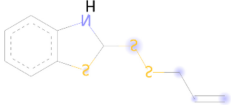cted from the unbound energy system. A strong binding affinity conformation indicates a stable protein-ligand complex. We also found that the compounds docked in similar conformations to both full-length PqsA structure as well as to the N-terminal structure only. As no significant contribution of the C-terminal on compounds binding at the N terminal was found, this study proceeded with the N-terminal structure only.

Enzyme Docking With Reported QS Inhibitors

To find the binding poses, interactions, and binding affinity of all the five known inhibitors, they were docked into the binding pocket of PqsA using the PyRx tool. Consequently, 300 poses for each inhibitor were generated, and inhibitors

with a strong binding affinity were selected for further analysis (**Figure 4**). The docking results showed that the binding affinity ranged from −4.4 to −8.5 kcal/mol. Docking analysis exhibited that the shared motif in compound 4 and 5 binds to the same residues (Val254, Phe209, Pro205, Phe252, Ala278, Tyr211, Gly279, Gly214, and Ile310), and the identical hydrogen bond of the shared motif in compound 4 and 5 was observed with the active site residues (Tyr211 and Gly214). There is a correlation between IC_{50} values of compounds and binding affinities of known inhibitors, except compound 4 that has a low IC_{50} value but high binding affinity. One possible reason for this could be that compound 4 has a large structure and more interactions, which produced a higher binding affinity, considers as a limitation of binding affinity-based inhibitor design. The interactions between each known inhibitor and PqsA are depicted in **Supplementary Figure S5**.

TABLE 1 | Binding affinity and interacting residues of the five reported compounds and two predicted analogs.

Known Inhibitors	Compounds	Interacting residues	Binding affinity (kcal/mol)
	Compound 1	Val254, Pro205, Phe209, Phe252, Met213, Gly210, Gly279, Gly214, Ala278, Thr304, Tyr211, Phe276, Gly300, Ala303, Gly302, Ile301, Val309, His308, Gly307	−4.4
	Compound 2	Lys206, Pro205, Phe209, Val254, Phe252, Met213, Gly210, Gly279, Gly212, Ala278, Gly214, Thr304, Tyr211, Asn215, Ala303, Gly302, His308, Val309, Gly307	−5.0
	Compound 3	Pro205, Met213, Gly214, Phe209, Phe252, Tyr211, Gly210, Ala278, His308, Thr304, Gly279, Val309, Ala303, Glu305, Gly302, Ser167, Ile301	−5.6
	Compound 4	Lys206, Pro205, Phe209, Val254, Gl253, Phe252, Met213, Gly212, Gly214, Gly210, Tyr211, Ala278, Gly279, Phe276, Gly300, Gly302, Ala303, Ile301, Val309, His308	−7.6
	Compound 5	Lys206, Pro205, Val254, Phe209, Phe252, Met213, Gly212, Gly214, Gly210, Tyr211, Ser165, Ala278, Ser280, Gly279, Gly166, Thr164, Thr304, Ser167, His308, Ala303, Gly302, Glu305, Ile301, Arg379	−8.5
Predicted inhibitors	Compound 1066	Pro205, Met213, Phe209, Gly214, Phe252, Gly210, Tyr211, Ala278, Thr304, His308, Gly279, Ser280, Val309, Ala303, Ser167, Glt307, Glu305, Lys172, Gly302, Thr168, Tyr362, His394, Arg397, Asp382, ly381, Gly381, Thr380, Tyr378, Arg379	−9.1
	Compound 1084	Pro205, Phe252, Met213, Gly214, Phe209, Ala278, Gly279, Tyr211, Gly210, Ser280, Val309, His308, Thr304, Ser165, Gly302, Ala303, Thr164, His394, Ser167, Glu305, Lys172, Thr168, Arg397, Asp382, Thr380	−9.3

FIGURE 8 | RMSD of PqsA protein in the presence of a compound. **(A)** Structure dynamics RMSD of PqsA protein in the presence of ligand **(B)** ligand dynamics RMSD in the PqsA protein pocket.

FIGURE 9 | Ligand movement from 1 to 50 ns. Ligand pose at 1 ns is represented by dim gray color while pose at 50 ns is in cyan. **(A)** Movement of compound 1084 and **(B)** movement of compound 1066.

Docking Analysis of Top Hits From Analog Library

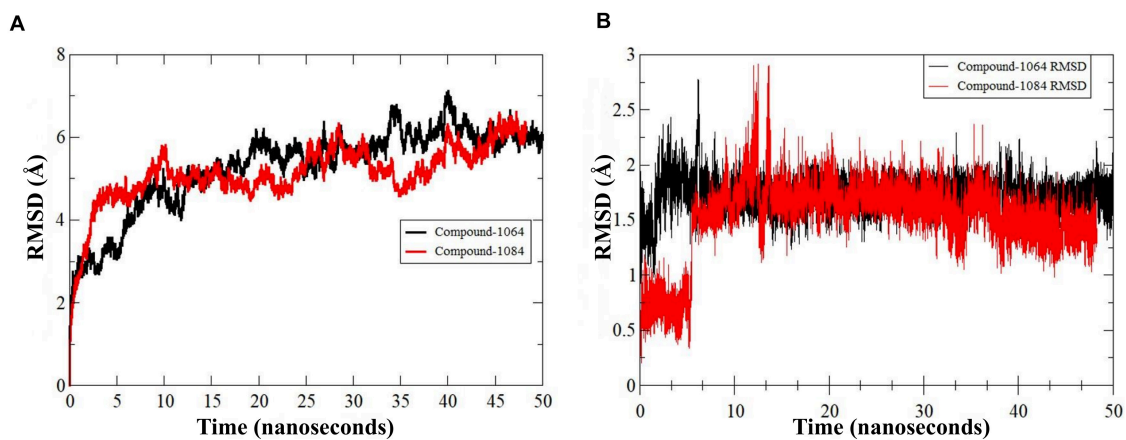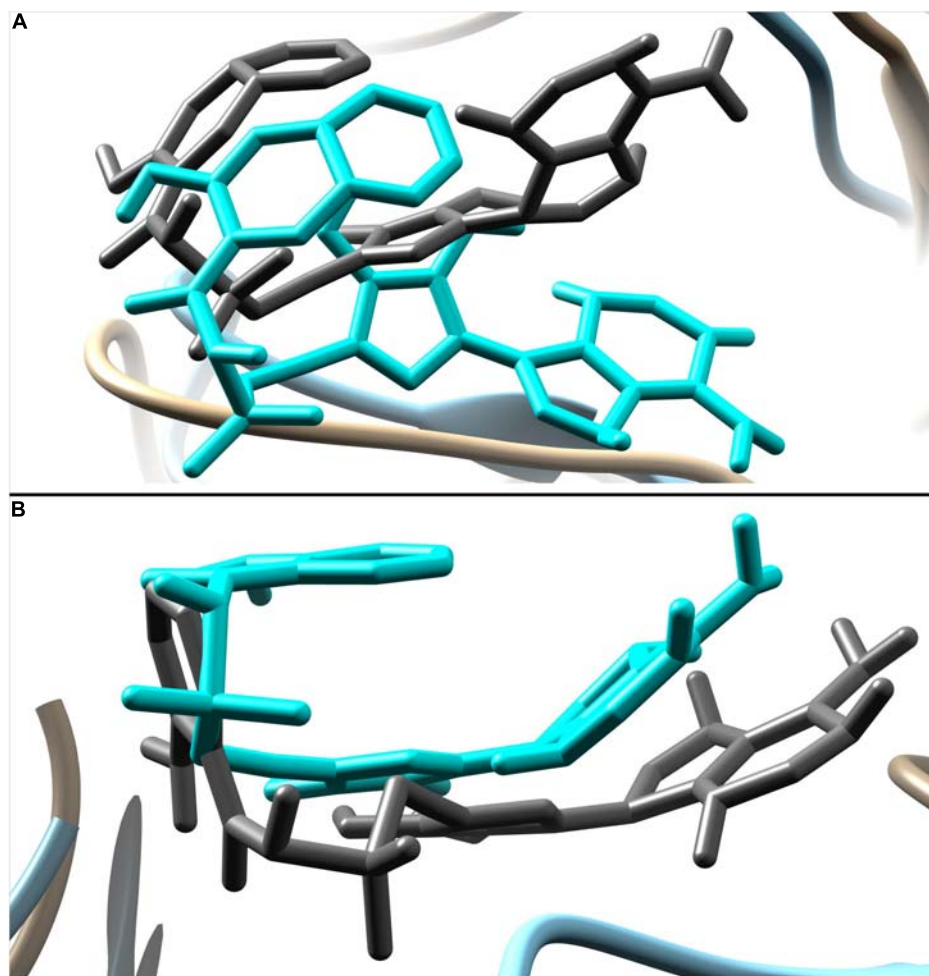The library containing 521 analog compounds was used for virtual screening to find more potent compounds to inhibit the QS system in *P. aeruginosa*. A brief binding analysis was performed for the top complexes with a high binding affinity to select the potent *pqs* QS-binder. As a result, the top two hit compounds were selected from the library, and these compounds were predicted to have greater potential than the previously reported compounds. Both top compounds (compound 1066 and 1084) have a significant structural similarity and also their shared motif docked in a similar position by making interactions with conserved active residues within the binding pocket. 2D interactions of the top two hit compounds are depicted in **Figure 5**.

The top two predicted compounds showed a stronger binding affinity with the conserved residues in the binding pocket (**Figure 6**). Both compounds contain one sulfur group in their structures. Generally, compounds with sulfur scaffolds, including antibiotics, antifungal agents, antitumor agents, and enzyme inhibitors, are related to a wide range of bioactivities (Feng et al., 2016).

Compound 1066 possessed seven conventional hydrogen bonds with the key residues, His308, Ala303, Arg397, Asp382, Thr380, and two bonds with Glu305; one carbon-hydrogen bonds with Gly279; one pi-alkyl interaction with Ala278; and

one pi-sulfur interaction with Phe209. The residues Tyr211, Gly210, Gly307, Thr304, Gly302, His394, Ser167, Lys172, Tyr378, Pro205, Phe252, Gly214, and Val309 took part in van der Waals interactions. The binding affinity of the compound 1066 with PqsA was −9.1 kcal/mol.

Compound 1084 also showed a very strong binding affinity (−9.3 kcal/mol) with PqsA by forming one conventional hydrogen bond with Ser280; one carbon-hydrogen bond with Asp382; one pi–pi stacked interaction with Phe209; and two pi-alkyl bonds with Ala278. The residues Glu305, Thr167, Ser165, Thr304, Ala303, Gly302, Gly210, Tyr211, His308, Gly214, Phe252, Pro205, Gly279, His394, and Arg397 were involved in van der Waals interactions with the compound, while one residue (Ser167) is involved in unfavorable donor-donor interaction.

The two hit compounds are analogs of compound 5 and share a significant portion of the structure. Compound 5 has been experimentally proved as a potent QS inhibitor exhibiting a Ki value of 16.5 ± 2.6 nM and a predicted binding affinity of −8.5 kcal/mol. Compared with the compound 5, the two hit compounds showed higher binding affinities of −9.3 kcal/mol and −9.1 kcal/mol, which represents that the two hits could be a high-affinity binder of PqsA enzyme. The hit compounds share a structural motif. The motifs of the hit compounds interact with the conserved active site residues in the same docking pose. RMSDs of the docked complexes with the two hits were 0.000Å,

TABLE 2 | Binding Free Energies of the complexes.

		MMGBSA				MMPBSA		
	Energy Component	**Average**	**Std. Dev.**	**Std. Err. of Mean**	**Energy Component**	**Average**	**Std. Dev.**	**Std. Err. of Mean**
Compound 1064	VDWAALS	−35.1814	3.2543	0.3254	VDWAALS	−35.1814	3.2543	0.3254
	EEL	−196.9409	10.4290	1.0429	EEL	−196.9409	10.4290	1.0429
	EGB	218.0214	8.8511	0.8851	EPB	203.9634	8.6698	0.8670
	ESURF	−4.3052	0.1455	0.0146	ENPOLAR	−3.2338	0.0879	0.0088
	—	—	—	—	EDISPER	0.0000	0.0000	0.0000
	∆G gas	−232.1223	9.5819	0.9582	∆G gas	−232.1223	9.5819	0.9582
	∆G solv	213.7161	8.8466	0.8847	∆G solv	200.7296	8.6597	0.8660
	∆TOTAL	−18.4062	3.2335	0.3233	∆ TOTAL	−31.3927	3.3925	0.3392
Compound 1084	VDWAALS	−57.6556	2.7124	0.2712	VDWAALS	−57.6556	2.7124	0.2712
	EEL	17.7792	8.4691	0.8469	EEL	17.7792	8.4691	0.8469
	EGB	17.4195	7.8428	0.7843	EPB	9.6471	10.1448	1.0145
	ESURF	−4.9013	0.1551	0.0155	ENPOLAR	−4.1481	0.0759	0.0076
	—	—	—	—	EDISPER	0.0000	0.0000	0.0000
	∆G gas	−39.8764	9.5508	0.9551	∆G gas	−39.8764	9.5508	0.9551
	∆G solv	12.5181	7.8326	0.7833	∆G solv	5.4990	10.1226	1.0123
	∆TOTAL	−27.3583	3.5233	0.3523	∆ TOTAL	−34.3774	5.1235	0.5124

TABLE 3 | Physicochemical properties of hit compounds.

Compound	MW (g/mol)	Lipinski's rule of five	PAINS	Lipophilicity	TPSA (Å²)	Synthetic accessibility	Water solubility	Bioavailability score	BBB Permeation	Skin permeability (cm/s)
Compound 1066	462.48	0 violations	0 alerts	0.07	187.51	4.69	Soluble	0.55	No	−8.91
Compound 1084	513	1 violation	0 alerts	1.33	181.72	4.84	Moderately soluble	0.17	No	−8.10

which indicates that there is no significant variability in enzyme structure and ligand binding pose.

In the docking analysis with known inhibitors, the residues (Pro205, Phe252, Met213, Gly214, Phe209, Ala278, Gly279, Tyr211, Gly210, His308, Gly302, and Ala303) were found to be involved in the interactions. We also found that Tyr211, Pro205, Phe252, Gly214, Phe209, Ala278, Gly279, Gly210, His308, Gly302, and Ala303 residues are also involved in the interactions with the compounds 1066 and 1084. Interestingly, Ala278 that is involved in the interactions of all inhibitors and is also involved in the interactions of compounds 1066 and 1084. This represents that the two hit compounds interact with the binding pocket in a similar manner with known inhibitors. This result provides important information in designing PqsA inhibitors. A brief graphical docked complex representation of both reported and predicted compounds has been depicted in **Figure 7** and their binding affinities are shown in **Table 1**.

Molecular Dynamics Simulation

The predicted two complexes comprised of a chemical compound and PqsA enzyme were used in molecular dynamics simulation to unveil the dynamics of the enzyme in the presence of a ligand as well as to confirm the conformation stability of ligands predicted by the docking simulation. For the two compounds, the initial conformations docked on the PqsA were stable and no major changes were observed in terms of ligand RMSD (**Figure 8**). The mean RMSD of compound 1064 was 1.7 Å, and the mean RMSD of compound 1064 was 1.51 Å. This result confirmed that the predicted binding modes by docking are stable and consistent. Ligand movements during the simulation from 1 to 50 ns are depicted in **Figure 9**.

MM-PBSA/GBSA Binding Free Energy Analysis

MM-PBSA method has been recently used to estimate the binding free energy of a given complex in rational drug discovery. As shown in **Table 2**, the gas phase energies (ΔG_{gas}) in both complexes are very high mainly due to the significant contribution of electrostatic energy (compound 1064) and the decent role of van der Waals energy (compound 1084). In contrast, the net solvation energy (ΔG_{solv}) is less favorable in both systems. Nevertheless, the total binding free energies of both complexes are promising: PqsA-compound 1064 has −18.4 kcal/mol in MM-GBSA and −31.4 kcal/mol in MM-PBSA, and PqsA-compound 1084 has −27.4 kcal/mol in MM-GBSA and −34.4 kcal/mol in MM-PBSA. These values reflect the stable complex conformation and high intermolecular affinity.

Bioavailability

The bioavailability of the two selected compounds is important in drug discovery because it determines the applicability of the compounds as a drug. Pharmacophoric properties, including Lipinski's rule of five and toxicological properties, are noteworthy to determine drug accessibility. Ideal drug candidates exhibit the following pharmacophoric properties. Drugs should satisfy Lipinski's rule of five (molecular weight < 500 g/mol, H-bond donor < 5, H-bond acceptor < 10, and cLog$P$ < 5). The topological polar surface area (TPSA) should range from 20 to 130 Å. Lipophilicity should range from −0.7 to +6.0, and more negative values indicate lower skin permeability. Online servers (SwissADME and PreADME) were used to evaluate the physicochemical properties of the top two hit compounds (**Table 3**). Important physicochemical properties of the hit compounds were predicted, which included Lipinski's rule of five, PAINS assay, lipophilicity, TPSA, and, more importantly, bioavailability. All these predicted physicochemical properties of the two compounds were suitable enough to attempt experimental evaluations. Therefore, these compounds could be new candidate agents for the efficacious management of bacterial infections.

CONCLUSION

Inhibition of the *P. aeruginosa* quinolone signaling system is an attractive and promising approach to impede infections by preventing biofilm formation. In this study, we present the applications of rational *in silico* drug discovery techniques to identify novel and more putative inhibitors for PqsA, an important enzyme in *P. aeruginosa* quinolone signaling. Based on virtual screening, we identified two compounds (compounds 1066 and 1084) as potent compounds showing a good affinity for the PqsA enzyme. Both these compounds have vital chemical moieties responsible for important chemical interactions with hotspot residues of the PqsA enzyme. The length of the compound also seems important because it provides a balanced network of chemical interactions at the docking site. The affinity and stability of the compounds binding mode were examined through molecular dynamic simulation and MMPB/GBSA assay both are in strong agreement of strong intermolecular affinity and formation of stable complexes. Furthermore, a suitable profile of drug-like properties and pharmacokinetics was revealed for both compounds, thereby increasing their chances of being good leads. Based on the findings, we believe that the compounds should be subjected to *in vitro* and *in vivo* investigations to affirm their potency and could be used in further structural optimization of new potent derivatives.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

BS, SA, and DN: conceptualization. BS: data curation. BS and S-iE: methodology. DN: supervision. BS and DN: manuscript writing. All authors have read and agreed to the published version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.577316/full#supplementary-material

REFERENCES

Ahmad, S., Raza, S., Uddin, R., and Azam, S. S. (2017). Binding mode analysis, dynamic simulation and binding free energy calculations of the MurF ligase from *Acinetobacter baumannii*. *J. Mol. Graph. Model.* 77, 72–85. doi: 10.1016/j.jmgm.2017.07.024

Alhede, M., Bjarnsholt, T., Jensen, P. Ø, Phipps, R. K., Moser, C., Christophersen, L., et al. (2009). *Pseudomonas aeruginosa* recognizes and responds aggressively to the presence of polymorphonuclear leukocytes. *Microbiology* 155, 3500–3508. doi: 10.1099/mic.0.031443-0

Ayers, M. (2012). ChemSpider: the free chemical database. *Reference Rev.* 26, 45–46. doi: 10.1108/09504121211271059

Bonomo, R. A., and Szabo, D. (2006). Mechanisms of multidrug resistance in Acinetobacter species and *Pseudomonas aeruginosa*. *Clin. Infec. Dis.* 43, S49–S56.

Calfee, M. W., Coleman, J. P., and Pesci, E. C. (2001). Interference with *Pseudomonas* quinolone signal synthesis inhibits virulence factor expression by *Pseudomonas aeruginosa*. *Proc. Nat. Acad. Sci. U.S.A.* 98, 11633–11637. doi: 10.1073/pnas.201328498

Cao, H., Krishnan, G., Goumnerov, B., Tsongalis, J., Tompkins, R., and Rahme, L. G. (2001). A quorum sensing-associated virulence gene of *Pseudomonas aeruginosa* encodes a LysR-like transcription regulator with a unique self-regulatory mechanism. *Proc. Nat. Acad. Sci. U.S.A.* 98, 14613–14618. doi: 10.1073/pnas.251465298

Case, D., Babin, V., Berryman, J., Betz, R., Cai, Q., Cerutti, D., et al. (2014). The FF14SB force field. *Amber* 14, 29–31.

Case, D., Cerutti, D., Cheatham, T., Darden, T., Duke, R., Giese, T., et al. (2016). *Others, AMBER16 Package*. San Francisco, CL: University of California.

Chiang, W.-C., Nilsson, M., Jensen, P. Ø, Høiby, N., Nielsen, T. E., Givskov, M., et al. (2013). Extracellular DNA shields against aminoglycosides in *Pseudomonas aeruginosa* biofilms. *Antimicrob. Agents Chemother.* 57, 2352–2361. doi: 10.1128/aac.00001-13

Daina, A., Michielin, O., and Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness, and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7, 42717.

Dallakyan, S., and Olson, A. J. (2015). "Small-Molecule Library Screening by Docking with PyRx," in *Chemical Biology*. New York, NY: Humana Press, 243–250.

Davies, M., Nowotka, M., Papadatos, G., Dedman, N., Gaulton, A., Atkinson, F., et al. (2015). ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucl. Acids Res.* 43, W612–W620.

De Beer, T. A., Berka, K., Thornton, J. M., and Laskowski, R. A. (2014). PDBSum additions. *Nucleic Acids Res.* 42, D292–D296.

Dickson, C. J., Rosso, L., Betz, R. M., Walker, R. C., and Gould, I. R. (2012). GAFFlipid: a general amber force field for the accurate molecular dynamics simulation of phospholipid. *Soft Matter* 8, 9617–9627. doi: 10.1039/c2sm26007g

Diggle, S. P., Lumjiaktase, P., Dipilato, F., Winzer, K., Kunakorn, M., Barrett, D. A., et al. (2006). Functional genetic analysis reveals a 2-alkyl-4-quinolone signaling system in the human pathogen Burkholderia pseudomallei and related bacteria. *Chem. Biol.* 13, 701–710. doi: 10.1016/j.chembiol.2006.05.006

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M., Eramian, D., Shen, M. Y., et al. (2007). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Prot. Sci.* 50, 2.9.1–2.9.31.

Feng, M., Tang, B., Liang, S. H., and Jiang, X. (2016). Sulfur containing scaffolds in drugs: synthesis and application in medicinal chemistry. *Curr. Top. Med. Chem.* 16, 1200–1216. doi: 10.2174/1568026616666150911741

Flemming, H.-C., and Wingender, J. (2001). Relevance of microbial extracellular polymeric substances (EPSs)-Part I: structural and ecological aspects. *Water Sci. Technol.* 43, 1–8. doi: 10.2166/wst.2001.0326

Flydal, M. I., Alcorlo-Pagés, M., Johannessen, F. G., Martínez-Caballero, S., Skjærven, L., Fernandez-Leiro, R., et al. (2019). Structure of full-length human phenylalanine hydroxylase in complex with tetrahydrobiopterin. *PNAS* 116, 11229–11234. doi: 10.1073/pnas.1902639116

Fong, J., Yuan, M., Jakobsen, T. H., Mortensen, K. T., Delos Santos, M. M. S., Chua, S. L., et al. (2017). Disulfide bond-containing ajoene analogues as novel quorum sensing inhibitors of *Pseudomonas aeruginosa*. *J. Med. Chem.* 60, 215–227. doi: 10.1021/acs.jmedchem.6b01025

Gaddy, J. A., and Actis, L. A. (2009). Regulation of *Acinetobacter baumannii* biofilm formation. *Future Microbiol.* 4, 273–278.

Gopalakrishnan, K., Sowmiya, G., Sheik, S., and Sekar, K. (2007). Ramachandran plot on the web (2.0). *Prot. Peptide Lett.* 14, 669–671. doi: 10.2174/092986607781483912

Grandclément, C., Tannières, M., Moréra, S., Dessaux, Y., and Faure, D. (2016). Quorum quenching: role in nature and applied developments. *FEMS Microbiol. Rev.* 40, 86–116. doi: 10.1093/femsre/fuv038

Gulick, A. M. (2009). Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* 4, 811–827. doi: 10.1021/cb900156h

Hentzer, M., Wu, H., Andersen, J. B., Riedel, K., Rasmussen, T. B., Bagge, N., et al. (2003). Attenuation of *Pseudomonas aeruginosa* virulence by quorum sensing inhibitors. *EMBO J.* 22, 3803–3815. doi: 10.1093/emboj/cdg366

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. doi: 10.1016/0263-7855(96)00018-5

Itoh, Y., Sekine, S.-I., and Yokoyama, S. (2015). Crystal structure of the full-length bacterial selenocysteine-specific elongation factor SelB. *Nucl. Acids Res.* 43, 9028–9038. doi: 10.1093/nar/gkv833

Izaguirre, J. A., Catarello, D. P., Wozniak, J. M., and Skeel, R. D. (2001). Langevin stabilization of molecular dynamics. *J. Chem. Phys.* 114, 2090–2098. doi: 10.1063/1.1332990

Ji, C., Sharma, I., Pratihar, D., Hudson, L. L., Maura, D., Guney, T., et al. (2016). Designed small-molecule inhibitors of the anthranilyl-CoA synthetase PqsA block quinolone biosynthesis in *Pseudomonas aeruginosa*. *ACS Chem. Biol.* 11, 3061–3067. doi: 10.1021/acschembio.6b00575

Jimenez, P. N., Koch, G., Thompson, J. A., Xavier, K. B., Cool, R. H., and Quax, W. J. (2012). The multiple signaling systems regulating virulence in *Pseudomonas aeruginosa*. *Microbiol. Mol. Biol. Rev.* 76, 46–65. doi: 10.1128/mmbr.05007-11

Kamal, A. A., Maurer, C. K., Allegretta, G., Haupenthal, J., Empting, M., and Hartmann, R. W. (2017). Quorum sensing inhibitors as path blockers for *Pseudomonas aeruginosa* infections: a new concept in anti-infective drug discovery. *Top Med. Chem.* 26, 185–210. doi: 10.1007/7355_2017_17

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2019). PubChem 2019 update: improved access to chemical data. *Nucl. Acids Res.* 47, D1102–D1109.

Kopec, K. K., Bozyczko-Coyne, D., and Williams, M. (2005). Target identification and validation in drug discovery: the role of proteomics. *Biochem. Pharmacol.* 69, 1133–1139. doi: 10.1016/j.bcp.2005.01.004

Kräutler, V., Van Gunsteren, W. F., and Hünenberger, P. H. (2001). A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* 22, 501–508. doi: 10.1002/1096-987x(20010415)22:5<501::aid-jcc1021>3.0.co;2-v

Lee, S., Chang, G., Lee, I., Chung, J., Sung, K., and No, K. (2004). The PreADME: Pc-based program for the batch prediction of ADME properties. *EuroQSAR* 9, 5–10.

Lemak, A., and Balabaev, N. (1994). On the Berendsen thermostat. *Mol. Simulat.* 13, 177–187. doi: 10.1080/08927029408021981

Lewis, K. (2007). Persister cells, dormancy, and infectious disease. *Nat. Rev. Microbiol.* 5, 48–56. doi: 10.1038/nrmicro1557

Mendelsohn, L. D. (2004). ChemDraw 8 ultra, windows and macintosh versions. *J. Chem. Inf. Comput. Sci.* 44, 2225–2226. doi: 10.1021/ci040123t

McKnight, S. L., Iglewski, B. H., and Pesci, E. C. (2000). The *Pseudomonas* quinolone signal regulates rhl quorum sensing in *Pseudomonas aeruginosa*. *J. Bacteriol.* 182, 2702–2708. doi: 10.1128/jb.182.10.2702-2708.2000

Miller, B. R., McGee, D. T., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012). MMPBSA. py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput* 8, 3314–3321. doi: 10.1021/ct300418h

Morris, G. M., and Lim-Wilby, M. (2008). *Molecular Docking in Molecular Modeling of Proteins.* Berlin: Springer, 365–382.

Mühlen, S., and Dersch, P. (2016). Anti-virulence strategies to target bacterial infections. in *Curr. Top Microbiol. Immunol.* 398, 147–183. doi: 10.1007/82_2015_490

Ng, W.-L., and Bassler, B. L. (2009). Bacterial quorum-sensing network architectures. *Annu. Rev. Genet.* 43, 197–222. doi: 10.1146/annurev-genet-102108-134304

Nosé, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52, 255–268. doi: 10.1080/00268978400101201

Page, M. G., and Heim, J. (2009). Prospects for the next anti-*Pseudomonas* drug. *Curr. Opin. Pharmacol.* 9, 558–565. doi: 10.1016/j.coph.2009.08.006

Pesci, E. C., Milbank, J. B., Pearson, J. P., Mcknight, S., Kende, A. S., Greenberg, E. P., et al. (1999). Quinolone signaling in the cell-to-cell communication system of *Pseudomonas aeruginosa*. *Proc. Nat. Acad. Sci. U.S.A.* 96, 11229–11234. doi: 10.1073/pnas.96.20.11229

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Sandmann, A., Dickschat, J., Jenke-Kodama, H., Kunze, B., Dittmann, E., and Müller, R. (2007). A type II polyketide synthase from the gram-negative bacterium *Stigmatella aurantiaca* is involved in aurachin alkaloid biosynthesis. *Angew. Chem.* 46, 2712–2716. doi: 10.1002/anie.200603513

Schühle, K., Jahn, M., Ghisla, S., and Fuchs, G. (2001). Two similar gene clusters coding for enzymes of a new type of aerobic 2-aminobenzoate (anthranilate) metabolism in the bacterium *Azoarcus evansii. J. Bacteriol.* 183, 5268–5278. doi: 10.1128/jb.183.18.5268-5278.2001

Shrivastava, S. R., Shrivastava, P. S., and Ramasamy, J. (2018). World health organization releases global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. *J. Med. Soc.* 32:76. doi: 10.4103/jms.jms_25_17

Strateva, T., and Yordanov, D. (2009). *Pseudomonas aeruginosa*–a phenomenon of bacterial resistance. *J. Med. Microbiol.* 58, 1133–1148. doi: 10.1099/jmm.0.009142-0

Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461.

Van Gennip, M., Christensen, L. D., Alhede, M., Phipps, R., Jensen, P. Ø, Christophersen, L., et al. (2009). Inactivation of the rhlA gene in *Pseudomonas aeruginosa* prevents rhamnolipid production, disabling the protection against polymorphonuclear leukocytes. *APMIS* 117, 537–546. doi: 10.1111/j.1600-0463.2009.02466.x

van Kessel, J. C. (2019). PQS signaling for more than a quorum: the collective stress response protects healthy *Pseudomonas aeruginosa* populations. *J. Bacteriol* 201:e00568-19. doi: 10.1128/JB.00568-19

Vial, L., Lépine, F., Milot, S., Groleau, M.-C., Dekimpe, V., Woods, D. E., et al. (2008). Burkholderia pseudomallei, B. *thailandensis, and B. ambifaria produce* 4-hydroxy-2-alkylquinoline analogues with a methyl group at the 3 position that is required for quorum-sensing regulation. *J. Bacteriol.* 190, 5339–5352. doi: 10.1128/jb.00400-08

Wagner, S., Sommer, R., Hinsberger, S., Lu, C., Hartmann, R. W., Empting, M., et al. (2016). Novel strategies for the treatment of *Pseudomonas aeruginosa* infections. *J. Med. Chem.* 59, 5929–5969.

Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng. Des. Sel.* 8, 127–134. doi: 10.1093/protein/8.2.127

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modeling of protein structures and complexes. *Nucl. Acids Res.* 46, W296–W303.

Waters, C. M., and Bassler, B. L. (2005). Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.* 21, 319–346. doi: 10.1146/annurev.cellbio.21.012704.131001

Williams, P., and Cámara, M. (2009). Quorum sensing and environmental adaptation in *Pseudomonas aeruginosa*: a tale of regulatory networks and multifunctional signal molecules. *Curr. Opin. Microbiol.* 12, 182–191. doi: 10.1016/j.mib.2009.01.005

Witzgall, F., Ewert, W., and Blankenfeldt, W. (2017). Structures of the N-terminal domain of PqsA in complex with anthraniloyl-and 6-fluoroanthraniloyl-AMP: substrate activation in *Pseudomonas* quinolone signal (PQS) biosynthesis. *ChemBioChem.* 18, 2045–2055. doi: 10.1002/cbic.201700374

Xiang, Z. (2006). Advances in homology protein structure modeling. *Curr. Protein Pept. Sci.* 7, 217–227. doi: 10.2174/138920306777452312

Xiao, G., Déziel, E., He, J., Lépine, F., Lesic, B., Castonguay, M. H., et al. (2006). MvfR, a key *Pseudomonas aeruginosa* pathogenicity LTTR-class regulatory protein, has dual ligands. *Mol. Microbiol.* 62, 1689–1699. doi: 10.1111/j.1365-2958.2006.05462.x

Zavascki, A. P., Carvalhaes, C. G., Picao, R. C., and Gales, A. C. (2010). Multidrug-resistant *Pseudomonas aeruginosa* and *Acinetobacter baumannii*: resistance mechanisms and implications for therapy. *Exp. Rev. Anti Infect. Ther.* 8, 71–93.

Interactions of FK506 and Rapamycin With FK506 Binding Protein 12 in Opportunistic Human Fungal Pathogens

Sandeep Vellanki, Alexis E. Garcia and Soo Chan Lee*

South Texas Center for Emerging Infectious Diseases, Department of Biology, The University of Texas at San Antonio, San Antonio, TX, United States

Over the past few decades advances in modern medicine have resulted in a global increase in the prevalence of fungal infections. Particularly people undergoing organ transplants or cancer treatments with a compromised immune system are at an elevated risk for lethal fungal infections such as invasive candidiasis, aspergillosis, cryptococcosis, etc. The emergence of drug resistance in fungal pathogens poses a serious threat to mankind and it is critical to identify new targets for the development of antifungals. Calcineurin and TOR proteins are conserved across eukaryotes including pathogenic fungi. Two small molecules FK506 and rapamycin bind to FKBP12 immunophilin and the resulting complexes (FK506-FKBP12 and rapamycin-FKBP12) target calcineurin and TOR, respectively in both humans and fungi. However, due to their immunosuppressive nature these drugs in the current form cannot be used as an antifungal. To overcome this, it is important to identify key differences between human and fungal FKBP12, calcineurin, and TOR proteins which will facilitate the development of new small molecules with higher affinity toward fungal components. The current review highlights FK506/rapamycin-FKBP12 interactions with calcineurin/TOR kinase in human and fungi, and development of non-immunosuppressive analogs of FK506, rapamycin, and novel small molecules in inhibition of fungal calcineurin and TOR kinase.

Keywords: FK506, FKBP12, rapamycin, calcineurin, TOR, fungi, antifungal

INTRODUCTION

Opportunistic invasive fungal pathogens cause significant morbidity and mortality in immunocompromised patients such as those undergoing stem cell or organ transplantation and in individuals with a dysfunctional immune system due to diabetes mellitus or AIDS (Husain et al., 2003; Enoch et al., 2006; Person et al., 2010; Low and Rotstein, 2011; Wang et al., 2017). While the most common opportunistic agents of fungal infections are *Aspergillus fumigatus*, *Cryptococcus neoformans,* and *Candida albicans*, there has been a rising number of cases caused by other fungi such as those belonging to Mucorales (Pfaller and Diekema, 2004). Unlike human cells, these fungi have a cell wall to maintain homeostasis which makes it an attractive target for

antifungals such as echinocandins (Cortés et al., 2019; Lima et al., 2019). Other antifungals such as azoles or polyenes target ergosterol in the cell membrane (Maertens, 2004). There is increasing evidence that fungi are either intrinsically resistant to antifungals or can develop resistance after prolonged exposure to antifungals (Wiederhold, 2017). There has also been less than optimistic progress in the development of antifungal drugs in the past decade. Therefore, there is an urgent need for the development of new antifungals to combat deadly fungal infections.

Small molecules are compounds targeting a specific biological component or a function. These small molecules which are either new compounds or approved drugs have been investigated for antifungal activity (Cully, 2018; Garcia et al., 2018; Bechman et al., 2019; Zarakas et al., 2019). One such small molecule is FK506 (tacrolimus; **Figures 1**, **2**), a 23-membered macrolide lactone first identified from the fermentation broth of the bacteria *Streptomyces tsukubaenis* by Fujisawa Pharmaceuticals (Kino et al., 1987). FK506 is an immunosuppressive drug used for the prevention of graft rejection in patients undergoing organ transplantation (Thomson, 1989; Meier-Kriesche et al., 2006; Ban et al., 2016). The immunosuppression by FK506 results from inhibition of T-cell activation and proliferation (Schreiber and Crabtree, 1992). On entering the cell, FK506 binds to a prolyl isomerase FKBP12 in the cytosol and the FK506-FKBP12 drug-immunophilin complex binds and inhibits the activity of calcineurin, a calcium-calmodulin dependent protein phosphatase. Calcineurin governs the activation and translocation of the nuclear factor of activated T cells (NFAT) from the cytoplasm to the nucleus to activate genes such as *Il-2* that are involved in T-cell activation and proliferation (Liu et al., 1991; Schreiber and Crabtree, 1992). When calcineurin activity is inhibited by FK506, the NFAT transcription factor cannot translocate into the nucleus (Schreiber and Crabtree, 1992). Interestingly, calcineurin is not only conserved across other eukaryotes like fungi but is also a key virulent factor in many pathogenic fungi, suggesting that FK506 can be effective against human pathogenic fungi (Steinbach et al., 2007). It was first reported in *C. neoformans* that the mutants lacking

Compound	Modification
L685, 818	C18 hydroxy, C21 ethyl
9D-FK506	C9 deoxo
9DP-FK506	C9 deoxoprolyl
31OD-FK506	C31 O-demethyl
9D31OD-FK506	C9 deoxo, C31 demethyl
Compound 7	C9 deoxo, C21 ethyl
APX879	C22 acetohydrazine

FIGURE 2 | The chemical structures of FK506 and its analogs discussed in this study. The numbers in red font indicate the modified positions to generate analogs.

calcineurin do not survive in conditions that mimic the host environment (Steinbach et al., 2007). Further studies have shown that calcineurin is also required for virulence and pathogenicity in *C. albicans* (Blankenship et al., 2003; Blankenship and Heitman, 2005), *A. fumigatus* (Steinbach et al., 2006), and *Mucor circinelloides* (Lee et al., 2013; Lee et al., 2015; Vellanki et al., 2020). Although, calcineurin is an attractive target for antifungal drug development, there are two key problems. First, the human and fungal calcineurin share 80% homology, and also fungal and human FKBP12s share about 48–58% sequence homology (Gobeil et al., 2019). The cross-reactivity of FK506 against human and fungal calcineurins represents a major hurdle. To overcome this, it is important to identify regions of calcineurin and FKBP12 that are present in fungi but not in humans. Secondly, FK506 exhibits antifungal activity by inhibiting fungal calcineurin (Odom et al., 1997a; Steinbach et al., 2004), but FK506 is an immunosuppressive drug, therefore it cannot be used in its current form to treat patients with fungal infections.

FIGURE 1 | Three dimensional structures of small molecules—FK506 **(left)** and rapamycin **(right)**.

The non-immunosuppressive analogs of FK506 offer a potential solution to overcome this problem. The first part of the review focuses on—(a) structural and molecular interactions between FK506-FKBP12-calcineurin in pathogenic fungi compared to human counterparts which in turn will inform the extent to which fungal FKBP12 and calcineurin are suitable antifungal candidates, (b) non-immunosuppressive analogs of FK506 as potential antifungal drugs.

Rapamycin (**Figures 1**, **5**) is a 29 membered macrolide lactam metabolite produced by the bacteria *Streptomyces hygroscopicus* isolated by Sehgal et al. (1975). Rapamycin was found to inhibit the growth of *S. cerevisiae* and, interestingly, FK506 antagonized the effect of rapamycin, suggesting that FK506 and rapamycin share the same cellular receptor, FKBP12 (Heitman et al., 1991). Indeed, Heitman et al. (1991, 1993) verified that rapamycin forms a complex with FKBP12 and inhibits TOR (target of rapamycin) kinase. Rapamycin was further investigated as a possible antifungal agent with potent activity against *C. albicans* however, it was later discovered that rapamycin possessed potent immunosuppressive activity (Martel et al., 1977; Hwang et al., 2017). This is attributed to rapamycin inhibiting the activation of T-cells and B-cells via the reduction of their sensitivity to interleukin-2 through inhibition of the mammalian Target of Rapamycin (mTOR) (Saunders et al., 2001). Thus, rapamycin has been previously used to prevent organ transplant rejection via immunosuppression (Martel et al., 1977; Saunders et al., 2001; Yoon et al., 2017). The mode of action by which rapamycin inhibits the immune system begins with binding to the prolyl isomerase FKBP12 (Saunders et al., 2001; Yoon et al., 2017). The rapamycin-FKBP12 complex will then bind with mTOR and inhibit its function (Saunders et al., 2001; Yoon et al., 2017). The result is the inhibition of T-cell and B-cell activation leading to immunosuppression. With regards to fungi, the TOR pathway is known to be involved in the response to nutrient resource availability (Cardenas et al., 1999; Bastidas et al., 2012; Nguyen et al., 2020). The serine/threonine kinase TOR is known to interact with two complexes known as TORC1 and TORC2, both of which are known to regulate their targets via phosphorylation (Cardenas et al., 1999; Bastidas et al., 2012). TORC1 is known to control various cellular processes such as protein synthesis, mRNA synthesis and degradation, autophagy, and nutrient transport; while TORC2 is involved in cell wall integrity and actin polarization (Kubota et al., 2003; Bastidas et al., 2012). TORC1 is the only of the two complexes to be sensitive to rapamycin via a similar pathway involving rapamycin binding to FKBP12 forming a rapamycin-FKBP12 complex that binds to TOR resulting in its inhibition (Bastidas et al., 2012; Nguyen et al., 2020). However, a chronic treatment with rapamycin can not only inhibit TORC1 but also disrupts TORC2 *in vivo* (Lamming et al., 2012). In the yeast fungus *Saccharomyces cerevisiae*, inhibition of TOR via rapamycin exposure results in multiple cellular responses such as protein synthesis inhibition and autophagy triggered by nutrient depletion (Cardenas et al., 1999; Bastidas et al., 2012; Nguyen et al., 2020). Therefore, TOR represents a critical target for the development of new antifungals. The second part of the review further highlights rapamycin-FKBP12-TOR interactions in pathogenic fungi.

INTERACTIONS BETWEEN FK506, FKBP12, AND CALCINEURIN IN PATHOGENIC FUNGAL SPECIES

Calcineurin is a heterodimer comprising of a catalytic A subunit and a regulatory B subunit; both subunits are required for its phosphatase activity (Rusnak and Mertz, 2000). The FK506-FKBP12 drug-receptor complex binds at the interface of the calcineurin B regulatory and catalytic A subunit and occludes substrate access (Liu et al., 1991). Thus, the formation of the FK506-FKBP12-calcineurin tertiary complex (**Figure 3A**)

FIGURE 3 | (A) A crystal structure of *A. fumigatus* calcineurin A and B, FKBP12, and FK506 [protein data bank (PDB): 6TZ7]. Calcineurin is a protein phosphatase comprising of a catalytic (green) and a regulatory subunit (yellow). The complex of FK506 (red) and FKBP12 (cyan) interacts with the calcineurin heterodimer to form a ternary complex and inhibit calcineurin function. **(B)** In the ribbon diagrams of human (left) and *A. fumigatus* (right) FKBP12-FK506, β sheets wrap around the α helix and the extended loops (such as the 80 and 40s loops) surround the FK506 binding pocket. Left: FK506-HsFKBP12 (PDB: 1FKJ), right: FK506-AfFKBP12 (PDB: 6TZ7). **(C)** Alignment of amino acids of the 40 and 80s loops of FKBP12. Hs, human; Af, *A. fumigatus*; Cn, *C. neoformans*; Ca, *C. albicans*; Mc, *M. circinelloides*. *Residue numbers are in respect to AfFKBP12.

is required for complete inhibition of calcineurin function. Here we detail how calcineurin plays a role in virulence and pathogenesis and how FK506, FKBP12, and calcineurin interact in pathogenic fungi.

Aspergillus fumigatus

Aspergillus fumigatus is the most frequent causative agent of invasive aspergillosis in immunocompromised patients (Brakhage et al., 2010; Baldin et al., 2015). Current treatment for aspergillosis primarily relies on the use of azoles, but polyenes and echinocandins are also used (Maschmeyer et al., 2007; Walsh et al., 2008). With rising antifungal resistance, demand for new antifungal drugs, and alternative treatment are on the rise (Howard et al., 2009; Wu et al., 2020). Previous studies have shown that calcineurin is required for cell wall integrity, hyphae formation, echinocandin resistance, virulence, and others in *A. fumigatus* (Steinbach et al., 2006; CramerJr., Perfect et al., 2008; Fortwendel et al., 2010; Juvvadi et al., 2014). Calcineurin inhibitor—FK506 exhibits antifungal activity *in vitro* (Steinbach et al., 2004). In a systemic model of murine aspergillosis, the infected mice treated with FK506 exhibited higher survival rates compared to the infected groups that did not receive FK506 treatment (High and Washburn, 1997). Together the data suggest that calcineurin is an attractive target for aspergillosis.

In *A. fumigatus*, FKBP12 is also the cellular receptor for FK506. Falloon et al. (2015) identified four orthologs of human FKBP12 (HsFKBP12) in *A. fumigatus*: *fkbp12-1, fkbp12-2, fkbp12-3, and fkbp12*-4. It was determined that *A. fumigatus* FKBP12-1 (AfFKBP12) was the binding partner for FK506. The AfFKBP12 shares 55% sequence similarity with HsFKBP12 (Falloon et al., 2015; Gobeil et al., 2019). The core structure of FKBP12 is conserved in both humans and fungi, and it comprises of five to six β sheets (β1–β6) that wrap around the central α helix (**Figure 3B**; Tonthat et al., 2016). Additionally, three extended loops surround the FK506 binding pocket—40s loop (between β2 and β3), the 50s loop (between β3′ and α1), and the 80s loop (between β4 and β5) (**Figure 3B**; Tonthat et al., 2016). Interestingly, when AfFKBP12 is present in the apo form, the 80s loop from one monomer docks into the active site of the second monomer and vice-versa to form a dimer possibly through self-catalysis function (Tonthat et al., 2016). Sequence alignment further revealed the presence of a specific proline residue at the tip of the 80s loop facilitates dimerization in AfFKBP12 but not in HsFKBP12 (**Figures 3B,C**). The self-substrate region overlaps with the FK506 binding region because alteration of the proline residue at the tip of the 80s loop to glycine conferred resistance to FK506 (Tonthat et al., 2016).

To evaluate the role of the 40s and 50s loop in AfFKBP12 interaction with FK506 and calcineurin, Juvvadi et al. (2020) induced amino acid point mutations in 40s (F37M/L) and 50s (W60V) loop. They found that in comparison to the 80s loop point mutations (P90G or V91C), alteration in the 40 and 50s loop increased contacts in the 80s loop resulting in decreased calcineurin binding and higher resistance to FK506 (Juvvadi et al., 2020). Although fungal and human FKBP12 share a striking similarity, there are non-conserved motifs in the 40s and 80s loop of fungal FKBP12 (**Figures 3B,C**; Juvvadi et al., 2019). Amino

acid substitutions in the 40s loop (F22T, Q50M, and R55E) and 80s loop (F88H) impacted the FK506-FKBP12 interaction with calcineurin. Particularly, replacing Phe with His (F88H) resulted in a significant increase in resistance to FK506 due to reduced binding of FK506-FKBP12 complex with calcineurin (Juvvadi et al., 2019). Interestingly, while H88 in the 80s loop is critical for mammalian FK506-FKBP12-calcineurin interaction, F88 is involved in *A. fumigatus* FK506-FKBP12-calcineurin interaction and inhibition of calcineurin function (**Figure 3C**; Juvvadi et al., 2019). Despite the differences, it is evident that alterations in the 80s loop can adversely affect the binding of both human and fungal FK506-FKBP12 complex to calcineurin (Tonthat et al., 2016; Juvvadi et al., 2019).

Candida albicans

Candida albicans is a commensal organism and can cause deep systemic infections in immunocompromised patients (Stelzner, 1990; Pfaller and Diekema, 2007; Low and Rotstein, 2011; Neville et al., 2015). Current treatments involve three different classes of antifungal drugs including azoles, echinocandins, and polyenes (Pierce et al., 2013; Vila et al., 2017). Antifungal drug resistance in *C. albicans* is not uncommon and is on the rise (Monroy-Pérez et al., 2016; Whaley et al., 2016).

Previous studies have shown that calcineurin is required for survival in serum and consequently, the calcineurin mutants failed to colonize tissues in infected murine animals (Blankenship et al., 2003; Blankenship and Heitman, 2005). Additionally, calcineurin is also required for tolerance to antifungal agents (Sanglard et al., 2003). Therefore, calcineurin represents a suitable target for candidiasis treatment, and understanding the structural differences between FKBP12 and calcineurin in *C. albicans* to humans is critical in developing non-immunosuppressive FK506 analogs.

It was previously reported that similar to *A. fumigatus*, the core structure of *C. albicans* FKBP12 (CaFKBP12) comprises of five β sheets that wrap around the central α helix and also has three extended loops (the 40s, 50s, and 80s) surrounding the FK506 binding pocket (Tonthat et al., 2016). Consistent with the observation made in *A. fumigatus*, the proline residue present at the tip (P104) of 80s loop docks into the active site of adjacent subunit and vice-versa facilitating FKBP12-FKBP12 interaction in apostate (Tonthat et al., 2016). The 80s loop bind in the same pocket as FK506. Interestingly, in *C. albicans* all the captured FKBP12s have P104 in *cis* conformation while in *A. fumigatus* the P90 in one subunit was bound in *cis* conformation while the P90 in other subunit was bound in a *trans*-state (**Figure 3C**; Tonthat et al., 2016).

Cryptococcus neoformans

Cryptococcus neoformans is the etiologic agent of cryptococcosis (Idnurm et al., 2005). Cryptococcosis primarily affects lungs in immunocompromised patients and can also cause lethal meningitis and encephalitis (Idnurm et al., 2005; Mednick et al., 2005; Bretaudeau et al., 2006). Currently, cryptococcosis can be treated with an azole such as fluconazole, but cryptococcal meningitis may require the use of polyenes such as amphotericin B (Saag et al., 2000). Clinically, the combination of amphotericin

B with flucytosine has resulted in better outcomes than treatment with amphotericin B alone (Larsen et al., 1990; Dromer et al., 2008). Regardless, there have been several reports of antifungal resistance in *C. neoformans* (Friese et al., 2001; Lortholary et al., 2006; Cheong and McCormack, 2013; Smith et al., 2015) indicating a need for the development of alternative treatment strategies.

FK506 is toxic to *C. neoformans* at 37°C, but not at 24°C (Odom et al., 1997a,b), thereby suggesting that calcineurin is required for growth at higher temperatures. As a consequence, the calcineurin mutants are not virulent in murine models of cryptococcal meningitis (Odom et al., 1997b). Cruz et al. identified the homolog of FK506 binding partner FKBP12 in *C. neoformans*. They found that CnFKBP12 share 59 and 57% sequence similarity with the HsFKBP12 and CaFKBP12, respectively. The amino acid residues that form the FK506-binding pocket were highly conserved across species. A single amino acid substitution in this conserved site—W60R destabilized CnFKBP12 and rendered resistance to FK506 and rapamycin (Cruz et al., 1999). Consistent with the findings in *C. albicans* and *A. fumigatus,* CnFKBP12 also has conserved phenylalanine in the 80s loop (F88) which is absent in HsFKBP12 (**Figure 3C**; Juvvadi et al., 2019).

Mucor circinelloides

Mucormycosis is an emerging lethal infection caused by fungi of order Mucorales (Petrikkos et al., 2012). *M. circinelloides* is one of the causative agents of this deadly disease (López-Fernández et al., 2018; Vellanki et al., 2018). When *M. circinelloides* is grown in the presence of FK506 it exhibits yeast growth instead of hyphae which shows that calcineurin regulates dimorphism in this fungus (hyphal to yeast transition) (Lee et al., 2013). Mutants lacking functional calcineurin are less virulent than the wild type which suggests that calcineurin is a key target for the treatment of mucormycosis (Lee et al., 2013, 2015; Vellanki et al., 2020). *M. circinelloides* encodes three calcineurin catalytic subunits (CnaA, CnaB, and CnaC) and one regulatory subunit (CnbR) (Lee et al., 2013). We have previously reported that mutations in the FK506-FKBP12 binding region of CnaA and the latch region of CnbR confer resistance to FK506 (Lee et al., 2013, 2015). Specifically, the CnbR mutation N125Y disrupted CnbR-FKBP12 interaction due to steric clashes with residue F47 and Q48 of McFKBP12 (Lee et al., 2015). The CnaA residue W377 is required for FK506-CnaA hydrophobic interaction and spontaneous mutation resulting in W377L was predicted to disrupt the hydrophobic pocket at the binding interface (Lee et al., 2015).

Gobeil et al. (2020) found that the amino acid sequence of McFKBP12 is 58% identical to HsFKBP12. The sequence variations were present in the 40, 50, and 80s extended loops surrounding the FK506 binding pocket. Interestingly, McFKBP12 sequence is 65% identical to AfFKBP12 and share structural similarity but functionally they are not equivalent in inhibiting calcineurin function (Gobeil et al., 2020). When *Affkbp12* was replaced with *Mcfkbp12* in *A. fumigatus*, FK506 was not able to inhibit calcineurin

function. As mentioned above, in the presence of FK506 the F88 residue in the 80s loop of AfFKBP12 is critical in the FK506-AfFKBP12-calcineurin interaction. Interestingly, McFKBP12 possesses a tyrosine at the 88th residue instead of phenylalanine because of which McFKBP12-FK506 failed to inhibit calcineurin function in *A. fumigatus* (**Figure 3C**; Gobeil et al., 2020).

Efficacy of FK506 Analogs Against Pathogenic Fungi

Fungal infections are prevalent in immunocompromised patients, therefore FK506 cannot be used as an antifungal to treat patients. To circumvent this problem several attempts have been made in generating FK506 analogs that have a lesser binding affinity toward human calcineurin (to reduce immunosuppression) while retaining antifungal properties.

L685,818 (**Figure 2**) is an FK506 analog and antagonist (Dumont et al., 1992). It is a C18 hydroxy, C21ethyl derivative of FK506. The presence of the C18 hydroxy group prevents L685,818-HsFKBP12 from inhibiting bovine calcineurin, and subsequently, L685,818 was found to be significantly less immunosuppressive (Dumont et al., 1992; Rotonda et al., 1993). Unlike, HsFKBP12 the yeast FKBP12 bound to L685,818 can inhibit bovine calcineurin function (Rotonda et al., 1993), and therefore its antifungal efficacy was evaluated. The analog L685,818 inhibited the growth of *C. neoformans* at 37°C, however, its potency was 10–100 times lower than FK506 (Odom et al., 1997a). Also, L685,818 alone or in combination with caspofungin did not exhibit antifungal activity against *A. fumigatus* (Kontoyiannis et al., 2003).

Nambu et al. (2017), synthesized FK506 antagonists which are permeable in mammalian cells but not in fungal cells, this way a dichotomy of FK506 and the antagonist will result in an inhibition of fungal calcineurin only. The antagonist compounds 15 and 18 (modifications shown in **Figure 2**) in combination with FK506 significantly reduced *A. fumigatus* proliferation compared to a challenge with antagonist alone.

FK506 analogs: 9-deoxo-FK506 (9D-FK506), 9-deoxo-prolyl-FK506 (9DP-FK506), 31-O-demethyl-FK506 (31OD-FK506), and 9-deoxo-31-O-demethyl-FK506 (9D31OD-FK506) (**Figure 2**) were generated by manipulating FK506 biosynthetic genes in *Streptomyces* spp. (Ban et al., 2013; Shinde et al., 2015). These analogs showed a significantly less immunosuppressive effect *in vitro* than FK506 (Lee et al., 2018). All the analogs except 9DP-FK506 exhibited antifungal activity against *C. neoformans*, *A. fumigatus*, and *C. albicans in vitro*, however, their efficacy was lower when compared to FK506. In a murine model of systemic cryptococcal infection, the infected mice treated with 9D31OD-FK506 and fluconazole exhibited prolonged survival when compared to the infected mice treated with either of the agents alone. This confirms that an *in vivo* synergistic interaction occurs between FK506 analog and fluconazole (Lee et al., 2018).

In another study, Beom et al. (2019) generated seven new FK506 analogs with modifications in either FKBP12 or

calcineurin binding sites. The 9-deoxo-21-ethyl-FK506 analog (named as compound 7 in the reference; **Figure 2**) exhibited significantly higher antifungal activity against *C. albicans* and *A. fumigatus* when compared to the previously described analog 9D31OD-FK506. This new analog also exhibited synergistic antifungal activity with fluconazole against *C. neoformans* (Beom et al., 2019).

A comparison of fungal and human FK506-FKBP12-calcineurin ternary structures revealed key differences between the species (Juvvadi et al., 2019). While a His 88 at the interface of FK506-HsFKBP12 is required for productive binding and inhibition of human calcineurin, Phe 88 on fungal FKBP12 is required to inhibit fungal calcineurin (**Figure 3C**). A Val 91 residue at the interface of fungal FK506-FKBP12 is also not conserved in mammalian counterparts. These differences were exploited to generate an FK506 analog—APX879, characterized by an acetohydrazide substitution at C22-ketone. The APX879 (**Figure 2**) interacts less favorably than FK506 with HsFKBP12 due to a higher steric clash with H88 residue (Juvvadi et al., 2019; Gobeil et al., 2020). Subsequently, APX879 was found to be 71-fold less immunosuppressive than FK506, as measured by their ability to induce IL-2 production from differentiated CD4$^+$ T cells. APX879 exhibits antifungal activity against *C. neoformans*, *C. albicans*, *A. fumigatus*, and *M. circinelloides*, however, its efficacy is lower than FK506 *in vitro*. Interestingly, APX879 is less toxic than FK506 in a murine model of cryptococcal infection. Additionally, infected animals treated with a combination of APX879 and fluconazole exhibited significantly extended median survival rates compared to infected animals treated with APX879 or fluconazole alone. The APX879 was not efficacious in murine models of aspergillosis, candidiasis, and mucormycosis (Juvvadi et al., 2019).

RAPAMYCIN—FKBP12 INTERACTION IN PATHOGENIC FUNGAL SPECIES

The interaction between rapamycin and FKBP12 (**Figure 4**) is known to result in the inhibition of the TOR pathway in fungi. This inhibition has been known to affect various cellular responses such as protein synthesis inhibition and autophagy (Cardenas et al., 1999; Bastidas et al., 2012). This rapamycin—FKBP12 interaction in pathogenic fungi has become the focus of various research groups as summarized below.

Aspergillus fumigatus
Little is known about the role of TOR in this filamentous fungus, although research has found that there may be a link with mitochondrial processes (Baldin et al., 2015). Specifically, a link between ornithine/arginine biosynthesis and an iron deficiency stress response of *A. fumigatus* have been studied (Baldin et al., 2015). The deletion of the single *tor* gene in *A. fumigatus* results in a lethal phenotype, thus researchers had to develop a conditional lethal *tor* mutant (Baldin et al., 2015). This was achieved by replacing the endogenous *tor* gene by the inducible *xylp-tor* gene cassette (Baldin et al.,

FIGURE 4 | Binding of rapamycin to the bovine FKBP12 (PDB: 1FKL).
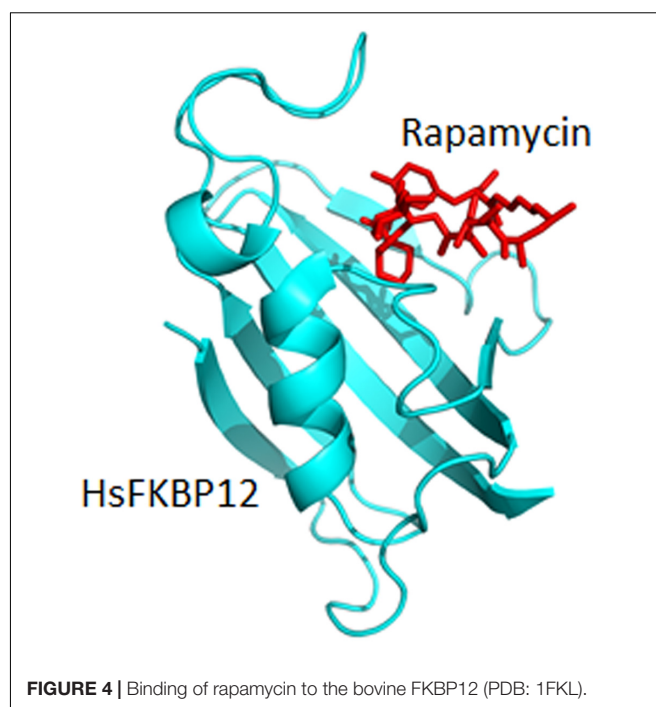
2015). Furthermore, the qRT-PCR analysis uncovered an increase in mRNA levels of ornithine biosynthesis genes under iron limitations, and iron regulation was lost when *tor* was repressed (Baldin et al., 2015). Furthermore, the iron regulator HapX was also investigated and researchers found that *hapx* expression was significantly reduced when *tor* was repressed (Baldin et al., 2015). This research concluded that the function of TOR may have two scenarios in *A. fumigatus*; TOR either acts upstream of HapX, or independently of HapX as a repressor of ornithine biosynthesis (Baldin et al., 2015). Finally, the effects of rapamycin on this inducible mutant were investigated. Rapamycin showed strong growth inhibition on the inducible mutant (Baldin et al., 2015). However, the increase of *tor* mRNA levels in the inducible mutant resulted in a slightly increased resistance against rapamycin (Baldin et al., 2015). Previous research also found a single *tor* gene in *Aspergillus nidulans*, a less frequent causal agent of aspergillosis in immunocompromised patients (Loewith and Hall, 2011).

Candida albicans
The necessity for new or alternative treatments for candidiasis is urgent due to emerging antifungal resistance. Previous studies have deduced that rapamycin presents antifungal activity against *C. albicans* (Sehgal et al., 1975). Rapamycin is known to diffuse into the cell and associate with FKBP12 (Cardenas et al., 1999; Bastidas et al., 2012). The newly formed FKBP12-rapamycin complex will target the two TOR kinases Tor1 and Tor2 (Cardenas et al., 1999; Bastidas et al., 2012). In *C. albicans*, the *RBP1* gene is known to encode a homolog of FKBP12 (Cruz et al., 2001). Interestingly, studies have shown that mutant strains lacking the *RBP1* and *TOR1* genes are viable and rapamycin resistant (Cruz et al., 2001). These

findings show that rapamycin antifungal activity is exerted via FKBP12 and Tor1 homologs in *C. albicans* (Cruz et al., 2001). One major contributor to antifungal drug resistance in *C. albicans* is loss of genes encoding ATP-binding cassette (ABC) transporters in the genomes (Lyons and White, 2000; Khandelwal et al., 2018). These transporters help export various substrates across the cell membrane including antifungal drugs (Khandelwal et al., 2018). Previous studies found that the loss of the transporter *CDR6/ROA1* results in a strain that is more resistant to antifungal treatment with azoles (Khandelwal et al., 2018). Researchers developed a *CDR6/ROA1* null strain and found that along with the increase in azole resistance, transcriptional profiling uncovered ribosome biogenesis genes were significantly upregulated (Khandelwal et al., 2018). This finding led to the discovery of a suspected link between ribosome biogenesis and TOR1 signaling (Khandelwal et al., 2018). Furthermore, the mutant strain was grown on media supplemented with rapamycin where TOR1 hyperactivation was observed (Khandelwal et al., 2018). This led to an Hsp90-dependent calcineurin stabilization resulting in an increase to azole resistance (Khandelwal et al., 2018). These findings were then replicated *in vivo* with a systemic infection mouse model, which resulted in a higher fungal load post fluconazole treatment in mice infected with this mutant strain (Khandelwal et al., 2018). This study uncovered a novel mechanism of azole resistance in *C. albicans* involving TOR signaling.

Cryptococcus neoformans

Although the TOR pathway is a highly conserved mechanism across eukaryotes, its role in *C. neoformans* is a topic of debate (So et al., 2019). Researchers have found two Tor-like kinases, Tor1 and Tlk1, in *C. neoformans* (So et al., 2019). This study discovered that *TLK1* is a dispensable gene resulting in the cultivation of a viable strain following its deletion. The opposite was concluded when *TOR1* was deleted (So et al., 2019). Further research explored the function of Tor1 by creating strains that overexpress the *TOR1* gene (So et al., 2019). It was concluded that Tor1 negatively regulates two crucial virulence factors, DNA damage response, and thermotolerance by reducing Rad53 and Mpk1 phosphorylation, respectively (So et al., 2019). When inhibiting TOR via rapamycin treatment actin depolarization in a Tor-1 dependent manner was observed (So et al., 2019). Finally, various rapamycin-sensitive and resistant strains were screened uncovering that the TOR pathway can possibly crosstalk with various stress signaling pathways (So et al., 2019).

Mucor circinelloides

One major hurdle encountered by physicians treating mucormycosis patients is that Mucorales are resistant to most antifungal drugs (Caramalho et al., 2017; Dannaoui, 2017). Thus, it has become of great interest to scientists to develop a greater understanding of how this fungus can evade antifungal drugs. One major target of interest is understanding the interaction between rapamycin and FKBP12 in this fungus. A recent study found that the target and mechanism of rapamycin in *M. circinelloides* is mediated via conserved

FKBP12 and Tor homologs (Bastidas et al., 2012). In this study researchers utilized spontaneous mutations that disrupted conserved residues in FKBP12, this resulted in rapamycin and FK506 resistance in these mutants (Bastidas et al., 2012). Total disruption of the FKBP12-encoding gene (*fkbpA*) also resulted in rapamycin and FK506 resistance (Bastidas et al., 2012). To further confirm their findings the expression of *M. circinelloides* FKBP12 was complemented in an *S. cerevisiae* mutant strain lacking FKBP12 resulting in the restoration of rapamycin sensitivity (Bastidas et al., 2012). Furthermore, it was found that the *M. circinelloides* FKBP12 and Tor interacted in a rapamycin-dependent fashion (Bastidas et al., 2012). Finally, *in vitro* studies found that rapamycin exhibited potent growth inhibition against *M. circinelloides* followed by an *in vivo* study with *Galleria mellonella* improving survival by 50% post-rapamycin treatment (Bastidas et al., 2012). These findings suggest that rapamycin can be an attractive drug target, and the development of rapamycin analogs with a lowered immunosuppressive activity can have the potential for novel antifungal treatment strategies.

Novel Small Molecules and Rapamycin Analogs Targeting TOR

We have previously established that the TOR pathway in fungi can regulate various cell stress pathways (Cardenas et al., 1999; Bastidas et al., 2012). This makes the TOR pathway an attractive target for the development of new antifungals. Researchers have previously identified a small molecule—beauvericin that not only potentiated antifungal treatment but was also tolerated by human cells (Shekhar-Guturja et al., 2016). Previous studies have found that beauvericin can enhance azole efficacy against the top leading fungal pathogens *C. albicans, C. neoformans*, and *A. fumigatus* (Shekhar-Guturja et al., 2016). Beauvericin has also shown the ability to block the emergence of resistance and even render resistant strains responsive to treatment (Shekhar-Guturja et al., 2016). The previously mentioned effects have been determined to be mediated via inhibition of multidrug efflux pumps and TORC1 signaling (Shekhar-Guturja et al., 2016). This activates the protein kinase CK2 resulting in the inhibition of the molecular chaperone Hsp90 (Shekhar-Guturja et al., 2016). Furthermore, this research found that substitutions in the efflux transporter Pdr5 that enable beauvericin efflux results in the impairment of azole efflux halting the resistance to this drug combination (Shekhar-Guturja et al., 2016). Similarly, an ATP-competitive TOR kinase inhibitor INK128 has been studied for the possible use of combination therapy with various azoles (Gao et al., 2016). In this study, 23 strains of *Aspergillus* were tested, all of which were clinical isolates derived from patients with invasive aspergillosis (Gao et al., 2016). The function of INK128 involves the inhibitor binding to the TOR catalytic domain and selectively inhibiting TOR (Gao et al., 2016). It was concluded that the efficacy of the azoles tested was significantly increased, although the same was not reported when tested with echinocandins and polyenes (Gao et al., 2016).

The synthesis of rapamycin analogs that present a lowered immunosuppressive activity in humans has been previously

explored (Cruz et al., 2001; Guo et al., 2019). One major research involved synthesizing over 45,000 rapamycin-inspired macrocycles (Guo et al., 2019). This massive undertaking was achieved by creating a rapamycin-like macromolecule library by replacing the effector domain of rapamycin with a combinatorial library of oligopeptides (Guo et al., 2019). By the use of ring-closing metathesis, these researchers were able to develop a robust macrocyclization method resulting in the synthesis of a 45,000-compound library of these hybrid macrocycles, named rapafucins, which utilized optomized FKBP-binding domains (Guo et al., 2019). Although this specific study was focused on making rapamycin-like molecules that can bind to new cellular proteins to aid with different types of diseases and disorders; we believe it is important to highlight the potential of synthesizing 45,000 rapamycin-like macrocycles and the possible therapeutic potentials such compounds may possess. Further studies have uncovered a series of rapamycin analogs that present a lowered immunosuppressive activity while presenting antifungal properties in *C. albicans* (Cruz et al., 2001). In this study, four analogs—analog 2, analog 18, analog 19, and analog 23 (**Figure 5**) were studied to determine their antifungal efficacy compared to rapamycin (Cruz et al., 2001). Previous studies

found that the FKBP12 homolog Rbp1 is required for rapamycin antifungal function (Cruz et al., 2001). This was confirmed by developing a *C. albicans rbp1/rbp1* mutant strain lacking a homolog of the FKBP12 protein (Cruz et al., 2001). It was found that these mutants were viable and resistant to both rapamycin and its analogs (Cruz et al., 2001). With regards to the fungicidal activity presented by these analogs, analogs 2 and 23 were found to be toxic and fungicidal to *C. albicans, C. neoformans,* and *S. cerevisiae,* while analogs 18 and 19 were weakly toxic to *C. neoformans* and had no activity against the other fungi tested. In this study, analogs 2 and 23 were found to be the most toxic to *C. albicans* and *C. neoformans* while sharing a common mechanism of action with rapamycin involving an FKBP12-dependent inhibition of the TOR kinases (Cruz et al., 2001).

CONCLUSION

Fungal calcineurin and TOR are key targets to develop new antifungals. In this review, we highlighted the structural differences between human and fungal FKBP12, calcineurin, and TOR components. FK506 targets both human and fungal calcineurin. However, several FK506 analogs, particularly which had modifications at C9, C21, C22, and C31 had increased affinity toward fungal FKBP12 and calcineurin resulting in lower immunosuppression while retaining antifungal activity. The sequences of FKBP12, calcineurin, and TOR components vary among fungal species. As a result, the antifungal efficacy of these analogs varies among the fungal pathogens tested. Nevertheless, these studies provide a proof of concept for developing non-immunosuppressive FK506 or rapamycin analogs with antifungal activity.

AUTHOR CONTRIBUTIONS

SL and SV conceived the study. SV, AG, and SL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FIGURE 5 | The chemical structures of rapamycin and its analogs in this study. The numbers in red font indicate the modified positions to generate analogs.

Compound	Modification
Analog 2	C40 (R)-OH
Analog 18	C7 (S)-2-Me-Thienyl
Analog 19	C7 (S)-NOHCOOiBu
Analog 23	C7 (S)-NOHCON-Piperidyl

ACKNOWLEDGMENTS

We would like to thank Courtney Smith for critical reading of the manuscript.
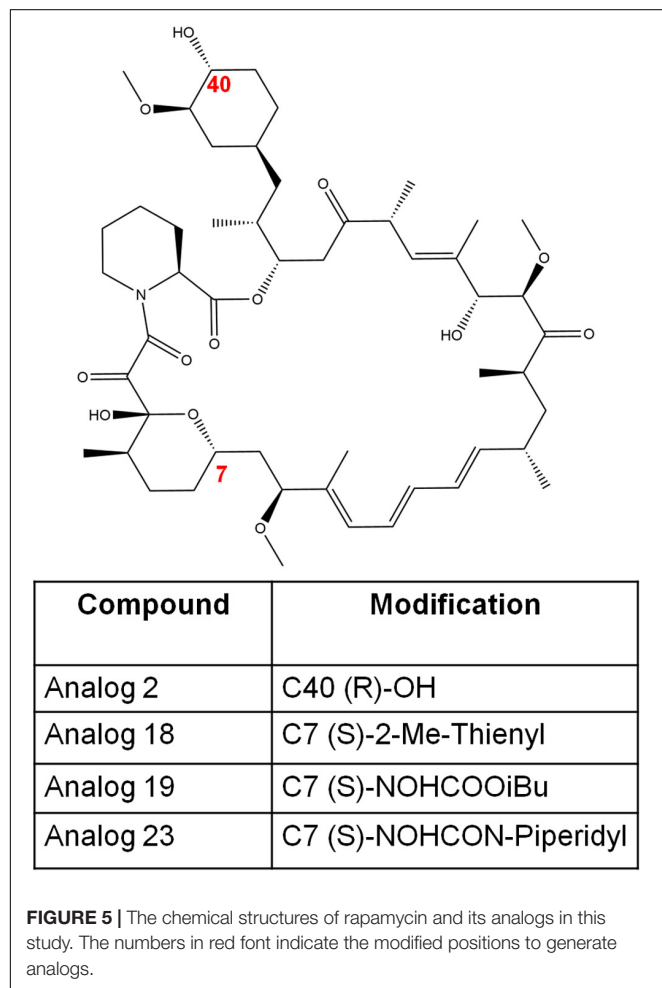
REFERENCES

Baldin, C., Valiante, V., Krüger, T., Schafferer, L., Haas, H., Kniemeyer, O., et al. (2015). Comparative proteomics of a TOR inducible *Aspergillus fumigatus* mutant reveals involvement of the Tor kinase in iron regulation. *Proteomics* 15, 2230–2243. doi: 10.1002/pmic.201400584

Ban, Y. H., Park, S. R., and Yoon, Y. J. (2016). The biosynthetic pathway of FK506 and its engineering: from past achievements to future prospects. *J. Ind. Microbiol. Biotechnol.* 43, 389–400. doi: 10.1007/s10295-015-1677-7

Ban, Y. H., Shinde, P. B., Hwang, J.-Y., Song, M.-C., Kim, D. H., Lim, S.-K., et al. (2013). Characterization of FK506 biosynthetic intermediates involved in post-PKS elaboration. *J. Nat. Prod.* 76, 1091–1098. doi: 10.1021/np4001224

Bastidas, R. J., Shertz, C. A., Lee, S. C., Heitman, J., and Cardenas, M. E. (2012). Rapamycin exerts antifungal activity in vitro and in vivo against Mucor circinelloides via FKBP12-dependent inhibition of TOR. *Eukaryot. Cell* 11, 270–281. doi: 10.1128/ec.05284-11

Bechman, K., Galloway, J. B., and Winthrop, K. L. (2019). Small-molecule protein kinases inhibitors and the risk of fungal infections. *Curr. Fungal Infect. Rep.* 13, 229–243. doi: 10.1007/s12281-019-00350-w

Beom, J. Y., Jung, J. A., Lee, K.-T., Hwangbo, A., Song, M. C., Lee, Y., et al. (2019). Biosynthesis of nonimmunosuppressive FK506 analogues with antifungal activity. *J. Nat. Prod.* 82, 2078–2086. doi: 10.1021/acs.jnatprod.9b00144

Blankenship, J. R., and Heitman, J. (2005). Calcineurin is required for Candida albicans to survive calcium stress in serum. *Infect. Immun.* 73, 5767–5774. doi: 10.1128/iai.73.9.5767-5774.2005

Blankenship, J. R., Wormley, F. L., Boyce, M. K., Schell, W. A., Filler, S. G., Perfect, J. R., et al. (2003). Calcineurin is essential for Candida albicans survival in serum and virulence. *Eukaryot. Cell* 2, 422–430. doi: 10.1128/ec.2.3.422-430.2003

Brakhage, A. A., Bruns, S., Thywissen, A., Zipfel, P. F., and Behnsen, J. (2010). Interaction of phagocytes with filamentous fungi. *Curr. Opin. Microbiol.* 13, 409–415. doi: 10.1016/j.mib.2010.04.009

Bretaudeau, K., Eloy, O., Richer, A., Bruneel, F., Scott-Algara, D., Lortholary, O., et al. (2006). Cryptococcal meningo-encephalitis in an apparently immunocompetent patient. *Rev. Neurol.* 162, 233–237. doi: 10.1016/s0035-3787(06)75005-6

Caramalho, R., Tyndall, J. D. A., Monk, B. C., Larentis, T., Lass-Flörl, C., and Lackner, M. (2017). Intrinsic short-tailed azole resistance in mucormycetes is due to an evolutionary conserved aminoacid substitution of the lanosterol 14α-demethylase. *Sci. Rep.* 7:15898. doi: 10.1038/s41598-017-16123-9

Cardenas, M. E., Cutler, N. S., Lorenz, M. C., Di Como, C. J., and Heitman, J. (1999). The TOR signaling cascade regulates gene expression in response to nutrients. *Genes Dev.* 13, 3271–3279. doi: 10.1101/gad.13.24.3271

Cheong, J. W. S., and McCormack, J. (2013). Fluconazole resistance in cryptococcal disease: emerging or intrinsic? *Med. Mycol.* 51, 261–269. doi: 10.3109/13693786.2012.715763

Cortés, J. C. G., Curto, M., Carvalho, V. S. D., Pérez, P., and Ribas, J. C. (2019). The fungal cell wall as a target for the development of new antifungal therapies. *Biotechnol. Adv.* 37, 107352. doi: 10.1016/j.biotechadv.2019.02.008

Cramer, R. A. Jr., Perfect, B. Z., Pinchai, N., Park, S., Perlin, D. S., Asfaw, Y. G., et al. (2008). Calcineurin target CrzA regulates conidial germination, hyphal growth, and pathogenesis of *Aspergillus fumigatus*. *Eukaryot. Cell* 7, 1085–1097. doi: 10.1128/ec.00086-08

Cruz, M. C., Cavallo, L. M., Görlach, J. M., Cox, G., Perfect, J. R., Cardenas, M. E., et al. (1999). Rapamycin antifungal action is mediated via conserved complexes with FKBP12 and TOR kinase homologs in Cryptococcus neoformans. *Mol. Cell. Biol.* 19, 4101–4112. doi: 10.1128/mcb.19.6.4101

Cruz, M. C., Goldstein, A. L., Blankenship, J., Del Poeta, M., Perfect, J. R., McCusker, J. H., et al. (2001). Rapamycin and less immunosuppressive analogs are toxic to Candida albicans and Cryptococcus neoformans via FKBP12-Dependent Inhibition of Tor. *Antimicrob. Agents Chemother.* 45:3162. doi: 10.1128/AAC.45.11.3162-3170.2001

Cully, M. (2018). Antifungal drugs: small molecules targeting a tertiary RNA structure fight fungi. *Nat. Rev. Drug Discov.* 17:864. doi: 10.1038/nrd.2018.205

Dannaoui, E. (2017). Antifungal resistance in mucorales. *Int. J. Antimicrob. Agents* 50, 617–621. doi: 10.1016/j.ijantimicag.2017.08.010

Dromer, F., Bernede-Bauduin, C., Guillemot, D., and Lortholary, O. (2008). Major role for amphotericin B-flucytosine combination in severe cryptococcosis. *PLoS One* 3:e2870. doi: 10.1371/journal.pone.0002870

Dumont, F. J., Staruch, M. J., Koprak, S. L., Siekierka, J. J., Lin, C. S., Harrison, R., et al. (1992). The immunosuppressive and toxic effects of FK506 are mechanistically related: pharmacology of a novel antagonist of FK506 and rapamycin. *J. Exp. Med.* 176, 751–760. doi: 10.1084/jem.176.3.751

Enoch, D. A., Ludlam, H. A., and Brown, N. M. (2006). Invasive fungal infections: a review of epidemiology and management options. *J. Med. Microbiol.* 55(Pt. 7), 809–818. doi: 10.1099/jmm.0.46548-0

Falloon, K., Juvvadi, P. R., Richards, A. D., Vargas-Muñiz, J. M., Renshaw, H., and Steinbach, W. J. (2015). Characterization of the FKBP12-encoding genes in *Aspergillus fumigatus*. *PLoS One* 10:e0137869. doi: 10.1371/journal.pone.0137869

Fortwendel, J. R., Juvvadi, P. R., Perfect, B. Z., Rogg, L. E., Perfect, J. R., and Steinbach, W. J. (2010). Transcriptional regulation of chitin synthases by calcineurin controls paradoxical growth of *Aspergillus fumigatus* in response to caspofungin. *Antimicrob. Agents Chemother.* 54, 1555–1563. doi: 10.1128/aac.00854-09

Friese, G., Discher, T., Füssle, R., Schmalreck, A., and Lohmeyer, J. (2001). Development of azole resistance during fluconazole maintenance therapy for AIDS-associated cryptococcal disease. *AIDS* 15, 2344–2345. doi: 10.1097/00002030-200111230-00026

Gao, L., Ding, X., Liu, Z., Wu, Q., Zeng, T., and Sun, Y. (2016). In vitro interactions between target of rapamycin kinase inhibitor and antifungal agents against Aspergillus Species. *Antimicrob. Agents Chemother.* 60:3813. doi: 10.1128/AAC.02921-15

Garcia, C., Burgain, A., Chaillot, J., Pic, É, Khemiri, I., and Sellam, A. (2018). A phenotypic small-molecule screen identifies halogenated salicylanilides as inhibitors of fungal morphogenesis, biofilm formation and host cell invasion. *Sci. Rep.* 8:11559. doi: 10.1038/s41598-018-29973-8

Gobeil, S. M.-C., Bobay, B. G., Juvvadi, P. R., Cole, D. C., Heitman, J., Steinbach, W. J., et al. (2020). Designing selective and non-immunosuppressive antifungal FK506 analogs: structures, biophysics and dynamics of fungal and human calcineurin-inhibitor complexes. *bioRxiv[Preprint]*. doi: 10.1101/2020.04.14.039800

Gobeil, S. M. C., Bobay, B. G., Spicer, L. D., and Venters, R. A. (2019). (15)N, (13)C and (1)H resonance assignments of FKBP12 proteins from the pathogenic fungi *Mucor circinelloides* and *Aspergillus fumigatus*. *Biomol. NMR Assign.* 13, 207–212. doi: 10.1007/s12104-019-09878-x

Guo, Z., Hong, S. Y., Wang, J., Rehan, S., Liu, W., Peng, H., et al. (2019). Rapamycin-inspired macrocycles with new target specificity. *Nat. Chem.* 11, 254–263. doi: 10.1038/s41557-018-0187-4

Heitman, J., Koller, A., Kunz, J., Henriquez, R., Schmidt, A., Movva, N. R., et al. (1993). The immunosuppressant FK506 inhibits amino acid import in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 13, 5010–5019. doi: 10.1128/mcb.13.8.5010

Heitman, J., Movva, N. R., and Hall, M. N. (1991). Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science* 253, 905–909. doi: 10.1126/science.1715094

High, K. P., and Washburn, R. G. (1997). Invasive aspergillosis in mice immunosuppressed with cyclosporin A, tacrolimus (FK506), or sirolimus (rapamycin). *J. Infect. Dis.* 175, 222–225. doi: 10.1093/infdis/175.1.222

Howard, S. J., Cerar, D., Anderson, M. J., Albarrag, A., Fisher, M. C., Pasqualotto, A. C., et al. (2009). Frequency and evolution of Azole resistance in *Aspergillus fumigatus* associated with treatment failure. *Emerg. Infect. Dis.* 15, 1068–1076. doi: 10.3201/eid1507.090043

Husain, S., Alexander, B. D., Munoz, P., Avery, R. K., Houston, S., Pruett, T., et al. (2003). Opportunistic mycelial fungal infections in organ transplant recipients: emerging importance of non-Aspergillus mycelial fungi. *Clin. Infect. Dis* 37, 221–229. doi: 10.1086/375822

Hwang, H.-Y., Cho, S. M., and Kwon, H. J. (2017). Approaches for discovering novel bioactive small molecules targeting autophagy. *Exp. Opin. Drug Discov.* 12, 909–923. doi: 10.1080/17460441.2017.1349751

Idnurm, A., Bahn, Y.-S., Nielsen, K., Lin, X., Fraser, J. A., and Heitman, J. (2005). Deciphering the Model Pathogenic Fungus Cryptococcus Neoformans. *Nat. Rev. Microbiol.* 3, 753–764. doi: 10.1038/nrmicro1245

Juvvadi, P. R., Bobay, B. G., Gobeil, S. M. C., Cole, D. C., Venters, R. A., Heitman, J., et al. (2020). FKBP12 dimerization mutations effect FK506 binding and differentially alter calcineurin inhibition in the human pathogen Aspergillus

fumigatus. *Biochem. Biophys. Res. Commun.* 526, 48–54. doi: 10.1016/j.bbrc.2020.03.062

Juvvadi, P. R., Fox, D. III, Bobay, B. G., Hoy, M. J., Gobeil, S. M. C., Venters, R. A., et al. (2019). Harnessing calcineurin-FK506-FKBP12 crystal structures from invasive fungal pathogens to develop antifungal agents. *Nat. Commun.* 10:4275. doi: 10.1038/s41467-019-12199-1

Juvvadi, P. R., Lamoth, F., and Steinbach, W. J. (2014). Calcineurin as a multifunctional regulator: unraveling novel functions in fungal stress responses, hyphal growth, drug resistance, and pathogenesis. *Fungal Biol. Rev.* 28, 56–69. doi: 10.1016/j.fbr.2014.02.004

Khandelwal, N. K., Chauhan, N., Sarkar, P., Esquivel, B. D., Coccetti, P., Singh, A., et al. (2018). Azole resistance in a *Candida albicans* mutant lacking the ABC transporter CDR6/ROA1 depends on TOR signaling. *J. Biol. Chem.* 293, 412–432.

Kino, T., Hatanaka, H., Hashimoto, M., Nishiyama, M., Goto, T., Okuhara, M., et al. (1987). FK506, a novel immunosuppressant isolated from a Streptomyces. I. Fermentation, isolation, and physico-chemical and biological characteristics. *J Antibiot* 40, 1249–1255. doi: 10.7164/antibiotics.40.1249

Kontoyiannis, D. P., Lewis, R. E., Osherov, N., Albert, N. D., and May, G. S. (2003). Combination of caspofungin with inhibitors of the calcineurin pathway attenuates growth in vitro in Aspergillus species. *J. Antimicrob. Chemother.* 51, 313–316. doi: 10.1093/jac/dkg090

Kubota, H., Obata, T., Ota, K., Sasaki, T., and Ito, T. (2003). Rapamycin-induced translational derepression of GCN4 mRNA involves a novel mechanism for activation of the eIF2 alpha kinase GCN2. *J. Biol. Chem.* 278, 20457–20460. doi: 10.1074/jbc.C300133200

Lamming, D. W., Ye, L., Katajisto, P., Goncalves, M. D., Saitoh, M., Stevens, D. M., et al. (2012). Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science* 335, 1638–1643. doi: 10.1126/science.1215135

Larsen, R. A., Leal, M. A., and Chan, L. S. (1990). Fluconazole compared with amphotericin B plus flucytosine for cryptococcal meningitis in AIDS. A randomized trial. *Ann. Intern. Med.* 113, 183–187. doi: 10.7326/0003-4819-113-3-183

Lee, S. C., Li, A., Calo, S., and Heitman, J. (2013). calcineurin plays key roles in the dimorphic transition and virulence of the human pathogenic zygomycete *Mucor circinelloides. PLoS Pathog* 9:e1003625. doi: 10.1371/journal.ppat.1003625

Lee, S. C., Li, A., Calo, S., Inoue, M., Tonthat, N. K., Bain, J. M., et al. (2015). Calcineurin orchestrates dimorphic transitions, antifungal drug responses and host–pathogen interactions of the pathogenic mucoralean fungus *Mucor circinelloides. Mol. Microbiol.* 97, 844–865. doi: 10.1111/mmi.13071

Lee, Y., Lee, K.-T., Lee, S. J., Beom, J. Y., Hwangbo, A., Jung, J. A., et al. (2018). In vitro and in vivo assessment of FK506 analogs as novel antifungal drug candidates. *Antimicrob. Agents Chemother.* 62, e1627–e1618. doi: 10.1128/aac.01627-18

Lima, S. L., Colombo, A. L., and de Almeida Junior, J. N. (2019). Fungal cell wall: emerging antifungals and drug resistance. *Front. Microbiol.* 10:2573. doi: 10.3389/fmicb.2019.02573

Liu, J., Farmer, J. D. Jr., Lane, W. S., Friedman, J., Weissman, I., Schreiber, S. L., et al. (1991). Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* 66, 807–815. doi: 10.1016/0092-8674(91)90124-h

Loewith, R., and Hall, M. N. (2011). Target of Rapamycin (TOR) in nutrient signaling and growth control. *Genetics* 189:1177. doi: 10.1534/genetics.111.133363

López-Fernández, L., Sanchis, M., Navarro-Rodríguez, P., Nicolás, F. E., Silva-Franco, F., Guarro, J., et al. (2018). Understanding *Mucor circinelloides* pathogenesis by comparative genomics and phenotypical studies. *Virulence* 9, 707–720. doi: 10.1080/21505594.2018.1435249

Lortholary, O., Poizat, G., Zeller, V., Neuville, S., Boibieux, A., Alvarez, M., et al. (2006). Long-term outcome of AIDS-associated cryptococcosis in the era of combination antiretroviral therapy. *AIDS* 20, 2183-2191.

Low, C. Y., and Rotstein, C. (2011). Emerging fungal infections in immunocompromised patients. *F1000 Med. Rep.* 3:14. doi: 10.3410/m3-14

Lyons, C. N., and White, T. C. (2000). Transcriptional analyses of antifungal drug resistance in Candida albicans. *Antimicrob. Agents Chemother.* 44, 2296–2303. doi: 10.1128/aac.44.9.2296-2303.2000

Maertens, J. A. (2004). History of the development of azole derivatives. *Clin. Microbiol. Infect.* 10, 1–10. doi: 10.1111/j.1470-9465.2004.00841.x

Martel, R. R., Klicius, J., and Galet, S. (1977). Inhibition of the immune response by rapamycin, a new antifungal antibiotic. *Can. J. Physiol. Pharmacol.* 55, 48–51. doi: 10.1139/y77-007

Maschmeyer, G., Haas, A., and Cornely, O. A. (2007). Invasive aspergillosis: epidemiology, diagnosis and management in immunocompromised patients. *Drugs* 67, 1567–1601. doi: 10.2165/00003495-200767110-00004

Mednick, A. J., Nosanchuk, J. D., and Casadevall, A. (2005). Melanization of *Cryptococcus neoformans* affects lung inflammatory responses during cryptococcal infection. *Infect. Immun.* 73, 2012. doi: 10.1128/IAI.73.4.2012-2019.2005

Meier-Kriesche, H. U., Li, S., Gruessner, R. W., Fung, J. J., Bustami, R. T., Barr, M. L., et al. (2006). Immunosuppression: evolution in practice and trends, 1994-2004. *Am. J. Transplant.* 6(5 Pt. 2), 1111–1131. doi: 10.1111/j.1600-6143.2006.01270.x

Monroy-Pérez, E., Paniagua-Contreras, G. L., Rodríguez-Purata, P., Vaca-Paniagua, F., Vázquez-Villaseñor, M., Díaz-Velásquez, C., et al. (2016). High virulence and antifungal resistance in clinical strains of *Candida albicans. Can J Infect Dis Med* 2016:5930489. doi: 10.1155/2016/5930489

Nambu, M., Covel, J. A., Kapoor, M., Li, X., Moloney, M. K., Numa, M. M., et al. (2017). A calcineurin antifungal strategy with analogs of FK506. *Bioorg. Med. Chem. Lett.* 27, 2465–2471. doi: 10.1016/j.bmcl.2017.04.004

Neville, B. A., d'Enfert, C., and Bougnoux, M.-E. (2015). Candida albicans commensalism in the gastrointestinal tract. *FEMS Yeast Res.* 15:fov081. doi: 10.1093/femsyr/fov081

Nguyen, D. V., Roret, T., Fernandez-Gonzalez, A., Kohler, A., Morel-Rouhier, M., Gelhaye, E., et al. (2020). Target Of Rapamycin pathway in the white-rot fungus Phanerochaete chrysosporium. *PLoS One* 15:e0224776. doi: 10.1371/journal.pone.0224776

Odom, A., Del Poeta, M., Perfect, J., and Heitman, J. (1997a). The immunosuppressant FK506 and its nonimmunosuppressive analog L-685,818 are toxic to Cryptococcus neoformans by inhibition of a common target protein. *Antimicrob. Agents Chemother.* 41, 156–161.

Odom, A., Muir, S., Lim, E., Toffaletti, D. L., Perfect, J., and Heitman, J. (1997b). Calcineurin is required for virulence of *Cryptococcus neoformans. EMBO J.* 16, 2576–2589. doi: 10.1093/emboj/16.10.2576

Person, A. K., Kontoyiannis, D. P., and Alexander, B. D. (2010). Fungal infections in transplant and oncology patients. *Infect. Dis. Clin. North Am.* 24, 439–459. doi: 10.1016/j.idc.2010.01.002

Petrikkos, G., Skiada, A., Lortholary, O., Roilides, E., Walsh, T. J., and Kontoyiannis, D. P. (2012). Epidemiology and clinical manifestations of mucormycosis. *Clin. Infect. Dis.* 54(Suppl. 1), S23–S34. doi: 10.1093/cid/cir866

Pfaller, M. A., and Diekema, D. J. (2004). Rare and emerging opportunistic fungal pathogens: concern for resistance beyond Candida albicans and Aspergillus fumigatus. *J. Clin. Microbiol.* 42, 4419–4431. doi: 10.1128/jcm.42.10.4419-4431.2004

Pfaller, M. A., and Diekema, D. J. (2007). Epidemiology of Invasive Candidiasis: a Persistent Public Health Problem. *Clin. Microbiol. Rev.* 20:133. doi: 10.1128/CMR.00029-06

Pierce, C. G., Srinivasan, A., Uppuluri, P., Ramasubramanian, A. K., and López-Ribot, J. L. (2013). Antifungal therapy with an emphasis on biofilms. *Curr. Opin. Pharmacol* 13, 726–730. doi: 10.1016/j.coph.2013.08.008

Rotonda, J., Burbaum, J. J., Chan, H. K., Marcy, A. I., and Becker, J. W. (1993). Improved calcineurin inhibition by yeast FKBP12-drug complexes. Crystallographic and functional analysis. *J. Biol. Chem.* 268, 7607–7609.

Rusnak, F., and Mertz, P. (2000). Calcineurin: form and function. *Physiol. Rev.* 80, 1483–1521. doi: 10.1152/physrev.2000.80.4.1483

Saag, M. S., Graybill, R. J., Larsen, R. A., Pappas, P. G., Perfect, J. R., Powderly, W. G., et al. (2000). Practice guidelines for the management of cryptococcal disease. *Clin. Infect. Dis* 30, 710–718. doi: 10.1086/313757

Sanglard, D., Ischer, F., Marchetti, O., Entenza, J., and Bille, J. (2003). Calcineurin A of Candida albicans: involvement in antifungal tolerance, cell morphogenesis and virulence. *Mol. Microbiol.* 48, 959–976. doi: 10.1046/j.1365-2958.2003.03495.x

Saunders, R. N., Metcalfe, M. S., and Nicholson, M. L. (2001). Rapamycin in transplantation: A review of the evidence. *Kidney Int.* 59, 3–16. doi: 10.1046/j.1523-1755.2001.00460.x

Schreiber, S. L., and Crabtree, G. R. (1992). The mechanism of action of cyclosporin A and FK506. *Immunol. Today* 13, 136–142. doi: 10.1016/0167-5699(92)90111-J

Sehgal, S. N., Baker, H., and Vézina, C. (1975). Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. *J. Antibiot.* 28, 727–732. doi: 10.7164/antibiotics.28.727

Shekhar-Guturja, T., Gunaherath, G. M. K. B., Wijeratne, E. M. K., Lambert, J.-P., Averette, A. F., Lee, S. C., et al. (2016). Dual action antifungal small molecule modulates multidrug efflux and TOR signaling. *Nat. Chem. Biol.* 12, 867–875. doi: 10.1038/nchembio.2165

Shinde, P. B., Ban, Y. H., Hwang, J.-Y., Cho, Y., Chen, Y.-A., Cheong, E., et al. (2015). A non-immunosuppressive FK506 analogue with neuroregenerative activity produced from a genetically engineered Streptomyces strain. *RSC Adv.* 5, 6823–6828. doi: 10.1039/C4RA11907J

Smith, K. D., Achan, B., Hullsiek, K. H., McDonald, T. R., Okagaki, L. H., Alhadab, A. A., et al. (2015). Increased antifungal drug resistance in clinical isolates of *Cryptococcus neoformans* in Uganda. *Antimicrob. Agents Chemother.* 59, 7197–7204. doi: 10.1128/aac.01299-15

So, Y. S., Lee, D. G., Idnurm, A., Ianiri, G., and Bahn, Y. S. (2019). The TOR Pathway plays pleiotropic roles in growth and stress responses of the fungal pathogen *Cryptococcus neoformans*. *Genetics* 212, 1241–1258. doi: 10.1534/genetics.119.302191

Steinbach, W. J., Cramer, R. A. Jr., Perfect, B. Z., Asfaw, Y. G., Sauer, T. C., Laura, K., et al. (2006). Calcineurin controls growth, morphology, and pathogenicity in *Aspergillus fumigatus*. *Eukaryot. Cell* 5, 1091–1103. doi: 10.1128/ec.00139-06

Steinbach, W. J., Reedy, J. L., Cramer, R. A., Perfect, J. R., and Heitman, J. (2007). Harnessing calcineurin as a novel anti-infective agent against invasive fungal infections. *Nat. Rev. Microbiol.* 5, 418–430. doi: 10.1038/nrmicro1680

Steinbach, W. J., Schell, W. A., Blankenship, J. R., Onyewu, C., Heitman, J., and Perfect, J. R. (2004). In vitro interactions between antifungals and immunosuppressants against *Aspergillus fumigatus*. *Antimicrob. Agents Chemother.* 48, 1664–1669. doi: 10.1128/aac.48.5.1664-1669.2004

Stelzner, A. (1990). F. *C. Odds, Candida and Candidosis, A Review and Bibliography (Second Edition).* X + 468 S., 97 Abb., 92 Tab. u. 22 Farbtafeln. London—Philadelphia—Toronto—Sydney—Tokyo 1988. Baillière Tindall (W. B. Saunders). £ 35.00. ISBN: 0–7020–1265–3. *J. Basic Microbiol.* 30, 382–383. doi: 10.1002/jobm.3620300522

Thomson, A. W. (1989). FK506–how much potential? *Immunol. Today* 10, 6–9. doi: 10.1016/0167-5699(89)90057-1

Tonthat, N. K., Juvvadi, P. R., Zhang, H., Lee, S. C., Venters, R., Spicer, L., et al. (2016). Structures of pathogenic fungal FKBP12s reveal possible self-catalysis function. *mBio* 7:e00492-16. doi: 10.1128/mBio.00492-16

Vellanki, S., Billmyre, R. B., Lorenzen, A., Campbell, M., Turner, B., Huh, E. Y., et al. (2020). A novel resistance pathway for calcineurin inhibitors in the human-pathogenic mucorales mucor circinelloides. *mBio* 11, e002949-19. doi: 10.1128/mBio.02949-19

Vellanki, S., Navarro-Mendoza, M. I., Garcia, A., Murcia, L., Perez-Arques, C., Garre, V., et al. (2018). Mucor circinelloides: growth, Maintenance, and Genetic Manipulation. *Curr. Protoc. Microbiol.* 49:e53. doi: 10.1002/cpmc.53

Vila, T., Romo, J. A., Pierce, C. G., McHardy, S. F., Saville, S. P., and Lopez-Ribot, J. L. (2017). Targeting Candida albicans filamentation for antifungal drug development. *Virulence* 8, 150–158. doi: 10.1080/21505594.2016.1197444

Walsh, T. J., Anaissie, E. J., Denning, D. W., Herbrecht, R., Kontoyiannis, D. P., Marr, K. A., et al. (2008). Treatment of Aspergillosis: clinical practice guidelines of the Infectious Diseases Society of America. *Clin. Infect. Dis.* 46, 327–360. doi: 10.1086/525258

Wang, R. J., Miller, R. F., and Huang, L. (2017). Approach to fungal infections in human immunodeficiency virus-infected individuals: pneumocystis and beyond. *Clin. Chest. Med.* 38, 465–477. doi: 10.1016/j.ccm.2017.04.008

Whaley, S. G., Berkow, E. L., Rybak, J. M., Nishimoto, A. T., Barker, K. S., and Rogers, P. D. (2016). Azole antifungal resistance in Candida albicans and emerging non-albicans Candida Species. *Front. Microbiol.* 7:2173. doi: 10.3389/fmicb.2016.02173

Wiederhold, N. P. (2017). Antifungal resistance: current trends and future strategies to combat. *Infect. Drug Resist.* 10, 249–259. doi: 10.2147/idr.S124918

Wu, C.-J., Liu, W.-L., Lai, C.-C., Chao, C.-M., Ko, W.-C., Wang, H.-C., et al. (2020). Multicenter Study of Azole-Resistant *Aspergillus fumigatus* Clinical Isolates, Taiwan(1). *Emerg. Infect. Dis.* 26, 804–806. doi: 10.3201/eid2604.190840

Yoon, S., Kim, J. H., Koh, Y., Tran, P. T., Ann, J., Yoon, I., et al. (2017). Discovery of simplified leucyladenylate sulfamates as novel leucyl-tRNA synthetase (LRS)-targeted mammalian target of rapamycin complex 1 (mTORC1) inhibitors. *Bioorg. Med. Chem.* 25, 4145–4152. doi: 10.1016/j.bmc.2017.06.002

Zarakas, M. A., Desai, J. V., Chamilos, G., and Lionakis, M. S. (2019). Fungal infections with ibrutinib and other small-molecule kinase inhibitors. *Curr. Fungal Infect. Rep.* 13, 86–98. doi: 10.1007/s12281-019-00343-9

Non-catalytic-Region Mutations Conferring Transition of Class A β-Lactamases Into ESBLs

Thinh-Phat Cao[1,2†], Hyojeong Yi[3†], Immanuel Dhanasingh[1], Suparna Ghosh[1], Jin Myung Choi[1], Kun Ho Lee[2,4], Seol Ryu[5], Heenam Stanley Kim[3]* and Sung Haeng Lee[1,2]*

[1] *Department of Cellular and Molecular Medicine, Chosun University School of Medicine, Gwangju, South Korea,*
[2] *Department of Biomedical Sciences, Gwangju Alzheimer's Disease and Related Dementia Cohort Research Center, College of Natural Sciences and Public Health and Safety, Chosun University, Gwangju, South Korea,* [3] *Division of Biosystems & Biomedical Sciences, College of Health Sciences, Korea University, Seoul, South Korea,* [4] *Aging Neuroscience Research Group, Korea Brain Research Institute, Daegu, South Korea,* [5] *Department of Chemistry, Chosun University, Gwangju, South Korea*

Despite class A ESBLs carrying substitutions outside catalytic regions, such as Cys69Tyr or Asn136Asp, have emerged as new clinical threats, the molecular mechanisms underlying their acquired antibiotics-hydrolytic activity remains unclear. We discovered that this non-catalytic-region (NCR) mutations induce significant dislocation of β3-β4 strands, conformational changes in critical residues associated with ligand binding to the lid domain, dynamic fluctuation of Ω-loop and β3-β4 elements. Such structural changes increase catalytic regions' flexibility, enlarge active site, and thereby accommodate third-generation cephalosporin antibiotics, ceftazidime (CAZ). Notably, the electrostatic property around the oxyanion hole of Cys69Tyr ESBL is significantly changed, resulting in possible additional stabilization of the acyl-enzyme intermediate. Interestingly, the NCR mutations are as effective for antibiotic resistance by altering the structure and dynamics in regions mediating substrate recognition and binding as single amino-acid substitutions in the catalytic region of the canonical ESBLs. We believe that our findings are crucial in developing successful therapeutic strategies against diverse class A ESBLs, including the new NCR-ESBLs.

Keywords: extended-spectrum β-lactamase, non-catalytic-region ESBL, ceftazidime, antibiotic resistance, X-ray crystallography

INTRODUCTION

Extended-spectrum β-lactamases (ESBLs) are a serious threat to human health due to their enhanced hydrolytic activity against third-generation cephalosporins such as ceftazidime (CAZ), a representative first-line drug for bacterial diseases (Dance, 2014). In general, class A β-lactamases, regardless of subfamilies, share the similar overall structural architecture of their catalytic region (**Supplementary Figure S1**), which can be separated into two major compartments: the conserved catalytic ensemble and the variable recognition ensemble. The conserved catalytic ensemble involving the hydrolysis of substrate consists of the reactive Ser70 which attacks the β-lactam amide bond; two general bases Lys73 and Glu166 which electrostatically activate the nucleophilic Ser70; the typical oxyanion hole formed by N atoms of Ser70 and Thr237 stabilizing the negative

transient acyl-intermediate; the catalytic (or hydrolytic) water coordinated to Glu166 and Asn170, which disrupts the acyl-bond (Page, 2008; Drawz and Bonomo, 2010). On the other hand, the recognition ensemble, consisting of three critical segments including Ω-loop (160–180 in the Ambler system) (Sirot et al., 1997; Celenza et al., 2008; Dobson et al., 2012; Levitt et al., 2012), lid (92–118) (Doucet et al., 2004; Bethel et al., 2006), and strands β3-β4 (230–251) (Du Bois et al., 1995; Giakkoupi et al., 2001; Shimizu-Ibuka et al., 2011; Ruggiero et al., 2017), is mostly responsible for the adaptability of class A β-lactamases toward different types of antibiotics (**Supplementary Figure S1**). Upon the innovation of drugs, the residues in this ensemble are further varied to accommodate novel substrates, by which canonical catalytic-region ESBLs are to be produced. These three segments surrounding reactive Ser70 in the active site are known to mediate substrate recognition and hydrolysis directly, and these ESBLs originate in hot spot mutations by single amino acid substitutions in one of the segments of the wild-type β-lactamases[1]. The mutations extend the substrate specificity of the enzyme by increasing the flexibility of the Ω-loop (Wang et al., 2002b), the charged state of the lid (Petit et al., 1995), and the interaction of the β3-β4 strands with substrates (Huletsky et al., 1993; Palzkill, 2018), indicating that the functionality and substrate specificity of the class A β-lactamases largely depend on the subtle changes in the size of the binding cleft.

Compared with these canonical ESBLs, the non-catalytic-region ESBLs (NCR-ESBLs) carry single amino-acid substitutions outside of the catalytic segments and potentially hydrolyze CAZ (Gniadkowski, 2001), while conserving residues for substrate binding (recognition ensemble) and hydrolysis (catalytic ensemble) virtually intact (Perez et al., 2007). To date, only two NCR mutations have been identified at the residue 69 (Chaibi et al., 1998; Sam et al., 2009; Dobson et al., 2012) in several class A β-lactamases, including PenL and at 136 (according to the Ambler system) in PenL alone (Dobson et al., 2012), which seem unlikely to involve substrate binding and hydrolysis (**Figure 1A**). The ESBLs carrying mutations on the position of 69 have been reported from several class A β-lactamases including PenI of *Burkholderia pseudomallei* (Sam et al., 2009), SHV (Giakkoupi et al., 1998; Helfand et al., 2002; Monica A. Totir et al., 2006) and TEM (Samy O. Meroueh et al., 2002; Wang et al., 2002b), wherein the substitutions of Cys69 into Tyr or Met69 into Val, Ile, Leu, Tyr, Phe, or Lys conferred resistance to large size CAZ. Although several structural and kinetics studies suggested that such substitutions on Met69 may disrupt, or perturb the oxyanion hole of the active site (Helfand et al., 2002; Wang et al., 2002b; Monica A. Totir et al., 2006), those studies were unable to address the effect of substitution from Cys69 into bulky Tyr or Phe in case of Pen-type β-lactamase in a molecular level. Another NCR-ESBL carrying a mutation on the position of 136(Asn136Asp) has been reported only in PenL, which displays higher MIC values of CAZ than PenL-WT for *B. thailandensis* (Dobson et al., 2012). PenL (previously called, PenA) from *B. thailandensis* is a class A β-lactamase, which has been extensively studied with regard to its transition into an extended-spectrum β-lactamase (ESBL). This enzyme can evolve *via* a simple nucleotide-substitution, deletion, or duplication mutation to an ESBL, which can hydrolyze third-generation cephalosporins, including ceftazidime. Asn136 locating on helix α4 likely stabilizes Ω-loop by forming hydrogen bonds with the backbone of Glu166 (Yi et al., 2016). Note that Glu166 is one of the critical catalytic residues on Ω-loop for substrate binding in many of class A β-lactamases (Drawz and Bonomo, 2010), thereby forming an energetically unfavorable non-proline *cis*-peptide Glu166 and Xaa167 (Thr167 in PenL, Xaa refer any amino acid on the position as Ambler system). Asn136 is thus crucial for the proper functional orientation of Glu166 proven by the loss of function by substitution of Asn to Ala in the position of 136 of TEM-1 (Banerjee et al., 1998). Therefore, the substitution of Asn136 into Asp in PenL likely disrupts the Ω-loop's stability, which may affect the substrate specificity of the newly emerged NCR-ESBL. The molecular mechanism of resistance against CAZ by 136 NCR mutation, however, has not been investigated.

In this study, we investigated the detailed molecular mechanisms underlying the acquired activity of the two NCR-ESBLs (PenL-Cys69Tyr and PenL-Asn136Asp) from *Burkholderia thailandensis* against CAZ, which do not directly involve with substrate binding and hydrolysis (Chaibi et al., 1998; Giakkoupi et al., 1998; Dobson et al., 2012; **Figure 1A**). Our results demonstrate that NCR associated mutations consequently induce the subtle rearrangements on three critical catalytic segments(ensemble) of the PenL, including distortion of the strands β3-β4 and alteration of the electrostatic potentials around the canonical oxyanion hole (Cys69Tyr), enhancement of flexibility of the Ω-loop and increased fluctuation of strands β3-β4 (Asn136Asp). As a result, the substrate-binding cleft of the enzymes is enlarged to accommodate large CAZ.

MATERIALS AND METHODS

Expression and Purification

Genes encoding for PenL-Cys69Tyr and PenL-Asn136Asp were isolated as described previously (Dobson et al., 2012) and sub-cloned into pET28a(+) expression vector. *E. coli* BL21(DE3)-competent cell strain was used to overexpress PenL-WT, PenL-Cys69Tyr, and PenL-Asn136Asp. Transformed cells were grown in Luria-Bertani (LB) media supplemented with 100 μg/mL of kanamycin at 37°C until the OD$_{600}$ reached ∼ 0.6 and induced by 0.5 mM of isopropyl-β-D thiogalactopyranoside (IPTG). After an additional 16 h of incubation at 18°C, cells were centrifuged at 5,000 × *g* and 4°C for 20 min. Cell pellets were resuspended in a buffer containing 50 mM Tris-HCl, pH 7.5, 500 mM NaCl and 10 mM imidazole, supplemented with 0.1 mM phenylmethane sulfonyl fluoride (PMSF), 1 mM dithiothreitol (DTT), and DNase I. The resuspension cocktails were disrupted by high-intensity sonication at 4°C, and the insoluble fractions were separated using high-speed centrifugation (20,000 × *g* at 4°C for 30 min). The soluble fractions, containing the desired proteins with N-terminal 6 × His-tag, were loaded onto Ni-agarose columns. Unbound proteins were washed out with excess

[1]https://www.lahey.org/Studies/

FIGURE 1 | NCR-ESBL-associated mutations and their kinetic analysis against CAZ. **(A)** The overall structure of PenL-WT containing the positions of the NCR-ESBL related mutations (Cys69Tyr and Asn136Asp) with three essential segments (lid, Ω-loop, and β3-β4) and CAZ hydrolysis-associated residues (black dots). **(B)** Kinetics of CAZ hydrolysis. The rate of CAZ hydrolysis by the two PenL-NCR-ESBLs plateaued, indicating the lower K_M value, increased catalytic efficiency K_{cat}/K_M (increased by 3–5-fold, respectively) and enhanced CAZ affinity. Data are averages of triplicate determination in three or four independent experiments from different preparations and presented with standard deviation. The lines are fits of the Michaelis-Menten equation to the data. **(C)** Chemical structure of ceftazidime (CAZ) and ceftazidime-like glycyl boronic acid (CBA). Negatively charged groups harboring oxyimino moiety are marked with red boxes, and asterisks indicate the acetamido hydroxyl groups (see **Supplementary Movie S1**).

buffer containing 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, and 20 mM imidazole. PenLs were eluted using the above buffer containing 250 mM imidazole. To remove the N-terminal 6 × His-tag, the eluted fractions were pulled and dialyzed against the buffer containing 20 mM Tris-HCl, 150 mM NaCl, and 2 mM CaCl₂, pH 7.5, followed by treatment with 10 U of human α-thrombin (HTI, United States) per 1 mg/mL protein. PenLs were further purified *via* size-exclusion chromatography using a HiLoad 16/60 Superdex 200 pg column (GE Healthcare, United States) saturated with buffer (20 mM Tris-HCl, pH 7.5, and 50 mM NaCl).

Crystallization and Structural Analysis

The purified PenL-Cys69Tyr and PenL-Asn136Asp were concentrated to 10 μg/μl using a 10,000 Da cut-off Vivaspin centrifugal concentrator (Sartorius). Crystallization was carried out using the hanging drop vapor diffusion method by mixing a 1.2 μl protein sample with a 1.2 μl reservoir solution. Initial screenings were set up using commercial crystallization kits obtained from Hampton (USA) and Rigaku (Japan). Crystals of PenL-Cys69Tyr were grown in a solution containing 200 mM sodium acetate trihydrate and 20% (w/v) polyethylene glycol 3,350 at 4°C. By contrast, crystals of PenL-Asn136Asp were grown in a solution consisting of 100 mM sodium acetate at pH 5.0, 200 mM sodium chloride, and 25% (w/v) polyethylene glycol 3,350 at 20°C. Crystals were transferred to the cryo-solutions comprising a growth solution supplemented with 15–20% glycerol for 30 s, followed by flash-freezing *via* immersion in liquid nitrogen.

To determine the complex structure of PenL-Cys69Tyr and PenL-Asn136Asp with ceftazidime-like glycylboronate (CBA), the cryo-solutions were supplemented with 2 mM CBA, followed

by soaking crystals overnight at 4°C before flash-freezing. All the crystals were diffracted at a maximum of ~1.3 Å resolution. X-ray diffraction and data collection were performed at Pohang Light Source (PLS) beamline 5C (Pohang, South Korea) using the ADSC Q315r CCD detector. Collected data were indexed, integrated, and scaled using HKL2000 (HKL Research Inc.). The structure of the two PenL variants was determined *via* molecular replacement using MolRep (Abergel, 2013) and the structure of PenL-WT (PDB code 5GL9) (Yi et al., 2016) as a reference model. The refinement was carried out with Refmac5 (Murshudov et al., 1997) and phenix.refine (Afonine et al., 2012; Echols et al., 2014). The coordinates and restraints of CBA were generated by eLBOW, and manually fitted to m*Fo-Fc* map employing Coot (Emsley and Cowtan, 2004). Details of data diffraction and structural refinement are shown in **Supplementary Table S1**.

Determination of Kinetic Parameters

Because of its poor spectroscopic property, ceftazidime (CAZ) can only be measured at a maximum of ~100 μM. Therefore, appropriate amounts of purified enzymes were mixed with various CAZ concentrations ranging from 5 to 100 μM in a reaction buffer comprising 50 mM potassium phosphate at pH 7.0 supplemented with 20 μg/mL of bovine serum albumin. The absorbance at 260 nm was immediately monitored at 25°C using a cuvette holder temperature controller. The initial velocity of PenL-WT was measured during the first 10 s using a standard procedure, and the velocities (v) were fitted to the Michaelis-Menten equation (**Supplementary Information**). The first-order persistence of reaction by PenL-WT prevented the calculation of K_M and k_{cat} within the experimental setup range; however, the catalytic efficiency can be estimated by reciprocal plotting, in which the slope of the regression curve is k_{cat}/K_M. Data fittings

were carried out using in-house Python scripts with the power of SciPy (Oliphant, 2007) library for parameter estimation, and Matplotlib (Hunter, 2007) module for data visualization.

Molecular Dynamics Simulation

PenL-WT, PenL-Cys69Tyr, and PenL-Asn136Asp were subjected to all-atom molecular dynamics (MD) simulations in explicit solvent using Gromacs 5.0.7 suite (Lemkul et al., 2015) and Gromos96-43a1 force field (Ramos et al., 2019). The coordinate of PenL-WT was obtained from the RCSB Protein Data Bank by fetching the ID 5GL9 (Yi et al., 2016). Molecules were solvated using the TIP3P water model and neutralized with a cubic boundary of 0.10 M of NaCl. Particle mesh Ewald method (Abraham and Gready, 2011) was used to determine the electrostatic interaction of systems, with real space and a *van der* Waals distance cut-off of 10 Å. After energy minimization by 500-step in steepest descent calculation, systems were heated to 300 K and simulated for 50 *ns*. The MD trajectory analysis was performed directly in the Gromacs package, and the RMSD and radius of gyration plotted with 100 frame intervals (**Supplementary Figure S8**).

RESULTS

Biochemical Properties of CAZ Hydrolysis by PenL Wild-Type and NCR-ESBLs

To elucidate the mechanism associated with CAZ hydrolysis by the two PenL NCR-ESBLs (PenL-Cys69Tyr and PenL-Asn136Asp), we first determined their kinetic parameters (**Table 1**). The two PenL NCR-ESBLs and wild-type PenL (PenL-WT) hydrolyzed CAZ over time in which the CAZ decay was linearly related to time for 10 s (**Supplementary Figure S2**). **Figure 1B** illustrates the comparative kinetics of CAZ hydrolysis by the two ESBLs and WT of PenL. The kinetics of CAZ hydrolysis were similar between the two ESBLs, with V_{max} and K_M values of 156.4 nmol/min per mg and 22.3 µM, respectively, for PenL-Cys69Tyr, and 155.8 nmol/min per mg and 12.5 µM, respectively, for PenL-Asn136Asp. The mutant's K_M values are 30–60 times lower than that of WT (**Table 1**), indicating that a stable acyl-enzyme complex may be formed after binding in the mutant ESBLs. Although the k_{cat} values of the two PenL-ESBLs were lower than that of PenL-WT, their catalytic efficiency (k_{cat}/K_M) was enhanced 3–5-fold compared with that of PenL-WT (**Table 1**, **Figure 1B**, and see **Supplementary Information**). Those biochemical results are likely to correspond to the higher

MIC values of CAZ than PenL-WT for *B. thailandensis* (Dobson et al., 2012). Overall, the changes in kinetic property, primarily with the decreased K_M, in the two PenL ESBL variants indicate that the two single substitutions led to alternative substrate recognition for the enzyme. Interestingly, the single mutations at non-canonical regions are involved neither in substrate binding or catalysis.

Next, to examine whether the enhanced CAZ affinity and catalytic efficiency were related to changes in the conformation of PenL-Cys69Tyr and PenL-Asn136Asp, Circular dichroism (CD) spectroscopy was used for comparison with that of wild-type (**Supplementary Information**). The CD spectra of their respective apo-forms (PenL-Cys69Tyr-apo and PenL-Asn136Asp-apo) showed no significant differences compared with the PenL-WT (**Supplementary Figure S3A**). However, the CD spectra of the two NCR-ESBLs in the presence of a non-hydrolyzable CAZ analog glycylboronate (CBA) slightly deviated from their apo-forms as well as the PenL-WT-apo form. The degrees of the CD spectral change of the mutants appeared similar to that of WT from CAZ presence. Moreover, the deviations in a spectral difference between the two ESBLs seemed even less than those with CBA, which might result from the rapid hydrolysis of CAZ accompanied by the slight conformational changes (**Supplementary Figure S3B**). These observations suggested that the conformation of the two NCR-ESBLs was similar to that of PenL-WT; however, the NCR-ESBLs may recognize the substrate CAZ (or CBA) by marginally changing their conformation of the side chain of residues in the active site or recognition ensemble.

Crystal Structures of PenL-Cys69Tyr NCR-ESBLs

To further investigate the conformational changes based on substrate recognition, we delineated four crystal structures of the two PenL NCR-ESBLs: two involving the apo-form and the other two associated with the CBA-bound form (**Supplementary Table S1** and **Supplementary Figure S4**). Consistent with CD spectra, PenL-Cys69Tyr-apo and PenL-Asn136Asp-apo were structurally similar to PenL-WT (PDB ID:5GL9) (Yi et al., 2016) with root-mean-square deviations (RMSDs) of 0.221 and 0.186 Å at their Cα atoms relative to the wild-type, respectively (**Figure 2A**). In particular, the configurations of the conserved catalytic residues and the catalytic water (W4) associated with hydrolysis coincided with those of PenL-WT (**Figures 3B**, **2B** and **Supplementary Figure S5**). These similarities suggest that the respective substitution does not structurally alter the active site involved in acylation/diacylation of the β-lactam

TABLE 1 | Kinetic parameters of ceftazidime hydrolysis.

Variant	K_M (µM)	V_{max} (nmol/min/mg)	k_{cat} (s^{-1})	k_{cat}/K_M (nM^{-1} s^{-1})
PenL-WT*	643.9 ± 1.07*	1657.0 ± 384.63*	2.0 ± 0.460*	3.0*
PenL-Cys69Tyr	22.3 ± 0.98	156.4 ± 2.51	0.2 ± 0.003	8.4
PenL-Asn136Asp	12.5 ± 0.54	155.8 ± 1.67	0.2 ± 0.002	14.9

*Data are means ± SD of triplicate determination in three or four independent experiments. *Estimated from linear regression.*
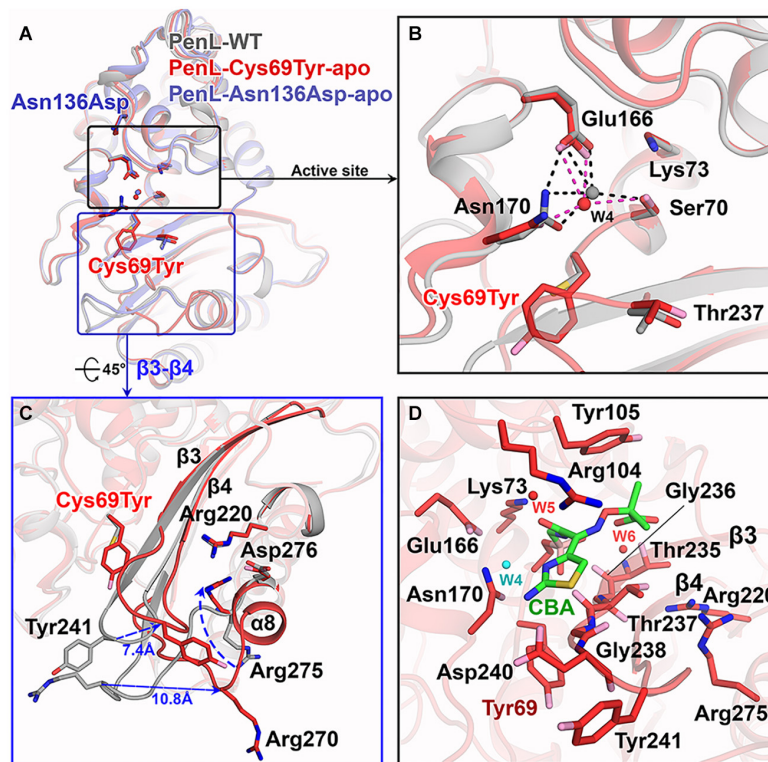
FIGURE 2 | Structure of NCR-ESBLs and Cys69Tyr mutation-induced β3-β4 changes. **(A)** Superposition of structures of the PenL NCR-ESBLs with PenL-WT. The structural similarity over the entire molecule is shown. Typically, the configuration of side chains for catalytic residues is well preserved among the proteins. **(B)** Comparison of active site between WT and Cys69Tyr variant. The catalytic residues from both proteins coordinate the catalytic waters (W4), which coincide in the active site. **(C)** Structural changes in β3-β4 of PenL-Cys69Tyr. Replacement of bulky Tyr at 69 increases the flexibility of τηε β3-β4 loop (∼7.4 Å at Tyr241) by disassembly of β4. The change induces the dislocation of a loop between β5 and α8 (∼10.8 Å at Arg270) and flip Arg275 on the β5-α8 loop toward the substrate-binding region, resulting in expansion of the active site and altered electronic properties of the oxyanion hole. **(D)** The active site structure of PenL-Cys69Tyr-CBA. Residues on b3 involved in the interaction with CBA, and Tyr105 also move upward for facilitating the entry of large substrate (see also **Supplementary Figure S6A**). Arg275 keeps its conformation projecting into the oxyanion hole, and Tyr241 moves back to that of WT to some extent.

backbone during hydrolysis (Lamotte-Brasseur et al., 1991; Page, 2008; Brown et al., 2009; Drawz and Bonomo, 2010; Papp-Wallace et al., 2013).

Interestingly, the PenL-Cys69Tyr-apo showed structural variation in the β3-β4 strands compared with PenL-WT-apo (**Figure 2C**). Notably, the C-terminus of β3 and the N-terminus of β4 in PenL-Cys69Tyr-apo were disrupted in three (Thr237 to Asp240) and two (Thr243 to Gly244) residues, respectively. The disruption in the β-strands was apparently induced by the steric hindrance of bulky Tyr69 against the β3-β4 sheet, resulting in enhanced flexibility and dislocation of the loop by 7.4 Å at Tyr241 (**Figure 2C**). The increased flexibility of the β4 N-terminus in PenL-Cys69Tyr-apo induced a loss of hydrogen bonds with adjacent C-terminus of β5 and subsequently dislocated the β5-α8 loop away from the corresponding position of Pen-WT-apo by 10.8 Å at Arg270 (**Figure 2C** and **Supplementary Movie S1**). Since the β3-β4 segment is one of the recognition ensembles, these changes likely expand the active site of the enzyme to accommodate large substrates such as CAZ. The structure of PenL-Cys69Tyr-CBA shows that the residues ranging from Thr237 to Asp240 of β3 mediated substrate recognition, and

the loop interacts with the aminothiazole ring and acetamido backbone, rather than the oxyimino group of CBA (**Figure 2D**). These results suggest that the NCR mutation induced similar effects as in the known canonical ESBLs. For instance, the dislocation of the loop between β3-β4 strands in TEM-52 ESBL (Gly238Ser on β3) enlarged the active site to bind CAZ without further loop extension and with minimal dislocation (∼2.9 Å) (Papp-Wallace et al., 2012). In addition to structural analysis, we calculated the electrostatic distribution on the surface of proteins (**Figure 3**). Surprisingly, we found a significant change in electrostatic property from negative to positive in the active site of PenL-Cys69Tyr-apo, which has never been reported in ESBLs. Upon distortion of strands β3-β4, the helix α8 was extended by four residues (Ala272-Arg275) (**Figure 2C** and **Supplementary Movie S1**). Accordingly, Arg275 rotated 180° to relocate closer to the Thr237 with a distance of ∼6.7 Å (between N atom of Arg275 and backbone N of Thr237) in the Arg220-Asp276-Thr237 cluster, which is critical for substrate binding and hydrolysis (Papp-Wallace et al., 2010, 2012; **Figures 2C,D**, **3C**). The Thr237 residue in the cluster mainly contributed to the oxyanion hole, stabilizing the tetrahedral intermediate formed
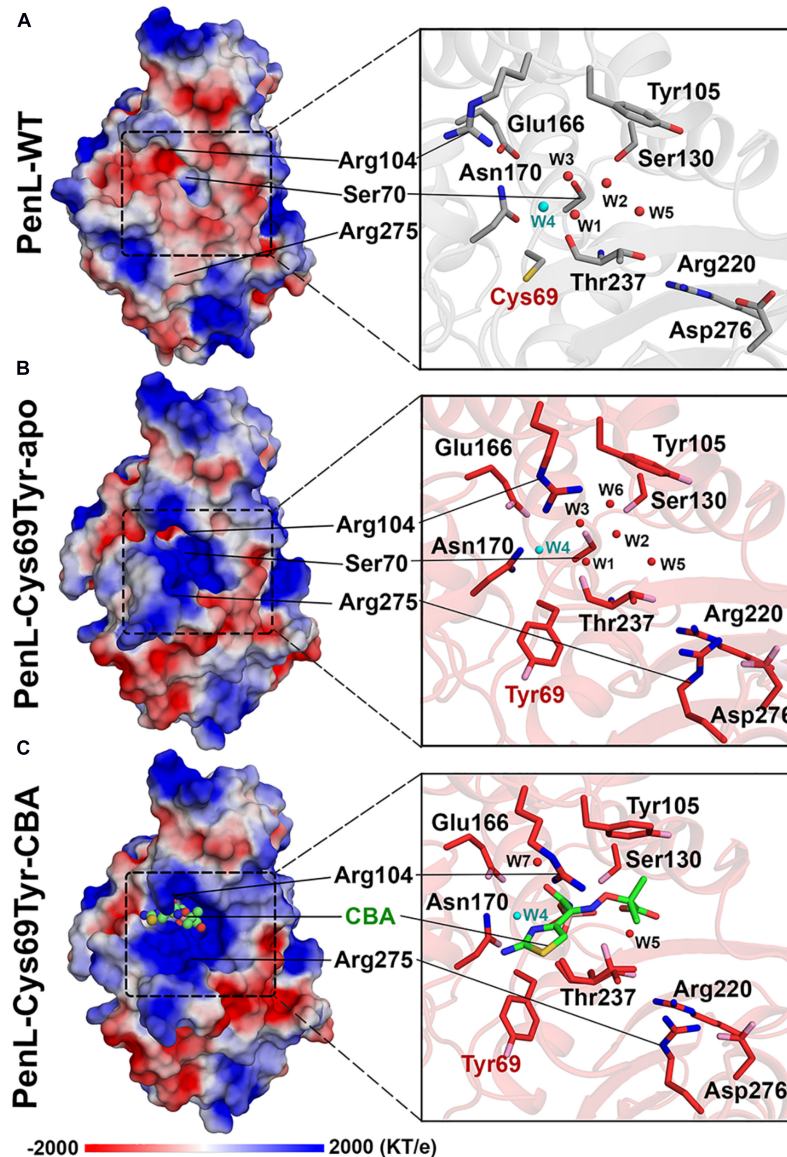
FIGURE 3 | Electrostatic potential changes in the catalytic cavity of PenL-Cys69Tyr. **(A)** Electrostatic property of the active site in PenL-WT. **(B,C)** Electrostatic distribution in the substrate-binding cleft of PenL-Cys69Tyr-apo and CBA-bound form (acylated). Residues such as Tyr69, Se70, Arg104, Glu166, Asn170, Arg275, and W4 catalytic water molecule are actively involved in substrate binding, acylation-deacylation step, and oxyanion hole formation (see **Supplementary Movie S2**).

via nucleophilic attack by catalytic Ser70 (**Figures 2C**, **3** and **Supplementary Figure S5**). Note that, based on the established mechanism of β-lactam hydrolysis (Drawz and Bonomo, 2010; Papp-Wallace et al., 2013), the N atom of Ambler residue 237 (i.e., Thr237 in this case) participates in the formation of the oxyanion hole, stabilizing the tetrahedral intermediate of a substrate with class A β-lactamase. Besides the residue 237 in the cluster, Arg220, whose configuration was not changed between PenL-WT and PenL-Cys69Tyr, is directly in close contact with Thr237 (~4.4 Å) (**Figures 2C,D**). Remarkably, the Arg275 in the Cys69Tyr mutant positions adjacent to Arg220 within ~3.6 Å, which may strengthen the positive charge around the oxyanion hole

(**Supplementary Figure S9C**). Consequently, the accumulation of additional basic amino acid (i.e., Arg275) increased the positive electrostatic potential around the oxyanion hole (**Figure 3** and **Supplementary Movie S2**). Notably, due to the presence of two negatively charged groups on CAZ (**Figure 1C**), the increased net positive charge in the PenL-Cys69Tyr active site may enhance the binding affinity with CAZ, which is consistent with the kinetic data (**Table 1**). Taken together, the findings indicate that the Cys69Tyr mutation in class A β-lactamase affects the structural interaction between β3–β4 strands and substrate by enlarging the substrate-binding cleft of the enzyme and altering electrostatic property around the oxyanion hole in the active

site. These changes in the NCR-ESBL may result in enhanced recognition of large CAZ without affecting the conserved catalytic residues of class A β-lactamases located in the active site for hydrolysis.

Crystal Structure of PenL-Asn136Asp NCR-ESBL

Similar to Cys69, the residue Asn136 is positioned outside the essential catalytic segments and not involved in the catalysis of antibiotics. However, a novel ESBL variant carrying the substitution of Asn136Asp was found only in PenL (Dobson et al., 2012) and also exhibited an increase in CAZ hydrolysis activity (**Table 1** and **Figure 1B**). Asn136 is located proximally to active site cleft so that its side chain forms a hydrogen bond to the backbone of Glu166 on Ω-loop and stabilizes the energetically unfavorable non-proline *cis*-peptide (Berg et al., 2012) between Glu166-Thr167 (**Figure 4A**). Notice that Glu166 is a critical catalytic residue that involved in both the activation of the Ser70 and deacylation step in β-lactam hydrolysis (Lamotte-Brasseur et al., 1991; Guillaume et al., 1997; Meroueh et al., 2005; Page, 2008; Drawz and Bonomo, 2010; Papp-Wallace et al., 2013). Asn136 thus appears to be responsible for the proper orientation of Glu166 and Ω-loop. Evidently, the loss of the stabilization by Asn136 (i.e., by mutation Asn136Ala) resulted in the functional deficiency of a class A β-lactamase (Banerjee et al., 1997, 1998). The replacement of Asn136 to aspartate abolishes a hydrogen bond formed with Glu166 in WT (**Figure 4A**), which probably leads to an increase in Ω-loop flexibility. To our surprise, the structure of PenL-Asn136Asp demonstrates no significant difference in comparison with PenL-WT (**Figure 4A**). Nonetheless, PenL-Asn136Asp can accommodate the CBA into its active site wherein the conformation does not change substantially (**Figure 4C** and **Supplementary Figure S6**). Residue Arg275 also adopts the same configuration as that of the wild-type enzyme, implying no such alteration of electrostatic distribution around active site cleft, as seen in PenL-Cys69Tyr (**Figure 4B**). However, the large-sized CAZ may bind to the active site with the help of the change in lid segment, although no such considerable distinction between PenL-Asn136Asp and PenL-WT, as well as between PenL-Asn136Asp-apo and PenL-Asn136Asp-CBA, is observed. Two residues, including Arg104 and Tyr105, in the lid of PenL-Asn136Asp-apo, appeared to move away from the active site than WT (**Figures 4B,C**). In PenL-Asn136Asp-apo, residue Arg104 swung away by ~4.0 Å from the corresponding position of WT, whereas Tyr105 moved upward by ~1.3 Å. Then, Arg104 returned toward the active site when the CBA bound. Instead, the Tyr105 moved further upward by another 2.0 Å in the CBA bound form (**Figure 4**, **Supplementary Figure S7**, and **Supplementary Movie S3**). These observations indicate that Asn136Asp mutation may affect the conformation of the lid segment to expand the size of the substrate-binding cleft and receive large size CAZ. Interestingly, the changes in the lid resemble that the effects of mutation on Ω-loop propagated into those residues to enlarge and accommodate CAZ (Yi et al., 2016). Therefore, the mutation at 136 may cause the instability of the Ω-loop.

An interesting question can then be raised here on how the relatively negative CAZ can be attracted to an electrostatically negative substrate-binding cleft of the PenL-Asn136Asp (**Supplementary Figure S6**) like that of PenL-WT. Taken together with the CD spectra data and kinetic analysis, these observations indicate that the binding of CAZ into the binding site might be governed by latent factors other than the electrostatic attraction, which was not be revealed in the static crystal structure. We thereby suggest two hypotheses. First, the higher degree of freedom would be induced to the Ω-loop of PenL-Asn136Asp due to the loss of one hydrogen bond by the mutation that occurred outside the Ω-loop. Indeed, this effect resembles cases in other class A β-lactamase ESBLs where the stable network on Ω-loop *per se* was devastated by the substitutions at Ambler position Arg164 or Asp179 located on Ω-loop (Orencia et al., 2001; Wang et al., 2002a). In this case, the resulting enhanced flexible motion of Ω-loop may induce the transient enlargement of active site cleft and facilitate the accommodation of third-generation cephalosporins like CAZ or CTX. Second, the mutation Asn136Asp may create intrinsic dynamic conformers that would efficiently accommodate CAZ. The viewpoint of functional promiscuity in protein conformation has been suggested formerly (Tokuriki and Tawfik, 2009), whereby poorly packed or disordered conformation by an accumulation of single mutations in proteins evolves conformationally diverse structures to adapting novel substrates.

Dynamics in Catalytic Regions of NCR-ESBLs

For the reasons, the molecular dynamics (MD) simulation was therefore conducted to compare the dynamic property of PenL-Asn136Asp with PenL-WT and PenL-Cys69Tyr on the CAZ binding (**Figure 5** and **Supplementary Figure S8**). The MD simulation indicates that the overall structure of PenL-Asn136Asp was stable through a 50 ns trajectory and roughly similar to PenL-WT. However, RMSD at three critical segments of PenL-Asn136Asp, similar to PenL-Cys69Tyr, was higher than PenL-WT and appeared to fluctuate, especially at the lid region (**Figure 5** and **Supplementary Figure 8**). Furthermore, the radius of gyration (Rgyr) plots also demonstrate a potential unfolding of two PenL-ESBLs, in contrast to the sustainable motion of PenL-WT (**Figure 5**). These observations strongly suggest that PenL-Asn136Asp (and PenL-Cys69Tyr) has a higher tendency for disorder than PenL-WT, which is correlated with the higher adaptability toward CAZ in term of *protein dynamism* (Tokuriki and Tawfik, 2009). In the molecular view, two regions appeared to involve in the accommodation of the large-size antibiotics followed by the transient disorder. First, the MD results showed the serial propagation of the mutation effect from Ω-loop into the β3-β4 element. The conformations of primary residues, including Asn136, Glu166, reactive Ser170, and the three essential catalytic segments involved in binding, remained unchanged in PenL-WT through MD trajectory (**Figure 5A**). By contrast, the mutant Asp136 side chain in PenL-Asn136Asp-apo shifted away from the Glu166 backbone of Ω-loop, leading to a large fluctuation of Glu166 (approximately 5.2 Å) and Ser70 (**Figures 5B,C**,
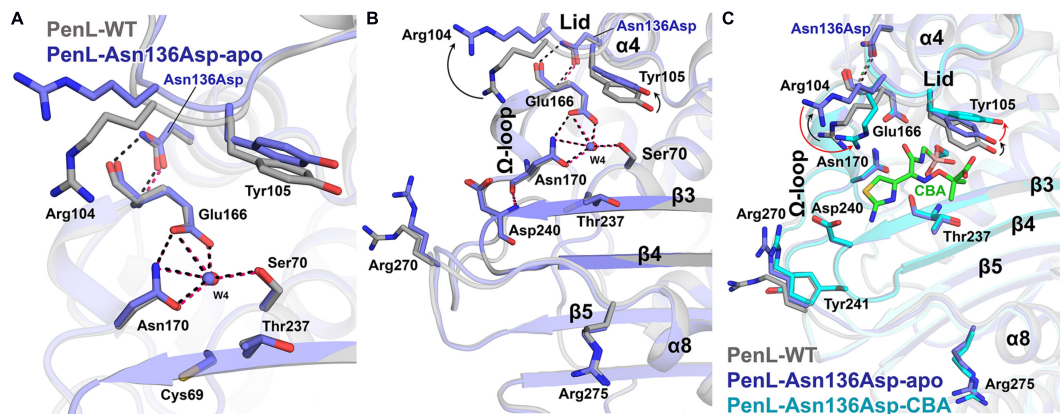
FIGURE 4 | Structure of Asn136Asp NCR-ESBL and changes on lid. **(A)** Superposition of structures of the PenL-Asn136Asp-apo with WT. The side chain configurations of catalytic residues are superimposed regardless of the mutation. **(B)** Conformational change on the lid segment. Asp104 and Tyr105 in apo-form were displaced from substrate-binding cleft compared to those in WT, resulting in the expansion of the active site's size. **(C)** Structure of PenL-Asn136Asp-CBA. The two residues, including Arg104 and Tyr105, underwent subsequent structural changes upon CBA binding. Moving Arg104 toward CBA (red arrow) and lifting Tyr105 upward further from the active site (red arrow) are likely to provide a large substrate like CAZ with adequate space. However, the β3-β4 and residues that underwent conformational changes in the Cys69Tyr variant remained unchanged (see **Supplementary Movie S3** and compare **Figure 2**).
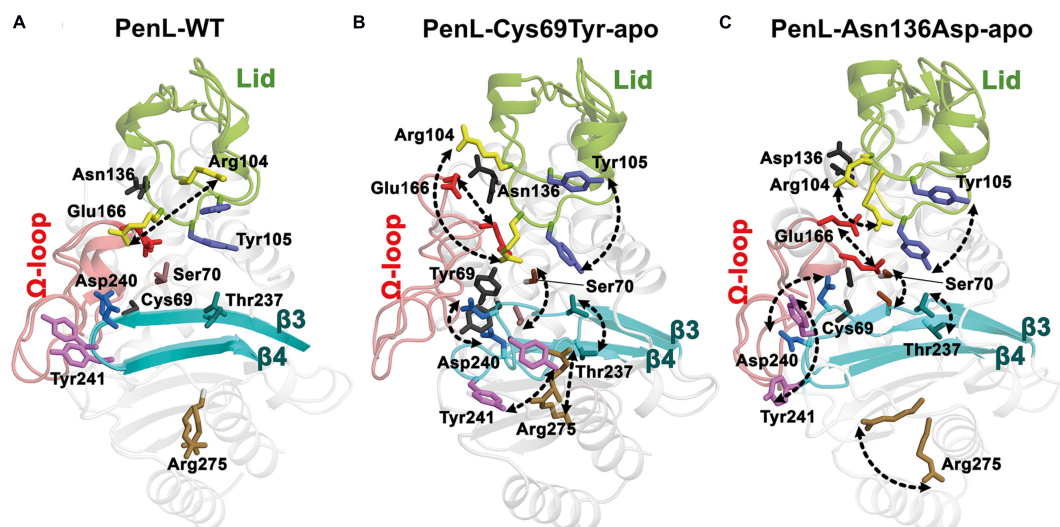

FIGURE 5 | Dynamics of catalytic regions in PenL-WT and NCR-ESBLs. The dynamics of the representative residues are shown from initial to expanded state with arrows. **(A)** Dynamics of PenL-WT. Most of the residues are static during simulation except for Arg104 that fluctuates horizontally. **(B)** PenL-Cys69Try and **(C)** PenL-Asn136Asp during the 50 ns MD simulation. The dynamics of substrate binding residues and the major catalytic regions, including Ω-loop, β3-β4 strands, and lid in the NCR-ESBLs, enlarge the opening of the substrate-binding active site (see **Supplementary Movie S4**).

Supplementary Figures S8B,C and **Supplementary Movie S4**). Surprisingly, the high degree of fluctuation involving the Ω-loop and Ser70 pushed the β3–β4 loop away *via* steric hindrance with a concurrent breakdown of the hydrogen bond between Asn170 and Asp240, located on the Ω-loop and β3, respectively (**Figures 4B**, **5C** and **Supplementary Movie S4**). The changes in the regions, which are intensively involved in substrate recognition, may result in the momentary expansion of the active site space and the improvement of CAZ binding, likewise observed from PenL-Cys69Tyr and Ω-loop tandem repeat ESBL of PenL (Yi et al., 2016). Second, the Arg104 residue in

PenL-Asn136Asp-apo penetrated the substrate-binding region compared with the Pen-WT, to facilitate the recognition and interaction with acetamido backbone of CBA through hydrogen bonding (**Supplementary Figure S6B**). Interestingly, the conformation of Arg104 in PenL-Asn136Asp-CBA was similar to those of PenL-Cys69Tyr-apo and -CBA, but further away than PenL-Asn136Asp-apo, suggesting that the entry of Arg104 into the substrate-binding cleft stabilized the acyl-form of PenL-Cys69Tyr *via* interaction with W5 in CBA (**Figure 5C** and **Supplementary Figure S6**). The increased affinity of PenL-Asn136Asp to CAZ also supported this result compared with

PenL-WT (**Table 1**). Besides, Tyr105 in PenL-Asn136Asp-CBA, a well-conserved residue for ligand interaction (Papp-Wallace et al., 2013), was dislocated further way than the PenL-WT and PenL-Asn136Asp-apo (**Figures 4C, 5C**). Indeed, the side-chain conformation of two residues in PenL-Asn136Asp dislocated further away from the substrate-binding site in crystal structures with respect to those of PenL-Cys69-Tyr and WT (**Figure 4** and **Supplementary Figure S7**). Therefore, the changes involving the lid region may induce momentary expansion of the active site and improve CAZ binding in the PenL-Asn136Asp and PenL-Cys69Tyr as in the canonical ESBLs (Patel et al., 2017).

DISCUSSION

We demonstrated that the effects of the two NCR mutations in ESBLs ultimately converged to the recognition ensemble that is one of the catalytic regions. The finding suggests a common evolutionary mechanism underlying the expansion of the substrate spectrum in β-lactamases, i.e., by widening the catalytic cleft for large-size antibiotics with high affinity *via* changes in local segments regardless of the location of mutations. Further, the electrostatic environment surrounding the oxyanion hole in PenL-Cys69Tyr can be modified by mutations enhancing substrate recognition. Although the Arg275 dislocation in PenL-Cys69Tyr accounts for the enhanced positive charge distribution around oxyanion hole, the Arg's substantial contribution to the charge transition is still in question due to the quite remote distance (∼6.7 Å) from Thr237 on β3-β4 (**Supplementary Figure S9A**) and the dynamic fluctuations (**Figure 5** and **Supplementary Figure S8C**). Of the MD simulation frames, several configurations regarding the Arg275 and Arg220 showed an augmented proximity to the active site with ∼4.0 Å and ∼4.4 Å, respectively (**Supplementary Figures S9B,C**). Furthermore, the disrupted regions (loop) on β3-β4 and β5-α8 (**Figure 2C**) by the Cys69Tyr mutation moves toward the active site as Arg275 approaches to Thr237, indicating the supply of N atoms nearer to the oxyanion hole (**Supplementary Figure S8C**). Therefore, the above arrangement of the basic residues and atoms in PenL-Cys69Tyr may induce a strong positive charge in the region around its oxyanion hole, which subsequently attracts the negative moiety-containing substrate, CAZ. The typical oxyanion hole, which is comprised of N atoms of reactive Ser70 and Thr237, and water W6 (**Figure 2** and **Supplementary Figures S5, S6**), is responsible for sustaining the carbonyl group of β-lactam backbone once the β-lactam is attacked by nucleophilic Ser70. Likewise, the tetrahedral intermediate of the substrate in both CBA-bound complex structures of PenL NCR-ESBLs is undoubtedly stabilized by the typical oxyanion hole as well studied in class A β -lactamases.

In addition to Arg275, Arg104 also contributed to the oxyanion hole (**Figures 3, 4**). For example, Arg104 present in the lid of PenL-Cys69Tyr, protruding into the ligand-binding cleft of the enzyme together with Asn132 and W5, increased the positive charge of the oxyanion hole for further stabilization of the Michaelis complex (**Supplementary Figure S7**). Interestingly, the

mode of interaction of Arg104 in PenL-Cys69Tyr resembled the oxyanion holes in multistep enzymes, such as thiolase (Kursula et al., 2002), which contain a "second oxyanion hole" with basic amino acids and water molecules. PenL-Asn136Asp showed relatively less perturbation of β3-β4, β5-α8, and penetration of Arg104; however, it may have transiently altered the electric potential around the oxyanion hole according to the protein dynamism theory (Tokuriki and Tawfik, 2009; **Supplementary Figure S8**). Indeed, MD simulations showed that Arg104 in PenL-Asn136Asp fluctuated similarly as in the PenL-Cys69Tyr (**Figure 5**, **Supplementary Figure S8**, and **Supplementary Movie S4**), where the flexible movement of Arg104 could momentarily modulate the electrostatic potential in the potential second oxyanion hole around substrate-binding cleft.

To the best of our knowledge, however, this potential second oxyanion hole has yet to be identified in ESBL catalysis, despite its possible existence in other ESBLs (Shimamura et al., 2002; Chen et al., 2005). The existence of the potential second oxyanion hole in the ESBLs may be attributed to their kinetic behaviors (Merilainen et al., 2009), as well as increased affinity. In other words, the increased affinity of PenL-Cys69Tyr for CAZ may facilitate the acylation step but simultaneously attenuate the deacylation process due to the formation of a stable acyl-enzyme complex, in which Arg104 on the flexible loop could be discharged from the interaction with Asn132 and W6 and subsequently could disrupt the potential "second oxyanion hole" before releasing a product (**Figure 3** and **Supplementary Figure S6**). The precise hydrolytic mechanism of CAZ mediated by NCR-ESBLs, however, needs to be further investigated.

In summary, we have described the structural mechanism underlying the hydrolysis of third-generation cephalosporins by novel NCR-ESBLs. Although the mutations were unconventional in that they occurred outside of the catalytic region for hydrolysis and substrate binding, the class A NCR-ESBLs exhibited altered substrate specificity and carried an expanded active site similar to the canonical class A ESBLs. We believe that the novel substrate-spectrum expansion mechanism of the class A β-lactamases described in this study will significantly enhance the current knowledge of the evolutionary trends of ESBLs against new antibiotics. Furthermore, from a practical perspective, such information will be crucial for developing novel and useful inhibitors or antibiotics targeting various class A ESBLs, including emerging NCR-ESBLs.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

TC and HY performed the experiments and generated the data for this study with an equal contribution. ID, SG, and

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.598998/full#supplementary-material

REFERENCES

Abergel, C. (2013). Molecular replacement: tricks and treats. *Acta Crystallogr. D Biol. Crystallogr.* 69(Pt 11), 2167–2173. doi: 10.1107/S0907444913015291

Abraham, M. J., and Gready, J. E. (2011). Optimization of parameters for molecular dynamics simulation using smooth particle−mesh Ewald in GROMACS 4.5. *J. Comput. Chem.* 32, 2031–2040. doi: 10.1002/jcc.21773

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., et al. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68(Pt 4), 352–367. doi: 10.1107/S0907444912001308

Banerjee, S., Pieper, U., Kapadia, G., Pannell, L. K., and Herzberg, O. (1998). Role of the ¿-loop in the activity, substrate specificity, and structure of class A β-lactamase. *Biochemistry* 37, 3286–3296.

Banerjee, S., Shigematsu, N., Pannell, L. K., Ruvinov, S., Orban, J., Schwarz, F., et al. (1997). Probing the non-proline Cis peptide bond in β-Lactamase from *Staphylococcus aureus* PC1 by the Replacement Asn136 → Ala. *Biochemistry* 36, 10857–10866. doi: 10.1021/bi970352r

Berg, J. M., Tymoczko, J. L., and Stryer, L. (2012). *Biochemistry*. Kate Ahr Parker: W. H. Freeman and Company.

Bethel, C. R., Hujer, A. M., Hujer, K. M., Thomson, J. M., Ruszczycky, M. W., Anderson, V. E., et al. (2006). Role of Asp104 in the SHV -Lactamase. *Antimicrob. Agents Chemother.* 50, 4124–4131. doi: 10.1128/aac.00848-06

Brown, N. G., Shanker, S., Prasad, B. V., and Palzkill, T. (2009). Structural and biochemical evidence that a TEM-1 beta-lactamase N170G active site mutant acts via substrate-assisted catalysis. *J. Biol. Chem.* 284, 33703–33712. doi: 10.1074/jbc.M109.053819

Celenza, G., Luzi, C., Aschi, M., Segatore, B., Setacci, D., Pellegrini, C., et al. (2008). Natural D240G Toho-1 mutant conferring resistance to ceftazidime: biochemical characterization of CTX-M-43. *J. Antimicrob. Chemother.* 62, 991–997. doi: 10.1093/jac/dkn339

Chaibi, E. B., Peduzzi, J., Farzaneh, S., Barthelemy, M., Sirot, D., and Labia, R. (1998). Clinical inhibitor-resistant mutants of the beta-lactamase TEM-1 at amino-acid position 69. Kinetic analysis and molecular modelling. *Biochim. Biophys. Acta* 1382, 38–46. doi: 10.1016/s0167-4838(97)00127-1

Chen, Y., Shoichet, B., and Bonnet, R. (2005). Structure, function, and inhibition along the reaction coordinate of CTX-M beta-lactamases. *J. Am. Chem. Soc.* 127, 5423–5434. doi: 10.1021/ja042850a

Dance, D. (2014). Treatment and prophylaxis of melioidosis. *Int. J. Antimicrob. Agents* 43, 310–318. doi: 10.1016/j.ijantimicag.2014.01.005

Dobson, R., Yi, H., Cho, K.-H., Cho, Y. S., Kim, K., Nierman, W. C., et al. (2012). Twelve positions in a β-lactamase that can expand its substrate spectrum with a single amino acid substitution. *PLoS One* 7:e37585. doi: 10.1371/journal.pone.0037585

Doucet, N., De Wals, P.-Y., and Pelletier, J. N. (2004). Site-saturation mutagenesis of Tyr-105 reveals its importance in substrate stabilization and discrimination in TEM-1 β-Lactamase. *J. Biol. Chem.* 279, 46295–46303. doi: 10.1074/jbc.M407606200

Drawz, S. M., and Bonomo, R. A. (2010). Three decades of beta-lactamase inhibitors. *Clin. Microbiol. Rev.* 23, 160–201. doi: 10.1128/CMR.00037-09

Du Bois, S. K., Marriott, M. S., and Amyes, S. G. (1995). TEM- and SHV-derived extended-spectrum beta-lactamases: relationship between selection, structure and function. *J. Antimicrob. Chemother.* 35, 7–22. doi: 10.1093/jac/35.1.7

Echols, N., Moriarty, N. W., Klei, H. E., Afonine, P. V., Bunkoczi, G., Headd, J. J., et al. (2014). Automating crystallographic structure solution and refinement of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* 70(Pt 1), 144–154. doi: 10.1107/S139900471302748X

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 12), 2126–2132. doi: 10.1107/S0907444904019158

Giakkoupi, P., Hujer, A. M., Miriagou, V., Tzelepi, E., Bonomo, R. A., and Tzouvelekis, L. S. (2001). Substitution of Thr for Ala-237 in TEM-17, TEM-12 and TEM-26: alterations in β-lactam resistance conferred on *Escherichia coli*. *FEMS Microbiol. Lett.* 201, 37–40. doi: 10.1016/s0378-1097(01)00239-7

Giakkoupi, P., Miriagou, V., Gazouli, M., Tzelepi, E., Legakis, N. J., and Tzouvelekis, L. S. (1998). Properties of mutant SHV-5 beta-lactamases constructed by substitution of isoleucine or valine for methionine at position 69. *Antimicrob. Agents Chemother.* 42, 1281–1283. doi: 10.1128/aac.42.5.1281

Gniadkowski, M. (2001). Evolution and epidemiology of extended-spectrum beta-lactamases (ESBLs) and ESBL-producing microorganisms. *Clin. Microbiol. Infect.* 7, 597–608. doi: 10.1046/j.1198-743x.2001.00330.x

Guillaume, G., Vanhove, M., Lamotte-Brasseur, J., Ledent, P., Jamin, M., Joris, B., et al. (1997). Site-directed mutagenesis of glutamate 166 in Two b-Lactamases. *J. Biol. Chem.* 272, 5438–5444. doi: 10.1074/jbc.272.9.5438

Helfand, M. S., Hujer, A. M., Sönnichsen, F. D., and Bonomo, R. A. (2002). Unexpected advanced generation cephalosporinase activity of the M69F variant of SHV β-Lactamase. *J. Biol. Chem.* 277, 47719–47723. doi: 10.1074/jbc.M207271200

Huletsky, A., Knox, J. R., and Levesque, R. C. (1993). Role of Ser-238 and Lys-240 in the hydrolysis of third-generation cephalosporins by SHV-type beta-lactamases probed by site-directed mutagenesis and three-dimensional modeling. *J. Biol. Chem.* 268, 3690–3697.

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90—-95.

Kursula, P., Ojala, J., Lambeir, A. M., and Wierenga, R. K. (2002). The catalytic cycle of biosynthetic thiolase: a conformational journey of an acetyl group through four binding modes and two oxyanion holes. *Biochemistry* 41, 15543–15556. doi: 10.1021/bi0266232

Lamotte-Brasseur, J., Dive, G., Dideberg, O., Charlier, P., Frere, J. M., and Ghuysen, J. M. (1991). Mechanism of acyl transfer by the class A serine beta-lactamase of *Streptomyces albus* G. *Biochem. J.* 279(Pt 1), 213–221. doi: 10.1042/bj2790213

Lemkul, J. A., Roux, B., van der Spoel, D., and MacKerell, A. D. Jr. (2015). Implementation of extended Lagrangian dynamics in GROMACS for polarizable simulations using the classical Drude oscillator model. *J. Comput. Chem.* 36, 1473–1479. doi: 10.1002/jcc.23937

Levitt, P. S., Papp-Wallace, K. M., Taracila, M. A., Hujer, A. M., Winkler, M. L., Smith, K. M., et al. (2012). Exploring the role of a conserved class a residue in the Ω-Loop of KPC-2 β-Lactamase. *J. Biol. Chem.* 287, 31783–31793. doi: 10.1074/jbc.M112.348540

Merilainen, G., Poikela, V., Kursula, P., and Wierenga, R. K. (2009). The thiolase reaction mechanism: the importance of Asn316 and His348 for stabilizing the enolate intermediate of the Claisen condensation. *Biochemistry* 48, 11011–11025. doi: 10.1021/bi901069h

Meroueh, S. O., Fisher, J. F., Schlegel, H. B., and Mobashery, S. (2005). Ab Initio QM/MM study of class A b-Lactamase acylation: dual participation of Glu166 and Lys73 in a concerted base promotion of Ser70. *J. Am. Chem. Soc.* 127, 15397–15407. doi: 10.1021/ja051592u

Meroueh, S. O., Roblin, P., Golemi, D., Maveyraud, L., Vakulenko, S. B., Zhang, Y., et al. (2002). Molecular dynamics at the root of expansion of function in the M69L inhibitor-resistant TEM b-lactamase from *Escherichia coli*. *J. Am. Chem. Soc.* 124, 9422–9430. doi: 10.1021/ja026547q

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53(Pt 3), 240–255. doi: 10.1107/S0907444996012255

Oliphant, T. E. (2007). Python for scientific computing. *Comput. Sci. Eng.* 9, 10–20. doi: 10.1109/Mcse.2007.58

Orencia, M. C., Yoon, J. S., Ness, J. E., Stemmer, W. P., and Stevens, R. C. (2001). Predicting the emergence of antibiotic resistance by directed evolution and structural analysis. *Nat. Struct. Biol.* 8, 238–242. doi: 10.1038/84981

Page, M. G. P. (2008). Extended-spectrum β-lactamases: structure and kinetic mechanism. *Clin. Microbiol. Infect.* 14, 63–74. doi: 10.1111/j.1469-0691.2007.01863.x

Palzkill, T. (2018). Structural and mechanistic basis for extended-spectrum drug-resistance mutations in altering the specificity of TEM, CTX-M, and KPC beta-lactamases. *Front. Mol. Biosci.* 5:16. doi: 10.3389/fmolb.2018.00016

Papp-Wallace, K. M., Taracila, M., Hornick, J. M., Hujer, A. M., Hujer, K. M., Distler, A. M., et al. (2010). Substrate selectivity and a novel role in inhibitor discrimination by residue 237 in the KPC-2 beta-lactamase. *Antimicrob. Agents Chemother.* 54, 2867–2877. doi: 10.1128/AAC.00197-10

Papp-Wallace, K. M., Taracila, M. A., Gatta, J. A., Ohuchi, N., Bonomo, R. A., and Nukaga, M. (2013). Insights into beta-lactamases from Burkholderia species, two phylogenetically related yet distinct resistance determinants. *J. Biol. Chem.* 288, 19090–19102. doi: 10.1074/jbc.M113.458315

Papp-Wallace, K. M., Taracila, M. A., Smith, K. M., Xu, Y., and Bonomo, R. A. (2012). Understanding the molecular determinants of substrate and inhibitor specificities in the Carbapenemase KPC-2: exploring the roles of Arg220 and Glu276. *Antimicrob. Agents Chemother.* 56, 4428–4438. doi: 10.1128/AAC.05769-11

Patel, M. P., Hu, L., Stojanoski, V., Sankaran, B., Prasad, B. V. V., and Palzkill, T. (2017). The drug-resistant variant P167S expands the substrate profile of CTX-M beta-Lactamases for Oxyimino-Cephalosporin Antibiotics by Enlarging the Active Site upon Acylation. *Biochemistry* 56, 3443–3453. doi: 10.1021/acs.biochem.7b00176

Perez, F., Endimiani, A., Hujer, K. M., and Bonomo, R. A. (2007). The continuing challenge of ESBLs. *Curr. Opin. Pharmacol.* 7, 459–469. doi: 10.1016/j.coph.2007.08.003

Petit, A., Maveyraud, L., Lenfant, F., Samama, J. P., Labia, R., and Masson, J. M. (1995). Multiple substitutions at position 104 of b-lactamase TEM-1: assessing the role of this residue in substrate specificity. *Biochem. J.* 305, 33–40. doi: 10.1042/bj3050033

Ramos, M. C., Horta, V. A. C., and Horta, B. A. C. (2019). Molecular dynamics simulations of PAMAM and PPI dendrimers using the GROMOS-Compatible 2016H66 Force Field. *J. Chem. Inform. Model.* 59, 1444–1457. doi: 10.1021/acs.jcim.8b00911

Ruggiero, M., Curto, L., Brunetti, F., Sauvage, E., Galleni, M., Power, P., et al. (2017). Impact of Mutations at Arg220 and Thr237 in PER-2 β-lactamase on conformation, activity, and susceptibility to inhibitors. *Antimicrob. Agents Chemother.* 61, e02193-16. doi: 10.1128/aac.02193-16

Sam, I. C., See, K. H., and Puthucheary, S. D. (2009). Variations in ceftazidime and amoxicillin-clavulanate susceptibilities within a clonal infection of *Burkholderia pseudomallei*. *J. Clin. Microbiol.* 47, 1556–1558. doi: 10.1128/JCM.01657-08

Shimamura, T., Ibuka, A., Fushinobu, S., Wakagi, T., Ishiguro, M., Ishii, Y., et al. (2002). Acyl-intermediate structures of the extended-spectrum class A beta-lactamase, Toho-1, in complex with cefotaxime, cephalothin, and benzylpenicillin. *J. Biol. Chem.* 277, 46601–46608. doi: 10.1074/jbc.M207884200

Shimizu-Ibuka, A., Oishi, M., Yamada, S., Ishii, Y., Mura, K., Sakai, H., et al. (2011). Roles of residues Cys69, Asn104, Phe160, Gly232, Ser237, and Asp240 in extended-spectrum β-Lactamase Toho-1. *Antimicrob. Agents Chemother.* 55, 284–290. doi: 10.1128/aac.00098-10

Sirot, D., Recule, C., Chaibi, E. B., Bret, L., Croize, J., Chanal-Claris, C., et al. (1997). A complex mutant of TEM-1 b-Lactamase with mutations encountered in both IRT-4 and extended-spectrum TEM-15 produced by an *Escherichia coli* clinical isolate. *Antimicrob. Agents Chemother.* 41, 1322–1325. doi: 10.1128/aac.41.6.1322

Tokuriki, N., and Tawfik, D. S. (2009). Protein dynamism and evolvability. *Science* 324, 203–207. doi: 10.1126/science.1169375

Totir, M. A., Padayatti, P. S., Helfand, M. S., Carey, M. P., Bonomo, R. A., Carey, P. R., et al. (2006). Effect of the inhibitor-resistant M69V substitution on the structures and populations of trans-enamine beta-lactamase intermediates. *Biochemistry* 45, 11895–11904. doi: 10.1021/bi060990m

Wang, X., Minasov, G., and Shoichet, B. K. (2002a). Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J. Mol. Biol.* 320, 85–95. doi: 10.1016/S0022-2836(02)00400-X

Wang, X., Minasov, G., and Shoichet, B. K. (2002b). The structural bases of antibiotic resistance in the clinically derived mutant β-Lactamases TEM-30, TEM-32, and TEM-34. *J. Biol. Chem.* 277, 32149–32156. doi: 10.1074/jbc.M204212200

Yi, H., Choi, J. M., Hwang, J., Prati, F., Cao, T.-P., Lee, S. H., et al. (2016). High adaptability of the omega loop underlies the substrate-spectrum-extension evolution of a class A β-lactamase, PenL. *Sci. Rep.* 6:36527. doi: 10.1038/srep36527

7,3′,4′-Trihydroxyisoflavone, a Metabolite of the Soy Isoflavone Daidzein, Suppresses α-Melanocyte-Stimulating Hormone-Induced Melanogenesis by Targeting Melanocortin 1 Receptor

Ji Hye Kim[1,2], Jae-Eun Lee[1], Taewon Kim[1], Myung Hun Yeom[3], Jun Seong Park[3], Eric di Luccio[4], Hanyong Chen[5], Zigang Dong[5], Ki Won Lee[6,7,8] and Nam Joo Kang[1]*

[1] School of Food Science and Biotechnology, Kyungpook National University, Daegu, South Korea, [2] Korean Medicine Application Center, Korea Institute of Oriental Medicine, Daegu, South Korea, [3] Amorepacific Corporation R&D Center, Skin Research Institute, Yongin, South Korea, [4] Department of Genetic Engineering, School of Life Sciences, College of Natural Sciences, Kyungpook National University, Daegu, South Korea, [5] The Hormel Institute, University of Minnesota, Austin, MN, United States, [6] World Class University Biomodulation Major, Department of Agricultural Biotechnology, Seoul National University, Seoul, South Korea, [7] Advanced Institutes of Convergence Technology, Seoul National University, Suwon, South Korea, [8] Research Institute of Bio Food Industry, Institute of Green Bio Science and Technology, Seoul National University, Pyeongchang, South Korea

7,3′,4′-Trihydroxyisoflavone (7,3′,4′-THIF) is a metabolite of daidzein which is a representative isoflavone found in soybean. Recent studies suggested that 7,3′,4′-THIF exerts a hypopigmentary effect in B16F10 cells, however, its underlying molecular mechanisms and specific target protein remain unknown. Here, we found that 7,3′,4′-THIF, but not daidzein, inhibited α-melanocyte-stimulating hormone (MSH)-induced intracellular and extracellular melanin production in B16F10 cells by directly targeting melanocortin 1 receptor (MC1R). Western blot data showed that 7,3′,4′-THIF inhibited α-MSH-induced tyrosinase, tyrosinase-related protein-1 (TYRP-1), and tyrosinase-related protein-2 (TYRP-2) expressions through the inhibition of Microphthalmia-associated transcription factor (MITF) expression and cAMP response element-binding (CREB) phosphorylation. 7,3′,4′-THIF also inhibited α-MSH-induced dephosphorylation of AKT and phosphorylation of p38 and cAMP-dependent protein kinase (PKA). cAMP and Pull-down assays indicated that 7,3′,4′-THIF strongly inhibited forskolin-induced intracellular cAMP production and bound MC1R directly by competing with α-MSH. Moreover, 7,3′,4′-THIF inhibited α-MSH-induced intracellular melanin production in human epidermal melanocytes (HEMs). Collectively, these results demonstrate that 7,3′,4′-THIF targets MC1R, resulting in the suppression of melanin production, suggesting a protective role for 7,3′,4′-THIF against melanogenesis.

Keywords: 7,3′,4′-Trihydroxyisoflavone, α-MSH, melanogenesis, depigmentation, melanocortin 1 receptor

INTRODUCTION

Melanin, synthesized in human melanocytes, plays an important role in protecting the skin from the harmful effects of ultraviolet (UV) radiation (Solano et al., 2006; Park et al., 2009). However, the accumulation of abnormal melanin can cause skin pigmentary disorders such as melasma, freckles and age spots and senile lentigines (Briganti et al., 2003).

The primary cause of hyperpigmentation is known to be exposure of the skin to UV radiation (Park et al., 2009). When exposed to UV radiation, keratinocytes secret α-melanocyte stimulating hormone (MSH). α-MSH stimulates melanocortin 1 receptor (MC1R) and subsequently induces intracellular cAMP production in melanocytes (Solano et al., 2006; Yamaguchi and Hearing, 2009; Gillbro and Olsson, 2011). Elevation of intracellular cAMP stimulates transcriptional factors such as microphthalmia-associated transcription factor (MITF) and cAMP response element-binding protein (CREB) via the cAMP-dependent protein kinase (PKA) and p38 MAPK signaling pathways (Busca and Ballotti, 2000; Smalley and Eisen, 2000; Tada et al., 2002; Saha et al., 2006; Solano et al., 2006; Yamaguchi and Hearing, 2009; Gillbro and Olsson, 2011). In contrast, inhibition of AKT signaling by cAMP blocks MITF degradation (Bellei et al., 2008; Jin et al., 2011). Consequently, the expression of melanogenic enzymes including tyrosinase and tyrosinase-related proteins (TYRPs) is upregulated, resulting in hyperpigmentation (Goding and Fisher, 1997; Sato et al., 1997; Yasumoto et al., 1997). Therefore, suppression of the MC1R signaling pathway may protect strategy against skin hyperpigmentation.

Epidemiologic studies have shown that the dietary consumption of soy may contribute to a reduced the risk of hyperpigmentation and several compounds found in soy have been studied in terms of their ability to promote skin health (Leyden and Wallo, 2011). Previous studies suggested that genistein induces cellular melanin synthesis and enhances tyrosinase activity (Yan et al., 1999). Whereas, daidzein isolated from *Maackia fauriei* inhibited tyrosinase activity, but the effect was weak (Kim et al., 2010). Unlike genistein, daidzein (**Figure 1A**) is converted to 7,3′,4′-trihydroxyisoflavone (7,3′,4′-THIF or 3′-hydroxydaidzein, **Figure 1B**) and other compounds in human liver microsomes (Kulling et al., 2001), which can lead to bioactivation. Recently, previous studies have reported that 7,3′,4′-THIF has a depigmenting effect on melanin production (Lee et al., 2006; Goh et al., 2011). Additionally, 7,3′,4′-THIF was shown to inhibit melanin production more effectively than genistein or daidzein in melan-a cells (Park et al., 2010). However, the underlying molecular mechanisms and specific target protein of 7,3′,4′-THIF remain unknown. Here, we report that 7,3′,4′-THIF attenuates α-MSH-induced melanogenesis by targeting MC1R.

MATERIALS AND METHODS

Chemicals

7,3′,4′-THIF was purchased from Indofine Chemical Co., Inc. (Hillsborough, NJ, United States). Dulbecco's modified Eagle's medium (DMEM), penicillin-streptomycin and 0.5% trypsin-EDTA were obtained from GIBCO§ Invitrogen (Auckland, NZ). Medium 254 and human melanocyte growth supplement (HMGS) purchased from Cascade Biologics (Portland, OR, United States). Fetal bovine serum (FBS), daidzein, α-MSH, arbutin, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), dimethyl sulfoxide (DMSO), gelatin, forskolin and β-actin antibody were purchased from Sigma Chemical Co. (St. Louis, MO, United States). The antibodies against tyrosinase, tyrosinase-related protein 1 (TYRP-1), tyrosinase-related protein 2 (TYRP-2), MITF, phosphorylated PKA α/β/γ (Thr-198), total PKA, MC1R, goat anti-mouse IgG-HRP, goat anti-rabbit IgG HRP, and donkey anti-goat IgG HRP-conjugated secondary antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, United States). Antibodies against phosphorylated Akt (Ser-473), phosphorylated mTOR (Ser-2448), phosphorylated GSK-3β (Ser-9), phosphorylated CREB (Ser-133), phosphorylated MKK3/6 (Ser-189/207), phosphorylated MSK1 (Ser-376), total PI3K, total Akt, total mTOR, total GSK-3β, total CREB, total MKK3/6, total MSK1, and total p38 were purchased from Cell Signaling Technology (Danvers, MA, United States). Antibodies against phosphorylated p38 (Tyr-180/Tyr-182) were purchased from BD Biosciences (San Jose, CA, United States). Skim milk was purchased from MB cell (LA, CA, United States). The cAMP immunoassay was purchased from Cayman (Ann Arbor, MI, United States). CNBr-Sepharose 4B and the chemiluminescence detection kit were purchased from Amersham Biosciences (Piscataway, NJ, United States). The protein assay kit was obtained from Bio-Rad Laboratories Inc. (Hercules, CA, United States). Recombinant MC1R was purchased from Millipore (Bedford, MA, United States).

Cell Culture

The murine melanoma cell line B16F10 was obtained from the Korean Cell Line Bank (Seoul, South Korea). B16F10 cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin at 37°C in a humidified atmosphere with 5% CO_2. Human epidermal melanocytes (HEMs) derived from moderately pigmented neonatal foreskins were purchased from Cascade Biologics. HEMs were cultured in Medium 254 supplemented with HMGS at 37°C in a humidified atmosphere with 5% CO_2.

Cell Viability

B16F10 cells (7×10^3) were cultured in a 96-well plate for 6 h. Cell culture media containing 7,3′,4′-THIF was added at final concentrations of 25, 50, and 100 μM, and then the cells were cultured for 96 h. HEMs (1×10^4) were cultured in a 96-well plate for 6 h. Cell culture media containing 7,3′,4′-THIF was added at final concentrations of 20, 40, and 80 μM, and then the cells were cultured for 72 h. MTT solution (5 mg/mL) were plused 20 μL/well and incubated 2 h. The media were removed and replaced 200 μL of dimethyl sulfoxide (DMSO) per well to dissolve the MTT formazan. After 2 h, absorbance was measured at 570 nm using a microplate reader (Sunrise-Basic Tecan; Grodig, Austria).
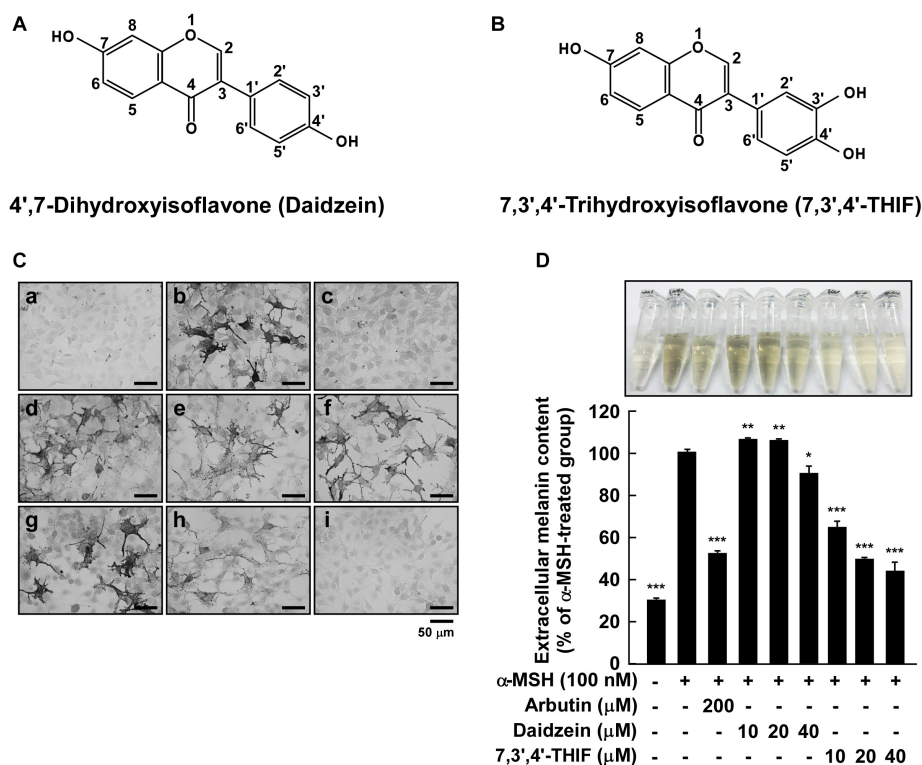
FIGURE 1 | Effect of 7,3′,4′-THIF or daidzein on the α-MSH-induced melanin content in B16F10 melanoma cells. **(A,B)** The chemical structures of daidzein **(A)** and 7,3′,4′-THIF **(B)**. **(C,D)** B16F10 cells were treated as described in "Materials and Methods," and the melanin contents were measured 3 days later. **(C)** Effects of 7,3′,4′-THIF or daidzein on α-MSH-induced intracellular melanin in untreated control cells (a) and in cells treated with α-MSH (b), α-MSH and 100 μM arbutin (c), α-MSH and 10 μM daidzein (d), α-MSH and 20 μM daidzein (e), α-MSH and 40 μM daidzein (f), α-MSH and 10 μM 7,3′,4′-THIF (g), α-MSH and 20 μM 7,3′,4′-THIF (h), and α-MSH and 40 μM 7,3′,4′-THIF (i). **(D)** Effects of 7,3′,4′-THIF or daidzein on the α-MSH-induced extracellular melanin level in the culture media of B16F10 melanoma cells. The cells were pretreated with samples at the indicated concentrations (10, 20, or 40 μM) for 1 h before being exposed to α-MSH (100 nM) for 3 days. The secreted melanin levels were determined as described in "Materials and Methods." Asterisks indicate a significant difference (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$) compared with α-MSH treated groups.

Fontana-Masson Staining

Intracellular melanin accumulation was visualized by Fontana-Masson staining with a slight modification (Joshi et al., 2007; An et al., 2010). Cells were fixed in 100% ethanol for 30 min at room temperature and stained for melanin using a Fontana-Masson staining kit from American Master*Tech Scientific, Inc. (Lodi, CA, United States), according to the manufacturer's instructions. In brief, cells were stained with ammoniacal silver for 60 min at 60°C, followed by incubation in 0.1% gold chloride and then in 5% sodium thiosulfate. Cell morphology and pigmentation were examined under a Nikon phase-contrast microscope (Tokyo, Japan). The images were analyzed using NIS-Elements 3.0 software.

Measurement of the Extracellular Melanin Content

The melanin content was measured using a slight modification of a previously reported method (Eisinger and Marko, 1982; Friedmann and Gilchrest, 1987; Gordon et al., 1989). Briefly, cells (8×10^3) were cultured in a six-well plate for 24 h. The culture media was replaced with the media containing daidzein or 7,3′,4′-THIF at the indicated concentrations (10, 20, or 40 μmol L^{-1}) for 1 h before being exposed to 100 nmol L^{-1} α-MSH and harvested 3 days later. After treatment, media were collected and the melanin levels therein were determined by measuring the absorbance at 405 nm using an ELISA reader.

Western Blotting

B16F10 cells (1.5×10^4) were cultured in a 6 cm dish for 48 h then starved in serum-free medium for an additional 24 h to eliminate the influence of FBS on kinase activation. HEMs (9.6×10^4) were cultured in 9 cm dish for 6 days. The culture media was replaced with the media containing 7,3′,4′-THIF (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH for 3 days. The harvested cells were disrupted and the supernatant fractions were boiled for 5 min. The protein concentration was determined using a dye-binding protein assay kit (Bio-Rad Laboratories Inc.), as described in the manufacturer's manual. Lysate protein (20–40 μg) was subjected to 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and electrophoretically transferred to a polyvinylidene fluoride membrane (Millipore). After blotting, the membrane was incubated with primary antibodies at 4°C overnight. After

hybridization with secondary antibodies, the protein bands were visualized using an ECL plus Western blotting detection system (Amersham™, Piscataway, NJ, United States). The relative intensities were quantified by Image J program.

cAMP Immunoassay

cAMP levels were measured using a cAMP immunoassay kit (Cayman). Briefly, B16F10 cells were treated with 7,3′,4′-THIF (10, 20, or 40 μM) for 1 h before being exposed to 1 μM forskolin for 30 min. Next, the cells were lysed in 0.1 M HCl to inhibit phosphodiesterase activity. The supernatants were then collected, neutralized, and diluted. After neutralization and dilution, a fixed amount of cAMP conjugate was added to compete with cAMP in the cell lysate for sites on rabbit polyclonal antibodies immobilized on a 96-well plate. After washing to remove excess conjugated and unbound cAMP, substrate solution was added to the wells to determine the activity of the bound enzyme. The color development was then stopped, after which the absorbance was read at 415 nm. The intensity of the color was inversely proportional to the cAMP concentration in the cell lysate.

In vitro and ex vivo Pull-Down Assays

Recombinant MC1R (25 μg) or a B16F10 cellular supernatant fraction (1,000 μg) was incubated with 7,3′,4′-THIF-Sepharose 4B beads (or Sepharose 4B alone as a control) (100 μL, 50% slurry) in reaction buffer [50 mM Tris (pH 7.5), 5 mM EDTA, 150 mM NaCl, 1 mM DTT, 0.01% Non-idet P-40, 2 μg/mL bovine serum albumin, 0.02 mM PMSF, and 1 × protease inhibitor mixture]. After incubation with gentle rocking overnight at 4°C, the beads were washed five times with buffer [50 mM Tris (pH 7.5), 5 mM EDTA, 150 mM NaCl, 1 mM DTT, 0.01% Non-idet P-40, and 0.02 mM PMSF], and the proteins bound to the beads were analyzed by immunoblotting.

α-MSH and 7,3′,4′-THIF Competition Assay

B16F10 cellular supernatant fraction (1,000 μg) was incubated with 100 μL of 7,3′,4′-THIF-Sepharose 4B beads or 100 μL of Sepharose 4B in a reaction buffer (see in vitro and ex vivo pull-down assays) for 24 h at 4°C, and α-MSH (0.1, 1, 10, or 100 μM) was added to a final volume of 500 μL and incubated for 24 h. The samples were washed, and proteins were then detected by western blotting.

Molecular Modeling and Energy Minimization

To investigate the detailed molecular basis of MC1R inhibition by 7,3′,4′-THIF, modeling study with 7,3′,4′-THIF and MC1R was performed twice. First, the sequence based search showed that MC1R has a similarity of 47% with a MC4R theoretical model (PDB entry: 2IQP) (Pogozheva et al., 2005), then the MC1R structure was built by using Prime v3.2 from Schrödinger suite 2013 (Jacobson et al., 2002, 2004) based this MC4R theoretical model, the loops of the MC1R homology structure were refined and minimized for docking studying. 7,3′,4′-THIF and daidzein were prepared under LigPrep with default parameter. The grid for docking was generated based on the binding sites that were predicted by SiteMap, finally, Flexible docking was performed in the extra precision (XP) mode. The number of poses per ligand was set to 10 in the post-docking minimization. Second, the sequence regions for the catalytic domains in human MC1R were identified using the PFAM profile database. Unfortunately, no experimental structure exists for the human MC1R. Therefore, models of the catalytic domains in human MC1R were built by homology modeling using closest templates available in the protein data bank (PDB). A BLAST search against the PDB identified the X-ray crystal structures 1CJK and 1AB8 with 67.8 and 38.8% sequence identity, respectively. After a careful multiple-sequence alignment with ClustalW V2.1, 100 models were generated with Modeller V9.10 (Sali and Blundell, 1993; Thompson et al., 1994). The modeled regions were 1-282 (MC1R). The best model, according to the intrinsic Modeller DOPE function, was chosen for docking studies (Krivov et al., 2009). Stereochemistry was assessed with PROCHECK (Laskowski et al., 1996). The stereochemistry of the best model was manually inspected with COOT (Emsley and Cowtan, 2004). Electrostatics calculations were performed with APBS V1.2.1; the molecular surfaces with the electrostatics properties (blue: positive charges; red: negative charges, with unit + 5/-5 kT/e) were rendered using VMD V1.8.9 and PyMOL V1.5 (Oberoi and Allewell, 1993; Holst et al., 1994; Wallace et al., 1995; Trott and Olson, 2010).

Docking

Prior to the docking calculation, the missing hydrogen atoms on the receptor were added using PDB2PQR in conjunction with the CHARMm force field (Oberoi and Allewell, 1993). AutoDock Vina was used to perform the dockings (Trott and Olson, 2010). Ligands were prepared with Spartan'10 (Wavefunction, Inc.) and the scripts provided by the AutoDock tools (v1.5.4 r29). The docking grid sizes and positions were set using AutoDock. The grid dimensions were large enough to cover the whole biological unit and centered onto previously defined relevant catalytic and biological areas. The grid size was 26 × 28 × 30 Å (MC1R). In MC1R, the loop region closing the receptor (a.a. 155–163) was set as the flexible region for docking. The lowest docking energy and predicted free binding energy for each compound were retained. These energies were used to sort the VLS results. Top docking solutions were manually inspected using COOT and PyMOL V1.5 to verify and confirm their compatibility with existing knowledge on the receptor (e.g., active site location, binding pockets, and conserved amino acid residues potentially involved in binding).

Model Analysis and Validation

The models were analyzed using Ligplot (Wallace et al., 1995). Confirmation of the interaction maps was performed manually by visual inspection of the models in COOT (Emsley and Cowtan, 2004).

Statistical Analysis

Where applicable, the data are expressed as means ± SD; Student's t-test was used for single statistical comparisons.

A probability value of $p < 0.05$ was used as the criterion for statistical significance.

RESULTS

Effect of 7,3′,4′-THIF and Daidzein on the α-MSH-Induced Melanin Content in B16F10 Melanoma Cells

To investigate the whitening effect of 7,3′,4′-THIF and daidzein, we first examined the melanin content in B16F10 melanoma cells using Fontana-Masson staining and a melanin content assay. 7,3′,4′-THIF showed cytotoxicity at 50 μM concentration when treated in B16F10 for 96 h, and 7,3′,4′-THIF were incubated with B16F10 under non-cytotoxic conditions for all experiments (**Supplementary Figure S1**). Treatment with 7,3′,4′-THIF, but not daidzein, significantly reduced the intracellular melanin content of B16F10 cells in a dose-dependent manner (**Figures 1Cg–i**). Consistent with this, 7,3′,4′-THIF inhibited

the α-MSH-induced extracellular melanin content in a dose-dependent manner (**Figure 1D**). Additionally, 7,3′,4′-THIF lightened the color of the extracellular culture media to a greater degree than did arbutin, a well-known whitening agent. In contrast, daidzein had no effect on the α-MSH-induced extracellular melanin content at all indicated concentrations (**Figure 1D**). These results suggest that 7,3′,4′-THIF, but not daidzein, exerts a strong depigmenting effect on B16F10 cells.

Effect of 7,3′,4′-THIF on the Expression of Melanogenic Enzymes and CREB Phsphorylation

Tyrosinase, TYRP-1, and TYRP-2 play important roles in melanogenesis (Park et al., 2009; Yamaguchi and Hearing, 2009; Gillbro and Olsson, 2011). Thus, we examined whether 7,3′,4′-THIF could suppress tyrosinase expression. 7,3′,4′-THIF suppressed α-MSH-induced tyrosinase, TYRP-1 and TYRP-2 expression in B16F10 cells (**Figure 2A**). Because α-MSH induces

FIGURE 2 | Effect of 7,3′,4′-THIF on melanogenic protein expression in B16F10 melanoma cells. **(A)** 7,3′,4′-THIF suppressed tyrosinase and TYRP-1 but not TYRP-2 expression in B16F10 melanoma cells. To determine the expression of melanogenic enzymes, the cells were pretreated with 7,3′,4′-THIF at the indicated concentrations (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH for 3 days. **(B)** 7,3′,4′-THIF suppressed MITF transcription factor expression and the phosphorylation of CREB in B16F10 melanoma cells. To examine transcription factor expression, the cells were pretreated with 7,3′,4′-THIF at the indicated concentrations (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH for 3 days or 30 min. The levels of indicated proteins were determined by Western blot analysis, as described in "Materials and Methods," using specific antibodies. The data are representative of more than two independent experiments that gave similar results. Asterisks indicate a significant difference (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$) compared with α-MSH treated group.

FIGURE 3 | Effect of 7,3′,4′-THIF on the α-MSH-induced dephosphorylation of AKT signaling or phosphorylation of p38 and PKA signaling in B16F10 melanoma cells. **(A)** 7,3′,4′-THIF activated the α-MSH-induced dephosphorylation of AKT. Cells were treated with 7,3′,4′-THIF (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH and harvested 1 h later. **(B)** 7,3′,4′-THIF inhibited the α-MSH-induced phosphorylation of p38. Cells were treated with 7,3′,4′-THIF (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH and harvested 15 min later. **(C)** 7,3′,4′-THIF inhibited the α-MSH-induced phosphorylation of PKA. Cells were treated with 7,3′,4′-THIF (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH and harvested 15 min later. The cells were disrupted and the levels of phosphorylated and total proteins were determined by Western blot analysis, as described in section "Materials and Methods," using specific antibodies against the respective phosphorylated and total proteins. The data are representative of three independent experiments that gave similar results. The levels of indicated proteins were determined by Western blot analysis, as described in section "Materials and Methods," using specific antibodies. The data are representative of more than two independent experiments that gave similar results. Asterisks indicate a significant difference (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$) compared with α-MSH treated group.

tyrosinase and TYRP expression mainly through cAMP/PKA signaling and subsequently induces both MITF expression and CREB phosphorylation, we next investigated whether 7,3′,4′-THIF inhibited MITF expression and CREB phosphorylation. 7,3′,4′-THIF strongly suppressed α-MSH-induced MITF expression and CREB phosphorylation (**Figure 2B**), suggesting that the whitening effect of 7,3′,4′-THIF is mediated by the suppression of melanogenic protein expression in B16F10 cells.

Effects of 7,3′,4′-THIF on α-MSH-Mediated Signal Pathways

Previous studies have shown that AKT, MAPK, and PKA signaling is involved in α-MSH-induced melanogenesis and tyrosinase expression in B16F10 cells (Busca and Ballotti, 2000; Smalley and Eisen, 2000; Tada et al., 2002; Saha et al., 2006; Solano et al., 2006; Bellei et al., 2008; Yamaguchi and Hearing, 2009; Gillbro and Olsson, 2011; Jin et al., 2011).

Thus, we next investigated the effect of 7,3′,4′-THIF on AKT, MAPK, and PKA signaling. 7,3′,4′-THIF activated the α-MSH-induced dephosphorylation of AKT, mTOR, and GSK3β in B16F10 cells at the concentrations tested (**Figure 3A**). 7,3′,4′-THIF also suppressed the α-MSH-induced phosphorylation of MKK3/6, p38, and MSK1 in a dose-dependent manner (**Figure 3B**). Additionally, 7,3′,4′-THIF suppressed the α-MSH-induced phosphorylation of PKA in B16F10 cells at 40 μM (**Figure 3C**). Taken together, these results indicate that 7,3′,4′-THIF inhibits tyrosinase expression by activating AKT signaling and suppressing MAPK and PKA signaling.

Binding of 7,3′,4′-THIF to MC1R

We next investigated whether 7,3′,4′-THIF affected adenylyl cyclase, an upstream effector of the AKT, MAPK, and PKA cascades. After the stimulation of MC1R by α-MSH, adenylyl cyclase converts ATP to cAMP, resulting in the activation of
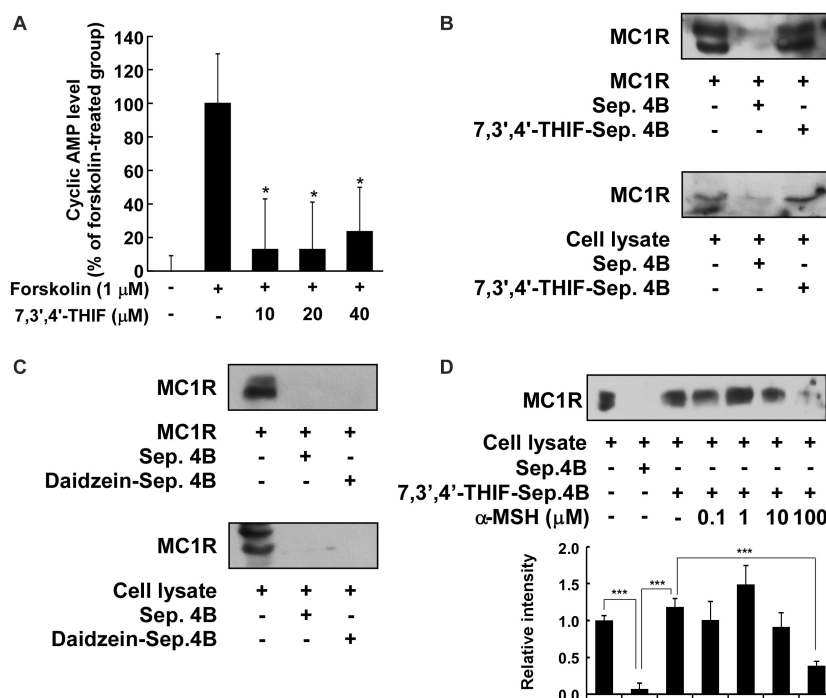

FIGURE 4 | Binding of 7,3′,4′-THIF to MC1R. **(A)** 7,3′,4′-THIF reduced the forskolin-induced cAMP level in B16F10 melanoma cells. The intracellular cAMP level was determined by a cAMP immunoassay as described in section "Materials and Methods." The results are expressed as the cAMP level relative to the forskolin-treated control. All data are presented as the mean ± SD of three independent determinations. Asterisks (*) indicate a significant difference ($p < 0.05$) compared with the forskolin-treated group. **(B)** 7,3′,4′-THIF binds MC1R directly *in vitro* and *ex vivo*. MC1R-7,3′,4′-THIF binding was confirmed by immunoblotting using antibodies against human MC1R (upper panel) or mouse MC1R (lower panel). Lane 1 (input control), human MC1R protein standard or whole B16F10 cell lysates; lane 2 (negative control), Sepharose 4B was used to pull-down MC1R, as described in section "Materials and Methods," or a B16F10 cell lysate precipitated with Sepharose 4B beads; and lane 3, 7,3′,4′-THIF-Sepharose 4B affinity beads were used to pull-down MC1R or whole B16F10 cell lysates precipitated with 7,3′,4′-THIF-Sepharose 4B affinity beads. **(C)** Daidzein did not bind with MC1R *in vitro* and *ex vivo*. MC1R-daidzein binding was confirmed by immunoblotting using antibodies against human MC1R (upper panel) or mouse MC1R (lower panel). Lane 1 (input control), human MC1R protein standard or whole B16F10 cell lysates; lane 2 (control), Sepharose 4B was used to pull-down MC1R, as described in section "Materials and Methods," or a B16F10 cell lysate precipitated with Sepharose 4B beads; and lane 3, daidzein-Sepharose 4B affinity beads were used to pull-down MC1R or whole B16F10 cell lysates precipitated with daidzein-Sepharose 4B affinity beads. **(D)** 7,3′,4′-THIF binds to MC1R competitively with α-MSH. B16F10 cellular supernatant fraction (1,000 μg) was incubated with α-MSH at the concentrations indicated (0, 0.1, 1, 10, or 100 μM) and 100 μM of 7,3′,4′-THIF-Sepharose 4B or Sepharose 4B (as a negative control) in a reaction buffer to a final volume of 500 μL. The pulled-down proteins were detected by western blot analysis as described in "Materials and Methods": lane 1 (input control), whole B16F10 cell lysates; lane 2 (negative control), indicating that neither MC1R binds with Sepharose 4B and lane 3 is the positive control, which indicates that MC1R binds with 7,3′,4′-THIF-Sepharose 4B. Each experiment was performed three times; representative blots are shown. Asterisks indicate a significant difference.
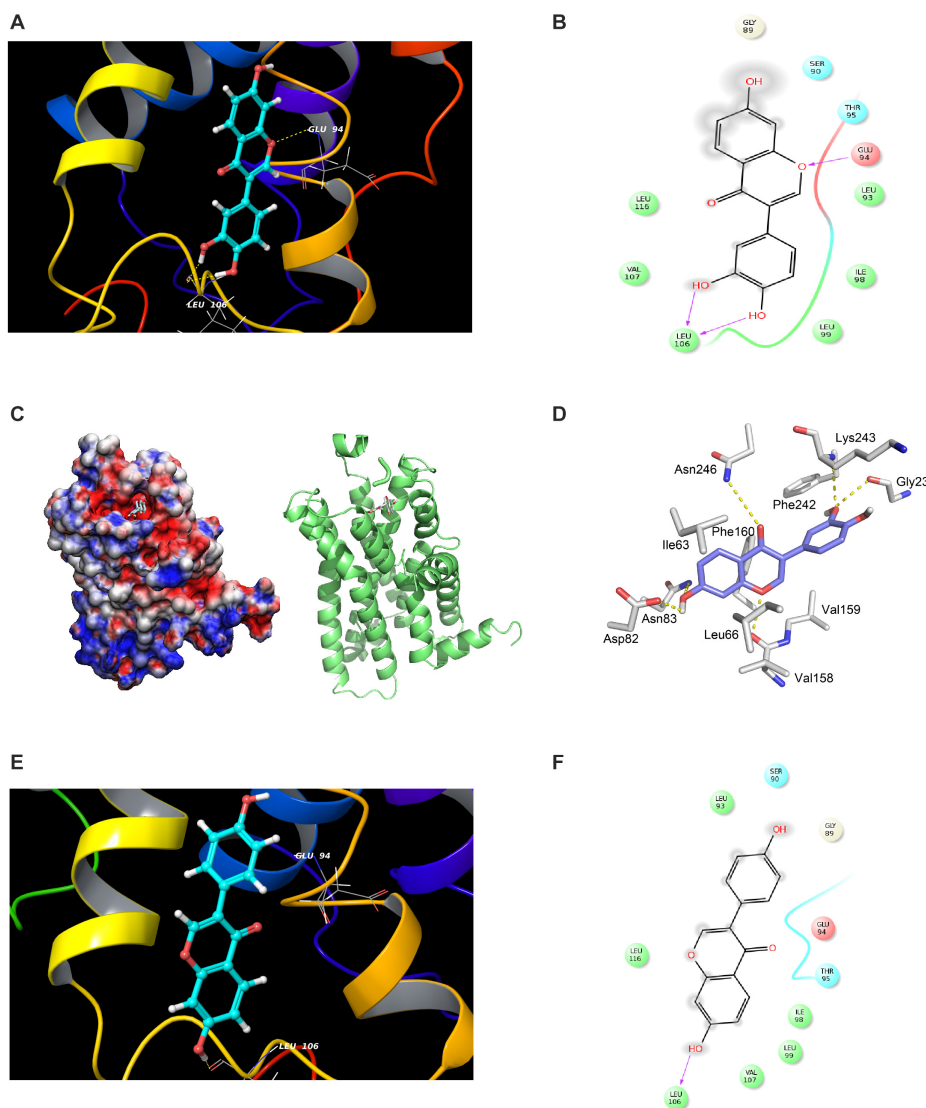
FIGURE 5 | Modeling study of MC1R binding to 7,3′,4′-THIF and daidzein. **(A–D)** 7,3′,4′-THIF binding with MC1R. 7,3′,4′-THIF binding (shown as stick) with MC1R formed hydrogen bonds (yellow dash line) at Glu94 and Leu106 **(A–C)**. Ligand interaction diagram of 7,3′,4′-THIF binding with MC1R **(D)**. The hydroxyl group at the 3′ position of 7,3′,4′-THIF binding with MC1R, formed hydrogen bonds (yellow dash line) with the backbone oxygen of Gly239 and the backbone nitrogen of Lys243 in the MC1R pocket. **(E,F)** Modeling study of MC1R binding to daidzein. Daidzein binding (shown as stick) with MC1R, formed hydrogen bonds (yellow dash line) at Leu106 **(E)** Ligand interaction diagram of daidzein binding with MC1R **(F)**.

TABLE 1 | Predicted docking energy.

Rank	Docking energy (kcal/mol)	Compound
1	−9.4 (−4.31)	7,3′,4′-Trihydroxyisoflavone (7,3′,4′-THIF)
2	−9.0 (−1.1)	4,7′-Dihydroxyisoflavone (Daidzein)
3	−6.8	α-MSH
4	0.1	Agouti Signaling protein (C-Term)

downstream signaling pathways (Solano et al., 2006; Yamaguchi and Hearing, 2009; Gillbro and Olsson, 2011). Therefore, to identify the effect of 7,3′,4′-THIF on the intracellular cAMP

level, we performed a cAMP immunoassay using forskolin, a direct activator of adenylyl cyclase. At 40 μM, 7,3′,4′-THIF reduced the forskolin-induced intracellular cAMP level by up to 23.8% (**Figure 4A**). These results indicate that the inhibition of tyrosinase expression by 7,3′,4′-THIF involves the inhibition of intracellular cAMP formation and adenylyl cyclase activity.

Accumulating data suggest that adenylyl cyclase or MC1R is the potential molecular target of 7,3′,4′-THIF, and that binding results in the inhibition of melanogenesis and tyrosinase expression. To confirm whether 7,3′,4′-THIF binds directly to MC1R, we performed an *in vitro* and *ex vivo* pull-down assay using 7,3′,4′-THIF-conjugated and non-conjugated Sepharose 4B beads. MC1R (**Figure 4B** and **Supplementary Figure S2A**,
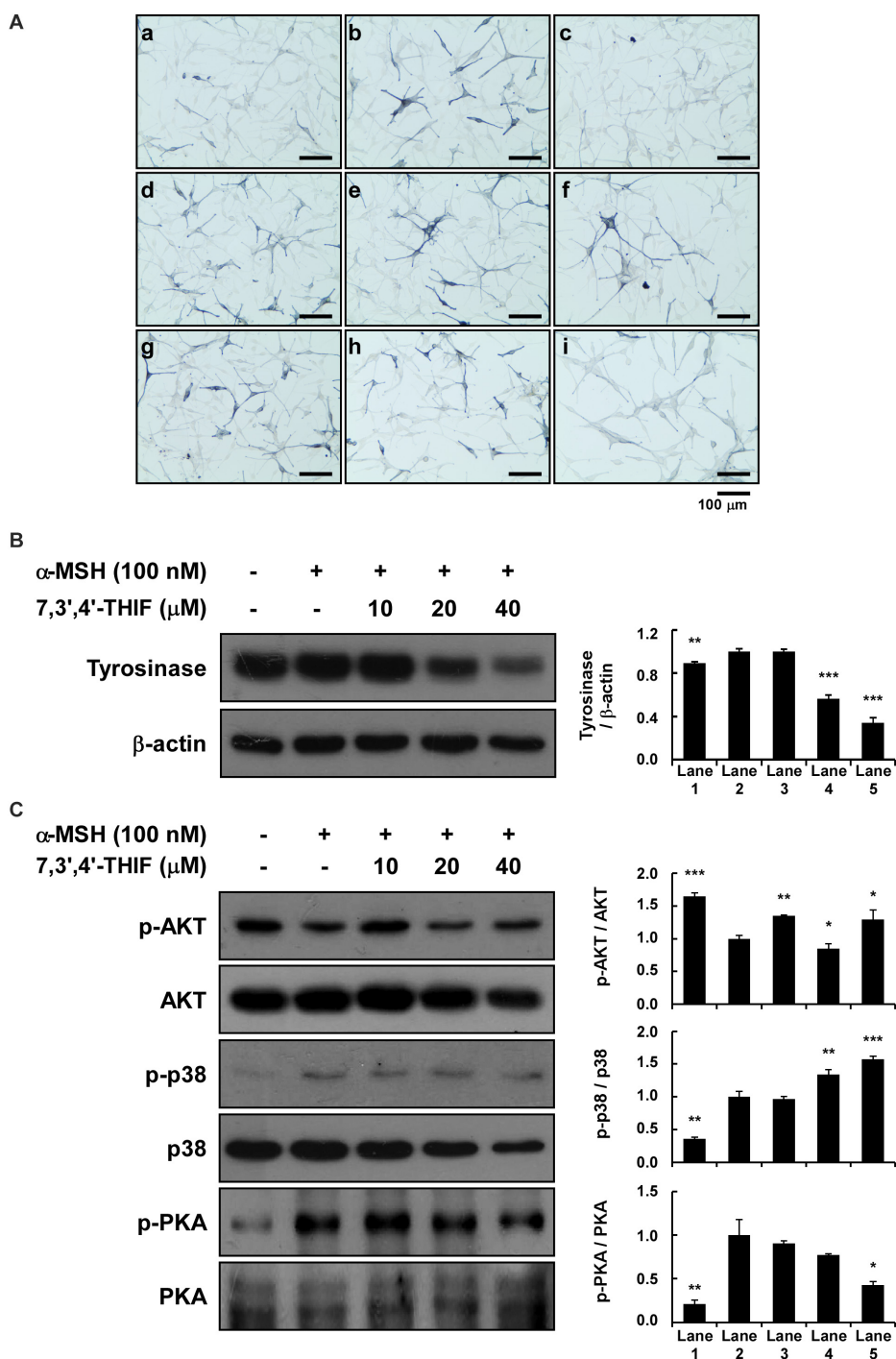
FIGURE 6 | Effect of 7,3′,4′-THIF and daidzein on the α-MSH-induced melanogenesis of HEMs. **(A)** Effects of 7,3′,4′-THIF on α-MSH-induced intracellular melanin in untreated control cells (a) and in cells treated with α-MSH (b), α-MSH and 100 μM arbutin (c), α-MSH and 10 μM daidzein (d), α-MSH and 20 μM daidzein (e), α-MSH and 40 μM daidzein (f), α-MSH and 10 μM 7,3′,4′-THIF (g), α-MSH and 20 μM 7,3′,4′-THIF (h), and α-MSH and 40 μM 7,3′,4′-THIF (i). **(B)** 7,3′,4′-THIF suppressed tyrosinase in HEMs. To determine the expression level, the cells were pretreated with 7,3′,4′-THIF at the indicated concentrations (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH for 3 days. **(C)** Effect of 7,3′,4′-THIF on the α-MSH-regulated phosphorylation of AKT, p38 and PKA. Cells were treated with 7,3′,4′-THIF (10, 20, or 40 μM) for 1 h before being exposed to 100 nM α-MSH and harvested 0.5 h later. The cells were disrupted and the levels of phosphorylated and total proteins were determined by Western blot analysis, as described in section "Materials and Methods," using specific antibodies against the respective phosphorylated and total proteins. The data are representative of more than two independent experiments that gave similar results. Asterisks indicate a significant difference (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$) compared with α-MSH treated group.

lane 3) bound to the 7,3′,4′-THIF-Sepharose 4B beads, but not to the non-conjugated Sepharose-4B beads (**Figure 4B** and **Supplementary Figure S2A**, lane 2), revealing that 7,3′,4′-THIF binds MC1R directly *in vitro* and *ex vivo*. However, daidzein did not bind to MC1R both *in vitro* and *ex vivo* (**Figure 4C** and **Supplementary Figure S2B**). The binding ability of 7,3′,4′-THIF with MC1R was altered in a concentration-dependent manner in the presence of α-MSH (**Figure 4D**), suggesting that 7,3′,4′-THIF competes with α-MSH for binding with MC1R. These results suggest that 7,3′,4′-THIF is an α-MSH-competitive inhibitor for suppressing MC1R activation. Computer modeling study also revealed that 7,3′,4′-THIF but not daidzein easily docked to the α-MSH-binding site of MC1R (**Figure 5** and **Table 1**).

Effect of 7,3′,4′-THIF on α-MSH-Induced Melanin Synthesis in HEMs

We next investigated the whitening effect of 7,3′,4′-THIF and daidzein on HEMs. 7,3′,4′-THIF at a concentration of 80 μM maintained cell viability above 80% for 72 h in HEMs (**Supplementary Figure S3**). 7,3′,4′-THIF, but not daidzein, significantly reduced the intracellular melanin content of HEMs in a dose-dependent manner (**Figure 6A**). It was also necessary to verify the mechanisms of 7,3′,4′,-THIF on melanogenesis in the human cells. 20 and 40 μM of 7,3′,4′-THIF significantly reduced α-MSH-induced tyrosinase expression in HEMs (**Figure 6B**). Furthermore, α-MSH-induced PKA phosphorylation was downregulation in HEMs cultured with 40 μM of 7,3′,4′-THIF, and phosphorylation of Akt inhibited by α-MSH treatment was partially ameliorated by 7,3′,4′-THIF treatment (**Figure 6C**). 7,3′,4′-THIF also slightly decreased the phosphorylated p38, but the level of p38 also showed a tendency to decrease, showing an increasing phospho-/total-form ratio. These results indicate that 7,3′,4′-THIF could exert the same depigmenting effect in both murine melanoma cells and human melanocytes.

DISCUSSION

Soybean, one of the most important foods in Asia, is important as a source of protein, as well as an important source of isoflavone (Genovese et al., 2007). In particular, genistein and daidzein are major active compounds of soybeans and their many pharmacological activities have been known. The isoflavones are metabolized in the body, which requires evaluation of physiological activity in the form of their metabolites. 7,3′,4′-THIF is considered one of the main oxidized metabolites of daidzein (Heinonen et al., 2003). Recently, 7,3′,4′-THIF, a metabolite of daidzein, was reported to have beneficial effects on hypopigmentation. Previous studies showed that 7,3′,4′-THIF inhibits melanin production in melan-a (Yan et al., 1999) and B16 melanoma (Goh et al., 2011) cells. Although accumulating evidence suggests that 7,3′,4′-THIF can inhibit hyperpigmentation, the underlying molecular mechanisms and its specific target protein have not been reported yet. Here, we report the marked depigmentation effect of 7,3′,4′-THIF on α-MSH-induced

hyperpigmentation and suggest the underlying molecular mechanism and targets.

To determine the effect of 7,3′,4′-THIF on α-MSH-induced melanogenesis, we first investigated its effect on intracellular and extracellular melanin levels in B16F10 melanoma cells. 7,3′,4′-THIF, but not daidzein, inhibited α-MSH-induced intracellular and extracellular melanin production in B16F10 cells. These results indicate that 7,3′,4′-THIF plays a more important role than daidzein in the depigmenting effect of soy.

Previous studies have established the role of tyrosinase and/or TYRPs in B16F10 melanoma cells (Goding and Fisher, 1997; Sato et al., 1997; Yasumoto et al., 1997). Upregulation of the level of several melanogenic enzymes, including tyrosinase and TYRPs, promotes melanin synthesis (Goding and Fisher, 1997; Sato et al., 1997; Yasumoto et al., 1997). Additionally, the expression of tyrosinase is primarily regulated by transcription factors such as MITF and CREB. Therefore, the inhibition of MITF and/or CREB leads to the suppression of melanin synthesis through reduced tyrosinase expression (Goding and Fisher, 1997; Sato et al., 1997; Yasumoto et al., 1997). Our results showed that 7,3′,4′-THIF inhibited the α-MSH-induced tyrosinase, TYRP-1, TYRP-2 and MITF expression and CREB phosphorylation. This transcriptional regulation of tyrosinase by 7,3′,4′-THIF was mediated by the activation of AKT and inhibition of p38 and PKA phosphorylation. 7,3′,4′-THIF also inhibited cAMP production by directly binding with MC1R both *in vitro* and *ex vivo*, competitively with α-MSH. Taken together, these results indicate that the inhibition of tyrosinase, TYRP-1, and TYRP-2 expression by 7,3′,4′-THIF was attributable to the suppression of

FIGURE 7 | Hypothetical mechanism of the depigmenting activity of 7,3′,4′-THIF.
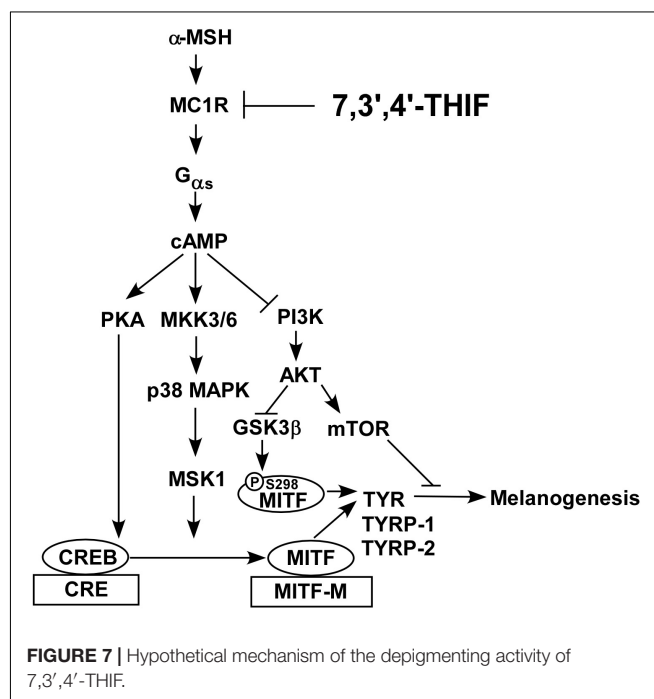
MC1R activity. The melanocytes for our western blot data had different treatment times of 7,3′,4′-THIF with α-MSH, depending on the proteins to be confirmed. Induction of tyrosinase, TYRPs, and MITF expression required a relatively long treatment of α-MSH for up to 3 days, and a relatively short treatment time (15–60 min) was required to regulate kinase phosphorylation. This is consistent with the sequence of processes in which the signaling pathways regulated by external stimulation regulate the transcriptional activity of MITF, followed by mRNA regulation, and finally the expression of TYR and TYRPs.

Considering the experimental result showing that 7,3′,4′-THIF binds to MC1R, we carried out modeling study to investigate the binding mode of 7,3′,4′-THIF to MC1R. In the previous study, Prusis et al. (1997) suggested that Glu94 of MC1R has great influence on ligand binding and receptor function. Based on our docking model results, only 7,3′,4′-THIF can bind at Glu94 with binding affinities of -4.31 kcal/mol (**Figures 5A–D** and **Table 1**), while daidzein can't, it formed a hydrogen at Leu106 with binding affinities of -1.1 kcal/mol (**Figures 5E,F** and **Table 1**). Additionally, in the another model structure of MC1R complexed with 7,3′,4′-THIF, the hydroxyl group at the 3′ position of 7,3′,4′-THIF can potentially form a hydrogen bond with the backbone oxygen of Gly239 and the backbone nitrogen of Lys243 in the MC1R pocket (**Figure 5D**). Because of the lack of a hydroxyl group at the 3′ position of daidzein, its interaction with the ligand binding site of MC1R would be weaker than that of 7,3′,4′-THIF; thus, daidzein is unable to effectively inhibit MC1R. These results suggested the possible reasons why only 7,3′,4′-THIF can inhibit the activity of MC1R, but not daidzein. Further investigation using X-ray crystallography to determine the structure of the inhibitor complex will elucidate the exact binding modes of 7,3′,4′-THIF with MC1R. Because primary human melanocytes are physiologically relevant to human skin, they have frequently been used for the *in vitro* screening of skin-whitening compounds. Consistent with the above result, 7,3′,4′-THIF significantly reduced intracellular melanin production in HEMs.

In summary, 7,3′,4′-THIF, a metabolite of daidzein, more effectively inhibited the α-MSH-induced hyperpigmentation of B16F10 cells than did daidzein. This inhibition is mediated primarily through suppressed MITF expression and CREB phosphorylation, and the subsequent reduced tyrosinase,

TYRP-1 and TYRP-2 expressions. 7,3′,4′-THIF inhibited the α-MSH-induced dephosphorylation of AKT and the phosphorylation of p38 and PKA. 7,3′,4′-THIF strongly inhibited the forskolin-induced intracellular cAMP level through binding with MC1R competing with α-MSH. 7,3′,4′-THIF also strongly suppressed α-MSH-induced intracellular melanin production in HEMs. Collectively, these results suggested that MC1R is the potent molecular targets that bind 7,3′,4′-THIF for suppressing hyperpigmentation (**Figure 7**). These results provide insight into the biological actions of 7,3′,4′-THIF and the molecular basis for the development of new skin whitening agent.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JK and NK contributed to conceptualization of the study. JK, J-EL, EL, HC, and ZD conducted experiment. MY, JP, and KL helped with analyzing the data. JK, J-EL, and NK wrote and edited the manuscript. J-EL, TK, and NK contributed to manuscript revision and approved the submitted version. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.577284/full#supplementary-material

REFERENCES

An, S. M., Koh, J. S., and Boo, Y. C. (2010). p-coumaric acid not only inhibits human tyrosinase activity in vitro but also melanogenesis in cells exposed to UVB. *Phytother. Res.* 24, 1175–1180. doi: 10.1002/ptr.3095

Bellei, B., Flori, E., Izzo, E., Maresca, V., and Picardo, M. (2008). GSK3beta inhibition promotes melanogenesis in mouse B16 melanoma cells and normal human melanocytes. *Cell Signal* 20, 1750–1761. doi: 10.1016/j.cellsig.2008.06.001

Briganti, S., Camera, E., and Picardo, M. (2003). Chemical and instrumental approaches to treat hyperpigmentation. *Pigment Cell Res.* 16, 101–110. doi: 10.1034/j.1600-0749.2003.00029.x

Busca, R., and Ballotti, R. (2000). Cyclic AMP a key messenger in the regulation of skin pigmentation. *Pigment Cell Res.* 13, 60–69. doi: 10.1034/j.1600-0749.2000.130203.x

Eisinger, M., and Marko, O. (1982). Selective proliferation of normal human melanocytes in vitro in the presence of phorbol ester and cholera toxin. *Proc. Natl. Acad. Sci. U S A.* 79, 2018–2022. doi: 10.1073/pnas.79.6.2018

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 12 Pt 1), 2126–2132. doi: 10.1107/S0907444904019158

Friedmann, P. S., and Gilchrest, B. A. (1987). Ultraviolet radiation directly induces pigment production by cultured human melanocytes. *J. Cell Physiol.* 133, 88–94. doi: 10.1002/jcp.1041330111

Genovese, M. I., Barbosa, A. C., Pinto Mda, S., and Lajolo, F. M. (2007). Commercial soy protein ingredients as isoflavone sources for functional foods. *Plant Foods Hum. Nutr.* 62, 53–58. doi: 10.1007/s11130-007-0041-0

Gillbro, J. M., and Olsson, M. J. (2011). The melanogenesis and mechanisms of skin-lightening agents - existing and new approaches. *Int. J. Cosmetic Sci.* 33, 210–221. doi: 10.1111/j.1468-2494.2010.00616.x

Goding, C. R., and Fisher, D. E. (1997). Regulation of melanocyte differentiation and growth. *Cell Growth Differ.* 8, 935–940.

Goh, M. J., Park, J. S., Bae, J. H., Kim, D. H., Kim, H. K., and Na, Y. J. (2011). Effects of Ortho-dihydroxyisoflavone Derivatives from Korean Fermented Soybean Paste on Melanogenesis in B16 Melanoma Cells and Human Skin Equivalents. *Phytother. Res.* 26, 1107–1112. doi: 10.1002/ptr.3682

Gordon, P. R., Mansur, C. P., and Gilchrest, B. A. (1989). Regulation of human melanocyte growth, dendricity, and melanization by keratinocyte derived factors. *J. Invest. Dermatol.* 92, 565–572. doi: 10.1111/1523-1747.ep12709595

Heinonen, S. M., Hoikkala, A., Wahala, K., and Adlercreutz, H. (2003). Metabolism of the soy isoflavones daidzein, genistein and glycitein in human subjects. Identification of new metabolites having an intact isoflavonoid skeleton. *J. Steroid Biochem. Mol. Biol.* 87, 285–299. doi: 10.1016/j.jsbmb.2003.09.003

Holst, M., Kozack, R. E., Saied, F., and Subramaniam, S. (1994). Protein electrostatics: rapid multigrid-based Newton algorithm for solution of the full nonlinear Poisson-Boltzmann equation. *J. Biomol. Struct. Dyn.* 11, 1437–1445. doi: 10.1080/07391102.1994.10508078

Jacobson, M. P., Friesner, R. A., Xiang, Z., and Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* 320, 597–608. doi: 10.1016/s0022-2836(02)00470-9

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J., Honig, B., Shaw, D. E., et al. (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins* 55, 351–367. doi: 10.1002/prot.10613

Jin, M. L., Park, S. Y., Kim, Y. H., Park, G., Son, H. J., and Lee, S. J. (2011). Suppression of alpha-MSH and IBMX-induced melanogenesis by cordycepin via inhibition of CREB and MITF, and activation of PI3K/Akt and ERK-dependent mechanisms. *Int. J. Mol. Med.* 29, 119–124. doi: 10.3892/ijmm.2011.807

Joshi, P. G., Nair, N., Begum, G., Joshi, N. B., Sinkar, V. P., and Vora, S. (2007). Melanocyte-keratinocyte interaction induces calcium signalling and melanin transfer to keratinocytes. *Pigment Cell Res.* 20, 380–384. doi: 10.1111/j.1600-0749.2007.00397.x

Kim, J. M., Ko, R. K., Jung, D. S., Kim, S. S., and Lee, N. H. (2010). Tyrosinase inhibitory constituents from the stems of Maackia fauriei. *Phytother. Res.* 24, 70–75. doi: 10.1002/ptr.2870

Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. Jr. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* 77, 778–795. doi: 10.1002/prot.22488

Kulling, S. E., Honig, D. M., and Metzler, M. (2001). Oxidative metabolism of the soy isoflavones daidzein and genistein in humans in vitro and in vivo. *J. Agric. Food Chem.* 49, 3024–3033. doi: 10.1021/jf0012695

Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.

Lee, M. H., Lin, Y. P., Hsu, F. L., Zhan, G. R., and Yen, K. Y. (2006). Bioactive constituents of Spatholobus suberectus in regulating tyrosinase-related proteins and mRNA in HEMn cells. *Phytochemistry* 67, 1262–1270. doi: 10.1016/j.phytochem.2006.05.008

Leyden, J., and Wallo, W. (2011). The mechanism of action and clinical benefits of soy for the treatment of hyperpigmentation. *Int. J. Dermatol.* 50, 470–477. doi: 10.1111/j.1365-4632.2010.04765.x

Oberoi, H., and Allewell, N. M. (1993). Multigrid solution of the nonlinear Poisson-Boltzmann equation and calculation of titration curves. *Biophys J.* 65, 48–55. doi: 10.1016/S0006-3495(93)81032-4

Park, H. Y., Kosmadaki, M., Yaar, M., and Gilchrest, B. A. (2009). Cellular mechanisms regulating human melanogenesis. *Cell Mol. Life Sci.* 66, 1493–1506. doi: 10.1007/s00018-009-8703-8

Park, J. S., Kim, D. H., Lee, J. K., Lee, J. Y., Kim, H. K., Lee, H. J., et al. (2010). Natural ortho-dihydroxyisoflavone derivatives from aged Korean fermented soybean paste as potent tyrosinase and melanin formation inhibitors. *Bioorg. Med. Chem. Lett.* 20, 1162–1164. doi: 10.1016/j.bmcl.2009.12.021

Pogozheva, I. D., Chai, B. X., Lomize, A. L., Fong, T. M., Weinberg, D. H., Nargund, R. P., et al. (2005). Interactions of human melanocortin 4 receptor with nonpeptide and peptide agonists. *Biochemistry* 44, 11329–11341. doi: 10.1021/bi0501840

Prusis, P., Schioth, H. B., Muceniece, R., Herzyk, P., Afshar, M., Hubbard, R. E., et al. (1997). Modeling of the three-dimensional structure of the human melanocortin 1 receptor, using an automated method and docking of a rigid cyclic melanocyte-stimulating hormone core peptide. *J. Mol. Graph. Model.* 15, 307–317. doi: 10.1016/s1093-3263(98)00004-7

Saha, B., Singh, S. K., Sarkar, C., Bera, R., Ratha, J., Tobin, D. J., et al. (2006). Activation of the Mitf promoter by lipid-stimulated activation of p38-stress signalling to CREB. *Pigment Cell Res.* 19, 595–605. doi: 10.1111/j.1600-0749.2006.00348.x

Sali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815. doi: 10.1006/jmbi.1993.1626

Sato, S., Roberts, K., Gambino, G., Cook, A., Kouzarides, T., and Goding, C. R. (1997). CBP/p300 as a co-factor for the Microphthalmia transcription factor. *Oncogene* 14, 3083–3092. doi: 10.1038/sj.onc.1201298

Smalley, K., and Eisen, T. (2000). The involvement of p38 mitogen-activated protein kinase in the alpha-melanocyte stimulating hormone (alpha-MSH)-induced melanogenic and anti-proliferative effects in B16 murine melanoma cells. *FEBS Lett.* 476, 198–202. doi: 10.1016/s0014-5793(00)01726-9

Solano, F., Briganti, S., Picardo, M., and Ghanem, G. (2006). Hypopigmenting agents: an updated review on biological, chemical and clinical aspects. *Pigment Cell Res.* 19, 550–571. doi: 10.1111/j.1600-0749.2006.00334.x

Tada, A., Pereira, E., Beitner-Johnson, D., Kavanagh, R., and Abdel-Malek, Z. A. (2002). Mitogen- and ultraviolet-B-induced signaling pathways in normal human melanocytes. *J. Invest. Dermatol.* 118, 316–322. doi: 10.1046/j.0022-202x.2001.01694.x

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461. doi: 10.1002/jcc.21334

Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* 8, 127–134. doi: 10.1093/protein/8.2.127

Yamaguchi, Y., and Hearing, V. J. (2009). Physiological factors that regulate skin pigmentation. *Biofactors* 35, 193–199. doi: 10.1002/Biof.29

Yan, C. H., Chen, X. G., Li, Y., and Han, R. (1999). Effects of genistein, a soybean-derived isoflavone, on proliferation and differentiation of B16-BL6 mouse melanoma cells. *J. Asian Nat. Prod. Res.* 1, 285–299. doi: 10.1080/10286029908039877

Yasumoto, K., Yokoyama, K., Takahashi, K., Tomita, Y., and Shibahara, S. (1997). Functional analysis of microphthalmia-associated transcription factor in pigment cell-specific transcription of the human tyrosinase family genes. *J. Biol. Chem.* 272, 503–509. doi: 10.1074/jbc.272.1.503

Functional Characterization of Primordial Protein Repair Enzyme M38 Metallo-Peptidase From *Fervidobacterium islandicum* AW-1

Jae Won La[1][†]**, Immanuel Dhanasingh**[2][†]**, Hyeonha Jang**[3]**, Sung Haeng Lee**[2] **and Dong-Woo Lee**[1]*****

[1] *Department of Biotechnology, Yonsei University, Seoul, South Korea,* [2] *Department of Cellular and Molecular Medicine, Chosun University School of Medicine, Gwangju, South Korea,* [3] *School of Applied Biosciences, Kyungpook National University, Daegu, South Korea*

The NA23_RS08100 gene of *Fervidobacterium islandicum* AW-1 encodes a keratin-degrading β-aspartyl peptidase (*Fi*BAP) that is highly expressed under starvation conditions. Herein, we expressed the gene in *Escherichia coli*, purified the recombinant enzyme to homogeneity, and investigated its function. The 318 kDa recombinant *Fi*BAP enzyme exhibited maximal activity at 80°C and pH 7.0 in the presence of Zn^{2+}. Size-exclusion chromatography revealed that the native enzyme is an octamer comprising a tetramer of dimers; this was further supported by determination of its crystal structure at 2.6 Å resolution. Consistently, the structure of *Fi*BAP revealed three additional salt bridges in each dimer, involving 12 ionic interactions that might contribute to its high thermostability. In addition, the co-crystal structure containing the substrate analog *N*-carbobenzoxy-β-Asp-Leu at 2.7 Å resolution revealed binuclear Zn^{2+}-mediated substrate binding, suggesting that *Fi*BAP is a hyperthermophilic type-I IadA, in accordance with sequence-based phylogenetic analysis. Indeed, complementation of a Leu auxotrophic *E. coli* mutant strain (ΔiadA and ΔleuB) with *Fi*BAP enabled the mutant strain to grow on isoAsp-Leu peptides. Remarkably, LC-MS/MS analysis of soluble keratin hydrolysates revealed that *Fi*BAP not only cleaves the C-terminus of isoAsp residues but also has a relatively broad substrate specificity toward α-peptide bonds. Moreover, heat shock-induced protein aggregates retarded bacterial growth, but expression of BAP alleviated the growth defect by degrading damaged proteins. Taken together, these results suggest that the viability of hyperthermophiles under stressful conditions may rely on the activity of BAP within cellular protein repair systems.

Keywords: M38 β-aspartyl peptidase, protein repair, starvation, type-I BAP, hyperthermophile stress responses, keratin degradation, *Fervidobacterium islandicum* AW-1

INTRODUCTION

Protein homeostasis (proteostasis), a balanced state between folded proteins and protein aggregates, is critical for cellular metabolism, physiology, and normal aging (Hutt and Balch, 2010). To maintain proteostasis, molecular chaperones control the folding of newly synthesized proteins, and the ubiquitin-proteasome

system degrades misfolded and damaged polypeptides (Hipp et al., 2019). In eukaryotes, damaged proteins destined for degradation are ubiquitinated and subsequently degraded by the 26S proteasome (Goldberg, 2003). Like the ubiquitin-proteasome pathway in eukaryotes, prokaryotes also degrade abnormal proteins through a quality control network (QCN) consisting of chaperones and proteases (Goldberg, 1972; Mogk et al., 2011). To ensure protein homeostasis, bacteria should monitor the folding of proteins and prevent the accumulation of misfolded proteins via coordinated refolding or degradation (Mogk et al., 2011). Misfolded and abnormal proteins are produced for several reasons, including somatic mutations in genes that result in proteins failing to adopt the native folded structure, errors during transcription or translation, failure of the chaperone machinery, mistakes during the post-translational modification of proteins, and structural modifications caused by environmental changes (heat, oxidative agents, pH, or osmotic conditions), and induction of protein misfolding through seeding and cross-seeding mechanisms (Visick and Clarke, 1995; Moreno-Gonzalez and Soto, 2011). One of the major modifications of proteins is the deamination of asparaginyl or the isomerization of aspartyl residues through a cyclic intermediate, resulting in aspartyl or isoaspartyl (isoAsp) residues that contribute to inactivation, aggregation, and aging of proteins in tissue (Cournoyer et al., 2005). Indeed, isoAsp formation is associated with abnormal functioning of various proteins, including calmodulin, epidermal growth factor, and ribonuclease (Aswad, 1995). The formation of isoAsp can also decrease the biological activity of a protein drug, alter its susceptibility to proteolytic degradation, and induce autoimmunity (Aswad et al., 2000).

Bacteria possess enzymatic repair systems to prevent conformational damage to proteins resulting from proline isomerization, methionine oxidation, and formation of isoAsp residues (Visick and Clarke, 1995). Protein L-isoAsp *O*-methyltransferase (PIMT), one of the QCN proteins, is found in most archaea and Gram-negative eubacteria, as well as eukaryotes (Ryttersgaard et al., 2002). PIMT minimizes accumulation of atypical isoAsp residues by methylating the free α-carboxyl group of isoAsp residues (Reissner and Aswad, 2003). PIMT-deficient mice accumulate high levels of damaged proteins containing isoAsp residues, resulting in growth retardation and sudden early death (Kim et al., 1997). A similar PIMT protein repair enzyme was identified in the hyperthermophilic bacterium *Thermotoga maritima*, which grows optimally at 90°C. Accumulation of proteins with altered aspartyl and asparaginyl residues is believed to be detrimental to cell survival at elevated temperatures. The rate of isoAsp residue formation increases 910-fold at 90°C compared with 23°C (Ichikawa and Clarke, 1998); hence, modified or damaged protein (peptide) aggregates should be rapidly digested to prevent cellular damage and to provide a potential nutrient source to support cell survival in harsh environments. However, PIMT alone does not appear to be sufficient to prevent accumulation of protein aggregates due to its low conversion activity (Reissner and Aswad, 2003). Rather, isoAsp peptidase hydrolyzes the peptide bond directly to generate the β-carboxylate group of Asp (Marti-Arbona et al.,

2005a), which is more efficient at handling peptides containing isoAsp residues.

The NA23_RS08100 gene encodes M38 β-aspartyl peptidase (BAP) in the feather-degrading bacterium *Fervidobacterium islandicum* AW-1 (Nam et al., 2002; Lee et al., 2015a). Among 57 genes encoding proteases in this organism, the NA23_RS08100 gene is most highly expressed when cells are grown under starvation conditions (Kang et al., 2020), suggesting that in addition to PIMT, the protein encoded by NA23_RS08100 might be involved in degrading isoAsp peptides under stressful conditions. To gain insight into the biological role of BAP as a QCN protein, we functionally and structurally characterized the protein product of the NA23_RS08100 gene in *F. islandicum* AW-1.

MATERIALS AND METHODS

Phylogenetic Analysis

A phylogenetic tree was generated using the maximum likelihood method in MEGA X version 10.0.5, based on multiple sequence alignment of M38 β-aspartyl peptidase from *Fervidobacterium islandicum* AW-1 (*Fi*BAP) and homologs sharing ≥ 40% amino acid sequence identity. A bootstrap consensus tree was inferred from 1,000 replicates. Evolutionary distances were computed using the Poisson correction method and expressed as the number of amino acid substitutions per site. Amino acid sequence alignment was performed using ClustalX software V2.1. Sequence similarities and secondary structure information from aligned sequences were generated using ESPript 3.0 (Robert and Gouet, 2014).

Cloning and Expression of the NA23_RS08100 Gene

Genomic DNA was isolated from *F. islandicum* AW-1, purified using a genomic DNA extraction kit according to the manufacturer's instructions (Qiagen, Hilden, Germany), and then used as template for PCR amplification. The NA23_RS08100 gene encoding *Fi*BAP was amplified by PCR using forward (5′-GCTAGCATGATAAAAATTATAAAGAACG-3′) and reverse (5′-CTCGAGTCAAAATTCAAAGTTTAAGTTC-3′) primers (underlined sequences represent the restriction sites for *Nhe*I and *Xho*I). The PCR product was digested with *Nhe*I and *Xho*I and cloned into the pET-28a(+) vector (Novagen, San Diego, CA), yielding pET-*Fi*BAP. The resulting plasmid encodes the target gene fused to a 6× His-tag at the N-terminus. For expression of recombinant *Fi*BAP, *E. coli* BL21 (DE3) cells transformed with pET-*Fi*BAP were grown at 37°C in 1 L of Luria-Bertani (LB) medium containing kanamycin (50 μg/mL) to an optical density at 600 nm (OD$_{600}$) of 0.5–0.8. After induction with 1 mM Isopropyl-1-thio-β-D-galactopyranoside (IPTG), cells were grown overnight at 37°C, then harvested by centrifugation (10,000 × g, 20 min, 4°C) and stored at −80°C.

Purification of *Fi*BAP

The harvested cells were resuspended in 20 mM Tris-HCl buffer containing 500 mM NaCl, 5 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride (PMSF) (pH 7.9) and disrupted

by sonication. After centrifugation at 10,000 × g for 20 min, the supernatant was heated at 60°C for 30 min then centrifuged at 10,000 × g for 20 min to remove denatured *E. coli* proteins. The supernatant was applied to Ni^{2+}-affinity resin (10 mL) equilibrated with the same buffer, and His-tagged protein was eluted with 20 mM Tris-HCl buffer containing 500 mM NaCl and 250 mM imidazole. Samples were loaded on a Superdex 200 10/300 GL column (GE Healthcare, USA) equilibrated with 25 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl (**Supplementary Table 1**). The major fractions were analyzed by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized using Coomassie Blue staining.

To monitor the recombinant *Fi*BAP, western blotting was performed using a 6×His tag monoclonal antibody (ThermoFisher Scientific, USA). Each sample was subjected to 12% SDS-PAGE, electrophoretically transferred to a polyvinylidene fluoride (PVDF) membrane (Bio-Rad, Hercules, CA, USA), blocked with 5% skim milk in TBST buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.05% Tween 20) at 4°C overnight, and incubated with horseradish peroxidase-conjugated 6×His epitope monoclonal antibody (1:1,000 dilution) for 2 h at room temperature. The membranes were then washed three times with TBST buffer for 10 min and were developed using a WESTSAVE Up western blotting detection system (AbFrontier, Seoul, Korea).

Preparation of Recombinant Feather Keratins Expressed in *E. coli* Cells

Recombinant feather keratins were prepared as described previously (Jin et al., 2017) and used as native substrates. Briefly, *E. coli* cells expressing recombinant feather keratins were resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 1 mM PMSF, pH 8.0) and disrupted by sonication. After centrifugation at 10,000 × g for 30 min, expressed keratins in the form of inclusion bodies were resuspended in lysis buffer containing 8 M urea and 1 mM PMSF, incubated on ice for 1 h, and centrifuged at 16,000 × g for 30 min. Supernatants were filtered through a 0.45 μm membrane, then applied to a 10 mL Ni-NTA agarose resin column (Qiagen, Germany) equilibrated with lysis buffer containing 8 M urea. Fractions containing unfolded keratins were eluted with 250 mM imidazole, concentrated using an Amicon Ultra-3K device (Millipore, USA), and buffer-exchanged by step-wise dialysis against 50 mM Tris-HCl (pH 8.0) at 4°C. Dialyzed samples were centrifuged at 10,000 × g for 30 min to remove insoluble material, and the resulting supernatants containing refolded keratins were concentrated using an Amicon Ultra-3K device (Millipore).

Enzyme Activity Assay

*Fi*BAP activity was determined by measuring the increase in free amino acids (Rosen, 1957). Reaction mixtures (0.16 mL) containing 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (pH 8.0), 500 ng/mL enzyme, and 1 mM β-Asp-Leu (β-DL) as a synthetic substrate were incubated at 80°C for 15 min. The reaction was stopped by adding trichloroacetic acid (TCA) reagent. After an additional 15 min incubation at 100°C for color development,

followed by cooling on ice, the absorbance was measured at 570 nm. One unit of *Fi*BAP activity was defined as the amount of enzyme that produced 1 μmol of product per min under the assay conditions (U/mg). Additionally, to assess the proteolytic activity of the purified enzyme, we measured the increase in free amino acids using the ninhydrin assay (Rosen, 1957) with casein and gelatin as substrates.

Biochemical Characterization of *Fi*BAP

The temperature dependence of *Fi*BAP activity was measured using the standard protocol but the reaction temperatures was varied from 40 to 98°C. To determine the effect of pH on *Fi*BAP activity, reaction mixtures (0.16 mL) were incubated at 80°C under standard assay conditions but HEPES buffer was replaced by 50 mM sodium acetate buffer (pH 4.0–6.0), HEPES buffer (pH 6.0–8.0), borate buffer (pH 8.0–10.0), or sodium bicarbonate buffer (pH 9.0–10.0). All pH values were adjusted at room temperature, and $\Delta pK_a/\Delta Ts$ (in which the latter term is the change in temperature) for each buffer was taken into account when the results were analyzed.

To determine the kinetic parameters, assays were performed in 50 mM HEPES (pH 8.0) containing 1 mM $CoCl_2$, 500 ng/mL enzyme, and 1–32 mM β-DL substrate. Reaction mixtures were incubated for 10 min at 80°C and stopped by cooling on ice. Kinetic parameters were obtained by fitting the empirical data to the Michaelis-Menten equation using Origin 8.0 software.

To analyze the effect of metals, purified *Fi*BAP was treated with 20 mM EDTA at room temperature for 2 h to remove metal ions, followed by overnight dialysis against 20 mM Tris-HCl (pH 7.5) at 4°C with three changes of buffer. The divalent metal ion content of both as-isolated and EDTA-treated samples were determined by high-resolution inductively-coupled plasma mass spectrometry (ICP-MS) on a PlasmaQuad 3 instrument. To assess the effects of various metal ions on *Fi*BAP activity, the EDTA-treated samples were preincubated with 1 mM $CoCl_2 \cdot 6H_2O$, $MnCl_2 \cdot 4H_2O$, $MgCl_2 \cdot 6H_2O$, $CaCl_2 \cdot 2H_2O$, $ZnCl_2 \cdot 6H_2O$, $CuCl_2 \cdot 2H_2O$, $FeCl_2 \cdot 6H_2O$, or $NiCl_2 \cdot 6H_2O$ for 15 min at 50°C. The residual activity was measured in triplicate under standard assay conditions.

Crystallization and Structure Determination of *Fi*BAP

Preliminary crystallization trials were performed using purified *Fi*BAP (10 mg/mL) in an appropriate buffer (20 mM Tris-HCl pH 7.4, 50 mM NaCl) mixed with protein crystallization solutions (Hampton Research and Wizard; 1:1 volume ratio) by the hanging drop vapor-diffusion method at 20°C (293 K) as described previously (Lee et al., 2015b). The initial *Fi*BAP crystals were further improved by modifying the Crystal Screen Lite 34 conditions (0.1 M sodium acetate pH 4.6 and 1 M sodium formate). *Fi*BAP was co-crystallized with substrate analog *N*-carbobenzoxy-β-Asp-Leu (Cbz-β-DL) at a molar ratio of 1:1.2, and crystals of the complex were obtained in modified PEG/Ion screen 2–32 conditions (2% v/v tacsimate pH 5.0, 0.1 M sodium citrate tribasic dihydrate, pH 5.6; 16% w/v polyethylene glycol 3350). Crystals were soaked for 30 s in a cryosolution containing

mother liquor plus 20% glycerol, then frozen in liquid nitrogen prior to synchrotron radiation diffraction experiments.

X-ray diffraction data were collected from *Fi*BAP crystals on beamline 7A at the Pohang Light source (Pohang, Korea) using an ADSC Q270 detector, with an oscillation of 1.0° and a 1 s exposure per frame over a 360° range at a wavelength of 0.97934 Å. Crystals of *Fi*BAP with and without β-DL as substrate grown using the above conditions diffracted to 2.6 Å and 2.7 Å, respectively. Diffraction data were processed using the HKL-2000 program (HKL Research Inc.). In the case of the ligand-free structure, molecular replacement (MR) was performed with *Ec*IadA (PDB: 1ONW) as the search model using the CCP4 program MOLREP (Vagin and Teplyakov, 2010). The initial model was further refined using the PHENIX (Afonine et al., 2010) and REFMAC5 (Murshudov et al., 2011) programs to achieve a model with R_{work} and R_{free} values of 16.9 and 22.1%, respectively. In the case of the Cbz-β-DL-bound *Fi*BAP diffraction dataset, to enhance the overall data quality, two native datasets of equivalent diffraction limits were combined using the Scale_and_Merge function of the PHENIX suite (Adams et al., 2010). MR was performed for the Cbz-β-DL-bound *Fi*BAP structure with the refined crystal structure of ligand-free *Fi*BAP as the search model using the Phaser program (McCoy, 2007). The chemical coordination file for β-DL was built using Coot (Emsley and Cowtan, 2004) and eLBOW (Moriarty et al., 2009). The ligand was fitted into the map, and several rounds of refinement were performed using the REFMAC5 and PHENIX programs, yielding R_{work} and R_{free} values of 19.7 and 27.5%, respectively. The structures of *Fi*BAP with and without Cbz-β-DL have been deposited in the Protein Data Bank (PDB) under accession codes 7CDH and 7CF6, respectively.

Construction of Leucine Auxotrophic *E. coli* Mutants

The leucine (Leu) auxotrophic *E. coli* BL21 (DE3) strain was constructed by deleting the *leuB* gene encoding 3-isopropyl malate dehydrogenase, which is involved in leucine biosynthesis, by the RED/ET recombination method using a Quick & Easy *E. coli* Gene deletion kit (Gene Bridges) according to the manufacturer's instructions. A functional cassette flanked with FRT and homology arms was generated by PCR using the FRTPGK-gb2-neo-FRT fragment as a template and primers 5′-AGTTGCAACGCAAAGCTCAACACAACGAAAACAACA-AGGAAACCGTGTGAAATTAACCCTCACTAAAGGGCGG-3′ (forward) and 5′-GTACACAACGTGAGCGTCGAACAAT-TTTTCGTATAACGTCTTAGCCATGATAATACGACTCACT-ATAGGGCTCG-3′ (reverse). The kanamycin selection marker was removed by transforming cells with FLP recombinase expression plasmid 706-FLP (Gene Bridges). After selecting mutants lacking the selection marker cassette by streaking each colony on LB agar plates with and without 50 μg/mL kanamycin, deletion of the *leuB* gene was confirmed by PCR with primers 5′-GCTAACTACAACGGTCGCCGCTTCCACGGCGTC-3′ (forward) and 5′-GGCGCGCAGACCATCGAACGCCTGCGG-TGAGG-3′ (reverse).

For functional complementation of *Fi*BAP, a functional cassette flanked with FRT and homology arms was generated to delete the *iadA* gene encoding *Ec*IadA using primers 5′-CATTGCTGTCGATCTGGGTTATGCAGCTTATTGTTTAA-CAAGGAGTTACCAATTAACCCTCACTAAAGGGCGG-3′ (forward) and 5′-TCAGCCGCCCTTGCGGGCATTCTACG-TCCATTCGGGCGGCTGACAACCGTTAATACGACTCACT-ATAGGGCTCG-3' (reverse). In addition, primers 5′-GTC-GGTCGCTGCCTGGGGACAGCCGAAGTG-3′ (forward) and 5′-CTGAGAGTTGCAGCGGCGTTACCTGGCGG-3′ (reverse) were used to confirm the *iadA* deletion mutants.

Complementation of *iadA* Deletion by *Fi*BAP

To monitor the growth phenotypes of the Leu auxotrophic *E. coli* strains harboring the pET28a-*Fi*BAP and the pET28a-*iadA* plasmids, each protein-overexpressing strain was streaked on appropriate minimal M9 agar plates containing 50 μg/mL kanamycin and 1 mM IPTG supplemented with different compositions of amino acids (19 amino acids with and without Leu) in the presence and absence of 1.6 mM β-DL. To investigate the effect of heat shock on bacterial growth retardation, each strain was grown at 37°C to an OD_{600} of 1.2–1.5 in LB medium containing kanamycin (50 μg/mL), and then incubated at 44°C for 0.5 h or longer until cells entered the stationary phase.

To determine the total protein content of soluble (supernatant) and insoluble (pellet) fractions, cells were harvested before (mid-exponential phase) and after (stationary phase) heat treatment. The same amount of cells, as judged by the OD_{600} value, was resuspended in 50 mM Tris-HCl buffer (pH 8.0) and disrupted by sonication. After centrifugation at 10,000 × g for 30 min, the protein concentration of fractionated samples was determined using the bicinchoninic acid (BCA) assay (Smith et al., 1985), and each subcellular fractions was visualized on the 12% acrylamide gel for SDS-PAGE analysis.

LC-MS/MS Analysis

Keratin hydrolysates were analyzed by reversed-phase HPLC-ESI-MS/MS using a NanoLC-2D Ultra system (Eksigent, Dublin, CA, USA) coupled to a LTQ-XL mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) in direct injection mode. Briefly, after a 5 μL injection, keratin hydrolysates were loaded onto a reversed-phase ProteoPepII C18 column (5 μm, 300 Å pore size, 0.15 × 25 mm; New Objective IntegraFrit, Scientific Instrument Services, Inc., Ringoes, NJ, USA) and eluted onto a Molex Polymicro Flexible Fused Silica Capillary Tubing (I.D. 75 μm, O.D. 375 μm) at a flow rate of 0.4 μL/min. The gradient consisted of mobile phase A [0.1% formic acid (v/v) in water], mobile phase B [0.1% formic acid (v/v) in acetonitrile], 2% solvent B (0–1 min), 5% solvent B (1–62 min), 35% solvent B (62–65 min), 60% solvent B (65–70 min), and 2% solvent B in A (70–90 min) with a total runtime of 90 min, including mobile phase equilibration.

MS analysis of peptide eluents was performed on a LTQ-XL system (ThermoFisher Scientific) in positive-ion mode with a nano ion spray voltage of 2.3 kV, a scan range of 300–1,800 (m/z),

a curtain gas pressure of 20 psi, a nebulizer gas pressure of 6 psi, and an interface heater temperature of 200°C. Full-scan MS spectra were acquired from 35 precursors selected for MS/MS analysis from the 300 to 1,800 m/z range, utilizing a dynamic exclusion of 30 s. The IDA collision energy (CE) parameter script, selecting up to 35 precursors with charge states of +2 to +3, was employed to automatically control the CE.

Data were processed using MSConvert (ProteoWizard). This software converts raw data (.RAW format) into peak lists (.mgf format). The FASTA database employed contained the *Gallus gallus* keratin sequence, and this afforded the opportunity to employ the target decoy database search strategy (Elias and Gygi, 2007). Data containing both MS and MS/MS information were uploaded into Xcalibar software (ThermoFisher Scientific) and used to generate MS-extracted ion chromatograms (XICs) for each identified peptide. For the MS-GF search, carbamidomethyl (C) was set as a fixed modification and oxidation (M) was set as a variable modification. The peptide tolerance was ±4.0 Da, and the MS/MS tolerance was ±1 Da. The software algorithm simultaneously searched all modifications listed in UniMod (http://www.unimod.org/) (Shilov et al., 2007). False discovery rate (FDR) analysis was also performed using integrated tools in PeptideShaker (CompOmics), which generated.mgf files that were subsequently searched against the current *G. gallus* keratins SwissProt database using MS-GF.

Differential Scanning Calorimetry (DSC) Measurement

Calorimetric measurements were performed using a VP-DSC microcalorimeter (Microcal Inc., GE Malvern, Worcestershire, UK). All scans were run at pH 7.4 in 10 mM potassium phosphate buffer containing 150 mM NaCl, over a temperature range from 10 to 130°C at a rate of 90°C/h. The cell volume was 0.8 mL. Potassium phosphate buffer was used for baseline scans, and apparent melting temperature (T_m) values of *Fi*BAP (25 μM) were determined.

Analysis of Isoaspartic Acid Residues

An ISOQUANT Isoaspartate Detection Kit (Promega) was used for the detection of isoaspartate (isoAsp) according to the manufacturer's instructions. Each cell lysate (2 μg) from heat-shocked strains was incubated at 30°C for 30 min in the presence of L-isoaspartyl (isoAsp) methyltransferase (PIMT) and S-adenosyl methionine (SAM) in the reaction buffer, and the stop solution was then added to halt the reaction. The reaction mixture (40 μl) was loaded onto a reversed-phase YMC-Triart C18 column (250 × 4.6 mm I.D. S-5 μm, 12 nm) equilibrated with 90% 50 mM potassium phosphate (pH 6.2) and 10% methanol. The amount of isoAsp was determined by quantifying eluted S-adenosyl homocysteine (SAH).

RESULTS

In silico-Based Annotation of the NA23_RS8100 Gene Product

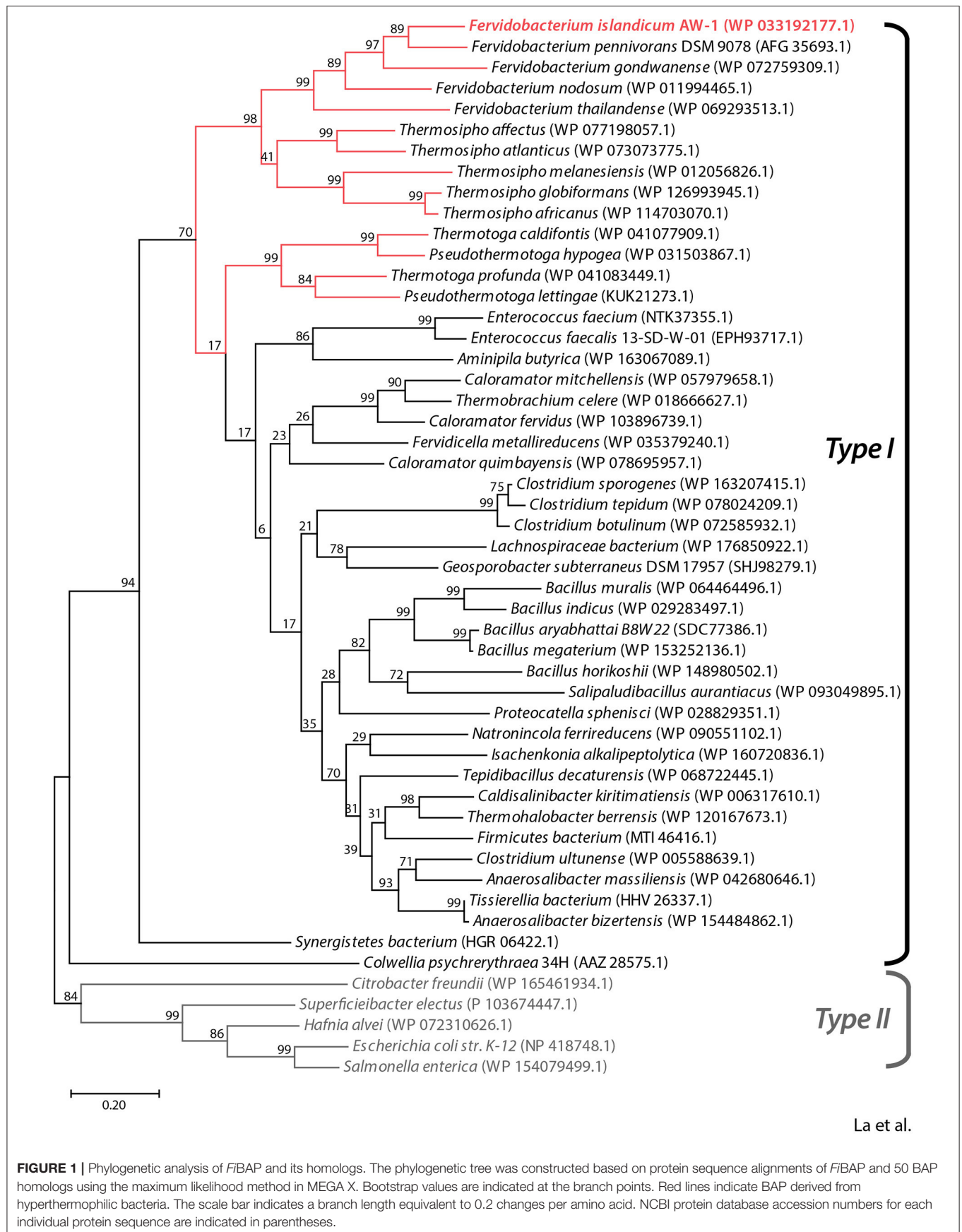To predict the functional role of the NA23_RS8100 gene product, we performed a BLASTP search using the deduced amino acid sequence to identify homologs. Multiple sequence alignment of the NA23_RS8100 gene product with β-aspartyl peptidases (BAPs) revealed that its amino acid sequence shares high levels of sequence identity with BAP from *Fervidobacterium* species (≥65%) and other genera including *Thermosipho* (≥55%), *Pseudothermotoga* (≥51%), *Thermotoga* (≥51%), *Clostridium* (≥49%), *Bacillus* (≥47%), and *Enterococcus* (≥49%; **Figure 1** and **Supplementary Figure 1**). In addition, the deduced amino acid sequence of the NA23_RS08100 gene product shares pronounced sequence identity with mesophilic BAPs from *Escherichia coli* (*Ec*IadA; 39% sequence identity; PDB:1ONW) and *Colwellia pshchrerythraea* 34H (*Cp*IadA; 41% sequence identity; PDB:5XGW), for which crystal structures are available. Remarkably, all BAP homologs derived from extremophiles belong to the Type I isoaspartyl dipeptidase (IadA) family, members of which possess a conserved Glu residue for metal binding in the active site (**Supplementary Figure 1**). Moreover, phylogenetic analysis showed that BAP from *F. islandicum* AW-1 (*Fi*BAP) has a highly close evolutionary relationship with homologs from extremophilic *Fervidobacterium* and *Thermosipho*, whereas extremophilic *Fi*BAP has diverged away from the Type II IadAs of mesophiles in the order *Enterobacteriales* and further away from *Ec*IadA and *Cp*IadA members of the BAP family (**Figure 1**). These results led us to tentatively conclude that the protein encoded by the NA23_RS8100 gene in *F. islandicum* AW-1 is β-aspartyl peptidase (subsequently named *Fi*BAP), representing an ancient form of IadA that has diverged from Type II IadA.

Biochemical Characterization of *Fi*BAP

We successfully expressed the NA23_RS08100 gene as a C-terminal hexahistidine (6×His)-tagged fusion protein in *E. coli* BL21 (DE3) cells. To facilitate subsequent purification, *E. coli* lysates were heat-treated to remove most of the endogenous proteins, and the resulting supernatants containing soluble *Fi*BAP were applied to a Ni^{2+}-affinity column. The purified 6×His-tagged *Fi*BAP fusion protein yields an intact recombinant protein with an estimated M_r of 42,000 by SDS-PAGE (**Figure 2A**). However, size exclusion chromatography using a Superdex 200 column suggested that the native form of *Fi*BAP is as an octameric protein with an apparent M_r of 318,000 (**Figure 2B**).

Unlike typical endo-proteases, *Fi*BAP displayed minimal activity toward casein and gelatin as substrates, but it was active against β-Asp-Leu (β-DL), with an apparent optimal temperature of 80°C (**Figure 2C**) and an apparent optimum pH of 7.0 at 80°C (**Figure 2D**). To further assess the structural stability of *Fi*BAP, we used DSC to determine the T_m, revealing a major midpoint of the thermal transition (T_m = 96.8 ± 0.5°C) and a minor one at 100.4 ± 1.44°C (**Figure 2E**).

Since *Ec*IadA is a metalloprotease, we performed ICP-MS analysis of as-isolated *Fi*BAP, demonstrating that it contained Zn^{2+} (1.14 ± 0.37 mol of metal per mol of monomer). To investigate the effects of inhibitors on *Fi*BAP activity, as-isolated enzyme was pre-incubated with various inhibitors and its residual activity was measured. Dithiothreitol, PMSF, and Pepstatin A had little inhibitory effect on *Fi*BAP, but 1%

FIGURE 1 | Phylogenetic analysis of *Fi*BAP and its homologs. The phylogenetic tree was constructed based on protein sequence alignments of *Fi*BAP and 50 BAP homologs using the maximum likelihood method in MEGA X. Bootstrap values are indicated at the branch points. Red lines indicate BAP derived from hyperthermophilic bacteria. The scale bar indicates a branch length equivalent to 0.2 changes per amino acid. NCBI protein database accession numbers for each individual protein sequence are indicated in parentheses.

FIGURE 2 | Biochemical properties of *Fi*BAP. **(A)** Purification of recombinant *Fi*BAP analyzed by 12% SDS-PAGE. **(B)** Size-exclusion chromatography of *Fi*BAP using a Superdex 200 pg 10/300 column and the enzyme activity of the eluents. The column was calibrated with blue dextran (2,000 kDa), thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), and ovalbumin (44 kDa) as standards (**Supplementary Figure 2**). The calibration curve shows Log M_r and K_{av} values for each standard protein. **(C)** Temperature dependence of *Fi*BAP activity. **(D)** Dependence of *Fi*BAP activity on pH. **(E)** DSC analysis of *Fi*BAP. **(F)** Effects of divalent metal ions and inhibitors on *Fi*BAP activity. Group differences were assessed using one-way analysis of variance (ANOVA), followed by Tukey HSD test (*$p <$ 0.05 and **$p <$ 0.01 vs. as-isolated). **(G)** Kinetic analysis of *Fi*BAP. The kinetic parameters for purified *Fi*BAP were estimated by steady-state kinetics analysis using β-DL as substrate.

SDS and 1 mM EDTA inhibited enzyme activity (**Figure 2F**). Subsequently, after a 15 min preincubation of EDTA-treated and dialyzed *Fi*BAP with various metal ions at 50°C, the residual activity was measured under standard assay conditions, demonstrating that *Fi*BAP activity was significantly increased in the presence of Zn^{2+} and Co^{2+} compared with that in the absence of divalent metal ions, suggesting that this enzyme may be a metalloenzyme.

Additionally, we performed steady-state kinetics analysis of *Fi*BAP with β-DL as substrate, yielding K_m and V_{max} values of 5.23 ± 1.59 mg/mL and 286.33 ± 65.7 μM/min, respectively (**Figure 2G**).

Notably, *Fi*BAP is the most abundant and highly expressed enzyme when *F. islandicum* AW-1 is grown on feathers under starvation conditions (Kang et al., 2020). Although a MEROPS (http://merops.sanger.ac.uk) database search provided some information on hierarchical, structure-based classification of the peptidases, it did not yield the predicted cleavage pattern for M38 BAP. To further investigate the hydrolysates resulting from *Fi*BAP degradation, we performed LC-MS/MS analysis using soluble keratins as native substrates (**Supplementary Table 2**). Remarkably, all peptides hydrolyzed by *Fi*BAP matched the keratin sequences derived from *G. gallus* sequences, indicating that *Fi*BAP can cleave not only the peptide bond between Asp and Leu, but also other α-peptide bonds between hydrophobic, aromatic, and hydrophilic amino acid residues at the C-terminus (**Figure 3**), as were the cases of the *Ec*IadA and BAP from rat tissue (Dorer et al., 1968; Gary and Clarke, 1995). This result suggests that *Fi*BAP is a beta-aspartyl

peptidase with a relatively broad substrate specificity toward α-peptide bonds.

Overall Architecture of *Fi*BAP

Crystal structures of the ligand-free form of *Fi*BAP and the ligand-bound form bound with *N*-carbobenzoxy-β-Asp-Leu (Cbz-β-DL) as a substrate analog were determined at resolutions of 2.6 and 2.7 Å, respectively (**Table 1**). Cbz-β-DL-free and -bound structures belonged to I422 and P22₁2 space groups, with one and four subunits in the asymmetric unit, respectively. The overall structure of each monomer can be divided into two distinct domains; an N-terminal β-sandwich domain (M1–G56; G344–E386) and a C-terminal catalytic domain (L57–K343), that are mainly involved in dimerization and catalysis, respectively (**Figure 4A**). The β-sandwich domain located at the N-terminus of the molecule mainly comprises nine β-sheets (β1–β6, β17–β19) arranged in two layers with β1, β3, β4, and β5 in one layer and β2, β6, and β17–β19 in the other layer (**Figure 4A**). Inter-layer interactions may stabilize the β-sandwich domain consisting of loops and β-strands through hydrophobic interactions and a salt bridge between K376 in β19 and E44 in β5. The catalytic domain at the C-terminus of the molecule folds into a (β/α)₈ triosephosphate isomerase (TIM)-barrel motif. Specifically, eight β-sheets (β7–β13) surrounded by nine α-helices (α2–α10) form the central core containing the substrate-binding site harboring two Zn^{2+} ions (**Figure 4A**).

A structural similarity search using the DALI server revealed that *Fi*BAP is very closely related to *Ec*IadA, a Type II IadA
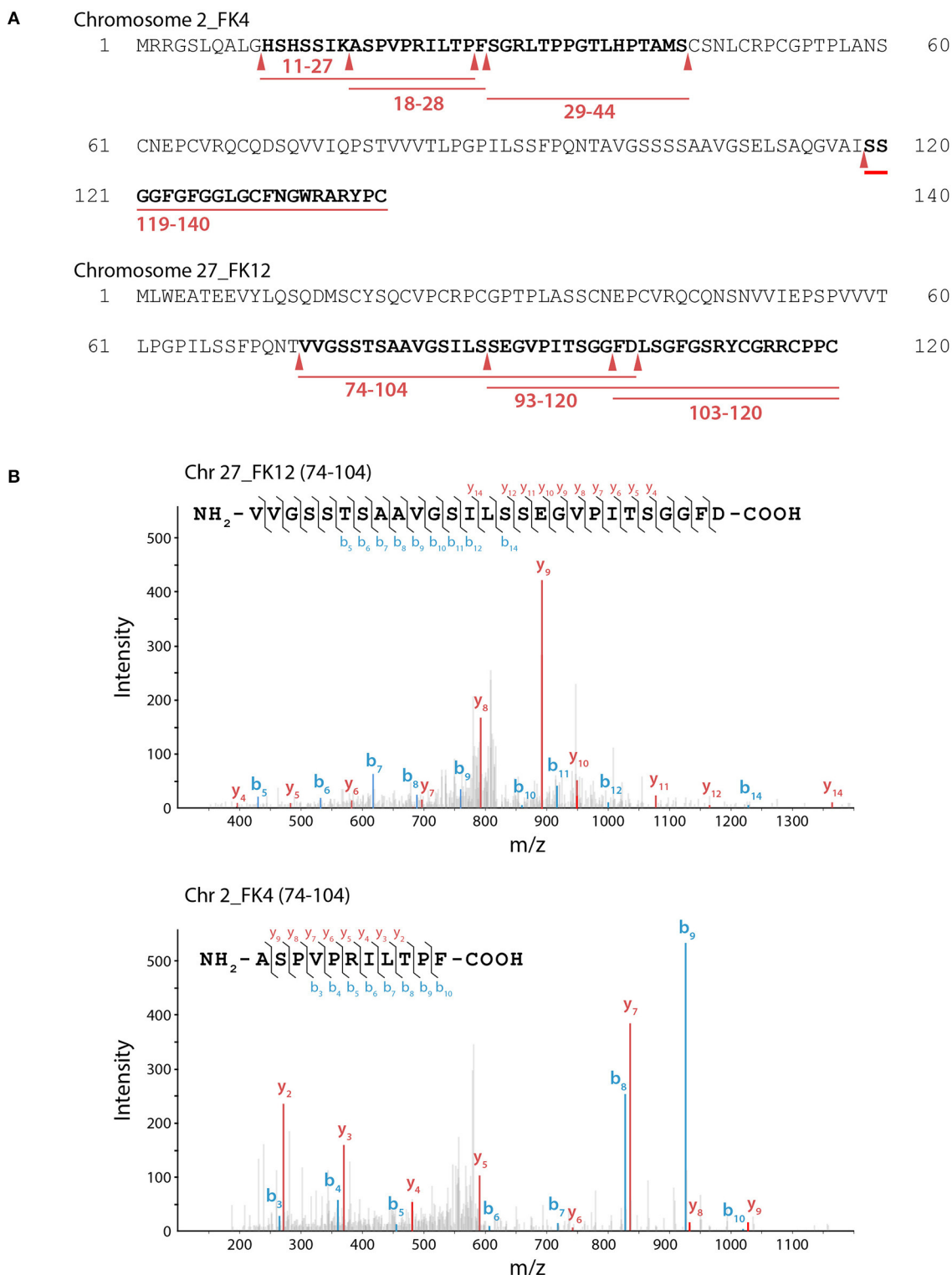
FIGURE 3 | LC-MS/MS analysis of keratin hydrolysates generated by *Fi*BAP. **(A)** Keratinolytic peptides matched with soluble Chr2_FK4 and Chr27_FK12 β-keratins are underlined in red, and their coverage sequences are depicted in bold. Red arrows indicate the putative cleavage site of *Fi*BAP from the identified peptide sequences. Purified recombinant feather keratin and *Fi*BAP were incubated for 18 h at 80°C in 50 mM Tris-HCl (pH 7.5). The enzyme/substrate ratio was 1:50 (w/w). **(B)** Representative MS spectra of keratinolytic peptides and fragment ions of detected peptides from recombinant feather keratins. The identified fragment ions from peptide sequences are labeled in the spectra and indicated on the peptide sequence above. The b (blue) and y (red) peaks correspond to the majority of the N- and C- terminal ions resulting from fragmentation of each peptide, respectively.

TABLE 1 | Crystallographic data collection statistics for *Fi*BAP.

	*Fi*BAP (ligand-free)	*Fi*BAP (ligand-bound)
PDB ID	**7CDH**	**7CF6**
Diffraction data statistics		
Beamline	PLS-7A	PLS-11C
Wavelength (Å)	0.97934	0.98461
Temperature (K)	100	100
Space group	I422	P2 2₁ 2
Cell parameters		
a, b, c (Å)	143.9, 143.9, 119.6	78.6, 150.4, 151.9
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Data resolution (Å)	30.0–2.6 (2.69–2.6)	47.6–2.7 (2.75–2.7)
Completeness (%)	97.7 (100.0)	99.7 (99.4)
Redundancy	22.9 (24.6)	10 (6.9)
Total reflections	439,332	466,992
Unique reflections	18,969 (1908)	36,967 (1041)
R_{merge}^a	7.3 (82.4)	7.4 (155.2)
$CC_{1/2}^b$	0.997 (0.963)	0.997 (0.368)
Matthew's coefficient (Å³ Da⁻¹)	3.66	2.67
Solvent content (%)	66.37	54.04
Average I/Sigma (I)	50.3 (5.54)	9.0 (1.0)
No. of chains per asymmetric unit	1	4
Refinement		
R_{work} (%)	19.9 (27.9)	19.7 (34.5)
R_{free} (%)	24.4 (36.4)	27.5 (37.7)
Protein residues/Ligands	360/0	1497/4
RMSD		
Angles (°)	1.35	1.97
Lengths (Å)	0.011	0.015
Average B-factors (Å²)	40.30	49.98
Ramachandran plot		
Most favored regions (%)	96	92.74
Allowed regions (%)	3.43	6.25
Outliers (%)	0.57	1.02

Values in parentheses correspond to the highest resolution shell.
$^a R_{merge} = \Sigma_{hkl} \, \Sigma_i \, | I_i(hkl) - <I(hkl)> | / \Sigma_{hkl} \, \Sigma_i \, I_i(hkl)$, *where $I_i(hkl)$ and $<I(hkl)>$ are the intensity of an individual reflection and the mean value of all measurements of an individual reflection, respectively.*
$^b CC_{1/2}$ *values are the correlations between intensities from random half-data sets.*

enzyme that has evolved in the direction of allowing post-translational modification (PTM) of the active site for metal binding and catalytic activity (Park et al., 2017). Structural neighbors identified in the *Fi*BAP query structure include IsoAsp dipeptidase, dihydropyrimidinase-related protein, enamidase, allantoinase, and dihydropyrimidinase, all of which are hydrolases. The structural architecture of *Fi*BAP was compared with previously determined structures of homologs (*Ec*IadA and *Cp*IadA). All monomers were superimposed with an average root mean square deviation (RMSD) of 0.705 Å for all Cα atoms (**Figure 4B**), suggesting that the overall scaffold of *Fi*BAP is similar to that of its homologs. Among these BAPs, structural elements are highly conserved, except for a minor difference in

the flexible loop connecting β4 and β5 (**Figures 4A,B**), which might be of no discernable biological significance.

Similar to *Ec*IadA (Elias and Gygi, 2007) and *Cp*IadA (Shilov et al., 2007), size-exclusion chromatography (**Figure 2B**) indicated that *Fi*BAP is an octamer in solution (S1–S8 in **Figure 4C**), usually referred to as a "tetramer of dimers" (**Figures 4C,D**), as supported by symmetry-related molecule analysis. The initial dimer is formed when the β-sandwich domain of each monomer (**Figure 4A**) interacts with an adjacent monomer, mainly through hydrophobic interactions involving several residues (I2, I4, V34, V35, V38, L39, F41, and I46) in the loop between β4 and β5 (upper red box in **Figure 4D**). Additionally, the substrate-binding cavity of the catalytic domain (**Figure 4A**) participates in dimerization by providing the second point of contact between adjacent partners through salt bridges. These four salt bridges are formed by E107 (α2), K114 (α2), R143 (α3), and D149 (α3) of each monomer of the dimer (lower red box in **Figure 4D**). Therefore, the hydrophobic interactions between β-sandwich domains, followed by hydrogen (H)-bonds and salt bridges between catalytic domains of adjacent subunits results in dimer formation (**Figure 4B** and **Supplementary Table 3**). Remarkably, the corresponding region in structural homologs such as *Ec*IadA and *Cp*IadA does not appear to contain some of these salt bridges because the distances between charged amino acids are >4 Å in all cases except E114-K150 of *Ec*IadA (**Supplementary Figure 3B**) (Thoden et al., 2003; Park et al., 2017). Hence, we propose that the strong salt bridges may contribute to the thermostability of *Fi*BAP, as reported previously for a hyperthermophilic PIMT, in which deletion of the interfacial residues that interact through salt bridges affected the overall thermostability of the enzyme (Tanaka et al., 2004).

Subsequently, each dimer interacts with two adjacent dimers at three contact points (boxed in **Figure 4E**) to form a biologically active octamer that includes two contact points (S2–S4 interface and S1–S3 interface in **Figure 4C**) between similar structural elements α4, α5 of one dimer, and loops T213–L220, G95–L106, and I159–L169 of the adjacent dimer. We identified additional salt bridges formed at this interface between residues R174 (α4), D178 (α4), and E214 (α5) with E171, R131, and R201, respectively (pink and purple boxes in **Figure 4E**). However, these corresponding residues are conserved and form salt bridges in both *Ec*IadA and *Cp*IadA, revealing their significance in the formation of a stable quaternary structure. In addition, the third contact point (S2–S3 interface) at the dimer-dimer interface is mainly comprised of H-bonds formed between interfacial residues (brown box in **Figure 4E**). These interactions at different contact points are required in BAPs for their overall stability and to maintain their high oligomeric state.

The ligand-free model included most amino acids except for residues D254–R259 and P290–L302, corresponding to the flexible loops in the molecule (**Figure 4A**). Intriguingly, the ligand-bound model included obvious electron density corresponding to loop P290–L302 between β14 and β15, which might have been stabilized upon ligand binding (**Figure 5A**). By contrast, loop D254–R259 of the Cbz-β-DL-bound *Fi*BAP structure remained disordered. Remarkably,
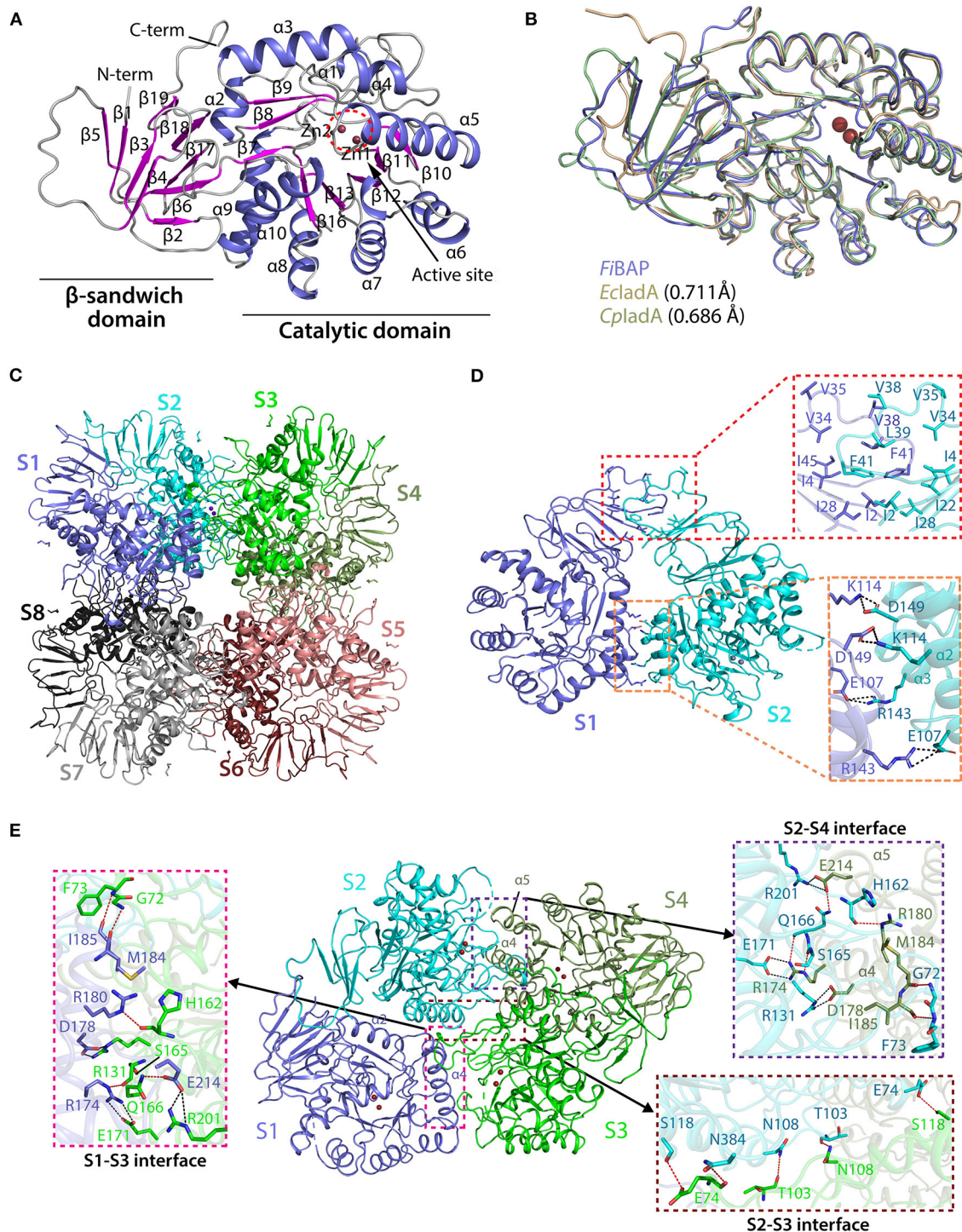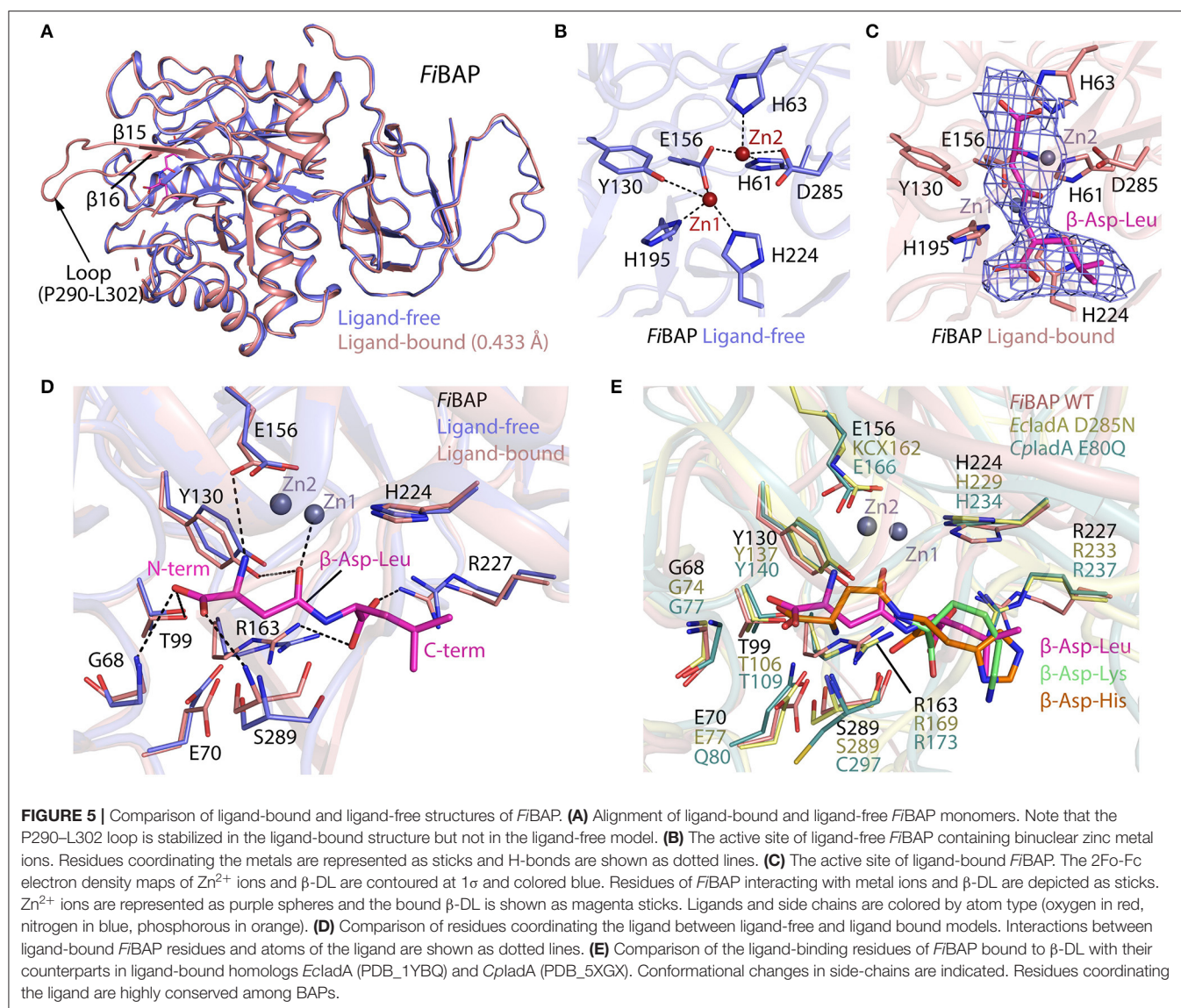
FIGURE 4 | Crystal structure of *Fi*BAP. **(A)** Domain structure of monomeric *Fi*BAP. The monomeric subunit of *Fi*BAP consists of two domains; an N-terminal β-sandwich domain and a C-terminal catalytic domain. The structural elements from the N- to C- termini of the protein are labeled. The active site (marked as a dotted red circle) contains two zinc metals (Zn1 and Zn2) located in the catalytic domain. **(B)** Structural alignments of *Fi*BAP with structural homologs *Ec*ladA (PDB ID 1ONW) and *Cp*ladA (5XGW). RMSD values are indicated in parentheses. **(C)** Quaternary structure of *Fi*BAP comprising dimers of two tetramers. Symmetry-related molecule analysis reveals *Fi*BAP to be an octamer of eight subunits (S1–S8, subunits colored differently). **(D)** Interactions at the interface between each subunit in a dimer. Dimers formed from two subunits (S1 and S2) are shown. The top red box shows hydrophobic interactions between β-sandwich domains, and the bottom salmon box shows salt bridges formed between catalytic domains of each subunit. **(E)** Dimer-dimer interactions. Interface interactions between the S1–S2 dimer (slate blue and cyan) and the S3–S4 dimer (green and smudge) are shown. The molecular interactions at three contact points, the S2–S4 interface (purple box), the S2–S3 interface (brown box), and the S1–S3 interface (magenta box), are shown in zoomed representation.

FIGURE 5 | Comparison of ligand-bound and ligand-free structures of *Fi*BAP. **(A)** Alignment of ligand-bound and ligand-free *Fi*BAP monomers. Note that the P290–L302 loop is stabilized in the ligand-bound structure but not in the ligand-free model. **(B)** The active site of ligand-free *Fi*BAP containing binuclear zinc metal ions. Residues coordinating the metals are represented as sticks and H-bonds are shown as dotted lines. **(C)** The active site of ligand-bound *Fi*BAP. The 2Fo-Fc electron density maps of Zn^{2+} ions and β-DL are contoured at 1σ and colored blue. Residues of *Fi*BAP interacting with metal ions and β-DL are depicted as sticks. Zn^{2+} ions are represented as purple spheres and the bound β-DL is shown as magenta sticks. Ligands and side chains are colored by atom type (oxygen in red, nitrogen in blue, phosphorous in orange). **(D)** Comparison of residues coordinating the ligand between ligand-free and ligand bound models. Interactions between ligand-bound *Fi*BAP residues and atoms of the ligand are shown as dotted lines. **(E)** Comparison of the ligand-binding residues of *Fi*BAP bound to β-DL with their counterparts in ligand-bound homologs *Ec*IadA (PDB_1YBQ) and *Cp*IadA (PDB_5XGX). Conformational changes in side-chains are indicated. Residues coordinating the ligand are highly conserved among BAPs.

the Cbz-β-DL-bound *Fi*BAP structure aligned with the ligand-free model with an RMSD of 0.433 Å, despite the good superimposition of the structural elements (**Figure 5A**). Such an unusually high RMSD value might be ascribed to the transformation of the missing loop (P290–L302) in the ligand-free structure into two stable antiparallel β-sheets (β14 and β15) in the ligand-bound structure (**Figure 5A**). Since this stabilized loop is located next to the substrate-binding site, it might function as a gate for substrate entry and/or product release (Park et al., 2017). Intriguingly, the crystal structure of *Fi*BAP bound to Cbz-β-DL included distinct electron density only for the β-DL moiety lacking carbamazepine (Cbz) in the active site (**Figure 5C**).

The Active Site of *Fi*BAP

The substrate-binding site of BAPs, referred to as the "binuclear zinc center," is located beneath the (β/α)8 TIM-barrel motif of

the catalytic domain, and comprises two Zn^{2+} (Zn1 and Zn2) ions 3.2 Å apart from each other (dotted circle in **Figure 4A**). Zn1 is coordinated by E156, H195, and H224 residues, while Zn2 is bound by H61, H63, E156, and D285 (**Figure 5B**). In the metal coordination network, Zn1 interacts with Zn2 *via* E156 (**Figure 5B**). By contrast, a carbamylated lysine residue (KCX162) in Type-II *Ec*IadA binds two positively charged Zn ions, similar to E156 in Type-I *Cp*IadA (**Figure 5E**) (Thoden et al., 2003; Park et al., 2017). In addition, E156 in *Fi*BAP forms a *cis*-peptide bond with G155 in the loop between α4 and β10, which is conserved in the corresponding loop of *Cp*IadA between G165 and E166. The *cis*-peptide bond favors loop bending in *Fi*BAP, which positions E156 toward the metal, as opposed to the carbamylated lysine (KCX162) in *Ec*IadA, suggesting that *Fi*BAP belongs to the Type-I IadAs.

There is no significant perturbation in the coordination of the binuclear zinc center in the Cbz-β-DL-bound *Fi*BAP structure
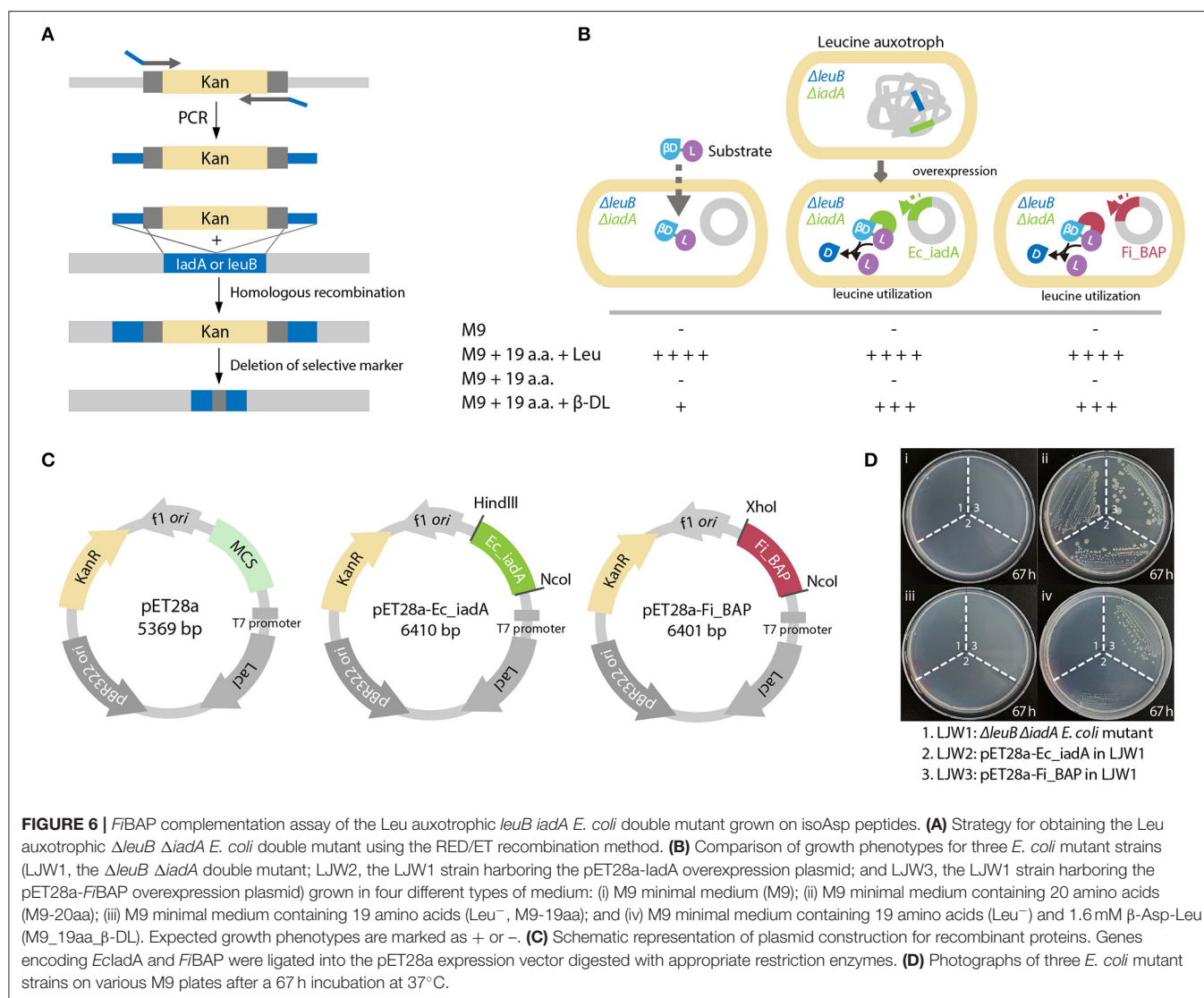
FIGURE 6 | *Fi*BAP complementation assay of the Leu auxotrophic *leuB iadA E. coli* double mutant grown on isoAsp peptides. **(A)** Strategy for obtaining the Leu auxotrophic Δ*leuB* Δ*iadA E. coli* double mutant using the RED/ET recombination method. **(B)** Comparison of growth phenotypes for three *E. coli* mutant strains (LJW1, the Δ*leuB* Δ*iadA* double mutant; LJW2, the LJW1 strain harboring the pET28a-IadA overexpression plasmid; and LJW3, the LJW1 strain harboring the pET28a-*Fi*BAP overexpression plasmid) grown in four different types of medium: (i) M9 minimal medium (M9); (ii) M9 minimal medium containing 20 amino acids (M9-20aa); (iii) M9 minimal medium containing 19 amino acids (Leu⁻, M9-19aa); and (iv) M9 minimal medium containing 19 amino acids (Leu⁻) and 1.6 mM β-Asp-Leu (M9-19aa_β-DL). Expected growth phenotypes are marked as + or −. **(C)** Schematic representation of plasmid construction for recombinant proteins. Genes encoding *Ec*IadA and *Fi*BAP were ligated into the pET28a expression vector digested with appropriate restriction enzymes. **(D)** Photographs of three *E. coli* mutant strains on various M9 plates after a 67 h incubation at 37°C.

(**Figure 5C**). However, structural comparison of the residues around the ligand-binding site reveals subtle changes of ∼1 Å in the side-chain conformations of some residues (G68, E70, Y130, R163, R227, and S289) between Cbz-β-DL-free and -bound forms (**Figure 5D**). Residues such as Y130, G68, and S289 are placed closer to the ligand than those in the ligand-free structure, thereby assisting coordination and stabilization, while R227 and R163 move away to provide space for ligand binding. Notably, the role of the conserved Y130, which interacts with the O06 atom of the substrate, has been studied, and the Y137F mutation in *Ec*IadA (corresponding to Y130 in *Fi*BAP) reduced the rate of catalysis by three orders of magnitude (Marti-Arbona et al., 2005b). Accordingly, the conserved Y130 in *Fi*BAP not only contributes to stabilization of the bound substrate, but also acts as a Lewis acid with the phenolic hydroxyl group of its side chain during hydrolysis of the peptide bond adjacent to Asp of the dipeptide (Marti-Arbona et al., 2005a; Park et al., 2017). The other surrounding residues that help to stabilize

the bound Cbz-β-DL are the amino groups of main-chain residues S289 and G68, while T99 interacts with O04 and O05 atoms of the side-chain of the Asp residue of β-aspartyl leucine to stabilize the N-terminus of the dipeptide during substrate recognition. Additionally, R163 and R227 interact with O01 and O03 atoms and stabilize the C-terminal end of the peptide. When the substrate is recognized and stabilized, other residues such as E156, Y130, and E70 partake in acid-base catalysis during cleavage of the peptide bond via zinc metal ions Zn1 and Zn2 (**Figure 5D**). As stated above, all residues involved in the recognition and catalysis of a substrate are highly conserved among homologs (**Figure 5E**). Intriguingly, the ligand-bound structures of *Ec*IadA and *Cp*IadA were only obtained following mutation of conserved residues involved in metal coordination and catalysis (D285N and E80Q, respectively). By contrast, the structure of Cbz-β-DL-bound *Fi*BAP was determined directly from wild-type *Fi*BAP. This observation may indicate that the crystallization conditions (temperature <20°C and pH < pH 5)
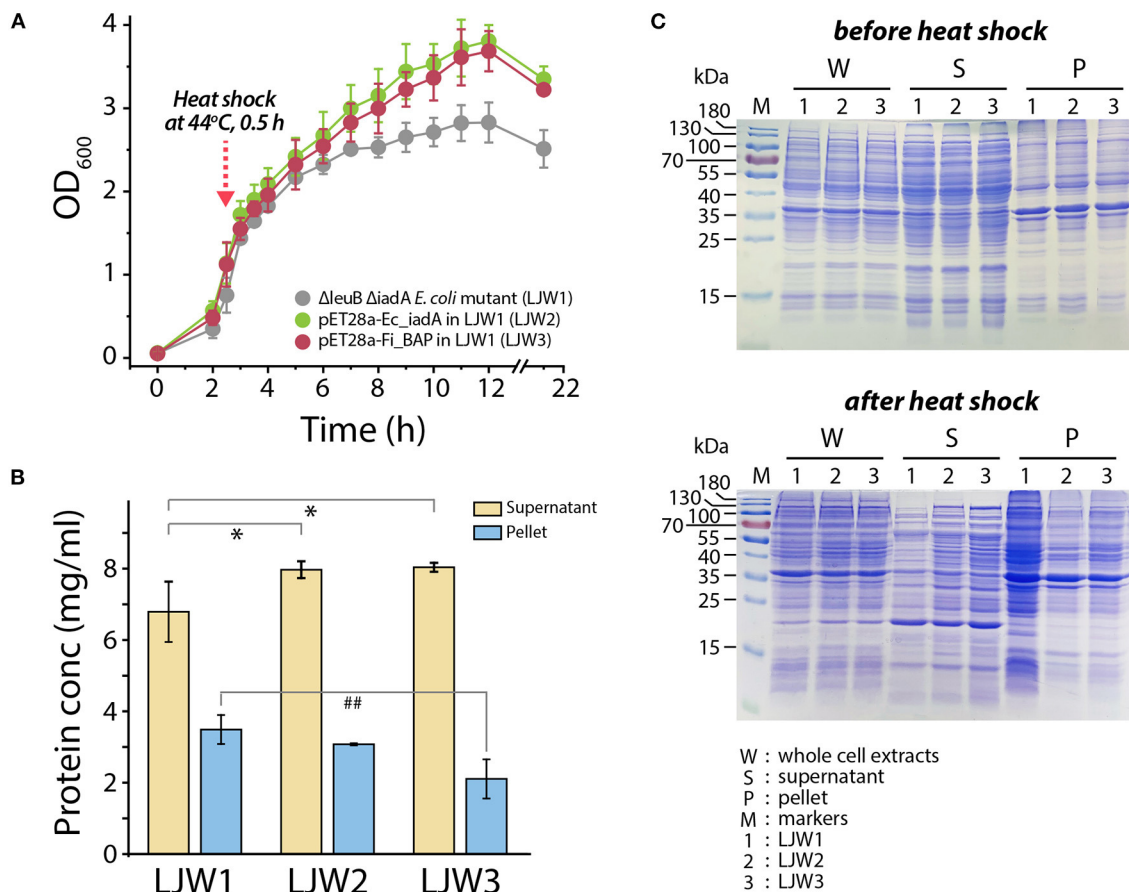
FIGURE 7 | Bacterial growth and protein pattern analysis after heat treatment. **(A)** Growth curve of *E. coli* mutant strains (LJW1, LJW2, and LJW3) in LB media at 37°C. Heat shock treatment was performed at 44°C for 30 min when cells reached the mid-exponential phase (OD$_{600}$ of 1.0–1.5). **(B)** Quantification of soluble (supernatant) and insoluble (pellet) protein in heat-treated cells. Group differences were assessed using one-way analysis of variance (ANOVA), followed by LSD test (*$p < 0.05$ vs. WT supernatant and ##$p < 0.01$ vs. WT pellet). **(C)** Comparison of whole protein, soluble (supernatant) protein, and insoluble (pellet) protein in heat-treated and non-heat-treated cells by SDS-PAGE analysis.

could have inactivated the thermostable *Fi*BAP enzyme, favoring formation of the substrate (β-DL)-bound complex structure without being cleaved, as seen in the electron density map (**Figure 5C**).

Complementation of a Leucine Auxotroph by *Fi*BAP

To validate whether *Fi*BAP can cleave β-DL *in vivo*, we constructed a Leu auxotrophic *E. coli* BL21 (DE3) strain by deleting the *leuB* gene encoding 3-isopropylmalate dehydrogenase and the *iadA* gene encoding BAP in *E. coli* (**Figure 6A**). We presumed that the hydrolysis of β-DL by *Fi*BAP could provide Leu as a nutrient for the Leu auxotrophic *E. coli* mutant, thereby supporting its growth on M9 minimal medium lacking Leu (**Figure 6B**). As expected, the Δ*leuB* and Δ*iadA* double mutant grew in M9 medium supplemented with 20 amino acids, but did not grow in M9 medium containing 19 amino acids without Leu (Leu⁻; **Figure 6D**). Next, to investigate the effect of overexpression of *Fi*BAP and *Ec*IadA on bacterial

growth, we constructed the pET28a-Ec_iadA and pET28a-Fi_BAP vectors (**Figure 6C**) and heterologously expressed the *FiBAP* and *iadA* genes in Δ*leuB* and Δ*iadA* double mutants of *E. coli* BL21 (DE3), respectively. The Δ*leuB* Δ*iadA* mutant did not grow on M9 medium supplemented with 19 amino acids (Leu⁻) in the presence of β-DL during 3 days of incubation, whereas the Δ*leuB* Δ*iadA* double mutant harboring the pET-*iadA* plasmid grew on M9 medium with 19 amino acids (Leu⁻) supplemented with β-DL, as described previously (Marti-Arbona et al., 2005a). Remarkably, the double mutant containing the pET-*Fi*BAP plasmid could also grow on M9 medium with 19 amino acids (Leu⁻) in the presence of β-DL even at a suboptimal temperature for *Fi*BAP (**Figure 6D**). This result strongly indicates that *Fi*BAP can also cleave the β-DL dipeptide to facilitate Leu utilization for bacterial growth, suggesting that *Fi*BAP activity enables *E. coli* to utilize β-aspartyl peptides as nutrient sources for cell growth.

Accordingly, we hypothesized that heat stress may cause accumulation of misfolded and/or aggregated polypeptides, including the formation of isoAsp residues, resulting in cellular
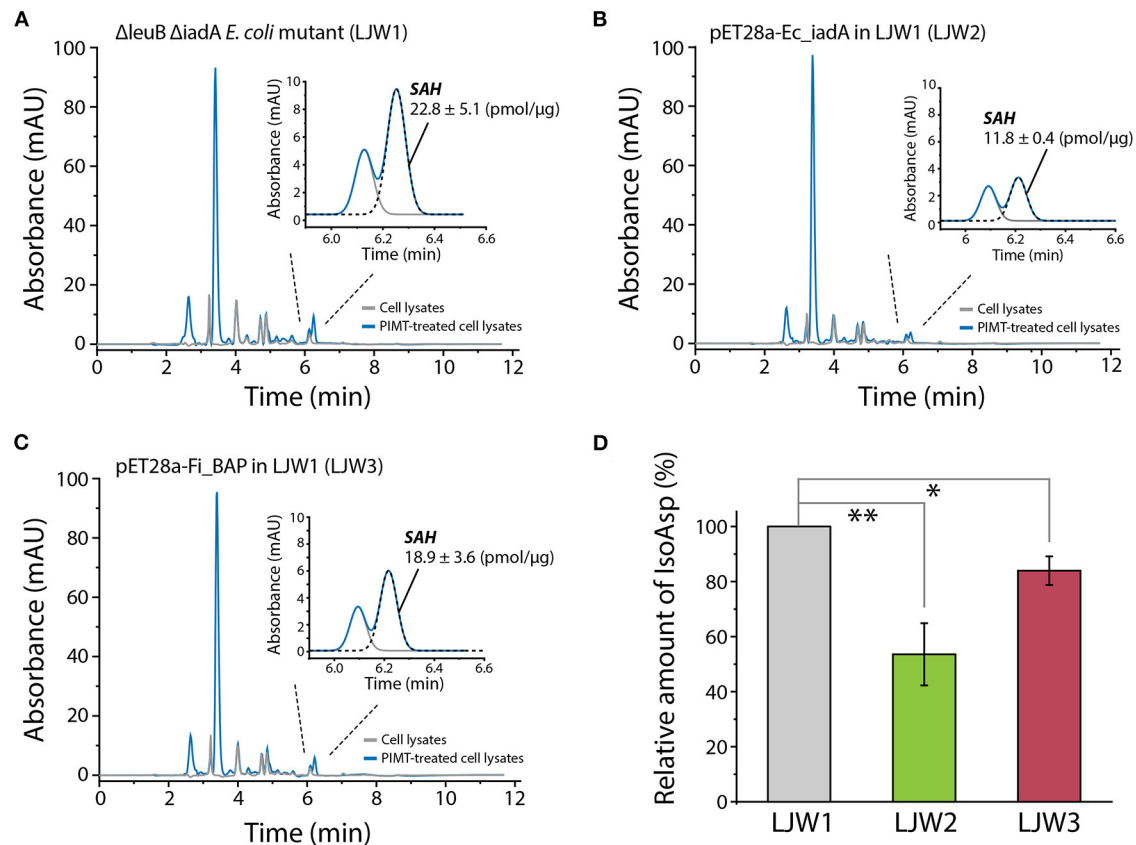
FIGURE 8 | RP-HPLC analysis of SAH for quantification of isoAsp residues in heat-shocked LJW1, LJW2, and LJW3 strains. **(A–C)** RP-HPLC analysis of SAH in LJW1, LJW2, and LJW3 strain and quantification of isoAsp by calculating the SAH peak area. **(D)** Relative amount of isoAsp in each heat- treated cell. Group differences were assessed using one-way analysis of variance (ANOVA), followed by LSD test [$**p < 0.01$ and $*p < 0.05$ vs. WT (LJW1)].

growth retardation. If so, then overexpression of *Fi*BAP may alleviate the bacterial growth defect. To validate this, we monitored the growth profiles of the Δ*leuB* Δ*iadA* double mutant grown on LB medium with and without *Fi*BAP upon heat shock at 44°C. Consequently, after a 30 min heat shock treatment, the growth rate and cellular production of the Δ*leuB* Δ*iadA* mutant were decreased by 30% (based on the OD_{600} value) compared with those of mutants expressing *Fi*BAP and *Ec*IadA, respectively (**Figure 7A**). In addition, when expressing BAP, the Δ*leuB* Δ*iadA* mutant exhibited a 15–40% decrease in the total protein concentrations of pellets, whereas the concentrations of soluble cytosolic proteins were increased by 20% (**Figure 7B**). Furthermore, differential protein pattern analysis by SDS-PAGE clearly indicated that expression of *Fi*BAP could affect the soluble cytosolic protein profiles of the mutant *E. coli* strains, suggesting that abnormal proteins including isoAsp residues could be efficiently degraded by BAP activity under heat stress conditions (**Figure 7C**).

Finally, to investigate whether overexpression of *Fi*BAP and *Ec*IadA can reduce the isoAsp-containing protein aggregates within cells, we determined the concentration of S-adenosyl homocysteine (SAH) as the PIMT-mediated by-product derived from the isoAsp-containing peptides in heat-treated mutant strains. Indeed, RP-HPLC analysis of cell lysates demonstrated that the concentrations of SAH in both mutants expressing *Fi*BAP and *Ec*IadA were lower by 20 and 50%, respectively, than that of the Δ*leuB* Δ*iadA E. coli* mutant (LJW1) negative control (**Figures 8A–C**). These results indicate that the relative amount of isoAsp-containing proteins was decreased significantly in mutants overexpressing *Ec*IadA (LJW2) or *Fi*BAP (LJW3) compared with WT (LJW1) (**Figure 8D**), suggesting that BAP can hydrolyze isoAsp peptide bonds in the protein aggregates, and thereby alleviate the bacterial growth defect under heat shock conditions.

DISCUSSION

Protein folding and misfolding are critical for correct cellular function and regulation. Indeed, compelling evidence suggests that misfolded and damaged protein aggregates contribute to cell dysfunction and tissue damage, leading to various diseases (Dobson, 2003; Goldberg, 2003). For example, misfolding of amyloidogenic proteins is strongly associated with the pathogenicity of neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's diseases, and AA amyloidosis.

Even metabolic syndromes such as type 2 diabetes can be triggered by misfolded protein aggregates (Moreno-Gonzalez and Soto, 2011). It has been reported that up to 30% of newly synthesized proteins may be misfolded and/or aggregated due to errors in translation or post-translational modifications, including the formation of abnormal amino acid residues (Schubert et al., 2000). In particular, unwanted protein aggregates resulting from the formation of isoAsp residues are an example of spontaneous age-related protein damage in cells, and their degradation and/or removal via deamination and dehydration is essential to prevent cellular cytotoxicity and aging (Ogé et al., 2008). To ensure cellular viability in response to starvation, organisms also tend to preserve protein homeostasis through protein repair systems using several strategies such as refolding, degrading, or sequestrating misfolded polypeptides (Goldberg, 1972; Finn and Dice, 2006; Chen et al., 2011).

In extremophiles thriving under harsh environments, in which proteins are vulnerable to protein inactivation and aggregation, cellular protein repair systems play a pivotal role in protein quality control (PQC) to support cellular integrity and survival (De Castro et al., 2006; Chen et al., 2011; Coker, 2019). In the present study, we functionally characterized the hyperthermophilic M38 β-aspartyl peptidase *Fi*BAP, a protein repair enzyme involved in hydrolysis of isoAsp residue-containing peptides in *F. islandicum* AW-1. Since the extremely thermophilic *F. islandicum* AW-1 belongs to the order Thermotogales, an ancient branch of the bacterial kingdom, this bacterium possesses primordial characteristics, making it an excellent model system for investigating how extremophiles have evolved mechanisms to prevent protein damage and/or denaturation under harsh environments (Singleton and Amelunxen, 1973). Although several functional studies on BAPs from mesophilic organisms have been reported, including human (Ryttersgaard et al., 2002), *E. coli* (Marti-Arbona et al., 2005b), and *Colwellia psychrerythraea* (Marti-Arbona et al., 2005b; Park et al., 2017), hyperthermophilic homologs remain uncharacterized.

Phylogenetic analysis based on the primary sequences of its homologs revealed that *Fi*BAP is distantly related to Type II IadA enzymes, suggesting that *Fi*BAP belongs to an ancient form of the BAP family in microorganisms (**Figure 1** and **Supplementary Figure 1**). Our detailed characterization revealed that *Fi*BAP is a hyperthermophilic BAP with a T_m value of 96.8°C in the presence of Zn^{2+} that forms a tetramer of dimers (**Figure 2**), consistent with the first crystal structures of octameric Cbz-β-DL-free and -bound BAPs, and size-exclusion chromatography and symmetry-related molecule analysis (**Figures 2B**, **4C**). The octameric state of *Fi*BAP was compared with the quaternary structures of structural homologs of *Ec*IadA and *Cp*IadA (**Supplementary Figure 3A**), indicating no significant difference in their overall architecture (**Supplementary Figure 3A**). Although the crystal structure of *Fi*BAP superimposed well with those of mesophilic *Ec*IadA and *Cp*IadA (**Figure 4B**), its mesophilic counterparts do not appear to include several of the strong salt bridges (**Supplementary Figure 3B**). Remarkably, the catalytic domain of *Fi*BAP participates in dimerization through salt-bridges via

R143, K114, E107, and D149 from each monomer, and this is largely responsible for the high thermostability (**Figure 4**). This observation is consistent with the study of a hyperthermophilic PIMT, in which mutation of the interfacial residues (D204 and D205) to alanine significantly affected thermostability (Gary and Clarke, 1995). Regarding substrate recognition and ligand stabilization, residues such as G68, T99, Y130, and S289 are highly conserved among all homologs (**Figure 5E**). Notably, residues T99 and S289 appear to be involved in recognition of the N-terminal iso-Asp side chain of the substrate, similar to the roles played by T57 and S59 in protein isoAsp methyltransferase (PIMT) that recognizes abnormal proteins containing isoAsp residues during protein repair (Chen et al., 2011). Consistently, the conserved Y130 residue, which serves the dual purpose of acting as a Lewis acid and supporting ligand stabilization, is also present in PIMT (Y55). Although, the amino acids mediating isoAsp recognition appear to be conserved with PIMT, the role of the binuclear zinc center in BAPs in cleaving the peptide bond is not conserved in PIMT. Taken together, the results indicate that *Fi*BAP is a thermostable alternative protein repair enzyme that possesses conserved residues for isoAsp residue recognition and for efficient Zn^{2+}-mediated peptide bond hydrolysis.

In contrast to the crystal structures of BAP homologs inactivated by mutations (Thoden et al., 2003; Park et al., 2017), the crystal structure of the ligand-bound form of *Fi*BAP was obtained without any significant cleavage of bound β-DL (**Figures 5A,C,E**). This could be explained by various reasons. Firstly, crystallization was carried out at below room temperature (20°C) in a low-pH buffer (pH < 5.0), which could have affected the enzyme activity (**Figures 2C,D**). Secondly, we synthesized Cbz-β-DL dipeptide as a substrate analog inhibitor in which benzyl chloroformate (Cbz) is a carbobenzyloxy protecting group that protects the substrate amino group. Previous studies have shown clearly that Cbz acts as an inhibitor of several enzymes systems including human alpha-thrombin (De Simone et al., 1997) and sortase cysteine transpeptidase (Jacobitz et al., 2014). Consistently, our substrate analog (Cbz-β-DL) containing Cbz inhibited enzyme activity in our enzyme assays (data not shown), which suggests that wt *Fi*BAP did not cleave the Cbz-β-DL dipeptide. Therefore, these modifications worked in our favor to obtain the crystal structure of the wild-type enzyme complexed with its substrate. However, the fate of Cbz in the substrate-bound enzyme is ambiguous due to the absence of electron density for the Cbz moiety in the structure. Interestingly, the location of an uncleaved β-DL based on the electron density map (**Figure 5C**) coincides precisely with the dipeptide (β-Asp-Lys or His) location described in previous studies (Thoden et al., 2003; Marti-Arbona et al., 2005a; Park et al., 2017) (**Figure 5E**). Thus, it seems probable that the Cbz entity may be cleaved by hydrogenolysis before the N-terminal end of the dipeptide enters the substrate-binding cavity, since the benzyl ring of Cbz (if uncleaved) might cause steric hindrance within the cavity. Furthermore, the second amino acid following Asp in the dipeptide (Leu in β-DL) is unlikely to interact with any surrounding residues of *Fi*BAP (**Figures 5D,E**).

The primordial *Fi*BAP is a unique metallopeptidase that is presumably involved in the sequestration of unwanted

forms of peptides. Notably, *Fi*BAP is the most abundant and highly expressed protein when *F. islandicum* AW-1 is grown on feather keratins under starvation conditions (Kang et al., 2020). Although this enzyme is not one of the major endo-type membrane proteases, but rather a protease involved in keratin degradation, its physiological role seems to be important for cellular survival due to the formation of isopeptides that contribute to thermostability, as observed in other anaerobic pathogens (Liu et al., 2016), as well as a providing a source of amino acids for new protein synthesis (Prouty and Goldberg, 1972). The formation of isoAsp residues has received much attention because it is a major structural modification that contributes to the inactivation, aggregation, and malfunction of proteins (Aswad, 1995; Cournoyer et al., 2005). Spontaneous isopeptide bond formation is accelerated at elevated temperatures, and intermolecular amide bonds help proteins to be more thermostable (Zakeri and Howarth, 2010; Si et al., 2016). However, such chemical modifications are not always favorable for cellular viability due to the accumulation of potentially harmful isoAsp peptides in the brain and other tissues (Cantor et al., 2009). Furthermore, balance between the formation and decomposition of isopeptide bonds via isoAsp peptidase is crucial to bacterial survival. Indeed, the growth profiles of the $\Delta leuB$ $\Delta iadA$ double mutant of *E. coli* BL21 clearly indicated that heat shock for 30 min significantly retarded bacterial growth, presumably due to an increase in the accumulation of abnormal proteins (**Figure 7A**). Notably, expression of *Fi*BAP, even at suboptimal temperatures ranging from 37 to 44°C, may be sufficient to support bacterial growth through efficient degradation of isoAsp residue-containing proteins. Although the amount of IsoAsp in mutant expressing *Fi*BAP was not decreased as much as in mutant expressing *Ec*IadA, levels of alleviation of growth defect were similar. This result suggests that enzymatic activity of *Fi*BAP is not limited toward isoAsp containing peptides but also includes a broad range of protein aggregates which is related to the interpretation of our MS data (**Figures 3, 7B,C, 8**).

Unlike BAPs, which are mainly present in extremophiles and several enteric bacteria, PIMT is a conserved and nearly ubiquitous enzyme present in all forms of life that catalyzes the transfer of the methyl group from AdoMet to α-carboxyl of an isoAsp site, resulting in the rapid decomposition of methyl esters to succinimide intermediates (Skinner et al., 2000). Therefore, like isoAsp peptidase, PIMT is a useful enzyme for repairing isoAsp damage in all organisms. Nevertheless, when decomposition of isopeptide bonds is a prerequisite for balancing nutrient availability and thermostability under extreme environments, PIMT is unlikely to be an efficient catalyst in terms of the rate of isopeptide decomposition and formation. Indeed, the formation of succinimide by PIMT resulting from isoAsp-containing peptides generates a mixture of isoAsp-Xaa and Asp-Xaa with a stoichiometric ratio of 7:3 (Aswad et al., 2000). However, BAP can directly hydrolyze isopeptide bonds via the β-carboxylate group of isoAsp, resulting in the release of Asp (Marti-Arbona et al., 2005a), and this might be a more efficient

repair mechanism for handing β-aspartyl-containing peptides. Furthermore, proteolytic pathways function during starvation to liberate amino acids, and this may be beneficial for dietary protein digestion and cellular protein turnover (Finn and Dice, 2006).

Regarding the functional and physiological roles of *Fi*BAP, we offer two suggestions explaining why this enzyme is conserved in all extremely thermophilic anaerobes, and why it is so abundant in *F. islandicum* AW-1 during keratin degradation. Firstly, it is one of the key enzymes involved in maintaining proteostasis in thermophiles thriving in harsh environments. By directly hydrolyzing peptide bonds in isoAsp-containing peptides, it can efficiently control aggregated, modified, or damaged proteins, thereby supporting the viability of thermophiles at elevated temperatures. Secondly, under starvation conditions, when insoluble keratins are the only nutrient available for cellular growth, this enzyme may utilize β-aspartyl peptides as a source of amino acids for protein biosynthesis. This hypothesis is supported by previous studies showing that *Fi*BAP is expressed at high levels under starvation conditions (Kang et al., 2020), suggesting that nutrient depletion acts as a stress condition that stimulates *Fi*BAP expression and enzyme activity. Furthermore, peptidase-deficient *E. coli* leucine auxotrophic strains did not grow in M9 minimal media containing 19 amino acids (Leu⁻) and β-DL. However, expression of *Ec*IadA or *Fi*BAP enabled the *E. coli* mutant to grow in this medium, implying that *Fi*BAP contribute to increasing the availability of nutrients (Viola, 2001), thereby assisting the maintenance of cellular growth under stressful conditions (**Figure 6**). Thus, *Fi*BAP may not be a membrane bound endo- or exo-type keratinolytic protease (Kang et al., 2020), but it appears to be strongly associated with the degradation of abnormal proteins as well as intracellular keratin peptides to provide a source of amino acids for protein biosynthesis under stress conditions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JL, ID, HJ, SL, and D-WL formulated the research plan, carried out experiments, analyzed and interpreted the data, and drafted the manuscript. JL, ID, SL, and D-WL participated in the design of the study and analyzed and interpreted the data. SL and D-WL conceived, planned, and supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66 (Pt 2), 213–221. doi: 10.1107/S0907444909052925

Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P., and Adams, P. D. (2010). Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 11), 1153–1163. doi: 10.1107/S0907444910026582

Aswad, D. W. (1995). *Deamidation and Isoaspartate Formation in Peptides and Proteins.* Boca Raton, FL: CRC Press.

Aswad, D. W., Paranandi, M. V., and Schurter, B. T. (2000). Isoaspartate in peptides and proteins: formation, significance, and analysis. *J. Pharm. Biomed. Anal.* 21, 1129–1136. doi: 10.1016/S0731-7085(99)00230-7

Cantor, J. R., Stone, E. M., Chantranupong, L., and Georgiou, G. (2009). The human asparaginase-like protein 1 hASRGL1 is an Ntn hydrolase with beta-aspartyl peptidase activity. *Biochemistry* 48, 11026–11031. doi: 10.1021/bi901397h

Chen, B., Retzlaff, M., Roos, T., and Frydman, J. (2011). Cellular strategies of protein quality control. *CSH Perspect. Biol.* 3:a004374. doi: 10.1101/cshperspect.a004374

Coker, J. A. (2019). Recent advances in understanding extremophiles. *F1000Research* 8:F1000 Faculty Rev-1917. doi: 10.12688/f1000research.20765.1

Cournoyer, J. J., Pittman, J. L., Ivleva, V. B., E., Fallows, L., Waskell, Costello, C. E., et al. (2005). Deamidation: differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Sci.* 14, 452–463. doi: 10.1110/ps.041062905

De Castro, R. E., Maupin-Furlow, J. A., Giménez, M. I., Herrera Seitz, M. K., and Sánchez, J. J. (2006). Haloarchaeal proteases and proteolytic systems. *FEMS Microbiol. Rev.* 30, 17–35. doi: 10.1111/j.1574-6976.2005.00003.x

De Simone, G., Balliano, G., Milla, G.allina, C., and Giordano, C., Tarricone, C., et al. (1997). Human α-thrombin inhibition by the highly selective compounds N-ethoxycarbonyl-D-Phe-Pro-α-azaLys p-nitrophenyl ester and N-carbobenzoxy-pro-α-azaLys p-nitrophenyl ester: a kinetic, thermodynamic and x-ray crystallographic study. *J. Mol. Biol.* 269, 558–569. doi: 10.1006/jmbi.1997.1037

Dobson, C. M. (2003). Protein folding and misfolding. *Nature* 426, 884–890. doi: 10.1038/nature02261

Dorer, F. E., Haley, E. E., and Buchanan, D. L. (1968). The hydrolysis of beta-aspartyl peptides by rat tissue. *Arch. Biochem. Biophys.* 127, 490–495. doi: 10.1016/0003-9861(68)90253-1

Elias, J. E., and Gygi, S. P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* 4, 207–214. doi: 10.1038/nmeth1019

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 12 Pt 1), 2126–2132. doi: 10.1107/S0907444904019158

Finn, P. F., and Dice, J. F. (2006). Proteolytic and lipolytic responses to starvation. *Nutrition* 22, 830–844. doi: 10.1016/j.nut.2006.04.008

Gary, J. D., and Clarke, S. (1995). Purification and characterization of an isoaspartyl dipeptidase from *Escherichia coli. J. Biol. Chem.* 270, 4076–4087. doi: 10.1074/jbc.270.8.4076

Goldberg, A. L. (1972). Degradation of abnormal proteins in *Escherichia coli* (protein breakdown-protein structure-mistranslation-amino acid analogs-puromycin). *Proc. Nat. Acad. Sci. U. S. A.* 69, 422–426. doi: 10.1073/pnas.69.2.422

Goldberg, A. L. (2003). Protein degradation and protection against misfolded or damaged proteins. *Nature* 426, 895–899. doi: 10.1038/nature02263

Hipp, M. S., Kasturi, P., and Hartl, F. U. (2019). The proteostasis network and its decline in ageing. *Nat. Rev. Mol. Cell Biol.* 20, 421–435. doi: 10.1038/s41580-019-0101-y

Hutt, D., and Balch, W. E. (2010). Cell Biology. The proteome in balance. *Science* 329, 766–767. doi: 10.1126/science.1194160

Ichikawa, J. K., and Clarke, S. (1998). A highly active protein repair enzyme from an extreme thermophile: the L-isoaspartyl methyltransferase from *Thermotoga maritima. Arch. Biochem. Biophys.* 358, 222–231. doi: 10.1006/abbi.1998.0830

Jacobitz, A. W., Wereszczynski, J., Yi, S. W., Amer, B. R., Huang, G. L., Nguyen, A. V., et al. (2014). Structural and computational studies of the *Staphylococcus aureus* sortase B-substrate complex reveal a substrate-stabilized oxyanion hole. *J. Biol. Chem.* 289, 8891–8902. doi: 10.1074/jbc.M113.509273

Jin, H.-S., Park, S. Y., Kim, K., Lee, Y.-J., Nam, G.-W., Kang, N. J., et al. (2017). Development of a keratinase activity assay using recombinant chicken feather keratin substrates. *PLoS ONE* 12:e0172712. doi: 10.1371/journal.pone.0172712

Kang, E., Jin, H. S., La, J. W., Sung, J. Y., Park, S. Y., Kim, W. C., et al. (2020). Identification of keratinases from *Fervidobacterium islandicum* AW-1 using dynamic gene expression profiling. *Microb. Biotechnol.* 13, 442–457. doi: 10.1111/1751-7915.13493

Kim, E., Lowenson, J. D., MacLaren, D. C., Clarke, S., and Young, S. G. (1997). Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice. *Proc. Nat. Acad. Sci. U. S. A.* 94, 6132–6137. doi: 10.1073/pnas.94.12.6132

Lee, Y. J., Dhanasingh, I., Ahn, J.-S., Jin, H.-S., Choi, J. M., Lee, S. H., et al. (2015b). Biochemical and structural characterization of a keratin-degrading M32 carboxypeptidase from *Fervidobacterium islandicum* AW-1. *Biochem. Bioph. Res. Commun.* 468, 927–933. doi: 10.1016/j.bbrc.2015.11.058

Lee, Y. J., Jeong, H., Park, G. S., Kwak, Y., Lee, S. J., Lee, S. J., et al. (2015a). Genome sequence of a native-feather degrading extremely thermophilic eubacterium, *Fervidobacterium islandicum* AW-1. *Stand. Genomic Sci.* 10:71. doi: 10.1186/s40793-015-0063-4

Liu, H., Ray, W. K., Helm, R. F., Popham, D. L., and Melville, S. B. (2016). Analysis of the spore membrane proteome in *Clostridium perfringens* implicates cyanophycin in spore assembly. *J. Bacteriol.* 198, 1773–1782. doi: 10.1128/JB.00212-16

Marti-Arbona, R., Fresquet, V., Thoden, J. B., Davis, M. L., Holden, H. M., and Raushel, F. M. (2005a). Mechanism of the reaction catalyzed by isoaspartyl dipeptidase from *Escherichia coli. Biochemistry* 44, 7115–7124. doi: 10.1021/bi050008r

Marti-Arbona, R., Thoden, J. B., Holden, H. M., and Raushel, F. M. (2005b). Functional significance of Glu-77 and Tyr-137 within the active site of isoaspartyl dipeptidase. *Bioorg. Chem.* 33, 448–458. doi: 10.1016/j.bioorg.2005.10.002

McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* 63(Pt 1), 32–41. doi: 10.1107/S0907444906045975

Mogk, A., Huber, D., and Bukau, B. (2011). Integrating protein homeostasis strategies in prokaryotes. *CSH Perspect. Biol.* 3:a004366. doi: 10.1101/cshperspect.a004366

Moreno-Gonzalez, I., and Soto, C. (2011). Misfolded protein aggregates: mechanisms, structures and potential for disease transmission. *Semin. Cell. Dev. Biol.* 22, 482–487. doi: 10.1016/j.semcdb.2011.04.002

Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009). Electronic ligand builder and optimization workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* 65(Pt 10), 1074–1080. doi: 10.1107/S0907444909029436

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., et al. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 67(Pt 4), 355–67. doi: 10.1107/S0907444911001314

Nam, G. W., Lee, D. W., Lee, H. S., Lee, N. J., Kim, B. C., Choe, E. A., et al. (2002). Native-feather degradation by Fervidobacterium islandicum AW-1, a newly isolated keratinase-producing thermophilic anaerobe. *Arch. Microbiol.* 178, 538–547. doi: 10.1007/s00203-002-0489-0

Ogé, L., Bourdais, G., Bove, J., Collet, B., Godin, B., Granier, F., et al. (2008). Protein repair L-isoaspartyl methyltransferase1 is involved in both seed longevity and germination vigor in *Arabidopsis*. *Plant Cell* 20:3022. doi: 10.1105/tpc.108.058479

Park, S. H., Lee, C. W., Lee, S. G., Shin, S. C., Kim, H. J., Park, H., et al. (2017). Crystal structure and functional characterization of an isoaspartyl dipeptidase (CpsIadA) from *Colwellia psychrerythraea* strain 34H. *PLoS ONE* 12:e0181705. doi: 10.1371/journal.pone.0181705

Prouty, W. F., and Goldberg, A. L. (1972). Fate of abnormal proteins in *E. coli* accumulation in intracellular granules before catabolism. *Nat. New Biol.* 240, 147–150. doi: 10.1038/newbio240147a0

Reissner, K. J., and Aswad, D. W. (2003). Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals? *Cell Mol. Life. Sci.* 60, 1281–1295. doi: 10.1007/s00018-003-2287-5

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–4. doi: 10.1093/nar/gku316

Rosen, H. (1957). A modified ninhydrin colorimetric analysis for amino acids. *Arch. Biochem. Biophys.* 67, 10–15. doi: 10.1016/0003-9861(57)90241-2

Ryttersgaard, C., Griffith, S. C., Sawaya, M. R., MacLaren, D. C., Clarke, S., and Yeates, T. O. (2002). Crystal structure of human L-isoaspartyl methyltransferase. *J. Biol. Chem.* 277, 10642–10646. doi: 10.1074/jbc.M200229200

Schubert, U., Antón, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bennink, J. R. (2000). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 404, 770–774. doi: 10.1038/35008096

Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., et al. (2007). The Paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteom.* 6:1638. doi: 10.1074/mcp.T600050-MCP200

Si, M., Xu, Q., Jiang, L., and Huang, H. (2016). SpyTag/SpyCatcher cyclization enhances the thermostability of firefly luciferase. *PLoS ONE* 11:e0162318. doi: 10.1371/journal.pone.0162318

Singleton, R. Jr., and Amelunxen, R. E. (1973). Proteins from thermophilic microorganisms. *Bacteriol. Rev.* 37, 320–342. doi: 10.1128/MMBR.37.3.320-342.1973

Skinner, M. M., Puvathingal, J. M., Walter, R. L., and Friedman, A. M. (2000). Crystal structure of protein isoaspartyl methyltransferase: a catalyst for protein repair. *Structure* 8, 1189–1201. doi: 10.1016/S0969-2126(00)00522-0

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., et al. (1985). Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150, 76–85. doi: 10.1016/0003-2697(85)90442-7

Tanaka, Y., Tsumoto, K., Yasutake, Y., Umetsu, M., Yao, M., Fukada, H., et al. (2004). How oligomerization contributes to the thermostability of an archaeon protein. Protein L-isoaspartyl-O-methyltransferase from *Sulfolobus tokodaii*. *J. Biol. Chem.* 279, 32957–32967. doi: 10.1074/jbc.M404405200

Thoden, J. B., Marti-Arbona, R., Raushel, F. M., and Holden, H. M. (2003). High-resolution X-ray structure of isoaspartyl dipeptidase from *Escherichia coli*. *Biochemistry* 42, 4874–4882. doi: 10.1021/bi034233p

Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 1), 22–25. doi: 10.1107/S0907444909042589

Viola, R. E. (2001). The central enzymes of the aspartate family of amino acid biosynthesis. *Acc. Chem. Res.* 34, 339–349. doi: 10.1021/ar000057q

Visick, J. E., and Clarke, S. (1995). Repair, refold, recycle: how bacteria can deal with spontaneous and environmental damage to proteins. *Mol. Microbiol.* 16, 835–845. doi: 10.1111/j.1365-2958.1995.tb02311.x

Zakeri, B., and Howarth, M. (2010). Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *J. Am. Chem. Soc.* 132, 4526–4527. doi: 10.1021/ja910795a

Antagonistic Control of Genetic Circuit Performance for Rapid Analysis of Targeted Enzyme Activity in Living Cells

Kil Koang Kwon[1], Haseong Kim[1,2], Soo-Jin Yeom[3], Eugene Rha[1], Jinju Lee[1,2], Hyewon Lee[1], Dae-Hee Lee[1,2] and Seung-Goo Lee[1,2*]

[1] Synthetic Biology and Bioengineering Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, South Korea, [2] Department of Biosystems and Bioengineering, KRIBB School of Biotechnology, University of Science and Technology (UST), Daejeon, South Korea, [3] School of Biological Sciences and Technology, Chonnam National University, Gwangju, South Korea

Genetic circuits have been developed for quantitative measurement of enzyme activity, metabolic engineering of strain development, and dynamic regulation of microbial cells. A genetic circuit consists of several bio-elements, including enzymes and regulatory cassettes, that can generate the desired output signal, which is then used as a precise criterion for enzyme screening and engineering. Antagonists and inhibitors are small molecules with inhibitory effects on regulators and enzymes, respectively. In this study, an antagonist and an inhibitor were applied to a genetic circuit for a dynamic detection range. We developed a genetic circuit relying on regulators and enzymes, allowing for straightforward control of its output signal without additional genetic modification. We used *para*-nitrophenol and alanine as an antagonist of DmpR and inhibitor of tyrosine phenol-lyase, respectively. We show that the antagonist resets the detection range of the genetic circuit similarly to a resistor in an electrical logic circuit. These biological resistors in genetic circuits can be used as a rapid and precise controller of variable outputs with minimal circuit configuration.

Keywords: inhibitor, antagonist, genetic circuit, phenolic compound, flow cytometry, resistor

INTRODUCTION

Designing a genetic circuit is a crucial step for programming of living organisms, a long-term aim of synthetic biology (Brophy and Voigt, 2014; Nielsen et al., 2016). Transcriptional regulators in the circuits can control downstream gene expression by recognizing and transmitting specific input signals (Eggeling et al., 2015; Mahr and Frunzke, 2016). Applications include metabolic flux regulation in metabolic engineering (Rogers et al., 2015) and circuit-based high-throughput enzyme screening (Choi et al., 2014). Of these, genetic circuits for screening enzymes have greatly evolved over the last decade by employing various transcriptional regulators including amidase (BenR), alcohol dehydrogenase (SoxR), phosphatase (DmpR), and lactam synthase (NitR) (Uchiyama and Miyazaki, 2010; Choi et al., 2014; Siedler et al., 2014; Yeom et al., 2018).

Fine tuning of gene expression is also necessary to enable complex and precise control of biological reactions. This fine tuning can be achieved by adjusting the output signal range of the genetic elements (Brophy and Voigt, 2014; Smanski et al., 2014). Despite significant advances in the engineering of tools for genetics, the development of efficient genetic circuits remains a time-consuming and labor-intensive process involving repetitive trial and error. In addition, synthetic circuits offset their sensitivity to balance the expression of regulator and reporter genes, which can lead to the loss of flexibility under various environmental conditions (Hausser et al., 2019). A highly sensitive genetic circuit is advantageous for detecting minute clues regarding enzyme activity from a natural sample or metagenome (Ngara and Zhang, 2018; Markel et al., 2020). However, once a genetic circuit reaches its maximum dynamic range for detecting the enhanced activity of engineered enzymes, it is essential to use another module with an appropriate operational range to further engineer enzyme activity. Therefore, like variable resistors in electric circuits, genetic circuit resistors (GCRs) should be developed to control the dynamic range of genetic circuits.

Inhibitors in enzymatic reactions and antagonists in allosteric proteins, such as transcriptional regulators, are molecules that act as negative feedback elements for each activity (Dixon, 1953; Kenakin, 2007). These molecules can bind to the active site of proteins and decrease their activity by interfering with enzyme-substrate or regulator-ligand complex formation. This property can be used to control the sensitivity and dynamic range of a genetic circuit by simply adding small molecules to the reaction system (Xie et al., 2014). By adjusting the concentration of the small molecules, they can be used as variable resistors for tunable signal production. These biological resistors may be implemented for switchable and precise monitoring of enzyme activity, while the genetic circuit maintains hypersensitivity without additional genetic modification.

We previously reported a genetic enzyme screening system (GESS) consisting of DmpR as a phenol-dependent transcriptional regulator and green fluorescent protein (GFP) as a reporter protein (Choi et al., 2014; Kwon et al., 2020). The quantitative range of a genetic circuit relying on an enzyme and regulator could be adjusted by using an enzyme-inhibitor pair and regulator-antagonist pair. As a proof of concept, DmpR-GESS and tyrosine phenol-lyase (TPL), which produces phenol from a tyrosine substrate, were used to model a genetic circuit and enzyme. Alanine and para-nitrophenol (pNP) were used as the enzyme inhibitor and antagonist of DmpR, respectively. We controlled the output of the genetic circuit using elements related to the enzyme and regulator without genetic modification. In various applications, such as enzyme evolution, the circuit could be controlled using bio-parts and related inhibitors or antagonists as GCRs.

MATERIALS AND METHODS

Materials
All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA). DNA polymerase and Gibson assembly kits were purchased from New England Biolabs (Ipswich, MA, USA). All oligonucleotides were synthesized by Macrogen (Daejeon, Korea). Plasmid DNA isolation and DNA extraction from agarose gels were performed using Qiagen kits (Hilden, Germany). DNA preparation and related techniques were performed according to the manufacturer's protocols.

Strains and Plasmids
Escherichia coli DH5α was used for cloning and genetic circuit experiments. Plasmids, pDmpR-GESS and pmDmpR-GESS were obtained from previous studies (Choi et al., 2014). The TPL gene from *Citrobacter freundii* and pAR plasmids (Kim et al., 2017) were amplified by PCR (TPL forward primer: 5′-TCA GCA GGA TCA CCA TAT GAA TTA TCC GGC AGA-3′, TPL reverse primer: 5′-TTG CGT TGC GCT TAG CTT TAG ATA TAG TCA AAG C-3′, pAR forward primer: 5′-GCT TTG ACT ATA TCT AAA GCT AAG CGC AAC GCA A-3′, pAR reverse primer: 5′-TCT GCC GGA TAA TTC ATA TGG TGA TCC TGC TGA A-3′). DNA fragments purified by agarose gel elution were ligated by Gibson assembly, and then transformed into DH5α cells to construct the pAR-TPL plasmid.

Analysis of Regulator-Antagonist Output Signal
Cells harboring pDmpR-GESS or pmDmpR-GESS were cultivated in lysogeny broth (LB) medium (10 g tryptone, 5 g yeast extract, and 5 g NaCl per liter) and M9 minimal medium (12.8 g $Na_2HPO_4 \cdot 7H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 0.01% (w/v) thiamine per later) supplemented with 4 g/L acetate as a carbon source and 50 μg/mL ampicillin. For the two-step phenol reaction, the cells were grown in LB at 37°C until an OD_{600} of 2.0 was reached, and then the culture media was changed to fresh M9 with 1 mM aromatic compounds and various concentrations of phenol by mild centrifugation (1,000 × g, 5 min) (Kwon et al., 2020). After 15 h of incubation at 37°C, the fluorescence intensities of the cells were measured using a FACSAriaIII (BD Biosciences, Franklin Lakes, NJ, USA) with a blue laser source (488 nm) and an FL1 (530/30 nm) photomultiplier tube. Data were acquired using BD CellQuest Pro (version 4.0.2, BD Biosciences) and analyzed using Flowjo software (Flowjo, Ashland, OR, USA).

To examine the antagonistic effect of pNP, cells harboring pDmpR-GESS were cultured in a two-step reaction in the presence of various concentrations of phenol and pNP. Fluorescence intensity and the optical density at 600 nm (OD_{600}) were analyzed with a multi-label reader (Victor V, PerkinElmer, Waltham, MA, USA).

Analysis of Enzyme Inhibitor Output Signals
To detect enzymatic activity related to inhibition, cells harboring DmpR-GESS and the TPL gene were cultured in a two-step reaction. After replacing the M9 media with 1 mM tyrosine, 10 μM pyridoxal 5′-phosphate (PLP), and various concentrations of alanine, cell growth, and fluorescence intensities were measured with a FACSAriaIII and an Infinite 200 PRO microplate reader (Tecan, Männedorf, Switzerland).

To detect enzymatic activity arising from the antagonistic effect of *p*NP, cells harboring pDmpR-GESS and the TPL gene were grown in LB with 20 μM L-rhamnose, 10 μM PLP, 50 μg/mL ampicillin, and 25 μg/mL chloramphenicol. After replacing the M9 media with 1 mM tyrosine, 10 μM PLP, and various concentrations of *p*NP, cell growth, and fluorescence intensity were measured with a microplate reader.

To detect the inhibitory effect of alanine in the solid phase, cells harboring TPL and the genetic circuit were cultured on LB agar plates with 1 mM tyrosine, 10 μM PLP, and 1 mM alanine at 37°C for 20 h. Fluorescence images were acquired using a fluorescence microscope (AZ100M, Nikon, Tokyo, Japan) with epifluorescence and diascopic DIC accessories. Images were acquired with a monochrome CCD camera (DS-Qi1Mc, Nikon) using a fluorescence filter set (GFP-HQ, Nikon) (Ex 455–485 nm, DM 495, BA 500–545). Images were processed and analyzed using Nikon's NIS-Elements AR 4.2 software.

RESULTS AND DISCUSSION

Antagonistic Ligand-Dependent Sensitivity Control in a Genetic Circuit

A phenol-responsive genetic circuit used as a DmpR-GESS consists of a regulator and ligand that form an AND logic gate. An AND logic gate recognizes two inputs simultaneously (Tabor et al., 2009). Our circuit is composed of two AND logic gates: the first logic gate has two inputs, the enzyme and its substrate, and the subsequent logic gate uses the output from the first AND gate and regulator as its inputs. Addition of an enzyme inhibitor and a regulator antagonist can act as resistors for the two AND gates. **Figure 1A** shows the strategy of the association between the antagonist and regulator in the genetic circuit. The antagonistic ligand, which binds to DmpR and inhibits transcriptional initiation, can be used as a GCR to suppress the output signal. In addition, the GCR can modulate the dynamic detection range of the genetic circuit by simple addition without genetic modification (**Figure 1B**).

Wild-type (WT) DmpR responds to phenol, triggering downstream expression of reporter genes. However, *p*NP does not affect WT DmpR activity, but only induces transcriptional initiation when a single point mutation of E135K is inserted into the WT DmpR (**Supplementary Figure 1**; Choi et al., 2014). The crystal structure of DmpR was recently solved and was found to possess a binding pocket containing Glu135 at a distance from the phenol binding site (**Supplementary Figure 2**; Park et al., 2020). The E135K mutation of DmpR may affect the position of Arg36, thus inducing a conformational change in the sensory domain dimer and activating transcriptional initiation. As *p*NP can bind to WT DmpR naturally without inducing transcriptional initiation, it may act as an antagonist.

In previous studies, *ortho*- substituted phenolic compounds were shown to strongly induce DmpR-based transcriptional initiation (O'Neill et al., 1999; Choi et al., 2014). In contrast, some *meta*- or *para*- substituted phenolic compounds do not activate DmpR, even when inhibiting its ATPase activity. In this experiment, benzene, 3,5-dimethylphenol, 2,4-dichlorophenol,
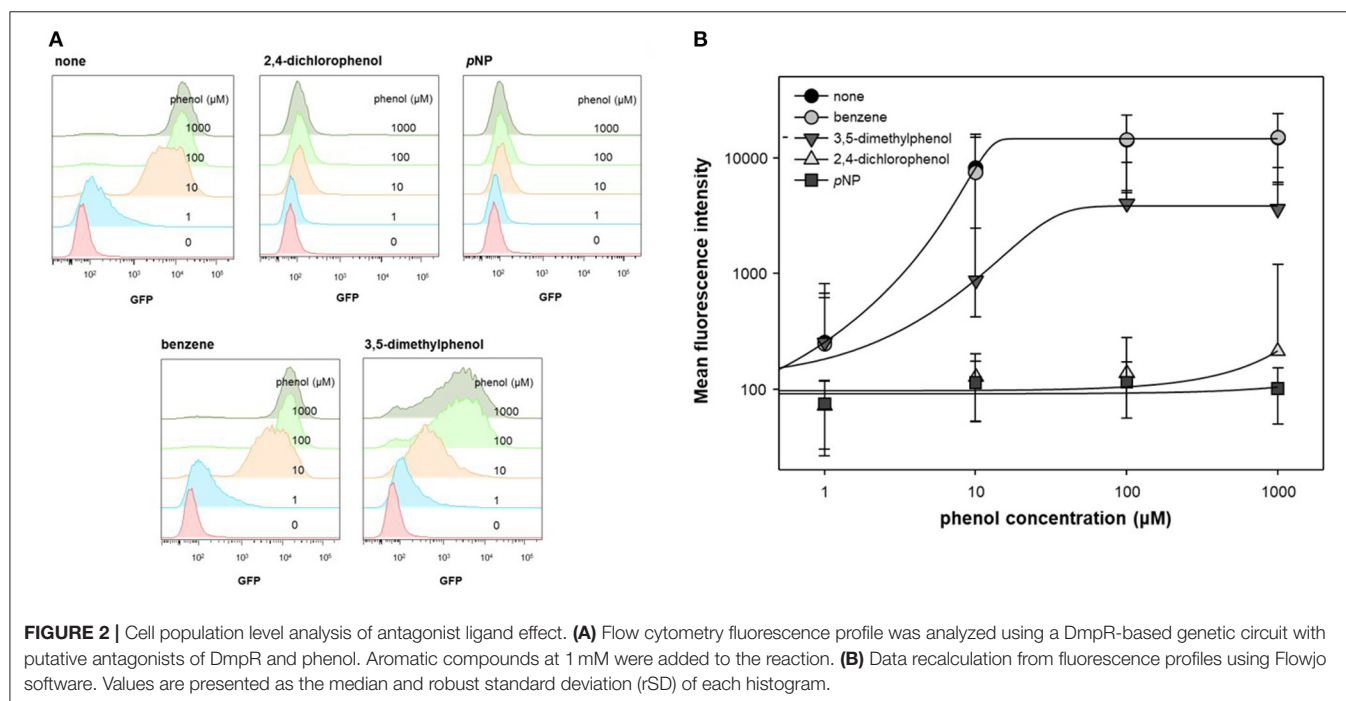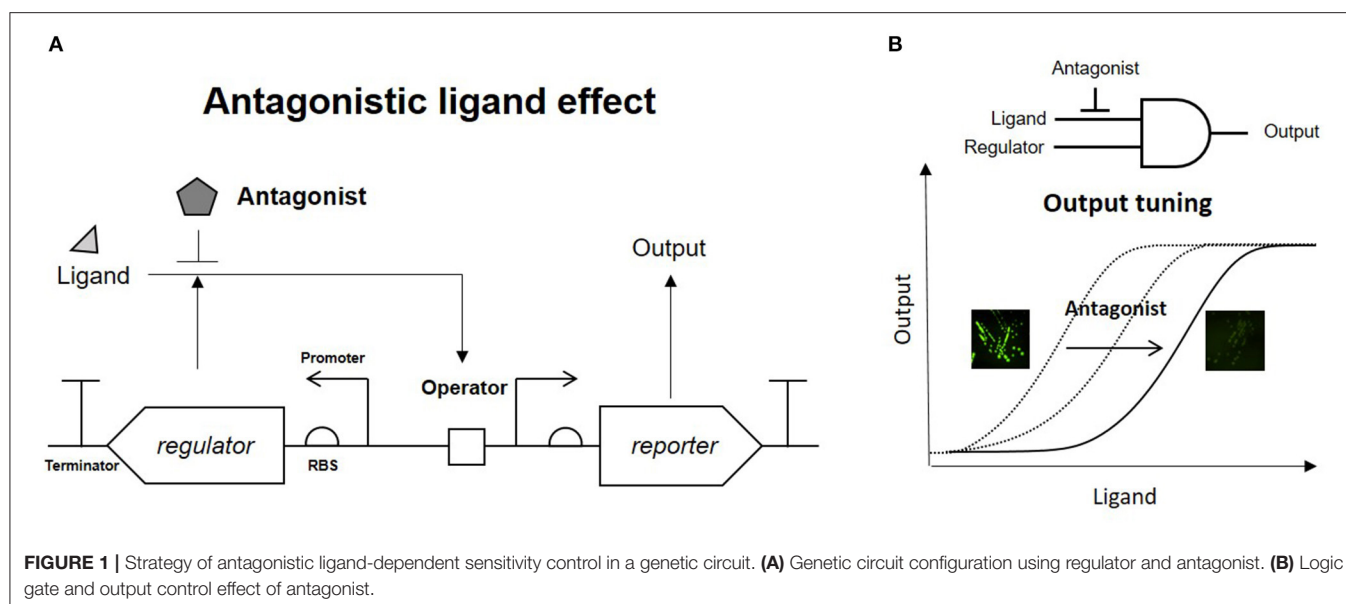
and *p*NP were administered along with phenol to test the inhibitory effect of DmpR. As DmpR is a σ54-dependent transcriptional regulator, it responds to metabolism of alternative carbon sources such as acetate (Kwon et al., 2020). To activate DmpR in the genetic circuit, cells harboring pDmpR-GESS were cultured in LB at and the culture media were changed to fresh M9 media containing 4 g/L acetate, 1 mM aromatic compounds, and various concentrations of phenol. Fluorescence intensity at a single-cell level was measured to evaluate the antagonist effect by flow cytometry (**Figure 2A**). Benzene had no inhibitory effect on phenol, and 2,4-dichlorophenol and *p*NP had stronger antagonistic effects toward the phenol-DmpR complex compared to 3,5-dimethylphenol (**Figure 2B**).

Analysis of the DmpR structure suggested that the hydroxyl group of phenol plays an important role in binding to DmpR (Park et al., 2020). No inhibitory effects were observed when benzene was added to the phenol and DmpR-GESS reaction system, suggesting that benzene does not bind to the active pocket of DmpR, including its aromatic ring moiety. Phenolic compounds containing functional groups in *meta* or *para* positions, such as 2,4-dichlorophenol or 3,5-dimethylphenol, may bind to the ligand binding site without inducing transcriptional activation, which can suppress the output signal. Among these compounds, *p*NP showed the greatest antagonistic effect and was applied to control the dynamic detection range of the genetic circuit as a resistor for the monitoring of enzyme activity.

Quantitative Control of the Dynamic Detection Range of a Genetic Circuit Using an Antagonist

To test the antagonistic effect of *p*NP, fluorescence intensities were measured at different concentrations of phenol and *p*NP in a two-step reaction (**Figure 3A**). *p*NP showed a significant antagonistic effect, even at low concentrations (1 μM). The maximum fluorescence intensity was maintained at concentrations of *p*NP up to 10 μM, but the intensity was lowered at concentrations higher than 10 μM, and the response to phenol was completely lost at 500 μM. The phenol $K_{1/2}$ of DmpR-GESS—the concentration of phenol at the half-maximal fluorescence signal—increased linearly with the *p*NP concentration (**Figure 3B**). The maximum phenol $K_{1/2}$ of the circuit along with *p*NP increased by ~30-fold compared to in its absence, with a wider dynamic detection range. By adding *p*NP as a GCR, the dynamic detection range of the genetic circuit can be controlled from several μM to hundreds of μM. Therefore, *p*NP is a strong antagonist of DmpR, and can be used to precisely control the genetic circuit.

Phenol and *p*NP are non-metabolites in *E. coli*, and thus they can be maintained at a constant intracellular concentration. Therefore, phenol as a ligand can exhibit high sensitivity, even at low concentrations, and *p*NP as an antagonist can exert a constant inhibitory effect over time. Many aromatic compounds can diffuse freely through the cell membrane via passive transport (Gallert and Winter, 1993; Chen and Fink, 2006). In addition, more than 200 enzymes can generate phenol or *p*NP from

FIGURE 1 | Strategy of antagonistic ligand-dependent sensitivity control in a genetic circuit. **(A)** Genetic circuit configuration using regulator and antagonist. **(B)** Logic gate and output control effect of antagonist.

FIGURE 2 | Cell population level analysis of antagonist ligand effect. **(A)** Flow cytometry fluorescence profile was analyzed using a DmpR-based genetic circuit with putative antagonists of DmpR and phenol. Aromatic compounds at 1 mM were added to the reaction. **(B)** Data recalculation from fluorescence profiles using Flowjo software. Values are presented as the median and robust standard deviation (rSD) of each histogram.
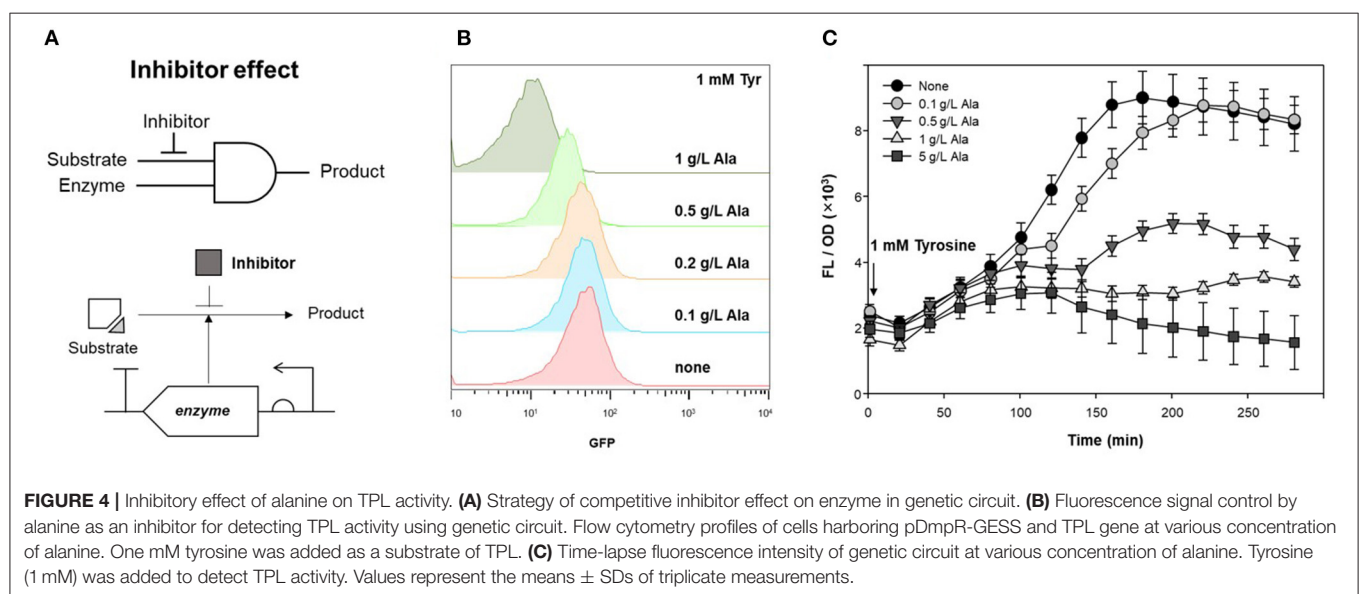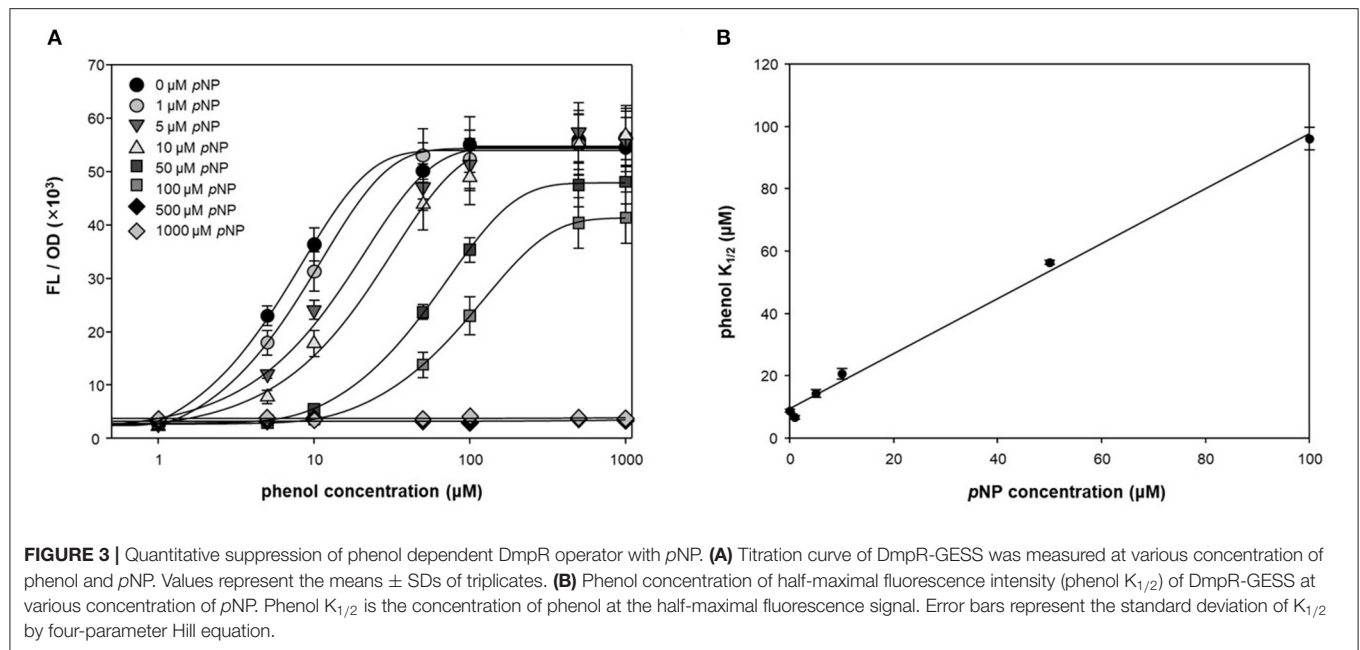
phenolic substrates through their catalytic reactions (Kim et al., 2015). Intermolecular release using phenolic substrates and heterologous enzymes in different cells can affect circuit signals in terms of pattern generation or edge detection. By making use of these properties, the circuit can be expanded to function in intercellular quorum sensing.

Application of Enzyme Inhibitory Effect in the Genetic Circuit

TPL, which produces phenol from tyrosine as a substrate, was applied to an AND gate of the genetic circuit. To control enzyme expression levels, pAR-TPL was constructed in an L-rhamnose expression system for tight regulation, as TPL can use tyrosine generated by the *E. coli* host's amino acid synthesis pathway (**Supplementary Figure 3**). Alanine, which is a competitive inhibitor of the TPL beta-elimination reaction, can be used as a GCR for TPL activity (Demidkina et al., 1987). **Figure 4A** shows the application of the enzyme inhibitor as the resistor in the AND logic gate using an enzyme and its substrate as inputs.

TPL was expressed in LB, after which the cells were transferred to M9 minimal media for two-step induction to maximize fluorescence intensity (Kwon et al., 2020). **Figure 4B** shows the

FIGURE 3 | Quantitative suppression of phenol dependent DmpR operator with pNP. **(A)** Titration curve of DmpR-GESS was measured at various concentration of phenol and pNP. Values represent the means ± SDs of triplicates. **(B)** Phenol concentration of half-maximal fluorescence intensity (phenol $K_{1/2}$) of DmpR-GESS at various concentration of pNP. Phenol $K_{1/2}$ is the concentration of phenol at the half-maximal fluorescence signal. Error bars represent the standard deviation of $K_{1/2}$ by four-parameter Hill equation.

FIGURE 4 | Inhibitory effect of alanine on TPL activity. **(A)** Strategy of competitive inhibitor effect on enzyme in genetic circuit. **(B)** Fluorescence signal control by alanine as an inhibitor for detecting TPL activity using genetic circuit. Flow cytometry profiles of cells harboring pDmpR-GESS and TPL gene at various concentration of alanine. One mM tyrosine was added as a substrate of TPL. **(C)** Time-lapse fluorescence intensity of genetic circuit at various concentration of alanine. Tyrosine (1 mM) was added to detect TPL activity. Values represent the means ± SDs of triplicate measurements.
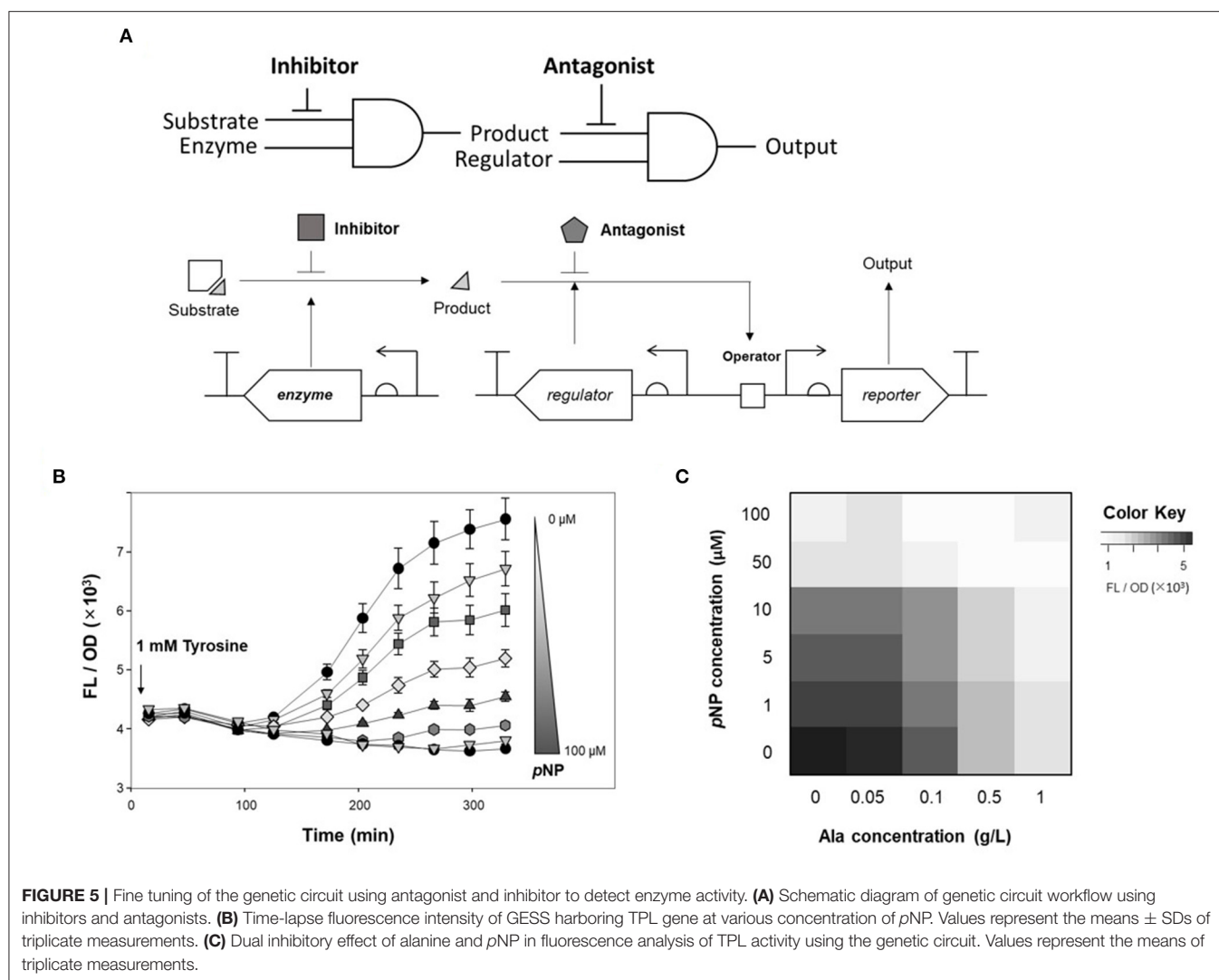
fluorescence intensity induced by different concentrations of alanine, as measured by flow cytometry. When 1 mM tyrosine was added to M9, the fluorescence intensity, which reflected the activity of TPL, was reduced at concentrations of alanine above 0.5 g/L. For solid phase assays, cells harboring the genetic circuit and TPL gene were incubated in LB agar plate containing 1 mM tyrosine and 1 g/L alanine at 37°C for 20 h. The fluorescence intensity of the colonies was suppressed in LB agar plates when both the substrate and inhibitor were present (**Supplementary Figure 4**). Thus, alanine can be used as an enzyme inhibitor of TPL and as the resistor in a genetic circuit.

Figure 4C shows the inhibitory effect of alanine in the genetic circuit, measured as the time-lapse fluorescence intensity. The fluorescence intensity was dependent on the inhibitor

concentration, and the fluorescence signal was restored at low concentrations of alanine (0.1 g/L). Alanine can be metabolized by the host and, over time, the inhibitory effect may be weakened. In a genetic circuit composed of enzymes, the output can be regulated by reducing the enzyme activity via addition of an inhibitor. If the inhibitor is a metabolite in the host, the intracellular concentration gradually decreases, resulting in a delayed-output signal until enzyme activity is restored.

Fine Tuning of Genetic Circuit Using Regulator Antagonist and Enzyme Inhibitor

To tune the genetic circuit, antagonists and inhibitors were applied to monitor enzyme activity. Of the two AND logic gates of DmpR-GESS (**Figure 5A**), the first AND gate, consisting of

FIGURE 5 | Fine tuning of the genetic circuit using antagonist and inhibitor to detect enzyme activity. (A) Schematic diagram of genetic circuit workflow using inhibitors and antagonists. (B) Time-lapse fluorescence intensity of GESS harboring TPL gene at various concentration of pNP. Values represent the means ± SDs of triplicate measurements. (C) Dual inhibitory effect of alanine and pNP in fluorescence analysis of TPL activity using the genetic circuit. Values represent the means of triplicate measurements.

TPL and tyrosine, was able to control the level of product, with alanine used as an inhibitor. As an antagonist, pNP can be used to modulate the output signal of the second AND gate using as inputs a regulator and the phenol produced by the first gate.

To quantify the use of pNP as a GCR for measuring enzyme activity, time-lapse fluorescence intensity of cells harboring TPL and DmpR-GESS was measured at various concentrations of pNP (**Figure 5B**). The fluorescence intensity decreased as the concentration of pNP increased. pNP, as a non-metabolite in the host, can maintain its inhibitory effect over time. The antagonist can act as a resistor in the genetic circuit, controlling the output signal through its effects on the regulator.

Finally, alanine and pNP as GCRs for each AND gates were applied to control the output of the genetic circuit (**Figure 5C**). Fine-tuning of the genetic circuit had a greater effect when the two GCRs were applied simultaneously than with a single application. As a result, the genetic circuit can control the output signal when GCRs are added and combined for each logic gates.

One of the great demands of enzyme engineering using genetic circuits is catalytic improvement. Various types of

genetic circuits have been developed over the past decade (Eggeling et al., 2015; Mahr and Frunzke, 2016). Many circuit studies investigated the control of the transcriptional signals by replacing genetic elements such as promoters and ribosome binding sites. The dynamic detection range of the genetic circuit is fixed for the original enzyme activity at the beginning of construction. Moreover, the detection range would not be revised during high-throughput screening rounds for better catalysts. Post-translational methods such as the use of protein degradation tags, incorporation of protein-protein interactions, and conformational changes have been used to control the expression of proteins (Pawson and Nash, 2000; Nussinov and Ma, 2012; Cameron and Collins, 2014). However, most of these post-translational methods rely on their specific characteristics for each new property of a genetic circuit.

To apply GCR to the design of genetic circuits, small molecules should be transported into the cell by passive or active transport systems. Given this basic condition, the output signal can be controlled by adjusting the activity of the enzyme or sensitivity of the regulator. The use of a GCR enables screening

for higher enzyme activity, or to control a dynamic sensing and regulation system which can generate multiple outputs using a single module.

CONCLUSIONS

This study demonstrates an approach for controlling genetic circuit properties using small molecules which interact with the enzyme and regulator which are parts of the genetic circuit. The enzyme and regulator generally work with specific small molecules, and these biochemical reactions are inhibited by other substances with similar molecular properties and act as inhibitors and antagonists. In this study, alanine was used to inhibit TPL and *p*NP was used as an antagonist of DmpR to produce a genetic circuit for monitoring enzyme activity. These inhibitory small molecules, by interacting with the enzyme or regulator, can be used as logical elements for circuit design. Our approach may be applied to a range of genetic circuit-mediated enzyme applications, such as the directed evolution of enzymes, without the need for reconstructing each different genetic circuit.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

S-GL conceived the study. ER performed the plasmid construction. KK performed the fluorescence analysis. JL performed the fluorescence image. S-GL, KK, HK, S-JY, HL, and D-HL wrote the manuscript. All authors have given approval to the final version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.599878/full#supplementary-material

REFERENCES

Brophy, J. A., and Voigt, C. A. (2014). Principles of genetic circuit design. *Nat. Methods* 11, 508–520. doi: 10.1038/nmeth.2926

Cameron, D. E., and Collins, J. J. (2014). Tunable protein degradation in bacteria. *Nat. Biotechnol.* 32, 1276–1281. doi: 10.1038/nbt.3053

Chen, H., and Fink, G. R. (2006). Feedback control of morphogenesis in fungi by aromatic alcohols. *Genes Dev.* 20, 1150–1161. doi: 10.1101/gad.1411806

Choi, S. L., Rha, E., Lee, S. J., Kim, H., Kwon, K., Jeong, Y. S., et al. (2014). Toward a generalized and high-throughput enzyme screening system based on artificial genetic circuits. *ACS Synth. Biol.* 3, 163–171. doi: 10.1021/sb400112u

Demidkina, T. V., Myagkikh, I. V., and Azhayev, A. V. (1987). Transamination catalysed by tyrosine phenol-lyase from citrobacter intermedius. *Eur. J. Biochem.* 170, 311–316. doi: 10.1111/j.1432-1033.1987.tb13701.x

Dixon, M. (1953). The determination of enzyme inhibitor constants. *Biochem. J.* 55, 170–171. doi: 10.1042/bj0550170

Eggeling, L., Bott, M., and Marienhagen, J. (2015). Novel screening methods–biosensors. *Curr. Opin. Biotechnol.* 35, 30–36. doi: 10.1016/j.copbio.2014.12.021

Gallert, C., and Winter, J. (1993). Uptake of phenol by the phenol-metabolizing bacteria of a stable, strictly anaerobic consortium. *Appl. Microbiol. Biotechnol.* 39, 627–631. doi: 10.1007/BF00205065

Hausser, J., Mayo, A., Keren, L., and Alon, U. (2019). Central dogma rates and the trade-off between precision and economy in gene expression. *Nat. Commun.* 10:68. doi: 10.1038/s41467-018-07391-8

Kenakin, T. (2007). Allosteric theory: taking therapeutic advantage of the malleable nature of GPCRs. *Curr. Neuropharmacol.* 5, 149–156. doi: 10.2174/157015907781695973

Kim, H., Kwon, K. K., Rha, E., and Lee, S.-G. (2015). "Genetic enzyme screening system: a method for high-throughput functional screening of novel enzymes from metagenomic libraries," in *Hydrocarbon and Lipid Microbiology Protocols*, eds T. McGenity, K. Timmis, and B. Nogales (Berlin; Heidelberg: Springer), 3–12. doi: 10.1007/8623_2015_65

Kim, S. K., Lee, D. H., Kim, O. C., Kim, J. F., and Yoon, S. H. (2017). Tunable control of an *Escherichia coli* expression system for the overproduction of membrane proteins by titrated expression of a mutant lac repressor. *ACS Synth. Biol.* 6, 1766–1773. doi: 10.1021/acssynbio.7b00102

Kwon, K. K., Yeom, S. J., Choi, S. L., Rha, E., Lee, H., Kim, H., et al. (2020). Acclimation of bacterial cell state for high-throughput enzyme engineering using a DmpR-dependent transcriptional activation system. *Sci. Rep.* 10:6091. doi: 10.1038/s41598-020-62892-1

Mahr, R., and Frunzke, J. (2016). Transcription factor-based biosensors in biotechnology: current state and future prospects. *Appl. Microbiol. Biotechnol.* 100, 79–90. doi: 10.1007/s00253-015-7090-3

Markel, U., Essani, K. D., Besirlioglu, V., Schiffels, J., Streit, W. R., and Schwaneberg, U. (2020). Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* 49, 233–262. doi: 10.1039/C8CS00981C

Ngara, T. R., and Zhang, H. (2018). Recent advances in function-based metagenomic screening. *Genomics Proteomics Bioinformatics* 16, 405–415. doi: 10.1016/j.gpb.2018.01.002

Nielsen, A. A., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., et al. (2016). Genetic circuit design automation. *Science* 352:aac7341. doi: 10.1126/science.aac7341

Nussinov, R., and Ma, B. (2012). Protein dynamics and conformational selection in bidirectional signal transduction. *BMC Biol.* 10:2. doi: 10.1186/1741-7007-10-2

O'Neill, E., Sze, C. C., and Shingler, V. (1999). Novel effector control through modulation of a preexisting binding site of the aromatic-responsive sigma(54)-dependent regulator DmpR. *J. Biol. Chem.* 274, 32425–32432. doi: 10.1074/jbc.274.45.32425

Park, K. H., Kim, S., Lee, S. J., Cho, J. E., Patil, V. V., Dumbrepatil, A. B., et al. (2020). Tetrameric architecture of an active phenol-bound form of the AAA(+) transcriptional regulator DmpR. *Nat. Commun.* 11:2728. doi: 10.1038/s41467-020-16562-5

Pawson, T., and Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes Dev.* 14, 1027–1047. doi: 10.1101/gad.14.9.1027

Rogers, J. K., Guzman, C. D., Taylor, N. D., Raman, S., Anderson, K., and Church, G. M. (2015). Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res.* 43, 7648–7660. doi: 10.1093/nar/gkv616

Siedler, S., Schendzielorz, G., Binder, S., Eggeling, L., Bringer, S., and Bott, M. (2014). SoxR as a single-cell biosensor for NADPH-consuming enzymes in *Escherichia coli*. *ACS Synth. Biol.* 3, 41–47. doi: 10.1021/sb40 0110j

Smanski, M. J., Bhatia, S., Zhao, D., Park, Y., Woodruff, L. B. A., Giannoukos, G., et al. (2014). Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* 32, 1241–1249. doi: 10.1038/nbt.3063

Tabor, J. J., Salis, H. M., Simpson, Z. B., Chevalier, A. A., Levskaya, A., Marcotte, E. M., et al. (2009). A synthetic genetic edge detection program. *Cell* 137, 1272–1281. doi: 10.1016/j.cell.2009.04.048

Uchiyama, T., and Miyazaki, K. (2010). Product-induced gene expression, a product-responsive reporter assay used to screen metagenomic libraries for enzyme-encoding genes. *Appl. Environ. Microbiol.* 76, 7029–7035. doi: 10.1128/AEM.00464-10

Xie, M., Ye, H., Hamri, G. C., and Fussenegger, M. (2014). Antagonistic control of a dual-input mammalian gene switch by food additives. *Nucleic Acids Res.* 42:e116. doi: 10.1093/nar/gku545

Yeom, S. J., Kim, M., Kwon, K. K., Fu, Y., Rha, E., Park, S. H., et al. (2018). A synthetic microbial biosensor for high-throughput screening of lactam biocatalysts. *Nat. Commun.* 9:5053. doi: 10.1038/s41467-018-07488-0

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership