

Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Sina Djafari Rouhani

April 4, 2022

Understanding the basis of substrate recognition by the RiPP biosynthetic enzyme SuiB

by

Sina Djafari Rouhani

Dr. Katherine Davis
Adviser

Department of Chemistry

Dr. Katherine Davis
Adviser

Dr. Khalid Salaita
Committee Member

Dr. Deepika Das
Committee Member

2022

Understanding the basis of substrate recognition by the RiPP biosynthetic enzyme SuiB

By

Sina Djafari Rouhani

Dr. Katherine Davis

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Chemistry

2022

Abstract

Understanding the basis of substrate recognition by the RiPP biosynthetic enzyme SuiB

By Sina Djafari Rouhani

Finding new ways to combat antibiotic resistant bacterial infections has become a serious concern in the scientific community. Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a newly classified group of natural products which have vast structural and chemical diversity correlated with antifungal, antibacterial, allelopathic, and antiviral activity. The modular and intrinsically tolerant biosynthetic pathways of RiPPs makes them a highly promising venture for bioengineering and downstream drug development. The RiPP recognition element (RRE) of RiPP biosynthetic enzymes has been implicated in the identification and transfer of ribosomally synthesized peptides to the active site for post-translational modification in a leader-dependent fashion. However, new questions about the role of the RRE domain in modifying enzymes and the function of the leader sequence in substrate peptides arose from the crystal structure of the RiPP peptide/enzyme pair of SuiA/SuiB which had the leader region of SuiA bound outside the RRE domain of SuiB. This study aims to analyze the interactions between the leader region of SuiA and the RRE domain of SuiB both to gain structural insights into the basis for precursor peptide recognition and clarify the role of the RRE. Molecular docking simulations were used in an attempt to build models of potential interactions between the SuiA-leader and the SuiB-RRE, however, the crystallographically-derived SuiA-leader structure failed to dock. Molecular dynamics simulations on the precursor peptide showed the α -helix of the leader region denaturing in solution, which suggested the binding conformation of the peptide leader sequence is non-helical. This result was verified by circular dichroism spectroscopy, indicating that interactions between the SuiA leader and the RRE domain of SuiB, occur with SuiA as a random coil as opposed to the α -helical conformation observed in the active site. X-ray crystallography was attempted to identify the specific interactions between the SuiA-leader and the SuiB-RRE. Crystals were obtained, but these experiments are ongoing. We are currently in the process of optimization to improve diffraction quality.

Understanding the basis of substrate recognition by the RiPP biosynthetic enzyme SuiB

By

Sina Djafari Rouhani

Dr. Katherine Davis

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Chemistry

2022

Acknowledgements

I would like to thank my committee members Dr. Deepika Das and Dr. Khalid Salaita for their profound role in my academic development in the Department of Chemistry of Emory University. My deepest appreciation goes to the Davis Lab, especially my research advisor, Dr. Katherine Davis, for her sincere mentorship and guidance. I would like to extend many thanks to my graduate student mentors, Stacey Jones and Tamra Blue, for building me into the scientist I am today. I would also like to recognize the contributions of fellow Undergraduate Research Assistant, Nithin Bagal, for his contributions to the computational studies of this thesis.

Table of Contents

1. INTRODUCTION	1
1.1 <i>Introduction to Ribosomally Synthesized and Post-translationally Modified Peptides</i>	1
1.2 <i>RiPP Biosynthesis</i>	2
1.3 <i>SuiA and SuiB: RiPP class of precursor peptide and radical SAM modifying enzyme</i>	3
1.4 <i>Research Overview</i>	7
2. METHODS	8
2.1 <i>Computational Methods</i>	8
2.2 <i>Mass Bacterial Expression of NusA-RRE</i>	10
2.3 <i>Circular Dichroism Spectroscopy</i>	15
2.4 <i>Crystallography</i>	16
3. RESULTS & ANALYSIS	17
3.1 <i>Molecular Docking of SuiA-leader with SuiB-RRE and Comparison with Analogous Homologs</i>	19
3.2 <i>Molecular Docking of a different class of RiPP peptide/modifying-enzyme pair</i>	27
3.3 <i>Circular Dichroism Spectroscopy of SuiA-leader and NusA-RRE</i>	30
3.4 <i>X-ray Protein Crystallography of SuiA-leader and NusA-RRE</i>	32
4. DISCUSSION	36
5. CONCLUSION	38
6. REFERENCES	41

Table of Figures

Figure 1. Diagram of RiPP biosynthesis.....	3
Figure 2. The <i>sui</i> gene cluster.....	4
Figure 3. Six RiPP modifying enzymes with their precursor peptides bound.....	6
Figure 4. Molecular docking computational workflow.....	9
Figure 5. NusA-RRE fusion protein plasmid.....	11
Figure 6. SDS-PAGE gels of NusA-RRE expression and purification.....	13
Figure 7. Size exclusion chromatograph of NusA-RRE.....	14
Figure 8. UV-vis absorption spectrum of NusA-RRE.....	15
Figure 9. Protein x-ray crystallography workflow.....	17
Figure 10. Representative model of SuiA docking with SuiB-RRE.....	20
Figure 11. Residue comparison of SuiB, AgaB, StrB.....	23
Figure 12. Residue comparison of AgaB and StrB RRE domains against SuiB.....	23
Figure 13. Hydrophobicity comparison between SuiB, AgaB, and StrB RRE regions.....	24
Figure 14. Leader sequence comparison of SuiA, AgaA, and StrA.....	24
Figure 15. Molecular docking of AgaA-leader/AgaB-RRE and StrA-leader/StrB-RRE.....	26
Figure 16. AgaA structural conformation in 30 ns MD simulation.....	26
Figure 17. Structural overlay of docked and crystalized CteA-leader/CteB-RRE models.....	28
Figure 18. Binding interactions in CteA-leader/CteB-RRE crystal structure.....	29
Figure 19. Binding interactions in CteA-leader/CteB-RRE docked structure.....	30
Figure 20. Circular dichroism spectra.....	32
Figure 21. Crystals grown in broad-range 96-well crystal screens.....	33
Figure 22. Crystals grown in optimization trays.....	34
Figure 23. TEV protease cleavage and new plasmid construct.....	36

1. INTRODUCTION

1.1 Introduction to Ribosomally Synthesized and Post-translationally Modified Peptides (RiPPs)

Antibiotic resistant bacterial infections have become prime concerns in the field of human health, and a high level of precedence has been established for finding new treatment options to address them. Natural products (NPs) have garnered a lot of attention to this end due to their potent biological activities and structurally unique pharmacophores.^{1, 2} Historically, microbes with a rich source of antibiotic compounds have been targeted with the goal of identifying therapeutically applicable small molecular metabolites.³ Vast resources have been allocated for engineering biosynthetic pathways responsible for those antibiotic compounds in the hope of generating novel analogs with appealing pharmacological properties. Four traditional classes of natural products have been identified in the 20th and early 21st centuries with prevalent antibiotic activity: terpenoids, alkaloids, polyketides, and non-ribosomal peptides.⁴⁻⁷ Regrettably, stringencies in the biosynthesis of these traditional classes of natural products, most specifically limitations in the capacity of non-ribosomal peptide and polyketide synthases to produce novel molecular analogs in large numbers, has prevented their use as effective therapeutics against antibiotic resistant bacterial infections.³

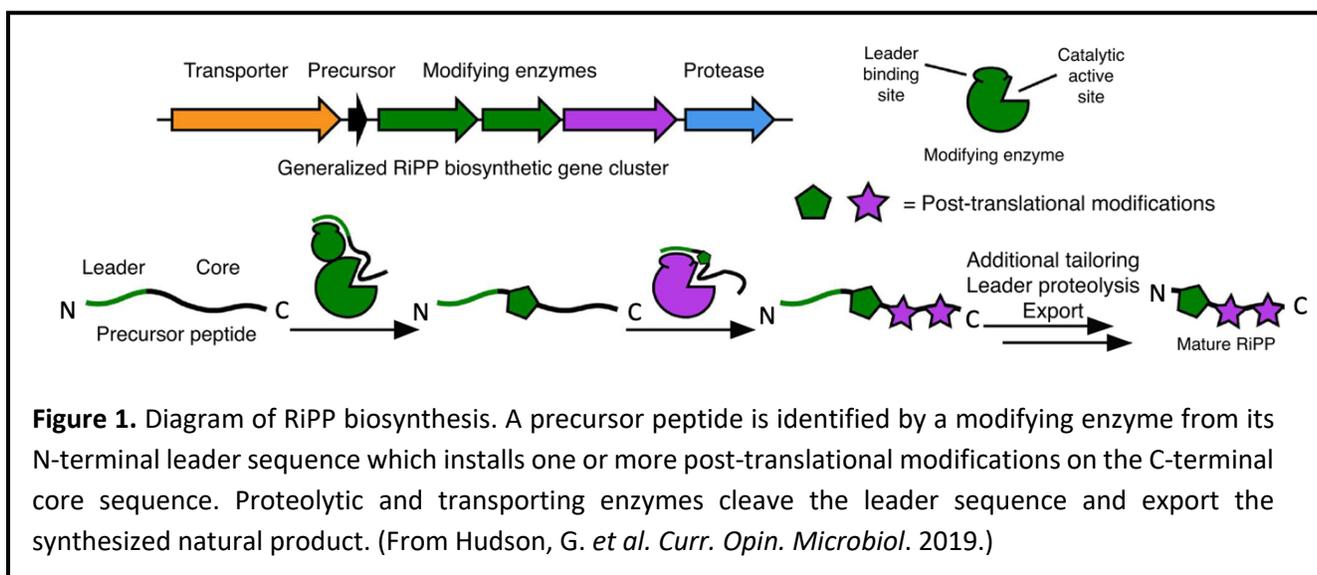
Nevertheless, genome sequencing has revealed a new class of natural products that display native antibiotic activity, the ribosomally synthesized and post-translationally modified peptides (RiPPs). RiPPs have been identified within all three domains of life with ubiquitous biosynthetic genes and vast structural diversity.⁴ In contrast to non-ribosomal peptides that use large multi-

modular enzyme complexes for integrating non-proteinogenic amino acids into a peptide backbone, RiPPs can achieve a similar degree of chemical diversity through sequential post-translational modifications of a ribosomally synthesized precursor peptide through a leader sequence dependent fashion.⁸ Examples of the chemical and structural diversity of RiPPs include carbon-carbon cross links, intramolecular thioether linkages, six-membered nitrogen-containing rings, and N-to-C macrocycles.⁹ This wide variety of structural and chemical features endow RiPPs with a range of activities such as antifungal, antibacterial, allelopathic, and antiviral.^{9, 10} Therefore, it is clear the chemical moieties of RiPPs coupled with their modular and intrinsically tolerant biosynthetic pathways, make the study of RiPPs a highly promising venture for discovering new antibiotics to combat antibiotic resistance.

1.2 RiPP Biosynthesis

RiPPs are characterized into subfamilies depending on their biosynthetic machinery and structural features, yet they all share a similar simple biosynthetic pathway (Fig. 1). RiPP biosynthesis initiates with the ribosomal synthesis of a precursor peptide typically containing an N-terminal leader region and a C-terminal core region.^{3, 9} The leader sequence of the precursor peptide is responsible for guiding the peptide through its natural product biosynthesis, while the core region harbors sites for post-translational modifications.¹¹ When a modifying enzyme is present, the recognition sequence of the N-terminal leader region of the precursor peptide binds to the enzyme, at which point the modifying enzyme installs one or more post-translational modifications in the C-terminal core regions of the precursor peptide. Finally, proteolytic and transporting enzymes from the biosynthetic gene cluster cleave the leader region of the peptide and export the natural product. The physical separation of the sites responsible for recognition

and modification of the precursor peptide enables RiPP biosynthetic pathways to be highly specific for recognition yet promiscuously variable for the modifications. This extraordinary biosynthetic malleability of RiPPs is what distinguishes them from non-ribosomal peptides and makes the promising sources for bioengineering and downstream drug development.



1.3 *SuiA* and *SuiB*: RiPP class of precursor peptide and radical SAM modifying enzyme

Understanding how RiPPs interact with their modifying enzymes, or more specifically, how RiPPs are identified and subsequently modified by associated tailoring enzymes is a core area of research that is fundamental for functionalizing RiPPs into effective antibiotics. In this regard, the structure of the RiPP modifying radical *S*-adenosylmethione (SAM) enzyme, *SuiB*, bound with its precursor peptide, *SuiA*, has revealed some interesting features which offer a new outlook for the role of the RRE in precursor peptide recognition.¹²

SuiB is a radical SAM enzyme containing three [4Fe-4S] clusters and is comprised of three functionally distinct domains (Fig. 2a). The N-terminal domain (residues 1-106) contains a winged helix-turn-helix (wHTH) topology that has been classified by bioinformatics studies as a RiPP precursor peptide recognition element (RRE).¹³ The RRE domain of RiPP biosynthetic modifying enzymes has been implicated in leader-dependent peptide recognition and recruitment, such that it both identifies a peptide and positions the core regions in the active site. Subsequent to the RRE domain is the canonical radical SAM domain (residues 107–310) which is connected by a short bridging region (residues 311–346). This domain is responsible for reductively cleaving SAM bound to a [4Fe-4S]⁺ cluster to generate a 5'-deoxyadenosyl radical (5'-dA·), which initiates turnover by abstraction of a hydrogen atom from the substrate.¹⁴ Finally, in the C-terminal of the enzyme there is the SPASM domain which contains a 7-cysteine motif (CX₉₋₁₅GX₄C-gap-CX₂CX₅CX₃C-gap-C) that supplements the binding of two additional “auxiliary” [4Fe-4S] clusters.

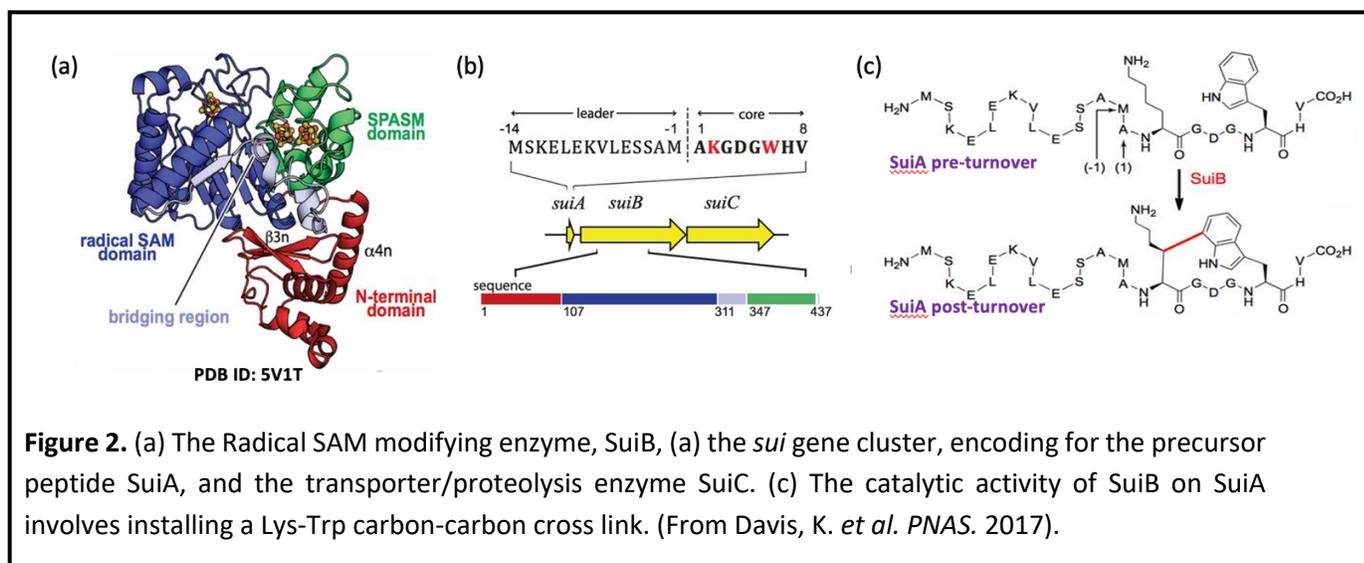
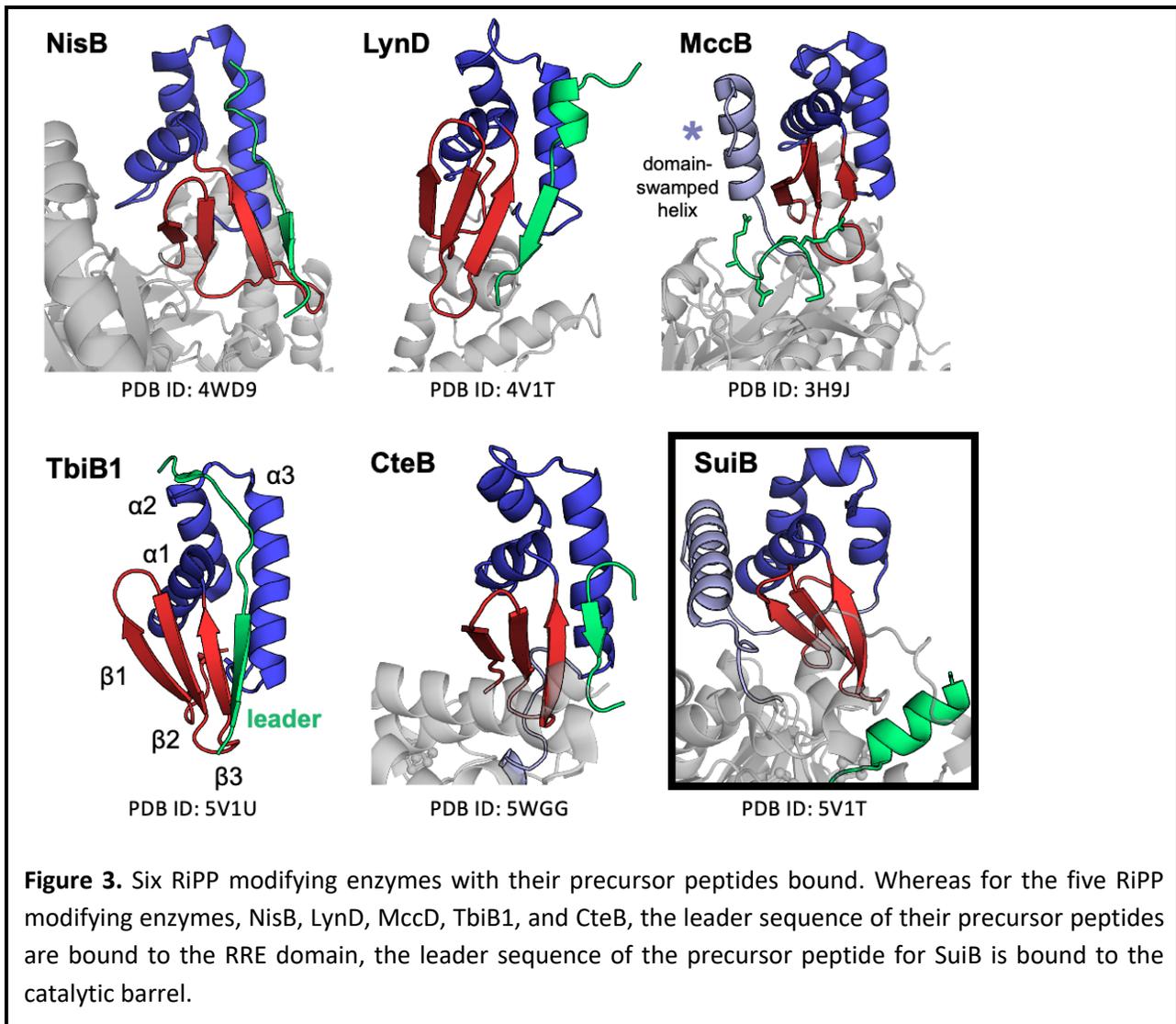


Figure 2. (a) The Radical SAM modifying enzyme, SuiB, (a) the *sui* gene cluster, encoding for the precursor peptide SuiA, and the transporter/proteolysis enzyme SuiC. (c) The catalytic activity of SuiB on SuiA involves installing a Lys-Trp carbon-carbon cross link. (From Davis, K. *et al.* PNAS. 2017).

More broadly, SuiB is part of the *sui* gene cluster for the biosynthesis of streptide (Fig. 2b). In the biosynthesis of streptide, the SuiA precursor peptide is ribosomally synthesized, which contains an N-terminal leader sequence and a C-terminal core sequence, typical of most RiPPs. The SuiA precursor peptide undergoes a post-translational modification after binding to the modifying enzyme, SuiB, in a leader dependent fashion. Specifically, turnover is achieved via a Lys-Trp carbon-carbon cross link in the core region of SuiA (Fig. 2c).¹⁵ The precise catalytic mechanism for post-translational modification on SuiA via SuiB follows a radical electrophilic aromatic substitution paradigm.^{16, 17} Finally, the transporter and protease enzyme SuiC cleaves the leader sequence of the post-translationally modified SuiA to form the natural product, streptide.

Previous studies have provided evidence to suggest that the leader sequence of precursor peptide directly interacts with the RRE of tailoring enzymes to act as a guide that facilitates the modification.¹⁸ Although the N-terminal domain in SuiB adopts an RRE fold primed to facilitate peptide delivery, the crystal structure shows the SuiA leader peptide primarily interacting with the catalytic barrel instead adjacent to both the bridging region and the SPASM domain (Fig. 3). This observation is unique among published RRE-containing structures, in which the RiPP-modifying enzymes have the leader region of their precursor peptides docked in the RRE domain (Fig. 3).¹⁸ Additionally, the α -helical nature of the SuiA leader is maintained within the catalytic barrel before it transitions into a loop that is adjacent to Aux I. While many RiPP leader peptides have been shown to adopt α -helical conformations in trifluoroethanol, the persistence of this secondary structure upon binding to the tailoring enzyme is a unique observation.



Recognition of the SuiA leader sequence is primarily achieved through interactions with the bridging region. These interactions orient the substrate helix and thereby facilitate proper arrangement of the core sequence in the active site. Thus, this discovery in SuiB not only elucidates leader peptide function but also provides new hypotheses into the role of the RRE domain during catalysis. Specifically, this unique SuiA-binding mode suggests that the RRE-like domain in SuiB is either vestigial or involved in an undetected interaction.

1.4 Research Overview

In order to understand the role of the RRE domain of RiPP-modifying enzymes, and characterize the function of the leader sequence in RiPP precursor peptides, this research project aimed to answer the question: what role, if any, does the RRE domain of SuiB play in identifying the SuiA leader sequence and shuttling it to the active site? The unique α -helical conformation of the leader region of SuiA, which was not bound to the RRE domain of SuiB, inspired a hypothesis in which the RRE domain both recognizes the peptide and delivers it to the active site but at a certain stage in the catalytic cycle releases the precursor peptide. Molecular docking studies were used for determining the interactions between the SuiA-leader peptide and the SuiB-RRE. Computational studies were supplemented with experimental methods, namely, circular dichroism spectroscopy was used to confirm binding between SuiA-leader and SuiB-RRE as well as for identifying the binding topology of SuiA-leader. X-ray crystallography was attempted for obtaining a novel crystal structure complexing SuiA-leader with SuiB-RRE domain expressed separately from the larger enzyme. Although these studies have not yet yielded diffracting crystals, this aspect is ongoing. Overall, our findings reveal new insights about SuiA-leader and SuiB-RRE interactions and, more broadly, about leader-dependent recognition of peptides in RiPP biosynthesis.

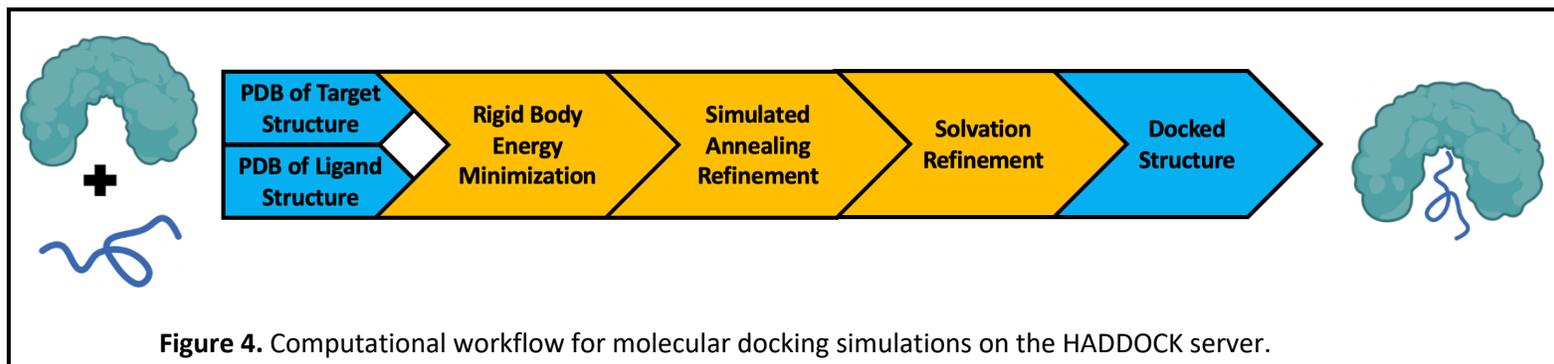
2. METHODS

2.1 Computational Methods

Molecular docking simulations were performed on the open-source molecular docking server, HADDOCK (high ambiguity driven protein-protein docking).¹⁹⁻²¹ An *ab initio* docking protocol was run on HADDOCK by providing PDB coordinates for the starting structures of each peptide-enzyme system studied (Fig. 4). The choice of HADDOCK as the platform for performing molecular docking simulations was motivated by HADDOCK's ability to allow conformational changes in the molecule during complex formation for both the side chains and the backbone. This feature in HADDOCK docking simulations is unique among the wide variety of docking servers.

From the user supplied PDB coordinates, HADDOCK automatically generates the topology of the molecules that are to be docked. At this point, HADDOCK follows a three-step protocol to deliver a set of highly accurate docked models showing the most likely interactions between the peptide-enzyme system: (i) randomization of orientations and rigid body energy minimization, (ii) semirigid simulated annealing in torsion angle space, and (iii) final refinement in the Cartesian space with an explicit solvent (Fig. 4). More specifically, inter- and intramolecular energies are evaluated using full electrostatic and van der Waals energy terms with an 8.5 Å distance cutoff using the OPLS nonbonded parameters. During the randomization process, the two partner proteins were placed 150 Å away from each other in space and allowed to randomly rotate around their center of masses. Rigid body energy minimizations were performed with four cycles of orientational optimization, where each protein was allowed to rotate to minimize the intermolecular energy function. This was followed by both translational and rotational

optimization, through which the two proteins were docked by rigid body energy minimization. For this stage, HADDOCK was programmed to find 1000 complex conformations, from which the 400 best structures in terms of intermolecular energies were chosen for three phases of simulated annealing refinement. In the first phase, 1000 steps from 2000 to 50 K at 8 fs time steps were simulated to optimize the orientation of the two proteins while considering them as rigid bodies. In the second phase, 4000 steps from 2000 to 50 K at 4 fs time steps were simulated to allow the side chains at the interface to move. In the third phase, 1000 steps from 500 to 50 K at 2 fs time steps were simulated to allow for conformational rearrangements where both the side chains and the backbone at the interface were allowed to move. The resulting structures from these simulated annealing refinements were then subjected to 200 steps of steepest descent energy minimizations. Finally, a gentle refinement of the structures was performed in an 8 Å shell of TIP3P water molecules. For this, molecular dynamics simulations over 5000 steps were performed on each structure that included a heating and cooling process with position restraints only on noninterface heavy atoms. HADDOCK uses pairwise backbone Root-Means Squared Deviation (RMSD) at the interface to report final structures as clusters (two or more conformations displaying an interface backbone RMSD less than 1.0 Å). These resulting clusters were then analyzed by the type of interactions they reported and their interaction energies.



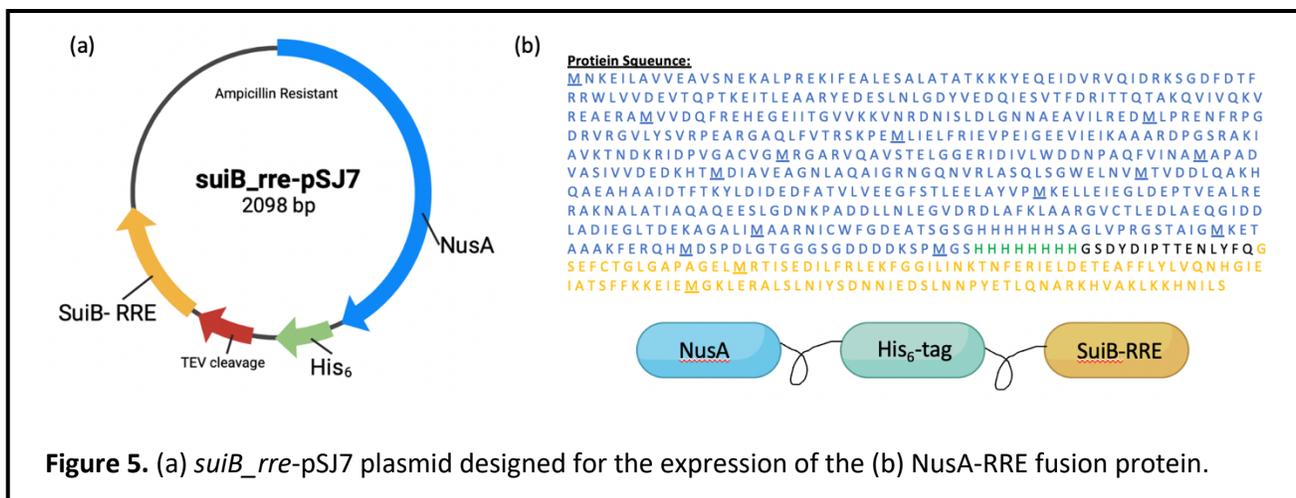
Molecular Dynamics (MD) simulations were run on GROMACS using the AMBER-99SB force field with a TIP3P cubic water model, which places the molecule at least 3 Å from the edge of the solvation box.^{22, 23} Initial structures were prepared in PDB format. Force field conversions were applied to these initial structures, from which the molecules were solvated. Energy minimization of the structures were run for 1 ns, followed by equilibration with and without position restraints. At this point, production MD simulations were run for 30 ns.

Moreover, bioinformatic multiple sequence alignments between different proteins were performed on Clustal Omega, which reported a percent sequence identity (defined as the percentage of residues with the exact identity at each point in the sequence of the proteins) and a percent sequence similarity (defined as the percentage of residues with similar biochemical properties at each point in the sequence of the proteins).^{24, 25} The sequence alignments were used to determine the viability of using analogous homologs of RiPPs and their modifying enzymes for comparative computational analysis. Visualization and modeling of the proteins was done with PyMOL Molecular Graphics System, VMD (Visual Molecular Dynamics), and Chimera.

2.2 Mass Bacterial Expression of NusA-RRE

The *suiB_rre*-pSJ7 plasmid was designed to create a fusion protein which incorporated the N-utilization substance A (NusA) solubility tag, His₆ affinity tag, and the RiPP Recognition Element (RRE) domain of the RiPP-modifying enzyme SuiB (Fig. 5). The NusA tag increases solubility and folding of the RRE fusion partner by acting as a “holdase” in preventing the aggregation of the fusion partner RRE protein such that the folding of this passenger protein can

occur spontaneously.²⁶ The His₆ tag was functionalized during purification via cobalt affinity chromatography due to the high affinity and strong coordination of His₆ for cobalt. The *suiB_rre-pSJ7* plasmid was transformed through electroporation into chemically competent BL21(DE3) *E. coli* cells for protein expression.



Large scale bacterial growth of the transformed *E. coli* cells was performed as follows: 100 mL of LB media supplemented with ampicillin (100 µg/mL) was added to a sterile 250 mL Erlenmeyer flask and inoculated with a single colony of the BL21(DE3) *E. coli* cells transformed with the *suiB_rre-pSJ7* plasmid. The 100 mL starter culture was grown at 37°C and 200 rpm for 12 hours. 10 mL of this starter culture was used to inoculate larger cultures containing 2 L of LB supplemented with ampicillin (100 µg/mL) for a 12 L total expression (6 x 2 L Erlenmeyer flasks). The large cultures were grown at 37°C, 200 rpm, to an OD_{600 nm} of ~0.6, at which point the large cultures were cooled in an ice bath for 5-10 minutes and induced with a final concentration of 0.2 mM of IPTG. After induction, the large cultures were incubated on a shaker for 24 hours at

18°C and 200 rpm. The *E. coli* cells were harvested by centrifugation (4°C, 8,000g, 60 minutes), frozen, and stored at -80°C.

Extraction and purification of the NusA-RRE protein was carried out at 4°C. The harvested *E. coli* cells were resuspended in a 250 mL beaker with lysis buffer (50 mM Tris, 50 mM NaCl, 5mM imidazole, 5% glycerol, pH 7.8, and 1mM β -mercaptoethanol) at 5 mL/g. The suspension was supplemented with protease inhibitor cocktail (0.1% V/V), PMSF (0.25 mM), lysozyme (1 mg/mL) and DNase (1 μ L/mL). The suspension was stirred and sonicated on ice for 60 minutes with a 15 s on/15 s off cycle at 30% power. Following sonication, the cell debris was pelleted via centrifugation at 4°C, 32,000g, for 60 minutes. The supernatant was loaded onto a cobalt metal affinity column (50 mL) which was equilibrated with lysis buffer. The column was washed with CV of lysis buffer and 4 CV of wash buffer (50 mM Tris, 50 mM NaCl, 30mM imidazole, 5% glycerol, pH 7.8, 1mM β -mercaptoethanol, and 0.25 mM PMSF). Finally, the NusA-RRE protein was eluted with 4 CV of elution buffer (50 mM Tris, 50 mM NaCl, 300mM imidazole, 5% glycerol, pH 7.8, 1mM β -mercaptoethanol, and 0.25 mM PMSF). The eluted NusA-RRE protein was concentrated to a volume of 10 mL, which was further purified through size exclusion chromatography on an FPLC (Fast Protein Liquid Chromatograph), buffer exchanged, and stored in 50 mM Tris, 50 mM NaCl, and 5% glycerol at pH 7.5.

The function of the NusA solubility tag in the fusion protein can be clearly seen in the solubility test done on the expression before and after induction with IPTG (Fig. 6a). The thick bands present in lane 3 of the SDS-PAGE gel clearly show the role of the NusA tag in increasing the solubility of the protein expressed by the plasmid. Moreover, lanes 5 and 6 (Fig. 6a)

demonstrate the ability of cobalt affinity chromatography to purify the protein solution, such that the NusA-RRE band is the thickest in the elution. However, impurities are still present in elution as can be seen with the bands below the NusA-RRE band in lane 6 (Fig. 6a). To remedy this, size exclusion chromatography proved to be an effective tool to further purify the elution from cobalt affinity chromatography (Fig. 6b). Size exclusion chromatography separated the different macromolecules present in the elution by order of molecular weight (Fig. 7). From the size exclusion chromatograph (Fig. 7), fractions 9-13 were representative of the NusA-RRE fusion protein. These fractions were run on an SDS-PAGE gel (Fig. 6b), from which fractions 11 and 12 were chosen for resuspension as the final NusA-RRE protein solution. The purity of the final NusA-RRE solution proved satisfactory for spectroscopy and crystallography in a conclusive SDS-PAGE gel (Fig. 6c).

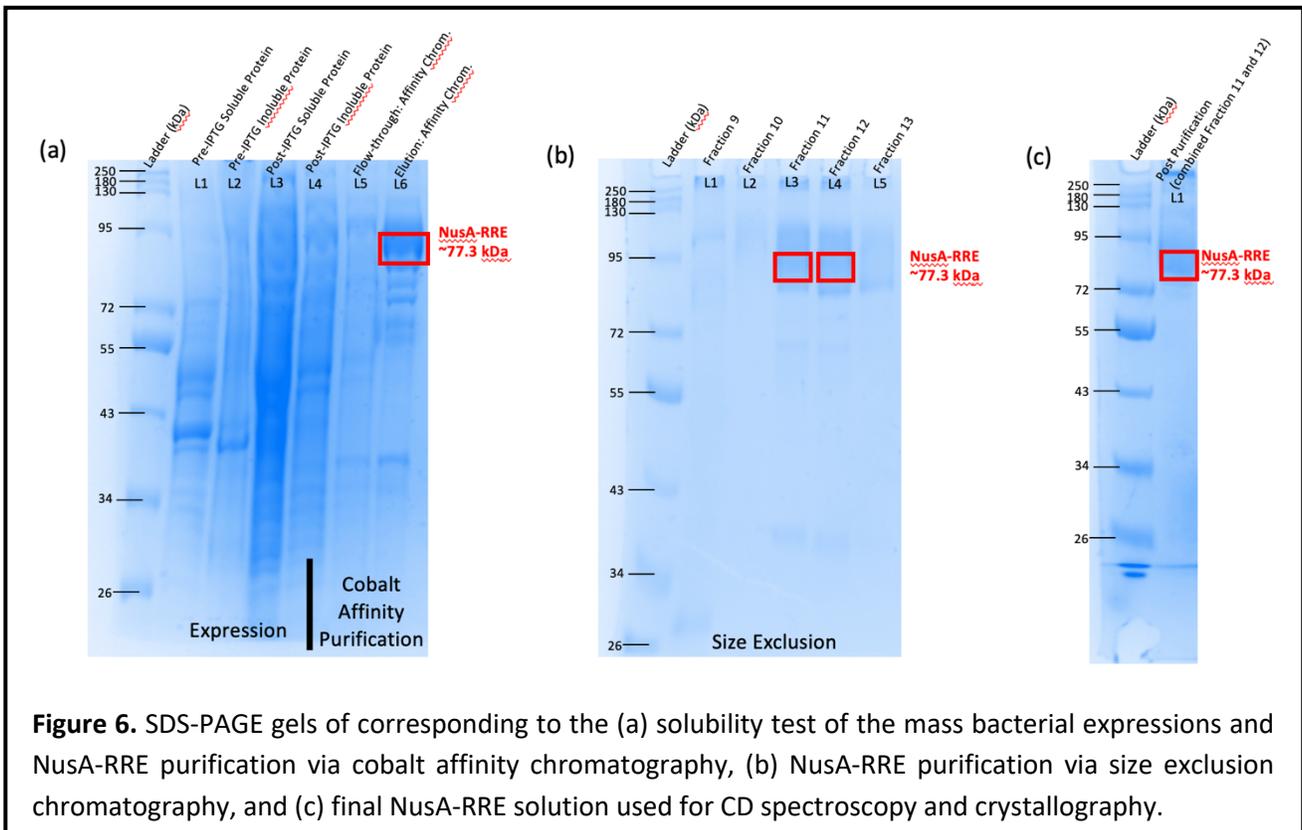
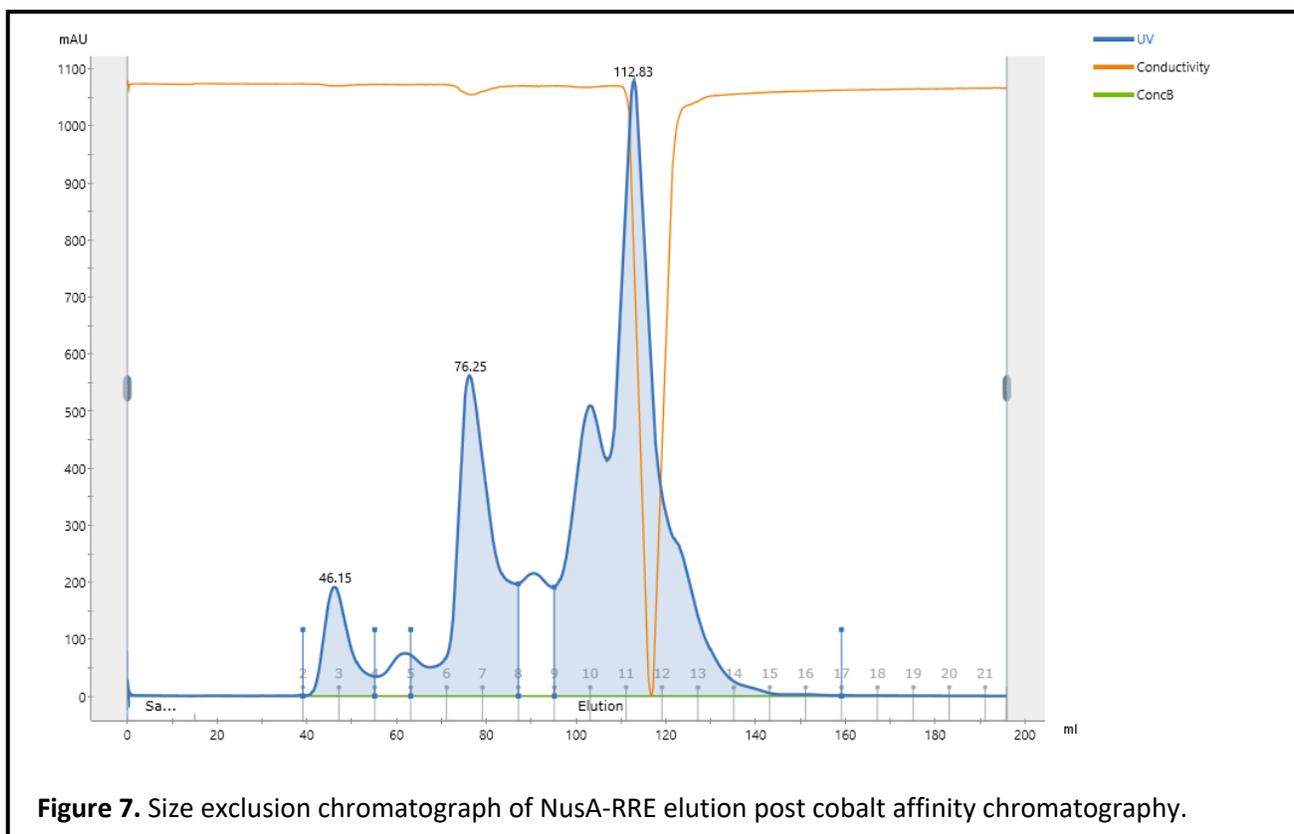
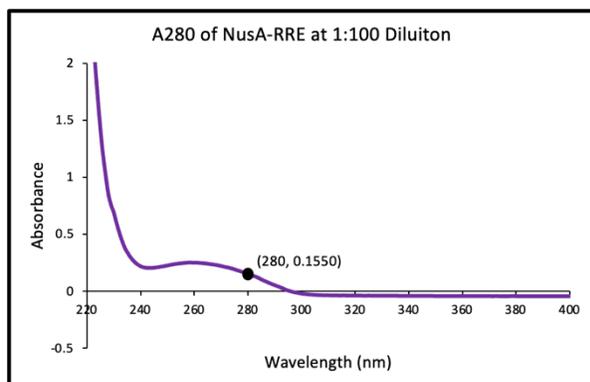


Figure 6. SDS-PAGE gels of corresponding to the (a) solubility test of the mass bacterial expressions and NusA-RRE purification via cobalt affinity chromatography, (b) NusA-RRE purification via size exclusion chromatography, and (c) final NusA-RRE solution used for CD spectroscopy and crystallography.



The concentration of the expressed and purified NusA-RRE was calculated using absorbance at 280 nm with Lambert-Beer's law (Fig. 8). It was determined that the purified NusA-RRE had a yield of 31.03 mg/mL. Additionally, the SuiA leader sequence was commercially synthesized with 95% purity at a concentration of 100 mg/mL. This expressed NusA-RRE complex and this commercially acquired SuiA-leader served the protein sources for experimental CD and crystallography studies.



From Lambert-Beer's Law:

$$A = \epsilon cl$$

Molar absorption coefficient of NusA-RRE = $38,640 \text{ M}^{-1}\text{cm}^{-1}$
 Optical path length = 1 cm

$0.1550 = (38,640)(1)C$
 Concentration = $4.01 \times 10^{-6} \text{ M} = 4.0114 \text{ } \mu\text{M}$
 Concentration Adjusted for Dilution = $401.14 \text{ } \mu\text{M} = 31.03 \text{ mg/mL}$

Figure 8. UV-vis absorption spectrum of NusA-RRE for measuring absorbance at 280 nm and calculating protein concentration via Lambert-Beer's law.

2.3 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy embodies an efficient and effective method for rapidly evaluating the secondary structure, folding, and binding of proteins.²⁷ In general, circular dichroism measures the unequal absorption of left-handed and right-handed circularly polarized light by chiral molecules. For proteins, the chromophores of the amides in the polypeptide backbone are aligned in arrays such that their optical transitions are shifted or split into multiple transitions due to exciton interactions. Therefore, varying structural elements and topologies of proteins show distinct characteristics on a CD spectrum. For example, α -helices are associated with negative bands at 222 nm and 208 nm with a positive band at 193 nm, β -sheets are associated with negative bands at 218 nm and positive bands at 195 nm, and random coils have low ellipticity above 210 nm and negative bands near 195 nm.²⁷ Additionally, binding interactions can also be identified from CD through shifts in the spectra. Subsequently, circular dichroism spectroscopy was used to determine whether interactions occurred in a complexed solution of

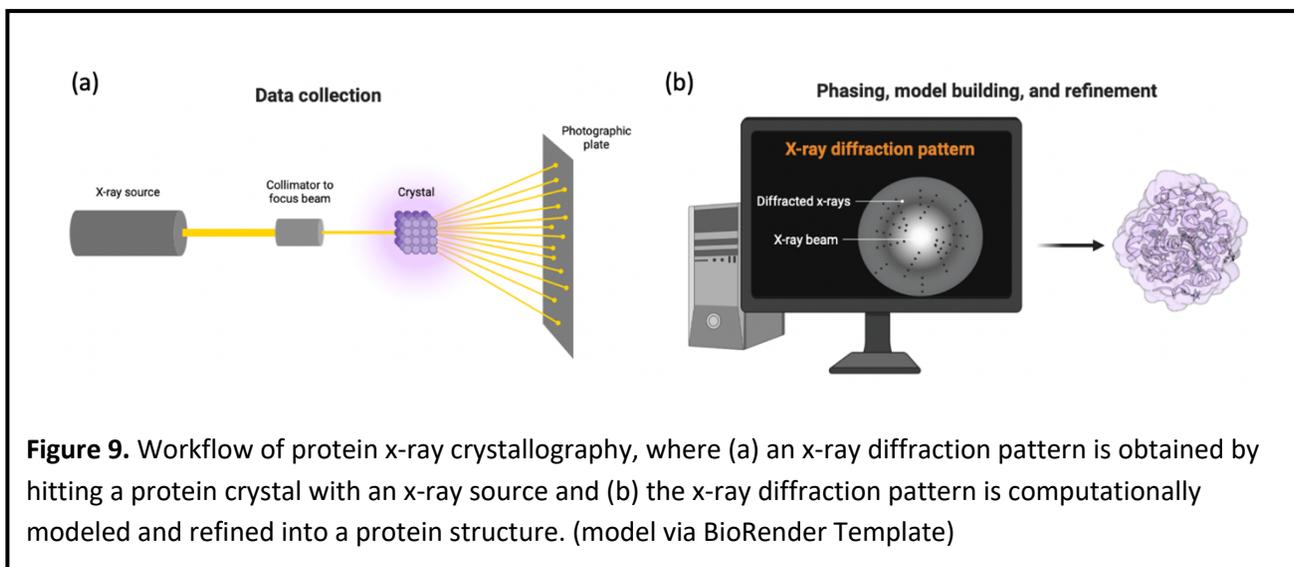
NusA-RRE with SuiA-leader, as well as to identify the secondary structure and topology of the SuiA-leader when present in such a complex.

2.4 Crystallography

Crystals of SuiA-leader bound to NusA-RRE were set up in broad-range sparse matrix screens and grown using the sitting well vapor diffusion method at room temperature (25°C). Protein crystals were grown by dissolving proteins in an aqueous environment and sample solution that induced transition of the protein into a supersaturated state. Supersaturation is a non-equilibrium state where some quantity of the protein is present in excess of the solubility limit. To re-establish favorable equilibrium conditions from a supersaturated state, solids, such as aggregation and crystals, form in solution. In the sitting well vapor diffusion method, a drop of the purified protein solution and reagent consisting of buffer and precipitant were placed in vapor equilibration with a larger liquid reservoir of the reagent. To achieve equilibrium, the aqueous solution diffused from the drop into the reservoir which caused supersaturation in the drop.

Upon collection of the crystals believed to represent complexes of SuiA-leader bound to NusA-RRE, X-ray diffraction patterns from the crystals were collected at the GM/CA beamline of the Advanced Photon Source (APS) at Argonne National Laboratory (Chicago, IL) (Fig. 9a). Although the crystals sent to the APS did not generate a diffraction pattern, in a typical X-ray crystallography workflow, diffraction data would be processed, indexed, integrated, and scaled using XDS software followed by merging with AIMLESS (Fig. 9b).^{28, 29} Model building and structural refinements would be done on COOT and Phenix, and the model quality would be

assessed with Molprobitry.³⁰⁻³² Figures and structural representations would be generated on PyMOL and VMD.



3. RESULTS & ANALYSIS

This project was designed with the goal of better understanding the structural and biochemical role of the RiPP Recognition Element (RRE) domain of RiPP-modifying enzymes, as well as characterizing the function of the leader sequence in RiPP precursor peptides. The leader sequence of SuiA and the RRE domain of SuiB were chosen as the RiPP peptide/modifying-enzyme model of study due to the unique leader sequence binding mode of SuiA that was observed when it was crystallized with SuiB.¹² Specifically, in observing the SuiA-leader obtain an unique α -helical conformation while it was bound to the SuiB catalytic barrel as opposed to expected binding with the SuiB RRE domain, new possibilities were formulated for the role of the RRE domain of RiPP-modifying enzymes and for the function of the leader sequence in RiPP precursor peptides. Given the vital role of the RRE domain in other crystalized structures of RiPP-

modifying enzymes, a most-likely scenario exists for which the RRE of SuiB both recognizes the peptide and delivers it to the active site but at a certain stage in the catalytic cycle must release the precursor peptide. In pursuing this hypothesis two sets of methods were pursued.

Computational experiments, namely via molecular docking simulations, were used for determining a premise for SuiA-leader interactions with SuiB-RRE. Molecular docking studies were not only performed on SuiA-leader and SuiB-RRE, but also on analogous homolog pairs AgaA/AgaB and StrA/StrB. The addition of the RiPP peptide/modifying-enzyme homologs were used to widen the scope of the molecular docking simulations. The high sequence similarity of the homologs suggested that integral structural and biochemical properties between the peptides and enzymes were retained so as to maintain the integrity of the comparative study. Additionally, the viability of the molecular docking experiments was tested on a control model which already had a crystalized structure depicting interactions between the leader sequence of its precursor peptide and the RRE domain of the modifying enzyme. The molecular docking server's ability to reproduce the interactions seen in the crystal structure was the means through which the viability of the computational experiments was assessed.

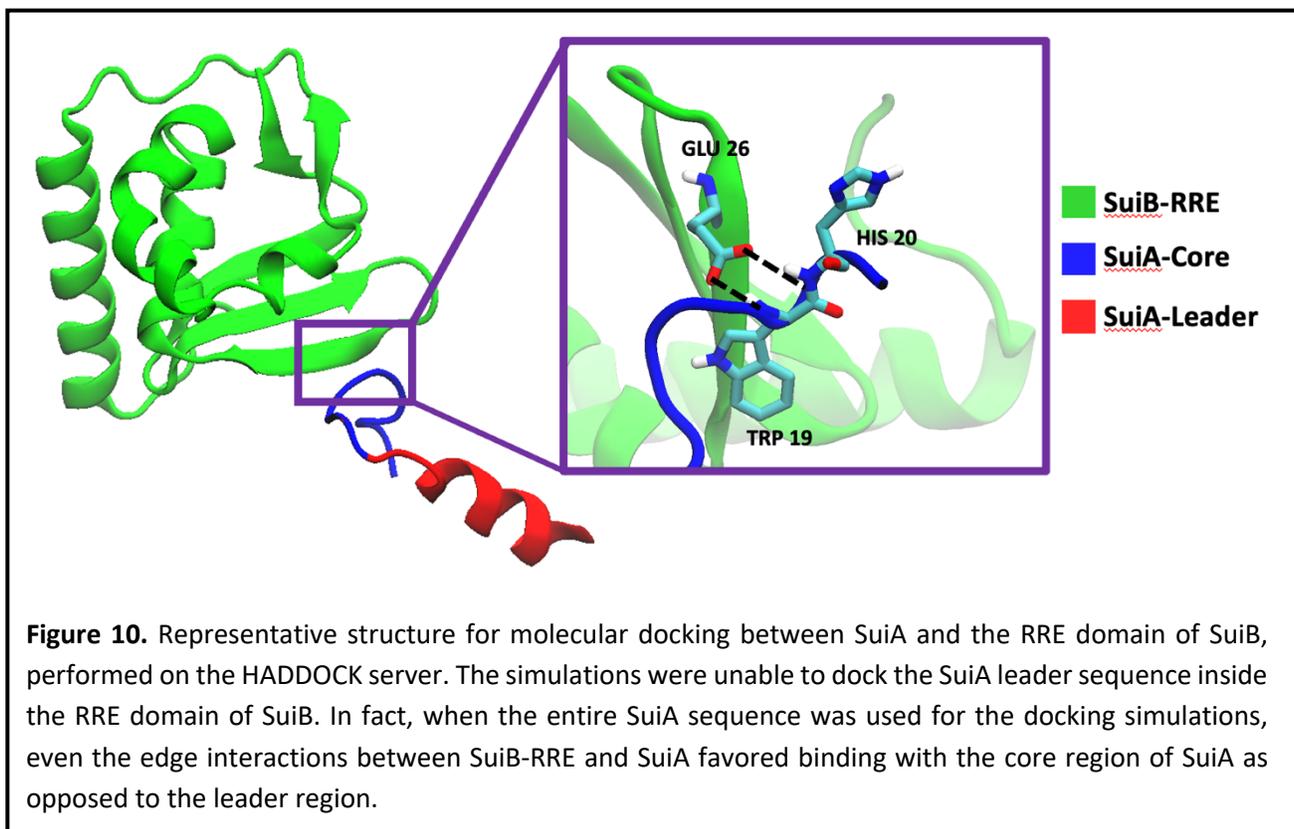
Experimental studies, aimed at elucidating more specific features of the SuiA-leader/SuiB-RRE model regarding binding interactions, were used to supplement findings from the computational studies. Circular dichroism spectroscopy was used to determine whether binding occurs between the SuiA-leader and SuiB-RRE, while also identifying the binding topology of the SuiA-leader. X-ray crystallography was used in an attempt to obtain highly accurate structural information pertaining to a SuiA-leader/SuiB-RRE complex. Even though the crystals

sent to the GM/CA beamline at Argonne National Laboratory failed to produce diffraction data, key insights into crystallization conditions were obtained for optimizing these crystals with a new biochemical basis.

3.1 Molecular Docking of SuiA-leader with SuiB-RRE and Comparison with Analogous Homologs

Molecular docking simulations were performed between SuiA and the RRE domain of SuiB on the HADDOCK server to obtain hypothetical, yet accurate, structural insight for the binding interactions of the leader sequence of SuiA within the RRE domain. The RRE domain of SuiB (PDB ID: 5V1T) was isolated from the other structural domains of the enzyme in order to maximize the conformational searches for binding interactions between the RRE domain and SuiA. Molecular docking simulations between SuiB-RRE and SuiA did not yield any model structures with the SuiA-leader sequence bound inside the RRE domain (Fig. 10). In fact, even the structural edge interactions which were observed between SuiB-RRE and SuiA favored binding with the core region of SuiA as opposed to the leader region. To obtain more directly correlated docking models between SuiB-RRE and SuiA-leader, the SuiA structure was truncated to only include the leader sequence. However, even molecular docking simulations between SuiB-RRE and the truncated SuiA-leader failed to produce models identifying a binding pocket of SuiB-RRE. Thus, from these results, three conjectures were made: (1) the SuiA-leader does not have biochemically viable interactions with the SuiB-RRE, (2) the SuiA-leader in its α -helix topology is not structurally viable for interacting with the SuiB-RRE, or (3) the computational platform for performing molecular

docking did not have a practical basis for identifying binding interactions between SuiA-leader and SuiB-RRE.

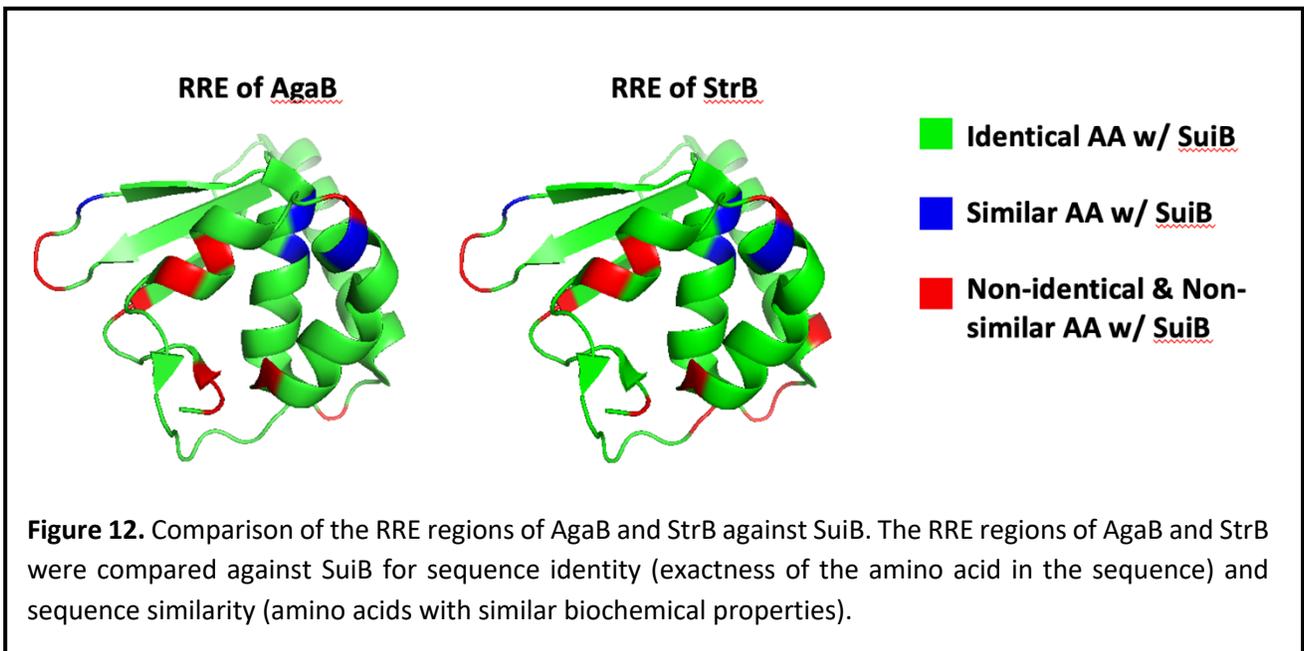
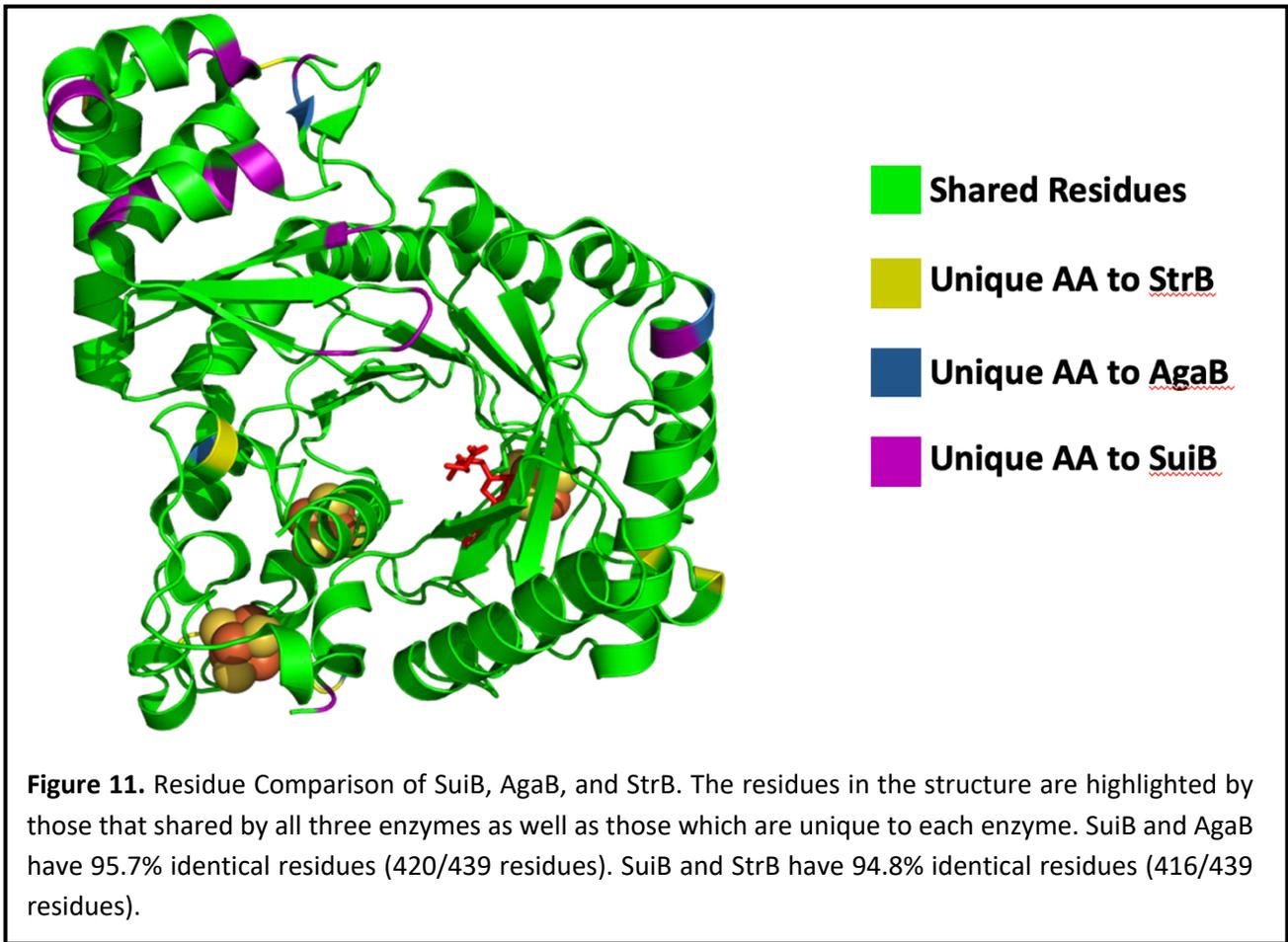


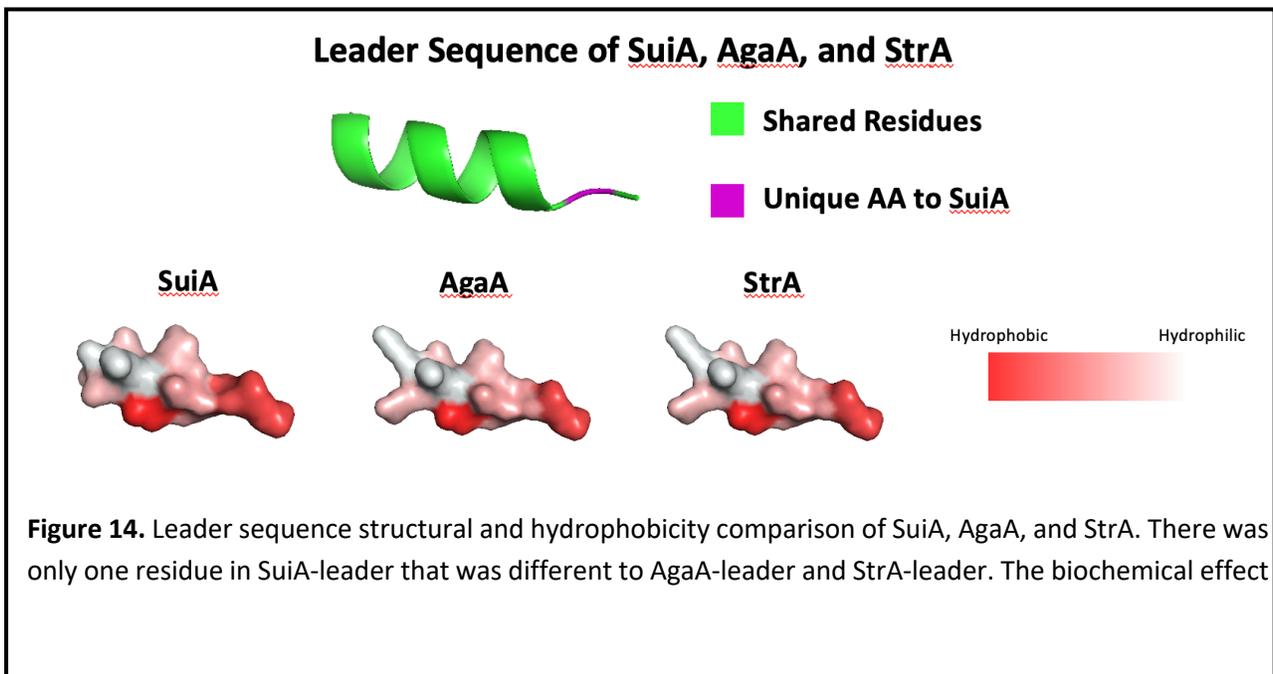
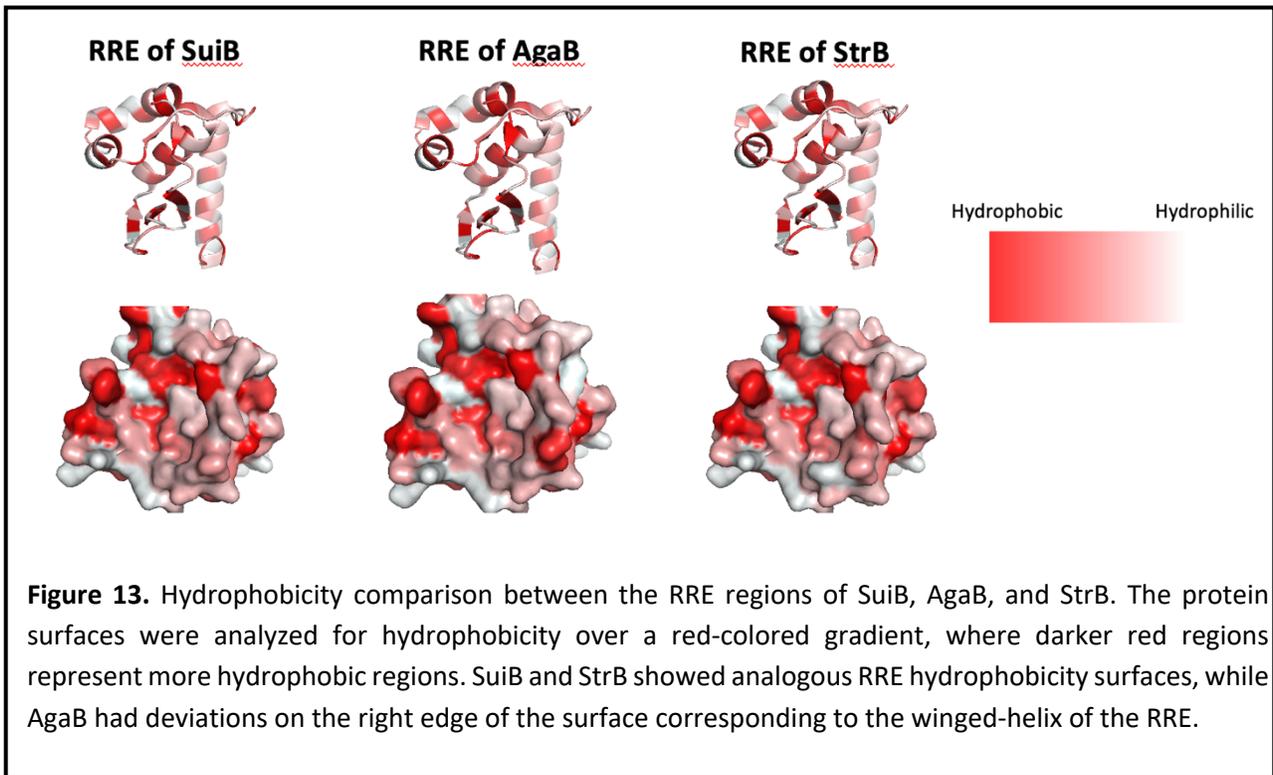
To address and minimize the wide-range of possibilities and inferences from the molecular docking simulations of SuiB-RRE to SuiA-leader, the computational study was expanded to incorporate two additional analogous homolog RiPP peptide/modifying-enzyme pairs for comparison with the study on SuiA/SuiB. Upon searching the microbial genome database, it was discovered two other *Streptococcus* pathogens also harbor the *sui*-like gene clusters for the RiPP-mediated biosynthetic production of streptide.^{15, 17} The *Streptococcus agalactiae* and *Streptococcus thermophilus* pathogens express radical SAM enzymes, AgaB and StrB, containing multiple [4Fe-4S] clusters responsible for catalyzing a Lys-Trp carbon-carbon

crosslink in their respective precursor peptides, AgaA and StrA, in a leader peptide-dependent fashion. The analogous homologs AgaA/AgaB and StrA/StrB served to supplement the molecular docking simulations on SuiA/SuiB.

However, before proceeding with molecular docking simulations on AgaA/AgaB and StrA/StrB, the structural and biochemical integrity of these SuiA/SuiB analogous homologs were studied through a sequence alignment on Clustal Omega. A structural comparison between SuiB, AgaB, and StrB revealed that SuiB and AgaB have 95.7% identical residues while SuiB and StrB have 94.8% identical residues (Fig. 11). A noteworthy observation was that there was no point in the sequences which had a different amino acid for all three enzymes. Thus, at each point of the sequences, at least two of the enzymes shared an amino acid. This was a promising observation in support of the structural homology between the three enzymes. However, the majority of the residue differences seen between the three enzymes occurred in the RRE region. Specifically, between SuiB and AgaB, of the nonidentical residues, about 74% of them were present in the RRE region. Between SuiB and StrB, of the nonidentical residues, about 70% of them were present in the RRE region. This prompted a more direct structural comparison between the RRE regions of SuiB, AgaB, and StrB (Fig. 12). The RRE regions of AgaB and StrB were compared against SuiB for sequence identity, exactness of amino acids in the sequence, and sequence similarity, biochemical similarity of amino acids in the sequence. AgaB and SuiB had an 86.8% sequence identity (92/106 residues) in the RRE region, and of the nonidentical residues, about 28.5% of them had sequence similarity (4/14 residues). StrB and SuiB had an 84.9% sequence identity (90/106 residues) in the RRE region, and of the nonidentical residues, 25% of them had sequence similarity (4/16 residues). Therefore, this result shows the structural homology of the RRE regions

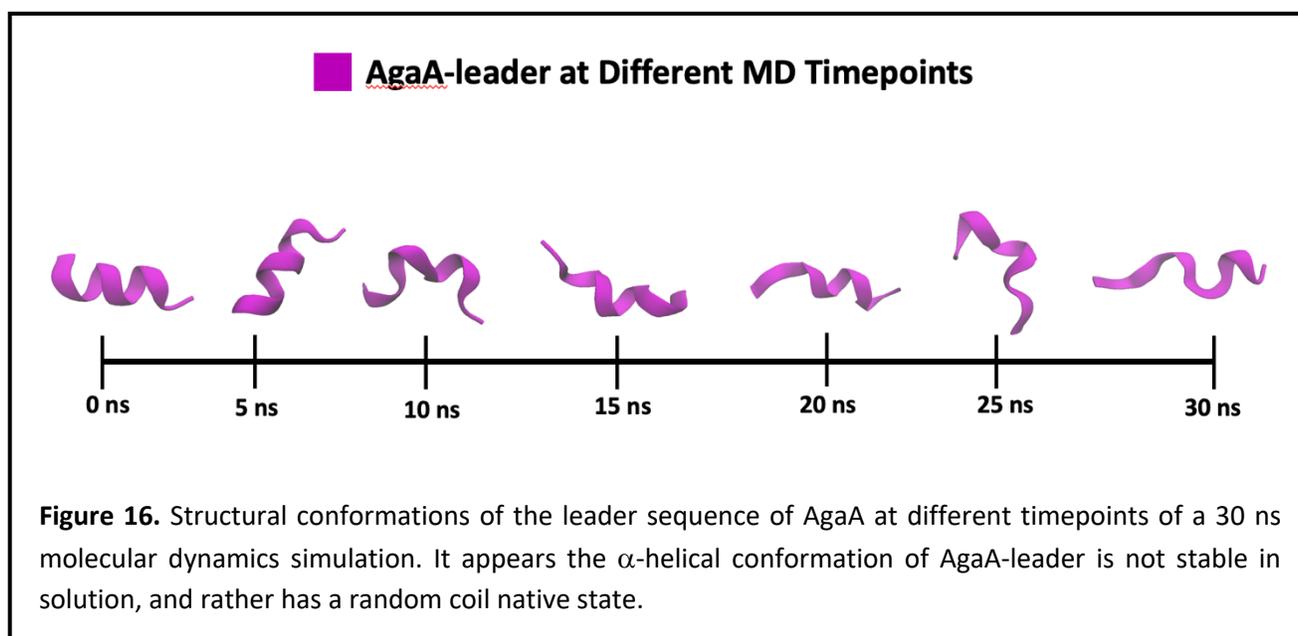
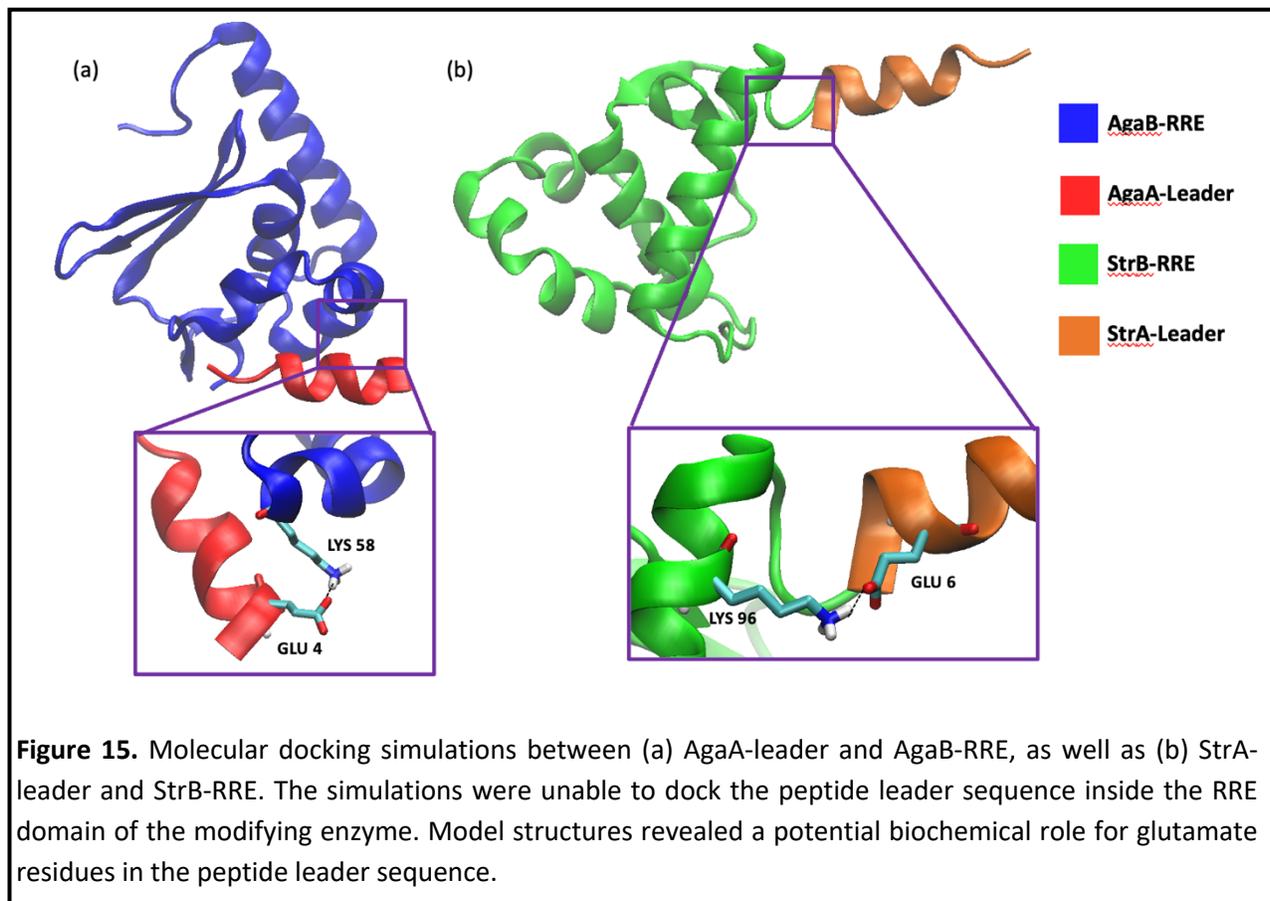
of the three enzymes is preserved not only through sequence identity but also sequence similarity. Still, to analyze the biochemical integrity of the RRE regions of the homologs due to the differences in the sequences, a hydrophobicity comparison was done (Fig. 13). The protein surfaces of the RRE regions of SuiB, AgaB, and StrB were highlighted over a color gradient to indicate the hydrophobicity of the protein based on amino acid sequences. SuiB and StrB showed analogous RRE hydrophobicity surfaces, while AgaB had deviations on the right edge of the surface corresponding to the winged-helix of the RRE. Yet, the deviations in the AgaB-RRE were only a minimal portion of the surface area. Finally, the leader sequences of SuiA, AgaA, and StrA were also compared for structural and biochemical homology (Fig. 14). It was revealed that AgaA and StrA had identical leader sequences, while SuiA only had a one residue difference between the other two peptides. However, the biochemical effect of that residue difference is negligible when placed into perspective of the leader sequence surface hydrophobicity. Although, the single residue difference in SuiA adds a more hydrophobic area to the protein surface, the overview of the surface proteins shows analogous hydrophobicity across all three peptides. Overall, the structural and hydrophobicity comparisons between the RiPP peptide/modifying-enzyme pairs SuiA/SuiB, AgaA/AgaB, and StrA/StrB, supports the notion of using AgaA/AgaB and StrA/StrB as homologs for expanding the computational molecular docking studies.





of that residue difference is negligible when placed into perspective of the leader sequence surface hydrophobicity.

Subsequently, molecular docking simulations were run on HADDOCK between AgaA-leader and AgaB-RRE as well as StrA-leader and StrB-RRE (Fig. 15). Similar to the SuiA-leader/SuiB-RRE model, the simulations were unable to dock the peptide leader sequence inside the RRE domain of the modifying enzyme. Still, model structures depicting edge interactions between the peptide leader sequence and the modifying enzyme RRE domain revealed a potential biochemical role for the glutamate residues in the peptide leader sequence. Based on the models generated from the docking simulations, GLU 4 and GLU 6 are the key residues involved in binding interactions. On a general basis, the failure of AgaA-leader/AgaB-RRE and StrA-leader/StrB-RRE to dock in a comparable manner to the SuiA-leader/SuiB-RRE model suggests support for the highly likely scenario where the SuiA-leader does not have an α -helix conformation when bound to SuiB-RRE. This hypothesis was tested by running the leader sequences of SuiA, AgaA, and StrA in a molecular dynamics simulation to observe the stability of the α -helix in solution (Fig. 16). Inherently, the peptide leader sequences denatured over time, showing preference toward a random coil conformation. Thus, the molecular docking simulations supplemented with molecular dynamics simulations endorses the assumption that the peptide leader sequences, like SuiA, are not α -helical when in solution, and the peptide leader sequences most likely bind to the RRE domain of their modifying enzymes when they are in their native conformation.



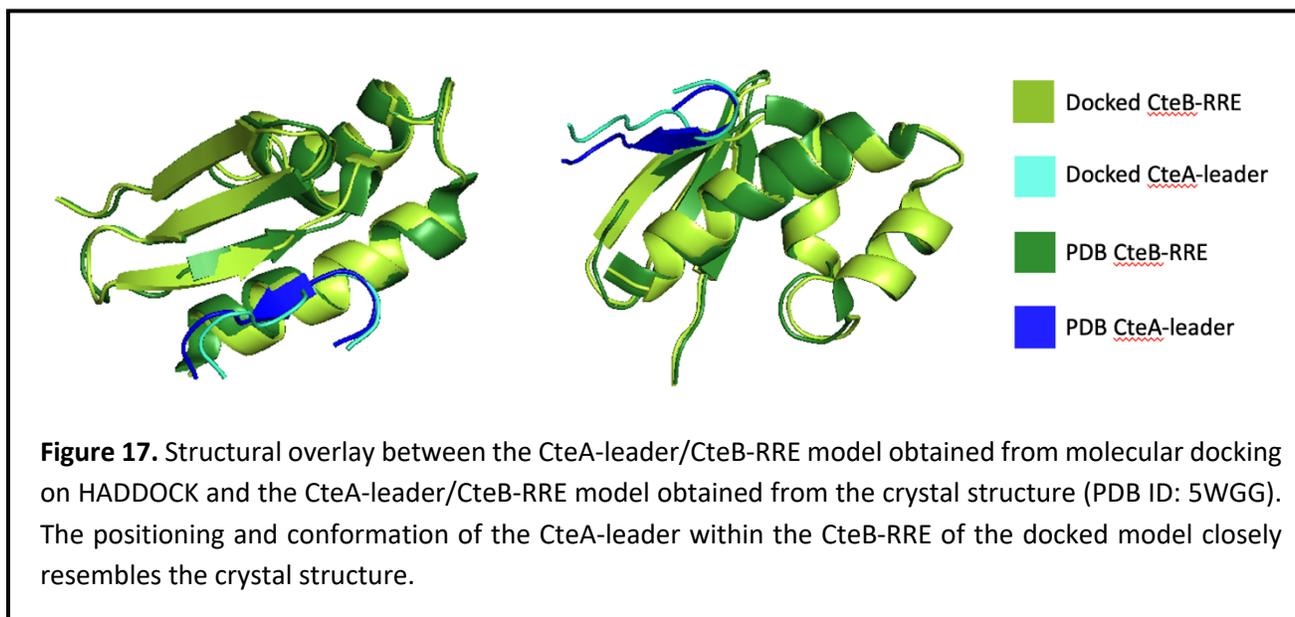
3.2 Molecular Docking of a different class of RiPP peptide/modifying-enzyme pair

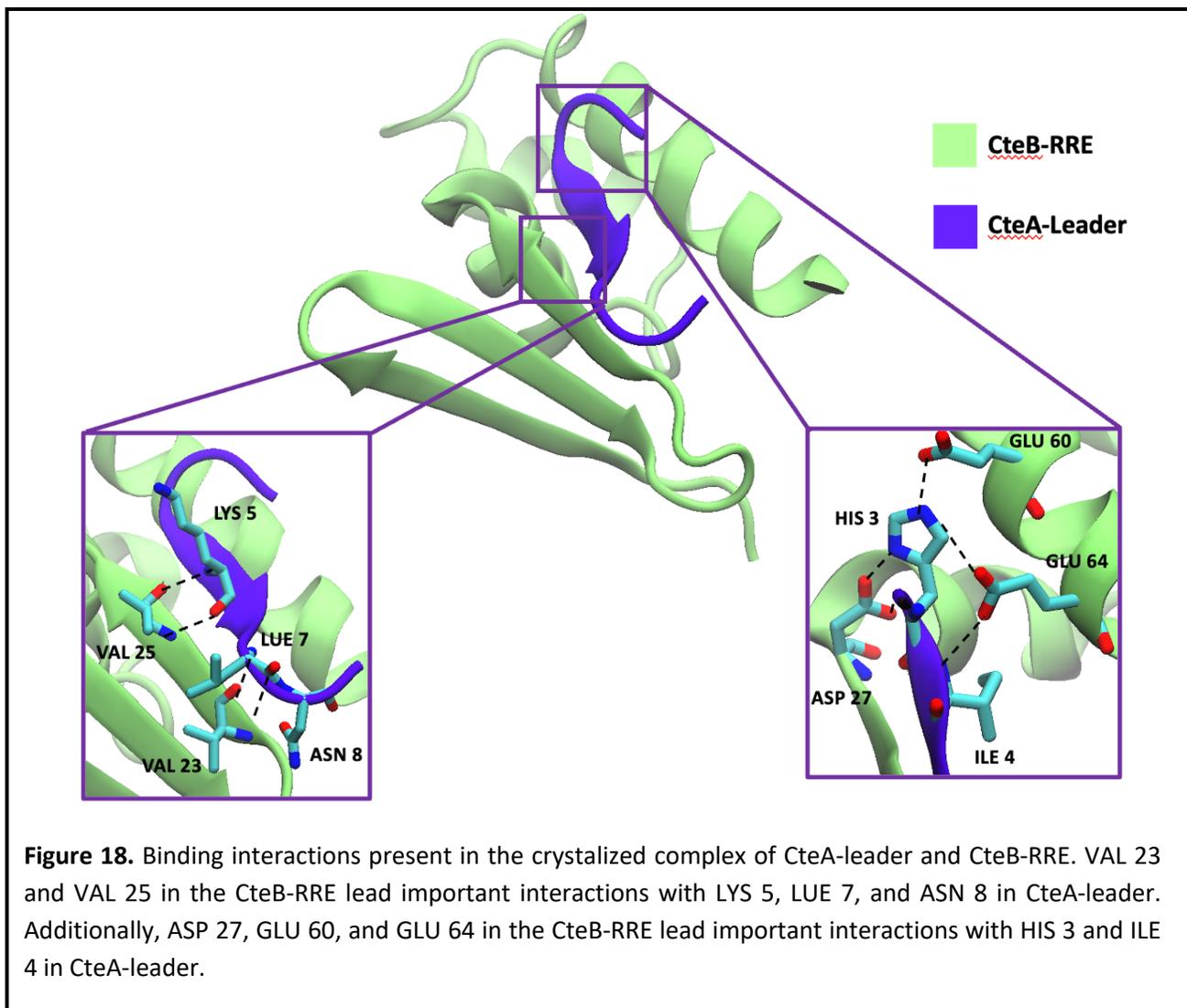
Despite the results obtained from expanding the molecular docking studies and supplementing them with molecular dynamics studies, the possibility of an incoherent computational platform for identifying RiPP peptide interactions with the RRE domain of modifying enzymes still existed. To verify the validity of the computational platform on the HADDOCK server used to perform the molecular docking simulations, a RiPP peptide/modifying-enzyme model was used in a control experiment for which a crystal structure of its peptide leader sequence bound to the RRE domain existed. The validity of the computational platform was assessed by its ability to re-create the bound conformation and interactions present in the crystal structure.

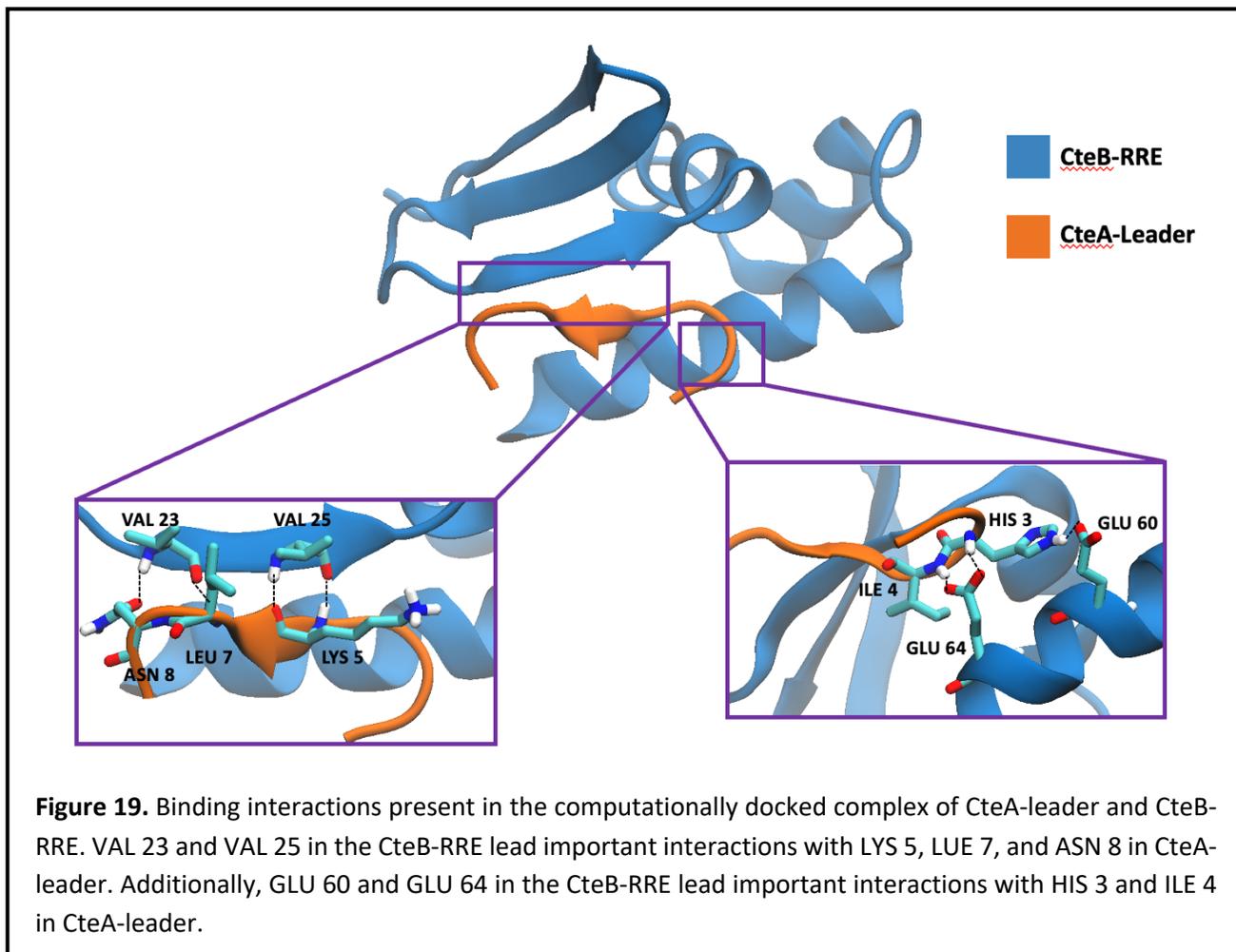
The RiPP peptide/modifying-enzyme pair CteA/CteB (PDB ID: 5WGG) were chosen as the model for this control experiment.³³ CteA and CteB are part of the gene cluster in *Clostridium thermocellum* ATCC 27405 responsible for the biosynthetic production of a sactipeptide natural product. CteB is a radical SAM enzyme that post-translationally installs a single sactioinine thioether linkage in the precursor peptide, CteA. The leader sequence of CteA has been crystallized in a complex with the RRE domain of CteB.

The CteA-leader was docked with the CteB-RRE on HADDOCK and the resulting model was structurally overlaid with the crystallized structure (Fig. 17). The positioning and conformation of the CteA-leader with respect to the CteB-RRE of the docked model closely resembled the crystal structure. The structural overlay qualitatively represents HADDOCK's

ability to accurately find and dock binding positions between RiPP precursor peptides and the RRE domain of modifying enzymes. Binding interactions between the docked model and the crystal structure were more closely analyzed for a conclusive comparison between the two (Fig. 18 & Fig. 19). The crystal structure showed VAL 23 and VAL 25 in the CteB-RRE forming important interactions with LYS 5, LUE 7, and ASN 8 in CteA-leader, as well as ASP 27, GLU 60, and GLU 64 in the CteB-RRE making important interactions with HIS 3 and ILE 4 in CteA-leader. The computationally docked complex of CteA-leader with CteB-RRE was able to replicate these interactions, except for that with ASP 27 in the CteB-RRE. Therefore, the ability of the HADDOCK computational platform to accurately re-create the binding interactions between CteA-leader and CteB-RRE demonstrates that molecular docking and the results it produces in relation to RiPP peptide/RRE interactions are valid sources of analysis.







3.3 Circular Dichroism Spectroscopy of SuiA-leader and NusA-RRE

Circular dichroism (CD) spectroscopy was used as the experimental method for confirming the SuiA-leader sequence does bind with the RRE domain of SuiB and for determining the binding conformation of the SuiA-leader in this complex. Since CD spectroscopy yields unique peaks and curvature for varying protein secondary structures, it represented an effective way for qualitatively characterizing SuiA-leader/SuiB-RRE binding interactions. The CD spectrum of the SuiA-leader/SuiB-RRE complex shows a clear rightward shift from the CD spectrum of the isolated SuiB-RRE (Fig. 20a). This shift is a strong indication that the SuiA-leader does in fact bind to the

SuiB-RRE. More evidence of this can be seen by subtracting the molar ellipticity of the isolated SuiA-leader from the complexed spectrum (Fig. 20b). The increased molar ellipticity of the peak in the complexed spectrum from 190 nm to 200 nm indicates a structural change in the RRE domain of SuiB which can be attributed to binding with the leader sequence of SuiA. Furthermore, the binding conformation of the SuiA-leader when bound with SuiB-RRE can be assessed by subtracting the molar ellipticity of the isolated SuiB-RRE from the complexed spectrum (Fig. 20c). The reduction of molar ellipticity from 210 nm to 240 nm as well as the negative molar ellipticity peak from 190 nm to 205 nm in the complexed spectrum depicts a representative state of a random coil. Hence, when SuiA-leader is bound to SuiB-RRE, it obtains a random coil topology. Additionally, the negative peak from 190 nm to 205 nm for the CD spectrum of isolated SuiA-leader suggests that SuiA-leader does not have an α -helix conformation in solution. Therefore, the results of the circular dichroism spectroscopy experiment support observations from the molecular docking and molecular dynamics studies, i.e., the α -helix of SuiA-leader denatures in solution and the binding conformation of SuiA-leader with SuiB-RRE is non-helical.

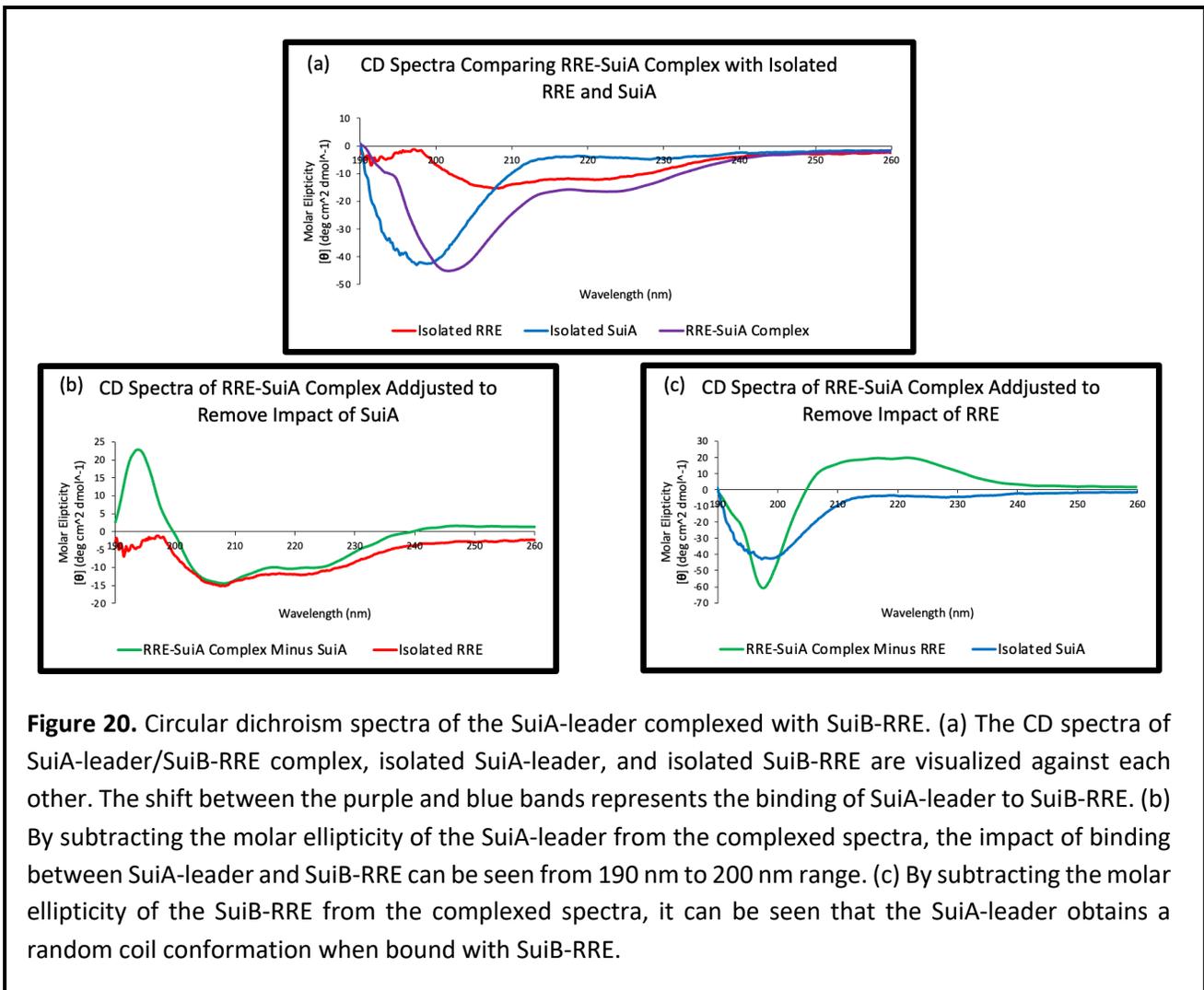


Figure 20. Circular dichroism spectra of the SuiA-leader complexed with SuiB-RRE. (a) The CD spectra of SuiA-leader/SuiB-RRE complex, isolated SuiA-leader, and isolated SuiB-RRE are visualized against each other. The shift between the purple and blue bands represents the binding of SuiA-leader to SuiB-RRE. (b) By subtracting the molar ellipticity of the SuiA-leader from the complexed spectra, the impact of binding between SuiA-leader and SuiB-RRE can be seen from 190 nm to 200 nm range. (c) By subtracting the molar ellipticity of the SuiB-RRE from the complexed spectra, it can be seen that the SuiA-leader obtains a random coil conformation when bound with SuiB-RRE.

3.4 X-ray Protein Crystallography of SuiA-leader and NusA-RRE

Crystals of the SuiA-leader complexed with the NusA-RRE were grown with the goal of using x-ray crystallography to obtain a novel structure highlighting specific interactions between the RiPP precursor peptide/modifying-enzyme pair. Across seven 96-well broad-range crystal screens, 18 crystals of SuiA-leader/NusA-RRE were grown (Fig. 21). Four of the crystallization conditions resulting in the best crystal growth were chosen for optimization. These four crystals were grown in different buffer and viscosity conditions with variations of Tris, HEPES, 1,6-

hexanediol, and polyethylene glycol in the reagents of the crystal screens. Though, notably, all four crystals were grown with the same counter ion, magnesium, in their crystal screen condition. Nine of the conditions from the four optimization screens resulted in crystal growth (Fig. 22). However, the crystals grown were particularly small, between 10-20 μm , and had poor morphology based on a lack of symmetry.

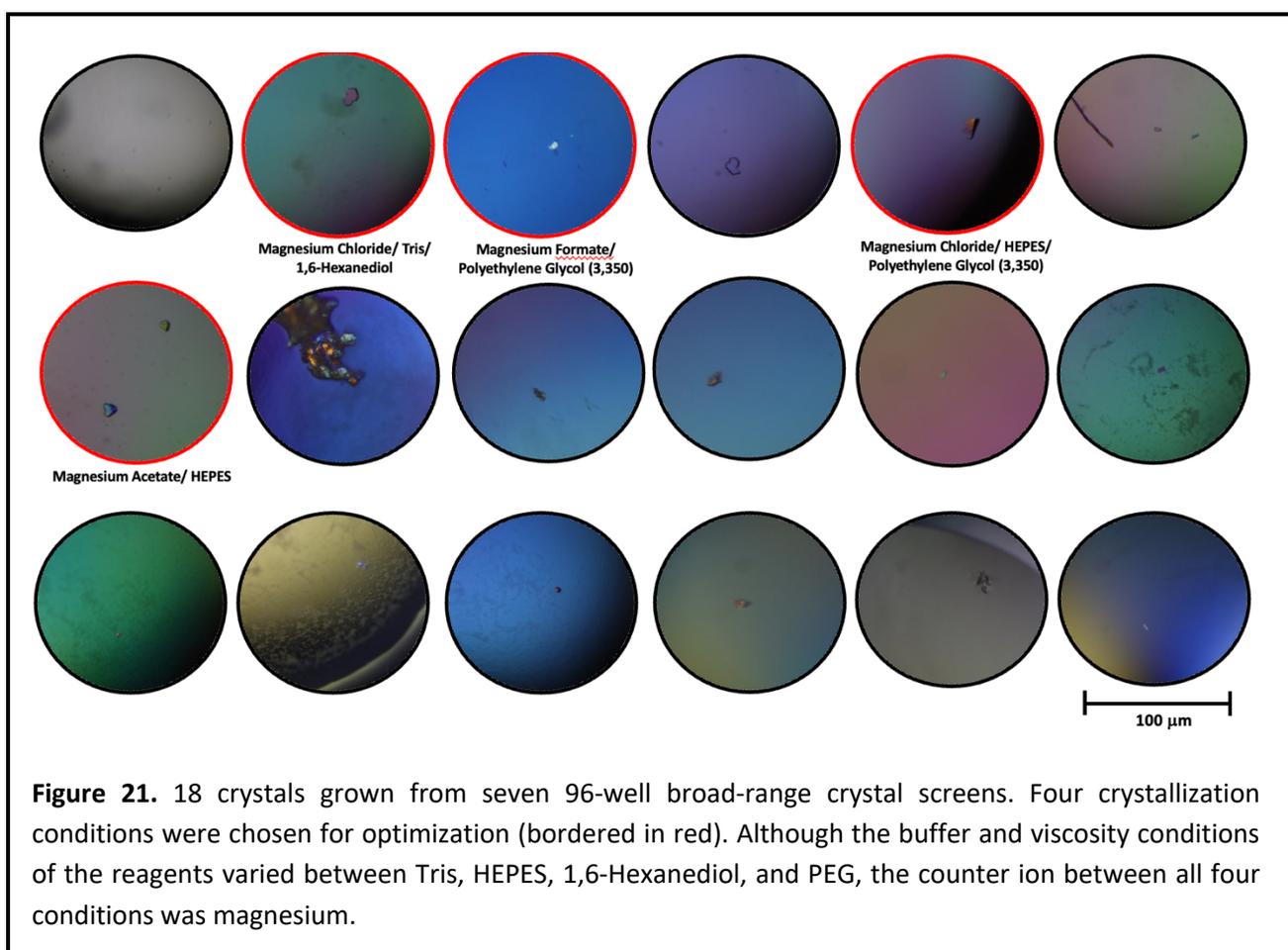
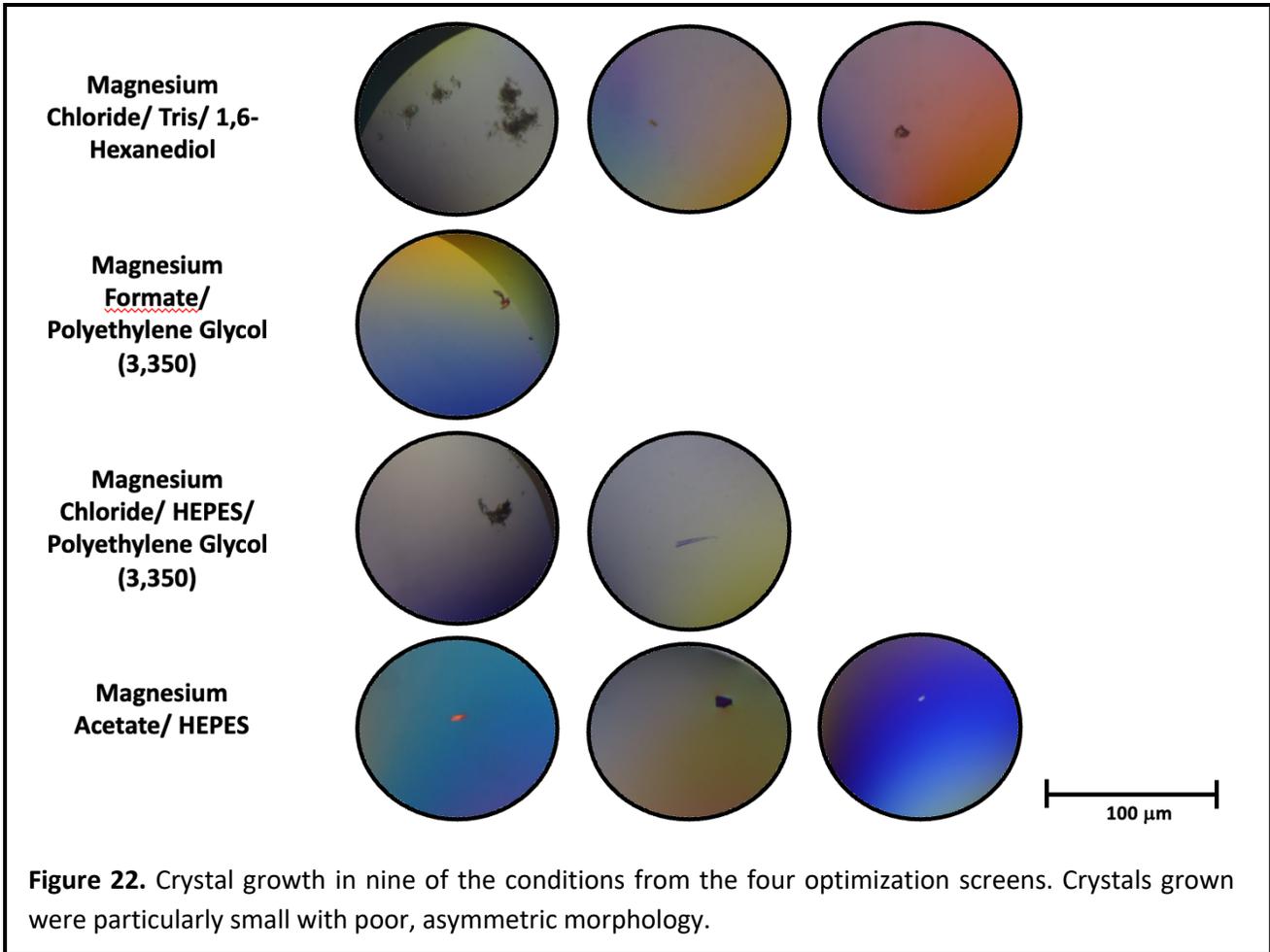


Figure 21. 18 crystals grown from seven 96-well broad-range crystal screens. Four crystallization conditions were chosen for optimization (bordered in red). Although the buffer and viscosity conditions of the reagents varied between Tris, HEPES, 1,6-Hexanediol, and PEG, the counter ion between all four conditions was magnesium.



Each of the 27 crystals grown across the broad-range and optimization screens were dyed with methylene blue to confirm they were protein crystals, and the best twelve size and shaped crystals were looped and sent to Argonne National Laboratory for x-ray crystallography. Unfortunately, all of the twelve crystals failed to yield a diffraction pattern when hit by the x-ray beam. The most probable cause for this failure is the small size of crystals and asymmetric morphology. Crystallization is the process of arranging atoms or molecules that are in a fluid or solution state into an ordered solid state, and this process occurs in two steps, nucleation and growth. Whereas nucleation of the SuiA-leader/NusA-RRE complex occurred, growth was hindered. This is most likely due to the fact that the NusA solubility tag was retained in the fusion

protein during crystallization. The large and dynamic nature of the NusA solubility tag probably prevents the protein complex from returning to an equilibrium state after reaching a supersaturated state.

The issue with the NusA solubility tag was anticipated, and a TEV protease cleavage site was designed in the sequence of the NusA-RRE fusion protein for cleaving the NusA portion. However, upon expressing and purifying the TEV protease with high purity and yield, it failed to perform the cleavage reaction (Fig. 23a). This can be seen between L6 and L7 of the SDS-PAGE gel. Although the TEV protease was present in the solution, it did not cleave the NusA solubility tag. The rationale behind this observation is that the cleavage site was not accessible for the TEV protease active site. The remedy to this dilemma was designing a new plasmid that would consider the limitations imposed by the previous plasmid (Fig. 23b). Thus, a plasmid was designed for expressing an MBP-SUMO-RRE fusion protein with two solubility tags and two Ulp1 protease cleavage sites. Since optimal crystallization conditions for nucleation have been determined from the work done on the NusA-RRE fusion protein, this new plasmid can be used to isolate the SuiB-RRE from the fusion protein solubility tags so as to induce crystal growth after nucleation in each of those conditions.

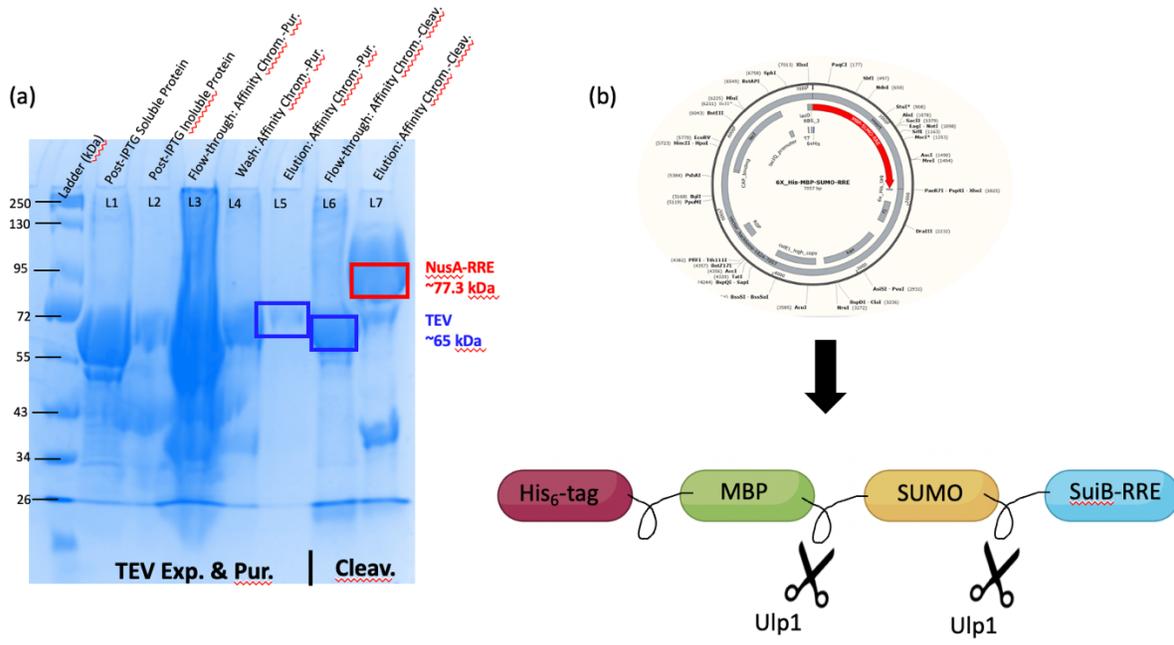


Figure 23. Failure of NusA cleavage from the RRE of SuiB. (a) Although TEV was expressed and purified in high yield and purity, it failed to cleave the NusA solubility tag from the expressed and purified NusA-RRE fusion protein. (b) A new plasmid was designed for SuiB-RRE expression incorporating the solubility tags MBP and SUMO with two Ulp1 protease cleavage sites.

4. DISCUSSION

Computational studies in the form of molecular docking simulations were used in an attempt to build models of potential interactions between the SuiA-leader and the SuiB-RRE. However, these docking simulations failed to produce a model with the SuiA-leader bound with the RRE domain of SuiB. In fact, when the entire sequence of SuiA was used in the docking simulations, the RRE domain of SuiB showed interactions with the core region of SuiA as opposed to the leader region. To assess the scope of these computational results, molecular docking simulations were run on the SuiA/SuiB analogous homologs of AgaA/AgaB and StrA/StrB. In these supplemental docking simulations, the leader regions of AgaA and StrA also failed to bind with

the RRE domains of AgaB and StrB, respectively. Though edge interactions involving GLU 4 and GLU 6 of the peptide leader regions suggested a possible biochemical role for these residues. Still, the inability to dock the peptide leader regions into the RRE domain of the modifying-enzymes suggests the α -helical conformation of the peptide leader region is not corroborative to binding with the RRE domain. Molecular dynamics simulations on the peptide leader sequence showed it denatures over time, supporting the notion that the peptide leader sequence does not exist in an α -helical topology in solution and obtains a different conformation when bound to the RRE domain. Moreover, the ability of the HADDOCK server to recreate the interactions between CteA-leader and CteB-RRE from the crystal structure validates the performance of the molecular docking simulations and the results it generated between SuiA-leader and SuiB-RRE.

Experimentally, circular dichroism (CD) spectroscopy was used to confirm analysis from the molecular docking studies. CD spectra of the complexed SuiA-leader/SuiB-RRE, isolated SuiA-leader, and isolated SuiB-RRE revealed that SuiA-leader does in fact bind with SuiB-RRE, while also the CD spectrum of the SuiA-leader when complexed with SuiB-RRE had a curvature representative of a random coil. Hence, the computational studies were supported by the CD experiment in that the α -helical conformation of SuiA-leader is not corroborative for binding with SuiB-RRE. To identify the specific interactions involved in the binding of SuiA-leader with SuiB-RRE x-ray protein crystallography was pursued as the method of choice. Disappointingly, although crystals of the SuiA-leader/SuiB-RRE were grown, they did not yield a diffraction pattern. The most likely reason is the small size and poor morphology of the crystals caused by the retention of the NusA solubility tag in the fusion protein of NusA-RRE. To address this, a new plasmid, which encodes more accessible cleavage sites, was designed with the hope of obtaining

an isolated yield of SuiB-RRE. The crystallization conditions identified from the NusA-RRE experiments will be used to regrow larger and more symmetric SuiA-leader/SuiB-RRE complexed crystals.

The RRE domain of RiPP modifying enzymes is an integral motif for the biosynthetic processing of ribosomally synthesized peptides. The evidence presented here confirms the role of the RRE domain in recognizing precursor peptides for post-translational modifications, while also increasing support for the role of the RRE domain in shuttling the core region of the peptide into the active site of the modifying enzyme. A crystal structure for the SuiA-leader/SuiB-RRE complex is still needed to prove this theory. Specifically, since the final location of the SuiA-leader in the SuiB modifying enzyme is known, by obtaining a crystal structure for the initial position of the SuiA-leader in the RRE-domain of SuiB, molecular dynamics simulations can be used to determine how the SuiA peptide is shuttled through the SuiB enzyme. Understanding the mechanism of action of RiPP biosynthetic pathways remains essential for functionalizing their substrate specificity and vast chemical moieties to develop new therapeutics.

5. CONCLUSION

Antibiotic resistant bacterial infections have become a serious concern for the scientific community and finding ways to combat them is essential for the preservation of human health. Natural products in the form of terpenoids, alkaloids, polyketides, and non-ribosomal peptides have traditionally been areas of research for addressing this concern. However, stringencies for biosynthetic production of novel analogs of these natural products has limited their capacity for therapeutic applications. Ribosomally synthesized and post-translationally modified peptides

(RiPPs) are a newly classified group of natural products which have vast structural and chemical diversity, while also showing promise in overcoming key limitations of other classes of natural products due to their modular and intrinsically tolerant biosynthetic pathways. In RiPP biosynthesis, it is widely believed that the RiPP recognition element (RRE) of a modifying enzyme identifies the leader sequence of a precursor peptide and shuttles the peptide to the active site of the enzyme for post-translational modifications in the core region of the peptide. However, the crystalized structure of the RiPP peptide/modifying-enzyme pair of SuiA/SuiB placed the leader sequence of SuiA outside the RRE of SuiB and bound to the catalytic barrel instead. This discovery posited new questions and hypotheses for the role of RRE domain in precursor peptide recognition and for the function of the leader sequence of the precursor peptide.

A combination of computational and experimental studies was performed to analyze the interactions between the leader region of precursor peptides and the RRE domains of modifying enzymes in RiPPs through the example of SuiA and SuiB. Molecular docking simulations were used in an attempt to build models of potential interactions between the SuiA-leader and the SuiB-RRE. However, the SuiA-leader failed to dock with the SuiB-RRE. Even supplemental molecular docking studies with homologous analogs of SuiA and SuiB showed a resistance for the α -helical leader region of the precursor peptide to bind to the RRE domain of the modifying enzyme. Molecular dynamics simulations on the precursor peptide showed the α -helix of the leader region denatures in solution, which suggested the binding conformation of the peptide leader sequence is nonhelical. This result was verified by circular dichroism spectroscopy, which revealed that the SuiA leader sequence does bind with the RRE domain of SuiB, but it does it in a random coil conformation as opposed to an α -helix conformation. An attempt to use x-ray

crystallography to obtain a novel structure for the SuiA-leader/SuiB-RRE complex, that would have identified the exact binding interactions between the peptide-enzyme pair, resulted in failure. However, insight into the crystallization conditions of the SuiA-leader/SuiB-RRE complex was gained for future use in this experiment. Overall, current results support the notion that the RRE domain both recognizes the peptide and delivers it to the active site but at a certain stage in the catalytic cycle releases the precursor peptide. However, obtaining a crystal structure of the SuiA-leader/SuiB-RRE complex remains imperative for proving this conjecture.

Moving forward, the immediate goal will be to regrow crystals of the SuiA-leader/SuiB-RRE complex with a refined plasmid for the expression of SuiB-RRE. A crystal structure that complexes the SuiA-leader with the SuiB-RRE will show the initial point of entry of the precursor peptide in the modifying enzyme. In addition to the existing crystal structure showing the terminal point of SuiA in the active site of SuiB, molecular dynamics simulations can be run to visualize how the RRE domain identifies and shuttles its precursor peptide to the catalytic barrel. Depending on these results, in the long term, post-translational modifications from RiPP modifying enzymes can be attempted on new synthetic peptides. All in all, natural products of RiPP biosynthetic pathways show promising signs of potential for important drug discovery.

6. REFERENCES

- 1) Falconer, S. B.; Czarny, T. L.; Brown, E. D. Antibiotics as probes of biological complexity. *Nat Chem Biol.* **2011**, *7*, 415–23.
- 2) Wender, P. A.; Miller, B. L. Synthesis at the molecular frontier. *Nature* **2009**, *460*, 197–201.
- 3) Hudson, G. A.; Mitchel D. A. RiPP antibiotics: biosynthesis and engineering potential. *Curr. Opin. Microbiol.* **2018**, *45*, 61-69.
- 4) Arnison, P. G.; *et al.* Ribosomally synthesized and post- translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, *30*, 108–160.
- 5) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106*, 3468– 3496
- 6) Felnagle, E. A.; *et al.* Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Mol. Pharm.* **2008**, *5*, 191–211.
- 7) Finking, R.; Marahiel, M. A. Biosynthesis of nonribosomal peptides1. *Annu. Rev. Microbiol.* **2004**, *58*, 453–488.
- 8) McIntosh, J. A.; Donia, M. S.; Schmidt, E. W. Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. *Nat. Prod. Rep.* **2009**, *26*, 537–559.
- 9) Ortega, M. A.; van der Donk, W. A. New insights into the biosynthetic logic of ribosomally synthesized and post-translation- ally modified peptide natural products. *Cell Chem. Biol.* **2016**, *23*, 31–44.
- 10) Walsh, C. T. Blurring the lines between ribosomal and nonribosomal peptide scaffolds. *ACS Chem. Biol.* **2014**, *9*, 1653–1661.
- 11) Oman, T. J.; van der Donk, W. A. Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **2010**, *6*, 9-18.
- 12) Davis, K. M.; Schramma, K. R.; Hansen, W. A; Bacik, J. P.; Khare, S. D.; Seyedsayamdost, M. R.; Ando, N. Structures of the peptide-modifying radical SAM enzyme SuiB elucidate the basis of substrate recognition. *PNAS* **2017**, *114* (39), 10420-10425.

- 13) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* **2015**, *11*, 564–570.
- 14) Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S-adenosylmethionine enzymes. *Chem. Rev.* **2014**, *114*, 4229–4317.
- 15) Schramma K. R.; Seyedsayamdost, M. R. Lysine-Tryptophan-Crosslinked Peptides Produced by Radical SAM Enzymes in Pathogenic Streptococci. *ACS Chem. Biol.* **2017**, *12*, 922-927.
- 16) Balo, A. R.; Caruso, A.; Tao, L.; Tantillo, D. J.; Seyedsayamdost, M. R.; Britt, R. D. Trapping a cross-linked lysine-tryptophan radical in the catalytic cycle of the radical SAM enzyme SuiB. *Biochemistry* **2021**, *118* (21), e2101571118.
- 17) Schramma K. R.; Forneris, C. C.; Caruso, A.; Seyedsayamdost, M. R. Mechanistic Investigations of Lysine-Tryptophan Cross-Link Formation Catalyze by Streptococcal Radical S-Adenosylmethionine Enzymes. *Biochemistry* **2018**, *57*, 461-468.
- 18) Chekan, J. R.; Ongpipattanakul, C.; Nair, S. K. Steric complementarity directs sequence promiscuous leader binding in RiPP biosynthesis. *Biochemistry* **2019**, *116* (48), 24049-24055.
- 19) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *7* (125), 1731-1737.
- 20) de Vries, S. J.; van Dijk, M.; Bonvin, A. M. J. J. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **2010**, *5*, 883-897.
- 21) van Zundert, G. C. P.; Rodrigues, J. P. G. L. M.; Trellet, M.; Schmitz, C.; Kastiris, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries S. J.; Bonvin, A. M. J. J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2015**, *4* (428), 720-725.
- 22) Honrnak, V.; *et al.* Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712-725.

- 23) Mahoney, M.; Jorgensen, W. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910-8922.
- 24) Sievers, F.; Wilm, A.; Dineen, DG.; Gibson, TJ; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, JD.; Higgins, D. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
- 25) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **2010**, *38* (2), 695-699.
- 26) Raran-Kurussi, S.; Waugh, D. S. Unrelated solubility-enhancing fusion partners MBP and NusA utilize a similar mode of action. *Biotechnol. Bioeng.* **2014**, *111* (12), 2407-2411.
- 27) Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **2006**, *1*, 2876-2890.
- 28) Kabsch, W. XCS. *Acta. Crystallogr. D* **2010**, *66* (2), 125-132.
- 29) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta. Crystallogr. D* **2013**, *69* (7), 1204-1214.
- 30) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of COOT. *Acta. Crystallogr. D* **2010**, *66* (4), 486-501.
- 31) Adams, P. D.; *et al.* PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta. Crystallogr. D* **2010**, *66* (2), 213-221.
- 32) Chen, V. B.; *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta. Crystallogr. D* **2010**, *66* (1), 12-21.
- 33) Grove, T. L.; Himes, P. M.; Hwang, S.; Yumerefendi, H.; Bonanno, J. B.; Kuhlman, B.; Almo, S. C.; Bowers, A. A. Structural Insights into Thioether Bond Formation in the Biosynthesis of Sactipeptides. *J. Am. Chem. Soc.* **2017**, *139* (34), 11734-11744.