EXPLORING THE BIOSYNTHESIS, FUNCTION, AND DETECTION OF THE NICKEL-PINCER NUCLEOTIDE COFACTOR

By

Jorge Luis Nevarez

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry—Doctor of Philosophy

2024

ABSTRACT

The nickel-pincer nucleotide (NPN) cofactor is an organometallic cofactor critical for the activity of lactate racemase (LarA) and other racemases and epimerases. This dissertation explores the biosynthesis, mechanistic roles, and detection strategies for the NPN cofactor, offering new insights into its biochemical and structural characteristics.

The first chapter introduces the pathway for NPN biosynthesis in *Lactiplantibacillus plantarum*, detailing the roles of LarB, LarC, and LarE enzymes. Subsequent chapters elucidate mechanisms underlying key steps of NPN assembly, including the identification of intermediates in LarB-catalyzed reactions and the exploration of sulfur transfer pathways mediated by LarE. The potential for IscS to enhance LarE activity in *Latilactobacillus sakei* is also examined, exploring alternative mechanisms for sustaining enzymatic activity.

In addressing challenges in studying NPN-dependent enzymes, novel methodologies were developed to enable detection and functional characterization. A circular dichroism spectroscopy assay was optimized for kinetic analysis of LarA and LarA homologs, while sulfonyl azide-based resins were employed to selectively isolate NPN-bound proteins from cell lysates. Additionally, efforts to detect the postulated nickel-hydride intermediate in LarA using ^{1}H-NMR spectroscopy highlight the complexities of probing transient intermediates in catalytic cycles.

This work advances our understanding of the biosynthesis, function, and detection of the NPN cofactor, providing a foundation for investigating its role in microbial metabolism and uncovering potential applications in biocatalysis and biotechnology.

ACKNOWLEDGEMENTS

Completing this dissertation has been one of the most challenging and rewarding journeys of my life, and it would not have been possible without the support, patience, and guidance of many individuals.

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Robert P. Hausinger, for his exceptional mentorship, patience, kindness, and unparalleled eye for detail. Your dedication to my growth as a scientist and your thoughtful guidance throughout my doctoral studies have been instrumental in shaping both my research and my career. I am also sincerely grateful to my co-advisor, Dr. Jian Hu, for his valuable feedback and support, which have contributed significantly to this work.

To my colleagues in the Hausinger Lab—Shramana Chatterjee, Santhosh Gatreddi, Joel Rankin, Yali Wang, and Aiko Turmo—thank you for your camaraderie, collaboration, and support. Working alongside you has been a privilege, and your input has been vital to the success of this work. I am also grateful to my undergraduate mentees, Aisha Narayan and Abigail Thompson, for their enthusiasm and dedication to our projects. I would like to acknowledge my dissertation committee members, Dr. Babak Borhan, Dr. Heedeok Hong, and Dr. Kin Sing Lee, for their constructive feedback, encouragement, and insightful suggestions. Your expertise and guidance have greatly enriched this research. I am deeply appreciative of my undergraduate mentors, Jim Horn and Dr. Dionne Griffin, who first inspired me to pursue a career in science. Your mentorship and encouragement were pivotal in igniting my passion for research, and I will always be grateful for your guidance.

A special and heartfelt thanks goes to my parents, Rafael and Maria Nevarez, whose unwavering support and inspired resilience have been my greatest source of strength. Your sacrifices, love, and belief in me have carried me through every challenge and made this achievement possible. I also extend my gratitude to my family and friends for their encouragement throughout this journey.

This dissertation is not only a reflection of my efforts but also the collective support, mentorship, and kindness I have received from everyone mentioned here. Thank you all for being a part of this milestone in my life.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ACN — acetonitrile

AMP — adenosine monophosphate

ATP — adenosine triphosphate

BME — 2-mercaptoethanol

BSA — benzenesulfonyl azide

CD — circular dichroism

CMP — cytidine mononucleotide

CMPylated — cytidinylylated

CoA — coenzyme A

CoA-S-S — coenzyme A persulfide

CTP — cytidine triphosphate

DaAD — dicarboxynicotinic acid adenine dinucleotide

DBU —1,8-diazabicyclo[5.4.0]undec-7-ene

DFT — Density functional theory

D-GPT — D-Glutamate-pyruvate transaminase

Dha — dehydroalanine

DIT — digital integration time

DTT — dithiothreitol

EDNMR — Electron-electron double resonance

EIC — extracted ion chromatograms

EPR — Electron paramagnetic resonance

ESI — electrospray ionization

FTIR — Fourier-transform infrared spectroscopy

HEPES — (2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid)

ICP-OES — inductively coupled plasma optical emission spectrometer

IPTG — isopropyl β-D-1-thiogalactopyranoside

$IscS_{Ec}$— cysteine desulfurase from *Escherichia coli*

$IscS_{Ls}$— cysteine desulfurase from *Lactococcus sakei*

KIE— kinetic isotope effect

$LarA_{Lp}$— LarA from *Lactiplantobacillus. Plantarum*

$LarB_{Lp}$— LarB from *Lactiplantobacillus. plantarum*

$LarC_{Lp}$— LarC from *Lactiplantobacillus. plantarum*

$LarE_{Mm}$— LarE from *Methanococcus maripaludis*

$LarE_{Ls}$— LarE from *Latilactobacillus sakei*

$LarE_{Tm}$— LarE from *Thermotoga maritima*

LC-MS — liquid chromatography-mass spectrometry

L-Cys — L-cysteine

LDH — lactate dehydrogenase

LRSA — lissamine rhodamine B sulfonyl azide

NaAD — nicotinic acid adenine dinucleotide

NaMN — nicotinic acid mononucleotide

NMN — nicotinamide mononucleotide

NMR — nuclear magnetic resonance

NPN — nickel-pincer nucleotide

P2CMN — pyridinium 3,5-biscarboxylic acid mononucleotide

P2TMN — pyridinium-3,5-bisthiocarboxylic acid mononucleotide

PBSA — polymer bound sulfonyl azide

PCET — proton-coupled electron transfer

PCHT — the proton-coupled hydride transfer

PCTMN — pyridinium-3-carboxy-5-thiocarboxylic acid mononucleotide

PLP — pyridoxal phosphate

PMSF — phenylmethylsulfonyl fluoride

PPi — pyrophosphate

SBSA — silica-bound sulfonyl azide

TFA — trifluoroacetic acid

TIC — total ion chromatogram

TOF — time-of-flight

CHAPTER 1

Introduction: Biosynthesis of the nickel-pincer nucleotide cofactor and the mechanism of lactate racemase, LarA

This chapter includes portions that were adapted with permission from text originally published in:

Nevarez, J. L., Turmo, A., Hu, J., Hausinger, R.P. (2020) Biological Pincer Complexes, *ChemCatChem* 12, 4242-4254.

Introduction

Biosynthesis of the nickel-pincer nucleotide cofactor

Genetic studies identified three accessory genes (*larB*, *larC*, and *larE*) that are required for lactate racemase activity as well as being co-localized and co-regulated with the enzyme structural gene, *larA*, in *Lactiplantibacillus plantarum*.[1] Subsequent investigations revealed the pathway by which the corresponding gene products function in the biosynthesis of the novel SCS-type Ni(II) nickel-pincer complex (**Figure 1-1**).

The NPN cofactor is derived from nicotinic acid adenine dinucleotide (NaAD) with LarB catalyzing the first step: carboxylation of C5 on the pyridinium ring while hydrolyzing the phosphoanhydride with release of adenosine monophosphate (AMP) to from pyridinium-3,5-dicarboxylic acid mononucleotide (P2CMN).[2] Carboxylation using CO_2 as the substrate was suggested to occur first, followed by hydrolysis of the dicarboxylated dinucleotide to promote P2CMN release. Conversely, direct hydrolysis of NaAD with release of nicotinamide

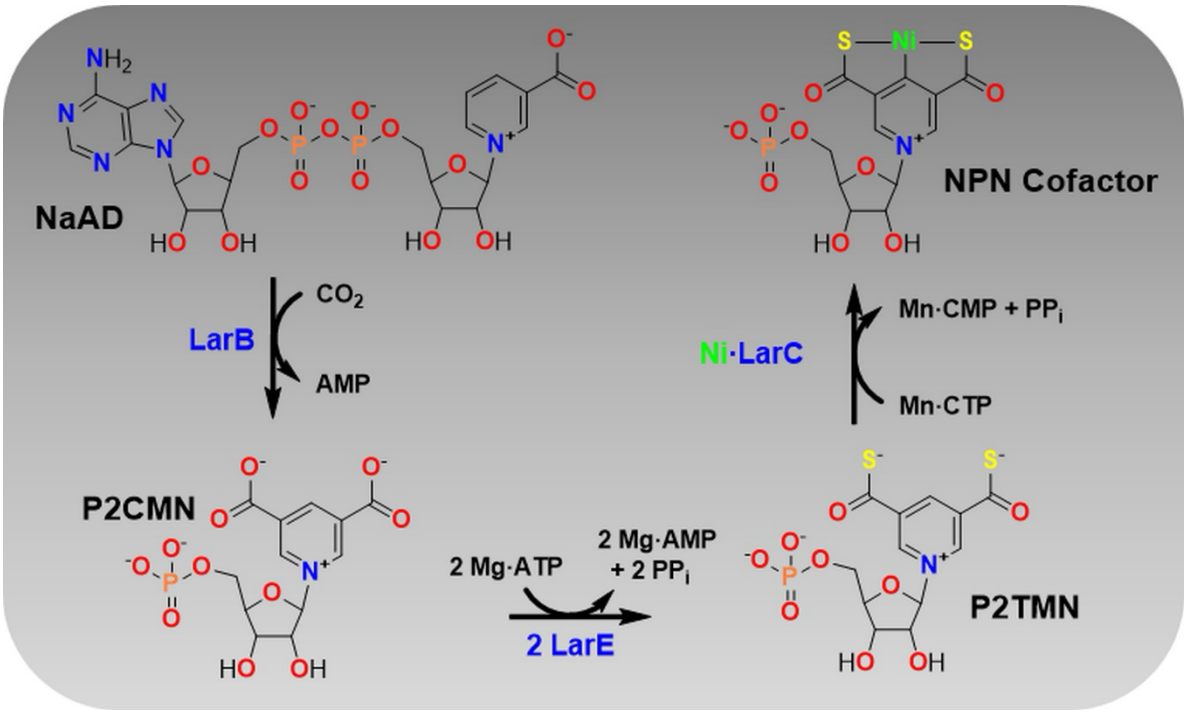

Figure 1-1. Biosynthesis of the NPN cofactor in *L. plantarum*. LarB initiates the pathway by catalyzing the carboxylation/hydrolysis reaction of NaAD to produce P2CMN. Two LarE subunits each sacrifice a cysteine residue sulfur atom while catalyzing an ATP-dependent reaction to generate P2TMN. Finally, LarC inserts nickel by a CTP-dependent reaction to form the NPN cofactor. NaAD = nicotinic acid adenine dinucleotide, P2CMN = pyridinium-3,5-dicarboxylic acid mononucleotide, P2TMN = pyridinium-3,5-dithiocarboxylic acid mononucleotide, NPN = nickel-pincer nucleotide.

mononucleotide (NMN) and AMP also occurs.[3] No external energy source is required for this reaction and it had been speculated that the energy released from the hydrolysis of NaAD is used for the carboxylation reaction.[2]

In the second stage of the NPN biosynthetic pathway, two molecules of *L. plantarum* LarE each sacrifice a cysteine side chain sulfur atom, forming dehydroalanine (Dha) residues, while sequentially converting the P2CMN carboxyl groups into thioacids, thus forming pyridinium-3-carboxy-5-thiocarboxylic acid mononucleotide (PCTMN, not depicted) and then pyridinium-3,5-dithiocarboxylic acid mononucleotide (P2TMN).[2,4] On the basis of structural and mechanistic studies, each sulfur transfer reaction involves (i) ATP-dependent activation of a substrate carboxyl group by adenylylation with release of pyrophosphate, (ii) cysteine residue attack on the activated substrate to form a thioester with release of AMP, (iii) deprotonation of the cysteine Cα position, and (iv) sulfur transfer to form the product thioacid. This type of sacrificial sulfur transfer resulting in a Dha residue is known to occur in only one other enzyme, thiamine thiazole synthase (THI4) from *Saccharomyces cerevisiae*.[5] The Dha-containing form of *L. plantarum* LarE ($LarE_{Lp}$) is capable of being recycled *in vitro* by incubation with the persulfide of coenzyme A (CoA-S-S^-) followed by addition of a reductant.[6] In this recovery reaction, the highly nucleophilic persulfide adds to the Dha residue yielding a CoA-S-S-LarE mixed disulfide that subsequently undergoes reduction. It is unclear whether Dha recycling is physiologically relevant; however, CoA-SH binds to and stabilizes LarE.[2,6] In addition to the LarE adducts with P2CMN and PCTMN formed during the sulfur transfer reactions, there is evidence for a LarE adduct of NPN, suggesting that nickel can insert into P2TMN, while covalently bound to LarE.[6] This result explains why large amounts of isolated *L. plantarum* LarE, when purified from cells that co-produce LarB and LarC, can activate LarA apoenzyme.[1]

In contrast $LarE_{Lp}$, most LarE homologs contain three conserved cysteine residues which are likely involved in binding a [4Fe-4S] cluster. Studies of LarE from *Thermotoga maritima* ($LarE_{Tm}$) and *Methanococcus maripaludis* ($LarE_{Mm}$) confirm that these cysteine residues bind to a [4Fe-4S] cluster, suggesting an alternative mechanism for sulfur insertion in these cases.[7,8] Mass spectrometry data revealed the formation of a [4Fe-5S] species on $LarE_{Tm}$ in the presence of L-cysteine and cysteine desulfurase from *Escherichia coli* ($IscS_{Ec}$). Additionally, it was shown that the [4Fe-5S] species diminishes in the presence of the substrate, supporting a model where

LarE accepts a sulfide onto the non-cysteinyl ligand-bound iron atom of the [4Fe-4S] cluster and transfers the noncore sulfide to the activated substrate.[7]

In the terminal step of NPN cofactor biosynthesis, LarC installs nickel by forming new nickel-carbon and nickel-sulfur sigma bonds.[9] This capability makes LarC the first enzyme identified to catalyze a cyclometalation reaction.[10–12] While the detailed molecular mechanism of the enzymatic reaction remains unclear, the recent detection of a cytidinylylated (CMPylated) P2TMN adduct supports a two-step process, where the substrate P2TMN is CMPylated, releasing pyrophosphate (PP_i) and facilitating the correct positioning of the pyridinium ring for nickel insertion.[13] Previous activity assays demonstrated that LarC is a CTP-dependent enzyme and x-ray crystallography of its C-terminal domain revealed a novel nucleotide-binding site.[9] Site-directed mutagenesis confirmed that the CTP-binding residues are essential for activity.[9] Interestingly, LarC appears to function stoichiometrically rather than catalytically, indicating that it operates as a single-turnover enzyme.[9]

Once synthesized, the NPN cofactor incorporates into the lactate racemase protein. In the case of *L. plantarum*, one thioacid of the cofactor covalently links to a lysine residue as a thioamide.[14] By contrast, the lactate racemase of *Thermoanaerobacter thermosaccharolyticum* binds the cofactor non-covalently to the protein.[1] The apoprotein of this species has been useful for monitoring NPN cofactor biosynthesis based on its ability to confer lactate racemase activity.[2,9]

Characterization of the NPN cofactor biosynthesis pathway has primarily focused on enzymes from *L. plantarum*; however, an analysis of over 1,000 bacterial and archaeal genomes reveals that approximately 9% contain genes that encode homologs of the LarA and NPN cofactor biosynthetic proteins.[1] Investigating these homologs might lead to the discovery of alternative enzymes with distinct catalytic properties for generating the NPN cofactor. Furthermore, genome analyses have identified natural fusions of some of the biosynthesis enzymes, possibly allowing for channeling of the pathway in nature. Interestingly, around 15% of the genomes lack a homolog of *larA*, but still possess homologs of *larB*, *larC*, and *larE*, suggesting that the NPN cofactor is synthesized for purposes other than lactate racemization.[1] Indeed, a superfamily of LarA homologs has been identified with several assayed for activity on a variety of simple 2-hydroxy acids.[15] Notably, 7 of the studied homologs have demonstrated the ability to racemize substrates such as L-malate, L-hydroxyglutarate, D-mannonate, and L-phenyllactate.[15] Through

genomic context and sequence analysis, these homologs have been classified into 25 distinct categories.[15]

Mechanistic role of the NPN cofactor in lactate racemization

The unique structure of the NPN cofactor raises the question of how it functions in racemase catalysis. Efforts to investigate the role of the NPN cofactor include reactivity analyses for synthetic model complexes, computational studies, and direct characterization of the enzyme.

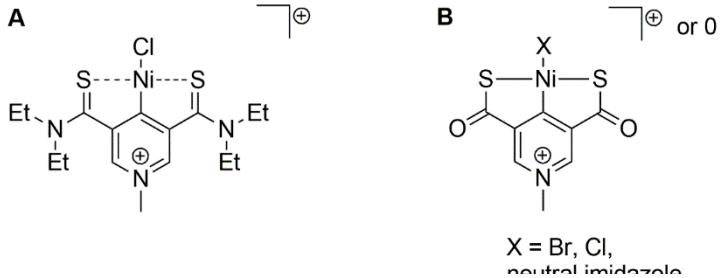

Figure 1-2. Structures of two synthetic nickel-pincer complexes resembling the NPN cofactor.

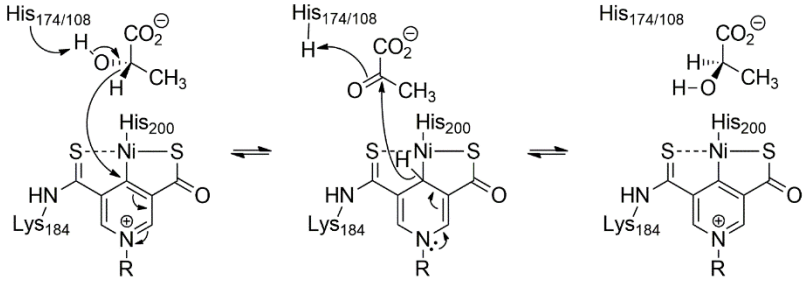

Figure 1-3. Postulated proton-coupled hydride transfer mechanism of lactate racemase.

The following paragraphs discuss each of these topics.

To gain insight into the function of the NPN cofactor of LarA, the reactivities of synthetic analogs were studied by Prof. Xile Hu and colleagues. The first-generation NPN analogue (Figure 1-2A), containing 3,5-bisdiethyl-thio-amide modification of an N-methylpyridinium ring, was capable of irreversibly dehydrogenating primary and secondary alcohols in the presence of a strong base (1,8-diazabicyclo[5.4.0]undec-7-ene, DBU), but was devoid of racemase activity.[16] The second-generation NPN mimic (**Figure 1-2B**), containing two thioacids instead of thioamides, successfully racemized lactic acid when combined with two equivalents of DBU.[17] Whereas LarA from *L. plantarum* exhibits a k_{cat} of 4,745 ± 544 s^{-1},[1] lactate racemization by the NPN mimic was very slow with a turnover number of 3 after 36 h.[17] The anionic sulfur ligands in the second-generation analog likely resulted in a more stable pincer complex

compared to the neutral sulfur atoms of the thioamides in the first-generation NPN mimic. The ability of an NPN analog to perform lactate racemization is quite astounding considering the lack of protein residues to orient the substrate and stabilize the transition states.

Density functional theory (DFT) calculations accompanying the above reports on the synthesis of nickel-pincer complexes[16,17] along with separate computational studies on 63-atom, 139-atom, 200-atom, or other models of the NPN cofactor[18] are consistent with a proton-coupled hydride transfer (PCHT) mechanism for lactate racemization. The PCHT mechanism envisions abstraction of the substrate hydroxyl proton by a general base concomitant with hydride transfer from the α-carbon of lactate to the cofactor, with the hydride returning to either face of the pyruvate intermediate (**Figure 1-3**). This mechanism was proposed when the NPN cofactor was discovered,[14] and that report noted two histidine residues flanking the active site; each of these residues may function as the general base depending on the starting enantiomer of lactate. Calculations using the B3LYP or similar level of theory for the synthetic complexes and models predict the transition state energy for lactate racemization ranges from 12.0 to 30 kcal/mol. For example, computations involving dehydrogenation of benzyl alcohol by the model compound shown in **Figure 1-2A** were consistent with DBU-catalyzed deprotonation of the alcohol simultaneous with hydride transfer to C2 or C4 of the pyridinium group with a transition energy

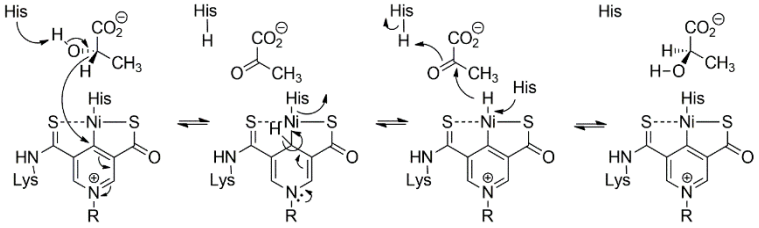

Figure 1-4. Alternative hypothesis for a PCHT mechanism in which the hydride can exist at two sites on the NPN cofactor and providing a rationale for the complexity of the NPN structure.

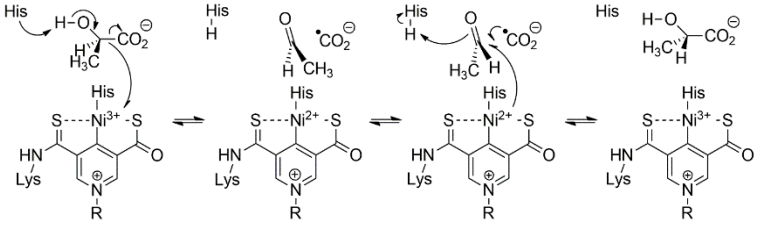

Figure 1-5. Hypothetical proton-coupled electron transfer (PCET) mechanism for LarA.

of 26 or 26.5 kcal/mol.[16] Similar results were obtained for DFT computations involving the functionally active nickel-pincer complex of **Figure 1-2B**; i.e. base-catalyzed deprotonation of the hydroxyl group combined with hydride transfer to C4 of the ring yielded a transition energy of 29.8 kcal/mol.[17] Similar analyses of lactate racemase models suggest the cofactor stabilizes the transition state and predicts lower activation barriers for transfer of the hydride to C4, while destabilizing the calculated pyruvate/NPN-hydride intermediate relative to the reactants.[18–20] Because nickel-hydride pincer complexes also are well known,[21] it is important to consider whether PCHT could provide an intermediate with hydride coordinating the metal. Calculations for the complexes in **Figure 1-2** ruled out a pathway involving a nickel-alkoxide intermediate leading to a nickel-hydride species because the transition state barrier was inaccessible (>40 kcal/mol).[16,17] Similarly, two of the model studies suggested a nickel-hydride intermediate was unlikely.[19,20] We note, however, that the metal remained coordinated to the histidine ligand in those calculations. By assuming the histidine residue can dissociate from the metal during the reaction, calculations revealed that hydride transfer to the nickel center was only 8.4 kcal/mol higher in energy than hydride transfer to C4.[22] A variation of the PCHT mechanism was proposed in which the cofactor can accept a hydride at either C4 or nickel, depending on the isomer of lactate provided, and exchange between these sites within the cofactor facilitates the racemization reaction (**Figure 1-4**). It may be possible to test this hypothesis by using ^{1}H-nuclear magnetic resonance (NMR) spectroscopy to detect the presence in LarA of a nickel-hydride, associated with a resonance upfield of tetramethylsilane.[21] A distinct mechanism for LarA, proton-coupled electron transfer (PCET), was proposed on the basis of quantum mechanics/molecular mechanics calculations.[23] This mechanism also proposes abstraction of the hydroxyl proton of lactate by a general base; however, it suggests this step occurs in concert with the homolytic cleavage of its C1-C2 bond and a 1-electron reduction of Ni(III) to Ni(II), yielding acetaldehyde and a carboxylate radical as intermediates (**Figure 1-5**). Rotation about the carbonyl bond of acetaldehyde followed by Ni(II) oxidation and reconstitution of the C1-C2 bond permits either isomer of lactate to form. The calculated energy barrier of hydroxyl group deprotonation and C1-C2 cleavage was 3.1 kcal/mol, whereas that for reconstituting the C1-C2 bond and reforming lactate was 15.5 kcal/mol.

Experimental studies provided evidence related to the PCHT and PCET mechanisms.[22] The key points distinguishing these proposals are the nickel redox state, the identity of the intermediate,

and the substrate bond that is broken and then reformed during the reaction. Whereas the PCHT mechanism maintains Ni in the +2 state, PCET stipulates that the resting Ni is +3 (doublet state), and reduced to +2 (singlet state) during catalysis. Electron paramagnetic resonance (EPR) spectroscopy was used to distinguish between these possibilities; no EPR signal is expected for Ni(II), whereas Ni(III) typically exhibits an EPR signal in the range of g = 2.4-2.0. Continuous-wave X-band EPR experiments conducted on purified LarA, with and without substrate, produced no signal in the 150-450 mT region scanned, thus favoring the PCHT mechanism. Further evidence supporting PCHT over PCET was obtained through efforts to detect the predicted intermediate of the reaction: pyruvate or acetaldehyde, respectively. Pyruvate was detected by coupling the peroxide produced via pyruvate oxidase with the horseradish peroxidase catalyzed oxidation of 10-acetyl-3,7-dihydroxyphenoxazine (Amplex Red) that was detected using fluorescence spectroscopy. In addition, pyruvate was shown to form by liquid chromatography-mass spectrometry (LC-MS) in quenched reactions of purified LarA$_{Lp}$ and pyruvate-free lactate after derivatization with 2,4-dinitrophenylhydrazine. Notably, the amount of pyruvate detected was 5-30 mole percent of the enzyme. To distinguish whether the C-H or C-C bond is broken, kinetic isotope effect (KIE) studies were carried out using lactic acid with deuterium incorporated at the C2 position. This experiment was performed originally using cell-free lysates from *Clostridium butylicum*, where lactate racemization was associated with KIEs of 2.16 and 2.14 for the D- and L-isomers, respectively.[24] More recent results compared the reaction rates of purified LarA$_{Lp}$ using 2-^{2}H-L-lactate versus non-labeled L-lactate, and resulted in a k_H/k_D of 3.11 ± 0.17.[22] This value is within the expected range (KIE of 3-5) of a hydride transfer in NAD^+ enzymes. Thus, the KIE evidence supports a PCHT mechanism, which predicts C-H cleavage, over the PCET mechanism, which predicts no C-H cleavage.

Conclusion and issues addressed in this thesis

In summary, the biosynthesis of the nickel-pincer nucleotide (NPN) cofactor and its role in lactate racemase activity exemplify a complex interplay of enzymatic processes and metal coordination chemistry. The intricate mechanisms of LarA, supported by both experimental and computational studies, highlight the unique functionality of the NPN cofactor in facilitating lactate racemization through proton-coupled hydride transfer. The evolutionary significance of the *lar* gene cluster, alongside the identification of homologs across diverse microbial taxa, underscores the potential for discovering alternative enzymatic pathways with distinct catalytic

properties. Continued exploration of these mechanisms not only enhances our understanding of lactate racemization but also opens avenues for biotechnological applications in metabolic engineering and synthetic biology. Future research will be pivotal in elucidating the full spectrum of NPN cofactor functions and its implications in broader biological contexts.

In Chapter 2, I address several hurdles that have impeded progress in studying NPN cofactor-containing enzymes. To avoid the need to purify such enzymes from their native hosts or from a heterologous expression system involving *Lactococcus lactis*, which exhibits inconsistent results, I participated with others to utilize the Duet expression system for producing active enzyme homologs in *Escherichia coli*. Of particular interest, I used this system to show that cyanobacterial homologs of LarB, LarE, and LarC can synthesize the NPN cofactor that activates $LarA_{Lp}$, even though these phototrophs lack a gene encoding a LarA homolog. This finding suggests that this group of microorganisms incorporates the NPN cofactor into a distinct protein architecture. To demonstrate the formation of NPN-derivatized protein in *E. coli*-expressed proteins, I applied an approach that had been used for identifying proteins containing a thioacid. Specifically, I used click chemistry to derivatize the protein with lissamine rhodamine B sulfonyl azide (LRSA) by reacting with the thioacid in the protein-bound NPN cofactor. Furthermore, to allow the monitoring of racemization reactions for a diverse range of 2-hydroxyacids, I demonstrated the utility of circular dichroism (CD) spectroscopy for such assays.

In Chapter 3, I provide evidence that dinicotinic acid adenine dinucleotide (DaAD) is formed as an intermediate during the reaction with NaAD of LarB from *L. plantarum* ($LarB_{Lp}$) and its S127A variant, which exhibits reduced hydrolytic activity.

In Chapter 4, I test the hypothesis that LarE from *Lactococcus sakei* ($LarE_{Ls}$) uses a third reaction mechanism for converting P2CMN to P2TMN. The sequence of $LarE_{Ls}$ does not possess three residues for ligating a [4Fe-4S] cluster as in $LarE_{Tm}$ and $LarE_{Mm}$, but the corresponding gene is encoded adjacent to a gene for a likely cysteine desulfurase IscS. This observation led to the notion that IscS either (i) donates a sulfur atom to a cysteine residue in $LarE_{Ls}$, forming a persulfide that could then be used for synthesis of the NPN cofactor while avoiding formation of a Dha residue, or (ii) rescues the Dha-containing form the $LarE_{Ls}$. My experimental efforts indicated no positive effect by IscS on $LarE_{Ls}$ catalysis. Rather, $LarE_{Ls}$ likely operates via the same mechanism as LarE from *L. plantarum*, resulting in a single turnover and a Dha residue.

In Chapter 5, I describe my attempts to identify NPN cofactor-containing proteins in cyanobacteria by using the LRSA labeling approach, and by using bioinformatics methods. I also summarize my attempts to trap the cofactor-containing proteins using derivatized resins.
In Chapter 6, I relate my unsuccessful attempt to detect a Ni-hydride intermediate in $LarA_{Lp}$ using ^{1}H-NMR spectroscopy.
Finally, in Chapter 7 I summarize the significant progress I have made in various facets of NPN cofactor biosynthesis and utilization. In addition, I enumerate several remaining topics that require further investigation.

REFERENCES

(1) Desguin, B.; Goffin, P.; Viaene, E.; Kleerebezem, M.; Martin-Diaconescu, V.; Maroney, M. J.; Declercq, J. P.; Soumillion, P.; Hols, P. Lactate Racemase Is a Nickel-Dependent Enzyme Activated by a Widespread Maturation System. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4615.

(2) Desguin, B.; Soumillion, P.; Hols, P.; Hausinger, R. P. Nickel-Pincer Cofactor Biosynthesis Involves LarB-Catalyzed Pyridinium Carboxylation and LarE-Dependent Sacrificial Sulfur Insertion. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (20), 5598–5603. https://doi.org/10.1073/pnas.1600486113.

(3) Rankin, J. A.; Chatterjee, S.; Tariq, Z.; Lagishetty, S.; Desguin, B.; Hu, J.; Hausinger, R. P. The LarB Carboxylase/Hydrolase Forms a Transient Cysteinyl-Pyridine Intermediate during Nickel-Pincer Nucleotide Cofactor Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (39), 1–7. https://doi.org/10.1073/pnas.2106202118.

(4) Fellner, M.; Desguin, B.; Hausinger, R. P.; Hu, J. Structural Insights into the Catalytic Mechanism of a Sacrificial Sulfur Insertase of the N-Type ATP Pyrophosphatase Family, LarE. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (34), 9074–9079. https://doi.org/10.1073/pnas.1704967114.

(5) Chatterjee, A.; Abeydeera, N. D.; Bale, S.; Pai, P. J.; Dorrestein, P. C.; Russell, D. H.; Ealick, S. E.; Begley, T. P. *Saccharomyces Cerevisiae* THI4p Is a Suicide Thiamine Thiazole Synthase. *Nature* **2011**, *478* (7370), 542–546. https://doi.org/10.1038/nature10503.

(6) Fellner, M.; Rankin, J. A.; Desguin, B.; Hu, J.; Hausinger, R. P. Analysis of the Active Site Cysteine Residue of the Sacrificial Sulfur Insertase LarE from *Lactobacillus plantarum*. *Biochemistry* **2018**, *57* (38), 5513–5523. https://doi.org/10.1021/acs.biochem.8b00601.

(7) Chatterjee, S.; Parson, K. F.; Ruotolo, B. T.; McCracken, J.; Hu, J.; Hausinger, R. P. Characterization of a [4Fe-4S]-Dependent LarE Sulfur Insertase That Facilitates Nickel-Pincer Nucleotide Cofactor Biosynthesis in *Thermotoga Maritima*. *J. Biol. Chem.* **2022**, *298* (7), 102131. https://doi.org/10.1016/j.jbc.2022.102131.

(8) Zecchin, P.; Pecqueur, L.; Oltmanns, J.; Velours, C.; Schünemann, V.; Fontecave, M.; Golinelli-Pimpaneau, B. Structure-Based Insights into the Mechanism of [4Fe-4S]-Dependent Sulfur Insertase LarE. *Protein Sci.* **2024**, *33* (2), 1–19. https://doi.org/10.1002/pro.4874.

(9) Desguin, B.; Fellner, M.; Riant, O.; Hu, J.; Hausinger, R. P.; Hols, P.; Soumillion, P. Biosynthesis of the Nickel-Pincer Nucleotide Cofactor of Lactate Racemase Requires a CTP-Dependent Cyclometallase. *J. Biol. Chem.* **2018**, *293* (32), 12303–12317. https://doi.org/10.1074/jbc.RA118.003741.

(10) Klein, A.; Sandleben, A.; Vogt, N. Synthesis, Structure and Reactivity of Cyclometalated Nickel(II) Complexes: A Review and Perspective. *Proc. Natl. Acad. Sci. India Sect. A - Phys. Sci.* **2016**, *86* (4), 533–549. https://doi.org/10.1007/s40010-016-0289-6.

(11) Albrecht, M. Cyclometalation Using D-Block Transition Metals: Fundamental Aspects and Recent Trends. *Chem. Rev.* **2010**, *110* (2), 576–623.

https://doi.org/10.1021/cr900279a.

(12) Dehand, J.; Pfeffer, M. Cyclometallated Compounds. *Coord. Chem. Rev.* **1976**, *18* (3), 327–352. https://doi.org/10.1016/S0010-8545(00)80431-2.

(13) Turmo, A.; Hu, J.; Hausinger, R. P. Characterization of the Nickel-Inserting Cyclometallase LarC from Moorella Thermoacetica and Identification of a Cytidinylylated Reaction Intermediate. *Metallomics* **2022**, *14* (3), 1–8. https://doi.org/10.1093/mtomcs/mfac014.

(14) Desguin, B.; Zhang, T.; Soumillion, P.; Hols, P.; Hu, J.; Hausinger, R. P. A Tethered Niacin-Derived Pincer Complex with a Nickel-Carbon Bond in Lactate Racemase. *Science (80-.).* **2015**, *349* (6243), 66–69. https://doi.org/10.1126/science.aab2272.

(15) Desguin, B.; Urdiain-Arraiza, J.; Da Costa, M.; Fellner, M.; Hu, J.; Hausinger, R. P.; Desmet, T.; Hols, P.; Soumillion, P. Uncovering a Superfamily of Nickel-Dependent Hydroxyacid Racemases and Epimerases. *Sci. Rep.* **2020**, *10* (1), 18123. https://doi.org/10.1038/s41598-020-74802-6.

(16) Xu, T.; Wodrich, M. D.; Scopelliti, R.; Corminboeuf, C.; Hu, X. Nickel Pincer Model of the Active Site of Lactate Racemase Involves Ligand Participation in Hydride Transfer. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (6), 1242–1245. https://doi.org/10.1073/pnas.1616038114.

(17) Shi, R.; Wodrich, M. D.; Pan, H. J.; Tirani, F. F.; Hu, X. Functional Models of the Nickel Pincer Nucleotide Cofactor of Lactate Racemase. *Angew. Chemie - Int. Ed.* **2019**, *58* (47), 16869–16872. https://doi.org/10.1002/anie.201910490.

(18) Qiu, B.; Yang, X. A Bio-Inspired Design and Computational Prediction of Scorpion-like SCS Nickel Pincer Complexes for Lactate Racemization. *Chem. Commun.* **2017**, *53* (83), 11410–11413. https://doi.org/10.1039/c7cc06416k.

(19) Zhang, X.; Chung, L. W. Alternative Mechanistic Strategy for Enzyme Catalysis in a Ni-Dependent Lactate Racemase (LarA): Intermediate Destabilization by the Cofactor. *Chem. - A Eur. J.* **2017**, *23* (15), 3623–3630. https://doi.org/10.1002/chem.201604893.

(20) Yu, M.-J.; Chen, S.-L. From NAD^+ to Nickel Pincer Complex: A Significant Cofactor Evolution Presented by Lactate Racemase. *Chem. - A Eur. J.* **2017**, *23* (31), 7545–7557. https://doi.org/10.1002/chem.201700405.

(21) Eberhardt, N. A.; Guan, H. Nickel Hydride Complexes. *Chem. Rev.* **2016**, *116* (15), 8373–8426. https://doi.org/10.1021/acs.chemrev.6b00259.

(22) Rankin, J. A.; Mauban, R. C.; Fellner, M.; Desguin, B.; McCracken, J.; Hu, J.; Varganov, S. A.; Hausinger, R. P. Lactate Racemase Nickel-Pincer Cofactor Operates by a Proton-Coupled Hydride Transfer Mechanism. *Biochemistry* **2018**, *57* (23), 3244–3251. https://doi.org/10.1021/acs.biochem.8b00100.

(23) Wang, B.; Shaik, S. The Nickel-Pincer Complex in Lactate Racemase Is an Electron Relay and Sink That Acts through Proton-Coupled Electron Transfer. *Angew. Chemie - Int. Ed.* **2017**, *56* (34). https://doi.org/10.1002/anie.201612065.

(24) Shapirof, S. S.; Dennis, D. Lactic Acid Racemization in *Clostridium Butylicum*. *Biochemistry* **1965**, *4* (11), 2283–2288.

CHAPTER 2

Overcoming barriers for investigating nickel-pincer nucleotide cofactor-related enzymes

Adapted with permission from:

Nevarez, J. L., Turmo, A., Gatreddi, S., Gupta, S., Hu, J., and Hausinger, R. P. (2024) Overcoming barriers for investigating nickel-pincer nucleotide cofactor-related enzymes, *mBio*. (in press as DOI: 10.1128/mbio.03404-24)

Author Contributions:

The work presented in this chapter highlights my contributions to the development and optimization of a circular dichroism (CD) spectroscopy assay for determining racemase kinetics. Additionally, I demonstrated that Lar protein homologs from *Synechocystis* sp. PCC 6803 can generate active LarA when coexpressed in *E. coli*. Furthermore, I conducted labeling experiments using lissamine rhodamine B sulfonyl azide to specifically target LarA in cell lysates.

Abstract

The nickel-pincer nucleotide (NPN) cofactor is a modified pyridinium mononucleotide that tri-coordinates nickel and is crucial for the activity of certain racemases and epimerases. LarB, LarC, and LarE are responsible for NPN synthesis, with the cofactor subsequently installed into LarA homologs. Hurdles for investigating the functional properties of such proteins arise from the difficulty of obtaining the active, NPN cofactor-loaded enzymes and in assaying their diverse reactivities. Here, we show that when the *Lactiplantibacillus plantarum lar* genes are cloned into the Duet expression system and cultured in *Escherichia coli*, they confer lactate racemase activity to the cells. By replacing *L. plantarum larA* with related genes from other microorganisms, this system allows for the generation of active LarA homologs. Furthermore, the Duet system enables the functional testing of LarB, LarC, and LarE homologs from other microorganisms. In addition to applying the Duet expression system for synthesis of active, NPN cofactor-containing enzymes in *E. coli*, we demonstrate that circular dichroism spectroscopy provides a broadly applicable means of assaying these enzymes. By selecting a wavelength of high molar ellipticity and low absorbance for a given 2-hydroxy acid substrate enantiomer, the conversion of one enantiomer/epimer into the other can be monitored for LarA homologs without the need for any coupling enzymes or reagents. The methods discussed here further our abilities to investigate the unique activities of Lar proteins.

Importance

Enzymes containing the NPN cofactor are prevalent in a wide range of microorganisms and catalyze various critical biochemical reactions, yet they remain underexplored due, in part, to limitations in current research methodologies. The two significant advancements described here, the heterologous production of active NPN-cofactor containing enzymes in *E. coli* and the use of a circular dichroism-based assay to monitor enzyme activities, expand our capacity to analyze these enzymes. Such additional detailed characterization will deepen our understanding of the diverse chemistry catalyzed by the NPN cofactor and potentially uncover novel roles for this organometallic species in microbial metabolism.

The nickel-pincer nucleotide (NPN) cofactor was first identified in lactate racemase from *Lactiplantibacillus plantarum* ($LarA_{Lp}$) and shown to be a pyridinium-3,5-dithiocarboxylic acid mononucleotide with nickel coordinated to the two sulfur atoms of the thiocarboxylates and C4 of the pyridinium ring.[1] The enzyme provides a histidyl residue to complete the square-planar

coordination geometry of the metal ion and a lysyl residue to covalently tether the cofactor by a thioamide linkage (**Figure 2-1A**). The cofactor does not always form an adduct with the enzyme as shown by LarA of *Thermoanaerobacterium thermosaccharolyticum* which loses the non-covalently bound NPN during enzyme purification.[2]

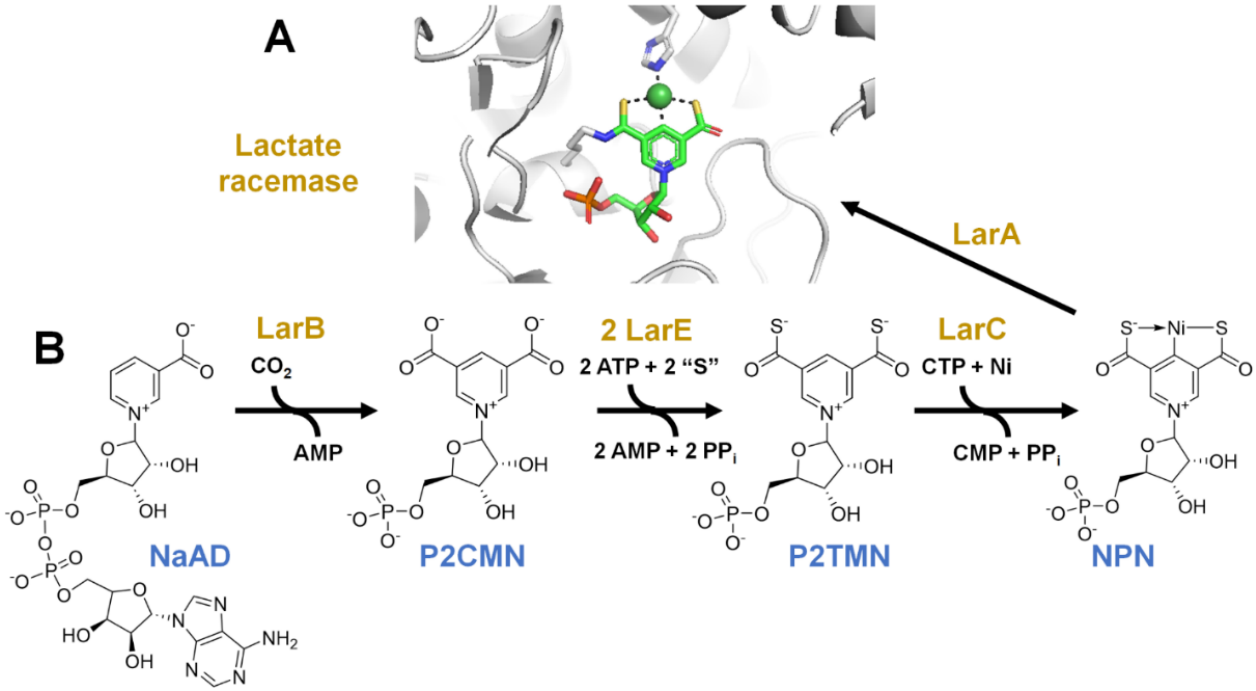

Figure 2-1. The structure and biosynthesis pathway for the NPN cofactor of lactate racemase. (A) Structure of the NPN cofactor in lactate racemase of *L. plantarum* (PDB: 5HUQ). (B) Pathway for biosynthesis of the NPN cofactor.

The pathway for biosynthesis of the NPN cofactor is established (**Figure 2-1B**). LarB adds CO_2 to C5 of nicotinic acid adenine dinucleotide (NaAD) and hydrolyzes the phosphoanhydride to release AMP and produce pyridinium-3,5-dicarboxylic acid mononucleotide (P2CMN).[3, 4] LarE from *L. plantarum* activates one of the P2CMN carboxylic acids by adenylylation while releasing pyrophosphate (PP_i), forms a covalent adduct with substrate using a cysteinyl residue while releasing AMP, and transfers the sulfur atom of this side chain to generate pyridinium-3-carboxy-5-thiocarboxylic acid mononucleotide (PCTMN) while forming a dehydroalanyl residue (Dha) in the protein.[3, 5] Another round of these reactions converts PCTMN to pyridinium-3,5-dithiocarboxylic acid mononucleotide (P2TMN) while hydrolyzing ATP to AMP and PP_i and forming Dha in a second LarE molecule.[6] Alternatively, LarE from *Thermotoga maritima* catalyzes similar transformations, but it possesses a [4Fe-4S] cluster bound by three cysteinyl residues that accepts a non-core sulfide at its open iron site from

cysteine desulfurase acting on the free amino acid L-cysteine, and this sulfur atom from the resulting [4Fe-5S] cluster is added to the adenylylated carboxyl groups of the substrates.[7] In the final biosynthetic step, *L. plantarum* LarC installs nickel ion into P2TMN to form the NPN cofactor in a CTP-dependent process.[8] Studies of LarC from *Moorella thermoacetica* showed that CTP is used to cytidinylylate the phosphate group of P2TMN, with the enzyme hydrolyzing the phosphoanhydride after nickel is incorporated.[9]

Bioinformatic analysis of over 1,000 eubacterial and archaeal genomes indicated that homologs of LarA and the NPN biosynthesis pathway enzymes are present in approximately 9% of this population [2]. Recent biochemical studies demonstrated that seven out of 13 potential *larA* homologs carried out a reaction distinct from lactate racemization; namely, racemization or epimerization of other 2-hydroxyacid substrates such as malate, 2-hydroxyglutarate, and the sugar D-gluconate [10]. Moreover, ~15% of the genomes analyzed contain homologs of *larB*, *larE*, and *larC*, but lack a *larA* homolog, suggesting that some microorganisms, including many cyanobacteria, synthesize the NPN cofactor using the usual pathway, but then incorporate the molecule into one or more non-LarA NPN-binding proteins.

Given the widespread appearance and diverse functionality of the LarA homologs and the likely presence of non-LarA NPN cofactor-binding proteins, it is important to both (i) develop methods that allow for the generation of cofactor-containing (active holoprotein) forms of these enzymes from diverse microorganisms and (ii) to devise routine assays for characterizing the properties of the different 2-hydroxyacid racemases/epimerases. Here, we describe methods to fulfill these needs. Notably, these approaches also can be used to confirm the reactivities and characterize the attributes of LarB-, LarE-, and LarC-like proteins. By adopting these methods to the study of Lar proteins, investigators will be able to better analyze the NPN utilizing and synthesizing enzymes and obtain a more complete understanding of the full catalytic potential of the NPN cofactor.

Materials and methods

Gene, plasmids, and cloning

Bacterial strains, plasmids, and primers used for this study are listed in **Table 2-A1**. PCR amplifications were performed using Q5 high-fidelity DNA polymerase following the manufacture's protocol (NEB, Ipswich, MA, USA) or using Vazyme Phanta Flash Super-Fidelity DNA polymerase (Nanjing, China). The primers used were purchased from IDT (Integrated

DNA Technologies, Coralville, IA, USA). DNA fragments were ligated using the *in vivo* subcloning assembly method[11, 12] or using Vazyme ClonExpress II One Step Cloning Kit. The transformations were performed using a standard chemical method in *E. coli* DH5α and BL21 (DE3) cells for plasmid amplification and protein expression purposes, respectively.[13]

The four *lar*-related genes from *L. plantarum* were incorporated into pETDuet (*larA* and *larB*) or pRSFDuet (*larE* and *larC*) plasmids (Merck KGaA, Darmstadt, Germany).[14] The homologs of *larA* from *Megasphaera elsdenii*[10] was exchanged for the version from *L. plantarum*. Similarly, genes encoding LarC of *Moorella thermoacetica* and homologs of the NPN cofactor biosynthesis enzymes from *Synechocystis* sp. PCC 6803 were switched with *larB*, *larC*, or *larE* of *L. plantarum*. All homologous genes were codon optimized for expression in *E. coli*, chemically synthesized (IDT), and sequentially swapped for the *L. plantarum* genes.

Purification of Isosphaera pallida LarA from Lactococcus lactis

The Strep-tagged LarA homolog from *I. pallida* ($LarA_{Ip}$, also named LarAH31) was expressed from pGIR210-LarAH31 [constructed from the pBAD-derived vector encoding LarAH2/Sar(10)] in *L. lactis* NZ3900 that was grown with shaking at 30 °C in M17 medium supplemented with 0.5% glucose and 7.5 μg/mL chloramphenicol. After reaching an OD_{600} of 0.3-0.4, 1 mM $NiCl_2$ and 5 μg/L of nisin A were added and the culture was grown for an additional 3-4 h. After overnight storage at 4 °C, the cells were collected by centrifugation (4000 x *g* for 15 min), resuspended in 100 mM Tris buffer, pH 7.5, containing 150 mM NaCl, 10 μg/mL lysozyme, and 2 μg/mL of DNaseI, and stirred at 4 °C for 1 h. To the suspension was added 1 mM phenylmethylsulfonyl fluoride (PMSF) and cell lysis was carried out by two passes through the French press at 16,000 psi. The supernatant was collected by centrifugation (18,000 x *g* for 70 min). $LarA_{Ip}$ was purified using StrepTactin XT resin (IBA, Göttingen, Germany) that was equilibrated with 100 mM Tris buffer, pH 7.5, containing 150 mM NaCl. After loading the sample, the column was washed with the same buffer, and then enzyme was eluted with this buffer containing 50 mM biotin and 50 mM NaOH. To improve the sample purity, the concentrated protein was chromatographed on a Superdex 200 increase 10/30 GL column equilibrated with 50 mM Tris buffer, pH 7.4, containing 300 mM NaCl. The monomeric peak fractions from the gel filtration column were used for further studies. Protein concentrations were determined by using a Nanodrop spectrophotometer and calculated with an extinction coefficient of 35,410 M^{-1} cm^{-1} matching the protein sequence.

Purification of active Strep-tagged LarA homologs from E. coli

In addition to using the above established method for purifying LarA homologs from recombinant *L. lactis* cells with nisin induction,[15] we developed a new approach to generate active LarA homologs in *E. coli* BL21 (DE3) using the Duet plasmid system.[14] The *E. coli* cells that had been co-transformed with pAT035 and pAT038, or their derivatives, were grown in autoinduction medium[16] or TB medium containing 50 μg/mL kanamycin and 100 μg/mL carbenicillin at room temperature while shaking at 220 RPM to an OD_{600} of 0.6-0.8. The cultures in TB medium were induced with 1 mM isopropyl-D-1-thiogalactopyranoside (IPTG). All cultures were incubated at room temperature for 20-24 h. When indicated, 1 mM of $NiCl_2$ and/or 1 mM of nicotinic acid (final concentrations) was added after 4 h of growth. Cell pellets were resuspended in 100 mM Tris, pH 7.5, buffer containing 150 mM NaCl and stored at -80 °C until use. Thawed cells were adjusted to contain 0.5 mM Na_2SO_3 (for cofactor stabilization), 1 mM PMSF, one tablet of cOmplete™ EDTA-free protease inhibitor cocktail (Roche, Basel, Switzerland), 1 mM lysozyme, 1 mM dithiothreitol (DTT), and 1 unit of benzonase. The cells were lysed by two passes through a French pressure cell at 16,000 psi. Strep-tagged LarA or its homologs were purified using a StrepTactin XT resin (IBA, Göttingen, Germany) with buffer that included 100 mM Tris at pH 7.5, 300 mM NaCl, and 0.05 mM Na_2SO_3, and the proteins were eluted with 50 mM biotin.[2] Protein concentrations were determined by the Bradford protein assay reagent (Bio-Rad, Hercules, CA, USA) using bovine serum albumin as the standard.

Coupled enzyme activity assays

The purified Strep-tagged $LarA_{Lp}$ protein was buffer exchanged to remove the Na_2SO_3 and biotin from the buffer using a PD-10 desalting column (Cytiva, Marlborough, MA, USA). To assess the Lar activity by the conventional assay, $LarA_{Lp}$ or its homologs (1 pmol) was mixed with 5-400 mM sodium L-lactate in 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7.0, for up to 12 min at 35 °C, then boiled for 10 min at 95 °C to inactivate the enzyme. The precipitated protein was removed by centrifugation at 17,000 x *g* and the supernatant was collected. The amount of D-lactate in the sample, produced by the lactate racemase activity of LarA, was measured using a commercial kit (Neogen, Lansing, MI, USA) as previously described.[2] The same kit was utilized when studying phenyllactate racemase, examining both L-lactate and L-phenyllactate as substrates.

Circular dichroism spectroscopy-based assay

An alternative assay that can be used with a wide assortment of NPN-containing racemases and epimerases made use of changes in molar ellipticity as monitored by circular dichroism (CD) spectroscopy. This assay was inspired by similar studies involving mandelate racemase,[17] which also utilizes a 2-hydroxyacid substrate. The L- or D-2-hydroxyacid sample was scanned while recording both the molar ellipticity and absorbance versus wavelength. By dividing the former by the latter, the wavelength providing the largest signal for molar ellipticity with low noise was identified. Changes in ellipticity at the selected wavelength allowed for real-time monitoring of racemization. The CD spectra were recorded using a Jasco J-815 CD spectrometer equipped with a Jasco CDF-426s/15 temperature control unit. For substrate CD spectra scans, the typical settings were as follows: wavelength range of 200-300 nm, data pitch of 0.5 nm, digital integration time (DIT) of 2 s, scanning speed of 100 nm/min, bandwidth of 1 nm, and 3 accumulations. For continuous wavelength monitoring, the typical settings included a data pitch of 1.0 s, a DIT of 4 s, and a bandwidth of 1 nm. All measurements were conducted using a 10 mm pathlength quartz cuvette.

The initial rates of LarA$_{Lp}$ activity were determined by monitoring the change in molar ellipticity at 232 nm. A standard curve was generated using known concentrations of D- and L-lactate, with the slope of the curve defining the relationship between molar ellipticity at 232 nm and lactate concentration. The rate of conversion of one isomer of lactate to the other was described by equation 1.

$$v = \left| \frac{\Delta\theta/\Delta t}{2l[\theta]} \right| \quad (1)$$

where v is the velocity of the reaction, $\Delta\theta/\Delta t$ represents the rate of change in observed ellipticity, l is the sample pathlength, [θ] is the ellipticity of lactate at the monitored wavelength, and the value 2 accounts for the equal and opposite contribution to the total ellipticity by the product.

Initial velocities were determined by substituting the measured values of $\Delta\theta/\Delta t$ and [θ] into equation 1, with $\Delta\theta/\Delta t$ being the slope of the linear portion of the reaction plot (from 0 to 12 s of the reaction) for each substrate concentration, and [θ] being 417.15 deg. mol^{-1} cm^2, the molar ellipticity at 232 nm for lactate. These initial velocities were then plotted against substrate concentrations to generate Michaelis-Menten plots. The data were fitted to the Michaelis-Menten equation using the "curve fit" function from SciPy in Python to obtain K_m and V_{max} values. To calculate the catalytic constant (k_{cat}), V_{max} was divided by the enzyme concentration (2.1 μM)

and converted to s^{-1}. Analogous approaches were utilized for monitoring catalysis involving other 2-hydroxyacids.

Lissamine rhodamine B sulfonyl azide (LRSA) labeling for detection of the protein bound NPN

To assess whether the LarA homologs had covalently incorporated the NPN cofactor, we investigated the ability of LRSA, a reagent that reacts with thiocarboxylic acids, to react with the protein bound NPN after dissociation of the nickel, i.e., protein-bound P2TMN.[18] Synthesis of LRSA was carried out as previously described.[19] The protocol for LRSA labeling of P2TMN-bound proteins was based on and modified from the procedure for labeling proteins that terminate in a thiocarboxylic acid at their carboxyl end.[19, 20] For this analysis, 0.5 g of cells were resuspended in 0.7 mL of 100 mM Tris-buffered saline containing 150 mM NaCl at pH 7.5 and transferred to a 2 mL tube to be lysed with a mini-beadbeater. The lysates were centrifuged, and the supernatant solutions were collected. The samples were roughly normalized based on the overall protein content using the absorbance at 280 nm, and buffer exchanged into 50 mM potassium phosphate, 300 mM NaCl, and 6 M urea, at pH 6.1. To each of the samples was added 10 µL of 15 mM LRSA in dimethyl sulfoxide and the vials were left to react in the dark at room temperature for 20 min. The protein portions of the samples were precipitated using the chloroform-methanol method[21] and resuspended in the phosphate urea buffer stated above. Each sample (20 µlL) was mixed with 5 µL of 5-fold concentrated sodium dodecyl sulfate (SDS)-loading buffer and 20 µL was loaded onto a 12% acrylamide gel, subjected to SDS-polyacrylamide gel electrophoresis (PAGE), and used for imaging the proteins with bound LRSA followed by staining with Coomassie brilliant blue. Analogous steps were used when examining purified proteins. The rhodamine-bound gel bands were excited at 530 nm while monitoring the emission at 580 nm and documented using the ChemiDoc MP imaging system (BioRad, Hercules, CA, USA).

Ultraviolet (UV)-visible spectroscopy

The spectra of the purified LarA proteins were measured using a quartz cuvette with a Shimadzu UV-2600 spectrophotometer (Kyoto, Japan) at room temperature. Sample volumes were 1 mL.

Nickel content analysis

Quantification of the LarA nickel content was carried out by using an Agilent 710 Series (Santa Clara, CA, USA) inductively coupled plasma optical emission spectrometer (ICP-OES). The samples were prepared by adjusting to 35% w/v HNO_3 and heating at 95 °C for one h to

mineralize the components. A final concentration of 0.1 ppm Yttrium (Sigma-Aldrich, St. Louis, MO, USA) was added to all samples as an internal standard. A nickel standard curve and a buffer control were used to account for background nickel contamination. Data were collected and analyzed using the ICP Expert II software.

Mass spectrometric analysis

Intact protein masses were analyzed by using a Waters G2-XS Q-TOF (time of flight) mass spectrometer by injecting 10 µL of sample onto a Thermo Hypersil Gold CN guard column (1.0 × 10 mm) for desalting. A gradient was run as follows using 0.1% formic acid in water (solvent A) and acetonitrile (solvent B) at a flow rate of 0.1 mL/min: initial conditions were 98% A/2% B, hold at 2% B to 5 min with the flow diverted to waste for the first 3 min, ramp to 75% B at 10 min and hold at 75% B to 12 min, return to 2% B at 12.01 min and hold to 15 min. Mass spectra were obtained using electrospray ionization in positive ion mode with a source temperature of 100 °C, cone voltage of 35 V, desolvation temperature of 350 °C, desolvation gas flow of 600 L/h, cone gas flow of 50 L/h, and capillary voltage of 3.0 kV. Data were acquired using a 0.5 s TOF MS scan across an *m/z* range of 200–2000. The spectra were deconvoluted in Masslynx using the maximum entropy (MaxEnt) I algorithm.

Results

Synthesis of active $LarA_{Lp}$ in E. coli

In prior studies, the four *L. plantarum lar* genes were placed under the control of a nisin-inducible promoter and transformed into *L. lactis* lacking the *lar* genes.[2] Homologs of *larA* from other microorganisms could then be substituted for the corresponding *L. plantarum* gene within this Gram-positive host resulting in a few active enzymes, but this method was not generally successful due to plasmid instability and other confounding issues.[10] In parallel with the *L. lactis* expression studies, earlier efforts expressed individual *larA* homologs in *E. coli* and the resulting apoproteins were mixed with biosynthetically produced NPN cofactor, a time-consuming and error-prone process that required the isolation and utilization of LarB, LarE, and LarC (**Figure 2-A1**).[10]

Here, we co-expressed the four *lar* genes of *L. plantarum* within *E. coli* using the pETDuet and pRSFDuet plasmids of the Duet vector system[14] and successfully generated lactate racemase activity in this genetically tractable and easily manipulated host microorganism. Importantly, these plasmids are stably maintained in the cells due to their compatible replication origins and different antibiotic resistance cassettes. The genes encoding $LarE_{Lp}$ and $LarC_{Lp}$ were selected for expression using the pRSFDuet plasmid (creating pAT038), a high copy number vector, since these proteins are thought to catalyze single turnover reactions, whereas *larA* and *larB* from *L. plantarum* were expressed from the pETDuet plasmid (forming pAT035) (**Figure 2-2**). Importantly, the gene encoding $LarA_{Lp}$ was cloned with a sequence for a Strep-tag on its C-terminus for easy purification prior to analyzing activity and testing for the presence of covalently bound NPN cofactor. Using this new system, we showed that *larA* homologs from other microorganisms can be swapped for the gene encoding $LarA_{Lp}$ to obtain NPN-containing forms of those alternative proteins. In addition, we demonstrated that genes encoding homologs to the NPN biosynthetic enzymes can be exchanged for the corresponding genes in this system to test their biosynthetic functionality. This new approach avoids the problems of the *L. lactis* cloning and expression system and the need to purify from *E. coli* each of the individual

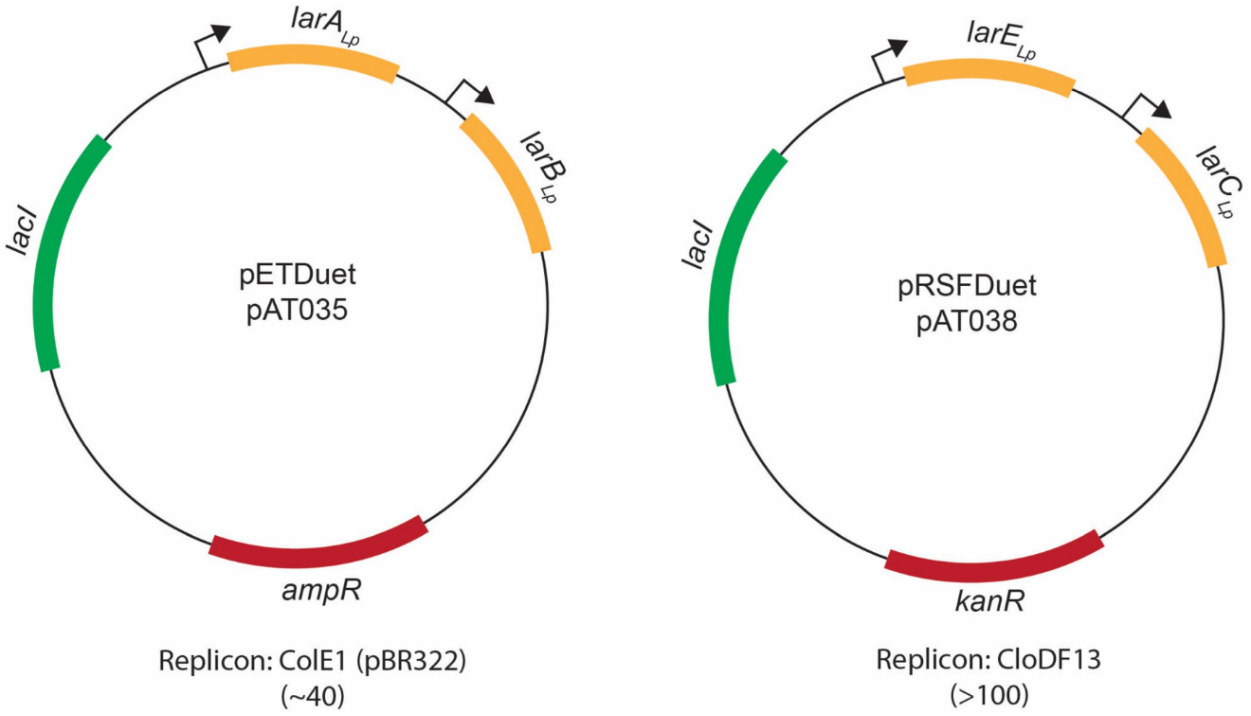

Figure 2-2. Plasmid design for expression of *lar* genes in *E. coli*. The number in parentheses indicates the plasmid copy number.

enzymes required for NPN biosynthesis. As described below, we demonstrated the production of active $LarA_{Lp}$ by (i) revealing the presence of covalently bound P2TMN using a fluorescent staining procedure that reacts with the novel thiocarboxylic acid, (ii) utilizing UV-visible spectroscopy to identify the chromophore associated with the cofactor, (iii) establishing the presence of nickel in the protein, (iv) validating the expected increase in protein size by mass spectrometry, and (v) directly assessing the enzyme activity.

Cell-free lysates derived from the *E. coli* Duet system expressing the *L. plantarum* genes were tested for the presence of NPN-bound $LarA_{Lp}$ by reacting the samples with LRSA, a reagent that specifically reacts with thiocarboxylic acids (**Figure 2-A2**). After resolving the proteins by SDS-PAGE, the protein bands were visualized by Coomassie blue staining (**Figure 2-3, left**), and the LRSA-reactive bands were identified by fluorescence imaging (**Figure 2-3, right**). We compared the intensity of labeling for samples of *E. coli* (pAT035/pAT038) cells that were grown without additive, with 1 mM nicotinic acid, with 1 mM $NiCl_2$, and with both 1 mM nicotinic acid and 1 mM $NiCl_2$. In addition, we examined a sample of *E. coli* (pAT035) as a negative control and investigated cell-free extracts of *L. lactis* (pGIR112, expressing the four *lar*

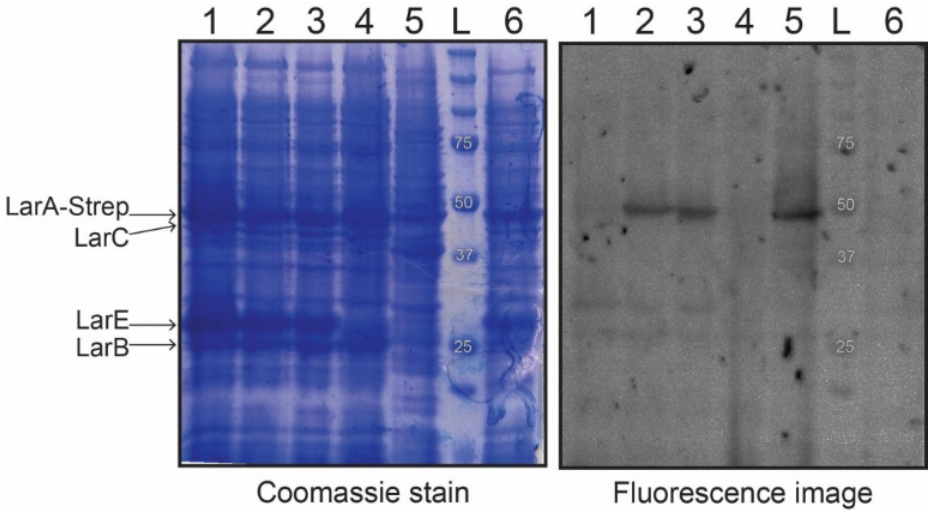

Figure 2-3. LRSA labeling of NPN-bound protein in cell-free *E. coli* lysates. LRSA-labeled protein samples were subjected to denaturing gel electrophoresis, imaged for fluorescence (right, with excitation and emission wavelengths of 530 nm and 580 nm), and stained with Coomassie brilliant blue (left). Lane 1: lysate of culture expressing pAT035 and pAT038 that was supplemented with 1 mM nicotinic acid; lane 2: lysate of this culture with 1 mM $NiCl_2$; lane 3: lysate of this culture with 1 mM nicotinic acid and $NiCl_2$; lane 4: lysate of culture expressing pAT035 alone with 1 mM nicotinic acid and $NiCl_2$; lane 5: *L. lactis* (pGIR112) lysate of cells supplemented with 1 mM $NiCl_2$ during growth; lane L: protein ladder; lane 6: culture with no additive.

genes)[2] as a positive control. A band corresponding in size to Strep-tagged LarA$_{Lp}$ (47.5 kDa) was fluorescently labeled when using samples derived from *E. coli* (pAT035/pAT038) or *L. lactis* (pGIR112) grown in the presence of nickel ions. In the absence of added nickel ions, the LarA band was not labeled, presumably due to a requirement for complete NPN cofactor biosynthesis prior to covalent attachment to LarA in the *E. coli* cells. No band was labeled for *E. coli* (pAT035) that produced LarA and LarB but was incapable of completing NPN cofactor biosynthesis because it lacked LarE and LarC. Analysis of the SDS-PAGE gel by Coomassie staining revealed the presence of the over-expressed LarC$_{Lp}$ (46.5 kDa), LarE$_{Lp}$ (30.5 kDa), and LarB$_{Lp}$ (25.3 kDa) proteins in the appropriate samples. Of additional interest, the relative intensities of several bands were altered in the *L. lactis* lysate which expresses a different set of genes. Although these results indicate that *E. coli* cells generated NPN-bound LarA$_{Lp}$, it does not allow for precise quantification of the labeling.

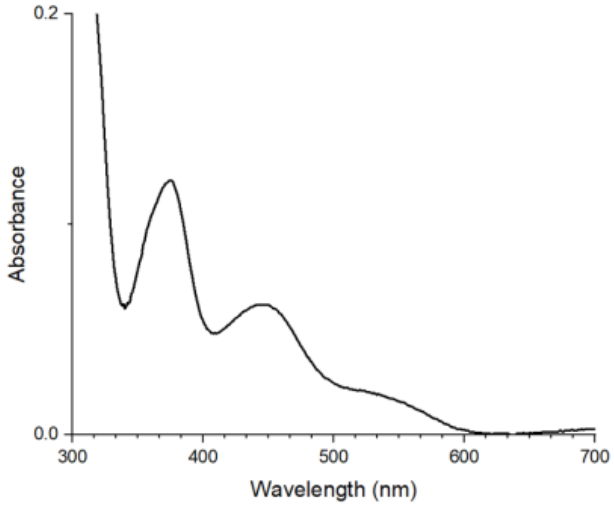

Figure 2-4. UV-visible difference spectrum obtained by subtracting the spectrum of LarA apoprotein from that of the LarA$_{Lp}$ holoprotein, derived from *E. coli* expression system. The absorbance at 280 nm was set to 1.0 for normalization.

As a second method to characterize LarA$_{Lp}$ that was produced in *E. coli*, we tested for the presence of the NPN cofactor chromophore in the purified sample. The UV-vis spectrum of the enzyme (**Figure 2-4**) revealed electronic transitions at ~380 nm and 440 nm along with a shoulder at 550 nm; these features are not present in the apoenzyme and are in agreement with prior findings for LarA$_{Lp}$ purified from *L. lactis*.[1] When standardized to the absorption values at 280 nm, the chromophore absorption intensities were approximately half of those reported previously for LarA$_{Lp}$.[1] This result demonstrates significant, but possibly incomplete, cofactor

incorporation and likely also reflects the partial loss of the nickel from the labile chromophore as described earlier.[2]

As another method to quantify the NPN cofactor content of the purified protein, we determined its metal content. ICP-OES results obtained within 24 h of purification demonstrated

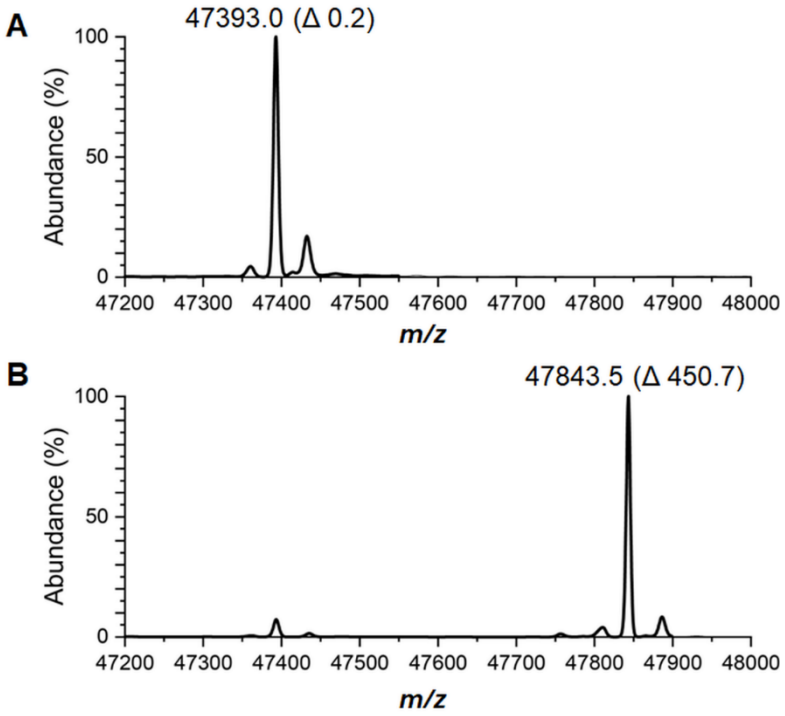

Figure 2-5. Mass spectra of LarA$_{Lp}$ purified from the *E. coli* Duet expression system. The mass spectra of (A) LarA$_{Lp}$ apoprotein and (B) LarA$_{Lp}$ holoprotein. The percent abundance is relative to the largest peak. In parenthesis are the mass differences relative to the theoretical mass of the LarA$_{Lp}$ apoprotein lacking the N-terminal methionine.

the presence of ~22 ± 2 % Ni atoms per LarA$_{Lp}$ subunit (SD, n = 3). The less than stoichiometric amount of metal associated with the protein is likely due to the instability of nickel within the NPN cofactor, as previously documented for sample that initially exhibited 69% Ni/subunit.[2]

A fourth technique to assess the NPN cofactor content of LarA$_{Lp}$ made use of protein mass spectrometry. **Figure 2-5** shows the results for apoprotein and freshly isolated holoprotein where each sample was shown to be nearly homogeneous and the *m*/*z* values are consistent with the masses expected for the C-terminally Strep-tagged proteins missing their N-terminal methionine residues, with an additional mass difference of 450.7 Da in the holoprotein. This result confirms that a substantial amount of NPN cofactor was covalently bound, presumably to

lysine 184, in the *E. coli*-derived *L. plantarum* LarA sample, as previously reported for *L. plantarum* LarA isolated from *L. lactis*.[1]

Finally, Lar activity was detected qualitatively in all lysate samples except for the negative control by using the conventional assay method. A newly developed assay for monitoring the activity of *E. coli*-produced LarA$_{Lp}$ and other racemase samples, described below in a separate section, confirmed the presence of active enzyme. A possible reason to explain the presence of activity in samples that were not labeled by LRSA labeling (from cells provided with only 1 mM nicotinic acid or no additive (lanes 1 and 6) could be due to insufficient levels of covalently bound NPN cofactor on LarA$_{Lp}$ to due to insufficient amounts of bound nickel.

Use of the Duet system for characterizing NPN-containing LarA homologs

We have successfully substituted genes for several LarA homologs into the pETDuet vector, but here we limit our discussion of this topic to a single illustrative example. *M. elsdenii* is an ecologically crucial rumen bacterium that metabolizes lactate and alleviates rumen acidosis that is generated by a high-grain diet.[22] Prior studies had reported the *E. coli* expression of a *larA* homolog from this microorganism followed by purification of the apoprotein (previously named LarAH10, but here termed LarA$_{Me}$).[10] Incubation of the apoprotein with biosynthesized NPN provided an enzyme that lacked Lar activity, but it exhibited very low levels of phenyllactate racemase activity (k_{cat} of 0.16 ± 0.03 s^{-1} and K_m of 0.4 ± 0.3 mM) as measured by capillary electrophoresis.[10] Those results suggested that phenyllactate was not the true substrate of the enzyme, that the protein was only partially loaded with cofactor, or that LarA$_{Me}$ was rapidly inactivated.

We subcloned the *M. elsdenii larA* from the previously described pBAD vector,[10] substituted it into the pETDuet expression vector, coexpressed it in *E. coli* with genes encoding the NPN-biosynthetic enzymes, and characterized the resulting LarA$_{Me}$ after isolation using a StrepTactin column. We observed LRSA reactivity of LarA$_{Me}$ both in cell-free extracts and with the purified protein obtained using the *E. coli* Duet system (**Figure 2-A3**). LarA$_{Me}$ also exhibited the expected weak chromophore spectrum corresponding to the NPN cofactor (**Figure 2-A4**). In addition, we showed that ~17% of the purified LarA$_{Me}$ derived from the *E. coli* system contained nickel, whereas the protein isolated from the *L. lactis* system did not appear to possess this metal ion.[10] Of great interest, analysis of the purified LarA$_{Me}$ by MS revealed several features, some of which are especially instructive for investigators who plan to utilize this method (**Figure 2-A5**).

A peak of m/z = 47,936 matches the expected size of the Strep-tagged apoprotein missing its amino terminal methionine residue. A peak corresponding to the intact holoprotein (expected m/z = 48,387) was not observed, but a peak of m/z = 48,344 was consistent with loss of nickel (-58.7 Da) and oxidation (+16 Da) of a sulfur atom; such a 408 Da mass difference was reported previously for LarA$_{Lp}$.[1] Of greatest interest are the two LarA$_{Me}$ features with m/z = 48,085 and 48,117. These values are 149 Da and 181 Da greater than the apoprotein and suggest that dicarboxylated nicotine and dithiocarboxylated nicotine are bound to the protein. Thus, in addition to the potential for losing nickel and undergoing oxidation of sulfur atoms, the NPN cofactor in LarA$_{Me}$ can be unstable due to cleavage of the glycosidic bond that joins the ribose phosphate to the pyridinium ring. Finally, we tested the activity of the *E. coli*-derived LarA$_{Me}$ using both L-lactate and L-phenyllactate. Whereas earlier studies had failed to detect Lar activity for LarA$_{Me}$, we observed an initial rate for conversion of L-lactate to D-lactate of ~2.5 s^{-1}, which is three orders of magnitude less that the value of 4,745 s^{-1} reported for LarA$_{Lp}$.[2] Our analysis of L-phenyllactate racemase activity relied on the ability of D-lactate dehydrogenase to use D-phenyllactate as an alternative (poor) substrate. We measured an initial rate of ~0.25 s^{-1}, which is 10% of the value observed for its lactate racemase activity. This result agrees with the previous report of a k_{cat} of 0.16 s^{-1}.[10] The prior study had quenched the reactions after 5 min, and we found the reactions catalyzed by LarA$_{Me}$ ceased after about 10 min, consistent with enzyme instability. Overall, our analyses of the *E. coli* purified LarA$_{Me}$ highlight the lability of the NPN cofactor in this protein and leave open the question of the authentic substrate of this enzyme.

Investigating the functionality of homologs of potential NPN biosynthetic enzymes

In addition to allowing the generation of functional NPN-containing LarA homologs from diverse microorganisms, the Duet plasmid system can be used to test the functionality of genes encoding the homologs of LarB, LarE, and LarC from varied sources. For example, *larC* from *M. thermoacetica* (encoding LarC$_{Mt}$), already shown to encode an enzyme with nickel insertase activity,[9] was swapped for *larC*$_{Lp}$ in pAT038 of the *E. coli* Duet system to create pAT039 and the resulting cells were shown to incorporate the NPN cofactor into LarA$_{Lp}$ (shown by the Δ450.2 peak) along with a form of the cofactor that was missing the Ni and potentially with a hydroxylated sulfur atom (based on the 409.7 Da increase in protein mass) (**Figure 2-A6A**). Furthermore, the substitution of LarC$_{Mt}$ for LarC$_{Lp}$ in the Duet system allowed the generation of lactate racemase activity. Similarly, swapping the *larC* homolog from *Synechocystis* sp. PCC

6803 (encoding LarC$_{Sc}$, UniProt number P72725) for *larC*$_{Lp}$ in the pAT038 plasmid yielded cells that derivatized most of its Lar$_{Lp}$ with bound NPN or a damaged cofactor (increasing the protein mass by 450.2 or 410.7 Da, respectively, **Figure 2-A6B**) and exhibited lactate racemase activity, confirming that the cyanobacterial LarC-like protein was capable of inserting nickel into P2TMN and creating a mature NPN cofactor. These findings demonstrate the versatility of the *E. coli* expression system that can be used to test the function of potential biosynthesis pathway enzymes identified only as being sequence homologs.

Showcasing an extreme example of how the Duet system can provide insights into the roles of genes that encode potential NPN cofactor biosynthetic enzymes, we exchanged the *Synechocystis* sp. PCC 6803 homologs of *larB* (encoding LarB$_{Sc}$), *larE* (encoding LarE$_{Sc}$), and *larC* for the corresponding *L. plantarum* genes in pAT035 and pAT038, while retaining the gene encoding LarA$_{Lp}$. The resulting cells were shown qualitatively to possess Lar activity, thus demonstrating that the cyanobacterial genes encode proteins that synthesize the NPN cofactor. Genes encoding LarB, LarE, and LarC are widely distributed in other strains of cyanobacteria, including *Synechocystis* sp. PCC 6803, where they also likely catalyze NPN cofactor biosynthesis. Notably, however, most cyanobacteria lack genes encoding LarA-like proteins. Thus, it is likely that these phototrophs synthesize the NPN cofactor for use in a yet-to-be-identified protein that possesses a sequence and fold distinct from LarA.

Development of a broadly applicable new assay for NPN-containing LarA homologs

Until recently, assays for racemases specific to different 2-hydroxyacids required the use of coupling enzymes that reacted with one enantiomer of the distinct substrates.[10] A newly developed approach to quantitatively monitor the two enantiomers of a range of 2-hydroxyacids utilizes a capillary electrophoresis approach that included vancomycin in the buffer as a chiral selector.[23] Here, we adapted a method using circular dichroism spectroscopy as a broadly applicable technique to monitor 2-hydroxyacid racemases and epimerases. This approach was inspired by similar studies carried out with mandelate racemase that acts on an aromatic 2-hydroxyacid.[17]

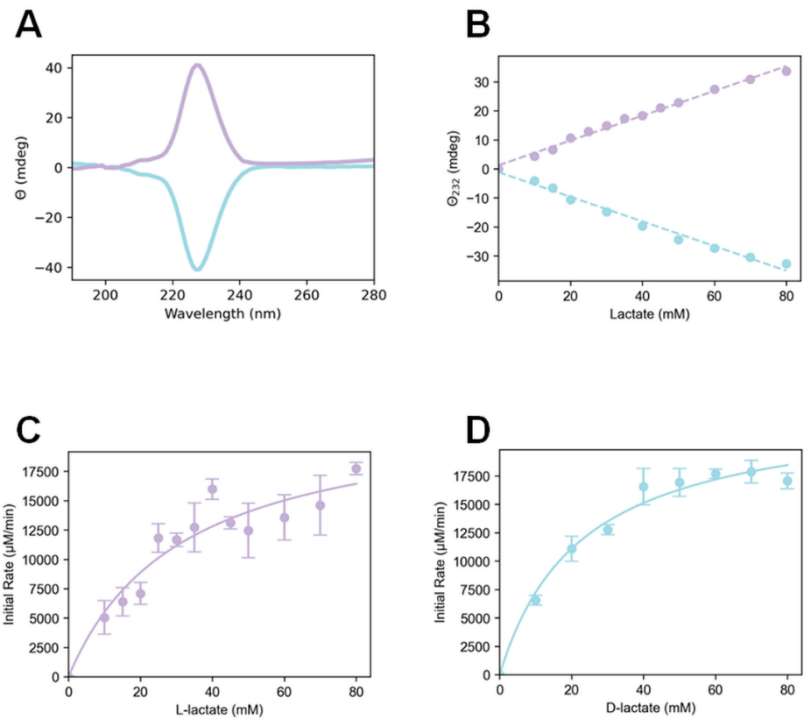

Figure 2-6. Kinetic analysis of *E. coli*-derived LarA$_{Lp}$ using a CD assay. (A) CD spectra of L- and D-lactate (purple and cyan lines, respectively). (B) Standard curves showing the changes in molar ellipticity as functions of L- or D-lactate concentrations. Changes in molar ellipticity versus time provided initial rates at varied concentrations of (C) L-lactate and (D) D-lactate. The data were obtained at 35 °C in phosphate-buffered saline (pH 7.4) using 2.13 μM LarA$_{Lp}$.

L- and D-lactate exhibit distinct CD spectra (**Figure 2-6A**), with the maximal difference in ellipticities occurring at a wavelength of 228 nm. To reduce background issues due to absorbance by the substrate, assays were carried out at 232 nm. Standard curves that monitored changes in ellipticity for varying concentrations of lactate yielded straight lines (**Figure 2-6B**). Using the *E. coli*-derived LarA$_{Lp}$, the initial rates of racemization were examined as a function of varied concentrations for each lactate enantiomer. Studies with L-lactate provided k_{cat} of 179 ± 25 s^{-1} and K_m of 32 ± 10 mM, whereas those using D-lactate yielded k_{cat} of 185 ± 12 s^{-1} and K_m of 23 ± 12 mM (**Figure 2-6C, 6D**). These rate constants are less than those obtained for LarA$_{Lp}$ purified from recombinant *L. lactis* cells (k_{cat} of 4,745 ± s^{-1} for L-lactate and 1,333 ± 131 s^{-1} for D-lactate),[2] perhaps due to incomplete incorporation of the NPN cofactor or to rapid inactivation of the *E. coli*-derived enzyme. In contrast, the above CD-derived Michaelis constants were

comparable to those in the prior report (K_m of 46 ± 20 mM for L-lactate and K_m of 11 ± 4 mM for D-lactate).

To demonstrate its versatility and to emphasize its limitations, this assay was applied to LarA$_{lp}$ from the thermophile *I. pallida*.[10] This enzyme, when purified from *L. lactis*, previously was shown to exhibit lactate racemase activity (k_{cat} of 4.7 ± 0.3 s^{-1} and K_m of 0.15 ± 0.04 mM for L-lactate, k_{cat} of 8.1 ± 0.5 s^{-1} and K_m of 0.56 ± 0.10 mM for D-lactate). In addition, the earlier study showed this enzyme was inhibited by D-2-hydroxyisovalerate, L-2-hydroxyisovalerate, D/L-2-hydroxyisocaproate, and L-2-hydroxyisocaproate (with K_i values of 0.15 ± 0.02 mM, 1.2 ± 0.2 mM, 2.1 ± 0.3 mM, and 14 ± 4 mM, respectively).[10] Using the CD assay, we confirmed that lactate is a substrate and newly demonstrated that the enzyme also racemizes 2-

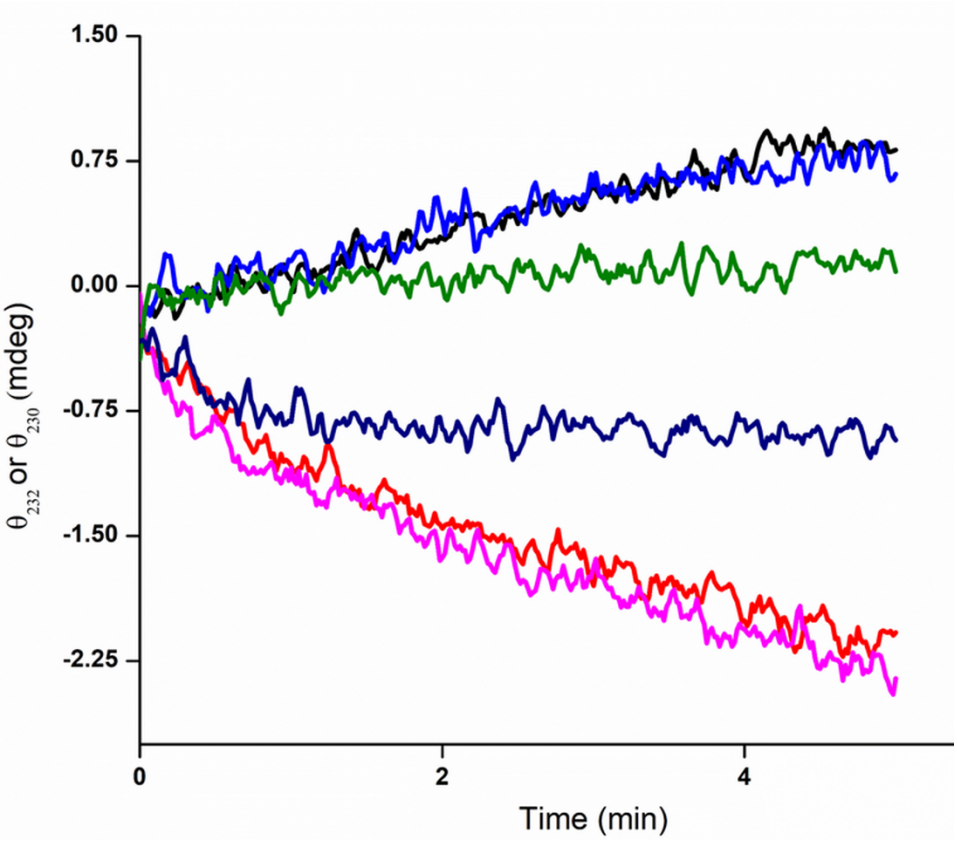

Figure 2-7. Racemase activity of LarA$_{lp}$ using lactate, 2-hydroxybutyrate, and 2-hydroxyisovalerate. 25 mM concentrations of D-lactate (black), L-lactate (pink), D-2-hydroxybutryate (blue), L-2-hydroxybutyrate (red), D-2-hydroxyisovalerate (green), and L-2-hydroxyisovalerate (dark gray) in 60 mM phosphate buffer (pH 7.4) were mi˙å®xed with 100 nM enzyme and the changes in ellipticities were monitored at 232 nm (lactate and 2-hydroxyisovalerate) or 230 nm (2-hydroxybutyrate) and 55 °C.

hydroxybutyrate and, very poorly, 2-hydroxyisovalerate (**Figure 2-7**). Unfortunately, we were not able to define the kinetic parameters for LarA$_{lp}$ utilization of these substrates because the assay method lacked sufficient sensitivity; i.e., the changes in ellipticity were too small at low substrate concentrations so that the K_m values were not reliably determined. While performing these studies we also noted that the signal amplitudes of these substrates were temperature dependent (**Figure 2-A7**), with the signal diminishing in intensity at the growth temperature of the microorganism compared to the intensities measured at 25 °C. The wavelengths associated with the maximal ellipticities also shift with temperature as most readily seen in the case of 2-hydroxyisovalerate. These points should be considered when designing assay experiments using this approach for other 2-hydroxyacid racemases/epimerases purified from thermophilic microorganisms. The CD spectra of 2-hydroxymandelate[17] and several other 2-hydroxyacids (**Figure 2-A8**) demonstrates the potential broad utility of this assay method.

Discussion

In this work we have made important progress to overcome two roadblocks that had hindered analysis of NPN cofactor-related proteins from diverse microorganisms. First, the ability to synthesize active NPN-containing LarA homologs encoded by the Duet plasmids in *E. coli*, a genetically tractable microorganism, greatly facilitates functional analysis of these enzymes. In addition, the Duet plasmid system allows for testing the functionality of the non-*L. plantarum* encoded proteins. Second, the circular dichroism assay is broadly applicable to 2-hydroxyacid racemases and avoids the need to develop a distinct coupled enzyme assay for each substrate.

The Duet plasmid system is a significant advance for generating active LarA homologs; however, these cofactor-containing proteins are labile and, even when freshly examined, they may not contain stochiometric levels of the NPN cofactor. This system potentially could be improved by growing the cells anaerobically, utilizing an *E. coli* strain lacking a nickel exporter (e.g., *rcnA*), or using a strain with altered nickel homeostasis (such as one with a defective NikR) that would lead to continuous importation of the metal ion. Modifications to the growth conditions (e.g., reduction in temperature, provision of NaAD precursors, and supplementation with greater levels of nickel ions) may further improve the extent of activation for the LarA-like proteins. Additional engineering of the strain could boost the production of NaAD in the cell by increasing expression of the genes in the Preiss-Handler pathway (i.e., genes encoding nicotinic

acid phosphoribosyl transferase and nicotinate-nucleotide adenylyltransferase.[24] The use of additional compatible plasmids with complementary antibiotic resistance markers from the Duet expression system[14] could allow the introduction of genes encoding other enzymes that could facilitate alternative NPN biosynthesis pathways. For example, one could incorporate *iscS* or other iron-sulfur biosynthetic genes to assist with synthesis of LarE versions containing a [4Fe-4S] cluster.[7]

The CD-based assay approach for direct and continuous monitoring of racemase and epimerase activities, described above, is well suited for use with a range of 2-hydroxyacids. This novel approach for monitoring this family of enzymes does, however, have limitations such as when trying to obtain kinetic parameters for enzymes with very low K_m values or when monitoring enzymes obtained from thermophiles. Temperature dependent CD spectra similar to what we describe have been reported previously for other compounds.[25, 26] In selected situations the continuous monitoring of racemase/epimerase reactions also is possible by use of coupled enzyme spectrophotometric assays, such as by linking lactate racemase with L- or D-lactate dehydrogenase activities;[15] however, these coupled assays are typically performed in a discontinuous manner.[10] Another discontinuous method to assess these activities uses capillary electrophoresis with the chiral selector vancomycin to separate the enantiomers.[23] Together, these assays offer investigators multiple methods to extend our understanding of the activities of NPN-containing enzymes.

REFERENCES

(1) Desguin B, Zhang T, Soumillion P, Hols P, Hu J, Hausinger RP. **2015**. A tethered niacin-derived pincer complex with a nickel-carbon bond in lactate racemase. *Science* 349:66-69.

(2) Desguin B, Goffin P, Viaene E, Kleerebezem M, Martin-Diaconescu V, Maroney MJ, Declercq J-P, Soumillion P, Hols P. **2014**. Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system. *Nat Commun* 5:3615.

(3) Desguin B, Soumillion P, Hols P, Hausinger RP. **2016**. Nickel-pincer cofactor biosynthesis involves LarB-catalyzed pyridinium carboxylation and LarE-dependent sacrificial sulfur insertion. *Proc Natl Acad Sci USA* 113:5598-5603.

(4) Rankin JA, Chatterjee S, Tariq Z, Lagishetty S, Desguin B, Hu J, Hausinger RP. **2021**. The LarB carboxylase/hydrolase forms a transient cysteinyl-pyridine intermediate during nickel-pincer nucleotide cofactor biosynthesis. *Proc Natl Acad Sci USA* 118:e2106202118.

(5) Fellner M, Desguin B, Hausinger RP, Hu J. **2017**. Structural insights into the catalytic mechanism of a sacrificial sulfur insertase of the N-type ATP pyrophosphatase family, LarE. *Proc Natl Acad Sci USA* 114:9074-9079.

(6) Fellner M, Rankin JA, Desguin B, Hu J, Hausinger RP. **2018**. Analysis of the active site cysteine residue of the sacrificial sulfur insertase LarE from *Lactobacillus plantarum*. *Biochemistry* 57:5513-5523.

(7) Chatterjee S, Parson KF, Ruotolo BT, McCracken J, Hu J, Hausinger RP. **2022**. Characterization of a [4Fe-4S]-dependent LarE sulfur insertase that facilitates nickel-pincer nucleotide cofactor biosynthesis in *Thermotoga maritima*. *J Biol Chem* 298:102131.

(8) Desguin B, Fellner M, Riant O, Hu J, Hausinger RP, Hols P, Soumillion P. **2018**. Biosynthesis of the nickel-pincer nucleotide cofactor of lactate racemase requires a CTP-dependent cyclometallase. *J Biol Chem* 293:12303-12317.

(9) Turmo A, Hu J, Hausinger RP. **2022**. Characterization of the nickel-inserting cyclometallase LarC from *Moorella thermoacetica* and identification of a cytidinylylated reaction intermediate. *Metallomics* 14:mfac014.

(10) Desguin B, Urdiain-Arraiza J, Da Costa M, Fellner M, Hu J, Hausinger RP, Desmet T, Hols P, Soumillion P. **2020**. Uncovering a superfamily of nickel-dependent hydroxyacid racemases and epimerases. *Sci Rep* 10:18123.

(11) Garcia-Nafria J, Watson JF, Gregor IH. **2016**. IVA cloning: A single-tube universal cloning system exploiting bacterial *in vivo* assembly. *Sci Rep* 6:27459.

(12) Huang F, Spangler JR, Huang AY. **2017.** *In vivo* cloning of up to 16 kb plasmids in *E. coli* is as simple as PCR. *PLoS One* 12:e0183974.

(13) Hanahan D, Jessee J, Bloom FR. **1991**. Plasmid transformation of *Escherichia coli* and other bacteria. *Meth Enzymol* 204:63-113.

(14) Tolia NH, Joshua-Tor L. **2006**. Strategies for protein expression in *Escherichia coli*. *Nature Methods* 3:55-64.

(15) Hausinger RP, Hu J, Desguin B. **2023**. The nickel-pincer coenzyme of lactate racemase: A case study of uncovering cofactor structure and biosynthesis. *Meth Enzymol* 685:in press.

(16) Studier FW. **2014**. Stable expression clones and auto-induction for protein production in *E. coli*. *In* Chen YW (ed), Structural Genomics: General Applications. Springer, New York, NY.

(17) Bearne SL, Hayden JA. **2023**. Application of circular dichroism-based assays to racemases and epimerases: Recognition and catalysis of reactions of chiral substrates by mandelate racemase. *Meth Enzymol* 685:127-169.

(18) Xie L. **2018**. Enzyme-catalyzed acylium ion formation and reagents for the detection of thiocarboxylates involved in the biosynthesis of the NAD derived pincer cofactor. Ph.D. Texas A & M University, College Station, TX.

(19) Krishnamoorthy K, Begley TP. **2010**. Reagent for the detection of protein thiocarboxylase in the bacterial proteome: Lissamine rhodamine B sulfonyl azide. J Am Chem Soc 132:11608-11612.

(20) Krishnamoorthy K, Begley TP. **2011**. Protein thiocarboxylate-dependent methionine biosynthesis in *Wolinella succinogenes*. J Am Chem Soc 133:379-386.

(21) Wessel D, Flugge UI. **1984**. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 138:141-143.

(22) Chen L, Shen Y, Wang C, Ding L, Zhao F, Wang M, Fu J, Wang H. **2019**. *Megasphaera elsdenii* lactate degradation pattern shifts in rumen acidosis models. *Front Microbiol* 10:162.

(23) Urdiain-Arraiza J, Desguin B. **2024**. Versatile capillary electrophoresis method for the direct chiral separation of aliphatic and aromatic a-hydroxy acids, b-hydroxy acids and polyhydroxy acids using vancomycin as chiral selector. *J Chromatogr A* 1715:464611.

(24) Yang L, Mu X, Nie Y, Xu Y. **2021**. Improving the production of NAD^+ via multi-strategy metabolic engineering in *Escherichia coli*. *Metab Eng* 64:122-133.

(25) Wiberg KB, Wang Y-g, Murphy MJ, Vaccaro PH. **2004.** Temperature dependence of optical rotation: a-Pinene, b-pinene, camphene, camphor and fenchone. *J Phys Chem A* 108:5559=5563.

(26) Mort BC, Autschbach J. **2007**. Temperature dependence of the optical rotation in six bicyclic organic molecules calculated by vibrational averaging. *ChemPhysChem* 8:605-616.

(27) de Ruyter PG, Kuipers OP, de Vos WM. **1996**. Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl Environ Microbiol* 62:3662-3667.

APPENDIX

Table 2-A1 List of strains, plasmids, and primers

Strain, plasmid or primer	**Characteristic(s) or sequence**		**Source or reference**
Strains			
Lc. Lactis			
NZ3900	MG1363 derivative		27
E. coli			
DH5α	F– φ80lacZΔ M15 Δ (*lacZYA-argF*) *U169 recA1 endA1 hsdR17* (rK– mK+) *phoA supE44* λ- *thi–1 gyrA96 relA1*		ThermoFisher
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS* λ *DE3 =* λ *sBamHIo* Δ*EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21* Δ*nin5*		NEB
Plasmids			
pETDuet	*Amp*^{r}		Novagen
pRSFDuet	*Kan*^{r}		Novagen
pAT035	*Amp*^{r}	$LarA_{Lp}$ and $LarB_{Lp}$ expression; $LarA_{Lp}$ purification	This study
pAT038	*Kan*^{r}	$LarE_{Lp}$ and $LarC_{Lp}$ expression	This study
pAT039	*Kan*^{r}	$LarE_{Lp}$ and $LarC_{Mt}$ expression	This study
pAT040	*Kan*^{r}	$LarE_{Lp}$ and $LarC_{Sc}$ expression	This study
pETDuet-LarA-Me	*Amp*^{r}	$LarA_{Me}$ and $LarB_{Lp}$ expression	This study
pETDuet-LarB-Sc	*Amp*^{r}	$LarA_{Lp}$ and $LarB_{Sc}$ expression	This study
pRSFDuet-LarEC-Sc	*Kan*^{r}	$LarE_{Sc}$ and $LarC_{Sc}$ expression	This study
pGIR112		For overexpression of LarA fused with a Strep-tag at the C-terminus from the *larABC1C2DE* operon	2
pGIR210-LarAH31	*Chl*^{r}	For overexpression of $LarA_{lp}$ fused with a Strep-tag at the C-terminus from the *larABC1C2DE* operon	This study
Primers	The orientation is 5’ to 3’ in all cases.		
LarALp-Strep_fw	CTTTAAGAAGGAGATATACCATGTCCGTTGCAATTGATTTACCATATGACAA		Subcloning
LarALp-Strep_rv	CCGCAAGCTTGTCGACCTACTTCTCAAATTGTGGATGACTCCAGC		Subcloning
pETDuet_MCS1_fw	CATCCACAATTTGAGAAGTAGCTTAAGTCGAACAGAAAGTAATCGTATTGTAC		Subcloning

Table 2-A1 (cont'd)

pETDuet_MCS1_rv	CATATGGTAAATCAATTGCAACGGACATGGTATATCTCCTTCTTAAAG	Subcloning
LarBLp_fw	GTTAAGTATAAGAAGGAGATATACATATGGCAACCACAGCAGAAATATTACAACAAGTG	Subcloning
LarBLp_rv	CCAGACTCGAGGGTACCTTACATTTGATTGACCATACTAGCTGAGTAGG	Subcloning
pETDuet_MCS2_fw	GTATGGTCAATCAAATGTAAGGTACCCTCGAGTCTGGTAAAG	Subcloning
pETDuet_MCS2_rv	CTGCTGTGGTTGCCATATGTATATCTCCTTCTTATACTTAACTAATATAC	Subcloning
LarAMe_fw	TTAAGAAGGAGATATACCATGGCAATGAAGACATT	Subcloning
LarAMe_rv	GATGACTCCAGCTAGCCTTTCGGACCGGAAA	Subcloning
pETDuet_MCS2_Sc_fw	CAGGGAGGATAACCTAGGCTGCTGCCAC	Subcloning
pETDuet_MCS2_Sc_rv	CCACTCACCATATGTATATCTCCTTCTTATACTTAAC	Subcloning
LarELp_fw	CTTTAATAAGGAGATATACCATGGCAACATTAGCAACAAAAAAAGCAACGTTAGTA	Subcloning
LarELp_rv	CTGTTCGACTTAAGCTAGGCGAAAGTGGCCAATTG	Subcloning
pRSFDuet_MCS1_fw	CCACTTTCGCCTAGCTTAAGTCGAACAGAAAGTAATCGTATTGTACA	Subcloning
pRSFDuet_MCS1_rv	GCTAATGTTGCCATGGTATATCTCCTTATTAAAGTTAAACAAAATTATTTC	Subcloning
LarCLp_fw	GTATAAGAAGGAGATATACATATGGGTGCTCAAACACTTTATTTAGACGCTTTTTC	Subcloning
LarCLp_rv	CCAGACTCGAGGGTACCTTACGCCTCCTCATCTAATTGATCTACCG	Subcloning
pRSFDuet_MCS2_fw	GATGAGGAGGCGTAAGGTACCCTCGAGTCTGGTAAAG	Subcloning
pRSFDuet_MCS2_rv	GCGTCTAAATAAAGTGTTTGAGCACCCATATGTATATCTCCTTCTTATACTTAAC	Subcloning
pRSF_MCS1_Sc_fw	CGTTTTGACTAAATGACTTAAGTCGAACAGAAAGTAATCG	Subcloning
pRSF_MCS1_Sc_rv	GGTTTTCTAATTTCGACTCTAACATGGTATATCTCCTTATTAAAGTTAAACAAAATTATT	Subcloning
LarCSc_fw	GAAGGAGATATACATATGGGTCTGATCGCC	Subcloning
LarCSc_rv	GACTCGAGGGTACCTTAGCTTTCCGG	Subcloning
LarBSc_fw	AGGAGATATACATATGGTGAGTGGGCCGGTC	Subcloning
LarBSc_rv	CCTAGGTTATCCTCCCTGACCACCGT	Subcloning
LarESc_fw	CTTTAATAAGGAGATATACCATGTTAGAGTCGAAATTAGAAAACC	Subcloning

Table 2-A1 (cont'd)

LarESc_rv	GTTCGACTTAAGTCATTTAGTCAAAACGCGATTG	Subcloning
pRSFDuet_MCS2_Sc_fw	CTGAGTCCGGAAAGCTAAGGTACCCTCGAG	Subcloning
pRSFDuet_MCS2_Sc_rv	CAAAATAGGCGATCAGACCCATATGTATATCTCC	Subcloning
LarCMt_fw	GAAGGAGATATACATATGAAGATCGCCTATTTTGATTGCTTTAGC	Subcloning
LarCMt_rv	GACTCGAGGGTACCTTAAAATGCTTTCAGTGCACGTGCCGC	Subcloning
pRSFDuet_MCS2_Mt_fw	GCACTGAAAGCATTTTAAGGTACCCTCGAGTCTGG	Subcloning
pRSFDuet_MCS2_Mt_rv	CAAAATAGGCGATCTTCATATGTATATCTCCTTCTTATACTTAAC	Subcloning

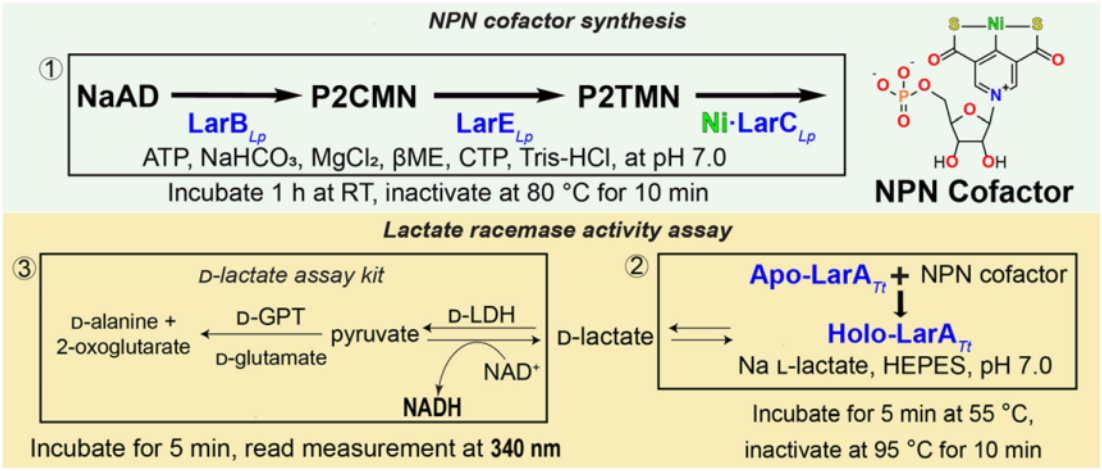

Figure 2-A1. Workflow of the conventional *in vitro* Lar assay. (1) The genes encoding LarB, LarE, and LarC from *L. plantarum* were separately expressed in *E. coli* and the purified proteins were used to convert NaAD via P2CMN and P2TMN intermediates to the NPN cofactor. (2) The cofactor was incorporated into the purified *T. thermosaccharolyticum* LarA apoprotein or into a LarA homolog from another microorganism. (3) The activated LarA was mixed with L-lactate and racemization was assayed spectrophotometrically by quantifying the amount of D-lactate produced based on the reduction of NAD^+ by D-lactate dehydrogenase (D-LDH). D-Glutamate-pyruvate transaminase (D-GPT) and added D-glutamate were used to drive the reaction to completion. βME, β-mercaptoethanol; RT, room temperature.

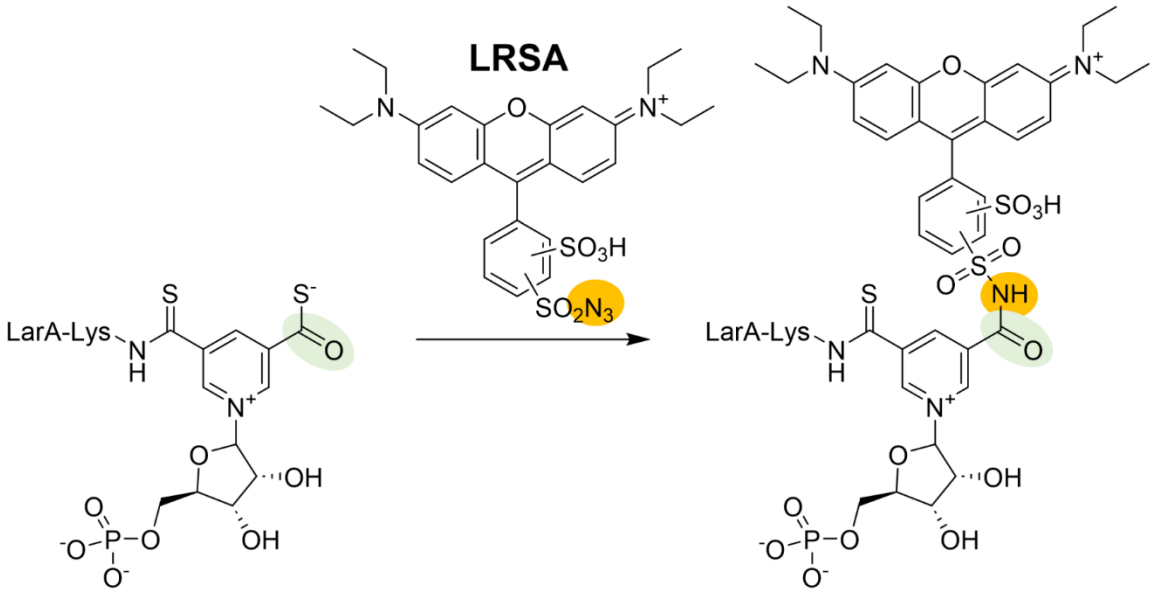

Figure 2-A2. LRSA labeling of the protein bound NPN cofactor. The "click chemistry" reaction between the sulfonyl azide and the thiocarboxylate occurs in the dark at room temperature.

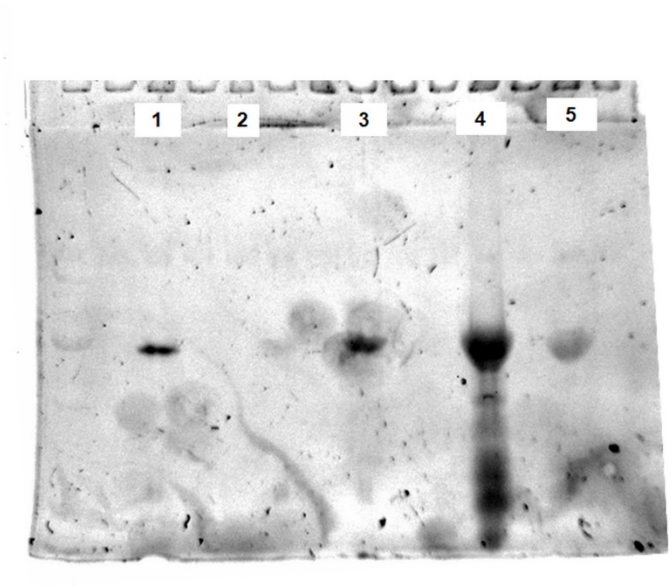

Figure 2-A3. LRSA labeling of LarA$_{Me}$ generated in the *E. coli* Duet system. Samples that were examined for LRSA reactivity include: Lane (1) purified control sample LarA$_{Lp}$ (~130 μg), (2) purified LarA$_{Me}$ apoprotein (~120 μg), (3) lysates containing holoprotein LarA$_{Me}$, (4) isolated holoprotein LarA$_{Me}$ (~180 μg), and (5) the same holoprotein LarA$_{Me}$ (~70 μg).

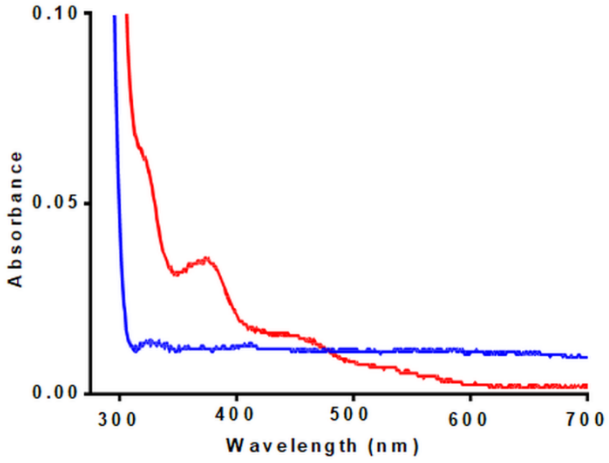

Figure 2-A4. UV-vis spectrum of LarA$_{Me}$ isolated from the *E. coli* Duet system. Spectra are compared for the Lar$_{Me}$ apoprotein (offset blue trace) and holoprotein (red trace) at 12 mg/mL.

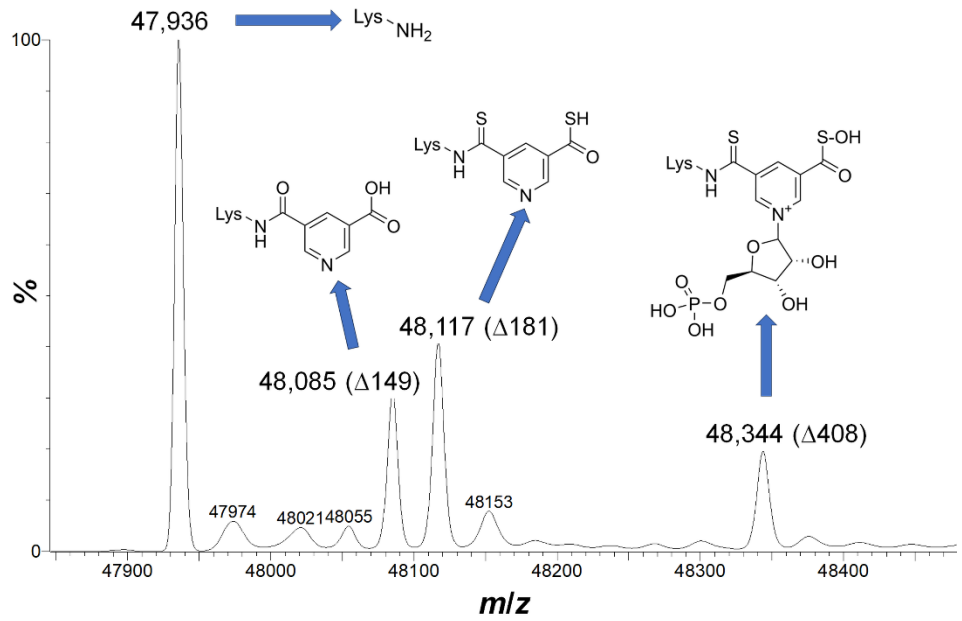

Figure 2-A5. Mass spectrum of LarA$_{Me}$ purified from the *E. coli* Duet expression system. Potential structures matching each feature are illustrated, where Lys is a side chain of the protein.

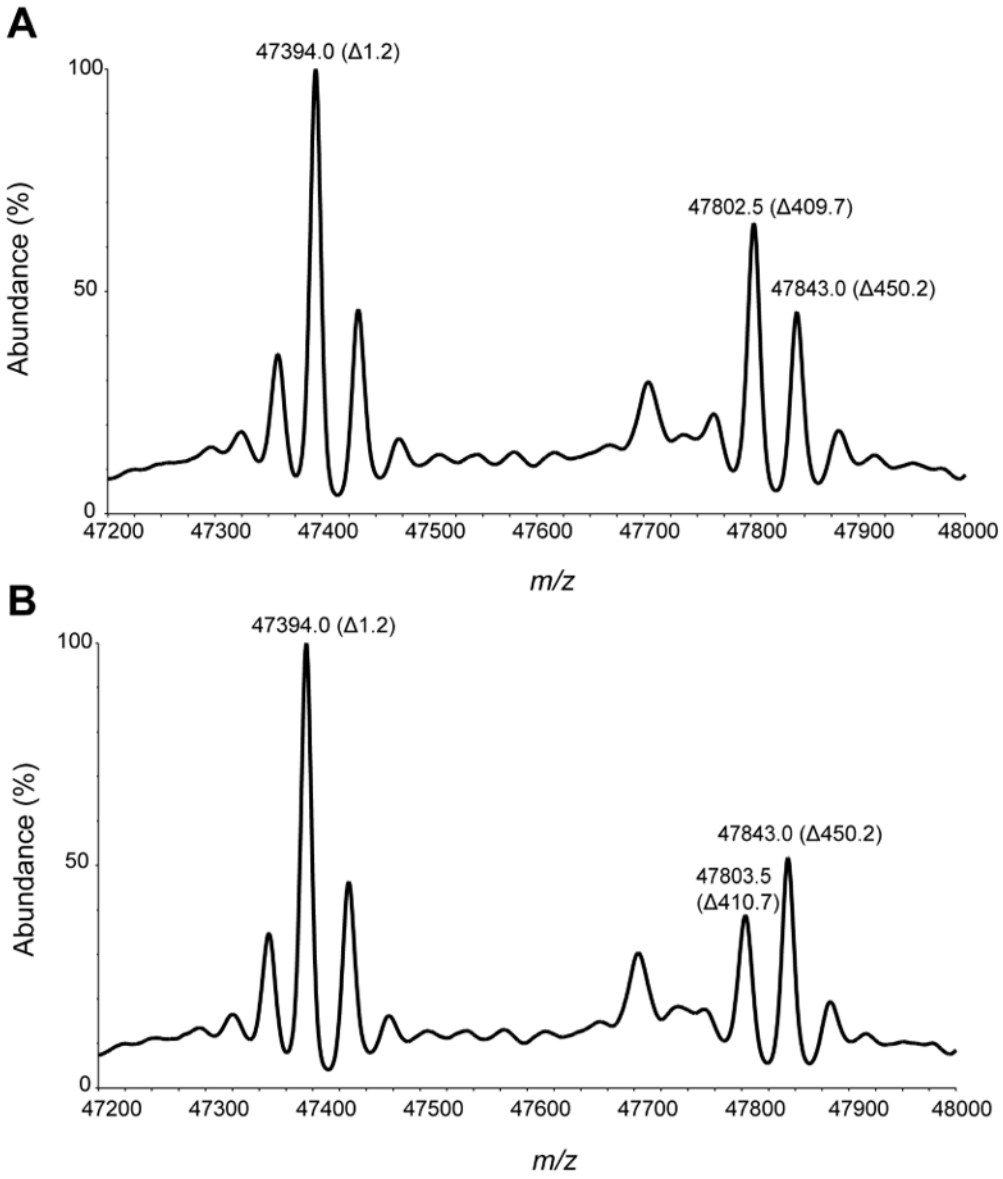

Figure 2-A6. Mass spectrum of LarA$_{Lp}$ purified from the *E. coli* expression system with the gene encoding LarC$_{Lp}$ replaced by genes encoding (A) LarC$_{Mt}$ and (B) LarC$_{Sc}$. The percent abundance is relative to the largest peak and values in parentheses are the mass differences to the theoretical mass of LarA$_{Lp}$ apoprotein lacking the N-terminal methionine.

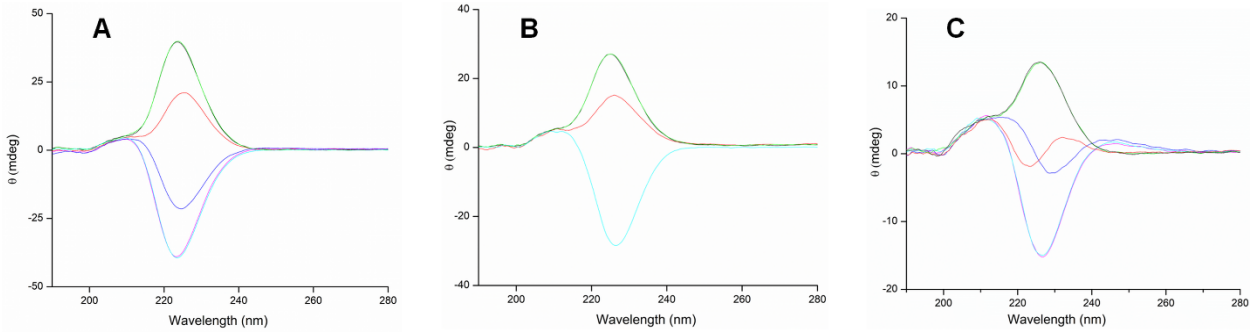

Figure 2-A7. Temperature dependence of the ellipticity for lactate, 2-hydroxybutyrate, and 2-hydroxyisovalerate. (A) Ellipticity of 25 mM L-lactate at 25 °C (black), 55 °C (red), and returned to 25 °C (green) and of 25 mM D-lactate treated similarly (cyan, blue, and magenta). (B) Ellipticity of 25 mM L-2-hydroxybutyrate at 25 °C (black), 55 °C (red), and returned to 25 °C (green) and of 25 mM D-2-hydroxybutyrate at 25 °C (cyan). The signal from D-2-hydroxybutyrate also decreases in intensity at higher temperature (not shown). (C) Ellipticity of 25 mM L-2-hydroxyisovalerate at 25 °C (black), 55 °C (red), and returned to 25 °C (green) and of 25 mM D-2-hydroxyisovalerate treated similarly (cyan, blue, and magenta). All spectra were collected using 60 mM phosphate buffer at pH 7.4.

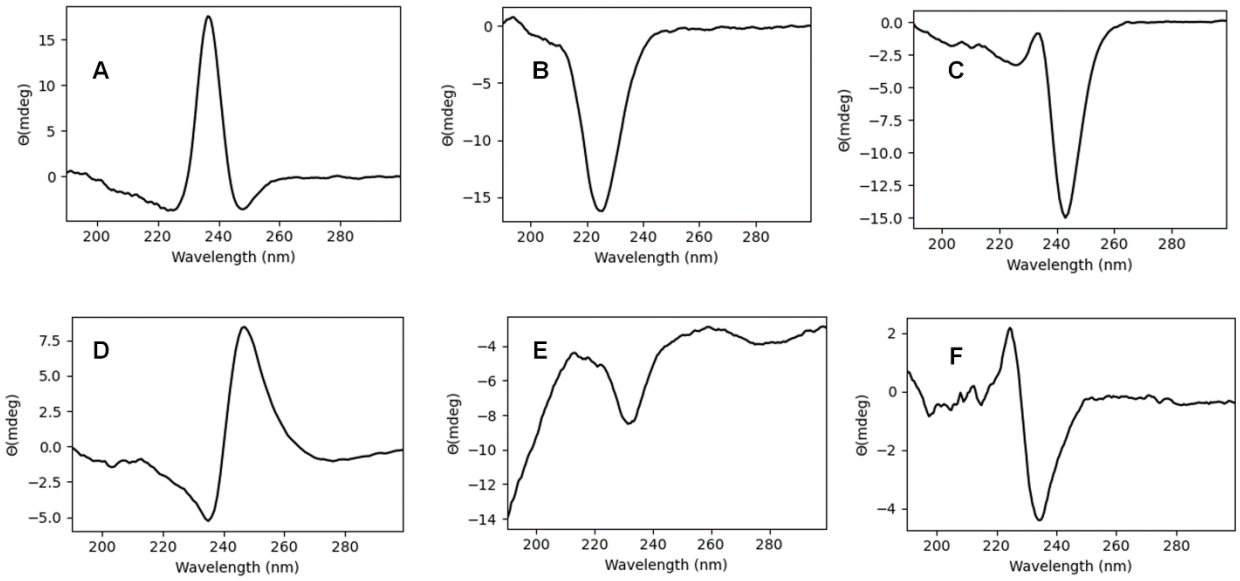

Figure 2-A8. CD spectra of selected other 2-hydroxyacids. (A) L-Malate. (B) D-Phosphoglycerate. (C) L-2-Hydroxy-3-methylbutyric acid. (D) L-2-Hydroxyisocaproate. (E) D-2-Hydroxyglutarate. (F) D-Gluconate. The samples contained 10 mg/mL of hydroxyacid in water and were examined at 35 °C.

CHAPTER 3

Structure of the LarB-substrate complex and identification of a reaction intermediate during nickel-pincer nucleotide cofactor biosynthesis

Adapted with permission from:

Chatterjee, S.; Nevarez, J. L.; Rankin, J. A.; Hu, J.; Hausinger, R. P. Structure of the LarB-Substrate Complex and Identification of a Reaction Intermediate during Nickel-Pincer Nucleotide Cofactor Biosynthesis. *Biochemistry* 2023, 62 (21), 3096–3104.

For this chapter, my primary contributions are the experimental design, execution, and data analysis of the mass spectrometric assays and UV-Vis absorbance experiments. Dr. Shramana Chatterjee carried out the structural studies under the guidance of Dr. Jian Hu.

Abstract

LarB catalyzes the first step of biosynthesis for the nickel-pincer nucleotide cofactor by converting nicotinic acid adenine dinucleotide (NaAD) to AMP and pyridinium-3,5-biscarboxylic acid mononucleotide (P2CMN). Prior studies had shown that LarB uses CO_2 for substrate carboxylation and reported the structure of a *Lactiplantibacillus plantarum* LarB·NAD^+ complex, revealing a covalent linkage between Cys221 and C4 of the pyridine ring. This interaction was proposed to promote C5 carboxylation, with C5-carboxylated-NaAD suggested to activate a magnesium-bound water, leading to phosphoanhydride hydrolysis. Here we extended the analysis of wild-type LarB by using ultraviolet-visible spectroscopy to obtain additional evidence for cysteinyl side chain attachment to the ring of NAD^+, thus demonstrating this linkage is not a crystallization artifact. Using the S127A variant of *L. plantarum* LarB, a form of the enzyme with a reduced rate of NaAD hydrolysis, we examined its interaction with authentic substrate. The intermediate arising from C5 carboxylation of NaAD, dinicotinic acid adenine dinucleotide (DaAD), was identified by use of mass spectrometry. S127A LarB exhibited spectroscopic evidence for a Cys221-NAD^+ adduct, but a covalent enzyme-NaAD linkage was not detectable. We solved the S127A LarB·NaAD structure, providing new insights into the enzyme mechanism, and tentatively identified the position and mode of CO_2 binding. The crystal structure revealed the location of the side chain for Glu180, which was previously disordered, but showed it is not well positioned to abstract the C5 proton in the adduct species to restore aromaticity as Cys221 is expelled. Based on these combined results, we propose a revised catalytic mechanism of LarB.

Introduction

LarB initiates the biosynthetic pathway of the nickel-pincer nucleotide (NPN) (**Figure 3-1**),[1] a required cofactor of lactate racemase and a large family of other racemases and epimerases.[2,3] In this pathway, LarB catalyzes the carboxylation of nicotinic acid adenine dinucleotide (NaAD) at C5 of the pyridinium ring and hydrolyzes the phosphoanhydride, resulting in formation of pyridinium-3,5-biscarboxylic acid mononucleotide (P2CMN) with the release of AMP.[4] LarE activates each carboxyl group of P2CMN by adenylylation and inserts sulfur atoms to release AMP and form pyridinium-3,5-bisthiocarboxylic acid mononucleotide (P2TMN) by either of two types of reactions, depending on the source of the enzyme.[5,6] LarC catalyzes the cytidinylylation of P2TMN while releasing pyrophosphate (PP_i), inserts nickel,

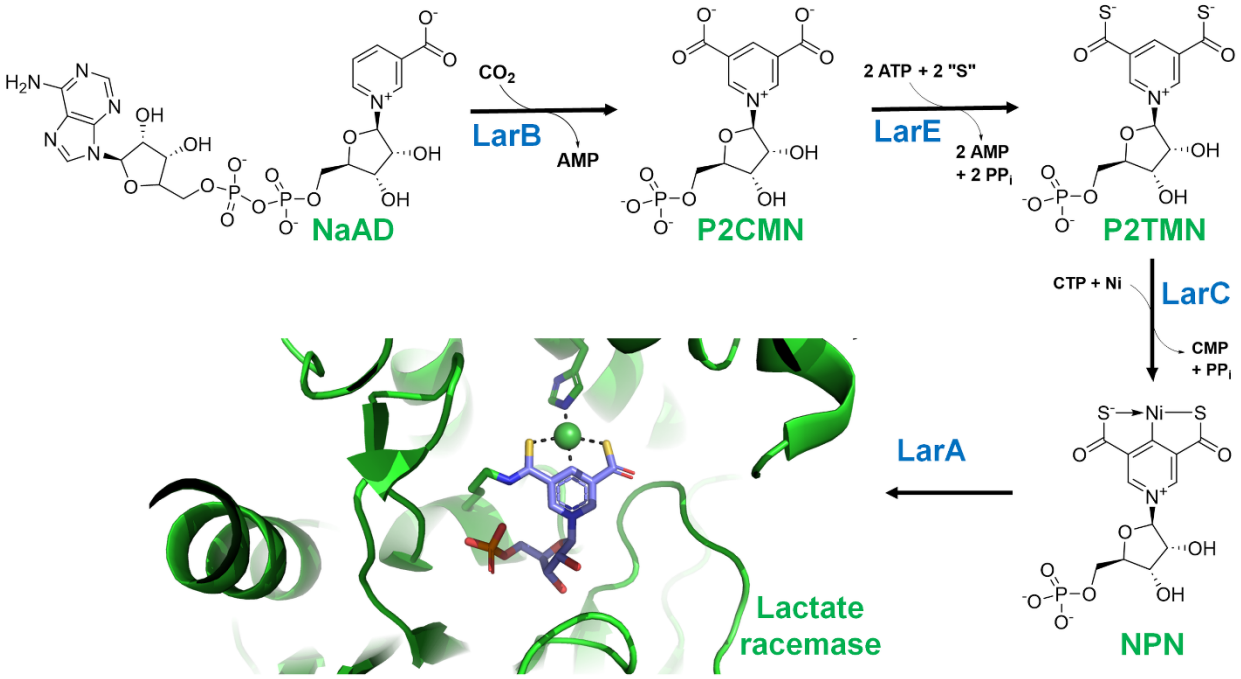

Figure 3-1. Biosynthetic pathway for the nickel-pincer nucleotide (NPN) cofactor. LarB carboxylates the C5 position of nicotinic acid adenine dinucleotide (NaAD) and hydrolyzes the phosphoanhydride to produce pyridinium-3,5-biscarboxylic acid mononucleotide (P2CMN) and AMP. LarE uses ATP to activate each carboxyl group by adenylylation while releasing pyrophosphate (PP_i) and inserts sulfur to form pyridinium-3,5-bisthiocarboxylic acid mononucleotide (P2TMN) as AMP is eliminated. LarC catalyzes the cytidinylylation of P2TMN while releasing PP_i, inserts nickel, then hydrolyzes the phosphoanhydride to produce the NPN cofactor and CMP. The NPN cofactor binds, in some cases covalently, to the active site of the target enzyme, as represented by lactate racemase of *L. plantarum* (PDB ID:5HUQ). A clip of the latter protein is shown in cartoon mode with two side chains depicted as sticks with green carbon atoms, the bound NPN is illustrated in sticks with blue carbon atoms, yellow sulfur atoms, red oxygen atoms, and an orange phosphorus atom, and nickel is illustrated as a green sphere.

then hydrolyzes the phosphoanhydride to produce the NPN cofactor and CMP.[7,8] The cofactor then incorporates into its target enzyme, in some cases, such as the lactate racemase of *Lactiplantibacillus* (formerly *Lactobacillus*) *plantarum*, becoming covalently attached to a lysyl side chain in a thioamide linkage.[2]

Structural and biochemical studies of LarB from *L. plantarum* have revealed several aspects of a previously proposed enzyme reaction mechanism (**Figure 3-2**).[9] By soaking the substrate analog NAD^+ into crystals of LarB, the structure of the LarB·NAD^+ complex was acquired. Notably, Cys221 covalently attaches to the C4 position of the pyridine ring. The

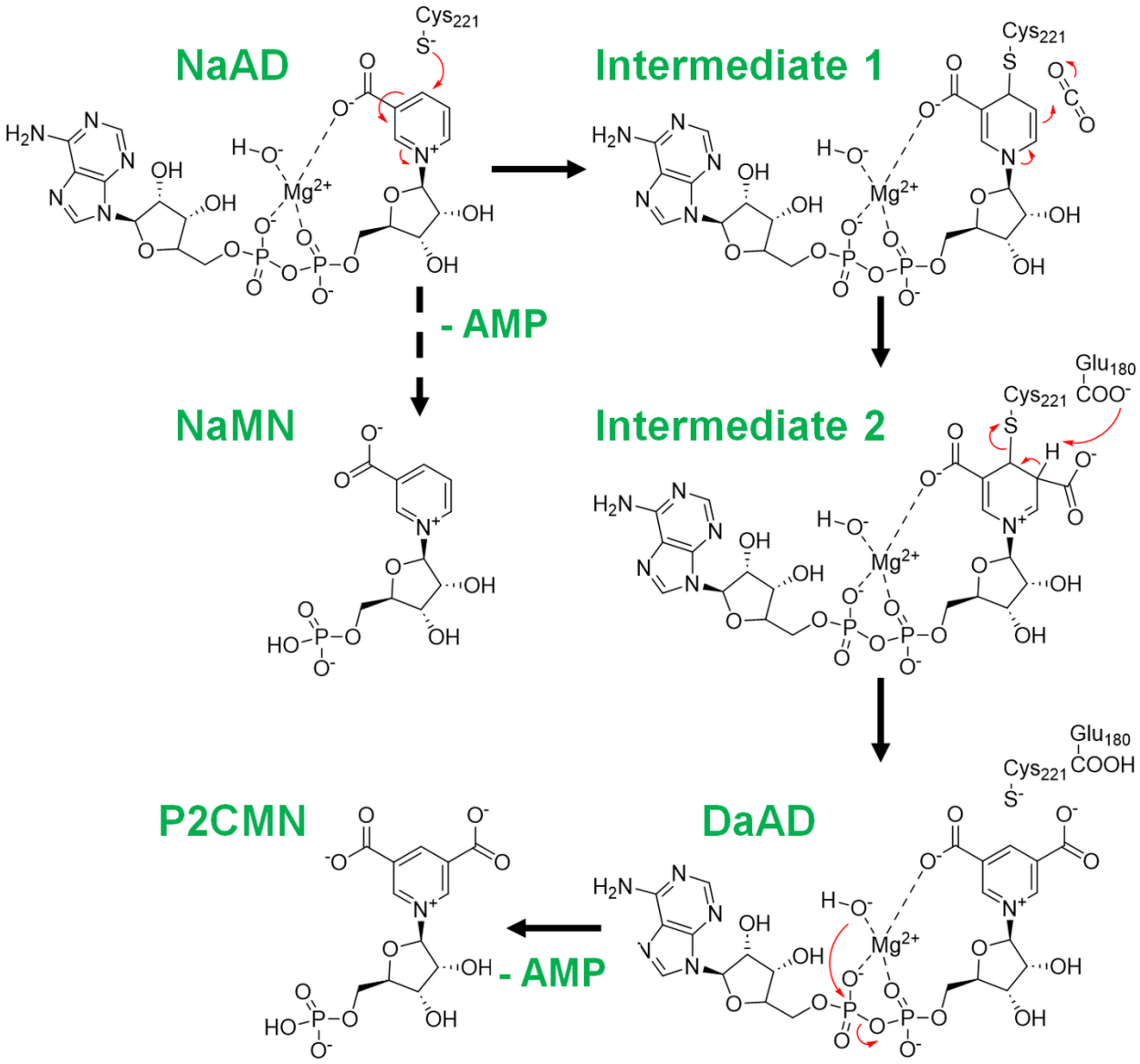

Figure 3-2. Previously proposed reaction mechanism for LarB.[9] Cys221 attacks C4 of NaAD to generate a covalent adduct, labeled intermediate 1. The resulting enhanced nucleophilicity of C5 leads to its carboxylation by CO_2, generating intermediate 2. The general base Glu180 abstracts the C5 proton, and the ring undergoes aromatization as Cys221 is expelled. Carboxyl group interaction with a nearby Mg^{2+} increases the reactivity of a bound solvent molecule. The hydroxide attacks the AMP-associated phosphate group of dinicotinic acid adenine dinucleotide (DaAD), with phosphoanhydride hydrolysis leading to release of AMP and P2CMN. A nonproductive hydrolysis of NaAD by LarB to produce AMP and nicotinic acid mononucleotide (NaMN) is illustrated by the dashed arrow.

enzyme-substrate adduct created by such chemistry would exhibit enhanced nucleophilicity at the C5 carbon atom, which could react with CO_2. Although the amide group of NAD^+ was found to be oriented away from the Mg^{2+} in the enzyme, we had predicted a flipped orientation of the ring in NaAD to allow for carboxylate interaction with the metal ion, providing a reasonable

explanation for why NAD^+ was not utilized as a substrate.[9] The nearby Glu180 was proposed to abstract the C5 proton leading to ring rearomatization as Cys221 is expelled; however, the side chain of this residue was disordered in the structure, so its position was uncertain. The Mg^{2+} coordinated carboxylate groups of dinicotinic acid adenine dinucleotide (DaAD) were suggested to enhance the reactivity of a bound water molecule and promote phosphoanhydride hydrolysis to form P2CMN and AMP. The enzyme also performs a non-productive hydrolysis of NaAD to produce AMP and nicotinic acid mononucleotide (NaMN). Significantly, the postulated DaAD intermediate has not been detected, direct evidence that a cysteinyl-pyridine intermediate participates in catalysis was not described, and no structures were reported for LarB complexes with the authentic substrate, NaAD.

Here, we extended our studies of the wild-type (WT) enzyme and intensively examined the S127A version of *L. plantarum* LarB. Ser127 is a conserved residue that forms a hydrogen bond to the pyridinium mononucleotide phosphate group of the substrate analogue, NAD^+. Therefore, this residue is likely to orient the phosphoanhydride region of NaAD to facilitate hydrolysis. We chose to study the S127A variant of LarB that exhibits a reduced rate of phosphoanhydride hydrolysis.[9] With this variant enzyme species we were able to use mass spectrometry (MS) to demonstrate the intermediacy of DaAD in the reaction and we successfully determined the structure of the S127A LarB·NaAD complex.

Materials and methods

Materials

NaAD, NaCl, $MgCl_2$, $ZnSO_4$, and magnesium formate were acquired from Sigma (St. Louis, MO, USA). Kanamycin, isopropyl β-d-1-thiogalactopyranoside (IPTG), and Tris were purchased from Gold Bio (St. Louis, MO, USA). $NaHCO_3$ was purchased from J. T. Baker. EDTA was purchased from Invitrogen. 50% PEG3350 and 50% ethylene glycol were purchased from Rigaku and Hampton, respectively. All other chemicals were reagent grade or better.

Purification of WT and S127A versions of LarB

We previously described the construction of overexpression plasmids encoding Strep II-tagged WT and S127A versions of LarB, the transformation of these plasmids into *E. coli* BL21 (DE3), and the purification of the two proteins.[9] The final buffer solution for the samples was 100 mM Tris-HCl (pH 7.5) containing 150 mM NaCl.

Mass spectrometric analysis

To assess reaction progress of WT and S127A LarB, we incubated the proteins (150 μM) with 50 mM $NaHCO_3$, 50 mM $MgCl_2$, and 0.2 mM NaAD in 100 mM Tris-HCl, pH 7.2, buffer at room temperature. The reactions were quenched either by adding four volumes of acetonitrile or by heat treatment at the times shown in **Figure 3-3**. The LarB reaction samples were analyzed using a Waters G2-XS Q-TOF mass spectrometer interfaced with a Waters Acquity UPLC. The 10-μL samples were injected onto a Waters Acquity UPLC BEH-C18 column (2.1 × 100 mm) that was held at 40 °C. The reaction components were separated by ion-pairing liquid chromatography (LC) using a binary gradient as follows: the initial conditions used 100% mobile phase A (10 mM tributylamine and 15 mM acetic acid in a 97:3 water/methanol (v/v) mixture) and 0% mobile phase B (methanol); this was held for 1 min, a linear ramp was used to increase to 99% B at 7 min, this was held until 8 min, then the solution was returned to 100% A at 8.01 min and held until 10 min. The flow rate was 0.3 mL/min. Spectra were obtained by time-of-flight electrospray ionization (ESI)-MS operating in negative ion mode with a capillary voltage of 2.0 kV, source temperature of 100 °C, cone voltage at 35 V, desolvation temperature of 350 °C, desolvation gas flow of 600 L/h, and cone gas flow of 50 L/h. Data were acquired using a data-independent MS^E method (scans with fast switching between no collision energy and using a collision energy ramp of 20–80 V) across an *m*/*z* range of 50–1500. The total ion chromatogram (TIC) and extracted ion chromatograms (EIC) for NaAD-derived metabolites were monitored. Daughter ion spectra were acquired for *m*/*z* 707.07 using an MS/MS method with selection in the quadrupole and fragmentation using a collision energy ramp of 10–60 V. Lockmass correction was performed in MassLynx software (Waters Corp.) using leucine enkephalin as the reference compound.

ESI-MS data were acquired for WT and S127A LarB proteins as purified (10 μM) and after short-term (approximately 10 s) incubation with NaAD (200 μM) using methods that were previously described.[10] All solutions were argon purged prior to mixing to minimize carboxylation of the substrate.

Ultraviolet-visible (UV-vis) Spectroscopy

Spectra were obtained for NAD^+ and NaAD and their $NaBH_4$-reduced forms. WT and S127A LarB were each titrated with NAD^+ and NaAD, after purging with argon to prevent carboxylation, and the spectra corrected for dilution.

Crystallization

Crystals of S127A LarB were obtained after mixing either 1 or 2 μL of protein samples (~12 mg/mL) with 1 μL of reservoir solution at a final pH of 7.5. The sitting drop reservoir contained 100 μL of 2 mM nitrilotriacetic acid, 0.7 mM $ZnSO_4$, 100 mM magnesium formate (not adjusted for pH), and 10-20% w/v PEG 3350. Zn^{2+} was previously shown to be necessary to obtain crystals of LarB,[9] but the metal must be removed in order for the active site to bind substrate. The S127A LarB crystals, which grew at 4 °C, were transferred to a solution of mother liquor containing 5 or 10 mM EDTA (pH 8), 10 mM of NaAD, 10-20% PEG3350, and 25 mM magnesium formate. After soaking overnight at 4 °C, the crystals were cryoprotected in 25% v/v ethylene glycol, 10 mM NaAD, 10-20% PEG3350, and 25 mM magnesium formate. In addition, some crystals were soaked with the same components of the cryoprotectant saturated with (i.e., 10-15%) CS_2 for not more than 20 min just before freezing. Crystals were flash frozen in liquid nitrogen.

Diffraction Data Collection, Structure Determination, and Analysis

Data sets were collected at a wavelength of 0.979 Å at beamline 21-ID-D in the Advanced Photon Source. To examine for any bound CS_2 in the EDTA-treated and NaAD-soaked LarB S127A crystal, diffraction datasets were collected at a wavelength of 1.549 Å to locate ordered sulfur atoms. Data sets were processed with HKL-2000[11] and iMosFilm,[12] with merging and scaling done using Aimless[13] and Scala.[14] The WT LarB structure (PDB: 7MJ2)[9] was used as the search model in Phaser[15] for molecular replacement The structures were refined using NCS-restrained least-squares refinement with isotropic B-factors. TLS refinement[16] was performed during the final stages. The test set for R_{free} was generated automatically in *Phenix.Refine*, with 5% of the reflections put aside during processing. Model building and refinement were conducted in Coot[17] and Phenix.[16] A simulated annealing composite omit map was generated in Phenix[16]to correct model bias and confirm ligand bindng. To better model NaAD, which was built in eLBOW, a polder map was generated. Data set statistics are listed in **Table 3-A1**. Pymol[18] were used to create structure figures.

Results

Comparison of NaAD-derived metabolites for WT and S127A LarB

The carboxylation kinetics of the S127A variant were previously shown to exhibit decreases in k_{cat} and K_m (for NaAD) compared to WT LarB (0.009 ± 0.006 s^{-1} versus 0.0103 ±

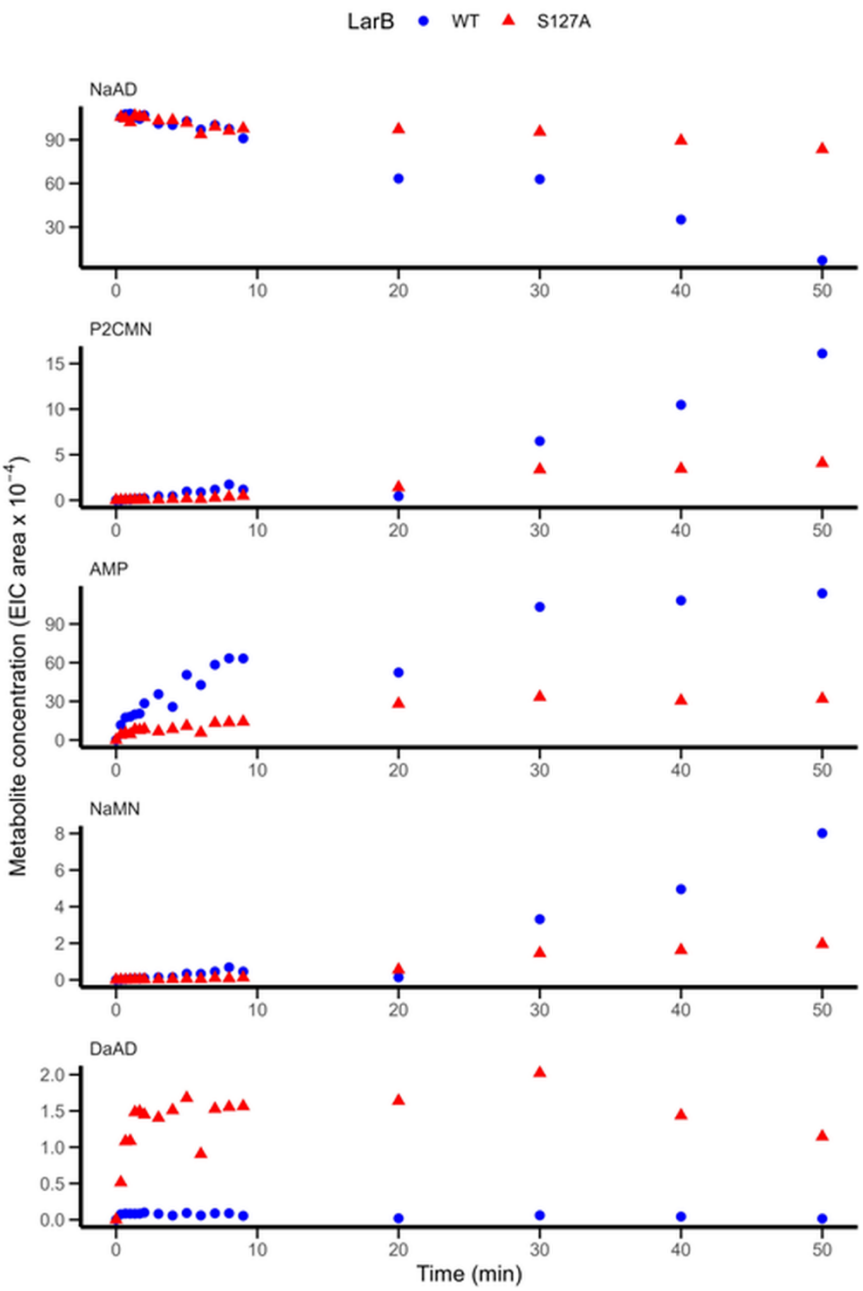

Figure 3-3. Time dependence of NaAD consumption and production of P2CMN, AMP, NaMN, and DaAD by WT (●) and S127A (Δ) LarB. Negative ionization mode LC-ESI-MS was used to generate extracted ion chromatograms of the indicated metabolites, and the corresponding integrated areas are shown for enzyme samples (150 μM) incubated with 50 mM $NaHCO_3$, 50 mM $MgCl_2$, and 0.2 mM NaAD in 100 mM Tris-HCl, pH 7.2, buffer at room temperature after the reaction times shown.

0.0005 s^{-1} and 3 ± 2 μM versus 14 ± 2 μM) with an overall ~40% reduction in the extent of carboxylation using standard assay conditions.[9] Of great interest, however, the S127A variant was shown to be capable of carboxylating ~75% of the NaAD substrate, whereas only ~20% of

the substrate was carboxylated by WT enzyme.[9] We attributed this difference to the nonproductive rapid hydrolysis of NaAD to form NaMN by WT LarB, whereas the S127A LarB exhibited ~90% reduction of this hydrolytic activity.

Consistent with these prior results, LC-ESI-MS analysis (**Figure 3-3** and **Figure 3-A1**) showed that WT LarB consumed NaAD (*m*/*z* 663.08) with an apparent second-order rate constant of 5.9 ± 0.2 x 10^{-5} mM s^{-1}, whereas the S127A variant of LarB exhibited an apparent second-order rate constant of 1.0 ± 0.1 mM s^{-1}, or a WT/S127A LarB ratio of rates of 5.9 ± 0.1. Notably, the linearity of the plots was retained over the entire 50 min of assay, as expected for the very low K_m values of these enzymes and the saturating concentrations of substrate (0.2 mM). We did not assess the precise rate constants of the products of the reactions because standards were not available for some of these compounds to determine actual concentrations; rather, integrated areas for the extracted ion chromatograms were monitored. In addition, we focused on the early time points (up to 540 s) to compare initial rates for most of the metabolites. The ratio of rates for WT LarB versus S127A LarB was 3.8 ± 1.0 for production of the desired product, P2CMN (*m*/*z* 378.02), in line with the previously noted lower production of this metabolite in single time point 10 min assays.[9] The production of AMP (*m*/*z* 346.05), derived from hydrolysis of both the starting substrate and from DaAD, was similarly greater for WT LarB compared to S127A LarB, with a ratio of initial rates of 5.6 ± 1.1. The overall production of AMP at longer time points was similarly greater for WT protein than for the S127A variant which is hindered in hydrolysis activity. In agreement with this finding, the rate of NaAD hydrolysis to produce the non-productive product NaMN (*m*/*z* 334.03) was 5.2 ± 0.9 greater for WT LarB compared to S127A LarB. In sum, these results demonstrate that the WT enzyme catalyzes both the productive and nonproductive hydrolysis reactions faster than the S127A protein, potentially providing the opportunity to visualize the suspected DaAD intermediate more readily in the S127A variant.

Identification of a LarB reaction intermediate

During the above analysis of WT and S127A LarB samples by LC-ESI-MS, we noted the production of another species (*m*/*z* 707.07) that increased in abundance during the first minute, was present in ~20-fold greater levels for S127A LarB compared to the WT protein and diminished at later time points (**Figure 3-3**). The mass and MS/MS fragmentation analysis of this species were consistent with that of DaAD, as shown for the 1 min sample that was

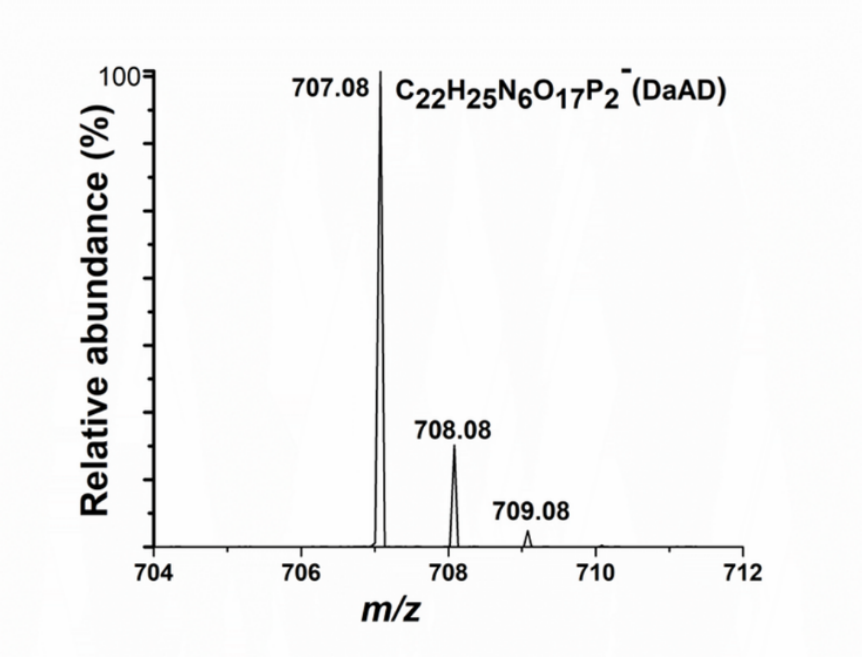

Figure 3-4. Identification of the DaAD intermediate. Negative ionization mode MS analysis of the novel species (*m*/*z* 707.07) formed after 1 min incubation of S127A LarB with NaAD, $MgCl_2$, and bicarbonate, and comparison to the molecular formula of DaAD.

quenched by acetonitrile addition (**Figure 3-4** and **Figure 3-A2**). Significantly, DaAD was not detected when quenching by heat treatment, consistent with the compound exhibiting thermal instability. In addition to detecting DaAD itself, the acetonitrile quenching and LC-ESI-MS approach may have released DaAD from any Intermediate 2 species (**Figure 3-2**) that may have been present.

Spectroscopic detection of a LarB cysteinyl adduct with NAD^+ but not NaAD

We have shown previously that Cys221, an invariant residue of LarB, covalently attaches to C4 of the pyridinium ring of NAD^+ in the crystal structure of the LarB·NAD^+ complex; i.e., we observed continuous electron density between S_γ and C4, separated by 1.8 Å, and noted a boat-like conformation of the pyridine ring consistent with its reduction.[9] To confirm that the

covalent adducts are not crystallization artifacts and to test whether they can be detected with the authentic NaAD substrate, we utilized UV-vis spectroscopy. Reduction of NAD^+ to yield NADH generates the well-known absorbance at 340 nm and an analogous absorbance occurs after hydride addition to NaAD (**Figure 3-A3A,B**). It is also established that thiols and a variety of nucleophiles including tris(alkyl)phosphine reductants can reversibly add to C4 of the pyridinium ring of nicotinamides to generate similar absorbances.[19-22] Thus, we reasoned that a Cys221-pyridine adduct (intermediates 1 and 2, **Figure 3-**2) would also be observed by UV-vis spectroscopy if formed at sufficient abundance. Indeed, we demonstrated that an absorbance with a maximum near 370 nm is generated when NAD^+ was titrated into WT or S127A forms of LarB, with all solutions argon purged to prevent carboxylation (**Figure 3-A3C,D**). In contrast, we were unable to detect any new absorbance feature when adding NaAD to either form of the protein. These results suggest that the protein-substrate adduct, assuming it is a required component of catalysis, represents a high energy state that is formed at very low levels for these conditions.

As an alternative approach to test for the presence of a covalent linkage formed between LarB and its substrate, we mixed WT and S127A LarB with NaAD and as quickly as possible (~10 s) analyzed the samples by ESI-MS. We were unable to detect any evidence for adduct formation in either sample when compared to the untreated protein (**Figure 3-A4**). We conclude that the Cys221-pyridine linkage is formed at undetectable levels and is too unstable for analysis by this method.

Structure of the S127A LarB·NaAD complex

Initial trials to crystallize WT LarB with NaAD were not successful, probably due to the significant hydrolytic activity of the sample. Thus, we used the hydrolysis-hindered S127A variant of the enzyme to solve the structure of the S127A LarB·NaAD complex. Zinc ions were included to generate the S127A LarB crystals, as previously described,[9] with the zinc subsequently was removed by EDTA treatment and the crystals washed then soaked with NaAD and $MgCl_2$. The structure was solved at a resolution of 3.1 Å (**Table 3-A1**) with the RMSDs in the range of 0.43-0.65 Å for each chain when compared to the counterpart in WT LarB (PDB: 7MJ1).[9] As previously noted for the WT enzyme,[9] the asymmetric unit of S127A LarB contained six molecules organized into three dimers (chains A & B, chains C & D, and chains E & F). In line with the prior study, the protein was organized into a tetramer of dimers with its symmetry

mates in the crystal lattice forming the physiologically relevant octamer containing a fourfold axis through the central tunnel. The interface of three protein chains combines to create each active site; e.g., chain C primarily binds the NaMN portion of the molecule, chain C′ (the symmetry mate of chain C) binds the AMP portion, and chain D supplies the critical Cys221 nucleophile. Importantly, we were able to model the authentic substrate NaAD into five out of six active sites of the S127A LarB·NaAD structure (**Figure 3-A5)**, which contrasts with the situation of the previously reported LarB·NAD^+ structure in which WT LarB bound NAD^+ at only two active sites. To minimize confusion, we name the distinct NaAD binding sites based on the chain interacting with the NaMN moiety and minimize mention of other chains at the active sites.

Overall, NaAD binds to the active site in a similar manner as the analog NAD^+;thus, the NaAD carboxylic acid group occupies the same position as the amide group in NAD^+ (**Figure 3-5A**, **Figure 3-A6** and **Figure 3-A7A**). This result was a surprise since we had anticipated the

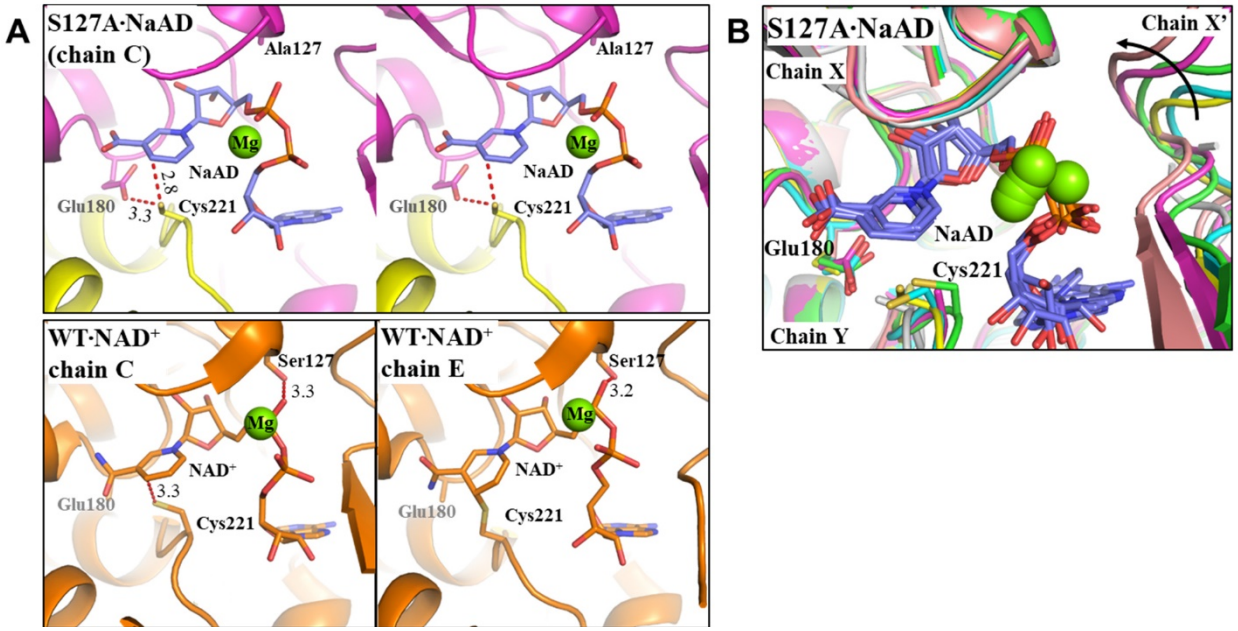

Figure 3-5. Binding of the authentic substrate NaAD to the active sites of S127A LarB. (A) Comparison of NaAD bound to chain C of S127 LarB (upper panel shown in stereoview) and two conformations (lower panel, not a stereoview) of NAD^+ in the WT LarB·NAD^+ structure (PDB: 7MJ1, chains C and E, where only the latter has Cys221 covalently linked to the substrate analog). Key interactions are shown by red dotted lines and the distances are expressed in angstroms. (B) Structural superimposition of the active sites of the six chains in one asymmetric unit. Chain X and chain X' represent the neighboring symmetry mates, and chain Y is the other monomer of the dimeric unit in the octamer. Six chain Xs are structurally aligned to show the different relative orientations of chain X' as indicated by the curved arrow. Mg^{2+} ions are depicted as green spheres.

carboxyl group of NaAD would be positioned on the opposite side of the ring from the amide of NAD^+ where it would coordinate the Mg^{2+} ion, as shown in the previously proposed mechanism of **Figure 3-2**. The active sites from the six chains are not identical due to the different orientations between the neighboring chains, reflecting the plasticity of the active site, with the NaAD molecules in chain A-E adopting different conformations (**Figure 3-5B** and **Figure 3-A6**). These apparent differences may alternatively be due to the relatively low resolution of the dataset. The chains that bind NaAD also bind Mg^{2+} (**Figure 3-A7B**).

Compared with the disordered Glu180 sidechain in the WT LarB·NAD^+ structure, the sidechain of this catalytic residue in the S127A LarB·NaAD structure was successfully resolved in three of the chains (**Figure 3-A7D**). Again inconsistent with the previously proposed mechanism (**Figure 3-2**), this residue is improperly positioned to serve as a general base to abstract the C5 proton (an average distance of 4.9 Å in the current structure) unless the ring flips during catalysis. We note, however, that Glu180 is within hydrogen bond distance of the thiol group of Cys221 (3.1-3.4 Å), another catalytic residue absolutely required for activity. This spatial relationship leads us to suggest that Cys221 and Glu180 form a catalytic dyad that facilitates nucleophilic attack of the thiolate anion on C4 of NaAD.

Unexpectedly, none of the bound NaAD molecules in the S127A LarB·NaAD structure form a covalent bond with Cys221, as was observed in the WT LarB·NAD^+ structure (PDB: 7MJ1), and the cysteinyl residues exhibit distinct conformations (**Figure 3-A7C**). The distances between the Cys221 S atom and C4 of the pyridinium ring occur in the range of 2.7-3.2 Å. It is plausible that the slightly different orientations of the pyridinium ring between the NaAD and NAD^+ bound structures (**Figure 3-A6**) may be partly responsible for the different reactivities, but the electronic differences in the ring from the anionic carboxylate versus the neutral amide species is more likely to affect the stability of adduct. The lack of covalent linkage between the protein and the authentic substrate is consistent with our UV-vis spectroscopy studies described earlier.

A putative CO_2 binding mode

To determine the CO_2 binding site, we soaked S127A LarB·NaAD crystals in different concentrations of $NaHCO_3$ (1 mM to 40 mM, pH 7.5) for different lengths of time (10 s to overnight); however, this process resulted in either destruction of crystals or provided only low-resolution diffraction. We also soaked crystals of the S127A LarB·NaAD complex in buffer

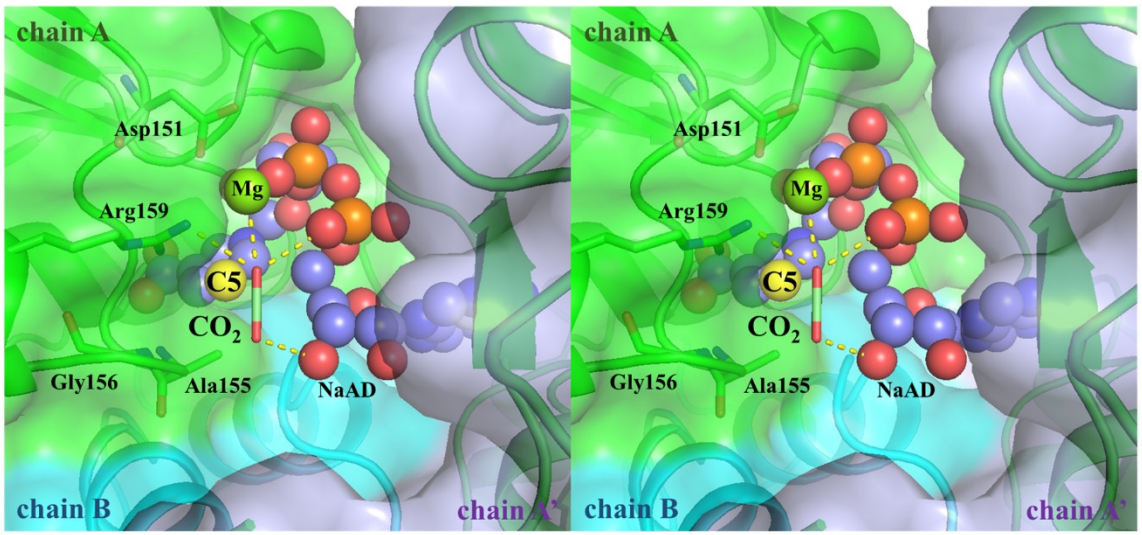

Figure 3-6. The proposed CO_2 binding mode shown in stereoview. CO_2 (stick mode) was modeled into the active site of chain A in the LarB·NaAD complex structure. The proteins (chain A in green, chain A' in light blue, and chain B in cyan) are depicted in surface mode and the residues forming the putative CO_2 channel are shown in stick mode. NaAD is drawn in sphere mode with C5 of the pyridinium highlighted in yellow. In this structural model, CO_2 forms electrostatic interactions with Mg^{2+}, Arg159, and the hydroxyl and phosphate groups of NaAD.

containing bicarbonate at lower pH and noted similar issues with crystal damage. The active site of the S127A·NaAD structure showed that there is only one narrow pathway allowing the CO_2 dissolved in solvent to reach C5 of NaAD for the carboxylation reaction (**Figure 3-6**). The Mg^{2+} bound to NaAD, together with three highly conserved residues (Ala155, Gly156, and Arg159), form the entrance of the narrow pathway. It is plausible that Arg159, which is essential for LarB activity,[9] uses its flexible long arm to recruit and then hand CO_2 to the bound Mg^{2+}. A structure of S127A LarB soaked with NaAD and CS_2 (PDB: 8STD, **Table 3-A1**, which unfortunately did not reveal binding of CS_2) clearly illustrated the mobile Arg159 residue (**Figure 3-A7E**). The Arg159 sidechain also exhibited mobility, but to a lesser extent, in the LarB·NAD^+, LarB·AMP, LarB·Zn, and S127A LarB·NaAD structures (**Figure 3-A8**). Like in ribulose bisphosphate carboxylase/oxygenase (RubisCO), a carboxylase that catalyzes CO_2 fixation by plants, CO_2 binding to Mg^{2+} will increase the electrophilicity of the carbon atom to facilitate nucleophilic attack by C5 of NaAD. A structural model depicting the gas entrance site and with bound CO_2 was generated based on this mechanism (**Figure 3-6**). We envision that a high-resolution structure of LarB crystals grown under further optimized conditions would be crucial to definitively determine the CO_2 binding site.

Discussion

By analysis of the changes in abundances of metabolites for the WT enzyme and S127A variant of LarB (**Figure 3-3**), we have confirmed that loss of the Ser127 side chain reduces the rate of consumption for NaAD and the rate of production of P2CMN, AMP, and NaMN. Also in agreement with our expectation, the use of the S127A variant allowed us to easily identify the intermediacy of DaAD during the reaction, and this metabolite was also detected at much lower concentration when using the WT enzyme. Finally, we have for the first time structurally characterized the enzyme in the presence of its authentic substrate, NaAD. Three features of this structure have led us to reconsider the previously proposed mechanism for the enzyme.[9] First, the carboxylate of NaAD and the amide of NAD^+ were found to be similarly positioned, even though we had assumed the pyridinium rings would flip relative to each other so that the NaAD carboxylate could coordinate Mg^{2+} (**Figure 3-2**) whereas the NAD^+ amide points away from the metal ion. Such a flip had provided a reasonable explanation for why NAD^+ does not serve as a substrate. The S127A LarB·NaAD structure allows this explanation only if a ring flip occurs during catalysis, a situation that seems unlikely. Second, the conserved Glu180 residue, positioned distant from the Mg^{2+} ion, was proposed to function as the general base that abstracts the C5 proton during catalysis (**Figure 3-2**); however, the orientation of the pyridinium ring in the NaAD bound structure precludes such a general base role unless the ring flips during catalysis. Third, the S127A LarB·NaAD structure reveals no Cys221-pyridine ring linkage as found in the WT LarB·NAD^+ structure (PDB: 7MJ1); such adduct formation was proposed to be critical to enhancing the nucleophilicity of C5 to allow for carboxylation (**Figure 3-2**). The lack of this adduct in the S127A LarB·NaAD structure is not due to the amino acid substitution in the variant as we showed by UV-vis spectroscopic analysis that such a linkage occurs when the variant protein mixed with NAD^+.

On the basis of these new experimental results, we propose a revised reaction mechanism for LarB (**Figure 3-7**). NaAD binds in a manner much like NAD^+, but small differences in orientation are present. The conserved Glu180 enhances the nucleophilicity of Cys221, rather than abstracting the C5 proton that is too distant from this residue. Cys221 adds to C4 of the

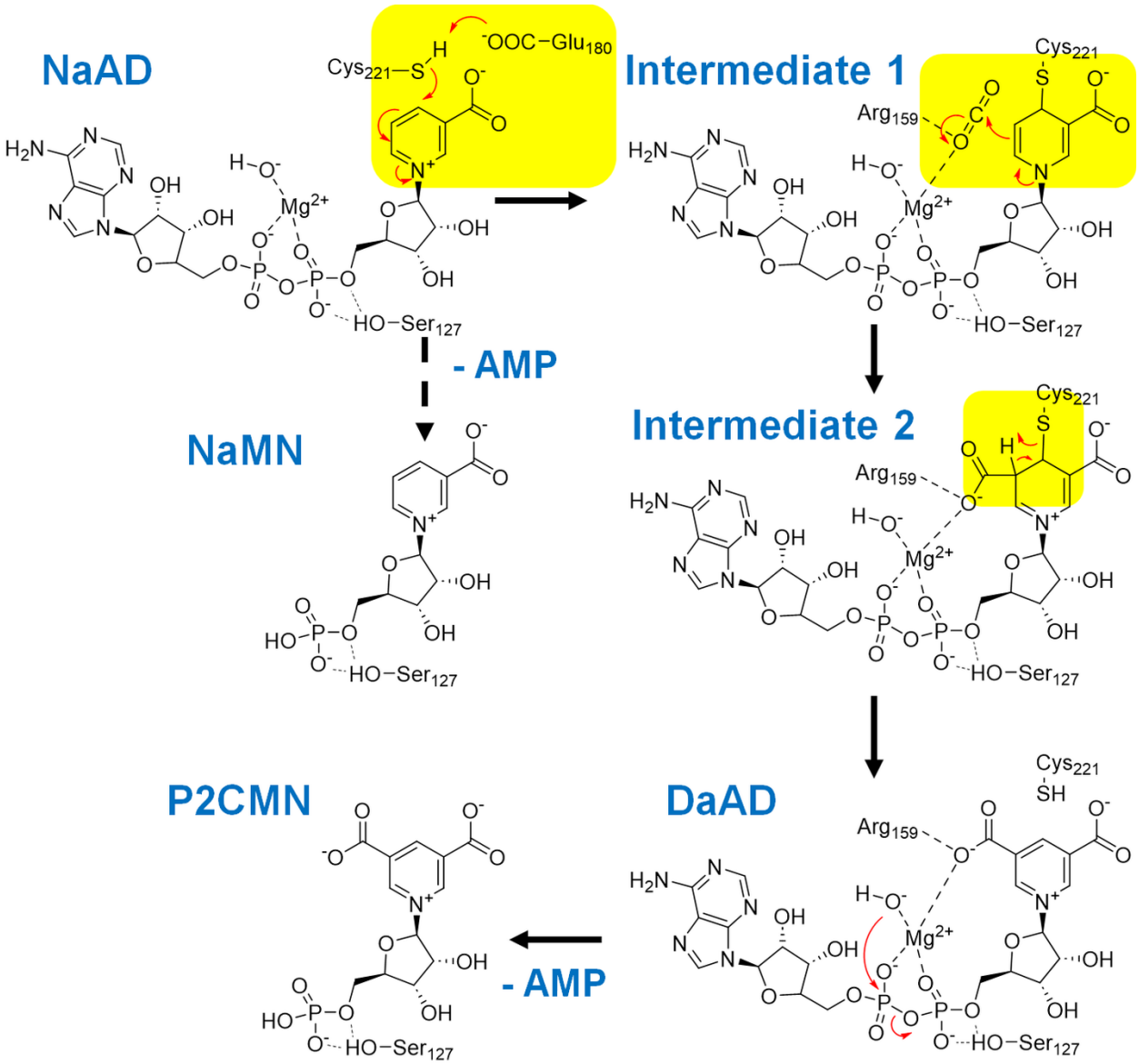

Figure 3-7. Revised proposal for the reaction mechanism of LarB. Differences from the previously proposed mechanism are highlighted in yellow. The carboxyl group of NaAD is oriented in the same manner as the amide of NAD^+, pointed away from the metal center. The conserved residue Glu180 is used to enhance the nucleophilicity of Cys221, which adds to C4 of the pyridinium ring. C5 of the ring attacks CO_2 that interacts with Mg^{2+} and Arg159. The carboxylation reaction completes via *syn* elimination with the expulsion of Cys221 coupled to its protonation. The productive and non-productive (dashed line) hydrolyses of DaAD and of NaAD are unchanged. Also illustrated is the important role of Ser127 in facilitating the hydrolysis reaction.

pyridinium ring to yield Intermediate 1, a low energy prominent intermediate in the case of NAD^+ but a high-energy undetectable intermediate in the case of NaAD. The stability of the adduct is suggested to be dictated by whether a neutral amide or an anionic carboxylate is present

in the bound nicotine derivative. The CO_2 substrate interacts with Mg^{2+} and Arg159, rather than near Glu180. The altered orientations of the pyridinium rings of NAD^+ and NaAD along with electronic effects of the carboxylate versus amide groups on the ring account for the carboxylation of NaAD but not NAD^+. Intermediate 2 decomposes without the assistance of Glu180 or any general base in the vicinity. We proposed this occurs by a *syn* elimination reaction with the coupled loss and protonation of Cys221. DaAD and NaAD are both hydrolyzed by a Mg^{2+}-coordinated water molecule. All reactions up to formation of DaAD are reversible, but LarB drives the overall reaction to completion by the hydrolysis reactions.

Acknowledgements

The authors thank Dr. Anthony Schilmiller of the MSU Mass Spectrometry Core of the Research Technology Support Facility for training and assistance with MS data interpretation. This work was supported by the National Science Foundation (CHE1807073) and the National Institutes of Health (GM128959) to R.P.H. and J.H.

REFERENCES

(1) Chatterjee, S., Gatreddi, S., Gupta, S., Nevarez, J. L., Rankin, J. A., Turmo, A., Hu, J., and Hausinger, R. P. (2022) Unveiling the mechanisms and biosynthesis of a novel nickel pincer enzyme, *Biochem. Soc. Trans. 50*, 1187-1196.

(2) Desguin, B., Zhang, T., Soumillion, P., Hols, P., Hu, J., and Hausinger, R. P. (2015) A tethered niacin-derived pincer complex with a nickel-carbon bond in lactate racemase, *Science 349*, 66-69.

(3) Desguin, B., Urdiain-Arraiza, J., Da Costa, M., Fellner, M., Hu, J., Hausinger, R. P., Desmet, T., Hols, P., and Soumillion, P. (2020) Uncovering a superfamily of nickel-dependent hydroxyacid racemases and epimerases, *Sci. Rep. 10*, 18123.

(4) Desguin, B., Soumillion, P., Hols, P., and Hausinger, R. P. (2016) Nickel-pincer cofactor biosynthesis involves LarB-catalyzed pyridinium carboxylation and LarE-dependent sacrificial sulfur insertion, *Proc. Natl. Acad. Sci. USA 113*, 5598-5603.

(5) Fellner, M., Desguin, B., Hausinger, R. P., and Hu, J. (2017) Structural insights into the catalytic mechanism of a sacrificial sulfur insertase of the N-type ATP pyrophosphatase family, LarE, *Proc. Natl. Acad. Sci. USA 114*, 9074-9079.

(6) Chatterjee, S., Parson, K. F., Ruotolo, B. T., McCracken, J., Hu, J., and Hausinger, R. P. (2022) Characterization of a (4Fe-4S)-dependent LarE sulfur insertase that facilitates nickel-pincer nucleotide cofactor biosynthesis in *Thermotoga maritima*, *J. Biol. Chem. 298*, 102131.

(7) Desguin, B., Fellner, M., Riant, O., Hu, J., Hausinger, R. P., Hols, P., and Soumillion, P. (2018) Biosynthesis of the nickel-pincer nucleotide cofactor of lactate racemase requires a CTP-dependent cyclometallase, *J. Biol. Chem. 293*, 12303-12317.

(8) Turmo, A., Hu, J., and Hausinger, R. P. (2022) Characterization of the nickel-inserting cyclometallase LarC from *Moorella thermoacetica* and identification of a cytidinylylated reaction intermediate, *Metallomics 14*, mfac014.

(9) Rankin, J. A., Chatterjee, S., Tariq, Z., Lagishetty, S., Desguin, B., Hu, J., and Hausinger, R. P. (2021) The LarB carboxylase/hydrolase forms a transient cysteinyl-pyridine intermediate during nickel-pincer nucleotide cofactor biosynthesis, *Proc. Natl. Acad. Sci. USA 118*, e2106202118.

(10) Fellner, M., Rankin, J. A., Desguin, B., Hu, J., and Hausinger, R. P. (2018) Analysis of the active site cysteine residue of the sacrificial sulfur insertase LarE from *Lactobacillus plantarum*, *Biochemistry 57*, 5513-5523.

(11) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Meth. Enzymol. 276*, 307-326.

(12) Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM, *Acta Crystallogr. D 67*, 271-281.

(13) Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?, *Acta Crystallogr. D 69*, 1204-1214.

(14) Evans, P. (2006) Scaling and assessment of data quality, *Acta Crystallogr. D 62*, 72-82.

(15) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J. Appl. Crystallogr. 40*, 658-674.

(16) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) *PHENIX*: A comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D 66*, 213-221.

(17) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of *Coot*, *Acta Crystallogr. D 66*, 486-501.

(18) The PyMOL Molecular Graphics System, Version 1.7.4 ed., Schrödinger, LCC.

(19) van Eys, J., and Kaplan, N. O. (1957) The addition of sulphydryl compounds to diphosphopyridine nucleotide analogues, *J. Biol. Chem. 228*, 305-314.

(20) Dittmer, D. C., and Kolyer, J. M. (1963) Addition compounds of thiols and 1-substituted nicotinamides, *J. Org. Chem. 28*, 1720-1722.

(21) Johnson, S. L., and Smith, K. W. (1977) Nucleophile and borate reactivity with nicotinamide adenine dinucleotide and its analogues, *J. Org. Chem. 42*, 2580-2589.

(22) Patel, S. M., Smith, T. G., Morton, M., Stiers, K. M., Seravalli, J., Mayclin, S. J., Edwards, T. E., Tanner, J. J., and Becker, D. F. (2020) Cautionary tale of using tris(alkyl)phosphine reducing agents with NAD^+-dependent enzymes, *Biochemistry 59*, 3285-3289.

APPENDIX

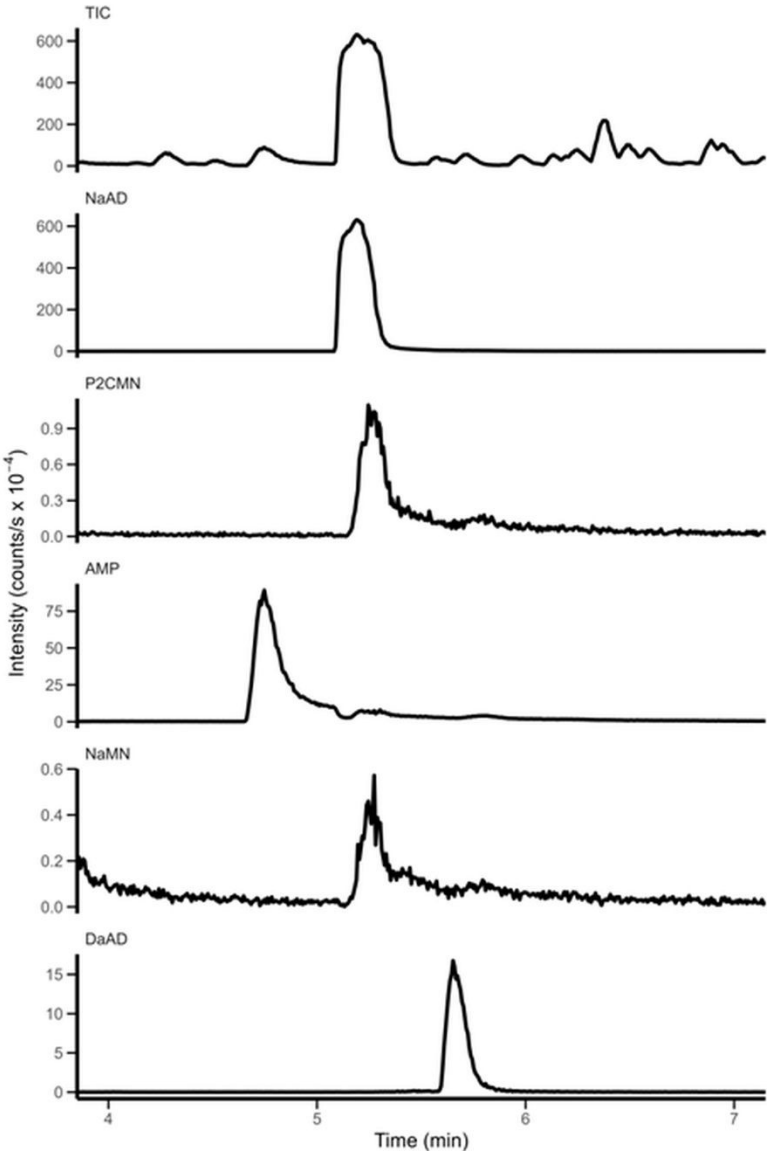

Figure 3-A1. Total ion chromatogram (TIC) and extracted ion chromatograms (EIC) for the NaAD, P2CMN, AMP, NaMN, and DaAD ions (*m*/*z* = 663.08, 378.02, 346.05, 334.03, and 707.07, respectively) after incubating S127A LarB (150 µM) with 0.2 mM NaAD, 50 mM $MgCl_2$, and 50 mM $NaHCO_3$ in 100 mM Tris-HCl, pH 7.2, for 300 s. Analogous TIC and EIC data were obtained by negative ionization mode LC-ESI-MS for S127A and WT LarB samples at all of the time points shown in **Figure 3-1**.

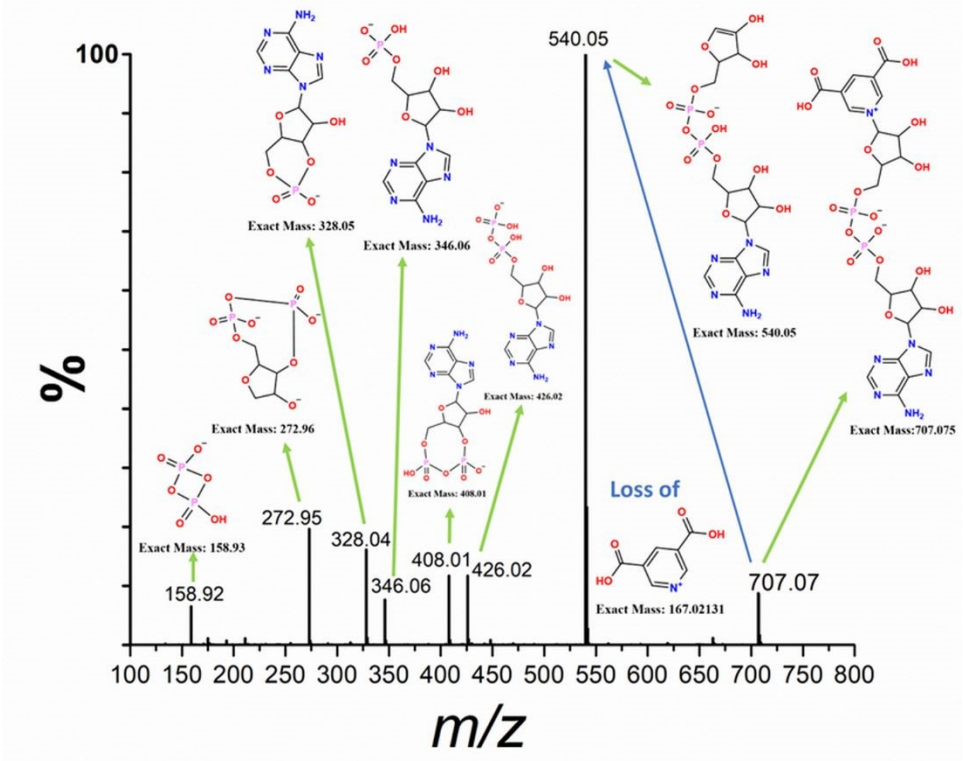

Figure 3-A2. MS/MS fragmentation analysis of the novel species generated by S127A LarB when incubated for 1 min in the presence of NaAD, $MgCl_2$, and bicarbonate.

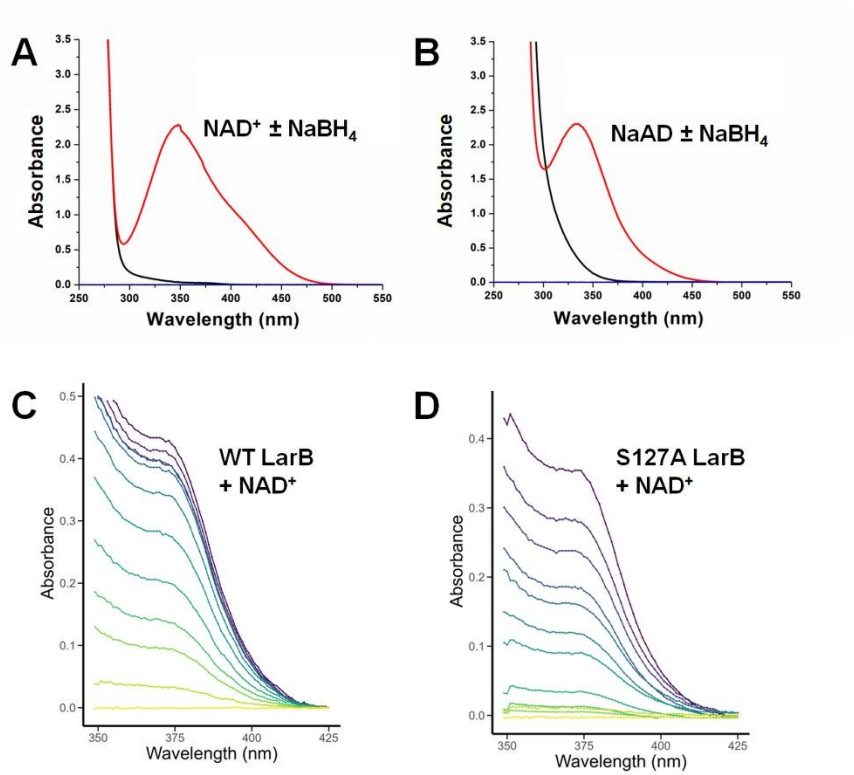

Figure 3-A3. UV-vis spectroscopic evidence for a covalent linkage of LarB Cys221 to NAD^+. Control experiments show 1 mM (A) NAD^+ and (B) NaAD (black lines) treated with 0.4 mM $NaBH_4$ (red lines) in 100 mM Tris, pH 7.5, containing 150 mM NaCl. Blue lines depict the buffer baseline. Using argon-purged solutions, (C) WT LarB (358 µM) and (D) S127A LarB (430 µM) were titrated with NAD^+ (0, 0.9, 1.7, 2.3, 2.9, 3.3, 3.8, 4.1, 4.4, 4.7, 5.0, and 5.5 mM, bottom to top) while monitoring changes in the spectra, which were corrected for dilution. Analogous studies using NaAD, rather than NAD^+, did not result in any discernable changes in the spectra of WT or S127A LarB. The buffer was 100 mM Tris, pH 7.5, containing 150 mM NaCl.

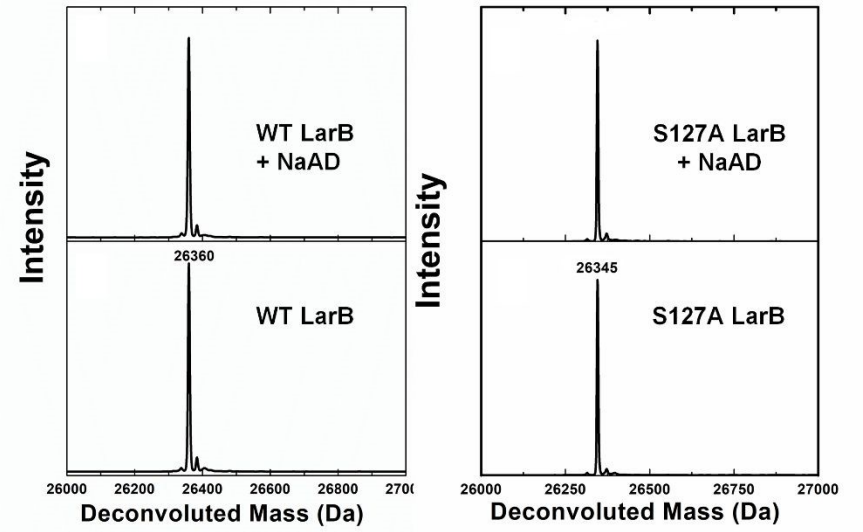

Figure 3-A4. ESI-MS data derived from 10 µM samples of WT LarB (left) or S127A LarB (right) in 100 mM Tris, pH 7.5, containing 150 mM NaCl (bottom) or the same Ar-purged samples treated with 200 µM NaAD immediately prior to analysis (top).

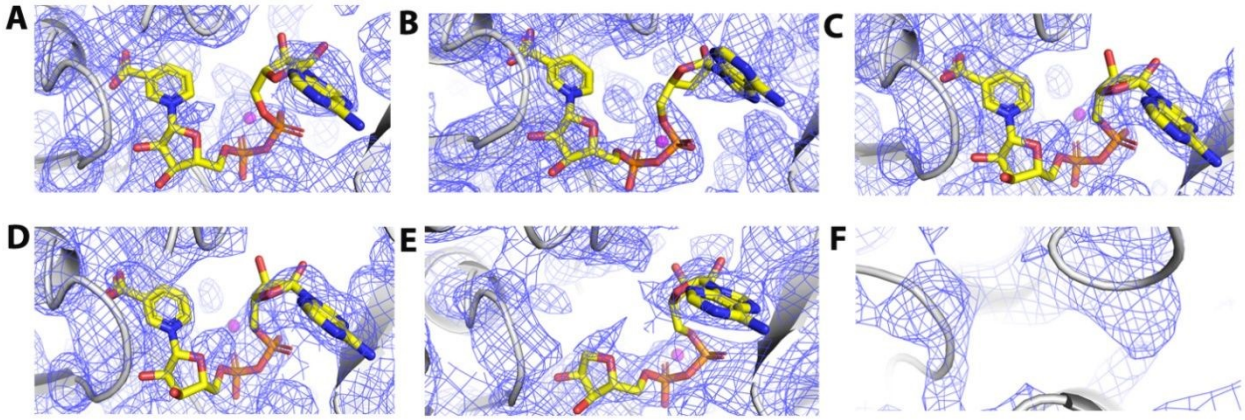

Figure 3-A5. Simulated annealing composite omit maps (2Fo-Fc, σ=1) of the LarB active sites of chains A-F. NaAD and Mg^{2+} are shown as yellow sticks and magenta spheres, respectively.

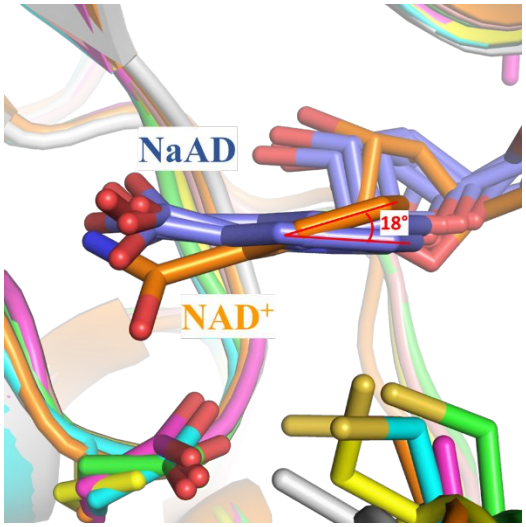

Figure 3-A6. Comparison of the orientations of the pyridinium ring in NAD^+ and NaAD bound structures. Chain A-F of the S127A LarB·NaAD structure and chain C of WT LarB·NAD^+ (PDB: 7MJ1) are structurally aligned, revealing that the NaAD molecules are consistently and distinctly tilted from NAD^+ with an approximately 18-degree angle.

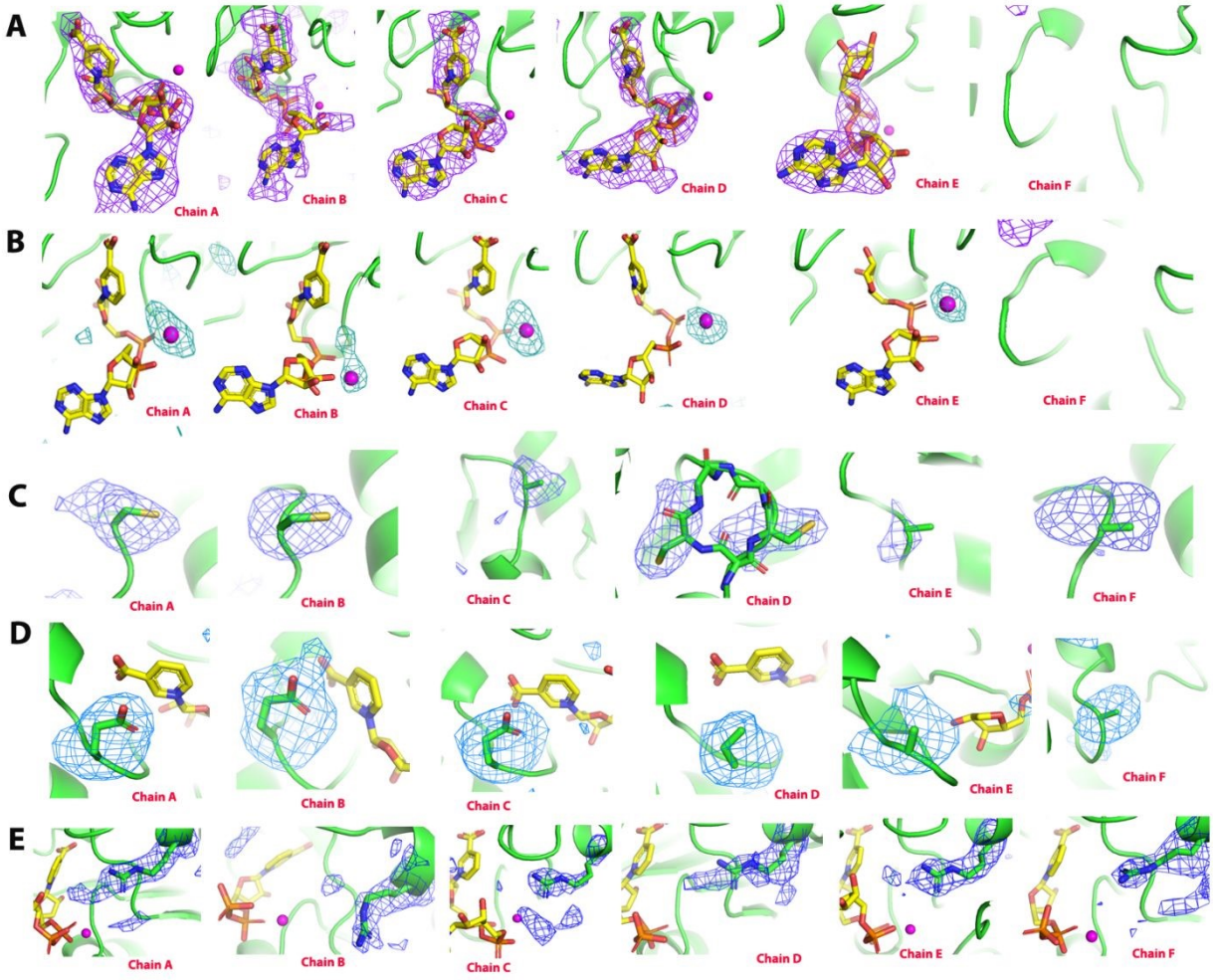

Figure 3-A7. Polder maps of the ligands and key residues in chains A-F (σ=3). (A) NaAD. (B) Mg^{2+}. (C) Cys221. (D) Glu180. (E) Arg159 in the CS_2-soaked structure.

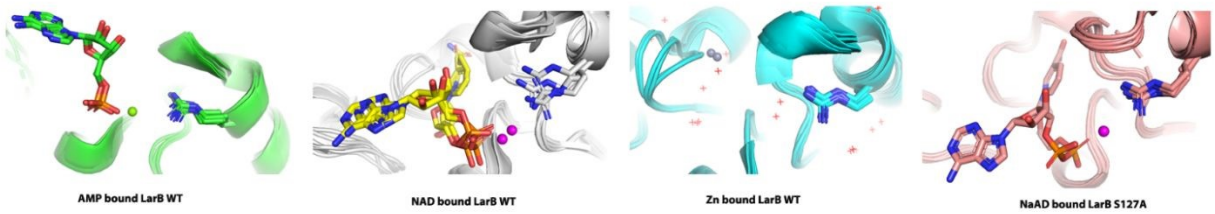

Figure 3-A8. Comparison of the mobility of Arg159 in various LarB structures.

Table 3-A1. Crystallization statistics for S127A LarB·NaAD complex and S127A LarB·NaAD complex soaked with CS_2

LarB_S127A crystals	**NaAD bound**	**NaAD bound & CS_2 soaked**
Data Collection		
Beamline	LS-CAT 21-ID-D	LS-CAT 21-ID-D
Wavelength (Å)	0.979	1.549
Space group	$P42_12$	$P42_12$
Unit cell a, b, c (Å) α, β, γ (°)	120.9, 120.9, 212.6 90, 90, 90	119.4, 119.4, 212.1 90, 90, 90
[a]Resolution (Å)	85.50 – 3.10 (3.29 – 3.10)	84.03 – 2.65 (2.75 – 2.65)
Unique reflections	29281 (4636)	44586 (4607)
[a]Redundancy	9.7 (9.9)	6.7 (6.9)
[a]Completeness (%)	99.8 (100)	99.7 (100.0)
[a] I/σI	8.2 (2.3)	6.5 (1.1)
a,bRmerge	0.180 (1.113)	0.196 (1.754)
a,cR_{pim}	0.088 (0.537)	0.118 (1.040)
dCC1/2	0.967 (0.767)	0.983 (0.380)
Refinement		
Protein atoms	8023	8401
Ligand molecules	5 DND	6 DND
Metal	7 Mg	8 Mg
eRwork/Rfree	0.236/0.286	0.242/0.273
B-factors (Å^2)	112.8	80.8
Protein	112.2	80.3
Ligand	136.7	99.1
Metal	124.1	65.8
R.m.s. deviation in bond lengths (Å)	0.010	0.018
R.m.s. deviation in bond angles (°)	1.22	0.74
Ramachandran plot (%) favored	99.22	98.67
Ramachandran plot (%) outliers	0.78	1.33
Rotamer outliers (%)	6.3	4.35
PDB ID	8SOQ	8STD

[a]Highest resolution shell is shown in parentheses. [b]$R_{merge} = \sum_{hkl} \sum_j |I_j(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_j Ij(hkl)$, where I is the intensity of reflection.
[c]$R_{pim} = \sum_{hkl} [1/(N-1)]^{1/2} \sum_j |I_j(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_j I_j(hkl)$, where N is the redundancy of the dataset. [d]$CC_{1/2}$ is the correlation coefficient of the half datasets.
[e]$R_{work} = \sum_{hkl} | |F_{obs}| - |F_{calc}| | / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} is the observed and the calculated structure factor, respectively. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total) omitted in model refinement.

CHAPTER 4

Exploration of potential IscS utilization by LarE in *Latilactobacillus sakei*

Introduction

LarE is a sulfur insertase involved in the biosynthesis of the nickel-pincer nucleotide (NPN) cofactor, which is crucial for the activity of lactic acid racemase (LarA) in many microorganisms including *Lactiplantibacillus plantarum*.[1] This enzyme is responsible for thiolating the carboxylic acids of pyridinium-3,5-dicarboxylic acid mononucleotide (P2CMN), a nicotinic acid-derived substrate, to form pyridinium-3,5-dithiocarboxylic acid mononucleotide (P2TMN).[2] Structurally, LarE is classified within the N-type ATP pyrophosphatase family, which is characterized by its involvement in adenylylation reactions that activate diverse substrates for subsequently reactivity, including for the incorporation of sulfur in some cases.[3]

Two mechanisms for LarE activity have been observed, one where LarE operates through a sacrificial sulfur transfer mechanism and one where LarE relies on a bound iron-sulfur cluster for sulfur transfer.[3–5] In LarE from *L. plantarum* ($LarE_{Lp}$), the enzyme donates a sulfur atom from its single cysteine residue (Cys176), resulting in the formation of a dehydroalanine (Dha) side chain. This process effectively sacrifices the enzyme's cysteinyl residue for the sulfur transfer.[3] In contrast to the *L. plantarum* mechanism, *Thermotoga maritima* ($LarE_{Tm}$) possesses a [4Fe-4S] cluster that is bound by three conserved cysteine residues with the non-ligated iron atom capable of coordinating sulfide and forming a [4Fe-5S] species, which then transfers the non-core sulfide to the activated substrate.[4] This mechanism reflects a more sustainable method of sulfur transfer compared to the sacrificial approach.

Analysis of bacterial genomes encoding LarE enzymes revealed an intriguing subset of *Latilactobacillus* species with a conserved *lar* operon immediately followed by an *iscS* gene (**Figure 4-1**), which encodes cysteine desulfurase (IscS). IscS is a pyridoxal phosphate (PLP)-dependent enzyme that transfers sulfur from L-cysteine (L-Cys) to other enzymes, such as those which synthesize iron-sulfur clusters.[6] Remarkably, none of the five *larE* genes in this subset

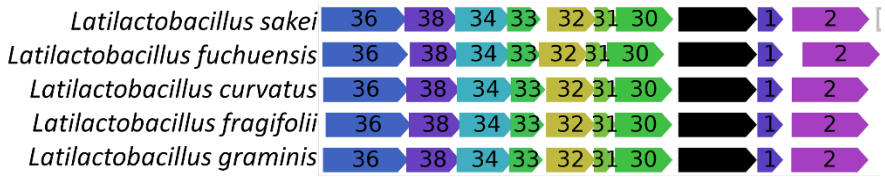

Figure 4-1. Genomic context of a group of Latilactobacilli that harbor a complete lar operon followed by iscS. The genes are larA (36), larB (38), larC1 (34), larC2 (33), larD (32), larE (30), iscS (black), DUF1831 (1), mnmA encoding a tRNA-specific 2-thiouridylase (2), and a gene encoding a hypothetical protein (31).

encode the three conserved cysteine residues necessary for iron-sulfur cluster binding as seen in $LarE_{Tm}$.[4] Given the known role of IscS in transferring sulfur to the iron-sulfur cluster of $LarE_{Tm}$ and other proteins throughout nature,[4,7] the proximity of *iscS* to *larE* in these genomes suggests two potential novel roles for IscS in which (i) it could assist in the regeneration of the inactive Dha-containing form of LarE or (ii) it could bypass the Dha-forming reaction by modifying a LarE Cys residue in these *Latilactobacillus* species to form a persulfide of that donates the sulfane sulfur to the activated substrate. These considerations challenge the traditional view of LarE as a costly single-turnover enzyme, proposing instead that its activity may be maintained through interaction with IscS.

In this work the proposed relationship between LarE and IscS in these *Latilactobacilli* was explored through coupled enzyme assays containing LarE from *L. sakei* ($LarE_{Ls}$) and IscS from *Escherichia coli* ($IscS_{Ec}$). The reactants, products and protein modifications from these reactions were monitored using HPLC-MS.

Experimental Procedures

Gene Cloning and Protein Expression

Genes encoding the target proteins were codon optimized for expression in *E. coli* and synthesized by Integrated DNA Technologies (IDT). Specifically, the *larE* and *iscS* genes from *L. sakei* were cloned into the pET23b(+) expression vector. The constructs included an N-terminal His_6-tag and a C-terminal Strep-tag to facilitate purification. Recombinant plasmids were transformed into *E. coli* BL21(DE3) competent cells via heat shock. Positive colonies were selected on lysogeny broth (LB) agar plates containing 100 μg/mL of ampicillin. Single colonies were inoculated into LB or terrific broth (TB) medium with 100 μg/mL of ampicillin and grown overnight at 37 °C with shaking at 200 rpm. Gene expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM when cultures reached an optical density of 600 nm (OD_{600}) of 0.6–0.8, and incubation was continued for 4 h at 37 °C or overnight at 18 °C.

Attempts were unsuccessful at producing $IscS_{Ls}$ in *E. coli*. Despite optimizing expression conditions—including varying induction temperatures (18 °C to 37 °C), IPTG concentrations (0.1–1 mM), and post-induction incubation times (4–20 h)—no soluble $IscS_{Ls}$ protein was detected. Expression of this *iscS* fused to the gene encoding the maltose binding protein (MBP) was examined as a means to enhance solubility; however, this strategy also did not yield

detectable levels of soluble $IscS_{Ec}$. Alternative approaches, such as co-expression with $larE_{Ls}$ and expression in different *E. coli* strains (C41, C43), were also explored without success. As an alternative to using the cognate $IscS_{Ls}$, I chose to investigate the potential use of $IscS_{Ec}$ for studies with $LarE_{Ls}$. Using a vector containing a gene encoding $IscS_{Ec}$, the protein was produced as previously described.[4]

Protein Purification

Harvested cells containing $LarE_{Ls}$, the S127A variant of $LarB_{Lp}$,[8] $IscS_{Ec}$,[4] LarC from *L. plantarum* ($LarC_{Lp}$),[9] and LarA from *Thermoanaerobacterium thermosaccharolyticum* ($LarA_{Tt}$)[10] were resuspended in lysis buffer containing 100 mM HEPES (pH 7.4) and 150 mM NaCl. Protease inhibitors, including 1 mM phenylmethylsulfonyl fluoride (PMSF), were added to prevent protein degradation. Cells were lysed by sonication on ice using a pulse program of 2 min on, 2 min off, for 3 cycles. The lysates were kept cold throughout the procedure to prevent thermal denaturation. The lysates were clarified by centrifugation at 15,000 × *g* for 30 min at 4 °C.

For $LarE_{Lp}$ purification, the supernatant was applied to a Strep-Tactin® XT 4Flow® high-capacity resin (IBA Lifesciences) pre-equilibrated with lysis buffer. The column was washed with the same lysis buffer to remove non-specifically bound proteins, and proteins were eluted with lysis buffer containing 50 mM biotin. Eluted fractions were collected for further analysis. Protein samples were concentrated using Amicon Ultra centrifugal filters with a 10 kDa molecular weight cut-off (Millipore) following the manufacturer's instructions. Protein concentrations were determined by first measuring the absorbance at 280 nm by means of a nano-drop spectrophotometer, then calculated using a molar extinction coefficient for LarE of 28,420 $M^{-1}cm^{-1}$. Samples were stored at 4 °C for short-term use and at –80 °C for long-term storage.

For the S127A variant of LarB, a variant that generates greater levels of P2CMN,[8] the same method was followed, except Strep-Tactin resin was used instead of Strep-Tactin XT. Due to the stronger binding to Strep-Tactin XT, which can prevent efficient elution, the regular Strep-Tactin resin was chosen for optimal elution. The protein concentration was determined using a molar extinction coefficient of 14,565 $M^{-1}cm^{-1}$.

$IscS_{Ec}$ was purified using Ni-NTA affinity chromatography with a similar HEPES buffer. The lysis buffer contained 5 mM imidazole for wash steps and 500 mM imidazole for elution.

Following elution, $IscS_{Ec}$ was buffer exchanged into imidazole-free HEPES buffer using spin concentrators to prepare the protein for downstream applications. The concentration of purified $IscS_{Ec}$ was determined using its molar extinction coefficient of 41,495 $M^{-1}cm^{-1}$. $LarC_{Lp}$ and the apoprotein of $LarA_{Tt}$ were produced and purified as previously described.[9,10]

LarE Activity Assays

The activity of LarE was assessed indirectly using a coupled enzyme assay by measuring the lactate racemase activity of $LarA_{Tt}$. LarB was used to generate the LarE substrate P2CMN from NaAD and $NaHCO_2$, then LarE catalyzed the conversion of P2CMN into P2TMN. P2TMN was subsequently converted by $LarC_{Lp}$ into the NPN cofactor that was incorporated into $LarA_{Tt}$. The lactate racemase activity of $LarA_{Tt}$ was measured using an L-lactate assay kit from Neogen, which includes L-lactate dehydrogenase, NAD^+, glutamate-pyruvate transaminase, and D-glutamate, allowing for the indirect assessment of LarE activity through the L-lactate detection assay when D-lactate was provided, as described previously.[11]

The reaction mixture for measuring the above activity contained 33 mM Tris-HCl buffer at pH 7.4, 0.7 mM L-Cys, 0.7 mM dithiothreitol (DTT), 0.7 mM adenosine triphosphate (ATP), 3.3 mM magnesium chloride ($MgCl_2$), 0.8 mM sodium nicotinate adenine dinucleotide (NaAD), 4.1 mM sodium bicarbonate ($NaHCO_3$), 0.5 mM cytidine triphosphate (CTP), and 0.9 mM nickel chloride ($NiCl_2$). Enzymes included in the mixture were 29 µM LarB S127A variant, 104 µM $LarE_{Ls}$, 70 µM LarC, and 11 µM $LarA_{Tt}$. Additionally, 19 µM $IscS_{Ec}$ was included in one set of reactions. Two conditions were tested: reactions with and without $IscS_{Ec}$. Buffer blanks were used as negative controls to account for any background signal in the enzyme assay.

All components except for $IscS_{Ec}$ were combined in a final volume of 410 µL and gently mixed then divided into two equal 200-µL portions. Equivalent amounts of either $IscS_{Ec}$ or buffer were added, and the reactions were split into two pairs with either D- or L-lactate added (10 mM) and incubated at 37 °C for 1 h. Following incubation, the mixtures were heated at 80 °C for 10 min to precipitate the proteins then pelleted and the supernatant was added to the L-lactate assay kit according to the manufacturer's instructions to measure the amount of L-lactate generated by each sample. The assay quantifies L-lactate levels through enzymatic reactions that result in the oxidation of NADH to NAD^+, leading to a decrease in absorbance at 340 nm. Absorbance readings were recorded, and relative LarA activity was determined using a control reaction.

Mass Spectrometry

Separate enzyme assays were prepared for the mass spectrometry analysis. First a reaction to generate P2CMN, the substrate of LarE was prepared. The reaction mixture containing LarB S127A at a final concentration of 0.07 mM, NaAD at 39 mM, $NaHCO_2$ at 67 mM, and $MgCl_2$ at 13 mM. The reaction was incubated overnight under optimal conditions to allow the enzymatic reaction to proceed. Following incubation, the protein was precipitated by heating the reaction mixture at 80 °C for 10 min. The precipitated protein was then pelleted by centrifugation at 13,000 rpm for 15 min using a benchtop centrifuge. The resulting supernatant, containing the reaction product P2CMN, unreacted NaAD and AMP as a byproduct of the reaction. This mixture was used as P2CMN in the $LarE_{Lp}$ assay.

The reaction assay aimed to evaluate the activity of LarE under multiple conditions by systematically varying the presence of specific reactants. The reaction mixtures included $LarE_{Lp}$, P2CMN, ATP, and $MgCl_2$ as core components, with additional conditions incorporating L-Cys DTT, BME, and IscS in specific combinations. Final concentrations of 0.05 mM $LarE_{Lp}$, 2 mM ATP, 5 mM $MgCl_2$, 0.014 mM IscS, and 2 mM of L-Cys, DTT, or BME, were used depending on the specific condition. P2CMN was included at a nominal concentration of 3.9 mM, which represents the theoretical concentration of P2CMN if 100% of the NaAD from the preceding reaction had been converted. However, as the conversion was incomplete, the actual concentration of P2CMN is lower than 3.9 mM but remains unknown

Proteins in the $LarE_{Ls}$ reactions were precipitated by heating the reaction mixtures at 95 °C for 5 min. The samples were then centrifuged to pellet the precipitated proteins, and the supernatants were transferred to new tubes and evaporated under vacuum. After evaporation, the samples were dissolved in buffer containing 8 mM DMHA (N,N-dimethylhexylamine) and 2.8 mM acetic acid.

Small molecules in the reaction samples were analyzed using a Waters G2-XS Q-TOF mass spectrometer interfaced with a Waters ACQUITY UPLC system. Ten µL samples were injected onto a Waters ACQUITY UPLC BEH-C18 column (2.1 × 100 mm) maintained at 40 °C. The reaction components were separated by ion-pairing liquid chromatography using a binary gradient. Mobile phase A consisted of 8 mM DMHA and 2.8 mM acetic acid in water, and mobile phase B was methanol. The initial conditions were 100% mobile phase A, held for 1 min, followed by a linear ramp to 99% mobile phase B at 7 min. This condition was held until 8 min,

then returned to 100% mobile phase A at 8.01 min and held until 10 min. The flow rate was 0.3 mL/min. Mass spectra were obtained using electrospray ionization (ESI) in negative ion mode with a capillary voltage of 2.0 kV, a source temperature of 100 °C, a cone voltage of 35 V, a desolvation temperature of 350 °C, a desolvation gas flow of 600 L/h, and a cone gas flow of 50 L/h. Data were acquired using a data-independent MS^E method (scans with fast switching between no collision energy and using a collision energy ramp of 20–80 V) across an *m*/*z* range of 50–1500. The total ion chromatogram (TIC) and extracted ion chromatograms (EIC) for NaAD-derived metabolites were monitored. Lockmass correction was performed in MassLynx software (Waters Corp.) using leucine enkephalin as the reference compound.

For protein mass spectrometry, samples were diluted to 1-10 μM protein. Mass spectrometric analysis was performed using a Waters G2-XS Q-TOF mass spectrometer. Ten μL protein samples were injected onto a Thermo Beta Basic CN column (1.0 × 10 mm) for desalting. A gradient of water with 0.1% formic acid (solvent A) and acetonitrile (solvent B) was applied at a flow rate of 0.1 mL/min. The initial conditions were 98% solvent A and 2% solvent B, held until 5 min with the flow diverted to waste for the first 3 min. A ramp to 75% solvent B was performed over 5 min, held at 75% solvent B until 12 min, then returned to 2% solvent B at 12.01 min and held until 15 min.

Mass spectra were acquired using ESI in positive ion mode with a source temperature of 100 °C, a cone voltage of 35 V, a desolvation temperature of 350 °C, a desolvation gas flow of 600 L/h, a cone gas flow of 50 L/h, and a capillary voltage of 3.0 kV. Data were collected using a 0.5 s time-of-flight (TOF) MS scan across an *m*/*z* range of 200–2000. The spectra were deconvoluted using the MaxEnt1 algorithm in MassLynx software to determine the molecular weight of the intact proteins. No protein digestion was performed prior to analysis, allowing for the detection of any post-translational modifications such as dehydroalanine formation.

Genomic Context

Complete *lar* operons were identified by using STRING (Search Tool for the Retrieval of Interacting Genes/Proteins), a comprehensive database that compiles known and predicted protein–protein interactions.[12] The protein sequence for $LarE_{Lp}$ (UniProtKB F9UST4) was used to search the STRING database. The Gene Neighborhood functionality was used to view the genomic context of various *lar* operons in different organisms. The *lar* operon in *L. sakei* was identified and was shown to precede a gene encoding a putative cysteine desulfurase

(aminotransferase, class V). The sequence of LarE$_{Ls}$ was compared to that of LarE$_{Tm}$, known to contain an [4Fe-4S] cluster bound by 3 conserved cysteine residues. LarE$_{Ls}$ lacks the 3 conserved cysteine residues required for Fe-S cluster binding, rather, the amino acid sequence of LarE$_{Ls}$ more closely resembles that of LarE$_{Lp}$ (**Figure A4-1**). The depiction of the genomic context of the *lar* operon in the *Latilactobacilli* was created using webFlaGs.[13]

Results

Gene expression and protein purification

The genes encoding LarE and IscS from *L. sakei* were cloned in *E. coli* BL21(DE3). LarE$_{Ls}$ was successfully inserted into the pET23b(+) expression vector. Transformation into *E. coli* BL21(DE3) cells resulted in high levels of protein expression under standard induction conditions. Whole-plasmid sequencing confirmed the correct insertion and integrity of the LarE$_{Ls}$ gene within the vector. Analysis of cell lysates via SDS-PAGE displayed prominent bands corresponding to the expected molecular weight of LarE$_{Ls}$, indicating successful expression. The protein was subsequently purified using Strep-tactin XT resin, yielding elution fractions containing a single band as visualized on SDS-PAGE corresponding to the correct molecular weight.

In contrast, although *iscS*$_{Ls}$ was successfully cloned into the pET23b(+) vector—as confirmed by whole plasmid sequencing—no expression of *iscS*$_{Ls}$ was detected in *E. coli* BL21(DE3) cells. To address this issue, various culture conditions were systematically tested, including the use of different media such as LB, TB, and TB auto-induction medium. Growth temperatures of 18°C, 25°C, and 37°C were explored, along with induction at these same temperatures. Initial screenings were conducted in 1 mL test cultures. Despite these efforts, SDS-PAGE analyses revealed no detectable levels of IscS$_{Ls}$ under any of the tested conditions. To enhance the expression of *iscS*$_{Ls}$ and improve the solubility of the encoded protein, the gene was subcloned into the pMCSG9 vector, introducing an N-terminal MBP fusion tag. Whole-plasmid sequencing verified the correct insertion of *iscS*$_{Ls}$ into pMCSG9; however, like previous attempts, no production of the MBP-IscS$_{Ls}$ fusion protein was observed under the same range of culture conditions. For these reasons, I used IscS$_{Ec}$ rather than IscS$_{Ls}$ for further studies.

Enzyme Assays

To investigate the potential utilization of IscS$_{Ec}$ by LarE$_{Ls}$ during catalysis, enzyme assays were conducted using purified apoprotein of LarA$_{Tt}$ along with LarB S1271A, LarC, and LarE$_{Ls}$ both

with and without the addition of $IscS_{Ec}$. The assays aimed to determine whether $IscS_{Ec}$ could enhance or otherwise affect the activity of $LarE_{Ls}$ by monitoring lactate racemase activity as a proxy.

The results from these enzyme assays were variable and often conflicting. In most experiments, the samples containing $IscS_{Ec}$ exhibited enzymatic activities that were less than or equal to those observed in the samples without $IscS_{Ec}$. This finding was unexpected, as IscS is known to facilitate sulfur transfer in various enzymatic reactions and was anticipated to potentially enhance $LarE_{Ls}$ activity.

The inconsistent results made it challenging to draw definitive conclusions about the effect of $IscS_{Ec}$ on $LarE_{Ls}$ activity. Possible reasons for the variability include differences in protein preparations, assay conditions, or interactions between the proteins that were not fully understood. Due to these inconsistencies, a different, more sensitive approach using mass spectrometry was employed.

Representative data from these enzyme assays are provided in the Appendix, illustrating the variability observed.

Mass Spectrometry Analysis of $LarE_{Ls}$ Reactants and Products

To gain a more detailed understanding of the enzymatic activity of $LarE_{Ls}$ and the influence of $IscS_{Ec}$, HPLC-MS was used to quantify the levels of reaction products and substrates under various conditions. The primary focus was on measuring the concentrations of P2TMN, P2CMN, NaAD, ATP, ADP, and AMP across different sample groups.

The MS analysis revealed that the highest levels of P2TMN, the primary product of $LarE_{Ls}$ activity, were observed in samples without the addition of dithiothreitol (DTT) or L-Cys, and no stimulation was observed when the assay was supplemented with $IscS_{Ec}$ (**Figure 4-A2**). Surprisingly, P2CMN levels were also elevated in this same group (**Figure 4-A3**), with all samples shown to contain substantial amounts of this LarB-generated metabolite. NaAD levels remained consistent across all samples (**Figure 4-A4**), reflecting that its concentration was unaffected by the presence or absence of these additives, as it was not utilized as a substrate by any of the enzymes in the reaction and was expected to remain relatively unchanged throughout ATP concentrations were found to be highest in the samples lacking L-Cys, DTT, or $IscS_{Ec}$ (**Figure 4-A5**). The levels decreased when either L-Cys and/or DTT were present. Notably, ATP levels were nearly completely depleted in samples containing $IscS_{Ec}$. AMP, an adenylylation

product and a product of NaAD hydrolysis, showed minimal variation across all conditions, remaining relatively constant (**Figure 4-A6**). ADP, presumably derived from ATP hydrolysis was present in all samples, with less present in samples containing $IscS_{Ec}$ (**Figure 4-A7**). These findings present a complex picture that deviates from the initial expectations. It was anticipated that as P2TMN levels increased, ATP levels would decrease due to substrate adenylylation by $LarE_{Ls}$ during each sulfur transfer reaction. Additionally, it was expected that P2CMN levels would decrease as P2TMN levels rose, reflecting the conversion of P2CMN to P2TMN. However, the observed data did not support these hypotheses. Instead, both P2TMN and P2CMN levels were highest in the same conditions, and ATP depletion did not correlate straightforwardly with P2TMN production. Histograms for each analyte are shown in the appendix (**Figures A4-2** through **A4-7**).

Protein Mass Spectrometry of $LarE_{Ls}$

To gain a better understanding of the potential impact of $IscS_{Ec}$ on $LarE_{Ls}$ catalysis, the structural modifications of $LarE_{Ls}$ under different assay conditions were examined by protein mass spectrometry. The $LarE_{Ls}$ protein spectra exhibited a peak at 34,681 Da (**Figure 4-A8**), corresponding to the full-length peptide containing a 6×His tag and a Strep tag, thus confirming the integrity of the expressed protein.

In the presence of P2CMN, Mg^{2+}, and ATP, an additional peak at 34,647 Da was detected (**Figure 4-A8**). This peak corresponds to a dehydroalanine sidechain modification caused by the loss of a sulfur and two hydrogen atoms during the sulfur insertion reaction catalyzed by $LarE_{Ls}$. Furthermore, a P2CMN-$LarE_{Ls}$ adduct was identified at 35,042 Da, indicating the formation of a covalent complex between P2CMN and $LarE_{Ls}$.

When L-Cys was added to the reaction, the native $LarE_{Ls}$ peak diminished, and an intense peak at 34,800 Da emerged (**Figure 4-A9**), corresponding to a cysteinyl disulfide adduct of $LarE_{Ls}$. A mixed disulfide was not observed when DTT substituted for the L-Cys (**Figure 4-A10**). Upon the addition of DTT to sample also containing L-Cys, the cysteinyl adduct peak disappeared completely, and the mass spectrum resembled that of native $LarE_{Ls}$, indicating that DTT effectively reduced the cysteinyl modification (**Figure 4-A11**). In contrast, when β-mercaptoethanol (BME) was used instead of DTT, the cysteinyl adduct remained (**Figure 4-A16**), suggesting that BME is not a sufficiently strong reductant to disrupt the L-Cys-$LarE_{Ls}$

interaction. The addition of $IscS_{Ec}$ to these samples had no effect on the species observed (**Figures A4-12** through **A4-15**).

Notably, a mass of 34,713 Da, which would correspond to a $LarE_{Ls}$ persulfide species resulting from sulfur transfer from $IscS_{Ec}$ to $LarE_{Ls}$, was not observed in any of the samples. This absence suggests that either the persulfide species is transient and below the detection limit of the current MS setup, or that the expected sulfur transfer does not occur under the tested conditions.

Discussion

The primary objective of this study was to investigate the potential of IscS in stimulating the turnover of the sulfur-insertase, $LarE_{Ls}$. Genomic analysis revealed a unique arrangement wherein the *lar* operon is followed by an *iscS* gene, suggesting a potential collaborative mechanism between LarE and IscS that does not rely on the traditional iron-sulfur [4Fe-4S] cluster-dependent regeneration observed in other LarE homologs, such as those from *Thermotoga maritima*.[4] This proposed mechanism aimed to challenge the existing paradigm of LarE as a single-turnover enzyme by introducing a system where IscS facilitates Dha-containing LarE regeneration to the native enzyme or using a cysteinyl persulfide intermediate rather than involving an Fe-S cluster.

$LarE_{Ls}$ was expressed in *E. coli* BL21(DE3) using the pET23b(+) vector and purified to homogeneity via Strep-tactin XT affinity chromatography. The purity and integrity of $LarE_{Ls}$ were confirmed through SDS-PAGE and mass spectrometry, which identified the expected molecular weight and specific post-translational modifications, including Dha formation and the P2CMN-LarE adduct. These modifications are consistent with LarE's known catalytic activity, affirming the functionality of the purified enzyme.

Despite multiple attempts to express $IscS_{Ls}$ in *E. coli*, including various expression conditions and fusion strategies, no soluble $IscS_{Ls}$ protein was detected. This persistent lack of expression necessitated the use of $IscS_{Ec}$ from *E. coli* in enzymatic assays. The inability to express $IscS_{Ls}$ poses a significant limitation for directly assessing its role in $LarE_{Ls}$ activity

Bioinformatics analysis revealed that the amino acid sequence of $IscS_{Ls}$ is predicted to be unstable, with an instability index of 43.22 (an instability index greater than 40 indicates instability).[14] This high instability index suggests that $IscS_{Ls}$ may inherently possess structural features that make it prone to degradation or misfolding when expressed in *E. coli*. Further analysis showed that all IscS homologs in the five *Latilactobacillus* species exhibiting this gene

distribution are similarly predicted to be unstable, except for IscS from *Latilactobacillus fuchuensis*, which has an instability index of 33.44, indicating greater stability. This finding suggests that IscS from *L. fuchuensis* may be a more suitable candidate for future studies aiming to investigate the interaction between IscS and LarE, as its predicted stability could facilitate successful expression and functional assays.

The enzymatic assays yielded variable and often conflicting results. Contrary to the initial hypothesis, the presence of $IscS_{Ec}$ did not enhance $LarE_{Ls}$ activity. In fact, in most cases, enzymatic activity was reduced or remained unchanged in the presence of $IscS_{Ec}$ compared to samples without it. This unexpected inhibition suggests that $IscS_{Ec}$ may interfere with $LarE_{Ls}$ activity rather than facilitate its regeneration.

Protein mass spectrometry provided valuable insights into the structural dynamics of $LarE_{Ls}$ under various assay conditions. The detection of a Dha modification and the formation of a P2CMN-LarE adduct corroborate the enzyme's catalytic activity and its interaction with substrates.[3] However, the absence of the $LarE_{Ls}$ persulfide species (34,713 Da) suggests that the proposed sulfur transfer from $IscS_{Ec}$ to $LarE_{Ls}$ does not occur under the tested conditions. This finding is pivotal as it undermines the hypothesized mechanism where IscS facilitates LarE regeneration through persulfide formation. Though not detected here, it is possible that the sulfur transfer to $LarE_{Ls}$ from $IscE_{Ec}$ occurs but is too short-lived to detect under the conditions used here. It may also be the case that $LarE_{Ls}$ is not compatible with IscS from *E. coli.* Future work on this project should use LarE and IscS pairs from the same organism.

A possible explanation for the perceived inhibitory effects observed in the $IscS_{Ec}$-containing samples could be the inhibition of $LarE_{Ls}$ by L-Cys. Mass spectrometry analysis identified a species with a mass of 34,800 Da, corresponding to a LarE-S-S-L-Cys disulfide, similar to the previously reported LarE-S-S-CoA disulfide.[3] The coenzyme A (CoA) disulfide species has been implicated in the regeneration of the cysteine sidechain from Dha. The same study found that L-Cys persulfide was capable of restoring some LarE to its native form, though at much lower levels compared to CoA persulfide.[3] Additionally, CoA was shown to bind LarE in a manner similar to ATP, inhibiting LarE activity by competing with ATP for binding.[3] The superior ability of CoA persulfide to regenerate native LarE may be attributed to the stronger interaction between LarE and CoA, relative to LarE and cysteine.

It is possible that the weak interaction between $LarE_{Ls}$ and $IscE_{Ec}$ contributes to the absence of increased activity in $IscS_{Ec}$-containing samples. This possibility raises the hypothesis that stronger protein-protein interactions may exist between LarE and IscS from the same organism, facilitating more efficient transfer of a persulfide sulfur and promoting LarE regeneration. Investigating the binding affinity and interaction dynamics of $LarE_{Ls}$ with $IscS_{Ls}$ could help elucidate whether these species-specific interactions are crucial for sulfur transfer and LarE catalysis.

In the *Latilactobacillus* species containing adjacent *lar* operon and *iscS* gene, two additional genes, DUF1831 and *mnmA*, are consistently found immediately downstream of *iscS*. This conserved arrangement suggests a functional relationship among these genes, potentially linking their roles in sulfur metabolism.

DUF1831 encodes a protein with a conserved domain of unknown function. However, its consistent presence downstream of *iscS* implies a supportive or regulatory role in sulfur transfer processes. DUF1831 may facilitate interactions between sulfur transfer enzymes or participate in stabilizing sulfur intermediates. Its proximity to *iscS* and *larE* raises the possibility that it could mediate or enhance the interaction between IscS and LarE.

The other downstream gene, *mnmA*, encodes tRNA 2-thiouridinesynthase, an enzyme responsible for adding a sulfur atom to the wobble position of certain tRNAs.[15] This modification is crucial for accurate translation during protein synthesis. MnmA depends on sulfur supplied by a donor, such as IscS, suggesting that the genomic co-localization of *iscS* and *mnmA* allows for efficient sulfur transfer.

The presence of the DUF1831 gene and *mnmA* alongside *iscS* and *larE* suggests a different role for IscS in these bacteria. It is possible that the neighboring *larE* and *iscS* genes have no functional relationship and their mutual involvement in sulfur chemistry is a mere coincidence.

Conclusion

This study aimed to explore the potential role of IscS in the regeneration or sustained activity of the sulfur-insertase LarE from *L. sakei*. While $LarE_{Ls}$ was successfully expressed and purified, attempts to express its native partner, $IscS_{Ls}$, were unsuccessful, necessitating the use of $IscS_{Ec}$ as a substitute. Enzymatic assays revealed that the presence of $IscS_{Ec}$ did not enhance Lar_{Ls} activity as hypothesized; instead, it often resulted in reduced or unchanged activity levels. Mass spectrometry provided evidence of a post-translational modification of $LarE_{Ls}$ involving

conversion of a cysteinyl residue to Dha, but it failed to detect the proposed persulfide intermediate, casting doubt on the envisioned IscS-facilitated persulfide mechanism of sulfur transfer or persulfide-dependent regeneration of the Dha-containing form of LarE.

These findings suggest that the genomic proximity of *larE* and *iscS* in *Latilactobacillus* species may not translate to a functional interaction as previously hypothesized, at least not under the experimental conditions tested. The absence of enhanced activity and the lack of detectable persulfide species indicate that alternative mechanisms or additional factors may be required for $LarE_{Ls}$ regeneration and sustained activity. Consequently, the traditional view of LarE as a single-turnover enzyme remains largely unchallenged by this study, highlighting the need for further investigation to uncover the true nature of LarE-IscS interactions in these bacterial systems.

REFERENCES

(1) Desguin, B.; Goffin, P.; Viaene, E.; Kleerebezem, M.; Martin-Diaconescu, V.; Maroney, M. J.; Declercq, J. P.; Soumillion, P.; Hols, P. Lactate Racemase Is a Nickel-Dependent Enzyme Activated by a Widespread Maturation System. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4615.

(2) Desguin, B.; Soumillion, P.; Hols, P.; Hausinger, R. P. Nickel-Pincer Cofactor Biosynthesis Involves LarB-Catalyzed Pyridinium Carboxylation and LarE-Dependent Sacrificial Sulfur Insertion. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (20), 5598–5603. https://doi.org/10.1073/pnas.1600486113.

(3) Fellner, M.; Rankin, J. A.; Desguin, B.; Hu, J.; Hausinger, R. P. Analysis of the Active Site Cysteine Residue of the Sacrificial Sulfur Insertase LarE from Lactobacillus Plantarum. *Biochemistry* **2018**, *57* (38), 5513–5523. https://doi.org/10.1021/acs.biochem.8b00601.

(4) Chatterjee, S.; Parson, K. F.; Ruotolo, B. T.; McCracken, J.; Hu, J.; Hausinger, R. P. Characterization of a [4Fe-4S]-Dependent LarE Sulfur Insertase That Facilitates Nickel-Pincer Nucleotide Cofactor Biosynthesis in Thermotoga Maritima. *J. Biol. Chem.* **2022**, *298* (7), 102131. https://doi.org/10.1016/j.jbc.2022.102131.

(5) Zecchin, P.; Pecqueur, L.; Oltmanns, J.; Velours, C.; Schünemann, V.; Fontecave, M.; Golinelli-Pimpaneau, B. Structure-Based Insights into the Mechanism of [4Fe-4S]-Dependent Sulfur Insertase LarE. *Protein Sci.* **2024**, *33* (2), 1–19. https://doi.org/10.1002/pro.4874.

(6) Black, K. A.; Dos Santos, P. C. Shared-Intermediates in the Biosynthesis of Thio-Cofactors: Mechanism and Functions of Cysteine Desulfurases and Sulfur Acceptors. *Biochim. Biophys. Acta - Mol. Cell Res.* **2015**, *1853* (6), 1470–1480. https://doi.org/10.1016/j.bbamcr.2014.10.018.

(7) Das, M.; Dewan, A.; Shee, S.; Singh, A. The Multifaceted Bacterial Cysteine Desulfurases: From Metabolism to Pathogenesis. *Antioxidants* **2021**, *10* (7). https://doi.org/10.3390/antiox10070997.

(8) Chatterjee, S.; Nevarez, J. L.; Rankin, J. A.; Hu, J.; Hausinger, R. P. Structure of the LarB-Substrate Complex and Identification of a Reaction Intermediate during Nickel-Pincer Nucleotide Cofactor Biosynthesis. *Biochemistry* **2023**, *62* (21), 3096–3104. https://doi.org/10.1021/acs.biochem.3c00242.

(9) Desguin, B.; Fellner, M.; Riant, O.; Hu, J.; Hausinger, R. P.; Hols, P.; Soumillion, P. Biosynthesis of the Nickel-Pincer Nucleotide Cofactor of Lactate Racemase Requires a CTP-Dependent Cyclometallase. *J. Biol. Chem.* **2018**, *293* (32), 12303–12317. https://doi.org/10.1074/jbc.RA118.003741.

(10) Desguin, B.; Goffin, P.; Viaene, E.; Kleerebezem, M.; Martin-Diaconescu, V.; Maroney, M. J.; Declercq, J. P.; Soumillion, P.; Hols, P. Lactate Racemase Is a Nickel-Dependent Enzyme Activated by a Widespread Maturation System. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4615.

(11) Rankin, J. A.; Mauban, R. C.; Fellner, M.; Desguin, B.; McCracken, J.; Hu, J.; Varganov, S. A.; Hausinger, R. P. Lactate Racemase Nickel-Pincer Cofactor Operates by a Proton-

Coupled Hydride Transfer Mechanism. *Biochemistry* **2018**, *57* (23), 3244–3251. https://doi.org/10.1021/acs.biochem.8b00100.

(12) Szklarczyk, D.; Gable, A. L.; Nastou, K. C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N. T.; Legeay, M.; Fang, T.; Bork, P.; et al. The STRING Database in 2021: Customizable Protein–Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res.* **2021**, *49* (18), 10800–10800. https://doi.org/10.1093/nar/gkab835.

(13) Saha, C. K.; Pires, R. S.; Brolin, H.; Delannoy, M.; Atkinson, G. C. FlaGs and WebFlaGs: Discovering Novel Biology through the Analysis of Gene Neighbourhood Conservation. *Bioinformatics* **2021**, *37* (9), 1312–1314. https://doi.org/10.1093/bioinformatics/btaa788.

(14) Guruprasad, K.; Reddy, B. V. B.; Pandit, M. W. Correlation between Stability of a Protein and Its Dipeptide Composition: A Novel Approach for Predicting in Vivo Stability of a Protein from Its Primary Sequence. *Protein Eng. Des. Sel.* **1990**, *4* (2), 155–161. https://doi.org/10.1093/protein/4.2.155.

(15) Ogunkola, M.; Wolff, L.; Fenteng, E. A.; Duffus, B. R.; Leimkühler, S. E. Coli MnmA Is an Fe-S Cluster-Independent 2-Thiouridylase. *Inorganics* **2024**, *12* (3). https://doi.org/10.3390/inorganics12030067.

APPENDIX

```
CLUSTAL O(1.2.4) multiple sequence alignment


LarE_Tm      ----MDKLQRISEAIKSKKKLVVMFSGGVDSTLLAKLAREVLGKNAVAL-TIDSPVIPRK      55
LarE_Ls      MNTLSEKKAMLAANLQQMKKVIVAFSGGIDSTLVLKMALETLGKENVLAVVANSELFTDE      60
LarE_Lp      MATLATKKATLVAALKDLQRVTVAFSGGIDSTLVLKMALDVLGRDNVTAVVANSELFTDE      60
                   *   :   ::. ::: * ****:****: *:* :.**:: *   . :* ::  :

LarE_Tm      EIEEAKNLANLIGIRHEFIELNELKSRHLIENPPDRCYLCRKLRDNIVKNWARENGFDVI      115
LarE_Ls      EFNKAIDLAHEMGANVETTTLNYLSDEHIANNTPETWYYSKKMFYTQLNQIAAEKGFDHV      120
LarE_Lp      EFDKAMSLAEELGANVQGTTLDYLSDDHIKNNTPDSWYYAKKMFYSRLNDIAANNGSAAV      120
             *:::* .**. :* . :   *: *.. *: :* *:  * .:*:  . ::: * ::*   :

LarE_Tm      ADGLNFSDLQDYRPGVKASTEDGIWHPFIDFEVTKEEIREYSKKLGLPTWDKPAMACLCS      175
LarE_Ls      LDGMIMDDNSDFRPGLRARDEAGAESVLQTAGFYKIDVRQLAKELNLNNWNKVASCSVSS      180
LarE_Lp      LDGMIKNDENDYRPGLKARSEAGARSLLQEADFFKTDVRALAQELGLTNWNKVASCSVSS      180
              **:  .* .*:***::*  * *    :    . * ::*  :::*.* .*:* * ..:.*

LarE_Tm      RFPYGFGLNEERVRMVEKAENFLRELGFREVRVRFFPYKTAVVEVGRDEMELLMGKRTDI      235
LarE_Ls      RFPYNTALTEDKIAQVMQSEKFLRDLGFATVRVRYHESI-ARVEVPEAQIADFLTHSEQI      239
LarE_Lp      RFPYGTTLTHDNIAQVMAAEKYLRSLGFPTVRVRFHNDI-ARIELPEARIGDFLVFNDRV      239
             ****.  *..:.:  *  :*::**.***  ****:.    * :*: . .:  ::     :

LarE_Tm      VLALQKIGFSFVTLDLEGFASGKLNRTIEGMSK------    268
LarE_Ls      NQALQQAGFEFVTVDLAGFKSGRMNESLSEQQKVALMTA    278
LarE_Lp      NRQLQSLGFRYVTLDLGGFRSGRMNDTLTKAQLATFA--    276
                **. ** :**:** ** **::* ::   .
```

Figure 4-A1. Sequence alignment of LarE$_{Tm}$, LarE$_{Ls}$, and LarE$_{Lp}$. Whereas LarE$_{Tm}$ has Cys residues at positions 92, 95, 172, and 174, LarE$_{Ls}$ and LarE$_{Lp}$ possess single Cys residue at positions 177 and 176, respectively.

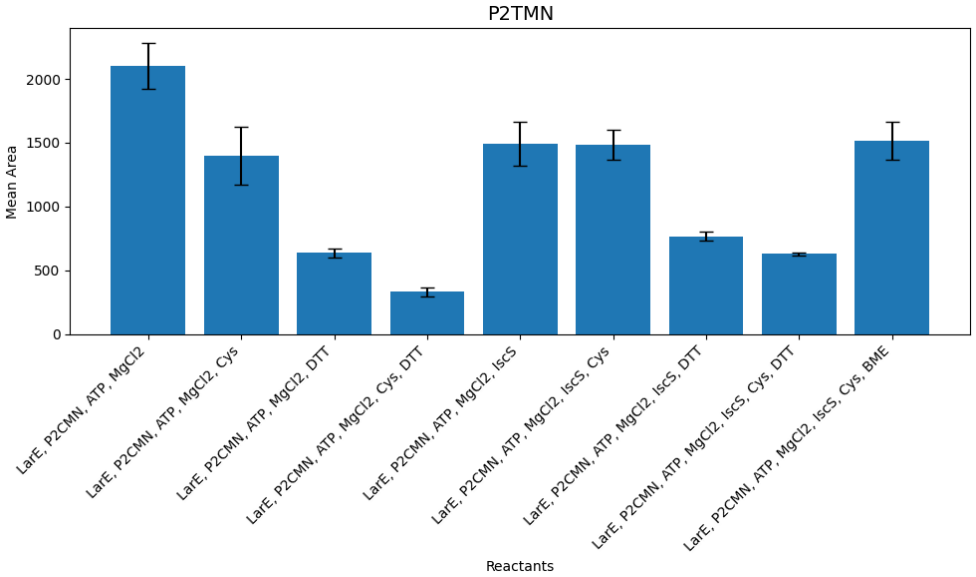

Figure 4-A2. Peak intensity for P2TMN with different reactants added to the LarE reaction. Error bars represent the standard error of the mean. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

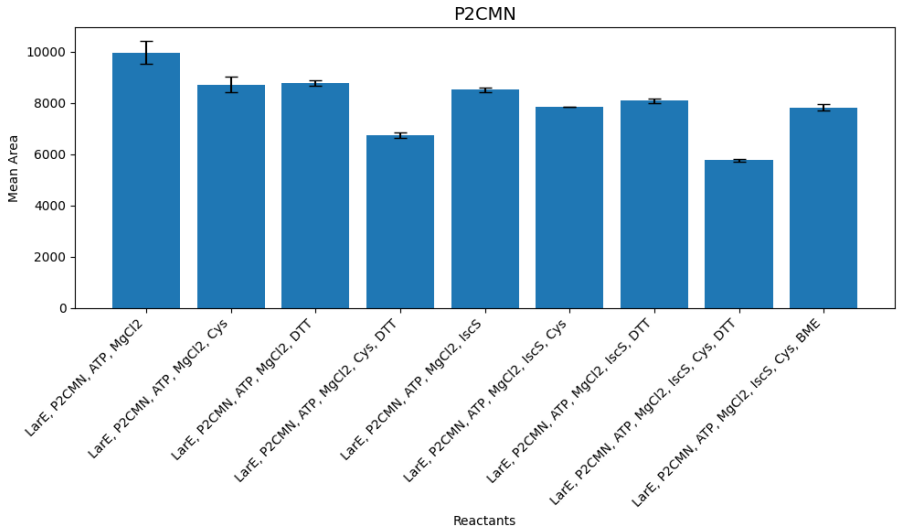

Figure 4-A3. Peak intensity for P2CMN with different reactants added to the LarE reaction. Error bars represent the standard error of the mean. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

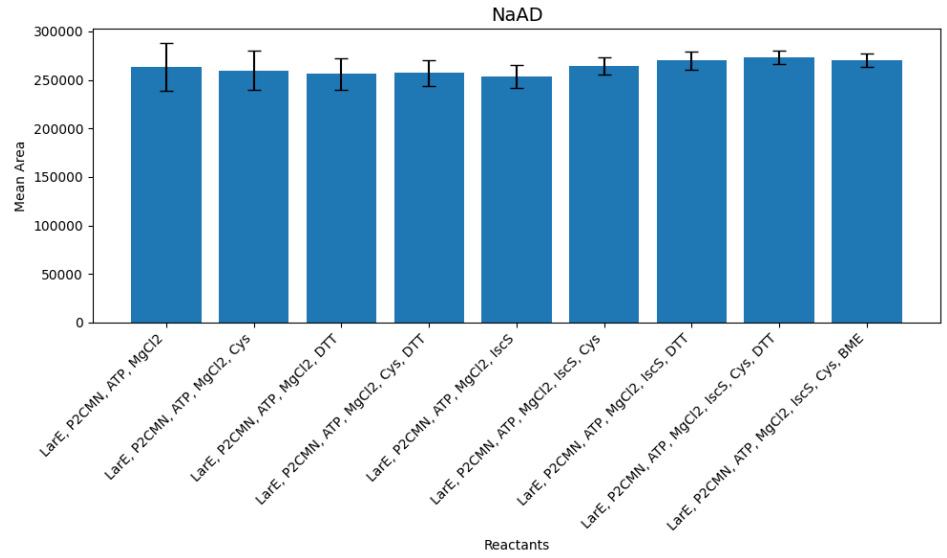

Figure 4-A4. Peak intensity for NaAD with different reactants added to the LarE reaction. Error bars represent the standard error of the mean. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

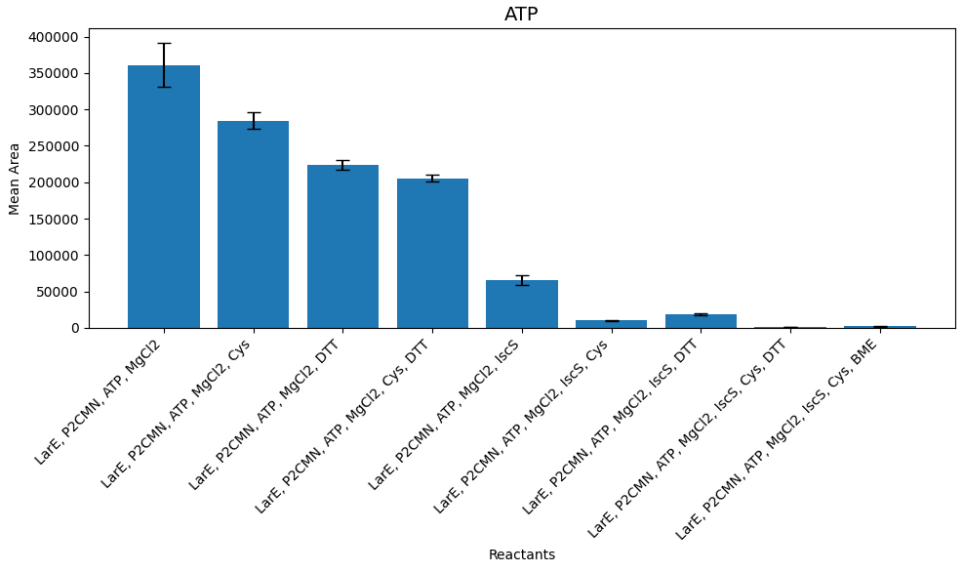

Figure 4-A5. Peak intensity for ATP with different reactants added to the LarE reaction. Error bars represent the standard error of the mean. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

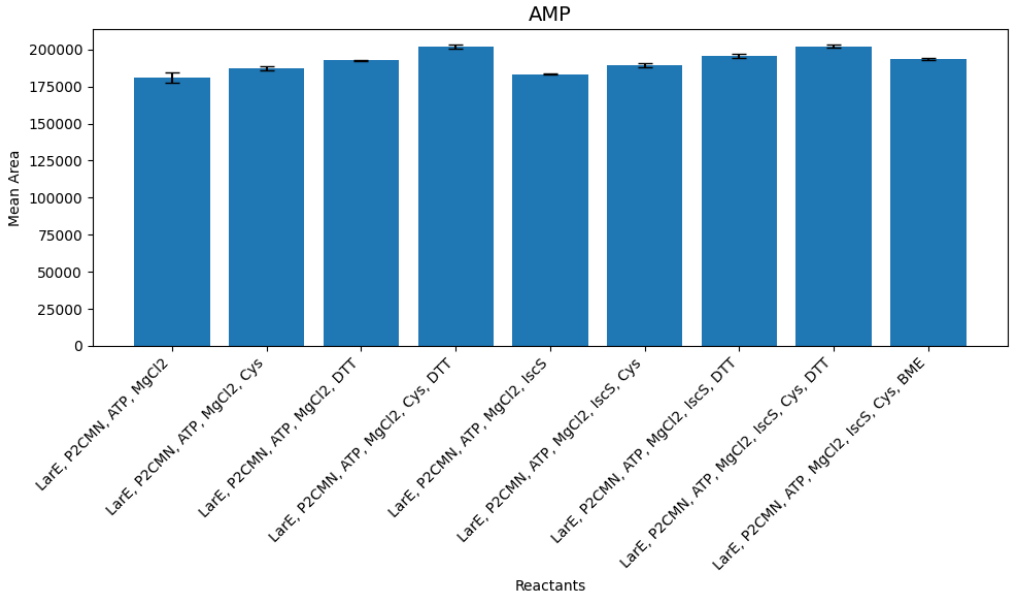

Figure 4-A6. Peak intensity for AMP with different reactants added to the LarE reaction. Error bars represent the standard error of the mean. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

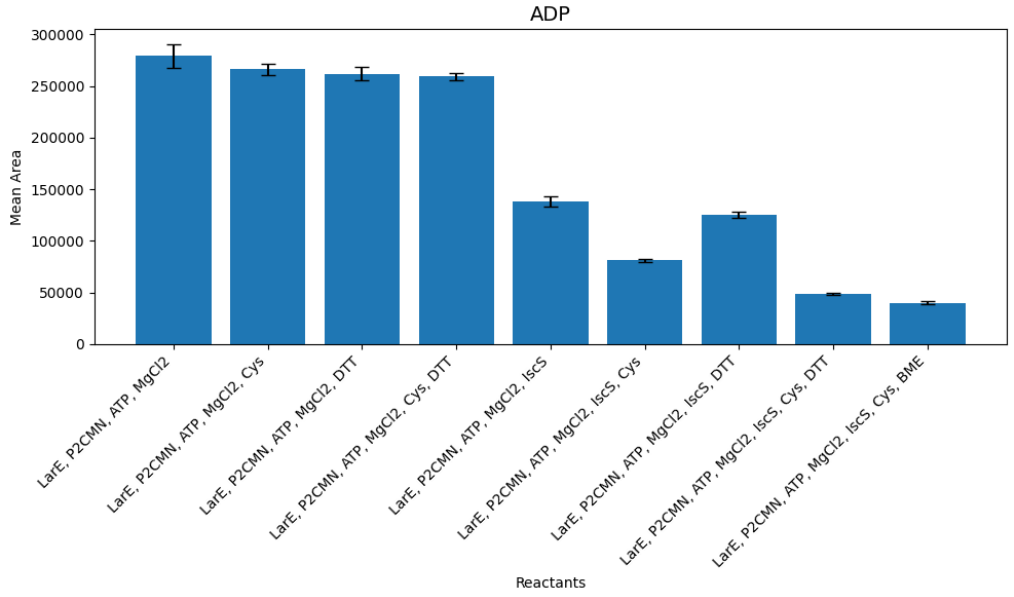

Figure 4-A7. Peak intensity for ADP with different reactants added to the LarE reaction. Error bars represent the standard error of the mean. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

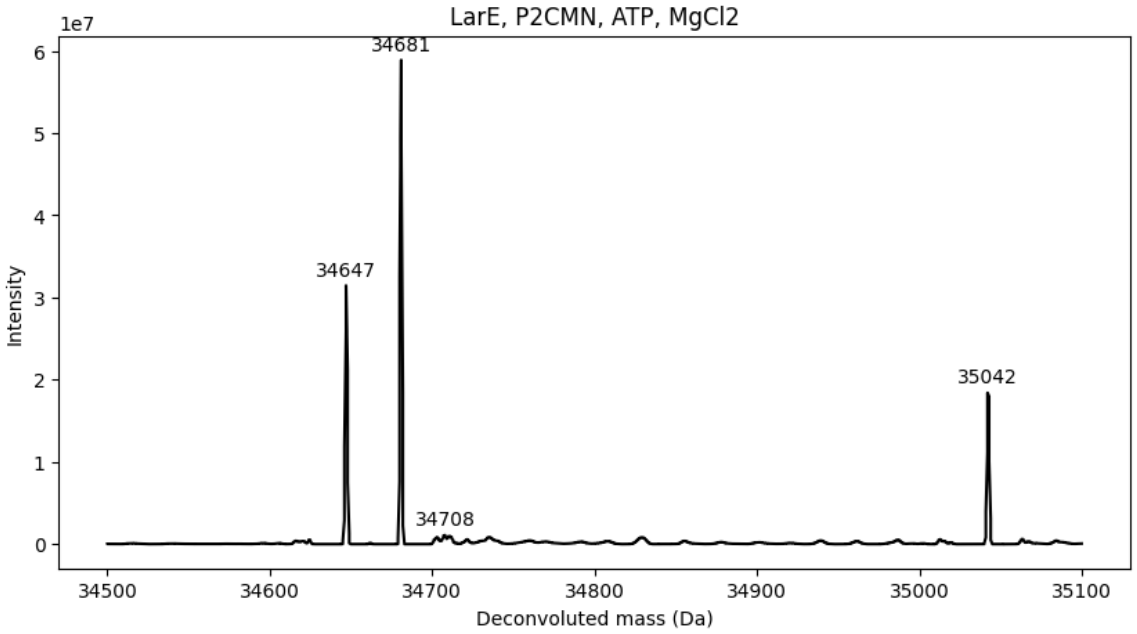

Figure 4-A8. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, and $MgCl_2$. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

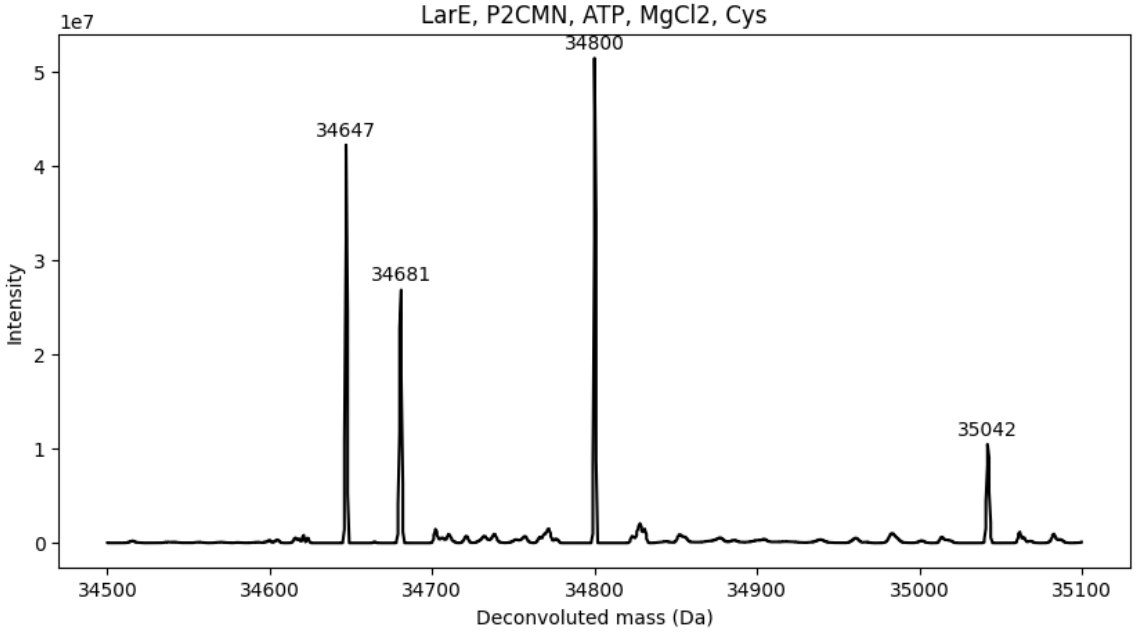

Figure 4-A9. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, and L-Cys. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

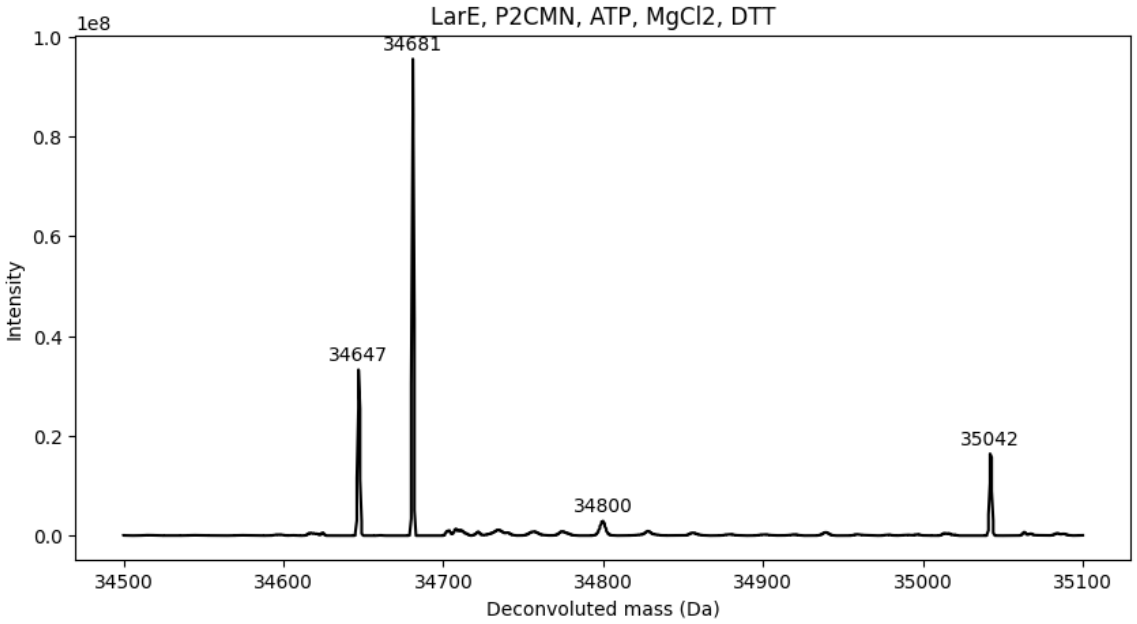

Figure 4-A10. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, and DTT. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

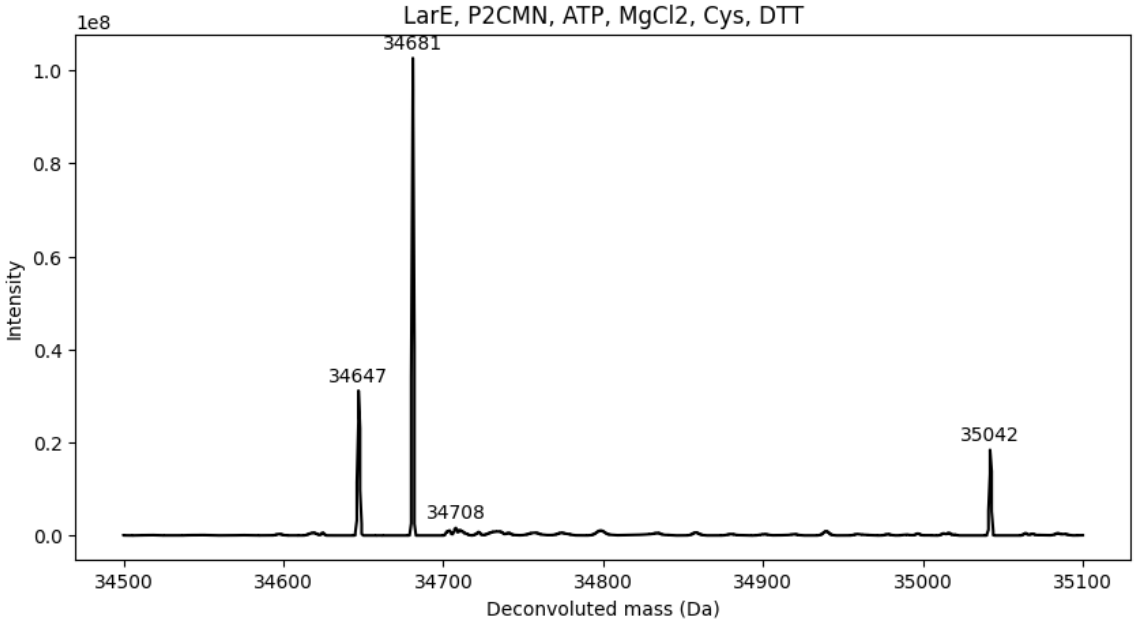

Figure 4-A11. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, L-Cys, and DTT. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

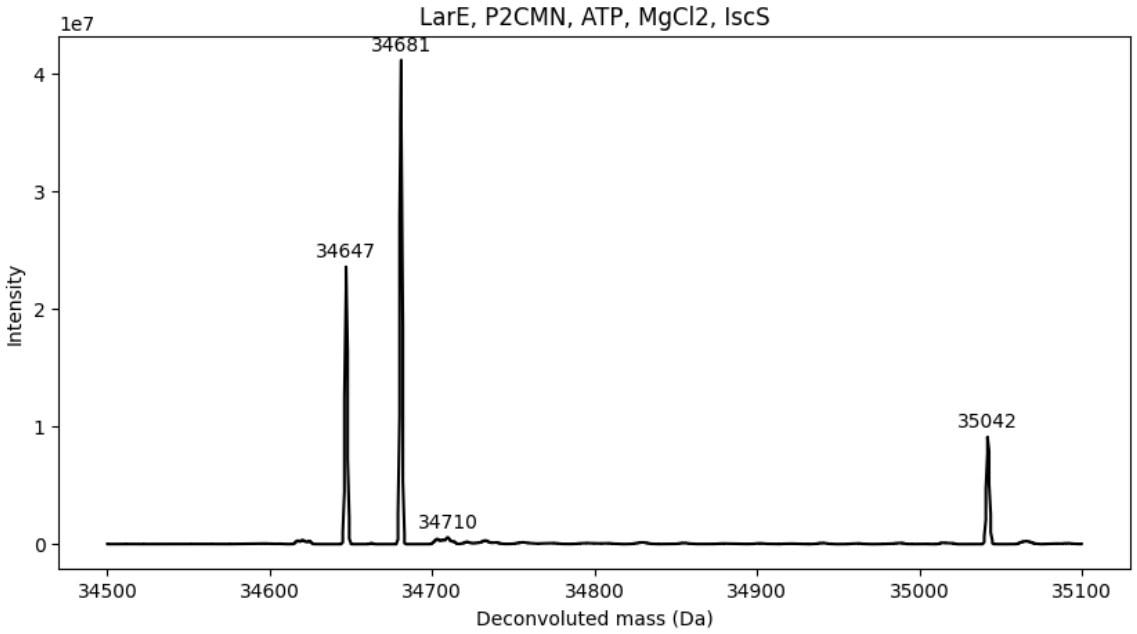

Figure 4-A12. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, and IscS. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

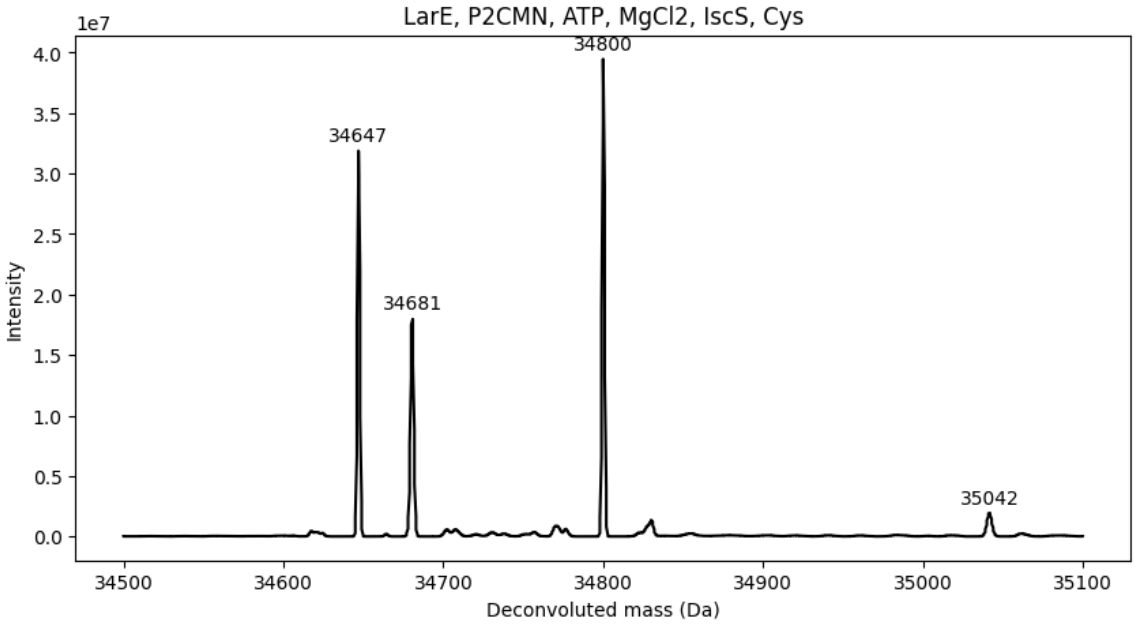

Figure 4-A13. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, IscS, and L-Cys. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

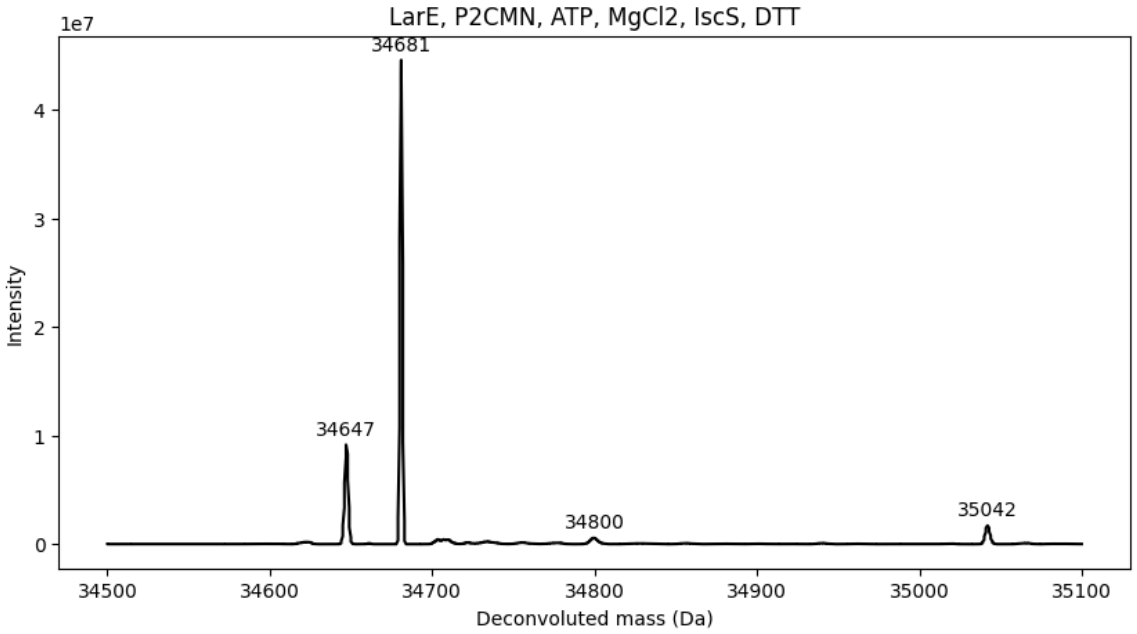

Figure 4-A14. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, IscS, and DTT. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

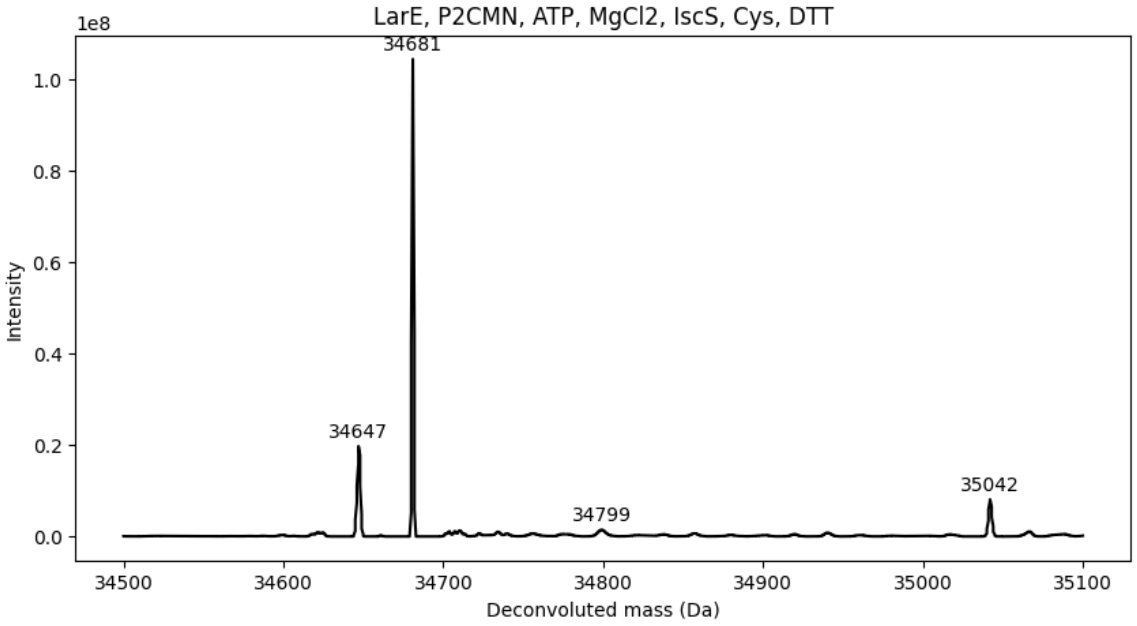

Figure 4-A15. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, IscS, L-Cys, and DTT. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

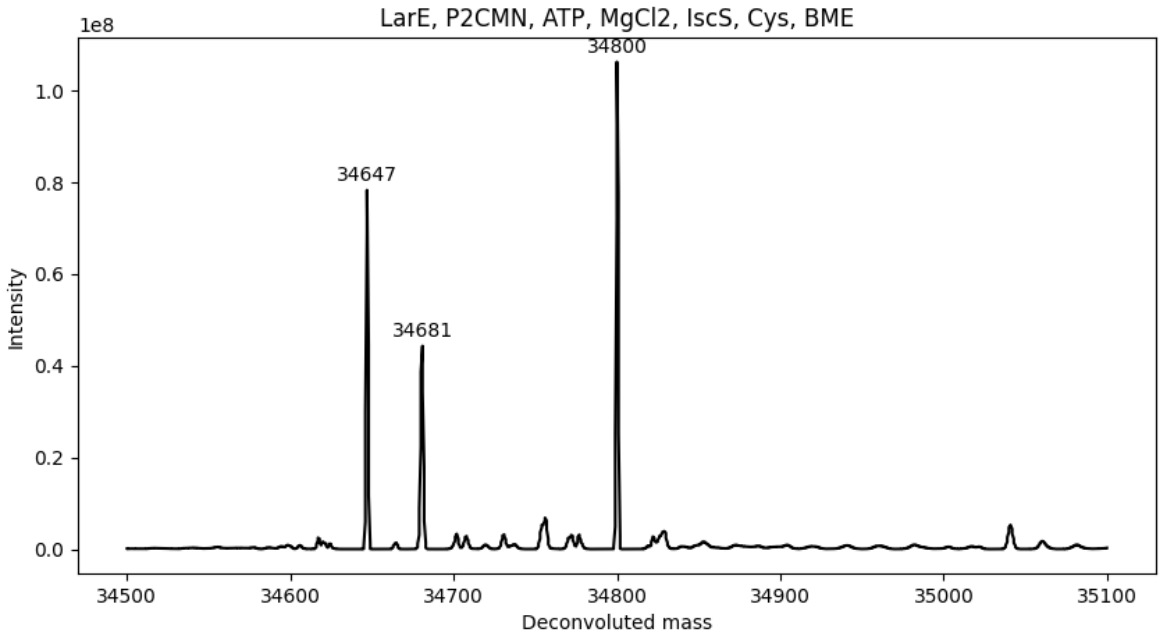

Figure 4-A16. Deconvoluted mass spectrum of the reaction containing LarE, P2CMN, ATP, $MgCl_2$, IscS, L-Cys, and BME. P2CMN refers to the supernatant obtained after precipitating the protein from a reaction mixture containing LarB S127A, NaAD, and bicarbonate.

CHAPTER 5

Sulfonyl azide resins for the sequestration of NPN cofactor-containing proteins from cell lysates

Introduction

Chapter 2 contains a discussion of utilizing lissamine rhodamine B sulfonyl azide (LRSA) to label lactate racemase from *Lactiplantibacillus plantarum* ($LarA_{Lp}$). LRSA selectively labeled LarA by covalently binding to its cofactor, the nickel-pincer nucleotide (NPN), within cell extracts from *E. coli* expressing the *larABCE* genes from *L. plantarum*. This labeling, confirmed via SDS-PAGE and fluorescence imaging, provides a powerful approach for detecting NPN cofactor-bound proteins.

Building on this method, this chapter also explores the application of LRSA to cell lysates of cyanobacteria, specifically targeting *Synechocystis* sp. PCC 6803 (*Synechocystis*)—a widely used model organism in cyanobacterial research due to its rapid growth, genetic tractability, and detailed physiological characterization.[1] Notably, while cyanobacteria possess homologs of *larB*, *larC*, and *larE*, no direct *larA* homolog has been identified,[2] suggesting that the NPN cofactor is associated with a yet unknown protein of indefinite function.

Further, this chapter examines the use of benzenesulfonyl azide (BSA) functionalized resins, such as polymer-bound (PBSA) and silica-bound (SBSA) versions, for isolating NPN cofactor-containing proteins from cell lysates of *Lactococcus lactis* expressing *larABCE* from *L. plantarum* and *Synechocystis*. The sulfonyl azide moiety in these resins selectively reacts with thioacids, forming stable thioamide bonds that facilitate efficient protein sequestration (**Figure 5-1**).

The combined application of LRSA and BSA-based resins provides an expanded toolkit for investigating NPN cofactor-dependent proteins. These techniques enable researchers to identify, tag, and isolate proteins that possess covalently bound NPN cofactor (and potentially for proteins

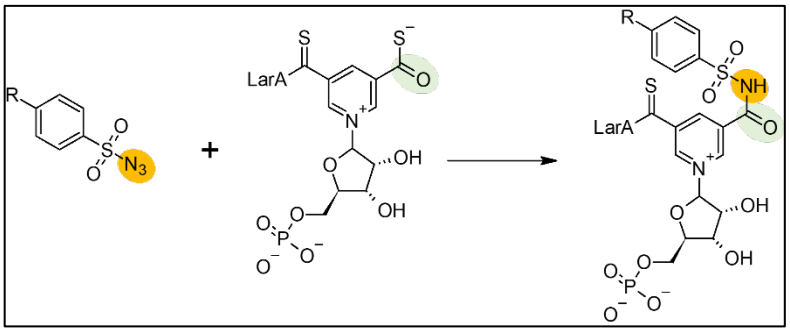

Figure 5-1. Scheme depicting the coupling between the resin bound sulfonyl azide and the free thioacid moiety of the NPN cofactor bound to LarA.

with non-covalently bound NPN cofactor) in diverse biological matrices, advancing our understanding of the role of NPN in microbial physiology. This chapter thereby seeks to uncover potential roles for NPN cofactor within cyanobacteria and other organisms by harnessing the selectivity and precision of these novel labeling and isolation tools.

Materials and Methods

Cell Culture

L. lactis growth was carried out as described previously.[3] Briefly, *larA* from *L. plantarum* was expressed in *L. lactis* strain NZ3900 containing the plasmid pGIR112. This plasmid contains the *larABCDE* genes under the control of the *nisA* promoter. Starter cultures were initially grown overnight at 30 °C in M17 medium supplemented with 0.5% (w/v) D-glucose and 10 mg/L chloramphenicol. These cultures were then used to inoculate (1% final volume) larger volumes of the same medium containing 5 mg/L chloramphenicol and grown at 28 °C with mild shaking. When the optical density at 600 nm (OD_{600}) reached 0.2–0.3, the cultures were induced with 10 µg/L nisin A, supplemented with 1 mM $NiCl_2$, and allowed to grow for an additional 4 h. Finally, cells were harvested by centrifugation at 5,000 x *g* for 10 min at 4 °C, the supernatant was discarded, and the cell pellets were stored at -80 °C.

Synechocystis sp. PCC 6803 was cultured in BG-11 medium that was prepared following the standard formulation.[4] Cultures were maintained at 30 °C under continuous illumination with white, fluorescent light at an intensity of approximately 50 µmol photons m^{-2} s^{-1}. Inoculation was performed to achieve an initial OD_{750} of 0.1. Growth was monitored by measuring OD_{750} at regular intervals until the culture reached the desired density, typically between OD_{750} 0.6 and 1.0. Finally, cells were harvested by centrifugation at 5,000 x *g* for 10 min at 4 °C. The cell pellet was then stored at -80 °C for later use.

Sulfonyl Azide Resin

The procedure for cell lysate preparation used here followed the same protocol as described in Chapter 2. Briefly, 0.5 g of cells were resuspended in 0.7 mL of Tris buffer (100 mM Tris-buffered saline with 150 mM NaCl at pH 7.5) and transferred to a 2 mL tube for lysis using a mini-beadbeater. Following lysis, the lysates were centrifuged at 13,000 rpm for 5 min, and the supernatant was collected. To obtain approximately uniform protein concentration, samples were normalized based on absorbance at 280 nm. Finally, using a desalting spin column,

the lysates were buffer exchanged into 50 mM potassium phosphate, 300 mM NaCl, and 6 M urea at pH 6.1, prior to reaction with LRSA and sulfonyl azide resins.

Approximately 50 mg of resin (either PBSA or SBSA) was equilibrated with urea buffer (50 mM potassium phosphate, 300 mM NaCl, and 6 M urea, pH 6.1) by performing a series of washes. For each wash, urea buffer was added to the resin in a 1.5 mL microfuge tube, briefly vortexed for 10 s, then centrifuged for 30 s at 800 rpm in a benchtop centrifuge. After centrifugation, the supernatant was removed and fresh urea buffer was added to the resin for the next wash. This process was repeated a total of three times, with no additional buffer added after the final wash, leaving the resin fully equilibrated.

Following equilibration, 200 µL of the lysates in urea buffer were added to the microfuge tube containing the resin. This mixture was gently rocked on a benchtop rocker for 30 min to allow thorough binding. After incubation, the tube was centrifuged at 13,000 rpm for 2 min, and the supernatant was collected for further analysis.

The resin was then subjected to a wash step similar to the initial equilibration procedure. First, the resin was washed three times with urea buffer, following the same method as the equilibration. Then the resin was washed three additional times using Tris buffer, to thoroughly remove nonspecifically binding proteins.

LRSA Labeling of Lysates

The procedure for LRSA labeling is described in detail in Chapter 2 of this dissertation. A concise protocol is described here. Lysates were labeled with 10 µL of 15 mM LRSA in dimethylsulfoxide and incubated in the dark at room temperature for 20 min. Proteins were precipitated via the chloroform-methanol method, then resuspended in 50 mM potassium phosphate, 300 mM NaCl, and 6 M urea, pH 6.1 Each sample was mixed with sodium dodecyl sulfate (SDS)-loading buffer, loaded onto a 12% acrylamide gel, and run on SDS-polyacrylamide gel electrophoresis (PAGE). Gels were imaged to detect rhodamine-labeled proteins and subsequently stained with Coomassie Brilliant Blue. In some cases, the fluorescence image was used as a reference guide for identifying gel bands that were excised using a sterile scalpel and later subjected to proteomics analysis.

Proteomics Analysis

Proteolytic digestion of gel bands and proteins bound to the azide resins was performed following the protocol by Shevchenko *et al*.,[5] with modifications. Samples were initially

dehydrated using 100% acetonitrile, then incubated with 10 mM dithiothreitol in 100 mM ammonium bicarbonate (pH ~8) at 56 °C for 45 min. After this reduction step, the samples were dehydrated again, followed by incubation in the dark with 50 mM iodoacetamide in 100 mM ammonium bicarbonate for 20 min to alkylate the cysteines. Subsequently, the samples were washed with ammonium bicarbonate and dehydrated once more. Sequencing-grade modified trypsin or chymotrypsin was prepared to a concentration of 0.005 μg/μL in 50 mM ammonium bicarbonate, and approximately 100 μL was added to each sample to ensure complete submersion. The samples were incubated overnight at 37 °C to facilitate proteolysis. Peptides were extracted by water bath sonication in a solution of 60% acetonitrile (ACN) and 1% trifluoroacetic acid (TFA), then vacuum-dried to a final volume of approximately 2 μL.

LC/MS/MS Analysis

Peptide samples were analyzed via liquid chromatography/dual mass spectrometry (LC/MS/MS), with 5 μL of sample injected using a Thermo EASYnLC 1200 system. Peptides were loaded onto a Thermo Acclaim PepMap RSLC 0.1 mm x 20 mm C18 trapping column and washed for approximately 5 min with Buffer A (99.9% water, 0.1% formic acid). Bound peptides were then eluted over 35 min onto a Thermo Acclaim PepMap RSLC 0.075 mm x 250 mm resolving column with a gradient increasing from 8% Buffer B to 40% Buffer B in 24 min, ramping to 90% Buffer B at 25 min, and held at 90% for the remainder of the run. Buffer B consisted of 80% acetonitrile, 0.1% formic acid, and 19.9% water. The flow rate was maintained at 300 nL/min, and the column temperature was kept constant at 50 °C using an integrated column oven (PRSO-V2, Sonation GmbH, Biberach, Germany).

Eluted peptides were ionized and introduced into a ThermoScientific Q-Exactive HF-X mass spectrometer via a FlexSpray ion source. Survey scans were acquired in the Orbitrap (60000 resolution, determined at m/z) followed by higher-energy collision-induced dissociation (HCD) of the top 15 ions from each survey scan, with fragment spectra recorded at 15,000 resolution. The resulting MS/MS spectra were converted to peak lists using Mascot Distiller (v2.7.0) and searched against a database containing *Lactococcus lactis*, *Lactiplantibacillus plantarum*, and *Synechocystis* sp. PCC 6803 protein sequences (downloaded from UniProt), along with common laboratory contaminants (cRAP project, www.thegpm.org) using the Mascot algorithm (v2.7). Mascot search results were analyzed with Scaffold (v4.11.1) to

probabilistically validate protein identifications, with assignments meeting the 1% false discovery rate (FDR) confidence filter considered reliable.

Mascot search parameters included allowance for up to two missed tryptic sites, a fixed modification of carbamidomethyl cysteine, a variable modification for methionine oxidation, a peptide tolerance of ±10 ppm, and an MS/MS tolerance of 0.02 Da. The FDR was calculated by performing a randomized database search. Only samples with greater than 95% confidence in detection were selected for further analysis.

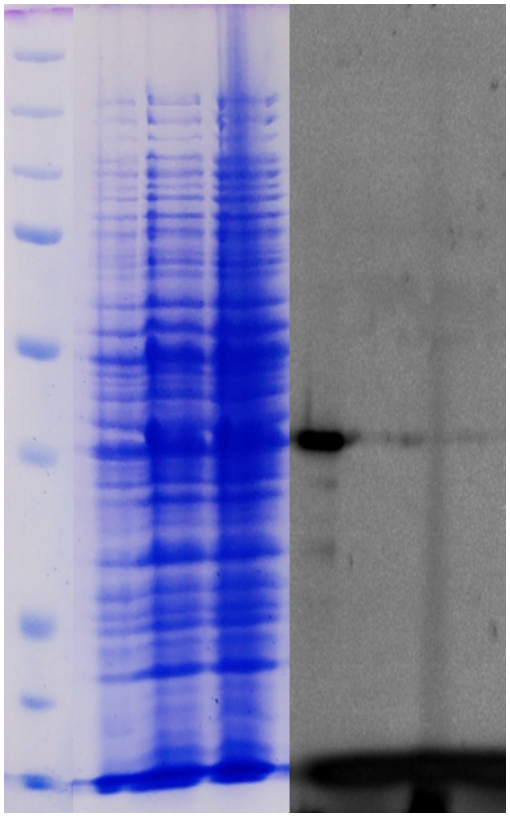

Figure 5-2. Coomassie brilliant blue stain (left side) and fluorescence image (right side) of L. lactis lysate: control (left lane), PBSA-treated (center lane), and SBSA-treated (right lane).

Results

Sulfonyl Azide Resin Sequesters LarA from L. lactis Lysate

To assess the ability of sulfonyl azide resins to selectively bind and sequester LarA from *L. lactis* lysate, an experiment was conducted with three lysate pools: one untreated and the other two incubated with either PBSA or SBSA resin. After incubation (or no treatment in the control),

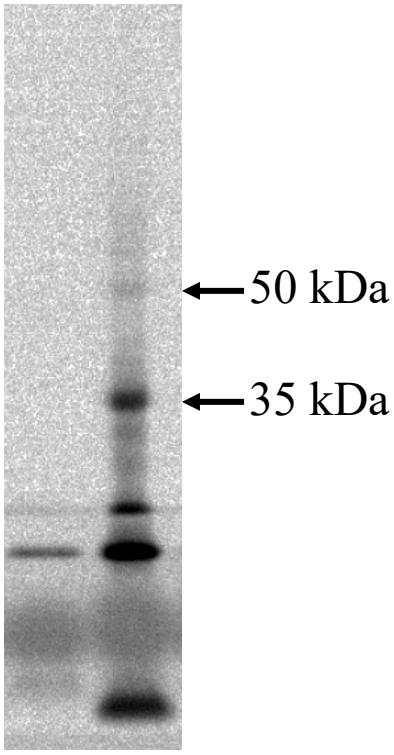

Figure 5-3. Unlabeled (left) and LRSA labeled (right) fluorescence image of Synechocystis lysates.

each pool was labeled with LRSA, followed by SDS-PAGE analysis, fluorescence imaging, and Coomassie Brilliant Blue staining (**Figure 5-2**).

The fluorescence image shows a clear LRSA label in the untreated lysate, as anticipated. In contrast, lysates treated with either PBSA or SBSA exhibit almost no detectable fluorescence signal corresponding to LarA, suggesting its removal from the lysate. Coomassie staining reveals comparable, if not greater, protein concentrations in the resin-treated samples, indicating that the absence of fluorescence is not due to differences in lysate concentration. These findings suggest that LarA likely reacts with the sulfonyl azide groups on the resins, forming a covalent bond and effectively removing it from the bulk lysate.

Application of LRSA and BSA to a Model Cyanobacterium

Following the promising results obtained with LRSA and sulfonyl azide resins for labeling and sequestering LarA in *L. lactis* cell lysates, these techniques were extended to cell extracts of *Synechocystis* sp. PCC 6803. LRSA labeling of the *Synechocystis* lysate revealed two distinct fluorescent bands at approximately 50 kDa and 35 kDa in SDS-PAGE analysis (**Figure**

5-3). These bands were excised for proteomic analysis. Following incubation with *Synechocystis* lysate, the PBSA resin was washed and treated with trypsin then analyzed by proteomics.

The resulting proteomes for the PBSA resin (**Table 5-A1**), the 50 kDa band (**Table 5-A2**), and the 35 kDa band (**Table 5-A3**) did not contain any proteins known to possess a thiocarboxylate group or show connections to Lar proteins or the NPN cofactor. Identified proteins were primarily involved in fundamental cellular functions such as photosynthesis (e.g., Photosystem II proteins, phycobiliproteins), transcription and translation (e.g., RNA polymerase subunits, ribosomal proteins, elongation factors), metabolic pathways (e.g., isocitrate dehydrogenase, fructose-bisphosphate aldolase), and nitrogen metabolism (e.g., glutamine synthetase, nitrogen regulatory protein P-II).

Discussion

The results presented here demonstrate the efficacy of sulfonyl azide resins in selectively sequestering LarA from *L. lactis* lysates. The loss of LarA's fluorescence signal in lysates treated with either PBSA or SBSA strongly supports the hypothesis that the free thiocarboxylate of the NPN cofactor in LarA forms a sulfonamide bond with the resin through reaction with the sulfonyl azide group. This covalent interaction likely results in the removal of LarA from the detectable pool, as confirmed by fluorescence imaging. These findings highlight the potential of sulfonyl azide-based resins as valuable tools for selective protein sequestration, especially for proteins such as LarA homologs that may carry the covalently bound unique NPN cofactor.

In extending these labeling and sequestration methods to *Synechocystis* sp. PCC 6803, the results reveal interesting, albeit unexpected, outcomes. LRSA labeling in *Synechocystis* lysate produced two distinct bands at approximately 50 kDa and 35 kDa. However, proteomic analysis of these bands, along with the proteins sequestered by PBSA resin from *Synechocystis* lysate, did not identify any proteins associated with known LarA homologs, thiocarboxylate modifications, or with NPN cofactor biosynthesis. Instead, the identified proteins were primarily involved in essential cellular processes, such as photosynthesis, transcription, translation, core metabolic pathways, and nitrogen metabolism.

The lack of detection of NPN cofactor-related Lar proteins in *Synechocystis* could also be attributed to the standard culturing conditions used in this study. In *L. plantarum*, lactate racemase activity—and consequently, NPN cofactor synthesis—is induced by an excess of L-lactate.[2] It is possible that NPN cofactor synthesis in cyanobacteria is similarly regulated by

environmental or metabolic conditions not represented in standard growth medium. Alternative expression conditions, such as nutrient stresses, varying carbon sources, or alterations in metal availability, may promote the synthesis of the NPN cofactor or Lar proteins in *Synechocystis*. Investigating these alternative conditions could help elucidate the regulatory factors influencing NPN cofactor synthesis and identify the specific conditions under which Lar-related pathways are activated in cyanobacteria.

The absence of Lar-related proteins or NPN cofactor-containing enzymes in the *Synechocystis* lysate results is notable, especially given the presence of *larB* (sll1489), *larC* (slr1411), and *larE* (slr1717) genes in the *Synechocystis* genome. These genes were shown to confer lactate racemase activity to $LarA_{Lp}$ when expressed together in *E. coli* (see Chapter 2). This discrepancy suggests that while *Synechocystis* possesses genes homologous to those enabling NPN cofactor synthesis, the absence of detectable LarA-related proteins in the lysate indicates either a regulatory mechanism preventing their expression under the tested conditions or possible differences in cofactor biosynthesis or stability within the *Synechocystis* cellular environment. Interestingly, many LarA homologs in other organisms contain a domain of unknown function (DUF362 domain), a region associated with LarA's unique functionality. *Synechocystis* does contain a gene (sll1697) encoding a DUF362 domain-containing protein, though it is relatively uncharacterized. Some annotations suggest it may be a thylakoid-associated protein.[6] There is speculation that NsiR4 (nitrogen stress-induced RNA 4), a cyanobacterial non-coding RNA which plays a role in the regulation of glutamine synthetase (GS) -- a key enzyme in biological nitrogen assimilation -- may be associated with DUF362 in many cyanobacteria.[7] This finding raises the possibility that the DUF362 protein could serve an alternative or adapted role in cyanobacteria, potentially linked to cellular functions distinct from those of LarA.

Conclusion

This chapter has demonstrated the efficacy of sulfonyl azide-functionalized resins in selectively sequestering NPN cofactor-containing proteins, specifically $LarA_{Lp}$, from *L. lactis* lysates. The successful removal of LarA, indicated by the absence of its fluorescence signal after resin treatment of lysates, strongly indicates that the sulfonyl azide moiety forms a covalent bond with the thiocarboxylate group of the covalently bound NPN cofactor. This interaction

effectively isolates LarA from the lysate, highlighting the potential of sulfonyl azide-based resins as valuable tools for the targeted sequestration of NPN cofactor-dependent enzymes.

When extending these techniques to the cyanobacterium *Synechocystis* sp. PCC 6803, LRSA labeling revealed two fluorescent proteins at approximately 50 kDa and 35 kDa. However, proteomic analyses did not identify any proteins related to LarA, thiocarboxylate modifications, or proteins that synthesize the NPN cofactor. Instead, the proteins detected were primarily involved in fundamental cellular processes.

The absence of NPN cofactor-related proteins in *Synechocystis* under standard culturing conditions suggests that the expression of these proteins may be tightly regulated and dependent on specific environmental factors not replicated in our experiments. In *L. plantarum*, NPN cofactor synthesis is induced by an excess of L-lactate, hinting that similar regulatory mechanisms could exist in cyanobacteria. It is plausible that alternative growth conditions—such as nutrient limitation, different carbon sources, or metal availability—are necessary to activate NPN cofactor biosynthesis and Lar protein expression in *Synechocystis*. Additionally, the presence of an uncharacterized DUF362 domain-containing protein, *sll1697*, raises the possibility of divergent or adapted roles for NPN cofactor-related proteins in cyanobacteria, potentially linked to nitrogen metabolism or stress responses.

These findings underscore the utility of sulfonyl azide resins and LRSA labeling as powerful methodologies for studying NPN cofactor-dependent proteins across diverse biological systems. They also highlight the need for optimized experimental conditions to detect and characterize such proteins in organisms where their expression may be condition dependent. Future research should focus on exploring various environmental and metabolic conditions that could induce NPN cofactor synthesis in cyanobacteria. Investigating the role of *sll1697* may also provide insights into alternative functions of the NPN cofactor.

In conclusion, the strategies developed and employed in this chapter expand the toolkit available for the identification and isolation of NPN cofactor-containing proteins. By applying these methods across different organisms and under varied conditions, we can enhance our understanding of the biological roles of the NPN cofactor.

REFERENCES

(1) Ikeuchi, M.; Tabata, S. Ikeuchi & Tabata, 2001. *Photosynth. Res.* **2001**, 73–83.

(2) Desguin, B.; Goffin, P.; Viaene, E.; Kleerebezem, M.; Martin-Diaconescu, V.; Maroney, M. J.; Declercq, J. P.; Soumillion, P.; Hols, P. Lactate Racemase Is a Nickel-Dependent Enzyme Activated by a Widespread Maturation System. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4615.

(3) Rankin, J. A.; Mauban, R. C.; Fellner, M.; Desguin, B.; McCracken, J.; Hu, J.; Varganov, S. A.; Hausinger, R. P. Lactate Racemase Nickel-Pincer Cofactor Operates by a Proton-Coupled Hydride Transfer Mechanism. *Biochemistry* **2018**, *57* (23), 3244–3251. https://doi.org/10.1021/acs.biochem.8b00100.

(4) Stanier, R. Y.; Kunisawa, R.; Mandel, M.; Cohen-Bazire, G. Purification and Properties of Unicellular Blue-Green Algae (Order Chroococcales). *Bacteriol. Rev.* **1971**, *35* (2), 171–205. https://doi.org/10.1128/mmbr.35.2.171-205.1971.

(5) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal. Chem.* **1996**, *68* (5), 850–858. https://doi.org/10.1021/ac950914h.

(6) Srivastava, R.; Pisareva, T.; Norling, B. Proteomic Studies of the Thylakoid Membrane of Synechocystis Sp. PCC 6803. *Proteomics* **2005**, *5* (18), 4905–4916. https://doi.org/10.1002/pmic.200500111.

(7) Klähn, S.; Schaal, C.; Georg, J.; Baumgartner, D.; Knippen, G.; Hagemann, M.; Muro-Pastor, A. M.; Hess, W. R. The SRNA NsiR4 Is Involved in Nitrogen Assimilation Control in Cyanobacteria by Targeting Glutamine Synthetase Inactivating Factor IF7. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (45), E6243–E6252. https://doi.org/10.1073/pnas.1508412112.

APPENDIX

Table 5-A1. List of proteins identified by MS after treating *Synechocystis* sp. PCC 6803 PSAR with trypsin.

Protein Name	**Accession Number**	**Gene Name**	**Molecular Weight**
dTDP-glucose 4-6-dehydratase	P74036	*rfbB*	37 kDa
Threonine--tRNA ligase	Q55806	*thrS*	69 kDa
Ribulose bisphosphate carboxylase large chain	P54205	*cbbL*	52 kDa
RNA polymerase sigma factor SigA	P74565	*sigA*	50 kDa
Protein translocase subunit SecA	Q55709	*secA*	107 kDa
Phycobiliprotein ApcE	Q55544	*apcE*	100 kDa
Photosystem II protein D1 2	P16033	*psbA2*	40 kDa
Photosystem II CP47 reaction center protein	P05429	*psbB*	56 kDa
Photosystem I P700 chlorophyll a apoprotein A2	P29255	*psaB*	81 kDa
OmpR subfamily	P74314	*slr0947*	26 kDa
Isocitrate dehydrogenase [NADP]	P80046	*icd*	52 kDa
Glutamine synthetase	P77961	*glnA*	53 kDa
Glutamate--ammonia ligase	P77970	*glnN*	79 kDa
Fructose-bisphosphate aldolase class 2	Q55664	*fbaA*	39 kDa
Elongation factor Tu	P74227	*tuf*	44 kDa
Elongation factor G 2	P74228	*fusB*	75 kDa
DNA-directed RNA polymerase subunit gamma	P74177	*rpoC1*	71 kDa
DNA-directed RNA polymerase subunit beta	P77965	*rpoB*	123 kDa
DNA gyrase subunit A	Q55738	*gyrA*	95 kDa
Cell division cycle protein	Q55589	*slr0374*	56 kDa
C-phycocyanin beta chain	Q54714	*cpcB*	18 kDa
C-phycocyanin alpha chain	Q54715	*cpcA*	18 kDa
ATP synthase subunit beta	P26527	*atpD*	52 kDa
50S ribosomal protein L22	P73315	*rplV*	14 kDa
50S ribosomal protein L19	P36239	*rplS*	14 kDa
50S ribosomal protein L14	P73310	*rplN*	13 kDa
50S ribosomal protein L6	P73306	*rplF*	20 kDa
50S ribosomal protein L4	P73319	*rplD*	23 kDa
50S ribosomal protein L1	P36236	*rplA*	26 kDa
30S ribosomal protein S12	P74230	*rpsL*	14 kDa
30S ribosomal protein S10	P74226	*rpsJ*	12 kDa
30S ribosomal protein S7	P74229	*rpsG*	17 kDa
30S ribosomal protein S3	P73314	*rpsC*	27 kDa
Circadian clock protein kinase KaiC	P74646	*kaiC*	58 kDa
Uncharacterized protein sll0936	P72861	*sll0936*	55 kDa
Photosystem II D2 protein	P09192	*psbD*	39 kDa
NAD(P)H-quinone oxidoreductase subunit H	P27724	*ndhH*	46 kDa

Table 5-A1 (cont'd)

Circadian clock protein KaiB	P74645	*kaiB*	12 kDa
ATP synthase subunit alpha	P27179	*atpA*	54 kDa
6-phosphogluconate dehydrogenase, decarboxylating	P52208	*gnd*	53 kDa
1-deoxy-D-xylulose-5-phosphate synthase	P73067	*dxs*	69 kDa
Nitrogen regulatory protein P-II	Q55247	*glnB*	12 kDa

Table 5-A2. List of proteins identified by MS after treating trypsin digest of a 50 kDa gel slice from LRSA labeled *Synechocystis sp.* PCC 6803 lysate.

Protein Name	**Accession Number**	**Gene Name**	**Molecular Weight**
Allophycocyanin beta chain	Q01952	*apcB*	17 kDa
Elongation factor Tu	P74227	*tuf*	44 kDa
Cluster of ATP-dependent Clp protease reg subunits	Q55662	*clpC*	91 kDa
ATP-dependent Clp protease regulatory subunit	Q55662	*clpC*	91 kDa
Chaperone protein ClpB 2	P74361	*clpB2*	98 kDa
DNA-directed RNA polymerase subunit beta	P77965	*rpoB*	123 kDa
C-phycocyanin beta chain	Q54714	*cpcB*	18 kDa
60 kDa chaperonin 1	Q05972	*groL1*	58 kDa
Chaperone protein dnaK2	P22358	*dnaK2*	68 kDa
6-phosphogluconate dehydrogenase, decarboxylating	P52208	*gnd*	53 kDa
Ribulose bisphosphate carboxylase large chain	P54205	*cbbL*	52 kDa
Elongation factor G 2	P74228	*fusB*	75 kDa
Nitrogen regulatory protein P-II	Q55247	*glnB*	12 kDa
ATP synthase subunit beta	P26527	*atpD*	52 kDa
C-phycocyanin alpha chain	Q54715	*cpcA*	18 kDa
Photosystem II CP47 reaction center protein	P05429	*psbB*	56 kDa
Photosystem I P700 chlorophyll a apoprotein A1	P29254	*psaA*	83 kDa
Glutamine synthetase	P77961	*glnA*	53 kDa
60 kDa chaperonin 2	P22034	*groL2*	58 kDa
S-adenosylmethionine synthase	P72871	*metK*	46 kDa
DNA-directed RNA polymerase subunit gamma	P74177	*rpoC1*	71 kDa
D-fructose 1,6-bisphosphatase class 2/sedoheptulose 1,7-bisphosphatase	P73922	*slr2094*	37 kDa
Glutamate--ammonia ligase	P77970	*glnN*	79 kDa
ATP-dependent zinc metalloprotease FtsH 4	P73437	*ftsH4*	68 kDa
ATP-dependent zinc metalloprotease FtsH 3	P72991	*ftsH3*	67 kDa
Cluster of Chaperone protein dnaK1	Q55154	*dnaK1*	75 kDa
Chaperone protein dnaK1	Q55154	*dnaK1*	75 kDa
Photosystem I P700 chlorophyll a apoprotein A2	P29255	*psaB*	81 kDa
DNA-directed RNA polymerase subunit beta'	P73334	*rpoC2*	145 kDa
Glucose-6-phosphate 1-dehydrogenase	P73411	*zwf*	58 kDa
Catalase-peroxidase	P73911	*katG*	84 kDa

Table 5-A3. List of proteins identified by MS after treating trypsin digest of a 35 kDa gel slice from LRSA labeled *Synechocystis* sp. PCC 6803 lysate.

Protein Name	**Accession Number**	**Gene Name**	**Molecular Weight**
Allophycocyanin beta chain	Q01952	*apcB*	17 kDa
Elongation factor Tu	P74227	*tuf*	44 kDa
ATP-dependent Clp protease regulatory subunit	Q55662	*clpC*	91 kDa
60 kDa chaperonin 1	Q05972	*groL1*	58 kDa
Elongation factor G 2	P74228	*fusB*	75 kDa
Chaperone protein ClpB 2	P74361	*clpB2*	98 kDa
DNA-directed RNA polymerase subunit gamma	P74177	*rpoC1*	71 kDa
C-phycocyanin beta chain	Q54714	*cpcB*	18 kDa
DNA-directed RNA polymerase subunit beta	P77965	*rpoB*	123 kDa
Allophycocyanin alpha chain	Q01951	*apcA*	17 kDa
60 kDa chaperonin 2	P22034	*groL2*	58 kDa
Ribulose bisphosphate carboxylase large chain	P54205	*cbbL*	52 kDa
30S ribosomal protein S1 homolog A	P73530	*rps1A*	37 kDa
Ferredoxin-dependent glutamate synthase 2	P55038	*gltS*	170 kDa
ATP synthase subunit alpha	P27179	*atpA*	54 kDa
Chaperone protein dnaK2	P22358	*dnaK2*	68 kDa
DNA-directed RNA polymerase subunit beta'	P73334	*rpoC2*	145 kDa
Glutamine synthetase	P77961	*glnA*	53 kDa
Chaperone protein ClpB 1	P74459	*clpB1*	101 kDa
ATP synthase subunit beta	P26527	*atpD*	52 kDa
Transaldolase	P72797	*tal*	43 kDa
Enolase	P77972	*eno*	47 kDa
Ribose-phosphate pyrophosphokinase	Q55848	*prs*	36 kDa
Alpha-1,4 glucan phosphorylase	P73546	*glgP*	98 kDa
Magnesium-chelatase subunit ChlI	P51634	*chlI*	39 kDa
Fructose-bisphosphate aldolase class 2	Q55664	*fbaA*	39 kDa
Photosystem I P700 chlorophyll a apoprotein A2	P29255	*psaB*	81 kDa
Carbon dioxide-concentrating mechanism protein CcmK homolog 1	P72760	*ccmK1*	12 kDa
Phycobilisome 32.1 kDa linker polypeptide, phycocyanin-associated, rod 2	P73204	*cpcC2*	31 kDa
Transketolase	P73282	*tktA*	72 kDa
D-alanine--D-alanine ligase	P73632	*ddl*	39 kDa
D-fructose 1,6-bisphosphatase class 2/sedoheptulose 1,7-bisphosphatase	P73922	*slr2094*	37 kDa
30S ribosomal protein S10	P74226	*rpsJ*	12 kDa
Slr0755 protein	P74640	*slr0755*	22 kDa
Isocitrate dehydrogenase [NADP]	P80046	*icd*	52 kDa
10 kDa chaperonin	Q05971	*groS*	11 kDa
ATP-dependent zinc metalloprotease FtsH 4	P73437	*ftsH4*	68 kDa

Table 5-A3(cont'd)

Photosystem II D2 protein	P09192	*psbD*	39 kDa
NAD(P)H-quinone oxidoreductase subunit J	P19125	*ndhJ*	21 kDa
Chorismate synthase	P23353	*aroC*	39 kDa
50S ribosomal protein L1	P36236	*rplA*	26 kDa
30S ribosomal protein S4	P48939	*rpsD*	23 kDa
Heat-inducible transcription repressor HrcA	P72795	*hrcA*	43 kDa
50S ribosomal protein L14	P73310	*rplN*	13 kDa
Acetylglutamate kinase	P73326	*argB*	32 kDa
Sll0867 protein	P73573	*sll0867*	49 kDa
Neopullulanase	P73757	*nplT*	58 kDa
D-3-phosphoglycerate dehydrogenase	P73821	*serA*	59 kDa
Slr1505 protein	P73943	*slr1505*	23 kDa
Probable cytosol aminopeptidase	P73971	*pepA*	52 kDa
30S ribosomal protein S7	P74229	*rpsG*	17 kDa
Alanine--tRNA ligase	P74423	*alaS*	96 kDa
Pyruvate dehydrogenase E1 component subunit alpha	P74490	*pdhA*	38 kDa
Circadian clock protein kinase KaiC	P74646	*kaiC*	58 kDa
Serine hydroxymethyltransferase	P77962	*glyA*	46 kDa
Glutamate--ammonia ligase	P77970	*glnN*	79 kDa
Chaperone protein dnaK1	Q55154	*dnaK1*	75 kDa
Acetyl-coenzyme A synthetase	Q55404	*acsA*	73 kDa
UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase	Q55469	*murE*	55 kDa
Slr0376 protein	Q55590	*slr0376*	13 kDa
Glutamate-1-semialdehyde 2,1-aminomutase	Q55665	*hemL*	46 kDa
Phosphomethylpyrimidine synthase	Q55894	*thiC*	51 kDa
Glyceraldehyde-3-phosphate dehydrogenase 2	P80505	*gap2*	37 kDa
Ssl1918 protein	P73031	*ssl1918*	11 kDa
Chaperone protein HtpG	P74702	*htpG*	75 kDa
CC-adding tRNA nucleotidyltransferase	Q55428	*sll0825*	46 kDa
2-isopropylmalate synthase	P48576	*leuA*	57 kDa
Probable RNA polymerase sigma-C factor	Q59996	*sigC*	47 kDa
Slr6050 protein	Q6YRV4	*slr6050*	129 kDa

CHAPTER 6

Efforts to detect the postulated Ni-hydride intermediate of the NPN cofactor by ^{1}H-NMR spectroscopy

Introduction

Nickel-dependent enzymes, despite being less common than their iron or copper counterparts, play crucial roles in various biological processes. They catalyze reactions essential for metabolism, energy production, and detoxification.[1,2] Notably, some of these enzymes utilize a unique nickel-pincer nucleotide (NPN) cofactor to facilitate catalysis.[3] The NPN cofactor, comprising a pyridinium-based ligand coordinating a nickel ion, has been identified in enzymes like lactate racemase (LarA), which catalyzes the interconversion of L- and D-lactate.[3] A key aspect of nickel-dependent enzymatic catalysis often involves the formation of transient nickel-hydride species,[2,4] and such a species has been proposed to exist during LarA catalysis.[5] Understanding the formation and behavior of nickel-hydride species is crucial for deciphering the catalytic mechanisms of these enzymes.

The inherent instability and short lifetimes of nickel-hydride intermediates pose significant challenges for their detection and characterization.[6] In LarA from *Lactiplantibacillus plantarum*, His200 is a critical active site residue that directly binds the Ni(II) ion, with the 3 other coordination sites provided by the cofactor. It is predicted computationally that His200 dissociates as the hydride binds to the nickel, though it remains energetically unfavorable by 8.4 kcal/mol compared to hydride binding at C4 of the cofactor.[7] Traditional spectroscopic methods may not be sufficiently sensitive to capture these fleeting species. I investigated whether proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy is an attractive analytical technique for detecting a nickel-hydride in LarA.

^{1}H-NMR spectroscopy is a powerful tool for studying metal-hydride complexes. It is particularly well-suited for detecting hydrogen nuclei bound to metals due to their unique magnetic environments.[4,8,9] The chemical shift of a hydride signal can provide valuable information about the electronic structure of the metal center and its coordination environment.

In the context of LarA, the proposed catalytic mechanism involves the formation of a transient nickel-hydride species on the NPN cofactor during the proton-coupled hydride transfer reaction.[7] Detecting this intermediate would provide substantial support for the proposed mechanism and enhance our understanding of LarA's catalytic cycle.

Materials and Methods

Expression and Purification of LarA

The protocol for *larA* expression and protein purification followed literature procedures.[7] LarA from *L. plantarum*, with a C-terminal Strep II tag, was produced in *Lactococcus lactis* strain NZ3900 transformed with plasmid pGIR112, which carries the modified *larA* and *larBCDE* genes under control of the *nisA* promoter. Overnight starter cultures were grown at 30 °C in M17 medium containing 0.5% (w/v) D-glucose and 10 mg/L chloramphenicol. These starter cultures were used to inoculate (1% final volume) larger volumes of the same medium with 5 mg/L chloramphenicol, then grown at 28 °C with mild shaking. At an OD_{600} of 0.2–0.3, cultures were induced with 10 μg/L nisin A, supplemented with 1 mM $NiCl_2$, and grown for an additional 4 h. Cells were harvested by centrifugation and stored at -80 °C.

Thawed cell pellets (5–10 g) were resuspended in 40 mL of 100 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl, 0.5 mM Na_2SO_3 (to stabilize enzyme activity), 1 mM phenylmethylsulfonyl fluoride (from a 100 mM ethanol stock), and 2.5 U/mL benzonase. The cell suspension was lysed by two passes through a French pressure cell at 16,000 psi, prechilled to 4 °C. The lysate was clarified by centrifugation at 27,000 *g* for 40 min at 4 °C.

The clarified lysate was applied at 4 °C to a 2 mL StrepTactin XT resin column (IBA) equilibrated with 10 mM phosphate, pH 7.4, containing 137 mM NaCl and 2.7 mM KCl (PBS). PBS was selected to minimize unwanted proton resonances from the buffer in the NMR spectra. After loading, the column was washed with three column volumes of PBS to remove non-specifically bound proteins. The enzyme was eluted using PBS containing 50 mM biotin.

Eluted fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), confirming that fractions contained bands corresponding to the correct molecular weight for LarA. Protein-containing fractions were pooled and concentrated using an Amicon Ultra Centrifugal Filter (10 kDa MWCO). To remove Na_2SO_3, the protein was washed by serial dilution with sulfite-free buffer, followed by reconcentration, repeating this cycle three times. A sample of protein was reserved to test for lactate racemase activity. LarA racemase activity was quantified by assessing the conversion of L-lactate to D-lactate using a D-lactate-specific lactate dehydrogenase (D-LDH) assay kit (Megazyme) as described previously.[10] The assay leverages the reduction of NAD^+ to NADH, which was monitored by measuring absorbance at 340 nm

NMR Spectroscopy Experiment Setup

The LarA was concentrated to approximately 0.5 mM as measured by absorbance at 280 nm using a Nano Drop UV-vis spectrophotometer. The final sample volume of 600 µL, contained in a 5 mm NMR tube, was sealed under an inert nitrogen atmosphere to prevent oxidative inactivation of the enzyme.

To initiate the catalytic reaction and promote the formation of the nickel-hydride species, L-lactate in D_2O was added to the NMR sample at a final concentration of 10 mM. The reaction was monitored at 4 °C. NMR spectra were acquired using a Bruker Ascend 600 MHz spectrometer. Water suppression was performed using a presaturation technique to minimize the intense solvent peak at δ 4.7 ppm. ^{1}H-NMR spectra were recorded with a total of 1,000 scans. The relaxation delay was set to 2 s.

Results

LarA Activity

Production and purification of LarA in *L. lactis* yielded plentiful (**Figure 6-A1**) amounts of active (**Figure 6-A2**) LarA.

NMR Spectral Analysis

The acquired ^{1}H NMR spectra showed well-resolved signals corresponding to lactate and peaks of much lower intensity corresponding to the protein resonances (**Figure 6-1**). No new signals attributable to a nickel-hydride species were observed in the expected chemical shift region (-5 to −25 ppm), where metal-hydride signals typically appear due to the high shielding environment of the nickel.[9]

Discussion

Approaches to improve chances of detecting the nickel-hydride of NPN in LarA by ^{1}H-NMR spectroscopy

The NMR experiments aimed at detecting the transient nickel-hydride species in LarA did not yield observable signals corresponding to the hydride intermediate. This outcome underscores the challenges inherent in characterizing these short-lived and often low-abundance species. Several factors could contribute to the lack of detection, such as insufficient protein concentration, incomplete cofactor loading, and the inherently transient nature of the hydride – with the lower energy state likely having the hydride located on C4 when the cofactor is in the

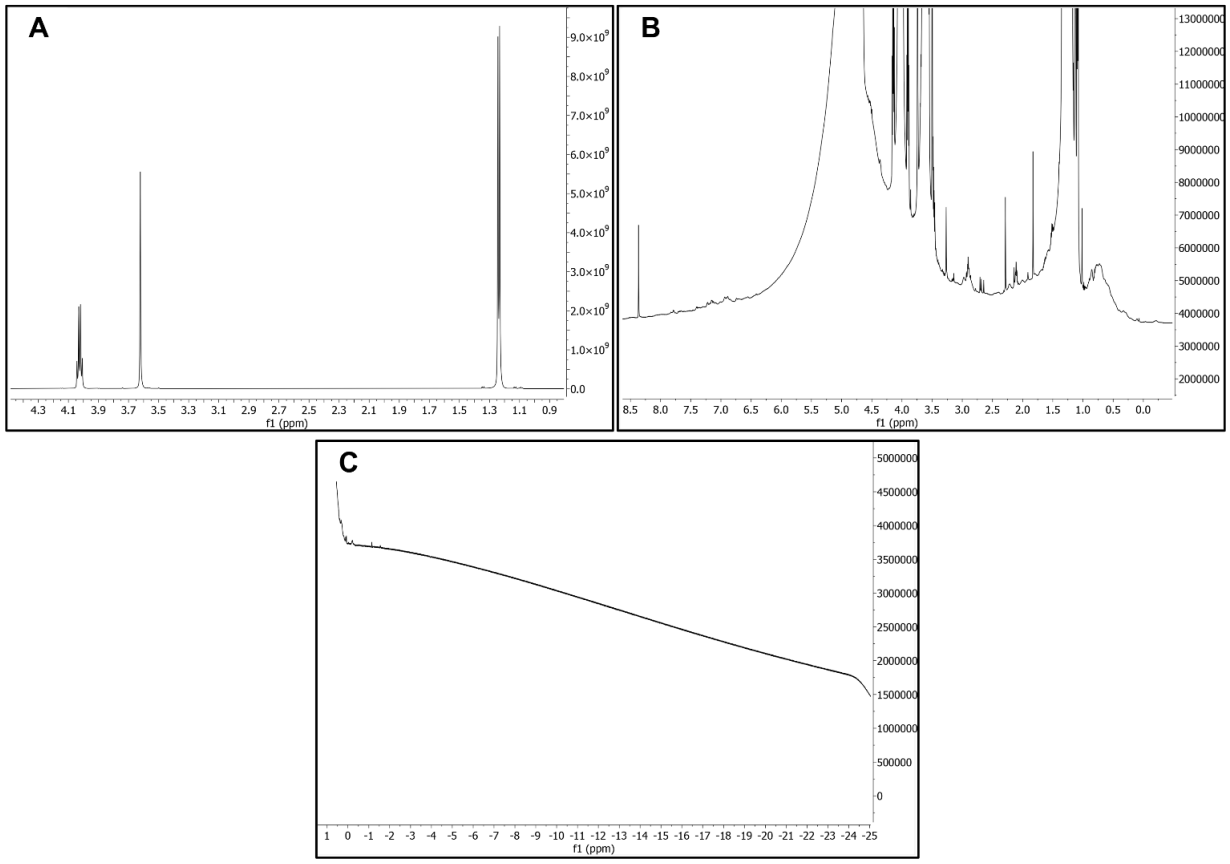

Figure 6-1. ^{1}H-NMR spectra of the reaction containing lactate and LarA from L. plantarum focused on the resonance corresponding to (A) a lactate control, (B) the enzyme, and (C) where the resonance for the nickel bound hydride is expected to appear.

reduced state. To overcome these challenges, a multi-faceted approach encompassing optimized protein production, advanced NMR techniques, and alternative versions of LarA is essential. Increasing protein concentration and ensuring complete cofactor loading are paramount for detecting the intermediate. Optimizing expression and purification protocols to maximize active enzyme yield with fully incorporated cofactor can significantly enhance the signal-to-noise ratio, making the hydride signal more discernible. A notable limitation here is the susceptibility of the NPN cofactor to oxidation in air.[7] Sodium sulfite has been used to prevent oxidation, but it binds tightly to the cofactor and causes inhibition. The Duet system has shown utility in producing high quantities of active LarA in *E. coli* BL21(DE3) as discussed in Chapter 2 of this dissertation. A strain capable of tolerating high nickel concentrations may be beneficial. Additionally, supplementing the growth medium with nicotinic acid has shown some success in increasing cofactor loading.

Using smaller homologs of LarA that retain the critical active site structure is an attractive option to simplify acquisition of the ^{1}H-NMR spectra and improve signal resolution. This approach not only improves resolution but also allows for a higher protein concentration in solution, thereby increasing detection sensitivity. Some LarA homologs may possess lower energy states associated with the nickel-hydride species, thus enhancing the ease of detecting this intermediate.

Lowering the reaction temperature during NMR experiments can slow down the catalytic cycle, potentially extending the lifetime of the hydride species, and thus increasing the likelihood of capturing the hydride signal. This approach, however, is complicated in aqueous solutions due to potential freezing and viscosity issues. These concerns can be partially mitigated by adding agents like glycerol, which lowers the freezing point and enhances solution stability without disrupting the protein's structure, but such supplementation also introduces additional proton resonances.

The use of parahydrogen, a spin isomer of dihydrogen, in conjunction with NMR spectroscopy can significantly boost hydride detection sensitivity.[11] Studies have successfully used parahydrogen-induced polarization (PHIP) to enhance signals in nickel(II) dihydrogen and hydride complexes, potentially enabling the detection of transient hydride species in LarA.[12] Coupling PHIP with substrate precursors made using parahydrogen, such as lactate generated by the activity of a hydrogenase and lactate dehydrogenase with pyruvate, could provide a feasible route to hyperpolarization, amplifying the hydride signal by several orders of magnitude.[13]

Incorporating higher-field NMR spectrometers, with magnetic fields above the 600 MHz used in this study, could significantly enhance spectral resolution and sensitivity. Higher-field NMR instruments improve signal dispersion, especially for complex biomolecules, making it easier to resolve signals from overlapping or transient species. Additionally, using a cryoprobe could further increase sensitivity, as cryogenically cooling the radiofrequency coils and preamplifiers dramatically reduces thermal noise. This enhancement would be particularly beneficial for detecting low-abundance intermediates like the nickel-hydride species in LarA, where signal strength is a limiting factor.

Alternative spectroscopic approaches for detection of the proposed nickel-hydride intermediate of NPN in LarA

Mössbauer spectroscopy is an effective tool for iron-containing systems but has limited applicability for systems with only nickel.[14] Although synchrotron-based ^{61}Ni Mössbauer spectroscopy has been employed to study magnetic properties and local geometries in nickel compounds, these methods are not readily adaptable to biological systems like LarA due to technical limitations and sensitivity issues.[15]

Fourier-transform infrared spectroscopy (FTIR) offers potential for detecting the Ni-hydride of LarA by observing the characteristic Ni-H stretch, which Eberhardt reports occurs between 1690 and 2000 cm^{-1}.[4] This method has been successfully applied in other contexts to identify Ni-H bonds, as cited in the literature.[16–18] However, applying FTIR to enzymes like LarA presents significant challenges, including the complexity of enzyme active sites, interference from overlapping vibrational modes of protein functional groups, and the potential for weak or obscured Ni-H signals within the broader protein matrix. These limitations make the direct observation of the Ni-hydride in enzymatic systems difficult and less straightforward than in simpler, model compounds.

While ^{1}H-NMR spectroscopy has been used for metal hydride detection, directly probing the nickel nucleus using ^{61}Ni-NMR spectroscopy could offer a more direct approach.[19] Loading LarA with enriched ^{61}Ni could enhance sensitivity; however, ^{61}Ni NMR remains challenging due to the nucleus's low natural abundance (1.1%) and inherently low sensitivity, which results in broad lines even in symmetric environments.[20] Although studies have successfully used ^{61}Ni solid-state NMR for diamagnetic nickel(0) complexes, the application of this technique in biological systems like LarA remains technically demanding and limited by sensitivity issues.[20] Computationally predicting the spectra for potential hydride configurations—such as hydride on C4 versus hydride on Ni—could complement experimental efforts by providing a reference for validating any detected signals.

Electron-electron double resonance (EDNMR) has also proven useful in probing nickel centers, as demonstrated in studies on the NiFe hydrogenase enzyme's Ni-B state (Ni(III)).[21] Although Ni(III) states are paramagnetic and easier to detect with EDNMR, it is worth investigating whether this approach might be adapted for transient species like the Ni-Hydride in LarA.

Conclusion

The detection and characterization of the nickel-hydride intermediates in LarA are essential for understanding the enzyme's catalytic mechanism. The initial 1H-NMR experiments did not detect the hydride intermediate, highlighting the challenges associated with transient and low-abundance species in complex biological systems. This outcome underscores the need for a multi-faceted approach that integrates optimized protein production, advanced NMR techniques, and complementary spectroscopic methods.

To improve detection, strategies such as increasing protein concentration, ensuring complete cofactor loading, and utilizing advanced NMR methodologies like PHIP, higher-field spectrometers, and cryoprobes are critical for enhancing sensitivity and resolution. Alternative techniques, including FTIR and ^{61}Ni NMR, offer additional avenues for studying the nickel-hydride, despite their technical challenges in biological systems. Computational predictions can further complement these efforts by providing references for experimental validation.
Future work should focus on integrating these approaches, exploring isotopic labeling, and investigating smaller LarA homologs to overcome detection barriers. By leveraging a synergistic combination of experimental and computational advancements, the characterization of the elusive nickel-hydride species in LarA can be achieved, offering valuable insights into its catalytic mechanism.

REFERENCES

(1) Alfano, M.; Cavazza, C. Structure, Function, and Biosynthesis of Nickel-dependent Enzymes. *Protein Sci.* **2020**, *29* (5), 1071–1089. https://doi.org/10.1002/pro.3836.

(2) Boer, J. L.; Mulrooney, S. B.; Hausinger, R. P. Nickel-Dependent Metalloenzymes. *Arch. Biochem. Biophys.* **2014**, *544*, 142–152. https://doi.org/10.1016/j.abb.2013.09.002.

(3) Desguin, B.; Zhang, T.; Soumillion, P.; Hols, P.; Hu, J.; Hausinger, R. P. A Tethered Niacin-Derived Pincer Complex with a Nickel-Carbon Bond in Lactate Racemase. *Science (80-.).* **2015**, *349* (6243), 66–69. https://doi.org/10.1126/science.aab2272.

(4) Eberhardt, N. A.; Guan, H. Nickel Hydride Complexes. *Chem. Rev.* **2016**, *116* (15), 8373–8426. https://doi.org/10.1021/acs.chemrev.6b00259.

(5) Hausinger, R. P.; Desguin, B.; Fellner, M.; Rankin, J. A.; Hu, J. Nickel–Pincer Nucleotide Cofactor Operates by a Proton-Coupled Hydride Transfer Mechanism. *Curr. Opin. Chem. Biol.* **2018**, *47*, 18–23. https://doi.org/https://doi.org/10.1016/j.cbpa.2018.06.019.

(6) Wang, H.; Yoda, Y.; Ogata, H.; Tanaka, Y.; Lubitz, W. A Strenuous Experimental Journey Searching for Spectroscopic Evidence of a Bridging Nickel-Iron-Hydride in [NiFe] Hydrogenase. *J. Synchrotron Radiat.* **2015**, *22*, 1334–1344. https://doi.org/10.1107/S1600577515017816.

(7) Rankin, J. A.; Mauban, R. C.; Fellner, M.; Desguin, B.; McCracken, J.; Hu, J.; Varganov, S. A.; Hausinger, R. P. Lactate Racemase Nickel-Pincer Cofactor Operates by a Proton-Coupled Hydride Transfer Mechanism. *Biochemistry* **2018**, *57* (23), 3244–3251. https://doi.org/10.1021/acs.biochem.8b00100.

(8) Barton, B. E.; Rauchfuss, T. B. Hydride-Containing Models for the Active Site of the Nickel-Iron Hydrogenases. *J. Am. Chem. Soc.* **2010**, *132* (42), 14877–14885. https://doi.org/10.1021/ja105312p.

(9) Wilson, A. D.; Shoemaker, R. K.; Miedaner, A.; Muckerman, J. T.; DuBois, D. L.; DuBois, M. R. Nature of Hydrogen Interactions with Ni(II) Complexes Containing Cyclic Phosphine Ligands with Pendant Nitrogen Bases. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (17), 6951–6956. https://doi.org/10.1073/pnas.0608928104.

(10) Desguin, B.; Goffin, P.; Viaene, E.; Kleerebezem, M.; Martin-Diaconescu, V.; Maroney, M. J.; Declercq, J. P.; Soumillion, P.; Hols, P. Lactate Racemase Is a Nickel-Dependent Enzyme Activated by a Widespread Maturation System. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4615.

(11) Eisenberg, R. Parahydrogen-Induced Polarization: A New Spin on Reactions with H2. *Acc. Chem. Res.* **1991**, *24* (4), 110–116.

(12) Kireev, N. V.; Kiryutin, A. S.; Pavlov, A. A.; Yurkovskaya, A. V.; Musina, E. I.; Karasik, A. A.; Shubina, E. S.; Ivanov, K. L.; Belkova, N. V. Nickel(II) Dihydrogen and Hydride Complexes as the Intermediates of H2 Heterolytic Splitting by Nickel Diazadiphosphacyclooctane Complexes. *Eur. J. Inorg. Chem.* **2021**, *2021* (41), 4265–4272. https://doi.org/10.1002/ejic.202100489.

(13) Nakada, N.; Okura, I.; Hasumi, F. Hydrogenation of Pyruvic Acid by Hydrogen with the Combination of Hydrogenase and Lactate Dehydrogenase. *J. Mol. Catal.* **1992**, *75* (1),

23–25. https://doi.org/10.1016/0304-5102(92)80097-Z.

(14) Siegbahn, P. E. M.; Chen, S. Theoretical Studies of Nickel-Dependent Enzymes. **2019**, 1–29.

(15) Gee, L. B.; Lin, C. Y.; Jenney, F. E.; Adams, M. W. W.; Yoda, Y.; Masuda, R.; Saito, M.; Kobayashi, Y.; Tamasaku, K.; Lerche, M.; et al. Synchrotron-Based Nickel Mössbauer Spectroscopy. *Inorg. Chem.* **2016**, *55* (14), 6866–6872. https://doi.org/10.1021/acs.inorgchem.5b03004.

(16) Schmidt, D.; Zell, T.; Schaub, T.; Radius, U. Si–H Bond Activation at {(NHC)2Ni0} Leading to Hydrido Silyl and Bis(Silyl) Complexes: A Versatile Tool for Catalytic Si–H/D Exchange{,} Acceptorless Dehydrogenative Coupling of Hydrosilanes{,} and Hydrogenation of Disilanes to Hydrosilanes. *Dalt. Trans.* **2014**, *43* (28), 10816–10827. https://doi.org/10.1039/C4DT01250J.

(17) Barabotti, P.; Diversi, P.; Ingrosso, G.; Lucherini, A.; Nuti, F. Regiospecific Hydride Abstraction from Metallacycles: Conversion of Metallacyclopentanes to Cationic π-Allylic Complexes. *J. Chem. Soc.{,} Dalt. Trans.* **1984**, No. 11, 2517–2523. https://doi.org/10.1039/DT9840002517.

(18) Chen, W.; Shimada, S.; Tanaka, M.; Kobayashi, Y.; Saigo, K. Reaction of [2-(SiH3)C6H4]2SiH2 with Ni(Et2PCH2CH2PEt2)(PEt3)2: Characterization of H2-(Si−H)Ni and NiIV−H Complexes. *J. Am. Chem. Soc.* **2004**, *126* (26), 8072–8073. https://doi.org/10.1021/ja039244k.

(19) Hao, N.; McGlinchey, M. J.; Sayer, B. G.; Schrobilgen, G. J. A 61Ni NMR Study of Some D10 Nickel Complexes. *J. Magn. Reson.* **1982**, *46* (1), 158–162. https://doi.org/10.1016/0022-2364(82)90176-7.

(20) Werhun, P.; Bryce, D. L. Structural and Crystallographic Information from 61Ni Solid-State NMR Spectroscopy: Diamagnetic Nickel Compounds. *Inorg. Chem.* **2017**, *56* (16), 9996–10006. https://doi.org/10.1021/acs.inorgchem.7b01536.

(21) Flores, M.; Agrawal, A. G.; Van Gastel, M.; Gärtner, W.; Lubitz, W. Electron-Electron Double Resonance-Detected NMR to Measure Metal Hyperfine Interactions: 61Ni in the Ni-B State of the [NiFe] Hydrogenase of Desulfovibrio Vulgaris Miyazaki F. *J. Am. Chem. Soc.* **2008**, *130* (8), 2402–2403. https://doi.org/10.1021/ja077976x.

APPENDIX

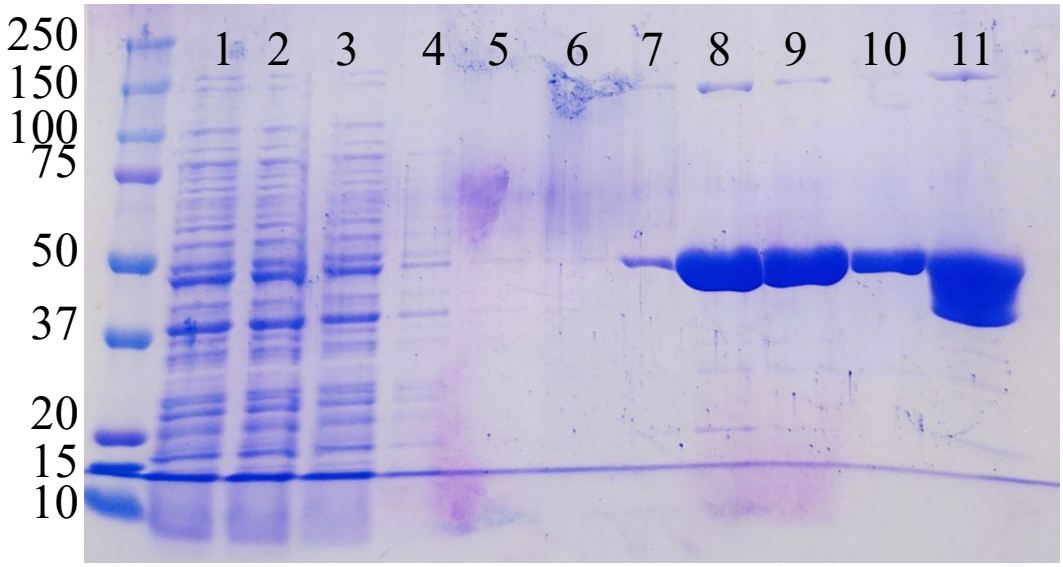

Figure 6-A1. SDS-PAGE analysis of LarA purification using StrepTactin XT resin. Lane 1-3: Flow-through samples; Lane 4: Wash fraction; Lanes 5-10: Elution fractions; 11: combined eluted fractions. A molecular weight ladder is shown in the leftmost lane.

substrate	no L-lactate	no L-lactate	1 mM L-lactate	10 mM L-lactate	100 mM L-lactate	no L-lactate	90 mM L-lactate
protein	buffer	purified LarA	purified LarA	purified LarA	purified LarA	pGIR112 lysate	pGIR112 lysate
ABS 340 nm	0.388	0.414	0.446	0.535	1.159	0.417	0.761
	0.387	0.415	0.43	0.455	1.211	0.406	0.755

Figure 6-A2. Absorbance at 340 nm corresponding to NADH production by lactate dehydrogenase following the reduction of NAD^+ by D-lactate. D-lactate was generated through LarA-catalyzed racemization of varying concentrations of L-lactate. Duplicate assays were performed using either purified LarA or cell lysate. Higher absorbance values at 340 nm indicate increased formation of D-lactate by LarA, as detected by D-lactate-specific lactate dehydrogenase.

CHAPTER 7

Conclusions and future studies

Conclusions from my studies

This dissertation represents a comprehensive exploration of the nickel-pincer nucleotide (NPN) cofactor, shedding light on its intricate biosynthesis, functional role in enzymatic catalysis, and the challenges associated with its detection and study. By combining biochemical, structural, and methodological approaches, the work presented here bridges critical gaps in our understanding of this unique organometallic cofactor and sets the stage for future advancements in microbial metabolism and enzyme engineering.

An introduction of the NPN

Chapter 1 contextualized the NPN cofactor as a central organometallic structure in lactate racemase (LarA). By outlining the biosynthetic pathway—spanning LarB, LarC, and LarE enzymes[1]—it established the foundation for this dissertation's inquiries. The introduction highlighted the cofactor's unique chemical properties and catalytic mechanism, the proton-coupled hydride transfer (PCHT),[2] while also identifying gaps in understanding its broader role across microbial taxa and its evolutionary significance. Future research should continue to explore the NPN cofactor's activity in homologs with diverse substrates and investigate structurally distant LarA homologs to uncover novel catalytic mechanisms and expand our understanding of their functional diversity.

Tools to advance the study of LarA and LarA homologs

To address the challenges of studying LarA and its homologs, this chapter focused on the development of innovative tools and assays for characterizing these systems. A circular dichroism assay was optimized as a broadly applicable method for monitoring racemization and epimerization reactions of LarA homologs with diverse 2-hydroxyacid substrates. By leveraging changes in molar ellipticity, this assay provided a real-time, coupling-enzyme-free approach, offering significant advantages in speed, simplicity, and versatility. Additionally, Lissamine Rhodamine B Sulfonyl Azide labeling was employed to selectively detect NPN-bound proteins via click chemistry, successfully demonstrating the presence of the cofactor in expressed proteins. The study also confirmed that LarA activity could be achieved using LarB, LarC, and LarE homologs from *Synechocystis* sp. PCC 6803, underscoring the functional conservation of these enzymes in Cyanobacteria who lack a gene encoding a LarA homolog. Future work should expand the application of these methodologies to a broader range of LarA homologs and apply labeling techniques to a diverse set of microbial extracts. These tools provide a strong foundation

for systematically exploring the biochemical properties and functional diversity of LarA and other potential NPN-containing proteins.

LarB's dicarboxylated, dinucleotide intermediate

The identification of dicarboxynicotinic acid adenine dinucleotide (DaAD) as a reaction intermediate in LarB-catalyzed processes provided a significant breakthrough in understanding NPN cofactor biosynthesis.[3] Structural studies revealed detailed insights into LarB's catalytic mechanism, including the pivotal role of substrate activation and hydrolysis in forming intermediates. By employing the S127A variant, which exhibits reduced hydrolytic activity, the study successfully trapped and characterized intermediates, enabling a deeper understanding of this enzyme's function. Complementary absorbance assays further validated these findings, underscoring the importance of combining structural and spectroscopic approaches to capture fleeting reaction states and to elucidate the mechanistic details of biosynthetic pathways. Future research should undergo comparative studies of LarB homologs from diverse species that could uncover variations in catalytic strategies and reveal new facets of NPN biosynthesis that may further our understanding of these unique biological processes.

LarE regeneration

Further research should delve into the biochemical roles of sulfur transfer proteins associated with LarE across diverse microbial systems, with a particular focus on the mechanism of cysteine regeneration from dehydroalanine (Dha). Prior studies demonstrated that LarE can be regenerated in vitro using the persulfide of coenzyme A (CoA),[4] but the physiological relevance of this process remains uncertain. Understanding the natural mechanism for LarE regeneration is critical for elucidating the full biosynthetic pathway of the NPN cofactor. Future work should investigate neighboring gene clusters or coexpressed genes that may encode proteins involved in this regeneration process, potentially revealing the physiological partners of LarE.

Efforts to selectively isolate NPN-bound proteins using sulfonyl azide resins provided proof-of-concept for the targeted identification of NPN cofactor-bound enzymes. This approach, integrated with bioinformatics, was applied to identify potential NPN-utilizing proteins in Synechocystis sp. PCC 6803; however, no definitive candidates were identified. Future work should optimize the resin's specificity and sensitivity. Coupling this method with advanced mass spectrometry could enhance its application in proteomics, enabling the discovery of novel NPN cofactor-utilizing enzymes and pathways.

Ni-hydride of the NPN

Efforts to detect the postulated Ni-hydride intermediate of the NPN cofactor using ^{1}H-NMR spectroscopy highlighted the significant challenges in capturing transient species within biocatalytic cycles. While these attempts were unsuccessful, they revealed opportunities for alternative strategies, such as enhancing sensitivity through parahydrogen-induced polarization or investigating smaller LarA homologs that may allow for more concentrated samples. Future studies should pair these approaches with cutting-edge NMR technology and computational modeling to better understand the intermediate's dynamics and refine existing mechanistic hypotheses.

Final Remarks and Future Perspectives

This dissertation has significantly advanced our understanding of the biosynthesis, function, and detection of the NPN cofactor, the unique organometallic pincer complex at the heart of the lactate racemase enzyme. Through the development of innovative tools and methodologies, this work has provided a robust platform for studying NPN-dependent systems. Investigations into Lar enzymes have further clarified the biosynthetic pathway, revealing key intermediates and mechanistic insights that enhance our understanding of how this complex cofactor is assembled and utilized. Additionally, efforts to elucidate the detailed catalytic mechanism of the NPN cofactor, including the pursuit of transient intermediates like the Ni-hydride species, have deepened our knowledge of its role in enzymatic processes.

Future research should aim to build on these findings by further exploring the functional diversity of NPN-dependent enzymes. A key objective will be to harness the unique chemistry of the NPN cofactor for novel biocatalytic applications, leveraging its distinct catalytic capabilities for processes such as stereoselective transformations or other challenging chemical conversions. This could open new avenues for engineering NPN-utilizing systems tailored for specific reactions, providing a foundation for translating these discoveries into industrial processes. By continuing to explore the unique properties and mechanisms of the NPN cofactor, future research can provide valuable insights into its broader biological roles and its potential applications in biocatalysis. These ongoing efforts will not only deepen our understanding of this remarkable cofactor but may also pave the way for its use in innovative and practical applications.

REFERENCES

(1) Desguin, B.; Goffin, P.; Viaene, E.; Kleerebezem, M.; Martin-Diaconescu, V.; Maroney, M. J.; Declercq, J. P.; Soumillion, P.; Hols, P. Lactate Racemase Is a Nickel-Dependent Enzyme Activated by a Widespread Maturation System. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4615.

(2) Hausinger, R. P.; Desguin, B.; Fellner, M.; Rankin, J. A.; Hu, J. Nickel–Pincer Nucleotide Cofactor Operates by a Proton-Coupled Hydride Transfer Mechanism. *Curr. Opin. Chem. Biol.* **2018**, *47*, 18–23. https://doi.org/https://doi.org/10.1016/j.cbpa.2018.06.019.

(3) Chatterjee, S.; Nevarez, J. L.; Rankin, J. A.; Hu, J.; Hausinger, R. P. Structure of the LarB-Substrate Complex and Identification of a Reaction Intermediate during Nickel-Pincer Nucleotide Cofactor Biosynthesis. *Biochemistry* **2023**, *62* (21), 3096–3104. https://doi.org/10.1021/acs.biochem.3c00242.

(4) Fellner, M.; Rankin, J. A.; Desguin, B.; Hu, J.; Hausinger, R. P. Analysis of the Active Site Cysteine Residue of the Sacrificial Sulfur Insertase LarE from Lactobacillus Plantarum. *Biochemistry* **2018**, *57* (38), 5513–5523. https://doi.org/10.1021/acs.biochem.8b00601.