NICKEL-PINCER NUCLEOTIDE COFACTOR: ITS FUNCTION AND SYNTHESIS

By

Joel Alexander Rankin

A DISSERTATION

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

Biochemistry and Molecular Biology—Doctor of Philosophy

2021

ABSTRACT

NICKEL-PINCER NUCLEOTIDE COFACTOR: ITS FUNCTION AND SYNTHESIS

By

Joel Alexander Rankin

The nickel-pincer nucleotide (NPN) is an organometallic cofactor that features Ni(II) tri-coordinated to a pyridinium mononucleotide in a planer configuration using two thioacid sulfur atoms and a covalent bond to a ring carbon atom. This arrangement, where a metal ion is held in a rigid planar conformation by two ligands attached to an aromatized ring and by an atom within the ring, is termed a pincer complex. Pincer complexes are well known in synthetic organic chemistry; however, NPN is one of only two known examples of such a complex in biology. NPN was first reported in 2015 in *Lactobacillus plantarum* lactate racemase (LarA), an enzyme which interconverts the two stereoisomers of lactate. This organism requires three gene products (LarB, LarE, and LarC) to synthesize the cofactor from nicotinic acid adenine dinucleotide (NaAD), typically considered a precursor to NAD.

Computational analyses had generated competing proposals for the reaction mechanism of lactate racemization. Through biochemical analysis of LarA, I demonstrated that it operates by a proton-coupled hydride transfer mechanism during which lactate is transiently oxidized to pyruvate as a hydride transfers onto NPN. Hydride return from NPN can occur to either face of the planar and pro-chiral pyruvate, restoring lactate while achieving racemization.

The first synthetic step in NPN synthesis is catalyzed by LarB, an NaAD carboxylase/hydrolase. LarB carboxylates C5 of the nicotinic acid moiety of NaAD and hydrolyzes the phosphoanhydride, releasing AMP and the product, pyridinium-3,5-biscarboxylic acid mononucleotide. I was able to crystallize and structurally characterize LarB and its

complexes with NAD^+ and AMP. My structural and biochemical experiments demonstrate that LarB utilizes CO_2 for carboxylation and my work suggests that the enzyme activates the substrate by forming a covalent bond between a cysteine residue and C4 of the nicotinic acid. The resulting dihydropyridine then undergoes carboxylation by nucleophilic addition to CO_2 followed by deprotonation and rearomatization with expulsion of the cysteine.

The second step in the synthesis pathway utilizes LarE, a sulfur insertase that uses ATP to activate the substrate by adenylylation then sacrificially donates a sulfur atom originating from a cysteine residue, leaving behind a dehydroalanine (Dha) residue in the protein. Two cycles of this process are needed to create pyridinium-3,5-bisthiocarboxylic acid, which is then metalated by LarC to form NPN. The requirement for two equivalents of the 31 kDa LarE protein for the synthesis of a single molecule of NPN represents an enormous energy investment by the cell. My work has suggested a route for regeneration of the cysteine residue from Dha. I have shown that CoA persulfide robustly reacts with Dha in the inactivated LarE, forming the LarE Cys-CoA form of the enzyme, and subsequent reduction restores the native form of LarE. Notably, LarE has a high affinity for CoA and the LarE Cys-CoA disulfide is known to be an active form of the enzyme *in vitro* when exogenous CoA is provided at physiologically relevant concentrations. Nevertheless, the *in vivo* relevance of this reaction is yet to be elucidated.

ACKNOWLEDGEMENTS

I first must thank my advisor and mentor, Robert Hausinger, who guided me on this journey to become a mature scientist. He has been supportive, patient and kind. I have had the pleasure of working with many fantastic colleagues. Tina Müller and Salette Martinez were immensely helpful when I first started in Hausinger Lab. Shramana Chatterjee, Mark Farrugia, Santhosh Gatreddi, Caitlyn Herr, Jorge Nevarez and Aiko Turmo, besides great scientific colleagues, are also a joy to work with. Stephanie Iorga, an undergraduate, was a great aide in many of my projects. I thank our collaborators Jian Hu and his lab, especially Matthias Fellner, who taught me the ropes of crystallography. Also, Benoît Desguin, a former postdoc in the lab and now collaborator. My committee members, Eric Hegg, Jian Hu, Erik Martinez-Hackart and Claire Vieille have been helpful and I thank them for their time. I thank my undergraduate mentor, Robert Kranz. Finally, I must thank my friends and family for being my support on this endeavor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

KEY TO ABBREVIATIONS

2-HIB—2-hydroxyisobutyric acid

AIR—aminoimidazole ribonucleotide

APAD—3-acetylpyridine dinucleotide

CAIR—carboxyaminoimidazole ribonucleotide

CAT—chloramphenicol acetyltransferase

CE—capillary electrophoresis

DFT—density functional theory

Dha—dehydroalanine

DM35PC—dimethyl pyridine-3,5-bicarboxylic acid

DNPH—2,4-dinitrophenylhydrazine

DTNB—5,5′-dithiobis(2-nitrobenzoic acid)

DTT—dithiothreitol

ESI—electrospray ionization

HRP—horseradish peroxidase

ICP—inductively coupled plasma

IDA—iminodiacetic acid

IPTG—isopropyl β-D-1-thiogalactopyranoside

LC—liquid chromatography

LDH—lactate dehydrogenase

MBP—maltose binding protein

MS—mass spectrometry

NaAD—nicotinic acid adenine dinucleotide

NPN—nickel-pincer nucleotide

NTA—nitrilotriacetic acid

OES—optical emission spectroscopy

P2CMN—pyridinium 3,5-biscarboxylic acid mononucleotide

P2TMN—pyridinium-3,5-bisthiocarboxylic acid mononucleotide

PCET—proton-coupled electron transfer

PCHT—proton-coupled hydride transfer

PMSF—phenylmethanesulfonyl fluoride

TCEP—tris(2-carboxyethyl)phosphine

XAS—X-ray absorption spectroscopy

CHAPTER 1

Introduction: The nickel-pincer nucleotide is an organometallic cofactor of lactate racemase that is activated by a novel maturation pathway

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

Hausinger, R. P., Desguin, B., Fellner, M., Rankin, J. A., and Hu, J. (2018) Nickel-pincer nucleotide cofactor, *Curr. Opin. Chem. Biol.* 47, 18-23.

Introduction

In 2015, a new organometallic cofactor was identified in lactate racemase (LarA of *Lactobacillus plantarum*), which interconverts the two stereoisomers of lactic acid (Figure 1-1).[1] This coenzyme was identified by mass spectrometry and X-ray crystallography and shown to contain nickel bound to pyridinium-3,5-dithiocarboxylic acid mononucleotide (P2TMN) (Figure 1-2). Inorganic chemists have synthesized many similar-appearing planar organic ligands that tri-coordinate metal ions and termed these compounds pincer complexes.[2, 3] The lactate racemase cofactor, henceforth called the nickel–pincer nucleotide (NPN), is the first biological example of such a molecule.

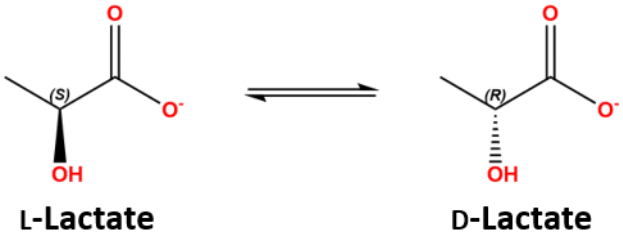


Figure 1-1. The reaction catalyzed by lactate racemase.

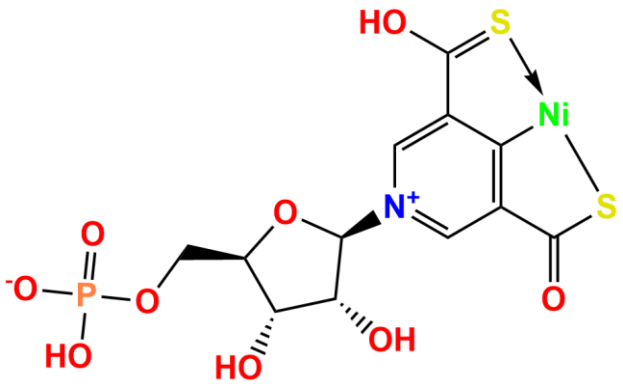


Figure 1-2. The nickel-pincer nucleotide cofactor (NPN). The NPN found in the LarA of *L. plantarum* features a thioamide bond to a lysine residue.

Early work on lactate racemase

The first demonstration of lactate racemase activity in a biological system was reported in 1936 in *Clostridium beijerinckii* (formerly *Clostridium butylicum*) and *Clostridium acetobutylicum*.[4] A follow-up study reported that the activity was unlikely to be a result of cellular lactate dehydrogenases (Ldh) on the basis of the stability of the enzyme and its inhibition by cyanide.[5] In 1952, Kitahara, *et al.* reported a crude, dehydrated cell-free extract of *L. plantarum* that possessed lactate racemase activity.[6] Importantly, this extract did not have

detectable L- or D-Ldh activity, demonstrating that the racemase activity does not arise from the sequential actions of those enzymes *via* soluble pyruvate.

Subsequent investigations of lactate racemase from diverse microorganisms began to characterize the mechanism of the reaction and to examine whether the enzyme possessed any necessary cofactors. In 1955, niacin auxotrophs of *L. plantarum* and other tested *Lactobacillus* were found to produce D-lactate in excess when nicotinic acid was limiting, *i.e.*, lactate was not fully racemized.[7] This result held when activity was assayed *in vitro* from cell lysates, suggesting that lactate racemase is an NADH-dependent enzyme.[8, 9] However, a later report from this group claimed that the racemase activity of *L. plantarum* may in fact result from the cooperation of NADH-dependent L- and D-Ldh, partly on the basis that the kinetics of L-Ldh and the supposed lactate racemase were very similar. These investigators speculated that a flavin-dependent D-Ldh, overlooked in their previous studies, led to the earlier conclusion that lactate racemase was a discrete enzyme.[10]

Lactate racemase from *C. acetobutylicum* was reported to have a dialyzable cofactor that was identified as pyridoxamine phosphate on the basis of its UV absorption band in the partly refined preparation.[11] This report also indicated that activity was recovered in dialyzed sample by adding ferrous ions and pyridoxamine phosphate to apoenzyme.

The racemase from *Lactobacillus sakei* (formerly *Lactobacillus sake*) was purified to near homogeneity and determined to be about 25 kDa in size.[12] The enzyme activity was not stimulated by supplementation with common organic cofactors; however, the protein was reported to contain two equivalents of iron. This group demonstrated a lack of detectable hydrogen exchange with the solvent during the course of the reaction. They also showed that added pyruvate was not converted to lactate by this enzyme.[12]

The most detailed early mechanistic work used lactate racemase from *C. beijerinckii*. The enzyme activity was not increased in the presence of iron or pyridoxamine phosphate (though the authors noted that glutathione in their assay may have bound some of the metal).[13] They demonstrated there is no hydrogen exchange with the solvent and determined, using 2-^{2}H-lactate, a normal substrate kinetic isotope effect of k_H/k_D = 2.14 and 2.16 for the L- and D- isomers, respectively. Based on these findings, the authors proposed an internal-hydride shift mechanism, where the α-hydrogen as a hydride shifts to the acid carbon to form a methyl glyoxal-like intermediate (possibly with a covalent linkage to the enzyme). The now symmetrical α-carbonyl would be prochiral, so the hydride could return to either side of the carbonyl achieving racemization of the compound.[14] Later studies lent support to this hypothesis: an ^{18}O label in 2-^{18}O-lactate was preserved during the reaction,[15] and the intermediate was shown to be a bound carbonyl.[16]

Recent work on LarA

Although LarA activity had been characterized in different organisms over several decades, little real progress had been made toward understanding the reaction mechanism underpinning the racemization.[8] A breakthrough was reported in 2014, when LarA from *L. plantarum* and *Thermoanaerobacterium thermosaccharolyticum* (both expressed recombinantly in *Lactococcus lactis* that natively lacks LarA activity) were found to contain nickel, with the nickel content and enzyme activity dependent on three gene products (LarB, LarC, and LarE) encoded in the same operon as *larA*. When the three accessory proteins were co-synthesized *in vivo*, only LarE was necessary for *in vitro* activation of LarA apoprotein, suggesting that the necessary Ni-containing cofactor was assembled on LarE.[17] A year later, the nature of this molecule was elucidated. It was found to be an organometallic cofactor featuring a square-planar

Ni center coordinated to two sulfur atoms and possessing an unusual Ni-carbon bond, the NPN cofactor (Figure 1-3).[1]

Role of the NPN cofactor in lactate racemase

When first identified, the LarA NPN cofactor was proposed to participate in a proton-coupled hydride transfer (PCHT) reaction in which (i) the D- or L-lactate hydroxyl groups are deprotonated by H108 or H174 (Figure 1-3a), (ii) pyruvate forms as hydride transfers to the cofactor, (iii) the same hydride returns to either face of pyruvate (known since 1965[14]), and (iv) the D- or L-alkoxide products are reprotonated.[1] Hydride was depicted as transferring to C4 of the pyridinium, akin to standard nicotinamide-type chemistry, but the alternative possibility of a nickel–hydride intermediate was mentioned.

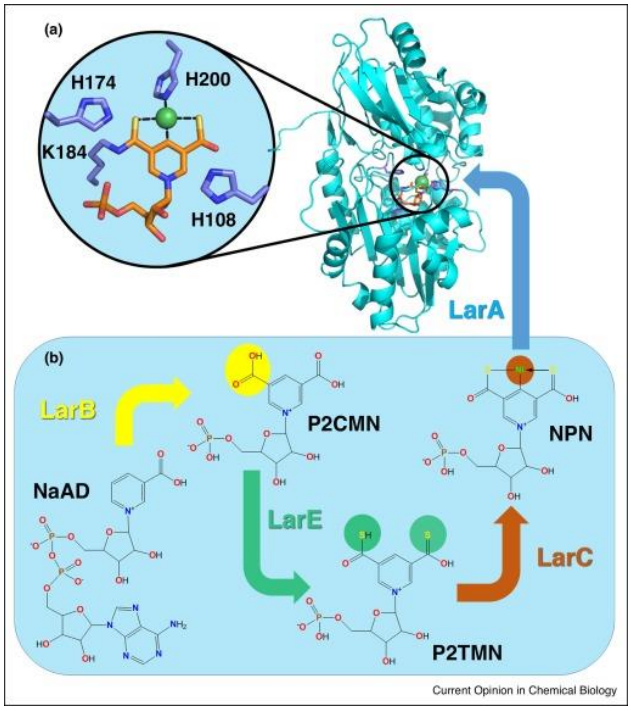


Figure 1-3. The lactate racemase cofactor and its synthesis. (a) The NPN cofactor forms an enzyme adduct with *L. plantarum* LarA by covalent attachment to K184. The square-planar nickel is tri-coordinated by the pincer ligand using two sulfurs and one carbon atom, with H200 also serving as a metal ligand. The well-positioned residues H108 and H174 likely facilitate catalysis. (b) The overall biosynthetic pathway for the NPN cofactor, starting with NaAD carboxylation/hydrolysis by LarB, sulfur insertion by LarE, and nickel insertion by LarC.

Synthetic models and computational studies have since provided important insights into the LarA mechanism. For example, a synthetic nickel-dithioamidepyridinium species resembling the NPN cofactor was shown to catalyze the dehydrogenation of alcohols, and accompanying density functional theory (DFT) calculations indicated hydride transfer to C4 of the pyridinium.[18] The authors did not observe racemization due to instability of the reduced pincer complex. Other groups carried out DFT calculations on 139-atom[19] or 200-atom[20] models of the LarA active site.

In both cases, the authors favored PCHT to the pyridinium C4 atom and ruled out hydride transfer to nickel. Calculations by Zhang and Chung indicated the energy barriers for the NPN cofactor were greater than those for the Ni-free species or for NAD^+, and these authors suggested the tethered, nickel-bound cofactor destabilizes the intermediate thus promoting racemization.[19] By contrast, Yu and Chen indicated the NPN cofactor exhibits greater hydride addition reactivity than NAD^+ for environments with medium to high polarity.[20] Qiu and Wang computationally predicted the racemization activities of scorpion-like nickel–pincer compounds with an appended imidazole group and found similar free energy barriers for racemization.[21]

In contrast to these DFT-based proposals suggesting a PCHT mechanism, Wang and Shaik used quantum mechanical/molecular mechanical calculations and proposed a proton-coupled electron transfer (PCET) mechanism.[22] These authors posited a resting enzyme with a Ni(III) redox state that is reduced to Ni(II) during the reaction. Furthermore, they invoked the transient cleavage of the C1–C2 bond of lactate during the reaction, yielding a CO_2 anion radical and acetaldehyde. Rotation of the acetaldehyde and reversal of the preceding steps provided for lactate racemization.

Biosynthesis of the NPN cofactor

The overall pathway for synthesis of the NPN cofactor (Figure 1-3b) was established using the components from *L. plantarum* expressed in *L. lactis*.[23] Biosynthesis begins with nicotinic acid adenine dinucleotide (NaAD), which undergoes C5 carboxylation and phosphoanhydride hydrolysis by LarB. Two sulfur atoms are incorporated into the resulting pyridininium-3,5-biscarboxylic acid mononucleotide (P2CMN) by LarE, forming P2TMN. LarC then inserts nickel, forming two S–Ni bonds and the unusual C-Ni bond. The NPN cofactor is covalently tethered to K184 of the *L. plantarum* lactate racemase enzyme as a thioamide;

however, the cofactor does not form an enzyme adduct with LarA from *T. thermosaccharolyticum*.[23]

LarB

LarB converts NaAD to P2CMN and AMP (Figure 1-4). This carboxylase/hydrolase is homologous to PurE, 5-(carboxyamino)imidazole ribonucleotide (N^5-CAIR) mutase, that catalyzes carboxyl transfer from the amine of aminoimidazole ribonucleotide (AIR) to its imidazole C4 (forming CAIR) during bacterial purine nucleotide synthesis (Figure 1-5).[24] PurE's partner enzyme, the carboxylase PurK, uses ATP hydrolysis to form N^5-CAIR. It was unclear whether LarB requires the energy from phosphoanhydride hydrolysis of NaAD to drive its carboxylation reaction. Kinetic analysis of LarB was complicated by the production of nicotinic acid mononucleotide and AMP, in what is presumed to be a side reaction. AMP release is independent of bicarbonate concentration, suggesting that hydrolysis occurs before carboxylation. When assayed by its ability to activate LarA, LarB exhibits a K_m of 33 ± 18 mM for bicarbonate. This is comparable to the K_m of 18.8 ± 3.9 mM for bicarbonate using PurK from *Staphylococcus aureus*.[25]

It was not known whether bicarbonate or CO_2 is the substrate for LarB. The PurE reaction is known to react CO_2 (generated by decarboxylation of N^5-CAIR) with the AIR

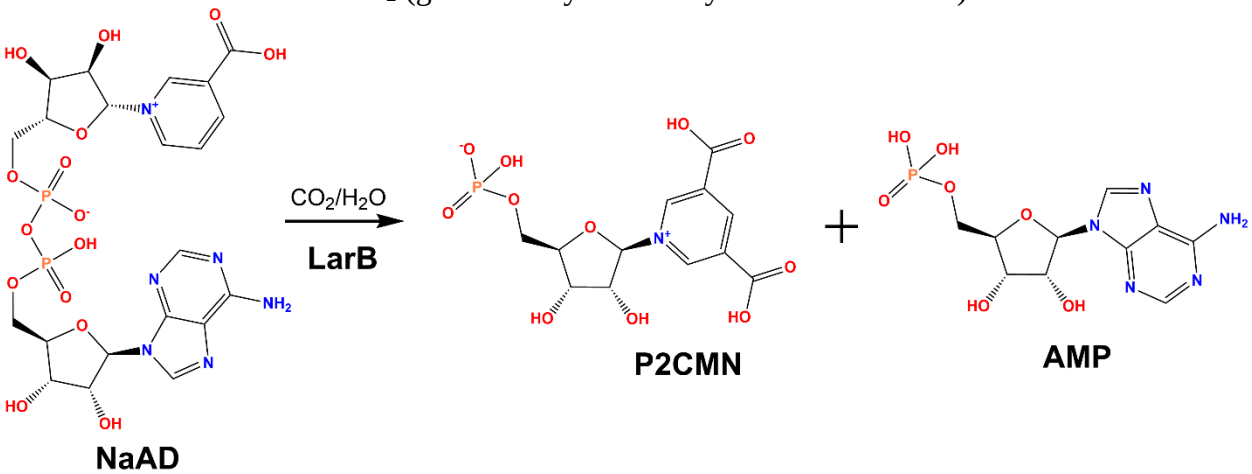


Figure 1-4. The P2CMN synthesis reaction catalyzed by LarB.

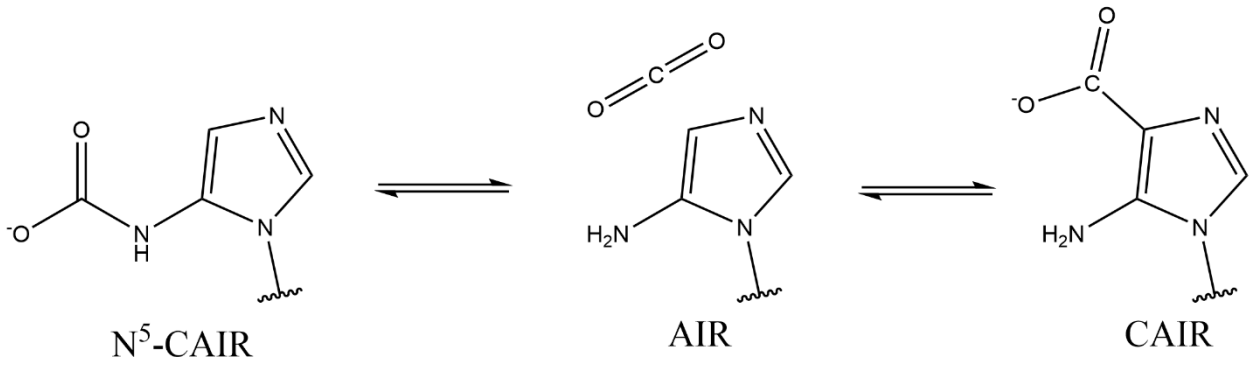


Figure 1-5.The reaction catalyzed by PurE (N^5-CAIR mutase).

intermediate, likely by C4 nucleophilic attack.[26] In contrast, C5 of the LarB substrate, NaAD, is not nucleophilic, possessing a formal positive charge and lacking the amino lone pair thought to play a role in the PurE mechanism. Furthermore, LarB does not require any cofactor or substrate commonly used by other carboxylases. Magnesium was the only identified cofactor necessary for LarB activity;[23] this metal seems more likely to play a role in nucleotide binding and hydrolysis than in carboxylation.

LarE

LarE has been identified as a sulfur transferase that converts P2CMN into P2TMN.[23] The inserted sulfur originates from C176 of LarE, rather than from typical sulfur sources such as free cysteine or sulfide. Structural characterization (Figure 1-6) identified LarE as a new member of the N-type pyrophosphatase family.[27] A conserved CSxxSR motif located in the C-terminal domain binds the phosphate moiety of P2CMN, positioning the substrate to react with Mg·ATP bound at the conserved PP-loop SGGxDS in the N-terminal pyrophosphatase domain. C176, located on an intervening loop,

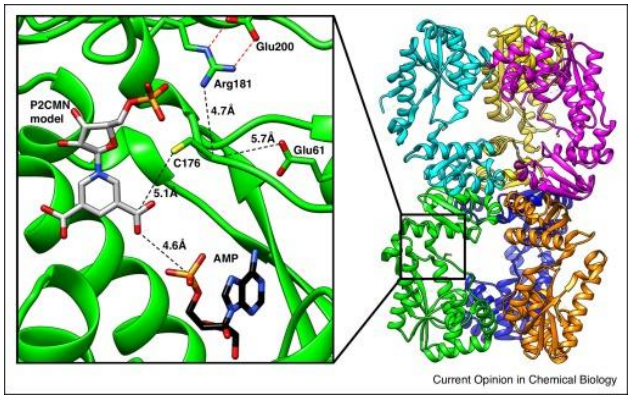


Figure 1-6. Structure of NPN cofactor biosynthetic enzyme LarE. Ribbon diagram of the hexameric protein (right) with an expanded view of one active site containing bound AMP and a model of bound P2CMN (left).

attacks the adenylylated substrate to release AMP and form the LarE-P2CMN adduct. Sacrificial sulfur transfer, possibly facilitated by a base-catalyzed β-elimination involving a nearby R181-E200 pair, leaves LarE with a dehydroalanine. The nearby E61 residue does not function as the general base according to mutagenesis studies. A conformation change not observed in the crystal structures cannot be excluded as a means to position another residue to facilitate the reaction.

LarE inserts two sulfur atoms into P2CMN, thus two monomers are required for P2TMN formation. The hexameric state observed in solution and in the crystal[27] does not reveal how an intermediate is transferred from one buried binding site to a neighboring site. Thus, the hexamer may simply enhance the local LarE concentration to facilitate recapture of the product of the first sulfur transfer.

Bioinformatic studies suggest that ~25% of all bacterial genomes contain a homolog of *larE*.[17, 28] It is costly to the cell if LarE catalyzes sacrificial sulfur transfer and undergoes degradation after a single turnover as in the *L. plantarum* protein, so a regeneration system may exist to add a sulfur atom to the dehydroalanine, a possibility that has not been explored. On the other hand, there are precedents for sacrificial protein substrates. For example, the 37 kDa THI4p from *Saccharomyces cerevisiae*, part of the thiamine biosynthesis pathway, has also been identified as a sacrificial sulfur insertase.[29] Alternatively, other LarE homologs might use a different sulfur transfer system. Most LarEs contain three cysteine residues in the vicinity of the active site, similar to the tRNA thiouridine synthetase N-type pyrophosphatase family member TtuA that transfers sulfur from a [4Fe4S] cluster.[30, 31] It is speculated that the majority of LarE homologs catalyze a cluster-based sulfur insertion.

In this thesis, I will fill in several gaps in our knowledge of lactate racemase and the biosynthesis of its NPN cofactor. In Chapter 2, I describe UV-visible and electron paramagnetic resonance spectroscopy, intermediate trapping, and substrate kinetic isotope effect studies to define the mechanism of *L. plantarum* lactate racemase. In Chapter 3, I detail the structural elucidation of LarB along with biochemical and mutagenic studies to establish its mechanism. In Chapter 4, I relate studies that demonstrate the ability to recycle the dehydroalanine form of LarE to the native enzyme form, and I describe results that explain the anomalous prior findings that only LarE is required for activation of LarA apoprotein. Finally, in Chapter 5 I highlight my major conclusions, summarize the results of a few other studies, and provide a perspective on remaining questions related to these topics.

REFERENCES

REFERENCES

(1) 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.

(2) van Koten, G., and Milstein, D. (2013) *Organometallic pincer chemistry*, Springer, Berlin.

(3) Szabó, K. J., and Wendt, O. F. (2014) *Pincer and Pincer-Type Complexes: Applications in Organic Synthesis and Catalysis*, John Wiley & Sons.

(4) Tatum, E. L., Peterson, W. H., and Fred, E. B. (1936) CCLXVI. Enzymic racemization of optically active lactic acid, *Biochem. J. 30*, 1892-1897.

(5) Christensen, W. B., Johnson, M. J., and Peterson, W. H. (1939) Properties of the lactic acid-racemizing enzyme of *Clostridium Butylicum*, *J. Biol. Chem. 127*, 421-430.

(6) Kitahara, K., Obayashi, A., and Fukui, S. (1952) Racemiase. Part I. Cell-free racemiase, *Enzymologia 15*, 259-266.

(7) Kitahara, K., and Obayashi, A. (1955) DL-Forming Lactic Acid Bacteria, *J. Gen. Appl. Microbiol. 1*, 237-245.

(8) Desguin, B., Soumillion, P., Hausinger, R. P., and Hols, P. (2017) Unexpected complexity in the lactate racemization system of lactic acid bacteria, *FEMS Microbiol. Rev. 41*, S71-S83.

(9) Kitahara, K., and Obayashi, A. (1959) Studies on the enzymes of lactic acid bacteria. Part XI, *Nippon Nōgeik Kaishi 33*, 497-499.

(10) Hiyama, T., Mizushima, S., and Kitahara, K. (1965) A racemizing enzyme system of *Lactobacillus Plantarum*, *J. Gen. Appl. Microbiol. 11*, 51-60.

(11) Katagiri, H., Sugimori, T., Imai, K., Komaki, S., Okuzumi, M., and Anyoji, H. (1960) On the metabolism of organic acids by *Clostridium acetobutylicum*. Part III. Purification and properties of racemiase., *Agric. Biol. Chem. 25*, 281-289.

(12) Hiyama, T., Fukui, S., and Kitahara, K. (1968) Purification and properties of lactate racemase from *Lactobacillus sake*, *J. Biochem. 64*, 99-107.

(13) Dennis, D., and Kaplan, N. O. (1963) Lactic acid racemization in *Clostridium butylicum*, *Biochem. Z. 338*, 485-495.

(14) Shapiro, S. S., and Dennis, D. (1965) Lactic acid racemization in *Clostridium butylicum*. Evidence for a direct internal hydride shift, *Biochemistry 4*, 2283-2288.

(15) Shapiro, S. S., and Dennis, D. (1966) Lactic acid racemization in *Cl.butylicum*: ^{18}O exchange studies, *Biochem. Biophys. Res. Commun. 22*, 635-639.

(16) Cantwell, A., and Dennis, D. (1974) Lactate racemase. Direct evidence for an α-carbonyl intermediate, *Biochemistry 13*, 287-291.

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

(18) Xu, T., Wodrich, M. D., Scopelliti, R., Corminboeuf, C., and Hu, X. (2017) Nickel pincer model of the active site of lactate racemase involves ligand participation in hydride transfer, *Proc. Natl. Acad. Sci. U. S. A. 114*, 1242-1245.

(19) Zhang, X., and Chung, L. W. (2017) Alternative mechanistic strategy for enzyme catalysis in a Ni-dependent lactate racemase (LarA): Intermediate destabilization by the cofactor, *Chem. Eur. j 23*, 3623-3630.

(20) Yu, M. J., and Chen, S. L. (2017) From NAD^+ to nickel pincer complex: A significant cofactor evolution presented by lactate racemase, *Chem. Eur. j 23*, 7545-7557.

(21) Qiu, B., and Yang, X. (2017) A bio-inspired design and computational prediction of scorpion-like SCS nickel pincer complexes for lactate racemization, *Chem. Commun. (Camb.) 53*, 11410-11413.

(22) Wang, B., and Shaik, S. (2017) The nickel-pincer complex in lactate racemase is an electron relay and sink that acts through proton-coupled electron transfer, *Angew. Chem. Int. Ed. Engl. 56*, 10098-10102.

(23) 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. U. S. A. 113*, 5598-5603.

(24) Goffin, P., Deghorain, M., Mainardi, J. L., Tytgat, I., Champomier-Verges, M. C., Kleerebezem, M., and Hols, P. (2005) Lactate racemization as a rescue pathway for supplying D-lactate to the cell wall biosynthesis machinery in *Lactobacillus plantarum*, *J. Bacteriol. 187*, 6750-6761.

(25) Brugarolas, P., Duguid, E. M., Zhang, W., Poor, C. B., and He, C. (2011) Structural and biochemical characterization of N^5-carboxyaminoimidazole ribonucleotide synthetase and N^5-carboxyaminoimidazole ribonucleotide mutase from *Staphylococcus aureus*, *Acta Crystallogr. D Biol. Crystallogr. 67*, 707-715.

(26) Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, II, Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E., and Stubbe, J. (2007) N^5-CAIR mutase: role of a CO_2 binding site and substrate movement in catalysis, *Biochemistry 46*, 2842-2855.

(27) 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. U. S. A. 114*, 9074-9079.

(28) Desguin, B., Soumillion, P., Hols, P., Hu, J., and Hausinger, R. P. (2017) Lactate racemase and its niacin-derived, covalently-tethered, nickel cofactor, In *The Biological Chemistry of Nickel* (Zamble, D. B., Rowinska-Zyrek, M., and Kozlowski, H., Ed.), pp 220– 236, Royal Society of Chemistry, Cambridge, U.K.

(29) Chatterjee, A., Abeydeera, N. D., Bale, S., Pai, P. J., Dorrestein, P. C., Russell, D. H., Ealick, S. E., and Begley, T. P. (2011) *Saccharomyces cerevisiae* THI4p is a suicide thiamine thiazole synthase, *Nature 478*, 542-U146.

(30) Arragain, S., Bimai, O., Legrand, P., Caillat, S., Ravanat, J. L., Touati, N., Binet, L., Atta, M., Fontecave, M., and Golinelli-Pimpaneau, B. (2017) Nonredox thiolation in tRNA occurring via sulfur activation by a [4Fe-4S] cluster, *Proc. Natl. Acad. Sci. U. S. A. 114*, 7355-7360.

(31) Chen, M., Asai, S. I., Narai, S., Nambu, S., Omura, N., Sakaguchi, Y., Suzuki, T., Ikeda-Saito, M., Watanabe, K., Yao, M., Shigi, N., and Tanaka, Y. (2017) Biochemical and structural characterization of oxygen-sensitive 2-thiouridine synthesis catalyzed by an iron-sulfur protein TtuA, *Proc. Natl. Acad. Sci. U. S. A. 114*, 4954-4959.

CHAPTER 2

Lactate racemase nickel-pincer cofactor operates by a proton-coupled hydride transfer mechanism

Adapted from:

Rankin, J. A., Mauban, R. C., Fellner, M., Desguin, B., McCracken, J., Hu, J., Varganov, S. A., and Hausinger, R. P. (2018) Lactate racemase nickel-pincer cofactor operates by a proton-coupled hydride transfer mechanism, *Biochemistry 57*, 3244-3251.

For this chapter, I carried out the enzymatic assays, intermediate identification, absorption spectroscopy, kinetic isotope effect studies, and prepared samples for electron paramagnetic resonance analysis. For the latter studies, Dr. John McCracken performed the data collection and analysis. Drs. Benoît Desguin and Matthias Fellner carried out the structural studies under the guidance of Dr. Jian Hu. Robert Mauban conducted the computational studies under the guidance of Dr. Sergey Varganov.

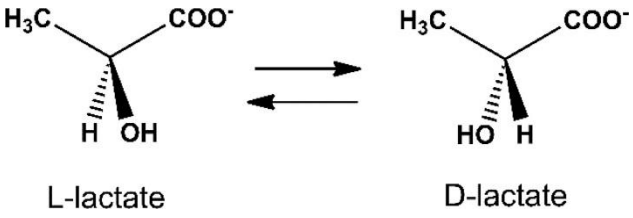
Abstract

Lactate racemase (LarA) of *Lactobacillus plantarum* contains a novel organometallic cofactor with nickel coordinated to a covalently tethered pincer ligand, pyridinium-3-thioamide-5-thiocarboxylic acid mononucleotide, but its function in the enzyme mechanism has not been elucidated. This study presents direct evidence that the nickel-pincer cofactor facilitates a proton-coupled hydride transfer (PCHT) mechanism during LarA-catalyzed lactate racemization. No signal was detected by electron paramagnetic resonance spectroscopy for LarA in the absence or presence of substrate, consistent with a +2 metal oxidation state and inconsistent with a previously proposed proton-coupled electron transfer mechanism. Pyruvate, the predicted intermediate for a PCHT mechanism, was observed in quenched solutions of LarA. A normal substrate kinetic isotope effect (k_H/k_D of 3.11 ± 0.17) was established using 2-^{2}H-D-lactate, further supporting a PCHT mechanism. UV–visible spectroscopy revealed a lactate-induced perturbation of the cofactor spectrum, notably increasing the absorbance at 340 nm, and demonstrated an interaction of the cofactor with the inhibitor sulfite. A crystal structure of LarA provided greater resolution (2.4 Å) than previously reported and revealed sulfite binding to the pyridinium C4 atom of the reduced pincer cofactor, mimicking hydride reduction during a PCHT catalytic cycle. Finally, computational modeling supports hydride transfer to the cofactor at the C4 position or to the nickel atom, but with formation of a nickel-hydride species requiring dissociation of the His200 metal ligand. In aggregate, these studies provide compelling evidence that the nickel-pincer cofactor acts by a PCHT mechanism.

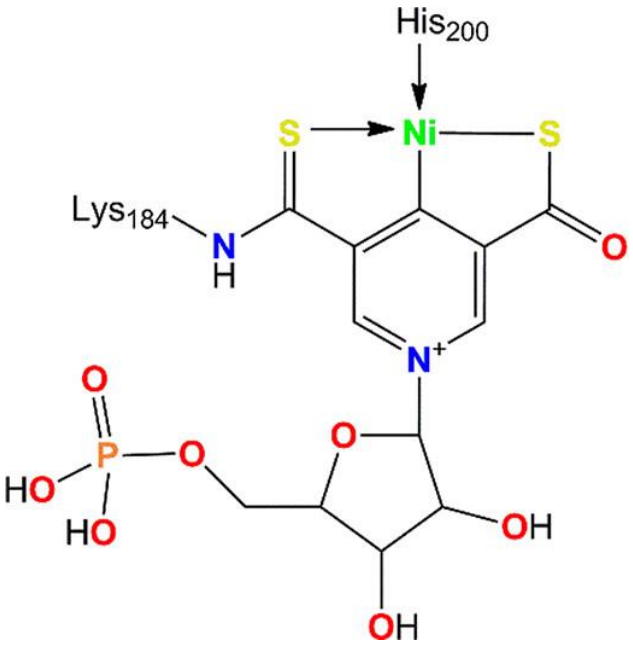
Introduction

Lactate racemase (LarA), a microbial enzyme that interconverts the D- and L-isomers of lactic acid (Scheme 2-1), was first reported for intact cells of two *Clostridium* species in 1936,[1] but only recently was it shown to possess a covalently tethered, nickel-containing cofactor derived from nicotinic acid adenine dinucleotide and synthesized by LarB, LarC, and LarE.[2-7] Structural characterization of LarA from *Lactobacillus plantarum* revealed a pyridinium-3-thioamide-5-thiocarboxylic acid mononucleotide attached to Lys184 of the protein, with the two sulfur atoms and the pyridinium C4 carbon atom bound to nickel as an (SCS) pincer complex with additional metal coordination by His200 (Scheme 2-2) . Two nearby histidine residues, His108 and His174, are well positioned to participate in catalysis by acting as general acids/bases, and their substitutions by alanine residues leads to enzyme inactivation.[2] The LarA structure is consistent with a proton-coupled hydride transfer (PCHT) mechanism (Scheme 2-3), in which one of the nearby histidine residues abstracts the hydroxyl group proton of lactate, a hydride transfers to the nickel-pincer cofactor with formation of bound pyruvate, and the hydride returns to either face of the ketoacid with

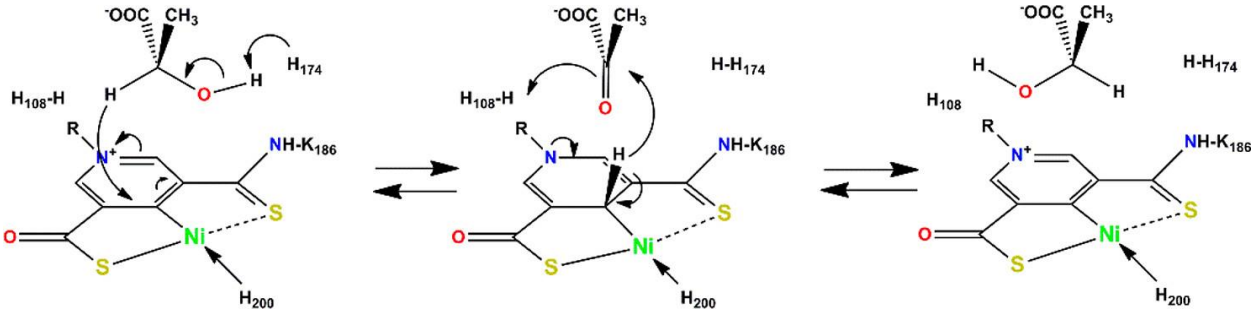
Scheme 2-1. Reaction Catalyzed by Lactate Racemase.



Scheme 2-2. Nickel-Pincer Cofactor of LarA.



Scheme 2-3. Potential Proton-Coupled Hydride Transfer (PCHT) Mechanism of LarA.



proton donation by the opposite histidine residue resulting in racemization. This scheme depicts hydride transfer to C4 of the pyridinium ring, similar to what is observed in NAD/NMN-dependent enzymes;[8] however, hydride transfer to the metal was also considered in the original report.[2] Indeed, the ability for hydride to localize either to the C4 position or to the nickel ion could potentially help explain the loss of stereospecificity for hydride return to pyruvate.[9] Direct experimental evidence for a PCHT mechanism has not been reported.

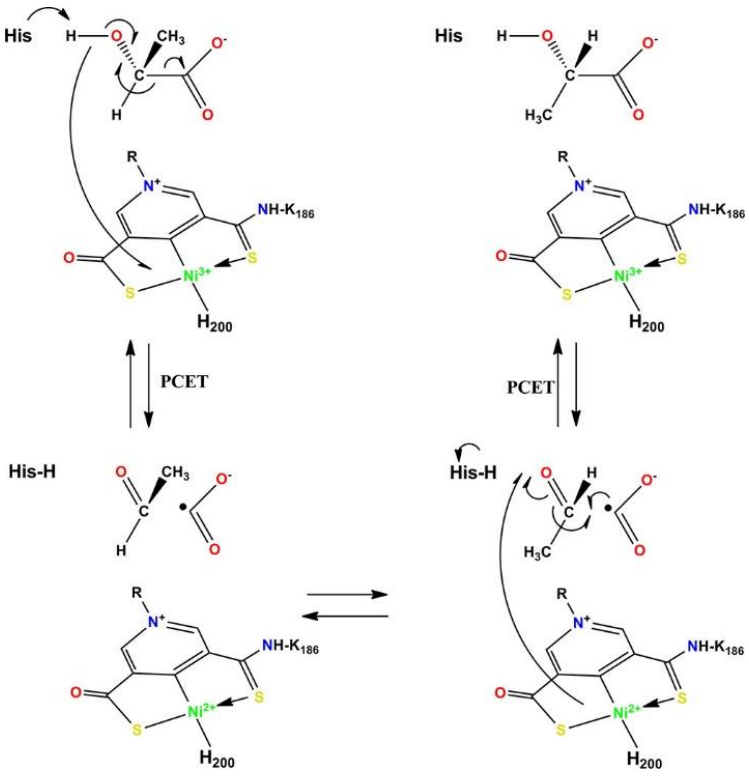
The elucidation of the LarA crystal structure inspired the synthesis of an (SCS) nickel-pincer model compound that closely resembles the enzyme metallocenter.[10] The nickel-containing 1-methyl-3,5-dithioamide-pyridinium species reacts with several primary and secondary alcohols to catalyze dehydrogenation reactions, forming the corresponding aldehydes and ketones. The authors carried out density functional theory (DFT) calculations indicating the reactions likely take place via hydride transfer to C4 of the pyridinium ring, and product analyses demonstrated the subsequent loss of nickel with rearomaticization of the ring.

Four other computational studies have addressed the mechanism by which the nickel-pincer cofactor facilitates the racemization of lactic acid. Zhang and Chung used DFT with M06, B3LYP-D3, and TPSSh-D3 methods to examine 63- and 139-atom models of the active site.[11] Their results are consistent with direct hydride transfer to the pyridinium C4 atom (Scheme 2-3), whereas their calculations are kinetically and thermodynamically inconsistent with hydride transfer to the nickel atom or to C8, and the authors report that His174 partially blocks hydride

transfer to C7. Replacement of the nickel atom with cobalt leads to higher energy barriers, while replacing the nickel with a proton or eliminating the tether lowers the energy barriers and increases the thermodynamic stability of the pyruvate intermediate. The authors proposed that destabilization of the intermediate is important to the racemization reaction. In addition, they suggested the involvement of His174 and His108 as alternate bases reacting with the different lactate isomers, as earlier proposed from the crystal structure.[2] Yu and Chen used the B3LYP-D3 method to study a 200-atom model of the active site.[12] Their calculations show no indication of a two-step mechanism (*i.e.*, it is a coupled reaction) and support a mechanism involving PCHT to C4 of the pyridinium ring (Scheme 2-3), while failing to transfer hydride to the nickel atom. In addition, these investigators found that the nickel pincer exhibited stronger hydride-addition reactivity compared to NAD^+ in environments of medium to high polarity. Qiu and Yang compared the predicted reactivities of several scorpion-like nickel-pincer complexes (independent of a protein environment) possessing tails that provide a reactive imidazole group.[13] Their calculations position the hydride on the pyridinium C4 atom and show increasing energy barriers as one shifts from a symmetric dithioamide pyridinium–Ni-Cl to asymmetric thioamide/thioacid pyridinium-Ni-CO to a dithioamide pyridinium–Pd-Cl to a dithiophosphine pyridinium–Pd-Br to a dithioamide benzyl–Ni-Cl species; thus, nickel is better than palladium and replacement of the pyridinium ring nitrogen by carbon is deleterious. In contrast to these three publications focused on PCHT, Wang and Shaik proposed an intriguing proton-coupled electron transfer (PCET) mechanism for the LarA reaction (Scheme 2-4).[14] These investigators used quantum mechanics/molecular mechanics (QM/MM) calculations in which the QM component used a hybrid UB3LYP functional while performing MM on the entire protein. They suggested the metal initially possesses a Ni(III) redox state that is reduced to Ni(II) during the

reaction, with the C1–C2 bond of lactate being transiently cleaved to yield a CO_2 radical anion and acetaldehyde that rotates, followed by reversal of the preceding steps. Taken together, it is clear that computational efforts have not established the mechanism.

Scheme 2-4 . Postulated Proton-Coupled Electron Transfer Mechanism of LarA.



Here, we provide experimental evidence in direct support of a PCHT mechanism of lactate racemase, we present a new crystal structure of LarA at better resolution that also supports a PCHT mechanism, and we provide computational support that the hydride can transfer to the nickel as well as the pyridinium C4 atom.

Experimental Procedures

LarA Purification, Activity, and Nickel Content

C-Terminal Strep II-tagged LarA of *L. plantarum* was isolated from *Lactococcus lactis* strain NZ3900 containing plasmid pGIR112 with the modified *larA* and *larBCDE* under control of the *nisA* promoter.[4] Overnight starter cultures were grown at 30 °C in M17 medium supplemented with 0.5% w/v D-glucose and 10 mg/L chloramphenicol, then were used to inoculate (at 1% final volume) larger volumes of the same medium with 5 mg/mL of chloramphenicol 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 4 h before collection by centrifugation and storage at −80 °C.

Cell pellets (5–10 g) were thawed and diluted to 40 mL of 100 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, 0.5 mM Na_2SO_3 (which stabilizes the enzyme activity),[2] 1 mM phenylmethylsulfonyl fluoride (using 100 mM stock dissolved in ethanol), and 2.5 U/mL benzonase, with, in some cases, 5 mM ascorbic acid. The cells were disrupted by two passes through a French pressure cell (16000 psi) that had been prechilled to 4 °C. Cell-free lysates were obtained after centrifugation at 27000*g* for 40 min at 4 °C. The lysates were applied at 4 °C to 2 mL of StrepTactin superflow or StrepTactin XT resin (IBA) equilibrated in 100 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, 0.05 mM Na_2SO_3, and, in some cases, 5 mM ascorbic acid. After the resins were washed with three column volumes of buffer, the enzyme was eluted with buffer containing either 5 mM desthiobiotin or 50 mM biotin, respectively. Alternatively, the LarA used in the crystallization experiments was purified as previously described.[2]

LarA activity was assayed by measuring the conversion of D- to L-lactate or L- to D-lactate using isomer specific dehydrogenases in a commercial kit (Megazyme) as previously described.[4] Protein concentrations were assessed by the Bradford assay.[15]

Electron Paramagnetic Resonance (EPR) Spectroscopy

Sulfite, a LarA inhibitor, was diluted to less than 100 nM in the protein sample by using an Ultra-4 10 kDa cutoff centrifugal filter (Amicon) to carry out two dilution/concentration buffer exchanges into 100 mM Tris-HCl, pH 7.5, containing 150 mM NaCl and 5 mM ascorbic acid. The sample was adjusted to 25% glycerol (w/v) in a volume of 160 μL containing 3.6 mg/mL (75 μM) of LarA and frozen in liquid N_2. After initial spectroscopic analysis, the sample

was thawed, 16 μL of 1 M D-lactate was added and mixed, and the modified sample was refrozen.

Continuous-wave EPR measurements were made at X-band on a Bruker E-680X spectrometer using a Bruker SHQ probe fitted with an Oxford ESR 900 liquid helium cryostat. Measurement conditions were microwave frequency, 9.407 GHz; sample temperature, 8 K; magnetic field scan range, 150–450 mT; magnetic field modulation amplitude, 1.0 mT, modulation frequency, 100 kHz. Microwave powers were varied from 0.1–1.0 mW and video gains ranged from 70–90 dB.

UV–Visible Spectroscopy

UV–visible absorbance spectra were obtained using a Shimadzu UV-2600 spectrophotometer. Purified LarA was separated from sulfite by using a PD-10 column (GE Healthcare) to exchange into 20 mM Tris-HCl, pH 7.5, containing 150 mM NaCl and the spectrum was collected for a 200-μL sample (0.94 mg/mL) using a microcuvette. To this sample was added either 20 μL of 1 M sodium D-lactate and then 20 μL of 100 mM Na_2SO_3 or just the sulfite with spectra collected after each addition corrected for dilution.

Pyruvate Detection

To eliminate contaminating pyruvate, stock solutions of lactate (either the l-isomer or d-isomer) were brought to high pH with NaOH, treated overnight with excess $NaBH_4$, and acidified with HCl.

LarA was buffer exchanged at 4 °C into deoxygenated, argon-equilibrated 20 mM Tris-HCl, pH 7.5, containing 150 mM NaCl by twice using a PD-10 column (GE Healthcare) followed by concentration using an Ultra-4 10-kDa centrifugal filter. The LarA samples (1–6 mg/mL) were mixed on ice with pyruvate-free lactate (200 mM final concentration) in 60 mM

MES-NaOH, pH 6.0, and incubated at 30 °C for 2 min. The reactions were quenched by adding an equal volume of ice-cold 0.5 M perchloric acid, and the mixtures were kept on ice for 5 min, neutralized with 250 mM K_2CO_3, and incubated on ice. The protein was removed by centrifugation (17000*g*, 1 min), and the supernatants were used in pyruvate assays as previously described.[16] Briefly, 20 μL of the supernatant was diluted to 200 μL in 100 mM potassium phosphate, pH 6.7, containing 1 mM Na-EDTA, 1 mM $MgCl_2$, 0.2 U/mL of pyruvate oxidase (MP Biomedical), 0.2 U/mL of horseradish peroxidase (HRP, Sigma Chemical Co.), and 50 μM 10-acetyl-3,7-dihydroxyphenoxazine (Cayman Chemical) in an opaque black microtiter plate. The assay was incubated at room temperature for 30 min, and the fluorescence was examined with a PerkinElmer EnVision 2104 plate reader using 530 nm excitation and 590 nm emission. The fluorescence intensities obtained for sample pyruvate concentrations were compared to those of a pyruvate standard curve.

As an independent approach to investigate the LarA intermediate, the protein samples used for the fluorescence assay were treated with perchloric acid (250 mM) as described earlier to precipitate the protein, and the remaining solution was derivatized with 2,4-dinitrophenyl-hydrazine (DNPH) by a method adapted from a previously described protocol.[17] In brief, 20 μL of the acidified solution was mixed with 10 μL formic acid, 100 μL of DNPH dissolved in water (0.1 mg/mL), and 180 μL of acetonitrile. The reaction mixture was incubated at 37 °C for 1 h, and then twice extracted with 600 μL of a 1:5 solution of ethyl acetate:hexanes. The combined uppermost layer was dried using a SpeedVac at 42 °C and dissolved in 100 μL of methanol. Liquid chromatographic separation of the sample (10 μL injection volume) was performed with a Waters Aquity UPLC system using a C18 column (Aquity BEH 1.7 μm, 2.1 × 100 mm) and mobile phases 0.15% formic acid in water (A) and acetonitrile (B). The gradient elution (0.3

mL/min) used the following program: 0–7 min, 1% B to 99% B; 7–9 min isocratic at 99% B; and 9–10 min, 1% B. Electrospray ionization-quantitative time-of-flight mass spectrometry used a XEVO G2-XS in negative ionization mode.

Substrate Kinetic Isotope Effect

The substrate kinetic isotope effect for 2-^{2}H-L-lactate was examined by incubating purified and sulfite-free LarA with 20 mM solutions of the sodium form of deuterated sample (CDN Isotopes) or non-deuterated L-lactate (Sigma) in 60 mM MES-NaOH, pH 6.0, for 10 min at 35 °C. The reactions were terminated by incubating at 95 °C for 10 min. The concentrations of D-lactate in the samples were determined as described earlier.

Structural Characterization of LarA

LarA was crystallized in 0.1 M Bis-Tris buffer (pH 6.5) containing 2 M ammonium sulfate. X-ray diffraction data were collected at the Advanced Photon Source GM/CA beamline 23-ID-D. Data sets were processed, merged, and scaled with HKL2000.[18] The phases were solved by Phaser molecular replacement in Phenix[19] using PDB entry 5HUQ[2] as the template. Model building and refinement were conducted in COOT[20] and Phenix.[19]

We started out modeling all chains with sulfates and the nonreduced cofactor as in our earlier structure.[2] Although this approach worked well in chain B, it did not fit the density in chains A and C and also led to clashing with the pyridinium C4 atom in both cases. Examining the $mF_o–DF_c$ map without a modeled ligand showed that the outward-facing density above the ring was more flat in chains A and C, while that in chain B appeared more round and the center of the density in chain B also was further away from the cofactor when comparing chain A to chain B. Lastly, the density of the ring and the molecule above it was apparently connected in chain A, but less so in chain B and again with chain C being in the middle. An attempt to model

sulfite and the nonreduced cofactor in chains A and C fit the density nicely; however, refining the sulfite correctly with the observed density remained challenging, especially in chain A. We therefore produced the model of the reduced cofactor and a covalent bond in A, which was refined very well and matched the density best. The reduced sulfite-modified version of the cofactor (pyridinium-3,5-bisthiocarboxylic acid mononucleotide or P2TMN) was created with Jligand.[21] The model and structure factors were deposited in the PDB under ID 6C1W. Statistics for the data set are listed in Table 2-A1. Structure analysis and figures were created with UCSF Chimera.[22]

Computational Methods

Two types of active site models were considered. Large models were used to calculate the reaction barriers for PCHT to the pyridinium C4 atom. The goal was to obtain the reaction barriers for different models using the same model size and level of theory. To this end, the model of Zhang and Chung[11] and the model of Yu and Chen[12] were reduced to 98 and 97 atoms, respectively. The 98-atom model contained the protonated Lys184 residue, while in the 97-atom model this residue was neutral, as in the original studies. To obtain these models, where necessary the single C–C bonds were cut and capped with hydrogen atoms. Only the positions of these hydrogen atoms were reoptimized, while positions of all other atoms were kept frozen. To predict the reaction barriers, the same procedure was applied to the reactant and transition state structures. Two small models with 34 atoms (with His200) and 22 atoms (without His200) were used to investigate the possibility of hydride binding to nickel. For all models except the small model with neutral His200, the total charge and electron spin were set to 0. The charge and spin of the model with neutral His200 were set to +1 and 0, respectively. Electronic structure calculations on the large models were carried out with the unrestricted B3LYP-D3 method and

def2-SVP basis set. Calculations on the small models were performed without D3 dispersion correction and with the slightly smaller def2-SV(P) basis set. In all calculations, the protein environment was simulated with the polarization continuum model using the dielectric constant ε = 4, the solvent radius r_{solv} = 1.20 Å, and the nickel radius r_{Ni} = 1.63 Å. All calculations were performed with the GAMESS suite of programs.[23]

Results and Discussion

EPR Analysis of LarA

A key distinction between the postulated PCHT and PCET mechanisms for LarA is the redox state of nickel in the enzyme. Whereas the mechanisms that invoke hydride transfer to the nickel-pincer cofactor assume the more stable Ni(II) state, the PCET mechanism invokes the more oxidized Ni(III) state for the starting enzyme, with intermediate states containing Ni(II). To test the redox state of the metal in LarA, we examined the sample lacking and containing substrate by EPR spectroscopy. A square planar nickel in the +2 oxidation state, as indicated by the crystal structure, would yield a diamagnetic species. By contrast, Ni(III) sites are typically octahedral or tetragonal and give rise to signals in the region from g = 2.4–2.0. Our measurements show no EPR signal over the 150–450 mT region scanned, consistent with diamagnetic Ni(II) in the samples.

UV–Visible Spectroscopic Analyses of LarA

The spectrum of sulfite-free LarA exhibited electronic transitions at 380 and 440 nm with a shoulder at 550 nm (Figure 2-1A). The addition of lactate led to absorbance increases at 340 and 400 nm along with a broad feature at 600 nm, as most easily visualized by the difference spectrum. The 340 nm transition associated with lactate addition to LarA (Figure 2-1A) matches the wavelength of the spectrum obtained by hydride addition to C4 of NAD^+ (with ∼10% of the

NADH extinction coefficient as expected for a partially reduced state), consistent with hydride addition to C4 of the nickel-pincer cofactor. The addition of sulfite to LarA containing lactate (Figure 2-1A) or directly added to LarA (Figure 2-1B) resulted in an intensification of the 380 nm electronic transition and a diminishment of the absorbance at 440 nm. The resulting 440 nm/380 nm absorbance ratio matches that of the previously reported chromophore for the enzyme that was in buffer containing 50 μM sulfite.[2] The 320 nm transition formed by sulfite addition to LarA (Figure 2-1B) is consistent with that formed by sulfite interaction with NAD^+ (Figure 2-1C), suggesting attack at C4 of the nickel-pincer cofactor. We cannot yet assign the longer wavelength transitions of the Ni-pincer cofactor.

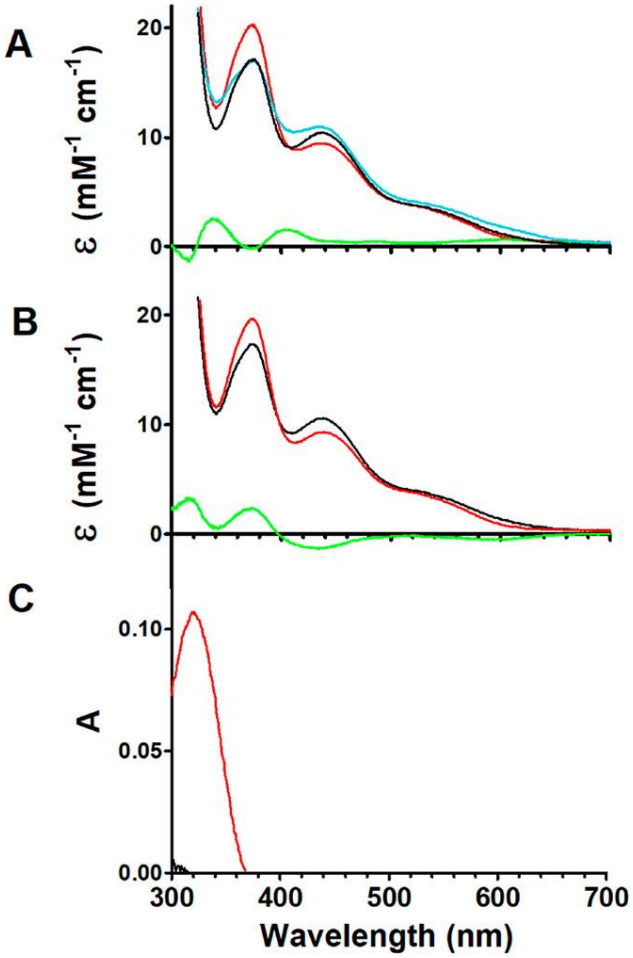


Figure 2-1. Lactate and sulfite alter the UV–visible spectrum of LarA. (A) Spectrum of LarA (black trace) and spectra obtained by sequential addition of lactate (cyan trace) then sulfite (red trace). Also shown is the difference spectrum for the changes when lactate is added (green trace). (B) Spectrum of LarA (black trace) and that obtained after adding sulfite (red trace), along with the difference spectrum (green trace). (C) Spectrum of NAD^+ before (black trace) and after addition of sulfite (red trace).

Characterization of the LarA Intermediate

A critical distinction between the PCHT and PCET mechanisms is the identity of the organic intermediate, pyruvate versus acetaldehyde, respectively. Pyruvate was present in enzyme samples that were mixed with borohydride-treated lactate substrate, quenched, and

assayed by using a pyruvate oxidase/HRP coupled system. The amount of pyruvate detected varied with the enzyme preparation and corresponded to a mole fraction of 5–30% compared to the amount of LarA. To take account of the enzyme specific activity for each preparation, thus circumventing the batch-to-batch variation, a consistent amount of pyruvate per unit of enzyme activity was observed (Figure 2-2). This result indicates that a significant fraction of LarA is in the reduced, pyruvate-bound state under near saturating conditions of substrate. Such a result argues against a destabilized intermediate state, as suggested by one computational study.[11]

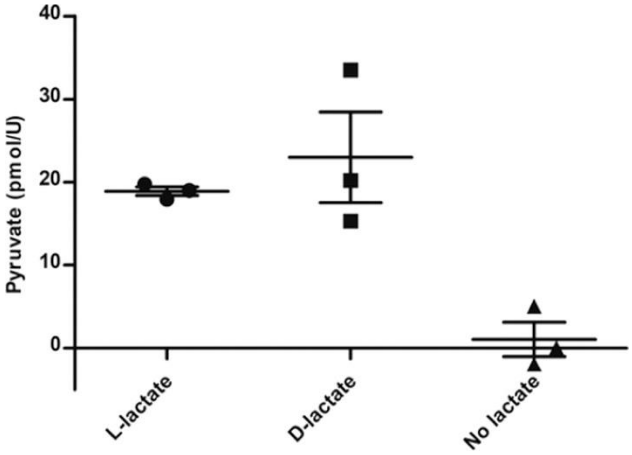


Figure 2-2. Detection of pyruvate as an intermediate in the LarA reaction. The amount of pyruvate was determined after quenching the reaction using 40–130 μM enzyme and 200 mM D- or L-lactate and a lactate-free control. The amount of pyruvate was quantified on the basis of the measured enzyme activity units (D- to L-conversion).

A mass spectrometry method involving DNPH derivatization of quenched LarA samples confirmed that pyruvate formed from D-lactate during the reaction (Figure 2-3). These findings are consistent with the intermediacy of pyruvate and consistent with a hydride transfer mechanism.

Substrate Kinetic Isotope Effect (KIE) Studies

Long before the elucidation of the LarA crystal structure, experimentalists had studied the mechanism of lactate racemase by examining substrate kinetic isotope effects using the enzyme from *Clostridium beijerinckii*. Specifically, Shapiro and Dennis examined the D- and L-isomers of 2-α-^{2}H-lactic acid to reveal k_H/k_D values of 2.16 and 2.14, respectively.[24] These KIEs are consistent with direct cleavage of the C–H bond in the substrate as expected for the hydride

transfer mechanism (Scheme 2-3, see page 18), but they are inconsistent with the PCET mechanism where this bond remains intact (Scheme 2-4, see page 20). These researchers also showed the product retained ^{2}H, consistent with either mechanism. Although *C. beijerinckii* possesses two genes encoding LarA-like proteins,[4] neither of these proteins have been connected to the cell's lactate racemase activity.

To directly examine the substrate KIE for LarA from *L. plantarum*, we compared the reaction rates using non-labeled and 2-α-^{2}H-lactic acid. The measured k_H/k_D of 3.11 ± 0.17 is consistent with a normal KIE (e.g., values of 3–5 typically are associated with hydride transfer in NAD^+ enzymes)[25] and supports the PCHT mechanism in which the C–H is broken and reformed during catalysis.

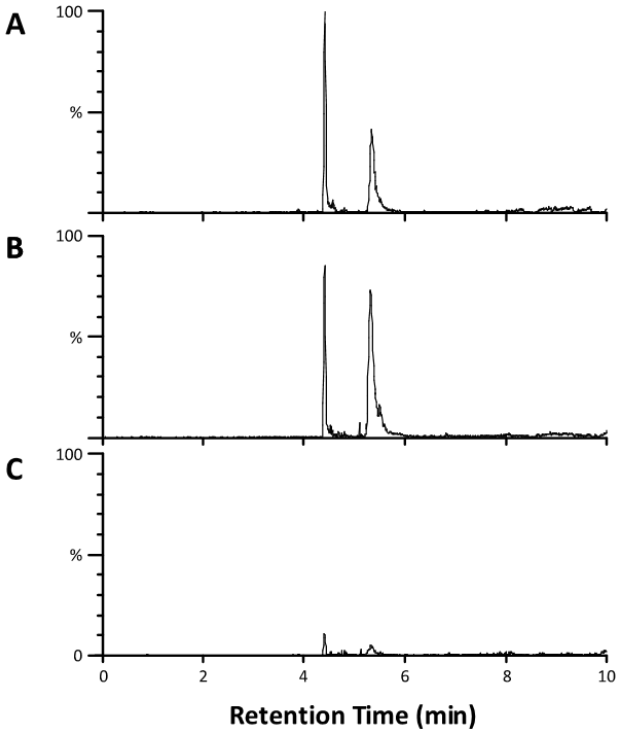


Figure 2-3. Analysis of DNPH-derivatized pyruvate by liquid chromatography-mass spectrometry. (A) LarA (4.4 mg/mL, 93 µM) was incubated with sodium D-lactate, precipitated with perchloric acid, derivatized with DNPH, extracted into organic solvent, dried, dissolved in methanol, and analyzed by liquid chromatography-mass spectrometry. (B) A standard of 10 µM sodium pyruvate was treated as in A. (C) Non-protein control. The profiles result from monitoring m/z of 267.04 ± 0.01. The y-axes are on the same scale

Structural Characterization of LarA in Complex with Sulfite

The published structure for LarA (5HUQ) solved at a moderate resolution (3 Å) revealed two chains per asymmetric unit with a closed conformation containing a buried nickel-pincer cofactor and a more open conformation with a metal-free pincer cofactor that was accessible to solvent.[2] A new crystal form of the same construct provided a structure with improved resolution (2.40 Å) where the three chains in one asymmetric unit occur in the closed

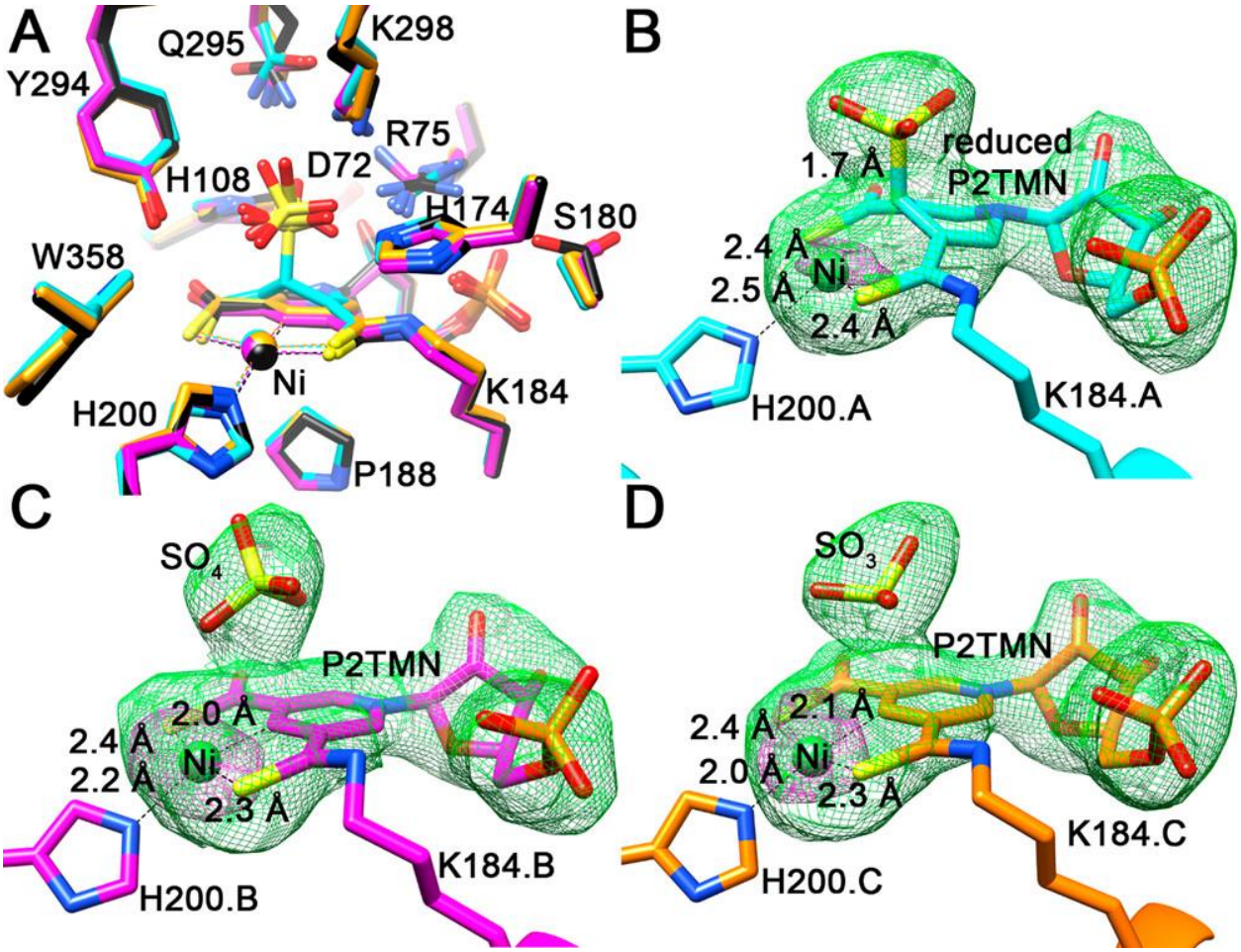


Figure 2-4. Comparison of the active sites of the LarA crystal structures. Carbon atoms are colored black for the closed conformation chain of the previous model (5HUQ); cyan for chain A, magenta for chain B, and light orange for chain C of the new structure (6C1W). (A) Active site based on C_α alignment of the four chains with active site residue side chains illustrated. Nickel atom and coordination colored as the corresponding carbon atoms. (B–D) Active sites of the three new chains with an $mF_o–F_c$ difference map shown for nickel, the cofactor and SO_3/SO_4 at 3σ in green. Anomalous map shown at 4σ in magenta. Nickel atoms are colored in light green. Nickel coordination is shown as black dashes with distances indicated. In B the C–S distance is labeled.

conformation, with RMSD of ~0.3–0.4 Å between the three chains and the closed chain of the previous model (5HUQ). The only overall backbone difference occurs at the C-terminal *Strep* II-tag.

As with the previously described closed conformation, all three chains in the current structure contain the nickel-pincer complex. Nearly all protein side chains at the active site pocket (residues Ser71, Asp72, Thr74, Tyr108, Ser174, Phe176, Pro188, Lys184, His200, Tyr294, Trp358) show the same side chain orientation as before (Figure 2-4A), with Lys184

again forming a covalent bond with the cofactor. Due to the higher resolution, Ser180 modeled slightly differently so that it now faces away from the phosphate moiety of the cofactor. The side chains of Arg75, Gln295, and Lys298 show minor variations among the three chains (Figure 2-4A). Several water molecules could now be placed with confidence, most importantly around the cofactor phosphate group.

The major novel findings derived from the new structure are observed adjacent to the nickel and the pyridinium ring. In the previously published structure for LarA, a large round density above the pyridinium ring was modeled as a sulfate molecule, potentially representing the lactate binding site.[2] In our new structure, we observe the same density; however, it occurs in three different variations in the three chains. In chain B the center of the density peak was ~3.4 Å away from the C4 of P2TMN and was again modeled as a sulfate molecule (Figure 2-4C), closely resembling the previous model. In chain A, however, the corresponding density was much closer to the ring, eliminating the possibility of modeling a sulfate which would lead to severe clashing in the structure. Considering that sulfite was used to protect LarA from oxidation and loss of nickel during purification and that sulfite can form a covalent adduct with C4 of NAD[26-28] and N5 of flavin,[29, 30] we believe that the density represents a sulfite which is covalently linked to C4 of the pyridinium of the pincer cofactor. When sulfite donates its lone electron pair to C4, P2TMN converts into a reduced form. We therefore generated a model with sulfite forming a covalent adduct with C4 of the reduced P2TMN, which nicely matched the $mF_o–F_c$ difference map of the omitted active site ligands after refinement (Figure 2-4B). The refined distance between the sulfur of sulfite and C4 of the cofactor (1.7 Å) is consistent with S–C bond formation. The position of the nickel atom also shifts slightly compared to the other chains. The resulting model indicates that the bond between nickel and the C4 atom is likely

broken or largely weakened, and the interaction between nickel and His200 is also weaker, as this bond is now longer by ~0.4 Å compared to other chains. It appears that nickel is less stable when P2TMN is in the reduced form. Indeed, when setting an occupancy of 1 for the nickel atom in chain A, the metal ion was refined to a B-factor ~20 $Å^2$ higher than the surrounding atoms, a value that is also greater than what was measured for the other chains and consistent with the observed weaker anomalous signal. This result indicates that the nickel atom in chain A has a greater thermal mobility and/or that it has reduced occupancy. Finally, chain C was modeled as P2TMN and a sulfite molecule, with the C4–S distance too far to be a covalent bond, but too close for a sulfate that would sterically clash (Figure 2-4D). This active site appears to be an intermediate step before P2TMN reduction and formation of the C–S bond. Of potential relevance to the enzyme mechanism, the observation of a covalent adduct between sulfite and the nickel-pincer cofactor is a plausible mimic of the interaction between a hydride and the active site.

Computational Analysis of LarA

Computational studies suggested the hydride can transfer either to the C4 position of the pyridinium ring or to the nickel, but the latter occurs only if His200 dissociates. If the protonated or neutral form of His200 is coordinated to nickel, the hydride binds only to the C4 position (Figure 2-5A,B). For the model with dissociated His200, the hydride can bind to nickel (Figure 2-5C); however, the binding to the C4 atom (Figure 2-5D) is more favorable by 8.4 kcal/mol. We investigated the hydride transfer to C4 in more detail using the large models obtained from the structures reported by Zhang and Chung[11] (Figure 2-5E) and Yu and Chen[12] (Figure 2-5F). The transition states were calculated to be 24.5 and 14.6 kcal/mol above the reactants with models E and F, respectively. These values are in good agreement with the 27.5 and 12.0 kcal/mol barriers

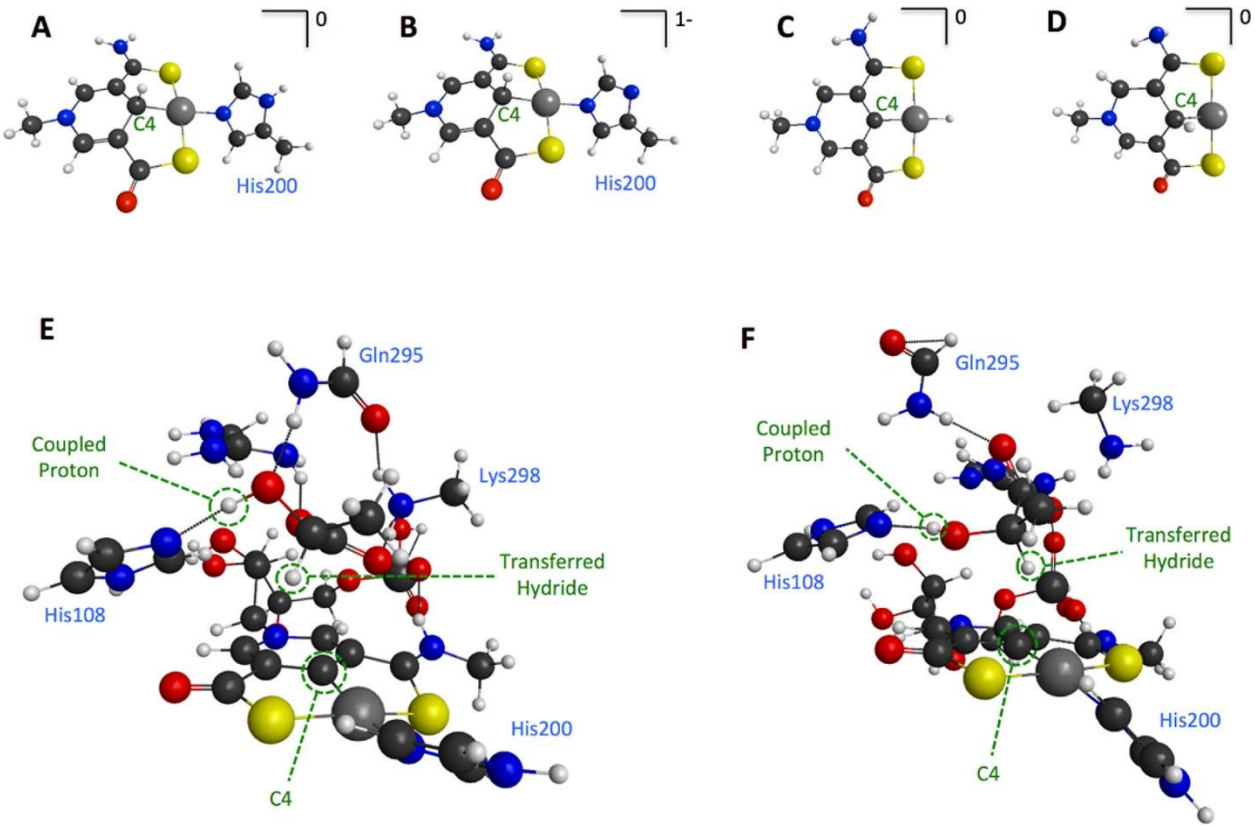


Figure 2-5. Models used to investigate hydride binding to the nickel-pincer cofactor. (A–D) Small models with hydride transfer to the C4 position and to the nickel. Large models (reactant geometries) obtained from the structures reported by (E) Zhang and Chung[11] or (F) Yu and Chen[12]. Colored as nickel, gray; sulfur, yellow; nitrogen, blue; oxygen, red; carbon, black; and hydrogen, white.

reported for similar models and level of theory.[11, 12] The significantly lower barrier predicted by model F is likely due to different orientations of Lys298 and Gln295, which stabilize the PCHT transition state.[12] The agreement between the predicted barrier for model F and the activation energy of 12.8 kcal/mol obtained from the experimental data,[14] as well as the lack of evidence for changing the nickel oxidation state provide strong support for the PCHT mechanism.

Conclusions

We have provided compelling evidence that LarA utilizes a PCHT rather than a PCET mechanism on the basis of (i) the lack of a Ni(III) EPR signal associated with the enzyme, (ii) the intermediacy of pyruvate upon quenching of the reaction, and (iii) the observation of a KIE when using 2-α-^{2}H-lactic acid thus demonstrating the breakage of this C–H bond. Consistent

with a hydride transfer to the nickel-pincer cofactor, we observed a perturbation of the UV–visible spectrum upon addition of substrate to the enzyme. Additionally, sulfite addition caused changes to the UV–vis spectrum of LarA. A sulfite adduct at the C4 position of the cofactor was identified in a higher resolution crystal structure of LarA, compatible with hydride addition at this site. Finally, we used computational methods to calculate the energies of transition states and intermediates in the reaction. In contrast to similar prior calculations, we note that hydride transfer to Ni(II) is feasible if the His200 ligand dissociates during catalysis. Significantly, synthetic Ni-pincer complexes with hydride in the same plane as the pincer are well-known.[31-35] Our calculations using active site models and these synthetic compound precedents suggest that further investigation is needed to test for a nickel-hydride intermediate in LarA.

APPENDIX

Table 2-A1. Crystal statistics for the LarA structure.

Data collection	**LarA**
Beamline	GM/CA 23-ID-D
Wavelength (Å)	0.979
Space group	$P2_12_12_1$
Unit cell a, b, c (Å); α, β, γ (°)	69, 158, 159; 90, 90, 90
[a]Resolution (Å)	40.00 – 2.40 (2.49 – 2.40)
Unique reflections	68,698 (6,669)
[a]Redundancy	13.2 (11.3)
[a]Completeness (%)	99.2 (97.7)
[a]$I/\sigma I$	17.3 (1.5)
[a,b]R_{merge}	0.103 (-)
[a,c]R_{pim}	0.043 (0.484)
[d]$CC_{1/2}$	0.998 (0.797)
Data refinement	
Protein atoms	10046
H_2O molecules	109
Ni atoms	3
SO_4 molecules	16
SO_3 molecules	1
4EY: P2TMN	2
ENJ: reduced P2TMN	1
[e]R_{work}/R_{free}	0.206/0.263
B-factors (Å^2)	53.9
Protein atoms	53.9
H_2O molecules	48.3
Ni atoms	61.9
SO_4 molecules	66.7
SO_3 molecules	57.6
4EY: P2TMN	46.7
ENJ: reduced P2TMN	48.8
R.m.s. deviation in bond lengths (Å)	0.008
R.m.s. deviation in bond angles (°)	0.94
Ramachandran plot (%) favored	96.7
Ramachandran plot (%) outliers	0
Rotamer outliers	0
PDB ID	6C1W

[a]Highest resolution shell is shown in parentheses.

[b]$R_{merge} = \sum_{hkl} \sum_j |I_j(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_j I_j(hkl)$, 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 (10% of the total) omitted in model refinement.

REFERENCES

REFERENCES

(1) Tatum, E. L., Peterson, W. H., and Fred, E. B. (1936) CCLXVI. Enzymic racemization of optically active lactic acid, *Biochem. J. 30*, 1892-1897.

(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., 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. U. S. A. 113*, 5598-5603.

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

(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. U. S. A. 114*, 9074-9079.

(6) Desguin, B., Soumillion, P., Hausinger, R. P., and Hols, P. (2017) Unexpected complexity in the lactate racemization system of lactic acid bacteria, *FEMS Microbiol. Rev. 41*, S71-S83.

(7) Fellner, M., Rankin, J. A., Hu, J., and Hausinger, R. P. (2017) Lactate Racemase, In *Encyclopedia of Inorganic and Bioinorganic Chemistry* (Scott, R. A., Ed.), Wiley & Sons, West Sussex, U.K.

(8) Belenky, P., Bogan, K. L., and Brenner, C. (2007) NAD^+ metabolism in health and disease, *Trends Biochem. Sci. 32*, 12-19.

(9) Desguin, B., Soumillion, P., Hols, P., Hu, J., and Hausinger, R. P. (2017) Lactate racemase and its niacin-derived, covalently-tethered, nickel cofactor, In *The Biological Chemistry of Nickel* (Zamble, D. B., Rowinska-Zyrek, M., and Kozlowski, H., Ed.), pp 220– 236, Royal Society of Chemistry, Cambridge, U.K.

(10) Xu, T., Wodrich, M. D., Scopelliti, R., Corminboeuf, C., and Hu, X. (2017) Nickel pincer model of the active site of lactate racemase involves ligand participation in hydride transfer, *Proc. Natl. Acad. Sci. U. S. A. 114*, 1242-1245.

(11) Zhang, X., and Chung, L. W. (2017) Alternative mechanistic strategy for enzyme catalysis in a Ni-dependent lactate racemase (LarA): Intermediate destabilization by the cofactor, *Chem. Eur. j 23*, 3623-3630.

(12) Yu, M. J., and Chen, S. L. (2017) From NAD^+ to nickel pincer complex: A significant cofactor evolution presented by lactate racemase, *Chem. Eur. j 23*, 7545-7557.

(13) Qiu, B., and Yang, X. (2017) A bio-inspired design and computational prediction of scorpion-like SCS nickel pincer complexes for lactate racemization, *Chem. Commun. (Camb.) 53*, 11410-11413.

(14) Wang, B., and Shaik, S. (2017) The nickel-pincer complex in lactate racemase is an electron relay and sink that acts through proton-coupled electron transfer, *Angew. Chem. Int. Ed. Engl. 56*, 10098-10102.

(15) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248-254.

(16) Zhu, A., Romero, R., and Petty, H. R. (2010) A sensitive fluorimetric assay for pyruvate, *Anal. Biochem. 396*, 146-151.

(17) Wang, Z., Tang, S., Jin, Y., Zhang, Y., Hattori, M., Zhang, H., and Zhang, N. (2015) Two main metabolites of gentiopicroside detected in rat plasma by LC-TOF-MS following 2,4-dinitrophenylhydrazine derivatization, *J. Pharm. Biomed. Anal. 107*, 1-6.

(18) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, In *Methods Enzymol.*, pp 307-326, Academic Press.

(19) Adams, P. D., Afonine, P. V., Bunkoczi, 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 Biol. Crystallogr. 66*, 213-221.

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

(21) Lebedev, A. A., Young, P., Isupov, M. N., Moroz, O. V., Vagin, A. A., and Murshudov, G. N. (2012) JLigand: a graphical tool for the CCP4 template-restraint library, *Acta Crystallogr. D Biol. Crystallogr. 68*, 431-440.

(22) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J. Comput. Chem. 25*, 1605-1612.

(23) Gordon, M. S., and Schmidt, M. W. (2005) Advances in Electronic Structure Theory: GAMESS a Decade Later, In *Theory and Applications of Computational Chemistry: The First Forty Years* (Dykstra, C. E., Frenking, G., Kim, K. S., and Scuseria, G. E., Ed.), pp 1167– 1189, Elsevier, Amsterdam.

(24) Shapiro, S. S., and Dennis, D. (1965) Lactic acid racemization in *Clostridium butylicum*. Evidence for a direct internal hydride shift, *Biochemistry 4*, 2283-2288.

(25) Fersht, A. (1977) *Enzyme Structure and Mechanism*, W. H. Freeman and Company, New York.

(26) Johnson, S. L., and Smith, K. W. (1976) The interaction of borate and sulfite with pyridine nucleotides, *Biochemistry 15*, 553-559.

(27) Parker, D. M., Lodola, A., and Holbrook, J. J. (1978) Use of the sulphite adduct of nicotinamide-adenine dinucleotide to study ionizations and the kinetics of lactate dehydrogenase and malate dehydrogenase, *Biochem. J. 173*, 959-967.

(28) Larsen, P. O., Sundahl, M., and Wieczorkowska, E. (1982) Hydride ion and sulfite ion attack on pyridine derivatives: Correspondence between biochemical and chemical reactions as demonstrated by UV spectroscopy, *J. Chem. Educ. 59*.

(29) Massey, V., Muller, F., Feldberg, R., Schuman, M., Sullivan, P. A., Howell, L. G., Mayhew, S. G., Matthews, R. G., and Foust, G. P. (1969) The reactivity of flavoproteins with sulfite. Possible relevance to the problem of oxygen reactivity, *J. Biol. Chem. 244*, 3999-4006.

(30) Tan, T. C., Pitsawong, W., Wongnate, T., Spadiut, O., Haltrich, D., Chaiyen, P., and Divne, C. (2010) H-bonding and positive charge at the N5/O4 locus are critical for covalent flavin attachment in *Trametes* pyranose 2-oxidase, *J. Mol. Biol. 402*, 578-594.

(31) Boro, B. J., Duesler, E. N., Goldberg, K. I., and Kemp, R. A. (2009) Synthesis, characterization, and reactivity of nickel hydride complexes containing 2,6-$C_6H_3(CH_2PR_2)_2$ (R = *t*Bu, *c*Hex, and *i*Pr) pincer ligands, *Inorg. Chem. 48*, 5081-5087.

(32) Mock, M. T., Potter, R. G., O'Hagan, M. J., Camaioni, D. M., Dougherty, W. G., Kassel, W. S., and DuBois, D. L. (2011) Synthesis and hydride transfer reactions of cobalt and nickel hydride complexes to BX_3 compounds, *Inorg. Chem. 50*, 11914-11928.

(33) Rossin, A., Peruzzini, M., and Zanobini, F. (2011) Nickel(II) hydride and fluoride pincer complexes and their reactivity with Lewis acids $BX_3 \cdot L$ (X = H, L = thf; X = F, L = Et_2O), *Dalton Trans 40*, 4447-4452.

(34) Chierotti, M. R., Rossin, A., Gobetto, R., and Peruzzini, M. (2013) Interaction between a transition-metal fluoride and a transition-metal hydride: water-mediated hydrofluoric acid evolution following fluoride solvation, *Inorg. Chem. 52*, 12616-12623.

(35) Wilson, G. L., Abraha, M., Krause, J. A., and Guan, H. (2015) Reactions of phenylacetylene with nickel POCOP-pincer hydride complexes resulting in different outcomes from their palladium analogues, *Dalton Trans 44*, 12128-12136.

CHAPTER 3

The LarB carboxylase/hydrolase forms a transient cysteinyl-pyridine intermediate during nickel-pincer nucleotide cofactor biosynthesis

Adapted from an early draft of:

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, *in preparation*

For this chapter, I conducted most of the kinetic characterization of LarB (except capillary electrophoresis) and all of the reported variant analysis. I also developed the crystallization conditions for LarB and conducted all of the crystallography experiments. I participated in model building and structural analysis. Zia Tariq and Dr. Benoît Desguin conducted the capillary electrophoresis experiments. Dr. Satyanarayana Lagishetty aided initial model building. Drs. Shramana Chatterjee and Jian Hu built the LarB models and conducted structural analysis. Drs. Benoît Desguin, Jian Hu and Robert Hausinger oversaw the studies.

Abstract

Enzymes possessing the nickel-pincer nucleotide (NPN) cofactor catalyze C2 racemization or epimerization reactions of α-hydroxyacid substrates. LarB initiates synthesis of the NPN cofactor from nicotinic acid adenine dinucleotide (NaAD) by performing dual reactions: pyridinium ring C5 carboxylation and phosphoanhydride hydrolysis. Here, I show that LarB uses carbon dioxide, not bicarbonate, as the substrate for carboxylation and activates water for hydrolytic attack on the AMP-associated phosphate of (carboxylated)-NaAD. Structural investigations show that LarB has an N-terminal domain of unique fold and a C-terminal domain related to that of aminoimidazole ribonucleotide carboxylase/mutase (PurE). Like PurE, LarB is octameric with four active sites located at subunit interfaces. The complex of LarB with NAD^+ reveals formation of a covalent adduct between the active site Cys221 and C4 of the substrate analog, resulting in a boat-shaped dearomatized pyridine ring. Formation of such an intermediate with NaAD would enhance the reactivity of C5 to facilitate carboxylation. Glu180 is well positioned to abstract the C5 proton, restoring aromaticity as Cys221 is expelled. The structure of as-isolated LarB and its complexes with NAD^+ and the product AMP identify potential residues important for substrate binding and catalysis. In combination with these findings, the results from structure-guided mutagenesis studies lead me to propose enzymatic mechanisms for both of the LarB reactions that are distinct from that of PurE.

Introduction

The nickel-pincer nucleotide (NPN), first identified as a coenzyme of lactate racemase from *Lactobacillus plantarum*,[1] is an organometallic cofactor with nickel covalently bonded in a planar arrangement to carbon and two thioacid sulfur atoms of a modified pyridinium mononucleotide (Figure 3-1A). This cofactor is synthesized from nicotinic acid adenine dinucleotide (NaAD) by the sequential actions of LarB, an NaAD carboxylase/hydrolase,[2] LarE, a sacrificial sulfur transferase,[3, 4] and LarC, a CTP-dependent nickel insertase or cyclometallase.[5] Homologs of the genes encoding these biosynthetic proteins are widespread within eubacteria and archaea.[6]

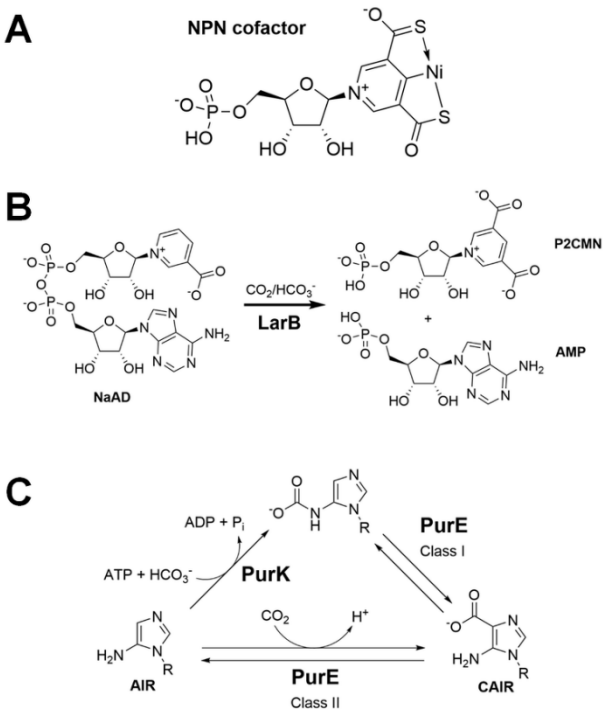


Figure 3-1. The nickel-pincer nucleotide (NPN) cofactor, LarB reaction, and comparison to the reactions of PurE. (A) The NPN cofactor possesses nickel bound to a carbon and two sulfur atoms of a modified pyridinium mononucleotide. (B) LarB catalyzes C5 ring carboxylation and phosphoanhydride hydrolysis of nicotinic acid adenine dinucleotide (NaAD) to release AMP and generate the NPN precursor pyridinium 3,5-biscarboxylic acid mononucleotide (P2CMN). (C) The carboxylation of aminoimidazole ribonucleotide (AIR) to form carboxyaminoimidazole ribonucleotide (CAIR) can occur by a two-step reaction in which an ATP-dependent PurK generates a carbamate using bicarbonate and this intermediate is converted to CAIR by a class I PurE mutase. Alternatively, Class II PurE carboxylases directly produce CAIR using carbon dioxide.

LarB, the subject of this work, acts on its substrate NaAD to carboxylate the C5 position of the nicotinic acid moiety and hydrolyze the phosphoanhydride bond, releasing AMP and forming the product pyridinium 3,5-biscarboxylic acid mononucleotide (P2CMN) (Figure 3-1B). In addition, the enzyme hydrolyzes NaAD to AMP and the noncarboxylated product nicotinic acid mononucleotide (NaMN) in what may be a futile side reaction. From the information

reported thus far, it remains unclear whether LarB uses phosphoanhydride hydrolysis to drive the carboxylation reaction.

The *L. plantarum* LarB sequence (246 amino acids, UniProtKB code: F9UST0) is 25.6% identical over 135 residues to that of N^5-carboxyaminoimidazole ribonucleotide (N^5-CAIR) mutase from the same microorganism (161 residues, UniProtKB code: F9URK5), a class I PurE that functions in bacterial purine nucleotide synthesis.[7] An ATP-dependent aminoimidazole ribonucleotide (AIR) carboxylase (PurK) first synthesizes N^5-CAIR, then PurE transfers the carboxyl group from the carbamate to the C4 position, forming CAIR (Figure 3-1C).[8, 9] In contrast to this two-step sequence, class II PurEs catalyze the direct carboxylation of AIR to CAIR without using ATP. For these enzymes, the reactions are driven by product consumption using the next enzyme in the pathway.[7] Notably, neither class of PurE is associated with any hydrolase activity. *L. plantarum* LarB possesses an N-terminal domain (114 amino acid residues) not found in the corresponding PurE, and the latter protein contains a 27-residue C-terminal region not found in LarB (Figure 3-2). The PurE C-terminus forms an α-helix that makes intermolecular contacts with the main domain of another subunit, contributing to its oligomeric structure,[7, 10] and the catalytic residues of PurE are not conserved in LarB suggesting a distinct carboxylation mechanism.

Here, I show that the substrate used for LarB carboxylation is carbon dioxide, not bicarbonate, and I clarify the mechanism of phosphoanhydride hydrolysis. The structures of as-isolated, NAD^+-bound, and AMP-bound *L. plantarum* LarB are reported and compared to the overall enzyme architecture of the PurE protein fold. I demonstrate the LarB carboxylation active site lies in a region analogous to the substrate-binding site of PurE, but with a different set of conserved residues used for catalysis. Significantly, the crystallographic data suggest a covalent

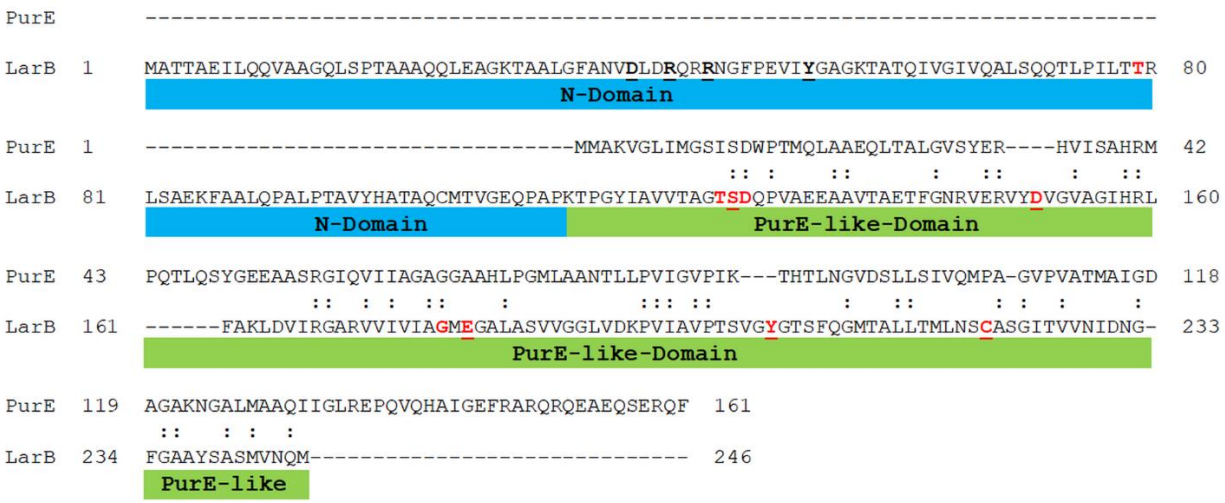


Figure 3-2 Alignment of LarB and PurE, both from *L. plantarum*.

adduct forms between NAD^+ and the invariable Cys221; a similar intermediate generated with NaAD would support its carboxylation. I show the conserved residues from the LarB N-terminal domain, not found in PurE, make contact with the adenosine moiety of the substrate, consistent with this domain being important for binding and hydrolyzing the dinucleotide substrate. I used site-directed mutagenesis to substitute key residues in the two catalytic sites, focusing on those conserved in LarB sequences and lacking in PurE. On the basis of these results, I propose mechanisms for the carboxylation and hydrolysis reactions catalyzed by LarB, which are fundamentally different from the chemistry of PurE.

Materials and Methods

Materials

Desthiobiotin, NaAD, and IDA (iminodiacetic acid) were acquired from Sigma (St. Louis, MO, USA). $H_2{}^{18}O$ was purchased from Sigma-Aldrich (Overijse, Belgium), $NaH^{14}CO_3$ (50-60 mCi/mmol) was purchased from American Radiolabeled Chemicals (St. Louis, MO, USA). Chloramphenicol, carbenicillin, kanamycin, gentamicin, and isopropyl β-D-1-

thiogalactopyranoside (IPTG) were purchased from Gold Bio (St. Louis, MO, USA). All other chemicals were reagent grade or better.

Vector Construction and Mutagenesis.

Plasmids pET:LarB and pBAD:LarB for production of C-terminally Strep II-tagged LarB in *Escherichia coli* were generated by cutting pGIR022[2] with NcoI and HindIII (NEB, Ipswich, MA, USA) to obtain the DNA fragment encoding LarB with the tag, purifying the fragment from an agarose gel slice, and using T4 ligase (Invitrogen, Waltham, MA, USA) to incorporate it into pET28a or pBADHisA that had been cut and purified likewise. The products were transformed by heat-shock into *E. coli* DH5α.[11] Site-directed mutagenesis was carried out using two-stage PCR,[12] gap-repair,[13] or the Q5 Site-Directed Mutagenesis kit (NEB), with mutations confirmed by sequence analysis. The strains, plasmids and primers used in this study are summarized in Table 3-A1.

LarB Purification

For crystallization and most other studies, C-terminally Strep II-tagged LarB from *L. plantarum* was produced in *E. coli* BL21 (DE3) [pBAD:LarB] cells. LB-Lennox medium was supplemented with 100 mg/L carbenicillin (1 L in 2.8 L Fernbach flask), inoculated (1%) with a fresh culture, and grown overnight at 37 °C with shaking at 225 RPM until reaching an OD_{600} of 0.5-1.0. The temperature of the culture was lowered to 25 °C, the cells were induced with 0.2% w/v L-arabinose, and growth continued overnight. Cultures were harvested by centrifugation and cell pellets were stored at -80 °C until use. Expression using *E. coli* BL21 (DE3) [pET:LarB] was performed similarly, supplementing the medium with 50 mg/L kanamycin, lowering the induction temperature to 16 °C, and inducing with 0.2 mM IPTG. For expression using *E. coli* BL21 (DE3) [pET:LarB] or *E. coli* BL21 (DE3) [pBAD:LarB] with the ArcticExpress plasmid

(Agilent), the overnight starter culture was supplemented with 25 mg/L kanamycin and 10 mg/L gentamicin. No antibiotics were used in the induction medium. The temperature was lowered to 12 °C, induced with 1 mM IPTG and grown for 24 h. The *E. coli* cell pellets were resuspended in 100 mM Tris-HCl, pH 7.5 (note: the pH values of all buffers were obtained at room temperature unless otherwise stated), containing 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (from a 100 mM stock in ethanol), and 1 U/mL Benzonase (EMD). Suspended cells were lysed by using a French pressure cell at 16000 psi at 4 °C. Lysates were clarified by centrifugation (either 40 min at 25000 *g* or 45 min at 146000 *g*) at 4 °C. Clarified lysates were loaded onto Streptactin-Superflow high-capacity resin (IBA; note, LarB with the Strep II tag could not be purified using Streptactin-XT because it did not elute from the resin) at 1 mL/min at 4 °C. The column was washed with 5 column volumes of 100 mM Tris-HCl, pH 7.5, containing 150 mM NaCl. LarB was eluted with 3 column volumes of the same buffer containing 5 mM desthiobiotin. Protein concentrations were determined by using the bicinchoninic acid[14] or Bradford (Bio-Rad) assay. Purification of LarB from *Lactococcus lactis* NZ3900 [pGIR026] cells was carried out as previously described.[2] LarB variants were purified using the same conditions as those for the wild-type enzyme and subjected to the standard assay conditions described below.

Physical Characterization of LarB

The subunit size of Strep II-tagged LarB was established by electrospray ionization-mass spectrometry (ESI-MS) analysis, as described previously.[4] The data were processed using UniDec[15] to determine molecular mass. For native size determination by size exclusion chromatography/multi-angle light scattering (SEC-MALS) analysis, LarB was chromatographed on a Sephadex 200 (preparative grade) 16/60 (GE Healthcare) column equilibrated in 10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl at a flow rate of 0.5 mL/min. A 0.2 μm filter was

placed in line before a miniDAWN TREOS MALS detector and TRex refractive index meter (Wyatt). The data were processed using ASTRA (Wyatt). Metal analysis of LarB was performed using both inductively coupled plasma-optical emission spectroscopy (ICP-OES) and inductively coupled plasma-mass spectrometry (ICP-MS). Samples for ICP-OES were prepared by boiling for 1 h in 35% (w/v) nitric acid before analysis using a Varian ICP-OES model 710. ICP-MS was conducted by the Center for Applied Isotope Studies at the University of Georgia.

LarB Enzyme Activity

Three methods were used to assay LarB activity. First, a previously described liquid chromatography (LC)-ESI-MS procedure[2] was used for monitoring the loss of NaAD (m/z = 663 after 3.4 min), the formation of P2CMN (m/z = 376 after 3.4 min), the generation of AMP (m/z = 346 at 2.9 min), and the aberrant production of the hydrolysis product NaMN (m/z = 334, 2.2 min). Secondly, capillary electrophoresis (CE) assays were carried out in buffer containing 75 mM Na_2HPO_4, 25 mM NaH_2PO_4, 100 mM glycine, 5 mM $MgCl_2$, 20 mM $NaHCO_3$, and 1 mM NaAD at a temperature of 25 °C. The assays were terminated by incubation at 90 °C for 5 min. CE analysis was performed using a Capel 105M Capillary Electrophoresis Instrument (Lumex) with a running buffer of 0.1 M sodium borate, pH 8.5, with a capillary diameter of 50 µm and a length of 48 cm (38 cm effective). Analysis was carried out at 20 °C. Prior to and after each run, the capillary was rinsed with water (1 min), 0.5 M NaOH (2 min), water (1 min), and running buffer (3 min) at 1000 psi and 0 V. Samples were injected (5 s) at 1000 psi and analyzed at 25 kV, with detection at 254 nm. Standards were analyzed in the same manner for comparison. The CE assay allowed for simultaneous monitoring of the loss of NaAD, the production of P2CMN, the release of AMP, and the formation of NaMN (Figure 3-3). Thirdly, I assayed substrate carboxylation based on ^{14}C incorporation from radiolabeled CO_2/bicarbonate. The standard

conditions were 50 mM MOPS-NaOH, pH 7.0, containing 4 mM $MgCl_2$, 10 mM $NaHCO_3$, 0.2 µCi $NaH^{14}CO_3$, and 0.2 mM NaAD in a volume of 100 µL. The reactions were incubated at 37 °C and quenched by addition of 10 µL of glacial acetic acid. Samples were dried in a vacuum centrifuge, dissolved in 100 µL of 1% v/v acetic acid, and dried again. The solids were dissolved in 100 µL of water and added to 5 mL of Safety-Solve scintillation solution (RPI). $NaH^{14}CO_3$ specific activity was measured using an alkaline scintillation fluid.[16] Scintillation samples were analyzed using a Beckman-Coulter LS6500 liquid scintillation counter.

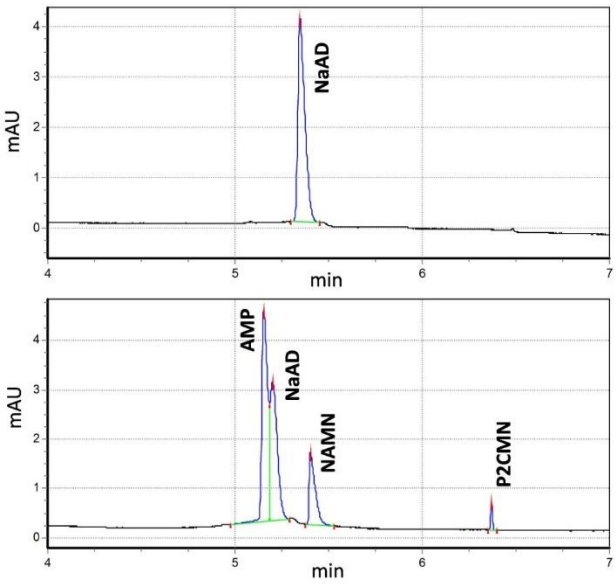


Figure 3-3. A LarB assay method using capillary electrophoresis allows for monitoring NaAD loss and production of AMP, P2CMN, and NaMN.

To explore the role of zinc on LarB activity, two studies were conducted. Using LarB purified by the typical protocol, a portion of the enzyme was incubated with 10 mM EDTA both overnight and after 1 h, desalted using a PD-10 column into wash buffer, and compared in activity to non-EDTA-treated LarB by using the standard assay. In addition, LarB was assayed in the standard assay buffer containing 0.5 mM IDA and varied concentrations (0-225 µM) of Zn sulfate. The apparent K_d of the IDA-Zn complex is difficult to predict given its dependence on pH and salinity, but IDA's useful buffering range has been described as 5 µM to 160 nM $[Zn^{2+}]_{free}$ at pH 7.4 and an ionic strength of 100 mM.[17]

To assess the LarB hydrolysis mechanism, the enzyme was incubated with NaAD in the presence of unlabeled or ^{18}O-water. For mass spectrometric analysis of the resulting AMP

product, 10-μL aliquots were injected into an Accela HPLC system coupled to a Q-Exactive (ThermoFisher) mass spectrometer. Separations were performed using a Kinetex C18 column (150 mm × 2.1 mm) in 0.1% formic acid with a gradient of acetonitrile. The run time for each sample was 25 min.

LarB Crystallization and Structure Determination

Prior to crystallization, 1-2 mL samples of Strep II-tagged LarB were further purified at room temperature by chromatography on a Sephadex 200 (preparative grade) 16/60 (GE Healthcare) column that had been equilibrated in 10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl. LarB was eluted at 1 mL/min and fractions containing the octameric species were combined and concentrated to ~12.5 mg/mL or 5 mg/mL according to UV absorption ($\varepsilon_{280\ nm}$ = 0.55 mL mg^{-1} cm^{-1} according to ProtParam).[18] Drops were set up as a 1:1 or 2:1 ratio of protein to reservoir solution containing 100 mM magnesium formate (not adjusted for pH) and 10-20% w/v PEG 3350. Notably, zinc promotes crystallization, so later drops included 2 mM nitrilotriacetic acid neutralized to a pH of ~7.5 with sodium hydroxide and 0.7 mM zinc sulfate. Multi-well plates were set up at 4 °C. Crystals typically appeared after 1-2 weeks.

The crystals were cryo-protected in reservoir solution containing 25% glycerol or ethylene glycol or 35% PEG 550 monomethyl ester. X-ray scattering data were collected at the Advanced Photon Source, Argonne National Labs, Lemont, IL, USA using the Life Sciences Collaborative Access Team (Table 3-A2). The crystal phase was solved by single-wavelength anomalous dispersion (SAD) after incorporating selenomethionine into the protein following published protocols.[19, 20] The phases of later datasets were solved by molecular replacement with the apoprotein structure using Phaser.[20] The crystals of LarB were transferred to a solution of mother liquor with 5 or 10 mM EDTA and 10 mM of either NAD^+ (disodium salt) or AMP

(monosodium salt). The crystals were allowed to soak overnight at 4 °C then cryoprotected in 25% v/v ethylene glycol.

The dataset of a SeMet crystal collected at the K-edge of selenium was indexed and scaled using HKL2000.[21] The initial phase was solved by SAD in Phenix.autosol at 4.9 Å and extended to 3.4 Å. The electron density map was good enough to allow for model building of the C-terminal PurE-like domain with the aid of a homology model generated by SWISSMODEL[22] using the PurE structure (PDB ID: 1D7A) as the template. The PurE-like domain homodimer was used to solve the structure of a native crystal diffraction dataset at 2.8 Å through molecular replacement by Phaser. Iterative model building in COOT[23] and refinement using Phenix.refine were conducted to gradually reduce R_{free} from 45% immediately after molecular replacement to below 30%, allowing for tracing and modeling most of the residues in the N-terminal domain, except for the first 45 less conserved residues. Eight and ten more residues at the N-terminal domain have been modeled in chain A and chain E in the Zn-bound and NAD^+-bound structures, respectively. Crystallographic statistics of data collection and refinement are listed in Table 3-A2. To confirm the identity of the bound metal as zinc in the as-isolated LarB crystal, diffraction datasets were collected at the wavelength of 1.277 Å and 1.305 Å, respectively.

Results

General Properties of LarB

I used LC-ESI-MS to determine the monomer size of LarB. The mass of 26357 Da (Figure 3-4A) is consistent with the expected size of 26360 Da for LarB missing its N-terminal methionine. SEC-MALS studies showed the molecular mass in solution is M_r = 202,700 Da

(Figure 3-4B). This is consistent with LarB forming an octamer (expected mass of 210880 Da), the quaternary structure also seen in the homolog, PurE.[7]

Using both the ^{14}C incorporation and CE assays, I determined a pH optimum of 6.5-7 (Figure 3-5) for the carboxylation reaction; thus, unless stated otherwise, all assays were carried out at pH 7.0. Nonproductive hydrolysis of NaAD by LarB exhibited a distinct, broad profile from pH 4 to 10 (Figure 3-5B, right). The isotope incorporation assay was used to determine a K_m for NaAD of 14.3 ± 1.8 µM and an apparent K_m for $NaHCO_3$ of 28 ± 16 mM (equivalent to a K_m for CO_2 of ~3.5 mM) (Table 3-1). These values are in good agreement with the parameters reported previously (13 ± 3 µM for NaAD and 33 ± 18 mM for bicarbonate) using an indirect lactate racemase activation assay.[2] These results are similar in magnitude to those reported for the representative PurE from *Treponema denticola* (K_m for AIR of 63-64 µM and K_m for CO_2 of 11-13.2 mM).[24] $MgCl_2$ is not a substrate, but upon varying its concentration I measured an apparent K_m for this cation of 0.42 ± 0.10 mM (Table 3-1), again similar to the value (0.55 ± 0.18 mM) previously measured with the indirect assay.[2]

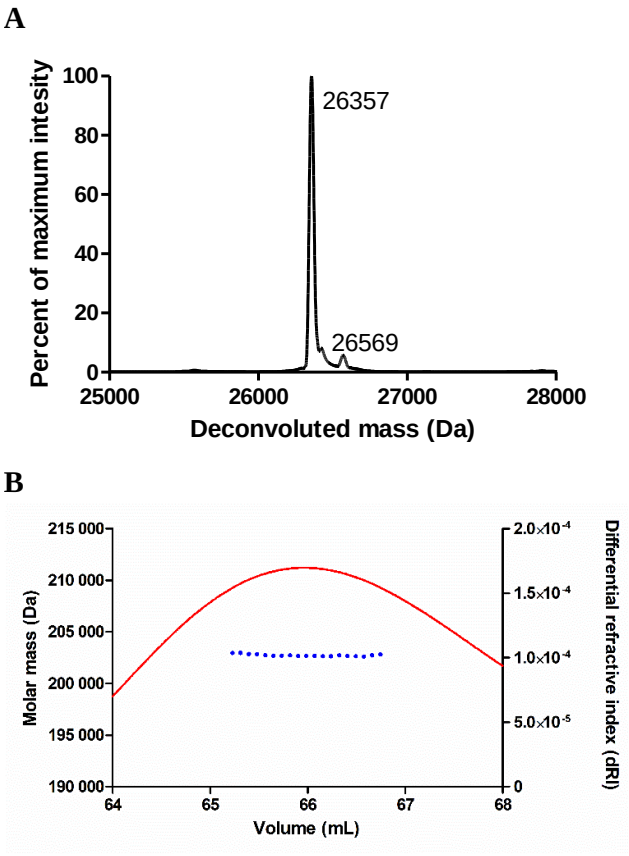


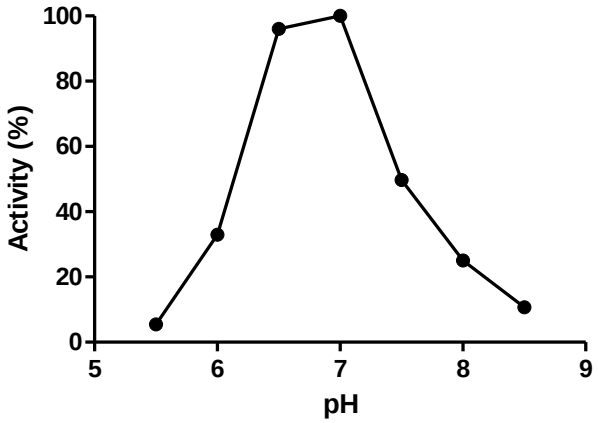
Figure 3-4. Subunit and native sizes of LarB. (A) ESI-MS profile of the Strep II-tagged LarB subunit (*m* = 26357 Da) that was eluted from a Streptactin column with desthiobiotin. The smaller peak (*m* = 26569 Da) may represent protein-bound desthiobiotin (214 Da in its protonated form). (B) SEC-MALS analysis of the octameric protein. The red line is the differential refractive index. The blue dots are the calculated molecular mass.

Table 3-1. Kinetic parameters of LarB[a]

Substrate	k_{cat} (s^{-1})	K_m
NaAD[b]	0.0103 ± 0.0005	14.3 ± 1.8 µM
$NaHCO_3$[c]	0.0354 ± 0.0034	28 ± 16 mM
$MgCl_2$[d]	0.0102 ± 0.0034	0.42 ± 0.10 mM

[a]Determined by using the ^{14}C incorporation assay, and includes 95% confidence intervals.
[b]Kinetic parameters measured using constant 10 mM $NaHCO_3$ and 4 mM $MgCl_2$ with varied [NaAD].
[c]Apparent kinetic parameters measured using constant 0.2 mM NaAD and 4 mM $MgCl_2$ with varied [$NaHCO_3$].
[d]Apparent kinetic parameters measured using constant 0.2 mM NaAD and 10 mM $NaHCO_3$ with varied [$MgCl_2$].

A



B

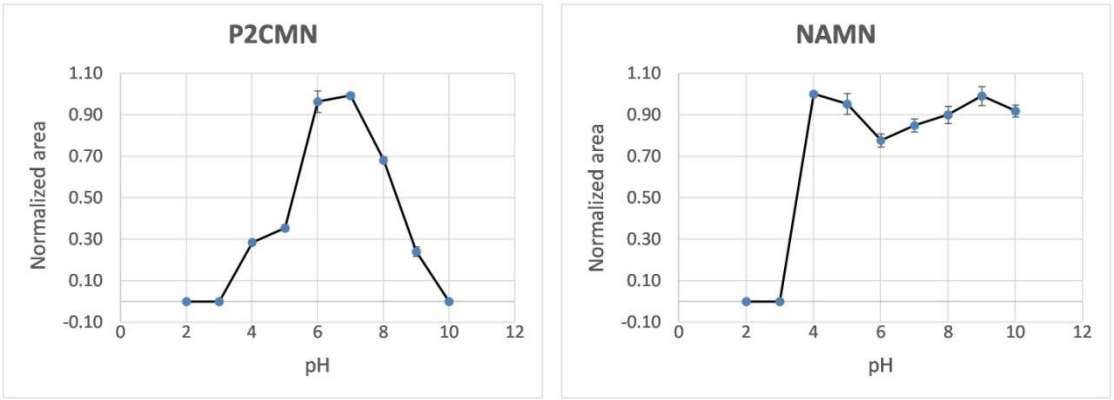


Figure 3-5. LarB exhibits a pH optimum of 7 for carboxylation. (A) LarB pH dependence monitored by the incorporation of ^{14}C from labeled CO_2/bicarbonate using the standard assay with the following buffers: Tris-HCl, pH 8.5, 8.0, and 7.5; MOPS-NaOH, pH 7.5, 7.0, and 6.5; MES-NaOH, pH 6.5, 6.0, and 5.5 (buffer pH values determined at 37 °C, n = 1). (B) pH dependence assessed by monitoring P2CMN or NaMN production using CE (n = 3).

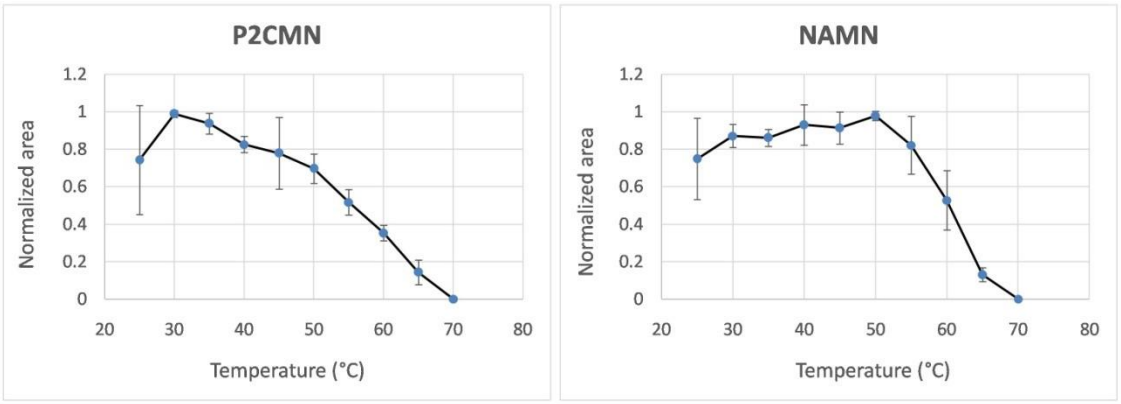


Figure 3-6. The LarB carboxylation reaction exhibits a temperature dependence that is distinct from the hydrolysis activity. CE was used to monitor: (left) P2CMN formation and (right) generation of the side product NaMN (n = 3).

Both the isotope incorporation and CE assays demonstrated that manganese, cobalt, and nickel can substitute for magnesium. In contrast, provision of calcium resulted in 12% activity and both copper- and zinc-treated samples were inactive. The temperature dependence of LarB was assessed by incubating enzyme with its substrates at various temperatures in 100 mM phosphate-glycine buffer, pH 7, and monitoring product formation by using the CE assay (Figure 3-6). The apparent optimum temperature for P2CMN formation was ~30 °C, whereas the apparent optimum temperature for production of NaMN and AMP was broader, from 30-50 °C. Thus, the carboxylation reaction was more sensitive to temperature than was the hydrolysis reaction and these activities are unlikely to be coupled.

Substrate Specificities of LarB for the Carboxylation and Hydrolysis Reactions

The radioisotope incorporation assay provided direct evidence that NaMN cannot serve as a LarB substrate. This assay also was used to test whether LarB could carboxylate several alternative nucleotides (NAD^+, NADH, NMN, and nicotinic acid), and only NAD^+ led to detectable ^{14}C incorporation (~0.26% of that measured using NaAD). I also used the ^{14}C incorporation assay to test whether various nucleotides could inhibit NaAD carboxylation by

LarB. The enzyme was slightly inhibited by 1 mM NAD^+, NADH, NMN, and NaMN (Figure 3-7). Neither the LarB product AMP nor nicotinic acid were not inhibitory to the enzyme at this concentration.

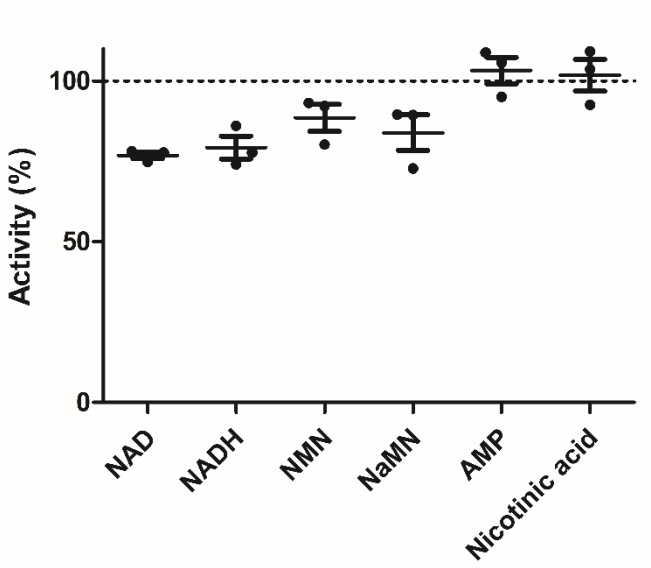


Figure 3-7. Carboxylation activity of LarB is only slightly inhibited by selected nucleotides at 1 mM concentrations (*n* =3, SEM).

LarB was incubated with NaAD in the absence of bicarbonate and the futile hydrolysis reaction allowed to go to completion, thus consuming all available NaAD. Subsequent addition of sodium bicarbonate failed to produce observable levels of P2CMN as monitored by the CE assay. Addition of fresh NaAD to this reaction mixture led to P2CMN formation, demonstrating the enzyme retained activity under these conditions. This result confirms the inability of LarB to carboxylate NaMN, as measured by ^{14}C incorporation (see above). Furthermore, prior studies had shown that NaMN is not utilized as a substrate for NPN synthesis based on a lactate racemase activation assay.[2]

In addition to NaAD, LarB was shown to hydrolyze NAD^+ and the acetylated analog of NaAD, 3-acetylpyridine adenine dinucleotide (Figure 3-8). NAD^+ and bicarbonate were added to the sample obtained after NaAD hydrolysis in bicarbonate-free solution; no P2CMN was observed. This result demonstrates that hydrolysis of NAD^+ cannot be coupled to the carboxylation of NaMN.

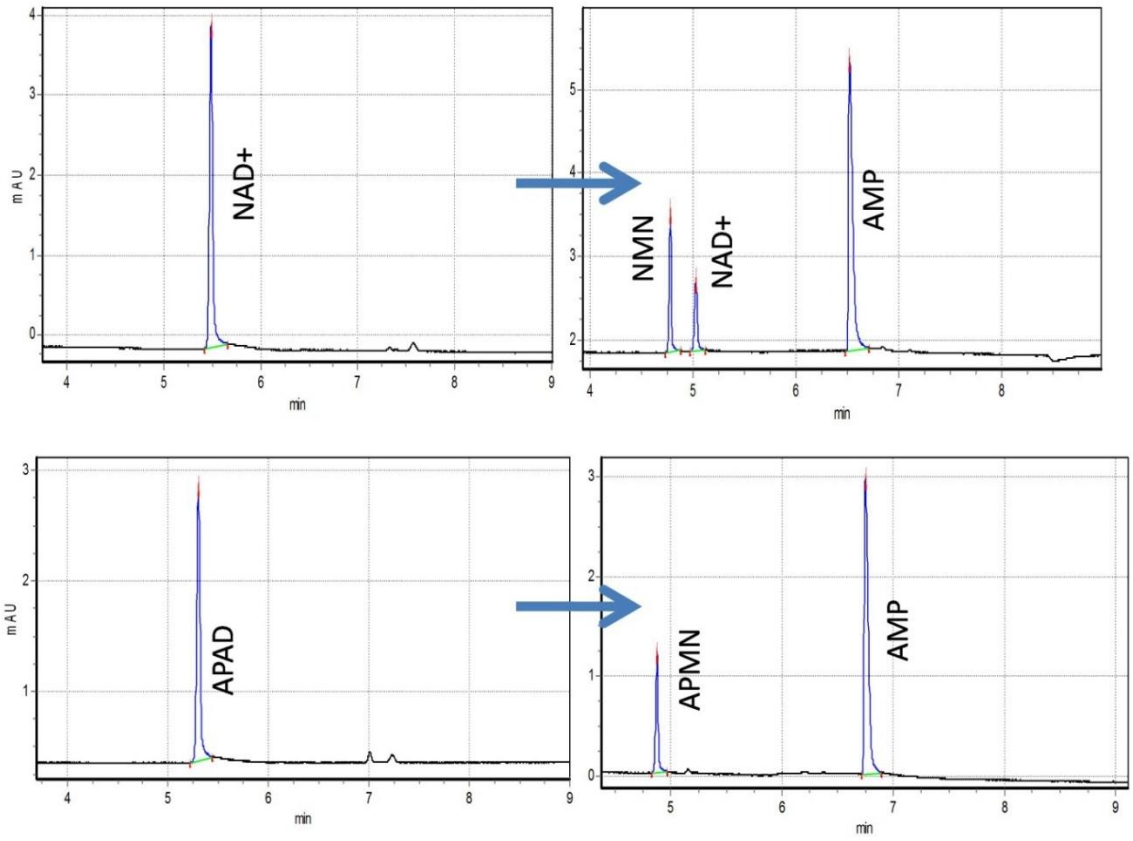


Figure 3-8. LarB uses NAD^+ and 3-acetylpyridine dinucleotide (APAD) as alternative substrates for hydrolysis. LarB (1 µM) was incubated overnight with NAD^+ or APAD (0.2 mM) at 30 °C in Tris-HCl buffer (100 mM, pH 8). Approximately 82% of NAD^+ was converted to NMN and AMP, whereas APAD was completely hydrolyzed into APMN and AMP. For comparison, analogous incubation conditions with NaAD completely converted the substrate to its products.

CO_2, not Bicarbonate, is the Carboxylation Substrate of LarB.

I investigated whether CO_2 or bicarbonate serves as the substrate of LarB using buffer at pH 8.0, where the equilibrium greatly favors bicarbonate. A ^{14}C isotope dilution experiment[25] demonstrated that the rate of radioactivity taken up from the bicarbonate/CO_2 mixture was transiently reduced when the sample was vigorously mixed with non-labeled CO_2 at 75 s (Figure 3-9). In contrast, when carbonic anhydrase was included in the assay the rate of incorporation was immediately reduced to the same final rate from the first experiment, as expected for the rapid conversion of CO_2 to bicarbonate and overall dilution of the inorganic carbon pool. The cessation of isotope incorporation was consistent with CO_2 being the authentic substrate.

Dinucleotide Hydrolysis Involves Water Attack on the AMP Phosphate

LarB-catalyzed hydrolysis of NaAD or its dicarboxylated derivative could occur by nucleophilic attack of water on the phosphate group that is proximal or distal to the pyridinium ribose moiety. To establish the chemistry of the hydrolytic reaction, we carried out the LarB reaction in buffer containing $H_2{}^{18}O$ and examined the products by ESI-MS. This approach allowed us to identify ^{18}O-containing AMP (Figure 3-10), demonstrating the reaction occurs by nucleophilic water attack on the AMP-associated phosphate group.

Structure of LarB

L. plantarum LarB isolated from *E. coli* was crystallized and structurally characterized. The phase was solved by SAD using a selenomethionine-substituted crystal, and the initial structural model was later used to solve the crystal structure of the native protein

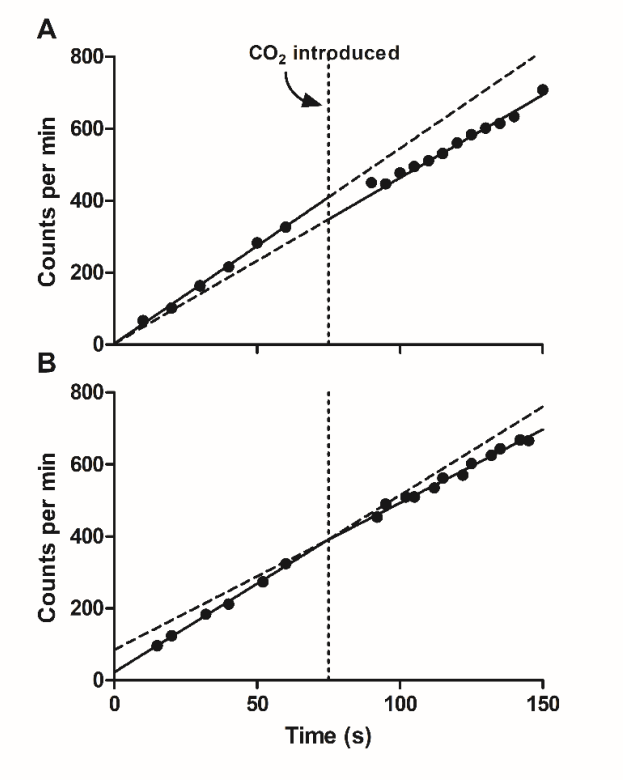


Figure 3-9. Carbon dioxide, not bicarbonate, is the substrate of LarB. (A) Short-term incorporation of ^{14}C into substrate using radiolabeled CO_2/bicarbonate, with unlabeled CO_2 added at 75 s for isotope dilution. (B) The same experiment in the presence of carbonic anhydrase.

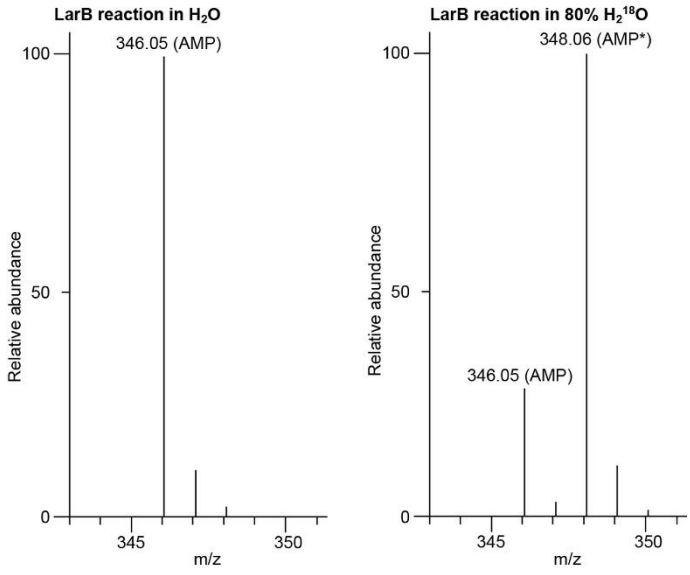


Figure 3-10. LarB incorporates ^{18}O from labeled solvent into AMP during NaAD hydrolysis. The figure depicts the mass spectra of LarB-generated AMP derived from NaAD in unlabeled water (left) and ^{18}O-labeled water (right).

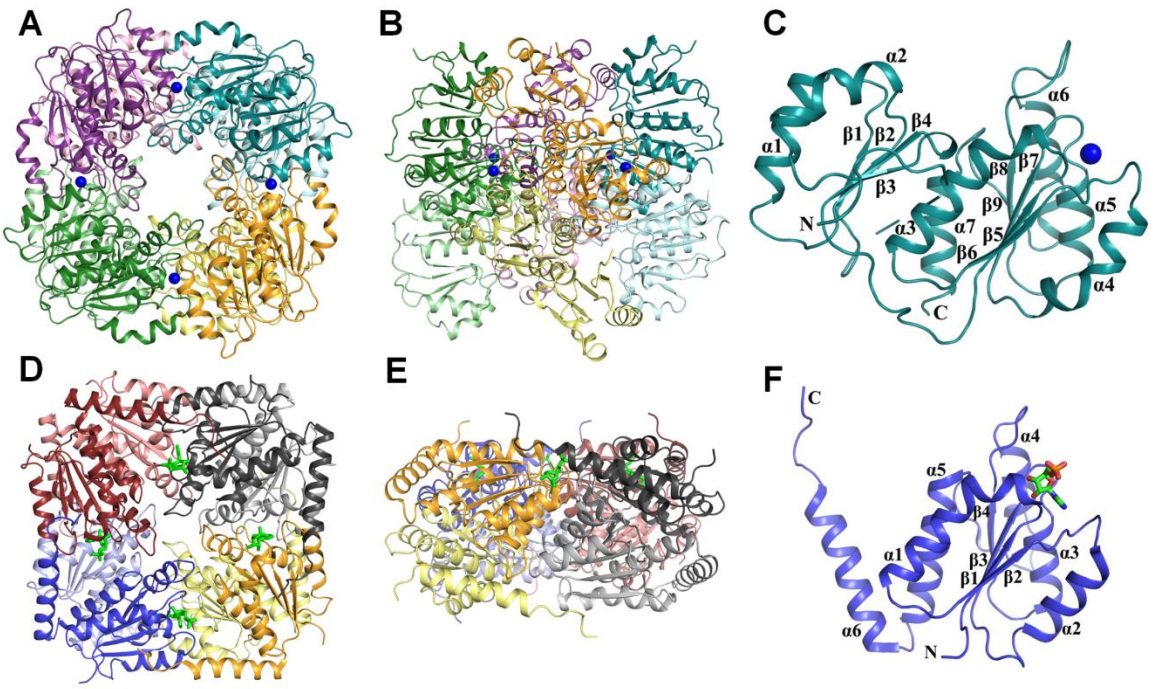


Figure 3-11. Structure of *L. plantarum* LarB and comparison to PurE. (A and B) Two views of the octameric LarB structure with each subunit depicted in cartoon mode using different colors (or different shades for the dimer substructures). The blue spheres represent the metal ion associated with the as-isolated enzyme. (C) LarB monomer. (D and E) Two views of the octameric PurE from *A. aceti* with bound AIR (green sticks, PDB 5cli). (F) PurE monomer with bound AIR.

at a resolution of 2.8 Å via molecular replacement (Table 3-A2). The asymmetric unit contains six LarB molecules organized into three dimers (chains A & B, chains C & D, and chains E & F) with overall root mean square deviation (RMSD) of 0.48 Å between the two monomers in each dimer and with 0.25 Å (A:B dimer vs. C:D dimer), 0.59 Å (A:B dimer vs. E:F dimer), and 0.66 Å (C:D dimer vs. E:F dimer), respectively. In the crystal lattice, each dimer is organized into a tetramer of dimers with its symmetry mates, forming an octamer containing a four-fold axis running through the central tunnel (two views are shown in Figure 3-11A and B), which is in agreement with the SEC-MALS results for LarB in solution (Figure 3-4B). These findings suggest that the octameric form is physiologically relevant. The overall octamer is like a square box with dimensions of ~75 Å along its two edges and 75 Å in thickness. The tetramers that form the top and bottom halves are related by pseudo-two-fold symmetry when considering the

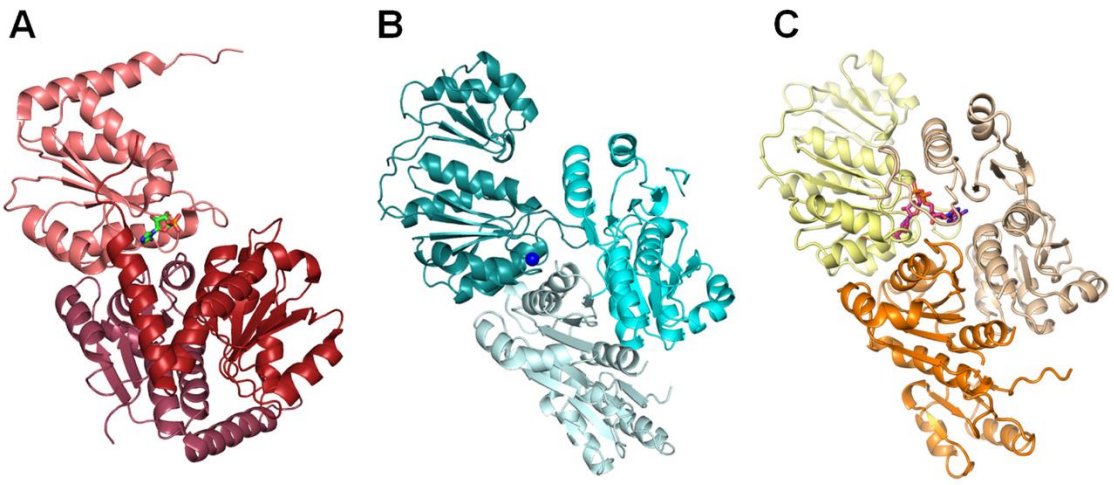


Figure 3-12. Active site location comparison of PurE and LarB. (A) The active site of PurE is located at the interface of three subunit chains (shown by different shades of red in cartoon view) as revealed by the location of AIR (depicted in stick mode with green carbon atoms). (B) The as-purified LarB has density corresponding to a metal ion (blue sphere) at the analogous intersection of three subunit chains (varied shades of blue). (C) The LarB (varied shades of yellow) with bound NAD^+ (sticks with magenta carbon atoms).

monomers as equivalent components. Compared with the other two dimers, the E:F dimer exhibits much poorer electron density, indicative of higher level of mobility in the lattice. The quaternary structure of LarB closely resembles that of PurE[7, 10] (shown for the protein from *Acetobacter aceti*, PDB access code: 5cli; Figure 3-11D and E), but the PurE octamer is thinner because of the lack of the N-terminal domain which is conserved in LarB proteins (Figure 3-2).

Each LarB monomer contains two domains (Figure 3-11C), designated as the N- and C-terminal domains, with a long cleft between them. The larger C-terminal domain consists of five parallel ß-sheets with two α-helices on one face and three on the other, representing the same fold as found in PurE (Figure 3-11F)[7, 8, 10] with an RMSD for Cα atoms of 1.6 Å despite the low (~20%) sequence identities between these proteins. Accordingly, the C-terminal domain is referred to as the PurE-like domain. In contrast to the PurE C-terminal α-helix that forms inter-monomer contacts within the octamer, the N-terminal domain occupies this region in the LarB structure and helps to facilitate intersubunit interactions. The DALI search engine[26] found no

compelling structural matches with the N-terminal domain of LarB; this α/β fold is similar to the d.68 IF3-like fold of the SCOPe database.[27] The initial 45 residues of the N-terminus are disordered in most chains (37 residues for chain A), and non-interpretable density was present. This region is not well conserved in LarB homologs (Figure 3-A1).

The octamer of LarB contains a cavity (maximum dimension of ~29 Å) at its center. The opening of the cavity is ~10 Å and lined by Gln209, whereas the interior of the cavity exhibits primarily hydrophobic residues, including Phe208 and Met211. Subunit interactions occur between the two halves (top and bottom) of the octamer. Analysis of the interactions for each pair of chains (Table 3-A3) indicates the stability of each dimer. The average interface area between monomers in each dimer is ~1165 $Å^2$ as calculated by *PDBe PISA* web server.[28]

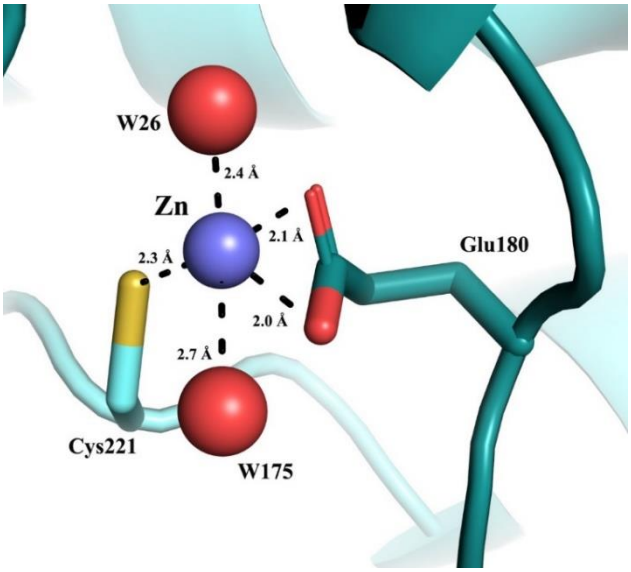


Figure 3-13. Close-up view of the interfacial metal bound to as-isolated LarB

The active site of LarB was tentatively identified by comparison to that of PurE (Figures 3-11C,F and 3-12A,B). The PurE active site binds its substrate AIR at the interface of three chains. The corresponding interfacial region of LarB possesses additional strong density at this position consistent with a metal ion. The metal is coordinated by the bi-

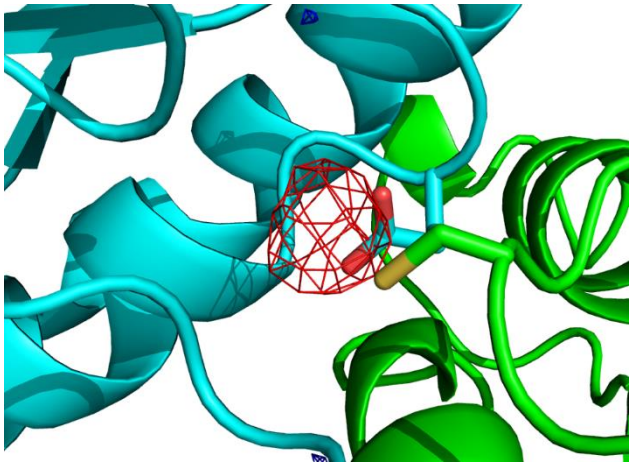


Figure 3-14. Identification of Zn in LarB. Anomalous difference map (red) using 1.28 Å and 1.31 Å (slightly greater and slightly below, respectively, the zinc K-edge), σ = 4.0.

dentate carboxylate of Glu180 from one monomer, the thiolate of Cys221 from another monomer, and two water molecules (Figure 3-13). Residues corresponding to Glu180 and Cys221 are not present at the active site of PurE proteins but are invariable in LarB sequences (Figures 3-2 and 3-A1). Most other residues at the putative LarB active site are hydrophobic and unlikely to be catalytically relevant. LarB activity requires magnesium,[2] although manganese, cobalt, or nickel can substitute, but the extra density was inconsistent with magnesium. The identity of the metal was established as zinc based on anomalous scattering of this density in the crystals (Figure 3-14). Furthermore, I demonstrated the presence of substoichiometric levels of Zn in purified LarB by using ICP-MS and ICP-OES. Subsequent experiments demonstrated that activity was not lost by EDTA treatment and that zinc was an inhibitor of the enzyme (see below). The observed zinc binding to LarB was consistent with a crystallization artifact where zinc ions were introduced into the protein sample during purification and/or crystallization.

Analysis of LarB Interaction with Zinc

To assess whether the zinc observed in the as-isolated LarB structure is necessary for the enzyme's activity, I treated a sample of purified protein overnight with EDTA (0.5 or 10 mM), removed the chelator by gel filtration chromatography, and assayed the remaining activity. Neither treatment reduced enzyme activity as measured by using the ^{14}C incorporation assay, thus demonstrating the metal is not necessary for catalysis. Of interest, EDTA-treated LarB did not generate crystals; however, this deficiency was

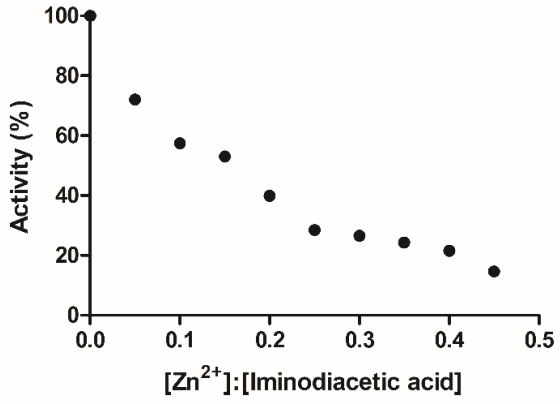


Figure 3-15. Zinc inhibits the carboxylase activity of LarB. The enzyme was assayed using standard buffer amended with 500 µM IDA and varied concentrations of Zn^{2+} (0-225 µM). The relative activities were compared on the basis of the incorporated radioactivity.

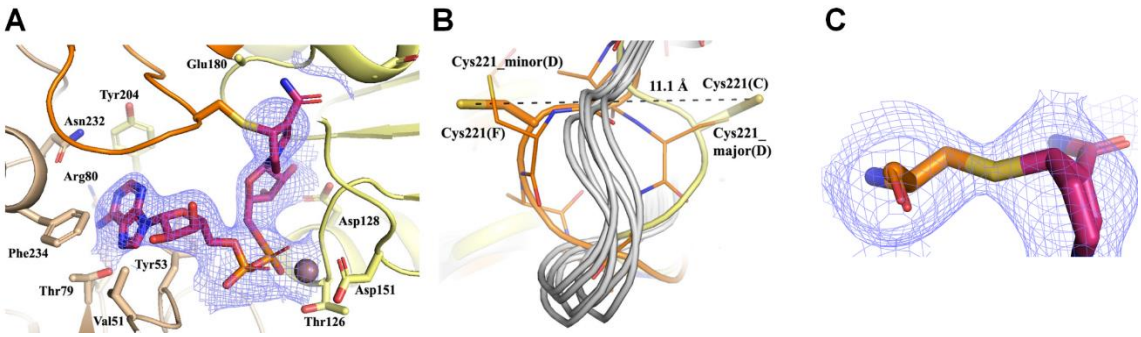


Figure 3-16. The NAD$^+$ binding site of LarB and evidence for a covalent linkage between C4 of the pyridinium ring and Cys221. (A) NAD$^+$ binds to LarB in the same location as Zn and analogous to the AIR binding-site in PurE. The surrounding region is depicted using cartoon view with different colors for the three subunits (chain E is light yellow and wheat, chain F is orange), Mg is shown as a sphere, and NAD$^+$ is shown as magenta sticks with the Fo-Fc omit map (σ = 3). (B) Dual conformation of Cs221 in the structure of the LarB-NAD$^+$ complex compared to PurE. The LarB region containing Cys221 (residues 219-223) is shown for chain F (orange cartoon, NAD$^+$ is bound), chain C (light yellow cartoon, lacking NAD$^+$), and chain D (orange sticks in two conformations, with 0.7 and 0.3 occupancies. Interchange of the two conformations requires ~11.1 Å flipping of this loop, associated with binding of the substrate analog. For comparison, the corresponding region is shown as gray cartoons for six PurE homologs (PDB: 5CLI, 2YWX, 1O4V, 3TRH, 3LP6, 4AY3). (C) Zoomed-in view of the covalent linkage between Cys221 and the pyridine ring derived from NAD$^+$ with Fo-Fc omit map (σ = 3).

overcome by providing the metal ion as a Zn-nitrilotriacetic acid complex. Significantly, after obtaining LarB crystals the metal was able to be removed by EDTA treatment without compromising diffraction.

To establish the effect of added zinc on LarB activity, I varied the concentration of this metal ion in buffer containing a uniform concentration of IDA and measured the enzyme's ability to incorporate ^{14}C-labeled bicarbonate into the substrate. Zinc was estimated to inhibit LarB with an IC_{50} in the range of 500 nM to 5 µM (Figure 3-15). This value is well above the estimated fM intracellular free zinc concentrations of *E. coli*,[29] so the inhibition of LarB by zinc is unlikely to be biologically relevant. Inhibition by manganese, cobalt, and nickel was not observed at these metal ion concentrations.

LarB-Ligand Complex Structures

Although a substrate-bound LarB structure was not obtained, structures were determined for complexes of enzyme with the substrate analog (and weak inhibitor) NAD$^+$ (at 3.4 Å) and

with the product AMP (3.0 Å) (Table 3-A2). In both cases, the structures were obtained by soaking the ligands into pre-formed crystals that had been treated with EDTA.

In the NAD^+-bound LarB octamer, the two tetramers adopt slightly different conformations. One tetramer is composed of monomers more tilted relative to the four-fold axis, creating more compact active sites and allowing for efficient NAD^+ binding. By contrast, the binding sites in the other tetramer are too loose to bind the ligand. NAD^+ is present at the interface of three chains, similar in location to the Zn-binding site in the as-isolated enzyme (Figure 3-12); however, only chains E/F and C/D bind NAD^+ (Figures 3-16 and 3-17). The nicotinamide moiety appears more dynamic, with a higher *B*-factor and less density fit compared to the rest of the ligand.

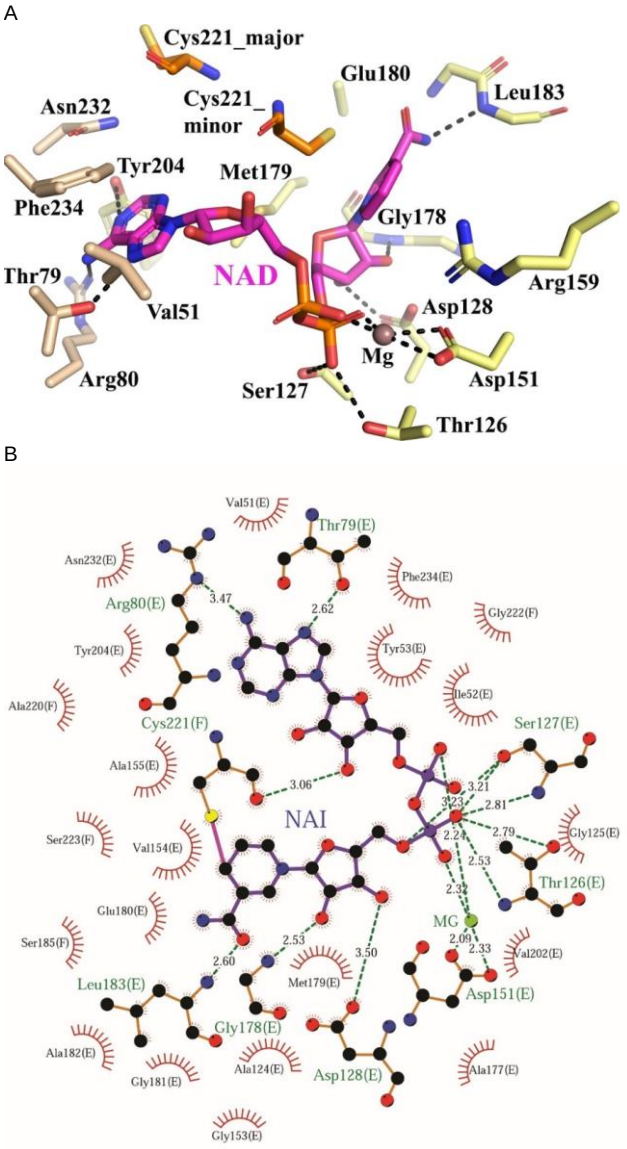


Figure 3-17. LarB residues interacting with bound NAD^+. (A) Close contacts of NAD^+ (shown as magenta sticks) bound at the interface of three subunits (chain C in wheat and light yellow; chain D in orange); Mg^{2+} shown as a sphere. (B) LigPlot of residues proximate to the covalently-bound Cys221-pyridine ring (NAI).

Due to disorder, we were unable to place the side chain of Glu180, except for chain B, where NAD^+ is not bound. The secondary structures for Zn-bound and NAD^+-bound LarB overlap well, except for the flexible loops, with RMSD of 0.86 Å. For NAD^+-bound LarB, chains E and F

exhibit reduced flexibility compared to the more disordered chains in the Zn-bound species, which allowed us to resolve an additional 10 residues at the N-terminus of chain E.

The LarB residues interacting with NAD^+ (Figures 3-16A and 3-17) are conserved in LarB homologs, but generally not in PurE sequences (Figure 3-2). The adenine-binding β8-α6 loop (containing Tyr204) of LarB has no counterpart in PurE, which lacks any aromatic side chain in this region. Thr79 forms a hydrogen bond with this nucleotide base. The adenine ribose is stabilized by a hydrogen bond with the backbone carbonyl of Cys221. The phosphate groups interact with Thr126, Ser127, and Asp151, whose conformation is fixed by a salt bridge to Arg159. Accompanying the phosphate groups of NAD^+ is a Mg ion, which is also coordinated by Asp151. The nicotinamide ribose is stabilized by Asp128 and the backbone carbonyl of Gly178. The pyridinium ring of NAD^+ occupies the same position and adopts a similar orientation in the active site of LarB as AIR in PurE (Figures 3-12 and 3-18), although the two enzymes do not share any catalytic residues. Overall, the LarB-NAD^+ structure and the sequence alignments reveal that LarB and PurE possess similar residues for binding the mononucleotide portion of the substrate, whereas the residues binding the AMP portion of the substrate are unique to LarB homologs. Importantly, the completely different sets of putative catalytic residues between the

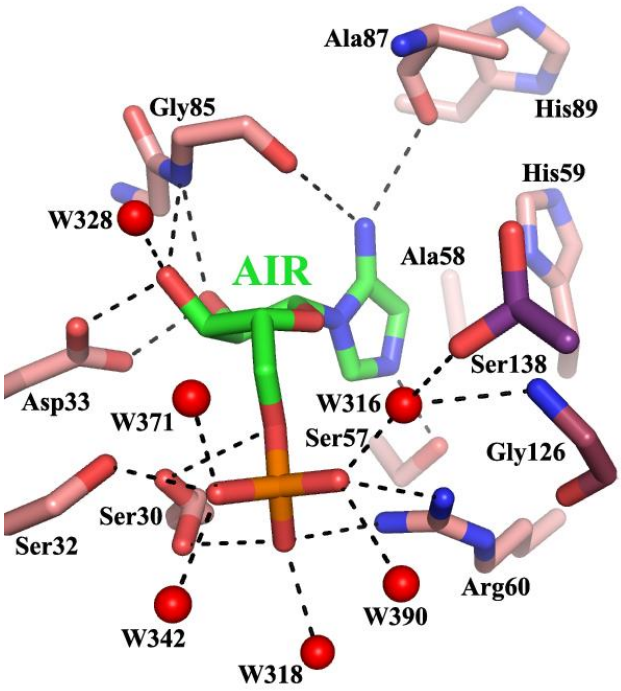


Figure 3-18. Close contacts of AIR bound at the interface of three PurE subunits (salmon, violet-purple, and raspberry). Waters are shown as red spheres.

two enzymes strongly indicate that LarB utilizes a catalytic mechanism distinct from PurE. Of special importance, Cys221 exhibits two positions in the LarB-NAD$^+$ complex (Figure 3-16B) with this residue pointing toward NAD$^+$ for chain F and away from the substrate analog for chain C. Indeed, the electron density map of the region containing Cys221 (residues 219 to 223) of chain D clearly shows an equilibrium of dual conformations – a minor conformation with Cys221 pointing towards NAD$^+$ and a major conformation with the 219-223 segment folding as a part of an α helix and Cys221 facing away from the active site (Figure 3-16B). The ~11.1 Å flipping of Cys221 is unlikely to be induced by ligand binding because the same dual conformations are also observed in the Zn-bound structure where only half of the active sites are occupied by Zn. For comparison, structural superposition of PurE homologs suggests that the presence of dual conformations is unique to LarB (Figure 3-16B). Of greatest interest is the situation for chain F where the electron density profile supports a linkage between the Sγ of Cys221 and C4 of the pyridine ring (Figure 3-16C). The S-C distance of ~1.8 Å and the boat-like configuration of the

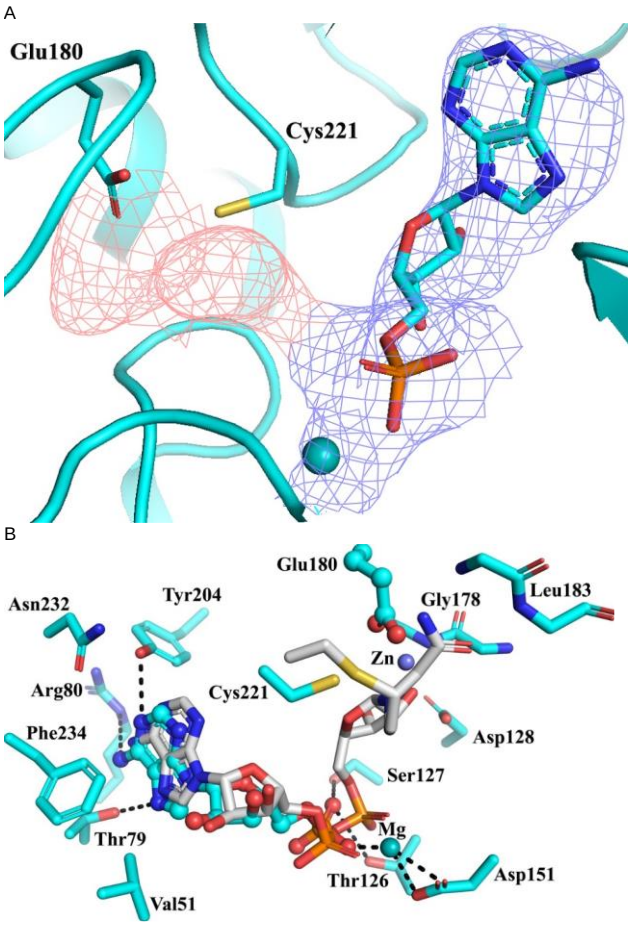


Figure 3-19. Structure of the LarB-AMP complex. (A) 2Fo-Fc map (σ = 1) indicating the position of Mg·AMP (blue mesh) and additional uninterpretable electron density (salmon mesh). (B) Overlay of the LarB-NAD$^+$ and LarB-AMP active site structures. AMP and the position of Glu180 in that structure are show as ball-and-sticks with cyan carbon atoms, and residues in close contact with AMP are shown as cyan sticks. NAD$^+$ is shown as sticks with gray carbon atoms (Glu180 is disordered in this structure). Mg and Zn are shown as teal and slate spheres, respectively.

pyridine ring required for better fitting the electron density are both consistent with a covalent bond. Formation of an adduct with the pyridinium ring stabilizes Cys221, thus explaining the improved *B*-factor of this chain with bound NAD^+ compared to chains C/D.

The AMP-bound structure of LarB was obtained by soaking LarB crystals with 10 mM AMP. The LarB-AMP complex possessed clear density for AMP along with an uninterpretable positive density between Glu180 and the phosphate group (Figure 3-19A). In addition, the LarB-AMP complex includes a Mg ion that is present in the same position as in the NAD^+-bound structure. The overall RMSD is 0.91 Å between the LarB-AMP and LarB-NAD^+ structures. Most of the secondary structures superpose well, with changes observed in the flexible loops. The structure of the LarB-AMP complex demonstrates that this ligand binds in a manner similar to the AMP portion of NAD^+ (Figure 3-19B). Unlike the NAD^+-bound LarB structure, where Glu180 is disordered, this residue exhibited continuous electron density in the AMP-bound chain.

LarB Variant Analyses

I targeted for substitution several conserved residues at the zinc- and ligand-binding sites of LarB along with residues in the N-terminal region and analyzed the variant proteins by monitoring the extent of ^{14}C incorporation after 10 min assays (Figure 3-20). When Cys221, the residue that coordinates zinc and covalently attaches to the pyridinium ring of NAD^+, was substituted with either serine or alanine, the carboxylation activity was abolished. I substituted Glu180, the other

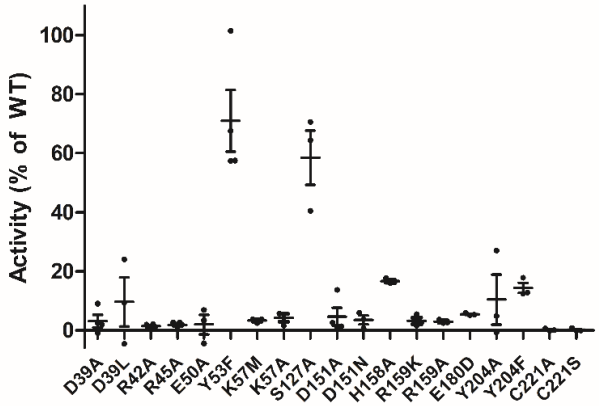


Figure 3-20. The specific activity of each variant is plotted as the percent of wild-type LarB activity. The lines represent the mean, and the error bars the SEM (n = 3 or more).

zinc-binding and potential carboxylation active-site residue, with alanine, glutamine, and aspartate; the E180A variant was poorly expressed and not further analyzed, the E180Q protein was predominantly insoluble with no activity for the soluble fraction, and the E180D enzyme retained ~5% of the activity of wild-type LarB. The D128A variant, affecting the residue responsible for binding the nicotinamide ribose moiety, was mainly insoluble, with no activity noted for the minor soluble portion (not included in the figure). Tyr204, involved in positioning the adenine portion of NAD^+, was substituted with phenylalanine leading to greatly diminished, but still partial, activity, consistent with a substrate-binding role for this aromatic group. In contrast, nearly full activity was observed for the Y53F variant affecting Tyr53 in the amino terminal domain that lies near the ribose in the AMP portion of NAD^+. Substitutions of other N-terminal domain residues Asp39, Arg42, and Arg45, positioned more distant to the NAD^+, led to the nearly complete loss of activity. Of particular interest for the studies described below, substituting alanine for Ser127, a conserved LarB residue that hydrogen bonds with the pyridine-linked phosphate group, resulted in only 40% loss of activity. A serine also is found at this position in PurE.

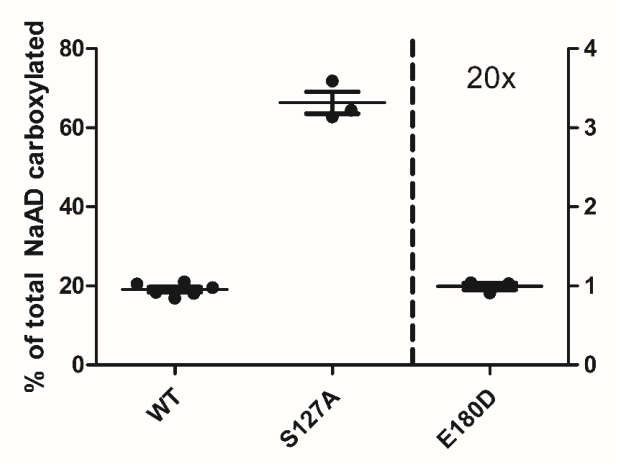


Figure 3-21. The S127A variant of LarB exhibits enhanced ^{14}C incorporation activity relative to the wild-type enzyme after overnight incubation. The overnight ^{14}C incubation for the E180D variant was only 5% of that for wild-type LarB, the same ratio of isotope incorporation as observed for the 10 min assay (Figure 3-20).

In addition to carrying out the above short-term ^{14}C incorporation assays, I assessed the amount of label incorporated into NaAD by selected LarB variants after overnight incubation. When LarB was incubated overnight with NaAD and ^{14}C-labeled bicarbonate/CO_2, the

incorporated radiolabel accounted for ~20% of the NaAD concentration (Figure 3-21). I attribute this low extent of labeling to the significant amount of non-productive enzymatic hydrolysis of NaAD to yield NaMN, which is not a substrate of LarB. The ratio of ^{14}C incorporation into substrate by the E180D variant and wild-type enzyme after overnight incubation (~1% vs. 20%, Figure 3-21) was similar to the ratio of incorporation after 10 min (~5% vs. 100%, respectively, Figure 3-20), perhaps indicating a nearly intact hydrolysis activity. Of keen interest, the amount of ^{14}C incorporated into NaAD after extended incubation with the S127A variant corresponded to a much greater portion of initial substrate (~66%); this result dramatically differs from the plateau of around 20% of the substrate observed in the wild-type enzyme (Figure 3-21). I attribute these results to the substitution of Ser127 by alanine having a modest effect on carboxylation activity while significantly decreasing the rate of phosphoanhydride hydrolysis, probably due to altered orientation of the pyrophosphate group caused by loss of the Ser127-phosphate interaction.

Discussion

My structural and biochemical studies of LarB have provided critical insights into the enzyme architecture and mechanism. Although zinc is coordinated to the protein as isolated, this metal ion is not required for activity and inhibits the carboxylation of NaAD. Indeed, it is inconceivable that NaAD could bind to the active site and be activated for carboxylation if Cys221 participated in binding zinc. Using the LarB-NAD^+ structure, we have modeled the binding of the authentic substrate into the protein (Figure 3-22). We suggest the pyridinium ring of NaAD is flipped by 180 ° compared to NAD^+, with the C3 carboxyl group coordinating the Mg^{2+} (along with bidentate Asp151 and oxygen atoms from both phosphates) while being

stabilized by the essential residue Arg159. The interaction between the substrate carboxyl group and the metal ion confers specificity on the enzyme so as to allow for C5 carboxylation, whereas NAD^+ is not carboxylated. The model includes a Mg^{2+}-bound water molecule that could be further activated for hydrolysis by coordination of the NaAD carboxylate to the metal. CO_2 is proposed to bind at the position of the amide group in the NAD^+-bound structure, with similar H-bonding interactions used to bind this triatomic substrate. In this model, Cys221 is proximate to C4 of the pyridinium ring and ready to facilitate catalysis.

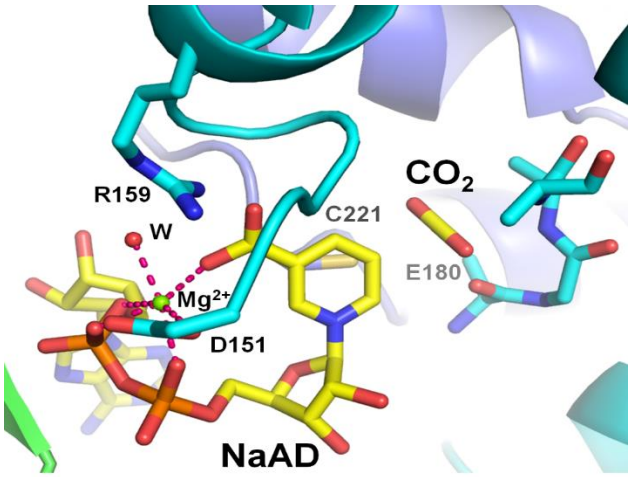


Figure 3-22. Structural model of the LarB active site in the NaAD- and CO_2-bound state. The model was built using the NAD^+-bound structure with NAD^+ replaced by NaAD in the same conformation, except that the pyridinium ring was flipped by 180 ° to allow for potential interactions of the carboxyl group Arg159 and Mg^{2+}. The position of the latter was shifted towards NaAD. A Mg^{2+}-bound water molecule was modeled and is presumably involved in hydrolysis of the phosphoanhydride bond. Coordination of Mg^{2+} by its ligands is indicated by the dashed lines. CO_2 occupies the position where the amide group of NAD^+ resides, and the residues forming H-bonds with the amide may be similarly used for CO_2 binding. The Glu180 side chain is truncated in the model as is the case for the NAD^+-bound structure.

The carboxylation reaction of LarB utilizes CO_2, not bicarbonate, analogous to what is found in the related enzyme, class II PurE.[8] Whereas PurE likely accomplishes the carboxylation reaction by direct nucleophilic attack of the aminoimidazole carbon atom on CO_2,[8, 10] the pyridinium C5 atom of NaAD is not nucleophilic; thus, LarB must use a distinct mechanism. Our structural results provide evidence that Cys221 attacks the pyridinium ring at C4 yielding a dihydropyridine, which is expected to enhance the nucleophilicity of the C5 carbon for attack on CO_2 (Figure 3-23A). A general base, likely the carboxylate of Glu180 – the only general base in the immediate vicinity according to the structure – would then complete the carboxylation

reaction by abstracting the proton on C5, resulting in expulsion of the thiolate and aromatization of the pyridine ring. Nucleophilic addition to CO_2 is a commonly observed mechanism of carboxylation for carboxylases that do not utilize the biotin cofactor,[30] but in this case the reaction must first be primed by forming a cysteinyl-pyridine adduct. The lack of carboxylation for NaMN provides compelling evidence that phosphoanhydride hydrolysis follows the carboxylation reaction with NaAD.

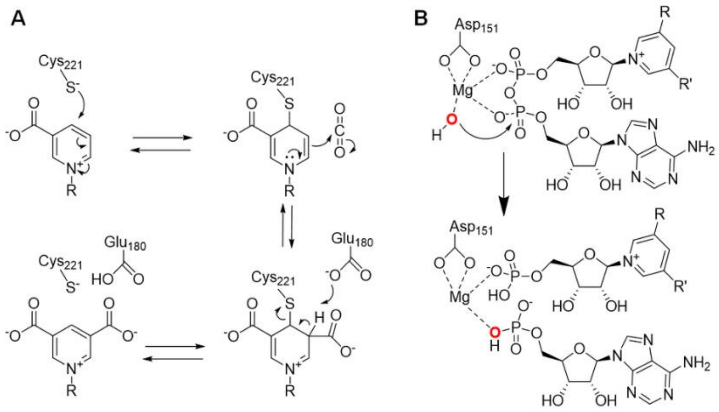


Figure 3-23. Proposed mechanisms for the LarB-dependent carboxylation and hydrolysis reactions. (A) Carboxylation. Cys221 adds to C4 of the NaAD pyridinium ring, thus enhancing the nucleophilicity of C5, which attacks carbon dioxide. Glu180 functions as a general base to abstract the C5 proton leading to rearomaticization and elimination of Cys221. (B) Hydrolysis. Magnesium activates a coordinated water that attacks the phosphate in the AMP portion of the substrate. In the desired reaction, both R and R′ = COO^-, but hydrolysis also occurs with the substrate NaAD for which R = COO^- and R′ = H. The hydrolysis reaction may facilitate product release.

As shown in the model of LarB with bound NaAD (Figure 3-22), the hydrolytic reaction is likely to involve a Mg-bound solvent molecule that attacks the phosphate distal to the pyridinium ring, releasing P2CMN (Figure 3-23B). The water is expected to be activated by the substrate carboxylate and by Mg-binding residue Asp151 that is nearly essential for enzymatic activity. Ser127, though not required for substrate carboxylation, appears to support the hydrolytic reaction on the basis of the increased total level of radioisotope incorporated into substrate when using the S127A variant. The serine residue may function to position for hydrolysis the phosphoanhydride of NaAD and its dicarboxylated analog. I speculate that phosphoanhydride hydrolysis facilitates product dissociation from the enzyme. My evidence does not provide support for an earlier proposal that energy associated with hydrolysis of the

phosphoanhydride is directly coupled to carboxylation;[2] rather, I propose that the two reactions are independent. Nevertheless, the low affinity of LarB for its product, P2CMN, prevents the enzyme from rapidly catalyzing the reverse reaction. In this way, phosphoanhydride hydrolysis indirectly drives the carboxylation reaction.

Data Availability

The atomic coordinates have been deposited in the PDB with accession codes7MJ2, 7MJ1, and 7MJ0, for the as-isolated (Zn-bound) LarB, LarB-NAD^+ complex, and LarB-AMP complex, respectively.

Acknowledgments

I thank Matthias Fellner for assistance during early structural efforts and the staff of the Mass Spectrometry and Metabolomics Core at MSU for their help with mass spectrometry studies. These studies were supported by grants to J.H. and R.P.H. from the National Science Foundation (CHE-1807073) and the National Institutes of Health (GM128959) and from grants to B.D. from the Fonds de la Recherche Scientifique - FNRS. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

APPENDIX

Table 3-A1. Strains, plasmids and primers used in this study.

Strain, plasmid, or primer	Characteristics or sequences	Source or methods
Strains		
E. coli		
DH10B	F- *endA*1 *recA*1 *galE*15 *galK*16 *nupG rpsL* Δ*lacX*74 Φ80*lacZ*ΔM15 *araD*139 Δ(*ara*,*leu*)7697 *mcrA* Δ(*mrr-hsdRMS-mcrBC*) λ-	Invitrogen
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] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	NEB
Plasmids		
pBAD:LarB	Ap^r; pBADHisA with with a 0.77-kb insert containing *larB* translationally fused to DNA encoding the StrepII-tag.	This study
pET:LarB	*Km^r, pET28a with a 0.77-kb insert containing larB translationally fused to DNA encoding the StrepII-tag*	
Primers		
LarB D39L-F	*GGTTTTGCCAATGTCCTGTTAGATCGTCAACGCCGCA*	Gap Repair
LarB D39L-R	*GACATTGGCAAAACCTAATGCCG*	Gap Repair
LarB R42A-F	*GTTTTGCCAATGTCGCTTTAGATCGTCAACG*	Two-stage PCR
LarB R42A-R	*CGTTGACGATCTAAAGCGACATTGGCAAAAC*	Two-stage PCR
LarB R42A-F	*CAATGTCGATTTAGATGCTCAACGCCGCAATGGC*	Two-stage PCR
LarB R42A-R	*GCCATTGCGGCGTTGAGCATCTAAATCGACATTG*	Two-stage PCR
LarB R45A-F	*GATTTAGATCGTCAACGCGCCAATGGCTTTCCGGAGG*	Two-stage PCR
LarB R45A-R	*CCTCCGGAAAGCCATTGGCGCGTTGACGATCTAAATC*	Two-stage PCR
LarB E50A-F	*GCAATGGCTTTCCGGCGGTCATCTACGGTGCTGGTAAGAC*	Gap Repair
LarB E50A-R	*CGGAAAGCCATTGCGGCG*	Gap Repair
LarB K57A-F	*ATCTACGGTGCTGGTGCGACGGCAACCCAAATTGTGGG*	Gap Repair
LarB K57M-F	*ATCTACGGTGCTGGTATGACGGCAACCCAAATTGTGGG*	Gap Repair
LarB K57X-R	*ACCAGCACCGTAGATGACCTC*	Gap Repair
LarB Y53F-F	*CGGAGGTCATCTTCGGTGCTGGTAAG*	Two-stage PCR
LarB Y53F-R	*CTTACCAGCACCGAAGATGACCTCCG*	Two-stage PCR
LarB S127A-F	*GGCGGGGACCGCAGATCAACCGG*	Q5[a]
LarB S127A-R	*GTCACCACAGCAATGTACCCTG*	Q5[a]
LarB D128A-F	*CGGGGACCTCCGCACAACCGGTTGCT*	Two-stage PCR
LarB D128A-R	*AGCAACCGGTTGTGCGGAGGTCCCCG*	Two-stage PCR
LarB D151N-F	*GTTGAACGAGTCTATAACGTGGGTGTTGCG*	Two-stage PCR
LarB D151N-R	*CGCAACACCCACGTTATAGACTCGTTCAAC*	Two-stage PCR
LarB D151L-F	*TGTTGAACGAGTCTATCTCGTGGGTGTTGCGGGAATC*	Gap Repair
LarB D151L-R	*ATAGACTCGTTCAACACGATTGCC*	Gap Repair
LarB D151A-F	*GAACGAGTCTATGCCGTGGGTGTTGC*	Two-stage PCR
LarB D151A-R	*GCAACACCCACGGCATAGACTCGTTC*	Two-stage PCR
LarB H158A-F	*GTGTTGCGGGAATCGCCCGACTGTTTGCCAAGTTGGATG*	Gap Repair
LarB R159K-F	*GTGTTGCGGGAATCCACAAACTGTTTGCCAAGTTGGATGTGATTC*	Gap Repair
LarB R159A-F	*GTGTTGCGGGAATCCACGCACTGTTTGCCAAGTTGGATGTGATTC*	Gap Repair
LarB H158/R159-R	*GATTCCCGCAACACCCAC*	Gap Repair
LarB E180D-F	*GATTGCGGGCATGGATGGTGCGCTGGC*	Two-stage PCR
LarB E180D-R	*GCCAGCGCACCATCCATGCCCGCAATC*	Two-stage PCR

Table 3-A1, (cont'd)

LarB E180Q-F	*GATTGCGGGCATGCAAGGTGCGCTGG*	Two-stage PCR
LarB E180Q-R	*CCAGCGCACCTTGCATGCCCGCAATC*	Two-stage PCR
LarB E180A-F	*ATTGCGGGCATGGCAGGTGCGCTGGC*	Two-stage PCR
LarB E180A-R	*GCCAGCGCACCTGCCATGCCCGCAAT*	Two-stage PCR
LarB Y204F-F	*CGACCAGCGTTGGTTTTGGTACTAGTTTTC*	Two-stage PCR
LarB Y204F-R	*GAAAACTAGTACCAAAACCAACGCTGGTCG*	Two-stage PCR
LarB Y204A-F	*CCGACCAGCGTTGGTGCTGGTACTAGTTTTC*	Two-stage PCR
LarB Y204A-R	*GAAAACTAGTACCAGCACCAACGCTGGTCGG*	Two-stage PCR
LarB C221S-F	*CCATGCTCAATAGCAGCGCGTCGGGGATTACC*	Two-stage PCR
LarB C221S-R	*GGTAATCCCCGACGCGCTGCTATTGAGCATGG*	Two-stage PCR
LarB C221A-F	*CCATGCTCAATAGCGCTGCGTCGGGGATTAC*	Two-stage PCR
LarB C221A-R	*GTAATCCCCGACGCAGCGCTATTGAGCATGG*	Two-stage PCR

[a]Q5, Site-directed mutagenesis via the NEB protocol.

Table 3-A2. Crystallization statistics for as-isolated LarB and the LarB-NAD^+ and LarB-AMP complexes

LarB crystals	As-purified	NAD^+ bound	AMP bound
Data collection			
Beamline	LS-CAT 21-ID-D	LS-CAT 21-ID-D	LS-CAT 21-ID-G
Wavelength (Å)	0.976	1.305	0.978
Space group	$P42_12$	$P42_12$	$P42_12$
Unit cell a, b, c (Å) α, β, γ (°)	120.3, 120.3, 212.9 90, 90, 90	120.4, 120.4, 213.5 90, 90, 90	121.3, 121.3, 213.4 90, 90, 90
[a]Resolution (Å)	34.03 - 2.80 (2.91 - 2.80)	39.94 – 3.40 (3.67 – 3.40)	39.79 - 3.01 (3.17 - 3.01)
Unique reflections	39138 (4327)	20627 (3997)	32347 (4619)
[a]Redundancy	22.4 (18.9)	4.5 (3.4)	9.7 (9.1)
[a]Completeness (%)	99.7 (99.8)	93.8 (90.3)	99.9 (99.9)
[a] I/σI	18.0 (3.6)	7.6 (2.1)	9.0 (2.2)
[a,b]R_{merge}	0.156 (0.964)	0.129 (0.612)	0.168 (1.168)
[a,c]R_{pim}	0.046 (0.308)	0.092 (0.484)	0.082 (0.590)
[d]$CC_{1/2}$	0.999 (0.891)	0.992 (0.674)	0.971 (0.699)
Refinement			
Protein atoms	8615	8257	8503
Ligand molecules	0	NAI, NAD	2 AMP
Metal	2 Zn, 3 Mg	4 Mg	2 Mg
Water	170	0	0
[e]R_{work}/R_{free}	0.2094/0.2426	0.2432/0.2689	0.2353/0.2746
B-factors (Å^2)	76.92	110.54	108.42
Protein	77.73	110.47	108.44
Ligand	-	116.70	71.34
Metal	43.18	86.86	24.13
H_2O	36.64	-	-
R.m.s. deviation in bond lengths (Å)	0.010	0.008	0.010
R.m.s. deviation in bond angles (°)	1.060	0.877	1.150
Ramachandran plot (%) favored	98.92	99.58	99.16
Ramachandran plot (%) outliers	1.08	0.42	0.84
Rotamer outliers (%)	7.10	9.00	9.04
PDB ID	7MJ2	7MJ1	7MJ0

[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.

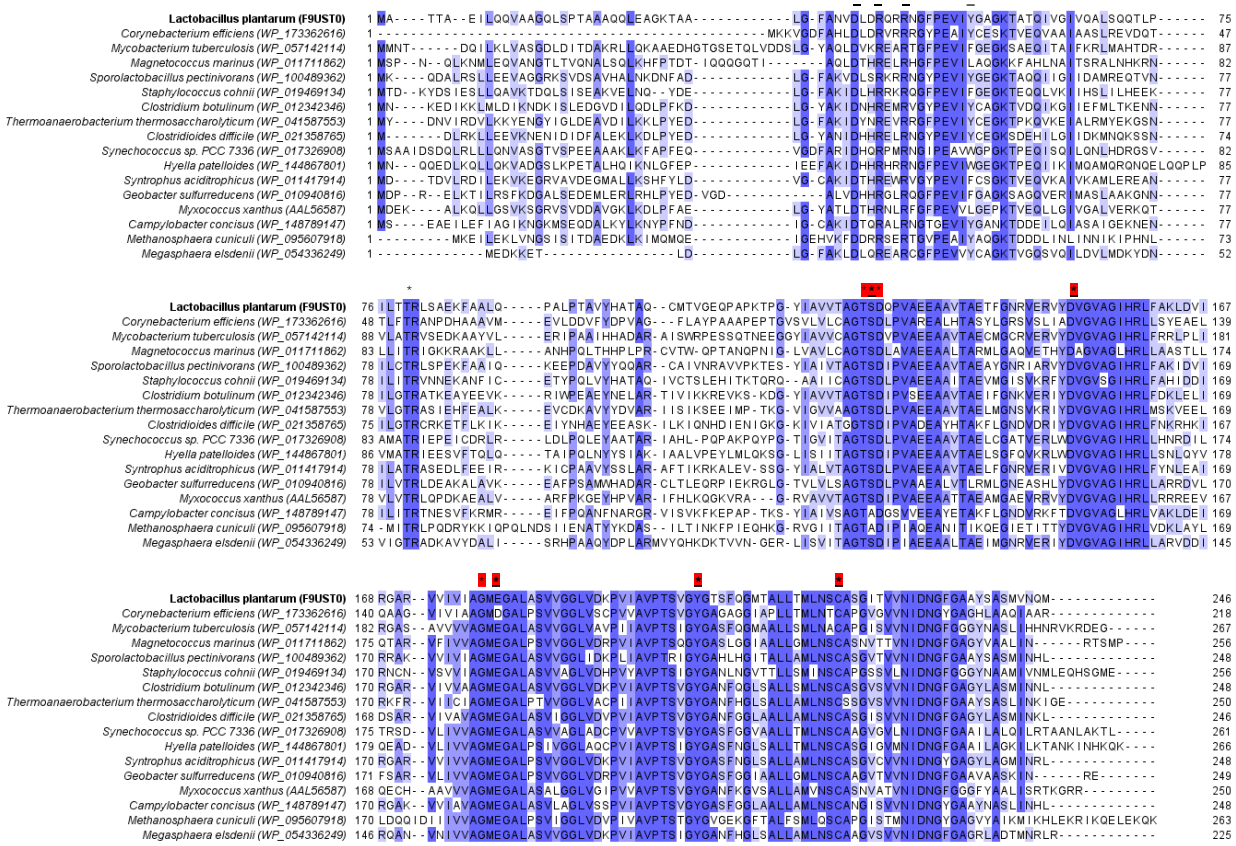


Figure 3-A1. Alignment of *L. plantarum* LarB to the corresponding protein of selected other bacteria (NCBI accession numbers are shown in parenthesis) identified using BLAST. The sequence alignment was produced using COBALT (NCBI). Residues are shaded in blue according to percent sequence identity. For both panels, the residues that directly interact with NAD+ are indicated in red (lettering or highlighted asterisks) and the positions of changes in variant proteins are underlined.

Table 3-A3. Analysis of chain interfaces in LarB.

Chain	Solvent accessible surface (Å^2)	Chain	Solvent accessible surface (Å^2)	Interface area (Å^2)	Δ^iG[a] (kcal/mol)	N_{HB}[b]	N_{SB}[c]	N_{DS}[d]
D	9611	C	9434	1243.4	-23.4	8	0	0
E	9273	F	9720	1160.3	-17.5	9	1	0
B	9309	A	9636	1094.2	-17.2	7	1	0

[a]Δ^iG, solvation free energy gain upon formation of the interface.

[b]N_{HB}, number of potential hydrogen bonds across the interface.

[c]N_{SB}, number of potential salt bridges across the interface.

[d]N_{DS}, number of potential disulfide bonds across the interface.

REFERENCES

REFERENCES

(1) 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.

(2) 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. U. S. A. 113*, 5598-5603.

(3) 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. U. S. A. 114*, 9074-9079.

(4) 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.

(5) 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.

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

(7) Zhang, Y., Morar, M., and Ealick, S. E. (2008) Structural biology of the purine biosynthetic pathway, *Cell. Mol. Life Sci. 65*, 3699-3724.

(8) Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, II, Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E., and Stubbe, J. (2007) N^5-CAIR mutase: role of a CO_2 binding site and substrate movement in catalysis, *Biochemistry 46*, 2842-2855.

(9) Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994) N^5-carboxyaminoimidazole ribonucleotide: evidence for a new intermediate and two new enzymatic activities in the de novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry 33*, 2269-2278.

(10) Mathews, I. I., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1999) Crystal structure of *Escherichia coli* PurE, an unusual mutase in the purine biosynthetic pathway, *Structure 7*, 1395-1406.

(11) Hanahan, D., Jessee, J., and Bloom, F. R. (1991) Plasmid transformation of *Escherichia coli* and other bacteria, *Methods Enzymol. 204*, 63-113.

(12) Wang, W., and Malcolm, B. A. (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis, *Biotechniques 26*, 680-682.

(13) Garcia-Nafria, J., Watson, J. F., and Greger, I. H. (2016) IVA cloning: A single-tube universal cloning system exploiting bacterial *In Vivo* Assembly, *Sci. Rep. 6*, 27459.

(14) Stoscheck, C. M. (1990) Quantitation of Protein, *Methods Enzymol. 182*, 50-68.

(15) Marty, M. T., Baldwin, A. J., Marklund, E. G., Hochberg, G. K., Benesch, J. L., and Robinson, C. V. (2015) Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles, *Anal. Chem. 87*, 4370-4376.

(16) Sudhani, H., Garca-Murria, M., Marn-Navarro, J., Garca-Ferris, C., Pearrubia, L., and Moreno, J. (2015) Assay of the Carboxylase Activity of Rubisco from *Chlamydomonas reinhardtii*, *Bio-Protocol 5*, e1672.

(17) Krezel, A., and Maret, W. (2016) The biological inorganic chemistry of zinc ions, *Arch. Biochem. Biophys. 611*, 3-19.

(18) Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I., and Stockinger, H. (2012) ExPASy: SIB bioinformatics resource portal, *Nucleic Acids Res. 40*, W597-603.

(19) Doublie, S. (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems, *Methods Mol. Biol. 363*, 91-108.

(20) Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J., and Adams, P. D. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, *Acta Crystallogr D Struct Biol 75*, 861-877.

(21) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, In *Methods Enzymol.*, pp 307-326, Academic Press.

(22) Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes, *Nucleic Acids Res. 46*, W296-W303.

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

(24) Tranchimand, S., Starks, C. M., Mathews, II, Hockings, S. C., and Kappock, T. J. (2011) *Treponema denticola* PurE Is a bacterial AIR carboxylase, *Biochemistry 50*, 4623-4637.

(25) Coombs, J., Maw, S. L., and Baldry, C. W. (1975) Metabolic regulation in C4 photosynthesis: The inorganic carbon substrate for PEP carboxylase, *Plant Sci. Lett. 4*, 97-102.

(26) Holm, L. (2020) DALI and the persistence of protein shape, *Protein Sci. 29*, 128-140.

(27) Fox, N. K., Brenner, S. E., and Chandonia, J. M. (2014) SCOPe: Structural Classification of Proteins--extended, integrating SCOP and ASTRAL data and classification of new structures, *Nucleic Acids Res. 42*, D304-309.

(28) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol. 372*, 774-797.

(29) Outten, C. E., and O'Halloran, T. V. (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis, *Science 292*, 2488-2492.

(30) Schada von Borzyskowski, L., Rosenthal, R. G., and Erb, T. J. (2013) Evolutionary history and biotechnological future of carboxylases, *J. Biotechnol. 168*, 243-251.

CHAPTER 4

Analysis of the active site cysteine residue of the sacrificial sulfur insertase LarE from *Lactobacillus plantarum*

Adapted from:

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.

In this chapter, I conducted the $LarE^{Dha}$ characterization and the LarE regeneration experiments. I also helped develop the LarE assay with Dr. Matthias Fellner. Dr. Fellner carried out the variant LarE analysis and all of the crystallography under the guidance of Dr. Jian Hu. Dr. Benoît Desguin created most of the strains and plasmids.

Abstract

LarE from *Lactobacillus plantarum* is an ATP-dependent sulfur transferase that sacrifices its Cys176 sulfur atom to form a dehydroalanine (Dha) side chain during biosynthesis of the covalently linked nickel-pincer nucleotide (NPN) cofactor (pyridinium 3-thioamide-5-thiocarboxylic acid mononucleotide) of lactate racemase. Coenzyme A (CoA) stabilizes LarE and forms a CoA-Cys176 mixed disulfide with the protein. This study presents the crystal structure of the LarE/CoA complex, revealing protein interactions with CoA that mimic those for binding ATP. CoA weakly inhibits LarE activity, and the persulfide of CoA is capable of partially regenerating functional LarE from the Dha176 form of the protein. The physiological relevance of this cycling reaction is unclear. A new form of LarE was discovered, an NPN-LarE covalent adduct, explaining prior results in which activation of the lactate racemase apoprotein required only the isolated LarE. The crystal structure of the inactive C176A variant revealed a fold essentially identical to that of wild-type LarE. Additional active site variants of LarE were created and their activities characterized, with all LarE variants analyzed in terms of the structure. Finally, the *L. plantarum* LarE structure was compared to a homology model of *Thermoanaerobacterium thermosaccharolyticum* LarE, predicted to contain three cysteine residues at the active site, and to other proteins with a similar fold and multiple active site cysteine residues. These findings suggest that some LarE orthologs may not be sacrificial but instead may catalyze sulfur transfer by using a persulfide mechanism or from a labile site on a [4Fe-4S] cluster at this position.

Introduction

Lactate racemase (Lar) interconverts the D and L isomers of lactic acid, a central metabolite of many microorganisms.[1] The enzyme allows for the production of D-lactate for cell wall biosynthesis in bacteria that possess only L-lactate dehydrogenase and permits the metabolism of both isomers from racemic mixtures of lactate in microorganisms containing a single form of lactate dehydrogenase.[2-4] The Lar enzyme from *Lactobacillus plantarum* includes the protein LarA (UnitProtKB entry F9USS9) and a covalently tethered (via Lys184) cofactor, pyridinium 3-thioamide-5-thiocarboxylic acid mononucleotide with nickel bound to C4 of the pyridinium ring, the two sulfur atoms of the pincer complex, and the His200 side chain (Figure 4-1).[5, 6] This organometallic complex will subsequently be termed the nickel-pincer nucleotide (NPN) cofactor.[7] The NPN cofactor is synthesized from nicotinic acid adenine dinucleotide (NaAD) by the sequential actions of LarB (UniProtKB entry F9UST0),[8] which carboxylates C5 of the pyridinium ring and hydrolyzes the phosphoanhydride bond, LarE (UniProtKB entry F9UST4), which converts the two carboxylates to thiocarboxylates,[9] and LarC (UniProtKB entry F9UST1), which installs the nickel.[8, 10] The process by which the NPN cofactor becomes covalently attached to LarA from *L. plantarum* has not been described, and activation of LarA

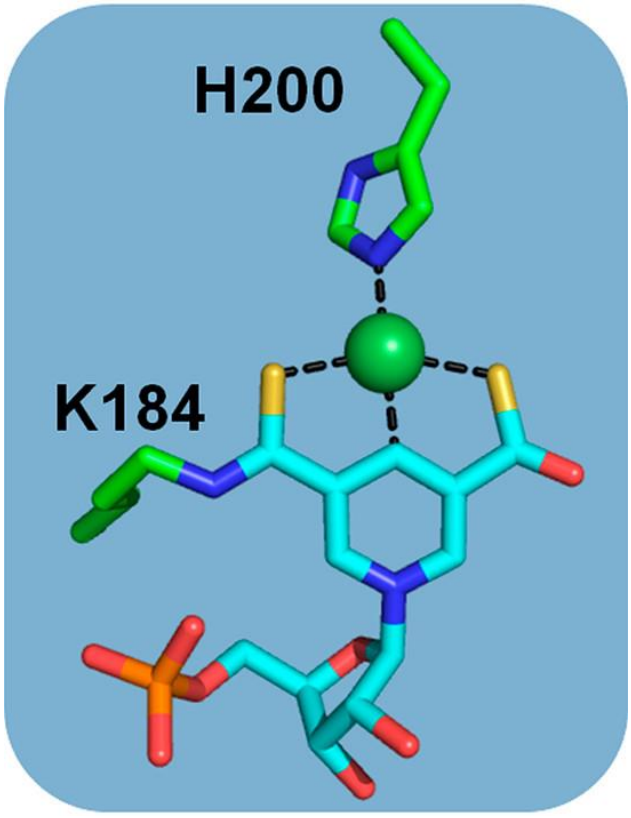


Figure 4-1. NPN cofactor in lactate racemase. Lys184 of LarA binds pyridinium 3-thioamide-5-thiocarboxylic acid mononucleotide [shown as sticks with cyan carbon atoms; Protein Data Bank (PDB) entry 5HUQ]. Two sulfur atoms and one carbon atom of the cofactor along with His200 coordinate nickel (green sphere).

from *Thermoanaerobacterium thermosaccharolyticum* (UniProtKB entry D9TQ02) does not require adduct formation. [8]

This study focuses on LarE, a member of the PP-loop pyrophosphatase family containing a PP-loop SGGxDS motif in its N-terminal region. Transformation of pyridinium 3,5-biscarboxylic acid mononucleotide (P2CMN) to pyridinium 3,5-bisthiocarboxylic acid mononucleotide (P2TMN) by LarE requires two cycles of carboxylate activation (involving ATP-dependent adenylylation) and sulfur insertion, with the sulfur originating from a cysteine residue (Cys176) of the protein, thus making LarE a sacrificial sulfur transferase that acquires a dehydroalanine (Dha) residue.[8, 9]

Here, we examine several unusual and interesting features of LarE with focus on its Cys176 residue. First, we investigate the unresolved nonsubstrate interaction of LarE with coenzyme A (CoA), a cellular component that stabilizes LarE and forms a disulfide with Cys176 but is not believed to be required for LarE activity,[8, 9] by presenting the crystal structure of LarE with bound CoA. Second, we provide evidence that CoA weakly inhibits LarE activity. Third, we demonstrate the ability of CoA-persulfide to regenerate Cys-containing LarE from the Dha176-containing protein ($LarE^{Dha}$). Fourth, we provide evidence of the presence of a novel NPN adduct of LarE, providing an explanation for earlier confounding observations related to the ability of purified LarE to activate the LarA apoprotein. Fifth, we investigate whether the loss of the Cys176 side chain significantly impacts the protein fold by characterizing the structure of the C176A variant. Sixth, we characterize the functions of several other key residues of LarE and summarize all LarE mutagenesis studies to better define the sacrificial sulfur insertion reaction.[9] Lastly, we speculate about whether LarE homologues containing additional active site cysteine residues may operate by a catalytic sulfur transfer process.

Materials and Methods

Genes, Plasmids, and Cloning

Site-directed mutagenesis of the gene encoding Strep II-tagged wild-type (WT) LarE from *L. plantarum* was performed using a QuikChange mutagenesis kit (Agilent) for creating the K101A, E128A, and E223A variants. All of the constructs were verified by DNA sequencing. Details, including instructions for creating the C176A variant, are summarized in our previous work.[9]

LarE Overexpression, Purification, and Characterization

WT, C176A, and the new variant forms of LarE were overexpressed and purified using our previously described pGIR076 *Escherichia coli* ArcticExpress-Strep-tactin (IBA) system.[9] The final buffer for all samples was 100 mM Tris-HCl (pH 7.5) containing 300 mM NaCl. The activities of WT and variant forms of LarE were assessed by monitoring samples for their abilities to transform P2CMN to P2TMN that was further metabolized by LarC and incorporated into LarA apoprotein from *T. thermosaccharolyticum*; this activated LarA enzyme was then assayed for lactate racemization as described elsewhere in detail.[9]

Crystallization

For both the CoA-bound WT protein and the C176A LarE variant, crystals were obtained after mixing 5 μL of protein samples with 5 μL of reservoir solutions. The hanging drop reservoir contained 100 μL of 30.0% (v/v) pentaerythritol ethoxylate (15/4 EO/OH), 50 mM Bis-Tris (pH 6.5), and 100 mM ammonium sulfate.

The CoA-bound sample contained 28 mg/mL WT LarE mixed with 0.9 mM CoA and was incubated for ~10 min on ice before setting up the drop. The C176A LarE sample contained 8 mg/mL variant protein. Using the same setup that was used for C176A, we also set up unsuccessful crystal drops for W97A (~1 mg/mL), S180A (~10, 12, and 15 mg/mL), R212A (~6 mg/mL), and D231R (~3, 11, and 24 mg/mL) variants of LarE using the same procedures.[9]

Diffraction Data Collection, Structure Determination, and Analysis

Data sets were collected at the Advanced Photon Source LS-CAT beamlines (21-ID-F and 21-ID-G). Data sets were processed with xdsapp[11] and iMos,[12] with merging and scaling done using aimless.[13] Phaser molecular replacement[14] was utilized using the WT apoprotein model 5UDQ. Model building and refinement were conducted in Coot[15] and Phenix.[14] Data set statistics are listed in Table 4-1. UCSF Chimera[16] was used to create structure figures.

Table 4-1. Crystallography statistics for the CoA-Bound and C176A Variant LarE Structures

	CoA-bound	C176A
Data Collection[a]		
beamline	LS-CAT 21-ID-F	LS-CAT 21-ID-G
wavelength (Å)	0.979	0.979
space group	$P4_122$	$P4_122$
unit cell dimensions		
a, b, c (Å)	107, 107, 320	109, 109, 329
α, β, γ (deg)	90, 90, 90	90, 90, 90
resolution (Å)	47.95−2.09 (2.12−2.09)	48.67−2.35 (2.39−2.35)
no. of unique reflections	111492 (5293)	82741 (4345)
redundancy	6.1 (6.0)	11.3 (11.2)
completeness (%)	99.7 (97.1)	99.5 (97.3)
$I/\sigma I$	14.7 (1.8)	18.9 (2.1)
R_{merge}[b]	0.084 (0.899)	0.102 (1.079)
R_{pim}[c]	0.055 (0.599)	0.046 (0.482)
$CC_{1/2}$[d]	0.998 (0.509)	0.999 (0.702)
Refinement		
no. of protein atoms	11949	11420
no. of CoA atoms	288	0
no. of phosphate atoms	30	30
no. of sulfate atoms	0	25
no. of waters	606	493
R_{work}/R_{free}[e]	0.195/0.234	0.187/0.244
B factor (Å^2)	43.8	56.3
protein	43.9	57.0
CoA	59.5	–
phosphate	38.5	47.7
sulfate	–	61.5
water	42.5	48.4
root-mean-square deviation in bond lengths (Å)	0.007	0.007
root-mean-square deviation in bond angles (deg)	0.975	0.922
Ramachandran plot, favored (%)	97.34	97.43
Ramachandran plot, outliers (%)	0	0
no. of rotamer outliers	0	0
PDB entry	6B2M	6B2O

[a]Data for the highest-resolution shell are shown in parentheses. [b]$R_{merge} = \sum_{hkl}\sum_j|I_j(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_j I_j(hkl)$, 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 data set. [d]$CC_{1/2}$ is the correlation coefficient of the half-data sets. [e]$R_{work} = \sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total) omitted during model refinement.

Purification of Dha-Containing LarE

LarE^{Dha} was purified from *Lactococcus lactis* NZ3900 containing pGIR172 that carries the *larA–larE* genes from *L. plantarum* with *larE* translationally fused to DNA encoding the Strep-tag II[5] as previously described.[8] The cells were grown overnight without being shaken in 15 mL of M17 medium (Oxoid or Difco) containing 0.5% (w/v) D-glucose and 10 mg/L chloramphenicol at 30 °C. This inoculum was added to 1.5 L of the same medium, but containing 5 mg/L chloramphenicol, and incubated at 28 °C while being shaken (40 rpm) until the OD_{600} reached ∼0.3. The culture was supplemented with 1 mM $NiCl_2$, induced with 10 μg/L nisin A (Sigma), and grown for 4 h before the cells were harvested by centrifugation at 5000*g* for 10 min. The cell pellet was washed using 100 mM Tris-HCl buffer (pH 7.5) containing 300 mM NaCl and stored until it was needed at −80 °C. The thawed cells were suspended in 40 mL of 100 mM Tris-HCl buffer (pH 7.5) containing 300 mM NaCl, 1 mM phenylmethanesulfonyl fluoride (PMSF, added as a 100 mM stock in ethanol), and 1 unit/mL Benzonase (Millipore) on ice and then lysed by two passes through a chilled French pressure cell at 16000 psi. The debris and intact cells were removed by centrifugation at 27000*g* and 4 °C for 40 min. Purification of LarE^{Dha} was performed by using a 1 mL bed volume of Strep-tactin-XT (IBA) resin equilibrated in 20 mM Tris-HCl buffer (pH 7.5) containing 300 mM NaCl at 4 °C. The lysates were loaded at a rate of 1 mL/min; the resin was washed by gravity flow with five 1 mL additions of 20 mM Tris-HCl buffer (pH 7.5) containing 300 mM NaCl, and the protein was eluted with six 0.5 mL additions of 20 mM Tris-HCl buffer (pH 7.5) containing 300 mM NaCl and 50 mM biotin (Santa Cruz Biotechnology) that had been neutralized with 50 mM NaOH.

Additional characterization of LarE^{Dha} included analysis of its reactivity with 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) and tris(2-carboxyethyl)phosphine (TCEP) in comparison

to that of LarE. To quantify accessible thiol groups, the protein samples were treated with 180 μM DTNB in 100 mM Tris-HCl (pH 7.5), allowed to react for 30 min, and monitored at 412 nm.[17] For testing the reactivity of LarE-associated Dha with TCEP,[18] the samples were treated with 5 mM TCEP in buffer and analyzed by mass spectrometry (see below).

LarE Regeneration Reactions

Lysates for use in LarE regeneration experiments were derived from cultures of *L. plantarum* NCIMB8826. Using a 2% inoculum, the cells were grown without being shaken at 37 °C in MRS medium (Sigma) supplemented with 0.1% (v/v) Tween 20. At an OD_{600} of ~1, the cultures were induced by adding 200 mM sodium L-lactate as a dry powder or left uninduced and incubated for a further 2 h. The cells were harvested by centrifugation at 5000*g*, and the pellets were washed in 100 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl and stored at −80 °C. The cells (~500 μL of cell paste) were mixed with glass beads (~500 μL, mean diameter of 100 μm) in 20 mM Tris-HCl (pH 7.5) containing 150 mM NaCl and, for maintaining anaerobic conditions,[19] 20 μg/mL glucose oxidase, 2 μg/mL catalase, and 0.3% (w/v) D-glucose in 1.5 mL screw cap tubes (deaerated by blowing argon over the surface) and lysed by using a Mini-Bead Beater 16 (BioSpec) with two 60 s pulses separated by 5 min on ice. The debris was removed by centrifugation at 20000*g* and 4 °C for 10 min.

Cysteine desulfurase (IscS) was purified from *E. coli* BL21 (DE3) containing pBH402, a plasmid with *E. coli iscS* fused to a sequence encoding a His_6 tag and under the control of the T7 promoter.[20] Overnight cultures of these cells were grown while being shaken (300 rpm) at 37 °C in LB containing 50 μg/mL kanamycin (Gold Biotechnology), then diluted 200-fold into 0.5 L of Lennox LB supplemented with 1 mM pyridoxine (Sigma) and 50 μg/mL kanamycin, and further grown while being shaken at the same temperature. When an OD_{600} of ~0.5 was reached, the

cultures were amended with 0.4 mM isopropyl β-D-1-thiogalactopyranoside and grown for an additional 3 h. The cells were collected by centrifugation and resuspended in 40 mL of buffer A [50 mM Na_2HPO_4 (pH 8.0) containing 300 mM NaCl and 20 mM imidazole] supplemented with 2.5 units/mL Benzonase and 1 mM PMSF. The suspension was chilled on ice and lysed by two passes through a chilled French press cell at 16000 psi. The cell debris was pelleted by centrifugation at 27000*g* and 4 °C for 40 min. The supernatant was applied to a column (10 mL) of Ni-nitrilotriacetic acid-Sepharose that had been equilibrated in buffer A. After the column had been washed twice with 50 mL of buffer A, IscS was eluted at a rate of 2 mL/min with a 100 mL linear gradient from buffer A to buffer A containing 500 mM imidazole. The 2.5 mL fractions were collected into tubes containing 2.5 μL of 10 mM EDTA (pH 8.0) to prevent precipitation. The IscS-containing fractions were combined and dialyzed overnight at 4 °C against 2 L of 50 mM potassium phosphate buffer (pH 7.5) containing 5 mM $MgCl_2$, 100 mM KCl, and 0.1 mM EDTA. The protein concentration was quantified using the published ε_{280} of 25400 M^{-1} cm^{-1},[20] and purified IscS was stored at 4 °C.

Organic persulfides were synthesized by incubating 5–20 mM disulfides with a 5-fold molar excess of NaHS in 300 mM Tris-HCl (pH 7.5, for making CoA and glutathione persulfides) or 100 mM NaOH (for the Cys persulfide) in an argon atmosphere.[21] The reaction mixtures were incubated at 30 °C for 30 min. Persulfide yields were quantified by using a cold cyanolysis procedure,[22] with yields typically of ~30% relative to the starting disulfide concentrations. Persulfide preparations were used immediately without further purification or frozen at −20 °C under argon for later use.

$LarE^{Dha}$ (2.8 μM, 0.087 mg/mL) was incubated in 20 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl, 20 μg/mL glucose oxidase, 2 μg/mL catalase, and 0.3% (w/v) D-

glucose (to ensure anoxic conditions), with various additives as specified. Selected mixtures included 10% cell-free lysates from *L. plantarum* cultures that were induced with lactic acid or non-induced. Some mixtures included 2 mM $MgCl_2$ and 1 mM ATP. Individual sulfur sources and a combined mixture of sulfur compounds (1 mM cysteine, 1 mM glutathione, 1 mM 3-mercaptopyruvate, 1 mM sodium thiosulfate, 1 mM sodium sulfide, and 200 μM CoA) were tested. In addition, IscS (4.6 μM, 0.21 mg/mL) with 1 mM cysteine was examined. For all samples, the reaction mixtures were layered with mineral oil and allowed to incubate at room temperature with 1 mL time points taken at 1 h or overnight. The samples lacking cell-free lysates were directly analyzed by mass spectrometry, but those amended with lysates were loaded by gravity flow at 4 °C onto columns containing 200 μL of Strep-tactin-XT resin; the columns were washed using three 1 mL additions of 20 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl, and the LarE protein was eluted using six 250 μL additions of the same buffer containing 50 mM biotin neutralized with 50 mM NaOH. These protein samples were concentrated to at least 1 μM LarE using 10 kDa cutoff Amicon centrifuge columns prior to further analysis.

Mass Spectrometry of LarE

The purified LarE protein samples (10 μL in buffer), including those from the LarE regeneration studies, were injected onto a cyano-chemistry high-performance liquid chromatography column that was equilibrated in 0.1% formic acid and eluted with an increasing gradient of acetonitrile. The fractions were analyzed by electrospray ionization mass spectrometry (ESI-MS) using a XEVO G2-XS instrument in positive ionization mode. The protein masses were derived from the MS data using MaxEnt (Waters Corp.). LarE samples were

also purified from *Lc. lactis* NZ3900 pGIR172 that had been induced for different lengths of time (1, 2, 3, and 4 h) and analyzed by the same ESI-MS approach.

LarE Enzyme Assay

The activity of LarE was assessed indirectly by a four-step process using purified components and assaying for Lar activity. The first step was the production of P2CMN by a 3 h incubation of LarB (3 μM) with 1 mM NaAD in 100 mM $NaHCO_3$ with quenching by heat (80 °C for 5 min). The second step involved LarE synthesis of P2TCM; the step 1 mixture (4% of the final volume) was incubated with 10 μM LarE, 2 mM reducing agent [Cys, dithiothreitol (DTT), CoA, β-mercaptoethanol, glutathione, or ascorbic acid], 2 mM ATP, and 20 mM $MgCl_2$ in 100 mM Tris-HCl (pH 7) for 5 min at room temperature, and then the reaction was stopped by heat. Step 3 involved synthesis of the NPN cofactor and generation of active LarA by reacting the step 2 mixture (10% of the final volume) with 1.5 μM LarC, 1.5 μM *T. thermosaccharolyticum* LarA apoprotein, 1 mM $MnCl_2$, 0.1 mM CTP, and 40 mM D-lactate in 100 mM Tris-HCl (pH 7) for 5 min at room temperature, with quenching by heat. The final step was to measure L-lactate with an enzymatic kit (Megazyme).

Results

Structural Investigation of Cys176/Coenzyme A Interaction

Prior studies had demonstrated that the addition of CoA stimulates the activation of LarA apoprotein in mixtures containing P2CMN (the product of LarB), Mg·ATP, LarE, LarC, and LarA apoprotein.[8] Unpublished observations showed a 765 Da adduct present in LarE purified from *Lc. lactis* cells, which we presumed to be CoA. When LarE (as the Strep II-tagged species lacking the N-terminal Met residue; 31550 Da) was purified in the presence of CoA, this adduct

(32316 Da) became the majority species and CoA was shown to form a disulfide with the single cysteine residue of LarE (Cys176). [8] Several thiol reductants (sulfide, Cys, DTT, and glutathione), including CoA, could reduce the disulfide and confer LarE activity in the Lar assay, but only CoA provided longer-term (20 h) stability to the protein (data not shown). An analogous CoA-protein disulfide (32274 Da) was formed with the D30A variant of LarE (31508 Da) but not with the C176A variant (31520 Da) (Figure 4-2). CoA is not required for *E. coli*-derived LarE to convert P2CMN to P2TMN.[9] The CoA-protein disulfide form of LarE converts to $LarE^{Dha}$ in the LarA apoprotein activation mixture. While it is clear that CoA interacts with LarE, further studies were required to establish the details of this interaction.

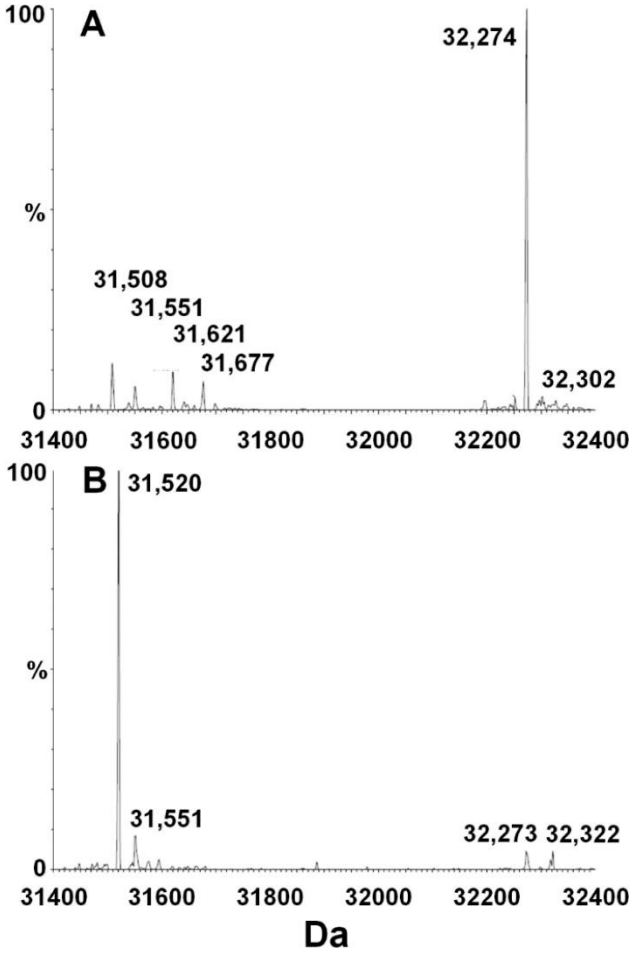


Figure 4-2. ESI-MS of (A) the CoA–S–S–Cys176 disulfide using D30A LarE and (B) the C176A variant of LarE, both purified in the presence of CoA.

LarE was isolated from both *Lc. lactis* and *E. coli* using purification buffers containing CoA, and each protein was crystallographically characterized. All LarE chains of the crystallographic hexamers had CoA bound. The phosphate group of the 3′-phosphoadenosine portion of the molecule was found at the PP loop (the ATP binding site) of LarE (Figure 4-3A). The adenine moiety forms two hydrogen bonds with the backbone atoms of Ala52 on strand β2.

The ribose is stabilized via interactions with the backbone atoms of Ala24 on β1 and Gly123 on β4. The diphosphate group of CoA was well-ordered in all structures but exhibited slightly different orientations between chain C and all other chains (Figure 4-3B). The differences were more notable for the remaining pantoic acid-alanine-cystamine portions of CoA; this part of the molecule is less ordered in all chains, does not show direct contacts with the protein, and shows slight variations in chains A, B, and D–F (F is shown as a representative example) compared to the very different orientation in chain C (Figure 4-3B). The distinct orientations of the coenzyme paralleled the differences in side chain orientations of the key residue, Trp97.[9] Although MS experiments had previously indicated a disulfide bond between CoA and Cys176 of LarE,[8] no disulfide bond was detected in the electron density maps. This result suggested that photoreduction of the disulfide may have occurred during data collection. Notably, the CoA sulfur atom was in the

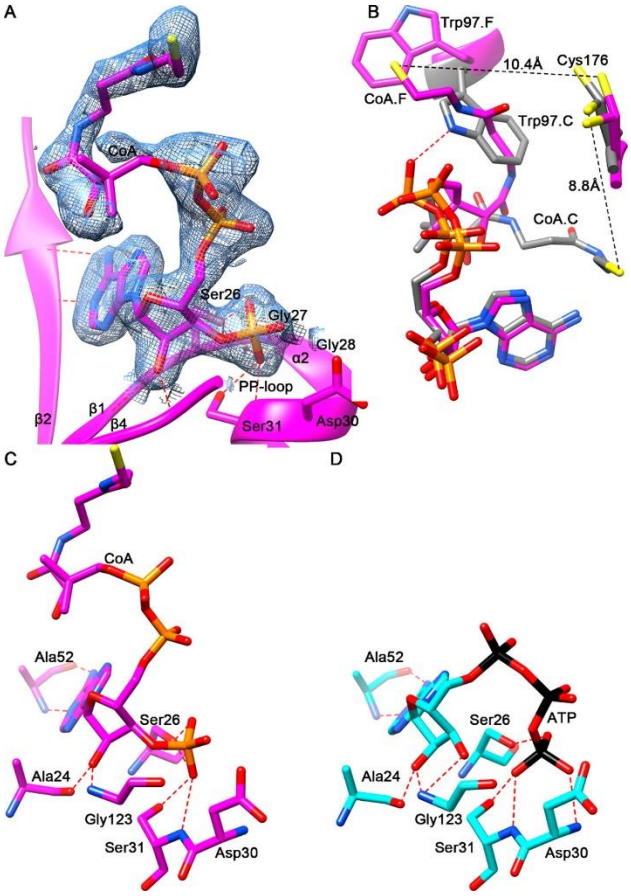


Figure 4-3. Binding of CoA to LarE (PDB entry 6B2M). All comparisons are based on a Cα alignment of the full length protein.[16] (A) The $2F_O - F_C$ map of CoA bound to chain F is shown as blue mesh at 1 σ. Interacting backbone residues are illustrated as ribbons with secondary structure features labeled, and the PP-loop residue side chains are shown. Hydrogen bonds are indicated as red dashes. (B) Comparison of CoA binding orientations in LarE. Chain F is colored magenta (representative for chains A, B, D, and E) and compared to chain C (gray). The corresponding side chains of Trp97 and Cys176 are illustrated in the same colors with distances between the CoA and Cys176 sulfur atoms shown as black dashes. (C and D) Comparison of CoA and ATP (PDB entry 5UDS)[9] binding to LarE. Interacting residues are shown in stick mode with hydrogen bonds indicated by red dashes. The carbon atoms of the CoA- and ATP-bound structures are colored magenta and cyan, respectively. The ligand phosphorus atoms of CoA are colored orange, and those of ATP are colored black.

vicinity of the Cys176 side chain, with the two different orientations having distances of 8.8 and 10.4 Å between the Cys176 and the CoA sulfur atoms. The transition between the two orientations of the least ordered portion of CoA would place the CoA sulfur close enough to the Cys176 side chain to form the disulfide.

The binding mode of the 3′-phosphoadenosine ribose portion of CoA is nearly identical to that of ATP, the difference being that the triphosphate of ATP is positioned to place the γ-phosphate into the PP loop, whereas it is the ribose phosphate group of CoA that is tightly associated with the PP-loop residues (Figure 4-3C,D).

CoA Weakly Inhibits LarE Activity

ATP likely has a binding affinity for LarE greater than that of CoA, compatible with its greater number of interactions involving the PP loop, thus overcoming any potent inhibitory effect of the coenzyme. Nevertheless, the apparent LarE activity was lower when using CoA compared to other reductants (Cys, DTT, β-mercaptoethanol, glutathione, or ascorbic acid) when the protein was subjected to short-term treatment with these reagents and analyzed using the indirect LarA apoprotein activation assay (Figure 4-4). This result suggests that CoA binding is partly inhibitory with respect to LarE activity.

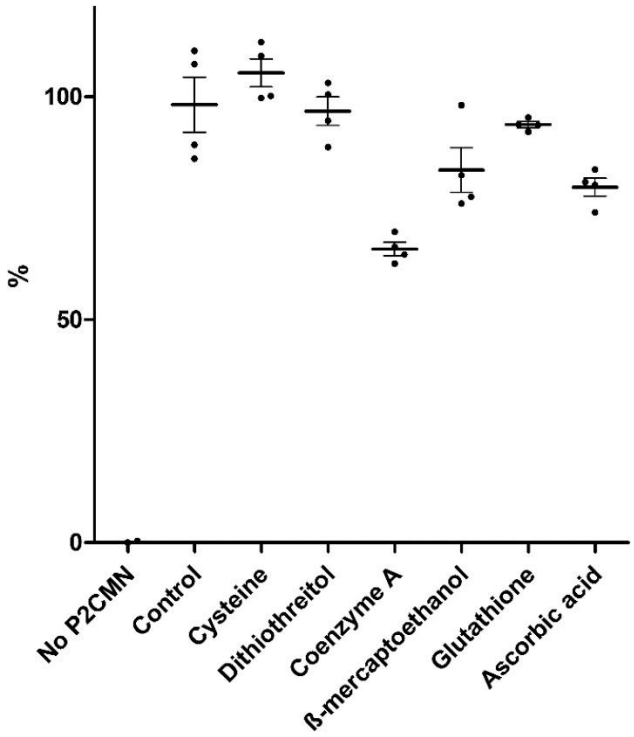


Figure 4-4. Inhibitory effect of CoA on the apparent LarE activity compared to other reductants. P2CMN was produced enzymatically from NaAD, converted to P2TMN by LarE in the presence of the indicated reductants (2 mM) along with 2 mM ATP and 20 mM $MgCl_2$ in 100 mM Tris-HCl buffer (pH 7), transformed into the NPN cofactor by LarC, incorporated into *T. thermosaccharolyticum* LarA apoprotein, and assayed for Lar activity, shown as the percent relative to the control without thiol.

Additional Characterization of LarE^{Dha}

Strep-tagged LarE was expressed in *Lc. lactis* in the context of the entire *lar* operon and purified. As previously described,[8] nearly all of the resulting LarE lacks its Cys176 sulfur atom according to ESI-MS (Figure 4-5A). This major form of the protein was observed for samples isolated from cultures that had been induced from 1 to 4 h (data not shown), so LarE^{Dha} is stable in the cells. Along with LarE^{Dha} (31517 Da) the spectrum reveals likely sodium complexes of the protein (≥31540 Da); the intensity of these forms was greatly reduced using an ammonium bicarbonate buffer. In contrast to the native LarE protein that exhibits DTNB reactivity accounting for 0.93 Cys thiol per protein, no 412 nm peak was detected using the LarE^{Dha} protein. As further support for the presence of Dha, the addition of TCEP to LarE^{Dha} led to partial modification to form a species at 31767 Da, or 250 Da larger than LarE^{Dha} (Figure 4-6), consistent with the

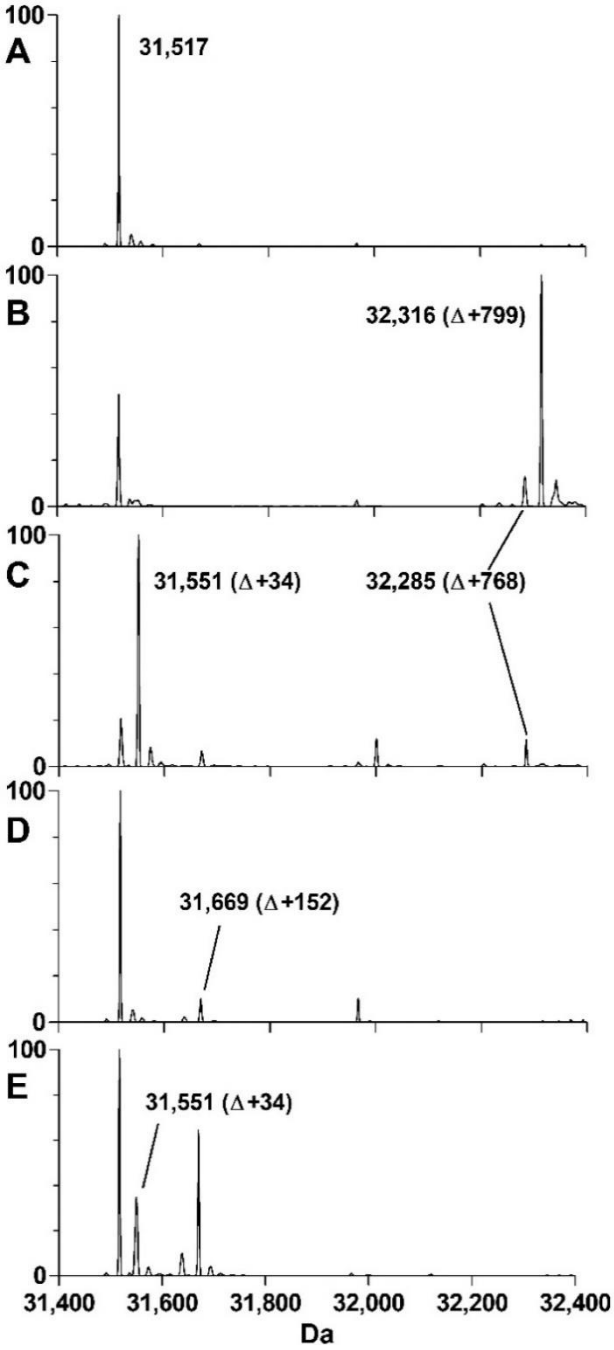


Figure 4-5. Regeneration of native LarE from LarE^{Dha}. (A) ESI-MS of LarE^{Dha}. (B) LarE^{Dha} treated with CoA persulfide (150 μM, 1 h at room temperature and then overnight at 4 °C), revealing conversion to the LarE–S–S–CoA disulfide. (C) Sample from panel B reduced with DTT (10 mM, 30 min at room temperature) to generate native LarE. (D) LarE^{Dha} treated with Cys persulfide. (E) Sample from panel D treated with DTT. The Δ values are relative to the 31517 Da value of LarE^{Dha}. The *y*-axes indicate the percent relative to the intensities of the maximum peaks.

known reactivity of TCEP toward dehydroalanine,[18] whereas this reagent had no effect on the mass spectrum of LarE.

Regeneration of Cys-Containing LarE from LarE^{Dha} by CoA Persulfide

Regeneration of native LarE from LarE^{Dha} was achieved by using the persulfide of CoA followed by reduction with DTT (Scheme 4-1). Incubation of LarE^{Dha} with CoA persulfide (Figure 4-5B) resulted in substantial conversion of the Dha-containing protein (31517 Da) to a species (32316 Da) consistent with the LarE–S–S–CoA disulfide.

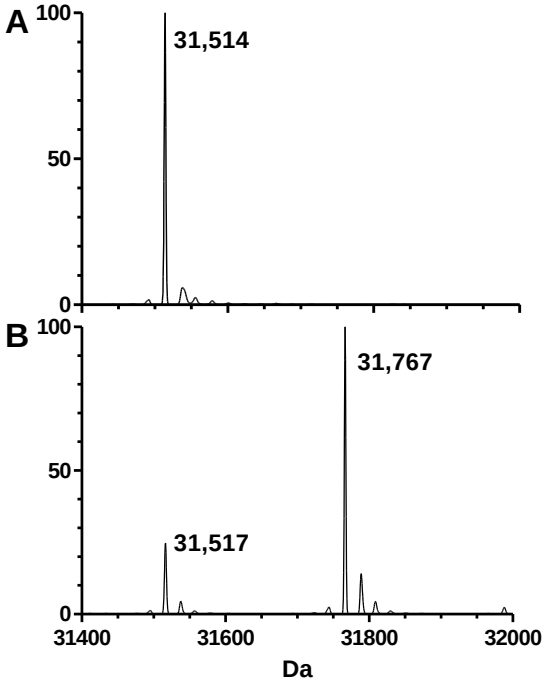
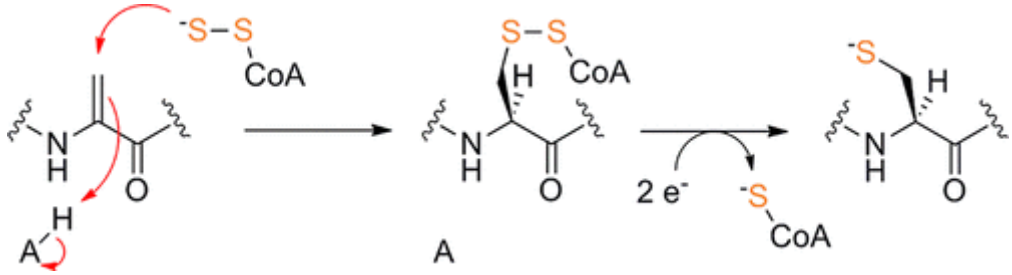


Figure 4-6. Interaction of TCEP with Dha-containing LarE. (A) LarE^{Dha} as purified. (B) LarE^{Dha} treated overnight with 5 mM TCEP (along with 1 mM Mg•ATP, and *L. plantarum* cell-free extracts).

In addition, we noted a minor species (32285 Da) suggesting the formation of the LarE–S–CoA thioether. Consistent with the thioether interpretation, the same species formed for LarE^{Dha} treated with CoA rather than the CoA persulfide (Figure 4-8). Subsequent reduction of the CoA persulfide mixture by DTT (Figure 4-5C) eliminated the putative disulfide species but not the proposed thioether and generated a peak (31551 Da) corresponding to native LarE. In contrast to the robust regeneration observed using the CoA persulfide, the Cys persulfide produced much

Scheme 4-1. Regeneration of LarE using the persulfide of CoA followed by reduction with DTT.



lower levels of the LarE–S–S–Cys disulfide [31669 Da (Figure 4-5D)] and corresponding smaller amounts of the native LarE after reduction (Figure 4-5E). DTT reduction of the Cys persulfide-treated sample led to an increase in the 31669 Da species that we attribute to formation of the thioether between DTT and LarE^{Dha}, which happens to possess the same mass as the LarE–S–S–Cys disulfide, as confirmed by a control reaction of DTT and the Dha-containing protein (Figure 4-7). No reactivity was detected between the glutathione persulfide and LarE^{Dha} (data not shown). These results suggest the CoA binding site of LarE is used to specifically bind the CoA persulfide during the sulfuration reaction. Although regeneration of native LarE from LarE^{Dha} was demonstrated by these *in vitro* experiments, it remains unclear whether a similar process occurs within *L. plantarum* cells.

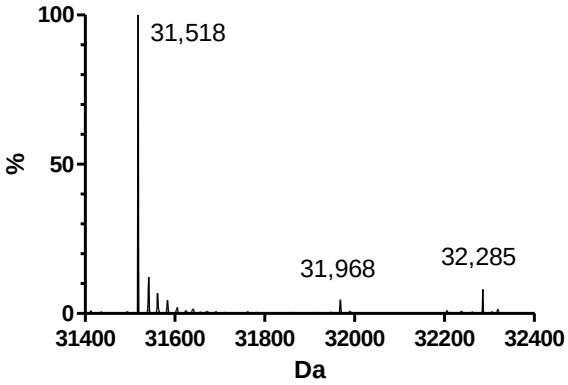


Figure 4-8. Formation of probable CoA-S-LarE thioether (32,285 Da) by incubation of LarE^{Dha} (31,518 Da) with 2 mM CoA (along with 1 mM Na_2S, and *L. plantarum* cell-free lysates) for 1 h at room temperature. Also shown is the presumed NPN adduct (31,968 Da)

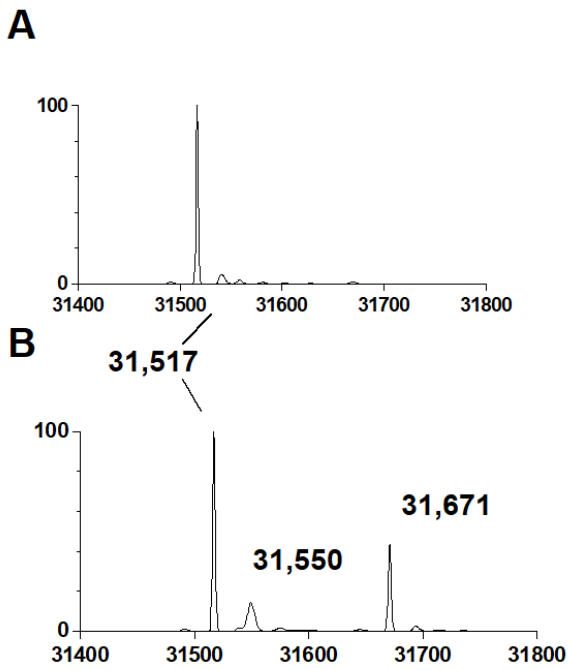


Figure 4-7. Effect of DTT on Dha-containing LarE. (A) LarE^{Dha} as purified. (B) LarE^{Dha} treated with 10 mM DTT for 30 min at room temperature.

Discovery of an NPN Adduct of LarE^{Dha}

Some preparations of LarE^{Dha} possessed a previously undescribed species (31968 Da) that was 450 Da larger than the Dha-containing protein (31518 Da) (Figure 4-9). Interestingly, this species is consistent with the formation of a covalent adduct between NPN and LarE. Such a species explains a previous confounding observation related to LarA activation. Namely, the incubation of purified LarA (1.4 pmol) with LarE (280 pmol, isolated from cells containing LarB and LarC) was shown to result in activation of ~30% of the lactate racemase apoprotein.[5] We interpret this prior result as arising from a secondary pathway for LarA activation that utilizes this new species with NPN covalently attached to LarE.

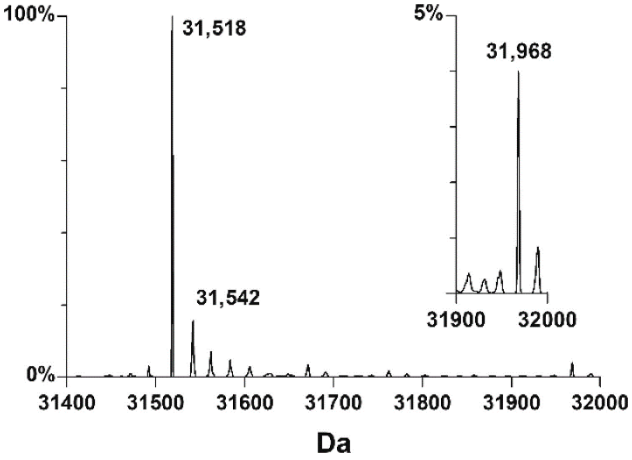


Figure 4-9. Example ESI-MS spectrum of LarE^{Dha} with an inset depicting an expanded view of the region from 31900 to 32000 Da.

Structural Characterization of C176A LarE and Functional Characterization of LarE Variants

We attempted to crystallize several LarE variants with largely reduced or abolished activity (W97A, C176A, S180A, and R212A) and the active D231R variant but had success only with the C176A LarE sample. The W97A variant, thought to be defective in its interaction with the P2CMN substrate,[9] was poorly produced, thus preventing crystallization. The S180A and R212A variants, considered to lack the P2CMN or inorganic phosphate binding site, were examined using many conditions, but the lack of successful crystallization suggests that a bound phosphate/sulfate molecule at this binding site is required for crystal formation. Although the D231R variant is as active as the WT enzyme, we observed no crystal formation, suggesting that the LarE trimer/hexamer interface where Asp231 is located may be disturbed, thus preventing

crystal growth. The successful case of the C176A variant protein yielded the same crystal morphology and size as the WT enzyme.

The overall fold of the C176A LarE structure (Figure 4-10A) was identical to that of the WT protein. Even the flexible linker between domains on which Cys176 is located matched in both structures. In addition, the bound phosphate previously noted in the WT enzyme[9] was also present in the C176A structure, indicating that hydrogen bonding with the sulfur atom (Figure 4-10B) is not crucial for binding the phosphate of P2CMN. The loss of activity of the C176A variant must be based solely on the functional role of the thiol group as a sulfur donor as the C176A structure shows no other structural difference from the WT enzyme.

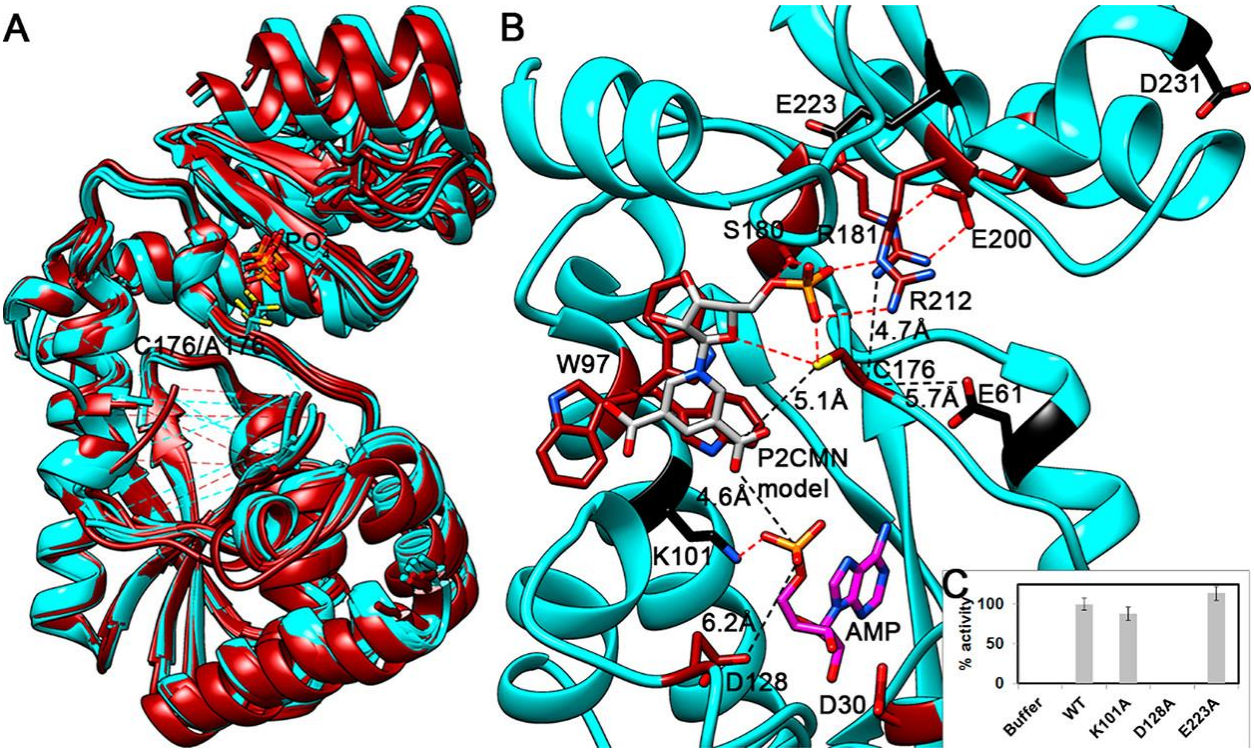


Figure 4-10. Analysis of LarE variants. (A) Comparison of the six chains of WT LarE colored cyan (PDB entry 5UDQ[9]) and those of the C176A LarE variant colored dark red (PDB entry 6B2O). The Cys176 side chain and nearby phosphate molecule are shown. (B) P2CMN/AMP-bound model of LarE based on chain A of NMN-bound LarE (PDB entry 5UDR) and chain C of AMP-bound LarE (PDB entry 5UDT).[9] Carbon atoms of the P2CMN model are colored gray, those of AMP magenta, and the side chains of residues that are altered in inactive variants dark red [D30A and C176A;[8] W97A, C176A, S180A, R181K, E200Q, and R212A;[9] and D128A (this publication)]. In addition, the side chains of residues that were substituted in active variants are colored black [E61Q and D231R[9] and K101A and E223A (this publication)]. Hydrogen bonds are shown as red dashes, and important interactions are illustrated as black dashes with the distances indicated. (C) Activity results of three variants characterized in this publication.

We previously characterized several LarE variants of highly conserved residues to show the importance of the active site cysteine (C176A), the PP loop (D30A), the P2CMN phosphate binding site (S180A and R212A), and several key residues in catalysis (W97A, R181K, and E200Q) while also ruling out the involvement of other residues (E61Q and D231R) (Figure 4-10B).[8, 9] Here, we characterized three additional variants to gain an even better understanding of LarE's function. Glu223 is buried in a manner similar to that of Glu200 in the C-terminal head domain of LarE, which forms trimer/hexamer interactions and binds the phosphate moiety of the substrate, P2CMN. Furthermore, Glu223 forms hydrogen bonds to the nearby Arg214 and Arg221 residues, with Arg214 being directly involved in phosphate binding. The E223A variant exhibited no significant change in activity compared to the WT enzyme. This finding shows that the interaction of Arg214 with phosphate is unlikely to require Glu223 so that phosphate binding requires only Ser180 and Arg212 (Figure 4-10B). The Arg214/Glu223 interaction is structurally similar to the nearby Arg181/Glu200 interaction, but in that case, the E200Q variant was inactive, highlighting the importance of the Glu200 side chain. Second, we examined Lys101 that is conserved in LarE sequences and forms a salt bridge with AMP (Figure 4-10B). Surprisingly, the K101A variant was as active as the WT enzyme, suggesting that the lysine is not critical for stabilizing AMP binding and that other residues on a nearby flexible loop might compensate for its absence through other AMP phosphate interaction residues.[9] Asp128 is another highly conserved residue on the flexible loop, but this residue does not directly interact with AMP in the crystal structure (Figure 4-10B). At a distance of 6.2 Å from the AMP phosphate, it is unclear what role this residue could play; nevertheless, it must be important as the D128A variant was inactive. This result demonstrates a key role for the flexible loop

containing Asp128, either in its disordered state found in most structures or as the helical fold in the AMP-bound structure.

Discussion

NPN Biosynthesis, LarE^{Dha} Regeneration, and a New Pathway for LarA Activation

We combine our new results with those from prior studies in a summary depiction of the LarA apoprotein activation process (Figure 4-11). In the primary pathway, LarB converts NaAD to P2CMN,[8] two molecules of LarE catalyze AMPylation of the pyridinium carboxyl groups followed by sacrificial sulfur transfers from Cys176 to form P2TMN (generating two copies of LarE^{Dha}),[8, 9] LarC inserts nickel by a CTP-dependent process,[8, 10] and the resulting free NPN activates LarA (Figure 4-11, blue arrow pathway).

We have now shown that LarE^{Dha} can be recycled to regenerate native LarE by *in vitro* incubation with CoA persulfide and reductant (Scheme 4-1 and Figure 4-11, orange arrows). The CoA persulfide is likely

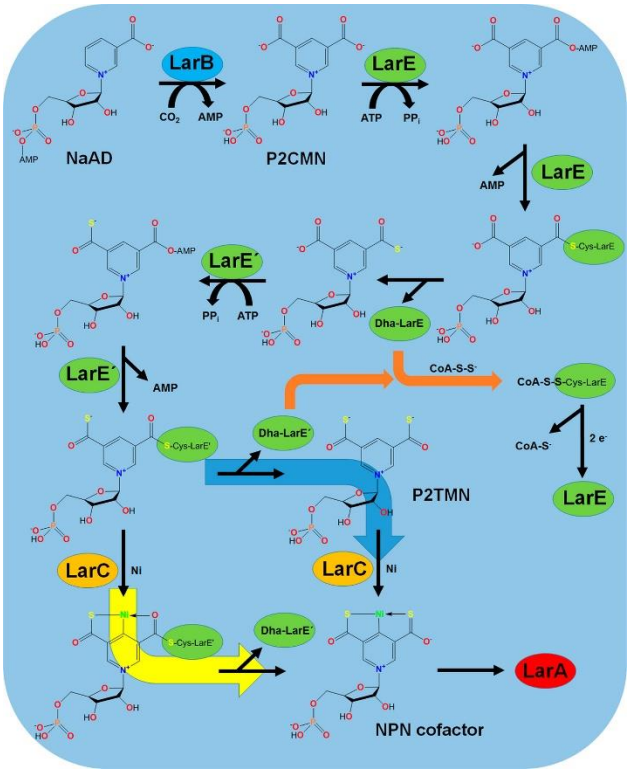


Figure 4-11. Biosynthesis of the NPN cofactor for LarA apoprotein activation, regeneration of LarE^{Dha}, and an alternate pathway to generate lactate racemase activity. In the predominant pathway, LarB carboxylates nicotinic acid adenine dinucleotide (NaAD) and hydrolyzes its phosphoanhydride linkage, thus forming P2CMN. LarE activates the substrate carboxyl group by adenylylation, forms a thioester linkage with the substrate, and sacrifices its Cys176 sulfur atom, and then a second LarE repeats the cycle to produce P2TMN (blue arrow). LarC inserts nickel into this species to generate the NPN cofactor that activates LarA. LarE^{Dha} can be regenerated to native LarE by incubation with CoA persulfide and reductant (orange arrows), but it is unclear whether these reactions are physiologically relevant. A newly identified NPN-LarE intermediate functions in an alternative, minor pathway (yellow arrow) in which LarC installs nickel into the pincer cofactor while it is still bound to LarE. NPN is released from this purified protein for LarA activation.

to bind at the CoA binding site of LarE for this reaction because other tested persulfides were inefficient or not effective. Such binding would perfectly position the persulfide to react with Dha. Persulfides are known to have increased nucleophilic character compared to thiols and are of increasing interest in biology.[23, 24] CoA persulfide is found in cells[25, 26] and has been shown to bind tightly to short chain acyl-CoA dehydrogenases.[27, 28] Although CoA persulfide is an attractive candidate for recycling $LarE^{Dha}$, the near exclusive presence of this form of the protein in *Lc. lactis* cells containing pGIR172 suggests this reaction is not physiologically relevant and LarE is a single-turnover enzyme.

Our finding of what appears to be an NPN-LarE adduct suggests that LarC can insert nickel into the incomplete pincer ligand that is still bound in the thioester linkage to LarE (Figure 4-11, yellow arrow). The flexibility of the LarE loop containing Cys176 could reasonably allow the pincer species to be metalated by LarC. The NPN-LarE species is presumed to release NPN, accounting for the previously reported activation of LarA apoprotein by just LarE,[5] resulting in the generation of the $LarE^{Dha}$ form of the protein. The *in vivo* relevance of the NPN-LarE adduct remains to be clarified; this species may be abundant within the cell but converts to $LarE^{Dha}$ as the protein is purified.

LarE Homologues: Potential Roles of Additional Cysteine Residues

We had previously noted that many LarE homologues have sequences in which Cys176 appears to be shifted by one residue and points toward two additional cysteine residues, not present in the *L. plantarum* LarE sequence, one of which substitutes for Trp97 (Figure 4-12A,B).[9] We speculate that these alternative LarE versions might catalyze a different mechanism of sulfur transfer that does not require sacrifice of a side chain sulfur atom as observed in the benchmark LarE. Comparison of the location for these cysteine residues in a

model of *T. thermosaccharolyticum* LarE (Figure 4-12B) to the structure of *Thermus thermophilus* TtuA,[29] another PP-loop pyrophosphatase family member that catalyzes sulfur transfer for 2-thiouridine synthesis, reveals the similarity of the LarE ortholog and the [4Fe-4S] cluster binding site bound by three cysteine residues in TtuA (Figure 4-12C). In that case, the sulfur transferred to uridine is suggested to derive from the non-[4Fe-4S] cluster sulfide atom (positioned where a fourth cysteine is found for many similar clusters), which must be replaced for each round of synthesis.[29] Structural similarities and sulfur transfer involving a labile extra cluster sulfide were noted in a study of TtuA from *Pyrococcus horikoshii* that also possesses a [4Fe-4S] cluster bound by three cysteines (Figure 4-12D).[30] Furthermore, the structure of the more distantly related ortholog adenosine 5′-phosphosulfate reductase from *Pseudomonas aeruginosa* also reveals a [4Fe-4S] cluster at this site;[31] however, in that case, four protein cysteine residues are present and the cluster is not thought to donate a sulfur atom (Figure 4-12E). Rather, this enzyme uses its [4Fe-4S] and flavin to reversibly transfer electrons from sulfite and AMP while forming adenylyl sulfate and a reduced electron acceptor. Interestingly, two other structurally related sulfur transferases, tRNA 2-thiouridylase or MnmA from *E. coli*

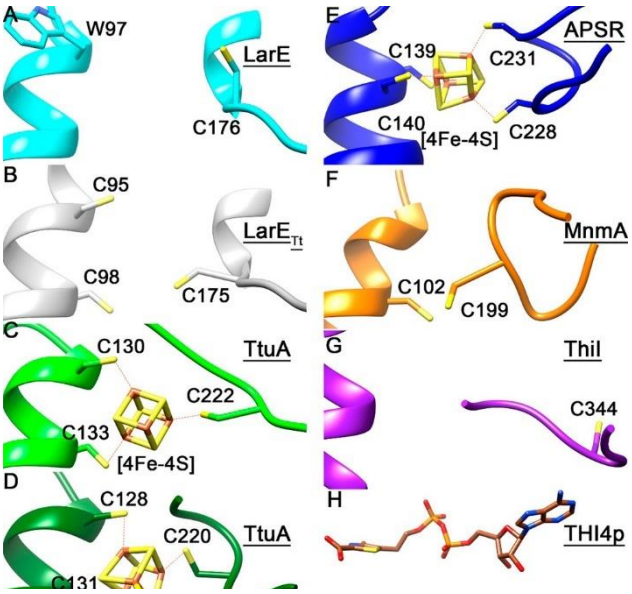


Figure 4-12. Comparison of active sites of structurally related sulfur transferases. (A) LarE from *L. plantarum* colored cyan (PDB entry 5UDR). (B) Homology model of LarE$_{Tt}$ from *T. thermosaccharolyticum* colored gray. (C) 2-Thiouridine synthetase (TtuA) from *Th. thermophilus* colored light green (PDB entry 5B4E). (D) TtuA from *P. horikoshii* colored dark green (PDB entry 5MKP). (E) Adenosine 5′-phosphosulfate reductase from *Ps. aeruginosa* colored blue (PDB entry 2GOY). (F) tRNA 2-thiouridylase (MnmA) from *E. coli* colored orange (PDB entry 2DEU). (G) tRNA 4-thiouridine synthetase (ThiI) from *Thm. maritima* colored purple (PDB entry 4KR6). (H) Functionally related sacrificial sulfur insertase (THI4p) from *S. cerevisiae* colored brown with bound adenylylated thiazole (PDB entry 3FPZ).

(Figure 4-12F)[32] and tRNA 4-thiouridine synthetase or ThiI from *Thermotoga maritima* (Figure 4-12G),[33] also contain their catalytic cysteine residues in the same region as LarE's Cys176. In these cases, cysteine persulfides are generated, with disulfide exchange involving another cysteine residue allowing for the insertion of a sulfur atom into the substrate. Although the structure has not been obtained, ThiI from *Methanococcus maripaludis* was shown to possess a [3Fe-4S] cluster that was needed for sulfur transfer.[34]

One precedent exists for a sacrificial sulfur transferase, THI4p that functions in the thiamine biosynthesis pathway of *Saccharomyces cerevisiae*.[35] Although related in function to LarE and similarly thought to catalyze a single turnover, THI4p possesses a distinct fold (Figure 4-12H). Of further interest, *Methanococcus jannaschii* possesses an ortholog of THI4p that uses exogenous sulfide for thiamin biosynthesis and is catalytic.[36] This example, like that mentioned above for ThiI, illustrates the situation in which orthologs of the same protein can exhibit distinct properties. In this case, one example is a sacrificial single-turnover enzyme while the other ortholog exhibits multiple-turnover catalytic activity.

LarE is structurally related to other sulfur transferases that contain [4Fe-4S] clusters or generate persulfides at their active sites, such as TtuA, MnmA, and ThiI. In addition, it is functionally related to sulfur transfer enzymes in thiamin synthesis that utilize two distinct mechanisms, one stoichiometric and one catalytic. Thus, we hypothesize that other LarE orthologs may bind a [4Fe-4S] cluster or generate persulfides and use these species for catalytic sulfur transfer. LarE homologues that use three active site cysteine residues to bind a [4Fe-4S] cluster could use the nonligated sulfur atom for attack of the AMPylated substrate without sacrificing a cysteine residue. Similarly, a persulfide could attack the adenylylated substrate with another cysteine forming a disulfide to allow for sulfur transfer. The active form of the enzyme

could then be reconstituted by the normal cellular cluster synthesis machinery or by persulfide formation to allow for additional rounds of sulfur transfer. It will be interesting to assess whether LarE functions catalytically in the multi-Cys proteins.

Conclusions

We elucidated the structure of the complex between CoA and LarE and unexpectedly observed strong similarities in binding modes for CoA and ATP. The sulfur atom of CoA can adopt multiple positions, including formation of the previously identified disulfide with Cys176 of LarE, and the structures provide insights into how CoA provides thermal stability to the protein. The persulfide of CoA was competent for restoring the sulfur atom to $LarE^{Dha}$, whereas other LarE regeneration efforts were poorly effective or unsuccessful. The low extent of LarE recycling *in vitro* and the lack of evidence for *in vivo* recycling are consistent with our prior conclusion that LarE is a sacrificial protein.[9] Only two LarE molecules are required for activation of each cofactor that becomes bound to LarA, so a stoichiometric mechanism may be sufficient for cellular needs. We identified a new species, the NPN adduct of LarE that explains the perplexing prior results in which purified LarE was sufficient to activate LarA apoprotein. Release of NPN from the small proportion of LarE with a bound cofactor accounted for LarA activation when using a vast excess of the accessory protein. The observation of this adduct shows that LarC can catalyze the metalation reaction of the organic ligand bound in a thioester linkage to LarE. We determined the structure of the C176A variant and found that the lack of the thiol side chain led to no conformational changes compared to the conformation of the WT enzyme. We also created three additional active site variants of LarE and discussed their activities in a structural context, along with those of previously described variants. Finally, we performed a structural comparison of the *L. plantarum* LarE to a homology model for an

alternative version of LarE with a predicted cluster of cysteine residues and other related enzymes that possess [4Fe-4S] clusters or a reactive persulfide in their active sites. From this analysis, we speculate that some orthologs of LarE may bind a [4Fe-4S] cluster providing them with the ability to act catalytically. Our supposition of two forms of an enzyme, one acting stoichiometrically to form $LarE^{Dha}$ and the other acting catalytically, has precedence in thiamin biosynthesis.

REFERENCES

REFERENCES

(1) Juturu, V., and Wu, J. C. (2015) Microbial production of lactic acid: The latest development, *Crit. Rev. Biotechnol. 36*, 967-977.

(2) Desguin, B., Soumillion, P., Hols, P., Hu, J., and Hausinger, R. P. (2017) Lactate racemase and its niacin-derived, covalently-tethered, nickel cofactor, In *The Biological Chemistry of Nickel* (Zamble, D. B., Rowinska-Zyrek, M., and Kozlowski, H., Ed.), pp 220– 236, Royal Society of Chemistry, Cambridge, U.K.

(3) Fellner, M., Rankin, J. A., Hu, J., and Hausinger, R. P. (2017) Lactate Racemase, In *Encyclopedia of Inorganic and Bioinorganic Chemistry* (Scott, R. A., Ed.), Wiley & Sons, West Sussex, U.K.

(4) Desguin, B., Soumillion, P., Hausinger, R. P., and Hols, P. (2017) Unexpected complexity in the lactate racemization system of lactic acid bacteria, *FEMS Microbiol. Rev. 41*, S71-S83.

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

(6) 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.

(7) Hausinger, R. P., Desguin, B., Fellner, M., Rankin, J. A., and Hu, J. (2018) Nickel-pincer nucleotide cofactor, *Curr. Opin. Chem. Biol. 47*, 18-23.

(8) 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. U. S. A. 113*, 5598-5603.

(9) 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. U. S. A. 114*, 9074-9079.

(10) 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.

(11) Krug, M., Weiss, M. S., Heinemann, U., and Mueller, U. (2012) XDSAPP: a graphical user interface for the convenient processing of diffraction data usingXDS, *J. Appl. Crystallogr. 45*, 568-572.

(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 Biol. Crystallogr. 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 Biol. Crystallogr. 69*, 1204-1214.

(14) Adams, P. D., Afonine, P. V., Bunkoczi, 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 Biol. Crystallogr. 66*, 213-221.

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

(16) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J. Comput. Chem. 25*, 1605-1612.

(17) Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid)—a reexamination, *Anal. Biochem. 94*, 75-81.

(18) Scian, M., Guttman, M., Bouldin, S. D., Outten, C. E., and Atkins, W. M. (2016) The myeloablative drug busulfan converts cysteine to dehydroalanine and lanthionine in redoxins, *Biochemistry 55*, 4720-4730.

(19) Englander, S. W., Calhoun, D. B., and Englander, J. J. (1987) Biochemistry without oxygen, *Anal. Biochem. 161*, 300-306.

(20) Mueller, E. G., Palenchar, P. M., and Buck, C. J. (2001) The role of the cysteine residues of ThiI in the generation of 4-thiouridine in tRNA, *J. Biol. Chem. 276*, 33588-33595.

(21) Kabil, O., and Banerjee, R. (2012) Characterization of patient mutations in human persulfide dioxygenase (ETHE1) involved in H_2S catabolism, *J. Biol. Chem. 287*, 44561-44567.

(22) Wood, J. L. (1987) Sulfane sulfur, *Methods Enzymol. 143*, 25-29.

(23) Cuevasanta, E., Moller, M. N., and Alvarez, B. (2017) Biological chemistry of hydrogen sulfide and persulfides, *Arch. Biochem. Biophys. 617*, 9-25.

(24) Park, C. M., Weerasinghe, L., Day, J. J., Fukuto, J. M., and Xian, M. (2015) Persulfides: current knowledge and challenges in chemistry and chemical biology, *Mol. Biosyst. 11*, 1775-1785.

(25) Shen, J., Keithly, M. E., Armstrong, R. N., Higgins, K. A., Edmonds, K. A., and Giedroc, D. P. (2015) *Staphylococcus aureus* CstB is a novel multidomain persulfide dioxygenase-

sulfurtransferase involved in hydrogen sulfide detoxification, *Biochemistry 54*, 4542-4554.

(26) Shen, J., Walsh, B. J. C., Flores-Mireles, A. L., Peng, H., Zhang, Y., Zhang, Y., Trinidad, J. C., Hultgren, S. J., and Giedroc, D. P. (2018) Hydrogen sulfide sensing through reactive sulfur species (RSS) and nitroxyl (HNO) in *Enterococcus faecalis*, *ACS Chem. Biol.*

(27) Shaw, L., and Engel, P. C. (1987) CoA-persulphide: a possible in vivo inhibitor of mammalian short-chain acyl-CoA dehydrogenase, *Biochim. Biophys. Acta, Lipids Lipid Metab. 919*, 171-174.

(28) Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W., Vockley, J., and Kim, J. J. (1997) Structure of human isovaleryl-CoA dehydrogenase at 2.6 Å resolution: structural basis for substrate specificity, *Biochemistry 36*, 8455-8464.

(29) Chen, M., Asai, S. I., Narai, S., Nambu, S., Omura, N., Sakaguchi, Y., Suzuki, T., Ikeda-Saito, M., Watanabe, K., Yao, M., Shigi, N., and Tanaka, Y. (2017) Biochemical and structural characterization of oxygen-sensitive 2-thiouridine synthesis catalyzed by an iron-sulfur protein TtuA, *Proc. Natl. Acad. Sci. U. S. A. 114*, 4954-4959.

(30) Arragain, S., Bimai, O., Legrand, P., Caillat, S., Ravanat, J. L., Touati, N., Binet, L., Atta, M., Fontecave, M., and Golinelli-Pimpaneau, B. (2017) Nonredox thiolation in tRNA occurring via sulfur activation by a [4Fe-4S] cluster, *Proc. Natl. Acad. Sci. U. S. A. 114*, 7355-7360.

(31) Chartron, J., Carroll, K. S., Shiau, C., Gao, H., Leary, J. A., Bertozzi, C. R., and Stout, C. D. (2006) Substrate recognition, protein dynamics, and iron-sulfur cluster in *Pseudomonas aeruginosa* adenosine 5'-phosphosulfate reductase, *J. Mol. Biol. 364*, 152-169.

(32) Numata, T., Ikeuchi, Y., Fukai, S., Suzuki, T., and Nureki, O. (2006) Snapshots of tRNA sulphuration via an adenylated intermediate, *Nature 442*, 419-424.

(33) Neumann, P., Lakomek, K., Naumann, P. T., Erwin, W. M., Lauhon, C. T., and Ficner, R. (2014) Crystal structure of a 4-thiouridine synthetase-RNA complex reveals specificity of tRNA U8 modification, *Nucleic Acids Res. 42*, 6673-6685.

(34) Liu, Y., Vinyard, D. J., Reesbeck, M. E., Suzuki, T., Manakongtreecheep, K., Holland, P. L., Brudvig, G. W., and Soll, D. (2016) A [3Fe-4S] cluster is required for tRNA thiolation in archaea and eukaryotes, *Proc. Natl. Acad. Sci. U. S. A. 113*, 12703-12708.

(35) Chatterjee, A., Abeydeera, N. D., Bale, S., Pai, P. J., Dorrestein, P. C., Russell, D. H., Ealick, S. E., and Begley, T. P. (2011) *Saccharomyces cerevisiae* THI4p is a suicide thiamine thiazole synthase, *Nature 478*, 542-U146.

(36) Eser, B. E., Zhang, X., Chanani, P. K., Begley, T. P., and Ealick, S. E. (2016) From suicide enzyme to catalyst: the iron-dependent sulfide transfer in *Methanococcus jannaschii* thiamin thiazole biosynthesis, *J. Am. Chem. Soc. 138*, 3639-3642.

CHAPTER 5

Additional studies

2-Hydroxyisobutyrate does not inhibit LarA

2-Hydroxyisobutyric acid (2-HIB) (Figure 5-1) was tested as a potential inhibitor of LarA. This compound contains two methyl groups on the α-carbon and is not optically active. As LarA accepts both D- and L-lactate as substrates, it is reasonable to postulate that the enzyme must be able to accept a methyl group at either position. 2-HIB lacks a hydrogen at the α-carbon; thus, LarA cannot utilize it as a substrate. Such a non-reactive substrate analog may be useful for crystallographic work and for electron paramagnetic resonance studies, as it might bind tightly to the active site in a single conformation and clarify how the authentic substrate interacts with NPN.

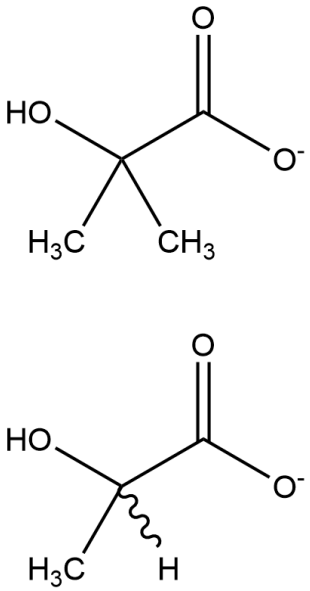


Figure 5-1. 2-HIB (top, shown in anion form) compared to lactate. Note that 2-HIB is symmetrical around the α-carbon.

Using the assay described in Chapter 2 (containing 20 mM lactate), I examined the extent of inhibition using purified LarA or *Lactococcus lactis* lysates containing LarA with varying amounts of 2-HIB that was neutralized with NaOH. No convincing inhibition was observed until reaching 200 mM of 2-HIB where ~50% of the control activity was retained (Figure 5-2). I deemed that further investigation of this compound as a LarA substrate analog was not warranted.

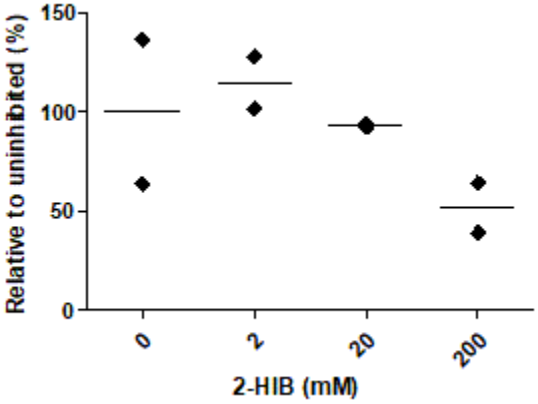


Figure 5-2. Inhibition of LarA by 2-HIB.

L. lactis cells containing pGIR100 and its derivatives are poorly resistant to chloramphenicol

Plasmid pGIR100[1] was formed by cloning the *lac* operon into pNZ8048, [2] which contains the nisin A inducible *nis* promotor from *L. lactis* and carries a constitutively expressed gene encoding chloramphenicol acetyltransferase (CAT), thus conferring resistance to the antibiotic chloramphenicol. When grown under standard conditions on M17 plates containing chloramphenicol (Chapter 2), *L. lactis* containing pGIR100 and its derivative, pGIR112,[1] exhibited significant growth defects compared to *L. lactis* containing pNZ8048. In addition, pGIR100 and its derivatives produced smaller colonies and developed colonies of various sizes. This defect was not present when chloramphenicol was excluded from the medium.

To investigate the source of this anomaly, I first eliminated the possibility that the non-uniform colony size was a contamination issue. When colonies of varying sizes were picked and streaked onto new plates, all had the same colony size distribution, regardless of the parent colony size.

Secondly, I sequenced the gene encoding CAT from the isolates to check for defects. I found that nearly all bases departing from the published sequence were synonymous substitutions or occurred after the end of the gene. The one exception was a change in the promotor; however, that change was also found in pNZ8048, so it cannot account for the growth defect. Furthermore, examples of the promoter variants were found in the BLAST database. The source of the growth defect has not been identified, but it has been ameliorated by using larger cultures and lower chloramphenicol concentrations when producing recombinant LarA.

Synthesis of 3,5-bis(methoxycarbonyl)-1-methylpyridin-1-ium

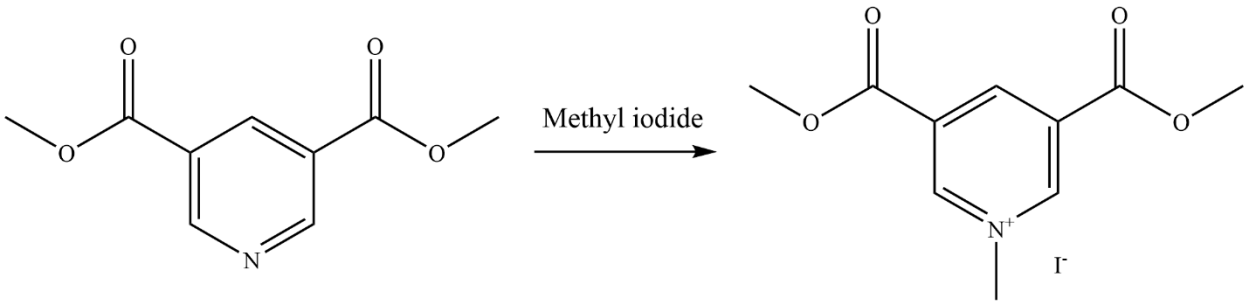


Figure 5-3. Proposed synthesis scheme of a putative LarE substrate analog. Methylation of dimethyl pyridine-3,5-bicarboxylic acid was carried out to synthesize the potential substrate analog, 3,5-bis(methoxycarbonyl)-1-methylpyridin-1-ium (shown on the right).

The native substrate of LarE, pyridinium 3,5-biscarboxylic acid mononucleotide (P2CMN), is not commercially available and has not yet been chemically synthesized. I hypothesized that a simpler compound lacking the ribose phosphate portion of the molecule may be useful for crystal soaking during structural studies and for enzyme kinetics analyses. I attempted to synthesize 3,5-bis(methoxycarbonyl)-1-methylpyridin-1-ium from the commercially available dimethyl pyridine-3,5-biscarboxylic acid (DM35PC) by exploiting methyl iodide, a commonly used methylating reagent (Figure 5-3). The methyl pyridinium species is reminiscent of P2CMN, the native LarE substrate, by possessing a positive charge on the nitrogen, as opposed to the uncharged pyridine. I judged that the methyl esters may not interfere with binding to the enzyme, since LarE accommodates an AMP phosphoester at these positions. Alternatively, the methyl esters would be straightforward to remove at a later step.

For a small trial, I dissolved 22.8 mg of DM35PC into 1 mL of methanol, added 14.5 µL of methyl iodide (a 2-fold molar excess), and allowed the mixture to react at room temperature for 2 h in the dark. The solution was dried under vacuum to yield a yellow product, in contrast to the off-white to light yellow color of the starting compound. The resulting mass (23.7 mg) was less than a 5% mass increase from the original 22.8 mg compared to the theoretical maximum yield of 39.4 mg of the iodide salt.

To increase the yield, I set up an overnight reaction with 18.3 mg of DM35PC dissolved in 1 mL of methanol with 11.7 µL of methyl iodide added. After 1 day in the dark at room temperature, the product was dried under vacuum, yielding a significantly yellow powder. This product was dissolved in 1 mL of methanol and the far-UV spectrum was obtained.

The profile below 300 nm was largely unchanged compared to that of DM35PC, however a new absorption band centered at 362 nm appeared (Figure 5-4). Interestingly, the 362 nm band did not form when the solvent was changed to water. This result suggests either that the desired product was not generated, or the compound is labile when using water. A mass spectrum was obtained using electrospray ionization in positive ionization mode and revealed the expected reaction product (210.08 *m/z*), consistent with the trimethylated species, but also demonstrated that remaining DM35PC (196.06 *m/z*) was the predominant species.

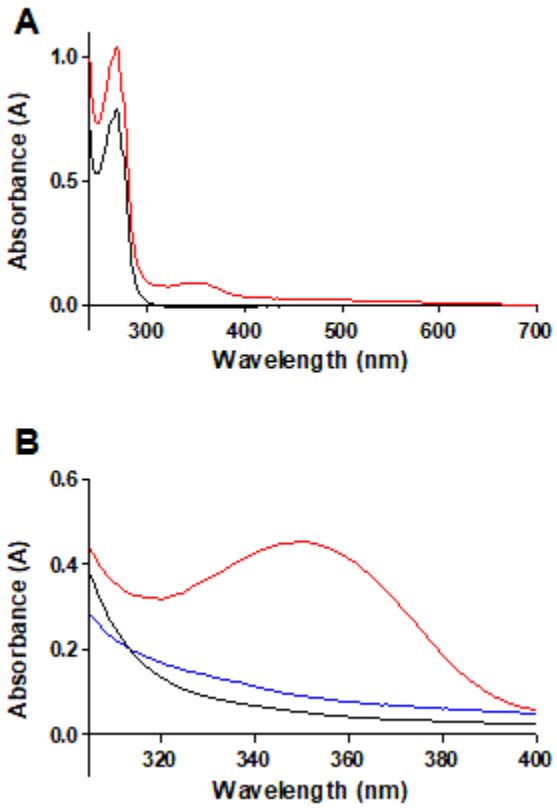


Figure 5-4. UV spectra of the synthesis product. A) Full UV-vis spectra of DM35PC (black) and the product (red) dissolved in methanol. Concentration is approximately 0.1 mg/mL. B) The product dissolved in methanol (red) has an absorbance band at 352 nm that is not seen when the product is dissolved in water (blue) or in DM35PC dissolved in methanol (black). Concentrations are approximately 2 mg/mL. All spectra were zeroed at 700 nm.

To separate the two species, I used carboxymethyl (CM)-Sepharose, a weak cation-exchange resin. This separation exploited the positive charge of the methylated pyridinium species, which should exhibit greater affinity for the resin than the starting DM35PC. The packed column (7 mm diameter by 51 mm in length)

was cleaned with 10 mM hydrochloric acid, equilibrated in 10 mM ammonium acetate, injected with ~20 µmol of the reaction product (dissolved in 150 µL of water), and eluted with 10 mM ammonium acetate at 200 µL/min. Fractions (0.5 mL) were collected while monitoring the UV absorbance at 254 nm. Two elution peaks were observed, a large early peak in fraction 3 and a smaller peak in fraction 12.

The UV spectra of these fractions support a preliminary assignment of fraction 3 as containing the starting material, DM35PC, and fraction 12 as containing the methyl pyridinium product. The three electronic transitions observed between 260 nm and 280 nm in fraction 3 are very similar to the bands observed in DM35PC (Figure 5-5 top, compare to Figure 5-4). In fraction 12, the strongest absorption band was red shifted by 2 nm compared to that for fraction 3 and DM35PC. In addition, the prominence of the 277 nm band was increased relative to the center electronic transition, while the high energy band was weakened. The 352 nm feature seen in earlier crude preparations (that vanished when dissolved in water) was observed in this preparation, albeit the absorption was weak (Figure 5-5, bottom). It seems

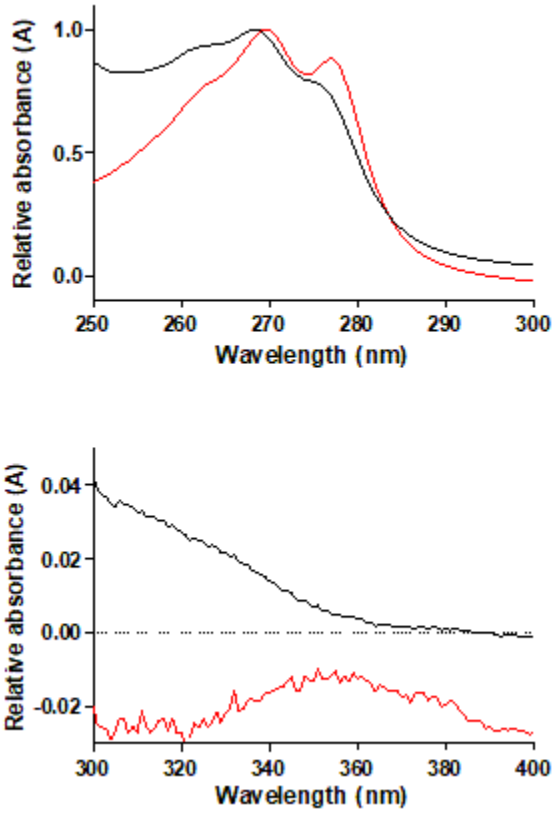


Figure 5-5. Two views of the UV spectra of the CM Sepharose elution fractions. Fraction 3 (putative DM35PC starting substrate) is in black and fraction 12 (putative product) is in red. Relative absorbance has been set to 1 for the 268-270 nm peak and zeroed at 700 nm.

probable that this compound is hydrolyzed slowly in neutral buffered water and methanol compared to unbuffered water.

Dr. Matthias Fellner used this compound in crystal soaking studies with LarE, but he did not detect any extra density that would be consistent with binding, and further efforts to use this substrate analog were not pursued.

Testing a methyl pyridinium analog of NPN

Dr. Xile Hu provided two samples, one of a higher purity than the other, of a potential NPN analog (Figure 5-6). Both samples consisted of large dark green crystals in a brown powder, though a noticeably greater portion of the dark green crystals were in the more refined sample. The more refined sample was mostly soluble to a calculated concentration of 10 mM in water, with a small amount of the brown powder remaining undissolved. The less refined sample was soluble to 5 mM in 50% dimethyl sulfoxide.

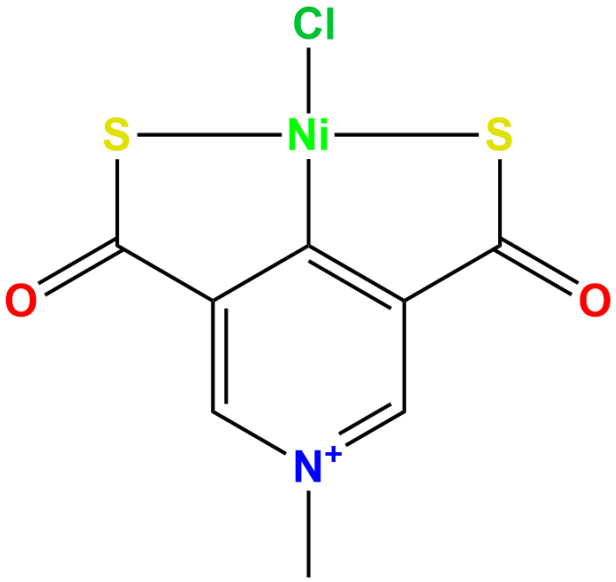


Figure 5-6. The NPN analog provided by Dr. Xile Hu. The compound was provided as the tetraphenylphosphonium salt

To test whether this compound can substitute for NPN in LarA, an assay for LarE activity was adapted.[3] In short, *L. lactis* cells harboring pGIR082, which encodes a *T. thermosaccharolyticum* LarA homolog, were induced and lysed using a bead mill. The clarified lysates were used as a source of LarA apoprotein. 5 µL of lysate was used in a total reaction volume of 50 µL in HEPES buffer, pH 7.0, and included 50 mM sodium L-lactate and 1 mM of the compound. These mixtures were incubated for 15 min at 50 °C and terminated by boiling.

The resulting solutions were then assayed for D-lactate. No significant conversion to D-lactate was observed. It is likely that the absence of ribose phosphate, a component of NPN, prevented binding of the analog to the enzyme.

Selenite and arsenite are not strong inhibitors of LarA

Sulfite was identified as a potent inhibitor of LarA and found to protect the cofactor from atmospheric oxygen, consistent with it interacting directly with the NPN cofactor as revealed later in crystallography studies.[4] To further probe how sulfite might interact with the nickel atom of NPN, I considered the possibility of using X-ray absorption spectroscopy (XAS) to deduce changes in the metal electronic environment and to assess any changes in the numbers and sizes of nearby scattering atoms. However, because the nickel center already has two sulfur ligands, any interaction with the sulfur atom of sulfite would be difficult to establish. As a potential approach to overcome this issue, I tested the interaction of LarA with arsenite and selenite, analogs that are isoelectronic with sulfite. The large electron density of these two atoms would facilitate investigations by XAS and may also aid in resolving an inhibitor-bound crystal structure.

I first verified that the lactate assay was not inhibited by selenite or arsenite by including them in a reaction with a known amount of D-lactate. Then, I conducted the LarA assay as described in chapter 2 with the sulfite concentration reduced to less than 100 nM and the putative inhibitors added as their sodium salts. My results indicate that the compounds

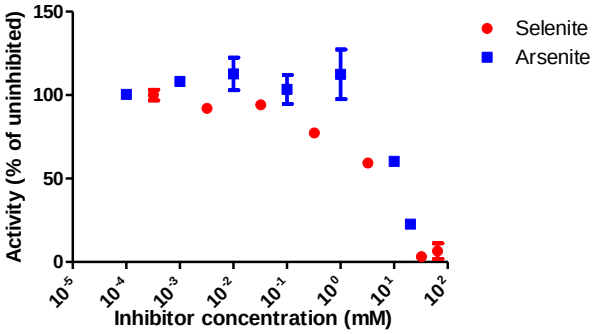


Figure 5-7. LarA inhibition by sodium selenite and sodium arsenite.

are not inhibitory until reaching the low mM range (Figure 5-7). On the basis of these results, and a later crystal structure of sulfite-bound LarA, this avenue was not pursued further.

LarB domain cloning

Table 5-5. List of oligonucleotide sequences

Name	Description	Sequence (5' to 3')
JR026	LarB LP C-term Rev	TAGGTACCCATGGTTAATTCCTCCTGTTAGC
JR027	LarB LP C-term Δ1-108 Fwd	CAGGTACCGAACAACCCGCACCGAAAAC
JR028	LarB LP N-term Δ109-246 Rev	CGTACTCGAGTCCCCGACCGTCATACACTG
JR029	LarB LP N-term Fwd	AGCACTCGAGCTGGAGTCATCCACAATTTG
JR030	LarB LP C-term Δ1-118 Fwd	CAGGTACCATTGCTGTGGTGACGGC
JR031	LarB LP N-term Δ119-246 Rev	CGTACTCGAGTTGTACCCTGGTGTTTTCGGTG
JR033	Forward for d1-102 LarB Tt	ATTAACCATGGATGTGGCGAGGATTATATC
JR034	Reverse for d1-102 LarB Tt	TCGCCACATCCATGGTTAATTCCTCCTGTTAG
JR035	Forward for d103-C LarB Tt	AGTTTATTACGCTAGCTGGAGTCATCC
JR036	Reverse for d103-C LarB Tt	TCCAGCTAGCGTAATAAACTGCTTTATCGCATACTTCC
JR037	Forward for dN -119 LarB Tt	ATTAACCATGGTTATCGGTGTTGTGGC
JR038	Reverse for dN-119 LarB Tt	CACCGATAACCATGGTTAATTCCTCCTGTTAG
JR039	Forward for d120-C LarB Tt	TACAAAAGGTGCTAGCTGGAGTCATCCAC
JR040	Reverse for d120-C LarB Tt	TCCAGCTAGCACCTTTTGTAGGCATAATCTCTTCTG
JR041	F-MBP all LarB Lp domain cloning	GGAATTAACCATGCACCATCATCATCATCATTC
JR042	R-vector for all LarB Lp MBP domain cloning	GATGGTGCATGGTTAATTCCTCCTGTTAGCC
JR043	R-MBP for MBP:LarB(Lp)ΔN-108	GTTCGGTACCGGCATTGGATTGGAAGTACAG
JR044	F-vector for MBP:LarB(Lp)ΔN-108	ATCCAATGCCGGTACCGAACAACCCG
JR045	F-vector for both ΔC MBP domains	ATCCAATGCCGCAACCACAGCAGAAATATTAC
JR046	R-MBP for both ΔC MBP domains	CTGTGGTTGCGGCATTGGATTGGAAGTACAG
JR047	R-MBP for MBP:LarB(Lp)ΔN-118	CAATGGTACCGGCATTGGATTGGAAGTACAG
JR048	F-vector for MBP:LarB(Lp)ΔN-118	ATCCAATGCCGGTACCATTGCTGTGGTG
JR075	LarBLp dN-45-F pET NcoI-F	TAACCATGGCAAATGGCTTTCCGGAGG
JR076	LarBLp dN-45 pET NcoI-R	TTGCCATGGTATATCTCCTTCTTAAAG

Although the initial crystals of LarB homologs from both *Lactobacillus plantarum* and *Thermoanaerobacterium thermosaccharolyticum* were visually appealing, they diffracted poorly. One potential strategy to enhance resolution is to crystallize individual domains of the protein with the hope that such domains provide better packing in the crystal. LarB was predicted to have two domains, so I attempted to express these individual domains. On the basis of homology to known PurE sequences and using domain prediction software (BLAST, InterPro), I identified the likely junction between the domains. LarB from *L. plantarum* was sectioned at residues 109 and 119 and LarB from *T. thermosaccharolyticum* at residues 102 and 119.

All methods used Q5 polymerase, KpnI, NcoI, DpnI (NEB), T4 ligase (Invitrogen) and PCR cleanup (Qiagen) and were used according to the manufacturer's recommendation. All clones were verified by sequencing.

The *L. plantarum* LarB variants were cloned with a cut and ligate strategy using pBAD:LarB$_{Lp}$ (Chapter 3) as the PCR substrate. pBAD:LarB(*Lp)*ΔN-108 was cloned using primers JR26 and JR27. The product was separated on an agarose gel, stained using 0.02% w/v methylene blue in Tris-borate-EDTA buffer, excised, and purified. After digestion with KpnI, the product was purified using a cleanup kit, ligated using T4 ligase, and transformed into *Escherichia coli* DH5α. The remaining *L. plantarum* LarB clones were not gel purified and were cut simultaneously using KpnI and DpnI. LarB(*Lp*)Δ109-C used primers JR28 and JR29, LarB(*Lp*)ΔN-118 used JR26 and JR30, and LarB(*Lp*)Δ119-C used JR29 and JR31. Expression was tested in *E. coli* BL21 (DE3) at 37 °C and 25 °C with induction using 0.2% w/v L-arabinose. Expression was also tested using the ArcticExpress (Agilent) plasmid in *E. coli* BL21 (DE3) at 12 °C. Expression was very poor at 37 °C and 25 °C and all products at 12 °C were associated with inclusion bodies.

The *T. thermosaccharolyticum* LarB variants were created using gap-repair cloning (also known as *in vitro* assembly).[5, 6] In short, PCR fragments of the constructs were made that excluded the deleted region and included *ca.* 20 bp of overlap at the ends. The fragments were transformed into chemically-competent *E. coli* DH5α, which combined the homologous ends. pBAD:LarB(*Tt*) (acquired from Dr. Benoît Desquin, unpublished) was used as the template. pBAD:LarB(*Tt*)ΔN-101 was created using primers JR33 and JR34, pBAD:LarB(*Tt*)Δ102-C used JR35 and JR36, pBAD:LarB(*Tt*)ΔN-119 used JR37 and JR38, and pBAD:LarB(*Tt*)Δ120-C used JR39 and JR40. I performed trial inductions in *E. coli* BL21(DE3) at 37 °C using 0.2% w/v L-

arabinose. The N-terminal domains did not express at detectable levels and the C-terminal domains had very poor expression.

To improve the solubility of the *L. plantarum* LarB domains, N-terminal maltose binding protein (MBP) fusions were constructed using Gap-repair cloning. In short, vector and insert PCR fragments were created with *ca.* 20 bp overlaps then transformed into chemically competent *E. coli* DH5α which recombined the fragments. The MBP insert was obtained using pMCSG9-MBP-EFE (from Dr. Salette Martinez-Martinez, unpublished) as a template. This construct included an N-terminal His_6 tag and a TEV cleavage site between MBP and the product of interest. All LarB domain expression PCR products used the expression construct described above. pBAD:MBP:LarB(*Lp*)ΔN-108 was created using primers JR41 and JR43 for the MBP sequence and JR42 and JR44 for the vector. pBAD:MBP:LarB(*Lp*)ΔN-118 was created using primers JR41 and JR47 for the MBP gene and JR42 and JR48 for the vector. Both pBAD:MBP:LarB(*Lp*)Δ109-C and pBAD:MBP:LarB(*Lp*)Δ119-C were created using primers JR41 and JR46 for the MBP DNA and JR42 and JR45 for the insert.

Trial inductions of these constructs showed robust production of all four proteins. The constructs had three tags, allowing purification by either StrepTactin, Ni-nitrilotriacetic acid (NTA), or amylose resins. Ni-NTA was used in a trial purification of MBP:LarB(*Lp*)ΔN-108. The protein displayed poor affinity for the resin with most flowing through, though some of the desired protein did bind and was then able to be eluted. This product was able to be cleaved by TEV protease to produce free LarB LarB(*Lp*)ΔN-108. To improve yields, I used StrepTactin-XT and amylose resins. These resins led to similar results as for the Ni-NTA purification: most of the protein did not bind, with only a portion able to be purified. Different salt conditions were used (2 M NaCl or 100 mM K_2SO_4) along with 0.1% v/v Triton X-100. Changing the salt

conditions did not help to increase affinity to the columns, but Triton X-100 improved the affinity of the protein for both the Ni-NTA and amylose resins.

The purified LarB(*Lp*)ΔN-108 was tested for LarB activity using the ^{14}C incorporation assay (Chapter 3). The truncated protein had no significant activity, although concentrated protein incubated overnight may have incorporated a small amount of isotope into the substrate.

I found that all four proteins had low affinity to different resins. Because the tags are differently positioned in the proteins, it is unlikely that the tags are all occluded – especially given that the products appear to have folded normally. Rather, the finding that addition of Triton X-100 improved the protein affinity for the resins may suggest misfolding or aggregation problems due to hydrophobic interaction. Because the goal of this project was to attempt crystallization of protein domains in lieu of the suitable full-length crystals, but a good crystallization condition for LarB from *L. plantarum* was identified, this effort was set aside.

The structure of $LarB_{Lp}$ was solved to a resolution of 2.7 Å (Chapter 3). However, the first 45 residues could not be determined from the electron density map. This deficiency may be a consequence of those residues having multiple conformations and/or being unstructured. Such an unstructured region may interfere with orderly packing within the crystal, potentially reducing the resolution, so I created a construct with the first 45 residues removed (a smaller deletion than detailed above). This construct was made using pET:LarB (Chapter 3). pET:LarBΔN-45 was created by amplifying pET:LarB using primers JR75 and JR 76, purifying the product, cutting with NcoI inactivating the restriction enzyme by heating at 80 °C for 20 min, ligating the desired fragment using T4 ligase, transforming the product into *E. coli* DH5α, and verifying by sequencing.

pET:LarBΔN-45 was co-transformed with the ArcticExpress plasmid (Agilent) into *E. coli* BL21 (DE3) and selected with 10 µg/mL gentamicin and 25 µg/mL kanamycin (half of standard concentrations). Cultures were grown and induced as described in chapter 3. About 2 mg of the truncated protein was purified from 1 L of culture; however, when separated on a Superdex 200 column, the dominant peak was near the void volume and no peak consistent with an octamer was observed.

REFERENCES

REFERENCES

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

(2) Kuipers, O. P., de Ruyter, P. G. G. A., Kleerebezem, M., and de Vos, W. M. (1998) Quorum sensing-controlled gene expression in lactic acid bacteria, *J. Biotechnol. 64*, 15-21.

(3) 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. U. S. A. 114*, 9074-9079.

(4) Desguin, B., Goffin, P., Bakouche, N., Diman, A., Viaene, E., Dandoy, D., Fontaine, L., Hallet, B., and Hols, P. (2015) Enantioselective regulation of lactate racemization by LarR in *Lactobacillus plantarum*, *J. Bacteriol. 197*, 219-230.

(5) Jacobus, A. P., and Gross, J. (2015) Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli*, *PLoS One 10*, e0119221.

(6) Walser, O. (2017) Personal communication.

CHAPTER 6

Conclusions and future studies

Conclusions from my studies

In this thesis, I have reported new discoveries concerning the nickel-pincer nucleotide (NPN), including mechanistic studies of an NPN-containing enzyme and analyses of two of the three proteins responsible for NPN synthesis. Lactate racemase (LarA), in which NPN was first discovered, inverts lactate chirality by transient oxidation of this substrate to pyruvate via proton-coupled hydride transfer to the coenzyme (Chapter 2). The now structurally-characterized LarB is responsible for the first step in the NPN synthesis pathway from nicotinic acid adenine dinucleotide (NaAD), catalyzing both the carboxylation of the nicotinic acid moiety and the hydrolysis of the phosphoanhydride (Chapter 3). In the next step (Chapter 4), LarE serves as a substrate, rather than as a catalyst, for insertion of sulfur into an NPN precursor, although my studies demonstrating *in vitro* recycling of the protein suggests that the *in vivo* catalytic potential is yet to be fully explored. Below, I expand on many potential avenues available for future studies.

NPN-utilizing enzymes

My work, along with that of my colleagues, has conclusively demonstrated that the lactate racemase from *Lactobacillus plantarum* forms a pyruvate intermediate during the racemization of lactate. My biochemical/biophysical characterizations of LarA (Chapter 2), along with computational studies by other groups, [1, 2] strongly suggest a proton-coupled hydride transfer mechanism couples the transient oxidation of lactate to the reduction of the NPN. A remaining gap in knowledge, however, concerns the precise chemical nature of the reduced coenzyme. Evidence for hydride transfer to C4 of NPN is based on the reduced cofactor UV absorption bands being comparable to those observed upon NAD^+ reduction to NADH and on computational analyses. My coworkers and I have also speculated that the hydride may bind to a

second site on the coenzyme; i.e., a Ni-hydride bond may form. Future work on the LarA mechanism should focus on demonstrating the chemical changes of NPN that occur during the enzyme reaction. Mass spectroscopy (MS) may be able to demonstrate the mass shift consistent with hydrogen transfer. The high abundance of pyruvate found in quenched reactions implies that a large fraction of the NPN may be in the reduced state at any point in time. This might enable the use of ^{1}H and ^{13}C-nuclear magnetic resonance to pinpoint which atom forms a bond with the incoming hydride (e.g., the pyridine ring positions and Ni-hydrides have distinct chemical shifts) and whether aromaticity of the cofactor is broken. Other potential studies could include further efforts to obtain lactate/pyruvate-bound crystal structures to better define substrate/intermediate binding and, perhaps, to view the reduced coenzyme state.

LarA homologs have been discovered that catalyze the chiral inversion of other α-hydroxy acids,[3] but the substrates for many of these homologs and other potential NPN-utilizing proteins have yet to be characterized. Additional studies are needed to identify the substrates of these enzymes by using knowledge gained from adjacent genes, metabolomics approaches, and other approaches, as well as high-throughput screening with a collection of α-hydroxy acids. There remains the outstanding question of what role NPN may play in organisms without an identified LarA homolog, a substantial fraction of organisms with NPN synthetic genes.[4] NPN has, so far, not been identified as a necessary cofactor in any enzyme not predicted to be a LarA homolog. Biology may potentially utilize this cofactor for chemistry that is unlike the LarA-catalyzed inversion of αC chirality. Investigation of non-LarA proteins that bind NPN is an important future research topic.

LarB catalyzes two different but coupled reactions

My efforts (Chapter 3) have demonstrated that the NaAD carboxylase/hydrolase (LarB) is a CO_2-utilizing carboxylase which requires no other substrates or enzymes to produce its product, pyridinium 3,5-biscarboxylic acid mononucleotide (P2CMN). While there are carboxylases that do not strictly need an energy source or partner enzyme, they need exceptionally high concentrations of CO_2 or bicarbonate to drive the reaction forward.[5] Neither class of PurEs, which synthesize carboxyaminoimidazole ribonucleotide (CAIR) and share homology with LarB, directly use a high energy molecule; their pathways drive the reaction forward by other means. The class II PurEs catalyze the reversible carboxylation of aminoimidazole ribonucleotide (AIR) to CAIR without the use of a high energy molecule; the flux through this enzyme is driven by product consumption by the ATP-utilizing PurC.[6] LarB is functional as a purified enzyme, so it cannot be driven forward by the subsequent enzyme in the pathway, LarE, for which P2CMN is a substrate. Class I PurEs do not catalyze a carboxylation, but rather they promote an isomerization reaction *via* a CO_2 intermediate. A different carboxylase PurK, not found in the class II pathway, utilizes ATP to carboxylate AIR at the N5 position to form N^5-CAIR, the substrate for the class I PurEs.[6] Nevertheless the class I PurE reaction is still readily reversible.[7]

The most obvious proposal for how LarB drives the NaAD carboxylation reaction lies in the enzyme's second activity, hydrolysis of the substrate phosphoanhydride high-energy bond, a reaction that is distinct from that catalyzed by PurE. Previous studies of LarA activation[8] and my results using a ^{14}C-based assay showed that nicotinic acid mononucleotide (NaMN), which lacks the AMP group and the hydrolysable phosphoanhydride, is not carboxylated by LarB to form P2CMN, suggesting that LarB's hydrolysis activity is essential to its carboxylation activity. The

reverse is not true: LarB readily hydrolyzes NaAD to AMP and NaMN without carboxylating the latter product. LarB also hydrolyzes NAD^+, but this hydrolysis cannot be coupled to the carboxylation of a separate molecule of NaMN. Additional evidence for interaction between the two reactions includes the lack of any carboxylated NaAD detected by either capillary electrophoresis or MS.

The question of how this coupling occurs is still outstanding. In parallel to the formation of carboxyphosphate by biotin-dependent enzymes,[9] one could imagine bicarbonate attacking the phosphoanhydride to yield an activated carboxy-AMP or carboxy-pyridinium-3-carboxylic acid mononucleotide that then participates in ring carboxylation; however, LarB uses CO_2 rather than bicarbonate as its substrate. Furthermore, the hydrolysis rate is independent of CO_2/HCO_3^- concentrations.[8] Another possibility is that hydrolysis may facilitate product release; i.e., the hydrolysis results in a conformational change providing the driving force for carboxylation. The simplest explanation, however, is that hydrolysis lowers the affinity of LarB for the product, preventing LarB from catalyzing the decarboxylation of P2CMN.

Ultimately, answering these questions may require more studies be carried out on the N-terminal region. The first 45 or so residues in our LarB structure remain unstructured, and a few of these residues appear to be important for catalysis based on my mutagenesis studies. I have attempted, without success, to produce a soluble truncated protein. Future studies can refine this approach to find the minimum necessary N-terminal truncation, and, with any luck, produce a crystal with a more stable N-terminus and visible electron density.

Further studies also should test the hypothesis that thiolate addition to C4 of the nicotinic acid activates the substrate for carboxylation. As I reported here, a cysteinyl sulfur forms a covalent linkage to NAD^+ in the nicotinamide-bound crystal structure of LarB; however,

additional evidence for such a species with the authentic substrate is needed. Future work can focus on confirming this proposed intermediate by demonstrating the presence of a covalent linkage using MS or NMR, or by chemically trapping the species. Techniques such as stop-flow UV-visible spectroscopy may also be useful to characterize this intermediate, including the kinetics of its formation.

LarE, a sacrificial sulfur insertase?

The enzyme responsible for thioacid formation on the nascent NPN cofactor, LarE, donates a cysteinyl sulfur atom as the sulfur source, generating dehydroalanine (Chapter 4), in one of the few known examples of a sacrificial "enzyme." [8, 10] Such a process, synthesizing and degrading two 31-kDa proteins for the creation of a single cofactor molecule, represents an extreme energy investment by the organism. Importantly, however, all of the studies on this enzyme have been conducted either *in vitro* or under overexpression conditions in heterologous systems. Thus, the single-turnover use of LarE may not accurately reflect the reality within the native *L. plantarum*, where *in vivo* systems may exist that can replenish the cysteinyl sulfur in LarE when cellular LarE content is in biologically relevant concentrations. Analysis of the LarE reaction in the native microorganism is an important direction for future studies.

The first report of LarE as a sacrificial protein also reported another unexpected finding: LarE has high affinity for CoA, readily forming a mixed disulfide at the reactive cysteine.[8] In fact, the CoA bound form of LarE was a major species found when the enzyme was purified without the addition of exogenous CoA. Furthermore, the authors found that the LarE-CoA mixed disulfide species was active *in vitro* when provided with exogenous CoASH. The high specific affinity of the enzyme for CoA, as opposed to other available intracellular thiols such as

cysteine, suggests that this binding is not adventitious. The authors suggested that CoA binding may help protect the reactive thiol.

Alternatively, CoA could be an intermediary in a LarE recycling system. I reported herein (Chapter 4) that CoA persulfide robustly reacts with the dehydroalanine to form the LarE-CoA disulfide, a species already known to be an active form of the protein. I have not, however, explored whether this reaction is plausible *in vivo*. CoA persulfide can be formed by the non-enzymatic reaction between CoA disulfide and free hydrosulfide, providing an avenue for its formation *in vivo*.[11] I have not found published references for the concentration of organic persulfides in *L. plantarum* or closely related organisms. In *Staphylococcus aureus*, CoA persulfide was found to make up about 2% of the total CoA pool, or approximately 20 μM of the persulfide.[12] This figure rose to over 10% of the CoA pool upon exposure to 0.2 mM sulfide. Similar concentrations were found in *Enterococcus faecalis*.[13] If similar concentrations of CoA persulfide exist in *L. plantarum*, it seems reasonably plausible that CoA-mediated LarE recycling is a viable strategy for the organism. Alternatively, *L. plantarum* may possess a cysteine desulfurase (IscS) that generates the persulfide of CoA. Although IscS-mediated formation of persulfides is a common mode of sulfur mobilization in biology,[14] I have demonstrated *in vitro* that *Escherichia coli* IscS is unable to regenerate LarE in the presence of cysteine.

Finally, as I have reported here, homology models of some LarE homologs contain three cysteine residues at the active site region and ideally positioned to form a [4Fe-4S] cluster. LarE is also structurally related to other sulfur transferases that utilize [4Fe-4S] clusters or form persulfides at the active site. My preliminary work shows that the LarE homolog from *Thermotoga maritima* has a UV-vis spectrum that is consistent with the presence of an iron

sulfur cluster, but this result does not speak to what role the cluster could be playing. Further efforts to characterize these forms of the LarE enzyme are needed.

Concluding remarks

In the six years since the discovery of NPN was first published,[15] headway has been made in understanding its chemistry related to lactate racemase and its broader role in nature for racemase and epimerase reactions. NPN and each of its biosynthetic enzymes have unique aspects and properties—many of which have since been discovered, but many questions about the synthesis and use of this coenzyme remain unanswered. My hope is that the work I have performed and reported here will further our understanding of how this incredible cofactor functions and how this unique coenzyme is made.

REFERENCES

REFERENCES

(1) Zhang, X., and Chung, L. W. (2017) Alternative mechanistic strategy for enzyme catalysis in a Ni-dependent lactate racemase (LarA): Intermediate destabilization by the cofactor, *Chem. Eur. j 23*, 3623-3630.

(2) Qiu, B., and Yang, X. (2017) A bio-inspired design and computational prediction of scorpion-like SCS nickel pincer complexes for lactate racemization, *Chem. Commun. (Camb.) 53*, 11410-11413.

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

(5) Schada von Borzyskowski, L., Rosenthal, R. G., and Erb, T. J. (2013) Evolutionary history and biotechnological future of carboxylases, *J. Biotechnol. 168*, 243-251.

(6) Zhang, Y., Morar, M., and Ealick, S. E. (2008) Structural biology of the purine biosynthetic pathway, *Cell. Mol. Life Sci. 65*, 3699-3724.

(7) Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, II, Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E., and Stubbe, J. (2007) N^5-CAIR mutase: role of a CO_2 binding site and substrate movement in catalysis, *Biochemistry 46*, 2842-2855.

(8) 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. U. S. A. 113*, 5598-5603.

(9) Erb, T. J. (2011) Carboxylases in Natural and Synthetic Microbial Pathways, *Appl. Environ. Microbiol. 77*, 8466-8477.

(10) Chatterjee, A., Abeydeera, N. D., Bale, S., Pai, P. J., Dorrestein, P. C., Russell, D. H., Ealick, S. E., and Begley, T. P. (2011) *Saccharomyces cerevisiae* THI4p is a suicide thiamine thiazole synthase, *Nature 478*, 542-U146.

(11) Kabil, O., and Banerjee, R. (2012) Characterization of patient mutations in human persulfide dioxygenase (ETHE1) involved in H_2S catabolism, *J. Biol. Chem. 287*, 44561-44567.

(12) Peng, H., Zhang, Y., Palmer, L. D., Kehl-Fie, T. E., Skaar, E. P., Trinidad, J. C., and Giedroc, D. P. (2017) Hydrogen sulfide and reactive sulfur species impact proteome S-

sulfhydration and global virulence regulation in *Staphylococcus aureus*, *ACS Infect Dis 3*, 744-755.

(13) Shen, J., Walsh, B. J. C., Flores-Mireles, A. L., Peng, H., Zhang, Y., Zhang, Y., Trinidad, J. C., Hultgren, S. J., and Giedroc, D. P. (2018) Hydrogen sulfide sensing through reactive sulfur species (RSS) and nitroxyl (HNO) in *Enterococcus faecalis*, *ACS Chem. Biol.*

(14) Black, K. A., and Dos Santos, P. C. (2015) Shared-intermediates in the biosynthesis of thio-cofactors: Mechanism and functions of cysteine desulfurases and sulfur acceptors, *Biochim. Biophys. Acta 1853*, 1470-1480.

(15) 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.