

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

Joel A. Rankin^a, Shramana Chatterjee^b, Zia Tariq^c, Satyanarayana Lagishetty^a, Benoît Desguin^c, Jian Hu^{a,d,1}, and Robert P. Hausinger^{a,b,1}

^aDepartment of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824; ^bDepartment of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824; ^cInstitute of Biomolecular Sciences and Technology, Université Catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium; and ^dDepartment of Chemistry, Michigan State University, East Lansing, MI 48824

Edited by Douglas C. Rees, California Institute of Technology, HHMI, Pasadena, CA, and approved August 23, 2021 (received for review March 31, 2021)

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, we 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 C5-carboxylated-NaAD. Structural investigations show that LarB has an N-terminal domain of unique fold and a C-terminal domain homologous to aminoimidazole ribonucleotide carboxylase/mutase (PurE). Like PurE, LarB is octameric with four active sites located at subunit interfaces. The complex of LarB with NAD^+ , an analog of NaAD, reveals the formation of a covalent adduct between the active site Cys221 and C4 of NAD^+ , resulting in a boat-shaped dearomatized pyridine ring. The 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 additional residues potentially important for substrate binding and catalysis. In combination with these findings, the results from structure-guided mutagenesis studies lead us to propose enzymatic mechanisms for both the carboxylation and hydrolysis reactions of LarB that are distinct from that of PurE.

nickel-pincer cofactor | carboxylation | hydrolysis | covalent adduct

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 (Fig. 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 cytidine triphosphate-dependent nickel insertase or cyclometallase (5). Homologs of the genes encoding these biosynthetic proteins are widespread within eubacteria and archaea (6).

LarB, the subject of this work, acts on its substrate NaAD by carboxylating the C5 position of the nicotinic acid moiety and hydrolyzing the phosphoanhydride bond, releasing AMP and forming the product pyridinium 3,5-biscarboxylic acid mononucleotide (P2CMN) (Fig. 1B). In addition, the enzyme catalyzes the nonproductive hydrolysis of NaAD to AMP and the non-carboxylated product nicotinic acid mononucleotide (NaMN) in what appears to 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 adenosine triphosphate (ATP)-dependent aminoimidazole ribonucleotide (AIR) carboxylase (PurK) first synthesizes N^5 -CAIR, and then PurE transfers the carboxyl group from the carbamate to the C4 position, forming CAIR (Fig. 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 (SI Appendix, Fig. S1A). 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, we show that the substrate used for LarB carboxylation is carbon dioxide, not bicarbonate, and we 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. We demonstrate that the LarB carboxylation active

Significance

Carboxylation of nonactivated carbon atoms in aromatic rings is of considerable synthetic importance, but these reactions typically require harsh conditions or highly reactive reagents. In this study, the mechanism of pyridinium ring carboxylation is established for LarB, an enzyme that is essential for biosynthesis of the recently discovered nickel-pincer nucleotide cofactor found in a diverse family of racemases or epimerases. Structural and mutagenesis methods were used to define a new paradigm for carboxylation of a pyridinium ring in which a cysteinyl-pyridine ring intermediate enhances the nucleophilicity of the adjacent carbon atom for reaction with carbon dioxide.

Author contributions: J.A.R., B.D., J.H., and R.P.H. designed research; J.A.R., Z.T., and B.D. performed research; J.A.R., S.C., Z.T., S.L., B.D., J.H., and R.P.H. analyzed data; and J.A.R., S.C., B.D., J.H., and R.P.H. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence may be addressed. Email: hausinger@msu.edu or hujian1@msu.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2106202118/-DCSupplemental>.

Published September 21, 2021.

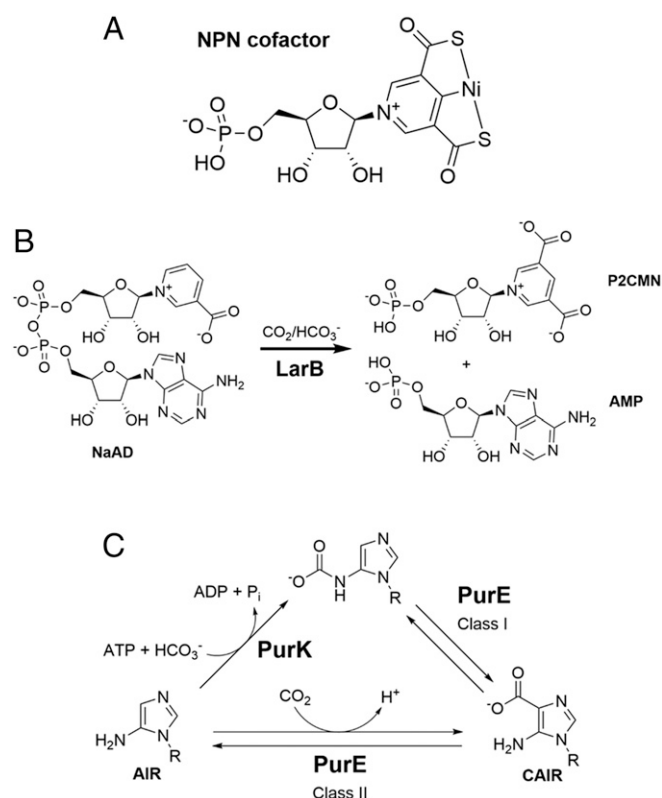


Fig. 1. NPN cofactor, LarB reaction, and comparison to the reactions of PurE. (A) The NPN cofactor possesses nickel bound to a carbon and two sulfur atoms of a modified pyridinium mononucleotide. (B) LarB catalyzes C5 ring carboxylation and phosphoanhydride hydrolysis of NaAD to release AMP and generate the NPN precursor P2CMN. (C) The carboxylation of AIR to form 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.

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 adduct forms between NAD^+ and the invariable Cys221 of LarB sequences; a similar intermediate generated with NaAD would support its carboxylation. We show that 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. We 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, we propose mechanisms for the carboxylation and hydrolysis reactions catalyzed by LarB that are fundamentally different from the chemistry used by PurE.

Results

CO_2 , Not Bicarbonate, Is the Carboxylation Substrate of LarB. *L. plantarum* LarB was produced in either *Escherichia coli* or *Lactococcus lactis* as the C-terminal Strep II-tagged protein and purified by Strep-Tactin chromatography as previously described (2). The general properties of the enzyme are described in *SI Appendix, Supplementary Text*, and we only note here that, given sufficient NaAD, bicarbonate, and MgCl_2 , LarB produces P2CMN and NaMN (the nonproductive NaAD hydrolysis product) in a linear manner for more than 100 min (*SI Appendix, Fig. S2*). We investigated whether CO_2 or bicarbonate is the substrate of LarB by using an isotope dilution experiment (11). The enzyme was

added to a solution containing NaAD, MgCl_2 , and ^{14}C -labeled bicarbonate/ CO_2 in buffer at pH 7.5, where the equilibrium greatly favors bicarbonate, and after the reaction had proceeded for 65 s, the sample was vigorously mixed with gaseous nonlabeled CO_2 and the assay continued. The incorporation of ^{14}C into product was briefly paused upon pulsing with CO_2 and then continued at a reduced rate due to equilibration of the added gas with the solution, demonstrating that CO_2 is the substrate (Fig. 2). In contrast, when carbonic anhydrase was included in the assay solution, the mixing with CO_2 led to an immediate reduction in the rate of ^{14}C incorporation into product, exhibiting the same final rate as in the first experiment, indicating rapid conversion of CO_2 to bicarbonate and overall dilution of the inorganic carbon pool. The transient cessation of isotope incorporation for multiple variations of the experiment lacking carbonic anhydrase 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 electrospray ionization mass spectrometry. This approach allowed us to identify ^{18}O -containing AMP (Fig. 3), demonstrating the reaction occurs by nucleophilic

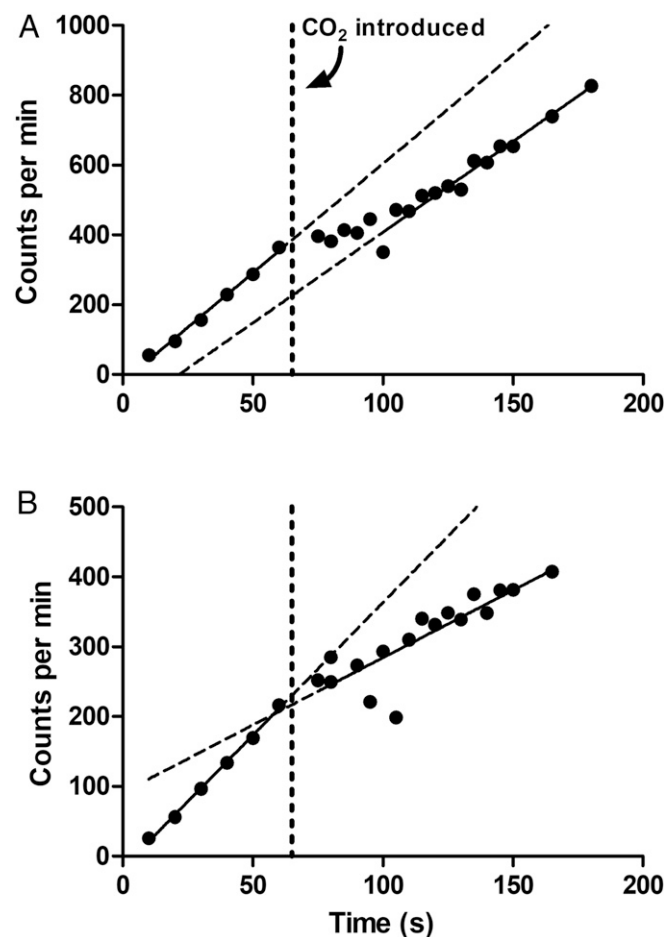


Fig. 2. Carbon dioxide, not bicarbonate, is the substrate of LarB. (A) Short-term incorporation of ^{14}C into substrate using radiolabeled CO_2 /bicarbonate, with unlabeled gaseous CO_2 added using vigorous mixing at 65 s for isotope dilution. (B) The same experiment in the presence of carbonic anhydrase.

water attack on the AMP-associated phosphate group. An experiment in which ^{18}O -labeled bicarbonate was added to the assay failed to produce detectable levels of ^{18}O -AMP, arguing against bicarbonate activation via reaction with the phosphoanhydride.

Structure of LarB. *L. plantarum* LarB isolated from *E. coli* was crystallized and structurally characterized. The phase was solved by single-wavelength anomalous dispersion using a selenomethionine-substituted crystal, and the initial structural model was later used to solve the crystal structure of the native protein at a resolution of 2.8 Å via molecular replacement (*SI Appendix, Table S1*). The asymmetric unit contains six LarB molecules organized into three dimers (chains A and B, chains C and D, and chains E and F), with an overall rmsd of 0.48 Å between the two monomers in each dimer and with cross-dimer rmsd values of 0.25 Å (A:B dimer versus C:D dimer), 0.59 Å (A:B dimer versus E:F dimer), and 0.66 Å (C:D dimer versus 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 fourfold axis running through the central tunnel (two views are shown in Fig. 4*A* and *B*), which is in agreement with the size exclusion chromatography-multiangle light scattering results for LarB in solution (*SI Appendix, Fig. S3*). 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-twofold symmetry when considering the monomers as equivalent components. Compared with the other two dimers, the E:F dimer exhibits much poorer electron density, indicative of a 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*, Protein Data Bank [PDB] access code 5CLI; Fig. 4*D* and *E*), but the PurE octamer has less volume because of the lack of the N-terminal domain which is conserved in LarB proteins (*SI Appendix, Fig. S1*).

Each LarB monomer contains two domains (Fig. 4*C*), designated as the N- and C-terminal domains, with a long cleft between them. The larger C-terminal domain consists of five parallel β -sheets with two α -helices on one face and three on the other, representing the same fold as found in PurE (Fig. 4*F*) (7, 9, 10) with an rmsd for C α atoms of 1.6 Å despite the low ($\sim 25\%$) 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 intermonomer 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 (12) found no compelling structural matches with the N-terminal domain of LarB; however, this α/β fold is similar to the d.68 IF3-like fold of the SCOPE database (13). The initial 45 residues of the N terminus are disordered in most chains, and noninterpretable density was present. This region is not well conserved in LarB homologs.

The active site of LarB was tentatively identified by comparison to that of PurE (Fig. 4*C* and *F* and *SI Appendix, Fig. S4 A* and *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 bidentate carboxylate of Glu180 from one monomer, the thiolate of Cys221 from another monomer, and two water molecules (*SI Appendix, Fig. S5*). Residues corresponding to Glu180 and Cys221 are invariable in LarB sequences but are not present at the active site of PurE proteins (*SI Appendix, Fig. S1*). Most other residues at the putative LarB active site are hydrophobic and unlikely to be catalytically relevant. LarB activity requires magnesium (2), although Mn, Co, or Ni can substitute; the extra density was inconsistent with magnesium. The identity of the metal was established as Zn based on anomalous scattering of this density in the crystals (*SI Appendix, Fig. S6*). Furthermore, we demonstrated the presence of substoichiometric levels of Zn in purified LarB by using inductively coupled plasma-mass spectrometry and inductively coupled plasma-optical emission spectroscopy. Subsequent experiments revealed that activity was not lost by ethylenediaminetetraacetic acid (EDTA) treatment and that Zn was an inhibitor of the enzyme (*SI Appendix, Supplementary Text*). The observed Zn binding to LarB was consistent with a crystallization artifact where Zn ions were introduced into the protein sample during purification and/or the formation of crystals.

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 (at 3.0 Å) (*SI Appendix, Table S1*). In both cases, the structures were acquired by soaking the ligands into preformed 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 fourfold 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 (*SI Appendix, Fig. S4*); however, only chains E/F and C/D bind NAD^+ (Fig. 5*A* and *SI Appendix, Fig. S7*). The nicotinamide moiety appears more dynamic, with a higher *B*-factor and less density fit compared to the rest of the ligand. Due to its 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 an 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^+ (Fig. 5*A* and *SI Appendix, Fig. S7*) are conserved in LarB homologs but generally not in PurE sequences (*SI Appendix, Fig. S1*). The adenine-binding $\beta 8$ - $\alpha 6$ 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,

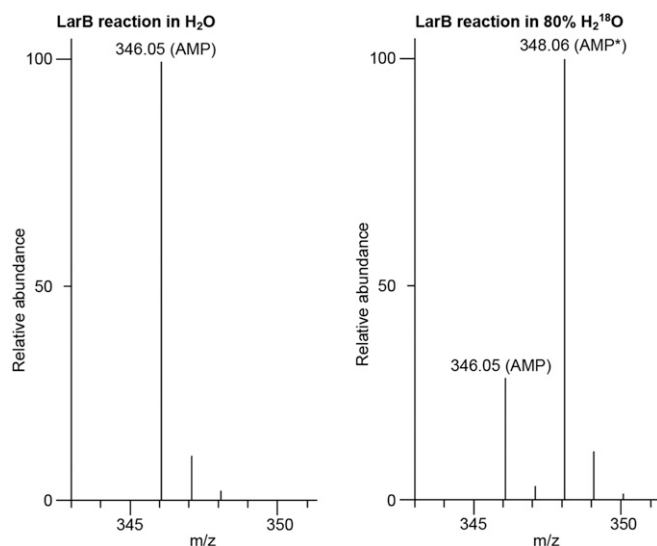


Fig. 3. LarB incorporates ^{18}O from labeled solvent into AMP during NaAD hydrolysis. The figure depicts the mass spectra of LarB-generated AMP derived from NaAD in unlabeled water (*Left*) and ^{18}O -labeled water (*Right*).

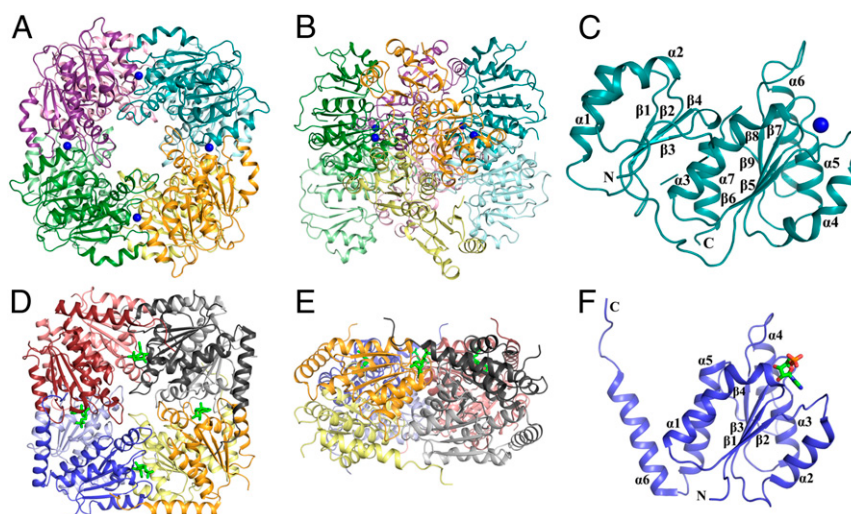


Fig. 4. Structure of *L. plantarum* LarB and comparison to PurE. (A and B) Two views of the octameric LarB structure, with each subunit depicted in cartoon mode using different colors (or different shades for the dimer substructures). The blue spheres represent the metal ion associated with the as-isolated enzyme. (C) LarB monomer. (D and E) Two views of the octameric PurE from *Acetobacter aceti* with bound AIR (green sticks, PDB 5CLI). (F) PurE monomer with bound AIR.

Ser127, and Asp151, whose conformation is fixed by a salt bridge to Arg159. Accompanying the phosphate groups of NAD^+ is an 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 (SI Appendix, Figs. S4 and S8), 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 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 (Fig. 5B), with this residue pointing toward NAD^+ for chain F and away from the substrate analog for chain C. Indeed, the electron density map of

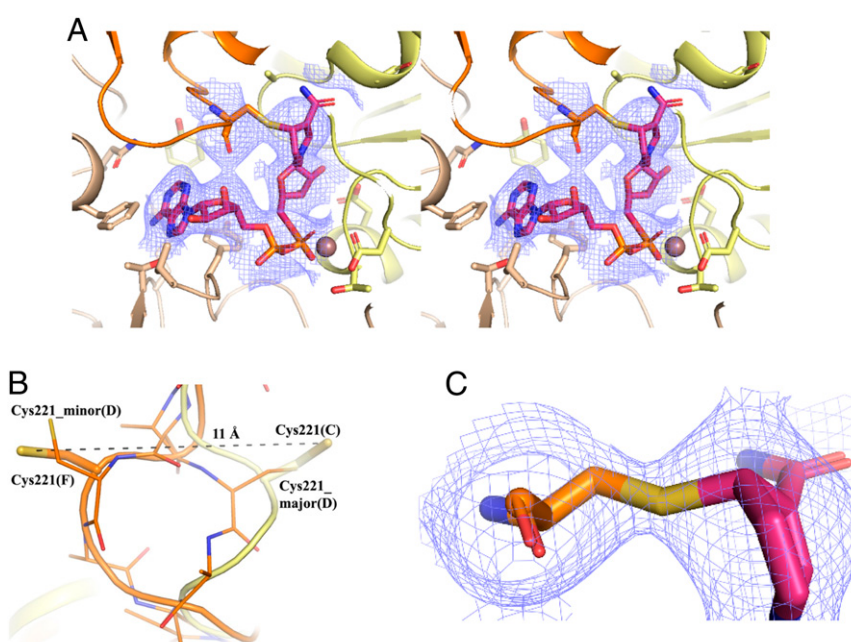


Fig. 5. The NAD^+ binding site of LarB and evidence for a covalent linkage between C4 of the pyridinium ring and Cys221. (A) Stereoview of NAD^+ bound to LarB at the same location as Zn and analogous to the AIR binding site in PurE. The surrounding region is depicted using cartoon view with different colors for the three subunits (chain E is light yellow and wheat; chain F is orange), Mg^{2+} is shown as a sphere, and NAD^+ is shown as magenta sticks with the $\text{F}_o - \text{F}_c$ omit map ($\sigma = 3$). (B) Dual conformation of Cys221 in the structure of the LarB- NAD^+ complex compared to PurE. The LarB region containing Cys221 (residues 219 to 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 [major] and 0.3 [minor] occupancies). Interchange of the two conformations requires ~ 11 Å flipping of this loop, associated with binding of the substrate analog. (C) Zoomed-in view of the covalent linkage between Cys221 and the pyridine ring derived from NAD^+ with $\text{F}_o - \text{F}_c$ omit map ($\sigma = 3$).

the region containing Cys221 (residues 219 to 223) of chain D clearly shows an equilibrium of dual conformations—a minor conformation with Cys221 pointing toward NAD⁺ and a major conformation with the 219 to 223 segments folding as a part of $\alpha 6$ and Cys221 facing away from the active site (Fig. 5B). The ~ 11 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 (*SI Appendix*, Fig. S9). Of greatest interest is the situation for chain F where the electron density profile supports a linkage between the Sy of Cys221 and C4 of the pyridine ring (Fig. 5C). The S-C distance of ~ 1.8 Å and the boat-like configuration of the pyridine ring required for better fitting the electron density are both consistent with a covalent bond. The 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 (*SI Appendix*, Fig. S10A). In addition, the LarB–AMP complex includes an 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⁺ (*SI Appendix*, Fig. S10B). Unlike the NAD⁺-bound LarB structure, where Glu180 is disordered, this residue exhibited continuous electron density in the AMP-bound chain.

LarB Variant Analyses. We targeted for substitution several conserved residues at the Zn- and ligand-binding sites of LarB along with residues in the N-terminal region and analyzed the variant proteins by monitoring the extent of ¹⁴C incorporation after 10-min assays (Fig. 6). When Cys221, the residue that coordinates Zn and covalently attaches to the pyridinium ring of NAD⁺, was substituted with either serine or alanine, the carboxylation activity was abolished. We substituted Glu180, the other Zn-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 $\sim 5\%$ 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 Fig. 6). 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 $\sim 50\%$ loss of carboxylation activity (with the putative C5-carboxylated NaAD subsequently hydrolyzed to AMP and P2CMN) but with greater than 90% loss of the nonproductive hydrolysis activity (*SI Appendix*, Fig. S2). Compared to the carboxylation kinetics of the wild-type enzyme (*SI Appendix*, Table S2), the S127A variant exhibited decreases in

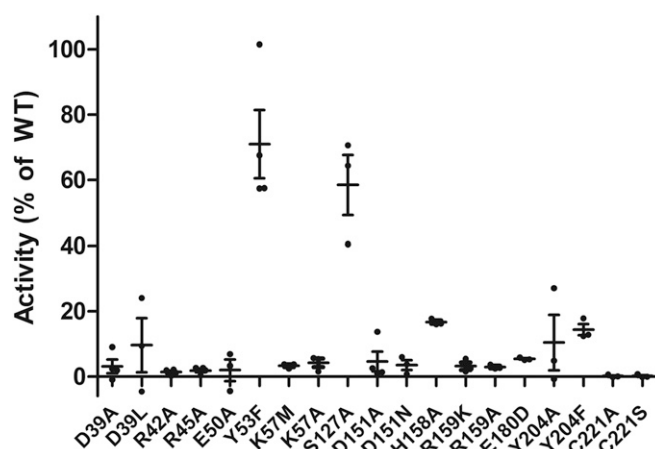


Fig. 6. The activity of each variant as measured using the 10-min ¹⁴C incorporation assay is plotted as the percent of wild-type LarB activity. The lines represent the mean and the error bars the SEM (*n* = 3 or more).

both k_{cat} (0.0089 s^{-1}) and K_m for NaAD ($3 \pm 2 \mu\text{M}$). A serine is also found at this position in PurE (*SI Appendix*, Fig. S1).

In addition to carrying out the above short-term ¹⁴C incorporation assays, we assessed the amount of label incorporated into NaAD by selected LarB variants after overnight incubation. When LarB was incubated overnight with NaAD and ¹⁴C-labeled bicarbonate/CO₂, the incorporated radiolabel accounted for $\sim 20\%$ of the NaAD concentration (*SI Appendix*, Fig. S11). We attribute this low extent of labeling to the significant amount of nonproductive enzymatic hydrolysis of NaAD to yield NaMN (*SI Appendix*, Fig. S2), which is not a substrate of LarB. The ratio of ¹⁴C incorporation into substrate by the E180D variant and wild-type enzyme after overnight incubation ($\sim 1\%$ versus 20% , *SI Appendix*, Fig. S11) was similar to the ratio of incorporation after 10 min ($\sim 5\%$ versus 100% , respectively, Fig. 6), perhaps indicating a nearly intact hydrolysis activity. Of keen interest, the amount of ¹⁴C incorporated into NaAD after extended incubation with the S127A variant corresponded to a much greater portion of initial substrate ($\sim 66\%$); this result dramatically differs from the plateau of around 20% of the substrate observed in the wild-type enzyme (*SI Appendix*, Fig. S11). We attribute these results to the substitution of Ser127 by alanine having a modest effect on carboxylation activity (likely associated with a small decrease in the rate of phosphoanhydride hydrolysis for C5-carboxy-NaAD versus the significant decrease in the nonproductive hydrolysis rate of NaAD for this variant seen in *SI Appendix*, Fig. S2), which may be due to an altered orientation of the pyrophosphate group caused by loss of the Ser127–phosphate interaction.

Discussion

Our structural and biochemical studies of LarB have provided critical insights into the enzyme architecture and mechanism. Although Zn is coordinated to the protein as-isolated, this metal ion is not required for activity; rather, it inhibits the carboxylation of NaAD. Indeed, it is difficult to see how NaAD could bind to the active site and be activated for carboxylation if Cys221 participates in binding Zn, as illustrated by the superposition of these two states (*SI Appendix*, Fig. S12). Using the LarB–NAD⁺ structure, we have modeled the binding of the authentic substrate into the protein (*SI Appendix*, Fig. S13). We suggest the pyridinium ring of NaAD is flipped by 180° compared to NAD⁺, with the C3 carboxyl group coordinating the Mg²⁺ (along with the 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

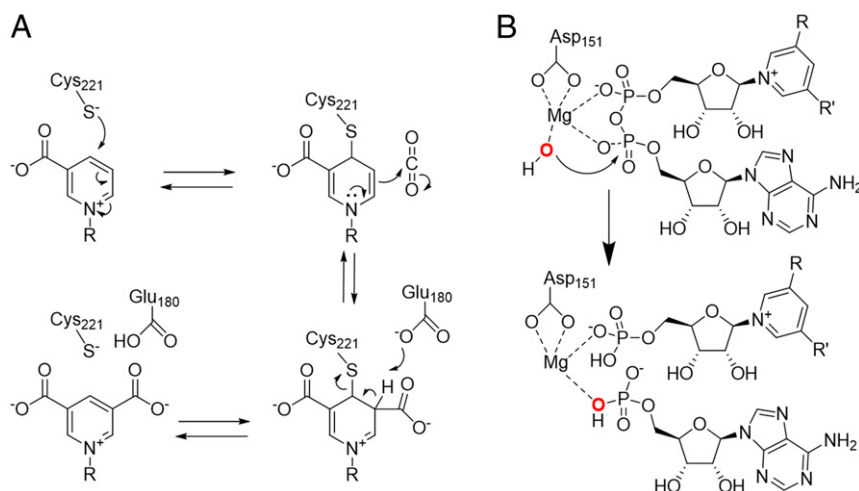


Fig. 7. Proposed mechanisms for the LarB-dependent carboxylation and hydrolysis reactions. (A) Carboxylation. Cys221 adds to C4 of the NaAD pyridinium ring, thus enhancing the nucleophilicity of C5, which attacks carbon dioxide. Glu180 functions as a general base to abstract the C5 proton, leading to rearomatization 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 nonproductive hydrolysis also occurs with the substrate NaAD for which R = COO[−] and R' = H. The hydrolysis of C5-carboxy-NaAD is proposed to facilitate product release.

confers specificity on the enzyme so as to allow for C5 carboxylation, whereas NAD⁺ is not carboxylated. The model includes an Mg²⁺-bound water molecule that could be further activated for hydrolysis by coordination of the NaAD carboxylate to the metal. CO₂ is shown at the position of the amide group in the NAD⁺-bound structure but without attempting to delineate specific interactions with the protein. We note the structural studies of PurE indicate that it binds CO₂ using backbone amides (9), whereas other CO₂-binding modes occur in 2-ketopropyl coenzyme M oxidoreductase/carboxylase (14), enoyl-CoA carboxylase/reductase (15), phosphoenolpyruvate carboxykinase (16), and other proteins. In the LarB model, Cys221 is proximate to C4 of the pyridinium ring and ready to facilitate catalysis.

The carboxylation reaction of LarB utilizes CO₂, not bicarbonate, analogous to what is found in the related enzyme, class II PurE (9). Whereas PurE likely accomplishes the carboxylation reaction by direct nucleophilic attack of the aminoimidazole carbon atom on CO₂ (9, 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 CO₂ (Fig. 7A). 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 CO₂ is a commonly observed mechanism of carboxylation for carboxylases that do not utilize the biotin cofactor (17), 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. The transient formation of a cysteinyl adduct to facilitate electrophilic addition to a heterocycle, as used in the LarB carboxylation mechanism, is an established strategy in enzymology. For example, thymidylate synthase (18) and DNA (cytosine-5)-methyltransferase (19) both add cysteine residues to nucleotide bases and thus increase the nucleophilicity of a ring carbon atom. Nevertheless, to our knowledge, LarB is the only example using this mechanism for carboxylation.

As shown in the model of LarB with bound NaAD (*SI Appendix*, Fig. S13), the hydrolytic reaction is likely to involve an Mg-bound solvent molecule that attacks the phosphate distal to the pyridinium ring, releasing P2CMN (Fig. 7B). The water is expected to be activated by the substrate carboxylate and by the 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 (*SI Appendix*, Fig. S11). The serine residue may function to position for hydrolysis the phosphoanhydride of NaAD and its dicarboxylated analog. We speculate that phosphoanhydride hydrolysis facilitates product dissociation from the enzyme. Our evidence does not provide support for an earlier proposal that energy associated with hydrolysis of the phosphoanhydride is directly coupled to carboxylation (2); rather, we propose that the two reactions are independent. Importantly, the low affinity of LarB for its product, P2CMN, prevents the enzyme from catalyzing the reverse reaction, as demonstrated by our observations of the LarB reaction (*SI Appendix*, *Supplementary Text*). In this way, phosphoanhydride hydrolysis indirectly drives the carboxylation reaction.

Materials and Methods

Details of reagents and experimental procedures for vector construction, mutagenesis, LarB purification, physical characterization, enzyme activity measurements, crystallization, and structure determination are described in *SI Appendix*. Also found there are general biophysical and kinetic properties of LarB.

Data Availability. The atomic coordinates have been deposited in PDB with accession codes 7MJ2, 7MJ1, and 7MJ0 for the as-isolated (Zn-bound) LarB, LarB-NAD⁺ complex, and LarB-AMP complex, respectively. All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Matthias Fellner for assistance during early structural efforts and the staff of the Mass Spectrometry and Metabolomics Core at Michigan State University for their help with mass spectrometry studies. These studies were supported by grants to J.H. and R.P.H. from the NSF (CHE-1807073) and the NIH (GM128959) and from grants to B.D. from the Fonds de la Recherche Scientifique. This research used resources of the Advanced Photon Source, a US 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 Life Sciences Collaborative Access Team Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

1. B. Desguin *et al.*, A tethered niacin-derived pincer complex with a nickel-carbon bond in lactate racemase. *Science* **349**, 66–69 (2015).
2. B. Desguin, P. Soumillion, P. Hols, R. P. Hausinger, 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 (2016).
3. M. Fellner, B. Desguin, R. P. Hausinger, J. Hu, 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 (2017).
4. M. Fellner, J. A. Rankin, B. Desguin, J. Hu, R. P. Hausinger, Analysis of the active site cysteine residue of the sacrificial sulfur insertase LarE from *Lactobacillus plantarum*. *Biochemistry* **57**, 5513–5523 (2018).
5. B. Desguin *et al.*, Biosynthesis of the nickel-pincer nucleotide cofactor of lactate racemase requires a CTP-dependent cyclometallase. *J. Biol. Chem.* **293**, 12303–12317 (2018).
6. B. Desguin *et al.*, Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system. *Nat. Commun.* **5**, 3615 (2014).
7. Y. Zhang, M. Morar, S. E. Ealick, Structural biology of the purine biosynthetic pathway. *Cell. Mol. Life Sci.* **65**, 3699–3724 (2008).
8. E. J. Mueller, E. Meyer, J. Rudolph, V. J. Davisson, J. Stubbe, N⁵-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 (1994).
9. A. A. Hoskins *et al.*, N5-CAIR mutase: Role of a CO₂ binding site and substrate movement in catalysis. *Biochemistry* **46**, 2842–2855 (2007).
10. I. I. Mathews, T. J. Kappock, J. Stubbe, S. E. Ealick, Crystal structure of *Escherichia coli* PurE, an unusual mutase in the purine biosynthetic pathway. *Structure* **7**, 1395–1406 (1999).
11. J. Coombs, S. L. Maw, C. W. Baldry, Metabolic regulation in C₄ photosynthesis: The inorganic carbon substrate for PEP carboxylase. *Plant Sci. Lett.* **4**, 97–102 (1975).
12. L. Holm, DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140 (2020).
13. N. K. Fox, S. E. Brenner, J.-M. Chandonia, SCOPe: Structural Classification of Proteins—extended, integrating SCOP and ASTRAL data and classification of new structures. *Nucleic Acids Res.* **42**, D304–D309 (2014).
14. A. S. Pandey, D. W. Mulder, S. A. Ensign, J. W. Peters, Structural basis for carbon dioxide binding by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *FEBS Lett.* **585**, 459–464 (2011).
15. G. M. M. Stoffel *et al.*, Four amino acids define the CO₂ binding pocket of enoyl-CoA carboxylases/reductases. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13964–13969 (2019).
16. J. J. H. Cotelesage *et al.*, How does an enzyme recognize CO₂? *Int. J. Biochem. Cell Biol.* **39**, 1204–1210 (2007).
17. L. Schada von Borzyskowski, R. G. Rosenthal, T. J. Erb, Evolutionary history and biotechnological future of carboxylases. *J. Biotechnol.* **168**, 243–251 (2013).
18. C. W. Carreras, D. V. Santi, The catalytic mechanism and structure of thymidylate synthase. *Annu. Rev. Biochem.* **64**, 721–762 (1995).
19. L. Chen *et al.*, Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry* **30**, 11018–11025 (1991).