ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

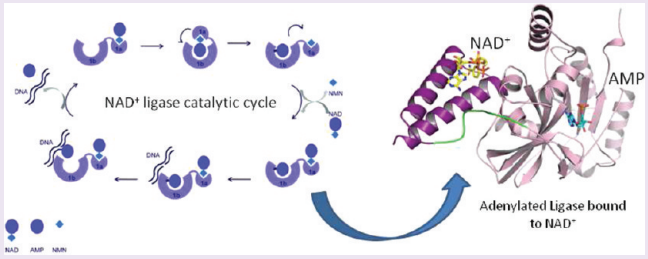
Structure Guided Understanding of NAD$^+$ Recognition in Bacterial DNA Ligases

Sushmita D. Lahiri,* Rong-Fang Gu, Ning Gao, Irene Karantzeni, Grant K. Walkup, and Scott D. Mills

Department of Bioscience, Infection Innovative Medicines Unit, AstraZeneca R&D Boston, Waltham, Massachusetts 02451, United States

Supporting Information

ABSTRACT: NAD$^+$-dependent DNA ligases (LigA) are essential bacterial enzymes that catalyze phosphodiester bond formation during DNA replication and repair processes. Phosphodiester bond formation proceeds through a 3-step reaction mechanism. In the first step, the LigA adenylation domain interacts with NAD$^+$ to form a covalent enzyme-AMP complex. Although it is well established that the specificity for binding of NAD$^+$ resides within the adenylation domain, the precise recognition elements for the initial binding event remain unclear. We report here the structure of the adenylation domain from *Haemophilus influenzae* LigA. This structure is a first snapshot of a LigA-AMP intermediate with NAD$^+$ bound to domain 1a in its open conformation. The binding affinities of NAD$^+$ for adenylated and nonadenylated forms of the *H. influenzae* LigA adenylation domain were similar. The combined crystallographic and NAD$^+$-binding data suggest that the initial recognition of NAD$^+$ is via the NMN binding region in domain 1a of LigA.



DNA ligases play an indispensable role in DNA replication, recombination, and repair processes in all organisms. DNA ligases catalyze the formation of phosphodiester bonds at single-stranded breaks in duplex DNA by employing a 3-step reaction mechanism. In the first step, a covalent enzyme-AMP intermediate is formed through a phosphoramidate bond to a conserved lysine side chain. In the second step, AMP is transferred from the active site lysine to the 5′-phosphate of DNA at the single-strand break. In the final step, formation of a phosphodiester bond seals the break, and AMP is liberated. DNA ligases form two distinct groups depending on the substrate used to adenylate the enzyme. Primarily, the eukaryotic, viral, and archael enzymes use ATP and the bacterial enzymes use NAD$^+$, although there are exceptions in both cases.[1] Distinct structural features in the two enzymes accompany this fundamental difference in substrate utilization.[2,3] The NAD$^+$-dependent DNA ligases are more elaborate, containing multiple domains that share minimal structural and sequence similarity with their eukaryotic counterpart. The differences in substrate specificity and structural exclusivity of the bacterial enzymes have enticed many to exploit this target for the discovery of novel antibacterial agents. NAD$^+$-competitive ligands with antibacterial activity have been reported to selectively inhibit bacterial DNA ligases with no inhibition of human DNA ligase I.[4−6] Crystal structures of LigA bound to representatives of each class of small molecule inhibitors show that the ligands occupy the AMP-binding site of the adenylation domain (PDB 3BA8, PDB 3BAA, and PDB 3PN1). Interestingly, the AMP-binding pocket of the adenylation domain is the most highly conserved region between the two classes of DNA ligases. Thus, the specificity of substrate recognition must be contributed by recognition elements that are away from the adenylation pocket.

NAD$^+$-dependent DNA ligases have a modular architecture consisting of four distinct structural domains: the N-terminal adenylation domain (further divided into subdomains 1a and 1b), the oligomer-binding domain (OBD), the zinc finger and helix-hairpin-helix domain (HhH domain), and the C-terminal BRCT domain for BRCA1 Carboxyl Terminus.[7,8] Numerous crystal structures of substrate-bound LigA complexes, coupled with biochemical and genetic studies, have helped to illuminate key structure−function relationships of the reaction mechanism.[3,7,9−16]

The adenylation domain was initially defined by limited proteolytic digestion and gene deletion analysis where large N-terminal fragments of LigA containing the conserved active-site lysine were shown to be capable of self-adenylation but were not active in DNA ligation.[17−19] Sriskanda and Shuman[10] convincingly demonstrated that subdomain 1a, which is unique to NAD$^+$-dependent enzymes, is essential for NAD$^+$-dependent adenylation of *E. coli* LigA. They further speculated that five conserved residues identified in the study are constituents of an NMN-binding site. As a result, a model for NAD$^+$ binding and catalysis was proposed.[7,8,10,11,15,16,20] The most current model for the LigA adenylation proposes that the reaction is initiated by the binding of NAD$^+$ to the AMP pocket in subdomain 1b.[2,15] This binding is accompanied by a significant conforma-

Received: September 25, 2011
Accepted: January 9, 2012
Published: January 9, 2012

tional movement of domain 1a to close over NAD^+. As a result of closure, conserved residues of domain 1a bind the NMN moiety of NAD^+ to orient the substrate for cleavage, leaving AMP covalently linked to the conserved lysine side chain in subdomain 1b and NMN noncovalently linked to subdomain 1a.

Structural data supports that a high degree of flexibility exists between subdomain 1b and subdomain 1a that appears to be critical for the adenylation reaction. All apo or product-bound forms exhibit an "open" form of the enzyme, where subdomain 1a is located at a significant distance away from the catalytic pocket within subdomain 1b.[16,21−23] A closed intermediate stage is represented by the NAD^+-bound *E. faecalis* structure, where a 180° swivel repositions subdomain 1a in close proximity to the adenylation pocket.[16] While there is structural information on the product (AMP and NMN) and the intermediate states (NAD^+ bound to the closed conformation), the structure of the ligase bound to NAD^+ prior to adenylation is still unknown. It is assumed that NAD^+ binds to subdomain 1b first, and subdomain 1a closes over the NAD^+ to orient the NMN leaving group onto subdomain 1a.[10] This reorientation may also be necessary to trigger the cleavage of the substrate. However, there is no direct evidence that the recognition of NAD^+ is by the AMP pocket.

Here, we report a 1.7 Å structure of *Haemophilus influenzae* LigA adenylation domain that is covalently bound to AMP through the active site lysine in subdomain 1b and is in complex with NAD^+ in subdomain 1a. The ligand occupancy of the structure solved here is reflective of the species obtained directly upon overexpression and purification from *E. coli* cells; no additional ligands were added during the crystallization experiment. The structure shows an alternative snapshot of NAD^+ bound to LigA. It is the first structure containing NAD^+ and AMP and suggests a new model for the mechanism of substrate binding during catalysis.

RESULTS AND DISCUSSION

Overall Structure of *H. influenzae* LigA Adenylation Domain. The adenylation domain (residues 1−324) of *H. influenzae* DNA ligase was expressed in *E. coli*, and the purified protein was used for crystallization without further addition of ligands or other treatments to alter the adenylation state of the protein. The overall structure is similar to those solved previously from other bacterial species and is composed of two subdomains.[7,15,16,23] The smaller subdomain 1a (residues 1−60) is a helix-turn-helix motif, while the larger subdomain 1b (residues 71−324) has two antiparallel β-sheets flanked by α-helices (Figure 1A). The two domains are connected by a linker region made up of approximately 10 residues. The relative disposition of the two domains makes the structure similar to the most commonly observed open conformation where subdomain 1a is farthest away from subdomain 1b. This is very different from the previously solved NAD^+ bound *E. faecalis* structure, which is in a closed form wherein subdomain 1a repositions itself near the active site of subdomain 1b.[16] Each subdomain of *H. influenzae* adenylation domain superimpose well with the respective subdomain of *E. faecalis* structure. However, due to the flexibility of the linker region, the relative orientation of the two subdomains in the *H. influenzae* ligase structure is distinct and corroborates the observed variability among all known structures (Figure 1B).

Subdomain 1b contains density for AMP indicating that it is contiguous with the electron density of the catalytic Lys117.

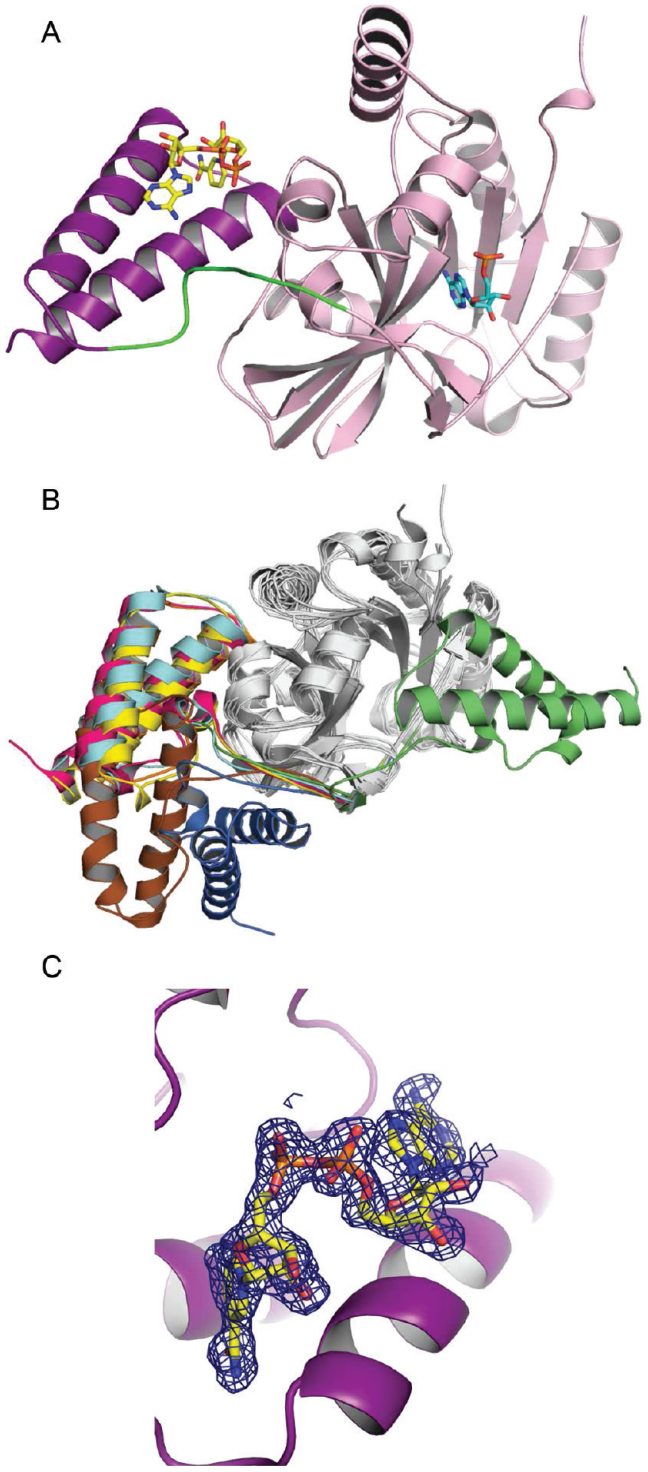


Figure 1. Structure of *H. influenzae* adenylation domain complexed with NAD^+. (A) Overall structure is depicted in ribbon diagram with the ligands shown in sticks. Subdomain 1a is shown in purple, subdomain 1b is colored pink, and the linker region is colored green. NAD^+ is depicted in yellow stick model, and covalently linked AMP is highlighted in cyan stick. (B) Overlay of the adenylation domains of different bacterial species to show the conformational differences in subdomain 1a with respect to subdomain 1b. Subdomain 1b for all is colored gray. Subdomain 1a from *E. faecalis* in cocomplex with NAD^+ is colored green (PDB 1TAE), *E. coli* full-length in blue (PDB 2OWO), *Mycobacterium tuberculosis* in brown (PDB 1ZAU), *Staphylococcus aureus* in yellow (PDB 3JSL) and *H. influenzae* in cyan. In case of E. *coli*, the remaining domains of the full-length protein have

dx.doi.org/10.1021/cb200392g | *ACS Chem. Biol.* 2012, 7, 571−580

Figure 1. continued

not been shown for clarity. (C) Electron density from a $F_o - F_c$ map at 2.5σ level is shown in blue cage around the NAD^+ molecule before inclusion of the molecule to the model.

The connected density suggests a covalent phosphoramide (P–N) bond between the AMP and the Lys that confirms the adenylated state of the protein. The overall AMP binding mode is similar to that observed in other AMP-bound adenylation domain structures, which is not surprising considering the similarity in the interactions with conserved residues. In addition to the AMP density in subdomain 1b, density was observed in subdomain 1a that clearly resembled an NAD^+ molecule (Figure 1c).The density for NAD^+ was observed in the previously described[16] NMN pocket of the protein. Thus, despite multiple column chromatography steps both covalently bound AMP as well as a noncovalently bound NAD^+ molecule were copurified with the protein.

Overexpressed *H. influenzae* DNA Ligase Adenylation Domain Is Adenylated and Copurifies with NAD^+. The protein preparation of adenylation domain construct for crystallographic studies as well as the full-length protein was checked for the degree of protein adenylation by mass spectral analysis, which showed that both protein forms are fully adenylated. In addition, it was specifically investigated whether the expressed proteins were binding to and copurifying with either NMN or NAD^+. Multiple-reaction monitoring (MRM) mass spectral methods were implemented to compare the total quantity of protein with the amount of NAD^+ or NMN recovered after heat-denaturation of the protein (Figure 2). The adenylation domain construct was found to have 0.13–0.17 equiv of NAD^+ present per protein molecule, as determined from the MRM area ion transition at 664.5/136.4 amu. The low amount of NAD^+-bound population is not unexpected, given the modest affinity of the molecule to the protein and the absence of extrinsic NAD^+ in the purification buffers. The amount of NAD^+ present in the full-length preparation was comparable, while no NMN was detected.

In addition, an *H. influenzae* DNA ligase adenylation domain construct was expressed and purified that was engineered to delete residues comprising subdomain 1a. A similar construct has been shown previously[27] to be incapable of self-adenylation. Upon purification, this construct was not adenylated, and mass spectrophotometeric analysis showed that neither NMN nor NAD^+ was copurified with the protein.

These results demonstrate that when subdomain 1a is present, the ligase is competent to bind NAD^+ in bacterial lysates and presumably in the cell. In addition, the presence of additional DNA binding domains in the full-length protein did not eliminate the ability of the adenylation domain to bind to NAD^+ under these conditions.

Binding Studies Corroborate the Relevance of NAD^+ Binding to Subdomain 1a in Solution. In order to confirm that the crystallographically observed NAD^+ binding to subdomain 1a has relevance within the solution phase, isothermal titration calorimetric binding studies were undertaken. Importantly, a standard binding buffer was selected to exclude Mg^{2+} ions and included 1 mM EDTA to avoid the complication of NAD^+ adenylation reaction (or its reverse) occurring during the studies. The suitability of excluding Mg^{2+} from the buffer was investigated by determining the dissociation constant of NAD^+ binding to adenylated full

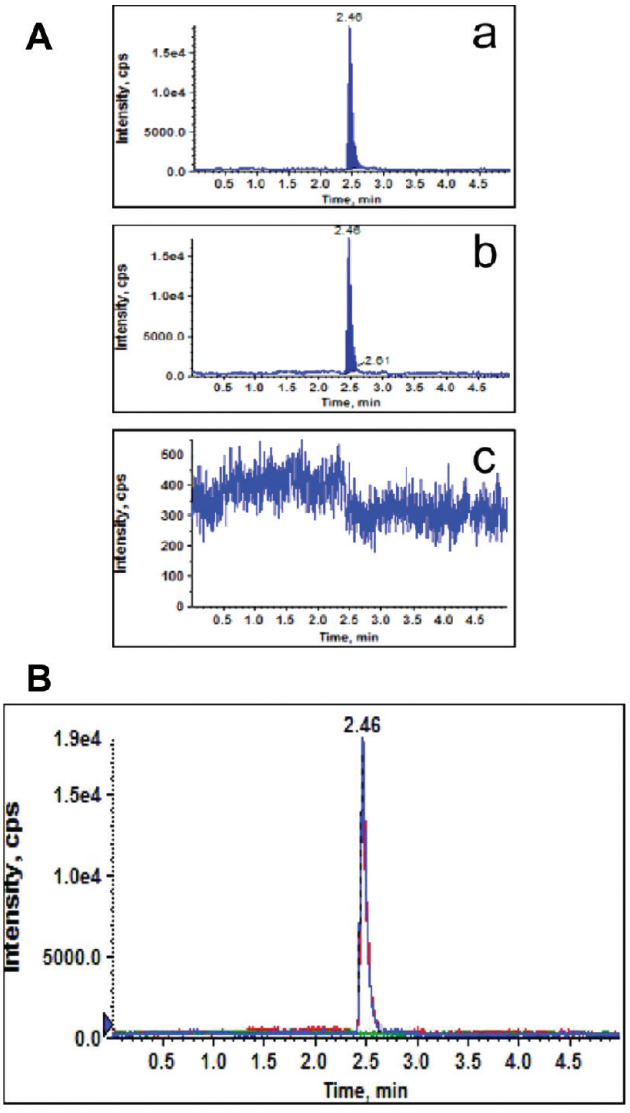


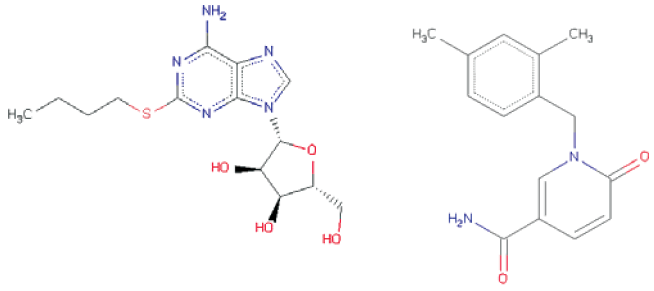
Figure 2. Quantitation of NAD^+ copurified with *H. influenzae* DNA ligase proteins by LC–MS/MS. (A) Individual MRM chromatogram of NAD^+: *a* is *H. influenzae* adenylation domain, *b* is *H. influenzae* full-length, *c* is *H. influenzae* adenylation domain with 1a deletion. (B) Overlay MRM chromatograms of NAD^+: blue, *H. influenzae* adenylation domain; red, *H. influenzae* full-length; green, *H. influenzae* adenylation domain with 1a deletion.

length *H. influenzae* ligase with and without magnesium ion present. In this case, further adenylation reaction cannot confound the noncovalent binding event of the ligand. It was observed that the affinity of the protein for NAD^+ in the presence of 8 mM $MgCl_2$, K_d = 40 ± 3 μM, was not substantially changed by the absence of Mg^{2+}, K_d = 45 ± 2 μM. Based on this result, additional studies were performed using the adenylation domain protein construct because of its direct relation to the crystal system and the greater convenience in producing large quantities (>100 mg) needed to conduct the studies.

Affinities of NAD^+, NMN, compound 1 (adenosine analogue) and compound 2 (nicotinamide analogue) to various forms of the *H. influenzae* adenylation domain (residues 1–324) were determined. Compounds 1 and 2 were chosen because they have previously been shown to bind simultaneously in the AMP binding pocket of domain 1b and in the

dx.doi.org/10.1021/cb200392g | *ACS Chem. Biol.* 2012, 7, 571–580

Table 1. *H. influenzae* Ligase Adenylation Domain Binding Dissociation Constants (μM) Determined by ITC



Compound 1 Compound 2

	direct binding cell contents		competitive binding cell contents	
syringe contents	adenylated	deadenylated	deadenylated protein and compound 1	deadenylated protein and compound 2
compound 1	no binding	4.4 ± 0.5		NT
compound 2	6.4 ± 0.2	7.6 ± 0.4	7.9 ± 0.3	
NMN	77 ± 11	76 ± 8	130 ± 40	competition
NAD	28 ± 3	28 ± 1	35 ± 2	competition

NMN binding pocket of domain 1a, respectively[6] (PDB 3PN1). Results of these studies are presented in Table 1. ITC data shows, as expected, that compound 1 binds only to the deadenylated construct. Additionally, compound 2 binding is indifferent to the adenylation state of the protein or the presence of compound 1. Likewise, NAD^+ and NMN each show negligible change in their binding affinity to the protein in different adenylation states. Of particular importance, both NAD^+ and NMN exhibit competitive behavior with compound 2 when binding to the deadenylated form of the protein (Figure 3). In addition, NAD^+ does not bind with measurable affinity to the *H. influenzae* adenylation domain construct in which domain 1a and the linker region (residues 2−69; Figure 4b) shown to make contact with NAD^+ have been deleted (Supp Figure 1). Conversely, compound 1 does bind to this construct with a K_d of 6 ± 1 μM and a stoichiometry of 0.96 ± 0.03 (Supplementary Figure 2). The positive result for the binding of compound 1 to the domain 1a deletion construct highlights that this construct remains properly folded to achieve binding through the domain 1b residues and further supports that NAD^+ binds with domain 1a.

NAD^+ Binding Pocket. A clear and extended density was observed in the previously described NMN binding pocket[16] that fits to a NAD^+ molecule. The electron densities of both the pyrimidine and the adenosine halves of the molecule are well-defined (Figure 1c). The conformation of the NAD^+ pocket observed here is distinct from the previously reported NAD^+ pocket in *E. faecalis* ligase. In the *E. faecalis* structure the NAD^+ was the captured product of the reversed "synthesis" reaction in the presences of NMN and AMP. In the structure described here, the subdomain 1a is in the open conformation at a significant distance from the AMP pocket. All residues involved in NAD^+ interactions are contributed by subdomain 1a with the exception of Arg 154, which comes from subdomain 1b (Figure 4a). Most of these residues are conserved among bacterial species (Figure 4b). The binding pocket forms a deep cleft around the pyrimidine recognition motif but is shallow and solvent-exposed around the rest of the NAD^+ binding site (Figure 4c). The pyrimidine binding mode observed here is very similar to those previously observed[16] in the presence of NAD^+ and NMN (Figure 5a and b). The nicotinamide ring is stacked by Tyr22 and Tyr35 on either side and is held tightly via specific polar interactions to the side chains of Tyr18, Glu19, His23, Pro28, Val30 and the main chains of Ser29, Asp32, and Glu19. The ribose moiety of the nicotinamide ring is recognized by Asp36 and Asp32, while Arg154 from subdomain 1b stabilizes the pyrophosphates.

In addition to the highly conserved interactions with the pyrimidine group of NAD^+ and NMN, the adenosine half of

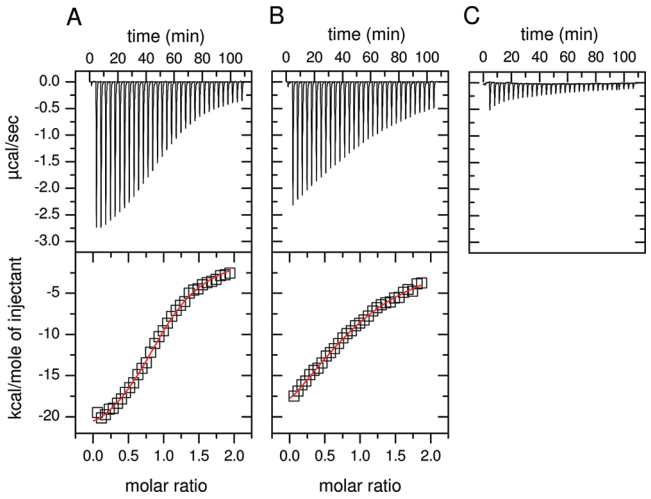


Figure 3. Binding of compound 2 and NAD^+ to deadenylated *H. influenzae* DNA ligase adenylation domain at 25 °C. Other conditions common to all plots are detailed in the Methods section. (A) Titration of 500 μM Compound 2 (syringe) into 50 μM protein (cell). ITC Dialysis Buffer was amended with 2.5% v/v DMSO in both cell and syringe to match solution composition. The dissociation constant calculated is 7.6 ± 0.4 μM. Other parameters values obtained for *N*, ΔH, and ΔS are 0.999 ± 0.008, −23.6 ± 0.3 kcal/mol, and −56 cal/(mol K), respectively. (B) Titration of 500 μM NAD^+ (syringe) into 50 μM protein (cell). The dissociation constant calculated is 28 ± 3 μM. Other parameters values obtained for *N*, ΔH, and ΔS are 1.00 ± 0.04, −28 ± 1 kcal/mol, and −73 cal/(mol K), respectively. (C) Titration of 500 μM NAD^+ (syringe) in into 50 μM protein with 75 μM compound 2 (cell). ITC dialysis buffer was amended with 2.5% v/v DMSO to match solution composition. The resulting heat signature could not be fit suitably to a one site binding model, consistent with NAD^+ and the more potent binding compound 2 competing for the same site.

dx.doi.org/10.1021/cb200392g | *ACS Chem. Biol.* 2012, 7, 571−580

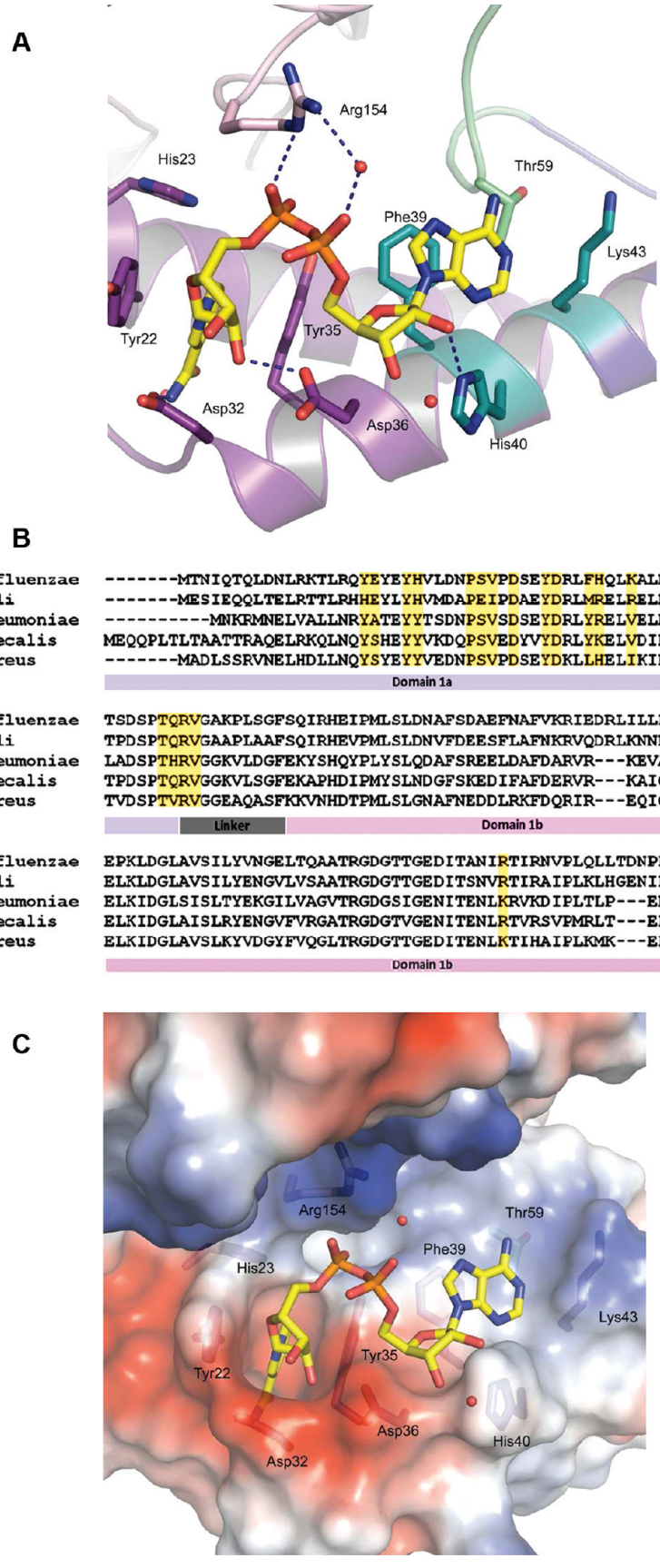


Figure 4. NAD^+ binding pocket. (A) A close view of the NAD^+-binding site on subdomain 1a. The residues involved in pyrimidine interaction are depicted in purple sticks, and those contacting the adenosine group are shown in cyan sticks. The residue from loop region is in green. (B) Alignment of N-terminus of LigA adenylation domain from *H. influenzae* Rd strain KW20, *E. coli* strain MG1655, *S. pneumoniae* strain R6, *E. faecalis* strain V583, and *S. aureus* strain Mu50. Residues involved in binding NAD^+ in yellow shaded boxes. Arg154 only residue from domain 1b involved in

dx.doi.org/10.1021/cb200392g | *ACS Chem. Biol.* 2012, 7, 571–580

Figure 4. continued

binding to NAD^+. (C) An electrostatic surface map of the NAD^+ pocket in subdomain 1a with the NAD^+ depicted in yellow sticks. The surface is colored red to blue from negative to positive charge (calculated using vacuum electrostatics in PyMol).

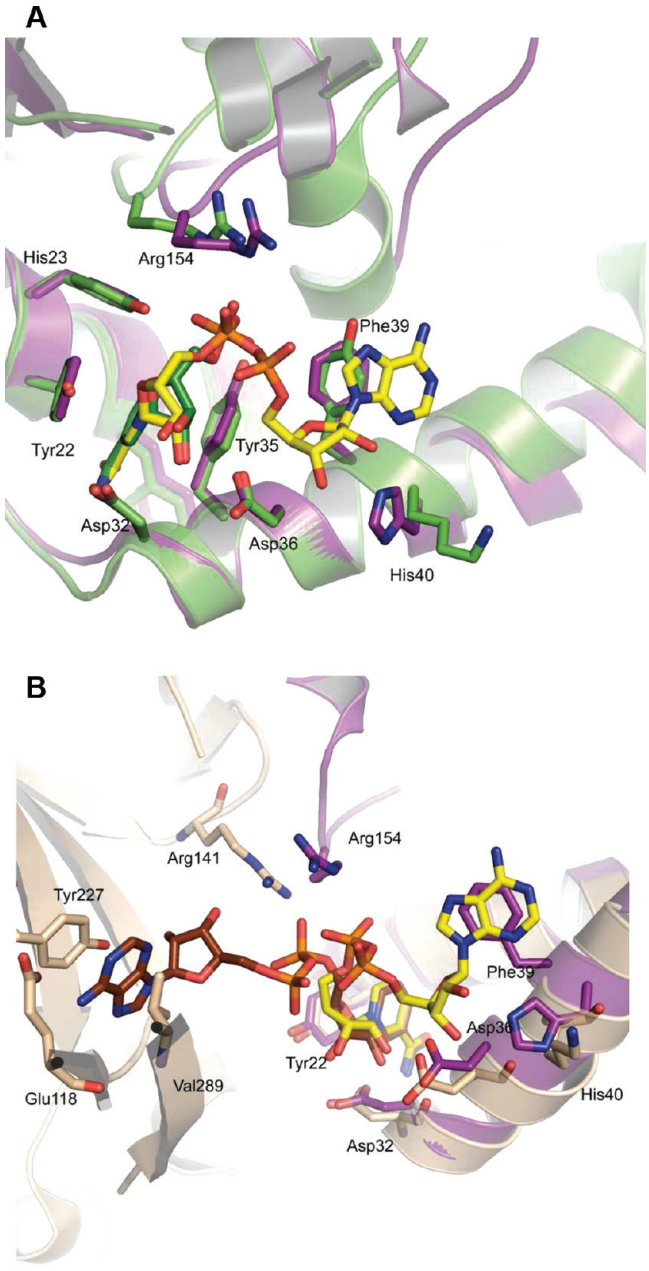


Figure 5. Comparison of the NAD^+ pocket. (A) Overlay of the *H. influenzae* NAD^+ pocket labled and in purple with NMN pocket of *E. faecalis* open structure in green (PDB 1TE8). The residues on the protein are highlighted in sticks with the respective colors, and the NAD^+ molecule is in yellow stick. (b) The NAD^+ pocket is overlaid on the closed NAD^+ structure from *E. faecalis* adenylation domain (PDB 1TAE). NAD^+ in *H. influenzae* is colored yellow (protein in pink), and the NAD^+ and protein in *E. faecalis* is depicted in shades of brown.

NAD^+ makes previously unobserved interactions with the residues on the shallow region of the binding pocket (Figure 4c). In particular, contacts are observed with the side chains of Lys43, His40, Thr59, Phe39, Val62, and Arg61, and the main-chains of Thr59, Gln60, Arg61, and Val62. Although many of these contacts are not predicted to make strong interactions, these residues help to form the pocket that holds the adenosine group in place. A key hydrogen bond is made by His40 to the 2′ OH of the adenosine ribose, which may explain the observed specificity of DNA-ligases for NAD^+ over NADP. Many of the residues involved in key interactions with NAD^+, such as Arg154 and Asp32, are also conserved across all bacterial species examined (Figure 4b). This suggests that the binding mode observed in this crystal structure is likely to be conserved among other NAD^+-dependent DNA ligases.

The conformations of NAD^+ bound to the open and closed pockets are very different (Figure 5b). When the NAD^+ molecules of the two structures are superimposed by their pyrimidine groups, the rest of the molecules rotate by nearly 90° relative to each other. This rotation is achieved by torsion around the C5 position of the pyrimidine-ribose group. The phosphate groups are more extended in the closed conformation than in the open conformation, which also implies a larger volume of the closed substrate pocket compared to the open form. The adenosine groups in both molecules remain in the *syn* conformation, although upon covalent addition to the lysine side chain, the conformation of the AMP changes to *anti*.

Comparison of Adenylation Domain and Full-Length Ligase Structures. The adenylation domain and full-length ligase structures were compared to understand the physiological relevance of the NAD^+ binding mode observed in this structure. The question asked specifically was if this binding mode is likely to be preserved in the full-length ligase and whether DNA-binding is likely to affect NAD^+ binding to subdomain 1a pocket. Because the NAD^+-dependent ligases are structurally and functionally very well conserved, it is possible to compare multiple structures of ligases from different bacterial species. The availability of the full-length structure of the apo protein from *Thermophilus filiformis*[7] and a full-length DNA bound structure from *E. coli*[15] allowed us to compare the relevant forms to answer these questions. Analysis of these structures suggests that NAD^+ should be able to bind to the adenylation domain of a full-length protein without any steric interference from the HhH and BRCT domains, which are further away from subdomain 1a in both forms. In the DNA-bound full-length structure, subdomain 1a is closer to the DNA and there is a shift in the relative orientation of subdomain 1a when compared to the adenylation domain structure. This conformation is distinct from both open and closed forms.[15] Comparison to the NAD^+ bound structure (Figure 6) shows that the relative positions of the residues involved in NAD^+ binding remain unchanged since subdomain 1a moves as a rigid body. However, if the NAD^+ were to bind to the subdomain 1a conformation observed in the *E. coli* DNA-bound structure, the close proximity of the pocket to DNA may cause some steric clash. It is more likely that the flexibility of subdomain 1a will readily allow it to adopt the conformation observed in this NAD^+ bound structure since there are minimal contacts of subdomain 1a to the DNA suggesting that it is not required for stabilizing DNA interactions. Further, functional role of subdomain 1a and hence any interaction of this domain to the DNA are dispensable for DNA binding and nick ligation.

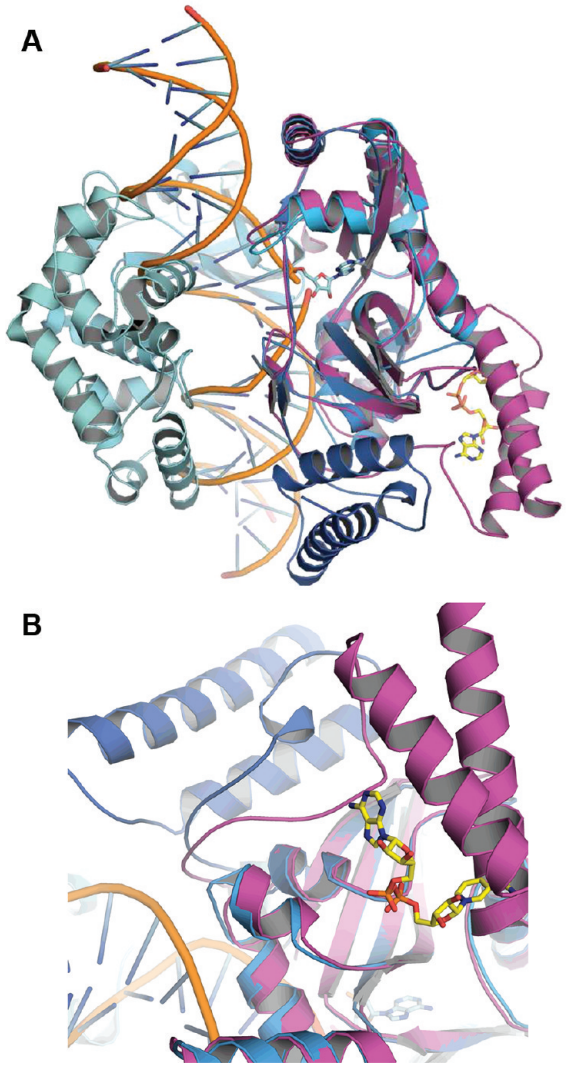


Figure 6. Comparison to full-length structure. Superposition of *H. influenzae* adenylation domain in purple to the *E. coli* DNA bound full-length structure in cyan (PDB 2OWO), shown (A) from a distance and (B) zoomed near the NAD^+ pocket. NAD^+ bound to *H. influenzae* is depicted in yellow, and AMP bound to *E. coli* is shown in cyan sticks. AMP molecule covalently bound to *H. influenzae* has not been shown for clarity.

Mechanistic Implication. It was first noted by Sriskanda and Shuman[10] that domain 1a is unique to bacterial ligase and contains crucial recognition elements responsible for NAD^+ specificity. However, it has been believed that both NAD^+, in the case of NAD^+-dependent ligases, and ATP, in case of ATP-dependent ligases, bind to the adenylation domain via adenosine recognition. In the NAD^+-dependent ligase, the binding of NAD^+ is then expected to be triggering a reorientation of the NMN pocket, resulting in subdomain 1a interactions with subdomain 1b. In both cases, the initial recognition motifs are suggested within the adenylation pocket, which is highly conserved between the two enzymes. Thus, if the adenosine group were to be the primary recognition motif for both, one would expect very little selectivity between the two substrates. However, biochemical and mutational data suggests very little cross reactivity between the substrates such that the NAD^+-dependent DNA ligase are nonfunctional in the presence of ATP.[28] Then how is the NAD^+-dependent ligase able to, exclusively, recognize NAD^+ and not ATP? Structural data in the past have highlighted the role of subdomain 1a for NAD^+ recognition. Multiple structures with NAD^+ and NMN with various isozymes have shown that residues in subdomain 1a are responsible for pyrimidine binding. In addition, mutational studies have demonstrated the essentiality of subdomain 1a residues for enzyme activity, where the adenylation reaction is completely abolished on deletion of key NMN interacting residues without any impairment to the ligation reaction.[10] Similar observations have also been made when the entire subdomain 1a was deleted from the protein.[29] Thus, the importance of NMN binding motifs for catalysis is apparent. Yet, to our knowledge no data exists that clearly shows the impact of subdomain 1a on NAD^+ binding. Specifically, does the subdomain 1a play a role in triggering catalysis or does it influence substrate binding?

The structural and binding studies described here demonstrate that subdomain 1a is competent to act as an initial recognition site for the substrate. In the cell, NAD^+ is able to compete off the product NMN since the protein comes copurified exclusively with NAD^+ and not NMN. The affinity of NAD^+ being higher than NMN for the deadenylated protein (Table 1) and the fact that NAD^+ is one of the more concentrated nucleotides in bacterial cells, present at ~3 mM,[30] and the total NMN concentration is approximately 5-fold lower[31] is also consistent with the expectation of NAD^+ exchanging with NMN in the cellular milieu. This binding is achieved by the presence of subdomain 1a, since no detectable NAD^+ was observed in the subdomain 1a-deleted protein purified under similar conditions. As expected, this construct is also incapable of self-adenylation in the presence of excess NAD^+. Similarly, the NAD^+ binding affinities for the adenylated versus deadenylated forms of adenylation domain are comparable, which suggests that the primary recognition elements for NAD^+ preceding enzyme adenylation are present in domain 1a and not 1b and that the binding properties of the two domains are independent. Assignment of this initial binding to site 1a is substantiated by the demonstration of competition between NAD^+ and compound 2, and indifference in binding between NAD^+ and compound 1.

We believe that the crystal structure observed here of an adenylated ligase bound to NAD^+ might be reflective of a relevant species within a bacterial cell and that NAD^+ binding being driven through initial interactions with domain 1a may contribute to the NAD^+ selectivity exhibited by bacterial DNA ligases. This idea is intellectually similar to and consistent with the prior suggestion[16,32] that NAD^+-induced formation of the covalent AMP phosphoramide likely induces the separation of subdomains 1a and 1b, thereby removing the product NMN from the active site. Specifically, our hypothesis (Figure 7) is that adenylated ligase remains competent to bind NAD^+ in subdomain 1a and that upon completion of DNA-processing steps the enzyme may already have NAD^+ bound and available for initiating readenylation via the closed conformation.

A feature of this hypothesis is that the recognition of NAD^+ in subdomain 1a, predominantly through its nicotinamide functionality, provides a rationale for how the enzyme attains some of the selectivity for the adenylation substrate. Finally, we interpret the structure of full length *E. coli* LigA bound to nicked DNA,[15] in which the NMN binding site of domain 1a is in an "open" conformation, to suggest that this subdomain remains available for NAD^+ binding. In combining these observations we find it a reasonable speculation that by forming

dx.doi.org/10.1021/cb200392g | *ACS Chem. Biol.* 2012, 7, 571–580

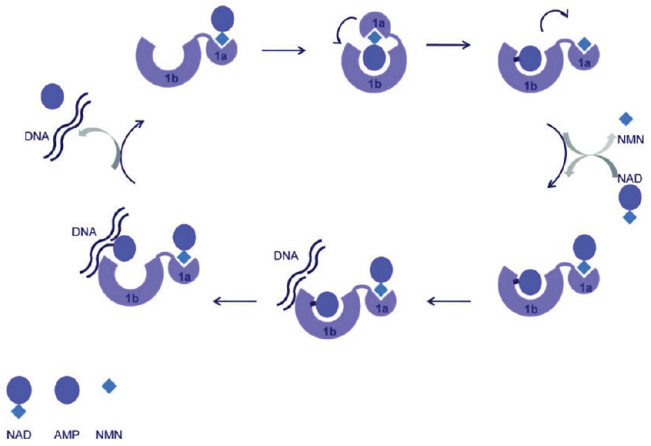


Figure 7. Proposal for the recruitment of NAD^+ to bacterial DNA ligase via initial interactions with subdomain 1a. A molecule of NAD^+ binds first to subdomain 1a (top left), followed by a conformational change that brings subdomain 1a in close proximity to subdomain 1b (top center). This recruitment and recognition of NAD^+ through the nicotinamide moiety followed by conformational change is proposed to provide some selectivity in the selection of the nucleotide substrate. This is followed by the adenylation reaction to covalently modify subdomain 1b. After enzyme adenylation, subdomains 1a and 1b are proposed to return to an open conformation (top right). Exchange of NMN for NAD^+ results in adenylated protein with a bound NAD^+ in domain 1a (bottom right), which is the structural species crystallized and described in this text. We speculate that the NAD^+ may remain bound through the proceeding DNA-processing steps (bottom center and left). Having the NAD^+ bound within subdomain 1a, the enzyme may then adenylate the protein for a subsequent round of catalysis.

the noncovalent subdomain 1a-NAD^+ complex and then undergoing a conformational change to bring the nucleotide into proximity with subdomain 1b for the adenylation reaction, elements of substrate selection are employed in advance of the final ordering of the protein conformation required for the adenylation reaction in the "closed" conformation.

METHODS

Cloning, Expression, and Purification. Details of cloning, expression, and purification along with the deadenylation procedures for ITC experiments are provided in the Supporting Information. In addition, the method for adenylation is described, which was used for ITC studies, where adenylated *H. influenzae* DNA ligase adenylation domain protein was prepared from the deadenylated stocks.

Isothermal Titration Calorimetry. ITC titrations were performed with a Microcal VP-ITC instrument (GE Healthcare, Piscataway, NJ). Experiments were conducted using vacuum-degassed solutions at a temperature of 25 °C, a reference power of 10 μcal s^{-1}, a stirring rate of 270 min^{-1}, and a delay between injections of 200 s. Experiments were run with the ligase protein receptor in the analysis cell and ligand in the injector. Injections were programmed with an initial 1 μL addition followed by 32 successive 8 μL additions. The buffer for all titrations consisted of ITC Dialysis Buffer consisting of 50 mM MOPS pH 7.5, 30 mM ammonium sulfate, 30 mM potassium chloride, 1 mM EDTA, 10% v/v glycerol, and 0.002% Brij-35. When ligand stocks for titration could not be made in ITC Dialysis Buffer, they were dissolved at high concentrations in DMSO (compounds 1 and 2) or in 0.5 M MOPS pH 7.5 (NAD^+ and NMN) such that the preparation of composition matched cell and syringe solutions required less than 2.5% dilution of ITC Dialysis Buffer. Typically, 50 μM protein was present in the cell, and ligand was added from the syringe at a concentration of 500 μM. For compound competitive studies, protein was present in the cell at 50 μM along with 75 μM of the competing compound. Data were analyzed using the software supplied by the manufacturer and fit to a single site binding model in all cases, and the errors listed are the results from the fitting process.

Table 2. Refinement Statistics for *H. influenzae* LigA Adenylation Domain Final Model

Data Collection	
wavelength	0.933
space group	$P4_32_12$
unit cell	a = 70.23, b = 70.23, c = 161.28
resolution limit (Å)	1.70
resolution range (Å)	34.9−1.70 (1.81−1.7)[a]
completeness overall (%)	97.7 (86.2)
redundancy	2.8 (1.5)
$Rmerge_{overall\ (\%)}$[b]	8.3 (20.6)
$I/SigI$	13.4 (2.1)
Refinement	
$Rvalue_{overall}$ (%)[2]	21.5
$Rvalue_{free}$ (%)	23.8
rms deviations from ideal values	
bond lengths (Å)	0.06
Bond angles (deg)	1.9
av B values (Å)	
protein main chain atoms	14.2
protein all atoms	15.5
ligand	
NAD^+	25.3
AMP	12.2
solvent	23.8
Φ, Ψ angle distribution for residues[c]	
in most favored regions (%)	92.9
in additional allowed regions (%)	6.7
in generously allowed regions (%)	0.4
in disallowed regions (%)	0

[a]Values in parentheses refer to the high-resolution shell. [b]$R_{merge} = S_{hkl}[(\Sigma_i\ |I_i - \langle I\rangle|)/\Sigma_i I_i]$ [2]$R_{value} = S_{hkl}\ ||F_{obs}| - |F_{calc}||/S_{hkl}\ |F_{obs}|$ [c]Ramachandran statistics as defined by PROCHECK

NAD^+ Quantitation. LC−MS was performed on a Shimadzu HPLC system connected to a Sciex API 4000 Mass Spectrometer equipped with a turbo ion-spray ionization source and operated in the positive mode. The protein was heated at 60 °C for 15 min and then centrifuged at 14,000 rpm. The supernatant was injected on an Atlantis T3 C18 column, 3 μm particle size, 2.1 mm × 30 mm. The mobile phase was 5 mM ammonium formate, pH 4.0 (A) and 100% methanol (B). The separation was performed by a linear gradient that began at 2% B (methanol) and ended after 2 min with 40% B (methanol). The mobile phase was then held at 90% for 1 min for a column wash followed by column regeneration for 1 min. The flow rate was kept constant at 0.2 mL min^{-1}. MRM of NAD^+ was used to monitor and quantify NAD^+. Mass spectrometer source conditions of curtain gas, Gas 1, Gas 2 were set at 30, 45, and 35, respectively. The ion spray voltage was 5000 and the temperature was 500 °C. The ion transition of NAD^+ was 664.5/136.4 amu. The optimum voltage settings for the declustering potential (DP), the collision energy (CE), and the collision exit potential (CXP) were 80, 65, and 8. A standard calibration curve of NAD^+ was generated and used to determine the concentrations of the unknown samples.

Crystallization. Purified *H. influenzae* DNA ligase adenylation domain was subjected to sparse matrix crystallization screening, using a protein concentration of about 40 mg mL^{-1} in 1 mM Tris-HCl pH 7.5 at a temperature 18 °C. Crystals were obtained by vapor diffusion using the hanging drop method. Purified adenylation domain was concentrated to 40 mg mL^{-1} in 1 mM Tris-HCl pH 7.5 and crystallized with the reservoir solution contained 16% (w/v) polyethylene glycol (PEG) 3500 and 350 mM sodium potassium tartrate. Hanging drops were set up at 18 °C by mixing 2 μL of protein solution with 2 μL of the reservoir solution and crystals were observed

dx.doi.org/10.1021/cb200392g | ACS Chem. Biol. 2012, 7, 571−580

within 3−4 days. The crystals were equilibrated in drops containing a cryoprotectant solution containing 20% glycerol in mother liquor before flash freezing in liquid nitrogen.

Structure Solution. Crystals diffracted to 1.7 Å resolution using the ESRF synchrotron radiation source at Grenoble. Data were integrated, scaled and merged using MOSFLM.[24] Crystals belonged to the tetragonal space group $P4_32_12$ and had cell parameters of $a = b = (70.23 \pm 0.7)$ Å, $c = (161.28 \pm 0.3)$ Å, and $\alpha = \beta = \gamma = 90.00°$. The details of the statistics are provided in Table 2. Analysis of the crystal unit cell dimensions indicated that the asymmetric unit likely contained a single molecule. The phase problem for the structure was solved by molecular replacement using the program Amore.[25] A search model was generated by threading the sequence of *H. influenzae* LigA adenylation domain onto the previously determined structures of *B. stearothermophilus* LigA adenylation domain (1B04) using Coot.[26] The entire data set was further refined using the rigid body refinement technique in CCP4i. Several rounds of interactive model building were then carried out using the program Coot. A significant peak in the difference Fourier ($F_o - F_c$) electron density map consistent with the presence of covalently linked AMP group was observed close to the conserved Lys116 with electron density linked to the residue. An additional significant density in the difference Fourier electron density map was found in the N terminal domain 1a that best fit the NAD^+ molecule. Adenosine and NAD^+ moieties were included once the R_{free}, the cross-validation R factor computed for the test set of 5% of unique reflections, dropped below 25%. The final statistics are shown in Table 2.

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

Coordinate and structure factors have been deposited in the Protein Data Base database with the code 3UQ8.

■ AUTHOR INFORMATION

Corresponding Author

*Ph: 781-839-4714. Fax: 781-839-4600. E-mail: sushmita.lahiri@astrazeneca.com.

■ REFERENCES

(1) Wilkinson, A., Day, J., and Bowater, R. (2001) Bacterial DNA ligases. *Mol. Microbiol. 40*, 1241−1248.
(2) Pascal, J. M. (2008) DNA and RNA ligases: structural variations and shared mechanisms. *Curr. Opin. Struct. Biol. 18*, 96−105.
(3) Shuman, S. (2009) DNA ligases: progress and prospects. *J. Biol. Chem. 284*, 17365−17369.
(4) Brotz-Oesterhelt, H., Knezevic, I., Bartel, S., Lampe, T., Warnecke-Eberz, U., Ziegelbauer, K., Habich, D., and Labischinski, H. (2003) Specific and potent inhibition of NAD-dependent DNA ligase by pyridochromanones. *J. Biol. Chem. 278*, 39435−39442.
(5) Meier, T. I., Yan, D., Peery, R. B., McAllister, K. A., Zook, C., Peng, S.-B., and Zhao, G. (2008) Identification and characterization of an inhibitor specific to bacterial NAD+-dependent DNA ligases. *FEBS J. 275*, 5258−5271.
(6) Mills, S. D., Eakin, A. E., Buurman, E. T., Newman, J. V., Gao, N., Huynh, H., Johnson, K. D., Lahiri, S., Shapiro, A. B., Walkup, G. K., Yang, W., and Stokes, S. S. (2011) Novel bacterial NAD+-dependent DNA ligase inhibitors with broad-spectrum activity and antibacterial efficacy in vivo. *Antibact. Agents Chemother. 55*, 1088−1096.
(7) Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J. K., Kim, H. K., Kwon, S. T., and Suh, S. W. (2000) Crystal structure of NAD-dependent DNA ligase: modular architecture and functional implications. *EMBO J. 19*, 1119−1129.
(8) Doherty, A. J., and Suh, S. W. (2000) Structural and mechanistic conservation in DNA ligases. *Nucleic Acids Res. 28*, 4051−4058.
(9) Sriskanda, V., Schwer, B., Ho, C. K., and Shuman, S. (1999) Mutational analysis of *Escherichia coli* DNA ligase identifies amino acids required for nick-ligation in vitro and for in vivo complementation of the growth of yeast cells deleted for CDC9 and LIG4. *Nucleic Acids Res. 27*, 3953−3963.
(10) Sriskanda, V., and Shuman, S. (2002) Conserved residues in domain Ia are required for the reaction of *Escherichia coli* DNA ligase with NAD+. *J. Biol. Chem. 277*, 9695−9700.
(11) Feng, H, Parker, J. M., Lu, J., and Cao, W. (2004) Effects of deletion and site-directed mutations on ligation steps of NAD +-dependent DNA ligase: a biochemical analysis of BRCA1 C-terminal domain. *Biochemistry 43*, 12648−12659.
(12) Wang, L. K., Zhu, H., and Shuman, S. (2009) Structure-guided mutational analysis of the nucleotidyltransferase domain of *Escherichia coli* DNA ligase (LigA). *J. Biol. Chem. 284*, 8486−8494.
(13) Wang, L. K., Nair, P. A., and Shuman, S. (2008) Structure-guided mutational analysis of the OB, HhH, and BRCT domains of *Escherichia coli* DNA ligase. *J. Biol. Chem. 283*, 23343−23352.
(14) Shuman, S. (1996) Closing the gap on DNA ligase. *Structure 4*, 653−656.
(15) Nandakumar, J., Nair, P. A., and Shuman, S. (2007) Last stop on the road to repair: structure of *E. coli* DNA ligase bound to nicked DNA-adenylate. *Mol. Cell 26*, 257−271.
(16) Gajiwala, K. S., and Pinko, C. (2004) Structural rearrangement accompanying NAD+ synthesis within a bacterial DNA ligase crystal. *Structure 12*, 1449−1459.
(17) Timson, D. J., and Wigley, D. B. (1999) Functional domains of an NAD-dependent DNA ligase. *J. Mol. Biol. 285*, 73−83.
(18) Kaczmarek, S. F., Zaniewski, R. P., Gootz, T. D., Danley, D. E., Mansour, M. N., Griffor, M., Kamath, A. V., Cronan, M., Mueller, J., and Sun, D. (2001) Cloning and functional characterization of an NAD(+)-dependent DNA ligase from *Staphylococcus aureus*. *J. Bacteriol. 183*, 3016−3024.
(19) Lim, J. H., Choi, J., Kim, W., Ahn, B. Y., and Han, Y. S. (2001) Mutational analyses of *Aquifex pyrophilus* DNA ligase define essential domains for self-adenylation and DNA binding activity. Arch. *Biochem. Biophys. 388*, 253−260.
(20) Pascal, J. M., Tsodikov, O. V., Hura, G. L., Song, W., Cotner, E. A., Classen, S., Tomkinson, A. E., Tainer, J. A., and Ellenberger, T. (2006) A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol. Cell 24*, 279−291.
(21) Singleton, M. R., Hakansson, K., Timson, D. J., and Wigley, B. (1999) Structure of the adenylation domain of an NAD-dependent DNA ligase. *Structure 7*, 35−42.
(22) Srivastava, S. K., Dube, D., Kukshal, V., Jha, A. K., Hajela, K., and Ramachandran, R. (2007) NAD+-dependent DNA ligase (Rv3014c) from *Mycobacterium tuberculosis*: novel structure-function relationship and identification of a specific inhibitor. *Proteins 69*, 97−111.
(23) Han, S., Chang, J. S., and Griffor, M. (2009) Structure of the adenylation domain of NAD(+)-dependent DNA ligase from *Staphylococcus aureus*. *Acta Crystallogr., Sect. F F65*, 1078−1082.
(24) Powell, H R. (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr., Sect. D 55*, 1690−1695.
(25) Trapani, S., and Navaza, J. (2008) AMoRe: classical and modern. *Acta Crystallogr., Sect. D 64*, 11−16.
(26) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D 60*, 2126−2132.
(27) Jeon, H. J., Shin, H. J., Choi, J. J., Hoe, H. S., Kim, H. K., Suh, S. W., and Kwon, S. T. (2004) Mutational analyses of the thermostable NAD+-dependent DNA ligase from *Thermus filiformis*. *FEMS Microbiol. Lett. 237*, 111−118.
(28) Modrich, P., and Lehman, I. R. (1970) Enzymatic joining of polynucleotides. IX. A simple and rapid assay of polynucleotide joining (ligase) activity by measurement of circle formation from linear deoxyadenylate-deoxythymidylate copolymer. *J. Biol. Chem. 245*, 3626−3631.

(29) Sriskanda, V., Moyer, R. W., and Shuman, S. (2001) NAD +-dependent DNA ligase encoded by a eukaryotic virus. *J. Biol. Chem. 276*, 36100−36109.
(30) Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol. 5*, 593−599.
(31) Heard, J. T., and Tritz, G. J. (1982) Isoniazind perturbation of the pyridine nucleotide cycle of *Escherichia coli*. *Microbios 35*, 169−178.
(32) Shuman, S. (2004) NAD+ specificity of bacterial DNA ligase revealed. *Structure 12*, 1335−1336.
(33) Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., and Ellenberger, T. (2006) DNA ligases: structure, reaction mechanism, and function. *Chem. Rev. 106*, 687−699.