DNA ligases are the sine qua non of genome integrity and essential for DNA replication and repair in all organisms. DNA ligases join 3'-OH and 5'-PO₄ ends via a series of three nucleotidyl transfer steps. In step 1, ligase reacts with ATP or NAD⁺ to form a covalent ligase-(lysyl-N⁴)–AMP intermediate and release pyrophosphate (PPᵢ) or nicotinamide mononucleotide. In step 2, AMP is transferred from ligase-adenylate to the 5'-mononucleotide. In step 3, ligase catalyzes attack by a DNA 3'-OH on the DNA-adenylate to seal the two ends via a phosphodiester bond and release AMP. All steps in the ligase pathway require a divalent cation cofactor.

The step1 auto-adenylation reaction of polynucleotide ligases relies on a conserved nucleotidyltransferase (NTase)² domain that includes six defining peptide motifs that form the nucleotide-binding pocket (1, 2). Motif I (KXXG) contains the lysine that becomes covalently attached to the AMP. RNA ligases and DNA ligases have different nucleic acid substrate preferences for step 2 of the end-joining pathway that are, to a first approximation, dictated by their distinctive structural modules appended to the NTase domain. DNA ligases comprise a coherent family insofar as they all have a core catalytic unit composed of an NTase domain fused at its C terminus to an OB domain.

ATP-dependent DNA ligases are present in all eukaryota and archaea as well as in many bacteria and DNA viruses. These DNA ligases play key roles in DNA replication, repair, and recombination. They vary greatly in their domain complexity, from minimal versions that consist of little more than the NTase–OB core to larger forms with various additional domains that abet DNA recognition and protein–protein interactions. Within a given taxon, two or more DNA ligase enzymes may coexist, often with a division of labor between them. Examples of minimal ATP-dependent ligases include bacteriophage T7 DNA ligase (3), Chlorella virus DNA ligase (4, 5), bacterial nonhomologous end-joining (NHEJ) ligases LigD and LigC (6–8), and bacterial DNA ligase LigE (9, 10). The more complex multidomain architectures of ATP-dependent DNA ligases are exemplified by mammalian enzymes Lig1, LigII, and LigIV (11–14).

Our understanding of DNA ligase mechanism and specificity has been advanced via crystal structures of exemplary enzymes, captured variously as the apoenzyme, the covalent ligase-(lysyl-N⁴)–AMP intermediate (step 1 product), ligase-AMP bound to a DNA nick (step 2 substrate), ligase bound to an AppDNA nick (step 2 product), ligase bound to a nick that was sealed in crystalllo (step 3 product), or ligase bound noncovalently to AMP (3–5, 8–20). A salient theme from the available structures is that the position of the OB domain (and other flanking

---

This work was supported by National Institutes of Health Grants GM126945 and AI64693. The Memorial Sloan Kettering Cancer Center structural biology core laboratory is supported by NCI, National Institutes of Health Grant P30-CA008748. The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The atomic coordinates and structure factors (codes 6NHX and 6NHZ) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 To whom correspondence should be addressed: Molecular Biology Program, Sloan Kettering Institute, 1275 York Ave., New York, NY 10065. E-mail: s-shuman@ski.mskcc.org.

2 The abbreviations used are: NTase, nucleotidyltransferase; PDB, Protein Data Bank; NHEJ, nonhomologous end-joining; DSB, double-strand break; POL, polymerase; aa, amino acid(s).
domains) relative to the NTase domain is variable according to the functional state in which the ligase is crystallized. This has prompted the suggestion, supported by NMR studies (21), that the ligase domains are flexible in solution and that specific movements accompany or orchestrate the steps of the catalytic cycle. In particular, the OB domain moves when ligase engages its DNA substrate such that the enzyme forms a C-shaped clamp at the nick in which the OB domain engages the minor groove (4, 10–12, 14, 15).

Two aspects of the ATP-dependent ligase mechanism remain poorly understood in structural terms: (i) the enzymatic interactions and catalytic roles of the essential divalent cation cofactor(s) and (ii) the nature of the step 1 Michaelis complex of DNA ligase with ATP and metals. Most structures of ATP-dependent DNA ligases have no divalent cation in the refined model, even in cases in which an active metal cofactor was included during crystallization. Structures have been reported for T7 DNA ligase and human LIGIV in complex with ATP (3, 13), but these are “off pathway” with respect to the lysine adenyllylation reaction because the PPi, leaving group (comprising the ATP β and γ phosphates) is oriented orthogonal to the motif I lysine nucleophile (Nζ–P–O3α angle of ~90°), a situation inimical to an in line-attack by lysine on the ATP α phosphate. In these off-pathway ligase-ATP complexes, which lack a metal cofactor, the OB domain is splayed out away from the NTase domain, and there are no enzymic contacts to the ATP β and γ phosphates. In the structure solved for crystals of *Sulfolobus* DNA ligase that had been soaked in ATP, the PPi, leaving group is apical to the lysine nucleophile (Nζ–P–O3α angle = 155°), but there is no reaction in *crystalllo* in the absence of a metal cofactor (17); in addition, the *Sulfolobus* OB domain is reflected away from the NTase domain. It is thought that step 1 catalysis of lysyl-NMP formation is driven by the presence of metal cofactor(s) and by adoption of a closed conformation of the OB domain in a step 1 Michaelis complex such that the OB domain engages the PPi, leaving group and helps place it apical to the lysine nucleophile (22). However, to our knowledge, there are as yet no structural snapshots of an ATP-dependent DNA ligase with ATP and metal cofactor in an on-pathway conformation.

Here we endeavored to address this issue by structural and biochemical studies of the ligase component of *Mycobacterium tuberculosis* LigD (MtuLigD). Bacterial LigD proteins are composed of three autonomous enzymatic modules fused in one polypeptide: an ATP-dependent DNA ligase (LigD-LIG) (8), a DNA-dependent RNA/DNA polymerase (LigD-POL) (23, 24), and a 3′ phosphoesterase (LigD-PE) (25). In league with Ku protein, bacterial LigD spearheads a low-fidelity NHEJ pathway of DNA double-strand break repair (26). Many bacterial proteomes also have LigC ligases, which are stand-alone homologs of the LigD-LIG domain (6, 27). A signature feature of LigD-LIG and LigC enzymes from *Pseudomonas aeruginosa* and *Agrobacterium tumefaciens* is their marked preference for sealing 3′-OH/5′-PO4 nicks containing a 3′-OH monoribonucleotide (29).

The structure of the MtuLigD-LIG domain, solved previously as the covalent ligase-(lysyl-Nζ)–AMP intermediate (8), resembles other ATP-dependent DNA ligases with respect to its component NTase and OB domains and the conservation of its catalytic motifs. By mutating the motif I lysine to methionine to prevent covalent adenyllylation, we now solve a 1.4 Å structure of a LigD-LIG-ATP complex and a 1.8 Å structure of a LigD-LIG-ATP-Mg complex. In both ATP-bound structures, the OB domain is in a closed conformation with respect to the NTase domain. Our results indicate that LigD-LIG employs a two-metal mechanism of lysine adenyllylation, driven by: (i) a catalytic Mg2+ that engages the ATP α phosphate and (ii) a second metal that bridges the ATP β and γ phosphates. Enzymic contacts to the ATP γ phosphate orient the PPi, leaving group favorably for in-line catalysis.

**Results**

*M. tuberculosis* LigD-LIG prefers to seal a 3′-OH monoribonucleotide nick

Many bacterial taxa (*e.g. Mycobacterium, Pseudomonas, Agrobacterium*) have a NHEJ system of DNA double-strand break (DSB) repair driven by Ku and one or two dedicated ATP-dependent DNA ligases (LigD and LigC) (30). LigD is a multi-functional enzyme composed of a ligase (LIG) domain fused to two other catalytic modules: a polymerase (POL) that preferentially adds ribonucleotides to DSB ends and a phosphoesterase that trims 3′ oligoribotracts until only a single 3′ ribonucleotide remains. LigD and LigC are conspicuously feeble at sealing 3′-OH/5′-PO4 DNA nicks *in vitro*. This property distinguishes them from LigA, the essential replicative bacterial ligase. Previous studies showed that efficient nick sealing by *Pseudomonas* and *Agrobacterium* LigD and LigC enzymes requires the presence of a single ribonucleotide at the broken 3′-OH end (29). The ribo effect on those LigD and LigC enzymes is specific for the 3′ terminal nucleotide and is either diminished or abolished when additional vicinal ribonucleotides are present. *In vitro* repair of a DSB by LigD requires the POL module and results in incorporation of an alkali-labile ribonucleotide at the repair junction (29). Those results illuminated an underlying logic for the organization of LigD, whereby POL and phosphoesterase can heal the broken 3′ end to generate the monoribo terminus favored by the NHEJ ligases.

It was of interest here to query whether the property of 3′ monoribonucleotide specificity applies to the LIG component of *M. tuberculosis* LigD. The autonomous LigD-LIG domain of the 759-amino acid MtuLigD protein extends from amino acids 452 to 759 (8). We reacted purified recombinant MtuLigD-LIG with two different nicked 36-bp duplexes (5′-32P-labeled on the nick 5′-PO4 strand) in which the unlabeled 18-mer 3′-OH strand consisted of either 18 deoxyribonucleotides (D18) or 17 deoxyribonucleotides and 1 ribonucleotide (D17R1) (Fig. 1). An enzyme titration experiment revealed that the specific activity of MtuLigD-LIG was 3.5-fold higher on the D17R1 nick (5.8 fmol nicks sealed per fmol LIG) than on the D18 nick (1.65 fmol nicks sealed per fmol LIG) (Fig. 1A). Both substrates were sealed to an equivalent extent at saturating enzyme. An analysis of the kinetics of nick sealing under conditions of enzyme excess is shown in Fig. 1B. The apparent ligation rate constant for D17R1 nick ligation (0.128 ± 0.0059 s−1) was 12-fold faster than the rate constant for D18 nick ligation (0.011 ± 0.00098 s−1). Thus, the introduction of a single 3′ ribonucleotide at the break stimulates nick sealing by MtuLigD-LIG.
DNA ligase mechanism

![Diagram of DNA ligase mechanism](image)

Table 1

Figs 3

The extent of nick sealing in plotted as a function of reaction time. Each datum is the average of three independent kinetic experiments. Reaction mixtures (10 μl) containing 50 mM Tris-HCl, pH 8.0, 5 mM MgCl₂, 0.2 mM ATP, 5 mM DTT, 1 pmol (0.1 μM) 5'-32P-labeled D17R1 or D18 nicked duplex, and increasing amounts of Ligated-D-LIG as specified were incubated at 37 °C. Aliquots (10 μl) were withdrawn at the times specified and quenched immediately with formamide/EDTA. The data were fit to a one-phase association in Prism.

Structure of MtuLigD-LIG in a complex with ATP

The structure of the MtuLigD-LIG domain, solved previously as the covalent ligase-(Lys481-N)–AMP intermediate (8), resembles prototypical ATP-dependent DNA ligases with respect to its core NTase and OB domains and the conservation of its NTase motifs (I, Ia, III, IIIa, IV, V, and VI), all of which are shared with NHEJ ligases D and C from other bacteria (Fig. 2). In the LigD-LIG–AMP structure, the NTase and OB domains are splayed wide apart in an “open” conformation that exposes the AMP phosphate on the surface of the NTase domain (Fig. 3B). An open conformation allows binding of the DNA substrate so that a reactive 5'-PO₄ DNA end is placed next to the AMP phosphate in preparation for step 2 adenylate transfer to form the AppDNA intermediate.

Our aim here was to capture a complex of MtuLigD-LIG with ATP in a state poised for catalysis of step 1 lysine adenylylation. To that end, we exploited a mutated version, K481M, in which the Lys₄⁸¹ nucleophile was replaced with methionine to preclude formation of the LIG-AMP intermediate. LIG-K481M was preincubated with 1 mM ATP and 5 mM MgCl₂ prior to crystallization by vapor diffusion against precipitant solution containing 0.1 M MES and 25% PEG 10000. Crystals diffracting to 1.4 Å resolution were in space group P2₁ with one LIG protomer in the asymmetric unit. The refined 1.4 Å structure (R_work/R_free = 0.188/0.198; Table 1) consisted of two polypeptide segments, from 453 to 652 and 659 to 759, punctuated by a disordered 6-amino acid surface loop. Electron density for ATP and MES, the buffer in the precipitant solution, was evident in the active site (Fig. 4); no density consistent with a metal ion was seen. In the ATP-bound ligase, the N-terminal NTase domain (aa 453–639) and the C-terminal OB domain (aa 640–759) adopt a “closed” conformation in which the OB domain undergoes a rigid body movement that brings it closer to the ATP phosphates (Fig. 3A). The conformational switch occurs around a pivot point at the interdomain junction (denoted by the arrow in Fig. 3A) and results in a 48 Å movement of the positions of Glu₆⁹⁳ and Asp₇⁰¹ and a 41 Å shift in the position of Arg₆⁸⁶.

Active site with ATP and MES

Fig. 5A shows a stereo view of the active site of the LIG-K₄⁸¹M-ATP-MES complex, highlighting atomic interactions with the ATP ribose and phosphates, the nearby MES ligand, and pertinent inter-NTase–OB domain contacts. The ATP ribose O2' receives a hydrogen bond from Arg₆⁸⁶ (motif I). Arg₇⁴⁵ (motif VI) makes bifurcated hydrogen bonds to the ATP α phosphate and a bidentate salt bridge to Asp₇⁰¹ (motif I). Arg₇⁴⁸ (motif VI) makes a bidentate interaction with the ATP β phosphate. Glu₇₂⁷ (a conserved residue in the OB domain; Fig. 2) forms a bifurcated salt bridge to Arg₇⁴⁵ and Arg₇⁴⁸. Lys₆³⁵ (motif V) makes bifurcated hydrogen bonds to the ATP α and γ phosphates. The γ phosphate also receives bidentate hydrogen bonds from Arg₅⁰¹ (motif Ia) and Arg₆₂⁹ (a conserved residue preceding motif V; Fig. 2) and a hydrogen bond from the ATP ribose 3'-OH. The MES morpholino ring makes van der Waals contacts with the ATP ribose and β phosphate, and the MES sulfonate oxygens receive hydrogen bonds from Arg₅⁰¹ and the Arg₄₈⁶ main-chain amide (Fig. 5A).

Structure of MtuLigD-LIG in a complex with ATP and magnesium

The unanticipated affinity of MES for the LigD-LIG active site (driven, we suspect by enzymic interactions with the MES
sulfonate) prompted us to cocrystallize LIG-K481M with ATP and magnesium under different conditions that might be more conducive to metal occupancy of the active site. LIG-K481M was preincubated with 1 mM ATP and 20 mM MgCl₂ prior to crystallization by vapor diffusion against precipitant solution containing 0.1 M ammonium fluoride and 20% PEG 3350. Crystals diffracting to 1.8 Å resolution were in space group P2₁ with one LIG protomer in the asymmetric unit. The refined model (R_work/R_free = 0.206/0.244; Table 1) included ATP and two magnesium ions in the active site. Superposition of the LIG/ATP/Mg and LIG/ATP/MES structures with respect to their NTase domains (Fig. 6A) showed that whereas LIG·ATP·Mg also adopted a closed NTase–OB conformation, the position of its OB domain was shifted in the LIG·ATP·MES structure, such that the OB domain was slightly further away from the NTase domain compared with LIG·ATP·MES. The extent of the movement was as little as 1 Å for some segments of the OB domain and as much as 5.3 Å for other OB domain elements.

**Figure 2. Bacterial NHEJ ligases.** The amino acid sequence of *M. tuberculosis* LigD ligase (MtuD) is aligned to the sequences of *P. aeruginosa* LigD ligase (PaeD), and *A. tumefaciens* LigD1 (AtuD1), LigD2 (AtuD2), LigC3 (AtuC3), and LigC2 (AtuC2). Gaps in the alignment are indicated by dashes. Positions of amino acid side chain identity/similarity in all six ligases are indicated by dots above the MtuD sequence. Nucleotidyltransferase motifs I, Ia, II, IIIa, IV, V, and VI are highlighted in blue boxes. Amino acids identified previously as essential for MtuLigD activity are shaded in blue. Conserved amino acids targeted for mutagenesis in the present study are shaded in gold.

**DNA ligase mechanism**
the “catalytic” metal ion for the lysine adenylylation reaction, directly engages one of the nonbridging α phosphate oxygens at a distance of 2.5 Å. Mg2 is a “noncatalytic” metal that bridges the ATP β and γ phosphates, at a distance of 2.6 Å to the β phosphate oxygen and 2.8 Å to the γ phosphate oxygen. Mg1 and Mg2 occupy active site positions that overlap the MES molecule in the K481M-ATP-MES complex (Fig. 5, compare A and B). Mg1 interacts, directly and via a water, with Glu11032 (motif IV). Mg2 interactions include, in addition to the β and γ phosphates oxygens, two waters at distances of 2.5 and 2.7 Å. One of these waters bridges Mg2 and Mg1 (Fig. 5B). Mg2 is situated 3.4 Å from the ATP ribose O3′.

Comparison of the ATP-Mg and ATP-MES complexes shows that there are notable changes, particularly with respect to contacts around the ATP α phosphate. Glu11032 severs its salt bridge to Arg729 (seen in the ATP-MES complex) and reorients to engage Mg1 that binds the α phosphate. The salt-bridge network involving Asp483, Arg724, Glu727, and Arg748 remains intact, but the direct contacts of Arg745 and Arg748 with the ATP α and β phosphates (seen in the ATP-MES complex) are relinquished, because of the slight movement of the OB domain away from the NTase domain. In the ATP-Mg complex, the α phosphate acquires new contacts (vis-à-vis ATP-MES), from Lys635 (motif V) and His665. The bidentate contacts of Arg486 and Arg629 with the ATP γ phosphate are retained in the ATP-Mg complex, as is the bifurcated interaction of Lys635 with the α and γ phosphates and the hydrogen bond from Arg486 to the ribose O2′.

Insights anent the mechanism of ligase adenylylation

Fig. 6B shows a laterally displaced superposition of the ATP substrate in the K481M-ATP-MES and K481M-ATP-Mg structures and the Lys481-AMP adduct in the WT LIG-AMP structure solved previously. The adenosine nucleoside is in the anti conformation in each case. As superimposed, the Lys481-Nζ is situated 3.6 Å from the ATP α phosphorus in the ATP-Mg complex and makes an Nζ–Pα–O3α angle of 137° with respect to the pyrophosphate leaving group. Consistent with a single-step in-line mechanism, we see that the α phosphate undergoes stereochemical inversion during the transition from ATP-Mg complex to lysyl-AMP intermediate (Fig. 6B). We envision that (i) the catalytic Mg1 metal ion stabilizes a pentavalent transition state of the ATP α phosphate during the lysine adenylylation reaction, abetted by the two conserved motif V lysines (Lys635 and Lys637) that contact the α phosphate oxygens, and (ii) the Mg2 metal ion engages the pyrophosphate leaving group and, with the assistance of basic side chains Arg501, Arg629, and Lys637 that coordinate the ATP γ phosphate, orients it for in-line displacement by the Lys481 nucleophile.

Structure-guided mutagenesis

Previous mutational analysis of MtuLigD had shown that alanine mutations of Asp483 (motif I), Glu530 (motif III), Glu613 (motif IV), and Lys637 (motif V) abolished DNA nick-sealing activity in vitro (8). Alanine mutation of Lys635 (motif V) reduced DNA ligation specific activity to 20% of the WT value (8). Here we conducted a new round of alanine scanning of LigD-LIG, guided by the structures of the ATP complexes, targeting the following conserved amino acids: Arg501 and Arg629 in the NTase domain that coordinate the ATP γ phosphate; and Arg745, Glu727, and Arg748 in the OB domain that form a salt-bridged network with Asp483, the arginines of which interact with the ATP phosphates in the ATP-MES complex. The recombinant LigD-LIG–Ala proteins were produced in Esche-
richia coli and purified from soluble bacterial lysates in parallel with the WT LigD-LIG as described under “Experimental procedures.” SDS-PAGE affirmed that the preparations were similarly enriched with respect to the LIG polypeptide (Fig. 7A). The WT LigD-LIG and LigD-LIG–Ala mutants were titrated for ligation of the D17R1 nick substrate (Fig. 7B). Specific activities were derived by linear regression fitting of the data and normalized to the WT (defined as 100%). The R501A and R629A proteins were 1.2 and 7.2% as active as WT, respectively. We surmise that the Arg501 interaction with the ATP γ phosphate is especially crucial for ligase activity. The E727A, R745A, and R748A mutants were 0.8, 4.3, and 6.2% as active as WT, respectively. It appears that the loss of the Glu727 salt bridge from Arg745 to Arg748 is more deleterious than subtraction of either of the two individual arginine side chains.

**Mutational effects on sealing at a preadenylylated nick**

WT LigD-LIG and the LigD-LIG–Ala mutants were assayed by enzyme titration for their ability to seal a preadenylylated D17R1 AppDNA nicked substrate in the absence of ATP (Fig. 8), the rationale being that mutational defects in either step 1 or step 2 of the ligase pathway might be bypassed by presenting the mutant enzyme with a preadenylylated nick. The 32P-labeled AppDNA strand was prepared reaction of the 5′ 32P-labeled 18-mer pDNA oligonucleotide with *E. coli* RtcA and ATP (31). The AppDNA strand was gel-purified and annealed with the D17R1 3′-OH strand and complementary 36-mer DNA template strand to form the substrate shown in Fig. 8. AppDNA ligation specific activities were derived by linear regression fitting of the data and normalized to the WT (defined as 100%) as follows: R501A (130%), R629A (140%), E727A (170%), R745A (100%), and R748A (72%). These results signify that Arg501, Arg629, Glu727, Arg745, and Arg748 are not critical for the isolated step 3 reaction of the LigD sealing pathway.

**Figure 4. Electron density for ATP and MES.** A stereo view of the simulated annealing omit Fo–Fc map of the ATP and MES ligands in the active site, contoured at 2σ, is shown in blue mesh. The density map was calculated with both ligands omitted from the model. ATP and MES are depicted as a stick models with beige and gray carbons, respectively.

**Figure 5. Active sites in the LIG-ATP-MES and LIG-ATP-Mg complexes.** A, stereo view of the active site of the LIG-ATP-MES complex. Amino acids and MES are shown as stick models with gray carbons. ATP is depicted as a stick model with beige carbons. Atomic interactions are indicated by black dashed lines. B, stereo view of the active site of the LIG-ATP-Mg complex, in the same orientation as in A. Amino acids and ATP are shown as stick models with gray and beige carbons, respectively. Mg2+ ions and associated waters are depicted as magenta and red spheres, respectively. Atomic contacts are indicated by dashed lines.

**Figure 6. Comparison of LigD-LIG structures.** A, stereo view of the LigD-LIG-ATP-MES structure (blue NTase domain and magenta OB domain) superimposed with respect to its NTase domain on the LigD-ATP-Mg structure (both domains colored beige) to highlight the slight movement of the OB domain. ATPs are rendered as stick models. B, the adenosine nucleotides and amino acid 481 of the LigD-ATP-MES, LigD-ATP-Mg, and LigD-AMP structures were superimposed and then offset horizontally.
DNA ligase mechanism

The covalent lysine nucleotidyltransferase superfamily embraces ATP-dependent DNA ligases, ATP-dependent RNA ligases, NAD\(^+\)-dependent DNA ligases, and GTP-dependent mRNA capping enzymes (1). The superfamily is defined by (i) the formation of a lysyl-NMP intermediate en route to the nucleotidylation of nucleic acid 5’ ends and (ii) a conserved NTase catalytic domain and a set of NTase motifs that form the NMP-binding pocket. Six distinct families of ATP-dependent RNA ligase (Rnl) have been characterized, each consisting of an NTase domain fused to a structurally unique flanking domain that defines the Rnl family (32–39). By contrast, ATP-dependent DNA ligases, NAD\(^+\)-dependent DNA ligases, and GTP-dependent mRNA capping enzymes share a core structure composed of a proximal NTase domain fused to a distal OB-fold domain (1). Crystal structures of DNA ligases and capping enzymes have revealed considerable variability in the position of the OB domain relative to the NTase domain. In the few cases where structures of the same enzyme have been captured at different stages along the reaction pathway, it seems that OB domain movements occur in sync, or at least correlate, with substrate binding and product release (1, 4, 5, 9, 10, 13–15, 22, 33). In general terms, it is thought that the DNA ligase OB domain is in an open conformation to allow ATP binding, undergoes domain closure for catalysis of lysine adenyllylation, and then reopens to allow release of the PPi product and exposure of the DNA nick-binding site overlying the lysyl-AMP phosphate. DNA binding is coupled to reorientation of the OB domain to form a C-shaped clamp around the nick. After step 3 phosphodiester synthesis, the clamp opens to allow release of the ligated DNA and AMP products.

The suite of available structures of LigD-LIG now includes two new complexes with ATP, both with the OB domain in a closed conformation, and a prior structure of the LIG-AMP covalent intermediate, wherein the OB domain is splayed widely open (8). It had been proposed previously, based on the structure of Chlorella virus capping enzyme in a reactive complex with GTP (22), that a consequence of OB domain closure upon NTP binding would be to bring NTase motif VI (located in the OB domain; Fig. 2) into the active site for lysine adenyllylation, wherein motif VI basic amino acids would engage the NTP/AMP and /H\(9253\) phosphates and orient the PPi leaving group for in-line catalysis. In the capping enzyme\(GTP\) complex, the /H\(9253\) phosphate is additionally coordinated by three amino acids in the NTase domain: Arg106 (in motif Ia), Lys234 (motif V), and Arg228 (located 6 amino acids upstream of the first motif V Lys) (Fig. 9A). A similar network of contacts to the ATP/AMP is seen presently in the LigD-LIG\(ATP\) complex and in the
Recent studies have highlighted a conserved mechanism of metal-dependent lysine adenylylation by ATP-dependent RNA ligases and NAD⁺-dependent DNA ligases, whereby an octahedral catalytic metal-(H₂O)₅ complex stabilizes the transition state of the ATP or NAD⁺ α phosphate (32, 33, 35, 38, 39). Three side chains of NTase motifs I, III, and IV bind the pentahydrated metal cofactor via waters. Structures of the Michaelis complex of RNA ligases from three different families reveal a two-metal mechanism of lysine adenylylation whereby a second octahedral noncatalytic metal complex bridges the ATP β and γ phosphates and orients the PPᵢ leaving group for in-line attack by the lysine nucleophile (32,33,39). By contrast, NAD⁺-dependent DNA ligase employs a one-metal mechanism of lysine adenylylation and relies on direct enzymic contacts to orient the nicotinamide mononucleotide leaving group (33).

An account of the metal requirement for lysine adenylylation by ATP-dependent DNA ligases was lacking. The present study of LigD-LIG points to a two-metal mechanism driven by a catalytic magnesium in the ATP α phosphate and a noncatalytic magnesium bridging the β and γ phosphates. Although the roles imputed to the two metals in the LigD-LIG-ATP structure are analogous to those of the ATP-dependent RNA ligases and although the positions of the motif I, motif III, and motif IV side chains that bind the catalytic metal in the Rnl5 family Michaelis complex (32) are virtually identical to the corresponding Asp⁴⁸³, Glu⁵₃₀, and Glu⁶¹₃ residues in the LigD-LIG-ATP-Mg complex, the specific ligD-LIG metal coordination complexes are different. The LigD catalytic Mg₁ coordination complex has only four occupants. The direct Mg₁ contact to the ATP α phosphate and a water-mediated contact to Glu⁶¹₃ are features shared with other ligases. The second Mg₁ contact to Glu⁶¹₃ is direct, rather than water-mediated as in RNA ligases and NAD⁺-dependent DNA ligase. The motif I Asp⁴⁸³ is absent from the LigD Mg₁ coordination complex, being instead engaged in a salt bridge to Arg⁷⁴⁵. Motif III Glu⁵₃₀ is also absent. These differences raise two possible scenarios: (i) the catalytic metal in LigD-LIG is engaged to the enzyme in a manner distinct from that employed by RNA ligases and NAD⁺-dependent DNA ligase; or (ii) the structure of the LigD-Mg₁-Mg complex, although solved at high resolution, approximates but does not fully reflect the state of the step 1 Michaelis complex.

In favor of the latter scenario is the structure of Pyrococcus furiosus DNA ligase (PfuLig) in a noncovalent complex with AMP, with a single magnesium ion in the active site and the OB domain in a closed conformation over the NTase domain (18). In the PfuLig-AMP structure, a hexahydrated magnesium complex with octahedral geometry is engaged to the enzyme via water-mediated contacts to motif I, motif III, and motif IV carboxylate side chains in a manner identical to that of the catalytic metal in the RNA ligase and NAD⁺-dependent DNA ligase.
DNA ligase mechanism

structures, the key difference being that the AMP phosphate in the PhuLig-AMP complex is not part of the metal coordination complex (its place being taken by a sixth water molecule).

Experimental procedures

Recombinant MtuLigD-LIG–K481M

The ORF encoding the LigD-LIG domain (aa 452–759) was PCR-amplified from M. tuberculosis genomic DNA with primers that introduced a BamHI site 5’ of the start codon and an XhoI site 3’ of the stop codon. The BamHI–XhoI fragment was inserted into pET28b–His10Smt3 to generate pET28b–His10Smt3–MtuLigD-LIG plasmid. The resulting expression plasmid encodes the MtuLigD-LIG polypeptide fused to an N-terminal His10Smt3 tag under the transcriptional control of a T7 RNA polymerase promoter. A K481M coding change was introduced into MtuLigD-LIG by PCR using mutagenic primers. The expression plasmid inserts were sequenced completely to exclude the acquisition of unwanted changes during PCR amplification and cloning. The pET28b–His10Smt3–MtuLigD-LIG–K481M plasmid was introduced into E. coli BL21(DE3) cells. A 2-liter culture was grown at 37 °C in Luria–Bertani medium containing 0.1 mg/ml kanamycin until the A600 reached 0.6. The culture was then adjusted to 0.5 mM isopropyl 1-thio-β-D-galactopyranoside, followed by incubation for 17 h at 17 °C with constant shaking. The cells were harvested by centrifugation and stored at −80 °C. All subsequent procedures were performed at 4 °C. The bacteria were suspended in 50 ml of buffer containing 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10% sucrose and then lysed by sonication for 10 min. Insoluble material was removed by centrifugation for 1 h at 16,000 rpm. The supernatant was applied to a 10-ml nickel–nitrilotriacetic acid–agarose column (Qiagen) that had been equilibrated with buffer A (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10% glycerol). The column was washed with buffer A, and bound material was eluted stepwise with 20, 100, 300, and 1000 mM imidazole in buffer A. The polypeptide compositions of the eluate fractions were monitored by SDS-PAGE. The His10Smt3–MtuLigD-LIG–K481M protein was recovered predominantly in the 300 mM imidazole eluate fractions, which were pooled, supplemented with Smt3-specific protease Ulp1 (at a MtuLigD-LIG:Ulp1 ratio of ~500:1), and then dialyzed overnight against buffer A with 20 mM imidazole. The LigD-LIG–K481M protein was separated from the cleaved His10Smt3 tag by passage over a second nickel–agarose column equilibrated in buffer A with 20 mM imidazole, whereby the tag-free LigD-LIG was recovered in the flow-through fraction. LigD-LIG–K481M was then concentrated by centrifugal ultrafiltration and further purified by gel filtration through a column of Superdex-200 equilibrated in 10 mM Tris HCl, pH 8.0, 100 mM NaCl. The peak fractions of LigD-LIG–K481M (which eluted as a monomer) were pooled, concentrated by centrifugal ultrafiltration, and stored at −80 °C.

Crystalization of LIG–K481M

A solution of LIG–K481M (10 mg/ml) was adjusted to 1 mM ATP and 5 mM MgCl2 and incubated for 15 min on ice before aliquots (1 μl) were mixed on a coverslip with an equal volume of precipitant solution containing 0.1 M MES, pH 6.5, 25% PEG 10000. Crystals were grown at room temperature by hanging-drop vapor diffusion against a reservoir of the same precipitant solution. Single crystals appearing after 2–3 days were harvested; cryoprotected by transfer to 0.1 M MES, pH 6.5, 1 mM ATP, 5 mM MgCl2, 25% PEG 10000, 15% PEG 400; and then flash-frozen in liquid nitrogen.

Alternatively, a solution of LIG–K481M (10 mg/ml) was adjusted to 1 mM ATP and 20 mM MgCl2, and incubated for 15 min on ice before aliquots (1 μl) were mixed on a coverslip with an equal volume of precipitant solution containing 0.1 M ammonium fluoride, 20% PEG 3350. Crystals were grown at room temperature by hanging-drop vapor diffusion against a reservoir of the same precipitant solution. Single crystals appearing after 2–3 days were harvested; cryoprotected by transfer to 0.1 M ammonium fluoride, 20% PEG 3350, 1 mM ATP, 20 mM MgCl2, 15% glycerol; and then flash-frozen in liquid nitrogen.

Diffraction data collection and structure determination

X-ray diffraction data were collected from single crystals at the Advanced Photon Source Beamline 24ID-C. Indexing and merging of the diffraction data were performed in HKL2000 (41). The phases were obtained by molecular replacement in MOLREP (42) using the NTase and OB domains of PDB entry 1VS0 separately as the search models. Interactive model building was performed in O (43). Refinement was accomplished with PHENIX (28). Data collection and refinement statistics are summarized in Table 1. The model of LIG–K481M-ATP-MES was refined to 1.4 Å resolution (Rwork/Rfree = 0.188/0.198) and the model of LIG–K481M-ATP-Mg to 1.8 Å resolution (Rwork/Rfree = 0.206/0.244). Both models comprised one LIG protomer in the asymmetric unit.

Accession numbers

Structure coordinates have been deposited in the Protein Data Bank database under accession codes 6NHX and 6NHZ.

Structure-guided mutagenesis

Alanine mutations were introduced into the LIG ORF of pET28b–His10Smt3–MtuLigD-LIG by PCR using mutagenic primers. The plasmid inserts were sequenced completely to exclude the acquisition of unwanted changes during PCR amplification and cloning. WT LigD-LIG and mutant LigD-LIG–Ala proteins were produced in E. coli and purified through the second nickel–agarose step as described above for LigD-LIG–K431M. Protein concentrations were determined by using the Bio-Rad dye reagent with BSA as the standard.

Author contributions—M.-C. U. and S. S. conceptualization; M.-C. U. and Y. G. data curation; M.-C. U., Y. G., and S. S. formal analysis; M.-C. U., Y. G., and S. S. investigation; M.-C. U. and Y. G. methodology; M.-C. U., Y. G., and S. S. writing-review and editing; S. S. supervision; S. S. funding acquisition; S. S. writing-original draft.

Acknowledgments—X-ray diffraction data were collected at synchrotron facilities supported by NIGMS, National Institutes of Health Grants P41GM103403, P41GM103473, and P41GM111244; ORIP, National Institutes of Health Grant HEI S10RR029205; and Department of Energy Contracts DE-AC02-06CH11357, DE-AC02-98CH110886, and E-SC0012704.
**DNA ligase mechanism**

References

1. Shuman, S., and Lima, C. D. (2004) The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidytransferases. *Curr. Opin. Struct. Biol.* **14**, 757–764 CrossRef Medline

2. Pascal, J. M. (2008) DNA and RNA ligases: structural variations and shared mechanisms. *Curr. Opin. Struct. Biol.* **18**, 96–105 CrossRef Medline

3. Subramanya, H. S., Doherty, A. J., Ashford, S. R., and Wigley, D. B. (1996) Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. *Cell* **85**, 607–615 CrossRef Medline

4. Odel, M., Sriskanda, V., Shuman, S., and Nikolov, D. (2000) Crystal structure of eukaryotic DNA ligase–adenylate illuminates the mechanism of nick sensing and strand joining. *Mol. Cell* **6**, 1183–1193 CrossRef Medline

5. Nair, P. A., Nadakumar, J., Smith, P., Odel, M., Lima, C. D., and Shuman, S. (2007) Structural basis for nick recognition by a minimal pluriportent DNA ligase. *Nat. Struct. Mol. Biol.* **14**, 770–778 CrossRef Medline

6. Gong, C., Martins, A., Bongiorno, P., Glickman, M., and Shuman, S. (2004) Biochemical and genetic analysis of the four DNA ligases of mycobacteria. *J. Biol. Chem.* **279**, 20594–20606 CrossRef Medline

7. Gong, C., Bongiorno, P., Martins, A., Stephanou, N. C., Zhu, H., Shuman, S., and Glickman, M. S. (2005) Mechanism of non-homologous end joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat. Struct. Mol. Biol.* **12**, 304–312 CrossRef Medline

8. Akey, D., Martins, A., Aniukwu, J., Glickman, M. S., Shuman, S., and Berger, J. M. (2006) Crystal structure and nonhomologous end joining function of the ligase component of *Mycobacterium* DNA ligase D. *J. Biol. Chem.** **281**, 13412–13423 CrossRef Medline

9. Williamson, A., Rothweiler, U., and Leiros, H. K. (2014) Enzyme-adenylate structure of a bacterial ATP-dependent DNA ligase with a minimized DNA-binding surface. *Acta Crystallogr. D* **70**, 3043–3056 CrossRef Medline

10. Williamson, A., Grigic, M., and Leiros, H. S. (2018) DNA binding with a minimal scaffold: structure-function analysis of Lig E DNA ligases. *Nucleic Acids Res.* **46**, 8616–8629 CrossRef Medline

11. Pascal, J. M., O’Brien, P. J., Tomkinson, A. E., and Ellenberger, T. (2004) Human DNA ligase I completely encircles and partially unwinds nicked DNA. *Nature* **432**, 473–478 CrossRef Medline

12. Cotner-Gohara, E., Kim, I. K., Hammel, M., Tainer, J. A., Tomkinson, A. E., and Ellenberger, T. (2010) Human DNA ligase III recognizes DNA ends by dynamic switching between two DNA-bound states, *Biochemistry* **49**, 6165–6176 CrossRef Medline

13. Ochi, T., Gu, X., and Blundell, T. L. (2013) Structure of the catalytic region of DNA ligase IV in complex with an Artemis fragment sheds light on double-strand break repair. *Structure* **21**, 672–679 CrossRef Medline

14. Kaminski, A. M., Tumbale, P. P., Schellenberg, M. J., Williams, R. S., Williams, J. G., Kunkel, T. A., Pedersen, L. C., and Bebenek, K. (2018) Structures of DNA-bound human ligase IV catalytic core reveal insights into substrate binding and catalysis. *Nat. Commun.* **9**, 2642 CrossRef Medline

15. Nadakumar, 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 CrossRef Medline

16. Shi, K., Bohl, T. E., Park, J., Zasada, A., Malik, S., Banerjee, S., Tran, V., Li, N., Yin, Z., Kurniawan, F., Orellana, K., and Aihara, H. (2018) T4 DNA ligase structure reveals a prototypical ATP-dependent ligase with a unique mode of sliding clamp interaction. *Nucleic Acids Res.* **46**, 10474–10488 CrossRef Medline

17. 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 CrossRef Medline

18. Nishida, H., Kiyonori, S., Ishino, M., and Ishikawa, K. (2006) The closed structure of an archaenal DNA ligase from *Pyrococcus furiosus*. *J. Mol. Biol.* **360**, 956–967 CrossRef Medline

19. Petrova, T., Bezsudnova, E. Y., Boyko, K. M., Mardanov, A. V., Polyakov, K. M., Volkov, V. V., Kozin, M., Ravin, N. V., Shabanil, I. G., Skrynin, K. G., Stekhanova, T. N., Kovalchuk, M. V., and Popov, V. O. (2012) ATP-dependent DNA ligase from *Thermococcus* sp. 1519 displays a new ar-

20. Wang, K. L., Nair, P. A., and Shuman, S. (2012) The adenylyltransferase domain of bacterial Pinp defines a unique RNA ligase family. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2296–2301 CrossRef Medline
DNA ligase mechanism

38. Wang, P., Chan, C. M., Christensen, D., Zhang, C., Selvadurai, K., and Huang, R. H. (2012) Molecular basis of bacterial protein Hen1 activating the ligase activity of bacterial protein Pnkp for RNA repair. Proc. Natl. Acad. Sci. U.S.A. 109, 13248–13253 CrossRef Medline

39. Banerjee, A., Ghosh, S., Goldgur, Y., and Shuman, S. (2018) Structure and two-metal mechanism of fungal tRNA ligase. Nucleic Acids Res., 10.1093/nar/gky1275 CrossRef Medline

40. Sawaya, R., and Shuman, S. (2003) Mutational analysis of the guanylyltransferase component of mammalian mRNA capping enzyme. Biochemistry 42, 8240–8249 CrossRef Medline

41. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 CrossRef Medline

42. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022–1025 CrossRef

43. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A 47, 110–119 CrossRef Medline