DNA Ligases: Structure, Reaction Mechanism, and Function

Alan E. Tomkinson,*,† Sangeetha Vijayakumar,‡ John M. Pascal,§ and Tom Ellenberger§

Radiation Oncology Research Laboratory and Marlene and Stewart Greenebaum Cancer Center, Molecular and Cellular Biology Graduate Program, University of Maryland School of Medicine, Baltimore, Maryland 21201, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received April 26, 2005

Contents

1. Introduction

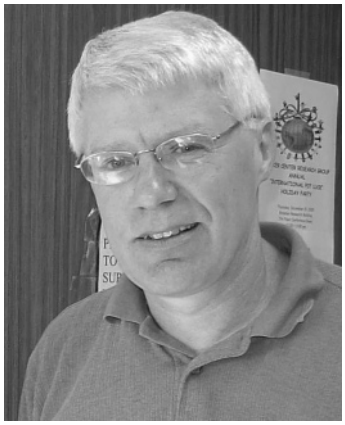
The joining of interruptions in the phosphodiester backbone of duplex DNA is critical for maintaining genomic integrity. DNA strand-breaks are produced as a consequence of normal DNA transactions in undamaged cells. For example, during replication of the human genome, approximately 2×10^7 Okazaki fragments are generated by discontinuous lagging strand DNA synthesis[1] and programmed site-specific DNA double-strand breaks are made in certain cell types such as immune cells undergoing immunoglobulin gene rearrangements.[2] In addition, DNA strand breaks can be generated either as a consequence of DNA lesion removal by excision repair pathways or directly by a DNA damaging agent. Approximately 2×10^5 DNA single-strand breaks are generated per mammalian cell per day by the processing of spontaneous DNA lesions such as abasic sites and oxidized DNA bases,[3] whereas exposure to 1 Gy of ionizing radiation causes 600−1000 single-strand breaks and 16−40 double-strand breaks.[4]

The enzymes that catalyze phosphodiester bond formation, DNA ligases, were discovered by several laboratories in 1967 and 1968 (reviewed in refs 5−7). In accord with their participation in multiple different DNA metabolic pathways, genetic inactivation of DNA ligases causes pleiotropic phenotypes including lethality and hypersensitivity to DNA damaging agents.[6,8−11] DNA ligases together with RNA ligases and mRNA capping enzymes constitute the nucleotidyl transferase superfamily.[12] Specifically, these enzymes interact with a nucleotide cofactor to form a covalent enzyme−nucleoside monophosphate. DNA ligases utilize either ATP or NAD^+ as the nucleotide cofactor. During the first of three chemical steps of ligation, a phosphoamide bond (P−N) forms between the α-amino group of an active site lysine and the 5′ phosphate of AMP for DNA and RNA ligases, or GMP for mRNA capping enzymes. The activated enzyme−NMP adduct is generated in the absence of a nucleic acid substrate. Inorganic pyrophosphate (PP_i) is released during step 1 by enzymes utilizing nucleoside triphosphates, whereas nicotinamide mononucleotide (NMN) is released by NAD^+-dependent ligases. In the second step of the reaction, the 5′ phosphate group of NMP is transferred from the active site lysine to a phosphorylated DNA 5′ end, forming a pyrophosphate linkage (5′P−5′P). The 5′ AMP activates the 5′ phosphate of a DNA substrate for phosphodiester bond formation. During step 3, the 3′ hydroxyl of an adjacent DNA strand attacks the 5′ phosphorylated DNA end to displace AMP and covalently join the DNA strands. In the first part of this review, we will discuss the reaction mechanism of DNA ligases in the context of the structural information obtained in recent years for bacterial, viral, and human DNA ligases.

Biochemical and molecular genetic studies with mammalian cells provided the first evidence that cells contain more than one species of DNA ligase.[13−16] Subsequent studies have revealed the existence of more than one DNA ligase in lower eukaryotic and prokaryotic organisms.[17−19] The presence of multiple DNA ligases suggested that these enzymes may have specific cellular functions. This notion was supported by the different phenotypes of mammalian cell lines deficient in either DNA ligase I, DNA ligase III, or DNA ligase IV.[11,20−22] A comparison of the polypeptides encoded by the mammalian *LIG1*, *LIG3*, and *LIG4* genes shows that these enzymes contain a conserved catalytic

* Corresponding author. Address: Radiation Oncology Research Laboratory and Greenebaum Cancer Center, Bressler Research Building BRB7-025, 655 West Baltimore Street, University of Maryland School of Medicine, Baltimore, MD 21201-1509. Telephone no.: 410-706-2795. Fax no.: 410-706-3000. E-mail: atomkinson@som.umaryland.edu.
† Radiation Oncology Research Laboratory and Greenebaum Cancer Center, University of Maryland School of Medicine.
‡ Molecular and Cellular Biology Graduate Program, University of Maryland School of Medicine.
§ Harvard Medical School.

10.1021/cr040498d CCC: $59.00
Published on Web 01/24/2006



Alan Tomkinson received his Ph.D. from the University of Newcastle upon Tyne, U.K., in 1983 before postdoctoral training at the University of California, Berkeley, and the Imperial Cancer Research Fund, U.K. His research focuses on the function of eukaryotic DNA ligases in DNA replication, DNA repair, and genetic recombination. He is a Professor in the Department of Radiation Oncology and Associate Director for Basic Research in the Marlene and Stewart Greenebaum Cancer Center at the University of Maryland School of Medicine.

Sangeetha Vijayakumar is a recent graduate from the University of Maryland School of Medicine, Molecular Medicine Ph.D program. She did her thesis on the interplay among DNA ligase I, RFC, and PCNA and its implications in Okazaki fragment processing under the mentorship of Dr. Alan Tomkinson.

domain flanked by unique sequences that presumably confer biological functions (Figure 1). In the last part of this review, we will describe protein partners of the DNA ligases and describe how these protein−protein interactions may coordinate the final ligation step with the earlier steps of DNA metabolic pathways.

2. Insights into the DNA Ligation Reaction Mechanism from the Structures of Nucleotidyl Transferases

2.1. Structures of the Catalytic Domain of ATP-Dependent Viral and Bacteriophage DNA Ligases

The simplest DNA ligases, encoded by viruses and bacteriophages, have a two-domain organization that is exemplified by the crystal structures of ATP-dependent DNA ligases from bacteriophage T7[23] and a *Chlorella spp*. virus.[24] The adjacent nucleotide-binding and OB-fold domains of these enzymes (Figure 2A) are structurally analogous to the catalytic core of more complex multidomain DNA ligases

John Pascal earned his B.S. in biochemistry from Clemson University (1995) and his Ph.D. in biochemistry from the University of Texas at Austin (2000). Under the mentorship of Jon D. Robertus at UT Austin, he determined crystal structures of RNA-binding and -modifying enzymes. In 2002, he moved to Harvard Medical School to work with Tom Ellenberger on the biochemical and structural analysis of enzymes involved in DNA replication, repair, and recombination.

Tom Ellenberger completed his doctorate in veterinary medicine in 1983 at Iowa State University (Ames, IA) before pursuing graduate studies at Harvard Medical School. His Ph.D. thesis work on mechanisms of gene amplification and multiple drug resistance in the protozoan parasite *Leishmania* marked the beginning of a long fascination with the enzymology of DNA metabolism. After postdoctoral studies at Harvard University (1989–1993) studying the structural basis of sequence-specific recognition of DNA by proteins, Dr. Ellenberger joined the faculty of Harvard Medical School in 1993 and was promoted to the rank of full Professor in 2000. As the Hsien Wu and Daisy Yen Wu Professor of Biochemistry at Harvard Medical School, Dr. Ellenberger's work focused on the structural enzymology of replication, DNA repair, and site-specific recombination. He has recently moved to Washington University School of Medicine (St. Louis, MO), where he is the Raymond H. Wittcoff Professor and Head of the Department of Biochemistry and Molecular Biophysics (http://ellenberger.wustl.edu).

found in bacteria[25,26] (Figure 2B) and higher eukaryotes[27] (Figure 2C), as well as the mRNA capping enzymes[28] (Figure 3A). The ATP-binding pocket within the amino-terminal domain of T7 and *Chlorella* virus ligases is formed by two antiparallel β-sheets flanked by α-helices (Figures 2A and 4A). Conserved residues of the nucleotidyl transferase family[12] line the binding pocket, where they participate in hydrogen bonding interactions with the ribose sugar hydroxyls of the adenylate cofactor. A flexible tether connects the nucleotide-binding domain to the carboxyl-terminal domain, which has the signature 5-stranded β-barrel structure of an oligosaccharide/oligonucleotide-binding fold (OB-fold).[29] The interdomain linker bears several residues that are conserved in the nucleotidyl transferase superfamily and

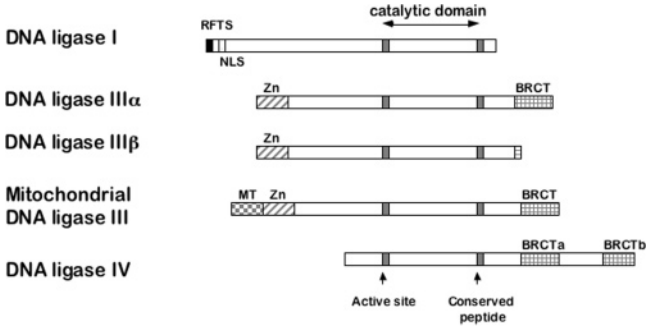


Figure 1. Schematic representation of human DNA ligases. The conserved catalytic domain of nucleotidyl transferases is located between motif 1 (active site, gray box) and motif VI (conserved peptide, gray box).[12] The N-terminal replication foci targeting (RFTS) and nuclear localization signals (NLS) of DNA ligase I are shown as the black and white boxes, respectively. Polypeptides encoded by the *LIG3* gene with different N- and C-terminal sequences are shown as DNA lig III α, DNA lig III *β*, and mitochondrial DNA lig III. The zinc finger motif located at the N-termini of DNA ligase III α and *β* is represented by diagonal stripes. The BRCT domain located at the C-terminus of DNA ligase III α is shown by grids. The unique C-terminus of DNA lig III *β* is shown by horizontal lines. The mitochondrial targeting signal (MT) is located at the N-terminus of mitochondrial DNA ligase III. The two C-terminal BRCT motifs of DNA ligase IV, BRCT a and BRCT b, are also shown by grids.

play a role in positioning these domains on DNA, as described below for the structure of human DNA ligase I.

2.2. Structures of NAD$^+$-Dependent Bacterial DNA Ligases

Crystal structures of several NAD$^+$-dependent bacterial ligases have been reported.[25,26,30] The structure of *Thermus filiformis* DNA ligase revealed a modular architecture consisting of a core [nucleotide-binding−OB-fold] module linked at its C-terminus to a zinc finger domain, a helix−hairpin−helix (HhH) domain, and a BRCA1 C-terminal repeat (BRCT) domain (Figure 2B).[25] Biochemical studies indicate that the HhH domain supports DNA-binding activity,[31] consistent with a proposal that the *T. filiformis* ligase might encircle DNA using the adenylation, OB-fold, and HhH domains to engage a substrate.[25] The function of the BRCT domain remains to be determined.

Two crystal structures of the structurally similar NAD$^+$-dependent DNA ligase from *Enterococcus faecalis* (Figure 3B) revealed an interesting conformational change in complexes with the substrate (NAD$^+$) or product (NMN) of the step 1 adenylation reaction.[30] Remarkably, the crystallized NMN complex could be converted to the NAD$^+$ complex (a reversal of step 1) by transient exposure of crystals to AMP in the absence of divalent metals, with an ensuing large-scale movement of two N-terminal helices of domain 1a flanking the adenylation domain that contain the binding pocket for the NMN moiety (Figure 3B, right panel). In the closed conformation with the NAD$^+$ substrate, the NMN-binding pocket is adjacent to the AMP-binding site (Figure 3B, left panel). Following the transfer of AMP from NAD$^+$ to an active site lysine, the NMN product bound to domain 1a rotates away from the active site, presumably to make room for a DNA substrate (Figure 3B, right panel). *T. filiformis* and *E. faecalis* DNA ligases have similar structures (Figures 2B and 3B), and it is likely this structural transition occurs in all bacterial DNA ligases.

2.3. Conformational Changes of Nucleotidyl Transferases Linked to Reaction Chemistry

Despite the similar core structures of NAD$^+$-dependent and NTP-dependent ligases and related capping enzymes, these enzymes employ different conformational states during the course of the reactions they catalyze. The ATP-dependent DNA ligases and GTP-dependent mRNA capping enzymes lack domain Ia of the NMN-binding site, and instead couple step 1 to step 2 by a rotation of the OB-fold with respect to the nucleotide-binding domain. This conformational transition was beautifully illustrated by crystal structures of the *Chlorella* virus mRNA capping enzyme.[28] Two molecules of the capping enzyme crystallized in different conformations, one in an inactive state (Figure 3A, left panel) and the other poised for step 1 chemistry (Figure 3A, right panel). In the active conformation, residues from motif VI on the surface of the OB-fold domain reach into the active site pocket of the nucleotide-binding domain **(**Figure 3A, right panel**)**. Motif VI residues orient the α- and *β*-phosphates of the GTP cofactor for an in-line attack of the active site lysine, forming the GMP−lysine intermediate. Motif VI residues are required for the step 1 reaction catalyzed by the *Chlorella* virus capping enzyme[32] as well as the *Chlorella* ATP-dependent ligase.[33] In the inactive form of the capping enzyme, crystal contacts prohibit the OB-fold domain from swiveling into the active conformation (Figure 3A, left panel). Step 2 chemistry requires a large-scale rotation of the OB-fold domain, consistent with biochemical studies[34,35] and illustrated by the crystal structure of the ATP-dependent human DNA ligase I in complex with a DNA substrate described below.

The crystal structure of human DNA ligase I bound to nicked, 5′ adenylated DNA offers the first view of a DNA ligase complexed to nucleic acid[27] (Figure 2C). In addition to the nucleotide-binding and OB-fold domains described above, DNA ligase I has another DNA-binding domain immediately N-terminal to the catalytic core (Figure 2C). The roles of these three domains in determining the catalytic efficiency and selectivity of human DNA ligase I are discussed below. Here, we focus on the reorientation of the OB-fold domain that is required for steps 2 and 3 of DNA nick joining. The OB-fold domains of human DNA ligase I and the *Chlorella* mRNA capping enzyme are structurally homologous, particularly around motif VI, strongly suggesting that the OB-fold domain is oriented similarly in both enzymes during step 1. However, the human DNA ligase I−DNA complex crystallized following the transfer of AMP to the 5′ phosphate end of DNA (step 2) reveals a strikingly different orientation of the OB-fold domain, with motif VI facing away from the active site (Figure 4B). In this orientation, the DNA-binding surface of the OB-fold points toward DNA bound in the active site of the nucleotide-binding domain. Arg871 from the OB-fold forms a salt bridge with Asp570 (motif I) of the nucleotide-binding domain, securing the OB-fold in this orientation, and Phe872 (OB-fold) contacts the ribose sugar at the 5′ end of the nick (Figure 5). Because the nucleotide-binding (motif VI) and DNA-binding residues are located on opposite surfaces of the OB-fold domain, a large-scale rotation of this domain is required during the transition from step 1 to step 2. The DNA nick joining reaction (step 3) is catalyzed by DNA ligase crystallized in the same conformation as the step 2 complex (J.M.P. and T.E., unpublished results). Thus, we conclude that swiveling of the OB-fold domain (Figure 3A) is an

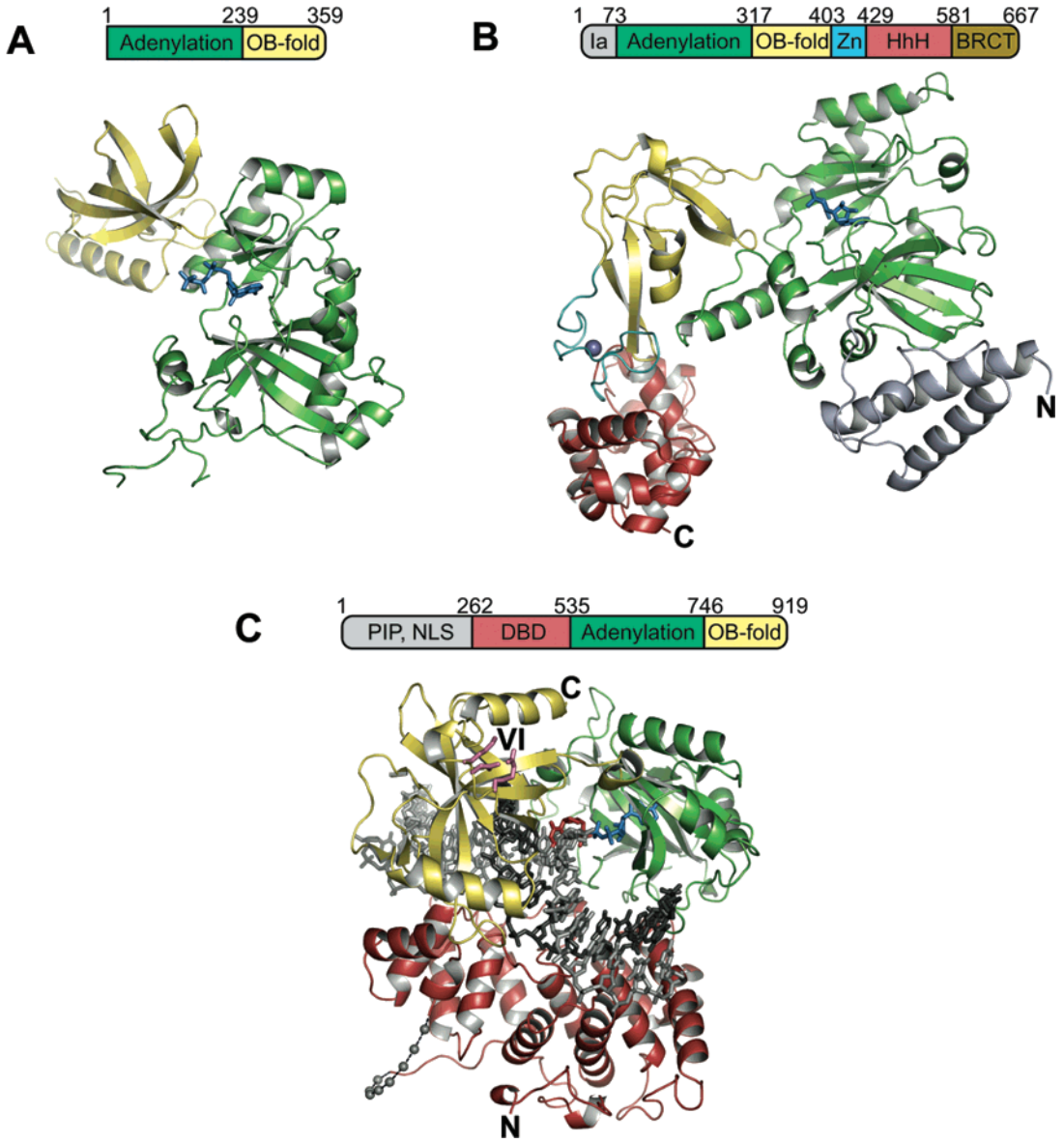


Figure 2. ATP- and NAD$^+$-dependent DNA ligases. (A) The ATP-dependent DNA ligase from bacteriophage T7 (PDB code 1a0i)[23] is a two-domain ligase: the adenylation or nucleotide-binding domain (green) binds ATP and is connected to an OB-fold domain (yellow) by a flexible linker. (B) The NAD$^+$-dependent DNA ligase from *Thermus filiformis* (PDB code 1dgs)[25] is a multidomain ligase. The basic folds for the nucleotide-binding domain and the OB-fold domain are similar to that found in T7 DNA ligase. Additionally, a zinc finger domain, a helix−hairpin−helix domain, and a BRCT domain (not in the structure) extend from the C-terminus of the OB-fold domain. Domain Ia (gray), which assists in the step 1 reaction, is N-terminal to the nucleotide-binding domain and is unique to the bacterial ligases. (C) Human DNA ligase I (PDB code 1x9n)[27] encircles nicked, adenylated DNA. The N-terminal region of DNA ligase I is missing from the crystal structure and contains the PCNA interacting peptide (PIP) and a nuclear localization signal (NLS). Three domains, DNA binding (DBD), adenylation, and OB-fold, contact and unwind the nicked DNA, exposing the nick termini to the adenylation domain. The nick termini are highlighted by a red 3′ hydroxyl nucleotide and a blue 5′ adenylated phosphate. Residues involved in the step 1 reaction (pink, motif VI) are distant from the active site in this conformation.

obligate step in DNA nick joining,[36] occurring after the DNA ligase is charged with AMP and before the adenylate cofactor is transferred to DNA.

2.4. Multidomain DNA Ligases Encircle Their DNA Substrates

The crystal structure of human DNA ligase I complexed to DNA not only revealed a novel DNA-binding domain (DBD) but also demonstrated that the enzyme encircles its DNA substrate (Figure 2C). The DBD domain, which contributes most of the DNA-binding affinity,[27] also directly contacts the nucleotide-binding and OB-fold domain, and can thereby stimulate the nick joining activity of the catalytic core *in trans*.[27] When contacting nicked DNA, the approximately twofold symmetric DBD domain engages the minor groove opposite the nicked strand, and it stabilizes the bound DNA in an underwound conformation by binding across the minor groove on either side of the nick (Figure 6A). Since the amino acid sequence of the DBD is conserved in other ATP-dependent eukaryotic DNA ligases[37] and this region is required for enzymatic activity in several multidomain ligases,[38−40] it appears likely that all these enzymes have a similar three-domain structure that allows them to encircle DNA. The structural similarity between the DBD of human DNA ligase I and the HhH domain of bacterial ligases (Figure 6B) and the DNA-binding activity of the HhH

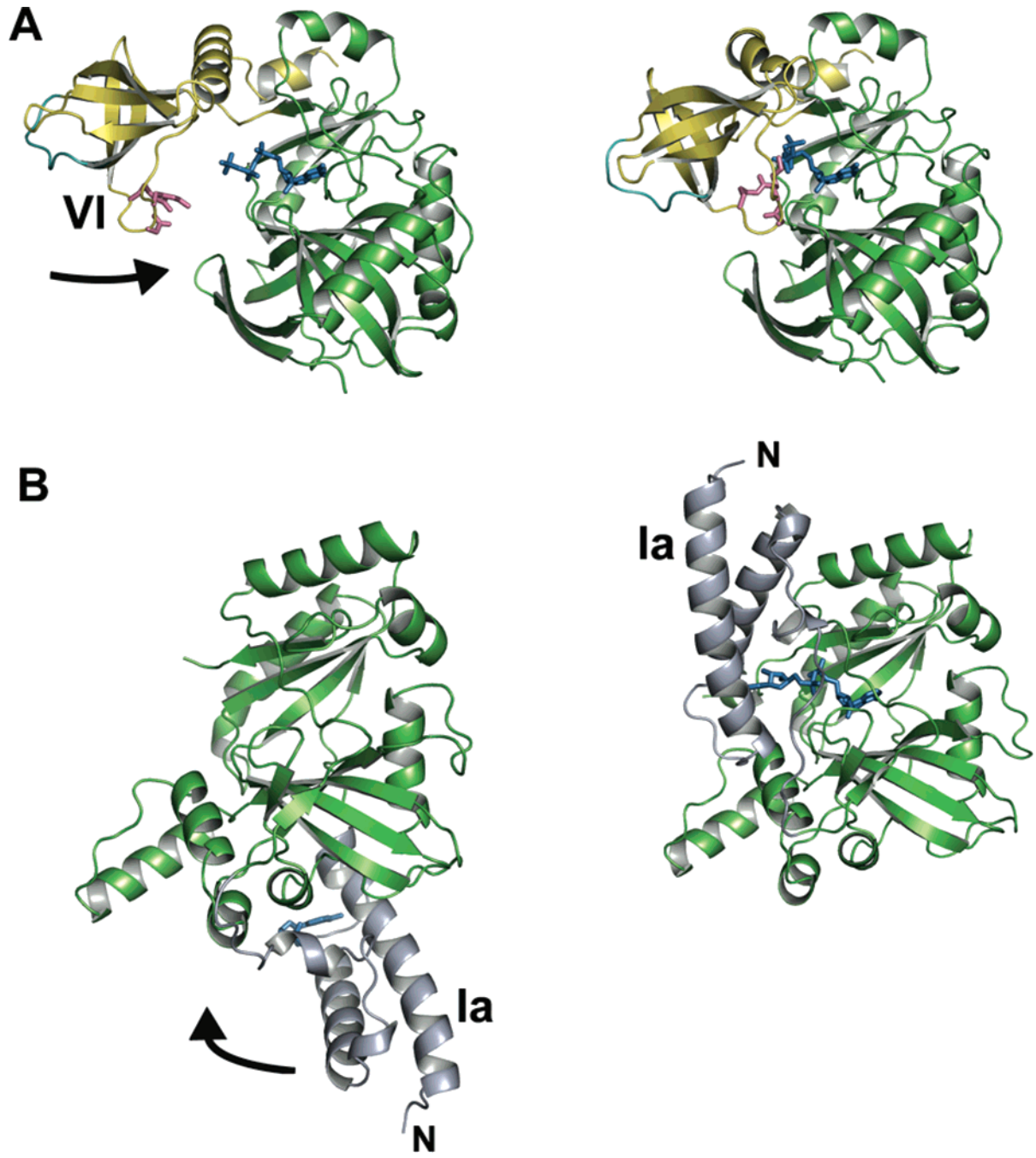


Figure 3. Active conformations during the step 1 reaction. (A) The *Chlorella* virus mRNA-capping enzyme has a nucleotide-binding domain (green) and an OB-fold domain (yellow) like those found in DNA ligases. The enzyme was crystallized in two conformations (PDB code 1ckm).[28] The inactive conformation (left panel) holds essential step 1 residues (motif VI, pink) away from the active site. The active conformation (right panel) allows the OB-fold domain to present motif VI residues to the active site of the nucleotide-binding site. Motif VI residues that are conserved among nucleotidyl transferases bind the α and *β* phosphates of GTP (blue) in order to position the α-phosphate for nucleophilic attack by the active site lysine. (B) The NAD^+-dependent DNA ligase from *Enterococcus faecalis* has been crystallized in two conformations (PDB codes 1ta8 and 1tae).[30] The left panel shows the enzyme in an inactive conformation, with the product of the step 1 reaction (NMN, in blue) bound to domain Ia. As shown in the right panel, domain Ia faces the nucleotide-binding domain during the step 1 reaction where it can properly orient the α-phosphate for attack by an active site lysine.

domain[31] together suggest that the NAD^+-dependent ligases also encircle their DNA substrates, with the HhH domain spanning between the nucleotide-binding and OB-fold domains.

2.5. Exposing the Ends of DNA for Ligation

All three domains of DNA ligase I stabilize the bound DNA in a highly distorted conformation, in which the axis of the double helix is displaced approximately 5 Å on either side of the nick (Figures 2C and 4B). This jog in the DNA axis has the effect of locally unrolling the double helix and widening both major and minor grooves immediately around the nicked ends of the DNA. The underwound DNA is slightly bent (∼15° bend), and its helical pitch changes from an A-form helix upstream of the nick (adjacent to the 3′ OH) to a B-form helix downstream of the nick (on the 5′ PO_4 side). Multiple interactions with the phosphodiester backbone and minor groove support the distorted structure of DNA in

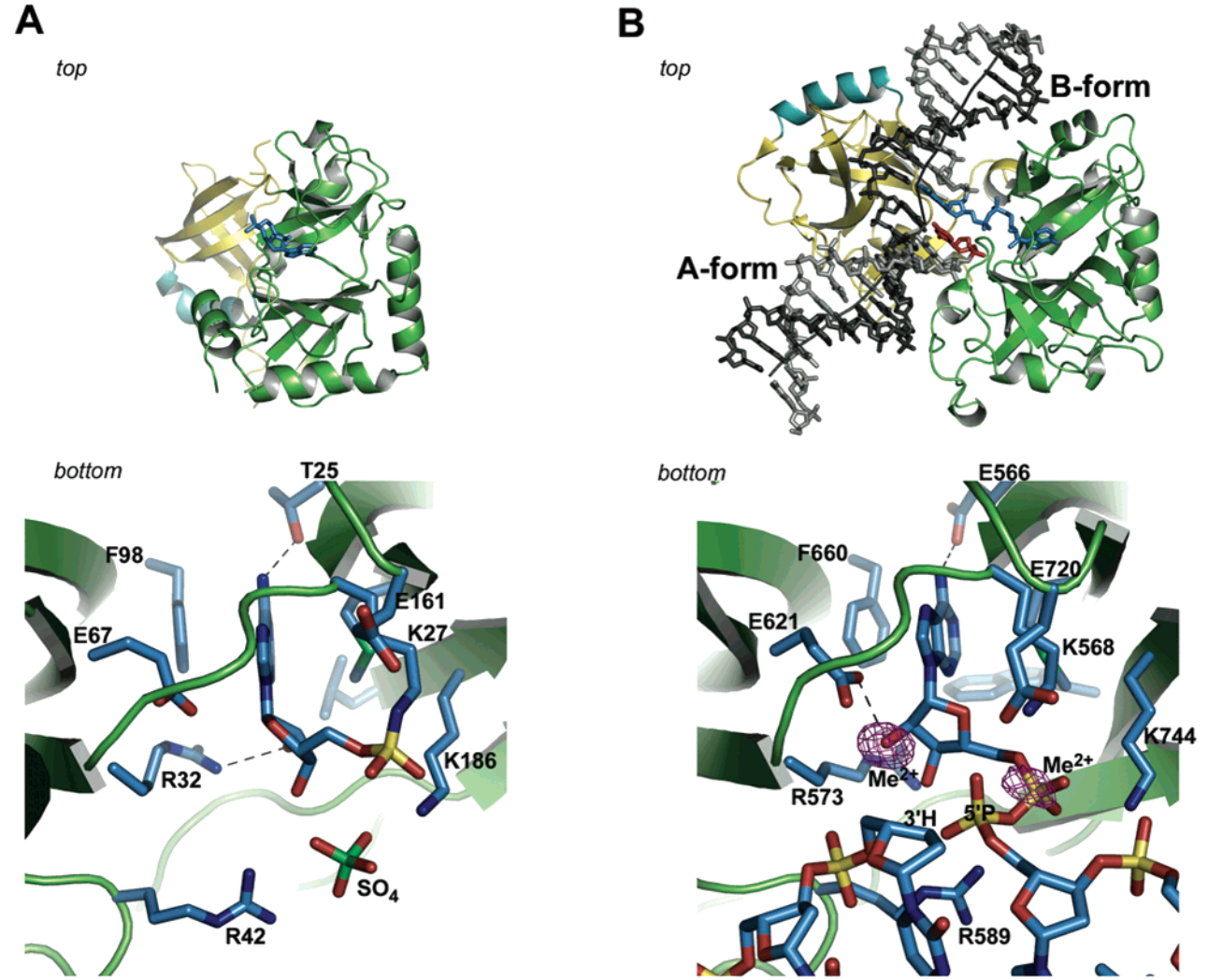


Figure 4. Covalent reaction intermediates. (A) The enzyme−adenylate intermediate from *Chlorella* virus DNA ligase (PDB code 1fvi).[24] (*top panel*) The OB-fold domain is positioned in an "open" conformation. This clears the nucleotide-binding surface of the activated enzyme for binding DNA substrate. Note the position of the cyan helix that caps the OB-fold β-barrel in comparison to that of the OB-fold domain of DNA ligase 1 in the top panel of part B. (*bottom panel*) The active site of *Chlorella* DNA ligase, highlighting the enzyme−adenylate intermediate. Conserved residues form the AMP-binding pocket, and arginine 32 positions the 2′ hydroxyl of ribose. Threonine 25 forms a base-specific contact with the N6 of adenine. A sulfate residue coordinated by arginine 42 binds in a site that is proposed to locate the incoming 5′ phosphate. (B) The catalytic core of human DNA ligase I. (*top panel*) The nick termini are highlighted by a blue 5′ adenylated end and a red 3′ hydroxyl nucleotide. The catalytic core distorts the DNA across the nick, shifting the helical axis approximately 5 Å. At the nick site, the substrate transitions from A-form to B-form DNA. The template strand (black) runs the length of the OB-fold β-barrel,[27] which partially unwinds the DNA. (*bottom panel*) The 5′ adenylated DNA intermediate in the DNA ligase I active site. Conserved residues glutamate 621 and arginine 573 position AMP through contacts with the hydroxyl groups of the ribose. Two potential metal-binding sites are located at the peaks of electron density from a $F_o - F_c$ difference Fourier map.

the complex. The template strand of the DNA substrate extends along the length of the β-barrel of the OB-fold domain, flattening the curvature of the phosphate backbone and extending the contour length of the DNA (Figure 4B). Two phenylalanines contributed by the nucleotide-binding domain (Phe635) and the OB-fold domain (Phe872) are inserted in the minor groove adjacent to the nick where they contribute to DNA bending and serve to position the ends of the nicked substrate in the active site (Figure 5). An α-helix that caps one end of the OB-fold domain abuts the 5′ PO_4 end of the nicked DNA strand as well as contacting the template strand (Figure 4B). Thus, the OB-fold domain contributes interactions with the phosphodiester backbone of both DNA strands as well as the minor groove surface of DNA opposite the nick. A structurally analogous DNA packaging protein Sac7d uses a β-barrel fold to bind nonspecifically to DNA, inserting hydrophobic residues into the minor groove to stabilize a bend centered on a transition from an A-form to B-form DNA helix at the binding site.[41]

The unwound conformation of the DNA substrate complexed with DNA ligase I positions the ends of the nicked strand into the active site for the end joining reaction. The majority of conserved residues interact with the 5′ adenylated PO_4 end of the DNA. Glutamate 621 and arginine 573 are within hydrogen bonding distance of the 2′ and 3′ hydroxyl groups, respectively, of the AMP ribose (Figure 4B). These interactions appear to orient the AMP cofactor and thereby position the 5′ phosphate for nick sealing during step 3. Notably, the amino acid in human DNA ligase IV (arginine 278) analogous to arginine 573 of DNA ligase I is mutated to histidine in a patient-derived cell line exhibiting defective DNA end joining.[11] Arginine 589 is located near the 5′ PO_4 end where it could contribute to binding affinity and alignment of the 5′ DNA end (Figure 4B). The analogous

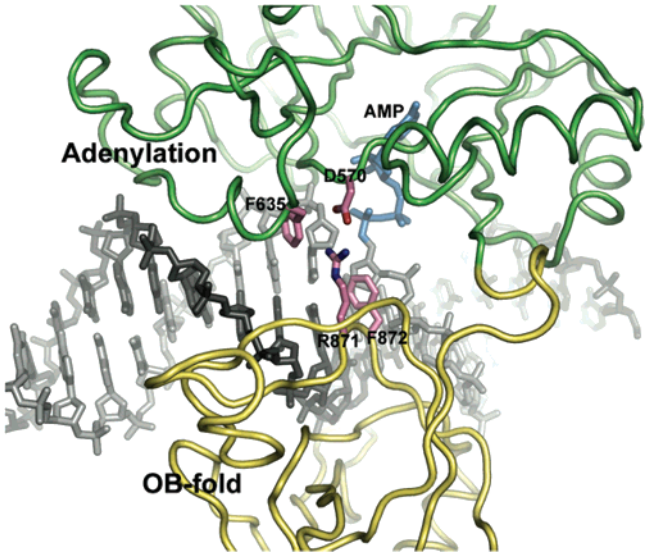


Figure 5. The DNA-binding interface of human ligase I. The adenylation domain and the OB-fold domain form an interface that positions conserved residues into the minor groove. A salt bridge between residues aspartate 570 and arginine 871 stabilizes the domain interface. Phenylalanines 635 and 872 enforce a bend in the DNA, widen the minor groove, and position the nick termini into the active site.

arginine in the *Chlorella* virus DNA ligase is essential for specific binding to phosphorylated nicks.[24] Lysine 568, which is covalently linked to AMP during step 1, is located near the 5′ phosphate of AMP during step 2 (Figure 4B). This lysine is well positioned to interact favorably with the phosphate of AMP or the 5′ PO_4, and thereby assist in step 3 reaction chemistry.[33] The chemical reversal of step 2 is energetically unfavorable, so the presence of lysine 568 near the 5′ PO_4 is unlikely to cause a futile interconversion between step 1 and step 2 reaction products. Thus, some active site residues perform double duty, serving different roles at each step of the ligation reaction.

2.6. Bound Metals Assist Ligation Chemistry

In addition to contacting the 5′ AMP cofactor, several conserved residues in the active site of DNA ligase I coordinate divalent metals. All three steps of the ligation reaction require divalent metals, and identifying the location(s) of the metal-binding sites is essential for understanding the reaction mechanism. The crystal structure of DNA ligase I complexed with a nicked DNA provides preliminary evidence for two metal-binding sites in the active site of the nucleotide-binding domain (Figure 4B). One metal site is located next to the AMP−DNA pyrophosphate linkage, in a position to be coordinated by the conserved residue glutamate 720 (Figure 4B). A metal at this site could counteract the developing negative charge on the AMP leaving group. A second metal-binding site is located near glutamate 621 and the 2′ hydroxyl of the AMP ribose (Figure 4B). It is possible that this metal orients and activates the 3′ OH of the nicked DNA strand for attack by the 5′ PO_4. However, in the crystal structure of the DNA ligase I-nicked DNA complex, the ends of the DNA are misaligned. This presumably reflects the presence of a 2′,3′-dideoxynucleoside at the 3′ terminus nick in the DNA substrate used to trap DNA ligase I for crystallization[27] that compromises the correct positioning of the nick termini in the active site by metal-mediated interactions.

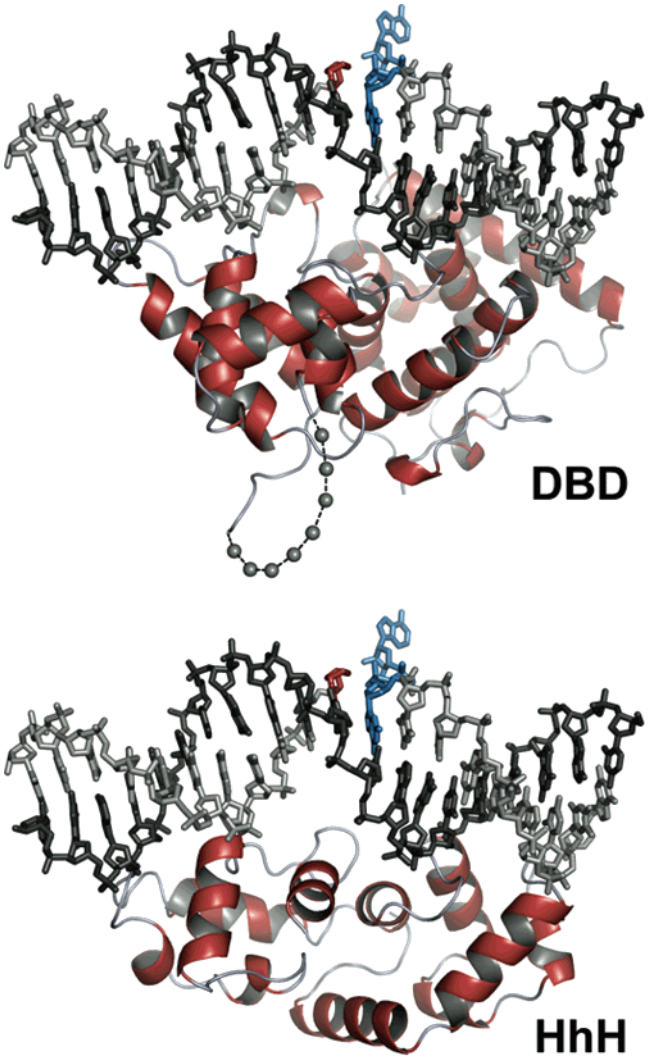


Figure 6. The DNA-binding domain (DBD) of human DNA ligase I and the helix−hairpin−helix (HhH) domain of *Thermus filiformis* DNA ligase. The top panel shows the DNA-binding domain (DBD) of human DNA ligase I bound to nicked DNA. The DNA-binding domain has an approximate twofold symmetry. The two contacts are made across the minor groove and are evenly spaced around the nick (top of the figure). The bottom panel shows that the HhH domain of the NAD^+-dependent DNA ligase from *Thermus filiformis*[25] has four helix−hairpin−helix motifs arranged on one surface with a similar twofold symmetry. The nicked DNA from the DNA ligase I structure is positioned above these DNA-binding elements, demonstrating that they are properly spaced to bind to the minor groove in a manner similar to that for ligase I. This suggests that the NAD^+-dependent DNA ligases might encircle their DNA substrates such as the mammalian ATP-dependent DNA ligases.

2.7. Nick Sensing and Fidelity of Joining

The *Chlorella* virus and *Vaccinia* virus DNA ligases bind selectively to nicked DNA substrates, in comparison to double-stranded DNA.[42,43] This nick recognition requires that the enzyme is charged with an AMP group and that the DNA nick has a 5′ PO_4 terminus. By contrast, the DBD of human DNA ligase I binds nonspecifically to nicked and double-stranded DNA, whereas the catalytic core fragment (nucleotide-binding and OB-fold domains) does not bind detectably to DNA. The absence of a specific nick-binding activity suggests that eukaryotic DNA ligase I may initially bind nonspecifically to DNA and then diffuse along the double

helix to the site of a nick, where the enzyme presumably folds into the ring-shaped active complex that encompasses the 3′ OH and 5′ PO_4 termini of the DNA nick. Although the enzymes may function this way *in vitro*, we will describe in the following section several different mechanisms that target these enzymes to their sites of action *in vivo*.

DNA ligases are very selective about the shapes and chemical characteristics of base pairs surrounding the DNA termini joined by ligation.[43−49] The catalytic selectivity of DNA ligase enzymes is referred to as the "fidelity" of DNA ligation, and it can have important biological consequences. For example, the discrimination of DNA ligase I against RNA-containing substrates may prevent the untimely ligation of unrepaired Okazaki fragments with a 5′ RNA primer present. Ligation fidelity may also contribute to the fidelity of excision repair of DNA damage. Mismatches generated by promiscuous polymerases functioning in DNA repair are poor substrates for ligation. The failure to quickly join the mispaired ends of DNA repair intermediates may provide an opportunity for exonucleolytic proofreading of a mismatched 3′ nucleotide by the repair enzymes such as the endonuclease APE1.[50]

The catalytic selectivities of mammalian DNA ligases have been assayed using homopolymer substrates consisting of short DNA or RNA oligomers annealed to a DNA or RNA template strand.[15,51] These studies revealed different end joining efficiencies of mammalian DNA ligases I, III, and IV that presumably reflect significant differences in how these enzymes engage the ends of DNA. In the crystal structure of the DNA ligase I−DNA complex, the OB-fold domain contacts the template strand and the 5′ adenylated strand of DNA. This interaction requires a B-form conformation of the DNA substrate (Figure 4B), supporting the notion that DNA ligases are sensitive to the shape of the nucleic substrate. Accordingly, human DNA ligase I is markedly less effective at joining substrates with RNA either in the template strand or in the 5′ adenylated strand that are more likely to adopt an A-form conformation.[27]

A second aspect of ligation fidelity is the strong discrimination against mismatched base pairs at the 3′ OH end of the nick and in upstream positions. As a general rule, mismatched base pairs are poorly tolerated upstream of the 3′ OH DNA end and better tolerated downstream of the 5′ PO_4 end of a nicked DNA substrate.[38,43] The 5′ phosphorylated nick terminus is positioned through multiple contacts with the attached AMP group (Figure 4B) and is therefore less likely to be affected by mispaired bases. In contrast, there are relatively fewer contacts with the 3′ OH end of the nick (Figure 4B); therefore, its correct positioning is more likely to depend on base pairing interactions with the (intact) template strand to correctly position the 3′ OH for end joining (step 3). A NAD^+-dependent DNA ligase from *Thermus thermophilus* HB8 is reported to probe the minor groove for proper hydrogen bond acceptors/donors, to ensure proper base pairing on the 3′ hydroxyl nick terminus.[52] For example, the α-helix from the nucleotide-binding domain that positions phenylalanine 635 of human DNA ligase I in the minor groove might contribute interactions that are sensitive to proper base pairing of the DNA (Figure 5). When the nucleotide-binding domain of the NAD^+-dependent DNA ligase from *Thermus* filiformis is superimposed on the ligase I−DNA complex, an α-helix (residues 199 to 209) inserts into the minor groove, where several arginine and glutamine residues are in intimate contact with the base pairs. This helix could function as a sensor that discriminates against binding to DNA substrates with irregular base pairs.

Table 1. Human DNA Ligases and Their Partner Proteins

gene	protein		protein partners
LIG1	DNA ligase I		PCNA, pol β, RFC
LIG3	DNA ligase III α	nuclear	XRCC1, PARP-1, NEIL-1
		mitochondrial	?
	DNA ligase III β		?
LIG4	DNA ligase IV		XRCC4

3. Protein Partners of Mammalian DNA Ligases

The participation of DNA ligases in different DNA transactions is mediated by specific protein−protein interactions. In the following sections, we will focus on the protein partners of mammalian DNA ligases I, III, and IV identified to date (Table 1). Where appropriate, the existence of similar interactions in lower eukaryotes and prokaryotes will be described since this conservation provides evidence for the functional signficance of the protein−protein interaction.

3.1. DNA Ligase I

The phenotype of cell lines deficient in either human DNA ligase I or Cdc9, its functional homologue in *S. cerevisiae*, provided compelling evidence linking these enzymes with DNA replication and DNA excision repair.[8,10,53] To identify replication factors that associate with DNA ligase I, a HeLa nuclear extract was fractionated by DNA ligase I-affinity chromatography. This approach led to the detection and characterization of a direct physical interaction between DNA ligase I, proliferating cell nuclear antigen (PCNA), a DNA sliding clamp,[54,55] and, more recently, an interaction between DNA ligase I and replication factor C (RFC), the sliding clamp loader.[56] Mapping studies and amino acid sequence comparisons demonstrated that DNA ligase I has an N-terminal PCNA-binding motif that was initially identified in the Cdk inhibitor p21 and has now been found in a growing number of proteins involved in DNA metabolism.[57,58] The PCNA-binding motif of DNA ligase I is necessary for the recruitment of this enzyme to replication foci.[57] Moreover, it is required for the efficient joining of Okazaki fragments and completion of long patch base excision repair in cell extract assays, suggesting that the PCNA−DNA ligase I interaction plays a key role in coordinating the final ligation step with earlier processing and DNA synthesis reactions.[55]

Since the PCNA homotrimer interacts with the replicative DNA polymerases δ and ϵ, RFC, and the FEN-1 nuclease, in addition to DNA ligase I, it is possible that PCNA coordinates their activities by forming ternary complexes with two or three replication factors.[59,60] Indeed there is evidence that a heterotrimeric PCNA-like complex from *Sulfolobus solfataricus* forms a ternary complex with the bacterial DNA polymerase, DNA ligase, and nuclease.[61] Furthermore, yeast PCNA coordinates the processing activities of FEN-1 and the DNA synthesis activities of DNA pol δ to generate a ligatable nick,[62] suggesting that PCNA ternary complexes may also mediate the reaction steps prior to ligation in eukaryotes. However, there is accumulating evidence that the binding of DNA ligase I to the interdomain connector loop of a PCNA monomer occludes the same binding site on other PCNA subunits in the homotrimer. First, the stoichiometry of PCNA−DNA ligase I complexes formed

on duplex DNA was 3:1.[54] Second, ternary complexes of FEN-1, PCNA, and Cdc9 could not be detected by co-immunoprecipitation (S.V. and A.E.T., unpublished result). Instead, FEN-1 and Cdc9 appeared to compete for PCNA binding, suggesting that their interactions with PCNA are mutually exclusive. Finally, the similar size and shape of the rings formed by PCNA and the catalytic domain of human DNA ligase I are consistent with DNA ligase I occluding the three binding sites on the PCNA trimer when the two rings are engaged.[27,63]

It is noteworthy that, in the *E. coli* DNA replication machinery, there is an interaction between the NAD^+-dependent replicative DNA ligase and the β protein, which forms a dimeric sliding clamp.[64] Furthermore, the structure of the *E. coli* NAD-dependent DNA ligase determined by X-ray crystallography[25] suggests that this enzyme may also encircle nicked DNA in a fashion similar to that for human DNA ligase I. Thus, the interaction between ring-shaped DNA ligases and sliding clamps is likely to be a common feature of prokaryotic and eukaryotic DNA replication (Figure 7). Although the relative orientation of the ligase and sliding clamp rings has yet to be determined, the similar size and shape of the PCNA and DNA ligase I rings (Figure 7) suggests that there may be extensive interactions between the surfaces of the rings. Indeed, there is evidence for an interaction between the putative DNA-binding domain of the *Sulfolobus* ATP-dependent DNA ligase and one of the subunits of the heterotrimeric *Sulfolobus* PCNA.[61] To date, the N-terminal PCNA-binding motif is the only region of DNA ligase I that has been shown to interact with PCNA.[55,57] While the docking of this motif to the interdomain connector loop of PCNA is important for protein−protein interactions in the absence of DNA, a C-terminal region of PCNA plays a critical role in the PCNA-dependent stimulation of FEN-1 nuclease activity.[65] In accord with these biochemical and genetic studies, it was shown recently that the amino acid sequence immediately N-terminal to the canonical PCNA-binding motif forms a β-zipper structure with the C-terminal region of PCNA that appears to lock the conformation of FEN-1 and the DNA substrate onto the PCNA ring.[66] It is possible that a similar mechanism is involved in the interaction between yeast Cdc9 and PCNA but, in human DNA ligase I, the PCNA-binding motif is located right at the N-terminus.[57]

As noted above, PCNA stimulates the nuclease activity of FEN-1.[65] By contrast, the effect of PCNA on DNA ligase I is less clear. Two laboratories have reported that PCNA does not stimulate DNA ligase I[54,67] whereas one laboratory observed a modest stimulation of DNA ligase I activity by high concentrations of PCNA.[68] Notably, the interaction between PCNA and DNA ligase I is distinct from the interactions with other replication proteins in that it mediates the final step of Okazaki fragment maturation. Thus, the apparent ability of DNA ligase I to block the interaction of other replication factors such as FEN-1 (S.V. and A.E.T., unpublished result) and DNA pol δ[54,69] to PCNA may prevent these factors from processing the nicked DNA intermediate. Although the use of ring proteins is an attractive mechanism to tether proteins to the DNA duplex, this creates a need to unload these proteins from duplex DNA. In the case of PCNA, the only activity that has been shown to unload PCNA is RFC, which is also the clamp loader.[70] Since neither PCNA nor DNA ligase I has a robust nick-binding activity, it is possible that RFC holds PCNA at the nick generated by the concerted action of pol δ and FEN-1.[62] The functional interaction of DNA ligase I with both PCNA and RFC not only may target this enzyme to the DNA nick but also may promote the subsequent RFC-mediated unloading of PCNA and dissociation of the protein complex assembled at the DNA nick. A potential problem with this model is that the structure of the yeast RFC−PCNA complex suggests that RFC would occlude the DNA nick.[71] However, it is possible that the functional interactions with PCNA are determined by the structure of the DNA substrate. For example, the switch between DNA pol α and DNA pol δ involves an interaction between RFC and a gapped template coated with RPA prior to the loading of PCNA onto the gapped template and assembly of the PCNA−pol δ complex.[72,73] Notably, the DNA-binding site of RFC predicted from the RFC−PCNA structure could accommodate a gapped duplex but not a nicked duplex.[71] Thus, the structure of the RFC−

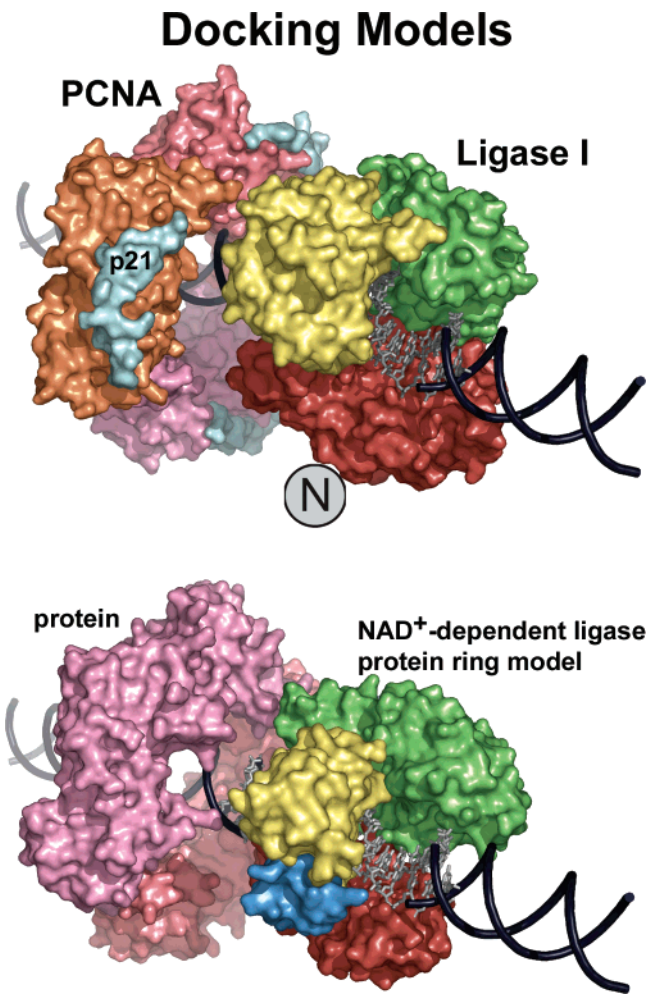


Figure 7. Theoretical docking models of prokaryotic and eukaryotic DNA ligases and sliding clamp proteins. In the top panel, the nicked DNA from the DNA ligase I structure has been extended as B-form DNA through the central channel of PCNA (PDB code 1axc).[63] The p21 peptide highlights the surface of PCNA where multiple proteins bind, including the N-terminus of DNA ligase I, which is missing from the structure.[27] In the bottom panel, the individual domains of the NAD^+-dependent DNA ligase from *Thermus filiformis*[25] have been superimposed onto to the domains of human DNA ligase I.[27] The superposition of the bacterial DNA ligase domains requires very little adjustment of the interdomain linkers. The shape of this modeled ligase complements the shape of the dimeric β clamp from *Escherichia coli*, suggesting that bacterial and eukaryotic systems utilize a similar mechanism of assembling toroids on DNA.

PCNA complex bound to nicked DNA may be significantly different, permitting DNA ligase I access to the DNA nick and promoting PCNA release from RFC.

3.2. DNA Ligase III

The *LIG3* gene is not found in lower eukaryotes such as *S. cerevisiae*. Interestingly, this gene encodes three different polypeptides. In somatic cells, the use of different in-frame alternative translation initiation sites generates two forms of DNA ligase III α that are specifically targeted to the mitochondria and the nucleus, respectively.[16,74,75] The shorter nuclear form lacks an obvious nuclear localization signal and may be recruited to the nucleus by complex formation with its partner protein, XRCC1,[21,76] which has a functional NLS and is described in more detail below. The additional N-terminal sequence of the longer mitochondrial version of DNA ligase III α targets this polypeptide to the mitochondria but is removed upon entry into the organelle. Although mitochondrial DNA ligase III α is critical for the mitochondrial function,[77] the role of this enzyme in mitochondrial DNA metabolism has yet to be elucidated. However, it does seem to function in this organelle without XRCC1,[78] suggesting that it may have novel mitochondrial partner proteins. A third form of DNA ligase III, DNA ligase III *β*, is generated by a germ cell-specific alternative splicing event that replaces the exon encoding the C-terminal region of DNA ligase III α that interacts with XRCC1 with an exon that encodes a short 17−18 amino acid peptide.[79] At the present time, the role of DNA ligase III *β* in gametogenesis is not known but the alternative splicing that occurs during the latter stages of meiotic recombination would be consistent with this isoform participating in either the completion of meiotic recombination events or as yet undefined postmeiotic DNA repair pathways.

The human XRCC1 gene was cloned based on its ability to complement the hypersensitivity of the mutant Chinese hamster ovary cell line EM9 to simple DNA alkylating agents.[76] Based on the phenotype of the EM9 cells, the XRCC1 gene product was implicated in base excision and DNA single-strand break repair. In subsequent studies, DNA ligase III α was identified as a partner protein of XRCC1.[21] Indeed XRCC1 is required for the stability and activity of nuclear DNA ligase III α.[22] Both DNA ligase III α and XRCC1 possess a C-terminal BRCT motif that was initially identified in the breast cancer susceptibility gene, the *BRCA1* gene.[79−82] Several laboratories demonstrated that these motifs interact in the DNA ligase III α−XRCC1 complex.[79,81−83] The structure of the C-terminal BRCT domain of XRCC1 has been determined and, as the prototype of this type of motif, was used to map the cancer-associated amino acid changes in the BRCT motif of BRCA1.[84] The XRCC1 protein has no catalytic activity, but there is a growing list of XRCC1-binding proteins involved in base excision and DNA single-strand break repair.[85−91] Recent studies provide evidence that defects in base excision and single-strand break repair may contribute to neurodegeneration. XRCC1 interacts with aprataxin, a protein implicated in a neurodegenerative disease.[92−95] Notably, this interaction is important for the stability of XRCC1 (and presumably DNA ligase III α). In addition, DNA ligase III α interacts with tyrosyl phopshodiesterase I, which is encoded by another gene that is mutated in a neurodegenerative disease.[96]

In base excision repair, XRCC1 interacts with enzymes involved at all the stages of the BER pathway from the initiating base excision step to the final ligation step.[21,22,85−91] This is difficult to reconcile with the "passing the baton" model, in which pathway intermediates are handed over to the next enzyme in the pathway. However, since there are more molecules of XRCC1 than DNA ligase III α,[97] it is formally possible that free XRCC1 participates in the earlier stages of the pathway with the DNA ligase III α−XRCC1 complex acting only at the final ligation step. In DNA single-strand break repair, the abundant nuclear protein PARP-1 binds to the breaks and initiates the recruitment of the repair machinery.[97−99] The binding of PARP-1 to DNA breaks activates its poly(ADP-ribose) polymerase activity, resulting in the addition of poly(ADP-ribose) chains to itself and nearby proteins such as histones.[100] Although the automodification of PARP-1 induces its dissociation from DNA breaks, it has been suggested that the network of negatively charged polymers in the vicinity of the break effectively hides the break from DNA metabolizing enzymes that would otherwise bind to the phosphodiester backbone of DNA. There is compelling cell biology and biochemical evidence that the DNA ligase III α−XRCC1 complex is recruited to DNA single-strand breaks by virtue of the preferential binding of both subunits of the complex to automodified PARP-1.[88,97] Thus, the synthesis of poly(ADP-ribose) serves as the signal for the recruitment of the single-strand break repair machinery that will then have to find the DNA strand break in the context of the negatively charged poly(ADP-ribose) polymer. Intriguingly, both PARP-1 and DNA ligase III α have similar zinc finger motifs at their N-termini.[101] In PARP-1, there are two adjacent zinc fingers that constitute the DNA-binding domain of this enzyme whereas, in DNA ligase III, there is only a single zinc finger. The structure of the DNA ligase III zinc finger has recently been determined by NMR, revealing a potential DNA-binding surface.[102] It has been suggested that the DNA ligase III zinc finger mediates the binding of the DNA ligase III α−XRCC1 to DNA single-strand breaks in the presence of poly(ADP-ribose).[97] Since the majority of DNA single-strand breaks generated by either reactive oxygen species or AP endonuclease will not be immediately ligatable, it is likely that DNA processing activities such as polynucleotide kinase, recruited via an interaction with XRCC1,[90] will act prior to ligation. A prediction of this model is that, upon the generation of a ligatable nick, DNA ligase III α will undergo a conformational change involving the displacement of the zinc finger from the DNA nick, enabling the nick to be encircled by the catalytic domain.

3.3. DNA Ligase IV

Similar to nuclear DNA ligase III α, DNA ligase IV has a partner protein, XRCC4, upon which it depends for stability and activity.[103−105] XRCC4 protein alone oligomerizes, forming dimers, tetramers, and higher order complexes.[106] The structure of the XRCC4 dimer[107,108] is reminiscent of SMC proteins with two globular head domains and two extended coiled-coil tails. DNA ligase IV consists of an N-terminal catalytic domain and a C-terminal region that contains two BRCT motifs.[16] The region between the BRCT motifs interacts with a region on the coiled-coils of XRCC4 and appears to block tetramer formation by XRCC4,[40,108] suggesting that the DNA ligase IV−XRCC4 complex is composed of one DNA ligase IV molecule and two XRCC4 molecules.

Genetic studies have implicated eukaryotic DNA ligase IV in the repair of DNA double-strand breaks by non-

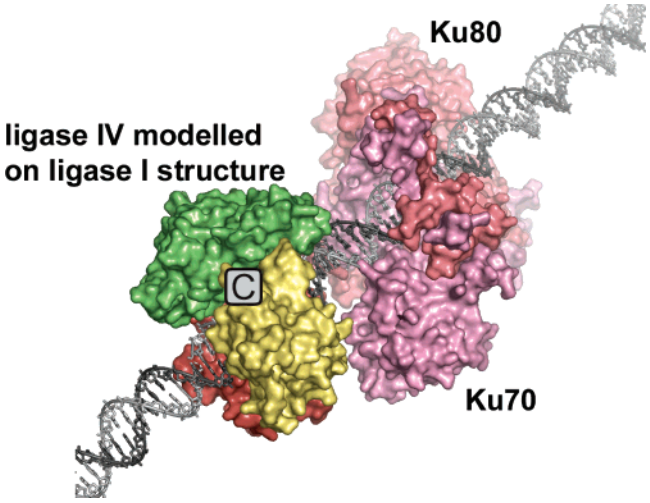


Figure 8. Docking of the predicted structure of the DNA ligase IV catalytic domain onto the DNA-bound Ku70−Ku80 heterodimer. The sequence for the catalytic domain of DNA ligase IV (residues 1−607) was modeled onto the DNA ligase I structure[27] using SWISS-MODEL.[129] B-form DNA was fitted onto the nicked DNA of the DNA ligase I structure and the duplex region of the DNA bound to Ku70−Ku80.[118] The position that the two BRCT domains extend from the C-terminus of the DNA ligase IV catalytic domain is designated.

homologous end joining (NHEJ) and in the completion of V(D)J recombination events in immune cells.[11,103,109−111] A key step in NHEJ is the juxtapositioning or synapsis of DNA ends. Although DNA ligase IV−XRCC4 binds to DNA ends,[112] the DNA-PKcs subunit of the DNA-dependent protein kinase (DNA-PK) appears to be the major factor bridging DNA ends in mammalian cells.[113] In accord with this function, DNA-PKcs interacts with and promotes intermolecular joining by DNA ligase IV−XRCC4.[112] Under physiological conditions, the stable interaction of DNA-PKcs with DNA ends is dependent upon the DNA-binding component of DNA-PK, the heterodimeric Ku70−Ku80.[114,115] DNA ligase IV−XRCC4 interacts with the Ku70−Ku80 complex in addition to DNA-PKcs, but there are conflicting published reports regarding the effect of Ku70−Ku80 on DNA ligase IV activity.[112,116,117] Interestingly, Ku70−Ku80 is also a ring protein that can slide along the DNA helix.[118] The size of the Ku70−Ku80 ring and the predicted ring formed by the catalytic domain of DNA ligase IV are similar (Figure 8), suggesting that, like Okazaki fragment joining, NHEJ may involve the functional interaction between a ring-shaped protein complex and the ring-shaped catalytic domain of DNA ligase IV. However, unlike the homotrimeric PCNA clamp that can be opened by the disruption of an interface between PCNA subunits, the intertwining of the Ku70 and Ku80 subunits suggests that another mechanism such as proteolysis may be required to remove Ku70−Ku80 complex from DNA duplexes after the completion of repair. Similarly, the structure of DNA-PKcs determined by electron microscopy suggests that this protein contains channels through which DNA is threaded.[119] At the present time, the architecture of the functional complexes formed by DNA-PK and DNA ligase IV−XRCC4 during the joining of DNA ends and the mechanisms by which these complexes are turned over after ligation are poorly understood.

During end-synapsis by DNA-PKcs, it appears that each DNA end has a DNA-PKcs molecule bound and it is protein−protein interactions between the end-bound DNA-PKcs molecules that juxtapose the DNA ends.[113] Since DNA ligase IV−XRCC4 associates with DNA-PKcs bound to DNA ends,[112] this suggests that there will be two DNA ligase IV−XRCC4 complexes at a DSB site. Although *Saccharomyces cerevisiae* lacks an obvious homologue of DNA-PKcs, it appears that the overall mechanism of NHEJ is conserved in eukaryotes. In yeast NHEJ, the yeast Rad50−Mre11−Xrs2 complex bridges DNA ends and specifically stimulates intermolecular DNA joining by the Dnl4−Lif1 complex, the yeast equivalent of mammalian DNA ligase IV−XRCC4.[120,121] The functional interaction between the Rad50−Mre11−Xrs2 and Dnl4−Lif1 complexes is mediated by a physical interaction between Xrs2 and Lif1 subunits of these complexes.[121]

The majority of DNA double-strand breaks generated *in vivo* will have damaged and/or noncomplementary termini that will require processing prior to ligation. DNA sequence analysis of the repaired break sites suggests that small regions (2−4 nucleotides) of complementary nucleotides are used to align the DNA ends during the processing reactions.[122−124] Elegant studies by the Lieber lab have provided evidence that DNA ligase IV−XRCC4 will ligate one strand even if the structure of the other broken strand is not ligatable,[125] suggesting that the processing reactions to generate ligatable termini can occur independently on the two broken strands.

The end processing reactions may involve enzymes that process damaged termini such as polynucleotide kinase, that remove mismatched nucleotides, and that fill in short gaps. There is growing evidence that DNA ligase IV plays a central role in coordinating the actions of the processing activities. The pol X family polymerases, μ and λ in mammals and Pol4 in *S. cerevisiae*, have been implicated in gap-filling reactions associated with NHEJ.[116,125−127] Notably there are conserved physical and functional interactions between these polymerases and DNA ligase IV that coordinate gap-filling DNA synthesis and DNA ligation. In addition, there is a similar interaction between DNA ligase IV and another pol X family member, terminal transferase, that is expressed in lymphoid cells and contributes to diversity during V(D)J recombination by adding nontemplated nucleotides to the broken coding ends.[116,125] In *S. cerevisiae*, DNA ligase IV interacts with FEN-1, a DNA structure-specific endonuclease that removes mismatched 5′ nucleotides, in addition to Pol4.[128] Since Pol4 also interacts with FEN-1, DNA ligase IV is involved in a series of species-specific pairwise protein−protein interactions that coordinate the end processing and ligation reactions that complete the repair of DSBs by NHEJ.[128] It is likely that a similar mechanism coordinates the activities of the mammalian factors involved in end processing.

4. Concluding Remarks

The recent determination of several DNA ligase structures by X-ray crystallography has provided an increasingly detailed understanding of the molecular mechanisms and dynamic conformational changes involved in the three-step ligation reaction. However, more experiments are needed to fully understand how DNA ligases catalyze the formation of phosphodiester bonds. In particular, it will be necessary to isolate all the different reaction intermediates for biochemical and crystallographic studies, to understand how the active site is remodeled at each step, and to guide mutational studies that address the functional importance of specific amino acid residues in the ligation reaction.

Genome sequencing studies have revealed that the presence of more than one DNA ligase gene occurs in pro-

karyotes as well as eukaryotes. The DNA ligase isoforms appear to have distinct cellular functions, and there is compelling evidence in eukaryotes that these enzymes are targeted to different DNA transactions by specific protein−protein interactions. Since human DNA ligase I and DNA ligase IV−XRCC4 both interact with ring proteins, PCNA and Ku70−Ku80, respectively, it is particularly intriguing that the DNA ligase catalytic domain forms a ring when it encircles nicked DNA. Further biochemical and structural studies are needed to elucidate the molecular mechanims by which DNA ligase-partner proteins recruit the DNA ligase catalytic domain to the DNA substrate and contribute to the DNA joining reaction.

5. Acknowledgments

Our work on DNA ligases is supported by the Structural Cell Biology of DNA Repair Program Grant (P01 CA92584) and other research grants awarded to A.E.T. (GM47251, GM57479, and ES12512) and T.E. (GM52504). J.M.P. is an NRSA postdoctoral fellow (GM 70151). T.E. is the Hsien Wu and Daisy Yen Wu Professor at Harvard Medical School.

6. References

(1) Yao, N.; Turner, J.; Kelman, Z.; Stukenberg, P. T.; Dean, F.; Shechter, D.; Pan, Z.-Q.; Hurwitz, J.; O'Donnell, M. *Genes Cells* **1996**, *1*, 101.
(2) Bassing, C. H.; Alt, F. W. *DNA Repair* **2004**, *3*, 781.
(3) Barnes, D. E.; Lindahl, T. *Annu. Rev. Genet.* **2004**, *38*, 445.
(4) Ward, J. F. *Prog. Nucleic Acids Res.* **1988**, *35*, 95.
(5) Engler, M. J.; Richardson, C. C. *DNA ligases*; Academic Press: New York, 1982.
(6) Lehman, I. R. *Science* **1974**, *186*, 790.
(7) Soderhall, S.; Lindahl, T. *FEBS Lett.* **1976**, *67*, 1.
(8) Johnston, L. H. *Mol. Gen. Genet.* **1979**, *170*, 89.
(9) Nasmyth, K. A. *Cell* **1977**, *12*, 1109.
(10) Barnes, D. E.; Tomkinson, A. E.; Lehmann, A. R.; Webster, A. D.; Lindahl, T. *Cell* **1992**, *69*, 495.
(11) Riballo, E.; Critchlow, S. E.; Teo, S.-H.; Doherty, A. J.; Priestly, A.; Broughton, B.; Kysela, B.; Beamish, H.; Plowman, N.; Arlett, C. F.; Lehmann, A. R.; Jackson, S. P.; Jeggo, P. A. *Curr. Biol.* **1999**, *9*, 699.
(12) Shuman, S.; Schwer, B. *Mol. Microbiol.* **1995**, *17*, 405.
(13) Barnes, D. E.; Johnston, L. H.; Kodama, K.; Tomkinson, A. E.; Lasko, D. D.; Lindahl, T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6679.
(14) Soderhall, S.; Lindahl, T. *Biochem. Biophys. Res. Commun.* **1973**, *53*, 910.
(15) Tomkinson, A. E.; Roberts, E.; Daly, G.; Totty, N. F.; Lindahl, T. *J. Biol. Chem.* **1991**, *266*, 21728.
(16) Wei, Y.-F.; Robins, P.; Carter, K.; Caldecott, K.; Pappin, D. J. C.; Yu, G.-L.; Wang, R.-P.; Shell, B. K.; Nash, R. A.; Schar, P.; Barnes, D. E.; Haseltine, W. A.; Lindahl, T. *Mol. Cell. Biol.* **1995**, *15*, 3206.
(17) Gong, C.; Martins, A.; Bongiorno, P.; Glickman, M.; Shuman, S. *J. Biol. Chem.* **2004**, *279*, 20594.
(18) Sriskanda, V.; Shuman, S. *Nucleic Acids Res.* **2001**, *29*, 493.
(19) Wilson, T. E.; Grawunder, U.; Lieber, M. R. *Nature* **1997**, *388*, 495.
(20) Barnes, D. E.; Tomkinson, A. E.; Lehmann, A. R.; Webster, A. D. B.; Lindahl, T. *Cell* **1992**, *69*, 495.
(21) Caldecott, K. W.; McKeown, C. K.; Tucker, J. D.; Ljungquist, S.; Thompson, L. H. *Mol. Cell Biol.* **1994**, *14*, 68.
(22) Caldecott, K. W.; McKeown, C. K.; Tucker, J. D.; Stanker, L.; Thompson, L. H. *Nucleic Acids Res.* **1996**, *23*, 4836.
(23) Subramanya, H. S.; Doherty, A. J.; Ashford, S. R.; Wigley, D. B. *Cell* **1996**, *85*, 607.
(24) Odell, M.; Sriskanda, V.; Shuman, S.; Nikolov, D. B. *Mol. Cell* **2000**, *6*, 1183−93.
(25) Lee, J. Y.; Chang, C.; Song, H. K.; Moon, J.; Yang, J. K.; Kim, H. K.; Kwon, S. T.; Suh, S. W. *EMBO J.* **2000**, *19*, 1119.
(26) Singleton, M. R.; Hakansson, K.; Timson, D. J.; Wigley, D. B. *Structure Fold Des.* **1999**, *7*, 35.
(27) Pascal, J. M.; O'Brien, P. J.; Tomkinson, A. E.; Ellenberger, T. *Nature* **2004**, *432*, 473.
(28) Hakansson, K.; Doherty, A. J.; Shuman, S.; Wigley, D. B. *Cell* **1997**, *89*, 545.
(29) Murzin, A. G. *EMBO J.* **1993**, *12*, 861.
(30) Gajiwala, K. S.; Pinko, C. *Structure (Cambridge)* **2004**, *12*, 1449.
(31) Jeon, H. J.; Shin, H. J.; Choi, J. J.; Hoe, H. S.; Kim, H. K.; Suh, S. W.; Kwon, S. T. *FEMS Microbiol. Lett.* **2004**, *237*, 111.
(32) Sawaya, R.; Shuman, S. *Biochemistry* **2003**, *42*, 8240.
(33) Sriskanda, V.; Shuman, S. *Nucleic Acids Res.* **1998**, *26*, 4618.
(34) Doherty, A. J.; Wigley, D. B. *J. Mol. Biol.* **1999**, *285*, 63.
(35) Doherty, A. J.; Dafforn, T. R. *J. Mol. Biol.* **2000**, *296*, 43.
(36) Doherty, A. J.; Suh, S. W. *Nucleic Acids Res.* **2000**, *28*, 4051.
(37) Martin, I. V.; MacNeill, S. A. *Genome Biol.* **2002**, *3*, Reviews 3005.
(38) Tomkinson, A. E.; Tappe, N. J.; Friedberg, E. C. *Biochemistry* **1992**, *31*, 11762.
(39) Sriskanda, V.; Schwer, B.; Ho, C. K.; Shuman, S. *Nucleic Acids Res.* **1999**, *27*, 3953.
(40) Grawunder, U.; Zimmer, D.; Leiber, M. R. *Curr. Biol.* **1998**, *8*, 873.
(41) Robinson, H.; Gao, Y. G.; McCrary, B. S.; Edmondson, S. P.; Shriver, J. W.; Wang, A. H. *Nature* **1998**, *392*, 202.
(42) Sekiguchi, J.; Shuman, S. *Nucleic Acids Res.* **1997**, *25*, 727.
(43) Sriskanda, V.; Shuman, S. *Nucleic Acids Res.* **1998**, *26*, 3536.
(44) Bhagwat, A. S.; Sanderson, R. J.; Lindahl, T. *Nucleic Acids Res.* **1999**, *27*, 4028.
(45) Luo, J.; Bergstrom, D. E.; Barany, F. *Nucleic Acids Res.* **1996**, *24*, 3071.
(46) Nakatani, M.; Ezaki, S.; Atomi, H.; Imanaka, T. *Eur. J. Biochem.* **2002**, *269*, 650.
(47) Pritchard, C. E.; Southern, E. M. *Nucleic Acids Res.* **1997**, *25*, 3403.
(48) Shuman, S. *Biochemistry* **1995**, *34*, 16138.
(49) Tong, J.; Barany, F.; Cao, W. *Nucleic Acids Res.* **2000**, *28*, 1447.
(50) Chou, K.-M.; Cheng, Y.-C. *Nature* **2002**, *415*, 655.
(51) Robins, P.; Lindahl, T. *J. Biol. Chem.* **1996**, *271*, 24257.
(52) Liu, P.; Burdzy, A.; Sowers, L. C. *Nucleic Acids Res.* **2004**, *32*, 4503.
(53) Johnston, L. H.; Nasmyth, K. A. *Nature* **1978**, *274*, 891.
(54) Levin, D. S.; Bai, W.; Yao, N.; O'Donnell, M.; Tomkinson, A. E. *Proc. Natl. Acad Sci. U.S.A.* **1997**, *94*, 1286.
(55) Levin, D. S.; McKenna, A. E.; Motycka, T. A.; Matsumoto, Y.; Tomkinson, A. E. *Curr. Biol.* **2000**, *10*, 91.
(56) Levin, D. S.; Vijayakumar, S.; Liu, X.; Bermudez, V. P.; Hurwitz, J.; Tomkinson, A. E. *J. Biol. Chem.* **2004**, *279*, 55196.
(57) Montecucco, A.; Rossi, R.; Levin, D. S.; Gary, R.; Park, M. S.; Motycka, T. A.; Ciarrocchi, G.; Villa, A.; Biamonti, G.; Tomkinson, A. E. *EMBO J.* **1998**, *17*, 3786.
(58) Warbrick, E. *BioEssays* **1998**, *20*, 195.
(59) Jonsson, Z.; Hubscher, U. *BioEssays* **1997**, *19*, 967.
(60) Kelman, Z. *Oncogene* **1997**, *14*, 629.
(61) Dionne, I.; Nookala, R. K.; Jackson, S. P.; Doherty, A. J.; Bell, S. D. *Mol. Cell* **2003**, *11*, 275.
(62) Garg, P.; Stith, C. M.; Sabouri, N.; Johansson, E.; Burgers, P. M. *Genes Dev.* **2004**, *18*, 2764.
(63) Gulbis, J. M.; Kelman, Z.; Hurwitz, J.; O'Donnell, M.; Kuriyan, J. *Cell* **1996**, *87*, 297.
(64) Lopez de Saro, F. J.; O'Donnell, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8376.
(65) Gomes, X. V.; Burgers, P. M. *EMBO J.* **2000**, *19*, 3811.
(66) Chapados, B. R.; Hosfield, D. J.; Han, S.; Qiu, J.; Yelent, B.; Shen, B.; Tainer, J. A. *Cell* **2004**, *116*, 39.
(67) Jonsson, Z. O.; Hindges, R.; Hubscher, U. *EMBO J.* **1998**, *17*, 2412.
(68) Tom, S.; Henricksen, L. A.; Park, M. S.; Bambara, R. A. *J. Biol. Chem.* **2001**, *276*, 24817.
(69) Mossi, R.; Ferrari, E.; Hubscher, U. *J. Biol. Chem.* **1998**, *273*, 14322.
(70) Cai, J.; Gibbs, E.; Uhlmann, F.; Phillips, B.; Yao, N.; O'Donnell, M.; Hurwitz, J. *J. Biol. Chem.* **1997**, *1997*, 18974.
(71) Bowman, G. D.; O'Donnell, M.; Kuriyan, J. *Nature* **2004**, *429*, 724.
(72) Waga, S.; Stillman, B. *Nature* **1994**, *369*, 207.
(73) Yuzhakov, A.; Kelman, Z.; Hurwitz, J.; O'Donnell, M. *EMBO J.* **1999**, *18*, 6189.
(74) Chen, J.; Tomkinson, A. E.; Ramos, W.; Mackey, Z. B.; Danehower, S.; Walter, C. A.; Schultz, R. A.; Besterman, J. M.; Husain, I. *Mol. Cell Biol.* **1995**, *15*, 5412.
(75) Lakshmipathy, U.; Campbell, C. *Mol. Cell Biol.* **1999**, *19*, 3869.
(76) Thompson, L. H.; Brookman, K. W.; Jones, N. J.; Allen, S. A.; Carrano, A. V. *Mol. Cell. Biol.* **1990**, *10*, 6160.
(77) Lakshmipathy, U.; Campbell, C. *Nucleic Acids Res.* **2001**, *29*, 668.
(78) Lakshmipathy, U.; Campbell, C. *Nucleic Acids Res.* **2000**, *28*, 3880.
(79) Mackey, Z. B.; Ramos, W.; Levin, D. S.; Walter, C. A.; McCarrey, J. R.; Tomkinson, A. E. *Mol. Cell. Biol.* **1997**, *17*, 989.
(80) Koonin, E. V.; Alschul, S. F.; Bork, P. *Nature Genetics* **1996**, *13*, 266.
(81) Nash, R. A.; Caldecott, K. W.; Barnes, D. E.; Lindahl, T. *Biochemistry* **1997**, *36*, 5207.
(82) Taylor, R. M.; Wickstead, B.; Cronin, S.; Caldecott, K. W. *Curr. Biol.* **1998**, *8*, 877.
(83) Taylor, R. M.; Moore, D. J.; Whitehouse, J.; Johnson, P.; Caldecott, K. W. *Mol. Cell. Biol.* **2000**, *20*, 735.

(84) Zhang, X. E.; Morera, S.; Bates, P. A.; Whitehead, P. C.; Coffer, A. I.; Hainbucher, A. I.; Nash, R. A.; Sternberg, M. J.; Lindahl, T.; Freemont, P. S. *EMBO J.* **1998**, *17*, 6404.
(85) Caldecott, K. W.; Aoufouchi, S.; Johnson, P.; Shall, S. *Nucleic Acids Res.* **1996**, *24*, 4387.
(86) Kubota, Y.; Nash, R. A.; Klungland, A.; Schar, P.; Barnes, D. E.; Lindahl, T. *EMBO J.* **1996**, *15*, 6662.
(87) Marsin, S.; Vidal, A. E.; Sossou, M.; Menissier-de Murcia, J.; Le Page, F.; Boiteux, S.; de Murcia, G.; Radicella, J. P. *J. Biol. Chem.* **2003**, *278*, 44068.
(88) Masson, M.; Niedergang, C.; Schreiebr, V.; Muller, S.; Menissier de Murcia, J.; de Murcia, G. *Mol. Cell. Biol.* **1998**, *18*, 3563.
(89) Vidal, A. E.; Boiteux, S.; Hickson, I. D.; Radicella, J. P. *EMBO J.* **2001**, *20*, 6530.
(90) Whitehouse, C. J.; Taylor, R. M.; Thistlethwaite, A.; Zhang, H.; Karimi-Busheri, F.; Lasko, D. D.; Weinfeld, M.; Caldecott, K. W. *Cell* **2001**, *104*, 107.
(91) Wiederhold, L.; Leppard, J. B.; Kedar, P.; Karimi-Busheri, F.; Rasouli-Nia, A.; Weinfeld, M.; Tomkinson, A. E.; Izumi, T.; Prasad, R.; Wilson, S. H.; Mitra, S.; Hazra, T. K. *Mol. Cell* **2004**, 209.
(92) Clements, P. M.; Breslin, C.; Deeks, E. D.; Byrd, P. J.; Ju, L.; Bieganowski, P.; Brenner, C.; Moreira, M. C.; Taylor, A. M.; Caldecott, W. *DNA Repair* **2004**, *3*, 1493.
(93) Date, H.; Igarashi, S.; Sano, Y.; Takahashi, T.; Takahashi, T.; Takano, H.; Tsuji, S.; Nishizawa, M.; Onodera, O. *Biochem. Biphys. Res. Commun.* **2004**, *325*, 1279.
(94) Gueven, N.; Becherel, O. J.; Kijas, A. W.; Chen, P.; Howe, O.; Rudolph, J. H.; Gatti, R.; Date, H.; Onodera, O.; Taucher-Sholz, G.; Lavin, M. F. *Hum. Mol. Genet.* **2004**, *13*, 1081.
(95) Luo, H.; Chan, D. W.; Yang, T.; Rodriguez, M.; Chen, B. P.; Leng, M.; Mu, J. J.; Chen, D.; Songyang, Z.; Wang, Y.; Qin, J. *Mol. Cell. Biol.* **2004**, *24*, 8356.
(96) El-Kamisy, S. F.; Saifi, G. M.; Weinfeld, M.; Johansson, F.; Helleday, T.; Lupski, J. R.; Caldecott, K. W. *Nature* **2005**, *434*, 108.
(97) Leppard, J. B.; Dong, Z.; Mackey, Z. B.; Tomkinson, A. E. *Mol. Cell. Biol.* **2003**, *23*, 5919.
(98) Okano, S.; Kanno, S.; Nakajima, S.; Yasui, A. *J. Biol. Chem.* **2000**, *275*, 32635.
(99) Okano, S.; Lan, L.; Caldecott, K. W.; Mori, T.; Yasui, A. *Mol. Cell. Biol.* **2003**, *23*, 3974.
(100) de Murcia, G.; de Murcia, J. M. *Trends Biochem. Sci*. **1994**, *19*, 172.
(101) Mackey, Z. B.; Niedergang, C.; Murcia, J. M.; Leppard, J.; Au, K.; Chen, J.; de Murcia, G.; Tomkinson, A. E. *J. Biol. Chem.* **1999**, *274*, 21679.
(102) Kulczyk, A. W.; Yang, J. C.; Neuhaus, D. *J. Mol. Biol.* **2004**, *341*, 723.
(103) Grawunder, U.; Zimmer, D.; Kulesza, P.; Lieber, M. R. *J. Biol. Chem.* **1998**, *273*, 24708.
(104) Grawunder, U.; Wilm, M.; Wu, X.; Kulesza, P.; Wilson, T. E.; Mann, M.; Lieber, M. R. *Nature* **1997**, *388*, 492.
(105) Critchlow, S. E.; Bowater, R. P.; Jackson, S. P. *Curr. Biol.* **1997**, *7*, 588.
(106) Modesti, M.; Junop, M. S.; Ghirlando, R.; van de Rakt, M.; Gellert, M.; Yang, W.; Kanaar, R. *J. Mol. Biol.* **2003**, *334*, 215.
(107) Junop, M. S.; Modesti, M.; Guarne, A.; Ghirlando, R.; Gellert, M.; Yang, W. *EMBO J.* **2000**, *19*, 5962.
(108) Sibanda, B. L.; Critchlow, S. E.; Begun, J.; Pei, X. Y.; Jackson, S. P.; Blundell, T. L.; Pellegrini, L. *Nat. Struct. Biol.* **2001**, *8*, 1015.
(109) Frank, K. M.; Sekiguchi, J. M.; Seidl, K. J.; Swat, W.; Rathbun, G. A.; Cheng, H. L.; Davidson, L.; Kangaloo, L.; Alt, F. W. *Nature* **1998**, *396*, 173.
(110) Grawunder, U.; Zimmer, D.; Fugmann, S.; Schwartz, K.; Lieber, M. *Mol. Cell* **1998**, *2*, 477.
(111) O'Driscoll, M.; Cerosaletti, K. M.; Girard, P. M.; Dai, Y.; Stumm, M.; Kysela, B.; Hirsch, B.; Gennery, A.; Palmer, S. E.; Seidel, J.; Gatti, R. A.; Varon, R.; Oettinger, M. A.; Neitzel, H.; Jeggo, P. A.; Concannon, P. *Mol. Cell.* **2001**, *8*, 1175.
(112) Chen, L.; Trujillo, K.; Sung, P.; Tomkinson, A. E. *J. Biol. Chem.* **2000**, *275*, 26196.
(113) DeFazio, L. G.; Stansel, R. M.; Griffith, J. D.; Chu, G. *EMBO J.* **2002**, *21*, 3192.
(114) Yaneva, M.; Kowaleski, T.; Lieber, M. R. *EMBO J.* **1997**, *16*, 5098.
(115) West, R. B.; Yaneva, M.; Lieber, M. R. *Mol. Cell. Biol.* **1998**, *18*, 5908.
(116) Nick McElhinny, S. A.; Snowden, C. M.; McCarville, J.; Ramsden, D. *Mol. Cell. Biol.* **2000**, *20*, 2996.
(117) Ramsden, A.; Gellert, M. *EMBO J.* **1998**, *17*, 609.
(118) Walker, J. R.; Corpina, R. A.; Goldberg, J. *Nature* **2001**, *412*, 607.
(119) Leuther, K. K.; Hammarsten, O.; Kornberg, R. D.; Chu, G. *EMBO J.* **1999**, *18*, 1114.
(120) Herrman, G.; Lindahl, T.; Schar, P. *EMBO J* **1998**, *17*, 4188.
(121) Chen, L.; Trujillo, K.; Ramos, W.; Sung, P.; Tomkinson, A. E. *Mol. Cell.* **2001**, *8*, 1105.
(122) Kramer, K. M.; Brock, J. A.; Bloom, K.; Moore, J. K.; Haber, J. E. *Mol. Cell. Biol.* **1994**, *14*, 1293.
(123) Moore, J. K.; Haber, J. *Mol. Cell. Biol.* **1996**, *16*, 2164.
(124) Roth, D. B.; Wilson, J. H. *Mol. Cell. Biol.* **1986**, *6*, 4295.
(125) Ma, Y.; Lu, H. B.; Tippin, B.; Goodman, M. F.; Shimazaki, N.; Koiwai, O.; Hsieh, C. L.; Schwartz, K.; Lieber, M. R. *Mol. Cell* **2004**, *16*, 701.
(126) Lee, J. W.; Blanco, L.; Zhou, T.; Garcia-Diaz, M.; Bebenek, K.; Kunkel, T. A.; Wang, Z.; Povirk, L. F. *J. Biol. Chem.* **2004**, *279*, 805.
(127) Tseng, H. M.; Tomkinson, A. E. *J. Biol. Chem.* **2002**, *277*, 45630.
(128) Tseng, H.-M.; Tomkinson, A. E. *J. Biol. Chem.* **2004**, *279*, 47588.
(129) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res.* **2003**, *31*, 3381−3385.

CR040498D