Published in final edited form as:

Subcell Biochem. 2012 ; 62: . doi:10.1007/978-94-007-4572-8_17.

DNA ligase I, the replicative DNA ligase

Timothy R.L. Howes¹ and Alan E. Tomkinson²

¹Biomedical Sciences Graduate Program and the Departments of Internal Medicine, University of New Mexico, Albuquerque, NM 87131

²Molecular Genetics and Microbiology, University of New Mexico Cancer Center, University of New Mexico, Albuquerque, NM 87131

Abstract

Multiple DNA ligation events are required to join the Okazaki fragments generated during lagging strand DNA synthesis. In eukaryotes, this is primarily carried out by members of the DNA ligase I family. The C-terminal catalytic region of these enzymes is composed of three domains: a DNA binding domain, an adenylation domain and an OB-fold domain. In the absence of DNA, these domains adopt an extended structure but transition into a compact ring structure when they engage a DNA nick, with each of the domains contacting the DNA. The non-catalytic N-terminal region of eukaryotic DNA ligase I is responsible for the specific participation of these enzymes in DNA replication. This proline-rich unstructured region contains the nuclear localization signal and a PCNA interaction motif that is critical for localization to replication foci and efficient joining of Okazaki fragments. DNA ligase I initially engages the PCNA trimer via this interaction motif which is located at the extreme N-terminus of this flexible region. It is likely that this facilitates an additional interaction between the DNA binding domain and the PCNA ring. The similar size and shape of the rings formed by the PCNA trimer and the DNA ligase I catalytic region when it engages a DNA nick suggest that these proteins interact to form a double-ring structure during the joining of Okazaki fragments. DNA ligase I also interacts with replication factor C, the factor that loads the PCNA trimeric ring onto DNA. This interaction, which is regulated by phosphorylation of the non-catalytic N-terminus of DNA ligase I, also appears to be critical for DNA replication.

Keywords

Lagging strand DNA synthesis; Okazaki fragments; genome instability; cancer predisposition; phosphodiester bond formation

17.1 Introduction

Since DNA polymerases only synthesize DNA in the 5' to 3' direction, one of the two antiparallel strands of duplex DNA is synthesized discontinuously as a series of short Okazaki fragments that are then joined by a DNA ligase to generate an intact strand. In 1967, several laboratories identified DNA ligase activity in extracts from uninfected *E. coli* cells and also from *E. coli* cells infected with bacteriophage T4 (Lehman, 1974). The following year DNA ligase activity was described in extracts from mammalian cells (Soderhall and Lindahl, 1976). Notably, the *Escherichia coli* DNA ligase is NAD⁺-dependent whereas the bacteriophage and mammalian DNA ligases are ATP-dependent (Lehman, 1974, Soderhall and Lindahl, 1976). Subsequent studies have revealed the existence of both NAD⁺- and ATP-dependent DNA ligases in prokaryotes. In contrast,

eukaryotic and viral DNA ligases are almost exclusively ATP-dependent (Ellenberger and Tomkinson 2008, Tomkinson, et al., 2006).

Apart from utilizing a different nucleotide co-factor, the reaction mechanisms of NAD⁺- and ATP-dependent are identical (Fig. 17.1). In the first step, the DNA ligases react with the nucleotide co-factor to form a covalent DNA ligase-adenylate complex in which the AMP moiety is linked to a specific lysine residue via a phosphoramidite bond. When the DNA ligase-adenylate engages a DNA nick with 3' HO and 5' phosphate termini, it transfers the AMP group to the 5' phosphate, forming a covalent DNA adenylate intermediate. Finally, the non-adenylated DNA ligase interacts with the DNA-adenylate, catalyzing phosphodiester bond formation and release of AMP as a result of nucleophilic attack on the 5' DNA adenylate by the 3' HO group.

The first eukaryotic DNA ligase genes were identified in screens for cell division cycle mutants in the yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Johnston and Nasmyth, 1978, Nasmyth, 1977). The DNA ligases encoded by the *CDC9* gene in *Saccharomyces cerevisiae* and the *cdc17⁺* gene in *Schizosaccharomyces pombe* are required for cell viability because of their essential role in DNA replication. Biochemical and immunological characterization of DNA ligase activity in mammalian cell extracts provided the first evidence that eukaryotes contain more than one species of DNA ligase (Soderhall and Lindahl, 1976). The presence of multiple DNA ligase enzymes suggested that these may have distinct cellular functions. In the remainder of this chapter, we will briefly describe the eukaryotic DNA ligases and then focus on DNA ligase I, the replicative DNA ligase, and the mechanisms that underlie the specific participation of this enzyme in DNA replication.

17.2 Eukaryotic DNA ligase genes

As mentioned above, the DNA ligases encoded by the *CDC9* and *cdc17⁺* genes of *S. cerevisiae* and *Schiz. pombe*, respectively, were the first eukaryotic DNA ligases to be identified (Johnston and Nasmyth, 1978, Nasmyth, 1977). Human cDNAs that complemented the temperature-sensitive phenotype of a yeast *cdc9* strain were isolated from a human cDNA library (Barnes, et al., 1990). Subsequent DNA sequencing revealed that these cDNAs encoded a polypeptide that was highly homologous with the yeast DNA ligases and contained sequences that were identical to those of peptides from purified mammalian DNA ligase I (Barnes, et al., 1990). Thus, human DNA ligase I and the yeast DNA ligases are functional homologs that belong to the DNA ligase I family of eukaryotes. Two other mammalian genes that encode DNA ligases, *LIG3* and *LIG4*, have been identified (Chen, et al., 1995, Wei, et al., 1995). Homologs of the *LIG4* gene have been found in all eukaryotes, whereas the *LIG3* gene appears to be restricted to vertebrates (Ellenberger and Tomkinson, 2008). The DNA ligases encoded by the three *LIG* genes share a conserved catalytic region that is flanked by unrelated amino- and/or carboxyl-terminal regions (Fig. 17.2). There is compelling evidence that interactions with specific protein partners mediated by these unique regions flanking the catalytic domain direct the participation of the DNA ligases in different DNA transactions (Ellenberger and Tomkinson, 2008).

17.3 DNA ligase I: molecular genetics and cell biology

The increased expression of the *LIG1* gene when quiescent cells are induced to proliferate and increased levels of DNA ligase I protein and activity in proliferating cells and tissues, implicated DNA ligase I in DNA replication (Petrini, et al., 1991, Soderhall and Lindahl, 1976). This linkage was strengthened by studies showing that DNA ligase I co-localized with replication foci in S-phase cells (Lasko, et al., 1990). The location of distinct amino acid sequences within the non-catalytic N-terminal region of DNA ligase I that function as

nuclear localization and replication foci targeting sequences (Cardoso, et al., 1997, Montecucco, et al., 1995) are shown in Figure 17.2. The mechanism underlying the recruitment of DNA ligase I to replication foci is described below.

A single case of human DNA ligase I-deficiency has been described. This individual, whose symptoms included retarded growth, delayed development, recurrent ear and chest infections and lymphoma, died at age 19 as result of complications following a chest infection (Webster, et al., 1992). Sequencing of genomic DNA revealed the presence of two different mutant *lig1* alleles. The maternally inherited *lig1* allele encodes a DNA ligase polypeptide with reduced catalytic activity whereas the other mutant allele, whose origin is not known, encodes a DNA ligase polypeptide with essentially no catalytic activity. In both mutant *lig1* alleles, the DNA sequence change results in a single amino acid substitution within the conserved catalytic region of DNA ligase I (Barnes, et al., 1992). The locations of the amino acid changes are described in the section below. Primary (46BR) and SV40-immortalized (46BR.1G1) fibroblasts established from the DNA ligase I-deficient individual exhibit defective joining of Okazaki fragments and sensitivity to a wide range of DNA damaging agents, particularly DNA alkylating agents (Teo, et al., 1983). Since both mutant *lig1* alleles are present in the primary fibroblasts whereas only the maternally inherited allele is present in the SV40-immortalized (46BR.1G1) fibroblasts, it appears that the maternal *lig1* allele is responsible for the individual's symptoms and the phenotype of the cell lines (Barnes, et al., 1992). As expected, both the DNA replication and repair defects of the 46BR.1G1 fibroblasts are complemented by expression of wild type DNA ligase I (Levin, et al., 2000).

Although the levels of DNA ligase I protein and activity are reduced by about 50% and 90%, respectively in the 46BR.1G1 fibroblasts compared with SV40-immortalized fibroblasts from a normal individual, there are no significant differences in cell cycle progression despite the defect in converting Okazaki fragments into high molecular weight DNA (Barnes, et al., 1992). In fact, the results of pulse-labeling studies indicate that the majority of Okazaki fragments are degraded rather than ligated (Henderson, et al., 1985, Levin, et al., 2004). Thus, it appears that when DNA ligase I is not available to ligate the nick between adjacent Okazaki fragments, the downstream fragment is displaced by DNA synthesis and then degraded. This model predicts that the lagging strand in 46BR.1G1 fibroblasts is synthesized as a series of longer fragments that are joined by either the defective DNA ligase I polypeptide or one of the other DNA ligases. Based on the results of genetic studies in the yeasts *S. cerevisiae* and *Schiz. pombe* (Johnston and Nasmyth, 1978, Nasmyth, 1977), it was expected that the mammalian *LIG1* gene would be essential. In accordance with this prediction, *lig1* null mouse embryonic stem cells could only be obtained when wild-type DNA ligase I cDNA was ectopically expressed (Petrini, et al., 1995). Surprisingly, *lig1* null embryos generated by crossing heterozygous mice were detectable until day 16 (Bentley, et al., 1996, Bentley, et al., 2002). Furthermore, it was possible to establish *lig1* null embryonic fibroblasts (MEFs) from these embryos, demonstrating that *LIG1* is not an essential gene in mouse somatic cells. Similar to the human 46BR.1G1 fibroblasts, the *lig1* null MEFs had a defect in converting Okazaki fragments into high molecular weight DNA but no defect in proliferation (Bentley, et al., 1996, Bentley, et al., 2002). In contrast to the 46BR.1G1 fibroblasts, the *lig1* null MEFs have no apparent DNA repair defect (Bentley, et al., 1996, Bentley, et al., 2002, Teo, et al., 1983). Thus, it appears that one of the other mammalian DNA ligases can substitute for DNA ligase I in DNA replication. The presence of additional DNA ligases in mammals encoded by the *LIG3* gene provides a possible explanation as to why the *LIG1* gene homolog is essential for cell viability in yeast but not in mammals. For example, DNA ligase III α and its partner protein XRCC1 are recruited to participate in the repair of DNA single strand breaks by an interaction with the poly(ADP-ribosylated) version of poly(ADP-ribose)

polymerase 1 (PARP-1), an abundant nuclear protein that binds to and is activated by DNA single strand breaks (Okano, et al., 2003, Okano, et al., 2005). Since the defect in Okazaki fragment processing caused by DNA ligase I deficiency is likely to result in relatively long-lived, single-strand interruptions on the lagging strand, it is possible that these breaks are recognized and joined by the PARP-1/DNA ligase III α single-strand break repair pathway. The sensitivity of human 46BR.1G1 fibroblasts and *lig1* null MEFs to a PARP-1 inhibitor (Lehmann, et al., 1988) is consistent with the single-strand break repair pathway joining single-strand breaks between Okazaki fragments that remain after lagging strand DNA synthesis.

It is likely that cases of DNA ligase I deficiency in humans will be extremely rare. A mouse model that reiterates the mutation in human 46BR.1G1 cells has been generated. These animals are small and have hematopoietic defects (Harrison, et al., 2002). Other notable features of these animals include increased genomic instability and an increased incidence of epithelial tumors (Harrison, et al., 2002). As with the *lig1* null MEFs, the MEFs harboring the equivalent mutation to that in the human 46BR.1G1 fibroblasts have a defect in replication but not repair, suggesting that the increased genome instability and cancer incidence is due to accumulation of abnormal replication intermediates (Harrison, et al., 2002). Together, these studies indicate that DNA ligase I plays an important role in DNA repair in human but not mouse cells (Harrison, et al., 2002, Teo, et al., 1983). It is conceivable that, while there is functional redundancy between DNA ligases I and III α in DNA repair as well as in DNA replication, the relative contribution of DNA ligase III α -dependent repair may be greater in mouse cells compared with human cells. However, recent studies with mouse *lig3* null MEFs do not support this model (Gao, et al., 2011, Simsek, et al., 2011).

17.4 DNA ligase I protein: structure and function

The 919 amino acid polypeptide encoded by human DNA ligase I cDNA has a highly asymmetric shape, which causes anomalous behavior during density gradient sedimentation and gel filtration experiments (Tomkinson, et al., 1990). Using limited proteolysis, it was found that catalytic activity resides within a relatively protease-resistant C-terminal fragment of about 78 kDa whereas the N-terminal fragment is extremely protease sensitive, indicative of an unstructured region (Tomkinson, et al., 1990). Notably, this N-terminal region is likely to have an extended, flexible conformation because it contains a large number of proline residues (Barnes, et al., 1990). In addition, the high proline content of DNA ligase I (approximately 9%) results in anomalous mobility during SDS-polyacrylamide gel electrophoresis, giving DNA ligase I an apparent molecular mass of 125 kDa (Tomkinson, et al., 1990).

The catalytic region of DNA ligase I contains six motifs (Fig. 17.2) that are conserved among the nucleotidyl transferase family that includes mRNA capping enzymes and RNA ligases in addition to DNA ligases (Shuman and Schwer, 1995). Motif I contains the active site lysine residue to which the AMP (or GMP) residue is covalently via a phosphoramidite bond (Fig. 17.2). This residue was initially identified by sequencing an adenylated tryptic peptide from bovine DNA ligase I (Tomkinson, et al., 1991). Using this sequence, it was possible to predict the position of the putative active site lysine residues in DNA ligases, RNA ligases and mRNA capping enzymes. As expected, substitution of Lys568, the lysine residue that binds the AMP moiety in human DNA ligase I (Fig. 17.2), prevents formation of the enzyme-AMP complex and therefore abolishes enzymatic activity (Kodama, et al., 1991, Tomkinson, et al., 1991). Interestingly, the mutated DNA ligase I enzyme identified in the DNA ligase I-deficient individual has a lysine residue instead of a glutamic acid at

position 566 (Barnes, et al., 1992). This amino acid change two residues away from the active site lysine markedly reduces formation of the enzyme-AMP intermediate.

The first nucleotidyl transferase structure to be determined was that of the DNA ligase encoded by bacteriophage T7 (Subramanya, et al., 1996). This was shortly followed by the structure of the RNA-capping enzyme from PCBV-1 (Hakansson, et al., 1997). These structures revealed the existence of the two domains: an adenylation/guanylylation domain, which contains conserved motifs I through V, and an oligomer binding-fold (OB) domain (OBD) containing motif VI. In 2004, Pascal *et al.* successfully crystallized the catalytic C-terminal region of DNA ligase I (residues 233-919) bound to a nicked DNA substrate (Pascal, et al., 2004). This structure revealed several novel features. Firstly, it showed that nicked DNA is encircled by DNA ligase I during catalysis (Fig. 17.3B), suggesting that the catalytic domain undergoes a large conformational change when it engages a nick. Secondly, it showed that the catalytic region of the larger eukaryotic DNA ligases contains a DNA binding domain in addition to the adenylation and OB-fold domains that make up the catalytic core (Fig. 17.2). Thus, the adenylation/guanylylation and OB-fold domains constitute the conserved catalytic core of nucleotidyl transferases with the DNA binding domain (DBD) being a characteristic feature of eukaryotic DNA ligases (Pascal, et al., 2004).

The DNA binding domain of DNA ligase I spans residues 262 to 534 (Fig. 16.2). As shown in Figure 17.3B, these residues fold into twelve α -helices that exhibit a two-fold symmetry (Pascal, et al., 2004). Due to the symmetry of the DBD, the interaction with DNA occurs via one tight reverse turn of two α -helices and an extended loop formation. These loops and helices arrange to create a relatively flat surface of approximately 2000 Å² that interacts almost exclusively with the phosphodiester backbone of the DNA substrate. The DBD interacts with the minor groove of the DNA backbone on both sides of the nick, explaining how DNA ligase I binds to DNA in a sequence-independent binding manner and why chemicals that bind to the minor groove of DNA, such as distamycin, inhibit DNA ligase activity (Montecucco, et al., 1991). Notably, the DBD stimulates the weak DNA joining activity of the DNA ligase I catalytic core containing the AdD and OBD when added *in trans*, indicating that contacts between the DBD and both AdD and the OBD observed in the crystal structure stabilize the folding of the catalytic core around the DNA nick (Pascal, et al., 2004).

The DNA ligase I adenylation domain, which spans from residue 535 to 747, contains conserved motifs I, III, IIIa, IV and V. These five motifs contribute to the surface of nucleotide binding pocket (Fig. 17.3B). Tryptophan 742 of motif V provides co-factor specificity by sterically excluding GTP. Furthermore, Arg573 and Glu621 of motifs I and III respectively, stabilize the hydroxyl groups on the ribose sugar of AMP via hydrogen bonding interactions (Pascal, et al., 2004). As mentioned above, one of the two mutant *ligI* alleles identified in the individual with DNA ligase I deficiency encodes a polypeptide in which Glu566 of motif I is replaced by a lysine residue (Barnes, et al., 1992). From the structure of DNA ligase I, it is evident that the Glu566 residue contributes to the specific interaction with ATP by forming a hydrogen-bond with the N6 of the adenine moiety (Pascal, et al., 2004). Replacement of Glu566 with a positively charged lysine residues disrupts this, providing an explanation as to why the mutant polypeptide is defective in the first step of the ligation reaction, formation of the covalent enzyme-adenylate intermediate.

The major structural feature of the OB-fold domain is a β -barrel and similar to the other two domains, it also interacts with the minor groove of the DNA (Pascal, et al., 2004). Contacts between AdD and OBD are critical for correctly positioning these domains when they engage a DNA nick. During catalysis, the AdD forms a salt bridge with the OBD via

Asp570 and Arg871, stabilizing the ligase catalytic domains in a conformation in which they fully encircle the DNA nick. This positioning of the AdD and OBD creates a surface that binds to and distorts the nicked DNA. Notably, the phenylalanine residues at positions 635 and 872 of the AdD and OBD are forced into the minor groove both 3' and 5' to the nick. As a result of these interactions, the DNA duplex upstream of the nick duplex assumes an A-form structure and the nick is opened up for ligation (Pascal, et al., 2004). Notably, the DNA binding site downstream of the nick is specific for B-form DNA, explaining why ligase I is not active on nicks within A-form duplexes formed by RNA duplexes and RNA-DNA hybrids (Pascal, et al., 2004). This ability to discriminate against duplexes containing ribonucleotides 5' to the nick presumably prevents premature joining of Okazaki fragments before the RNA primer has been removed. The maternally inherited mutant *lig1* allele in the individual with DNA ligase I-deficiency encodes a polypeptide in which the arginine 771 within the OBD is replaced by a tryptophan residue (Barnes, et al., 1992). This mutant enzyme has markedly reduced catalytic activity and is, as expected, defective in step 2 of the ligation reaction (Fig. 17.1), transfer of the AMP moiety from the ligase to the 5' phosphate termini of the DNA nick (Prigent, et al., 1994).

Although eukaryotic DNA ligases have not been crystallized without nicked DNA, the structure of an ATP-dependent DNA ligase from the archaeal organism *Sulfolobus solfataricus* has been determined in the absence of DNA (Pascal, et al., 2006). As shown in Figure 17.3A, the catalytic region of the archaeal enzyme has the same three domain organization as eukaryotic DNA ligases but, in the absence of DNA, the three domains are arranged in an extended conformation. The major difference between the extended (Fig. 17.3A) and closed (Fig. 17.3B) conformations of the three domains is the position of the OBD domain relative to the other two domains (Pascal, et al., 2006). Thus, the OBD undergoes a large change in conformation during the nicked DNA-dependent transition from the extended to the closed form with interactions between the OBD and DBD playing key roles in stabilizing the closed form. Based on structures of smaller DNA ligases, it appears likely that the OBD also undergoes conformational changes when the enzyme interacts with ATP to form the enzyme-adenylate, possibly reorienting the OBD to expose a DNA binding surface (Odell, et al., 2000, Subramanya, et al., 1996). While the unstructured N-terminal region of DNA ligase I is dispensable for catalytic activity *in vitro*, it does appear to be essential for cellular function and viability (Mackenney, et al., 1997, Petrini, et al., 1995). As mentioned previously, this region contains a bipartite nuclear localization signal located between residues 111 and 179 and a sequence that is required for targeting to replication factories, residues 2 to 9 (Cardoso, et al., 1997, Montecucco, et al., 1995). In addition, the N-terminal region is phosphorylated on several serine/threonine residues (Fig. 17.2) by casein kinase II and cyclin-dependent kinases during cell cycle progression (Ferrari, et al., 2003, Frouin, et al., 2002, Prigent, et al., 1992). This results in a hyperphosphorylated form of DNA ligase I in M-phase cells. It appears likely that these phosphorylation events regulate the participation of DNA ligase I in DNA replication because phosphorylation site mutants fail to correct the DNA replication defect of DNA ligase I-deficient 46BR cells (Soza, et al., 2009, Vijayakumar, et al., 2009).

17.5 DNA ligase I: protein interactions

Proteins are directed to participate in complex DNA transactions, such as DNA replication, by specific protein-protein interactions. Proliferating cell nuclear antigen (PCNA), the eukaryotic homotrimeric DNA sliding clamp (see Chapter 15, this volume) that functions as a processivity factor for the replicative DNA polymerases, was the first DNA ligase I-interacting protein to be identified (Levin, et al., 1997), and mediates DNA ligase I interaction with the rest of the eukaryotic replisome. Residues 2 to 9 within the non-catalytic N-terminal region of DNA ligase I constitute the major PCNA binding site within DNA

ligase I (Montecucco, et al., 1998). Notably, this same sequence, which is homologous to a PCNA-interacting protein motif, or 'PIP box,' that has been identified in a large number of proteins, is required for the recruitment of DNA ligase I to replication factories (Montecucco, et al., 1998). Amino acid changes that disrupt PCNA binding abolish both the recruitment of DNA ligase I to replication factories and the correction of the DNA replication defect in 46BR cells, demonstrating the critical role of this interaction in the subnuclear targeting of DNA ligase I and the efficient joining of Okazaki fragments (Levin, et al., 2000, Montecucco, et al., 1998). Since the PIP box motif binds to the interdomain connector loop of PCNA (Vijayakumar, et al., 2007), it appears likely that, when the flexible N-terminal region of DNA ligase I initially binds to the interdomain connector loop of PCNA trimer, the catalytic region remains in an extended conformation (Fig. 17.4A) (Pascal, et al., 2006).

DNA ligase I stably interacts with PCNA trimers that are topologically linked to duplex DNA with only one molecule of DNA ligase I bound per PCNA trimer (Levin, et al., 1997). This suggests that the other potential binding sites are occluded because of dynamic conformational changes in DNA ligase I as a consequence of its flexible, extended structure. Alternatively, it is possible that the initial docking of DNA ligase I with a PCNA trimer via the PIP box facilitates lower affinity interactions that extend the protein-protein interaction interface. In support of this idea, the DBD binds weakly to the subunit-subunit interface region of homotrimeric PCNA (Song, et al., 2009). Interestingly, the DBD also mediates the interaction with human Rad9-Rad1-Hus1, a heterotrimeric clamp involved in cell cycle checkpoints (Song, et al., 2009) and heterotrimeric PCNA from *Sulfolobus solfataricus* (Pascal, et al., 2006). Given the similarity in size and shape between the PCNA ring and the ring structure formed when the catalytic region of DNA ligase I engages a DNA nick (Fig. 17.4B), the PCNA ring may facilitate the transition of the catalytic region of DNA ligase I from the extended conformation to the compact ring structure. It has been proposed that the DBD, which provides the majority of the DNA binding affinity, serves as a pivot during this transition after initial docking via the PIP box (Fig. 17.4A). Unlike studies on the interaction between DNA ligase I and the heterotrimeric Rad9-Rad1-Hus1 checkpoint DNA sliding clamp (Song, et al., 2007, Wang, et al., 2006), there are contradictory reports as to whether the interaction with PCNA stimulates nick joining by DNA ligase I (Levin, et al., 1997, Levin, et al., 2004, Tom, et al., 2001).

DNA ligase I also functionally interacts with two other DNA replication proteins, replication protein A (RPA), the heterotrimeric complex that binds to single stranded DNA (see Chapter 10, this volume, and Ranalli, et al., 2002), and replication factor C (RFC), a heteropentameric complex that loads PCNA onto DNA (see Chapter 14, this volume, and Levin, et al., 2004). Although a direct physical interaction between RPA and DNA ligase I has not been demonstrated, RPA specifically stimulates the rate of catalysis by DNA ligase I (Ranalli, et al., 2002). In contrast to RPA, RFC inhibits DNA joining by DNA ligase I (Levin, et al., 2004). This interaction and inhibition, which involves the large subunit of RFC, p140/RFC1, is abolished by replacement of the four serine residues that are phosphorylated in DNA ligase I with glutamic acid residues (Vijayakumar, et al., 2009). Notably, the inhibition of DNA ligase I by RFC can also be alleviated by inclusion of PCNA in the reaction, providing that DNA ligase I has a functional PIP box (Vijayakumar, et al., 2009). Unlike the interaction with RFC, DNA ligase I binding to PCNA is not modulated by phosphorylation (Vijayakumar, et al., 2009). Thus, the failure of the phosphorylation site mutant of DNA ligase I to complement the replication defect in DNA ligase I-deficient cells may be due to the disrupted interaction with RFC (Vijayakumar, et al., 2009). Although these studies indicate that physical and functional interactions among DNA ligase I, RFC and PCNA are critical for DNA replication, the mechanisms by which these interactions contribute to Okazaki fragment processing and joining are not fully understood.

17.6 Concluding remarks

The catalytic region of DNA ligase I encircles and ligates the nicks between adjacent Okazaki fragments during lagging strand DNA synthesis. Although there is compelling evidence that interactions with PCNA and almost certainly RFC are critical for the specific participation of DNA ligase I in DNA replication, further studies are needed to delineate the precise molecular mechanisms by which these interactions contribute to the co-ordinated processing and joining of Okazaki fragments and how these interactions are regulated by phosphorylation of DNA ligase I. The cancer predisposition exhibited by a mouse model of DNA ligase I-deficiency highlight the importance of this co-ordination and regulation in preventing genome instability during DNA replication.

References

- Barnes DE, Johnston LH, Kodama K, Tomkinson AE, Lasko DD, Lindahl T. Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*. 1990; 87:6679–6683. [PubMed: 2204063]
- Barnes DE, Tomkinson AE, Lehmann AR, Webster AD, Lindahl T. Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell*. 1992; 69:495–503. [PubMed: 1581963]
- Bentley D, Selfridge J, Millar JK, Samuel K, Hole N, Ansell JD, Melton DW. DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability. *Nat Genet*. 1996; 13:489–491. [PubMed: 8696349]
- Bentley DJ, Harrison C, Ketchen AM, Redhead NJ, Samuel K, Waterfall M, Ansell JD, Melton DW. DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability. *J Cell Sci*. 2002; 115:1551–1561. [PubMed: 11896201]
- Cardoso MC, Joseph C, Rahn HP, Reusch R, Nadal-Ginard B, Leonhardt H. Mapping and use of a sequence that targets DNA ligase I to sites of DNA replication in vivo. *J Cell Biol*. 1997; 139:579–587. [PubMed: 9348276]
- Chen J, Tomkinson AE, Ramos W, Mackey ZB, Danehower S, Walter CA, Schultz RA, Besterman JM, Husain I. Mammalian DNA ligase III: molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination. *Mol Cell Biol*. 1995; 15:5412–5422. [PubMed: 7565692]
- Ellenberger T, Tomkinson AE. Eukaryotic DNA ligases: structural and functional insights. *Annu Rev Biochem*. 2008; 77:313–338. [PubMed: 18518823]
- Ferrari G, Rossi R, Arosio D, Vindigni A, Biamonti G, Montecucco A. Cell cycle-dependent phosphorylation of human DNA ligase I at the cyclin-dependent kinase sites. *J Biol Chem*. 2003; 278:37761–37767. [PubMed: 12851383]
- Frouin I, Montecucco A, Biamonti G, Hubscher U, Spadari S, Maga G. Cell cycle-dependent dynamic association of cyclin/Cdk complexes with human DNA replication proteins. *EMBO J*. 2002; 21:2485–2495. [PubMed: 12006500]
- Gao Y, Katyal S, Lee Y, Zhao J, Rehgle JE, Russell HR, McKinnon PJ. DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. *Nature*. 2011; 471:240–244. [PubMed: 21390131]
- Hakansson K, Doherty AJ, Shuman S, Wigley DB. X-ray crystallography reveals a large conformational change during guanylyl transfer by mRNA capping enzymes. *Cell*. 1997; 89:545–553. [PubMed: 9160746]
- Harrison C, Ketchen AM, Redhead NJ, O'Sullivan MJ, Melton DW. Replication failure, genome instability, and increased cancer susceptibility in mice with a point mutation in the DNA ligase I gene. *Cancer Res*. 2002; 62:4065–4074. [PubMed: 12124343]
- Henderson LM, Arlett CF, Harcourt SA, Lehmann AR, Broughton BC. Cells from an immunodeficient patient (46BR) with a defect in DNA ligation are hypomutable but hypersensitive to the induction of sister chromatid exchanges. *Proc Natl Acad Sci U S A*. 1985; 82:2044–2048. [PubMed: 3856882]

- Johnston LH, Nasmyth KA. *Saccharomyces cerevisiae* cell cycle mutant *cdc9* is defective in DNA ligase. *Nature*. 1978; 274:891–893. [PubMed: 355897]
- Kodama K, Barnes DE, Lindahl T. *In vitro* mutagenesis and functional expression in *Escherichia coli* of a cDNA encoding the catalytic domain of human DNA ligase I. *Nucleic Acids Res*. 1991; 19:6093–6099. [PubMed: 1956768]
- Lasko DD, Tomkinson AE, Lindahl T. Mammalian DNA ligases. Biosynthesis and intracellular localization of DNA ligase I. *J Biol Chem*. 1990; 265:12618–12622. [PubMed: 2197279]
- Lehman IR. DNA ligase: structure, mechanism, and function. *Science*. 1974; 186:790–797. [PubMed: 4377758]
- Lehmann AR, Willis AE, Broughton BC, James MR, Steingrimsdottir H, Harcourt SA, Arlett CF, Lindahl T. Relation between the human fibroblast strain 46BR and cell lines representative of Bloom's syndrome. *Cancer Res*. 1988; 48:6343–6347. [PubMed: 3180052]
- Levin DS, Bai W, Yao N, O'Donnell M, Tomkinson AE. An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. *Proc Natl Acad Sci U S A*. 1997; 94:12863–12868. [PubMed: 9371766]
- Levin DS, McKenna AE, Motycka TA, Matsumoto Y, Tomkinson AE. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Curr Biol*. 2000; 10:919–922. [PubMed: 10959839]
- Levin DS, Vijayakumar S, Liu X, Bermudez VP, Hurwitz J, Tomkinson AE. A conserved interaction between the replicative clamp loader and DNA ligase in eukaryotes: implications for Okazaki fragment joining. *J Biol Chem*. 2004; 279:55196–55201. [PubMed: 15502161]
- Mackenney VJ, Barnes DE, Lindahl T. Specific function of DNA ligase I in simian virus 40 DNA replication by human cell-free extracts is mediated by the amino-terminal non-catalytic domain. *J Biol Chem*. 1997; 272:11550–11556. [PubMed: 9111070]
- Montecucco A, Fontana M, Focher F, Lestingi M, Spadari S, Ciarrocchi G. Specific inhibition of human DNA ligase adenylation by a distamycin derivative possessing antitumor activity. *Nucleic Acids Res*. 1991; 19:1067–1072. [PubMed: 1708493]
- Montecucco A, Rossi R, Levin DS, Gary R, Park MS, Motycka TA, Ciarrocchi G, Villa A, Biamonti G, Tomkinson AE. DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. *EMBO J*. 1998; 17:3786–3795. [PubMed: 9649448]
- Montecucco A, Savini E, Weighardt F, Rossi R, Ciarrocchi G, Villa A, Biamonti G. The N-terminal domain of human DNA ligase I contains the nuclear localization signal and directs the enzyme to sites of DNA replication. *EMBO J*. 1995; 14:5379–5386. [PubMed: 7489727]
- Nasmyth KA. Temperature-sensitive lethal mutants in the structural gene for DNA ligase in the yeast *Schizosaccharomyces pombe*. *Cell*. 1977; 12:1109–1120. [PubMed: 597860]
- Odell M, Sriskanda V, Shuman S, Nikolov DB. Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining. *Mol Cell*. 2000; 6:1183–1193. [PubMed: 11106756]
- Okano S, Lan L, Caldecott KW, Mori T, Yasui A. Spatial and temporal cellular responses to single-strand breaks in human cells. *Mol Cell Biol*. 2003; 23:3974–3981. [PubMed: 12748298]
- Okano S, Lan L, Tomkinson AE, Yasui A. Translocation of XRCC1 and DNA ligase III α from centrosomes to chromosomes in response to DNA damage in mitotic human cells. *Nucleic Acids Res*. 2005; 33:422–429. [PubMed: 15653642]
- Pascal JM, O'Brien PJ, Tomkinson AE, Ellenberger T. Human DNA ligase I completely encircles and partially unwinds nicked DNA. *Nature*. 2004; 432:473–478. [PubMed: 15565146]
- Pascal JM, Tsodikov OV, Hura GL, Song W, Cotner EA, Classen S, Tomkinson AE, Tainer JA, Ellenberger T. A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol Cell*. 2006; 24:279–291. [PubMed: 17052461]
- Petrini JH, Huwiler KG, Weaver DT. A wild-type DNA ligase I gene is expressed in Bloom's syndrome cells. *Proc Natl Acad Sci U S A*. 1991; 88:7615–7619. [PubMed: 1881902]
- Petrini JH, Xiao Y, Weaver DT. DNA ligase I mediates essential functions in mammalian cells. *Mol Cell Biol*. 1995; 15:4303–4308. [PubMed: 7623824]

- Prigent C, Lasko DD, Kodama K, Woodgett JR, Lindahl T. Activation of mammalian DNA ligase I through phosphorylation by casein kinase II. *EMBO J.* 1992; 11:2925–2933. [PubMed: 1639065]
- Prigent C, Satoh MS, Daly G, Barnes DE, Lindahl T. Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. *Mol Cell Biol.* 1994; 14:310–317. [PubMed: 8264597]
- Ranalli TA, DeMott MS, Bambara RA. Mechanism underlying replication protein a stimulation of DNA ligase I. *J Biol Chem.* 2002; 277:1719–1727. [PubMed: 11698410]
- Shuman S, Schwer B. RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. *Mol Microbiol.* 1995; 17:405–410. [PubMed: 8559059]
- Simsek D, Furda A, Gao Y, Artus J, Brunet E, Hadjantonakis AK, Van Houten B, Shuman S, McKinnon PJ, Jasin M. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature.* 2011; 471:245–248. [PubMed: 21390132]
- Soderhall S, Lindahl T. DNA ligases of eukaryotes. *FEBS Lett.* 1976; 67:1–8. [PubMed: 782919]
- Song W, Levin DS, Varkey J, Post S, Bermudez VP, Hurwitz J, Tomkinson AE. A conserved physical and functional interaction between the cell cycle checkpoint clamp loader and DNA ligase I of eukaryotes. *J Biol Chem.* 2007; 282:22721–22730. [PubMed: 17561505]
- Song W, Pascal J, Ellenberger T, Tomkinson AE. The DNA binding domain of human DNA ligase I interacts with both nicked DNA and the DNA sliding clamps, PCNA and hRad9-hRad1-hHus1. *DNA Repair (Amst).* 2009; 8:912–919. [PubMed: 19523882]
- Soza S, Leva V, Vago R, Ferrari G, Mazzini G, Biamonti G, Montecucco A. DNA ligase I deficiency leads to replication-dependent DNA damage and impacts cell morphology without blocking cell cycle progression. *Mol Cell Biol.* 2009; 29:2032–2041. [PubMed: 19223467]
- Subramanya HS, Doherty AJ, Ashford SR, Wigley DB. Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. *Cell.* 1996; 85:607–615. [PubMed: 8653795]
- Teo IA, Arlett CF, Harcourt SA, Priestley A, Broughton BC. Multiple hypersensitivity to mutagens in a cell strain (46BR) derived from a patient with immuno-deficiencies. *Mutat Res.* 1983; 107:371–386. [PubMed: 6408472]
- Tom S, Henricksen LA, Park MS, Bambara RA. DNA ligase I and proliferating cell nuclear antigen form a functional complex. *J Biol Chem.* 2001; 276:24817–24825. [PubMed: 11331287]
- Tomkinson AE, Lasko DD, Daly G, Lindahl T. Mammalian DNA ligases. Catalytic domain and size of DNA ligase I. *J Biol Chem.* 1990; 265:12611–12617. [PubMed: 1695631]
- Tomkinson AE, Totty NF, Ginsburg M, Lindahl T. Location of the active site for enzyme-adenylate formation in DNA ligases. *Proc Natl Acad Sci U S A.* 1991; 88:400–404. [PubMed: 1988940]
- Tomkinson AE, Vijayakumar S, Pascal JM, Ellenberger T. DNA ligases: structure, reaction mechanism, and function. *Chem Rev.* 2006; 106:687–699. [PubMed: 16464020]
- Vijayakumar S, Chapados BR, Schmidt KH, Kolodner RD, Tainer JA, Tomkinson AE. The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase. *Nucleic Acids Res.* 2007; 35:1624–1637. [PubMed: 17308348]
- Vijayakumar S, Dziegielewska B, Levin DS, Song W, Yin J, Yang A, Matsumoto Y, Bermudez VP, Hurwitz J, Tomkinson AE. Phosphorylation of human DNA ligase I regulates its interaction with replication factor C and its participation in DNA replication and DNA repair. *Mol Cell Biol.* 2009; 29:2042–2052. [PubMed: 19223468]
- Wang W, Lindsey-Boltz LA, Sancar A, Bambara RA. Mechanism of stimulation of human DNA ligase I by the Rad9-Rad1-Hus1 checkpoint complex. *J Biol Chem.* 2006; 281:20865–20872. [PubMed: 16731526]
- Webster AD, Barnes DE, Arlett CF, Lehmann AR, Lindahl T. Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene. *Lancet.* 1992; 339:1508–1509. [PubMed: 1351188]
- Wei YF, Robins P, Carter K, Caldecott KW, Papin DJC, Yu G-L, Wang R-P, Shell BK, Nash RA, Schar P, Barnes DE, Haseltine WA, Lindahl T. Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and genetic recombination. *Mol Cell Biol.* 1995; 15:3206–3216. [PubMed: 7760816]

Figure 17.1. Three-step mechanism of DNA ligation

(1) DNA ligase I binds and hydrolyses adenosine triphosphate (ATP), releasing pyrophosphate (PP_i) and a covalent ligase-adenosine monophosphate (AMP) intermediate. (2) The AMP group is subsequently transferred from the ligase polypeptide to the 5' phosphate termini of a nick in duplex DNA. (3) The non-adenylyated ligase catalyzes phosphodiester bond formation in a reaction involving nucleophilic attack by the 3'HO group and release of AMP.

2

A

Human DNA ligase I

Human DNA ligase III

B

Lig1 DBD

Lig3 DBD

Lig4 DBD

Figure 17.2. Alignment of the polypeptides encoded by the three human *LIG* genes

The conserved catalytic regions of the human DNA ligases each contain a DNA binding (DBD, red), adenylation (Add, green) and OB-fold (OBD, yellow) domain. The Add and OBD, which make up the catalytic core of the nucleotidyl transferase family that also includes RNA ligases and mRNA capping enzymes, contain six highly conserved motifs (I, III, IIIa, IV, V and VI). The position of the active site lysine residue within motif I that forms the covalent bond with AMP is indicated for each DNA ligase. The non-catalytic regions that flank the DNA ligase catalytic region determine the subcellular distribution and cellular functions of the DNA ligases. The positions of the nuclear localization signals (NLS, blue) of DNA ligase I and DNA ligase III β , and the mitochondrial leader sequence (MLS, cyan) of DNA ligase III α are shown. The replication factory targeting sequence (RFTS), which is also a PCNA interacting peptide (PIP) box that targets DNA ligase I replication and interacts with PCNA, is indicated in grey. Sites of phosphorylation on serine residues within the non-catalytic N-terminal region of DNA ligase I are indicated. Amino acids, glutamine 566 and arginine 771, are replaced with lysine and tryptophan, respectively in the polypeptides encoded by the mutant *ligI* alleles of the only known DNA ligase I deficient human identified to date. All the DNA ligases encoded by the *LIG3* gene have an N-terminal zinc finger (orange) that is involved in DNA binding. Both DNA ligase III α and DNA ligase IV contain a breast and ovarian cancer susceptibility protein-1 C-terminal motifs (BRCT, dark green) domain that are involved in protein-protein interactions.

Figure 17.3. DNA binding induces a large conformational change in the DNA ligase catalytic region

(**A**) In the crystal structure of *Sulfolobus solfataricus* DNA ligase obtained in the absence of DNA (Pascal, et al., 2006), the DNA binding (DBD, shown in red), adenylation domain (AdD, green), and OB-fold (OBD, yellow) domains are in an extended conformation with relatively few contacts between the domains. (**B**) In the crystal structure of human DNA ligase I in complex with nicked DNA (Pascal, et al., 2004), the DNA binding (DBD, red), adenylation domain (AdD, green) and OB-fold (OBD, yellow) domains encircle the nicked DNA forming a ring structure that is stabilized by each domains interacting with the DNA and by contacts between the DNA binding and OB-fold domains.

Figure 17.4. Models for the interaction between PCNA and DNA ligase I on nicked DNA
(A) DNA ligase I initially engages a PCNA trimer (orange) by an interaction between the DNA ligase I PIP box (gray) and the interdomain connector loop (IDCL) region of the one of PCNA subunits. This docking facilitates a lower affinity interaction between the DNA ligase I DNA binding domain (red) and the PCNA trimer. At this stage, the DNA ligase I catalytic region remains in an extended conformation. This complex of PCNA and DNA ligase I slides freely along the duplex until it encounters a nick. The catalytic region of DNA ligase I then encircles the DNA nick. Interactions with the face of the PCNA trimer are likely to facilitate the transition of the catalytic region from the extended conformation into the compact ring structure. Domains of DNA ligase I are coloured as Figure 17.3. **(B)** A theoretical space-filling model of the double ring structure formed by PCNA (shown predominantly in orange but with the IDCL highlighted in blue; PDB: 1AXC) and the catalytic region of DNA ligase I (PDB: 1X9N) on nicked duplex DNA.