

Human DNA Ligase IV and the Ligase IV/XRCC4 Complex: Analysis of Nick Ligation Fidelity[†]

Yu Wang,[‡] Brandon J. Lamarche,[‡] and Ming-Daw Tsai*

Departments of Chemistry and Biochemistry, The Biophysics Program, and The Ohio State Biochemistry Program, The Ohio State University, Columbus, Ohio 43210, and Genomics Research Center, Academia Sinica, Taiwan

Received October 16, 2006; Revised Manuscript Received February 16, 2007

ABSTRACT: In addition to linking nicked/fragmented DNA molecules back into a contiguous duplex, DNA ligases also have the capacity to influence the accuracy of DNA repair pathways via their tolerance/intolerance of nicks containing mismatched base pairs. Although human DNA ligase I (Okazaki fragment processing) and the human DNA ligase III/XRCC1 complex (general DNA repair) have been shown to be relatively intolerant of nicks containing mismatched base pairs, the human DNA ligase IV/XRCC4 complex has not been studied in this regard. Ligase IV/XRCC4 is the sole DNA ligase involved in the repair of double strand breaks (DSBs) via the non-homologous end joining (NHEJ) pathway. During the repair of DSBs generated by chemical/physical damage as well as the repair of the programmed DSB intermediates of V(D)J recombination, there are scenarios where, at least conceptually, a capacity for ligating nicks containing mismatched base pairs would appear to be advantageous. Herein we examine whether ligase IV/XRCC4 can contribute a mismatched nick ligation activity to NHEJ. Toward this end, we (i) describe an *E. coli*-based coexpression system that provides relatively high yields of the ligase IV/XRCC4 complex, (ii) describe a unique rate-limiting step, which has bearing on how the complex is assayed, (iii) specifically analyze how XRCC4 influences ligase IV catalysis and substrate specificity, and (iv) probe the mismatch tolerance/intolerance of DNA ligase IV/XRCC4 via quantitative *in vitro* kinetic analyses. Analogous to most other DNA ligases, ligase IV/XRCC4 is shown to be fairly intolerant of nicks containing mismatched base pairs. These results are discussed in light of the biological roles of NHEJ.

In addition to sealing broken DNA back into a contiguous duplex, DNA ligases also have the capacity to influence the accuracy of the repair pathways they are involved in. This is well illustrated by the example of mammalian base excision repair (BER¹), where nicks containing 3'-OH mismatched base pairs, resulting from aberrant gap-filling by DNA polymerase β (Pol β), are sealed inefficiently by the DNA ligase III/XRCC1 complex (1). This delay in mismatch ligation is expected to provide a greater window of opportunity for nick editing by the 3' \rightarrow 5' exonuclease activity of APE1 (1, 2), the predominant mammalian AP endonuclease (APE).

Though a large body of qualitative and quantitative data suggests that DNA ligases from a broad spectrum of organisms are intolerant of non-Watson–Crick base pairing

schemes at the 3'-OH terminus of a nick (3 and references therein), our recent analysis of the African swine fever virus (ASFV)-encoded DNA ligase demonstrated that analogous to the recently discovered error-prone DNA polymerases, there are exceptions to this rule (3). The ASFV DNA ligase is not only tolerant of numerous 3' mismatched base pairing schemes but also actually displays higher catalytic efficiency for sealing a 3' C:T² mismatch than it does for sealing nicks containing Watson–Crick base pairs. Whether the error-tolerance of this viral DNA ligase has biological relevance remains to be determined, but we have suggested that by efficiently competing with the nick-editing activity of ASFV APE for the 3' mismatched nicks synthesized by the extremely error-prone ASFV DNA polymerase X, the ASFV ligase could contribute to the sequence diversity that is known to exist between different isolates of the virus (3–5).

Are there other scenarios in nature where, at least conceptually, error-tolerant DNA ligation might be (i) useful for generating genetic diversity and/or (ii) critical to surviving forms of DNA damage where accuracy at ligation junctions is less important than the rapid reassembly of a contiguous duplex? If the above scenarios can in fact be envisioned, does mismatch ligation actually occur *in vivo*? If mismatch

[†] This work was supported by NIH Grant GM43268. B.J.L. was supported in part by a predoctoral NIH CBIP fellowship (2T32 GM08512).

* To whom correspondence should be addressed. Phone: (614) 292-3080. Fax: (614) 292-1532. E-mail: tsai@chemistry.ohio-state.edu.

[‡] Department of Chemistry, The Ohio State University.

¹ Abbreviations: adenylyl, adenosine monophosphate (AMP); AEB-SF, 4-(2-aminoethyl)-benzenesulfonyl fluoride; AP, apurinic/apyrimidinic (abasic); APE, AP endonuclease; ASFV, African swine fever virus; ATP, adenosine triphosphate; BER, base excision repair; BSA, bovine serum albumin; DEAE, diethylaminoethyl; DSB, double strand break; DTT, dithiothreitol; EDTA, ethylene diaminetetraacetic acid; NHEJ, non-homologous end joining; PMSF, phenylmethylsulfonyl fluoride; PNK, polynucleotide kinase; Pol, DNA polymerase; ROS, reactive oxygen species; SHM, somatic hypermutation; XRCC4, X-ray cross-complementation protein 4.

² Herein, the identity of the base pair on the 3'-OH side of a nick is described as X:Y, where X represents the templating nucleotide, and Y is the nucleotide that would have been inserted by a polymerase in the preceding gap-filling step.

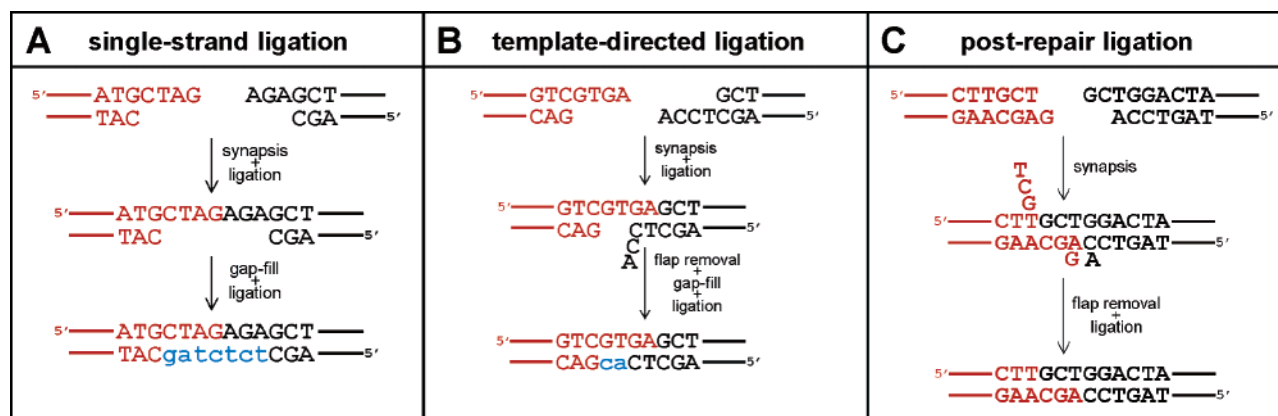


FIGURE 1: NHEJ can proceed via multiple routes. By transforming mammalian cells with linearized plasmids containing different combinations of termini (blunt, sticky, compatible, and incompatible) and then sequencing the rejoined products, both the activities employed during NHEJ and the order in which these activities are utilized were previously surmised. Importantly, though a particular cell type will typically repair a given set of termini preferentially using one route of NHEJ, multiple NHEJ routes are often used. Moreover, each route can yield an array of products that vary in the precise location at which joining has occurred. A single example from each of the three main routes of NHEJ, as described by Roth and Wilson (14), is illustrated here. Whereas the red and black capital letters denote the terminal sequences of the DNA fragments, the blue lower case letters represent nucleotides incorporated by a DNA polymerase during NHEJ. (A) Single-strand ligation. The defining feature of this route of NHEJ is the direct ligation of protruding single strands of opposite polarity or of a protruding single strand to a blunt-end fragment. Though only one rejoined product sequence is illustrated, this particular set of substrate termini would be expected to yield additional products, where one or more nucleotides had been removed from one or both of the protruding single stranded regions prior to the first ligation event (giving rise to a series of truncated products). (B) Template-directed ligation. In this route of NHEJ, the synapsis machinery employs regions of microhomology to facilitate ligation of the first strand. Subsequently, the second strand must be processed before ligation can occur. (C) Post-repair ligation. When an initial ligation event is not favorable because (i) single stranded regions of opposite polarity (as in part A) are not present or (ii) significant complementarity (as in part B) does not exist within the single stranded regions of the two termini, the NHEJ machinery will frequently search for regions of microhomology distal to the break. Once microhomology has been found and synapsis effected, neither strand can be ligated until end processing (removal of flaps, filling in gaps, etc.) has occurred.

ligation does not occur, is this due to the inherent fidelity of the ligase, or is it a consequence of a robust competing mismatched nick-editing activity? Because human DNA ligase IV repairs double strand breaks (DSBs) within multiple contexts where a capacity for mismatched nick ligation could potentially be advantageous, the above questions are pertinent to this enzyme. The work described here evaluates the theoretical advantages of error-tolerant nick ligation by human ligase IV, assesses the actual 3' mismatch ligation capabilities of this enzyme, and then evaluates these findings in light of previous *in vitro* and *in vivo* analyses of ligase IV.

DSBs are produced by ionizing radiation (6), select chemotherapy drugs (7), metabolic byproduct reactive oxygen species (ROS) (8), and are also generated as intermediates during the combinatorial construction of immunoglobulin genes in V(D)J recombination (9). Whether generated by chemical/physical damage or programmed enzymatic activity, in mammals DSBs are predominantly repaired by non-homologous end joining (NHEJ) (9). In short, NHEJ utilizes a DNA end-binding/DSB synapsis-effecting protein complex (DNA-PK, which consists of the Ku heterodimer and the DNA-dependent protein kinase catalytic subunit), a nuclease (Artemis), DNA polymerases (Pols λ , μ , and terminal deoxynucleotidyl transferase), an accessory/stimulatory protein (XLF), and a DNA ligase (the ligase IV/XRCC4 complex) (9–13). The precise manner in which the above activities are employed varies depending on the nature of the DNA termini (Figure 1) (14, 15). Moreover, for a single pair of DNA termini, multiple joining products are often generated (i.e., the location of the ligation junction varies) (15).

An ability for DNA ligase IV to seal 3' mismatched nicks could potentially be advantageous for two reasons. (i) During V(D)J recombination, diversity within immunoglobulin loci is created not only by randomly recombining the V, D, and J coding segments but also by effecting heterogeneity at the sites where these sub-exons are joined together (16, 17). The introduction of point mutations within these joints by way of mismatch ligation could enhance the sequence diversity of immunoglobulin complementary-determining regions; note that nicks containing 3' mismatches could be created both by the synapse of strands that are not fully complementary and by aberrant gap filling by Pols λ or μ (Figure 2A). (ii) In the repair of chemically/physically generated DSBs containing incompatible termini, there are scenarios where rather than subjecting the termini to nuclease degradation in order to utilize regions of microhomology that are distal to the break, it would appear to be both easier and less costly (in terms of information loss³) for ligase IV to simply directly join sticky ends that lack full complementarity (Figure 2B). Despite its potential advantages, the utilization of 3' mismatch ligation during NHEJ cannot be proven unequivocally by an analysis of joining product sequences. As illustrated in Figure 2C, repair joints that can be formed

³ When the DNA termini being processed by NHEJ have lost their original sequence complementarity (because of nuclease degradation), there is always the possibility that they will be rejoined in a manner that creates a frameshift mutation. Relative to the NHEJ mechanisms shown in Figure 1, a mechanism involving mismatch ligation (Figure 2) neither increases nor decreases the probability of generating a frameshift. Accordingly, one potential advantage of mismatch ligation over micro-homology mediated repair is that the former generates point mutations, whereas the latter generates deletion mutations, which are on average more deleterious than point mutations.

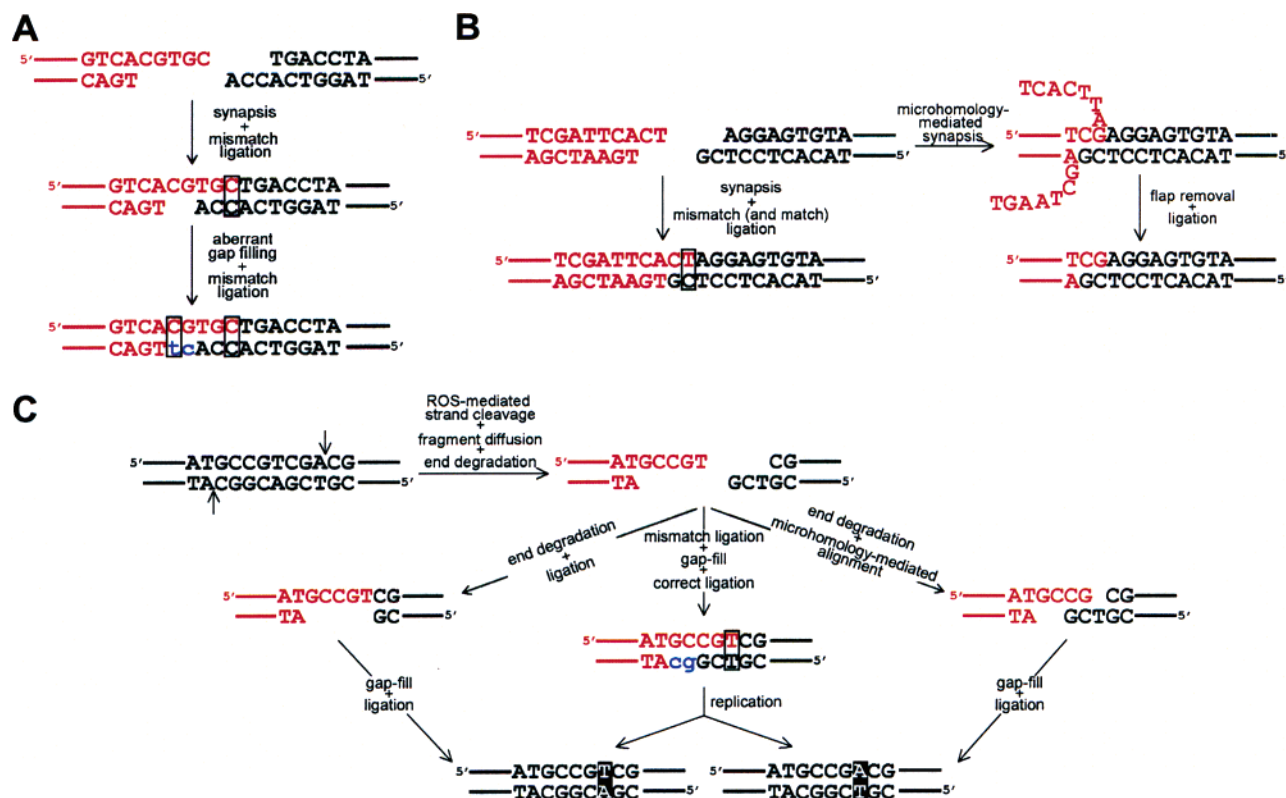


FIGURE 2. 3' mismatch ligation during NHEJ may be beneficial but cannot be unequivocally demonstrated. The red and black capital letters denote the terminal sequences of the DNA fragments; the blue lower case letters represent nucleotides incorporated by a DNA polymerase during NHEJ; mismatches are enclosed by filled boxes; and the two unique product sequences resulting from replication of the mismatch are enclosed by open boxes. (A) 3' mismatch ligation could enhance the sequence diversity generated during V(D)J recombination. During NHEJ-mediated repair of V(D)J recombination DSB intermediates, nicks containing 3' mismatches could be formed either by the synapsis of termini that are not fully complementary (middle) or by aberrant gap filling (bottom). If ligated, these 3' mismatches would generate point mutations within the regions linking the V, D, and J subexons; this increase in nucleotide sequence diversity would yield a greater repertoire of immunoglobulin complementarity-determining regions. (B) 3' mismatch ligation would minimize the number of nucleotides lost during NHEJ. DNA termini are occasionally degraded by nucleases to the point where extensive homology can no longer mediate the synapsis of the two fragments. Though synapsis can be mediated by regions of microhomology distal to the break, this can yield products containing large deletions (right). Alternatively, the direct ligation of termini that lack full complementarity minimizes the number of nucleotides deleted, but introduces a point mutation (left). For a discussion about frameshift mutations as they relate to the above two pathways, see footnote 3. (C) 3' mismatch ligation during NHEJ cannot be unequivocally demonstrated. 3' mismatch ligation and replication of the resultant mismatched product yields two progeny sequences that could have been generated via alternative routes of NHEJ.

via 3' mismatch ligation can always be generated by other routes of NHEJ.

Is 3' mismatched nick ligation ever employed during NHEJ? If not, is the absence of mismatch ligation a consequence of mismatch intolerance by ligase IV, efficient mismatched nick editing by Artemis (or some other nuclease), a simple inability of the NHEJ machinery to synapse thermodynamically unstable mismatched ends, or some combination of these?

The purpose of the work described here was to evaluate whether or not human DNA ligase IV/XRCC4 can contribute a 3' mismatched nick ligation activity to NHEJ. Because a quantitative kinetic analysis of ligase IV/XRCC4 had not previously been undertaken, requisite to the above objective were the initial optimization of techniques for the expression, purification, and assay of the protein complex. Here, we first describe protocols for expressing/purifying ligase IV/XRCC4 in yields higher than those previously published. Subsequently, we describe a unique rate-limiting step in the catalytic cycle of ligase IV/XRCC4 that has bearing on how the complex is assayed. After optimizing conditions so that assays can accurately yield information about nick ligation fidelity, we provide a kinetic analysis of the ability of ligase

IV, both in the absence and presence of XRCC4, to ligate nicked substrates containing different base pair combinations at their 3'-OH termini.

EXPERIMENTAL PROCEDURES

Materials. A plasmid containing human DNA ligase 4 cDNA (derived from hippocampus; NCBI accession no. AAH37491) was obtained from Open Biosystems. Note that this gene differs from the one derived from the human prostate (NCBI accession no. X83441); whereas the hippocampal gene specifies phenylalanine at residue 246, the prostate gene specifies serine at this position. The pRSF-Duet-1 plasmid was a gift from T. Ellenberger (Washington University School of Medicine). pET30-*xrcc4*, which contains full length *xrcc4* inserted between the NdeI and SalI sites of pET30a (18), was a gift from S. Jackson (University of Cambridge). The pACYCDuet-1 plasmid as well as BL21-(DE3) and Rosetta(DE3)pLysS cells were from Novagen. TOP10 cells were from Invitrogen. Pfu DNA polymerase was from Stratagene. Complete, EDTA-free Protease Inhibitor Cocktail Tablets and bovine serum albumin (BSA) were from Roche. 4-(2-Aminoethyl)-benzenesulfonyl fluoride (AEBSF) was from BioShop. Phenylmethylsulfonyl fluoride

Table 1: Schemes for the Individual Expression and Coexpression of Human DNA Ligase IV and XRCC4

scheme	cell type	plasmid 1 ^a	plasmid 2 ^a	yield ^b (mg/L)
1	Rosetta(DE3)pLysS	pET30- <i>lig4</i> -His ₆		0.10
2	Rosetta(DE3)pLysS	pET30- <i>xrcc4</i> -His ₆		0.30
3	BL21(DE3)	pET30- <i>xrcc4</i>	pACYC- <i>lig4</i> -His ₆	0.10
4	BL21(DE3)	pET30- <i>lig4</i> -His ₆	pACYC- <i>xrcc4</i>	0.40
5	Rosetta(DE3)pLysS	pRSFDuet- <i>lig4</i> (His ₆)- <i>xrcc4</i>		1.4 ^c

^a For descriptions of plasmid construction and nomenclature, see Experimental Procedures. ^b Protein yields are for post purification and were calculated on the basis of absorption at 280 nm using the extinction coefficients listed in Experimental Procedures. The values reported for coexpression schemes 3–5 reflect the yield of the complex. ^c The highest yield of the ligase IV/XRCC4 complex was achieved by simultaneously expressing both proteins from the same vector.

(PMSF), Tween 20, and Diethylaminoethyl (DEAE) Sephacel were from Sigma. Talon resin was from BD Biosciences. Amicon ultrafiltration membranes (regenerated cellulose) and centrifugal filter devices were from Millipore. Hydroxyapatite was from BioRad. NP-40 was from USB. Oligonucleotides were from Integrated DNA Technologies (IDT). T4 DNA ligase and T4 polynucleotide kinase (PNK) were from New England Biolabs. Sep-Pak C18 columns were from Waters. ATP, MicroSpin G-25 columns, and Sephacryl 200 (S-200) were from Amersham Biosciences. Materials and reagents that are not listed here were of standard molecular biology grade.

Cloning. The *ligase 4* gene was amplified from its cDNA using Pfu DNA polymerase and the following primers: 5'-TGGCTGCCTCACAACTTC-3' and 5'-GCCGCTCGAGTTAGTGATGGTGATGGTGATGAATCAAATACTGGTTTTCTTCTTGTAATTCAC-3'. These primers specify full-length ligase IV containing an unmodified N-terminus and a C-terminal His₆ tag. Digestion of this PCR product with XhoI alone yielded a sticky end downstream of the gene and left a blunt end upstream of the gene. pET30a was prepared to accommodate this insert by (i) digestion with NdeI, (ii) incubation with Pfu DNA polymerase to fill in the sticky ends, and (iii) digestion with XhoI. Ligation of the above *ligase 4* blunt/sticky insert into the blunt/sticky pET30a was with T4 DNA ligase. The sequence of this expression plasmid, referred to as pET30-*lig4*-His₆, was verified at the OSU Plant and Microbe Genomics Facility, as were the sequences of all other expression plasmids used in this study.

The pACYC-*lig4*-His₆ expression plasmid, which encodes full-length ligase IV containing an unmodified N-terminus and a C-terminal His₆ tag, was prepared in a manner similar to that described above except that a 5' NcoI site (which was filled in to give a blunt end) and a 3' PacI site (which was left sticky ended) were used. The primers used for amplifying *ligase 4* here were 5'-GCTGCCTCACAACTTCACAAAC-3' and 5'-CCTTAATTAATTAGTGATGGTGATGGTGATGAATCAAATACTGGTTTTCTTCTTGTAATTCAC-3'.

To generate the pET30-*xrcc4*-His₆ expression plasmid, pET30-*xrcc4* (described in the Materials section) was subjected to the Quick Change method to insert a C-terminal His₆ tag. The primers used here were 5'-GACCTCTTTGATGAGATTCATCA CCATCACCATCACTAAGTCGACAAGCTTGCG-3' and 5'-CGCAAGCTTGTCGAC TTAGTGATGGTGATGGTGATGAATCTCATCAAAGAGGTC-3'.

To generate the pACYC-*xrcc4* expression plasmid, the *xrcc4* gene was amplified from pET30-*xrcc4* (described in

the Materials section) and inserted between the NcoI and XhoI sites of pACYCDuet-1. The primers used for amplifying *xrcc4* here were 5'-CATGCCATGGAGAGAAAAAT-AAGCAGAATCCAC-3' and 5'-GCCGCTCGAGTT AAATCTCATCAAAGAGGTCTTCTGGG-3'.

A vector for coexpressing C-terminally His₆-tagged ligase IV and non-tagged XRCC4 was prepared by first inserting *xrcc4* between the NdeI and XhoI sites of multiple cloning site 2 of pRSFDuet-1 and subsequently inserting *ligase 4* between the NcoI and EcoRI sites of multiple cloning site 1. The latter step was accomplished by (i) digesting the *xrcc4*-containing pRSFDuet-1 with NcoI, (ii) incubating the linearized plasmid with Pfu DNA polymerase to fill in the sticky ends, (iii) digesting both the insert and the linearized plasmid with EcoRI, and (iv) ligating the blunt/sticky insert into the blunt/sticky plasmid with T4 DNA ligase. The primers used for amplifying *xrcc4* here were 5'-GGCATATGGAG AGAAAAATAAGCAGAATCCAC-3' and 5'-GCCGCTCGAGT TAAATCTCATCAAAGAGGTCTTCTGGG-3'. The primers used for amplifying *ligase 4* here were 5'-GCTGCCTCACAACTTCACAAAC-3' and 5'-CCGAATTCTTAGTG ATGGTGATG GTGATGAATC-3'. This coexpression plasmid will be referred to as pRSFDuet-*lig4*(His₆)-*xrcc4*.

Protein Expression, Purification, and Quantitation. Schemes outlining the plasmids and cells used for both the coexpression and the individual expression of ligase IV and XRCC4 are listed in Table 1. Cells were grown to mid-log phase in LB medium at 37 °C. The temperature was then decreased to 16 °C and protein expression induced with 1 mM isopropyl-β-D-thiogalactopyranoside. Incubation/shaking was continued for 16 h at 16 °C before harvesting by centrifugation and flash freezing the cell pellet in liquid nitrogen. The one exception to this protocol was for scheme 2 (see Table 1), where the expression of XRCC4 was induced for 8 h at 30 °C.

All purification steps were performed at 4 °C, unless noted otherwise. Buffers used for the Talon cobalt column were as follows. Buffer A: 50 mM sodium phosphate, 300 mM NaCl, 10% glycerol (v/v) at pH 8.0, 5 mM imidazole, and 1 mM β-mercaptoethanol. Buffer B: same as buffer A but supplemented with 1% NP-40. Buffer C: same as buffer B but with pH 7.0 and 10 mM imidazole. Buffer D: same as buffer C but with 1 M NaCl. Buffer E: same as buffer D but lacking NP-40. Buffer F: same as buffer E but with 300 mM NaCl. Buffer G: same as buffer F but with 80 mM imidazole. Buffer H: same as buffer G but with 170 mM imidazole.

Cell pellets were resuspended in buffer A (~5 mL of buffer/g of cell pellet) that had been supplemented with

Complete, EDTA-free Protease Inhibitor Cocktail, AEBSF (100 μ M), and PMSF (100 μ M). Lysis was by sonication. Following the removal of cellular debris by centrifugation, the clarified lysate was applied to a Talon column (\sim 5 mL bed volume) that had been equilibrated with buffer A. Washing was with \sim 100 mL each of buffers A, B, C, D, E, and F, in that order. Elution was with \sim 90 mL of buffer G followed by \sim 50 mL of buffer H, collecting 10 mL fractions. Free ligase IV, free XRCC4, and the ligase IV/XRCC4 complex each eluted predominantly in buffer G (80 mM imidazole).

Fractions containing the target protein (or the protein complex) were pooled and concentrated using an Amicon ultrafiltration cell with a 10,000 NMWL membrane. The concentrated sample was loaded onto a DEAE column (15 cm \times 2.5 cm (inside diameter)) that had been equilibrated with buffer consisting of 50 mM Tris-Borate (pH 7.0), 10% glycerol (v/v), and 5 mM DTT. After washing with the same buffer, protein was eluted using a 400 mL (total volume) linear KCl gradient (30–300 mM for free ligase IV, 240–430 mM for free XRCC4, and 75–500 mM for the ligase IV/XRCC4 complex) at a flow rate of 1 mL/min, collecting 6 mL fractions. Free ligase IV eluted at \sim 130–180 mM KCl, free XRCC4 eluted at \sim 270–350 mM KCl, and the ligase IV/XRCC4 complex eluted at \sim 150–240 mM KCl.

Appropriate fractions were pooled, concentrated in an Amicon centrifugal filter device (10,000 NMWL), and loaded onto an S-200 column (200 cm \times 3 cm (inside diameter)) that had been equilibrated with buffer consisting of 50 mM Tris-Borate (pH 7.0), 100 mM KCl, 5% glycerol, and 5 mM DTT. S-200 fractions containing the target protein (or the protein complex) were pooled, concentrated in a centrifugal filter device, and then tested for the presence of contaminant endo/exonucleases. The latter was accomplished by subjecting an aliquot of the protein to a hot DNA ligation assay (see below for details), wherein the degradation of nicked DNA substrate (or ligated DNA product) can be readily detected. At this point, free XRCC4 did not contain nuclease contaminants and was therefore supplemented with glycerol to a final concentration of 50% (v/v), aliquoted, flash frozen in liquid nitrogen, and stored at -80°C . Though after S-200 chromatography free ligase IV and the ligase IV/XRCC4 complex both appeared to be pure on a silver-stained SDS–polyacrylamide gel, both typically contained trace quantities of nuclease contaminants.

To remove nuclease contaminants, free ligase IV and the ligase IV/XRCC4 complex were subjected to further purification on a hydroxyapatite column (\sim 5 mL bed volume) that had been equilibrated with 10 mM sodium phosphate, 2% glycerol (v/v), 10 mM DTT, and 0.5% Tween 20. For washing and elution, a step gradient was employed, consisting of 20, 50, 80, 100, 160, 320, and 500 mM sodium phosphate; Tween 20 was excluded from the latter two steps. Free ligase IV and the ligase IV/XRCC4 complex both eluted in the 320 mM step. Note that in order to prevent the sodium phosphate from crystallizing on the column, the 160, 320, and 500 mM elution steps were performed at 20°C . Appropriate fractions were pooled, concentrated in a centrifugal filter device at 20°C , and then subjected to buffer exchange (into the S-200 buffer described above) by diluting the sample \sim 12-fold (inside the concentrator) and then reconcentrating it; the dilution/reconcentration cycle was

performed a total of three times. Free ligase IV was then supplemented with glycerol to a final concentration of 50% (v/v), aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C . The ligase IV/XRCC4 complex was stored in an analogous manner but was not supplemented with glycerol because this was found to cause precipitation.

It is important to note that when stored at -80°C , both free ligase IV and the ligase IV/XRCC4 complex are stable for many months. However, when stored at -20°C , both the free protein and the complex show significant loss in activity; whereas the half-life of the complex at -20°C is \sim 15 days⁴, the half-life of free ligase IV is even shorter. Accordingly, once an aliquot of protein had been thawed from -80°C and used to run an assay, any remaining volume was simply thrown away.

The concentration of purified protein (or the protein complex) was determined by UV absorbance using extinction coefficients calculated by the ProtParam tool of the ExPASy Proteomics Server; the extinction coefficients for free ligase IV, free XRCC4, and the ligase IV/XRCC4 complex are 96580, 24990, and 147880 $\text{M}^{-1}\text{cm}^{-1}$, respectively. Note that in order to determine the extinction coefficient for the ligase IV/XRCC4 complex, a virtual protein was employed consisting of one ligase IV molecule fused to two XRCC4 molecules (to reflect the fact that the complex is expected to consist of one ligase IV molecule bound to an XRCC4 dimer (20, 21)).

The values reported in Table 1 represent the yield of protein (or the protein complex) after purification and are, therefore, not as direct an assessment of expression levels as could be obtained via Western blotting prior to cell lysis. Although slight variations from one purification to the next (such as which fractions are pooled or discarded) can influence the final yield of protein (or the protein complex), the expression/purification protocols reported here were used many times with the final results being reproducible. On the basis of this reproducibility, we assert that the values in Table 1 accurately reflect differences in the relative levels of protein expression and are not merely artifacts of the protein purification protocol.

Preparation of DNA Substrates. Oligonucleotides were purified by denaturing polyacrylamide electrophoresis under standard conditions (3). After gel extraction, oligos were desalted on Sep-Pak C18 columns, dried in a speed-vac, and resuspended in TE buffer (10 mM Tris-HCl and 1 mM EDTA at pH 7.5). Concentrations were determined by UV absorbance at 260 nm using extinction coefficients calculated by the oligo analyzer tool on IDT's website.

The appropriate oligonucleotides were labeled using T4 polynucleotide kinase (PNK) and [γ - ^{32}P]ATP. After heat-inactivating PNK, free ATP was removed on a MicroSpin G-25 column. Nicked substrates (Figure 3) were assembled at room temperature by combining the upstream 26-mer with the 45-mer template and the 5'-phosphorylated downstream 19-mer in a ratio of 1:1.2:1.4, respectively. The rationale for using nicked rather than DSB substrates is addressed in

⁴ In contrast, when purified from HeLa cell nuclei, Robins and Lindahl found DNA ligase IV/XRCC4 to be stable at -20°C for many months (19). This difference in stability could be due to differences in storage buffer composition or, perhaps, stabilizing post-translational modifications that are not possible in our *E. coli* expression system.

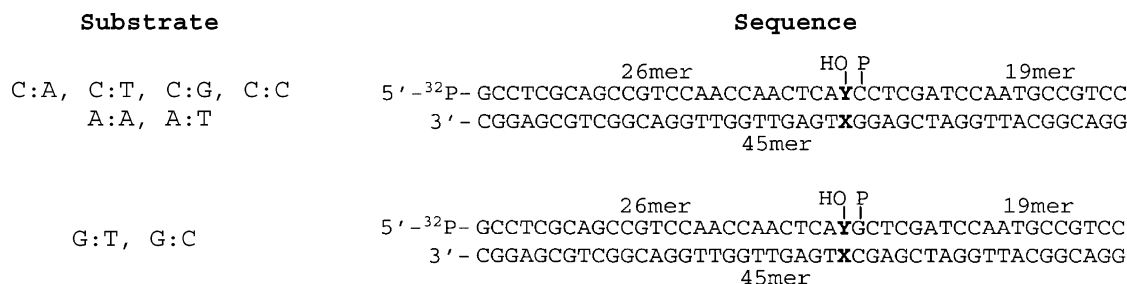


FIGURE 3: Nicked DNA substrates. The substrates are named according to the base pair that is 3' (upstream) to the nick. This base pair is described as X:Y, where X denotes the templating nucleotide, and Y represents the nucleotide that would have been incorporated by a polymerase in the preceding gap-filling step. OH is the 3'-hydroxyl, and P is the 5'-phosphate.

the Discussion section. Heating and slow cooling the oligonucleotide mixture to promote annealing/assembly of the nicked duplex was found to be unnecessary. Substrates were exclusively constructed with the ³²P-label at the 5'-terminus of the upstream 26-mer. On average, only 3% of the substrate molecules were labeled. Prior to storage at 4 °C, substrates were diluted to a concentration of 6–30 μM using the following buffer: 50 mM Tris-borate, 100 mM KCl and 15% glycerol (v/v), pH adjusted to 7.8 with KOH while at 37 °C.

Steady State Assays. Steady state DNA ligation assays were performed in a buffer/cosolvent mix that we previously determined to be optimal for ASFV DNA ligase (3); this consisted of 50 mM Tris-borate (pH adjusted to 7.8 with KOH while at 37 °C), 100 mM KCl, and 15% glycerol (v/v), supplemented with 15 mM MgCl₂, 1 mM ATP, 300 μM DTT, and 1.5 mg/mL BSA (all concentrations are final). The concentrations of enzyme and DNA substrate used varied from one assay to the next and are, therefore, specified in the Results section. Steady state analysis of ligation activity versus pH was conducted in the same buffer/cosolvent system described above but at pH 6.6, 7.1, 7.6, 8.0, 8.5, or 8.9.

All steady state assays were conducted manually as follows: After a 5 min preincubation at 37 °C, a solution containing enzyme, MgCl₂, ATP, DTT, and BSA at twice the intended concentrations was added to an equal volume of nicked DNA substrate, also at twice the intended concentration. Generally, 10 μL aliquots were removed at the appropriate time points and quenched in 10 μL of formamide containing xylene cyanol and bromophenol blue. Reaction products were resolved on 15% denaturing polyacrylamide gels and visualized by phosphor screen autoradiography using a STORM840 scanner from Molecular Dynamics. Band intensity quantitation and data plotting were performed with ImageQuant (GE Healthcare) and SigmaPlot 9.0 (Systat Software, Inc.), respectively. When appropriate, biphasic steady state time courses were fit to the burst equation: [product] = A[1 - exp(-k_{obs}t)] + ct, where A is the burst amplitude, k_{obs} is the apparent first-order rate constant, and c is the steady state rate.

Pre-Steady-State Assays. Pre-steady state (single turnover) DNA ligation assays were performed manually as described above for the steady state assays, except that ATP was excluded from the assay buffer. As is described in detail in the Results section, the first step of the DNA ligation mechanism involves enzyme adenylylation (i.e., “charging” a catalytic lysine with adenosine monophosphate (AMP)). When expressed in *E. coli* as described above, both free ligase IV and the ligase IV/XRCC4 complex exist predominantly in

the adenylylated state. When assayed in the absence of ATP, the adenylylated ligase molecules catalyze one round of nick sealing but are subsequently unable to recharge themselves. These time courses were fit to a single exponential: [product] = A[1 - exp(-k_{obs}t)], where A is the single turnover amplitude, and k_{obs} is the apparent first-order rate constant.

The influence of [KCl] on the pre-steady state activity of the ligase IV/XRCC4 complex was analyzed in a buffer/cosolvent system similar to that described above, except that ATP was excluded, and [KCl] was adjusted to be either 0, 40, 80, 100, 120, 160, 200, or 300 mM.

Analysis of Ligation Fidelity. Fidelity analyses were performed in the pre-steady state (as described above), using enzyme concentrations⁵ of 8.5–15 nM (free ligase IV) or 50–80 nM (the ligase IV/XRCC4 complex), and six or seven different concentrations of the nicked DNA substrate. Saturation curves were obtained by plotting the apparent rate constant (k_{obs}) of single turnover reactions as a function of substrate concentration and fitting the data to the Michaelis–Menten equation: k_{obs} = k_{lig}[S]/(K_{d,app} + [S]), where k_{lig} is the catalytic rate constant, and K_{d,app} is the apparent equilibrium dissociation constant. Each saturation curve was repeated at least twice, and the reported data represents the single best independent trial. Herein, fidelity is reported as (k_{lig}/K_{d,app})_{correct}/(k_{lig}/K_{d,app})_{incorrect}, where correct and incorrect denote Watson–Crick and non-Watson–Crick base pairs, respectively.

RESULTS

Protein Expression and Purification. First identified as a unique, previously unrecognized cDNA in 1995 (22), human DNA ligase IV was purified for the first time from HeLa cell nuclei in 1996 (19). In that first study, ligase IV was found to copurify with another protein (19), suggesting that it exists *in vivo* as a stable hetero-multimeric protein complex. The ligase IV binding partner was eventually determined to be XRCC4 (X-ray cross-complementation protein 4) (18, 23). Analogous to the interaction between human DNA ligase III and XRCC1 (24, 25), the interaction between ligase IV and XRCC4 is crucial (26, 27). The presence of XRCC4 both protects ligase IV from proteolysis (in the absence of XRCC4, virtually no full length ligase IV can be detected in mammalian cells (28)) and enhances the efficiency of sealing nicked (29) and sticky-ended (23)

⁵ These values do not represent the total enzyme concentration but rather the concentration of active adenylylated enzyme, which was determined by the amplitude of single-turnover assays (in the absence of ATP).

substrates. Moreover, ablation of the interaction between ligase IV and XRCC4 results in impaired V(D)J recombination and hypersensitivity to DSBs (26, 27).

Though others have had success with the individual expression of ligase IV (in mammalian (30) and bacterial cells (29)) and XRCC4 (in insect (23) and bacterial cells (18, 29)), and with the coexpression of ligase IV and XRCC4 (in insect (31) and bacterial cells (20, 32)), yields have been low. Using the strategies outlined in Table 1, we attempted to develop *E. coli*-based expression systems with higher yields. Because we were interested in analyzing specifically how XRCC4 influences the activity and fidelity of ligase IV, we initially expressed and purified these two proteins separately (so that they could be reconstituted in different ratios). After expressing ligase IV and XRCC4 individually from the medium copy number pET30 plasmid in Rosetta(DE3)pLysS cells and then individually subjecting the proteins to multiple purification steps, we obtained yields of each protein that were similar to, or lower than, those reported in a previous study (schemes 1 and 2 in Table 1) (29). Note that although XRCC4 expressed well enough to display a distinct induction band in the crude cell lysate (analyzed on a silver-stained SDS–polyacrylamide gel), the presence of ligase IV was not discernible until after the first purification step. It is also worthwhile to note a difficulty that we encountered when assembling the ligase IV/XRCC4 complex from the individually expressed/purified proteins. Though structural (21) and biophysical (20) data indicate that the complex consists of one ligase IV molecule bound to an XRCC4 dimer, when titrating ligase IV with XRCC4, maximal steady state ligation activity was observed at a ligase IV/XRCC4 ratio of ~2:1 (rather than the expected 1:2) (data not shown). One explanation for this is simply that the ligase IV concentration, which for this particular experiment had been determined by UV absorbance, was an overestimate of the true concentration of active ligase IV molecules.

The difficulty in assembling the ligase IV/XRCC4 complex in the proper stoichiometry could potentially be avoided by combining a cell pellet containing overexpressed ligase IV with a cell pellet containing overexpressed XRCC4 and then proceeding with cell lysis (at which point the complex would assemble) and protein purification as usual; because the ligase IV/(XRCC4)₂ complex is so stable [$K_d \sim 10^{-10}$ M (20)], the stoichiometry of the complex is expected to be preserved throughout the course of the purification. However, the XRCC4-mediated stabilization of ligase IV that is observed in mammalian cells (28) suggested that coexpressing the two proteins in *E. coli* might similarly enhance the yield of ligase IV and, therefore, of the ligase IV/XRCC4 complex. Accordingly, we attempted to coexpress ligase IV and XRCC4 both from separate plasmids (using BL21(DE3) cells) and from the same plasmid (using Rosetta(DE3)pLysS cells). When ligase IV is expressed from a low copy number plasmid (pACYC = ~20 copies per cell) while XRCC4 is simultaneously being expressed from a medium copy number plasmid (pET30 = ~40 copies per cell), the yield of the purified complex is a mere 0.1 mg/L (scheme 3 in Table 1). When this scenario is reversed, with *ligase 4* in the medium copy plasmid and *xrcc4* in the low copy plasmid, the yield of the purified complex increases 4-fold (scheme 4 in Table 1); this is probably a consequence of

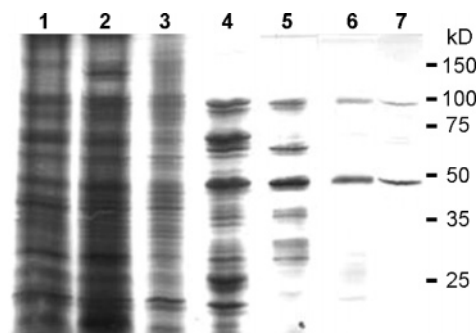


FIGURE 4: Silver-stained gel illustrating purification of the ligase IV/XRCC4 complex. Lane 1, uninduced cell culture; lane 2, induced cell culture; lane 3, clarified cell lysate (i.e., after centrifugation to remove insoluble material); lane 4, eluate from the Talon cobalt column; lane 5, eluate from the DEAE column; lane 6, eluate from the S-200 column; lane 7, eluate from the hydroxyapatite column. The position of molecular weight markers is shown to the right of the gel. Note that the apparent molecular weight of XRCC4 (~48 kD) far exceeds the actual molecular weight (38 kD); this phenomenon has been observed by other groups (31) and has been speculated to be a consequence of XRCC4's amino acid composition (18).

the medium copy plasmid generating a higher concentration of *ligase 4* mRNA (which should be helpful because transcription, translation, and folding of the 104 kD ligase IV are not expected to be as robust as they are for the 38 kD XRCC4).

Because both *ligase 4* and *xrcc4* contain codons that are used infrequently in *E. coli*, it seemed worthwhile to attempt coexpression in Rosetta(DE3)pLysS cells (which synthesize tRNAs that are otherwise rare in *E. coli*). The difficulty in using two expression plasmids that are compatible both with each other and the tRNA-expressing plasmid was circumvented by placing both *ligase 4* and *xrcc4* into a single dual expression vector. This coexpression system (scheme 5 in Table 1) proved to be superior to the others examined here and enabled us to obtain highly pure ligase IV/XRCC4 complex (Figure 4) at a yield of 1.4 mg/L. When the 1:2 stoichiometry of the ligase IV/XRCC4 complex is considered, along with the molecular weights of the two proteins, the 1.4 mg/L yield of the complex corresponds to ~0.8 mg/L of ligase IV. This is a significant enhancement compared to what we and others (29) observed when expressing ligase IV by itself (scheme 1 in Table 1) and is consistent with an XRCC4-mediated stabilization/protection of ligase IV in *E. coli*.

Steady State Assays. Our previous analysis of DNA ligation fidelity, focusing on the DNA ligases encoded by ASFV and bacteriophage T4, was conducted in the steady state (3). It was our intention to analyze the human ligase IV/XRCC4 complex in an analogous manner so that the data for these different enzymes could be directly compared. However, our initial efforts to assay the ligase IV/XRCC4 complex in the steady state yielded unexpected, nonlinear time courses. In the typical time course shown in Figure 5A, where 35 nM complex was incubated with 400 nM nicked substrate, there is an obvious drop in the turnover rate before even 15% of the substrate has been consumed, suggesting that this apparent loss in activity is not due to product inhibition. Thermal inactivation of the enzyme over the course of the reaction cannot account for this behavior either because preincubating the complex at 37 °C for 5,

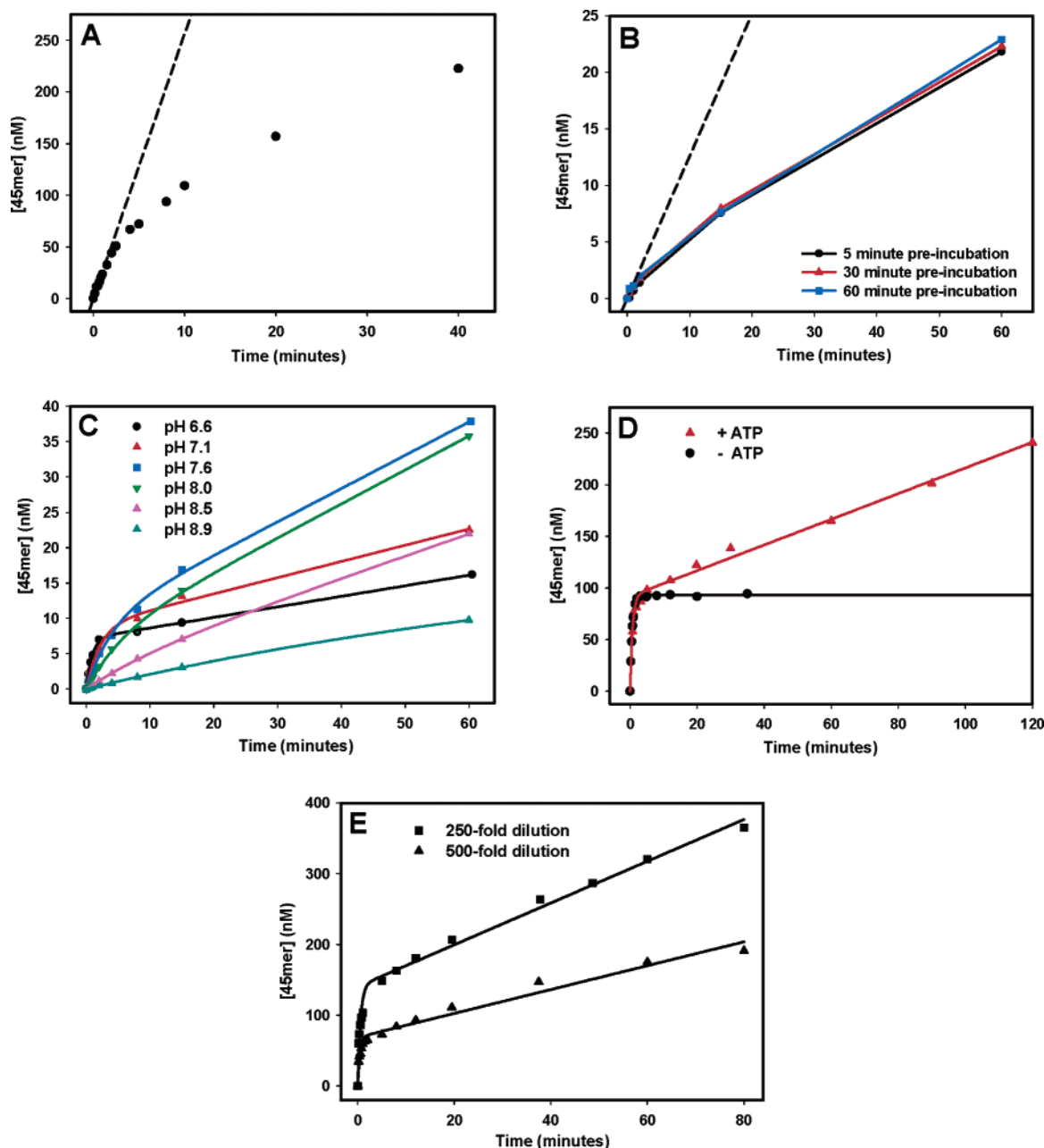


FIGURE 5: Ligase IV/XRCC4 complex displays biphasic behavior in the steady state. (A) Typical nonlinear time course for the ligase IV/XRCC4 complex in the steady state. The ligase IV/XRCC4 complex (35 nM) was incubated with 400 nM nicked C:G at pH 7.8 in the presence of ATP. The dashed reference line is a linear fit extending from the origin through the first four time points. (B) Nonlinear steady state behavior of the ligase IV/XRCC4 complex is not a consequence of enzyme inactivation. The complex (5 nM) was incubated with 100 nM nicked C:G at pH 7.8 in the presence of ATP. Prior to initiating the reaction, the solution containing the enzyme and cosubstrates was preincubated at 37 °C for 5 min, 30 min, or 60 min. The dashed reference line is a linear fit extending from the origin through the first two time points. (C) pH dramatically influences the shape of the ligase IV/XRCC4 steady state time course. The complex (7 nM) was incubated with 100 nM nicked C:G in the presence of ATP at pH 6.6, 7.1, 7.6, 8.0, 8.5, or 8.9. Note that at pH 6.6, the time course consists of an initial rapid exponential phase followed by a slower linear phase (canonical burst behavior). However, as pH is increased, the rates of these two phases become increasingly similar, leading to the nonlinear behavior observed above in parts A and B. At pH ≥ 8.0 , the enzyme complex begins losing activity. (D) Pre-steady state burst likely corresponds to the turnover of the pre-adenylylated enzyme. At pH 6.6, the ligase IV/XRCC4 complex was incubated with 600 nM nicked G:C substrate in either the presence (\blacktriangle , fit to the burst equation) or the absence (\bullet , fit to a single exponential) of ATP. While the pre-steady state burst amplitude and rate constant were 92 nM and 1.4 min^{-1} , respectively, the single turnover amplitude and rate constant were 93 nM and 1.5 min^{-1} , respectively. (E) Burst amplitudes and steady state rates observed at pH 6.6 correlate with the extent of enzyme dilution. 250 (\blacksquare)- and 500 (\blacktriangle)-fold dilutions of the ligase IV/XRCC4 complex were incubated with 750 nM nicked G:C substrate in the presence of ATP, yielding burst amplitudes of 141 and 74 nM, respectively. Whereas the fast phase of these reactions was independent of enzyme concentration ($1.9 \pm 0.07 \text{ min}^{-1}$), the slower linear phase was proportional to enzyme concentration: 2.95 nM/min (250-fold dilution) vs 1.42 nM/min (500-fold dilution).

30, or 60 min prior to initiating the reaction had no effect on the time course (Figure 5B).

Because the nonlinear time courses of the ligase IV/XRCC4 complex are, to our knowledge, unprecedented

among this class of enzymes⁶, a mechanistic explanation for this behavior was not immediately obvious, and accordingly, the best method for fitting the data was initially unclear; neither linear fits of the earliest time points nor single

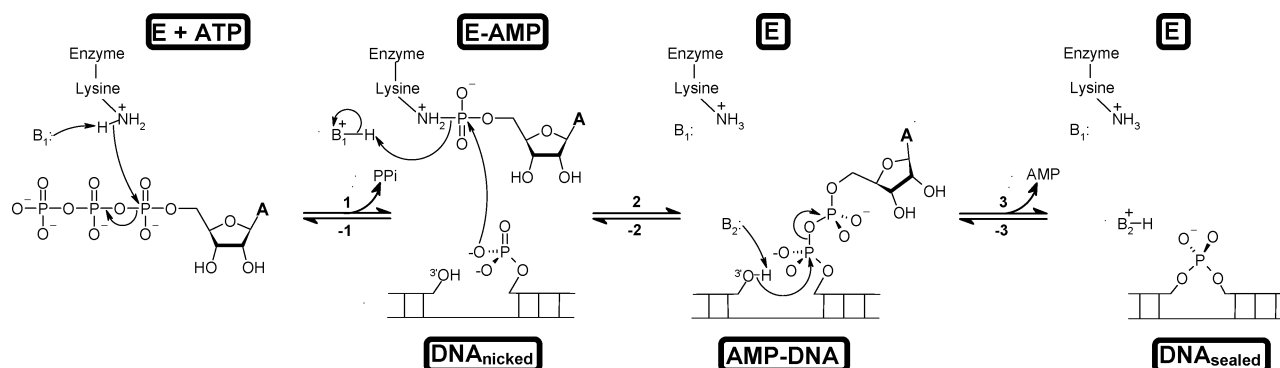


FIGURE 6: Universal ATP-dependent DNA ligase mechanism. In the initial charging step, a catalytic lysine attacks the α -phosphate of ATP, generating a lysyl-AMP adduct (E-AMP) and releasing pyrophosphate (step 1). Adenylylated ligase then binds nicked DNA and transfers the AMP from lysine to the 5'-phosphate of the nick (step 2). The 3'-hydroxyl of the nick then attacks the adenylylated 5'-phosphate, forming a phosphodiester bond and eliminating the AMP (step 3).

exponential or burst fits of the entire time course were of good quality. With the hope of finding an assay condition that would yield data that was (i) more amenable to quantitative analysis and (ii) that might provide a mechanistic explanation for the nonlinear behavior, we varied the assay pH. As shown in Figure 5C, at pH 6.6, the ligase IV/XRCC4 complex displays canonical burst behavior consisting of an initial rapid exponential phase followed by a slower linear phase. As pH is increased, the rate of the two phases converge, with the rate of the exponential phase decreasing and the rate of the linear phase increasing so that at pH ≥ 7.6 , the biphasic nature of the curve is less distinct and much more difficult to fit accurately using the burst equation. At pH ≥ 8.0 , the complex begins to lose activity.

The unique biphasic nick ligation time course displayed by the ligase IV/XRCC4 complex provides information about the relative rates of the different steps in the reaction mechanism. As outlined in Figure 6, in the initial charging step, a catalytic lysine of the DNA ligase attacks the α -phosphate of ATP, generating a lysyl-AMP adduct (E-AMP) and releasing pyrophosphate (step 1). Adenylylated DNA ligase (E-AMP) then binds nicked DNA and transfers AMP from lysine to the 5'-phosphate of the nick (step 2). The 3'-hydroxyl of the nick then attacks the adenylylated 5'-phosphate, forming a phosphodiester bond and eliminating AMP (step 3). In light of this universal DNA ligase mechanism (34, 35), we hypothesized that (i) the initial fast phase of the biphasic time course reflects one round of nick sealing (encompassing both steps 2 and 3) catalyzed by enzyme that is pre-adenylylated⁷ and that (ii) the slow linear phase reflects the subsequent cycling of the enzyme, which is limited by the rate of enzyme recharging (step 1). Both facets of this hypothesis are addressed below.

If the pre-steady state phase of the reaction corresponds to the turnover of pre-adenylylated enzyme, then both the amplitude and the rate constant of this phase should be equivalent to the amplitude and rate constant of a single turnover reaction performed under similar conditions but in

the absence of ATP. As shown in Figure 5D, this is in fact the case. At pH 6.6, while the amplitude and rate constant of the pre-steady state phase in the biphasic time course are 92 nM and 1.4 min^{-1} , the single turnover amplitude and rate constant are 93 nM and 1.5 min^{-1} , respectively. That the pre-steady state phase of the biphasic time course reflects the initial turnover of the pre-adenylylated enzyme is further supported by the fact that the burst amplitude correlates very well with the extent to which the ligase IV/XRCC4 complex has been diluted: burst amplitude = 141 nM at 250 \times dilution and 74 nM at 500 \times dilution (Figure 5E). As would be expected, the rate constant of the pre-steady state phase is independent of enzyme concentration ($1.9 \pm 0.07 \text{ min}^{-1}$), whereas the rate of the steady state phase varies in proportion to enzyme concentration: 2.95 nM/min at 250 \times dilution and 1.42 nM/min at 500 \times dilution (Figure 5E).

Because the above data suggests that the pre-steady state phase of the biphasic time course corresponds to the turnover of the pre-adenylylated enzyme, the steady state phase must be limited in rate by a chemical or physical step occurring after step 3 in the ligase mechanism (Figure 6). That the steady state phase is limited in rate by the dissociation of the enzyme/product complex (as is the case for DNA polymerases) seems unlikely because the enzyme's affinity for non-nicked DNA is expected to be extremely poor (36, 37). Another alternative, that a post-chemistry protein-reopening conformational change limits the rate of the steady state phase, is also unlikely because this phase displays a strong and incremental pH dependence (Figure 5C), which is generally more characteristic of a chemical step than a macromolecular rearrangement. The best explanation, therefore, seems to be that the steady state phase reflects the rate of enzyme recharging (step 1 in Figure 6). Efforts to demonstrate that the steady state rate of the biphasic time course is equivalent to the rate of enzyme adenylylation have so far been hampered by technical challenges. Because our preparations of the ligase IV/XRCC4 complex consist predominantly of the pre-adenylylated species (E-AMP), the concentration of the uncharged enzyme that is accessible to the reaction $\text{E} + \text{ATP} \rightarrow \text{E-AMP} + \text{PPi}$ is extremely low, and this has made accurate rate measurements difficult even when using coupled assays; efforts to monitor ligase adenylylation after a pre-incubation with nicked DNA (to effect decharging) are being pursued.

⁶ Note that we (3) and others (33) have been able to effect burst behavior during DNA ligation but only by using substrates or reaction conditions that are unnatural.

⁷ DNA ligases exist *in vivo* as a mixture of the adenylylated and non-adenylylated species. When expressed in *E. coli* and purified by the protocol described above, we have consistently found preparations of human ligase IV/XRCC4 to consist predominantly of the adenylylated species.

Though we presently lack direct empirical evidence that the steady state phase of the biphasic time course corresponds to step 1 of the ligase mechanism, this interpretation is consistent with the pH dependence of this phase. Charging of the ligase catalytic lysine requires the deprotonation of the lysine ϵ -amino group so that it can attack the α -phosphate of ATP (step 1 in Figure 6). This step of the reaction is therefore expected to be enhanced at high pH and mitigated at low pH, which is consistent with the data in Figure 5C. That steady state assays of the ligase IV/XRCC4 complex always display some degree of biphasic behavior suggests that within the pH range examined here (6.6–8.9), enzyme adenylation is always the slowest step in the complete catalytic cycle of this enzyme.

Collectively, the data in Figure 5C–E and the qualitative analysis of pH dependence described above suggest that regardless of pH, enzyme recharging is most likely the rate-limiting step in the complete catalytic cycle of the ligase IV/XRCC4 complex; to our knowledge, this is unique among all previously studied DNA ligases. Whether this atypical behavior has functional importance is an open question. However, this finding is of great practical value because it indicates that information about how the ligase IV/XRCC4 complex discriminates between matched versus mismatched base pairs is confined to the pre-steady state phase of the reaction. To obtain fidelity information, we were, therefore, faced with two options: (i) assay the complex in the steady state at pH 6.6 (where the distinct biphasic behavior is amenable to accurate fitting), fit time courses to the burst equation, and use only the pre-steady-state rate constants for comparing matched versus mismatched base pairs, or (ii) assay the complex at a pH that is more physiologically relevant, but instead of trying to fit the data to the burst equation (which gives poor fits at neutral pH), simply use single turnover reactions (in the absence of ATP) that can be accurately fit to a single exponential. Because both of these assay methods would consume similar quantities of enzyme and because both would yield pre-steady state kinetic parameters (rather than the steady-state parameters we had initially intended to generate), it seemed best to use the second option and work at the more physiologically relevant pH.

Influence of [KCl] on the Pre-Steady State Activity of the Ligase IV/XRCC4 Complex. Having decided, on the basis of the above discussion, to conduct nick ligation single turnover assays at pH 7.8 (the same pH at which our previous fidelity analyses were conducted), we next examined the activity of the ligase IV/XRCC4 complex as a function of ionic strength. Consistent with the report of Robins and Lindahl (19) but contrary to the report of Grawunder and co-workers (23), nick ligation activity decreases dramatically with increasing [KCl] (Figure 7). Though activity is maximal at 0 mM KCl, this condition is expected to yield artificially low fidelity values (i.e., high error-tolerance). Human DNA ligase I and the human DNA ligase III/XRCC1 complex both discriminate against mismatched nicks more strongly when assayed at high KCl concentrations that mimic the physiological state (1); similar findings have also been reported for T4 DNA ligase (38). Though the ligase IV/XRCC4 complex functions at only 20% of its maximal activity at 100 mM KCl, we decided to conduct assays at this ionic strength because (i) it represents a compromise between high activity and physiologically relevant fidelity values, and (ii)

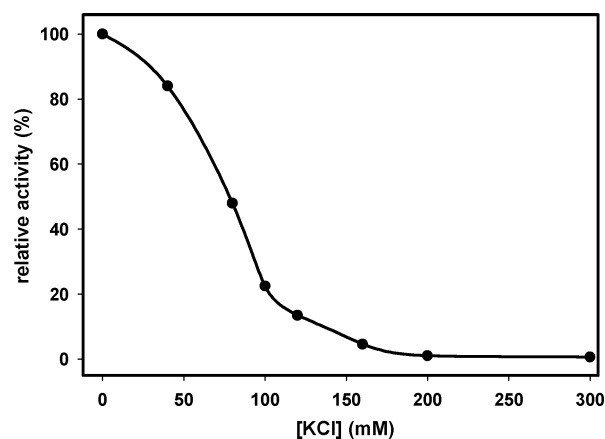


FIGURE 7: Influence of [KCl] on the pre-steady state activity of the ligase IV/XRCC4 complex. The ligase IV/XRCC4 complex (30 nM) was incubated with 160 nM nicked C:G substrate in the absence of ATP at KCl concentrations spanning from 0 to 300 mM.

this is the same [KCl] at which our previous fidelity analyses were conducted, making comparisons between the different enzymes as straightforward as possible (3).

Fidelity of Ligase IV and the Ligase IV/XRCC4 Complex. To elucidate the fidelity spectrum of the ligase IV/XRCC4 complex, we attempted to assay it against substrates containing eight different base pair combinations on the upstream (3'-OH) side of the nick. Though the G:G and G:A mismatches proved intractable, the two other base pairs containing G in the templating position (G:C and G:T) and the four base pairs containing C in the templating position (C:G, C:T, C:C, and C:A) all yielded usable data. A subset of these six base pairs was also assayed with free ligase IV in order to determine the specific catalytic and fidelity contributions of XRCC4. Whereas the representative time courses and saturation curves are shown in Figure 8, the complete set of kinetic parameters are displayed in Table 2. Salient features are described below.

(i) Consistent with what we previously observed for the ASFV and T4 DNA ligases, the C:T, C:C, C:A, and G:T mismatches do not significantly perturb the ligation rate constant of free ligase IV or the ligase IV/XRCC4 complex. The greatest mismatch-induced reduction in k_{lig} is a mere 5.6-fold (for the ligation of C:A vs C:G by the ligase IV/XRCC4 complex). Moreover, for the C:T mismatch, both the free and the complexed protein display k_{lig} values that are $\sim 25\%$ higher than the k_{lig} values for the corresponding Watson–Crick C:G base pair.

(ii) The inability, or minimal ability, of ligase IV and the ligase IV/XRCC4 complex to discriminate against mismatched base pairs through the rate of chemistry is compensated for by a reduction in substrate affinity. For all of the mismatches examined, both the free and the complexed protein display increased $K_{d,app}$ values relative to the corresponding correctly matched base pairs.

(iii) The above two properties combine to give rise to modest fidelity values ranging from 3.3 to 510. Consistent with what we previously observed for the ASFV and T4 DNA ligases (3), these fidelity values are significantly lower than those of DNA polymerases. Compared with ASFV and T4 ligases, ligase IV (in both its free and complexed states) displays higher fidelity for some mismatches and lower fidelity for other mismatches (Figure 9).

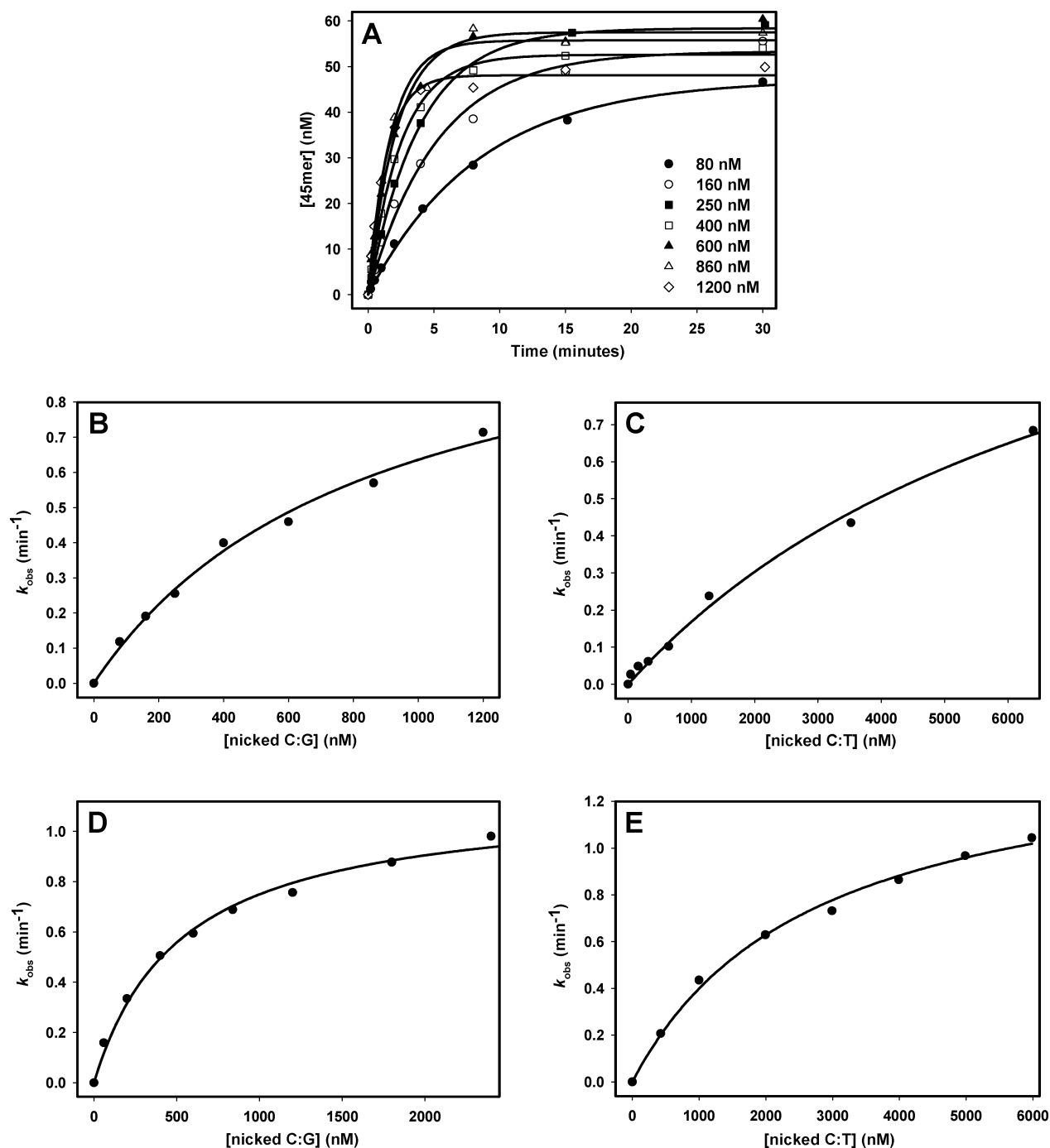


FIGURE 8: Representative time courses and saturation curves for nick ligation by free ligase IV and the ligase IV/XRCC4 complex. (A) Time courses for ligation of nicked C:G by the ligase IV/XRCC4 complex. Substrate concentrations are shown within the figure. (B and C) Saturation curves for ligation of nicked C:G and nicked C:T by the ligase IV/XRCC4 complex. (D and E) Saturation curves for ligation of nicked C:G and nicked C:T by free ligase IV.

(iv) How does XRCC4 influence the activity and fidelity of ligase IV? While XRCC4 has virtually no influence on the k_{lig} parameter for the C:G, C:T, and G:C base pairs, it reduces k_{lig} by $\sim 20\%$ for the G:T base pair. In contrast, XRCC4 significantly perturbs the $K_{\text{d,app}}$ parameter, causing the ligase IV/XRCC4 complex to always display lower affinity for substrate than free ligase IV. Importantly, the presence of XRCC4 reduces ligase IV's affinity for mismatched substrates more extensively than it reduces ligase IV's affinity for correctly matched substrates. Collectively, the above XRCC4-induced changes result in the ligase IV/XRCC4 complex being slightly less catalytically efficient

in the pre-steady state but also being of modestly higher fidelity than free ligase IV.

DISCUSSION

Expression/Purification of the Ligase IV/XRCC4 Complex.

An obstacle to the physical/kinetic/structural characterization of ligase IV/XRCC4 has been the inability to obtain large quantities of the pure complex. The coexpression and purification protocol described here (scheme 5 in Table 1) improves the yield of ligase IV ~ 8 -fold relative to what we and others (29) observe when expressing ligase IV alone. Because of the large size of ligase IV, coexpressing it with

Table 2: Kinetic Parameters for Nick Ligation by Free Ligase IV and the Ligase IV/XRCC4 Complex as a Function of 3' Base Pair Identity

base pair ^a	k_{lig} (s ⁻¹)	$K_{\text{d,app}}$ (nM)	$k_{\text{lig}}/K_{\text{d,app}}$ (M ⁻¹ s ⁻¹)	fidelity ^b
ligase IV/XRCC4				
C:G	0.020 ± 0.0026	850 ± 210	2.3×10^4	7.4
C:T	0.025 ± 0.0038	8000 ± 1900	3.1×10^3	
C:C ^c	0.011 ± 0.00050	15000 ± 1300	7.3×10^2	
C:A	0.0036 ± 0.00060	7200 ± 2200	5.0×10^2	
G:C	0.019 ± 0.0010	310 ± 56	6.1×10^4	
G:T ^c	0.0020 ± 0.00030	17000 ± 3300	1.2×10^2	510
ligase IV				
C:G	0.019 ± 0.00070	520 ± 58	3.6×10^4	3.3
C:T	0.024 ± 0.0049	2200 ± 780	1.1×10^4	
G:C	0.019 ± 0.00070	160 ± 20	1.2×10^5	
G:T	0.0025 ± 0.0011	2800 ± 1900	8.9×10^2	

^a In the base pair notation X:Y, X refers to the templating nucleotide, and Y is the nucleotide that would have been inserted by a polymerase in the preceding gap-filling step (see Figure 3 for structures). ^b Fidelity is defined as $(k_{\text{lig}}/K_{\text{d,app}})_{\text{correct}}/(k_{\text{lig}}/K_{\text{d,app}})_{\text{incorrect}}$. ^c The low affinity that the ligase IV/XRCC4 complex displays for these substrates made it impossible to achieve saturation of the enzyme.

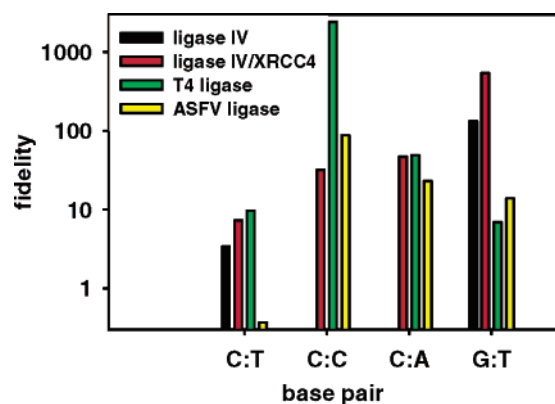


FIGURE 9: Comparison of the fidelities of four different DNA ligases. Note the logarithmic scale for the y-axis. Fidelity values for free ligase IV against C:C and C:A were not determined, hence the absence of the back bar for these base pairs. Fidelity is defined as $(k_{\text{lig}}/K_{\text{d,app}})_{\text{correct}}/(k_{\text{lig}}/K_{\text{d,app}})_{\text{incorrect}}$ (for free ligase IV and the ligase IV/XRCC4 complex) or $(k_{\text{cat}}/K_{\text{M}})_{\text{correct}}/(k_{\text{cat}}/K_{\text{M}})_{\text{incorrect}}$ (for ASFV and T4 DNA ligases), where correct and incorrect denote Watson–Crick and non-Watson–Crick base pairs, respectively. The data for the ASFV and T4 ligases is from ref 3(3).

XRCC4 in a chaperone-overexpressing strain would be worth consideration in any future effort to enhance yields. Our experience indicates that, at least for the expression schemes outlined in Table 1, the intracellular concentration of XRCC4 always exceeds that of ligase IV; purification protocols employing affinity tags should take this into consideration.

In mammalian cells, a large percentage of ligase IV molecules exist in the adenylylated state (19, 23). Similarly, when expressed in *E. coli*, we have consistently found ligase IV (in both its free and XRCC4-complexed forms) to purify predominantly in the adenylylated state. This could be a consequence of (i) the almost quantitative charging of the protein *in vivo* and/or (ii) a greater stability/resilience of the charged form during cell lysis and purification. Consistent with the former, even at early stages of ligase IV/XRCC4 purification, only small quantities of labeled protein are detected upon incubation with Mg^{2+} and $[\alpha\text{-}^{32}\text{P}]\text{ATP}$; consistent with the latter, we have found the unadenylylated form of ligase IV (in both its free and XRCC4-complexed

forms) to be less stable than the adenylylated form during extended periods of incubation at 37 °C (data not shown). Studies with other DNA ligases have demonstrated the presence of a conformational change upon enzyme adenylylation (39–42); in some cases, this was found to enhance thermostability (41, 42). In light of the above, the inclusion of a $\text{Mg}^{2+}/\text{ATP}$ incubation step at an early phase of protein purification could conceivably enhance the yield of ligase IV.

Unusual Rate-Limiting Step. At neutral pH, different preparations of free ligase IV and the ligase IV/XRCC4 complex invariably display nonlinear steady state time courses, suggesting this to be an inherent property of the enzyme. Our data for the ligase IV/XRCC4 complex indicates that this behavior is a consequence of a step after nick sealing (step 3 in Figure 6) being rate-limiting in the complete catalytic cycle. Though it has not been unequivocally proven, our data is consistent with this rate-limiting step being enzyme adenylylation.

The interaction of the ligase IV/XRCC4 complex with other components of the NHEJ machinery at synapsed DSBs may alter the kinetic properties of ligase IV. Accordingly, the biphasic behavior described here may not be biologically relevant; even if this behavior persists *in vivo*, whether it serves a functional purpose is questionable. However, knowledge of this behavior is of practical value because it demonstrates that unlike other DNA ligases but analogous to DNA polymerases, fidelity analyses of ligase IV/XRCC4 must be performed in the pre-steady state.

Use of Nicked Rather than DSB Substrates. *In vivo* ligase IV/XRCC4 appears to exclusively participate in the repair of DSBs (i.e., it is not involved in pathways such as BER that only involve single strand breaks). Despite this, our use of nicked substrates is appropriate for two reasons. (i) *Nicked substrates represent the intermediates that are generated after sealing the first strand of a DSB, and ligase IV is the best candidate for sealing these.* *In vivo*, ligase IV/XRCC4 is absolutely required for NHEJ (43–45); *in vitro*, ligase IV/XRCC4 can reconstitute NHEJ activity (15, 46). On the basis of these facts, the sealing of both DNA strands by ligase IV is implicit to models of NHEJ. However, this need not be the case. The initial ligation event in NHEJ is unique compared to ligation in other forms of DNA repair because it frequently utilizes a nick that is not flanked on both sides by extensive duplex regions (Figure 1A and B). In contrast, ligation of the second strand of a DSB is analogous to what occurs in other repair pathways. Accordingly, it is conceivable that each strand of a DSB could be sealed sequentially by a different human DNA ligase. However, the fact that ligase IV is indispensable for NHEJ *in vivo* (43–45) suggests that this ligase alone is equipped for cooperating with the other NHEJ factors in sealing the first and/or the second strand of a DSB. Moreover, the fact that mammalian cell lines with deficiencies in DNA ligases I and III do not display a deleterious NHEJ phenotype (47–50) suggests that it is ligase IV that is responsible for sealing both strands of a DSB *in vivo*. Accordingly, the nicked substrates used here represent true ligase IV substrates that are present in the latter half of the DSB repair reaction. (ii) *Nicked substrates eliminate the need for synapsis.* The critical features of NHEJ-mediated DSB repair are synapsis, end processing, and the ligation of each strand. Though ligase IV/XRCC4

can, in the absence of other NHEJ factors, inefficiently synapse and ligate cohesive DSBs *in vitro* (23, 51), *in vivo* ligase IV/XRCC4 is expected to act on nicks that are synapsed/stabilized by the NHEJ machinery as a whole. By obviating the necessity of synapsis, the nicked substrates used here enable a focused analysis of the nick sealing step alone.

Influence of XRCC4. Using a qualitative steady state assay, Gellert and co-workers demonstrated an XRCC4-induced enhancement of ligase IV nick sealing (29). However, in our pre-steady state assays, the presence of XRCC4, for every base pair analyzed, decreases the catalytic efficiency of ligase IV nick sealing, predominantly by reducing substrate affinity (Table 2). The most reasonable explanation for XRCC4 reducing the efficiency of the pre-steady state (steps 2 and 3) and enhancing the efficiency of the steady state is that XRCC4 stimulates ligase IV adenylation (step 1). This interpretation is consistent with multiple reports demonstrating that XRCC4 increases the quantity of adenylylated ligase IV formed during a given length of incubation (29, 52).

Though somewhat deleterious to the efficiency of ligase IV-catalyzed nick sealing, XRCC4 provides a modest enhancement in nick ligation fidelity, achieved predominantly through differential substrate affinity. Although XRCC4 reduces the affinity of ligase IV for all substrates, this reduction is more severe for mismatched nicks than it is for correctly base paired nicks (Table 2). How exactly is this accomplished? Does any portion of the XRCC4 dimer contact the nick? Does the presumed XRCC4-induced ligase IV conformational change produce a more discriminating active site? Answers to these questions await the structural characterization of ligase IV in its free and complexed forms.

Considering the modest impact that XRCC4 has on ligase IV activity and fidelity during nick sealing, evolution of the extremely tight association between these two proteins ($K_d \sim 10^{-10}$ M (20)) must have been driven by other selective pressures. Besides XRCC4's ability to stabilize/protect ligase IV (28) and stimulate ligase IV adenylation (29, 52), it can significantly enhance ligase IV-mediated DSB ligation in an *in vitro* single turnover assay (23). Because XRCC4-mediated stimulation of enzyme adenylation cannot account for this behavior (ATP was absent), a likely explanation would appear to be that XRCC4 facilitates DSB synapsis. Whether this is a consequence of XRCC4 interacting directly with the DNA fragments or XRCC4 enhancing the synapsis activity of ligase IV is presently unclear but is worthy of future investigation.

Fidelity of Ligase IV and Its Biological Implications. Our finding that analogous to the recently discovered error-prone DNA polymerases, the ASFV DNA ligase does not rigorously adhere to the canonical Watson–Crick base pairing rules prompted a consideration of how this relaxed substrate specificity might enhance viral fitness/diversity and whether error-tolerant DNA ligation might be useful in other pathways/organisms. We speculated that 3' mismatch tolerance in the NHEJ-specific human DNA ligase IV could potentially be beneficial by (i) enhancing the subexon junctional diversity that is created during V(D)J recombination (Figure 2A) and (ii) minimizing the number of large deletions that result from having to use regions of microhomology that are distal to the DSB termini (Figure 2B). We were unable to assess the extent to which mismatch ligation may have contributed to joining events in previously published analyses of NHEJ

because products that could have been the result of mismatch ligation may also have been formed by other non-mismatched pathways of NHEJ (Figure 2C). Accordingly, we sought to examine the 3'-mismatch-utilizing potential of NHEJ via a reductionistic analysis of DNA ligase IV. The data in Table 2 and the fidelity comparison with other DNA ligases in Figure 9 indicate that ligase IV, in both its free and XRCC4-complexed forms, is not especially adept at ligating nicks containing 3' mismatched base pairs. The ligase IV/XRCC4 complex seals nicks containing the 3' C:T, C:C, C:A, and G:T mismatches 7.4-, 32-, 46-, and 510-fold, respectively, less efficiently than the corresponding Watson–Crick base pairs. Although these fidelity values are quite low relative to a standard DNA polymerase⁸, they fall within a range that is typical of DNA ligases (3).

The relative inefficiency of 3' mismatch ligation by ligase IV/XRCC4 suggests that during NHEJ, 3' mismatched intermediates are not likely to be consistently converted directly to product. The NHEJ-specific DNA Pols λ and μ do not possess 3' \rightarrow 5' exonuclease activity; therefore, nicks containing 3' mismatches are presumably either (i) simultaneously held in synapse and processed by Artemis or some other as of yet unidentified nuclease to generate a gap that can subsequently be filled (this would resemble the nick editing and resynthesis effected by APE1 and Pol β in BER) or (ii) released from the synapse so that more extensive end modification and/or microhomology searching can occur.

Nicks containing mismatches on the 3'-OH side, rather than the 5'-phosphate side, were examined in this study for two reasons. First, because they can be generated both by polymerase error and by the synapsis of strands that are not fully complementary, mismatches on the 3'-side of a nick are more likely to occur than are mismatches on the 5'-side of a nick⁹. Second, numerous studies have uniformly demonstrated that 3' mismatches are more difficult for DNA ligases to cope with than are mismatches on the 5'-side. Accordingly, 3' mismatches were deemed to constitute the most rigorous test of NHEJ's capacity for mismatch tolerance. Despite the relative intolerance of ligase IV for non-canonical base pairs on the 3'-side of a nick, it may still be worthwhile to examine how efficiently 5' mismatches are utilized.

As has been pointed out elsewhere (46), NHEJ is utilized for the repair of DSBs within different contexts that have opposing objectives. Although accuracy (minimization of nucleotides lost and/or substituted) during the repair of chemically/physically generated DSBs at random positions in the genome is critical for stability/longevity, it is equally important that the repair of programmed DSBs during V(D)J recombination yields junctional diversity. Though, as with any reductionistic biochemical analysis, it is possible that the nick ligation fidelity of ligase IV/XRCC4 alone does not accurately reflect the fidelity of the full NHEJ complex, in order for 3' mismatch ligation to be biologically relevant, the other NHEJ factors would need to considerably relax the substrate specificity of ligase IV. This does not seem likely. Accordingly, it appears that (i) the mismatch ligation

⁸ Using a similar definition of fidelity, the average fidelity of the moderately error-prone eukaryotic Pol β is $\sim 33,000$ (53).

⁹ Owing to the inherent polarity of DNA synthesis, mismatches on the 5'-side of a nick can be generated by the synapsis of incompatible strands but not by polymerase synthesis errors.

scenarios that we posited are not particularly advantageous, (ii) extreme mismatch tolerance by ligase IV is too dangerous/difficult to control outside the special context of V(D)J recombination, or (iii) it simply has not been possible to evolve ligase IV toward extremely low fidelity while maintaining sufficient activity.

Our findings have implications for DNA transactions other than NHEJ. When coupled with Lindahl and coworkers' analysis of human DNA ligases I and III (1), our data suggests that in human cells nicks containing 3' mismatched base pairs will always be sealed inefficiently, regardless of the repair pathway in which they occur in human cells. Somatic hypermutation, wherein immunoglobulin loci are further diversified (beyond what was previously accomplished by V(D)J recombination) via base substitutions, was previously hypothesized to proceed, at least in part, via mutagenic patch repair (54, 55). Although mutation of G:C base pairs now appears to occur by replication over abasic sites (generated by deamination of cytosine and the subsequent removal of uracil) (56, 57), the mechanism of mutation at A:T base pairs awaits elucidation. The collective body of data on human DNA ligase fidelity now indicates that if mutagenic patch repair is utilized during SHM, gaps must be larger than one nucleotide in length, and any mismatches within the gap must occur upstream of the nucleotide that constitutes the 3'-OH side of the nick.

ACKNOWLEDGMENT

We thank T. Ellenberger for the gift of pRSFDuet-1 and S. Jackson for providing us with pET30-*xrcc4*.

REFERENCES

- Bhagwat, A. S., Sanderson, R. J., and Lindahl, T. (1999) Delayed DNA joining at 3' mismatches by human DNA ligases, *Nucleic Acids Res.* 27, 4028–4033.
- Chou, K. M., and Cheng, Y. C. (2002) An exonucleolytic activity of human apurinic/apyrimidinic endonuclease on 3' mispaired DNA, *Nature* 415, 655–659.
- Lamarche, B. J., Showalter, A. K., and Tsai, M.-D. (2005) An error-prone viral DNA ligase, *Biochemistry* 44, 8408–8417.
- Showalter, A. K., Byeon, I. J., Su, M. I., and Tsai, M. D. (2001) Solution structure of a viral DNA polymerase X and evidence for a mutagenic function, *Nat. Struct. Biol.* 8, 942–946.
- Lamarche, B. J., and Tsai, M. D. (2006) Contributions of an endonuclease IV homologue to DNA repair in the African swine fever virus, *Biochemistry* 45, 2790–2803.
- Sutherland, B. M., Bennett, P. V., Sidorkina, O., and Laval, J. (2000) Clustered DNA damages induced in isolated DNA and in human cells by low doses of ionizing radiation, *Proc. Natl. Acad. Sci. U.S.A.* 97, 103–108.
- Bradley, M. O., and Kohn, K. W. (1979) X-ray induced DNA double strand break production and repair in mammalian cells as measured by neutral filter elution, *Nucleic Acids Res.* 7, 793–804.
- Demple, B., Johnson, A., and Fung, D. (1986) Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H₂O₂-damaged *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A.* 83, 7731–7735.
- Bassing, C. H., and Alt, F. W. (2004) The cellular response to general and programmed DNA double strand breaks, *DNA Repair* 3, 781–796.
- Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. (2003) Mechanism and regulation of human non-homologous DNA end-joining, *Nat. Rev. Mol. Cell Biol.* 4, 712–720.
- Ahnesorg, P., Smith, P., and Jackson, S. P. (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining, *Cell* 124, 301–313.
- Buck, D., Malivert, L., de Chasseval, R., Barraud, A., Fondaneche, M. C., Sanal, O., Plebani, A., Stephan, J. L., Hufnagel, M., Le Deist, F., Fischer, A., Durandy, A., de Villartay, J. P., and Revy, P. (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly, *Cell* 124, 287–299.
- Hentges, P., Ahnesorg, P., Pitcher, R. S., Bruce, C. K., Kysela, B., Green, A. J., Bianchi, J., Wilson, T. E., Jackson, S. P., and Doherty, A. J. (2006) Evolutionary and functional conservation of the DNA non-homologous end-joining protein, XLF/Cernunnos, *J. Biol. Chem.* 281, 37517–37526.
- Roth, D. B., and Wilson, J. H. (1986) Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction, *Mol. Cell. Biol.* 6, 4295–4304.
- Ma, Y., Lu, H., Tippin, B., Goodman, M. F., Shimazaki, N., Koiki, O., Hsieh, C. L., Schwarz, K., and Lieber, M. R. (2004) A biochemically defined system for mammalian nonhomologous DNA end joining, *Mol. Cell* 16, 701–713.
- Max, E. E., Seidman, J. G., and Leder, P. (1979) Sequences of five potential recombination sites encoded close to an immunoglobulin kappa constant region gene, *Proc. Natl. Acad. Sci. U.S.A.* 76, 3450–3454.
- Gerstein, R. M., and Lieber, M. R. (1993) Extent to which homology can constrain coding exon junctional diversity in V(D)J recombination, *Nature* 363, 625–627.
- Critchlow, S. E., Bowater, R. P., and Jackson, S. P. (1997) Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV, *Curr. Biol.* 7, 588–598.
- Robins, P., and Lindahl, T. (1996) DNA ligase IV from HeLa cell nuclei, *J. Biol. Chem.* 271, 24257–24261.
- Modesti, M., Junop, M. S., Ghirlando, R., Rakt, M. v. d., Gellert, M., Yang, W., and Kanaar, R. (2003) Tetramerization and DNA Ligase IV interaction of the DNA double-strand break repair protein XRCC4 are mutually exclusive, *J. Mol. Biol.* 334, 215–228.
- Sibanda, B. L., Critchlow, S. E., Begun, J., Pei, X. Y., Jackson, S. P., Blundell, T. L., and Pellegrini, L. (2001) Crystal structure of an XRCC4-DNA ligase IV complex, *Nat. Struct. Biol.* 8, 1015–1019.
- 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., and Lindahl, T. (1995) Molecular cloning and expression of human cDNAs encoding a novel DNA Ligase IV and DNA Ligase III, an enzyme active in DNA repair and recombination, *Mol. Cell. Biol.* 15, 3206–3216.
- Grawunder, U., Wilm, M., Wu, X., Kulesza, P., Wilson, T. E., Mann, M., and Lieber, M. R. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells, *Nature* 388, 492–495.
- Caldecott, K. W., McKeown, C. K., Tucker, J. D., Ljungquist, S., and Thompson, L. H. (1994) An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III, *Mol. Cell. Biol.* 14, 68–76.
- Mortusewicz, O., Rothbauer, U., Cardoso, M. C., and Leonhardt, H. (2006) Differential recruitment of DNA Ligase I and III to DNA repair sites, *Nucleic Acids Res.* 34, 3523–3532.
- Grawunder, U., Zimmer, D., Kulesza, P., and Lieber, M. R. (1998) Requirement for an interaction of XRCC4 with DNA Ligase IV for wild-type V(D)J recombination and DNA double-strand break repair in vivo, *J. Biol. Chem.* 273, 24708–24714.
- O'Driscoll, M., Cerosaletti, K. M., Girard, O. M., Dai, Y., Stumm, M., Kysela, B., Hirsch, B., Gennerly, A., Palmer, S. E., Seidel, J., Gatti, R. A., Varon, R., Oettinger, M. A., Neitzel, H., Jeggo, P. A., and Concannon, P. (2001) DNA Ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency, *Mol. Cell* 8, 1175–1185.
- Bryans, M., Valenzano, M. C., and Stamato, T. D. (1999) Absence of DNA ligase IV protein in XR-1 cells: evidence for stabilization by XRCC4, *Mutat. Res.* 433, 53–58.
- Modesti, M., Hesse, J. E., and Gellert, M. (1999) DNA binding of XRCC4 protein is associated with V(D)J recombination but not with stimulation of DNA ligase IV activity, *EMBO J.* 18, 2008–2018.
- Grawunder, U., Zimmer, D., and Lieber, M. R. (1998) DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains, *Curr. Biol.* 8, 873–876.
- Lee, K. J., Huang, J., Takeda, Y., and Dynan, W. S. (2000) DNA Ligase IV and XRCC4 form a stable mixed tetramer that functions synergistically with other repair factors in a cell-free end-joining system, *J. Biol. Chem.* 275, 34787–34796.

32. Jones, J. M., and Gellert, M. (2001) Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends, *Proc. Natl. Acad. Sci. U.S.A.* 98, 12926–12931.
33. Cherepanov, A. V., and de Vries, S. (2003) Kinetics and thermodynamics of nick sealing by T4 DNA ligase, *Eur. J. Biochem.* 270, 4315–4325.
34. Lehman, I. R. (1974) DNA ligase: structure, mechanism, and function, *Science* 186, 790–797.
35. Shuman, S., and Lima, C. D. (2004) The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases, *Curr. Opin. Struct. Biol.* 14, 757–764.
36. Sriskanda, V., and Shuman, S. (1998) Chlorella virus DNA ligase: nick recognition and mutational analysis, *Nucleic Acids Res.* 26, 525–531.
37. Wilkinson, A., Smith, A., Bullard, D., Lavesa-Curto, M., Sayer, H., Bonner, A., Hemmings, A., and Bowater, R. (2005) Analysis of ligation and DNA binding by *Escherichia coli* DNA ligase (LigA), *Biochim. Biophys. Acta* 1749, 113–122.
38. Wu, D. Y., and Wallace, R. B. (1989) Specificity of the nick-closing activity of bacteriophage T4 DNA ligase, *Gene* 76, 245–254.
39. Odell, M., Sriskanda, V., Shuman, S., and Nikolov, D. B. (2000) Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining, *Mol. Cell* 6, 1183–11893.
40. Gajiwala, K. S., and Pinko, C. (2004) Structural rearrangement accompanying NAD⁺ synthesis within a bacterial DNA ligase crystal, *Structure* 12, 1449–1459.
41. Georlette, D., Blaise, V., Bouillenne, F., Damien, B., Thorbjarnardottir, S. H., Depiereux, E., Gerday, C., Uversky, V. N., and Feller, G. (2004) Adenylation-dependent conformation and unfolding pathways of the NAD⁺-dependent DNA ligase from the thermophile *Thermus scotoductus*, *Biophys. J.* 86, 1089–1104.
42. Georlette, D., Blaise, V., Dohmen, C., Bouillenne, F., Damien, B., Depiereux, E., Gerday, C., Uversky, V. N., and Feller, G. (2003) Cofactor binding modulates the conformational stabilities and unfolding patterns of NAD(+)–dependent DNA ligases from *Escherichia coli* and *Thermus scotoductus*, *J. Biol. Chem.* 278, 49945–49953.
43. Wilson, T. E., Grawunder, U., and Lieber, M. R. (1997) Yeast DNA ligase IV mediates non-homologous DNA end joining, *Nature* 388, 495–498.
44. Teo, S. H., and Jackson, S. P. (1997) Identification of *Saccharomyces cerevisiae* DNA ligase IV: involvement in DNA double-strand break repair, *EMBO J.* 16, 4788–4795.
45. Grawunder, U., Zimmer, D., Fugmann, S., Schwarz, K., and Lieber, M. R. (1998) DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes, *Mol. Cell* 2, 477–484.
46. Nick McElhinny, S. A., Havener, J. M., Garcia-Diaz, M., Juarez, R., Bebenek, K., Kee, B. L., Blanco, L., Kunkel, T. A., and Ramsden, D. A. (2005) A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining, *Mol. Cell* 19, 357–366.
47. Taccioli, G. E., Rathbun, G., Oltz, E., Stamato, T., Jeggo, P. A., and Alt, F. W. (1993) Impairment of V(D)J recombination in double-strand break repair mutants, *Science* 260, 207–210.
48. Hsieh, C. L., Arlett, C. F., and Lieber, M. R. (1993) V(D)J recombination in ataxia telangiectasia, Bloom's syndrome, and a DNA ligase I-associated immunodeficiency disorder, *J. Biol. Chem.* 268, 20105–20109.
49. Petriani, J. H., Donovan, J. W., Dimare, C., and Weaver, D. T. (1994) Normal V(D)J coding junction formation in DNA ligase I deficiency syndromes, *J. Immunol.* 152, 176–183.
50. Kemp, L. M., Sedgwick, S. G., and Jeggo, P. A. (1984) X-ray sensitive mutants of Chinese hamster ovary cells defective in double-strand break rejoining, *Mutat. Res.* 132, 189–196.
51. Smith, J., Riballo, E., Kysela, B., Baldeyron, C., Manolis, K., Masson, C., Lieber, M. R., Papadopoulos, D., and Jeggo, P. (2003) Impact of DNA ligase IV on the fidelity of end joining in human cells, *Nucleic Acids Res.* 31, 2157–2167.
52. Girard, P. M., Kysela, B., Harer, C. J., Doherty, A. J., and Jeggo, P. A. (2004) Analysis of DNA ligase IV mutations found in LIG4 syndrome patients: the impact of two linked polymorphisms, *Hum. Mol. Genet.* 13, 2369–2376.
53. Ahn, J., Kraynov, V. S., Zhong, X., Werneburg, B. G., and Tsai, M. D. (1998) DNA polymerase beta: effects of gapped DNA substrates on dNTP specificity, fidelity, processivity and conformational changes, *Biochem. J.* 331, 79–87.
54. Martomo, S. A., and Gearhart, P. J. (2006) Somatic hypermutation: subverted DNA repair, *Curr. Opin. Immunol.* 18, 243–248.
55. Bertocci, B., Quint, L., Delbos, F., Garcia, C., Reynaud, C. A., and Weill, J. C. (1998) Probing immunoglobulin gene hypermutation with microsatellites suggests a nonreplicative short patch DNA synthesis process, *Immunity* 9, 257–265.
56. Di Noia, J., and Neuberger, M. S. (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase, *Nature* 419, 43–48.
57. Neuberger, M. S., Di Noia, J. M., Beale, R. C., Williams, G. T., Yang, Z., and Rada, C. (2005) Somatic hypermutation at AT pairs: polymerase error versus dUTP incorporation, *Nat. Rev. Immunol.* 5, 171–178.