Structures of LIG1 uncover a lack of sugar discrimination against a ribonucleotide at the 3'-end of nick DNA

Qun Tang[1], Mitchell Gulkis[1], and Melike Çağlayan[1,*]

[1]Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610, USA

*To whom correspondence should be addressed. Tel.: +1 352-294-8383; E-mail: caglayanm@ufl.edu

Ribonucleotides can be incorporated by DNA polymerases and the subsequent joining of 3'-OH and 5'-P ends in the phosphodiester backbone at the nick by DNA ligase during DNA replication and repair is critical for maintaining genome stability. Although it has been extensively studied for DNA polymerases across families, the sugar discrimination mechanism of a human DNA ligase at atomic resolution is entirely missing. Here, for the first time, we determine X-ray structure of DNA ligase I (LIG1) in complex with nick DNA containing rG:C at the 3'-end and capture the ligase at the final phosphodiester bond formation step of the ligation reaction involving an adenylate (AMP) release. Moreover, we show mutagenic end joining of the nick DNA substrate with preinserted 3'-rG:C by LIG1 *in vitro*. Our findings reveal an important source of ribonucleotides embedded in genomic DNA, which could come from the failure of LIG1 to discriminate against a ribonucleotide at the 3'-end during nick sealing step of DNA replication and repair.

Keywords: Ribonucleotide, rNTP, DNA polymerase, DNA ligase I, DNA ligase IIIα/XRCC1 complex, mutagenic ligation, DNA repair, DNA replication, genome stability

Introduction

The concentrations of ribonucleotide triphosphates (rNTPs) are much more abundant than those of deoxyribonucleotide triphosphates (dNTPs) within a cell, and therefore, their

26 misincorporation into genomic DNA by DNA polymerases occur at higher frequencies than
27 mismatch nucleotide insertions (1-4). It has been estimated that >10,000 or >1,000,000
28 ribonucleotides can be incorporated into yeast and mouse genomes, respectively, making
29 genomic ribonucleotides the most prevalent source of cellular DNA damage (5-13).
30 Ribonucleotides embedded into genomic DNA can confound structural and chemical integrity
31 of duplex DNA, increase its susceptibility to endogenous or exogenous damage, and lead to
32 several types of genome instability such as replication blockage, mutagenesis, aberrant
33 recombination, protein-DNA crosslinks, double-strand breaks (DSBs), and chromosome
34 alterations if not repaired by ribonucleotide excision repair (RER) (14-20).
35 In human cells, ribonucleotide incorporation can occur during almost all cellular DNA
36 transactions such as DNA replication by replicative polymerases (pols α, δ, and ε), non-
37 homologous end joining (NHEJ) of DSB repair by pol μ and base excision repair (BER) by
38 pols β and λ, translesion synthesis (TLS) by pol ι, and mitochondrial repair by pol γ (6-11,21-
39 26). Due to the high cellular rNTP/dNTP imbalance, discrimination against ribonucleotide
40 incorporation by these DNA polymerases is essential for the maintenance of genomic integrity
41 (1-4). It has been extensively shown by structure/function studies that most DNA polymerases
42 use a side chain named the "steric gate" (*i.e.*, B-, and Y- family pols) or a protein backbone
43 segment in minor groove nucleotide binding pocket (*i.e*, X-family pols) to govern
44 ribonucleotide exclusion (6-26). Replicative DNA polymerases effectively remove mismatches
45 through their 3'-5' exonuclease domains, which is found to be relatively weaker at proofreading
46 ribonucleotides during DNA synthesis (27). The kinetic basis of inefficient rNTP incorporation
47 by DNA polymerases is mainly governed by weaker binding and slower rate of incorporation
48 compared to dNTPs (21).
49 DNA polymerases and DNA ligases have fundamental roles in maintaining genome stability
50 during the downstream steps of DNA replication and repair (35). Indeed, the efficiency of DNA

51 ligation by DNA ligase at the final step relies on the accuracy of DNA synthesis that requires
52 a Watson-Crick base pair between 5'-phosphate (P) and 3'-hydoxyl (OH) termini of the nick
53 after DNA polymerase incorporates a correct dNTP (36). In our previous studies, we reported
54 the importance of the coordination between DNA polymerase (pol) β and DNA ligase I (LIG1)
55 or DNA ligase IIIα/XRCC1 (LIGIIIα/XRCC1) complex at the downstream steps of BER
56 pathway to maintain accurate repair (37-47). Our biochemical studies have revealed that, in
57 certain situations, BER responses can lead to mutagenic outcomes and the DNA ligation step
58 of the BER pathway can be compromised in case of damaged (8-oxodGTP) or mismatched
59 nucleotide incorporation by polβ (35-52). Furthermore, at atomic resolution, we recently
60 reported that the LIG1 discriminates against the mutagenic mismatches distinctly depending
61 on the architecture of 3'-terminus at the nick DNA. Our structures of LIG1/nick DNA
62 complexes with G:T and A:C mismatches revealed the ligase strategies that favor or deter the
63 ligation of base substitution errors at the final step of BER pathway (53).
64 Although the distinct features of the sugar selection mechanism have been extensively reported
65 for DNA polymerases across families by X-ray crystallography studies (7-34), the mechanism
66 of ribonucleotide discrimination by human DNA ligases at atomic resolution is entirely
67 missing. Given the fact of high cellular rNTP/dNTP ratio, the mutagenic ligation of
68 ribonucleotides by a DNA ligase at the final step following their incorporation by a DNA
69 polymerase at a non-negligible rate during DNA replication or repair could constitute a major
70 threat to genome stability. In the present study, we aimed to elucidate the mechanism by which
71 human LIG1 facilitates or hinders sealing ribonucleotide at the 3'-end of nick DNA. Our
72 LIG1/nick DNA structure uncovers a lack of a proficient sugar discrimination and
73 demonstrates that the ligase active site can accommodate 3'-riboG opposite template base C.
74 For the first-time, we showed a phosphodiester bond formation between 3'-OH and 5'-P ends
75 of the nick DNA, which refers to the final step of the ligation reaction involving an adenylate

76 (AMP) release. Moreover, we demonstrated the mutagenic ligation of nick DNA substrate
77 containing 3'-preinserted rG:C by LIG1 and LIGIIIα/XRCC1 complex in the ligation assays *in*
78 *vitro*. Overall, our findings provide a novel insight into the characterization of ribonucleotide
79 selectivity on the downstream events of DNA repair and replication as well as a lack of sugar
80 discrimination by LIG1, demonstrating its distinctive ability to seal nick DNA with 3'-rG:C at
81 atomic resolution.

82 **Results**

83 **Structures of LIG1/rG:C demonstrate a failure to discriminate against a ribonucleotide**
84 **at the 3'-end of nick DNA**

85 To elucidate the sugar discrimination mechanism of LIG1 at atomic resolution, we determined
86 the structure of LIG1 in complex with nick DNA containing rG:C at the 3'-end (Table 1). Our
87 LIG1/3'-rG:C structure shows density for the 2'-hydroxyl (2'-OH) group on the ribose ring and
88 demonstrates the formation of a phosphodiester bond between adjacent 3'-OH and 5'-P ends of
89 DNA termini at the nick (Figure 1). This LIG1 structure represents the first structure of human
90 DNA ligase that engages with a ribonucleotide-containing nick DNA during the final step of
91 the ligation reaction where 3'-OH terminus of the nick attacks the 5'-P terminus downstream of
92 the nick to form a phosphodiester bond that is coupled to a release of adenylate (AMP).

93 In our previously solved LIG1/A:T structure, we reported that the active site can accommodate
94 a cognate A:T base pair where DNA-AMP intermediate is formed (53). The comparison of the
95 2Fc-Fo maps at 3'-OH and 5'-P ends of the nick DNA demonstrated continuous and
96 discontinuous maps in the LIG1 structures of rG:C and A:T, respectively (Figure 2A-B). This
97 means that both DNA ends are positioned for nick sealing in the LIG1/A:T structure, which
98 refers to the step 2 of the ligation reaction, and a phosphodiester bond between 3'-OH and 5'-P
99 ends is formed in the LIG1/3'-rG:C structure. The Fo-Fc map at 3σ (Figure 1B) and the 2Fo-
100 Fc map at 1σ (Figure 2B) demonstrate that the AMP is incomplete in the LIG1/rG:C and shows

101 less density than the LIG1/A:T structure. This observation indicates that AMP is still linked to
102 the 5'-P end to form DNA-AMP intermediate in the LIG1/A:T structure, while it's already
103 released in the LIG1/3'-rG:C structure, which refers to the step 3 of the ligation reaction. The
104 overlay of both LIG1 structures shows a superimposed Cα root mean square deviation of 0.398
105 Å (Figure 2C) and no significant conformational difference in the domain organization of the
106 LIG1 catalytic core (Supplementary Figure 1).
107 Our previously solved LIG1/mismatch structures demonstrated that the A:C is positioned at
108 the initial step (step 1) where the ligase active site stays adenylated (LIG1-AMP) at its active
109 site lysine residue (K568), and that the LIG1/G:T mismatch is captured during step 2 when an
110 adenylate is transferred to the 5'-end of the nick (53), while the LIG1/rG:C structure shows a
111 phosphodiester bond formation coupled to AMP release at step 3 of the ligation reaction
112 (Supplementary Figure 2). We also compared the map of LIG1/rG:C (EE/AA) with the
113 previously solved structure of LIG1/C:G (wild-type) (54,55). The overlay of both LIG1
114 structures demonstrated the 2Fc-Fo discontinuous map at 3'-OH and 5'-P ends of the nick for
115 LIG1/C:G and a continuous map for LIG1/rG:C (Supplementary Figure 3). We observed a
116 Watson-Crick conformation for rG:C and sugar pucker analysis of LIG1/3'-rG:C structure
117 showed that the ribose adopts a C3'-endo conformation, which is similar to that of the
118 LIG1/A:T structure (Supplementary Figure 4).
119 **Mutagenic ligation of nick DNA containing a ribonucleotide at 3'-end by LIG1**
120 To investigate the sugar discrimination mechanism of LIG1 against a ribonucleotide at the 3'-
121 end of the nick *in vitro*, we tested the ligation efficiency of wild-type and EE/AA mutant of
122 LIG1 using the nick DNA substrate with 3'-preinserted rG opposite template base C that could
123 be formed during DNA replication or repair when DNA polymerase incorporates a rNTP
124 (Supplementary Scheme 1). In the control reactions, we used the nick DNA substrate with 3'-
125 dG:C.

126 Our results demonstrated the mutagenic ligation of nick DNA with 3'-rG:C by wild-type
127 (Figure 3A, lines 9-14) and EE/AA mutant (Figure 3C, lines 9-14) of LIG1. This was
128 comparable with the ligation products of nick DNA containing a cognate 3'-dG:C (Figure 3A
129 and 3C, lines 2-7). We observed a similar amount of ligation products for both DNA substrates
130 (3'-rG:C *versus* 3'-dG:C) with wild-type (Figure 3B) and the low fidelity mutant of the ligase
131 (Figure 3D). Similarly, our results demonstrated no significant difference between the
132 mutagenic ligation efficiencies of both LIG1 proteins (Supplementary Figures 5-6).
133 In our structure studies, we use the EE/AA mutant of LIG1 carrying Glu(E)346 and Glu(E)592
134 mutations at the Mg^{2+}-dependent high-fidelity site to crystallize the ligase in the conditions
135 lacking Mg^{2+} (53,55). In the reaction conditions that mimic the LIG1 crystals, we also observed
136 efficient nick sealing of the DNA substrate with 3'-rG:C in the absence of Mg^{2+} for longer time
137 points of incubation (Supplementary Figure 7).
138 In our previous studies (38,42,45,53), we reported that LIG1 can ligate nick DNA containing
139 a Watson-Crick like mismatch (G:T) and oxidative damage (8-oxoG:A). We next compared
140 the mutagenic ligation efficiency of LIG1 for the nick DNA substrates containing a
141 mismatched, damaged, and ribonucleotide-containing ends (Supplementary Figure 8). Our
142 results showed that the ligation of 3'-8-oxoG:A and 3'-G:T yields nick sealing products along
143 with the accumulation of DNA-AMP intermediates (Supplementary Figure 8A, lines 2-7 and
144 9-14, respectively), while there was only mutagenic ligation products in the presence of the
145 nick DNA with 3'-rG:C (Supplementary Figure 8A, lines 16-21). These results demonstrate
146 more efficient nick sealing of a ribonucleotide than a mismatched or damaged base at the 3'-
147 end, which could explain why the ligase active site engages with those DNA ends at different
148 steps of the ligation reaction (Supplementary Figure 8B).
149 **Role of LIG1 active site residues for sugar discrimination against 3'-rG:C at the nick**
150 **DNA**

151 We previously reported that the LIG1 active site shows distinct DNA conformations depending
152 on the identity of the 3'-terminus (cognate *versus* mismatched) (53). The overlay of LIG1
153 structures in complex with nick DNA containing 3'-rG:C and 3'-A:T showed no conformational
154 change at the positions of the LIG1 active site residues F635, R871, and F872, which reside in
155 the close proximity to the 2'-OH of the ribose (Figure 4A). Furthermore, we did not observe a
156 shift in the position of F872 at the ligase active site engaging with rG:C as we reported
157 previously in the LIG1/A:C structure (Figure 4B).
158 To further investigate this observation, we tested the importance of the ligase active site residue
159 F872 for the end joining efficiency of the nick DNA substrate with 3'-rG:C by LIG1 in the
160 ligation assays (Supplementary Scheme 1). For this purpose, we used the LIG1 mutants
161 carrying a F872A mutation in the wild-type (F872A) and low fidelity mutant (EE/AA)
162 background. Our results demonstrated that LIG1 F872A (Figure 5A, lines 2-7) and LIG1
163 EE/AA/F872A (Figure 5A, lines 8-13) can also efficiently seal nick DNA with 3'-rG:C. There
164 were slightly less ligation products at the initial time points of the reaction (Figure 5B and
165 Supplementary Figure 9) in comparison with the wild-type and EE/AA mutant of LIG1 (Figure
166 3). These findings demonstrate that the active site residue F872 that is located upstream of the
167 nick and positioned close to the deoxyribose moiety of the nucleotide at the 5'-end has no role
168 in sugar discrimination by LIG1.

169 **Discussion**

170 Ribonucleotides (rNTPs) are found at a concentration several orders of magnitude higher than
171 dNTPs in most cells (1-3). It has been shown that the eukaryotic replicative polymerases insert
172 around 10,000 rNTPs per round of DNA replication and repair polymerases are also capable
173 of inserting rNTPs during DNA synthesis with some being better at discriminating against
174 rNTP insertion than others (6,14). Therefore, rNMPs are the most common type of DNA lesion
175 found in the genomic DNA with critical effects on overall genome stability (27). Although

176 DNA polymerases have evolved selectivity against rNTPs favoring dNTP insertion and have
177 an intrinsic exonuclease proofreading function to remove rNTPs, this is not sufficient to
178 prevent large amounts of rNTPs remaining in the DNA backbone (6-26). The RER pathway,
179 involving RNase H2 function, can efficiently remove ribonucleotides from RNA:DNA hybrids
180 by cleaving at the 5'-end of the ribonucleotide, which is followed by strand displacement
181 synthesis by pol δ and flap removal by Flap Endonuclease 1 (FEN1) and/or Dna2, and finally
182 ligation step by LIG1 (56-58).
183 In addition to its role in the RER pathway, LIG1 is responsible for the ligation of over 10
184 million Okazaki fragments during each round of DNA replication and nick sealing of the repair
185 products at the final step of all DNA repair pathways which process DNA damage that could
186 be formed due to a variety of endogenous and exogenous sources (59-61). The nick repair
187 products are indeed potentially toxic intermediates, and this may lead to DNA strand breaks
188 that could trigger harmful nuclease activities, cytotoxicity, double strand breaks in replicating
189 DNA, and ultimately genomic instability (61-63). Therefore, DNA ligases could be challenged
190 by various genomic insults such as oxidative damage, non-canonical mismatches, and
191 ribonucleotides embedded into genomic DNA due to their incorporations by DNA polymerases
192 on the downstream steps of DNA repair and replication (35,36). DNA ligases have evolved a
193 tolerance for ribonucleotides at 3'-end while heavily discriminating against a ribonucleotide at
194 5'-end (54,59-63). On the other hand, the mechanism of sugar discrimination during nick
195 sealing reaction by DNA ligase at atomic resolution remains unknown. ATP-dependent DNA
196 ligases catalyze the ligation reaction including three sequential steps (64-69): (i) enzyme
197 adenylation at the active site lysine residue (*i.e,* K568) and formation of the ligase-AMP
198 intermediate, (ii) transfer of AMP group to 5'-P end of nick DNA resulting in the formation of
199 DNA-AMP intermediate, and finally (iii) phosphodiester bond formation between adjacent 3'-
200 OH and 5'-P DNA termini at a nick.

201 Our previous and present LIG1 structures revealed the features of DNA substrate and LIG1
202 interaction that dictate accurate *versus* mutagenic outcomes (53). Our structures contribute to
203 the understanding of the mechanism by which LIG1 discriminates against non-canonical DNA
204 intermediates during nick sealing at each individual step of the ligation reaction as well as how
205 the ligase active site engages with nick DNA distinctly depending on the architecture of
206 cognate (step 2), mismatched (step 1 or 2), or ribonucleotide (step 3)-containing ends
207 (Supplementary Table 1). Overall LIG1 structures in complex with nick DNA substrates have
208 provided considerable mechanistic details about steps 1 and 2 of the ligation reaction while
209 relatively less information has been reported for step 3 (53-55,70,71). In addition to our present
210 study of the LIG1/rG:C structure showing the human ligase active site at step 3 in the crystal
211 conditions lacking Mg^{2+} using EDTA pre-treated LIG1 protein, another step 3 structure has
212 been shown in ATP-dependent ligase from *Prochlorococcus marinus* (Pmar-Lig) (72). In this
213 structure, Pmar-Lig (step 2) crystals were immersed in cryoprotectant buffer containing $MgCl_2$
214 and then frozen immediately to get the step 3 Pmar-Lig/DNA structure, which is poised for
215 phosphodiester bond formation with partially occupied AMP in the active site. Since it is well
216 known that Mg^{2+} is required for step 3 of the ligation reaction, it is at first puzzling why our
217 LIG1/rG:C without Mg^{2+} shows step 3 (73). However, this is likely because of the 3'-endo
218 sugar pucker that ribonucleotides adopt. In B-form DNA, the sugar moieties usually adopt the
219 2'-endo sugar pucker conformation, while in A-form DNA or RNA duplexes the sugar moieties
220 usually adopt the 3'-endo sugar pucker conformation (74). The ligase active site forces the
221 DNA upstream of the nick to adopt an A-form conformation which causes the sugar at the 3'-
222 end to assume the 3'-endo conformation, both of which are necessary for ligation (54). However,
223 Mg^{2+} plays an important role in this step. In previous studies, it has been shown that the
224 catalytic Mg^{2+} can activate the 3'-OH for nucleophilic attack on the 5'-P (73). However, this
225 catalytic Mg^{2+} plays a dual role where the interaction between Mg^{2+} and the 3'-OH also

226 stabilizes the 3'-endo conformation of the 3'-end. When we compare the Mg^{2+}-containing
227 structure of LIG1 wild-type/3'-ddC:G (dideoxycytidine), which lacks a 3'-OH group, as well
228 as the structure of Pmar-Lig/3'-dC:G with non-active Mn^{2+}, the divalent metal atom is in almost
229 the same position, which demonstrates that the position of the divalent metal is conserved for
230 the ligases. If we imagine the position of the 3'-OH in the LIG1/3'-ddC:G structure, it is also in
231 the same position as Pmar-Lig/3'-dC:G; additionally, the divalent metal interacts with both 3'-
232 OH and 5'-P (Supplementary Figure 10). Since rG at the 3'-end adopts a stable 3'-endo
233 conformation due to its identity as a ribonucleotide, it does not require the rigidity enforced by
234 the catalytic Mg^{2+} to properly align the 3'-OH for catalysis. We assume that this is the reason
235 LIG1/rG:C shows step 3 of ligation reaction even without Mg^{2+}. Indeed, this is further
236 supported biochemically by the ligation of nick DNA substrate with 3'-rG:C in the absence of
237 Mg^{2+} (Supplementary Figure 7).
238 To understand the mechanism of sugar discrimination by human DNA ligases more
239 comprehensively, further structural and biochemical studies are required with other repair
240 ligases such as DNA ligase IIIα and DSB repair ligase IV (75-81). We previously reported that
241 the BER DNA ligases, LIG1 and DNA ligase IIIα/XRCC1 complex, are compromised by
242 subtle changes in all 12 possible mismatched base pairs at the 3'-end of the nick repair
243 intermediate (42). In the present study, our ligation assays with DNA ligase IIIα/XRCC1
244 complex also demonstrated mutagenic nick sealing of DNA substrate with 3'-rG:C as efficient
245 as a cognate 3'-dG:C ligation (Supplementary Figure 11), and we did not observe any
246 significant difference between the mutagenic ligation efficiencies of LIG1 and DNA ligase
247 IIIα/XRCC1 complex (Supplementary Figure 6). Although these findings indicate a lack of
248 sugar discrimination against a ribonucleotide by LIG1 for 3'-rG:C nick DNA only, further
249 structure/function characterizations of LIG1 against the nick DNA intermediates containing all
250 possible ribonucleotides mismatches at either 3'- or 5'-end at a nick are necessary to provide a

251 comprehensive insight into how LIG1 bestows substrate specificity during DNA replication
252 and repair, which is important to understand the disease mechanisms and to develop potential
253 therapeutic targets and strategies (82-84).

254 **Methods**

255 **Preparation of nick DNA substrates**

256 DNA substrates with a 6-carboxyfluorescein (FAM) label were prepared as described
257 previously (35-53). Nick DNA substrates containing 3'-preinserted dG:C, dG:T, 8oxodG:A,
258 and rG:C were used in the ligation assays including DNA ligase I (wild-type, and F872A,
259 EE/AA, and EE/AA/F872A mutants) or DNA ligase IIIα/XRCC1 complex (Supplementary
260 Table 2). Nick DNA substrates for LIG1 X-ray crystallography were prepared by annealing
261 upstream, downstream, and template primers (Supplementary Table 3).

262 **Protein purifications**

263 His-tag (pET-24b) human C-terminal (Δ261) wild-type and E346A/E592A (EE/AA) mutant
264 of LIG1 proteins were purified as described previously (35-53). Briefly, they were
265 overexpressed in Rosetta (DE3) *E. coli* cells in Terrific Broth (TB) media with kanamycin
266 (50 $\mu g ml^{-1}$) and chloramphenicol (34 $\mu g ml^{-1}$) at 37 °C. The cells were induced with 0.5 mM
267 isopropyl β-D-thiogalactoside (IPTG) when the OD_{600} was reached to 1.0, and the
268 overexpression was continued for overnight at 28 °C. Cells were lysed in the lysis buffer
269 [50 mM Tris-HCl (pH 7.0), 500 mM NaCl, 20 mM imidazole, 10% glycerol, 1 mM PMSF, an
270 EDTA-free protease inhibitor cocktail tablet] by sonication at 4 °C and the cell lysate was
271 pelleted at 31,000 x g for 1 h at 4 °C. Proteins were purified by a HisTrap HP column with an
272 increasing imidazole concentration (20-300 mM) after being equilibrated in the binding buffer
273 [50 mM Tris-HCl (pH 7.0), 500 mM NaCl, 20 mM imidazole, 10% glycerol] at 4 °C. The
274 collected fractions were subsequently loaded onto a HiTrap Heparin column that was
275 equilibrated with the binding buffer [50 mM Tris-HCl (pH 7.0), 50 mM NaCl, 1.0 mM EDTA,

276 and 10% glycerol], and then eluted with a linear gradient of NaCl up to 1 M. LIG1 proteins
277 were further purified by Superdex 200 10/300 columns in the buffer [20 mM Tris-HCl (pH
278 7.0), 200 mM NaCl, and 1 mM DTT]. The plasmid DNA constructs encoding human C-
279 terminal (Δ261) LIG1 genes carrying a single F872A and triple EE/AA/F872A mutations were
280 generated by side-directed mutagenesis using the plasmid DNA of LIG1 wild-type and
281 E346A/E592A (EE/AA) mutant, respectively, as previously described (35-53). His-tag LIG1
282 mutant proteins F872A and EE/AA/F872A were overexpressed and purified as described
283 above. Human his-tag DNA ligase IIIα/XRCC1 complex (pDUET1) was purified as previously
284 described (35-53). Briefly, the complex protein was overexpressed in BL21(DE3) *E. coli* cells
285 in LB media at 37 °C and the cell lysis was prepared as described above. The supernatant was
286 loaded onto a HisTrap HP column (GE Health Sciences) and purified with an increasing
287 imidazole gradient (20-300 mM) elution at 4 °C. The collected fractions were subsequently
288 loaded onto a HiTrap Heparin column that was equilibrated with the binding buffer [25 mM
289 Tris-HCl (pH 7.5), 50 mM NaCl, and 10% glycerol], and then eluted with a linear gradient of
290 NaCl up to 1 M. The protein complex was further purified by Superdex 200 Increase 10/300
291 chromatography (GE Healthcare) in the buffer [25 mM Tris-HCl (pH 7.0), 200 mM NaCl,
292 glycerol 5%, and 1 mM DTT]. All proteins purified in this study were dialyzed against storage
293 buffer including 25 mM Tris-HCl (pH 7.0), 200 mM NaCl, concentrated, frozen in liquid
294 nitrogen, and stored at -80 °C.

295 **Crystallization and structure determination**

296 LIG1 C-terminal (Δ261) EE/AA mutant that was pre-treated with EDTA before using in the
297 crystallization experiments in the absence of Mg^{2+} as described previously (53). All LIG1-nick
298 DNA complex crystals were grown at 20 °C using the hanging drop method. LIG1 (at 27 $mgml^{-}$
299 ^{1})/DNA complex solution was prepared in 20 mM Tris-HCl (pH 7.0), 200 mM NaCl, 1 mM
300 DTT, 1 mM EDTA and 1 mM ATP at 1:4:1 DNA:protein molar ratio and then mixed with 1

301 μl reservoir solution containing 100 mM MES (pH 6.1), 100 mM lithium acetate, and 14%
302 (w/v) polyethylene glycol PEG3350. All crystals grew in one day. They were then washed in
303 the reservoir solution with 20% glycerol and flash cooled in liquid nitrogen for data collection.
304 The crystals were maintained at 100° K during X-ray diffraction data collection using the
305 beamline 7B2 at Cornell High Energy Synchrotron Source (CHESS). The diffraction images
306 were indexed and integrated using HKL2000. All structures were solved by the molecular
307 replacement method using PHASER with PDB entry 7SUM as a search model (53,85,86).
308 Iterative rounds of model building were performed in COOT and the final models were refined
309 with PHENIX or REFMAC5 (86-89). 3DNA program was used for sugar pucker analysis (90).
310 All structural images were drawn using PyMOL (The PyMOL Molecular Graphics System,
311 V0.99, Schrödinger, LLC). Detailed crystallographic statistics are provided in Table 1.

312 **DNA ligation assays**

313 The nick sealing efficiency of DNA ligase (DNA ligase I or DNA ligase IIIα/XRCC1 complex)
314 was investigated to determine the sugar discrimination mechanism against 3'-preinserted rG
315 opposite template base C in the ligation assays (Supplementary Scheme 1) as described
316 previously (35-53). Briefly, the reaction mixture containing 50 mM Tris-HCl (pH 7.5), 100
317 mM KCl, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 100 μgml^{-1} BSA, 10% glycerol, and 500
318 nM nick DNA substrate was initiated by the addition of 100 nM DNA ligase, incubated at 37
319 °C, and stopped at the time points indicated in the figure legends. The reaction products were
320 then quenched with an equal amount of gel loading buffer [95% formamide, 20 mM
321 ethylenediaminetetraacetic acid, 0.02% bromophenol blue and 0.02% xylene cyanol]. After
322 incubation at 95 °C for 3 min, the reaction products were separated by electrophoresis on an
323 18% denaturing polyacrylamide gel. The gels were scanned with a Typhoon Phosphor Imager
324 (Amersham Typhoon RGB), and the data were analyzed using ImageQuant software. The
325 ligation assays were performed similarly with nick DNA substrates containing 3'-dG:C, 3'-

326 8oxodG:A, or 3'-dG:T. In addition, nick sealing of DNA with 3'-rG:C by EE/AA mutant of
327 LIG1 in the absence of Mg^{2+} was performed as described above in the reaction buffer lacking
328 10 mM $MgCl_2$.

329 **Data availability**

330 Atomic coordinates and structure factors for the reported crystal structure of LIG1 in complex
331 with nick DNA/3'-riboG:C has been deposited in the RCSB Protein Data Bank under accession
332 number 8DVT. All relevant data are available from the authors upon reasonable request.

333 **References**

334 1. Klein, H. L. Genome instabilities arising from ribonucleotides in DNA. *DNA Repair*
335 **56**, 26-32 (2017).
336 2. Potenski, C. J. & Klein, H. L. How the misincorporation of ribonucleotides into
337 genomic DNA can be both harmful and helpful to cells. *Nucleic Acids Res*. **42**, 10226-
338 10234 (2014).
339 3. Cerritelli, S. M. & Crouch, R. J. The balancing act of ribonucleotides in DNA. *Trends*
340 *Biochem. Sci.* **41**, 434-445 (2016).
341 4. Caldecott, K. W. Ribose-an internal threat to DNA. *Science* **343**, 260-261 (2014).
342 5. Yao, N. Y., Schroeder, J. W., Yurieva, O., Simmons, L. A. & O'Donnell, M. E. Cost
343 of rNTP/dNTP pool imbalance at the replication fork. *Proc. Natl. Acad. Sci*. **110**,
344 12942-12947 (2013).
345 6. Williams, J. S., Lujan, S. A. & Kunkel, T. A. Processing ribonucleotides incorporated
346 during eukaryotic DNA replication. *Nat. Rev. Mol. Cell Biol.* **17**, 350-363 (2016).
347 7. Clausen, A. R., Murray, M. S., Passer, A. R., Pedersen, L. C., Kunkel, T. A., Clausen,
348 A. R., Murray, M. S., Passer, A. R., Pedersen, L. C. & Kunkel, T. A. Structure-function
349 analysis of ribonucleotide bypass by B family DNA replicases. *Proc. Natl. Acad. Sci.*
350 *U.S.A.,* **110**, 16802-16807 (2013).

351 8. Clausen, A. R., Zhang, S., Burgers, P. N., Lee, M. Y. & Kunkel, T. A. Ribonucleotide
352 incorporation, proofreading and bypass by human DNA polymerase delta. *DNA Repair*
353 **12**, 121-127 (2013).
354 9. Nick McElhinny, S. A., Watts, B. E., Kumar, D., Watt, D. L., Lundstrom, E. B.,
355 Burgers, P. M., Johansson, E., Chabes, A. & Kunkel, T. A. Abundant ribonucleotide
356 incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. U. S.*
357 *A*. **107**, 4949-4954 (2010).
358 10. Gosavi, R. A., Moon, A. F., Kunkel, T. A., Pedersen, L. C. & Bebenek, K. The catalytic
359 cycle for ribonucleotide incorporation by human DNA pol lambda. *Nucleic Acids Res.*
360 **40**, 7518-7527 (2012).
361 11. Goksenin, A. Y. *et al*. Human DNA polymerase epsilon is able to efficiently extend
362 from multiple consecutive ribonucleotides. *J. Biol. Chem.* **287**, 42675-42684 (2012).
363 12. Watt, D. L., Johansson, E., Burgers, P. M., Kunkel, T. A., Watt, D. L., Johansson, E.,
364 Burgers, P. M. & Kunkel, T. A. Replication of ribonucleotide-containing DNA
365 templates by yeast replicative polymerases. *DNA Repair*, **10**, 897-902 (2011).
366 13. Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B. & Kunkel, T. A.
367 Ribonucleotides are signals for mismatch repair of leading-strand replication errors.
368 *Mol. Cell* **50**, 437-443 (2013).
369 14. Nick McElhinny, S. A., Kumar, D., Clark, A. B., Watt, D. L., Watts, B. E., Lundstrom,
370 E. B., Johansson, E., Chabes, A. & Kunkel, T. A. Genome instability due to
371 ribonucleotide incorporation into DNA. *Nat. Chem. Biol*. **6**, 774-781 (2010).
372 15. Sassa, A., Yasui, M. & Honma, M. Current perspectives on mechanisms of
373 ribonucleotide incorporation and processing in mammalian DNA. *Genes Environ*. **41**,
374 3 (2019).

375 16. DeRose, E. F., Perera, L., Murray, M. S., Kunkel, T. A. & London, R. E. Solution
376 structure of the Dickerson DNA dodecamer containing a single ribonucleotide.
377 *Biochemistry* **51**, 2407-2416 (2012).
378 17. Jaishree, T. N., van der Marel, G. A., van Boom, J. H. & Wang, A. H Structural
379 influence of RNA incorporation in DNA: quantitative nuclear magnetic resonance
380 refinement of d(CG)r(CG)d(CG) and d(CG)r(C)d(TAGCG). *Biochemistry* **32**, 4903-
381 4911 (1993).
382 18. Hovatter, K. R. & Martinson, H. G. Ribonucleotide-induced helical alteration in DNA
383 prevents nucleosome formation. *Proc. Natl. Acad. Sci. U. S. A*. **84**, 1162-1166 (1987).
384 19. Conover, H. N., Lujan, S. A., Chapman, M. J., Cornelio, D. A., Sharif, R., Williams, J.
385 S., Clark, A. B., Camilo, F., Kunkel, T. A. & Argueso, J. L. Stimulation of
386 chromosomal re- arrangements by ribonucleotides. *Genetics* **201**, 951-961 (2015).
387 20. Huang, S. N., Williams, J. S., Arana, M. E., Kunkel, T. A. & Pommier, Y.
388 Topoisomerase I-mediated cleavage at unrepaired ribonucleotides generates DNA
389 double-strand breaks. *EMBO J*. **36**, 361-373 (2016).
390 21. Joyce, C. M. Choosing the right sugar: how polymerases select a nucleotide substrate.
391 *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1619-1622 (1997).
392 22. Brown, J. A., Fiala, K. A., Fowler, J. D., Sherrer, S. M., Newmister, S. A., Duym, W.
393 W. & Suo, Z. A novel mechanism of sugar selection utilized by a human X-family
394 DNA polymerase. *J. Mol. Biol*. **395**, 282-290 (2010).
395 23. Cavanaugh, N. A., Beard, W. A. & Wilson, S. H. DNA polymerase β ribonucleotide
396 discrimination. *J. Biol. Chem*. **285**, 24457-24465 (2010).
397 24. Donigan, K. A., McLenigan, M. P., Yang, W., Goodman, M. F. & Woodgate, R. The
398 steric gate of DNA polymerase iota regulates ribonucleotide incorporation and
399 deoxyribonucleotide fidelity. *J. Biol. Chem*. **289**, 9136-9145 (2014).

400 25. Moon, A. F., Pryor, J. M., Ramsden, D. A., Kunkel, T. A., Bebenek, K. & Pedersen, L.
401 C. Structural accommodation of ribonucleotide incorporation by the DNA repair
402 enzyme polymerase μ. *Nucleic Acids Res*. **45**, 9138-9148 (2017).
403 26. Kasiviswanathan, R. & Copeland, W. C. Ribonucleotide discrimination and reverse
404 transcription by the human mitochondrial DNA polymerase. *J. Biol. Chem*. **286**, 31490-
405 31500 (2011).
406 27. Williams, J. S. & Kunkel, T. A. Ribonucleotides in DNA: Origins, repair and
407 consequences. *DNA Repair* **19**, 27-37 (2014).
408 28. Brown, J. A. & Suo, Z. Unlocking the sugar steric gate of DNA polymerases.
409 *Biochemistry* **50**, 1135-1142 (2011).
410 29. DeLucia, A. M., Grindley, N. D. & Joyce, C. M. An error-prone family Y DNA
411 polymerase (DinB homolog from Sulfolobus solfataricus) uses a 'steric gate' residue
412 for dis- crimination against ribonucleotides. *Nucleic Acids Res*. **31**, 4129-4137 (2003).
413 30. Brown, J. A. & Suo, Z. Unlocking the sugar 'steric gate' of DNA polymerases.
414 *Biochemistry* **50**, 1135-1142 (2011).
415 31. Ruiz, J. F., Juarez, R., Garcia-Diaz, M., Terrados, G., Picher, A. J., Gonzalez-Barrera,
416 S., Fernandez de Hanestrosa, A. R. & Blanco, L. Lack of sugar discrimination by
417 human pol μ requires a single glycine residue. *Nucleic Acids Res*. **31**, 4441-4449 (2014).
418 32. Reijns, M. A. M. *et al.* Enzymatic removal of ribonucleotides from DNA is essential
419 for mammalian genome integrity and development. *Cell* **149**, 23 (2012).
420 33. Vaisman, A. & Woodgate, R. Ribonucleotide discrimination by translesion synthesis
421 DNA polymerases. *Crit. Rev. Biochem. Mol. Biol*. **53**, 382-402 (2018).
422 34. Bonnin, A., Lazaro, J. M.M., Blanco, L. & Salas, M. A single tyrosine prevents
423 insertion of ribonucleotides in the eukaryotic-type φ29 DNA polymerase. *J. Mol. Biol*.
424 **290**, 241-251 (1999).

425 35. Çağlayan, M. Interplay between DNA polymerases and DNA ligases: Influence on
426 substrate channeling and the fidelity of DNA ligation. *J. Mol. Biol*. **431**, 2068-2081
427 (2019).
428 36. Çağlayan, M. & Wilson, S. H. Oxidant and environmental toxicant-induced effects
429 compromise DNA ligation during base excision DNA repair. *DNA Repair* **35**, 85-89
430 (2015).
431 37. Çağlayan, M., Batra, V. K., Sassa, A., Prasad, R. & Wilson, S. H. Role of polymerase
432 β in complementing aprataxin deficiency during abasic-site base excision repair. *Nat.*
433 *Struct. Mol. Biol*. **21**, 497-499 (2014).
434 38. Çağlayan, M. *et al.* Oxidized nucleotide insertion by pol β confounds ligation during
435 base excision repair. *Nat. Commun*. **8**, 14045 (2017).
436 39. Çağlayan, M. *et al.* Complementation of aprataxin deficiency by base excision repair
437 enzymes. *Nucleic Acids Res*. **43**, 2271-2281 (2015).
438 40. Çağlayan, M. *et al.* Complementation of aprataxin deficiency by base excision repair
439 enzymes in mitochondrial extracts. *Nucleic Acids Res.* **45**, 10079-10088 (2017).
440 41. Çağlayan, M. & Wilson, S. H. Role of DNA polymerase β oxidized nucleotide insertion
441 in DNA ligation failure. *J. Radiat. Res*. **58**, 603-607 (2017).
442 42. Çağlayan, M. The ligation of pol β mismatch insertion products governs the formation
443 of promutagenic base excision DNA repair intermediates. *Nucleic Acids Res*. **48**, 3708-
444 3721 (2020).
445 43. Çağlayan, M. Pol β gap filling, DNA ligation and substrate-product channeling during
446 base excision repair opposite oxidized 5-methylcytosine modifications. *DNA Repair*
447 **95**,102945 (2020).

448 44. Tang, Q., Kamble, P. & Çağlayan, M. DNA ligase I variants fail in the ligation of
449 mutagenic repair intermediates with mismatches and oxidative DNA damage.
450 *Mutagenesis* **35**, 391-404 (2020).
451 45. Kamble, P., Hall, K., Chandak, M., Tang, Q. & Çağlayan, M. DNA ligase I fidelity the
452 mutagenic ligation of pol β oxidized and mismatch nucleotide insertion products in base
453 excision repair. *J. Biol. Chem*. **296**, 100427 (2021).
454 46. Tang, Q. & Çağlayan, M. The scaffold protein XRCC1 stabilizes the formation of
455 polβ/gap DNA and ligase IIIα/nick DNA complexes in base excision repair. *J. Biol.*
456 *Chem*. **297**, 101025 (2021).
457 47. Horton, J. K., Stefanick, D. F., Çağlayan, M., Zhao, M. L., Gassman, N. R. & Wilson,
458 S.H. XRCC1 phosphorylation affects aprataxin recruitment and DNA deadenylation
459 activity. *DNA Repair* **64**, 26-33 (2018).
460 48. Prasad, R., Çağlayan, M., Da-Peng, D., Nadalutti, C. A., Gassman, N. R., Zhao, M.,
461 Stefanick, D. F., Horton, J. K., Krasich, R., Longley, M. J., Copeland, W. C., Griffith,
462 J. D. & Wilson, S. H. DNA polymerase β: The missing link of the base excision repair
463 machinery in mammalian mitochondria. *DNA Repair* **60**, 77-88 (2017).
464 49. Sassa, A., Çağlayan, M., Rodriguez, Y., Beard, W. A., Wilson, S. H., Nohmi, T.,
465 Honma, M. & Yasui, M. Impact of ribonucleotide backbone on translesion synthesis
466 and repair of 7,8-Dihydro-8-oxoguanine. *J. Biol. Chem*. **291**, 24314-24323 (2016).
467 50. Sassa, A., Çağlayan, M., Dyrkheeva, N. S., Beard, W. A. & Wilson, S. H. Base excision
468 repair of tandem modifications in a methylated CpG dinucleotide. *J. Biol. Chem*. 289,
469 13996-4008 (2014).
470 51. Çağlayan, M. & Wilson, S. H. Pol μ dGTP mismatch insertion opposite T coupled with
471 ligation reveals a promutagenic DNA intermediate during double strand break repair.
472 *Nat. Commun*. **9**, 4213 (2018).

473 52. Çağlayan, M. Pol μ ribonucleotide insertion opposite 8-oxodG facilitates the ligation
474 of premutagenic DNA repair intermediate. *Sci. Rep*. **10**, 940 (2020).
475 53. Tang Q., Gulkis, M., McKenna R. & Çağlayan, M. Structures of LIG1 that engage with
476 mutagenic mismatches inserted by polβ in base excision repair. *Nat. Commun*. **13**, 3860
477 (2022).
478 54. Pascal, J. M. *et al.* Human DNA ligase I completely encircles and partially unwinds
479 nicked DNA. *Nature* **432**, 473-478 (2004).
480 55. Tumbale, P. P. *et al.* Two-tiered enforcement of high-fidelity DNA ligation. *Nat.*
481 *Commun.* **10**, 5431 (2019).
482 56. Sparks, J. L. *et al.* RNase H2-initiated ribonucleotide excision repair. *Mol. Cell* **47**, 980-
483 986 (2012).
484 57. Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: the enzymes in eukaryotes. *FEBS J.*
485 **276**, 1494-1505 (2009).
486 58. Hiller, B., Achleitner, M., Glage, S., Naumann, R., Behrendt, R. & Roers, A.
487 Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome
488 integrity. *J. Exp. Med*. **209**, 1419-1426 (2012).
489 59. Timson, D. J., Singleton, M. R. & Wigley, D. B. DNA ligases in the repair and
490 replication of DNA. *Mutat. Res*. **460**, 301-318 (2000).
491 60. Shuman, S. DNA ligases: progress and prospects. *J. Biol. Chem*. **284**, 17365-17369
492 (2009).
493 61. Tomkinson, A. E., Vijayakumar, S., Pascal, J. M. & Ellenberger, T. DNA ligases:
494 structure, reaction mechanism, and function. *Chem. Rev*. **106**, 687-699 (2006).
495 62. Doherty, A. J. & Suh, S. W. Structural and mechanistic conservation in DNA ligases.
496 *Nuc. Acids Res*. **28**, 4051-4058 (2000).

497 63. Ellenberger, T. & Tomkinson, A. E. Eukaryotic DNA ligases: Structural and functional
498 insights. *Annu. Rev. Biochem*. **77**, 313-338 (2008).
499 64. Unciuleac, M., Goldgur, Y. & Shuman, S. Two-metal versus one-metal mechanisms of
500 lysine adenylylation by ATP-dependent and NAR^+-dependent polynucleotide ligases.
501 *Proc. Natl. Acad Sci. U S A* **114**, 2592-2597 (2017).
502 65. Tomkinson, A. E., Tappe, N. J. & Friedberg, E. C. DNA Ligase I from *Saccharomyces*
503 *cerevisiae*: Physical and biochemical characterization of the CDC9 gene product.
504 *Biochemistry* **31**, 11762-11771(1992).
505 66. Nishida, H., Kiyonari, S., Ishino, Y. & Morikawa, K. The closed structure of an
506 archaeal DNA ligase from *Pyrococcus furiosus*. *J. Mol. Biol*. **360**, 956-967 (2006).
507 67. Chen, Y. *et al.* Structure of the error-prone DNA ligase of African swine fever virus
508 identifies critical active site residues. *Nat. Commun*. **10**, 387 (2019).
509 68. Unciuleac, M., Goldgur, Y. & Shuman, S. Structures of ATP-bound DNA ligase D in
510 a closed domain conformation reveal a network of amino acid and metal contacts to the
511 ATP phosphates. *J. Biol. Chem.* **294**, 5094-5104 (2019).
512 69. Odell, M., Sriskanda, V., Shuman, S. & Nikolov, D. Crystalstructure of eukaryotic
513 DNA ligase–adenylate illuminates the mechanism of nick sensing and strand joining.
514 *Mol. Cell* **6**, 1183-1193 (2000).
515 70. Williams, J. S. *et al.* High-fidelity DNA ligation enforces accurate Okazaki fragment
516 maturation during DNA replication. *Nat. Commun*. **12**, 482 (2021).
517 71. Jurkiw, T. J. *et al.* LIG1 syndrome mutations remodel a cooperative network of ligand
518 binding interactions to compromise ligation efficiency. *Nucleic Acids Res*. **49**, 1619-
519 1630 (2021).

520 72. Williamson, A. & Leiros, H. K. S. Structural intermediates of a DNA-ligase complex
521 illuminate the role of the catalytic metal ion and mechanism of phosphodiester bond
522 formation. *Nucleic Acids Res*. **47**, 7147-7162 (2019).
523 73. Unciuleac, M., Goldgur, Y. & Shuman, S. Two-metal versus one-metal mechanisms of
524 lysine adenylylation by ATP-dependent and NAR^+-dependent polynucleotide ligases.
525 *Proc. Natl. Acad Sci. U S A* **114**, 2592-2597 (2017).
526 74. Dejaegere, A. P. & Case, D. A. Density functional study of ribose and deoxyribose
527 chemical shifts. *J. Phys. Chem*. **102**, 5280-5289 (1998).
528 75. Cotner-Gohara, E., Kim, I., Hammel, M., Tainer, J. A., Tomkinson, A. E. &
529 Ellenberger, T. Human DNA ligase III recognizes DNA ends by dynamic switching
530 between two DNA-bound states. *Biochemistry* **49**, 6165-6176 (2010) .
531 76. Kukshal, V., Kim, I., Hura, G. L., Tomkinson, A. E., Trainer, J. A. & Ellenberger, T.
532 Human DNA ligase III bridges two DNA ends to promote specific intermolecular DNA
533 end joining. *Nucleic Acids Res*. **43**, 7021-7031 (2015) .
534 77. Cuneo, M. J., Gabel, S. A., Krahn, J. A., Ricker, M. A. & London, R. E. The structural
535 basis for partitioning of the XRCC1/DNA ligase IIIα BRCT-mediated dimer
536 complexes. *Nucleic Acids Res*. **39**, 7816-7827 (2011).
537 78. Conlin, M. P., Reid, D. A., Small, G. W., Chang, H. H., Watanabe, G., Lieber, M. R.,
538 Ramsden, D. A. & Rothenberg, E. DNA Ligase IV guides end-processing choice during
539 nonhomologous end joining. *Cell Rep*., **20**, 2810-2819 (2017).
540 79. Ochi, T., Gu, X. & Blundell, T. L. Structure of the catalytic region of DNA ligase IV
541 in complex with an Artemis fragment sheds light on double-strand break repair.
542 *Structure*, **21**, 672-679 (2013).

543 80. Ochi, T., Wu, Q., Chirgadze, D. Y., Grossmann, J. G., Bolanos-Garcia, V. M. &
544 Blundell, T. L. Structural insights into the role of domain flexibility in human DNA
545 ligase IV. *Structure*, **20**, 1212-1222 (2012).
546 81. Kaminski, A., Tumbale, P. P., Schellenberg, M. J., Williams, R. S., Williams, J. G.,
547 Kunkel, T. A., Pedersen, L. C. & Bebenek, K. Structures of DNA-bound human ligase
548 IV catalytic core reveal insights into substrate binding and catalysis. *Nat. Commun*., **9**,
549 2642 (2018).
550 82. Tomkinson, A. E., Howes, T. R. & Wiest, N. E. DNA ligases as therapeutic targets.
551 *Transl. Cancer Res.* **2**, 31 (2013).
552 83. Zhong, S. *et al.* Identification and validation of human DNA ligase inhibitors using
553 computer-aided drug design. *J. Med. Chem*. **51**, 4553-4562 (2008).
554 84. Chen, X. *et al.* Rational design of human DNA ligase inhibitors that target cellular DNA
555 replication and repair. *Cancer Res*. **68**, 3169-3177 (2008).
556 85. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in
557 oscillation mode. *Methods Enzymol*. **276**, 307-326 (1997).
558 86. McCoy, A. J. *et al*. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674
559 (2007).
560 87. Emsley, P. *et al.* Features and development of Coot. *Acta Crystallogr. D. Biol.*
561 *Crystallogr.* **66**, 486-501 (2010).
562 88. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for
563 macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213-221
564 (2010).
565 89. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal
566 structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355-367 (2011).

567 90. Li, S., Olson, W. K. & Lu, X. J. Web 3DNA 2.0 for the analysis, visualization, and
568 modeling of 3D nucleic acid structures. *Nucleic Acids Res*. **47**, W26-W34 (2019).

569 **Funding**

570 This work was supported by a grant 1R35GM147111-01 from the National Institute of General
571 Medical Sciences (NIGMS).

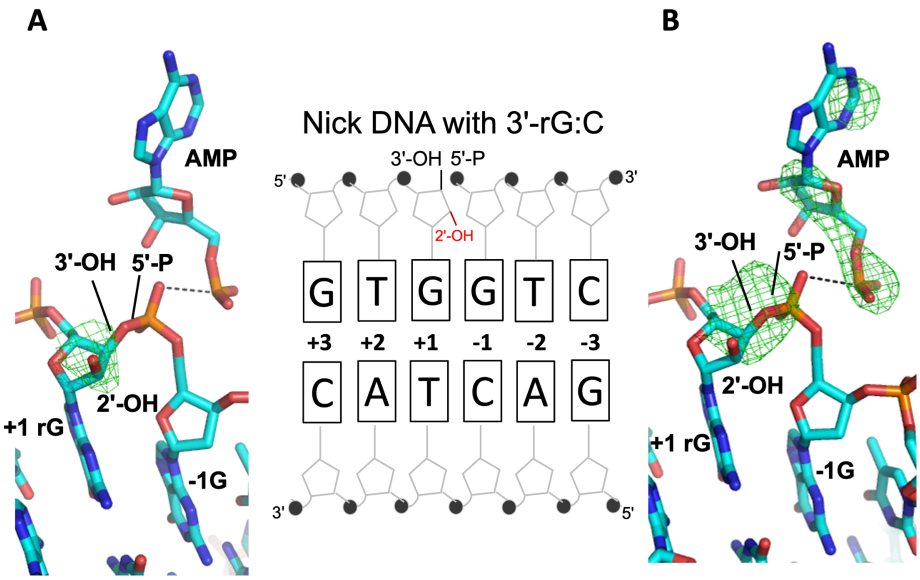
572 **Acknowledgements**

573 This work is based upon research conducted at the Center for High Energy X-ray Sciences
574 (CHEXS), which is supported by the National Science Foundation under award DMR-
575 1829070, and the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is
576 supported by award 1-P30-GM124166-01A1 from the National Institute of General Medical
577 Sciences, National Institutes of Health, and by New York State's Empire State Development
578 Corporation (NYSTAR). The authors thank Dr. Robert McKenna and Jacob Combs
579 (University of Florida) for beamline 7B2 at CHESS, and an assistance with crystal shipment
580 and data collection.

581 **Author contributions**

582 Conceptualization M.Ç., methodology and investigation T.Q., M.G., M.Ç.; writing-original
583 draft T.Q., M.G., M.Ç.; writing-reviewing and editing T.Q., M.G., M.Ç.; funding acquisition
584 M.Ç.

585



586 **Figure 1. Structure of LIG1 bound to nick DNA duplex containing 3'-rG:C. A,** X-ray
587 crystal structure of LIG1/nick DNA duplex containing a ribonucleotide (rG) at 3'-terminus of
588 the nick DNA. Simulated annealing omit map (Fo-Fc) of LIG1/rG:C (cyan) is displayed for 2'-
589 OH of 3'-terminus contoured at 3σ. Schematic view of DNA including DNA:RNA hybrid used
590 in the LIG1/nick DNA duplex crystallization. **B,** Simulated annealing omit map (Fo-Fc) of
591 LIG1/rG:C is displayed for 3'-OH terminus and AMP contoured at 3σ. DNA and AMP are
592 shown as sticks and the map is depicted in green.

593

594

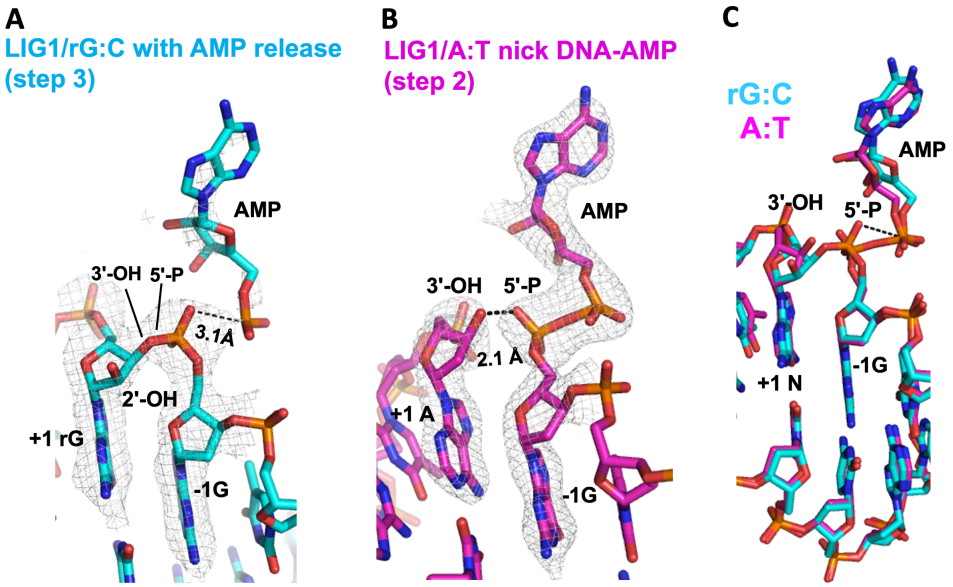


Figure 2. LIG1/rG:C structure demonstrates a phosphodiester bond formation and AMP release during the final ligation step at atomic resolution. A-B, The 2Fo - Fc density map (1σ) is displayed for 3'- and 5'-terminus of the nick in the LIG1/rG:C (cyan) and LIG1/A:T (magenta) structures. AMP is incomplete and shows lower density in the LIG1/rG:C (step 3) than the LIG1/A:T (step 2). DNA and AMP are shown as sticks and the map is depicted in grey. **C**, Overlay of LIG1 structures bound to the nick DNA duplexes containing rG:C (cyan) and A:T (PDB: 7SUM, magenta) ends shows the difference at the positions of 3'- and 5'-terminus of the nick.

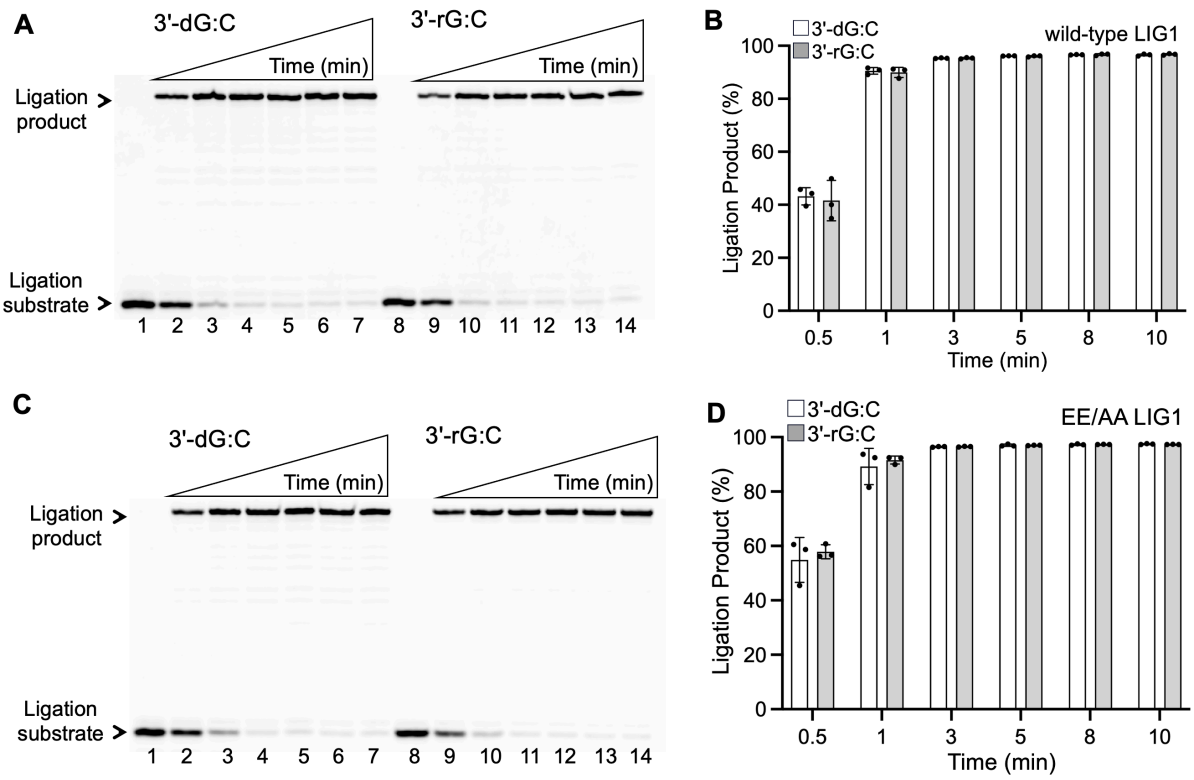
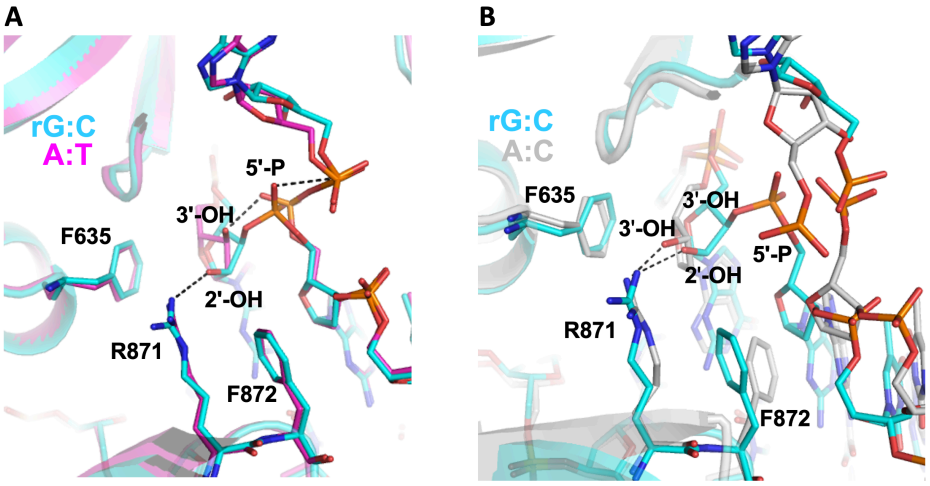
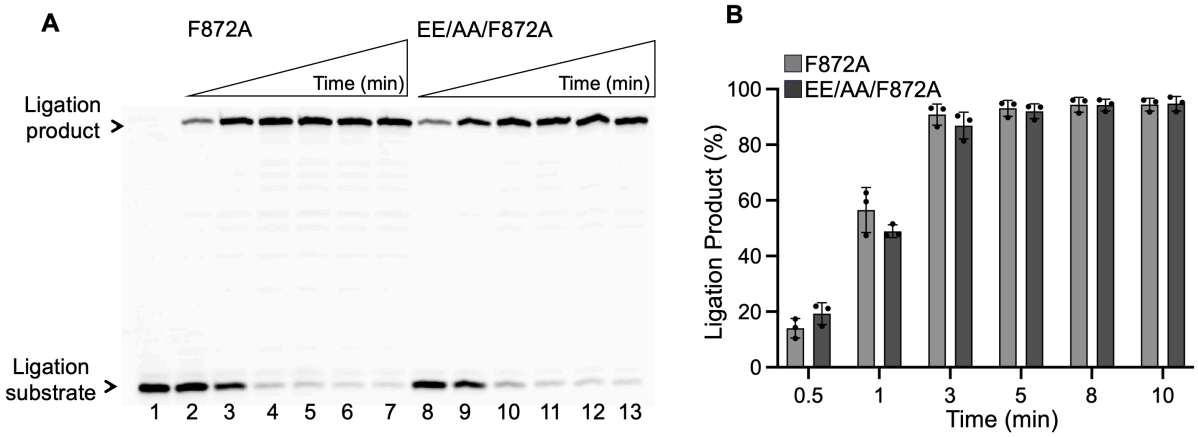


Figure 3. Mutagenic ligation of the nick DNA with 3'-ribonucleotide by LIG1. A-C, Lanes 1 and 8 are the negative enzyme controls of the nick DNA substrates with 3'-dG:C and 3'-rG:C, respectively. Lanes 2-7 and 9-14 are the ligation products in the presence of 3'-dG:C and 3'-rG:C, respectively, by wild-type (A) and EE/AA mutant (C) of LIG1, and correspond to time points of 0.5, 1, 3, 5, 8, and 10 min. **B-D,** Graphs show time-dependent change in the amount of ligation products. The data represent the average from three independent experiments ± SD.



613

614 **Figure 4. Comparison of LIG1 active site residues between the structures of LIG1/nick**
615 **DNA complexes with rG:C versus cognate and mismatched ends.** Overlay of LIG1
616 structures bound to the nick DNA duplexes containing rG:C (cyan), A:T (PDB: 7SUM,
617 magenta), and A:C (PDB: 7SX5, grey) ends shows the differences in the positions of 3'- and
618 5'-terminus at the nick DNA and no difference in the LIG1 active site F635, R871, and F872
619 residues around 2'-OH of the ribose.



620

621 **Figure 5. Role of LIG1 active side residue F872 for mutagenic ligation of the nick DNA**
622 **with 3'-rG:C. A,** Lane 1 is the negative enzyme control of the nick DNA substrate with 3'-
623 rG:C. Lanes 2-7 and 8-13 are the ligation reaction products by LIG1 F872 and EE/AA/F872A
624 mutants, respectively, and correspond to time points of 0.5, 1, 3, 5, 8, and 10 min. **B,** Graph
625 shows time-dependent change in the amount of ligation products. The data represent the
626 average from three independent experiments ± SD.

627 **Table 1. X-ray data collection and refinement statistics of LIG1/nick DNA duplex with a**

628 **ribonucleotide (rG:C) at the 3'-end.**

629

630

	$LIG1^{EE/AA}$ Nick DNA with 3'-rG:C
PDB entry ID	
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	65.065 115.633 126.338
a, b, g (°)	90
Resolution (Å)	20-3.0 (3.05-3.0)
R_{sym}	0.181 (1.451)
$I / \sigma (I)$	17.7 (1.7)
$CC_{1/2}$	1.000 (0.655)
CC^*	1.000 (0.890)
Completeness (%)	96.3 (100)
Redundancy	12.5 (13.6)
Refinement	
Resolution (Å)	20-3.0
No. reflections	17321
R_{work}/R_{free}	0.172/0.206
Non-H atoms:	
Protein/DNA	5665
AMP	21
Average B-factors ($Å^2$):	60.1
R.M.S.D	
Bond lengths (Å)	0.009
Bond angles (°)	1.174