Polγ coordinates DNA synthesis and proofreading to ensure mitochondrial genome integrity

Joon Park[1,2], **Geoffrey K. Herrmann**[1,2], **Patrick G. Mitchell**[3], **Michael B. Sherman**[1,2], **Y. Whitney Yin**[1,2,4,✉]

[1]Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX, USA

[2]Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, TX, USA

[3]Division of CryoEM and Bioimaging, Stanford Synchrotron Radiation Lightsource, Stanford Linear Accelerator Center National Accelerator Laboratory, Stanford University, Menlo Park, CA, USA

[4]Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, TX, USA

Abstract

Accurate replication of mitochondrial DNA (mtDNA) by DNA polymerase γ (Polγ) is essential for maintaining cellular energy supplies, metabolism, and cell cycle control. To illustrate the structural mechanism for Polγ coordinating polymerase (pol) and exonuclease (exo) activities to ensure rapid and accurate DNA synthesis, we determined four cryo-EM structures of Polγ captured after accurate or erroneous incorporation to a resolution of 2.4–3.0 Å. The structures show that Polγ employs a dual-checkpoint mechanism to sense nucleotide misincorporation and initiate proofreading. The transition from replication to error editing is accompanied by increased dynamics in both DNA and enzyme, in which the polymerase relaxes its processivity and the primer–template DNA unwinds, rotates, and backtracks to shuttle the mismatch-containing primer

terminus 32 Å to the exo site for editing. Our structural and functional studies also provide a foundation for analyses of Polγ mutation-induced human diseases and aging.

---

mtDNA integrity is critical for cellular energy supply, metabolism, and cell cycle control. Reduced mtDNA-replication fidelity, arising from genetic or hereditary factors, oxidative stress, or mtDNA-replication errors, results in mutations that have been associated with multiple systemic human diseases and aging–. Human mtDNA replication is performed by the high-fidelity DNA Polγ primarily by an asymmetrical strand-displacement mechanism in which the two strands are synthesized asynchronously[1,2], distinct from the canonical leading and lagging strand coupled DNA synthesis in the nucleus. The trimeric Polγ holoenzyme consists of a catalytic subunit, PolγA, which contains 5′–3′ polymerase (pol) and 3′–5′ exonuclease (exo) active sites for DNA synthesis and proofreading, respectively, and a dimeric accessory subunit, PolγB. Although PolγA belongs to the Pol I polymerase family, which typically has low processivity, upon association with the subunit PolγB, the holoenzyme gains sufficient processivity for a replicase[3,4]. In the holoenzyme, PolγA forms a constant interaction with the proximal PolγB monomer but only interacts with the distal PolγB monomer in replication mode for processivity enhancement[5–7]. PolγB, albeit having no intrinsic enzymatic activity, regulates both pol and exo activities of the holoenzyme[8,9].

The ratio of pol/exo activities has been used to define polymerases' mutator and antimutator phenotypes[10–13]. Imbalanced pol/exo activity in Polγ not only directly affects replication fidelity but is also implicated in human diseases and chronological aging. Patients carrying a PolγA mutation that disrupts the PolγA and distal PolγB interface exhibited diminished DNA synthesis and elevated exo activity, clinically manifesting as neurological disorders[14,15]. Mice carrying an exonuclease-deficient (exo⁻) Polγ variant displayed increased mtDNA mutations and premature aging syndromes[16]. Similar observations were made in *Drosophila* carrying Polγ exo⁻, but knocking in polymerase-deficient Polγ reversed the aging characteristics[17], establishing a direct connection between Polγ proofreading deficiency and aging. While extensive biochemical and structural studies have been conducted on pol activity, structural studies on proofreading and communication between the pol and exo sites are lacking.

Here we report structures of human Polγ captured in the reaction pathway from replication to proofreading using both correctly matched and mismatched primer–template (p/t) DNA. While Polγ is predominantly in a single conformation during accurate DNA replication, it exhibits high structural heterogeneity after a misincorporation. We identified three Polγ conformations during proofreading, representing mutagenetic synthesis, mismatch correction, and a transition state between replication and editing modes. These structures provide unprecedented, comprehensive insights into Polγ's maintenance of mitochondrial genome integrity.

Results

Substrates for Polγ replication and proofreading complexes

Among DNA mismatches, the G•T mismatch is the most frequent error made by many high-fidelity polymerases[18–21] and is also more likely to escape proofreading in Polγ than other mismatches[22], probably because the G•T mismatch is able to form an enolic Watson–Crick (W-C)-like base pair in addition to a wobble base pair[23]. To gain insights into how Polγ distinguishes between correct and incorrect incorporation and initiates proofreading, we designed two identical 24–28-nucleotide (nt) p/t DNA duplexes with the exception that the 3′-terminal nucleoside forms either a G•C W-C base pair or a G•T mismatch with the template (Fig. 1a). The Polγ ternary replication complex was formed with the G•C p/t substrate and the Polγ proofreading ternary complex was formed with the G•T p/t substrate, both with a correct incoming nucleotide, dCTP. An inhibitory Ca^{2+} ion was included to halt pol and exo reactions. Structural studies of the G•T mismatch could elucidate not only a general proofreading mechanism but also the mutagenic nature of the G•T mismatch in Polγ.

Structure of an accurately replicating Polγ ternary complex

The cryo-EM structure of the Polγ replication ternary complex with G•C p/t was reconstructed and refined to 2.46-Å resolution by estimation with a Fourier shell correlation (FSC) threshold at 0.143 between independently processed half-datasets (Table 1 and Extended Data Fig. 1)[24]. The dataset consists of a primarily homogeneous single conformation, and the cryo-EM structure is nearly identical to that determined by crystallography[6], except that the accessory subunit PolγB is wild type in the cryo-EM structure and a deletion mutant (PolγB- I4) in the crystal structure.

The trimeric holoenzyme consists of a catalytic subunit, PolγA, which forms an interface with both proximal and distal monomers of the accessory subunit PolγB (Fig. 1b). PolγA contains pol, exo, and spacer domains, and the pol domain adopts the canonical right-hand configuration with fingers, thumb, and palm subdomains[5]. The pol and exo sites are separated approximately by a linear distance of 32 Å.

Structures of proofreading Polγ

In contrast to the homogeneous replication complex, Polγ in complex with G•T p/t is structurally heterogeneous, resulting in three major conformations (Fig. 1c and Extended Data Fig. 2). The statistics of refinement and geometry of the structures are detailed in Table 1. The structures were captured with Polγ bound to the G•T mismatch in the enolic form, in excising the mismatch, and in an intermediate of transitioning from replication to proofreading at 2.65-Å, 2.91-Å, and 3.04-Å resolutions, respectively (Extended Data Fig. 2). Note that the primer 3′ end locations in pol and exo sites are mutually exclusive (Fig. 1c).

Replication conformer.—The polymerase and G•T p/t in the replication (R) conformer assume conformations identical to that of the replication complex with G•C DNA, and the G•T mismatch adopts the enolic, W-C-like form in the pol site (Fig. 2a,b). Superposition of 13,043 atoms between the two structures yielded a root-mean-square deviation value of

0.606 Å, and the enolic G•T is superimposable to the G•C pair (Extended Data Fig. 3a). The correct incoming nucleotide dCTP forms a W-C pair with the templating G (Extended Data Fig. 3b,c). EM densities of G•T mismatch is similar to that of G•C pair (Extended Data Fig. 3d,e). An inhibitory Ca^{2+} ion is coordinated by the catalytic residues Asp890 and Asp1135 (Extended Data Fig. 3f). If Mg^{2+} ions were present, Polγ would extend the primer and bury the enolic G•T mismatch, resulting in a G-C to A-T transition mutation.

Error-editing conformer.—The $3'$ mismatched nucleoside (position n–1) and two adjacent correctly matched nucleosides (n–2 and n–3) in the primer are completely unwound from the template, and the $3'$ terminus of the primer has been evacuated from the pol site and switched to the exo site (Fig. 2d,h). The $3'$ mismatch is coordinated with two Ca^{2+} ions, which are mediated by exonucleolytic catalytic residues, Asp198 and Glu200, in a nucleolytic-competent configuration analogous to the editing complex of the Pol I Klenow fragment[25]. The incoming nucleotide dCTP remains bound to the fingers subdomain but is no longer paired to any templating residue (Extended Data Fig. 4d).

Intermediate conformer.—The $3'$ end of the primer is not situated in a catalytically competent position for either pol or exo site in the intermediate (I) conformer (Fig. 2c,g), and the conformations of the polymerase and DNA differ from the R or error-editing (E) conformer: the G•T mismatch pair is deviated from the W-C base pair by ~6.5 Å. The incoming nucleotide dCTP is bound to the O helix of the fingers subdomain but is not paired with the template (Extended Data Fig. 4c).

Because the Polγ proofreading complex was formed under identical conditions as that with correctly matched p/t, the conformational diversity is most likely induced by the mismatch, reflecting the various stages of mismatch processing. These structures suggest that Polγ can only process mismatched base pairs in wobble form, and those in enolic form will likely result in a replication error.

Dual fidelity checkpoints and initiation of proofreading

A single-nucleotide-incorporation cycle includes incoming nucleotide binding in the N site (insertion site), a phosphodiester bond-formation reaction that extends the $3'$ end of the primer into the N site, and translocation that moves the primer terminus from the N site to the P site (post-insertion site) (Extended Data Fig. 4e). The Polγ proofreading structures clearly show dCTP occupying the N site in all conformers (Extended Data Fig. 4a–d), the $3'$ mismatch of the primer occupying the P site, and the polymerase adopting the post-translocation conformation. The N site acting as a fidelity checkpoint for correct incoming nucleotides has been well documented[26–29]; our structures show an additional checkpoint in the P site that recognizes and edits a misincorporated nucleotide at a post-translocated state that escapes the N site checkpoint.

In the replication complex with correctly matched DNA, the nascent G•C (n–1) base pair in the P site interacts with Arg853 and Gln1102, a pair of amino acids conserved in the Pol I family, where Arg853 NH_η^+ interacts with O2 of the primer C_{n-1}, and Gln1102 O_ζ interacts with N3 of the template G_{n-1}, respectively (Fig. 2a). Because the distance between O2 and N3 is invariable in all W-C base pairs but variable in wobble pairs[30],

the Arg853–Gln1102 interaction can distinguish correct from incorrect base pairs. Arg853 also forms a bipartite coordination with His1134 juxtaposed to the catalytic Asp1135 (Fig. 2a,b), forming an interaction network of an Arg853–Gln1102–His1134 triad with the primer 3′-OH and correctly positioning the catalytic loop (His1134–Asp1135–Glu1136) for the nucleotidyl-transfer reaction (Extended Data Fig. 5a). The interaction bends three base pairs near the primer 3′ terminus by 16° from the upstream B-form DNA. Similar observations are made in the proofreading R conformer, proving the mutagenic nature of the enolic form of the G•T mismatch.

In the I conformer, Arg853–Gln1102 no longer interact with the mismatched n–1 base pair, which results in primer 3′ terminus migration from the pol site and less interaction with His1134 (Fig. 2c). Loss of the interaction network shifted the Cα of Asp1135 by 1.3 Å from the catalytic-competent position (Extended Data Fig. 5b) and could reduce DNA-synthesis activity. In the E conformer, distances between the primer 3′ mismatch T_{n-1} and the template G_{n-1} to Arg853 and Gln1102 have been extended to 26.4 Å and 17.2 Å, respectively. The results suggest that Arg853–Gln1102 interaction with the nascent base pair associates W-C complementarity with catalysis and abolishes the interaction network that will initiate primer shuttling. Arg853–Gln1102 performs functions as the second checkpoint in the P site that eradicates misincorporations that evade the first checkpoint in the N site (Extended Data Fig. 4e). We termed Arg853–Gln1102 the 'fidelity switch'.

To test the function of the fidelity switch in regulation of replication and proofreading, we constructed Polγ^{R853A} exo-proficient (exo$^+$) and exo$^-$ variants and examined the activities of the holoenzymes. The exo$^-$ variant contains catalytic residue (D198A and E200A) double mutations. On correctly matched p/t, Polγ^{R853A} (exo$^+$) is unable to extend the primer DNA in the presence of dNTP and instead degrades the primer efficiently (Fig. 2i, lanes 10–16). To test whether the lack of DNA synthesis in the R853A variant is outcompeted by exo activity, we repeated the same assay using Polγ^{R853A} exo$^-$. On correctly matched p/t, Polγ^{R853A} exo$^-$ only synthesized two nucleotides in 1 h, while Polγ exo$^-$ formed a full-length product (20-nt addition) in less than 1 min (Fig. 2j). The results show that the mutant has severely reduced intrinsic pol catalysis and exhibits elevated exo activity. The enzymatic assays are in good agreement with the structural conclusion that Arg853 is critical in positioning correctly matched p/t template in the pol site and ejects mismatches toward the exo site. Moreover, mutation of Arg853 ejects even correctly matched p/t and jeopardizes its synthesis.

Polymerase structural changes during the replication–proofreading transition

Fingers opening.—The fingers subdomain constitutes the binding site for incoming nucleotides. Only when bound to a correct nucleotide does it undergo an open-to-closed conformational change to catalyze phosphodiester bond formation. In the R conformer, the fingers subdomain is closed. In the E and I conformers, even though the correct incoming nucleotide, dCTP, is present, the fingers subdomain is fully open in the E conformer and 'ajar' in the I conformer, relative to that in the R and E conformers (Fig. 3a and Supplementary Video 1). The results suggest that a mismatch can hinder the fingers subdomain from closing, thus preventing misincorporation. The presence of nucleotides did

not hinder Polγ from proofreading the nascent mismatched base pair, which is in agreement with a previous single-molecule study[31]. The conclusion is sensible, as proofreading in vivo always functions in the presence of nucleotides.

Thumb rotation.—The thumb subdomain binds to the upstream region of the p/t. The tip of the thumb subdomain rotates 22° from the R conformer to the E conformer, thus directing the p/t from the pol site to the exo site (Fig. 3b). The conformation of the thumb in the I conformer is midway between the R and E conformers.

Trigger loop extending.—A trigger loop (858-TWLTASN-864) situated in the DNA path and juxtaposed to Arg853 undergoes structural changes concerted with the RQH triad interaction with the 3′ terminus of the primer (Supplementary Video 2). The trigger loop in the R conformer contracts to enable the p/t to enter the pol site but extends in the I and E conformers to preclude the mismatched p/t from entering the pol site (Extended Data Fig. 5d–i). The extended trigger loop would sterically clash with the catalytically competent primer location, implying that it contributes to ejecting the primer 3′ terminus from the pol site.

Spacer domain relaxing.—The spacer domain that interacts with the upstream portion of the p/t moves 2.7 Å from its position in the R conformer to that in the I conformer and then another 10 Å from its position in the I conformer to that in the E conformer (Fig. 3c), coordinated with the thumb and DNA rotation.

Subunit interface reduction.—The holoenzyme subunit interface between the PolγA spacer and thumb domains with the PolγB proximal monomer is maintained in all conformers, but the interaction with the PolγB distal monomer occurs only in the R conformer (Fig. 3d) and the I conformer (Fig. 3e), where PolγA Arg232 and the Glu394/ Glu449 PolγB distal monomer region form a tight junction, indicating that the polymerase remains in high-processivity mode. The junction in the E conformer is extended to 17 Å (Fig. 3f), similar to that in the apo enzyme (Extended Data Fig. 6d), suggesting that the E conformer is in low-processivity mode. Prior studies demonstrated that the PolγA substitutions R232G and R232H, which would eliminate the specific subunit interaction, render Polγ with decreased pol and increased exo activities[14], confirming the importance of this region in balancing the two activities. Infant patients carrying these Polγ mutations clinically manifested a combination of progressive neurological and hepatic failure[15].

The overall consequences of Polγ conformational changes during the transition from replication to proofreading are the stepwise reversion to a less-processive, apo-like structure and reduction in protein–DNA contacts (Supplementary Video 3).

DNA strand separation, rotation and backtracking

Because the DNA in all mismatch conformers is identical in length and sequence, p/t interaction with the polymerase can be used to illustrate DNA conformational changes during the transition from replication to proofreading. In the R conformer, Arg802 and Arg807 of the thumb subdomain are in close proximity to the template residues A_{n-8} and G_{n-5} (Fig. 4a), whereas, in the E conformer, they are in close proximity to the primer G_{n-7}

and C_{n-5} (Fig. 4e); we calculated that the p/t duplex rotated 118.9° and backtracked by 4.1 Å, a distance of about 1 bp, from the R to the E conformer (Extended Data Fig. 7c,d). In the I conformer, the p/t duplex rotates 15° relative to that in the R conformer without backtracking (Extended Data Fig. 7b) and maintains a similar interaction with the template strand as in the R conformer (Fig. 4c,d). The most noticeable change in the p/t duplex is the strand separation from T_{n-1} to T_{n-3} positions.

In the E conformer, three nucleosides (Tn, A_{n-1}, and T_{n-2}) in the primer 3′ terminal region are unwound from the template and transferred to the exo site. The primer strand in the E conformer is also moved 32 Å from its location in the R conformer (Extended Data Fig. 8c). The frayed p/t strands are stabilized by a helical splitter subdomain (residues 301–365), in which Phe307 stacks with the template strand T_{n-2} and Arg309 stacks with the primer strand A_{n-4}, stabilizing the first unwound base pair (Extended Data Fig. 9e). The splitter subdomain does not make contact with DNA in the R or I conformer (Extended Data Fig. 8a,b).

As the primer strand will be one nucleoside shorter after the excision, backtracking of the p/t will position the edited 3′ terminus of the primer into the P site of the pol site, ready to resume DNA synthesis. The 105° p/t rotation from the I to the E conformer is close to unwinding three base pairs in B-form DNA ($3 \times 36° = 108°$) while fixing the template strand, which could induce positive supercoiling in the upstream DNA and result in DNA rotation.

Stepwise changes of p/t DNA conformation and polymerase interactions during the transition from replication to proofreading are illustrated in Fig. 4b,d,f.

Mismatch removal

The primer 3′ terminus is coordinated by two Ca^{2+} ions mediated by the carboxyl group of Asp198 and Glu200 in the exo site. Asp198 interacts with a nonbridging oxygen of the scissile phosphodiester bond between T_{n-1} and A_{n-2} nucleosides, and Glu200 interacts with the 3′-OH of T_{n-1} (Fig. 5a). The organization of the Polγ exo site is consistent with an SN_2 nucleolytic reaction mechanism in which Asp198 catalyzes phosphodiester bond hydrolysis and Glu200 stabilizes the leaving group, analogous to the *Escherichia coli* DNA polymerase (DNAP) editing complex about which the two-metal catalysis mechanism was proposed[25].

Reverse transfer of primer after proofreading

In addition to metal coordination by negatively charged residues (Asp198, Glu200, Asp274, and Asp399), the frayed single-stranded DNA is primarily held in the exo site at the 3′ terminus by T-shaped π stacking (Phe273) and hydrophobic residues (Leu203, Met295, Phe377, and Val378) (Fig. 5b). Phosphates of A_{n-2} and T_{n-3} also interact with the main chain residues (Asn354, Ser355, and Leu356) for additional stabilization. Once the 3′ terminal nucleoside is excised, the negatively charged exo site will reduce its affinity to the remainder of the single-stranded primer strand and repel the primer after excision (Fig. 5c). The structure suggests that reverse transfer of primer to the pol site is likely to be spontaneous following the nucleolytic reaction, and the exonuclease reaction is likely non-processive.

Discussion

Replication fidelity is achieved by the polymerase selecting the correct nucleotide in the pol site and removing misincorporated nucleosides in the exo site. We report here an ensemble of Polγ ternary complex structures captured along the reaction pathway from replication to proofreading. Our study provides a structural basis for Polγ mismatch recognition, primer shuttling and mismatch excision as well as the structural mechanism for the mutagenic nature of the G•T mismatch in high-fidelity DNA polymerases.

We concluded that Polγ proofreading is a multistep process that minimally includes (1) the fidelity switch recognizes mismatches in the pol site, halting DNA synthesis, (2) the polymerase's fingers subdomain partially opens, the thumb rotates, and the trigger loop extends, ejecting mismatched primer 6.6 Å from the pol site, and (3) the fingers subdomain opens completely, the thumb rotates another 12°, and the p/t DNA unwinds, backtracks, and rotates by another 105°. The concerted conformational changes move the 3′ primer terminus 32.7 Å from its expelled position to the exo site in a non-linear fashion (Fig. 6). Once a misincorporation occurs, the polymerase transitions from a processive, pro-synthesis mode to a non-processive, pro-nucleolytic mode.

The structural changes that accompany the transition from replication to proofreading occur in the pol domain while the exo domain remains unchanged. We thus hypothesize that the pol site plays an active role in proofreading and the exo site plays a passive role, excising any 3′ terminal nucleoside from the primer in the exo site, regardless of its base pair complementary. Therefore, Polγ proofreading initiation is achieved solely by the enzyme and is independent of the presence of incoming nucleotides and catalytic metal ions.

Pathway of transition between replication and proofreading

In this study, in addition to the replication R conformer and editing E conformer, we also identified an intermediate I conformer. We believe that the I conformer is a functionally relevant intermediate during Polγ transition between two operative states based on the following evidence: first, the I conformer was detected in substantial quantity (~30%) only in the G•T mismatch DNA ternary complex but was found in diminished quantity (<5%) in the same complex with the G•C correctly matched p/t DNA, indicating that the increased population of or conformational change to the I conformer is induced by the mismatch; second, an 'ajar' fingers configuration that is consistent with the Polγ I conformer was observed in DNA Pol I in complex with mismatch-containing DNA in single-molecule Förster resonance energy-transfer (smFRET) studies[31,32], in a crystal structure of mismatch-containing DNA Pol I complex[27], and in T7 DNA polymerase by optical tweezer studies[33]. In the optical tweezer study, the intermediate state between pol and exo sites was catalytically incompetent but still retained its 'memory' on DNA[33]. This observation is consistent with our findings, as the I conformer does not backtrack on DNA and maintains a similar interaction with the p/t duplex as the R conformer.

Based on the structural results, we propose two mechanisms for transition from replication to proofreading in high-fidelity polymerases: an intramolecular model in which the primer strand is 'guided' between the pol and exo sites following a defined pathway and

an intermolecular model in which the polymerase rapidly relaxes the bound DNA in replication mode and rebinds it in editing mode, or a 'catch-and-release' model[34,35]. Both intermolecular and intramolecular primer shuttling has been observed in high-fidelity DNA polymerases by single-molecule studies[33,35].

Resolving a stable Polγ intermediate structure suggests that the Polγ transition pathway from replication to editing has a low-energy point and is accomplished in at least two phases. In phase I, the polymerase transitions from an R conformer to an I conformer, which is induced by mismatch detection and ejection of the primer 3′ terminus from the pol site. In phase II, the polymerase transitions from an I conformer to an E conformer, shuttling the primer 3′ terminus to the exo site (Fig. 6). The 3′ terminus of the primer shuttling from the pol site to the exo site does not follow the shortest linear distance; instead, the primer terminus first moves 9.5 Å and then 33.9 Å for a total of 43.4 Å in a triangular fashion to reach the exo site that is 32.7 Å away (Extended Data Fig. 7).

As relatively small conformational changes in DNA polymerase occur between the R and I conformers, we suggest that the phase I transition is accomplished via an intramolecular pathway, similar to the computational analyses of *E. coli* DNA Pol III in which the movement of the 3′ end of the primer follows a defined pathway[36]. These analyses suggest that the two-phase proofreading pathway may be common among different polymerases.

Because we did not detect any conformer between the I and E conformers, the pathway for phase II cannot be unambiguously defined. As the structure of the editing complex is nearly identical to the Polγ apo form (Extended Data Fig. 6) and substantial structural changes occur in the p/t DNA, the polymerase may reach the E conformer by transiently relaxing its grip of DNA in the I conformer, reversing itself to the apo form and then rebinding to the frayed DNA in a rotated manner. Thus, this second segment of the replication–proofreading transition could be achieved either by an intermolecular or an intramolecular pathway (phases IIa and IIb, Fig. 6). Rapid release of DNA was also observed by smFRET in the DNA Pol I Klenow fragment bound to mismatch-containing DNA[35].

A unified mechanism for mismatch recognition

The frayed p/t configuration has been observed in different polymerase families of which editing complex structures have been determined thus far (RB69 DNA Pol (B family)[37] and DNA Pol III (C family)[38]), and p/t backtracking and rotation have been predicted computationally for T7 DNA Pol (A family)[39]. Therefore, our structural conclusion derived from Polγ could be a general proofreading mechanism for high-fidelity polymerases.

Proofreading begins with error recognition by the second checkpoint, the fidelity switch. The Polγ fidelity switch, Arg853–Gln1102 (Extended Data Fig. 9a), that plays a critical role in coupling mismatch recognition is highly conserved in sequence and function in A family polymerases, for example, Arg573–Gln754 in *Taq* DNA polymerase[40], Arg429–Gln615 in T7 DNA Pol[41] and Arg668–Gln849 in *E. coli* DNA Pol I[42] (Extended Data Fig. 9b). The fidelity switch appears to be functionally conserved in members of other polymerase families, such as Lys706–Asp621 of RB69 DNA Pol (B family)[43] and Arg424–Asp618 of *E. coli* DNA Pol III (C family)[44] (Extended Data Fig. 9c,d). Wobble mismatch-induced pol

site distortion has been observed in BF DNAP (A family)[45] and RB69 DNAP (B family)[26], whereas enolic-form mismatch has been seen in BF DNAP (A family)[46] and Polλ (X family)[28].

Functional importance of the human Polγ fidelity switch is elucidated by their mutations implicated in human diseases. Biochemically, substitutions in the human Polγ fidelity switch (R853A, R853W, R853Q) reduced DNA-synthesis proficiency by more than 1,000-fold relative to wild type (ref. 47 and this study) with elevated exonuclease activity; the Polγ[Q1102A] variant showed a similar trend with reduced pol and elevated exo activities[17], confirming the regulatory role in balancing the pol and exo ratio. Clinically, Polγ fidelity switch mutants manifest as mitochondrial diseases. Patients carrying R853W and P587L substitutions developed progressive external ophthalmoplegia[48] and Parkinsonism[49]; substitutions R853Q and T251I were found in individuals with myocerebrohepatopathy disorders[50].

A duplex DNA containing a terminal mismatch (n–1) can increase the dynamics of itself and two adjacent (n–2 and n–3) nucleosides[38]. It is conceivable that sufficient proofreading of a polymerase depends on its ability to stabilize the transiently frayed p/t. In the Polγ editing complex, the frayed 3′ mismatch primer and the template were stabilized by Phe307 and Arg309 of the splitter subdomain at the DNA fork (Extended Data Fig. 9e). Polγ Arg309 is functionally important. Patients carrying the R309L substitution in trans with T251I displayed profound proximal and distal neurological muscle weakness and wasting[51]; the R309H substitution compounded with G737R resulted in slowed psychomotor development and progressive balance difficulties and motor functions[52,53]; and the R309H substitution in trans with G1051R were found in patients with epilepsy and Alpers disease in trans with R627Q[52–54].

The concept of stabilizing the frayed DNA appears to be conserved among different polymerase families, even though the Polγ splitter subdomain is not. In the A family *E. coli* Klenow fragment, Arg455 interacts with the n–3 nucleoside in the primer strand (Extended Data Fig. 9f), but the aromatic residue is missing. In the B family RB69 polymerase editing complexes, Phe123 and Arg260 interact with the frayed primer n–2 nucleoside and the template n–3 nucleoside, respectively (Extended Data Fig. 9g). In the C family *E. coli* DNA Pol III editing complex, Tyr453 and Arg447 interact with the n–4 nucleosides of the primer and template strands, respectively (Extended Data Fig. 9h). Substitution of Tyr453 with alanine reduces exo activity[38], confirming its role in proofreading. The analyses suggest that arginine plays a critical role in stabilization of frayed DNA, as it is conserved in A, B, and C family polymerases, while the aromatic residue plays a supportive role, as it is missing in some DNA polymerases.

The structures presented here are an ensemble showing a high-fidelity DNA polymerase at various stages of proofreading, owing to the advantage of cryo-EM over other methods that may capture a single structure at a time. Our study not only enables us to dissect the structural mechanism of maintenance of mitochondrial genome integrity but also provides a blueprint to understand many Polγ substitutions implicated in human diseases that were previously inexplicable.

Methods

Protein purification

His-tagged PolγA and PolγB subunits were expressed in *Spodoptera frugiperda* (Sf9) cells and *E. coli* Rosetta (DE3) cells, respectively, and purified as previously described[5]. The wild-type PolγA subunit construct lacks the N-terminal 25 residues (the putative mitochondrial-localization sequence) as well as ten of the 13 sequential glutamines (residues 43–52). The sequences for PolγA^{R853A} exo$^+$ and exo$^-$ were constructed by site-directed mutagenesis on the pBacPak9 plasmid, and baculovirus was constructed via homologous recombination with BacPAK6 DNA (Takara Bio). The wild-type and mutant PolγA proteins were expressed in Sf9 cells transiently infected with the baculovirus and purified sequentially using TALON Superflow cobalt resin (Cytiva) and a Superdex 200 gel filtration column (Cytiva) following an established procedure[5]. The construct of wild-type PolγB lacking the N-terminal 25 residues (the putative mitochondrial-localization sequence) was expressed in *E. coli* BL21 RIL cells and purified using Ni-NTA affinity (Qiagen) and Mono S cation exchange chromatography. Purified PolγA and PolγB were complexed at a molar ratio of 1:2, respectively, on ice and then applied to a Superdex 200 gel filtration column (Cytiva). Fractions containing the Polγ holoenzyme were pooled, concentrated and stored at −80 °C in small aliquots. Protein concentrations were determined based on the absorbance at 280 nm using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

Oligonucleotide preparation

All oligonucleotides were synthesized by Integrated DNA Technologies and purified using high-performance liquid chromatography. For structural studies, a 28-nt template (5′-CGAGGTATGGCACTGGCCGTCGTTTTCG-3′) was annealed with a correctly matched 24-nt oligonucleotide (5′-CGAAAACGACGGCCAGTGCCATAC-3′) to generate G•C p/t substrate for the replication complex or a 24-nt mismatched oligonucleotide (5′-CGAAAACGACGGCCAGTGCCATAT-3′) to generate G•T p/t substrate for the proofreading complex. Annealing of oligonucleotides was performed at a 1:1 primer/template ratio by heating at 95 °C for 5 min followed by slow cooling to room temperature overnight in 50 mM Tris (pH 7.8), 50 mM NaCl, and 1 mM EDTA (pH 8.0).

For activity assays, a 25-nt primer (5′-CGAAAACGAGGGCCAGTGCCATACC-3′) was annealed to a 45-nt template (5′-TACGAGCCTGCCTGACGTGCGGTATGGCACTGGCCCTCGTTTTCG-3′) at a 1:1.1 primer/template ratio in a boiling water bath and allowed to cool to room temperature overnight in 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 1 mM EDTA.

Cryo-electron microscopy sample preparation

Polγ (2 μM) was incubated with 2 μM correctly matched or mismatched DNA and 1 mM dCTP on ice in buffer containing 20 mM HEPES (pH 7.5), 140 mM KCl, 1 mM EDTA (pH 8.0), 10 mM CaCl$_2$, 10 mM 2-β-mercaptoethanol, and 0.01% (wt/vol) octyl-β-glucoside. Sample (4 μl) was applied onto a plasma-cleaned QUANTIFOIL R 2/1 Cu 200 grid (Electron Microscopy Sciences) and rapidly frozen in liquid ethane using the Vitrobot Mark IV system at 22 °C and 100% humidity (Thermo Fisher Scientific).

Cryo-electron microscopy data acquisition

Frozen grids of the Polγ replicating complex were loaded into a Titan Krios G3i microscope (Thermo Fisher Scientific) equipped with a K3 direct electron detector with a GIF Quantum energy filter (20-eV energy slit) (Gatan) and operating at 300 keV, with a condenser lens aperture of 50 μm, a spot size of 4, and a parallel beam with an illuminated area 1.0 μm in diameter. Cryo-EM data were automatically acquired using EPU software in counted super-resolution mode at a nominal magnification of 105,000× (corresponds to 0.425 Å per pixel) with nominal defocus range between −1.5 and −2.5 μm. Fifty-frame movie stacks were collected over an exposure time of 1.10 s with a total dose of ~50 e$^-$ Å^{-1}. A total of 9,828 movie stacks were collected.

Similar movies were collected from frozen grids of the Polγ proofreading complex using the same microscope under the same conditions. Forty-four-frame movie stacks were collected over an exposure time of 1.0 s with a total dose of ~44 e$^-$ Å^{-1}. A total of 14,122 movie stacks were collected.

Cryo-electron microscopy data processing

For the Polγ–DNA$_{G•C}$–dCTP replication ternary complex, the acquired movie frames were imported into cryoSPARC[55] for image processing (Extended Data Fig. 1a). Movie data were motion corrected for beam-induced motion and 2× binned on the fly to 0.85 Å per pixel using Patch Motion Correction. The contrast transfer function (CTF) was estimated using Patch CTF Estimation. Micrographs with CTF fit worse than 4 Å were excluded, leaving 9,697 micrographs for further processing. A total of 13,138,434 particles were selected using Blob Picker and extracted with 4× binning (80 pixels). After three rounds of two-dimensional (2D) classification, 200,000 of the 7,698,592 remaining particles were used to create three initial three-dimensional (3D) volumes using Ab Initio. The resulting 3D volumes were used as input for two iterative rounds of heterogeneous refinement, resulting in two good classes: class 2 (2,521,940 particles) and class 3 (3,679,400 particles). Particles belonging to class 1 in the first round of heterogeneous refinement were discarded before starting the second round. Particles were recentered and re-extracted without binning (320 pixels), resulting in 2,504,778 and 3,656,462 particles for class 2 and class 3, respectively. Class 2 and class 3 were first refined using homogeneous refinement, resulting in resolutions of 2.73 Å and 2.48 Å, respectively. Class 2 was subjected to another round of heterogeneous refinement with three repeated input volumes from the previous homogeneous refinement. Class 3 (3,656,462 particles) was combined with class 2-3 (838,529 particles) and was further refined by homogeneous refinement, followed by cryoSPARC's implementation of local and global CTF refinements[56] and a final round of non-uniform refinement[57], yielding a 2.46-Å-resolution reconstruction according to gold-standard FSC (GSFSC) at 0.143 (Extended Data Fig. 1b).

For the Polγ-DNA$_{G•T}$–dCTP proofreading ternary complex, the acquired movie frames were imported into cryoSPARC[55] and preprocessed similarly as the replication ternary complex (Extended Data Fig. 2a). Micrographs with CTF fit worse than 4 Å were excluded, leaving 12,256 micrographs for further processing. Particles were picked on 12,256 micrographs using Blob Picker, and 17,463,536 particles were extracted with 4× binning (80 pixels).

After three iterative rounds of 2D classification, 200,000 of the 5,212,165 remaining particles were used to create five initial 3D volumes using Ab Initio. The resulting 3D volumes were used as input for two iterative rounds of heterogeneous refinement, resulting in class 2 (1,504,248 particles) and class 3 (1,546,413 particles). Particles belonging to classes 4 and 5 in the first round of heterogeneous refinement were discarded before starting the second round. Particles belonging to class 1 were initially kept for the second round of heterogeneous refinement. Particles were recentered and re-extracted without binning (320 pixels), resulting in 1,494,729 and 1,537,285 particles for class 2 and class 3, respectively. Class 2 and class 3 were first refined using homogeneous refinement, resulting in resolutions of 2.89 Å and 2.74 Å, respectively. Class 2 was subjected to another round heterogeneous refinement with three repeated input volumes from the previous homogeneous refinement. Class 2-1 and class 2-2 were combined and further refined using homogeneous refinement, followed by local and global CTF refinements[56] and a final round of non-uniform refinement[57], yielding a 2.91-Å-resolution reconstruction according to GSFSC at 0.143 (Extended Data Fig. 2b, left). Class 3-1 was refined identically, yielding a 3.08-Å-resolution reconstruction (Extended Data Fig. 2b right). Class 3-2 and Class 3-3 were combined and refined similarly, yielding a 2.66-Å-resolution reconstruction according to GSFSC at 0.143 (Extended Data Fig. 2b, right).

For the E conformer, local refinement was performed in cryoSPARC[55] to improve the electron density. Two masks were created using UCSF Chimera[58]: one covering the PolγA subunit and duplex DNA and the other covering the PolγB dimer. Masks were imported into cryoSPARC with dilation and soft padding. Local refinement was performed using each mask without particle subtraction. Local refinement of the map covering the PolγA subunit and duplex DNA refined to a resolution of 2.76 Å according to GSFSC at 0.143. Local refinement of the map covering the PolγB dimer refined to a resolution of 2.76 Å according to GSFSC at 0.143. These two maps were independently post-processed using DeepEMhancer[59] or LocSpiral[60] and combined with 'combine_focused_maps' in Phenix[61].

Local resolutions of final reconstructions were estimated by cryoSPARC's implementation of blocres[62] and were colored by a resolution range from 2 Å to 4 Å in UCSF ChimeraX[63] (Extended Data Figs. 1b and 2b).

Model building, refinement and analysis

Cryo-EM maps were post-processed with DeepEMhancer[59] and LocSpiral[60]. Crystal structures of the Polγ–DNA–dCTP ternary complex (PDB 4ZTZ)[6] and wild-type PolγB (PDB 2G4C)[64] were used to build initial models by rigid body fitting in Chimera[58] and manual adjustments in Coot[65] into the DeepEMhancer-processed map. Models were refined iteratively by real-space refinement in Phenix[61] and by manual adjustments and improvements in Coot[65] and ISOLDE[66]. The final refinement was carried out with the LocSpiral-processed map because it gave better rotamer conformations than the DeepEMhancer-processed map. The local grid search option was turned off. Refined models were validated with MolProbity[67] and Q-score analysis[68]. Protein–DNA interaction was analyzed using DNAproDB[69]. Figures were prepared using UCSF Chimera[58], UCSF ChimeraX[63], Segger[70] and PyMOL (Schrödinger).

Polγ activity assays

DNA-synthesis assays were carried out with 50 nM Polγ incubated with 100 nM DNA at 37 °C for 5 min in binding buffer (25 mM HEPES (pH 7.5), 140 mM KCl, 5% (vol/vol) glycerol, 1 mM β-mercaptoethanol, and 50 μg ml^{-1} BSA). Reactions were started by the addition of dNTPs (50 μM each) and 10 mM MgCl$_2$. At the indicated times, aliquots were removed and quenched in stop buffer (80% formamide, 50 mM EDTA, 0.1% (wt/vol) SDS, 5% glycerol, and 0.02% bromophenol blue). Reaction mixtures were heated at 95 °C for 5 min and then immediately placed on ice. Reaction products were separated on 20% polyacrylamide–7 M urea gels. Following electrophoresis, the gels were soaked in 50% methanol and 20% glycerol solution, dried overnight under heat and vacuum, autoradiographed and imaged using an Amersham Typhoon Biomolecular Imager (Cytiva).

Extended Data
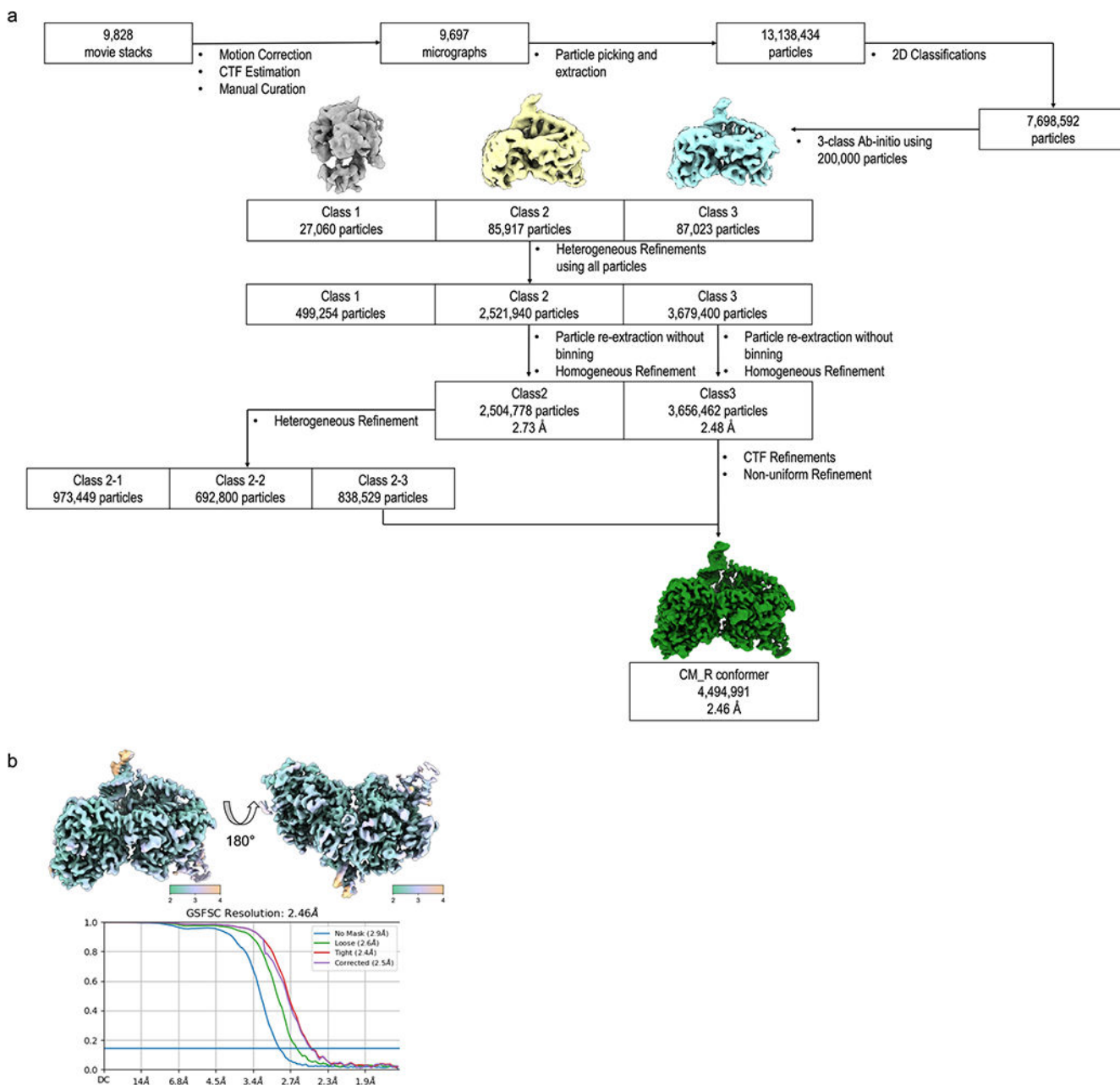
Extended Data Fig. 1 |. Cryo-EM single-particle data processing pipeline for replication ternary complex.

a, Data were imported into cryoSPARC for 2D and 3D image analysis. See Methods for details. **b**, EM map of replication complex colored by local resolution estimated by cryoSPARC's implementation of blocres at 0.5 FSC threshold and FSC resolution at 0.143 GSFSC threshold reported by cryoSPARC.
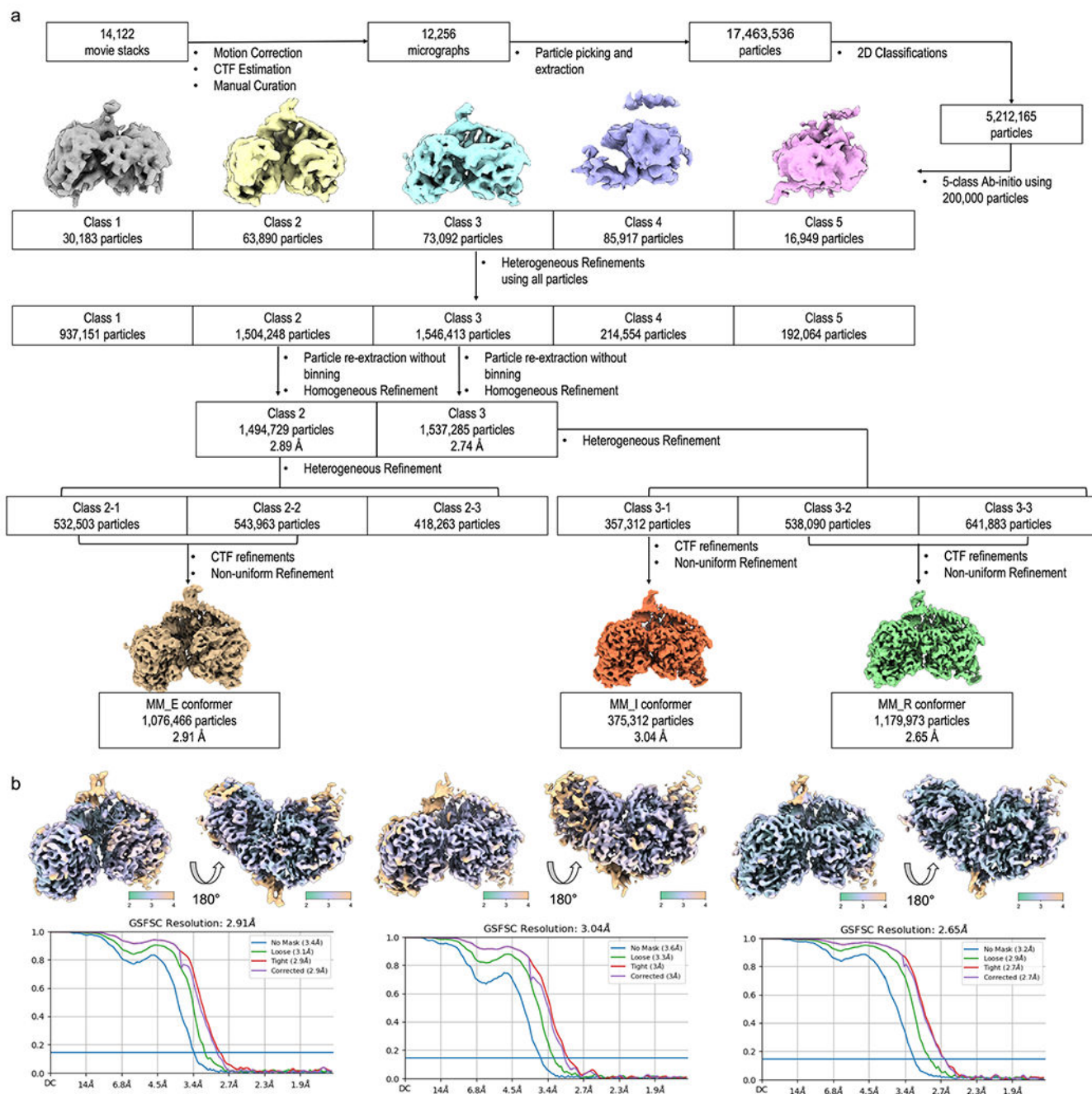
Extended Data Fig. 2 |. Cryo-EM single-particle data processing pipeline for proofreading ternary complex.

a, Data were imported into cryoSPARC for 2D and 3D image analysis. See Methods for details. **b**, EM map of proofreading E-conformer (Left), I-conformer (Middle), and R-conformer (Right) colored by local resolution estimated by cryoSPARC's implementation of blocres at 0.5 FSC threshold and FSC resolution at 0.143 GSFSC threshold reported by cryoSPARC.
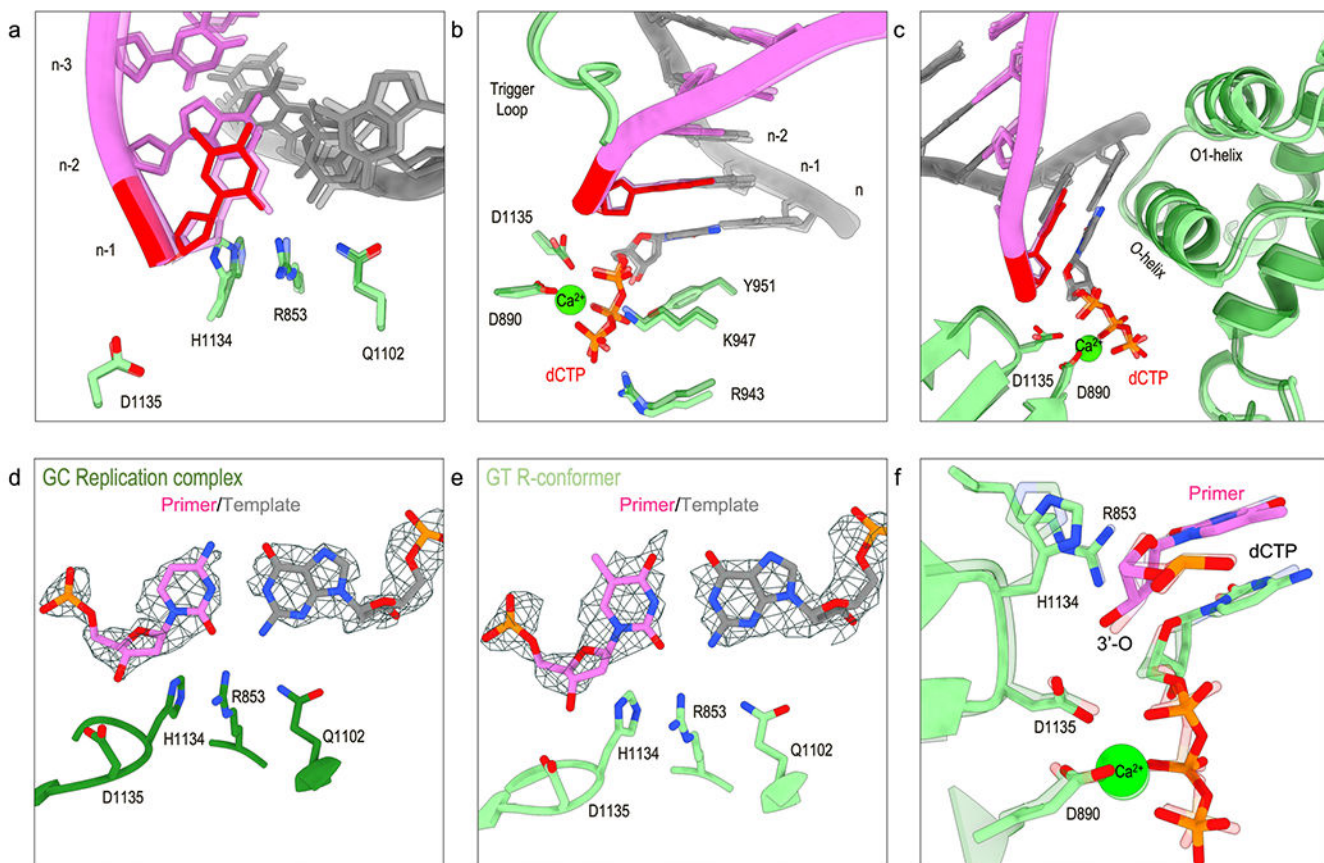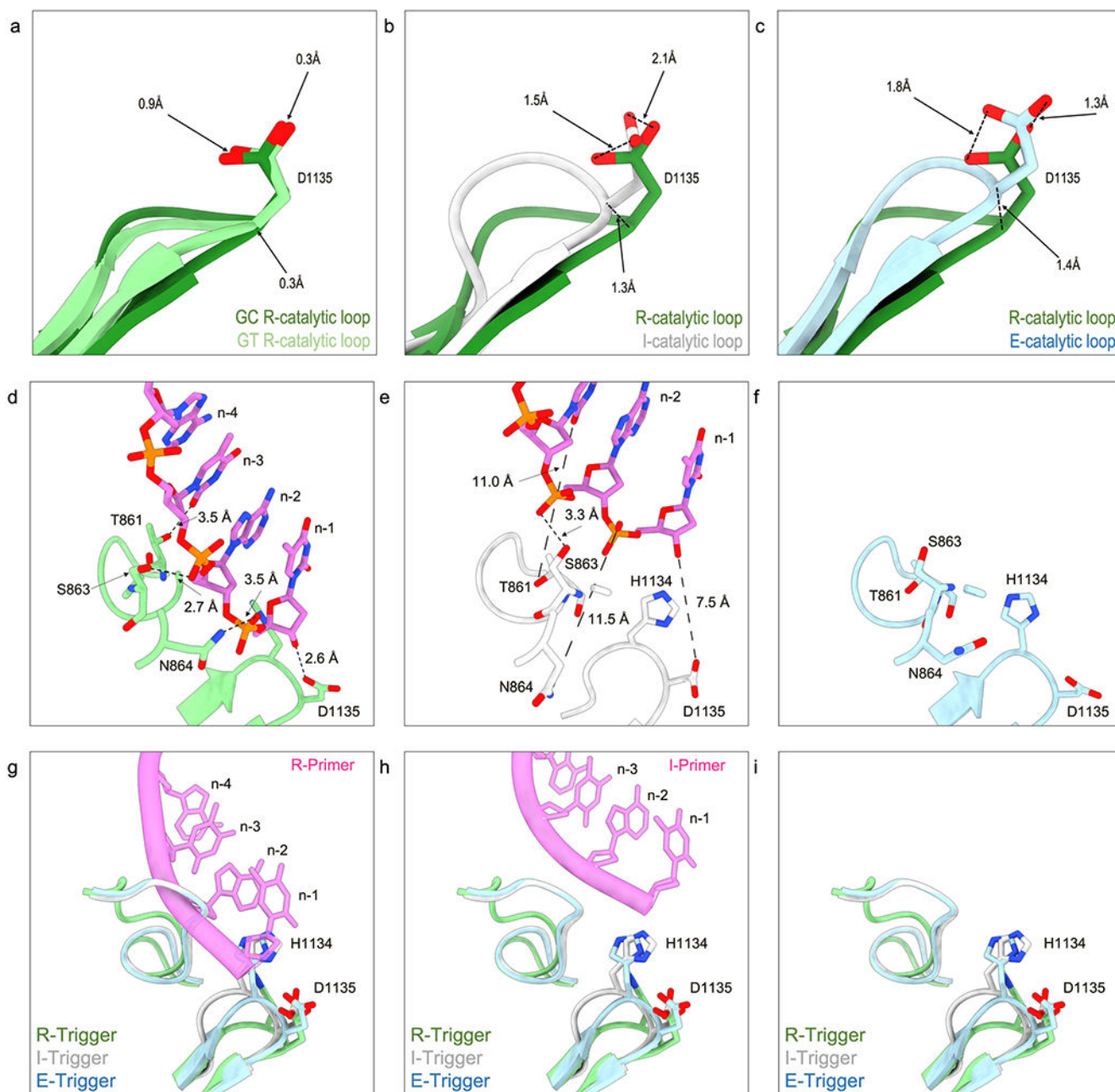
Extended Data Fig. 3 |. Superposition of G·C replication complex and G·T mismatch R-conformer in the *pol* active site.

Superposition of (**a**) DNA, catalytic residue (Asp[1135]), and RQH triad (Arg[853]Gln[1102]His[1134]), (**b**) trigger loop, and (**c**) catalytic loop and Fingers subdomain. G·C replication complex and G·T R-conformer are shown in transparent and light green, respectively. Primer strand is shown in pink, template strand in gray, and mismatched nucleoside in red. Overlay of primer-template base pair in the P-site with corresponding EM density in (**d**) G·C replication complex and in (**e**) G·T R-conformer. (**f**) Superposition of 3′-end of the primer in the *pol* active site of G·C replication complex (transparent) and G·T R-conformer (light green) showing the positioning of 3′-OH. Residues Asp[890], Asp[1135], Arg[853], and His[1134] as well as the incoming nucleotide, dCTP, are displayed.
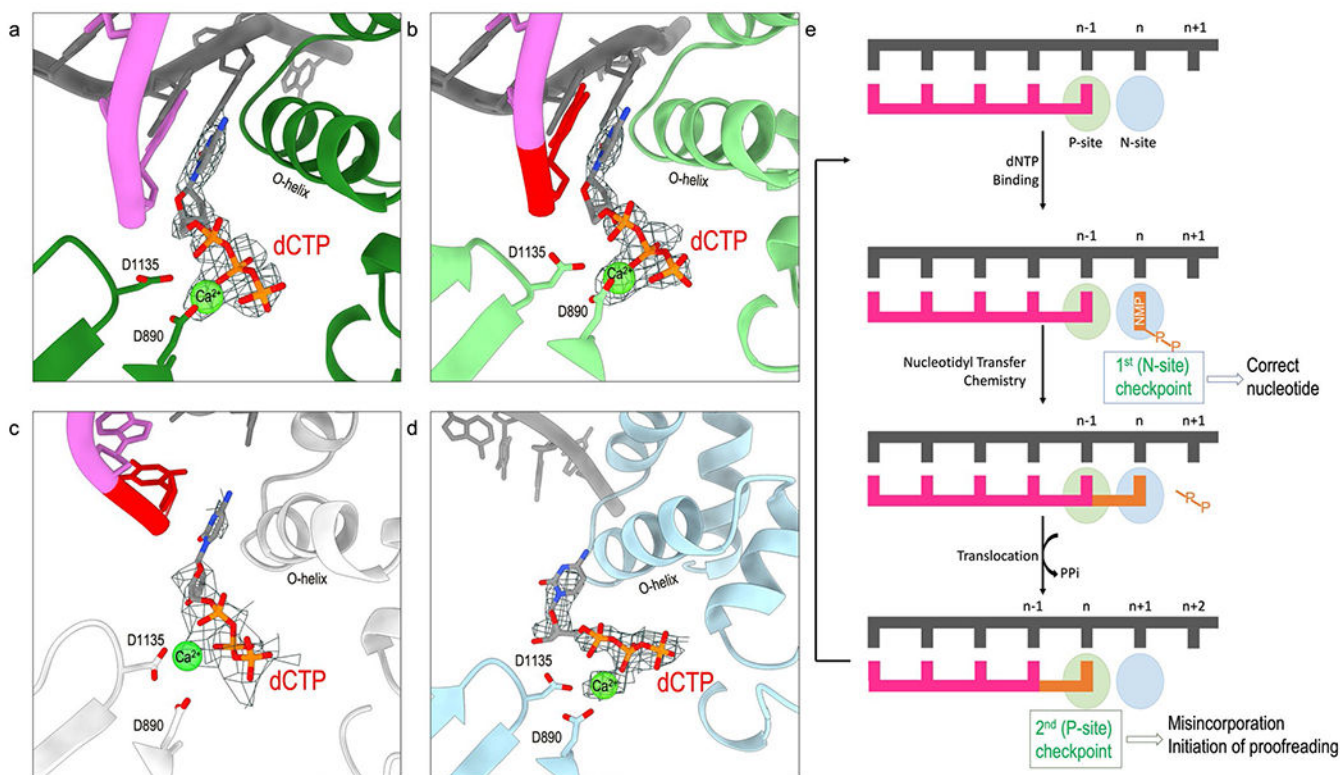
Extended Data Fig. 4 |. Incoming nucleotide in Pol γ ternary complexes and dual fidelity checkpoint.

Incoming nucleotide dCTP is in the *pol* active site of (**a**) the replication complex and (**b-d**) proofreading complex. In R-conformers (**a, b**), dCTP forms W-C pair with the template but does not form any base pair in the I-conformer (**c**) and in the E-conformer (**d**). Electron density for dCTP and Ca^{2+} shown in light blue mesh. Proteins are colored based on their conformation, primer strand in pink, template strand in gray, and mismatched nucleoside in red. **e**, Scheme of DNA synthesis pathway. Fidelity checkpoint at pre- and post-nucleotide incorporation stages in Pol γ. In the N-site, where incoming nucleotide binds, Watson-Crick base checking first chooses the correct nucleotide triphosphate to incorporate. After the incorporation, Arg^{853} and Gln^{1102} checks the nascent base pair for correct Watson-Crick geometry.

Extended Data Fig. 5 |. Catalytic loop positioning and trigger loop interaction in Pol γ ternary complexes.

a–c, Comparison of catalytic loops between G·C replicating complex and G·T R-conformer (**a**), I-conformer (**b**), and G·T E-conformer (**c**). **d–f,** Detailed interaction between trigger loop and primer strand in R- (**d**), I- (**e**), and E-conformers (**f**). **g–i,** Superposition of trigger loops in three proofreading structures. Primer strand from R- (**g**), I- (**h**), and E- (**i**) conformers are shown.
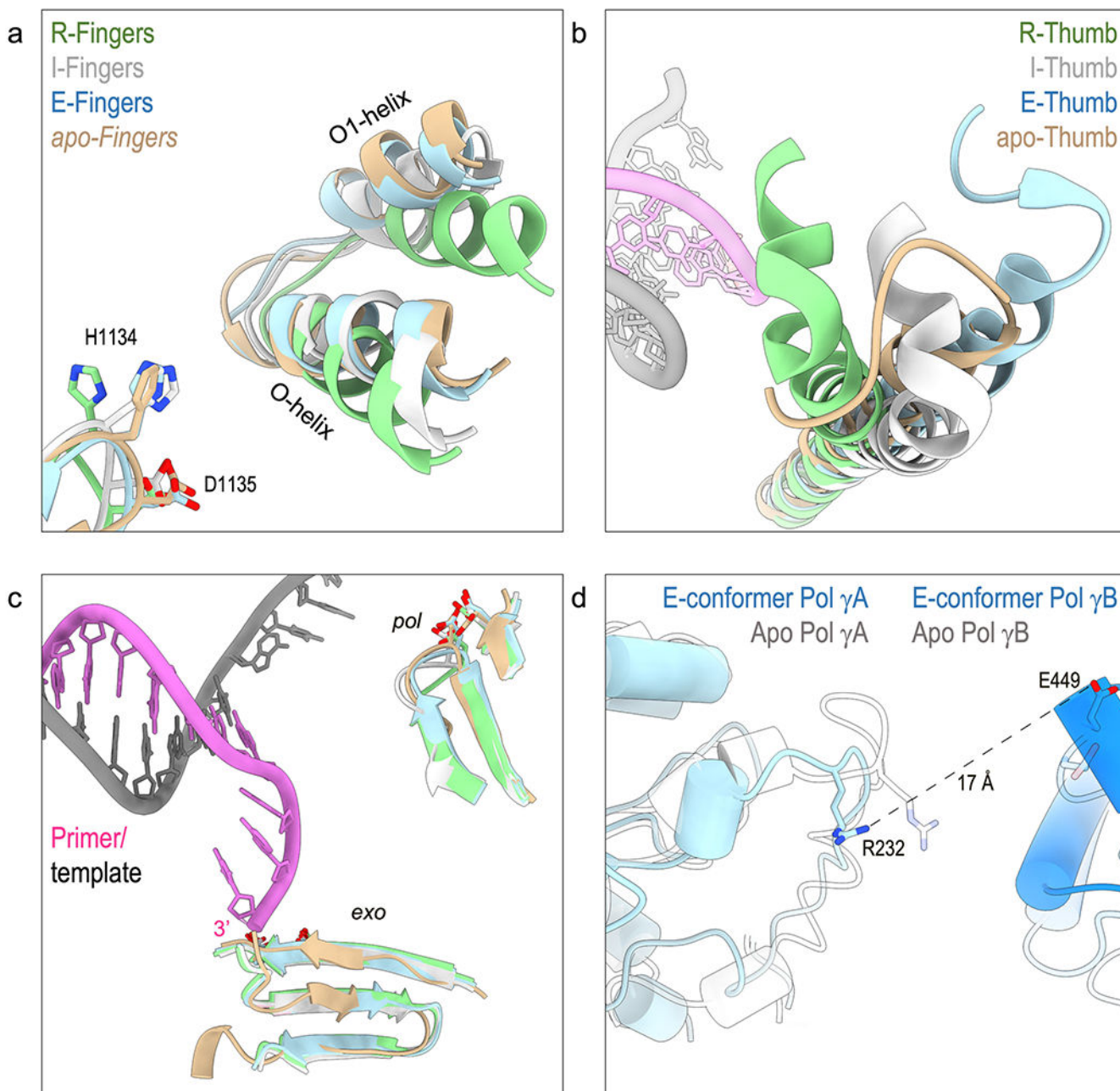
Extended Data Fig. 6 |. Structural comparison of proofreading structures with *apo* Pol γ.
a-c, Superposition of Fingers subdomain (**a**), Thumb subdomain (**b**), and polymerase and exonuclease active sites (**c**) of proofreading structures and *apo* Pol γ (tan). Proofreading structures are colored according to conformation: R (light green), I (white), and E (light blue). **d**, Superimposed subunit interface of E-conformer and the *apo* enzyme.
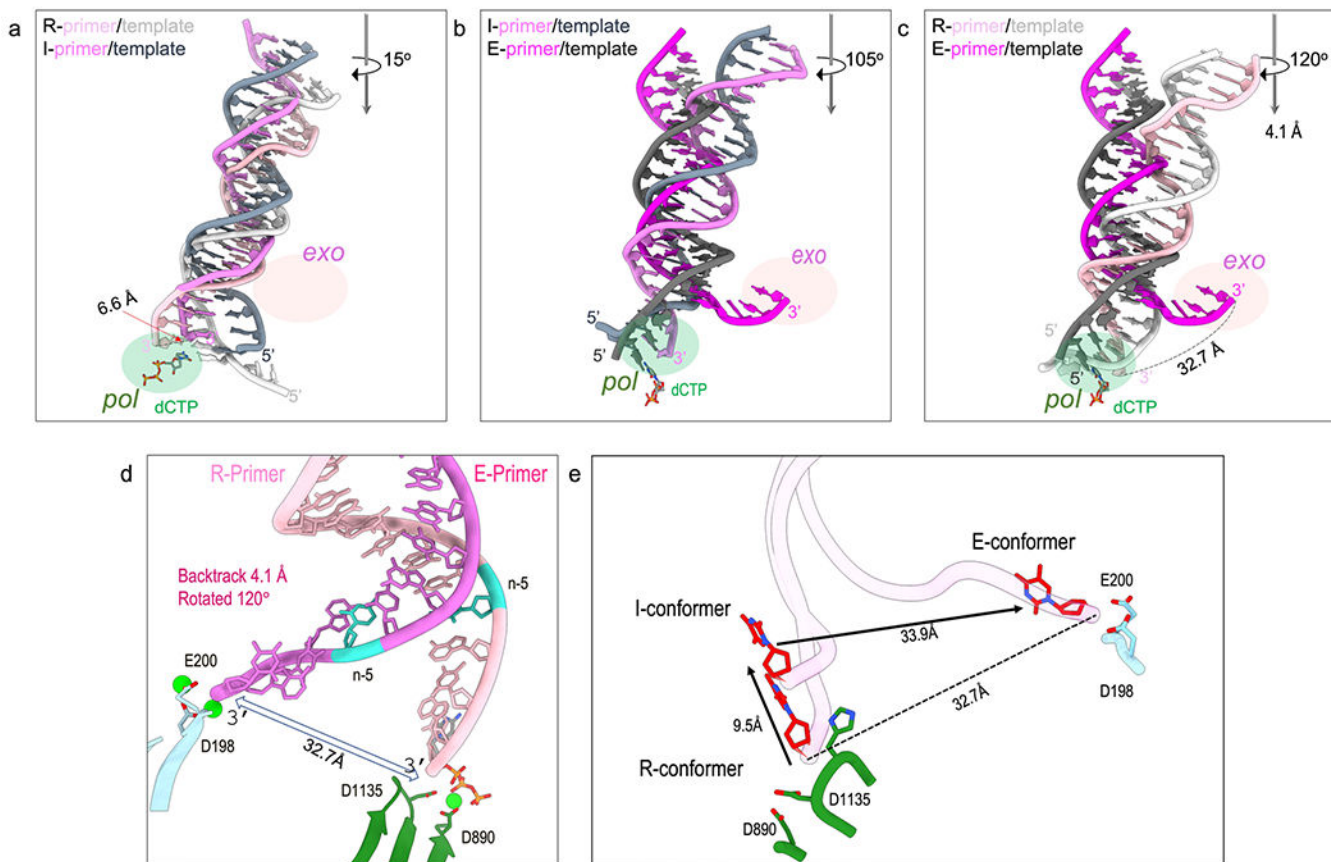
Extended Data Fig. 7 |. DNA movement between conformational states in Pol γ.
a-c, Overall DNA movement associated with conformational changes are shown between
R- and I- (**a**), I- and E- (**b**), and R- and E-conformers (**c**). **d**, Superposition of the primer
strand in the correctly matched G·C replication complex and mismatch G·T E-conformer
after superimposing the invariable *exo* active site, nucleoside *n-5* is marked as a point of
reference for comparison. **e**, Triangular primer 3′-end shuttling from the *pol* site to the *exo*
site via an intermediate of the pathway. The primer moves 9.5 Å from R-conformer *pol* site
to I-conformer and another 33.9 Å to the *exo* site in the E-conformer to shuttle between *pol*
and *exo* sites that are 32.7 Å apart. Catalytic residues for *pol* (green) and *exo* (light blue) are
shown.
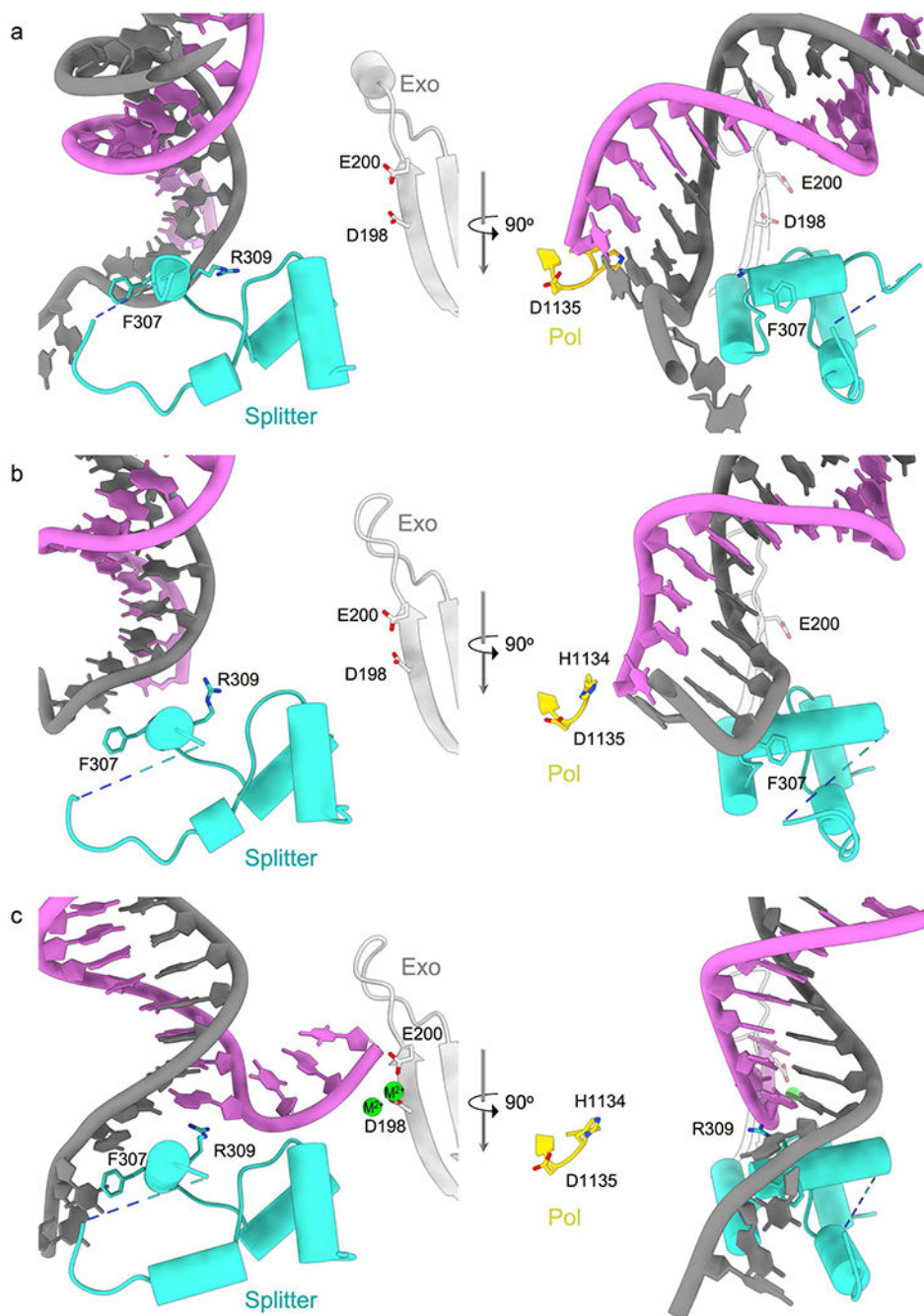
Extended Data Fig. 8 |. DNA position in proofreading structures relative to exonuclease site and splitter helix.
Top (left) and 90° rotated (right) views of R- (**a**), I- (**b**), and E- (**c**) conformers. In R-conformer (**a**), DNA sits in polymerase active site and does not contact splitter helix. In I-conformer (**b**), DNA is lifted out of the polymerase active site, but does not reach the splitter helix. In E-conformer (**c**), DNA is completely out of the polymerase active site and is stabilized by the splitter helix at the fork junction.
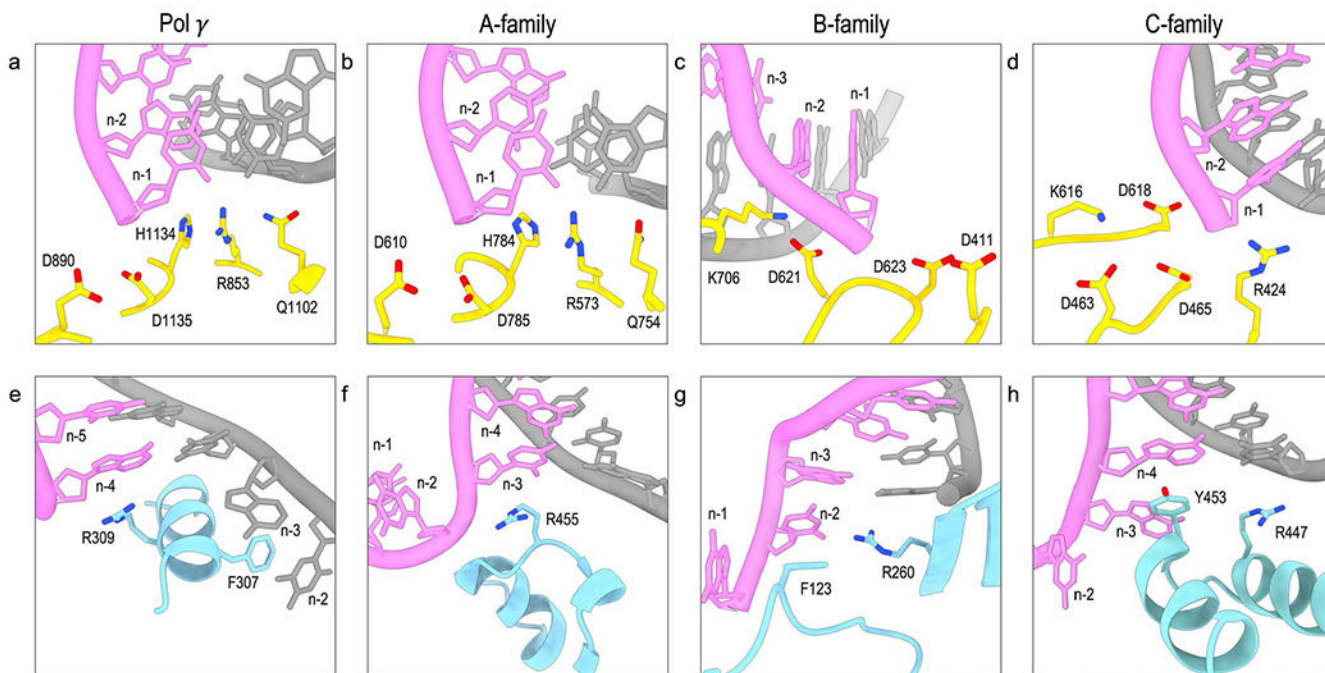
Extended Data Fig. 9 |. Conserved mechanism of proofreading among different DNA polymerases.

a-d, Side-by-side comparison of polymerase active site of Pol γ G·C replication complex (**a**), *Taq* Klenow Fragment (PDB:3KTQ)(**b**), RB69 DNA polymerase (PDB: 3NCI) (**c**), and *Taq* DNA polymerase III α subunit (PDB: 3E0D) (**d**) from A-,B-, and C-family DNA polymerases, respectively. **e-h**, Side-by-side comparison of DNA fork junction in Pol γ G·T E-conformer (**e**), *E. coli* Klenow Fragment (PDB: 1KLN) (**f**), RB69 DNA polymerase (PDB: 1CLQ) (**g**), and *E. coli* DNA polymerase III (PDB: 5M1S) (**h**) from A-,B-, and C-family DNA polymerases, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data availability

Cryo-EM density maps and corresponding coordinates have been deposited in the Worldwide Protein Data Bank (PDB) OneDep System under the following accession codes: G·C replication complex (EMD-27154) (PDB 8D33), G·T R conformer (EMD-27155) (PDB 8D37), G·T I conformer (EMD-27163) (PDB 8D3R), and G«T E conformer (EMD-27172)

(PDB 8D42). Accession codes for consensus and local refinement maps of the G·T E conformer are EMD-27169 (consensus), EMD-27170 (local refinement of subunit A and p/t DNA) and EMD-27171 (local refinement of subunit B).

References

1. Robberson DL, Kasamatsu H & Vinograd J Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. Proc. Natl Acad. Sci. USA 69, 737–741 (1972). [PubMed: 4501588]

2. Berk AJ & Clayton DA Mechanism of mitochondrial DNA replication in mouse L-cells: asynchronous replication of strands, segregation of circular daughter molecules, aspects of topology and turnover of an initiation sequence. J. Mol. Biol 86, 801–824 (1974). [PubMed: 4473554]

3. Lim SE, Longley MJ & Copeland WC The mitochondrial p55 accessory subunit of human DNA polymerase γ enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. J. Biol. Chem 274, 38197–38203 (1999). [PubMed: 10608893]

4. Yin YW Structural insight on processivity, human disease and antiviral drug toxicity. Curr. Opin. Struct. Biol 21, 83–91 (2011). [PubMed: 21185718]

5. Lee YS, Kennedy WD & Yin YW Structural insights into human mitochondrial DNA replication and disease-related polymerase mutations. Cell 139, 312–324 (2009). [PubMed: 19837034]

6. Szymanski MR et al. Structural basis for processivity and antiviral drug toxicity in human mitochondrial DNA replicase. EMBO J. 34, 1959–1970 (2015). [PubMed: 26056153]

7. Sohl CD et al. Probing the structural and molecular basis of nucleotide selectivity by human mitochondrial DNA polymerase γ. Proc. Natl Acad. Sci. USA 112, 8596–8601 (2015). [PubMed: 26124101]

8. Johnson AA & Johnson KA Exonuclease proofreading by human mitochondrial DNA polymerase. J. Biol. Chem 276, 38097–38107 (2001). [PubMed: 11477094]

9. Lee YS et al. Each monomer of the dimeric accessory protein for human mitochondrial DNA polymerase has a distinct role in conferring processivity. J. Biol. Chem 285, 1490–1499 (2010). [PubMed: 19858216]

10. Wu P, Nossal N & Benkovic SJ Kinetic characterization of a bacteriophage T4 antimutator DNA polymerase. Biochemistry 37, 14748–14755 (1998). [PubMed: 9778349]

11. Reha-Krantz LJ Regulation of DNA polymerase exonucleolytic proofreading activity: studies of bacteriophage T4 'antimutator' DNA polymerases. Genetics 148, 1551–1557 (1998). [PubMed: 9560374]

12. Hadjimarcou MI, Kokoska RJ, Petes TD & Reha-Krantz LJ Identification of a mutant DNA polymerase δ in Saccharomyces cerevisiae with an antimutator phenotype for frameshift mutations. Genetics 158, 177–186 (2001). [PubMed: 11333228]

13. Foury F & Szczepanowska K Antimutator alleles of yeast DNA polymerase γ modulate the balance between DNA synthesis and excision. PLoS ONE 6, e27847 (2011). [PubMed: 22114710]

14. Lee YS, Johnson KA, Molineux IJ & Yin YW A single mutation in human mitochondrial DNA polymerase Pol γA affects both polymerization and proofreading activities of only the holoenzyme. J. Biol. Chem 285, 28105–28116 (2010). [PubMed: 20513922]

15. Ferrari G. et al. Infantile hepatocerebral syndromes associated with mutations in the mitochondrial DNA polymerase-γA. Brain 128, 723–731 (2005). [PubMed: 15689359]

16. Trifunovic A. et al. Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature 429, 417–423 (2004). [PubMed: 15164064]

17. Bratic A. et al. Complementation between polymerase- and exonuclease-deficient mitochondrial DNA polymerase mutants in genomically engineered flies. Nat. Commun 6, 8808 (2015). [PubMed: 26554610]

18. Kunkel TA & Alexander PS The base substitution fidelity of eucaryotic DNA polymerases. Mispairing frequencies, site preferences, insertion preferences, and base substitution by dislocation. J. Biol. Chem 261, 160–166 (1986). [PubMed: 3941068]

19. Fortune JM et al. *Saccharomyces cerevisiae* DNA polymerase δ: high fidelity for base substitutions but lower fidelity for single- and multi-base deletions. J. Biol. Chem 280, 29980–29987 (2005). [PubMed: 15964835]

20. Bebenek A. et al. Interacting fidelity defects in the replicative DNA polymerase of bacteriophage RB69. J. Biol. Chem 276, 10387–10397 (2001). [PubMed: 11133987]

21. Wong I, Patel SS & Johnson KA An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. Biochemistry 30, 526–537 (1991). [PubMed: 1846299]

22. Johnson AA & Johnson KA Fidelity of nucleotide incorporation by human mitochondrial DNA polymerase. J. Biol. Chem 276, 38090–38096 (2001). [PubMed: 11477093]

23. Kimsey IJ et al. Dynamic basis for dG·dT misincorporation via tautomerization and ionization. Nature 554, 195–201 (2018). [PubMed: 29420478]

24. van Heel M & Schatz M Fourier shell correlation threshold criteria. J. Struct. Biol 151, 250–262 (2005). [PubMed: 16125414]

25. Freemont PS, Friedman JM, Beese LS, Sanderson MR & Steitz TA Cocrystal structure of an editing complex of Klenow fragment with DNA. Proc. Natl Acad. Sci. USA 85, 8924–8928 (1988). [PubMed: 3194400]

26. Xia S, Wang J & Konigsberg WH DNA mismatch synthesis complexes provide insights into base selectivity of a B family DNA polymerase. J. Am. Chem. Soc 135, 193–202 (2013). [PubMed: 23214497]

27. Wu EY & Beese LS The structure of a high fidelity DNA polymerase bound to a mismatched nucleotide reveals an 'ajar' intermediate conformation in the nucleotide selection mechanism. J. Biol. Chem 286, 19758–19767 (2011). [PubMed: 21454515]

28. Bebenek K, Pedersen LC & Kunkel TA Replication infidelity via a mismatch with Watson–Crick geometry. Proc. Natl Acad. Sci. USA 108, 1862–1867 (2011). [PubMed: 21233421]

29. Lin P. et al. Incorrect nucleotide insertion at the active site of a G:A mismatch catalyzed by DNA polymerase β. Proc. Natl Acad. Sci. USA 105, 5670–5674 (2008). [PubMed: 18391201]

30. Seeman NC, Rosenberg JM & Rich A Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl Acad. Sci. USA 73, 804–808 (1976). [PubMed: 1062791]

31. Berezhna SY, Gill JP, Lamichhane R & Millar DP Single-molecule Förster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase I. J. Am. Chem. Soc 134, 11261–11268 (2012). [PubMed: 22650319]

32. Hohlbein J. et al. Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. Nat. Commun 4, 2131 (2013). [PubMed: 23831915]

33. Hoekstra TP et al. Switching between exonucleolysis and replication by T7 DNA polymerase ensures high fidelity. Biophys. J 112, 575–583 (2017). [PubMed: 28256218]

34. Reha-Krantz LJ DNA polymerase proofreading: multiple roles maintain genome stability. Biochim. Biophys. Acta 1804, 1049–1063 (2010). [PubMed: 19545649]

35. Lamichhane R, Berezhna SY, Gill JP, Van der Schans E & Millar DP Dynamics of site switching in DNA polymerase. J. Am. Chem. Soc 135, 4735–4742 (2013). [PubMed: 23409810]

36. Dodd T et al. Polymerization and editing modes of a high-fidelity DNA polymerase are linked by a well-defined path. Nat. Commun 11, 5379 (2020). [PubMed: 33097731]

37. Shamoo Y & Steitz TA Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell 99, 155–166 (1999). [PubMed: 10535734]

38. Fernandez-Leiro R et al. Self-correcting mismatches during high-fidelity DNA replication. Nat. Struct. Mol. Biol 24, 140–143 (2017). [PubMed: 28067916]

39. Dangerfield TL, Kirmizialtin S & Johnson KA Substrate specificity and proposed structure of the proofreading complex of T7 DNA polymerase. J. Biol. Chem 298, 101627 (2022). [PubMed: 35074426]

40. Li Y, Korolev S & Waksman G Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. EMBO J. 17, 7514–7525 (1998). [PubMed: 9857206]

41. Doublie S, Tabor S, Long AM, Richardson CC & Ellenberger T Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature 391, 251–258 (1998). [PubMed: 9440688]

42. Beese LS, Derbyshire V & Steitz TA Structure of DNA polymerase I Klenow fragment bound to duplex DNA. Science 260, 352–355 (1993). [PubMed: 8469987]

43. Franklin MC, Wang J & Steitz TA Structure of the replicating complex of a Pol a family DNA polymerase. Cell 105, 657–667 (2001). [PubMed: 11389835]

44. Fernandez-Leiro R, Conrad J, Scheres SH & Lamers MH Cryo-EM structures of the *E. coli* replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and τ, eLife 4, e11134 (2015). [PubMed: 26499492]

45. Johnson SJ & Beese LS Structures of mismatch replication errors observed in a DNA polymerase. Cell 116, 803–816 (2004). [PubMed: 15035983]

46. Wang W, Hellinga HW & Beese LS Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. Proc. Natl Acad. Sci. USA 108, 17644–17648 (2011). [PubMed: 22006298]

47. Kasiviswanathan R, Longley MJ, Chan SS & Copeland WC Disease mutations in the human mitochondrial DNA polymerase thumb subdomain impart severe defects in mitochondrial DNA replication. J. Biol. Chem 284, 19501–19510 (2009). [PubMed: 19478085]

48. González-Vioque E et al. Association of novel *POLG* mutations and multiple mitochondrial DNA deletions with variable clinical phenotypes in a Spanish population. Arch. Neurol 63, 107–111 (2006). [PubMed: 16401742]

49. Davidzon G et al. Early-onset familial parkinsonism due to *POLG* mutations. Ann. Neurol 59, 859–862 (2006). [PubMed: 16634032]

50. Wong LJ et al. Molecular and clinical genetics of mitochondrial diseases due to *POLG* mutations. Hum. Mutat 29, E150–E172 (2008). [PubMed: 18546365]

51. Lamantea E et al. Mutations of mitochondrial DNA polymerase γA are a frequent cause of autosomal dominant or recessive progressive external ophthalmoplegia. Ann. Neurol 52, 211–219 (2002). [PubMed: 12210792]

52. Phillips J et al. *POLG* mutations presenting as Charcot–Marie–Tooth disease. J. Peripher. Nerv. Syst 24, 213–218 (2019). [PubMed: 30843307]

53. Horvath R et al. Phenotypic spectrum associated with mutations of the mitochondrial polymerase γ gene. Brain 129, 1674–1684 (2006). [PubMed: 16621917]

54. Da Pozzo P et al. Novel *POLG* mutations and variable clinical phenotypes in 13 Italian patients. Neurol. Sci 38, 563–570 (2017). [PubMed: 28130605]

55. Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296 (2017). [PubMed: 28165473]

56. Zivanov J, Nakane T & Scheres SHW Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. IUCrJ 7, 253–267 (2020).

57. Punjani A, Zhang H & Fleet DJ Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat. Methods 17, 1214–1221 (2020). [PubMed: 33257830]

58. Pettersen EF et al. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612 (2004). [PubMed: 15264254]

59. Sanchez-Garcia R et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. Commun. Biol 4, 874 (2021). [PubMed: 34267316]

60. Kaur S et al. Local computational methods to improve the interpretability and analysis of cryo-EM maps. Nat. Commun 12, 1240 (2021). [PubMed: 33623015]

61. Liebschner D et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol 75, 861–877 (2019). [PubMed: 31588918]

62. Cardone G, Heymann JB & Steven AC One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. J. Struct. Biol 184, 226–236 (2013). [PubMed: 23954653]

63. Pettersen EF et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82 (2021). [PubMed: 32881101]

64. Fan L et al. A novel processive mechanism for DNA synthesis revealed by structure, modeling and mutagenesis of the accessory subunit of human mitochondrial DNA polymerase. J. Mol. Biol 358, 1229–1243 (2006). [PubMed: 16574152]

65. Casañal A, Lohkamp B & Emsley P Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. Protein Sci. 29, 1069–1078 (2020). [PubMed: 31730249]

66. Croll TI ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr. D Struct. Biol 74, 519–530 (2018). [PubMed: 29872003]

67. Williams CJ et al. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. 27, 293–315 (2018). [PubMed: 29067766]

68. Pintilie G et al. Measurement of atom resolvability in cryo-EM maps with Q-scores. Nat. Methods 17, 328–334 (2020). [PubMed: 32042190]

69. Sagendorf JM, Markarian N, Berman HM & Rohs R DNAproDB: an expanded database and web-based tool for structural analysis of DNA–protein complexes. Nucleic Acids Res. 48, D277–D287 (2020). [PubMed: 31612957]

70. Pintilie G, Chen DH, Haase-Pettingell CA, King JA & Chiu W Resolution and probabilistic models of components in cryoEM maps of mature P22 bacteriophage. Biophys. J 110, 827–839 (2016). [PubMed: 26743049]
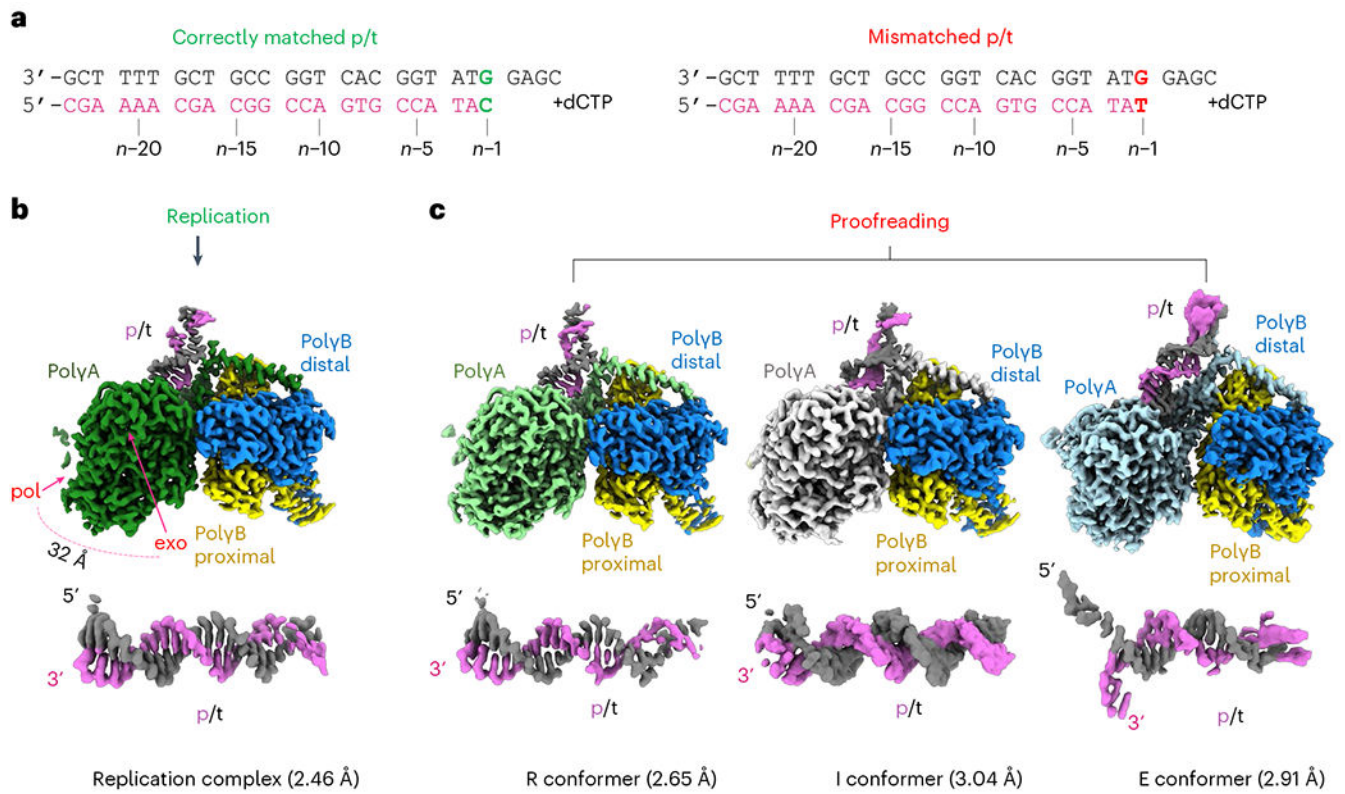
Fig. 1 |. Cryo-EM reconstructions of human mtDNA Polγ ternary complexes.
a, DNA p/t sequences containing a correctly matched G·C and a mismatched G·T nascent base pair, respectively. **b**, Cryo-EM map of the Polγ replication complex with the correctly matched G·C p/t DNA. **c**, Cryo-EM maps of G·T mismatch R, I, and E conformers, respectively. The density of p/t DNA from the corresponding complexes is presented below.
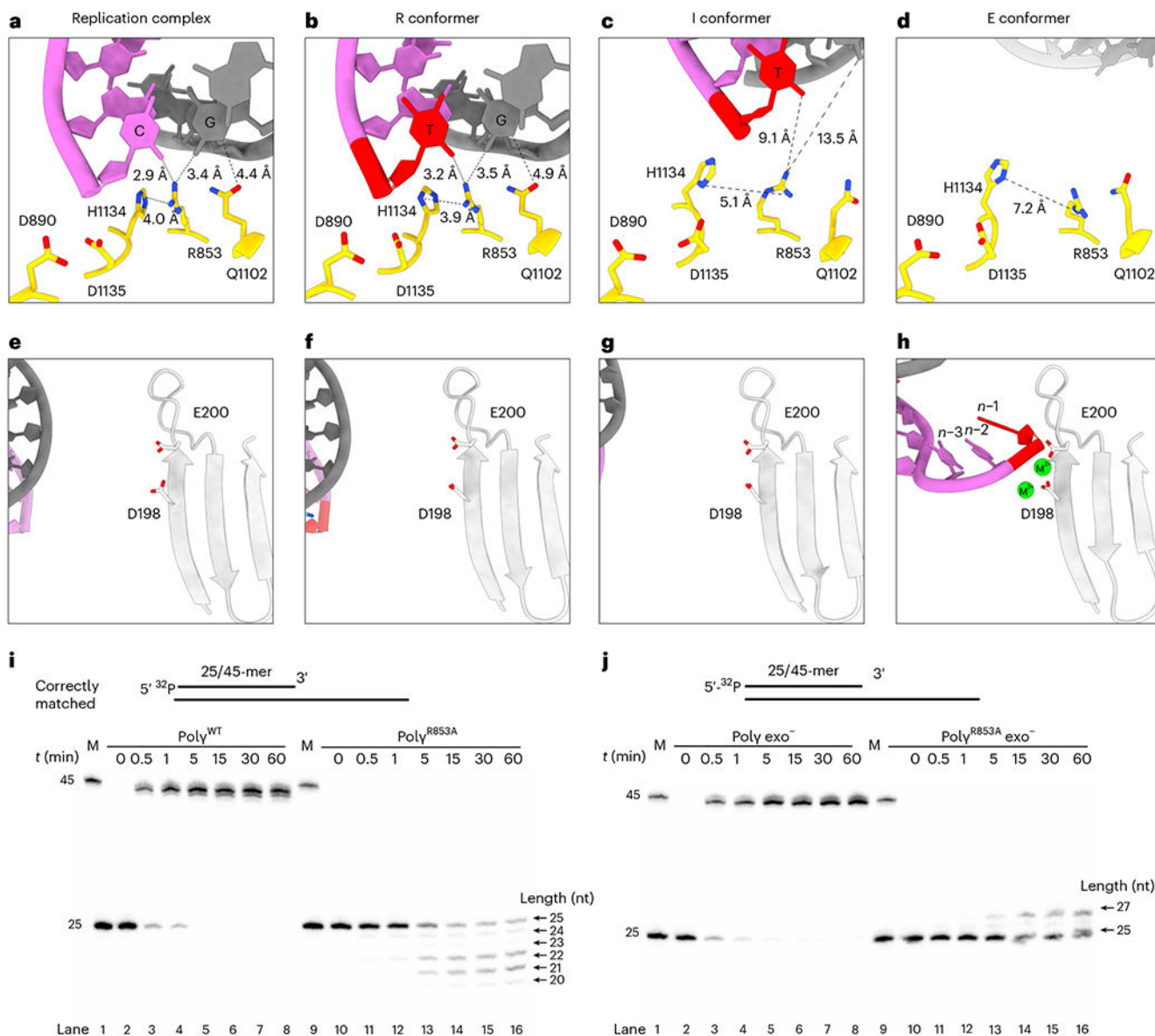
Fig. 2 |. Locations of the primer and template in Polγ replication and proofreading complexes.
The pol (**a–d**) and exo (**e–h**) active sites of Polγ in complex with correctly matched G·C p/t (**a,e**), mismatched G·T p/t in the R conformer (**b,f**), the I conformer (**c,g**), and the E conformer (**d,h**). The location of the 3′ terminus of the primer (pink) in pol and exo site is mutually exclusive, marking the polymerase in replication or proofreading mode. Only the primer in the E conformer is located in the exo site (**e–h**), M^{2+} in **h** denotes the divalent metal ion. The fidelity switch Arg853–Gln1102 coordinates the nascent base pair (n–1) and catalytic loop with W-C and enolic base pairs (**a,b**) but not with non-enolic mismatches in I and E conformers (**c,d**). DNA-synthesis activities of exo-proficient Pol³ wild type (WT) and R853A mutant (**i**) and exo-deficient enzymes (**j**) measured on a correctly matched 5′ ^{32}P 25-nt primer annealed to a 45-nt template in the presence of a dNTP mixture, where M
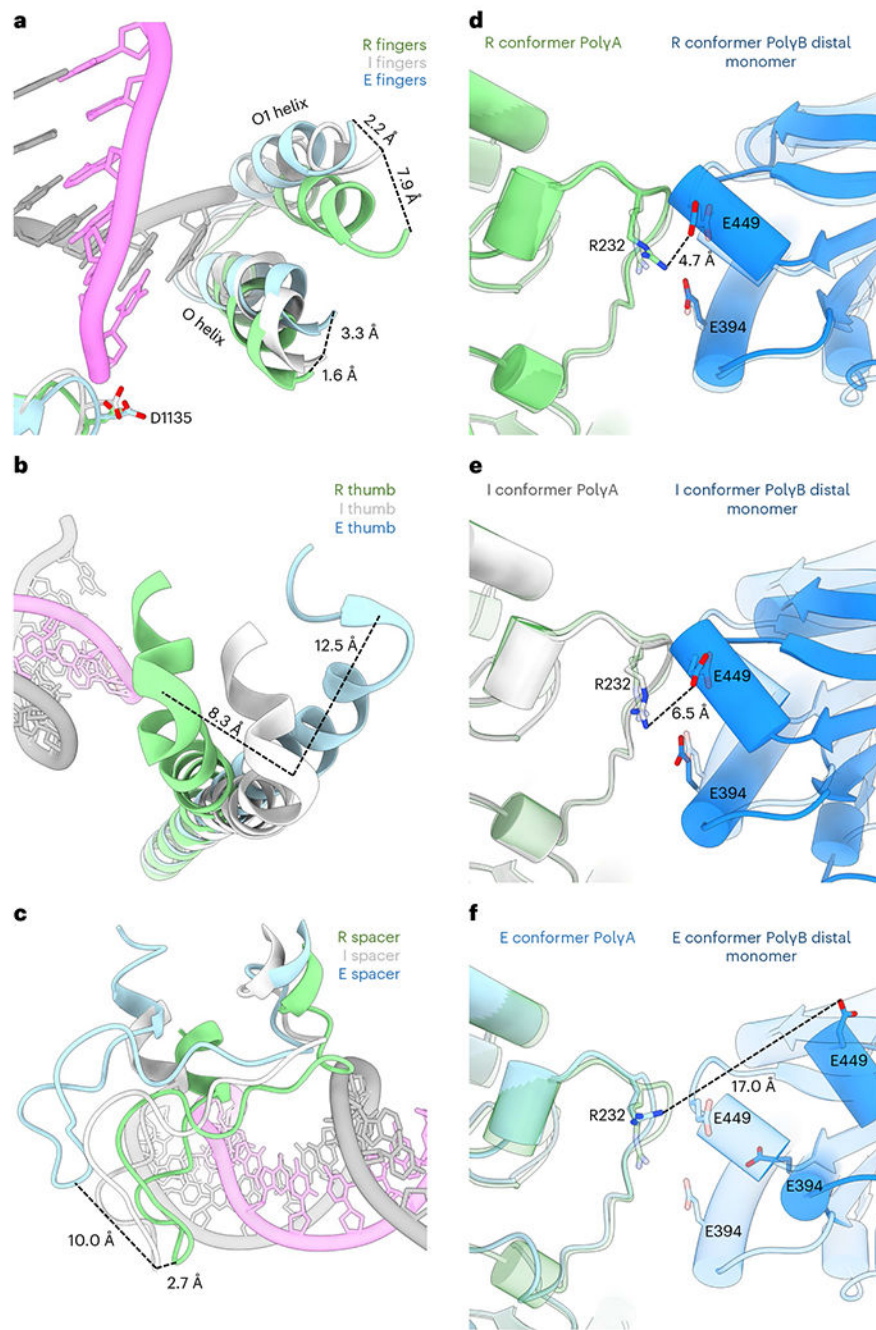
in lane 1 denotes the DNA marker. Gels in **i,j** are representative results from triplicate experiments.

Fig. 3 |. Protein conformational changes associated with primer shuttling.
Overlaid proofreading R, I, and E conformers O and O1 helices of the fingers subdomain
(**a**), the thumb subdomain (**b**), and the spacer domain (**c**). **d–f**, Subunit interaction of the
catalytic subunit PolγA and the accessory subunit PolγB distal monomer of the mismatched
R conformer (**d**), the I conformer (**e**), and the E conformer (**f**) superimposed on the correctly
matched replication complex (transparent).
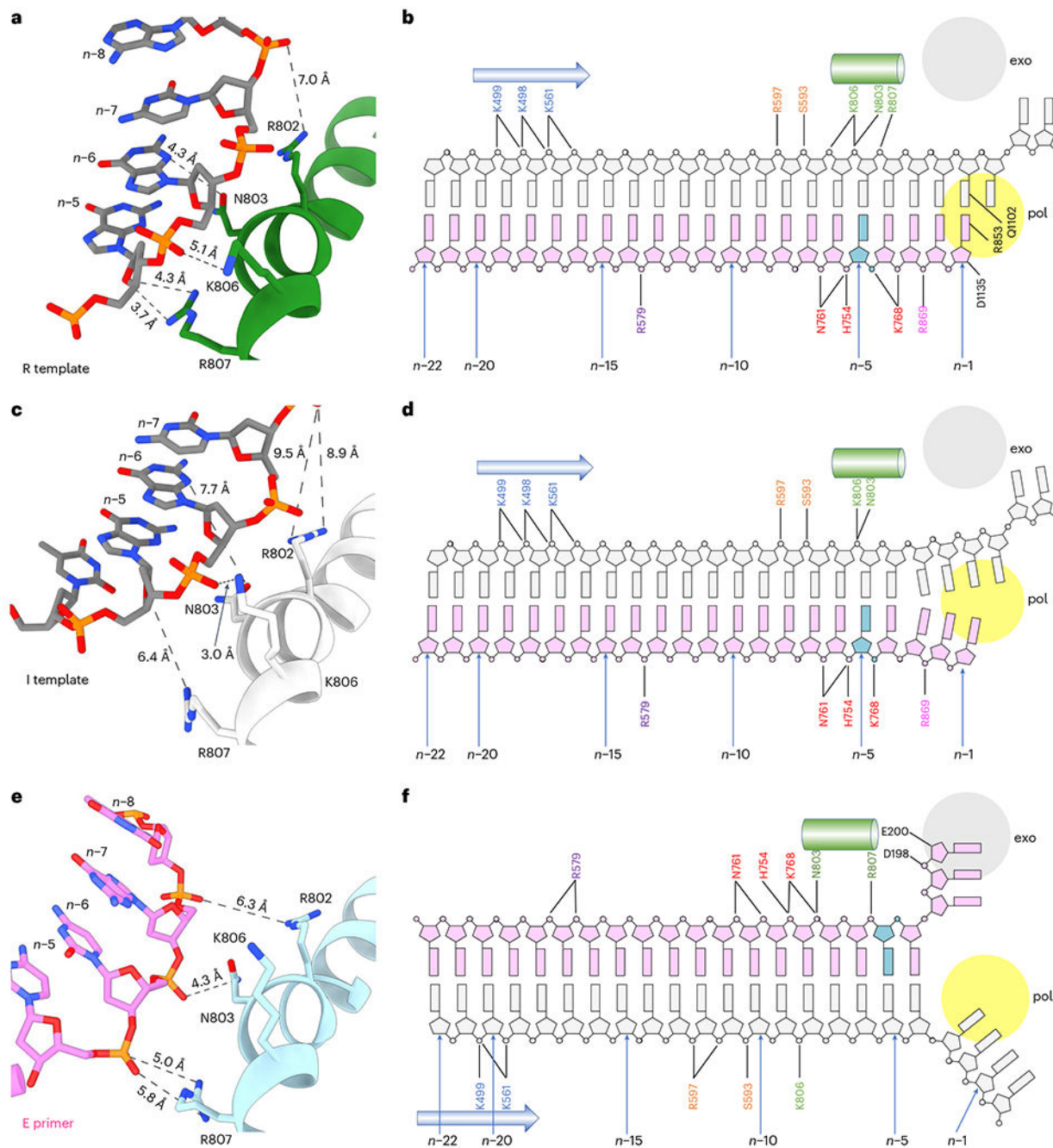
Fig. 4 |. Changes in DNA conformation and interaction with polymerase in the transition from replication to proofreading.

Interaction of the thumb subdomain with the template strand in the G·C replication complex (**a**) and the G·T mismatch I conformer (**c**) but with the primer strand in the E conformer (**e**). **b,d,f**, Schematics of overall protein–DNA contact maps of the three structures. The subset of residues that interact with the DNA backbone is selected for display. Residues colored in black only interact with p/t DNA in one of the structures. The green cylinder represents the portion of the thumb subdomain in **a,c,e,** and the blue arrow represents the spacer domain K

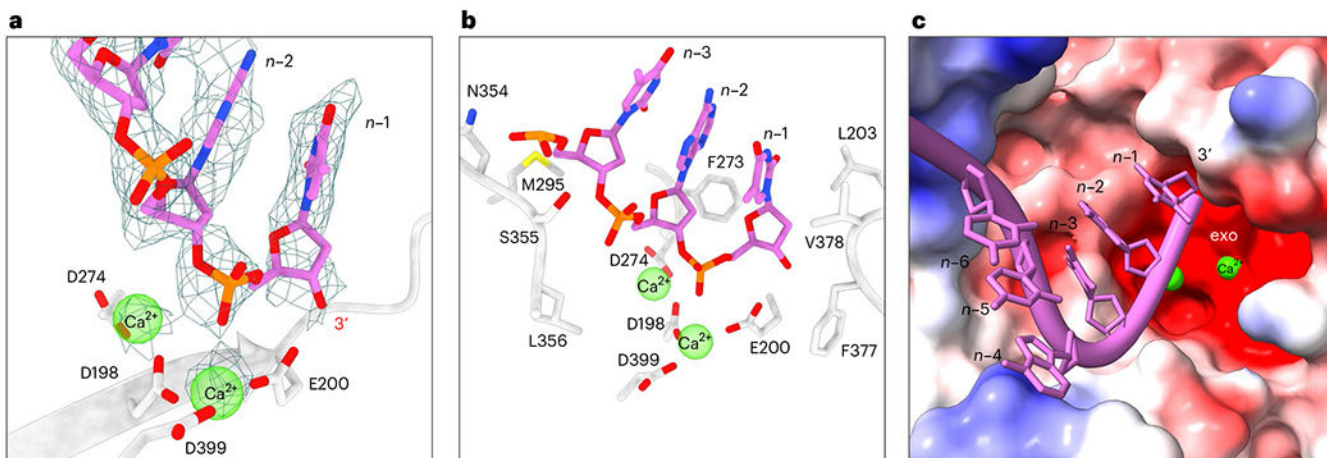tract that interacts with the upstream p/t duplex. The primer and template strands are colored in pink and gray, respectively, and the nucleotide n–5 is colored differently in all primers as a marker for comparison.

Fig. 5 |. DNA mismatch processing.
a, Coordination of the 3′ mismatched primer with Ca^{2+} ions (superimposed with electron density) and catalytic residues Asp198 and Glu200 in the exo site. **b**, Primer stabilization in the exo site by metal coordination as well as acidic and hydrophobic amino acids. **c**, Electrostatic surface representation of the negatively charged exo site surrounding the primer terminus. Ca^{2+} ions are shown as a reference point.
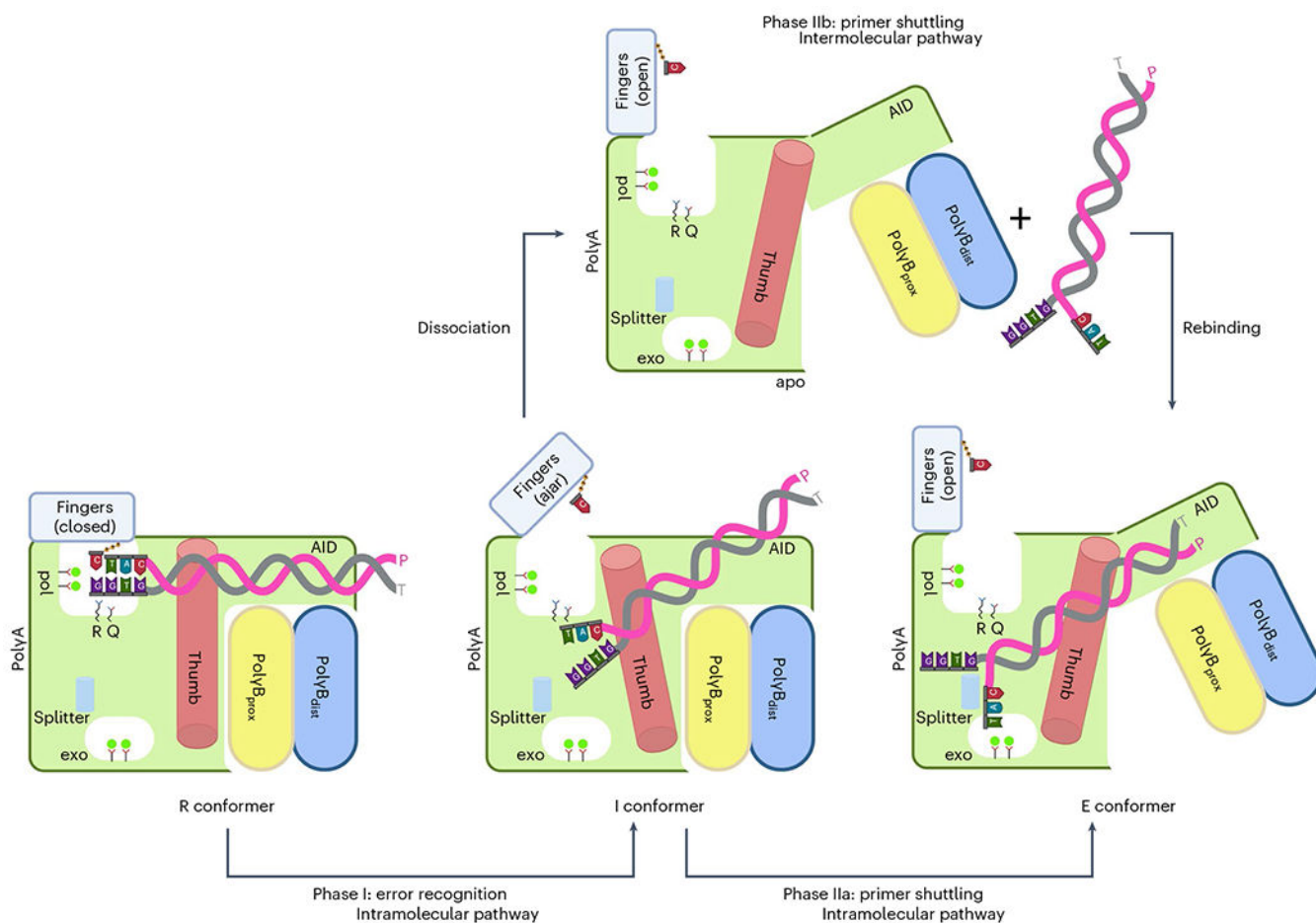
Fig. 6 |. Proposed two-phase Polγ replication–proofreading transition pathways.
Phase I, an R conformer to I conformer intramolecular transition; phase II, I conformer to E conformer transition, which can be intramolecular (phase IIa) or intermolecular (phase IIb) or a combined pathway. In phase I, Polγ detects non-W-C geometry of the nascent base pair in the pol active site via the Arg853–Gln1102 fidelity switch and ejects the 3′ mismatch-containing primer strand out of the pol site, the fingers subdomain opens, and the thumb subdomain rotates. This will result in the I conformer. In phase IIa, p/t stays bound to Pol³ during the primer shuttling process. Fingers and thumb subdomains further open and rotate to situate p/t in the E conformer. The splitter helix stabilizes the split p/tat the fork junction, and the PolγA AID (accessory interacting determinant) subdomain and the PolγB proximal (prox) and distal (dist) monomers rotate outwards as a rigid-body. In phase IIb, p/t is transiently released from Polγ. Three base pairs from the 3′ end of the primer unwind, and Polγ, now in apo form, will bind to the frayed DNA into the E conformer. In either pathway, subunit interaction between the catalytic PolγA and the PolγB distal monomer is decreased, making Polγ less processive during the nucleoside-excision process.

Table 1 |

Cryo-EM data collection, refinement and validation statistics of Polγ replication and proofreading complexes

	G·C replication complex (EMD-27154) (PDB 8D33)	G·T R conformer (EMD-27155) (PDB 8D37)	G·T I conformer (EMD-27163) (PDB 8D3R)	G·T E conformer (EMD-27172) (PDB 8D42)
Data collection and processing				
Magnification	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300
Electron exposure (e−/Å²)	49	44	44	44
Defocus range (μm)	1.5–2.5	1.5–2.5	1.5–2.5	1.5–2.5
Pixel size (Å)	0.425	0.425	0.425	0.425
Symmetry imposed	C1	C1	C1	C1
Number of micrographs				
Collected	9,828	14,122	14,122	14,122
Used	9,697	12,256	12,256	12,256
Initial particle images (no.)	15,053,689	17,463,536	17,463,536	17,463,536
Final particle images (no.)	3,656,462	1,179,973	357,312	1,076,466
Map resolution (Å)				
Corrected, FSC = 0.143	2.46	2.65	3.08	2.91
Refinement				
Initial model used (PDB code)	4ZTZ; 2G4C	4ZTZ; 2G4C	4ZTZ; 2G4C	4ZTZ; 2G4C
Model resolution (Å)	2.6	3.0	3.4	3.1
FSC threshold	0.5	0.5	0.5	0.5
Map-sharpening B factor (Å²)				
Model composition				
Non-hydrogen atoms	15,275	15,317	15,106	15,231
Protein/DNA residues	1,784/46	1,784/48	1,768/45	1,786/43
dCTP/Ca²⁺	1/1	1/1	1/1	1/3
B factors (Å²)				
Protein	42.77	54.66	80.09	67.70
DNA	41.53	59.65	97.59	52.19

	G-C replication complex (EMD-27154) (PDB 8D33)	G-T R conformer (EMD-27155) (PDB 8D37)	G-T I conformer (EMD-27163) (PDB 8D3R)	G-T E conformer (EMD-27172) (PDB 8D42)
Ligand	38.53	46.10	107.67	65.38
R.m.s. deviations				
Bond lengths (Å)	0.002	0.002	0.002	0.002
Bond angles (°)	0.480	0.461	0.548	0.497
Validation				
MolProbity score	1.41	1.47	1.71	1.50
Clashscore	3.94	4.50	6.75	4.85
Poor rotamers (%)	0.39	0.32	0.07	0.84
Cβ outliers (%)	0	0	0	0
Ramachandran plot				
Favored (%)	96.47	96.30	95.18	96.31
Allowed (%)	3.47	3.70	4.82	3.69
Disallowed (%)	0.06	0	0	0
Q score				
Protein	0.681	0.584	0.585	0.611
DNA	0.610	0.575	0.369	0.493
Ca^{2+}	0.867	0.831	0.746	0.832
dCTP	0.694	0.698	0.580	0.654