Published in final edited form as:

Nat Struct Mol Biol. 2017 February ; 24(2): 140–143. doi:10.1038/nsmb.3348.

Self-correcting mismatches during high-fidelity DNA replication

Rafael Fernandez-Leiro¹, Julian Conrad^{1,2}, Ji-Chun Yang¹, Stefan M. V. Freund¹, Sjors H. W. Scheres¹, and Meindert H. Lamers¹

¹MRC laboratory of Molecular Biology, Cambridge, United Kingdom

Abstract

Faithful DNA replication is essential to all forms of life and depends on the action of 3'-5' exonucleases that remove misincorporated nucleotides from the newly synthesized strand. However, how the DNA is transferred from the polymerase to exonuclease active site is not known. Here we present the cryo-EM structure of the editing mode of the catalytic core of the *E. coli* replisome, revealing a dramatic distortion of the DNA whereby the polymerase thumb domain acts as a wedge that separates the two DNA strands. Importantly, NMR analysis of the DNA substrate shows that the presence of a mismatch increases the fraying of the DNA, enabling it to reach the exonuclease active site. It is therefore the mismatch that corrects itself, whereas the exonuclease subunit plays a passive role. Hence, our work provides unique insights into high fidelity replication and establishes a new paradigm for correction of misincorporated nucleotides.

Introduction

High-fidelity DNA polymerases such as PolIII α , the α subunit of the *E. coli* replisome, employ a narrow active site to prevent the insertion of the wrong nucleotide¹. In the rare occasions that an incorrect nucleotide is incorporated, the distorted geometry of the mismatched base pair prevents further extension of the DNA strand². Therefore, to continue DNA synthesis, all high-fidelity DNA polymerases contain a 3'-5' exonuclease that removes the misincorporated nucleotide.

The *E. coli* replicative DNA polymerase Pol III α uses a separate exonuclease, ϵ , for the removal of misincorporated nucleotides³. In turn, the exonuclease binds the small protein θ , with no known function⁴. Together, PolIII α , ϵ and θ are termed 'core'⁵. PolIII α furthermore binds to the β sliding clamp⁶ that provides processivity, and to the C-terminal domain of the clamp loader subunit τ that is required for the repeated release of the polymerase at the lagging strand^{7,8}. Once assembled, the pentameric polIII α -exonuclease- θ -clamp- τ complex catalyzes DNA synthesis at speeds of up to 1000 nucleotides/second^{9,10} and with

Correspondence should be addressed to M.H.L. (mlamers@mrcmb.cam.ac.uk).

²Present address: Science for Life Laboratory, Solna, Sweden

Author Contributions

R.F.L. and M.H.L. designed and directed experiments. R.F.L. and J.C. collected and processed cryo-EM data. S.H.W.S. assisted in data processing. R.F.L. purified proteins and performed biochemical assays. S.M.F. and J.C.Y. collected and processed NMR data. M.H.L. and R.F.L. wrote the manuscript.

Competing Financial Interests

The authors declare no competing financial interests.

an error rate of ~1 per million¹¹. Because the exonuclease active site is ~60 Å away from the polymerase active site¹², it is not clear how the DNA moves to the exonuclease site when a mismatch is encountered.

To better understand how the mismatches are removed during high fidelity DNA replication, we have determined the cryo-electron microscopy (cryo-EM) structure of the *E. coli* PolIII α -exonuclease- θ -clamp complex bound to a mismatched DNA substrate. The structure reveals a dramatic distortion of the DNA supported by the polymerase thumb domain that acts as a wedge that separates the two DNA strands. In addition, the DNA is kinked and pushed into the inner rim of the β sliding clamp that stabilizes the distorted DNA conformation. Importantly, NMR analysis of the DNA substrate shows that the presence of a mismatch increases the fraying of the DNA by one base pair. This enables the DNA to adopt the distorted conformation and reach the exonuclease active site. It is therefore the mismatch that corrects itself, whereas the exonuclease subunit plays a passive role and only serves to remove the terminal nucleotide and does not "proofread". As the mismatch induced fraying of the DNA termini is the same for all forms of life, it is possible that the self correction of the DNA may be a universal mechanism of high-fidelity DNA polymerases.

Results

Overall structure of the editing complex

We determined the cryo-EM structure of the *E. coli* PolIII α -exonuclease- θ -clamp editing complex bound to a mismatched DNA substrate to a resolution of 6.7 Å (Fig. 1, Supplementary Video 1, Supplementary Fig. 1, and Supplementary Table 1). In brief, the polymerase is tethered to the clamp via its internal binding motif (residues 920-924) and indirectly via the exonuclease that is located between the clamp and polymerase thumb domain. θ is bound to the exonuclease, facing away from the DNA, similar to the structure of θ bound to the isolated catalytic domain of the exonuclease¹³. Although present in the protein complex, the polymerase tail (residues 925-1160) and τ_{500} (the C-terminal domain of the τ subunit: residues 500-643) are not visible in this structure, indicating that they are flexible in the editing mode.

Previously, we have determined the cryo-EM structure of the PolIII α -exonuclease-clamp complex in the DNA synthesis mode¹². Between the DNA synthesis and editing mode, there is surprisingly little movement in the protein subunits, except for a ~6 Å outward movement of the polymerase thumb domain, and a ~6 Å inward movement of the exonuclease towards the DNA (Fig. 1d). In contrast, the DNA undergoes a dramatic rearrangement, characterized by three distinct movements: a ~90° co-axial screw rotation towards the clamp, a 30° in-plane tilt, and the fraying of the DNA end by three base pairs (Fig. 1c and Supplementary Video 2). As a result, the 3' terminus of the primer strand travels ~55 Å from the polymerase to the exonuclease active site.

Multiple DNA interactions stabilize the distorted DNA conformation

The distorted conformation of the DNA appears to be supported by four distinct interactions. First, the thumb domain of the polymerase functions as a wedge that separates the two DNA

strands. The last base pair before the junction of the DNA substrate comes close to tyrosine 453 of the thumb domain (Fig. 2a). Aromatic residues are frequently used to stack onto bases and stabilize the interaction between protein and DNA¹⁴, suggesting that tyrosine 453 may play a similar role. Indeed, an aromatic residue is found conserved at this position in the replicative C-family polymerases, but not in the predicted¹⁵ error prone C-family polymerase DnaE2 (Supplementary Fig. 2a). The separation of the two DNA strands by the thumb domain is unique to the C-family of DNA polymerases, as A-family DNA polymerases such as *E. coli* Pol I, and B-family polymerases, such as *E. coli* Pol II and the eukaryotic polymerase δ and ϵ , use a very different mechanism to separate the primer and template strand¹⁶ (Supplementary Fig. 3). Second, the template strand occupies the polymerase active site in a non-canonical manner, extending past the polymerase active site and the binding site of the incoming nucleotide (Fig. 2b). This position of the template strand is incompatible with DNA synthesis and appears to serve to stabilize the open conformation of the DNA. Third, the DNA no longer travels through the center of the clamp as observed in the DNA synthesis mode (Fig. 2c). Instead the 30° tilt of the DNA pushes it into the inner rim of the clamp that appears to act as a "lock" on the DNA that prevents the protein from sliding. Here, clear density shows an interaction between the DNA backbone and loops in the inner rim of the clamp, likely involving glutamine 143, histidine 148, and arginine 197 of the clamp (Fig. 2c). Arginine 197 also interacts with the DNA during DNA synthesis, whereas glutamine 143 and histidine 148 appear unique to the editing mode. Interestingly, the angle of the DNA, as well as the contacts between clamp and DNA are distinct from those observed in the crystal structure of the free *E. coli* β clamp and DNA¹⁷ (Fig. 2f). Finally, the exonuclease moves inwards by ~6 Å (Fig. 1d) and interacts with the DNA backbone via a loop (residues 137 to 144) close to the exonuclease active site (Fig. 2d). Two basic residues in this loop, lysine 141 and arginine 142, are likely to be involved in this interaction. These two residues are conserved in ϵ homologs from α , β , γ *Proteobacteria* (which include *E. coli*) that use a separate exonuclease for removal of misincorporated nucleotides (Supplementary Fig. 2b) but not in ϵ homologs from bacterial species that use a Polymerase and Histidinol Phosphatase (PHP) domain¹⁸ as their main replicative exonuclease^{19,20}. Together, these interactions stabilize the distorted DNA conformation that enables the primer strand to reach the exonuclease active site (Fig. 2e). The position of the terminal nucleotides in the exonuclease active site is similar to that observed in the crystal structure of Pol I¹⁶ and the Pol II homolog PolB21 even though these polymerases are structurally dissimilar and use very different mechanisms to separate the mismatched primer strand from template strand (Supplementary Fig. 3d).

Next, we used an exonuclease assay to validate the interactions described above (Fig. 2g). Both the mutation of the polymerase thumb contact (Pol^{Y453A}) as well as the mutation of the clamp contacts (clamp^{Q143A/H148A}) result in reduced exonuclease activity, while a combination of the two mutants reduces exonuclease activity even further. Mutation of either lysine 141 (exo^{K141A}) or arginine 142 (exo^{R142A}) in the exonuclease both strongly reduce the exonuclease activity, while deletion of the single stranded overhang in the template strand almost completely abolishes activity (Fig. 2h).

The terminal mismatch increases DNA fraying

The moderate changes in the protein part of the complex contrast with the dramatic rearrangement of the DNA. We therefore wondered if the distortion of the DNA is not induced by the protein, but rather by the DNA itself. To explore this further we used nuclear magnetic resonance (NMR) to determine the solvent exchange rates of the imino protons located between complementary bases²² (see Methods and Supplementary Fig. 4). The terminal base pair of the DNA duplex is highly dynamic and rapidly exchanges with the solvent²³ and therefore could not be assigned. Also the second base pair shows a much faster exchange rate compared to the other base pairs at positions 3, 4, 7, 10, and 11 (Fig. 3a,b). The presence of a terminal mismatch (C:T) dramatically increases the exchange rate of the third base pair, whereas the subsequent base pairs remain unchanged. Removal of the mismatched nucleotide (C:–) restores the fraying of the DNA duplex back to two base pairs. Hence, on a matched substrate the last two base pairs will fray, while three base pairs will fray on a mismatched substrate. This is in agreement with UV DNA-melting experiments where both the presence of a 3' terminal mismatch or the absence of a terminal base pair decrease the melting temperature (T_m) by $\sim 10^\circ\text{C}$ (Fig. 3c). The three base pair fraying of the mismatched substrate matches perfectly with the three unpaired base pairs observed in the cryo-EM structure suggesting that only a mismatched DNA substrate can reach the exonuclease active site.

The increased fraying of the mismatched DNA duplex also provides a simple mechanism for how the primer strand is returned to the polymerase active site. Once the mismatch is removed, the primer terminus can no longer reach the exonuclease active site, preventing processive exonuclease activity. Indeed, the non-processive action of the exonuclease is evident in a primer extension assay, where the extension from a mismatched substrate only occurs after removal of the mismatch. Here the activity of the exonuclease does not extend past the mismatch, as no additional bands are apparent below that position (Fig. 3d). The non-processive behavior of the exonuclease is furthermore consistent with its low affinity for DNA with a measured K_d of $\sim 30\ \mu\text{M}$ (Fig. 3e) in agreement with earlier findings²⁴. In contrast, the core-clamp complex binds DNA ~ 200 -fold better, with a K_d of $\sim 0.15\ \mu\text{M}$ (Fig. 3e).

Discussion

Our work reveals that in the editing mode of the bacterial replicase, the DNA undergoes a dramatic transformation that is stabilized by a unique set of interactions. This distortion enables the mismatched nucleotide to reach the exonuclease that is strategically placed far away such that only a three base pair frayed substrate can reach its active site. Our work furthermore suggests that it is the DNA that corrects itself, by making use of an alternative binding mode within the complex that can only be reached with a mismatch at the primer terminus. In contrast, the exonuclease that has historically been termed the "proofreader" plays a passive role and only serves to remove the terminal nucleotide of the presented DNA primer strand. Importantly, a similar three base pair fraying is also observed in the A-family and B-family DNA polymerases, even though these are structurally divergent polymerases and their exonuclease domains are located in very different positions (Supplementary Fig.

3). This suggests that the three base pair fraying induced self correction of the DNA may be a universal mechanism of high-fidelity DNA polymerases.

Accession codes

The cryo-EM maps have been deposited in the Electron Microscopy Data Bank under accession code EMD-4141 and EMD-4142; the corresponding atomic model has been deposited in the Protein Data Bank under accession code 5M1S.

Data accessibility

Cryo-EM density maps and associated atomic model have been deposited in the EMDB (EMD-4141 and EMD-4142) and in the Protein Data Bank (PDB 5M1S), respectively. Source data for Figure 2 and 3 are available with the paper online (Supplementary Source Data Set 1 and 2 respectively). Other data are available from the corresponding author upon request.

Online Methods

Protein expression and purification

All proteins were expressed in *E. coli* (DE3) BL21 and purified as described before^{12,25}. The θ subunit was purified as described²⁶.

Cryo-EM

Sample preparation, cryo-EM data collection and data analysis were performed in a similar manner to the complex in DNA polymerase mode¹², with the following exceptions. The DNA oligo was shortened by 4 base pairs and the 3' terminal base of the primer strand was altered to a T to create a C:T mismatch. A phosphorothioate bond was used between the last two bases of the primer strand to prevent hydrolysis by the exonuclease (see (Supplementary Table 2)). All protein subunits were mixed in equimolar ratio and purified by gel filtration. A ten-fold excess of the DNA was added before preparation of the cryo-EM sample grid. Data were collected using a Titan Krios electron microscope (FEI) operated at 300 kV equipped with a K2 Summit direct electron detector (Gatan) mounted after a Gatan Imaging Filter (GIF) using a 20eV slit to remove inelastic scattered electrons. 20 frame image stacks were collected in electron counting mode using a flux of $2 \text{ e}/\text{\AA}^2/\text{sec}$ and a total dose of $40 \text{ e}/\text{\AA}^2$. Frames were aligned and averaged using MOTIONCORR²⁷. Contrast transfer function parameters were calculated using Gctf²⁸. All subsequent particle picking and data processing was done in a pre-release version of Relion-2.29, using the polymerase mode structure¹² as initial reference. For details on the classification strategy see (Supplementary Fig. 1. Particle-based movement correction and per-frame B-factor weighting³⁰ was performed in Relion-2. FSC-curves were corrected for the convolution effects of a soft mask using high-resolution noise-substitution³¹ and reported resolutions are based on the gold-standard FSC-0.143 criterion³². Maps were sharpened using an automatically estimated negative B-factor³³. It was noted that between different 3D classes, the relative orientation of the polymerase, exonuclease and clamp varied to a limited degree (see (Supplementary Video 3)). This conformational heterogeneity hampers the accurate alignment of the whole-particle images and therefore reduces the number of particles that can be used for the final

reconstruction. In order to improve the resolution we used signal subtraction and focused classification³⁴ to independently align the polymerase, exonuclease, θ , and DNA (Supplementary Fig. 1c-e). With this procedure we improved the resolution of the polymerase-exonuclease- θ -DNA region of the complex to 6.1 Å. Local resolution maps were calculated using RESMAP³⁵. The cryo-EM structure of the PolIII α -clamp-exonuclease complex in the polymerase mode (PDB code: 5FKW)¹² was used as a starting model, and the NMR structure of θ bound to the ϵ catalytic domain (PDB code: 2XY8)¹³ was used to place θ into the cryo-EM map. The positioning of the primer strand in the exonuclease active site was guided by the crystal structure of the isolated exonuclease bound to a di-nucleotide³⁶, as well as the crystal structures of the exonuclease domains from DNA Pol I from *E. coli*¹⁶ and Pol B from *Pyrococcus abyssi*²¹ that belong to the same family of DEDDh exonucleases. Details of the DNA model and maps are shown in Supplementary Fig. 1g. The model was manually adjusted in Coot³⁷ and geometry of the protein optimized in REFMAC³⁸ using DNA-specific restraints generated in LibG38.

Comparison of A, B, C-type DNA polymerases

The following structures were used for the comparison of DNA polymerases in DNA synthesis and editing mode: A-type (Pol I-like): 1QTM39, 1KLN16, B-type (Pol II-like): 3K5740, 4FLW21, C-type (Pol III-like): 5FKW12, and this work. Structures in polymerase mode were manually aligned using the DNA as a guide. Structures in the editing mode were subsequently aligned to their respective structures in the polymerase mode. D357 in Pol I (shown in white sticks) was mutated to alanine in the structure and was mutated back to the original aspartate for clarity.

NMR spectroscopy

For NMR studies, a matched 20-mer (NMR-20/20C:G), a 20-mer with a C:T terminal mismatch (NMR-20/20C:T), and a matched 20/19-mer lacking the 3' terminal base on the primer strand (NMR-20/19C:-) were used (Supplementary Table 2). The DNA sequences were similar to that of the DNA oligo used for cryo-EM with omission of the single stranded overhang. DNA substrates were annealed, purified by gel filtration over a Superdex 75 3.2/30 column in 10 mM sodium phosphate pH 7.0 and 200 mM NaCl. NMR samples were of a final DNA concentration of 100 μ M in the same buffer. All data were acquired on a Bruker Avance 800MHz spectrometer with TCI cryoprobe (Bruker Inc, Billerica). Sequential assignments of imino proton resonances for the three substrates were achieved via through-space contacts observed in 2D NOESY spectra collected at 278K. At this temperature, solvent exchange is minimized and NOE_{HN-HN} (through-space) cross-peaks between imino protons at a mixing time of 300 ms are well resolved. Datasets acquired at 288, 298 and 310 K allowed to transfer obtained assignments within this temperature range. The temperature-dependent substantial line-broadening of terminal imino proton resonances immediately indicated a substantial increase in their exchange with the solvent. At 310 K, based on a 2D NOESY experiment with a mixing time of 130 ms, only very faint NOE cross-peaks were observed, indicating that solvent exchange is dominating the decay of the diagonal peaks. The decay of the diagonal peaks for the imino protons at position 2, 3, 4, 7, 10, and 11 was measured with a set of ten 2D NOESY experiments collected at 310 K with mixing times ranging from 5-130 ms. Apparent solvent exchange rates were based on the

diagonal peak decay measured according to 22 and analyzed using Kaleidagraph (Synergy Software). Apparent exchange rates were normalized to that of the internal base pair 11, which shows the lowest exchange rate. Also, the exchange rate of this base pair was similar in all three DNA substrates. The assumption was made that the dominant contribution to this decay is from the fast solvent exchange process. The required R_1 rates were obtained from non-selective saturation recovery experiments⁴¹ with a variable delay ranging from 25-250 ms.

UV melting

DNA substrates UV-16/11^{C:G}, UV-16/11^{C:T} and UV-16/11^{C:G} (Supplementary Table 2) were annealed at a final concentration of 5 μ M in a buffer containing 10 mM Tris pH 8, 150 mM NaCl and 5 mM MgCl₂. UV absorbance at 260 nm was recorded while performing a 1°C/min temperature ramp from 20 to 80°C using a Cary Varian 6000i spectrophotometer. Data was normalized to minimum and maximum values for each dataset. A Boltzmann sigmoidal curve was calculated to fit the normalized data using Graphpad Prism.

Exonuclease activity assay

Detection of exonuclease activity was performed using DNA substrates 37/26^{C:T} and 26/26^{C:T} (Supplementary Table 2) with a 5' fluorescein (FAM) label in the primer strand. Reactions (10 μ l) contained 5 nM DNA, 10 μ M PolIII α , clamp (dimer), exonuclease (ϵ), and θ in the a buffer containing 20 mM Tris pH 7.5, 50 mM Potassium Glutamate, 8 mM MgAcetate, 2 mM DTT and 30 μ g/ml BSA. Reactions were incubated at 25°C and quenched with 35 mM EDTA in 65% formamide at different times (as stated in figures). Samples were separated on a 20% acrylamide denaturing urea-PAGE with 6M urea gel for 70 minutes at 30W. Primer strand was visualized on a Typhoon Imager (GE).

Primer extension assay

Assays were performed as described for the exonuclease activity assays using DNA substrates 37/26^{G:C} and 37/26^{G:A} (Supplementary Table 2) with a 5' fluorescein (FAM) label in the primer strand and in the presence of 200 μ M dNTP in the reaction mixture.

Fluorescence anisotropy

DNA binding studies were performed using 5' fluorescein labeled DNA substrates 37/25^{G:C} (see Supplementary Table 2). Reactions (20 μ l) contained 2.5 nM DNA and of either PolIII α -exonuclease- θ -clamp (2 nM - 10 μ M) or exonuclease- θ (2nM – 95 μ M) in a buffer containing 20 mM Tris pH 7.5, 50 mM Potassium Glutamate, 8 mM MgAcetate, 2 mM DTT and 30 μ g/ml BSA. Reactions were incubated at 25°C for 5 minutes and measured in a PHERAstar plate reader (BMG LABTECH). Data were normalized to the maximum and minimum anisotropy values and fitted to a single site binding model using nonlinear regression in Graphpad Prism.

Sequence alignment. Non-redundant protein sequences for PolIII α and ϵ homologs were retrieved from the MPI Bioinformatics Toolkit⁴² and aligned using MAFFT⁴³. Sequence logos were generated using Weblogo 344.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank David Neuhaus for suggestions and David Neuhaus and Roger Williams for reading of the manuscript. This work was supported by the UK Medical Research Council through grants U105197143 to MHL and MC_UP_A025_1013 to SHWS.

References

1. Johnson KA. The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta*. 2010; 1804:1041–1048. [PubMed: 20079883]
2. Johnson SJ, Beese LS. Structures of mismatch replication errors observed in a DNA polymerase. *Cell*. 2004; 116:803–816. [PubMed: 15035983]
3. Scheuermann R, Tam S, Burgers PM, Lu C, Echols H. Identification of the epsilon-subunit of *Escherichia coli* DNA polymerase III holoenzyme as the dnaQ gene product: a fidelity subunit for DNA replication. *Proc Natl Acad Sci USA*. 1983; 80:7085–7089. [PubMed: 6359162]
4. Studwell-Vaughan PS, O'Donnell M. DNA polymerase III accessory proteins. V. Theta encoded by holE. *J Biol Chem*. 1993; 268:11785–11791. [PubMed: 8505306]
5. McHenry CS, Crow W. DNA polymerase III of *Escherichia coli*. Purification and identification of subunits. *J Biol Chem*. 1979; 254:1748–1753. [PubMed: 368075]
6. Stukenberg PT, Studwell-Vaughan PS, O'Donnell M. Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J Biol Chem*. 1991; 266:11328–11334. [PubMed: 2040637]
7. Leu FP, Georgescu R, O'Donnell M. Mechanism of the *E. coli* tau processivity switch during lagging-strand synthesis. *Mol Cell*. 2003; 11:315–327. [PubMed: 12620221]
8. Georgescu RE, et al. Mechanism of polymerase collision release from sliding clamps on the lagging strand. *EMBO J*. 2009; 28:2981–2991. [PubMed: 19696739]
9. Mok M, Mariani KJ. The *Escherichia coli* preprimosome and DNA B helicase can form replication forks that move at the same rate. *J Biol Chem*. 1987; 262:16644–16654. [PubMed: 2824502]
10. McInerney P, Johnson A, Katz F, O'Donnell M. Characterization of a triple DNA polymerase replisome. *Mol Cell*. 2007; 27:527–538. [PubMed: 17707226]
11. Bloom L, et al. Fidelity of *Escherichia coli* DNA polymerase III holoenzyme - The effects of beta, gamma complex processivity proteins and epsilon proofreading exonuclease on nucleotide misincorporation efficiencies. *J Biol Chem*. 1997; 272:27919–27930. [PubMed: 9346941]
12. 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*. 2015; 4:459.
13. Schmitz C, Bonvin AMJJ. Protein-protein HADDOCK using exclusively pseudocontact shifts. *J Biomol NMR*. 2011; 50:263–266. [PubMed: 21626213]
14. Wilson KA, Kellie JL, Wetmore SD. DNA-protein π -interactions in nature: abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar. *Nucleic Acids Res*. 2014; 42:6726–6741. [PubMed: 24744240]
15. Boshoff HIM, Reed MB, Barry CE, Mizrahi V. DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell*. 2003; 113:183–193. [PubMed: 12705867]
16. Beese LS, Derbyshire V, Steitz TA. Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science*. 1993; 260:352–355. [PubMed: 8469987]
17. Georgescu RE, et al. Structure of a sliding clamp on DNA. *Cell*. 2008; 132:43–54. [PubMed: 18191219]
18. Aravind L, Koonin EV. Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res*. 1998; 26:3746–3752. [PubMed: 9685491]

19. Barros T, et al. A structural role for the PHP domain in *E. coli* DNA polymerase III. *BMC Struct Biol.* 2013; 13:8. [PubMed: 23672456]
20. Rock JM, et al. DNA replication fidelity in *Mycobacterium tuberculosis* is mediated by an ancestral prokaryotic proofreader. *Nat Genet.* 2015; 47:677–681. [PubMed: 25894501]
21. Gouge J, Ralec C, Henneke G, Delarue M. Molecular recognition of canonical and deaminated bases by *P. abyssi* family B DNA polymerase. *J Mol Biol.* 2012; 423:315–336. [PubMed: 22902479]
22. Dobson CM, Lian L-Y, Redfield C, Topping KD. Measurement of hydrogen exchange rates using 2D NMR spectroscopy. *Journal of Magnetic Resonance.* 1986; 69:201–209.
23. Leijon M, Gräslund A. Effects of sequence and length on imino proton exchange and base pair opening kinetics in DNA oligonucleotide duplexes. *Nucleic Acids Res.* 1992; 20:5339–5343. [PubMed: 1331987]
24. Miller H, Perrino FW. Kinetic mechanism of the 3'→5' proofreading exonuclease of DNA polymerase III. Analysis by steady state and pre-steady state methods. *Biochemistry.* 1996; 35:12919–12925. [PubMed: 8841137]
25. Toste Rêgo A, Holding AN, Kent H, Lamers MH. Architecture of the Pol III-clamp-exonuclease complex reveals key roles of the exonuclease subunit in processive DNA synthesis and repair. *EMBO J.* 2013; 32:1334–1343. [PubMed: 23549287]
26. Keniry MA, et al. Structure of the theta subunit of *Escherichia coli* DNA polymerase III in complex with the epsilon subunit. *J Bacteriol.* 2006; 188:4464–4473. [PubMed: 16740953]
27. Li X, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods.* 2013; 10:584–590. [PubMed: 23644547]
28. Zhang K. Gctf: Real-time CTF determination and correction. *J Struct Biol.* 2016; 193:1–12. [PubMed: 26592709]
29. Scheres SHW. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol.* 2012; 180:519–530. [PubMed: 23000701]
30. Scheres SH. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife.* 2014; 3:e03665. [PubMed: 25122622]
31. Chen S, et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy.* 2013; 135:24–35. [PubMed: 23872039]
32. Scheres SHW, Chen S. Prevention of overfitting in cryo-EM structure determination. *Nat Methods.* 2012; 9:853–854. [PubMed: 22842542]
33. Rosenthal PB, Henderson R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol.* 2003; 333:721–745. [PubMed: 14568533]
34. Bai X-C, Rajendra E, Yang G, Shi Y, Scheres SHW. Sampling the conformational space of the catalytic subunit of human γ -secretase. *Elife.* 2015; 4:1485.
35. Kucukelbir A, Sigworth FJ, Tagare HD. Quantifying the local resolution of cryo-EM density maps. *Nat Methods.* 2014; 11:63–65. [PubMed: 24213166]
36. Hamdan S, Carr PD, Brown SE, Ollis DL, Dixon NE. Structural basis for proofreading during replication of the *Escherichia coli* chromosome. *Structure.* 2002; 10:535–546. [PubMed: 11937058]
37. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:486–501. [PubMed: 20383002]
38. Brown A, et al. Tools for macromolecular model building and refinement into electron cryomicroscopy reconstructions. *Acta Crystallogr D Biol Crystallogr.* 2015; 71:136–153. [PubMed: 25615868]
39. Li Y, Mitaxov V, Waksman G. Structure-based design of Taq DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc Natl Acad Sci USA.* 1999; 96:9491–9496. [PubMed: 10449720]
40. Wang F, Yang W. Structural insight into translesion synthesis by DNA Pol II. *Cell.* 2009; 139:1279–1289. [PubMed: 20064374]

41. Markley JL, Horsley WJ, Klein MP. Spin-Lattice Relaxation Measurements in Slowly Relaxing Complex Spectra. *The Journal of Chemical Physics*. 1971; 55:3604–3605.
42. Alva V, Nam S-Z, Söding J, Lupas AN. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res*. 2016; 44:W410–W415. [PubMed: 27131380]
43. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013; 30:772–780. [PubMed: 23329690]
44. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. *Genome Res*. 2004; 14:1188–1190. [PubMed: 15173120]

Figure 1.

Cryo-EM structure of the *E. coli* PolIII α -clamp-exonuclease- θ DNA complex in editing mode. (a) Cryo-EM map of the PolIII α , clamp, exonuclease and θ complex bound to a mismatched DNA substrate. Colors indicate the position of the different proteins. (b) Cartoon representation of PolIII α , clamp, exonuclease, θ and DNA fitted into the cryo-EM map. See also Supplementary Video 1. (c) Comparison of the DNA conformation in polymerization (left) and editing (right) mode. Template strand nucleotide 18 is colored in magenta to better visualize the screw motion of the DNA. (d) Movement of the exonuclease

and the polymerase thumb domain in between the polymerization (grey) and editing (color) mode. See also Supplementary Video 2.

with a blunt ended DNA (**h**) (Supplementary Table 2). Assays were performed without deoxynucleotides (dNTPs). Uncropped gels are shown in Supplementary Data Set 1.

Figure 3.

A terminal mismatch increases fraying of the DNA. **(a)** Sequence of the three DNA substrates used for the NMR studies: top, middle, and bottom substrates show the matched (C:G), mismatched (C:T) or unpaired (C:-) substrates. The normalized exchange rates of the measured base pairs are indicated in background color. **(b)** 1D imino proton resonances for these DNA substrates at 310K. The numbers on the top spectrum indicate the base pair position. Note, no signal was observed for the 20/20^{C:T} and 20/19^{C:-} substrates at position 2 (indicated with “*”). **(c)** UV DNA-melting experiments show a ~10 degrees shift in the melting temperature for the non-matched DNA substrates. **(d)** Denaturing urea-PAGE showing primer extension activity in the presence of nucleotides on a matched and mismatched substrate (see Supplementary Table 2 for substrate details). Uncropped gels are shown in Supplementary Data Set 1 **(e)** Fluorescence anisotropy analysis of the DNA binding affinities of the full complex and the exonuclease-θ complex.