



OPEN

Efficient discrimination against RNA-containing primers by human DNA polymerase ϵ

Alisa E. Lisova^{1,2}, Andrey G. Baranovskiy^{1,2}, Lucia M. Morstadt¹, Nigar D. Babayeva¹ & Tahir H. Tahirov¹✉

DNA polymerase ϵ (Pol ϵ) performs bulk synthesis of DNA on the leading strand during genome replication. Pol ϵ binds two substrates, a template:primer and dNTP, and catalyzes a covalent attachment of dNMP to the 3' end of the primer. Previous studies have shown that Pol ϵ easily inserts and extends ribonucleotides, which may promote mutagenesis and genome instability. In this work, we analyzed the mechanisms of discrimination against RNA-containing primers by human Pol ϵ (hPol ϵ), performing binding and kinetic studies at near-physiological salt concentration. Pre-steady-state kinetic studies revealed that hPol ϵ_{CD} extends RNA primers with approximately 3300-fold lower efficiency in comparison to DNA, and addition of one dNMP to the 3' end of an RNA primer increases activity 36-fold. Likewise, addition of one rNMP to the 3' end of a DNA primer reduces activity 38-fold. The binding studies conducted in the presence of 0.15 M NaCl revealed that human hPol ϵ_{CD} has low affinity to DNA (K_D of 1.5 μ M). Strikingly, a change of salt concentration from 0.1 M to 0.15 M reduces the stability of the hPol ϵ_{CD} /DNA complex by 25-fold. Upon template:primer binding, the incoming dNTP and magnesium ions make hPol ϵ discriminative against RNA and chimeric RNA–DNA primers. In summary, our studies revealed that hPol ϵ discrimination against RNA-containing primers is based on the following factors: incoming dNTP, magnesium ions, a steric gate for the primer 2'OH, and the rigid template:primer binding pocket near the catalytic site. In addition, we showed the importance of conducting functional studies at near-physiological salt concentration.

Abbreviations

hPol ϵ	Human DNA polymerase ϵ
hPol ϵ_{CD}	Catalytic domain of human DNA polymerase ϵ
CMG	Cdc45-MCM-GINS
EMSA	Electrophoretic mobility shift assay
TCEP	Tris(2-carboxyethyl)phosphine

DNA polymerase ϵ (Pol ϵ) is one of the main eukaryotic replicases and is responsible for extension of the leading DNA strand during replication of genetic material¹. It belongs to the B-family of DNA polymerases, which also includes Pol δ , Pol α , and Pol ζ . Human Pol ϵ (hPol ϵ) consists of four subunits: the catalytic subunit (p261) and the auxiliary subunits p59, p17, and p12. p261 is a bi-lobe protein comprised of two tandem exonuclease/polymerase domains, the first of which is active, while the second is inactive and plays an essential structural role^{2–4}. Two small subunits, p17 and p12, form a heterodimer, which bridges the two lobes of p261⁵. Pol ϵ is tethered to the Cdc45-MCM-GINS (CMG) helicase via the N-terminus of p59 and the C-terminal part of p261^{4,6,7}.

The accuracy of DNA replication is important for genome stability. Pol ϵ achieves the high fidelity of DNA replication due to its accurate polymerase and proofreading exonuclease, which removes incorrect nucleotides from the primer 3' end. Instability of hPol ϵ holoenzyme drives replication stress, tumorigenesis, and developmental abnormalities^{8,9}. Somatic mutations affecting the exonuclease domain of hPol ϵ often occur in hypermutated colorectal and endometrial tumors^{10–12}. Many of them cause a mutator or ultra-mutator phenotype upon modelling in yeast system^{13–15}.

The vast majority of previous kinetic and DNA binding studies of human and yeast Pol ϵ were conducted at low salt concentration or in the absence of it^{16–20}, which might affect the interpretation of obtained results. In

¹Eppley Institute for Research in Cancer and Allied Diseases, Fred and Pamela Buffett Cancer Center, University of Nebraska Medical Center, Omaha, NE 68198, USA. ²These authors contributed equally: Alisa E. Lisova and Andrey G. Baranovskiy. ✉email: ttahirov@unmc.edu

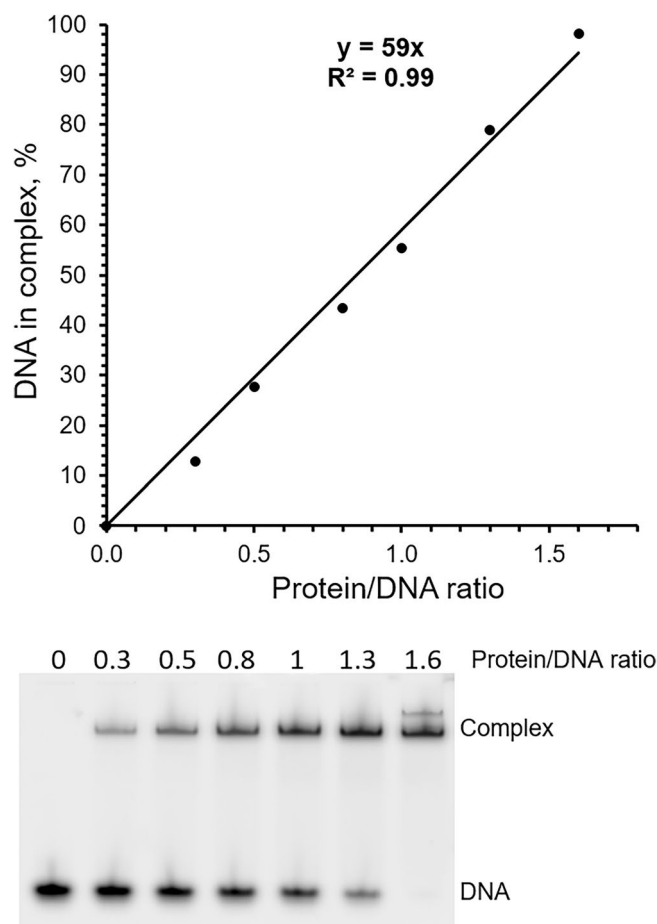


Figure 1. Estimation of the level of active molecules by EMSA. The Cy3-labeled DNA (0.5 μ M) was incubated with varying amounts of hPole_{CD}. Samples were separated by electrophoresis in 5% acrylamide gel and visualized by Typhoon FLA 9500. The percentage of DNA in the complex is plotted against the protein/DNA ratio, and the generated trend line shows the percentage of active molecules capable of DNA binding. The result of one representative experiment is shown.

this work, we performed functional studies of hPole at near-physiological salt concentration by employing pre-steady-state and binding kinetics. We studied the discrimination of hPole against RNA-containing primers at template:primer binding and dNTP incorporation steps. In addition, we examined the effect of dNTP, magnesium ions, and salt concentration on the stability of hPole/template:primer complexes.

Results

Salt concentration has a dramatic effect on the stability of the hPole/DNA complex. The catalytic domain of human DNA polymerase ϵ (hPole_{CD}; residues 27-1172) with inactivated exonuclease activity was expressed in *Escherichia coli* and purified to near homogeneity (Suppl. Figure S1). The concentrated sample had a brownish color due to the presence of a [4Fe-4S] cluster^{21,22}. Before starting functional studies with the obtained hPole_{CD} sample, we estimated the level of molecules possessing DNA binding activity. Increasing amounts of protein were added to reactions containing 0.5 μ M DNA, and the level of DNA in complex with a protein was estimated by electrophoretic mobility shift assay (EMSA; Fig. 1). It was found that a hPole_{CD} sample with homogeneity greater than 90% (Suppl. Figure S1) has a level of active molecules of $55.4 \pm 4.3\%$. All data described below were obtained using the active enzyme concentration.

The analysis of Pole_{CD} interaction with DNA (Suppl. Table S1) was performed at near-physiological salt concentration using an Octet K2, which employs Bio-Layer Interferometry (BLI) technology to monitor molecular interactions in real time. This approach allows obtaining the rate constants of complex formation (k_{on}) and dissociation (k_{off}) as well as the dissociation constant (K_D). A 23-mer DNA template with biotin at the 5'-end was primed by an 11-mer DNA primer and loaded on a streptavidin-coated sensor. Primer length selection was based on structural data showing that ten base-pairs in a duplex is sufficient to form all necessary contacts with the conserved catalytic domain of yeast Pole_{CD}^{23,24}.

Binding studies revealed that in the presence of 0.15 M NaCl hPole_{CD} binds a DNA duplex with low affinity ($K_D = 1.5 \mu$ M, Table 1). The obtained k_{off} value of 0.46 s^{-1} indicates that the half-life of the hPole_{CD}/DNA complex is $\sim 1.5 \text{ s}$ on average. Surprisingly, previous reports showed one to two orders of magnitude stronger interaction

[NaCl] mM	dTTP ^a Mg ²⁺	$k_{\text{off}} \times 10^{-3} \text{ s}^{-1}$	$k_{\text{on}} (\text{mM}^{-1} \text{ s}^{-1})$	K_D^b (nM)
150	–	463 ± 59	304 ± 20	1519 ± 91
	+	910 ± 29	235 ± 36	3790 ± 321
125	–	138 ± 18	258 ± 7.6	532 ± 53
100	–	21.8 ± 3.9	356 ± 7.2	61.4 ± 12
	+	31.9 ± 5.6	264 ± 24	120 ± 11

Table 1. Effect of salt concentration and dTTP/Mg²⁺ on interaction of hPole_{CD} with DNA. Data are presented as mean ± SD (n = 3). ^adTTP was added at a concentration of 50 μM, together with 5 mM MgCl₂. ^bK_D values are obtained by dividing k_{off} by k_{on} .

Primer sequence	$k_{\text{off}} \times 10^{-3} \text{ s}^{-1}$		$k_{\text{on}} (\text{mM}^{-1} \text{ s}^{-1})$		K_D^a (nM)	
	–	+ dTTP	–	+ dTTP	–	+ dTTP
GCCUGGAGCGC (0)	21.9 ± 2.2	–	336 ± 80	–	65.8 ± 8.8	–
GCCUGGAGCG/ddC/ (1) ^a	19.1 ± 2.5	88.6 ± 11	269 ± 38	210 ± 15	73.0 ± 14	424 ± 22
GCCUGGAGCG/ddC/ (3)	19.8 ± 0.1	110 ± 18	318 ± 29	250 ± 24	62.5 ± 6.1	439 ± 28
GCCUGGAGCG/ddC/ (5)	18.2 ± 1.8	58.5 ± 1.7	264 ± 51	269 ± 24	69.8 ± 6.4	218 ± 13
GCCUGGAGCG/ddC/ (7)	15.0 ± 2.1	22.6 ± 2.7	257 ± 31	273 ± 22	58.6 ± 1.3	83 ± 3.3
GCCTGGAGCG/ddC/ (11)	21.8 ± 3.9	31.9 ± 5.6	356 ± 7.2	264 ± 24	61.4 ± 12	120 ± 11
GCCTGGAGCG/3dC/ (10) ^b	17.1 ± 1	54.9 ± 0.2	378 ± 27	244 ± 13	45.6 ± 6.1	225 ± 13

Table 2. Effect of primer structure and dTTP/Mg²⁺ on interaction of human Pole_{CD} with template:primer. 2'-Deoxy- and 2',3'-dideoxy- nucleotides are underlined and their quantity is indicated in parenthesis. The provided values for DNA primers at 0.1 M NaCl are from Table 1. Data are presented as mean ± SD (n = 3). ^a/ddC/—2',3'-dideoxy-cytosine. ^b/3dC/—3'-deoxy-cytosine.

with DNA for human and yeast Pole. For example, the reported K_D values for the complexes of hPole and hPole_{CD} with DNA are 22 nM and 79 nM, respectively^{16,18}. For yeast Pole and Pole_{CD}, the corresponding K_D values are 11.6 nM and 15.6 nM, respectively¹⁹. The discrepancy with our data might be explained by the absence of salt in those studies. Of note, the obtained k_{on} value of 304 mM⁻¹ s⁻¹ (Table 1) is close to the previously reported k_{on} value of 270 mM⁻¹ s⁻¹ acquired by a different approach based on the measurement of DNA polymerase activity¹⁶.

To analyze the effect of salt concentration on the stability of the hPole_{CD}/DNA complex, we also conducted binding experiments in the presence of 0.1 M NaCl. Strikingly, a decrease in salt concentration from 0.15 M to 0.1 M reduced K_D by 25-fold (from 1.52 μM to 61 nM), which is mainly due to the 21-fold reduction in k_{off} value (Table 1). Consistently, in the presence of 125 mM NaCl, the hPole/DNA complex showed a K_D of 532 nM, which is 8.7-fold higher than the K_D value obtained at 100 mM NaCl. The dramatic effect of the ionic strength on the stability of the hPole_{CD}/DNA complex will be discussed below.

We analyzed how the incoming dNTP at near-physiological salt concentration²⁵ affects the interaction of hPole_{CD} with DNA. DNA polymerases use divalent metals to coordinate the triphosphate moiety of dNTP, so magnesium chloride was added to the reaction as well. Addition of 50 μM dTTP and 5 mM MgCl₂ in the presence of 0.15 M NaCl resulted in a 2.5-fold reduction in affinity of hPole_{CD} to DNA, with an almost two-fold increase in k_{off} value (Table 1). A similar effect was observed at 0.1 M NaCl, with K_D and k_{off} being increased by two- and 1.5-fold, respectively. The following exploration revealed that Mg²⁺ ions make a major impact on Pole interaction with a template:primer (see below).

The incoming dNTP and Mg²⁺ aid hPole in discrimination against RNA and chimeric primers upon template:primer binding. We analyzed how hPole_{CD} discriminates against RNA and chimeric RNA–DNA primers upon binding them in the presence and absence of dTTP/Mg²⁺. Without them, similar K_D values were obtained for all types of primers (Table 2; Fig. 2). The addition of 50 μM dTTP and 5 mM MgCl₂ reduced the stability of all complexes but to varying extents depending on primer structure. For a chimeric primer with seven dNMPs at the 3' end, K_D increased only 1.5-fold, which is close to the two-fold increase observed for the DNA primer. In the presence of dTTP/Mg²⁺, the chimeric primers with one or three dNMPs at the 3' end demonstrate K_D values that are 3.5-fold higher compared to the DNA primer (Table 2; Fig. 2). In comparison, the discrimination factor reduced ~two-fold for a primer with five dNMPs at the 3' end.

In order to estimate how insertion of one rNMP affects interaction of hPole with the template:primer, we used a DNA primer with a 3'-deoxy-cytosine. This primer has a 2'-hydroxyl to imitate ribonucleotide and lacks 3'OH to prevent primer extension. Upon addition of dTTP/Mg²⁺, the affinity of hPole_{CD} to the template:primer decreased five-fold, which is similar to the 5.8-fold reduction observed for an RNA primer with a dideoxy-cytosine at the 3'-end (Table 2; Fig. 2). Direct comparison of affinity to a DNA primer and to a primer with 3'-rNMP demonstrates a 1.9-fold higher selectivity for DNA. Therefore, hPole senses the primer 2'-hydroxyl only in the presence of incoming dNTP and Mg²⁺.

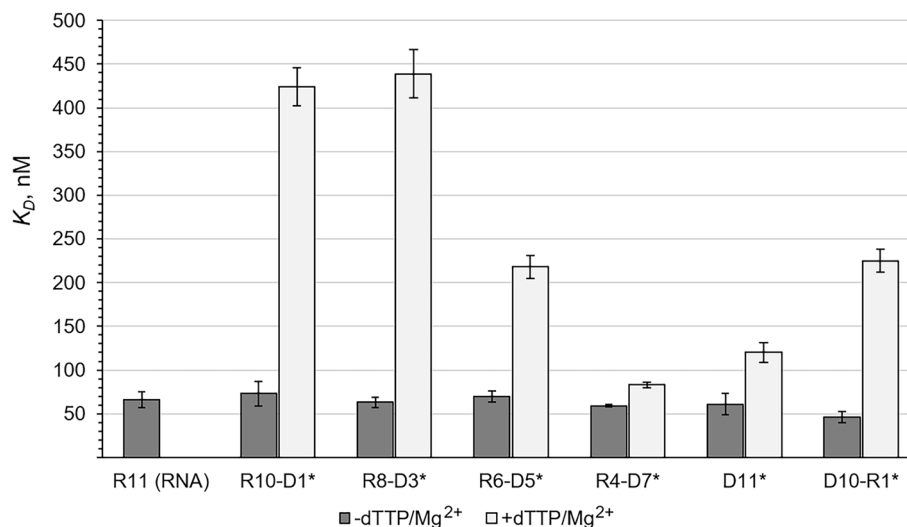


Figure 2. Effect of primer structure and dTTP/Mg²⁺ on hPole_{CD} affinity to a template:primer. The K_D values from Table 2 are presented as bar graphs showing the mean \pm SD (n = 3). The primer type is indicated below the bar. Asterisk indicates the absence of primer 3'-hydroxyl.

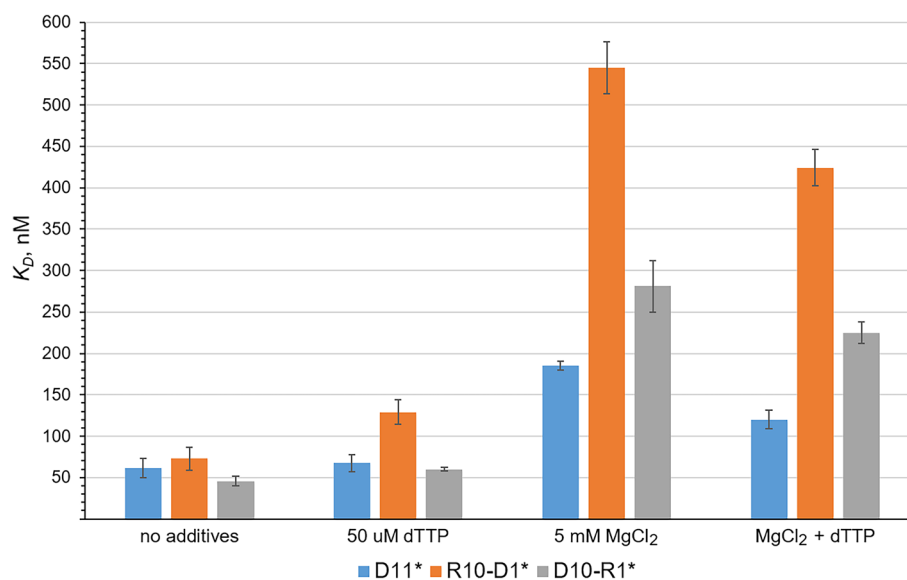


Figure 3. Effect of dTTP and Mg²⁺ on hPole_{CD} affinity to a template:primer. The data are presented as bar graphs showing the mean \pm SD (n = 3). The primer type is indicated below the additives. Asterisk indicates the absence of primer 3'-hydroxyl.

The effect of either 5 mM MgCl₂ or 50 μ M dTTP on hPole_{CD} affinity to a template:primer was analyzed using the following primers: DNA, R10-D1, and D10-R1 (Fig. 3). Interestingly, these additives added separately or together had a stronger impact on hPole interaction with RNA-containing primers. The most impressive effect was observed for R10-D1, where addition of either dTTP or Mg²⁺ increased K_D values by 1.8- and 7.5-fold, respectively (Fig. 3). Thus, in the presence of dTTP, hPole has a 1.9-fold higher affinity to a DNA primer versus R10-D1; in the presence of Mg²⁺, the difference is three-fold. Of note, discrimination against both RNA-containing primers is slightly higher in the presence of dTTP/Mg²⁺ than Mg²⁺ (Fig. 3). These data indicate that, upon ternary complex formation, the incoming dNTP and Mg²⁺ help hPole to discriminate against RNA-containing primers until they are extended with several dNMPs.

hPole extends RNA-containing primers with greatly reduced activity. It was previously shown that Pole is able to incorporate ribonucleotides into a growing DNA strand and to extend them^{26,27}. Moreover, Pole can extend R-loops, which present the DNA template primed by RNA as well as a displaced complemen-

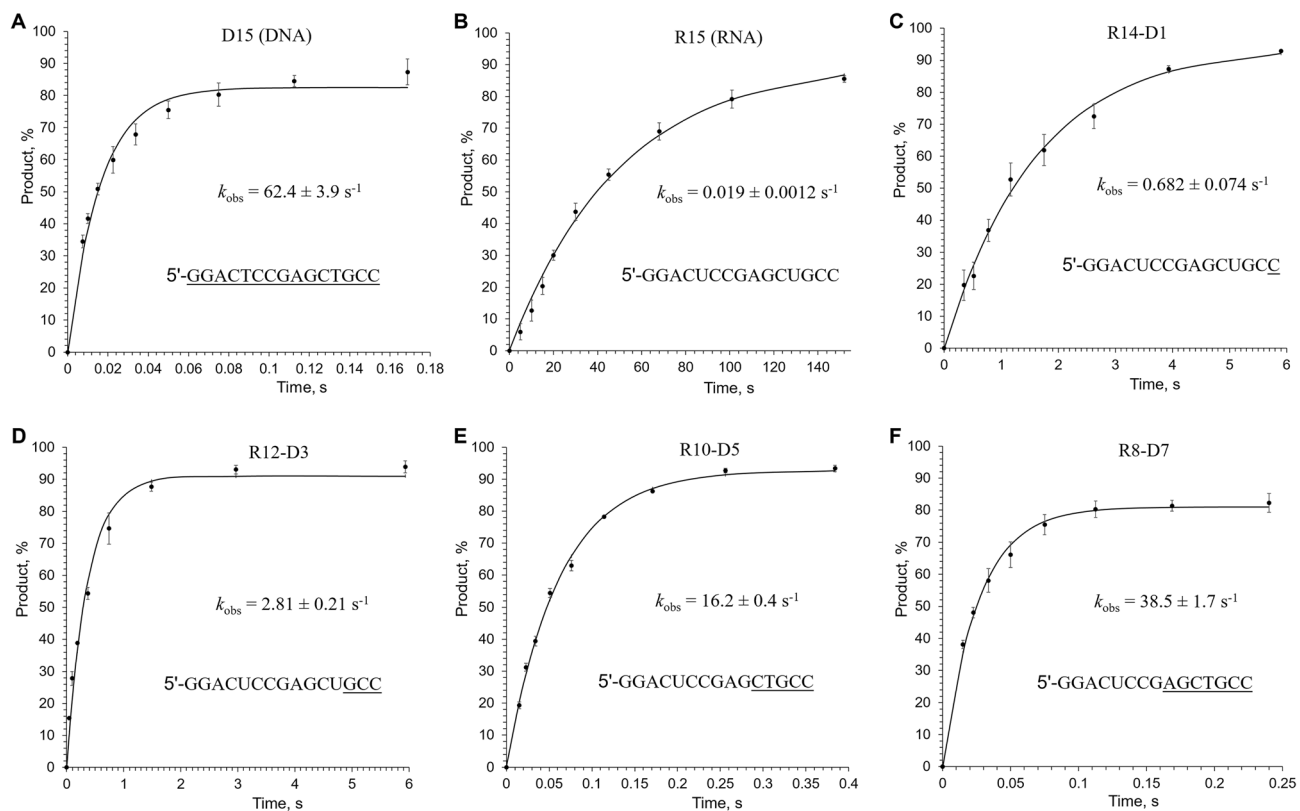


Figure 4. Effect of primer structure on its extension by hPol ϵ_{CD} . Activity of hPol ϵ_{CD} in extension of a DNA primer (A), an RNA primer (B), and chimeric RNA–DNA primers with one (C), three (D), five (E), and seven (F) deoxynucleotides at the primer 3′-end was analyzed. Primer sequences are indicated on the graphs; DNA stretches in chimeric primers are underlined. Percentage of extended primer was plotted against time and the data fit to a single-exponential equation (1). Reactions, containing 0.44 μ M hPol ϵ_{CD} , 0.2 μ M template:primer, and 1 mM dTTP, were incubated at 35 °C at the indicated time points. The mean k_{obs} values and standard deviations were calculated from three independent experiments using GraphPad Prizm software.

tary DNA strand²⁸. Accordingly, we decided to explore how the presence of ribonucleotides in a primer affects its extension by hPol ϵ_{CD} by employing the pre-steady-state kinetic approach. Single-nucleotide incorporation experiments were performed under single-turnover conditions by providing a two-fold excess of hPol ϵ_{CD} over DNA (Suppl. Table S1). hPol ϵ_{CD} (0.88 μ M) was incubated with a Cy3-labeled DNA (0.4 μ M) and quickly mixed with an equal volume of 2 mM dTTP and 10 mM MgCl₂ under rapid chemical quench conditions. The fraction of extended primer was plotted against time (Fig. 4) and the data were fit to a single-exponential equation (Eq. 1).

hPol ϵ_{CD} extends the DNA primer with a rate of 62.4 s⁻¹ obtained at 35 °C in the presence of 0.1 M NaCl (Fig. 4A). In comparison, the catalytic domain of hPol α has a k_{pol} of 33.8 s⁻¹ at conditions similar to those used in this work²⁹. These results are in line with the recently reported average rate of replication fork movement in yeast cells, which is ~50 nucleotides per second³⁰. A significantly higher k_{pol} value of 248 s⁻¹ was obtained for the hPol ϵ_{CD} at 20 °C in the absence of salt¹⁶. The authors of that study noted that the rate of polymerization was >500 s⁻¹ at 37 °C. The almost tenfold difference in activity with our data is probably due to the inhibitory effect of salt concentration as previously demonstrated³¹.

Activity of hPol ϵ_{CD} in extension of RNA and chimeric primers was analyzed at the same conditions as with the DNA primer. Strikingly, the rate of RNA primer extension is 3300-fold lower compared to DNA (Fig. 4B). Such a severe effect on catalysis indicates that pre-catalytic positioning of the primer 3′-hydroxyl relative to the α -phosphate of dNTP and/or the catalytic metal is changed. Upon addition of dNMPs to the 3′-end of an RNA primer, the rate of primer extension gradually increases (Fig. 4C–F). Thus, the rates of dNMP incorporation on the chimeric RNA–DNA primers with one, three, five, and seven dNMPs at the 3′-end are 92-, 22-, 3.8-, and 1.6-fold lower, respectively, than on DNA (Fig. 5).

Interestingly, addition of just one dNMP to the 3′-end of an RNA primer increases activity 36-fold, from 0.019 s⁻¹ to 0.682 s⁻¹, while the subsequent addition of every two dNMPs increases activity only several-fold (Fig. 4). Moreover, addition of one ribonucleotide to the 3′-end of a DNA primer reduces the rate of dTTP incorporation 38-fold, from 62.4 s⁻¹ to 1.63 s⁻¹ (Fig. 6). A similar effect on Pol ϵ activity upon addition of either 3′-rNMP to a DNA primer or 3′-dNMP to an RNA primer indicates the presence of a steric gate for 2′OH. Indeed, analysis of the structure of yeast Pol ϵ_{CD} in complex with DNA and dNTP^{23,24} revealed a conserved Thr876 (Suppl. Figure S2) located under the sugar and sensing the 2′-hydroxyl of a primer (Fig. 7). The other steric gate

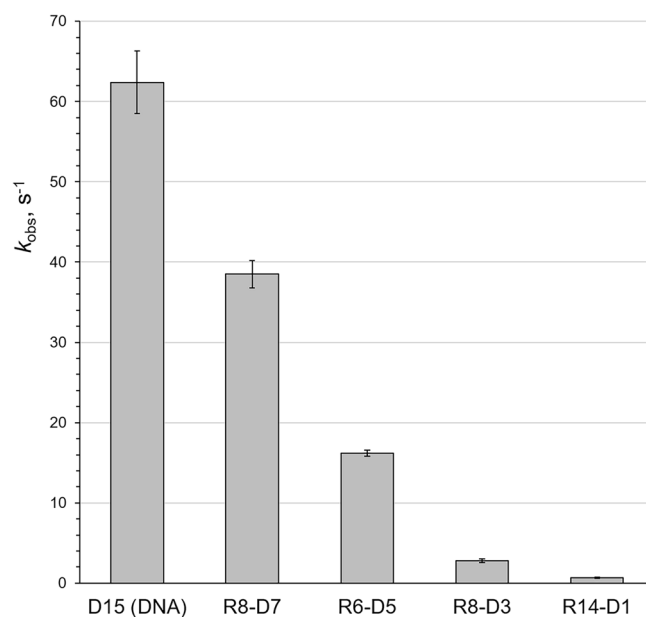


Figure 5. Efficiency of RNA–DNA primer extension by hPol ϵ_{CD} depends on the length of DNA tract at the 3′-end. The k_{obs} values are presented as bar graphs showing the mean \pm SD ($n = 3$). The type of primer is indicated below the bar.

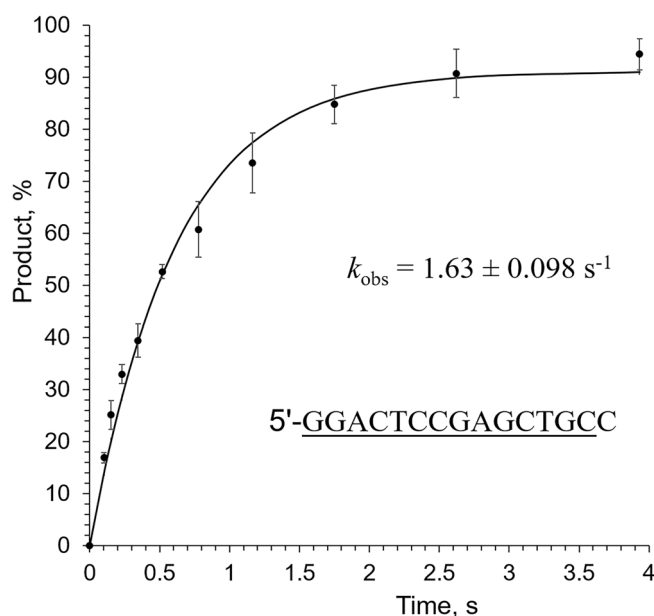


Figure 6. Kinetics of dNMP incorporation into a DNA primer with a ribonucleotide at the 3′-end. Percentage of extended primer was plotted against time and the data fit to a single-exponential equation (Eq. 1). Primer sequence is indicated on the graph with the DNA tract underlined. Reactions, containing $0.44 \mu\text{M}$ hPol ϵ_{CD} , $0.2 \mu\text{M}$ template:primer, and 1 mM dTTP, were incubated at 35°C at the indicated time points. The mean k_{obs} values and standard deviations were calculated from three independent experiments by using GraphPad Prism software.

is reserved for preventing rNTPs incorporation, where a conserved tyrosine residue clashes with a 2′-hydroxyl of an incoming nucleotide triphosphate³².

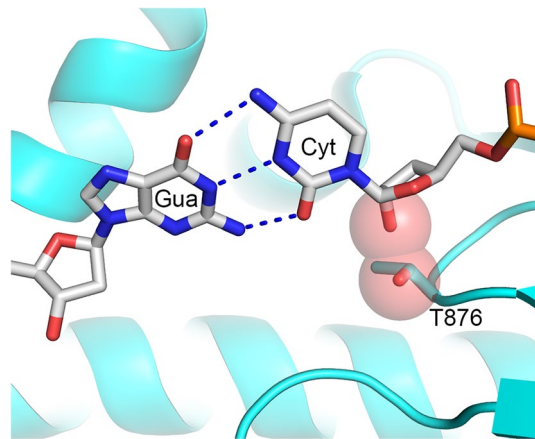


Figure 7. Pol ϵ has a steric gate for the 2'-hydroxyl of a primer. The protein is represented as cartoon and colored cyan; Thr876 is shown as sticks. The template and primer are shown as sticks and colored blue for nitrogen, red for oxygen, orange for phosphorus, and grey for carbon. The O1 of Thr876 and modeled O2 of primer Cyt are presented as spheres with 60% transparency and colored red. The interception of spheres indicates steric hindrance. The crystal structure of the ternary complex of yeast Pol ϵ_{CD} with DNA and dCTP (pdb code 6hiv (22)) was used for this presentation prepared with the PyMOL Molecular Graphics System.

Discussion

Protein–DNA interactions are mainly electrostatic in nature, which makes them sensitive to ionic strength. If the inhibitory effect of salt is due to competition of ions with charged groups of Pol ϵ and DNA involved in interaction, it should be proportional to the salt concentration. Despite this prediction, we have shown that salt severely affects the interaction of hPol ϵ_{CD} with DNA. In particular, a 1.5-fold increase in salt concentration from 0.1 M to 0.15 M results in a 25-fold reduction in the stability of the hPol ϵ_{CD} /DNA complex (Table 1). Such a dramatic effect of salt concentration on affinity to DNA can be attributed to changes in protein solvation and elevated conformational dynamics of the flexible parts/domains. The catalytic domains of B-family DNA polymerases have the universal “right-hand” DNA polymerase fold³³ consisting of five subdomains: N-terminal, exonuclease, palm, fingers, and thumb. Among them, the fingers are the most flexible and change the conformation from “open” to “closed” upon dNTP binding. The thumb domain grips the distant part of the DNA duplex and demonstrates significant flexibility according to structural studies of hPol α ²⁹. An ordered movement of the palm domain during polymerase translocation along the duplex was proposed from structural studies of yeast Pol α ³⁴.

This study has shown that hPol ϵ_{CD} exhibits low affinity to the template:primer at physiological salt concentration, with a 19-fold higher K_D value than the 79 nM obtained in the absence of salt¹⁶. Interestingly, hPol α shows a similar with Pol ϵ affinity to DNA in the presence of 100 mM ($K_D = 41$ nM) and 150 mM NaCl ($K_D = 1.4$ μ M), despite the absence of the so-called processivity domain that is unique to Pol ϵ ³⁵. According to structural data^{22–24}, this domain makes three to four additional hydrogen bonds with a template:primer (a detailed comparison of hPol α and yPol ϵ interactions with a template:primer are provided in³⁶). In other work, we have shown that hPol ϵ_{CD} with an intact exonuclease active site and without internal truncations also exhibits a strong sensitivity to ionic strength upon template:primer binding³⁷.

The results of our binding studies support the idea that PCNA is required for processive synthesis of the leading strand, according to the current model of the eukaryotic replication fork³⁸. The recently published structure of yeast Pol ϵ holoenzyme revealed that accessory subunits do not interact with the DNA duplex⁵, indicating that the holoenzyme footprint on a DNA duplex is only ten base-pairs, the same as for the catalytic domain^{22,24}. Pol ϵ is flexibly attached to the CMG helicase through interaction with the C-terminal domain of the catalytic subunit and the N-terminal domain of the second accessory subunit^{4,7}. In contrast to PCNA, interaction with CMG does not prevent the Pol ϵ catalytic domain from occasional dissociation from the growing primer end.

The main role of Pol ϵ is a synthesis of the leading DNA strand, but it may handle RNA and chimeric RNA–DNA strands in certain circumstances. One example is a ribonucleotide at the growing primer end, which Pol ϵ can insert and extend^{126,27}. Another example is the ability of Pol ϵ to extend the 3'-ends of R-loops, serving as a possible way to restart DNA replication²⁸. In addition, it is possible that Pol α sometimes preliminarily terminates DNA synthesis due to generation of a mismatch, like purine–purine, which is difficult to extend. In this case, Pol ϵ may bind the chimeric RNA–DNA primer, proofread the non-cognate nucleotide, and extend the corrected primer with dNMPs until the length of a primer is enough for RFC to load PCNA. A similar mechanism was suggested for the lagging strand where Pol δ proofreads the Pol α -generated mismatches³⁹. Intriguingly, some portion of hPol ϵ is associated with RNA polymerase II independently of the cell cycle⁴⁰. Moreover, according to results of UV cross-linking studies, hPol ϵ is located in close proximity to the newly synthesized RNA strand⁴⁰. These data suggest that Pol ϵ could play a currently unknown role in RNA transcript processing.

As we found, hPol ϵ_{CD} shows approximately 3,000-fold lower activity in extension of RNA versus DNA primers, making it unlikely that Pol ϵ plays a role in restarting the replication fork from the R-loops. The k_{obs} value of 0.019 s⁻¹ indicates that half of RNA primers will be extended with one dNMP in ~ 40 s. The best candidate

for R-loop extension would be Pol α , which displays a similar rate of DNA and RNA primer extension²⁹. On the other hand, upon extension of an RNA primer with dNMPs, hPol ϵ_{CD} activity gradually increases, and the chimeric primer with seven dNMPs is extended fairly well (Fig. 4). Our studies have shown that, in comparison to template:primer binding, the rate of dNMP incorporation is more sensitive to the presence of ribonucleotides in the primer. It is interesting that despite their similarly organized DNA binding sites²⁹, Pol α and Pol ϵ demonstrate such a significant difference in selectivity to a DNA primer upon catalysis of phosphodiester bond formation.

Incorporation of ribonucleotides to the nascent DNA strand by replicative DNA polymerases could be a challenge for genome integrity^{41,42}. It was reported previously that hPol ϵ_{CD} readily inserts ribonucleotides, and almost half of them escape proofreading by intrinsic exonuclease²⁷. Interestingly, activity of hPol ϵ was reduced only two- to three-fold when one–three consecutive ribonucleotides were added to the primer 3'-end²⁷. Such a low effect of 3'-rNMPs on primer extension may be due to the absence of salt in reaction and to the type of assay where the rate of enzyme/DNA complex formation limits the reaction rate due to low concentrations of DNA and Pol ϵ and the presence of 10% glycerol. Our study, performed at single-turnover conditions and in the presence of 0.1 M NaCl, demonstrates much stronger sugar selectivity, with a 38-fold reduction in activity by 3'-rNMP (Fig. 6). Due to the balance between DNA polymerase and exonuclease activities⁴³, the hampered extension of 3'-ribonucleotides will result in increased probability of their excision.

A notable finding of the current study is the role of incoming dNTP and magnesium ions in the discrimination of hPol ϵ against primers containing ribonucleotides. When dNTP and Mg²⁺ are not present, hPol ϵ binds DNA, RNA, and chimeric primers with similar affinity. In the presence of dTTP/Mg²⁺, selectivity for DNA primers increases two- to four-fold. Interestingly, Mg²⁺ ions play the main role in discrimination against RNA-containing primers upon template:primer binding (Fig. 3). Magnesium ions most likely interact with DNA phosphates and compete with DNA-binding residues of hPol ϵ_{CD} . In the case of RNA-containing primers, a stronger effect of Mg²⁺ on the hPol ϵ /template:primer interaction can be due to steric hindrances generated by the wide helix of DNA:RNA duplexes and by the primer 2'-hydroxyl. Of note, dTTP alone has a stronger impact on hPol ϵ affinity to the R10-D1 primer than to DNA (Fig. 3). This is an interesting example of how one substrate modulates enzyme selectivity for another. Noteworthy, there is no discrimination against RNA primers extended with seven or more dNMPs.

Available structural data provide clues as to why, upon template:primer binding, hPol ϵ discriminates against the chimeric primers with one or three dNMPs with similar efficiency and does not discriminate against primers containing seven dNMPs at the 3' end. Pol ϵ interacts with a ten base-pair duplex, where the first four base pairs from the growing primer end interact with the rigid DNA-binding pocket established mainly by the palm subdomain with significant contributions by the main-chain atoms^{22–24}. The remainder of the template:primer interacts with the flexible thumb, which can adjust its conformation to accommodate a wider DNA:RNA duplex. The absence of discrimination against RNA when dNTP and Mg²⁺ are not added may be due to the slightly different mode of duplex binding, which reduces steric hindrance between the wide DNA:RNA duplex and the rigid DNA-binding cleft. dNTP and Mg²⁺ binding and the fingers closing likely forces hPol ϵ interaction with the template:primer in a specific way, which results in steric hindrance and reduced affinity to RNA and chimeric primers.

Materials and methods

Cloning, expression, and purification. The catalytic domain of hPol ϵ (hPol ϵ_{CD}), with the proofreading 3' \rightarrow 5' exonuclease inactivated by the double mutation D275N/E277Q, was cloned into a pASHSUL-1 plasmid⁴⁴ to produce the recombinant protein with a cleavable N-terminal His₆-Sumo tag. To increase protein solubility and yield, we deleted the external loop spanning amino acids 185–209. The corresponding region is distant from the active center and is mostly disordered in all yeast Pol ϵ_{CD} structures.

hPol ϵ_{CD} were expressed in *E. coli* strain Rosetta-2 (DE3) at 18 °C for 16 h following induction with 0.2 μ g/ml anhydrotetracycline. Afterward, the cells were harvested by centrifugation at 4000 g for 15 min, washed with PBS, aliquoted and maintained at –80 °C.

The purification protocol includes chromatography on a Ni-IDA column (Bio-Rad), His-Sumo-tag digestion during overnight dialysis, and chromatography on a Heparin HP HiTrap column (GE Healthcare) as well as on a size-exclusion column in buffer containing 25 mM Tris-HEPES (pH 7.8), 0.15 M NaCl, 1% glycerol, and 2 mM tris(2-carboxyethyl)phosphine (TCEP). Finally, samples were concentrated to 30–60 μ M and flash-frozen in aliquots. Protein concentrations were estimated by measuring the absorbance at 280 nm and using extinction coefficients of 154 mM⁻¹ cm⁻¹. The extinction coefficients were calculated with ProtParam⁴⁵.

Estimation of the level of DNA-binding (active) molecules by electrophoretic mobility shift assay. Reactions containing 0.5 μ M DNA (Suppl. Table S1) and varying amounts of protein were incubated in 10 μ l for 5 min at room temperature in the buffer containing 20 mM Tris-Hepes (pH 7.8), 100 mM NaCl, 2% glycerol, 2 mM TCEP, and 0.2 mg/mL BSA; 5 μ l was then loaded on 5% native PAGE. Samples labeled with Cy3-dye were visualized using the Typhoon 9410 imager (GE Healthcare) and quantified using ImageJ software (version 1.53 s, National Institutes of Health). The mean value and standard deviation are calculated from three independent experiments.

Binding studies. Analysis of binding kinetics was done at 23 °C on an Octet K2 (Sartorius AG). This device uses Bio-Layer Interferometry technology to monitor molecular interactions in real time. A template with a biotin-TEG at the 5'-end was annealed to the primers (Suppl. Table S1) and immobilized on a streptavidin-coated biosensor (SAX, Sartorius AG). Primers were added at two-fold molar excess over the template. SAX sensors were loaded with oligonucleotide-biotin at 50 nM concentration for 7 min at 500 rpm. Then sensors were blocked by incubating for 2 min in 10 μ g/ml biocytin. In the first row of a 96-well microplate (Greiner Bio-One), the first

six wells contained the buffer, consisting of 30 mM Tris-Hepes, pH 7.8, 100 mM or 150 mM NaCl, 2 mM TCEP, and 0.002% Tween 20. The next six wells contained two-fold dilutions of hPol ϵ_{CD} in the same buffer. All wells in the adjacent row contained only the buffer for reference. Examples of binding curves are shown in Suppl. Figure S3, where the effect of either 5 mM MgCl $_2$ or 50 μ M dTTP on the stability of the hPol ϵ_{CD} /DNA complex was analyzed. Data Analysis HT software (ver. 11.1, Sartorius AG) was used to calculate binding constants (k_{on} , k_{off} and K_D) by using global fitting. The average value and standard deviation were calculated from three independent experiments.

Kinetic studies. Pre-steady-state kinetic studies were performed on the QFM-4000 rapid chemical quench apparatus (BioLogic, France) at 35 °C. Reactions contained 0.44 μ M hPol ϵ_{CD} (active molecules), 0.2 μ M DNA, 1 mM dTTP, 25 mM Tris-HEPES, pH 7.8, 0.1 M NaCl, 5 mM MgCl $_2$, 2 mM TCEP, and 0.2 mg/mL BSA. hPol ϵ_{CD} was incubated with a Cy3-labeled 15-mer primer annealed to a 25-mer DNA template (Suppl. Table S1), to allow formation of the binary complex, and rapidly mixed with a solution containing 2 mM dTTP and 10 mM MgCl $_2$ followed by quenching with 0.3 M EDTA. Products were collected in a tube containing 15 μ l 100% formamide and separated by 20% Urea-PAGE. The Cy3-labeled products were visualized using the Typhoon FLA 9500 (GE Healthcare) and quantified by ImageJ, version 1.53 (NIH). The extended primer fraction was calculated by dividing the amount of extended primer by the amount of primer added in reaction. The percent of extended primer was plotted against time and the data fit to a single-exponential equation:

$$[\text{Product}] = A \times (1 - e^{-k_{obs}t}) \quad (1)$$

where A is the amplitude, k_{obs} is the observed rate for dNTP incorporation, and t is the time.

Data availability

The data that support the findings of this study are included in the Supplementary Information file or available from the corresponding author on request.

Received: 30 March 2022; Accepted: 9 June 2022

Published online: 17 June 2022

References

- Lujan, S. A., Williams, J. S. & Kunkel, T. A. DNA polymerases divide the labor of genome replication. *Trends Cell Biol.* **26**, 640–654 (2016).
- Tahirov, T. H., Makarova, K. S., Rogozin, I. B., Pavlov, Y. I. & Koonin, E. V. Evolution of DNA polymerases: An inactivated polymerase-exonuclease module in Pol epsilon and a chimeric origin of eukaryotic polymerases from two classes of archaeal ancestors. *Biol Direct* **4**, 11 (2009).
- Dua, R., Levy, D. L. & Campbell, J. L. Analysis of the essential functions of the C-terminal protein/protein interaction domain of *Saccharomyces cerevisiae* pol epsilon and its unexpected ability to support growth in the absence of the DNA polymerase domain. *J. Biol. Chem.* **274**, 22283–22288 (1999).
- Zhou, J. C. *et al.* CMG-Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. *Proc. Natl. Acad. Sci. USA* **114**, 4141–4146 (2017).
- Yuan, Z., Georgescu, R., Schauer, G. D., O'Donnell, M. E. & Li, H. Structure of the polymerase epsilon holoenzyme and atomic model of the leading strand replisome. *Nat. Commun.* **11**, 3156 (2020).
- Sengupta, S., van Deursen, F., de Piccoli, G. & Labib, K. Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr. Biol.* **23**, 543–552 (2013).
- Jones, M. L., Baris, Y., Taylor, M. R. G. & Yeeles, J. T. P. Structure of a human replisome shows the organisation and interactions of a DNA replication machine. *EMBO J.* **40**, e108819 (2021).
- Pachlopnik Schmid, J. *et al.* Polymerase epsilon1 mutation in a human syndrome with facial dysmorphism, immunodeficiency, livedo, and short stature (“FILS syndrome”). *J. Exp. Med.* **209**, 2323–2330 (2012).
- Bellelli, R. *et al.* Polepsilon instability drives replication stress, abnormal development, and tumorigenesis. *Mol. Cell* **70**(707–721), e707 (2018).
- Barbari, S. R. & Shcherbakova, P. V. Replicative DNA polymerase defects in human cancers: Consequences, mechanisms, and implications for therapy. *DNA Repair (Amst)* **56**, 16 (2017).
- Park, V. S. & Pursell, Z. F. POLE proofreading defects: Contributions to mutagenesis and cancer. *DNA Repair (Amst)* **76**, 50–59 (2019).
- Rayner, E. *et al.* A panoply of errors: Polymerase proofreading domain mutations in cancer. *Nat. Rev. Cancer* **16**, 71–81 (2016).
- Kane, D. P. & Shcherbakova, P. V. A common cancer-associated DNA polymerase epsilon mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Res.* **74**, 1895–1901 (2014).
- Barbari, S. R., Kane, D. P., Moore, E. A. & Shcherbakova, P. V. Functional analysis of cancer-associated DNA polymerase epsilon variants in *Saccharomyces cerevisiae*. *G3 (Bethesda)* **8**, 1019–1029 (2018).
- Herzog, M. *et al.* Mutagenic mechanisms of cancer-associated DNA polymerase alleles. *Nucleic Acids Res.* **49**, 3919–3931 (2021).
- Zahurancik, W. J., Klein, S. J. & Suo, Z. Kinetic mechanism of DNA polymerization catalyzed by human DNA polymerase epsilon. *Biochemistry* **52**, 7041–7049 (2013).
- Zahurancik, W. J., Baranovskiy, A. G., Tahirov, T. H. & Suo, Z. Comparison of the kinetic parameters of the truncated catalytic subunit and holoenzyme of human DNA polymerase varepsilon. *DNA Repair (Amst)* **29**, 16–22 (2015).
- Zahurancik, W. J. & Suo, Z. Kinetic investigation of the polymerase and exonuclease activities of human DNA polymerase epsilon holoenzyme. *J. Biol. Chem.* **295**, 17251–17264 (2020).
- Ganai, R. A., Osterman, P. & Johansson, E. Yeast DNA polymerase catalytic core and holoenzyme have comparable catalytic rates. *J. Biol. Chem.* **290**, 3825–3835 (2015).
- Eddy, S. *et al.* Evidence for the kinetic partitioning of polymerase activity on G-quadruplex DNA. *Biochemistry* **54**, 3218–3230 (2015).
- Jain, R. *et al.* An iron-sulfur cluster in the polymerase domain of yeast DNA polymerase epsilon. *J. Mol. Biol.* **426**, 301–308 (2014).
- Ter Beek, J. *et al.* Structural evidence for an essential Fe-S cluster in the catalytic core domain of DNA polymerase. *Nucleic Acids Res.* **47**, 5712–5722 (2019).

23. Hogg, M. *et al.* Structural basis for processive DNA synthesis by yeast DNA polymerase varepsilon. *Nat. Struct. Mol. Biol.* **21**, 49–55 (2014).
24. Jain, R. *et al.* Crystal structure of yeast DNA polymerase epsilon catalytic domain. *PLoS ONE* **9**, e94835 (2014).
25. Ferraro, P., Franzolin, E., Pontarin, G., Reichard, P. & Bianchi, V. Quantitation of cellular deoxynucleoside triphosphates. *Nucleic Acids Res.* **38**, e85 (2010).
26. Nick McElhinny, S. A. *et al.* Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. USA* **107**, 4949–4954 (2010).
27. Goksenin, A. Y. *et al.* Human DNA polymerase epsilon is able to efficiently extend from multiple consecutive ribonucleotides. *J. Biol. Chem.* **287**, 42675–42684 (2012).
28. Mazina, O. M. *et al.* Replication protein A binds RNA and promotes R-loop formation. *J. Biol. Chem.* **295**, 14203–14213 (2020).
29. Baranovskiy, A. G., Siebler, H. M., Pavlov, Y. I. & Tahirov, T. H. Iron-sulfur clusters in DNA polymerases and primases of eukaryotes. *Methods Enzymol.* **599**, 1–20 (2018).
30. Raghuraman, M. K. *et al.* Replication dynamics of the yeast genome. *Science* **294**, 115–121 (2001).
31. Chui, G. & Linn, S. Further characterization of HeLa DNA polymerase epsilon. *J. Biol. Chem.* **270**, 7799–7808 (1995).
32. Brown, J. A. & Suo, Z. Unlocking the sugar “steric gate” of DNA polymerases. *Biochemistry* **50**, 1135–1142 (2011).
33. Wang, J. *et al.* Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087–1099 (1997).
34. Perera, R. L. *et al.* Mechanism for priming DNA synthesis by yeast DNA polymerase alpha. *Elife* **2**, e00482 (2013).
35. Baranovskiy, A. G., Babayeva, N. D., Lisova, A. E., Morstadt, L. M. & Tahirov, T. H. Structural and functional insight into mismatch extension by human DNA polymerase alpha. *Proc. Natl. Acad. Sci. USA* **119**, e2111744119 (2022).
36. Baranovskiy, A. G. *et al.* Activity and fidelity of human DNA polymerase alpha depend on primer structure. *J. Biol. Chem.* **293**, 6824–6843 (2018).
37. Lisova, A. E. *et al.* The iron-sulfur cluster is critical for DNA binding by human DNA polymerase ϵ . *BioRxiv* <https://doi.org/10.1101/2022.1105.1105.490830> (2022).
38. Burgers, P. M. J. & Kunkel, T. A. Eukaryotic DNA replication fork. *Annu. Rev. Biochem.* **86**, 417–438 (2017).
39. Pavlov, Y. I. *et al.* Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Curr. Biol.* **16**, 202–207 (2006).
40. Rytkonen, A. K. *et al.* DNA polymerase epsilon associates with the elongating form of RNA polymerase II and nascent transcripts. *FEBS J.* **273**, 5535–5549 (2006).
41. Henninger, E. E. & Pursell, Z. F. DNA polymerase epsilon and its roles in genome stability. *IUBMB Life* **66**, 339–351 (2014).
42. Williams, J. S. & Kunkel, T. A. Ribonucleotides in DNA: origins, repair and consequences. *DNA Repair (Amst)* **19**, 27–37 (2014).
43. Ganai, R. A., Bylund, G. O. & Johansson, E. Switching between polymerase and exonuclease sites in DNA polymerase epsilon. *Nucleic Acids Res.* **43**, 932–942 (2015).
44. Weeks, S. D., Drinker, M. & Loll, P. J. Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr. Purif.* **53**, 40–50 (2007).
45. Gasteiger, E. *et al.* Protein identification and analysis tools on the ExpASY server. In *The Proteomics Protocols Handbook* (ed. Walker, J. M.) 571–607 (Humana, 2005).

Acknowledgements

We thank P. Shcherbakova and B. Preston for critical reading and K. Jordan for editing this manuscript.

Author contributions

A.E.L. and A.G.B. purified hPole_{CD} samples and performed functional studies. L.M.M. and N.D.B. prepared the plasmids encoding different hPole constructs and analyzed their expression. T.H.T. initiated and supervised the project. A.G.B. and A.E.L. wrote the manuscript, with contributions and critical comments from the other authors.

Funding

This work was supported by the National Institute of General Medical Sciences (NIGMS) grant R35GM127085 to T.H.T. The University of Nebraska Medical Center (UNMC) Genomics Core receives partial support from the National Institute for General Medical Science (NIGMS) INBRE P20GM103427 grant and the Fred & Pamela Buffett Cancer Center Support Grant P30CA036727. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-14602-2>.

Correspondence and requests for materials should be addressed to T.H.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022