A mechanistic model of primer synthesis from catalytic structures of DNA polymerase α–primase

Elwood A. Mullins[1,2], Lauren E. Salay[2,3,4], Clarissa L. Durie[5,6,7], Jane E. Jackman[8,9], Melanie D. Ohi[5,6], Walter J. Chazin[2,3,10,*], and Brandt F. Eichman[1,2,3,*]

[1]Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, USA
[2]Center for Structural Biology, Vanderbilt University, Nashville, Tennessee, USA
[3]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee, USA
[4]Present address: Department of Biochemistry, University of Washington, Seattle, Washington, USA
[5]Department of Cell and Developmental Biology, University of Michigan School of Medicine, Ann Arbor, Michigan, USA
[6]Life Sciences Institute, University of Michigan, Ann Arbor, Michigan, USA
[7]Present address: Department of Biochemistry, University of Missouri, Columbia, Missouri, USA
[8]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio, USA
[9]Center for RNA Biology, The Ohio State University, Columbus, Ohio, USA
[10]Department of Chemistry, Vanderbilt University, Nashville, Tennessee, USA

*Address correspondence to these authors: walter.j.chazin@vanderbilt.edu, brandt.eichman@vanderbilt.edu

Abstract

The mechanism by which polymerase α–primase (polα–primase) synthesizes chimeric RNA-DNA primers of defined length and composition, necessary for replication fidelity and genome stability, is unknown. Here, we report cryo-EM structures of polα–primase in complex with primed templates representing various stages of DNA synthesis. Our data show how interaction of the primase regulatory subunit with the primer 5′-end facilitates handoff of the primer to polα and increases polα processivity, thereby regulating both RNA and DNA composition. The structures detail how flexibility within the heterotetramer enables synthesis across two active sites and provide evidence that termination of DNA synthesis is facilitated by reduction of polα and primase affinities for the varied conformations along the chimeric primer/template duplex. Together, these findings elucidate a critical catalytic step in replication initiation and provide a comprehensive model for primer synthesis by polα–primase.

Introduction

DNA polymerase α–primase (polα–primase) is unique among the replicative polymerases by virtue of its ability to perform *de novo* synthesis from a DNA template. Polα–primase initiates DNA synthesis by generating short (~30-nucleotide) primers [1-3], which are required for leading and lagging strand replication by polε and polδ [4]. Generation of primers is particularly important on the lagging strand, where discontinuous DNA synthesis is repeatedly initiated during genome replication and telomere elongation [5]. As such, polα–primase is essential for replication of the genome, and defects in priming have been implicated in cellular ageing and an impaired DNA damage response [6-10].

Polα–primase possesses both DNA-dependent RNA polymerase (primase) and DNA polymerase (polα) activities. Primase initiates synthesis from nucleoside triphosphates (NTPs) to generate a 7–10-ribonucleotide primer [11,12], which is extended by ~20 deoxyribonucleotides by polα (**Fig. 1A**) [1,2]. The resulting chimeric RNA-DNA primers are extended by polδ to generate each Okazaki fragment. Polα lacks proofreading capability and is therefore mutagenic [13]. The potentially deleterious effects of polα generated errors and incorporation of large amounts of RNA into the nuclear genome by primase are avoided by enzymatic removal of primers during Okazaki fragment maturation [14-19]. Thus, the composition and short length of the primer synthesized by polα–primase is an important factor in genomic integrity.

Polα–primase is a heterotetramer composed of two polymerase subunits—catalytic POLA1 (also referred to as p180) and regulatory POLA2 (p70, p68)—and two primase subunits—catalytic PRIM1 (p49, p48) and regulatory PRIM2 (p58) (**Fig. 1B**) [20]. The primase dimer, which is necessary for normal RNA synthesis [21], is formed by interactions between PRIM1 and the N-terminal domain of PRIM2 ($PRIM2_N$), with the C-terminal domain of PRIM2 ($PRIM2_C$) tethered by a flexible linker [22-29]. POLA1, which is functional on its own, relies on POLA2 for heterotetramer formation [30-33]. Structural analysis of the yeast enzyme revealed a bilobal structure in which the catalytic domain of POLA1 ($POLA1_{cat}$) is flexibly tethered via its C-terminal domain ($POLA1_C$) to the tetramer core, consisting of PRIM1, $PRIM2_N$, $POLA1_C$, and the C-terminal domain of POLA2 ($POLA2_C$) [26,30]. In the absence of nucleic acid substrates or protein partners, the human enzyme has only been observed in a compact, non-catalytic ("auto-inhibitory" or AI) configuration in which the active site of $POLA1_{cat}$ is occluded by the tetramer core, as visualized by both X-ray crystallography and electron microscopy (EM) [34,35]. The human enzyme has also been visualized in a catalytically accessible configuration as a complex with the telomere maintenance CTC1–STN1–TEN1 (CST) complex and single-stranded DNA, which

provided a model for how CST orients the active sites of polα and primase for RNA synthesis and intramolecular handoff of the primer [36].

Several models have emerged in recent years that largely attribute the counting ability of polα–primase to the $PRIM2_C$ regulatory domain. Despite the catalytic center of primase residing in PRIM1, $PRIM2_C$ is critical for all aspects of RNA synthesis, including regulation of RNA length [23-25,37,38]. $PRIM2_C$ has also been implicated in transfer of the RNA primer to polα by virtue of its interaction with the primer 5′-triphosphate and the template 3′-overhang [21,34,39,40]. These interactions are believed to dictate RNA counting by spatially limiting the length of RNA synthesized by PRIM1, either via the PRIM2 interdomain linker [40,41] or via steric collision between $PRIM2_C$ and $PRIM2_N$ [33,34]. While these models explain the upper limit of RNA length (i.e., 10 nucleotides), neither explains the range of RNA primers (7–10 nucleotides) extended by polα [12]. $PRIM2_C$ and the primer 5′-triphosphate have also been shown to be important for DNA synthesis [42,43], although the role of $PRIM2_C$ in the transition from RNA to DNA synthesis and in DNA counting by polα is unknown. However, a recent model for DNA synthesis proposed a similar spatial limitation imposed by the POLA1 and PRIM2 interdomain linkers [43]. Alternatively, structures of $POLA1_{cat}$ bound to RNA/DNA and DNA/DNA duplexes afforded a model for DNA counting whereby preferential binding to relatively A-form duplexes by polα leads to decreased affinity and eventual dissociation from the primer/template as a result of a progressive increase in B-form character as DNA is synthesized [32,44].

The molecular details by which polα–primase synthesizes chimeric primers of defined length and composition have remained a matter of debate because there are no structures of the heterotetramer bound to primer/template. Here, we present cryo-EM structures of polα–primase alone and bound to primed templates representing various stages of DNA synthesis, beginning at the point immediately following transfer of the RNA primer to polα. Our structures reveal that dissociation of $POLA1_{cat}$ and $PRIM2_C$ from the AI complex enables a series of configurational and conformational changes that span a remarkable range of motion, both in the absence and the presence of nucleic acid substrates. At the onset of DNA synthesis, we observe a direct interaction between $POLA1_{cat}$ and $PRIM2_C$ that provides an explanation for how polα can extend RNA primers ranging in length from 7–10 nucleotides, and that suggests primer transfer is largely defined by competition between PRIM1 and $POLA1_{cat}$ for the primer 3′-end. We also observe association of $PRIM2_C$ with chimeric RNA-DNA substrates corresponding to intermediate and late stages of DNA synthesis. Complementary biochemical experiments comparing the synthesis activities of polα–primase and various subcomplexes are consistent with $PRIM2_C$ acting as a limited DNA processivity factor, reducing the proportion of both shorter

and longer primers [43]. Additionally, we provide direct structural evidence linking primer composition, duplex conformation, and substrate affinity. Together, these findings establish a comprehensive model for chimeric primer synthesis by polα–primase.

Results

Polα–primase exists in an equilibrium between compact and extended configurations

Prior structural studies have observed polα–primase in one of two states [26,34-36,45,46], suggesting the heterotetramer undergoes a dramatic structural transition from a compact, auto-inhibitory state to an extended, catalytically competent configuration upon binding nucleic acid substrates or protein partners. To better understand the likely equilibrium between the two states, we first analyzed the protein alone using single-particle cryo-EM. *Xenopus laevis* polα–primase was selected based on the stability and the quality of the purified complex (**Fig. 2A**) and the high sequence conservation with human polα–primase (69% identity and 90% similarity) (**Supplemental Figs. S1–S4**). In the absence of nucleic acid substrates, 2D classification sorted 69% of particles into classes with density for all four subunits in the compact AI configuration and the remaining 31% into classes lacking density for POLA1$_{cat}$ (**Fig. 2B and Supplemental Figs. S5 and S6**). We initially refined a partial 3D reconstruction of the AI complex to a resolution of 2.8 Å that exhibited poor density for PRIM1 (**Supplemental Figs. S5 and S7**). To improve the interpretability of the map in this region, we selected a subset of particles that yielded a 3.2-Å reconstruction with clear density for all four subunits (**Fig. 2C and Supplemental Figs. S5 and S7**). Only the small N-terminal domain of POLA2 (POLA2$_N$) was not well defined. As previously observed in lower-resolution structures [34,35,45], the AI complex is stabilized by extensive contacts between POLA1$_{cat}$ and every domain in the tetramer core, as well as PRIM2$_C$. Conversely, PRIM2$_C$ contacts only POLA1$_{cat}$, PRIM2$_N$, and the interdomain linkers in POLA1 (residues 1242–1275) and PRIM2 (residues 262–278). These interactions sterically occlude substrate recognition by both POLA1$_{cat}$ and PRIM2$_C$. To better understand the structural rearrangements necessary to bind nucleic acid substrates, dissociation of POLA1$_{cat}$ and PRIM2$_C$ from the AI complex was examined by 3D variability analysis, using all particles initially aligned on the AI complex during preliminary 3D refinement (**Supplemental Figs. S5 and S8**). Two principal motions were observed: (1) a hinging motion in PRIM2$_N$ that pulls PRIM1 away from POLA1$_{cat}$, and (2) a bending/twisting motion in POLA1$_{cat}$ that pulls the exonuclease domain away from POLA2$_C$ and that begins to pull the palm domain away from PRIM2$_N$.

Interactions involving the thumb domain were unchanged, as were all interactions involving $PRIM2_C$.

Without the conformational restrictions imposed by engagement of $POLA1_{cat}$ in the AI complex, conformational heterogeneity is increased within the tetramer core. The subset of particles grouped into 2D classes that lacked density for $POLA1_{cat}$ could be further divided into classes with clear, partial, or no density for $PRIM2_C$ (**Fig. 2B and Supplemental Figs. S5 and S6**). 3D reconstructions generated from these particles revealed three distinct conformations characterized by hinging motions in which the PRIM1 and $POLA1_C$/$POLA2_C$ lobes move relative to one another by pivoting around flexible loops in $PRIM2_N$ (residues 153–157 and 216–219) and $POLA1_C$ (residues 1439–1443), respectively (**Fig. 2D,E and Supplemental Fig. S5**). Together, these data demonstrate that in the absence of nucleic acid substrates, polα–primase exists predominantly in a compact auto-inhibitory state, but also adopts a series of extended configurations in which $POLA1_{cat}$ and $PRIM2_C$ are disengaged from the other domains in the complex. The 3D variability analysis of the AI complex, along with the varying degrees of $PRIM2_C$ engagement with the tetramer core, suggests that dissociation of $POLA1_{cat}$ precedes dissociation of $PRIM2_C$, and initiates the structural transition away from an auto-inhibitory state.

<u>$PRIM2_C$ remains bound to the primer during DNA synthesis</u>

Polα–primase possesses dual enzymatic activities that enable synthesis of chimeric RNA-DNA primers. To investigate the molecular bases for the transition from RNA to DNA synthesis and the low processivity of polα, we designed substrates to visualize polα–primase at two key points during DNA synthesis—immediately after intramolecular handoff of the RNA primer from primase to polα (DNA initiation, DI) and after polα has elongated the primer with DNA (DNA elongation, DE). The DI primer comprises nine ribonucleotides, and the chimeric DE primer comprises nine ribonucleotides followed by 11 deoxyribonucleotides (**Fig. 3A and Supplemental Fig. S9**). Both primers include the biologically relevant 5′-triphosphate group that would be introduced by primase during RNA initiation. Ternary complexes containing polα–primase, primer/template, and incoming nucleotide were trapped using 2′,3′-dideoxycytidine to terminate the primer strand and prevent phosphodiester bond formation with dGTP.

Relative to polα–primase in the absence of substrate, the 2D class averages determined in the presence of the DI and DE substrates showed increased configurational heterogeneity and dramatic configurational rearrangements (**Fig. 3B and Supplemental Figs. S10–S13**). The vast majority of the class averages lacked clear density for part of the complex, indicating a large degree of flexibility and conformational/configurational heterogeneity remains when

substrates are bound. These class averages could be divided into two groups: those with density for POLA1$_{cat}$ and PRIM2$_C$ bound to substrate (hereafter referred to as the DNA synthesis assembly, or DSA), and those with density for everything else (i.e., the tetramer core). With the DI substrate, 2D classes showed 67% of particles with clear density for only the DSA, 26% for only the tetramer core, and ~6% for polα–primase in various configurations. With the DE substrate, 2D classes showed a similar distribution, but with fewer particles aligned on the DSA and more particles aligned on the tetramer core. Only 0 and 3% of particles with the DI and DE substrates, respectively, were found to be in the AI configuration.

To better characterize interactions with the substrates, we performed focused 3D refinement using masks around POLA1$_{cat}$, PRIM2$_C$, and the primer/template (**Fig. 3C and Supplemental Figs. S10 and S12**). With both the DI and DE substrates, the 3.3-Å and 4.4-Å reconstructions showed POLA1$_{cat}$ and PRIM2$_C$ bound to opposing ends of the duplexes (**Fig. 3C**), indicating that interactions with PRIM2$_C$ are maintained during transfer of the primer to polα and throughout DNA synthesis. Focused 3D refinement of only POLA1$_{cat}$ and the DNA/DNA portion of the DE substrate improved the resolution of the reconstruction to 3.5 Å (**Supplemental Figs. S12 and S14**). In both DSA structures, POLA1$_{cat}$ is in a closed ternary complex with the primer 3′-end and the incoming dGTP, and PRIM2$_C$ is bound to the primer 5′-triphosphate. With the DI substrate, POLA1$_{cat}$ and PRIM2$_C$ are in contact with one another, albeit with the entire subunit interface consisting of a single interaction between two loops (**Fig. 3C and Supplemental Fig. S15**). With the longer DE substrate, the two domains are separated by a full turn of the RNA-DNA/DNA duplex. 3D variability analysis of these DI and DE subcomplexes revealed dissociation of PRIM2$_C$ from a subset of the particles, as well as bending of the oligonucleotide duplexes (**Supplemental Fig. S16**). With the shorter DI substrate, this bending altered the protein-protein interface between POLA1$_{cat}$ and PRIM2$_C$. Prompted by the seeming malleability and small size of the subunit interface, we modeled POLA1$_{cat}$ and PRIM2$_C$ with shorter RNA primers (**Supplemental Fig. S15**). With 8-mer and 7-mer primers, we observed only minor clashes between side chains, which appear unlikely to prevent DNA initiation. Conversely, we saw severe steric overlap between POLA1$_{cat}$ and PRIM2$_C$ with a 6-mer primer. These models explain the previous observation that polα can extend RNA primers as short as seven nucleotides in length [12]. Additionally, binding of the nascent primer by PRIM2$_C$ during DNA elongation explains the earlier observation that the primer 5′-triphosphate affects not only RNA synthesis but also DNA synthesis [42], as well as the recent finding that PRIM2$_C$ increases the processivity of polα–primase [43].

For simultaneous binding of the primer ends to be maintained throughout DNA synthesis, as the primer increases in length, large-scale protein conformational changes are necessary. Because of the extreme conformational and configurational heterogeneity inherent in the polα–primase complex, consensus reconstructions were not possible. Nonetheless, the structures of the DSA and tetramer core subcomplexes enabled us to model the complete polα–primase complex bound to DI and DE substrates (**Fig. 3D,E**). Using the 2D class averages as a guide to manually dock the subcomplexes, 3D models were generated for the two predominant configurations observed with each substrate (**Fig. 3D**). Comparison of the various polα–primase models showed that the DSA rotates by as much as 180° with respect to the tetramer core, with the pivot point defined by weak interactions between $PRIM2_C$ and $PRIM2_N$ (**Fig. 3E**). In addition to this swiveling, the DSA rotates by nearly 180° with respect to the helical axis of the oligonucleotide duplexes. In all models of the four-subunit complex, the substrate-bound configurations appear to be stabilized by interactions between $POLA1_{cat}$ and the tetramer core. However, the variability of these interactions, as well as the large proportion of 2D class averages showing density for only one or the other subcomplex, suggests these protein-protein contacts are weak and transient. Thus, in the presence of primer/template substrates representing two key stages of DNA synthesis, polα–primase adopts a wide range of configurations characterized by concerted movement of $POLA1_{cat}$, $PRIM2_C$, and the oligonucleotide duplex relative to the tetramer core.

Primase enhances an intrinsic counting ability in polα

Whether counting of DNA nucleotides is an inherent property of polα or only a property of the complete polα–primase complex is unclear [32,34,43,47,48]. Our finding that $PRIM2_C$ remains bound to the primer after transfer to polα reveals a means by which primase might participate in counting during DNA synthesis. We theorized that binding of the substrate by $PRIM2_C$ would alter DNA elongation by indirectly tethering $POLA1_{cat}$ to the primer via the tetramer core. To investigate this theory, we compared the RNA and DNA synthesis activities of the polα–primase complex with a complex in which the interdomain linker between $POLA1_{cat}$ and $POLA1_C$ was severed, constructed by mixing purified $POLA1_{cat}$ and $\Delta POLA1_{cat}$ (PRIM1/PRIM2/$POLA1_C$/POLA2) (**Fig. 4A and Supplemental Fig. S17**). As a control, we measured the RNA and DNA synthesis activities of $POLA1_{cat}$ and $\Delta POLA1_{cat}$ separately. Primer extension was performed by adding radiolabeled ATP or dATP and unlabeled NTPs and dNTPs to a 5′-triphosphorylated 6-ribonucleotide primer annealed to a 99-deoxyribonucleotide template (**Fig. 4B**). $POLA1_{cat}$ alone generated extension products of predominantly 17–33 nucleotides,

with the main product being 30 nucleotides in length (**Fig. 4A,C**). Similar to previous observations with primase alone [21,29], ΔPOLA1$_{cat}$ rapidly formed products of nine nucleotides before abruptly terminating at a 10-mer primer (**Fig. 4A**). The severed complex generated by combining POLA1$_{cat}$ and ΔPOLA1$_{cat}$ afforded products representing an admixture of the individual control reactions (**Fig. 4A,C and Supplemental Fig. S18**). In contrast, intact polα–primase generated fewer extension products outside of 30 nucleotides (**Fig. 4A,C**), and produced the 30-mer product at a dramatically faster rate (**Fig. 4D**). These data demonstrate an intrinsic—albeit approximate—counting ability in polα that is improved by the physical linkage between POLA1$_{cat}$ and the remainder of the complex, reinforcing the preference of polα–primase to generate a 30-nucleotide primer.

Primer/template conformation modulates substrate recognition

The intrinsic counting ability of POLA1$_{cat}$ has previously been attributed to conformation-dependent recognition of RNA/DNA and DNA/DNA duplexes [32,44]. Because of the relatively high resolutions of the DSA reconstructions, we were able to determine the conformations of the biologically relevant 5′-triphosphorylated RNA/DNA and RNA-DNA/DNA substrates and to evaluate the details of their recognition by POLA1$_{cat}$ and PRIM2$_{C}$. In both the DI and DE subcomplexes, POLA1$_{cat}$ was observed in a closed conformation in which the fingers domain was fully engaged with an incoming dGTP and active-site residues within the palm domain (Asp859, Phe860, and Asp1000) coordinated a catalytic Mg^{2+} ion (**Fig. 5A,B and Supplemental Fig. S19**). The thumb domain was also fully engaged with both the RNA/DNA and RNA-DNA/DNA duplexes (**Fig. 5A,C**). To our knowledge, the DE subcomplex represents the first structure of POLA1$_{cat}$ with the thumb and fingers domains engaged with a DNA/DNA duplex and an incoming dNTP (**Supplemental Fig. S20**). In both the DI and DE subcomplexes, POLA1$_{cat}$ induces a local A-form conformation in the duplexes that positions and orients the 2′,3′-dideoxycytidine at the 3′-end of the primers for phosphodiester bond formation with dGTP (**Fig. 5B,E,F**). Similar catalytic arrangements have been observed in structures of polδ and polε [49,50]. However, POLA1$_{cat}$ possesses a wider substrate-binding cleft that results from structural differences in the thumb domain (**Supplemental Fig. S21**). This wider cleft has been proposed to complement the wider diameter of an A-form RNA/DNA duplex [32,44]. In both of our structures, we observed intermediate AB-form conformations [51,52], or nearly intermediate AB-form conformations, over most of the region of the duplexes bound by POLA1$_{cat}$ (**Fig. 5E,F**).

Importantly, small but significant differences in the interactions between the thumb domain and the DI and DE substrates illustrate how POLA1$_{cat}$ accommodates different minor groove

geometries (**Fig. 5A,C,E,F**). With the DI substrate, three aspartate residues (Asp998, Asp1079, and Asp1144) form hydrogen-bonding interactions with 2′-hydroxyl groups along the RNA primer. With the DE substrate, conformational changes at the tip of the thumb domain (residues 1142–1148) pull Asp1144 away from the chimeric primer while directing Lys1145 and Lys1146 toward the relatively narrow minor groove in the same DNA/DNA region of the duplex. These new hydrogen-bonding interactions likely partially offset the loss of RNA-specific hydrogen bonds with the DE substrate. While RNA/DNA and DNA/DNA duplexes generally adopt A-form and B-form geometries [53], respectively, the intermediate AB-form conformations observed with both the DI and DE substrates suggest $POLA1_{cat}$ remodels the primer/template during DNA synthesis. The local propensity of the duplex to adopt this intermediate conformation is likely correlated with substrate affinity and, consequently, with processivity [32,44].

$PRIM2_C$ has been shown to increase the processivity of polα–primase during DNA synthesis [43]. The two DSA structures indicate that this increased processivity is the result of continued association of $PRIM2_C$ with the primer/template during DNA elongation. In both the DI and DE subcomplexes, interactions between $PRIM2_C$ and the primer strand are identical (**Fig. 5A,D**). Three positively charged residues (His306, Arg308, and Arg312) at the N-terminus of helix α15 surround the 5′-triphosphate and offset the negative charge, and His309 stacks against the 5′-terminal nucleobase. Surprisingly, interactions between $PRIM2_C$ and the template strand are markedly different in the two subcomplexes, despite $PRIM2_C$ binding the same RNA/DNA portion of both substrates. In the DI structure, the DNA template is positioned such that only the last paired nucleotide, which is not coplanar with the 5′-terminal nucleotide in the RNA primer, and the first unpaired nucleotide form extensive contacts with $PRIM2_C$. Rotation of the phosphodeoxyribose backbone swings the second and third unpaired nucleotides away from $PRIM2_C$ and toward $POLA1_{cat}$. In the DE structure, the DNA template is positioned such that the last paired nucleotide, which is coplanar with the 5′-terminal nucleotide in the RNA primer, and the first, second, and third unpaired nucleotides form extensive contacts with $PRIM2_C$. Relative to the DI substrate, interactions with several residues (Tyr353, His357, Lys364, and Tyr368) in the flexible template-binding loop (residues 359–372) are shifted by one nucleotide toward the 5′-end of the template strand. Virtually identical interactions were observed in an X-ray crystal structure of human $PRIM2_C$ bound to a short RNA/DNA duplex [34] (**Supplemental Fig. S22**). The unique interactions that we observe with the DI substrate appear to result from local duplex remodeling by $POLA1_{cat}$, which causes the RNA/DNA duplex to have less A-form character than either the crystal structure or the RNA/DNA portion of the DE substrate (**Fig. 5D,E,F and Supplemental Fig. S22**). With reduced A-form character, the DI

substrate has less base pair inclination, which is pronounced in A-form duplexes but greatly reduced in B-form conformations [53], pulling the 3′-overhang away from $PRIM2_C$. The affinity of $PRIM2_C$ for the primer/template was previously shown to depend on both the 5′-triphosphate and the 3′-overhang [34,40]. Thus, tight binding of the substrate appears to require substantial inclination of the base pairs at the end of duplex. Together, these data indicate that like $POLA1_{cat}$, the affinity of $PRIM2_C$ for the primer/template is modulated by the nucleic acid conformation, which is imposed by both the DNA composition of the primer and binding by $POLA1_{cat}$.

Loose association of full-length primers allows for reversion to the auto-inhibitory configuration

The finding that substrate recognition by both $POLA1_{cat}$ and $PRIM2_C$ is dependent on duplex conformation has implications for the mechanism of DNA counting. To investigate how duplex conformation affects the preference of polα–primase to generate 30-nucleotide primers, we designed a substrate to visualize the complex at the end of DNA synthesis (DNA termination, DT). The DT primer comprises nine ribonucleotides followed by 20 deoxyribonucleotides (**Fig. 6A and Supplemental Fig. S9**) and includes the same 5′-triphosphate group as the DI and DE primers. A ternary DT complex containing polα–primase, primer/template, and incoming nucleotide was trapped using 2′,3′-dideoxycytidine to terminate the primer strand and prevent phosphodiester bond formation with dGTP.

As with the DI and DE substrates, a substantial proportion of 2D class averages showed density for only the DSA or the tetramer core (**Fig. 6B and Supplemental Figs. S23 and S24**), consistent with large-scale conformational and configurational dynamics. Unlike with the shorter substrates, however, no class averages showed density for polα–primase bound to the DT substrate. This could indicate an even greater degree of heterogeneity and/or the lack of possible protein-protein contacts between $POLA1_{cat}$ and the remainder of the complex when bound to the longer oligonucleotide duplex. Regardless, this level of heterogeneity suggests that the interdomain linkers in POLA1 and PRIM2 remain slack when bound to a full-length primer. Surprisingly, 63% of particles were grouped into 2D classes that are clearly representative of the AI complex. Protein-protein contacts in this configuration sterically occlude substrate recognition by both $POLA1_{cat}$ and $PRIM2_C$, preventing binding of the DT substrate. With the DI and DE substrates, only 0 and 3% of particles, respectively, were found to be in the AI configuration, while in the absence of substrate, 69% of particles were found to be in this configuration. If the percentage of particles found to be in the AI configuration is used as a

metric for substrate affinity, this trend suggests progressively weaker binding of longer substrates by both $POLA1_{cat}$ and $PRIM2_C$.

To better understand the structural basis for seemingly weaker association of the DT substrate, we performed homogeneous refinement using the particles aligned on the DSA subcomplex (**Fig. 6C and Supplemental Fig. S23**). A preliminary 3.9-Å reconstruction showed clear density for $POLA1_{cat}$ in a closed conformation, with the fingers and thumb domains fully engaged with an incoming dGTP and an RNA-DNA/DNA duplex. As with the DE substrate, the DNA/DNA portion of the duplex bound by $POLA1_{cat}$ appears to have an intermediate AB-form conformation. However, the central DNA/DNA portion of the duplex outside the substrate-binding cleft of $POLA1_{cat}$ has a distinctly B-form conformation with a wide, shallow major groove and a narrow, deep minor groove. The conformation becomes noticeably more intermediate AB-form, and possibly A-form, near and in the RNA/DNA portion of the duplex. Density for $PRIM2_C$ in the reconstruction is weak, likely because of bending of the duplex and/or dissociation of $PRIM2_C$ from the substrate. Relative to the DE subcomplex, density for $PRIM2_C$ in the DT subcomplex is clearly much weaker, suggesting a greater extent of $PRIM2_C$ dissociation. We speculate that weaker association of $PRIM2_C$ is the result of a less A-form RNA/DNA portion of the duplex, caused by proximity to the distinctly B-form DNA/DNA portion of the DT substrate.

Discussion

Multiple, seemingly conflicting models have been proposed to explain the counting abilities of primase and polα [32-34,40,41,43,44]. Our structures of polα–primase create a framework to reconcile and integrate these models, providing a comprehensive mechanistic understanding of nucleotide counting (**Fig. 7A,B**). We show that in the absence of nucleic acid substrates or protein partners, polα–primase exists primarily in a compact, auto-inhibitory state, but spontaneous dissociation of $POLA1_{cat}$ initiates a series of dramatic structural rearrangements necessary for chimeric primer synthesis across two spatially separate active sites. In all current models of RNA synthesis, primer formation is initiated by PRIM1 and $PRIM2_C$ bringing together two NTPs on a single-stranded DNA template. In the linker model, elongation of the primer continues until halted by full extension of the PRIM2 interdomain linker [40,41]. In the collision model, elongation is ceased by a steric clash between $PRIM2_N$ and $PRIM2_C$ [34]. As previously discussed [34], the linker model is not consistent with domain arrangement in the tetramer core subcomplex. Conversely, the collision model is supported by structural data and structure-based

modeling [34], and provides a convincing rationale for termination of RNA synthesis, but strictly at 10 nucleotides. Both models fail to explain previous observations showing extension of RNA primers as short as seven nucleotides in length by polα [12]. PRIM2$_C$ has been suggested to anchor the primer during its transfer between primase and polα active sites by remaining bound to the primer 5′-end [34,36,43]. Moreover, the highly distributive nature of RNA and DNA synthesis by PRIM1 and POLA1$_{cat}$ in the absence of PRIM2$_C$ [29,43,54] suggests that PRIM2$_C$ acts as a hub to facilitate re-association of PRIM1 and POLA1$_{cat}$ with the primer. Our structures confirm that PRIM2$_C$ remains bound to the 5′-end of the primer after transfer to polα, and provide evidence that a dynamic protein-protein interface allows PRIM2$_C$ and POLA1$_{cat}$ to simultaneously bind RNA primers of seven or more nucleotides, while transfer of RNA primers less than seven nucleotides in length to POLA1$_{cat}$ would be sterically blocked by PRIM2$_C$. Thus, the likely driving factor in the primase-to-polymerase switch is a competition between PRIM1 and POLA1$_{cat}$ for the 3′-end of the primer, rather than termination of RNA synthesis at a maximal length of 10 nucleotides. In this competition model, steric collision between PRIM2$_N$ and PRIM2$_C$ would serve only as a failsafe to terminate RNA synthesis, ensuring longer stretches of RNA (> 10 nts) are not incorporated into the genome, should POLA1$_{cat}$ not outcompete PRIM1 for the primer.

Our finding that PRIM2$_C$ remains bound to the primer after transfer to polα has far-reaching implications for DNA synthesis, including the preference of polα–primase to generate chimeric primers of ~30 nucleotides. Simultaneous association of both PRIM2$_C$ and POLA1$_{cat}$ is a prerequisite for the linker model of DNA termination, which posits that the POLA1 and PRIM2 interdomain linkers prevent separation of PRIM2$_C$ and POLA1$_{cat}$ beyond the distance associated with an ~30-mer primer [43]. However, we observed seemingly weaker binding of chimeric primers by both POLA1$_{cat}$ and PRIM2$_C$ as the DNA composition of the primer increased (as measured by the proportion of particles in the AI configuration). Additionally, the linker model is unable to provide a rationale for the intrinsic counting ability of POLA1$_{cat}$, which we and others have observed [32,47]. By contrast, the affinity model of DNA termination, which proposes that POLA1$_{cat}$ preferentially recognizes A-form duplexes typical of RNA/DNA hybrids, and loses affinity for the substrate as the duplex becomes more B-form throughout DNA synthesis [32,44], provides a mechanistic explanation for the innate counting ability of the polymerase catalytic core. Our data expand upon the affinity model by showing preferential binding of A-form duplexes by PRIM2$_C$, and by providing a structural basis for this specificity. The relatively high affinity that PRIM2$_C$ possesses for A-form RNA/DNA duplexes [34,40] is consistent with its roles in RNA synthesis and intramolecular handoff from primase to polα, as described above. However, the relatively low affinity that PRIM2$_C$ possesses for chimeric primers with increasingly greater proportions of DNA

limits the range of products for which it is likely to act as a processivity factor during DNA elongation. Because the linker model requires both $POLA1_{cat}$ and $PRIM2_C$ to remain bound to the substrate, the interdomain linkers are likely to play, at most, a failsafe role similar to that proposed above for RNA synthesis, restricting chimeric primer length to reduce the mutagenic potential of polα.

The dramatic structural rearrangements that we observe are consistent with the idea that flexible tethering of $POLA1_{cat}$ and $PRIM2_C$ to the tetramer core is necessary for transfer of the nascent primer from primase to polα [26,34]. As we demonstrate here, these rearrangements are also necessary for simultaneous binding of the primer ends by $POLA1_{cat}$ and $PRIM2_C$ during DNA elongation, as the primer increases in length. Presumably, this plasticity also facilitates interaction with the many protein partners with which polα–primase is known to interact at replication forks and telomeres. At replication forks, polα–primase interacts with multiple proteins associated with the replisome, including MCM10, AND-1, RPA, RFC, polδ, and the CMG helicase [19,55], as well as several nucleosome core proteins [56,57]. Despite this extensive network of contacts, polα–primase appears to remain conformationally dynamic [58,59]. However, it is unclear if polα–primase simultaneously interacts with all or only a subset of these proteins, as well as what initiates the transition from the compact auto-inhibitory configuration to an extended catalytic state. Many of these interactions involve contacts with the unstructured N-terminal tail of POLA1, leading us to wonder if these interactions may stimulate disengagement of $POLA1_{cat}$ from the AI complex, enabling POLA1 to act as a sensor for activation of polα–primase.

Polα–primase appears to remain similarly dynamic at telomeres, as three distinct configurations of the polα–primase/CST complex bound to single-stranded DNA have been observed [36,45,46]. The first structure, termed a recruitment complex, showed polα–primase in the AI configuration with CST primarily associated via interactions with $POLA1_C$ [45]. The structural element in $POLA1_C$ involved in these contacts was determined to be specific to metazoans. Consistent with this finding, the AI configuration has only been observed with homologs of polα–primase from higher eukaryotes, including *Homo sapiens* [34,35] and now *Xenopus laevis*, while *Saccharomyces cerevisiae* and *Tetrahymena thermophila* polα–primase have only been observed in extended configurations [26,46]. The second structure of polα–primase/CST, termed a pre-initiation complex, showed polα–primase in an extended configuration in which $POLA1_{cat}$ and the tetramer core are scaffolded by CST [36]. In this configuration, the PRIM1 and $POLA1_{cat}$ active sites are locked in opposing orientations, with $PRIM2_C$ free to move between them. This arrangement was proposed to facilitate RNA synthesis and transfer of the RNA primer from

primase to polα. Moreover, CST was shown to enhance primer synthesis, without affecting primer length [36,60].

Docking of our DI subcomplex structure into the pre-initiation complex provided further insight into primer transfer (**Supplemental Fig. S25**). The model revealed additional interactions between the thumb domain of $POLA1_{cat}$ and the CTC1 and STN1 subunits of CST, which may stabilize the closed, substrate-bound conformation of the polymerase catalytic core, supporting $POLA1_{cat}$ in its competition with PRIM1 for the RNA primer. When bound to the DI substrate, the close proximity of $PRIM2_C$ to $PRIM2_N$ and PRIM1 seems likely to prevent DNA elongation beyond incorporation of the first dNTP. This steric hindrance suggests rearrangement or partial dissociation of the polα–primase/CST complex is required prior to DNA elongation. Consistent with this observation, docking of our DE subcomplex structure introduced a severe steric clash between $PRIM2_C$ and $PRIM2_N$ (**Supplemental Fig. S25**), which seems unlikely to be avoided by bending of the RNA-DNA/DNA duplex. The third structure of polα–primase/CST showed $POLA1_{cat}$ associated with CST, as in the pre-initiation complex, but with the tetramer core disengaged [46]. Together, these observations suggest the tetramer core dissociates prior to DNA elongation by $POLA1_{cat}$. Based on the evolutionarily relationship between CST and RPA [61], we believe RPA likely scaffolds polα–primase in a similar manner, explaining RPA-dependent stimulation of primer synthesis by polα–primase [62]. How these interactions with RPA and CST are coordinated with those of other proteins at replication forks and telomeres, as well as the effects these interactions may have on the conformational and configurational dynamics of polα–primase and the mechanism of chimer primer synthesis, remain to be determined.

Conclusion

Determining the first structures of polα–primase with catalytic substrates, along with performing a structure-based biochemical analysis, has enabled us to propose a comprehensive model for primer synthesis. This model is built upon the large body of available structural and biochemical data while unifying and refining prevailing understanding of the mechanistic basis for polα–primase function. Our structural data reveal that polα–primase both remodels and reads the conformation of its primer/template substrate, which explains its unique ability to regulate primer length (i.e., count). These data also reveal that polα–primase retains a remarkable degree of conformational and configurational flexibility as it synthesizes primers that is crucial to the intricate functions of this complex and dynamic protein machine. A greater appreciation of the type and degree of intrinsic flexibility that we have observed for polα–

primase is anticipated as active eukaryotic replisomes and other cellular machinery come into view at ever increasing resolution.

Methods

Cloning and expression. The genes encoding *Xenopus laevis* PRIM1 (NCBI accession NP_001080819.1), PRIM2 (NCBI accession XP_018121015.1), POLA1 (NCBI accession NP_001082055.1), and POLA2 (NCBI accession NP_001086972.1) were initially cloned into pFastBac-HTa or pFastBac-HTb expression vectors (Invitrogen). PRIM1 was then subcloned into a modified pET-19b expression vector (Novagen) encoding an HRV 3C-cleavable decahistidine fusion tag. PRIM2 was subcloned into a modified pRSFDuet-1 expression vector (Novagen). And, POLA1 and POLA2 were subcloned into an extensively modified, polygenic pFastBac-1 expression vector (Invitrogen) encoding TEV-cleavable StrepII-GFP (POLA1) and hexahistidine-MBP (POLA2) fusion tags. Protein expression and purification were facilitated by truncating the gene encoding POLA1 to remove the N-terminal tail (residues 1–334). To generate a $POLA1_{cat}$ construct, the catalytic core of POLA1 (residues 335–1254) was subcloned into a modified pET-27b(+) expression vector (Novagen) encoding an HRV 3C-cleavable hexahistidine-SUMO fusion tag. To generate a $\Delta POLA1_{cat}$ construct, the C-terminal domain of POLA1 (residues 1262–1458) and POLA2 were subcloned into a modified, polycistronic pCDFDuet-1 expression vector (Novagen) encoding an HRV 3C-cleavable hexahistidine ($POLA1_C$) fusion tag. Production of baculovirus and expression of POLA1/POLA2 in *Spodoptera frugiperda* Sf9 cells were performed by the Expression and Microbiology Core at UC Berkeley. PRIM1/PRIM2, $POLA1_{cat}$, and $POLA1_C$/POLA2 were expressed in *Escherichia coli* Rosetta 2(DE3) cells. Cultures were grown at 37°C in LB medium supplemented with 34 mg/L chloramphenicol, 100 mg/L ampicillin (PRIM1/PRIM2), 30 mg/L kanamycin (PRIM1/PRIM2 and $POLA1_{cat}$), and 70 mg/L streptomycin ($POLA1_C$/POLA2). Upon reaching mid-log phase, cultures were further supplemented with 100 mg/L ammonium iron(III) citrate (PRIM1/PRIM2) and 5 µM $ZnSO_4$ (PRIM1/PRIM2 and $POLA1_C$/POLA2) and cooled to 16°C. Protein overexpression was then induced by addition of IPTG to a final concentration of 0.5 mM. After overnight incubation, cells were harvested by centrifugation. Cell pellets were stored at −80°C. *Acanthamoeba castellanii* TLP2 was expressed as previously described [63-65].

Protein purification. Cell pellets were thawed in lukewarm water and resuspended in buffer L [50 mM sodium phosphate pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, 10 mM imidazole, and cOmplete EDTA-free protease inhibitor cocktail (Roche)]. Both insect cells and bacterial cells were lysed by Dounce homogenization followed by gentle sonication. The two lysates were then combined and cleared by centrifugation prior to application onto a Ni-NTA Agarose column (Qiagen) equilibrated in buffer L lacking protease inhibitor cocktail. The column was washed with 15 column volumes (cv) of buffer N1 [50 mM sodium phosphate pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 20 mM imidazole], followed by 5 cv of buffer N2 [50 mM HEPES•NaOH pH 7.5, 300 mM NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 20 mM imidazole]. Protein was eluted by linearly transitioning over 15 cv to buffer N3 [50 mM HEPES•NaOH pH 7.5, 300 mM NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 500 mM imidazole]. Fractions containing polα–primase were pooled and supplemented with 1 mM TCEP before fusion tags were removed by overnight cleavage at 4°C. The following day, protein was diluted three-fold in buffer H1 [50 mM HEPES•NaOH pH 7.5, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 1 mM TCEP] and passed through a 0.22-µm filter before being applied to a heparin Sepharose column (Cytiva) equilibrated in a 9:1 mixture of buffer H1 and buffer H2 [50 mM HEPES•NaOH pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 1 mM TCEP]. The column was then washed with 5 cv of the same 9:1 mixture before protein was eluted by linearly transitioning over 10 cv to only buffer H2. Fractions containing polα–primase were pooled and again applied to a Ni-NTA Agarose column equilibrated in buffer L lacking protease inhibitor cocktail. Protein was eluted in 10 cv of the same buffer, and fractions containing polα–primase were pooled and supplemented with 1 mM TCEP. Protein was then concentrated by diafiltration, cleared by centrifugation, and applied to a Superose 6 Increase column (Cytiva) equilibrated in buffer S [20 mM HEPES•NaOH pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM TCEP]. Protein was eluted in 1.5 cv of the same buffer, and fractions containing polα–primase were pooled and concentrated by diafiltration. Aliquots were used immediately or frozen in liquid nitrogen and stored at −80°C. $POLA1_{cat}$ and $\Delta POLA1_{cat}$ were purified in the same manner as polα–primase. AcaTLP2 was purified as previously described [63-65].

Oligonucleotide synthesis and purification. RNA, DNA, and chimeric RNA-DNA oligonucleotides were purchased from Integrated DNA Technologies and resuspended in buffer A (10 mM Mes•NaOH pH 6.5 and 40 mM NaCl). To generate 5′-triphosphorylated primers, 500 µM 5′-monophosphorylated primer and 550 µM template were annealed by heating to 85°C and then slowly cooling to 22°C, producing a double-stranded substrate for the tRNA-modifying

enzyme AcaTLP2 [65,66]. AcaTLP2 reaction mixtures were prepared with 50 μM enzyme, 100 μM primer/template, 150 μM ATP, 1 mM GTP, 25 mM Hepes•NaOH pH 7.5, 125 mM NaCl, 10 mM $MgCl_2$, and 3 mM DTT and incubated at 22°C. After 1 h, an equal volume of quenching/loading buffer containing 80% (v/v) formamide, 100 mM EDTA pH 8.0, and 220 μM competitor oligonucleotide was added. Samples were then heated to 85°C and slowly cooled to 22°C to denature the primer and the template. Oligonucleotides were separated by denaturing urea-polyacrylamide gel electrophoresis (PAGE) as previously described [67], except electrophoresis was performed at 30 W for 9 h. For each reaction, the band corresponding to the triphosphorylated primer was located by UV-shadowing and cut from the gel. The primer was then extracted into 24 mL of 300 mM NaOAc pH 5.5 using a crush-and-soak procedure in which the sample was placed at −80°C until frozen, thawed in a water bath at 42°C, and then rocked overnight at 4°C. The following day, the supernatant was decanted and overnight extraction was repeated using a second 24-mL volume of 300 mM NaOAc pH 5.5. The two volumes were combined, passed through a 0.22-μm filter, and split into four 12-mL aliquots. 36 mL of cold 100% ethanol was then added to each aliquot. Samples were thoroughly mixed, incubated at −80°C for 2 h, warmed to −20°C, and centrifuged at 12,000×g and 4°C for 1 h. After decanting the supernatant, pellets were air-dried for ~15 min and resuspended in 1 mL of 300 mM NaOAc pH 5.5, using the same 1 mL of buffer for all pellets. The sample was then split into two 0.5-mL aliquots, and 1.5 mL of cold 100% ethanol was added to each. Samples were again thoroughly mixed, incubated at −80°C for 2 h, warmed to −20°C, and centrifuged at 12,000×g and 4°C for 1 h. After decanting the supernatant, pellets were air-dried for ~15 min and then gently rinsed with 200 μL of cold 95% (v/v) ethanol. The ethanol was removed and the pellet was air-dried for an additional ~15 min. The triphosphorylated primer was then resuspended in 100 μL of buffer A, using the same 100 μL of buffer for both pellets, and stored at −80°C. Oligonucleotide sequences are provided in **Supplemental Table S3**.

Cryo-EM sample preparation and data collection. Polα–primase was concentrated to ≥ 3 μM by diafiltration. To generate RNA-primed templates, 110 μM 5′-triphosphorylated RNA primer and 100 μM DNA template were annealed by heating to 85°C and then slowly cooling to 4°C. Substrate complexes were prepared by incubating 3 μM polα–primase, 3.6 μM primer/template, and 1 mM dGTP on ice for 30 min. Using a Vitrobot Mark IV (FEI), 2.5 μL of sample was applied to glow-discharged, holey-carbon R1.2/1.3 200-mesh copper grids (Quantifoil) at 4°C and 100% humidity. Grids were immediately blotted with filter paper (Ted Pella) and flash-frozen in liquid ethane. Grids prepared without substrate or with DT substrate were loaded into a Titan Krios G4

electron microscope (Thermo Scientific) operated at 300 kV for automated data collection using SerialEM [68]. Movies of dose-fractionated frames were acquired with a K3 BioQuantum direct electron detector (Gatan) functioning in counting mode. Each of the datasets was collected from a single grid in a single imaging session. Grids prepared with DI or DE substrates were loaded into a Talos Arctica electron microscope (Thermo Scientific) operated at 200 kV for automated data collection using Leginon [69]. Movies of dose-fractionated frames were acquired with a K2 Summit direct electron detector (Gatan) functioning in counting mode. Each of the datasets was collected in three to four imaging sessions, using a different grid for each session. Additional details for each of the datasets are provided in **Supplemental Tables S1 and S2**.

Cryo-EM data processing. For pre-processing of all datasets, motion-correction of raw movies was performed using MotionCor2 [70] and patch CTF estimation was performed for motion-corrected micrographs using cryoSPARC [71]. All initial processing steps were also performed in cryoSPARC. Particles from each dataset were picked by a combination of blob picking, template picking, and neural-network picking using Topaz [72]. Templates for template picking were selected from the highest-resolution 2D class averages resulting from blob picking, while Topaz was trained from a small set of manually picked particles. Particles were extracted with a box size of either 400 (substrate-free and DT datasets) or 360 (DI and DE datasets) pixels and then downsampled to 120 pixels (pixel size of 2.73 Å) by Fourier truncation. 2D classification was first performed separately for particles picked by each method. Selected particles were then combined and duplicate and triplicate particles were removed before final 2D classification was performed. For the DT dataset, particles were sorted manually before initial 3D reconstructions were generated by ab initio reconstruction. For all other datasets, initial 3D reconstructions were generated by multiclass ab initio reconstruction followed by heterogeneous refinement. Satisfactory 3D reconstructions could only be obtained for the AI complex (substrate-free and DT datasets), the DSA (DI, DE, and DT datasets), and the tetramer core (all datasets). Because of extreme configurational and conformational heterogeneity, satisfactory 3D reconstructions could not be obtained for the full polα–primase complex bound to a primer/template substrate. For reconstructions of the tetramer core, conformational heterogeneity was reduced through additional iterations of multiclass ab initio reconstruction and heterogeneous refinement. For all 3D reconstructions, homogeneous refinement was initially performed using downsampled particles and then repeated after particles were recentered and re-extracted either without downsampling or with reduced downsampling. For the AI complex from the substrate-free dataset and the DSA from the DI and DE datasets, all subsequent processing steps were

performed in RELION-3.1 [73]. Particle alignments were first improved by focused 3D refinement, before heterogeneity was reduced through focused, alignment-free 3D classification (following particle subtraction, as necessary). CTF refinement was also performed to correct for per-particle defocus differences, as well as for imperfect microscope alignment during data collection. Using refined CTF parameters, focused 3D refinement was repeated for each final set of particles. To improve high-resolution features in the reconstructions, despite local quality differences, maps were sharpened using DeepEMhancer [74]. Data-processing workflows are outlined in **Supplemental Figs. S5, S10, S12, and S23**.

Model building and refinement. Homology models of *X. laevis* polα–primase subunits were generated from corresponding *Homo sapiens* X-ray crystal structures using SWISS-MODEL [75]. POLA1, POLA2$_C$, and PRIM2 homology models (from PDB accession 5EXR) were docked as rigid bodies into the high-resolution, partial reconstruction of the AI complex. Following manual modification, atomic coordinates and isotropic temperature factors were refined against the cryo-EM density for all non-hydrogen atoms using Ramachandran, secondary structure, and rotamer restraints. Hydrogen atoms were placed in riding positions and were not refined against the density. Both the refined model and a PRIM1 homology model (from PDB accession 4BPU) were then docked as rigid bodies into the complete reconstruction of the AI complex. After manual adjustment, atomic coordinates and isotropic temperature factors were refined as described above, except reference restraints were applied based on the *X. laevis* PRIM2/POLA1/POLA2$_C$ structure and the *H. sapiens* PRIM1 structure. POLA1$_{cat}$ and PRIM2$_C$ homology models (from PDB accessions 4QCL and 5F0Q) were then docked as rigid bodies into the reconstruction of the DI subcomplex. Following manual modification of the protein and addition of the substrate, atomic coordinates and isotropic temperature factors were refined as described above, except base pairing and base stacking restraints were applied to the oligonucleotide duplex. The refined POLA1$_{cat}$ model was then docked as a rigid body into the high-resolution, partial reconstruction of the DE subcomplex. After manual adjustment of the protein and addition of the substrate, atomic coordinates and isotropic temperature factors were refined as described above for the DI subcomplex. The refined models of POLA1$_{cat}$ from the partial DE structure and PRIM2$_C$ from the DI structure were then docked as rigid bodies into the complete DI subcomplex reconstruction. Following manual modification of the protein and addition of the substrate, atomic coordinates and isotropic temperature factors were refined as described above for the DI subcomplex, except reference restraints were applied based on the *X. laevis* POLA1$_{cat}$ and PRIM2$_C$ structures. For all structures, manual protein modification and

substrate addition were performed in Coot [76], while docking and refinement were performed in PHENIX [77]. Final models were validated using MolProbity [78]. Refinement and validation statistics are provided in **Supplemental Table S1 and Supplemental Figs. S7 and S14**. Figures were prepared in PyMOL (https://www.pymol.org), UCSF Chimera [79], or UCSF ChimeraX [80].

Primer elongation. To generate an RNA-primed template for elongation assays, 110 μM 5′-triphosphorylated RNA primer and 100 μM DNA template were annealed by heating to 85°C and then slowly cooling to 4°C. Reaction mixtures were then prepared with 50 nM $POLA1_{cat}$, $\Delta POLA1_{cat}$, $POLA1_{cat}/\Delta POLA1_{cat}$, or polα–primase; 500 nM primer/template; 0.1 μM [α-^{32}P]-ATP (3,000 Ci/mmol; PerkinElmer) or [α-^{32}P]-dATP (3,000 Ci/mmol; PerkinElmer); 9.9 μM ATP or dATP; 200 μM GTP, CTP, and UTP; 20 μM dGTP, dCTP, and dTTP; 20 mM HEPES•NaOH pH 7.5; 150 mM NaCl; 5 mM $MgCl_2$; and 1 mM TCEP and incubated at 22°C. 8-μL aliquots were periodically withdrawn, quenched with 2 μL of 100 mM EDTA pH 8.0, and incubated at 65°C for 15 min. To denature the primer and the template, 10 μL of loading buffer [80% (w/v) formamide, 10 mM EDTA pH 8.0, 1 mg/mL bromophenol blue, 1 mg/mL xylene cyanol, and 1.5 μM competitor oligonucleotide] was added and samples were incubated at 85°C for 10 min and then slowly cooled to 22°C. Elongation products were separated by denaturing urea-PAGE as previously described [67], except 18% polyacrylamide (19:1 acrylamide/bisacrylamide) gels were run at 20 W for 7 h and visualized by phosphorimaging using an Amersham Typhoon biomolecular imager (Cytiva). Experiments were performed in triplicate.

Data availability. Structures were deposited in the Protein Data Bank under accession codes 8G99 (AI complex, $POLA1/POLA2_C/PRIM2$), 8G9F (AI complex, $POLA1/POLA2_C/PRIM1/PRIM2$), 8G9L (DI subcomplex, $POLA1_{cat}/PRIM2_C$), 8G9N (DE subcomplex, $POLA1_{cat}$), and 8G9O (DE subcomplex, $POLA1_{cat}/PRIM2_C$). Maps were deposited in the Electron Microscopy Data Bank under accession codes EMD-29862 (AI complex, partial), EMD-29864 (AI complex, complete), EMD-29871 (DI subcomplex), EMD-29872 (DE subcomplex, partial), EMD-29862 (DE subcomplex, complete), EMD-29886 (DT subcomplex), EMD-29888 (tetramer core subcomplex, substrate free, conformation 1), EMD-29889 (tetramer core subcomplex, substrate free, conformation 2), and EMD-29891 (tetramer core subcomplex, substrate free, conformation 3).

References

1. Nethanel, T., Reisfeld, S., Dinter-Gottlieb, G. & Kaufmann, G. An Okazaki piece of simian virus 40 may be synthesized by ligation of shorter precursor chains. *J Virol* **62**, 2867-73 (1988).
2. Bullock, P.A., Seo, Y.S. & Hurwitz, J. Initiation of simian virus 40 DNA synthesis in vitro. *Mol Cell Biol* **11**, 2350-61 (1991).
3. Eki, T., Matsumoto, T., Murakami, Y. & Hurwitz, J. The replication of DNA containing the simian virus 40 origin by the monopolymerase and dipolymerase systems. *Journal of Biological Chemistry* **267**, 7284-7294 (1992).
4. Garg, P. & Burgers, P.M. DNA polymerases that propagate the eukaryotic DNA replication fork. *Crit Rev Biochem Mol Biol* **40**, 115-28 (2005).
5. Stodola, J.L. & Burgers, P.M. Mechanism of Lagging-Strand DNA Replication in Eukaryotes. *Adv Exp Med Biol* **1042**, 117-133 (2017).
6. Fukuda, M., Taguchi, T. & Ohashi, M. Age-dependent changes in DNA polymerase fidelity and proofreading activity during cellular aging. *Mech Ageing Dev* **109**, 141-51 (1999).
7. Levy, N. et al. XRCC1 interacts with the p58 subunit of DNA Pol alpha-primase and may coordinate DNA repair and replication during S phase. *Nucleic Acids Res* **37**, 3177-88 (2009).
8. Srivastava, V.K. & Busbee, D.L. Replicative enzymes and ageing: importance of DNA polymerase alpha function to the events of cellular ageing. *Ageing Res Rev* **1**, 443-63 (2002).
9. Taricani, L., Shanahan, F. & Parry, D. Replication stress activates DNA polymerase alpha-associated Chk1. *Cell Cycle* **8**, 482-9 (2009).
10. Yan, S. & Michael, W.M. TopBP1 and DNA polymerase alpha-mediated recruitment of the 9-1-1 complex to stalled replication forks: implications for a replication restart-based mechanism for ATR checkpoint activation. *Cell Cycle* **8**, 2877-84 (2009).
11. Kuchta, R.D., Reid, B. & Chang, L.M. DNA primase. Processivity and the primase to polymerase alpha activity switch. *J Biol Chem* **265**, 16158-65 (1990).
12. Sheaff, R.J. & Kuchta, R.D. Mechanism of calf thymus DNA primase: slow initiation, rapid polymerization, and intelligent termination. *Biochemistry* **32**, 3027-37 (1993).
13. Thomas, D.C. et al. Fidelity of mammalian DNA replication and replicative DNA polymerases. *Biochemistry* **30**, 11751-9 (1991).
14. Maga, G. et al. Okazaki fragment processing: modulation of the strand displacement activity of DNA polymerase delta by the concerted action of replication protein A, proliferating cell nuclear antigen, and flap endonuclease-1. *Proc Natl Acad Sci U S A* **98**, 14298-303 (2001).
15. Pavlov, Y.I. et al. Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Curr Biol* **16**, 202-7 (2006).
16. Rossi, M.L. & Bambara, R.A. Reconstituted Okazaki fragment processing indicates two pathways of primer removal. *J Biol Chem* **281**, 26051-61 (2006).
17. Stith, C.M., Sterling, J., Resnick, M.A., Gordenin, D.A. & Burgers, P.M. Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. *J Biol Chem* **283**, 34129-40 (2008).
18. Williams, J.S. & Kunkel, T.A. Ribonucleotides in DNA: origins, repair and consequences. *DNA Repair (Amst)* **19**, 27-37 (2014).
19. Burgers, P.M.J. & Kunkel, T.A. Eukaryotic DNA Replication Fork. *Annu Rev Biochem* **86**, 417-438 (2017).
20. Pellegrini, L. The Pol alpha-primase complex. *Subcell Biochem* **62**, 157-69 (2012).
21. Zerbe, L.K. & Kuchta, R.D. The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. *Biochemistry* **41**, 4891-900 (2002).
22. Lao-Sirieix, S.H., Nookala, R.K., Roversi, P., Bell, S.D. & Pellegrini, L. Structure of the heterodimeric core primase. *Nat Struct Mol Biol* **12**, 1137-44 (2005).
23. Vaithiyalingam, S., Warren, E.M., Eichman, B.F. & Chazin, W.J. Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. *Proc Natl Acad Sci U S A* **107**, 13684-9 (2010).
24. Sauguet, L., Klinge, S., Perera, R.L., Maman, J.D. & Pellegrini, L. Shared active site architecture between the large subunit of eukaryotic primase and DNA photolyase. *PLoS One* **5**, e10083 (2010).

25. Agarkar, V.B., Babayeva, N.D., Pavlov, Y.I. & Tahirov, T.H. Crystal structure of the C-terminal domain of human DNA primase large subunit: implications for the mechanism of the primase-polymerase alpha switch. *Cell Cycle* **10**, 926-31 (2011).
26. Nunez-Ramirez, R. et al. Flexible tethering of primase and DNA Pol alpha in the eukaryotic primosome. *Nucleic Acids Res* **39**, 8187-99 (2011).
27. Kilkenny, M.L., Longo, M.A., Perera, R.L. & Pellegrini, L. Structures of human primase reveal design of nucleotide elongation site and mode of Pol alpha tethering. *Proc Natl Acad Sci U S A* **110**, 15961-6 (2013).
28. Vaithiyalingam, S. et al. Insights into eukaryotic primer synthesis from structures of the p48 subunit of human DNA primase. *J Mol Biol* **426**, 558-69 (2014).
29. Baranovskiy, A.G. et al. Crystal structure of the human primase. *J Biol Chem* **290**, 5635-46 (2015).
30. Klinge, S., Nunez-Ramirez, R., Llorca, O. & Pellegrini, L. 3D architecture of DNA Pol alpha reveals the functional core of multi-subunit replicative polymerases. *EMBO J* **28**, 1978-87 (2009).
31. Kilkenny, M.L., De Piccoli, G., Perera, R.L., Labib, K. & Pellegrini, L. A conserved motif in the C-terminal tail of DNA polymerase alpha tethers primase to the eukaryotic replisome. *J Biol Chem* **287**, 23740-7 (2012).
32. Perera, R.L. et al. Mechanism for priming DNA synthesis by yeast DNA polymerase alpha. *Elife* **2**, e00482 (2013).
33. Attali, I., Botchan, M.R. & Berger, J.M. Structural Mechanisms for Replicating DNA in Eukaryotes. *Annual Review of Biochemistry* **90**, 77-106 (2021).
34. Baranovskiy, A.G. et al. Mechanism of concerted RNA-DNA primer synthesis by the human primosome. *J Biol Chem* **291**, 10006-10020 (2016).
35. Kilkenny, M.L. et al. Structural basis for the interaction of SARS-CoV-2 virulence factor nsp1 with DNA polymerase α-primase. *Protein Sci* **31**, 333-344 (2022).
36. He, Q. et al. Structures of the human CST-Polalpha-primase complex bound to telomere templates. *Nature* (2022).
37. Klinge, S., Hirst, J., Maman, J.D., Krude, T. & Pellegrini, L. An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nat Struct Mol Biol* **14**, 875-7 (2007).
38. Weiner, B.E. et al. An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *J Biol Chem* **282**, 33444-33451 (2007).
39. Arezi, B., Kirk, B.W., Copeland, W.C. & Kuchta, R.D. Interactions of DNA with human DNA primase monitored with photoactivatable cross-linking agents: implications for the role of the p58 subunit. *Biochemistry* **38**, 12899-907 (1999).
40. Baranovskiy, A.G. et al. Insight into the Human DNA Primase Interaction with Template-primer. *J Biol Chem* (2016).
41. Xie, P. A model for dynamics of primer extension by eukaryotic DNA primase. *Eur Biophys J* **40**, 1157-65 (2011).
42. Thompson, H.C., Sheaff, R.J. & Kuchta, R.D. Interactions of calf thymus DNA polymerase alpha with primer/templates. *Nucleic Acids Res* **23**, 4109-15 (1995).
43. Baranovskiy, A.G., Lisova, A.E., Morstadt, L.M., Babayeva, N.D. & Tahirov, T.H. Insight into RNA-DNA primer length counting by human primosome. *Nucleic Acids Res* **50**, 6264-6270 (2022).
44. Coloma, J., Johnson, R.E., Prakash, L., Prakash, S. & Aggarwal, A.K. Human DNA polymerase alpha in binary complex with a DNA:DNA template-primer. *Sci Rep* **6**, 23784 (2016).
45. Cai, S.W. et al. Cryo-EM structure of the human CST-Polalpha/primase complex in a recruitment state. *Nat Struct Mol Biol* **29**, 813-819 (2022).
46. He, Y. et al. Structure of Tetrahymena telomerase-bound CST with polymerase alpha-primase. *Nature* (2022).
47. Zhang, Y., Baranovskiy, A.G., Tahirov, T.H. & Pavlov, Y.I. The C-terminal domain of the DNA polymerase catalytic subunit regulates the primase and polymerase activities of the human DNA polymerase α-primase complex. *J Biol Chem* **289**, 22021-34 (2014).
48. Zhang, Y., Baranovskiy, A.G., Tahirov, E.T., Tahirov, T.H. & Pavlov, Y.I. Divalent ions attenuate DNA synthesis by human DNA polymerase alpha by changing the structure of the template/primer or by perturbing the polymerase reaction. *DNA Repair (Amst)* **43**, 24-33 (2016).

49. Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S. & Aggarwal, A.K. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat Struct Mol Biol* **16**, 979-86 (2009).
50. Hogg, M. et al. Structural basis for processive DNA synthesis by yeast DNA polymerase varepsilon. *Nat Struct Mol Biol* **21**, 49-55 (2014).
51. El Hassan, M.A. & Calladine, C.R. Conformational characteristics of DNA: empirical classifications and a hypothesis for the conformational behaviour of dinucleotide steps. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **355**, 43-100 (1997).
52. Lu, X.J. & Olson, W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **31**, 5108-21 (2003).
53. Neidle, S. *Nucleic acid structure and recognition*, xi, 188 p. (Oxford University Press, Oxford ; New York, 2002).
54. Schneider, A. et al. Primase activity of human DNA polymerase alpha-primase. Divalent cations stabilize the enzyme activity of the p48 subunit. *J Biol Chem* **273**, 21608-15 (1998).
55. Salay, L.E., Cordoba, J.J. & Chazin, W.J. DNA Recogntition/Processing | DNA Polymerase Alpha-Primase: Biochemical and Structural Mechanisms. in *Encyclopedia of Biological Chemistry* 3rd edn Vol. III (ed. Jez, J.) 431-444 (Elsevier, 2021).
56. Evrin, C., Maman, J.D., Diamante, A., Pellegrini, L. & Labib, K. Histone H2A-H2B binding by Pol alpha in the eukaryotic replisome contributes to the maintenance of repressive chromatin. *EMBO J* **37**(2018).
57. Li, Z. et al. DNA polymerase alpha interacts with H3-H4 and facilitates the transfer of parental histones to lagging strands. *Sci Adv* **6**, eabb5820 (2020).
58. Sun, J. et al. The architecture of a eukaryotic replisome. *Nat Struct Mol Biol* **22**, 976-82 (2015).
59. 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).
60. Zaug, A.J., Goodrich, K.J., Song, J.J., Sullivan, A.E. & Cech, T.R. Reconstitution of a telomeric replicon organized by CST. *Nature* **608**, 819-825 (2022).
61. Lim, C.J. et al. The structure of human CST reveals a decameric assembly bound to telomeric DNA. *Science* **368**, 1081-1085 (2020).
62. Braun, K.A., Lao, Y., He, Z., Ingles, C.J. & Wold, M.S. Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry* **36**, 8443-54 (1997).
63. Abad, M.G., Rao, B.S. & Jackman, J.E. Template-dependent 3'-5' nucleotide addition is a shared feature of tRNAHis guanylyltransferase enzymes from multiple domains of life. *Proc Natl Acad Sci U S A* **107**, 674-9 (2010).
64. Abad, M.G. et al. A role for tRNA(His) guanylyltransferase (Thg1)-like proteins from Dictyostelium discoideum in mitochondrial 5'-tRNA editing. *Rna* **17**, 613-23 (2011).
65. Rao, B.S., Mohammad, F., Gray, M.W. & Jackman, J.E. Absence of a universal element for tRNAHis identity in Acanthamoeba castellanii. *Nucleic Acids Res* **41**, 1885-94 (2013).
66. Smith, B.A. & Jackman, J.E. Saccharomyces cerevisiae Thg1 uses 5'-pyrophosphate removal to control addition of nucleotides to tRNA(His.). *Biochemistry* **53**, 1380-91 (2014).
67. Mullins, E.A. et al. An HPLC-tandem mass spectrometry method for simultaneous detection of alkylated base excision repair products. *Methods* **64**, 59-66 (2013).
68. Mastronarde, D.N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51 (2005).
69. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41-60 (2005).
70. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
71. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
72. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160 (2019).
73. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**(2018).

74. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biol* **4**, 874 (2021).
75. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-w303 (2018).
76. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
77. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
78. Davis, I.W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-83 (2007).
79. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
80. Goddard, T.D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).

Acknowledgments

This work was funded by National Institutes of Health grants R35GM136401 (B.F.E.), R35GM118089 (W.J.C.), R01GM087543 (J.E.J.), and PO1CA92584. We thank Miaw-Sheue Tsai and members of the Expression and Microbiology Core at UC Berkeley for baculovirus production and insect cell protein expression, and members of the Eichman and Chazin research groups for helpful discussions. Cryo-EM data were collected at the Center for Structural Biology Cryo-EM Facility at Vanderbilt University and the Life Sciences Institute Cryo-EM Facility at the University of Michigan. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

B.F.E. conceived of the study; B.F.E. and W.J.C. oversaw the study; E.A.M., L.E.S., and B.F.E. designed experiments; J.E.J. provided reagents; E.A.M., L.E.S., and C.L.D. collected data; E.A.M., L.E.S., M.D.O., W.J.C., and B.F.E. analyzed data; E.A.M., L.E.S., W.J.C., and B.F.E. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Figure Legends

Figure 1. Chimeric primer synthesis. A. Stages of primer synthesis by polα–primase. Triphosphate moieties are depicted as brown circles. Primase and polymerase α synthesize RNA and DNA, respectively, to generate a short chimeric primer. **B**. Schematic of subunit and domain organization in the heterotetrameric polα–primase complex.

Figure 2. Conformational and configurational dynamics. A. Characterization of purified *Xenopus laevis* polα–primase by denaturing gel electrophoresis. Subunits are colored as in **Fig. 1B** (POLA1, blue; POLA2, cyan; PRIM1, dark green; PRIM2, light green). The POLA1 construct lacks residues 1–334. **B**. Representative 2D class averages and corresponding subunit schematics. The percentage of particles belonging to the AI complex or the tetramer core subcomplex is indicated below the class averages. Class averages were selected to show $PRIM2_C$ at various degrees of association with the tetramer core. All final 2D class averages are shown in **Supplemental Fig. S6**. **C**. 3D reconstruction of polα–primase in the auto-inhibitory configuration. *Left*, cryo-EM density map. *Right*, molecular model refined against the density map. $POLA1_{cat}$ is outlined in blue. **D**. Distinct conformations of the tetramer core. *Top*, cryo-EM density maps. *Bottom*, molecular models docked into the density maps. Models were docked as three rigid bodies extracted from the auto-inhibitory complex. The points at which the complex were split and then reassembled are indicated with arrowheads. **E**. Schematic overlay of the three tetramer core conformations shown in panel **D**. The range of motion resulting from conformational dynamics is represented with curved arrows.

Figure 3. Configurational rearrangements during DNA initiation and DNA elongation. In each panel, DI and DE (sub)complexes are positioned on the left and the right, respectively. **A**. Substrate schematics. Triphosphate moieties are depicted as brown circles. In both substrates, the primers terminate in a 2′,3′-dideoxycytidine to prevent incorporation of dGTP. **B**. Representative 2D class averages and corresponding subunit schematics. The percentage of particles belonging to the DI or DE subcomplexes, the tetramer core subcomplex, or one of the consensus complexes (DI-1, DI-2, DE-1, or DE-2) is indicated below the class averages. 1% of particles from the DI dataset could not be reliably assigned to a group. 0% and 3% of particles from the DI and DE datasets, respectively, are consistent with the AI complex. All final 2D class averages are shown in **Supplemental Figs. S11 and S13**. **C**. 3D reconstructions and molecular models of $POLA1_{cat}$ and $PRIM2_C$ bound to the 3′- and 5′-ends of the primers, respectively. The

positions of the interdomain linkers are indicated with arrowheads. **D,E**. DI and DE consensus models. Models were constructed from two rigid bodies, comprised of the $POLA1_{cat}$/$PRIM2_C$ models shown in panel C and the corresponding tetramer core models, and guided by the 2D class averages shown in panel B. **D**. Consensus models oriented as in the 2D class averages. **E**. Consensus models aligned on the tetramer core.

Figure 4. Primer elongation. A. Separation and detection of RNA and RNA-DNA elongation products by denaturing gel electrophoresis and autoradiography. Intermediate and doublet bands result from a mixture of RNA and DNA incorporation at positions 7–10 (**Supplemental Fig. S18**). **B**. Substrate schematic. Triphosphate moieties are depicted as brown circles. [α-^{32}P]-labelled ATP and [α-^{32}P]-labelled dATP are only incorporated at position 7 in the primer. **C,D**. Quantitation of the gel in panel A. **C**. Relative intensities of the elongation products normalized to the intensity of the 30-mer. **D**. Kinetics of 30-mer formation. Experiments were performed in triplicate.

Figure 5. Substrate recognition and duplex remodeling. In each panel, the DI and DE structures are positioned on the left and the right, respectively. **A**. Schematic of substrate-binding contacts in $POLA1_{cat}$ (blue) and $PRIM2_C$ (green). **B**. Recognition of an incoming dGTP by $POLA1_{cat}$. Phosphodiester bond formation is prevented by a 2′,3′-dideoxycytidine (ddC) at the end of the primers. **C**. Recognition of RNA/DNA and DNA/DNA duplexes through substrate-adaptive conformational changes in $POLA1_{cat}$. **D**. Recognition of distinct primer/template conformations by $PRIM2_C$. **E**. RNA/DNA and RNA-DNA/DNA duplexes extracted from the DI and DE structures. Major (M) and minor (m) grooves are indicated. Numbers to the left of the duplexes indicate the base step starting from the 5′-end of the primers**. F**. Analysis of duplex conformation. Z_P is the interstrand phosphate-phosphate distance after projection onto the helical (Z) axis. The regions of the duplexes bound by $POLA1_{cat}$ (blue) and $PRIM2_C$ (green) are shown below the plots.

Figure 6. Configurational rearrangements upon DNA termination. A. Substrate schematic. Triphosphate moieties are depicted as brown circles. The primer terminates in a 2′,3′-dideoxycytidine to prevent incorporation of dGTP. **B**. Representative 2D class averages and corresponding subunit schematics. The percentage of particles belonging to the polα subcomplex, the primase subcomplex, or the auto-inhibitory complex is indicated below the class averages. All final 2D class averages are shown in **Supplemental Fig. S24**. **C**. 3D

reconstruction of $POLA1_{cat}$ and $PRIM2_C$ bound to primer/template. The approximate position and size of $PRIM2_C$ are indicated with a dashed oval.

Figure 7. Structural rearrangements during chimeric primer synthesis. A. Stages of RNA and DNA synthesis. $PRIM2_C$ remains bound to the 5′-end of the primer during all stages of synthesis. The transition from RNA elongation to DNA initiation occurs when the 3′-end of the primer is transferred from PRIM1 to $POLA1_{cat}$. After generation of a full-length primer, DNA elongation is terminated by dissociation of polα–primase from the primer/template. Protein-protein contacts between $PRIM2_C$ and $POLA1_{cat}$ are represented by a dashed line. **B**. Subunit schematics showing configurational rearrangements throughout the catalytic cycle of primer synthesis. Schematics are aligned on the tetramer core (PRIM1/$PRIM2_N$/$POLA1_C$/$POLA2_C$) subcomplex. The RNA initiation and RNA elongation configurations were derived from hypothetical models in which the positions of $POLA1_{cat}$ are indeterminate and the positions of $PRIM2_C$ are dictated by the geometric requirements of phosphodiester bond formation in the active site of PRIM1. The DNA initiation and DNA elongation configurations correspond to DI-2 and DE-2, respectively (**Fig. 3**). The DNA termination configuration was derived from a hypothetical model in which the primer and the template in DE-2 were extended while maintaining the position of $PRIM2_C$. Loose association of $POLA1_{cat}$ and $PRIM2_C$ with the full-length primer/template is indicated with wobble lines. At each stage of primer synthesis, the configuration depicted is representative of a broad configurational/conformational distribution.

Figure 1

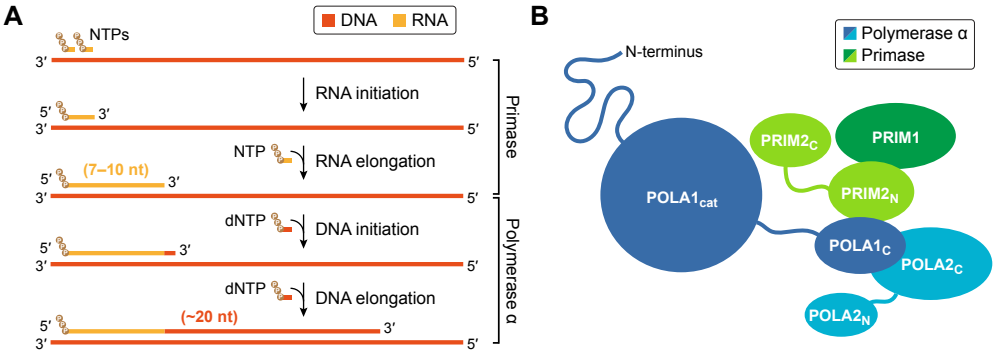

A
DNA
RNA
NTPs
3′
5′
RNA initiation
5′
3′
3′
5′
NTP
RNA elongation
5′
(7–10 nt)
3′
3′
5′
dNTP
DNA initiation
5′
3′
3′
5′
dNTP
DNA elongation
5′
(~20 nt)
3′
3′
5′
Primase
Polymerase α
B
Polymerase α
Primase
N-terminus
POLA1cat
PRIM2C
PRIM1
PRIM2N
POLA1C
POLA2C
POLA2N

Figure 2

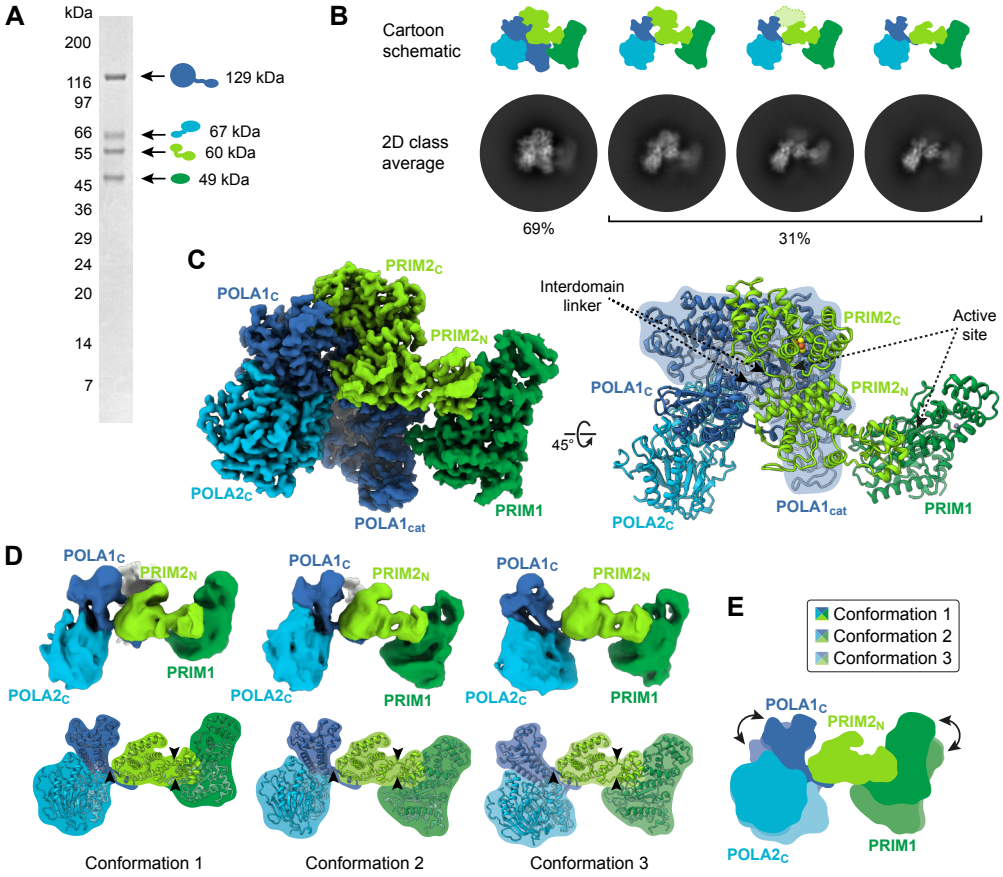

A
kDa
200
116
97
66
55
45
36
29
24
20
14
7
129 kDa
67 kDa
60 kDa
49 kDa
B
Cartoon schematic
2D class average
69%
31%
C
POLA1C
PRIM2C
PRIM2N
POLA2C
POLA1cat
PRIM1
Interdomain linker
Active site
45°
D
POLA1C
PRIM2N
PRIM1
POLA2C
Conformation 1
Conformation 2
Conformation 3
E
Conformation 1
Conformation 2
Conformation 3
POLA1C
PRIM2N
POLA2C
PRIM1

Figure 3

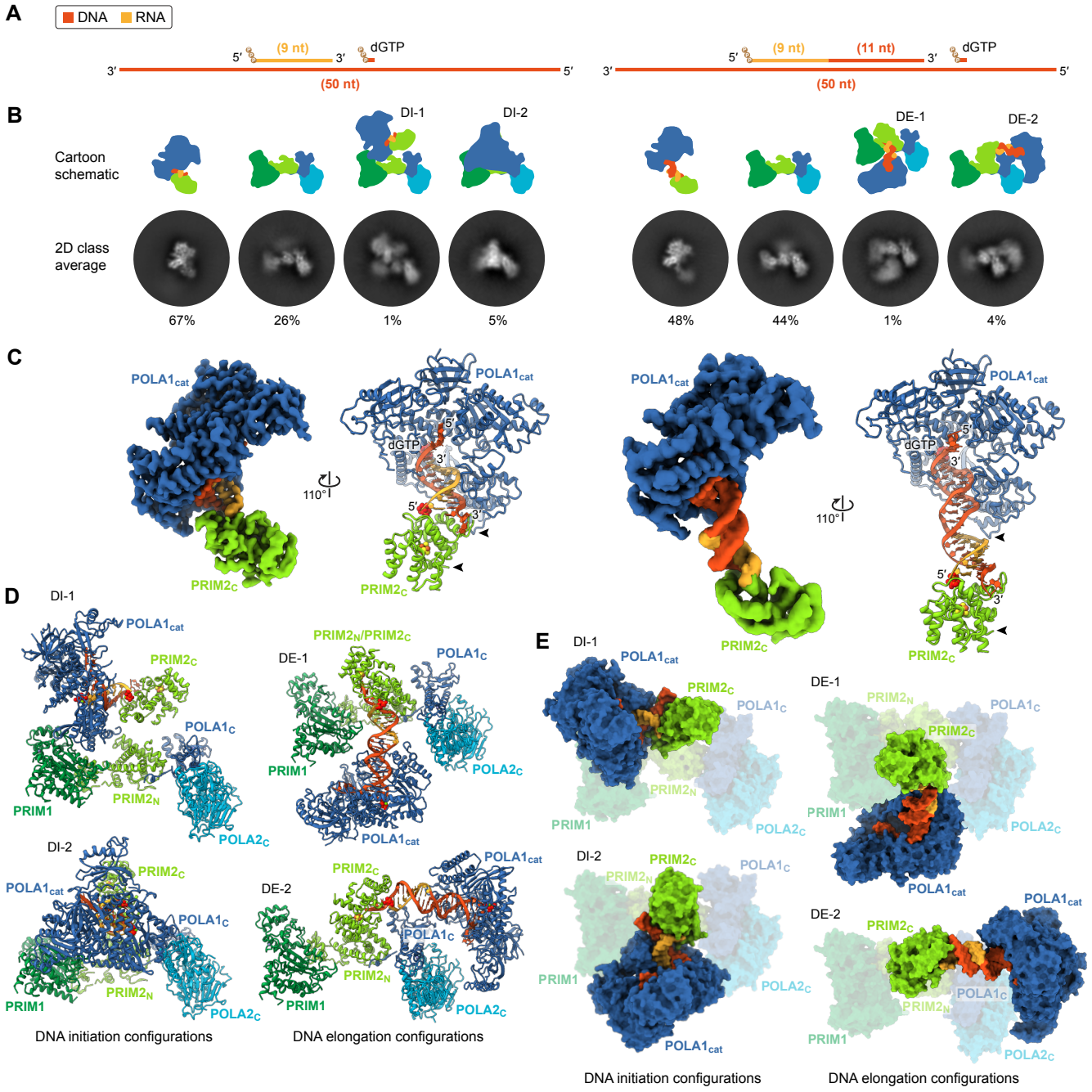

A
DNA
RNA
5′
(9 nt)
3′
dGTP
3′
5′
(50 nt)
5′
(9 nt)
(11 nt)
3′
dGTP
3′
5′
(50 nt)
B
Cartoon schematic
2D class average
DI-1
DI-2
DE-1
DE-2
67%
26%
1%
5%
48%
44%
1%
4%
C
POLA1cat
PRIM2C
dGTP
5′
3′
110°
D
DI-1
DI-2
DE-1
DE-2
POLA1cat
PRIM2C
PRIM2N/PRIM2C
POLA1C
PRIM1
PRIM2N
POLA2C
DNA initiation configurations
DNA elongation configurations
E
DI-1
DE-1
DI-2
DE-2
POLA1cat
PRIM2C
POLA1C
PRIM2N
PRIM1
POLA2C
DNA initiation configurations
DNA elongation configurations

Figure 4

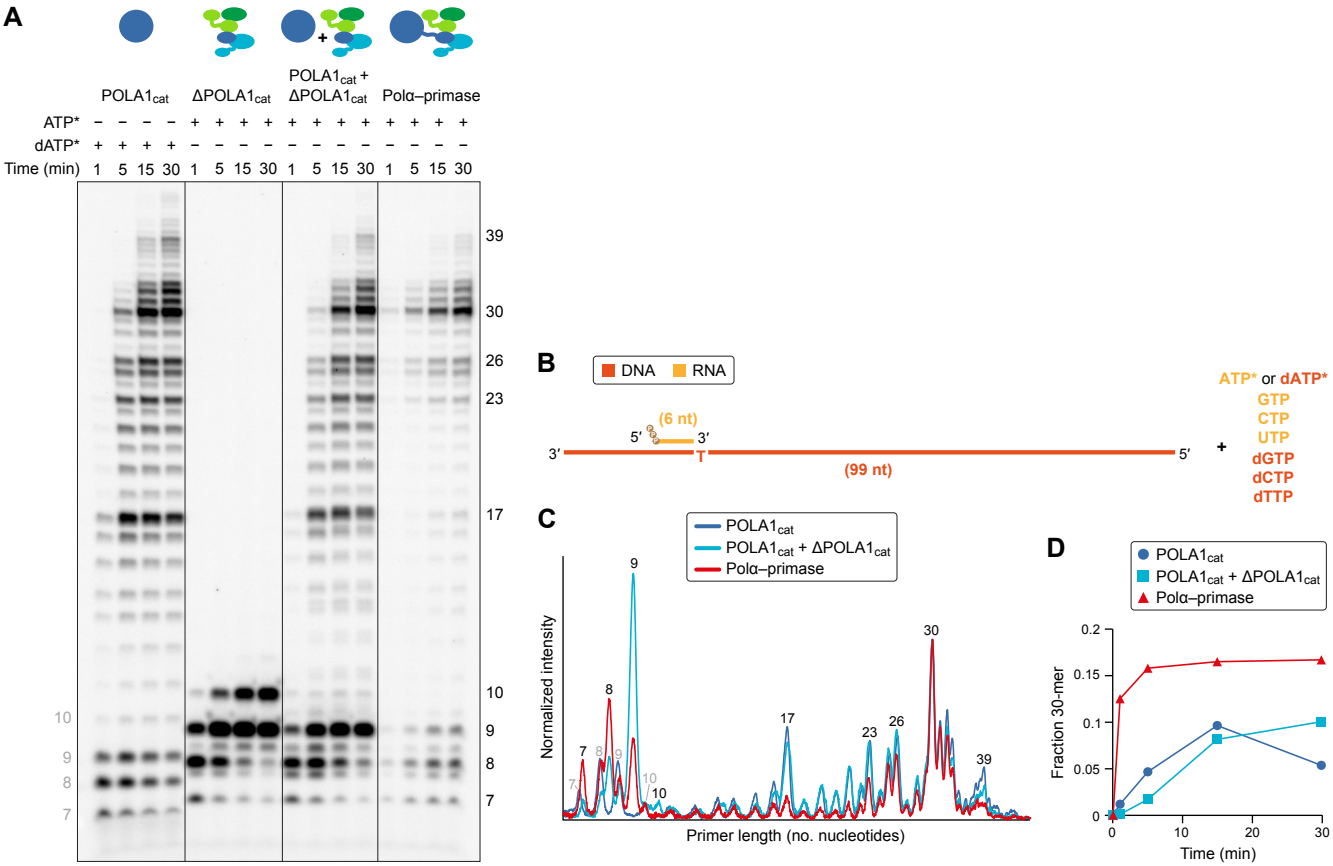

A
POLA1cat
ΔPOLA1cat
POLA1cat + ΔPOLA1cat
Polα–primase
ATP*
dATP*
Time (min)
1 5 15 30 1 5 15 30 1 5 15 30 1 5 15 30
39
30
26
23
17
10
9
8
7
B
DNA
RNA
(6 nt)
5′
3′
T
(99 nt)
ATP* or dATP*
GTP
CTP
UTP
dGTP
dCTP
dTTP
C
POLA1cat
POLA1cat + ΔPOLA1cat
Polα–primase
Normalized intensity
Primer length (no. nucleotides)
D
Fraction 30-mer
Time (min)

Figure 5

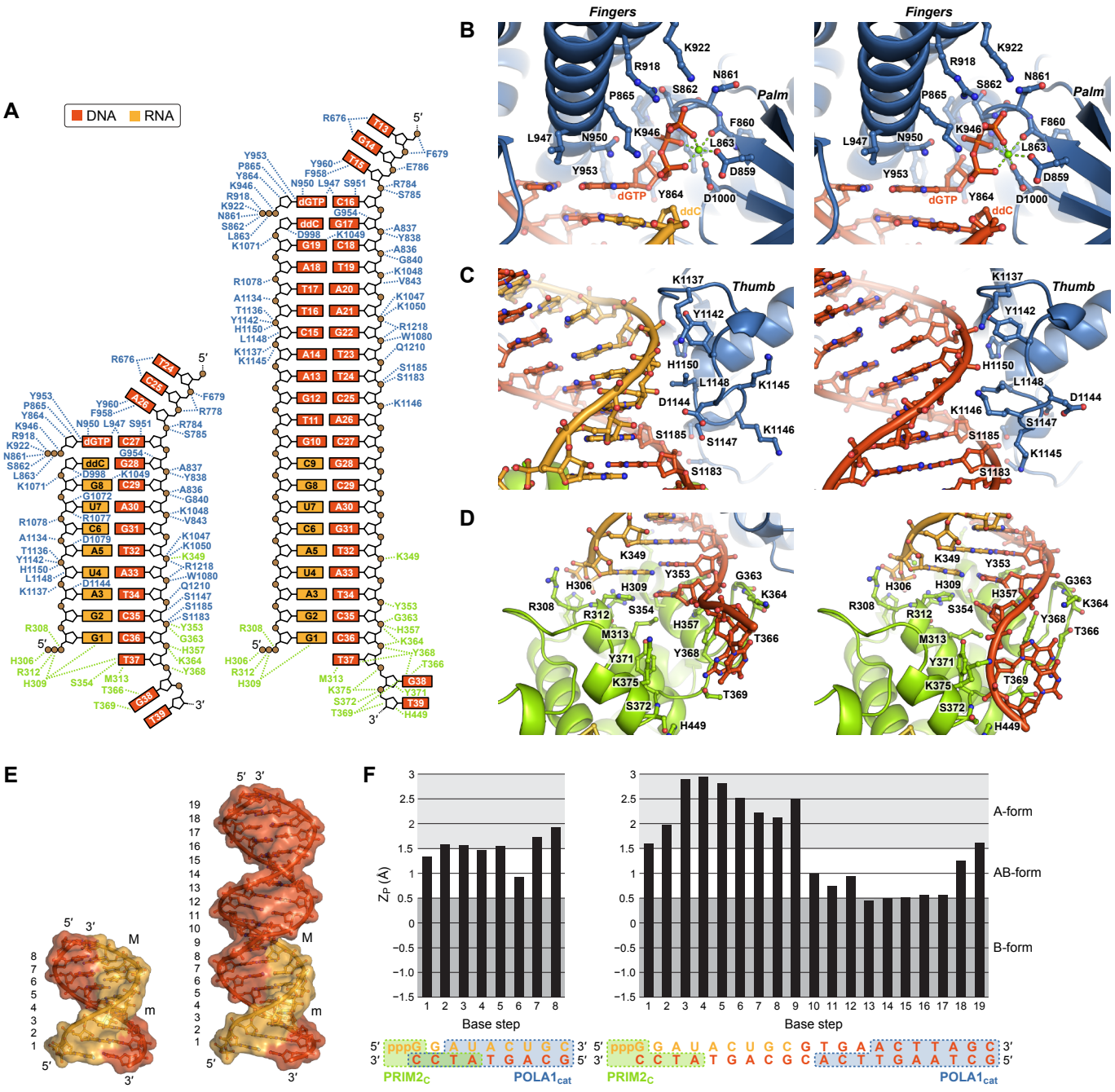

A
DNA
RNA
B
Fingers
Palm
C
Thumb
D
E
M
m
F
A-form
AB-form
B-form
Base step
PRIM2C
POLA1cat

Figure 6

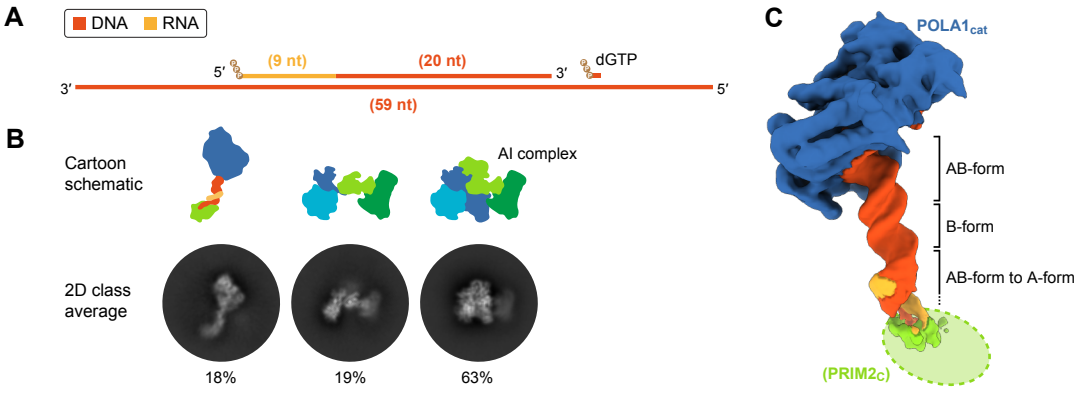
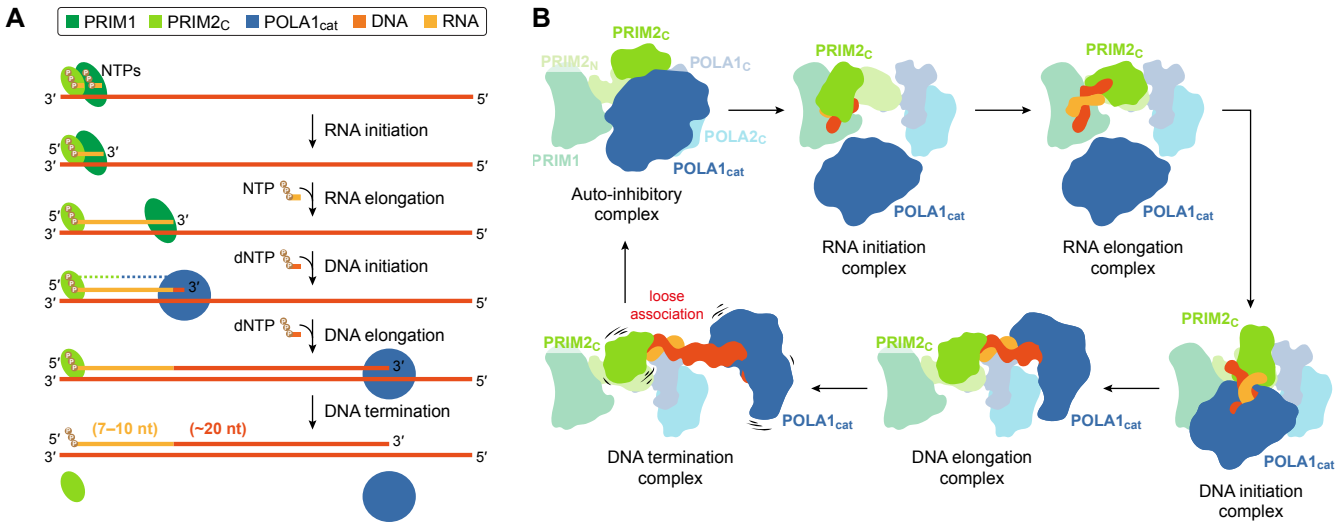

A
DNA
RNA
5′
(9 nt)
(20 nt)
3′
dGTP
3′
5′
(59 nt)
B
Cartoon schematic
AI complex
2D class average
18%
19%
63%
C
POLA1cat
AB-form
B-form
AB-form to A-form
(PRIM2C)

Figure 7


A
PRIM1
PRIM2C
POLA1cat
DNA
RNA
NTPs
3′
5′
RNA initiation
NTP
RNA elongation
dNTP
DNA initiation
DNA elongation
DNA termination
(7–10 nt)
(~20 nt)
B
PRIM2C
PRIM2N
POLA1C
POLA2C
PRIM1
POLA1cat
Auto-inhibitory complex
RNA initiation complex
RNA elongation complex
DNA initiation complex
DNA elongation complex
DNA termination complex
loose association