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], Noah P. Bradley[1,2,8], Jane E. Jackman[9,10], Melanie D. Ohi[5,6], Walter J. Chazin[2,3,11,*], 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]Presesnt address: Department of Molecular Biosciences, Northwestern University, Evanston, Illinois, USA
[9]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio, USA
[10]Center for RNA Biology, The Ohio State University, Columbus, Ohio, USA
[11]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 (TC), 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 catalytically accessible configurations

77 associated with the replisome [36] and the telomere maintenance CTC1–STN1–TEN1 (CST)
78 complex [37]. However, neither structure showed polα–primase engaged with primer/template.
79 Several models have emerged in recent years that largely attribute the counting ability of
80 polα–primase to the $PRIM2_C$ regulatory domain. Despite the catalytic center of primase residing
81 in PRIM1, $PRIM2_C$ is critical for all aspects of RNA synthesis, including regulation of RNA length
82 [23-25,38,39]. $PRIM2_C$ has also been implicated in transfer of the RNA primer to polα by virtue of its
83 interaction with the primer 5′-triphosphate and the template 3′-overhang [21,34,40,41]. These
84 interactions are believed to dictate RNA counting by spatially limiting the length of RNA
85 synthesized by PRIM1, either via the PRIM2 interdomain linker [41,42] or via steric collision
86 between $PRIM2_C$ and $PRIM2_N$ [33,34]. While these models explain the upper limit of RNA length
87 (i.e., 10 nucleotides), neither explains the range of RNA primers (7–10 nucleotides) extended by
88 polα [12]. $PRIM2_C$ and the primer 5′-triphosphate have also been shown to be important for DNA
89 synthesis [43,44], although the role of $PRIM2_C$ in the transition from RNA to DNA synthesis and in
90 DNA counting by polα is unknown. However, a recent model for DNA synthesis proposed a
91 similar spatial limitation imposed by the POLA1 and PRIM2 interdomain linkers [44]. Alternatively,
92 structures of $POLA1_{cat}$ bound to RNA/DNA and DNA/DNA duplexes afforded a model for DNA
93 counting whereby preferential binding to relatively A-form duplexes by polα leads to decreased
94 affinity and eventual dissociation from the primer/template as a result of a progressive increase
95 in B-form character as DNA is synthesized [32,45].
96 The molecular details by which polα–primase synthesizes chimeric primers of defined
97 length and composition have remained a matter of debate because there are few structures of
98 the heterotetramer bound to primer/template [46,47], and no structures showing the intermediate
99 and final stages of DNA synthesis. Here, we present cryo-EM structures of polα–primase alone
100 and bound to primed templates representing DNA synthesis at early, intermediate, and final
101 stages, beginning at the point immediately following transfer of the RNA primer to polα. Our
102 structures reveal that dissociation of $POLA1_{cat}$ and $PRIM2_C$ from the AI complex enables a
103 series of configurational and conformational changes that span a remarkable range of motion,
104 both in the absence and the presence of nucleic acid substrates. At the onset of DNA synthesis,
105 we observe a direct interaction between $POLA1_{cat}$ and $PRIM2_C$ that provides an explanation for
106 how polα can extend RNA primers ranging in length from 7–10 nucleotides, and that suggests
107 primer transfer is largely defined by competition between PRIM1 and $POLA1_{cat}$ for the primer 3′-
108 end. We also observe association of $PRIM2_C$ with chimeric RNA-DNA substrates corresponding
109 to intermediate and late stages of DNA synthesis. Complementary biochemical experiments
110 comparing the synthesis activities of the full polα–primase complex and various subcomplexes

111 are consistent with $PRIM2_C$ acting as a limited DNA processivity factor, reducing the proportion
112 of both shorter and longer primers [44]. Additionally, we provide direct structural evidence linking
113 primer composition, duplex conformation, and substrate affinity, with central roles for $POLA1_{cat}$
114 and $PRIM2_C$. Together, these findings establish a comprehensive model for chimeric primer
115 synthesis by polα–primase.
116
117
118 **Results**
119
120 <u>Polα–primase exists in an equilibrium between compact and extended configurations</u>
121 Prior structures in the absence of nucleic acid substrates or protein partners have shown
122 polα–primase in one of two states [26,34,35,37,47-49], suggesting the heterotetramer undergoes a
123 dramatic structural transition from a compact, auto-inhibitory state to an extended, catalytically
124 competent configuration prior to primer synthesis. To better understand the pre-existing
125 equilibrium between the two states, we first analyzed the protein alone using single-particle
126 cryo-EM. *Xenopus laevis* polα–primase was selected based on the stability and the quality of
127 the purified complex (**Fig. 2A**) and the high sequence conservation with human polα–primase
128 (69% identity and 90% similarity) (**Supplemental Figs. S1–S4**). We removed the N-terminal tail
129 of POLA1 (residues 1–334) to further improve the stability of the complex. In the absence of
130 nucleic acid substrates, 2D classification sorted 69% of particles into classes with density for all
131 four subunits in the compact AI configuration, 30% into classes lacking density for $POLA1_{cat}$,
132 and the remaining 1% into a class with density for only $POLA1_{cat}$ (**Fig. 2B and Supplemental**
133 **Figs. S5 and S6**). We initially refined a partial 3D reconstruction of the AI complex to a
134 resolution of 2.8 Å that exhibited poor density for PRIM1 (**Supplemental Figs. S5 and S7 and**
135 **Supplemental Table S1**). To improve the interpretability of the map in this region, we identified
136 a subset of particles that yielded a 3.2-Å reconstruction with clear density for all four subunits
137 (**Fig. 2C, Supplemental Figs. S5 and S7, and Supplemental Table S1**). Only the small N-
138 terminal domain of POLA2 ($POLA2_N$) was not well defined. As previously observed in lower-
139 resolution structures [34,35,47,48], the AI complex is stabilized by extensive contacts between
140 $POLA1_{cat}$ and every domain in the tetramer core, as well as $PRIM2_C$. Conversely, $PRIM2_C$
141 contacts only $POLA1_{cat}$, $PRIM2_N$, and the interdomain linkers in POLA1 (residues 1242–1275)
142 and PRIM2 (residues 262–278). These interactions sterically occlude substrate recognition by
143 both $POLA1_{cat}$ and $PRIM2_C$. To better understand the structural rearrangements necessary to
144 bind nucleic acid substrates, dissociation of $POLA1_{cat}$ and $PRIM2_C$ from the AI complex was

145 examined by 3D variability analysis, using all particles initially aligned on the AI complex during
146 preliminary 3D refinement (**Supplemental Figs. S5 and S8**). Two principal motions were
147 observed: (1) a hinging motion in $PRIM2_N$ that pulls PRIM1 away from $POLA1_{cat}$, and (2) a
148 bending/twisting motion in $POLA1_{cat}$ that pulls the exonuclease domain away from $POLA2_C$ and
149 that begins to pull the palm domain away from $PRIM2_N$. Interactions involving the thumb domain
150 were unchanged, as were all interactions involving $PRIM2_C$.
151 Without the conformational restrictions imposed by engagement of $POLA1_{cat}$ in the AI
152 complex, conformational heterogeneity is increased in the tetramer core. The subset of particles
153 grouped into 2D classes that lacked density for $POLA1_{cat}$ could be further divided into classes
154 with clear, partial, or no density for $PRIM2_C$ (**Fig. 2B and Supplemental Figs. S5 and S6**). 3D
155 reconstructions of the tetramer core were generated from particles selected through multiclass
156 ab initio reconstruction and heterogeneous refinement. These reconstructions revealed three
157 distinct conformations characterized by hinging motions in which the PRIM1 and
158 $POLA1_C$/$POLA2_C$ lobes move relative to one another by pivoting around flexible loops in
159 $PRIM2_N$ (residues 153–157 and 216–219) and $POLA1_C$ (residues 1439–1443), respectively
160 (**Fig. 2D,E and Supplemental Fig. S5**). Together, these data demonstrate that in the absence
161 of nucleic acid substrates, polα–primase exists predominantly in a compact auto-inhibitory state,
162 but also adopts a series of extended configurations in which $POLA1_{cat}$ and $PRIM2_C$ are
163 disengaged from the other domains in the complex. Based on 3D variability analysis of the AI
164 complex, along with inspection of the 2D class averages, we speculate that dissociation of
165 $POLA1_{cat}$ precedes dissociation of $PRIM2_C$, and initiates the structural transition away from an
166 auto-inhibitory state.
167
168 <u>$PRIM2_C$ remains bound to the primer during DNA synthesis</u>
169 Polα–primase possesses dual enzymatic activities that enable synthesis of chimeric RNA-
170 DNA primers. To investigate the molecular bases for the transition from RNA to DNA synthesis
171 and the low processivity of polα, we designed substrates to visualize polα–primase at two key
172 points during DNA synthesis—immediately after intramolecular handoff of the RNA primer from
173 primase to polα (DNA initiation, DI) and after polα has elongated the primer with DNA (DNA
174 elongation, DE). The DI primer is comprised of nine ribonucleotides and the chimeric DE primer
175 is comprised of nine ribonucleotides followed by 11 deoxyribonucleotides (**Fig. 3A,**
176 **Supplemental Fig. S9, and Supplemental Table S2**). Both primers include the critically
177 important 5′-triphosphate group that would be introduced by primase during RNA initiation.
178 Ternary complexes containing polα–primase, primer/template, and incoming nucleotide were

179 trapped using 2′,3′-dideoxycytidine to terminate the primer strand and prevent phosphodiester
180 bond formation with dGTP.
181 Relative to polα–primase in the absence of substrate, the 2D class averages determined
182 in the presence of the DI and DE substrates showed increased configurational heterogeneity
183 and dramatic configurational rearrangements (**Fig. 3B and Supplemental Figs. S10–S13**). The
184 vast majority of the class averages lacked clear density for part of the complex, indicating a
185 large degree of flexibility and conformational/configurational heterogeneity remains when
186 substrates are bound. These class averages could be divided into two groups: those with
187 density for $POLA1_{cat}$ and $PRIM2_C$ bound to substrate (hereafter referred to as the DNA
188 synthesis assembly, or DSA), and those with density for everything else (i.e., the tetramer core).
189 With the DI substrate, 2D classes showed 67% of particles with clear density for only the DNA
190 synthesis assembly, 26% for only the tetramer core, and ~6% for polα–primase in various
191 configurations. With the DE substrate, 2D classes showed a similar distribution, but with fewer
192 particles aligned on the DNA synthesis assembly and more particles aligned on the tetramer
193 core. Only 0 and 3% of particles with the DI and DE substrates, respectively, were found to be
194 in the AI configuration.
195 To better characterize interactions with the substrates, we performed focused 3D
196 refinement using masks around $POLA1_{cat}$, $PRIM2_C$, and the primer/template (**Fig. 3C,**
197 **Supplemental Figs. S10, S12, S14, and S15, and Supplemental Tables S3 and S4**). With
198 both the DI and DE substrates, the 3.3-Å and 4.4-Å reconstructions showed $POLA1_{cat}$ and
199 $PRIM2_C$ bound to opposing ends of the duplexes (**Fig. 3C**), indicating that interactions with
200 $PRIM2_C$ are maintained during transfer of the primer to polα and throughout DNA synthesis.
201 Focused 3D refinement of only $POLA1_{cat}$ and the DNA/DNA portion of the DE substrate
202 improved the resolution of the reconstruction to 3.5 Å (**Supplemental Figs. S12 and S15 and**
203 **Supplemental Table S4**). In both DSA structures, $POLA1_{cat}$ is in a closed ternary complex with
204 the primer 3′-end and the incoming dGTP, and $PRIM2_C$ is bound to the primer 5′-triphosphate.
205 With the DI substrate, $POLA1_{cat}$ and $PRIM2_C$ are in contact with one another, albeit with the
206 entire subunit interface consisting of a single interaction between two loops (**Fig. 3C and**
207 **Supplemental Fig. S16**). With the longer DE substrate, the two domains are separated by a full
208 turn of the RNA-DNA/DNA duplex. 3D variability analysis of these DI and DE subcomplexes
209 revealed dissociation of $PRIM2_C$ from a subset of the particles, as well as bending of the
210 oligonucleotide duplexes (**Supplemental Figs. S17 and S18**). With the shorter DI substrate,
211 this bending altered the protein-protein interface between $POLA1_{cat}$ and $PRIM2_C$.

212 Prompted by the seeming malleability and small size of the subunit interface, we modeled
213 POLA1$_{cat}$ and PRIM2$_C$ with shorter RNA primers (**Supplemental Fig. S16**). With 8-mer and 7-
214 mer primers, we observed only minor clashes between side chains, which appear unlikely to
215 prevent DNA initiation. Conversely, we saw severe steric overlap between POLA1$_{cat}$ and
216 PRIM2$_C$ with a 6-mer primer. These models explain the previous observation that polα can
217 extend RNA primers as short as seven nucleotides in length [12]. Additionally, binding of the
218 nascent primer by PRIM2$_C$ during DNA elongation explains the earlier observation that the
219 primer 5′-triphosphate affects not only RNA synthesis but also DNA synthesis [43], as well as the
220 recent finding that PRIM2$_C$ increases the processivity of polα–primase [44].
221 For simultaneous binding of the primer ends to be maintained throughout DNA synthesis,
222 as the primer increases in length, large-scale protein conformational changes are necessary.
223 Because of the extreme conformational and configurational heterogeneity inherent in the polα–
224 primase complex, we were unable to generate reconstructions of the full substrate-bound
225 complex by multiclass ab initio reconstruction and heterogeneous refinement. However, we
226 were able to generate reconstructions from subsets of particles manually selected from 2D class
227 averages (**Supplemental Figs. S19 and S20 and Supplemental Tables S3 and S4**). Rigid-
228 body docking of the DSA and TC subcomplexes into these low-resolution reconstructions
229 enabled us to model two configurations of the complete polα–primase complex bound to DI
230 substrate and two configurations bound to DE substrate (**Fig. 3D,E, Supplemental Figs. S19**
231 **and S20, and Supplemental Tables S3 and S4**). Comparison of the four substrate-bound
232 configurations showed that the DNA synthesis assembly rotates by as much as 180° with
233 respect to the tetramer core, with the pivot point occurring between PRIM2$_C$ and PRIM2$_N$ (**Fig.**
234 **3E**). In addition to this swiveling, the DNA synthesis assembly rotates by nearly 180° with
235 respect to the helical axis of the oligonucleotide duplexes. In all models of the four-subunit
236 complex, the substrate-bound configurations appear to be stabilized by interactions between the
237 DNA synthesis assembly and the tetramer core. However, the variability of these interactions,
238 as well as the large proportion of 2D class averages showing density for only one or the other
239 subcomplex, suggests these protein-protein contacts are weak and transient. Thus, in the
240 presence of primer/template substrates representing two key stages of DNA synthesis, polα–
241 primase adopts a wide range of configurations characterized by concerted movement of
242 POLA1$_{cat}$, PRIM2$_C$, and the oligonucleotide duplex relative to the tetramer core.
243
244 <u>Primase enhances an intrinsic counting ability in polα</u>

245 Whether counting of DNA nucleotides is an inherent property of polα or only a property of
246 the complete polα–primase complex is unclear [32,34,44,50,51]. Our finding that PRIM2$_C$ remains
247 bound to the primer after transfer to polα reveals a means by which primase might participate in
248 counting during DNA synthesis. We theorized that binding of the substrate by PRIM2$_C$ would
249 alter DNA elongation by indirectly tethering POLA1$_{cat}$ to the primer via the tetramer core. To
250 investigate this theory, we compared the RNA and DNA synthesis activities of the polα–primase
251 complex with a complex in which the interdomain linker between POLA1$_{cat}$ and POLA1$_C$ was
252 severed, constructed by mixing purified POLA1$_{cat}$ and ΔPOLA1$_{cat}$
253 (PRIM1/PRIM2/POLA1$_C$/POLA2) (**Fig. 4A and Supplemental Fig. S21**). As a control, we
254 measured the RNA and DNA synthesis activities of POLA1$_{cat}$ and ΔPOLA1$_{cat}$ separately. Primer
255 extension was performed by adding radiolabeled ATP or dATP and unlabeled NTPs and dNTPs
256 to a 5′-triphosphorylated 6-ribonucleotide primer annealed to a 99-deoxyribonucleotide template
257 (**Fig. 4B and Supplemental Table S2**). POLA1$_{cat}$ alone generated extension products of
258 predominantly 17–33 nucleotides, with the main product being 30 nucleotides in length (**Fig.
259 4A,C**). Similar to previous observations with primase alone [21,29], ΔPOLA1$_{cat}$ rapidly formed
260 products of nine nucleotides before abruptly terminating at a 10-mer primer (**Fig. 4A**). The
261 severed complex generated by combining POLA1$_{cat}$ and ΔPOLA1$_{cat}$ afforded products
262 representing an admixture of the individual control reactions (**Fig. 4A,C and Supplemental Fig.
263 S22**). In contrast, intact polα–primase generated fewer extension products outside of 30
264 nucleotides (**Fig. 4A,C**), and produced the 30-mer product at a dramatically faster rate (**Fig.
265 4D**). However, the total activity of polα–primase was notably less than that of the other
266 constructs. We speculate this reduced activity results from a substantial proportion of the polα–
267 primase molecules being in the AI configuration, and therefore unable to perform primer
268 elongation without undergoing a dramatic configurational transition. Collectively, these data
269 demonstrate an intrinsic—albeit approximate—counting ability in polα that is improved by the
270 physical linkage between POLA1$_{cat}$ and the remainder of the complex, reinforcing the
271 preference of polα–primase to generate a 30-nucleotide primer.
272 To better understand the approximate counting ability of polα, we performed binding
273 assays with POLA1$_{cat}$ and DNA/DNA, RNA/DNA, and two RNA-DNA/DNA duplexes (**Fig. 4E
274 and Supplemental Table S5**). Consistent with previous findings [32,45], we observed significantly
275 tighter binding of an RNA/DNA substrate than a corresponding DNA/DNA substrate. Addition of
276 11 nucleotides of DNA to the RNA primer, creating an RNA-DNA/DNA substrate, had no effect
277 on binding affinity. However, a further addition of nine nucleotides of DNA, producing a longer
278 RNA-DNA/DNA duplex, resulted in significantly weaker binding, approaching that of the

279 DNA/DNA substrate. For both RNA-DNA/DNA duplexes, the DNA/DNA regions were sufficiently
280 long to fully accommodate the binding footprint of the polymerase catalytic core. Thus, these
281 findings indicate the affinity of $POLA1_{cat}$ for chimeric primers is dependent on not only the length
282 of the DNA/DNA region of the duplex but also the proximity of the duplex end to the RNA/DNA
283 region, providing a basis for counting and the termination of DNA synthesis.
284
285 Primer/template conformation modulates substrate recognition
286 Consistent with the results from our binding assays, the intrinsic counting ability of polα
287 has been proposed to result from conformation-dependent recognition of RNA/DNA and
288 DNA/DNA duplexes [32,45]. Because of the relatively high resolutions of the DSA reconstructions,
289 we were able to determine the conformations of the 5′-triphosphorylated RNA/DNA and RNA-
290 DNA/DNA substrates in the DI and DE subcomplexes and to evaluate the details of their
291 recognition by $POLA1_{cat}$ and $PRIM2_C$. In both structures, $POLA1_{cat}$ was observed in a closed
292 conformation in which the fingers domain was fully engaged with an incoming dGTP and active-
293 site residues within the palm domain (Asp859, Phe860, and Asp1000) coordinated a catalytic
294 Mg^{2+} ion (**Fig. 5A,B and Supplemental Fig. S23**). The thumb domain was also fully engaged
295 with both the RNA/DNA and RNA-DNA/DNA duplexes (**Fig. 5A,C**). To our knowledge, the DE
296 subcomplex represents the first structure of $POLA1_{cat}$ with the thumb and fingers domains
297 engaged with a DNA/DNA duplex and an incoming dNTP (**Supplemental Fig. S24**). In both
298 subcomplexes, $POLA1_{cat}$ induces a local A-form conformation in the duplexes that positions and
299 orients the 2′,3′-dideoxycytidine at the 3′-end of the primers for phosphodiester bond formation
300 with dGTP (**Fig. 5B,E,F**). Similar catalytic arrangements have been observed in structures of
301 polδ and polε [52,53]. However, $POLA1_{cat}$ possesses a wider substrate-binding cleft that results
302 from structural differences in the thumb domain (**Supplemental Fig. S25**). This wider cleft has
303 been proposed to complement the wider diameter of an A-form RNA/DNA duplex [32,45]. In both of
304 our structures, we observed intermediate AB-form conformations [54,55], or nearly intermediate
305 AB-form conformations, over most of the region of the duplexes bound by $POLA1_{cat}$ (**Fig. 5E,F**).
306 Importantly, small but significant differences in the interactions between the thumb domain
307 and the DI and DE substrates illustrate how $POLA1_{cat}$ accommodates different minor groove
308 geometries (**Fig. 5A,C,E,F**). With the DI substrate, three aspartate residues (Asp998, Asp1079,
309 and Asp1144) form hydrogen-bonding interactions with 2′-hydroxyl groups along the RNA
310 primer. With the DE substrate, conformational changes at the tip of the thumb domain (residues
311 1142–1148) pull Asp1144 away from the chimeric primer while directing Lys1145 and Lys1146
312 toward the relatively narrow minor groove in the same DNA/DNA region of the duplex. These

313 new hydrogen-bonding interactions likely partially offset the loss of RNA-specific hydrogen
314 bonds with the DE substrate. While RNA/DNA and DNA/DNA duplexes generally adopt A-form
315 and B-form geometries [56], respectively, the intermediate AB-form conformations observed with
316 both the DI and DE substrates suggest $POLA1_{cat}$ remodels the primer/template during DNA
317 synthesis. The local propensity of the duplex to adopt this intermediate conformation is likely
318 correlated with substrate affinity and, consequently, with processivity [32,45].
319 $PRIM2_C$ has been shown to increase the processivity of polα–primase during DNA
320 synthesis [44]. The two DSA structures suggest that this increased processivity is the result of
321 continued association of $PRIM2_C$ with the primer/template during DNA elongation. In both the DI
322 and DE subcomplexes, interactions between $PRIM2_C$ and the primer strand are identical (**Fig.**
323 **5A,D**). Three positively charged residues (His306, Arg308, and Arg312) at the N-terminus of
324 helix α15 surround the 5′-triphosphate and offset the negative charge, and His309 stacks
325 against the 5′-terminal nucleobase. Surprisingly, interactions between $PRIM2_C$ and the template
326 strand are markedly different in the two subcomplexes, despite $PRIM2_C$ binding the same
327 RNA/DNA portion of both substrates. In the DI structure, the DNA template is positioned such
328 that only the last paired nucleotide, which is not coplanar with the 5′-terminal nucleotide in the
329 RNA primer, and the first unpaired nucleotide form extensive contacts with $PRIM2_C$. Rotation of
330 the phosphodeoxyribose backbone swings the second and third unpaired nucleotides away
331 from $PRIM2_C$ and toward $POLA1_{cat}$. In the DE structure, the DNA template is positioned such
332 that the last paired nucleotide, which is coplanar with the 5′-terminal nucleotide in the RNA
333 primer, and the first, second, and third unpaired nucleotides form extensive contacts with
334 $PRIM2_C$. Relative to the DI substrate, interactions with several residues (Tyr353, His357,
335 Lys364, and Tyr368) in the flexible template-binding loop (residues 359–372) are shifted by one
336 nucleotide toward the 5′-end of the template strand as a result of increased base pair inclination
337 in the DE substrate. Virtually identical interactions were observed in an X-ray crystal structure of
338 human $PRIM2_C$ bound to a short RNA/DNA duplex [34] (**Supplemental Fig. S26**).
339 The unique interactions that we observe with the DI substrate appear to result from local
340 duplex remodeling by $POLA1_{cat}$, which causes the RNA/DNA duplex to have less A-form
341 character, as quantified by the Z_P parameter [54,55,57], than either the RNA/DNA portion of the DE
342 substrate or the RNA/DNA duplex in the crystal structure (**Fig. 5D,E,F and Supplemental Fig.**
343 **S26**). In addition to reduced A-form character, the DI substrate has lower base pair inclination,
344 which pulls the 3′-overhang away from $PRIM2_C$. The affinity of $PRIM2_C$ for the primer/template
345 was previously shown to depend on both the 5′-triphosphate and the 3′-overhang [34,41]. Thus,
346 tight binding of the substrate appears to require substantial inclination of the base pairs at the

347 end of duplex. Together, these data indicate that like POLA1$_{cat}$, the affinity of PRIM2$_C$ for the
348 primer/template is modulated by the nucleic acid conformation, which is imposed by both the
349 DNA composition of the primer and binding by the polymerase catalytic core.
350
351 Loose association of full-length primers allows for reversion to the auto-inhibitory configuration
352 The finding that substrate recognition by both POLA1$_{cat}$ and PRIM2$_C$ is dependent on
353 duplex conformation has implications for the mechanism of DNA counting. To investigate how
354 duplex conformation affects the preference of polα–primase to generate 30-nucleotide primers,
355 we designed a substrate to visualize the complex at the end of DNA synthesis (DNA
356 termination, DT). The DT primer is comprised of nine ribonucleotides followed by 20
357 deoxyribonucleotides (**Fig. 6A, Supplemental Fig. S9, and Supplemental Table S2**) and
358 includes the same 5′-triphosphate group as the DI and DE primers. A ternary DT complex
359 containing polα–primase, primer/template, and incoming nucleotide was trapped using 2′,3′-
360 dideoxycytidine to terminate the primer strand and prevent phosphodiester bond formation with
361 dGTP.
362 As with the DI and DE substrates, a substantial proportion of 2D class averages showed
363 density for only the DNA synthesis assembly or the tetramer core (**Fig. 6B and Supplemental**
364 **Figs. S27 and S28**), consistent with large-scale conformational and configurational dynamics.
365 Unlike with the shorter substrates, however, no class averages showed density for the complete
366 polα–primase complex bound to substrate. This level of heterogeneity suggests that the
367 interdomain linkers in POLA1 and PRIM2 remain slack when bound to a full-length primer, as
368 taught linkers would seemingly reduce the overall flexibility of the DT complex. Surprisingly,
369 63% of particles were grouped into 2D classes with density clearly representative of the AI
370 complex. Protein-protein contacts in this configuration sterically occlude substrate recognition by
371 both POLA1$_{cat}$ and PRIM2$_C$, indicating the DT substrate is not bound in most particles. With the
372 DI and DE substrates, only 0 and 3% of particles, respectively, were found to be in the AI
373 configuration, while in the absence of substrate, 69% of particles were found to be in this
374 configuration. While this type of comparative quantitation of cryo-EM data has the potential for
375 bias, this dramatic increase in AI particles in the presence of DT substrate suggests
376 substantially weaker binding of the longer chimeric primer by both POLA1$_{cat}$ and PRIM2$_C$.
377 To better understand the structural basis for weaker association of the DT substrate, we
378 performed focused 3D refinement of the DT subcomplex. Using a mask around POLA1$_{cat}$ and
379 the DNA/DNA region of the primer/template, we generated a 2.9-Å reconstruction with clear
380 density for POLA1$_{cat}$ in a closed conformation, with the thumb and fingers domains fully

381 engaged with the duplex and an incoming dGTP (**Fig. 6C, Supplemental Figs. S27 and S29,**
382 **and Supplemental Table S6**). 3D variability analysis of the complete DT subcomplex, including
383 POLA1$_{cat}$, PRIM2$_C$, and the full-length primer/template, revealed duplex bending similar to that
384 seen in the DE subcomplex, as well as dissociation of PRIM2$_C$ from the substrate
385 (**Supplemental Fig. S30**). However, the degree to which PRIM2$_C$ is disengaged from the
386 primer/template appears to be substantially greater with the DT substrate. Consistent with this
387 observation, 3D classification sorted most particles aligned on the DT subcomplex into classes
388 with little to no density for PRIM2$_C$ (**Supplemental Fig. S27**). Following classification, focused
389 3D refinement provided 3.8- and 3.6-Å reconstructions of the DT subcomplex with (DT
390 subcomplex 1, or DTS-1) and without (DT subcomplex 2, or DTS-2) PRIM2$_C$ engaged with the
391 substrate (**Fig. 6D, Supplemental Figs. S27 and S29, and Supplemental Table S6**). The
392 overall conformations of the duplexes in these structures, as quantified by the Z_P parameter, are
393 quite similar (**Fig. 6E**). However, the base pair inclination in the RNA/DNA portion of DTS-2
394 appears to be reduced relative to DTS-1 (**Fig. 6F and Supplemental Fig. S31**). While the
395 degree of certainty associated with this observation is limited by the local resolution of the
396 reconstructions (**Supplemental Fig. S29**), this conformational difference is consistent with our
397 conclusion that tight substrate binding by PRIM2$_C$ is dependent on the base pair inclination at
398 the end of the duplex. Conversely, the origins for weaker binding of the DT substrate by
399 POLA1$_{cat}$ are not apparent in the structures. In particular, all structures of the DNA synthesis
400 assembly reveal POLA1$_{cat}$ has remodeled the bound region of the duplex—whether it is
401 RNA/DNA or DNA/DNA—to achieve similar intermediate AB-form conformations
402 (**Supplemental Fig. S31**). However, remodeling of this type requires an input of free energy,
403 which in turn reduces the free energy associated with binding affinity. Thus, primer/template
404 duplexes predisposed to adopt intermediate AB-form conformations, like those in the remodeled
405 substrates, require less free energy to be remodeled, and therefore will be bound with higher
406 affinity.
407
408
409 **Discussion**
410
411 Multiple, seemingly conflicting models have been proposed to explain the counting abilities
412 of primase and polα [32-34,41,42,44-47]. Our structures of polα–primase create a framework to
413 reconcile and integrate these models, providing a comprehensive mechanistic understanding of
414 nucleotide counting (**Fig. 7A,B**). We show that in the absence of nucleic acid substrates or

415 protein partners, polα–primase exists primarily in a compact, auto-inhibitory state, but
416 spontaneous dissociation of $POLA1_{cat}$ and $PRIM2_C$ initiates a series of dramatic structural
417 rearrangements necessary for chimeric primer synthesis across two spatially separate active
418 sites. In all current models of RNA synthesis, primer formation is initiated by PRIM1 and $PRIM2_C$
419 bringing together two NTPs on a single-stranded DNA template. In the linker model, elongation
420 of the primer continues until halted by full extension of the PRIM2 interdomain linker [41,42,47]. In
421 the collision model, elongation is ceased by a steric clash between $PRIM2_N$ and $PRIM2_C$ [34]. As
422 previously discussed [34], the linker model is not consistent with domain arrangement in the
423 tetramer core. Conversely, the collision model is supported by structural data and structure-
424 based modeling [34], and provides a convincing rationale for termination of RNA synthesis, but
425 strictly at 9–10 nucleotides. Both models fail to explain previous observations showing extension
426 of RNA primers as short as seven nucleotides in length by polα [12].
427 Based on our structure of the DI subcomplex, as well as previous biochemical data, we
428 propose a competition model for RNA counting and primer transfer. $PRIM2_C$ has been
429 suggested to anchor the primer during its transfer between primase and polα active sites by
430 remaining bound to the primer 5′-end [34,37,44]. Moreover, the highly distributive nature of RNA and
431 DNA synthesis by PRIM1 and $POLA1_{cat}$ in the absence of $PRIM2_C$ [29,44,58] suggests that $PRIM2_C$
432 acts as a hub to facilitate re-association of PRIM1 and $POLA1_{cat}$ with the primer. Our structures
433 confirm that $PRIM2_C$ remains bound to the 5′-end of the primer after transfer to polα, and
434 provide evidence that a dynamic protein-protein interface allows $PRIM2_C$ and $POLA1_{cat}$ to
435 simultaneously bind RNA primers of seven or more nucleotides, while transfer of RNA primers
436 less than seven nucleotides in length to $POLA1_{cat}$ would be sterically blocked by $PRIM2_C$. Thus,
437 the likely driving factor in the primase-to-polymerase switch is a competition between PRIM1
438 and $POLA1_{cat}$ for the 3′-end of the primer, rather than termination of RNA synthesis at a
439 maximal length of 10 nucleotides. In this competition model, steric collision between $PRIM2_N$
440 and $PRIM2_C$ would serve only as a failsafe to terminate RNA synthesis, ensuring longer
441 stretches of RNA (> 10 nts) are not incorporated into the genome, should $POLA1_{cat}$ not
442 outcompete PRIM1 for the primer.
443 Our finding that $PRIM2_C$ remains bound to the primer after transfer to polα has far-
444 reaching implications for DNA synthesis, including the preference of polα–primase to generate
445 chimeric primers of ~30 nucleotides. Simultaneous association of both $PRIM2_C$ and $POLA1_{cat}$ is
446 a prerequisite for the linker model of DNA termination, which posits that the POLA1 and PRIM2
447 interdomain linkers prevent separation of $PRIM2_C$ and $POLA1_{cat}$ beyond the distance associated
448 with an ~30-mer primer [44,47]. However, we observed seemingly weaker binding of chimeric

449 primers by both $POLA1_{cat}$ and $PRIM2_C$ as the DNA composition of the primer increased (as
450 measured by direct binding assays using $POLA1_{cat}$ and implied by the proportion of particles in
451 the AI configuration). Moreover, if termination were defined only by the length of the interdomain
452 linkers in POLA1 and PRIM2, we would expect to observe a stable complex between polα–
453 primase and the DT substrate, likely with well-ordered, taught linkers. Additionally, the linker
454 model is unable to provide a rationale for the intrinsic counting ability of $POLA1_{cat}$, which we and
455 others have observed [32,50]. By contrast, the affinity model of DNA termination, which proposes
456 that $POLA1_{cat}$ preferentially recognizes A-form duplexes typical of RNA/DNA hybrids, and loses
457 affinity for the substrate as the duplex becomes more B-form throughout DNA synthesis [32,45],
458 provides a mechanistic explanation for the innate counting ability of the polymerase catalytic
459 core. Our data expand upon the affinity model by suggesting a role for preferential binding of A-
460 form duplexes by $PRIM2_C$, and by providing a possible structural basis for this specificity. Thus,
461 we propose that DNA termination is the result of both $POLA1_{cat}$ and $PRIM2_C$ losing affinity for
462 the primer/template as the B-form propensity of the duplex increases during DNA synthesis.
463 Because the linker model requires both $POLA1_{cat}$ and $PRIM2_C$ to remain bound to the substrate,
464 the interdomain linkers are likely to play, at most, a failsafe role in restricting primer length, and
465 thereby limiting the mutagenic potential of polα.
466 The dramatic structural rearrangements that we observe are consistent with the idea that
467 flexible tethering of $POLA1_{cat}$ and $PRIM2_C$ to the tetramer core is necessary for transfer of the
468 nascent primer from primase to polα [26,34]. As we demonstrate here, these rearrangements are
469 also necessary for simultaneous binding of the primer ends by $POLA1_{cat}$ and $PRIM2_C$ during
470 DNA elongation, as the primer increases in length. Presumably, this plasticity also facilitates
471 interaction with the many protein partners with which polα–primase is known to interact at
472 replication forks and telomeres. At replication forks, polα–primase interacts with several proteins
473 associated with the replisome, including MCM10, AND-1, RPA, RFC, polδ, and the CMG
474 helicase [19,36,59], as well as multiple nucleosome core proteins [60,61]. Despite this extensive
475 network of contacts, polα–primase appears to remain conformationally dynamic [62,63]. Moreover,
476 recent structures of polα–primase bound to a replisome complex revealed a manner of flexible
477 association [36] that is unlikely to restrict even the extraordinarily broad range of motion that we
478 observed in our structures.
479 Polα–primase appears to remain similarly dynamic at telomeres, as three distinct
480 configurations of the polα–primase/CST complex bound to single-stranded DNA have been
481 observed [37,48,49]. The first structure, termed a recruitment complex, showed polα–primase in the
482 AI configuration with CST primarily associated via interactions with $POLA1_C$ [48]. The second

483 structure, designated a pre-initiation complex, showed polα–primase in an extended
484 configuration in which $POLA1_{cat}$ and the tetramer core are scaffolded by CST [37]. In this
485 configuration, the PRIM1 and $POLA1_{cat}$ active sites are locked in opposing orientations, with
486 $PRIM2_C$ free to move between them. This arrangement was proposed to facilitate RNA
487 synthesis and transfer of the RNA primer from primase to polα. Consistent with this notion, CST
488 has been shown to enhance primer synthesis, without substantially affecting primer length
489 [37,64,65]. Moreover, docking of our DSA structures into the pre-initiation complex revealed
490 additional interactions between the thumb domain of $POLA1_{cat}$ and the CTC1 and STN1
491 subunits of CST (**Supplemental Fig. S32**), which may stabilize the closed, substrate-bound
492 conformation of the polymerase catalytic core, supporting $POLA1_{cat}$ in its competition with
493 PRIM1 for the RNA primer. The positions of our structures in these models, however, suggest
494 DNA elongation would lead to minor and then severe steric clashes between $PRIM2_C$ and the
495 tetramer core as the primer reaches lengths of 10 and 18 nucleotides, respectively. In this
496 configuration, these clashes are likely to prevent synthesis of a full-length primer. This
497 conclusion is supported by the third structure of polα–primase/CST, which showed $POLA1_{cat}$
498 associated with CST, as in the pre-initiation complex, but with the tetramer core disengaged [49].
499 Together, these observations suggest the tetramer core is not tightly bound to CST and
500 dissociates prior to or during DNA elongation by $POLA1_{cat}$. Based on the evolutionarily
501 relationship between CST and RPA [66], we believe RPA likely scaffolds polα–primase in a similar
502 manner, explaining RPA-dependent stimulation of primer synthesis by polα–primase [67]. How
503 these interactions with RPA and CST are coordinated with those of other proteins at replication
504 forks and telomeres, as well as the effects these interactions may have on the conformational
505 and configurational dynamics of polα–primase and the mechanism of chimeric primer synthesis,
506 remain to be determined.
507
508 Conclusion
509 Determining structures of polα–primase with multiple catalytic substrates, along with
510 performing structure-based biochemical and biophysical analyses, has enabled us to propose a
511 comprehensive model for primer synthesis. This model is built upon the large body of available
512 structural and biochemical data while unifying and refining prevailing understanding of the
513 mechanistic basis for polα–primase function. Our structural data reveal that polα–primase both
514 remodels and reads the conformation of its primer/template substrate, which explains its unique
515 ability to regulate primer length (i.e., count). These data also reveal that polα–primase retains a
516 remarkable degree of conformational and configurational flexibility that is crucial to the intricate

517 functions of this complex and dynamic protein machine. A greater appreciation of the type and
518 degree of intrinsic flexibility that we have observed for polα–primase is anticipated as active
519 eukaryotic replisomes and other cellular machinery come into view at ever increasing resolution.
520
521
522 **Methods**
523
524 **Cloning and expression.** The genes encoding *Xenopus laevis* PRIM1 (NCBI accession
525 NP_001080819.1), PRIM2 (NCBI accession XP_018121015.1), POLA1 (NCBI accession
526 NP_001082055.1), and POLA2 (NCBI accession NP_001086972.1) were initially cloned into
527 pFastBac-HTa or pFastBac-HTb expression vectors (Invitrogen). PRIM1 was then subcloned
528 into a modified pET-19b expression vector (Novagen) encoding an HRV 3C-cleavable
529 decahistidine fusion tag. PRIM2 was subcloned into a modified pRSFDuet-1 expression vector
530 (Novagen). And, POLA1 and POLA2 were subcloned into an extensively modified, polygenic
531 pFastBac-1 expression vector (Invitrogen) encoding TEV-cleavable StrepII-GFP (POLA1) and
532 hexahistidine-MBP (POLA2) fusion tags. Protein expression and purification were facilitated by
533 truncating the gene encoding POLA1 to remove the N-terminal tail (residues 1–334). To
534 generate a $POLA1_{cat}$ construct, the catalytic core of POLA1 (residues 335–1254) was subcloned
535 into a modified pET-27b(+) expression vector (Novagen) encoding an HRV 3C-cleavable
536 hexahistidine-SUMO fusion tag. To generate a $\Delta POLA1_{cat}$ construct, the C-terminal domain of
537 POLA1 (residues 1262–1458) and POLA2 were subcloned into a modified, polycistronic
538 pCDFDuet-1 expression vector (Novagen) encoding an HRV 3C-cleavable hexahistidine
539 ($POLA1_C$) fusion tag. Production of baculovirus and expression of POLA1/POLA2 in *Spodoptera*
540 *frugiperda* Sf9 cells were performed by the Expression and Microbiology Core at UC Berkeley.
541 PRIM1/PRIM2, $POLA1_{cat}$, and $POLA1_C$/POLA2 were expressed in *Escherichia coli* Rosetta
542 2(DE3) cells. Cultures were grown at 37°C in LB medium supplemented with 34 mg/L
543 chloramphenicol, 100 mg/L ampicillin (PRIM1/PRIM2), 30 mg/L kanamycin (PRIM1/PRIM2 and
544 $POLA1_{cat}$), and 70 mg/L streptomycin ($POLA1_C$/POLA2). Upon reaching mid-log phase, cultures
545 were further supplemented with 100 mg/L ammonium iron(III) citrate (PRIM1/PRIM2) and 5 μM
546 $ZnSO_4$ (PRIM1/PRIM2 and $POLA1_C$/POLA2) and cooled to 16°C. Protein overexpression was
547 then induced by addition of IPTG to a final concentration of 0.5 mM. After overnight incubation,
548 cells were harvested by centrifugation. Cell pellets were stored at −80°C. *Acanthamoeba*
549 *castellanii* TLP2 was expressed as previously described [68-70]. Protein constructs are depicted in
550 **Supplemental Fig. S33**.

551

552 **Protein purification.** Cell pellets were thawed in lukewarm water and resuspended in buffer L
553 [50 mM sodium phosphate pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, 10 mM imidazole,
554 and cOmplete EDTA-free protease inhibitor cocktail (Roche)]. Both insect cells and bacterial
555 cells were lysed by Dounce homogenization followed by gentle sonication. The two lysates were
556 then combined and cleared by centrifugation prior to application onto a Ni-NTA Agarose column
557 (Qiagen) equilibrated in buffer L lacking protease inhibitor cocktail. The column was washed
558 with 15 column volumes (cv) of buffer N1 [50 mM sodium phosphate pH 7.5, 1 M NaCl, 5% (v/v)
559 glycerol, 1 mM $MgCl_2$, and 20 mM imidazole], followed by 5 cv of buffer N2 [50 mM
560 HEPES•NaOH pH 7.5, 300 mM NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 20 mM imidazole].
561 Protein was eluted by linearly transitioning over 15 cv to buffer N3 [50 mM HEPES•NaOH pH
562 7.5, 300 mM NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 500 mM imidazole]. Fractions containing
563 polα–primase were pooled and supplemented with 1 mM TCEP before fusion tags were
564 removed by overnight cleavage at 4°C. The following day, protein was diluted three-fold in
565 buffer H1 [50 mM HEPES•NaOH pH 7.5, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 1 mM TCEP] and
566 passed through a 0.22-µm filter before being applied to a heparin Sepharose column (Cytiva)
567 equilibrated in a 9:1 mixture of buffer H1 and buffer H2 [50 mM HEPES•NaOH pH 7.5, 1 M
568 NaCl, 5% (v/v) glycerol, 1 mM $MgCl_2$, and 1 mM TCEP]. The column was then washed with 5 cv
569 of the same 9:1 mixture before protein was eluted by linearly transitioning over 10 cv to only
570 buffer H2. Fractions containing polα–primase were pooled and again applied to a Ni-NTA
571 Agarose column equilibrated in buffer L lacking protease inhibitor cocktail. Protein was eluted in
572 10 cv of the same buffer, and fractions containing polα–primase were pooled and supplemented
573 with 1 mM TCEP. Protein was then concentrated by diafiltration, cleared by centrifugation, and
574 applied to a Superose 6 Increase column (Cytiva) equilibrated in buffer S [20 mM
575 HEPES•NaOH pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM TCEP]. Protein was eluted in 1.5
576 cv of the same buffer, and fractions containing polα–primase were pooled and concentrated by
577 diafiltration. Aliquots were used immediately or frozen in liquid nitrogen and stored at −80°C.
578 $POLA1_{cat}$ and $\Delta POLA1_{cat}$ were purified in the same manner as polα–primase. AcaTLP2 was
579 purified as previously described [68-70].
580

581 **Oligonucleotide synthesis and purification.** RNA, DNA, and chimeric RNA-DNA
582 oligonucleotides were purchased from Integrated DNA Technologies and resuspended in buffer
583 A (10 mM Mes•NaOH pH 6.5 and 40 mM NaCl). To generate 5′-triphosphorylated primers, 500
584 µM 5′-monophosphorylated primer and 550 µM template were annealed by heating to 85°C and

585 then slowly cooling to 22°C, producing a double-stranded substrate for the tRNA-modifying
586 enzyme AcaTLP2 [70,71]. AcaTLP2 reaction mixtures were prepared with 50 μM enzyme, 100 μM
587 primer/template, 150 μM ATP, 1 mM GTP, 25 mM Hepes•NaOH pH 7.5, 125 mM NaCl, 10 mM
588 $MgCl_2$, and 3 mM DTT and incubated at 22°C. After 1 h, an equal volume of quenching/loading
589 buffer containing 80% (v/v) formamide, 100 mM EDTA pH 8.0, and 220 μM competitor
590 oligonucleotide was added. Samples were then heated to 85°C and slowly cooled to 22°C to
591 denature the primer and the template. Oligonucleotides were separated by denaturing urea-
592 polyacrylamide gel electrophoresis (PAGE) as previously described [72], except electrophoresis
593 was performed at 30 W for 9 h. For each reaction, the band corresponding to the
594 triphosphorylated primer was located by UV-shadowing and cut from the gel. The primer was
595 then extracted into 24 mL of 300 mM NaOAc pH 5.5 using a crush-and-soak procedure in which
596 the sample was placed at −80°C until frozen, thawed in a water bath at 42°C, and then rocked
597 overnight at 4°C. The following day, the supernatant was decanted and overnight extraction was
598 repeated using a second 24-mL volume of 300 mM NaOAc pH 5.5. The two volumes were
599 combined, passed through a 0.22-μm filter, and split into four 12-mL aliquots. 36 mL of cold
600 100% ethanol was then added to each aliquot. Samples were thoroughly mixed, incubated at
601 −80°C for 2 h, warmed to −20°C, and centrifuged at 12,000×g and 4°C for 1 h. After decanting
602 the supernatant, pellets were air-dried for ~15 min and resuspended in 1 mL of 300 mM NaOAc
603 pH 5.5, using the same 1 mL of buffer for all pellets. The sample was then split into two 0.5-mL
604 aliquots, and 1.5 mL of cold 100% ethanol was added to each. Samples were again thoroughly
605 mixed, incubated at −80°C for 2 h, warmed to −20°C, and centrifuged at 12,000×g and 4°C for 1
606 h. After decanting the supernatant, pellets were air-dried for ~15 min and then gently rinsed with
607 200 μL of cold 95% (v/v) ethanol. The ethanol was removed and the pellet was air-dried for an
608 additional ~15 min. The triphosphorylated primer was then resuspended in 100 μL of buffer A,
609 using the same 100 μL of buffer for both pellets, and stored at −80°C. Oligonucleotide
610 sequences are provided in **Supplemental Table S2**.
611
612 **Cryo-EM sample preparation and data collection.** Polα–primase was concentrated to ≥ 3 μM
613 by diafiltration. To generate RNA-primed templates, 110 μM 5′-triphosphorylated RNA primer
614 and 100 μM DNA template were annealed by heating to 85°C and then slowly cooling to 4°C.
615 Substrate complexes were prepared by incubating 3 μM polα–primase, 3.6 μM primer/template,
616 and 1 mM dGTP on ice for 30 min. Using a Vitrobot Mark IV (FEI), 2.5 μL of sample was applied
617 to glow-discharged, holey-carbon R1.2/1.3 200-mesh copper grids (Quantifoil) at 4°C and 100%
618 humidity. Grids were immediately blotted with filter paper (Ted Pella) and flash-frozen in liquid

619 ethane. Grids prepared without substrate or with DT substrate were loaded into a Titan Krios G4
620 electron microscope (Thermo Scientific) operated at 300 kV for automated data collection using
621 SerialEM [73]. Movies of dose-fractionated frames were acquired with a K3 BioQuantum direct
622 electron detector (Gatan) functioning in counting mode. Each of the datasets was collected from
623 a single grid in a single imaging session. Grids prepared with DI or DE substrates were loaded
624 into a Talos Arctica electron microscope (Thermo Scientific) operated at 200 kV for automated
625 data collection using Leginon [74]. Movies of dose-fractionated frames were acquired with a K2
626 Summit direct electron detector (Gatan) functioning in counting mode. Each of the datasets was
627 collected in three to four imaging sessions, using a different grid for each session. Additional
628 details for each of the datasets are provided in **Supplemental Tables S1, S3, S4, and S6**.
629
630 **Cryo-EM data processing.** For pre-processing of all datasets, motion-correction of raw movies
631 was performed using MotionCor2 [75] and patch CTF estimation was performed for motion-
632 corrected micrographs using cryoSPARC [76]. All initial processing steps were also performed in
633 cryoSPARC. Particles from each dataset were picked by a combination of blob picking, template
634 picking, and neural-network picking using Topaz [77]. Templates for template picking were
635 selected from the highest-resolution 2D class averages resulting from blob picking, while Topaz
636 was trained from a small set of manually picked particles. Particles were extracted with a box
637 size of either 400 (substrate-free and DT datasets) or 360 (DI and DE datasets) pixels and then
638 downsampled to 120 pixels (pixel size of 2.73 Å) by Fourier truncation. 2D classification was
639 first performed separately for particles picked by each method. Selected particles were then
640 combined and duplicate and triplicate particles were removed before final 2D classification was
641 performed. For the DT dataset, particles were sorted manually before initial 3D reconstructions
642 were generated by ab initio reconstruction. For all other datasets, initial 3D reconstructions were
643 generated by multiclass ab initio reconstruction followed by heterogeneous refinement.
644 Satisfactory 3D reconstructions could only be obtained for the AI complex (substrate-free and
645 DT datasets), the DNA synthesis assembly (DI, DE, and DT datasets), and the tetramer core (all
646 datasets). Because of extreme configurational and conformational heterogeneity, satisfactory
647 3D reconstructions could not be obtained for the full polα–primase complex bound to a
648 primer/template substrate. For reconstructions of the tetramer core, conformational
649 heterogeneity was reduced through additional iterations of multiclass ab initio reconstruction
650 and heterogeneous refinement. For all 3D reconstructions, homogeneous refinement was
651 initially performed using downsampled particles and then repeated after particles were
652 recentered and re-extracted either without downsampling or with reduced downsampling. For

653 the AI complex from the substrate-free dataset and the DNA synthesis assembly from the DI,
654 DE, and DT datasets, all subsequent processing steps were performed in RELION-3.1 [78].
655 Particle alignments were first improved by focused 3D refinement, before heterogeneity was
656 reduced through focused, alignment-free 3D classification (following particle subtraction, as
657 necessary). CTF refinement was also performed to correct for per-particle defocus differences,
658 as well as for imperfect microscope alignment during data collection. Using refined CTF
659 parameters, focused 3D refinement was repeated for each final set of particles. To improve
660 high-resolution features in the reconstructions, despite local quality differences, maps were
661 sharpened using DeepEMhancer [79]. Data-processing workflows are outlined in **Supplemental**
662 **Figs. S5, S10, S12, and S27**.
663 A modified workflow was employed to obtain reconstructions of the full polα–primase
664 complex bound to primer/template substrates. Particles corresponding to configurations DI-1,
665 DI-2, DE-1, and DE-2 were manually selected by examination of 2D class averages and
666 downsampled to 64 pixels (pixel size of 5.12 Å) by Fourier truncation. Ab initio reconstruction
667 and non-uniform refinement were then performed to generate low-resolution 3D reconstructions.
668 For particles initially assigned to configuration DI-2, multiclass ab initio reconstruction and
669 heterogeneous refinement revealed a possible third configuration of the DI complex. All
670 processing steps in the modified workflow were performed in cryoSPARC. Modified data-
671 processing workflows are outlined in **Supplemental Figs. S19 and S20**.
672
673 **Model building and refinement.** Homology models of *X. laevis* polα–primase subunits were
674 generated from corresponding *Homo sapiens* X-ray crystal structures using SWISS-MODEL [80].
675 For the AI complex, POLA1, POLA2$_C$, and PRIM2 homology models (from PDB accession
676 5EXR) were docked as rigid bodies into the high-resolution partial reconstruction. Following
677 manual modification, atomic coordinates and isotropic temperature factors were refined against
678 the cryo-EM density for all non-hydrogen atoms using Ramachandran, secondary structure, and
679 rotamer restraints. Hydrogen atoms were placed in riding positions and were not refined against
680 the density. Both the refined model and a PRIM1 homology model (from PDB accession 4BPU)
681 were then docked as rigid bodies into the complete reconstruction. After manual adjustment,
682 atomic coordinates and isotropic temperature factors were refined as described above, except
683 reference restraints were applied based on the *X. laevis* PRIM2/POLA1/POLA2$_C$ structure and
684 the *H. sapiens* PRIM1 structure. For the DI subcomplex, POLA1$_{cat}$ and PRIM2$_C$ homology
685 models (from PDB accessions 4QCL and 5F0Q) were docked as rigid bodies into the
686 corresponding reconstruction. Following manual modification of the protein and de novo addition

687 of the substrate, atomic coordinates and isotropic temperature factors were refined as described
688 above, except base pairing and base stacking restraints were added for the oligonucleotide
689 duplex. For the DE subcomplex, the refined $POLA1_{cat}$ model from the DI subcomplex was
690 docked as a rigid body into the high-resolution partial reconstruction. After manual adjustment of
691 the protein and de novo addition of the substrate, atomic coordinates and isotropic temperature
692 factors were refined as described above for the DI subcomplex. The refined models of $POLA1_{cat}$
693 from the partial DE structure and $PRIM2_C$ from the DI structure were then docked as rigid bodies
694 into the complete reconstruction. The oligonucleotide substrate from the partial DE subcomplex
695 and the 3′-overhang and first base pair from the *H. sapiens* $PRIM2_C$ crystal structure (PDB
696 accession 5F0Q) were also docked as rigid bodies. Following manual modification of the
697 protein, including adjustment of the template-binding residues in $PRIM2_C$ to match those in the
698 *H. sapiens* crystal structure, and de novo addition of the intervening region of the substrate,
699 atomic coordinates and isotropic temperature factors were refined as described above for the DI
700 subcomplex, except reference restraints were applied based on the *X. laevis* $POLA1_{cat}$ and
701 $PRIM2_C$ structures. For the DT subcomplex, the refined $POLA1_{cat}$ model from the DE
702 subcomplex was docked as a rigid body into the high-resolution partial reconstruction. After
703 manual adjustment of the protein and de novo addition of the substrate, atomic coordinates and
704 isotropic temperature factors were refined as described above for the DI subcomplex. The
705 refined models of $POLA1_{cat}$ from the partial DT structure and $PRIM2_C$ from the complete DE
706 structure were then docked as rigid bodies into the reconstruction of DTS-1. The oligonucleotide
707 substrate from the partial DT subcomplex and the 3′-overhang and first nine base pairs from the
708 DE subcomplex were also docked as rigid bodies. Following manual modification of the protein
709 and de novo addition of the intervening region of the substrate, atomic coordinates and isotropic
710 temperature factors were refined as described above for the complete DE subcomplex.
711 Similarly, the refined model of $POLA1_{cat}$ from the partial DT structure, the oligonucleotide
712 substrate from the partial DT structure, and the 3′-overhang and first nine base pairs from the
713 complete DE structure were docked as rigid bodies into the reconstruction of DTS-2. Following
714 manual modification of the protein and de novo addition of the intervening region of the
715 substrate, atomic coordinates and isotropic temperature factors were refined as described
716 above for the complete DE subcomplex. For all structures, manual protein modification and de
717 novo substrate addition were performed in Coot [81], while docking and refinement were
718 performed in PHENIX [82]. If possible, rotamers were determined from the cryo-EM density. In
719 instances in which the correct conformation of the side chain was ambiguous, rotamers were
720 selected to match those in higher-resolution cryo-EM structures from this study or those in

721 previously published crystal structures. In instances in which both the correct conformation was
722 ambiguous and an appropriate reference structure was unavailable, rotamers were manually
723 selected to avoid steric clashes, and if possible, form favorable electrostatic interactions. Final
724 models were validated using MolProbity [83].
725 Structures of the complete polα–primase complex in configurations DI-1, DI-2, DE-1, and
726 DE-2 and of the TC subcomplex in configurations 1, 2, and 3 were generated from rigid bodies.
727 For all structures, the TC subcomplex was reconstituted from three rigid bodies taken from the
728 AI complex, while the DI and DE subcomplexes were each maintained as a single rigid body.
729 Docking into low-resolution reconstructions was performed in ChimeraX [84].
730 Refinement and validation statistics are provided in **Supplemental Figs. S7, S14, S15,**
731 **S19, S20, and S29 and Supplemental Tables S1, S3, S4, and S6**. Figures were prepared in
732 PyMOL (https://www.pymol.org), UCSF Chimera [85], or UCSF ChimeraX.
733
734 **Primer elongation.** To generate an RNA-primed template for elongation assays, 110 μM 5′-
735 triphosphorylated RNA primer and 100 μM DNA template were annealed by heating to 85°C
736 and then slowly cooling to 4°C. Reaction mixtures were then prepared with 50 nM $POLA1_{cat}$,
737 $\Delta POLA1_{cat}$, $POLA1_{cat}/\Delta POLA1_{cat}$, or polα–primase; 500 nM primer/template; 0.1 μM [α-^{32}P]-ATP
738 (3,000 Ci/mmol; PerkinElmer) or [α-^{32}P]-dATP (3,000 Ci/mmol; PerkinElmer); 9.9 μM ATP or
739 dATP; 200 μM GTP, CTP, and UTP; 20 μM dGTP, dCTP, and dTTP; 20 mM HEPES•NaOH pH
740 7.5; 150 mM NaCl; 5 mM $MgCl_2$; and 1 mM TCEP and incubated at 22°C. 8-μL aliquots were
741 periodically withdrawn, quenched with 2 μL of 100 mM EDTA pH 8.0, and incubated at 65°C for
742 15 min. To denature the primer and the template, 10 μL of loading buffer [80% (w/v) formamide,
743 10 mM EDTA pH 8.0, 1 mg/mL bromophenol blue, 1 mg/mL xylene cyanol, and 1.5 μM
744 competitor oligonucleotide] was added and samples were incubated at 85°C for 10 min and then
745 slowly cooled to 22°C. Elongation products were separated by denaturing urea-PAGE as
746 previously described [72], except 18% polyacrylamide (19:1 acrylamide/bisacrylamide) gels were
747 run at 20 W for 7 h and visualized by phosphorimaging using an Amersham Typhoon
748 biomolecular imager (Cytiva). Experiments were performed in triplicate.
749
750 **Primer/template binding.** Binding reactions were performed in 20 mM HEPES•NaOH pH 7.5,
751 150 mM NaCl, 5 mM $MgCl_2$, 1 mM TCEP, 20 μM dGTP, and 0.05% (v/v) Tween-20 with 10 nM
752 Cy5-labeled primer/template and 0.15–5,120 nM $POLA1_{cat}$. Immediately after mixing, samples
753 were centrifuged at 20,000×g and 4°C for 5 min, incubated on ice for 10 min, and then
754 incubated at 20°C for 10 min. Microscale thermophoresis data were collected using a Monolith

NT.115 instrument (NanoTemper). All experiments were performed using the red fluorescence laser ($\lambda_{excitation}$ = 600–640 nm; $\lambda_{emission}$ = 660–720 nm) at 80% power and the infrared laser (λ = 1480 nm) at medium power. Temperature-related intensity changes were monitored at 20°C with a pre-IR phase of 3 s, followed by an IR-on phase of 20 s, and a post-IR phase of 1 s. Data quality was assessed by monitoring fluorescence quenching and photobleaching, as well as protein aggregation and adsorption. Experiments were performed in triplicate. For each protein-substrate binding titration, normalized fluorescence values [$F_{norm} = (F_{1\ (IR\text{-}hot)})/(F_{0\ (IR\text{-}cold)})$] were tracked as parts per thousand and converted to ΔF_{norm} [$F_{norm\ (n)} - F_{norm\ (initial)}$]. ΔF_{norm} values for individual isotherms were normalized to generate the fraction of bound protein (f_B). Dissociation constants (K_d) were derived from fits to the data using a one-site binding model [$f_B = x/(K_d+x)$, where x is the total concentration of $POLA1_{cat}$].

Data availability. Structures were deposited in the Protein Data Bank under accession codes 8G99 (AI complex, partial), 8G9F (AI complex, complete), 8G9L (DI subcomplex), 8G9N (DE subcomplex, partial), 8G9O (DE subcomplex, complete), 8UCU (DT subcomplex, partial), 8UCV (DT subcomplex 1), and 8UCW (DT subcomplex 2). 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-42140 (DT subcomplex, partial), EMD-42141 (DT subcomplex 1, complete), EMD-42142 (DT subcomplex 2, complete), EMD-29888 (TC subcomplex, conformation 1), EMD-29889 (TC subcomplex, conformation 2), and EMD-29891 (TC subcomplex, 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. Jones, M.L., Aria, V., Baris, Y. & Yeeles, J.T.P. How Pol α-primase is targeted to replisomes to prime eukaryotic DNA replication. *Molecular Cell* **83**, 2911-2924.e16 (2023).
37. He, Q. et al. Structures of the human CST-Polalpha-primase complex bound to telomere templates. *Nature* (2022).
38. 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).
39. 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).
40. 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).
41. Baranovskiy, A.G. et al. Insight into the Human DNA Primase Interaction with Template-primer. *J Biol Chem* (2016).
42. Xie, P. A model for dynamics of primer extension by eukaryotic DNA primase. *Eur Biophys J* **40**, 1157-65 (2011).
43. 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).
44. 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).
45. 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).
46. He, Q. et al. Structures of human primosome elongation complexes. *Nature Structural & Molecular Biology* **30**, 579-583 (2023).
47. Yuan, Z., Georgescu, R., Li, H. & O'Donnell, M.E. Molecular choreography of primer synthesis by the eukaryotic Pol α-primase. *Nature Communications* **14**, 3697 (2023).
48. 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).
49. He, Y. et al. Structure of Tetrahymena telomerase-bound CST with polymerase alpha-primase. *Nature* (2022).
50. 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).
51. 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).
52. 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).
53. Hogg, M. et al. Structural basis for processive DNA synthesis by yeast DNA polymerase varepsilon. *Nat Struct Mol Biol* **21**, 49-55 (2014).
54. 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).

55. 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).
56. Neidle, S. *Nucleic Acid Structure and Recognition*, (Oxford University Press, New York, 2002).
57. Lu, X.J., Shakked, Z. & Olson, W.K. A-form conformational motifs in ligand-bound DNA structures. *J Mol Biol* **300**, 819-40 (2000).
58. 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).
59. 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 (2021).
60. 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).
61. 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).
62. Sun, J. et al. The architecture of a eukaryotic replisome. *Nat Struct Mol Biol* **22**, 976-82 (2015).
63. 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).
64. Lue, N.F., Chan, J., Wright, W.E. & Hurwitz, J. The CDC13-STN1-TEN1 complex stimulates Pol α activity by promoting RNA priming and primase-to-polymerase switch. *Nature Communications* **5**, 5762 (2014).
65. 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).
66. Lim, C.J. et al. The structure of human CST reveals a decameric assembly bound to telomeric DNA. *Science* **368**, 1081-1085 (2020).
67. 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).
68. 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).
69. 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).
70. 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).
71. 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).
72. 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).
73. Mastronarde, D.N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51 (2005).
74. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41-60 (2005).
75. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
76. 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).
77. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160 (2019).
78. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**(2018).
79. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biol* **4**, 874 (2021).
80. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-w303 (2018).

81. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
82. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
83. Davis, I.W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-83 (2007).
84. Goddard, T.D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).
85. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).

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., N.P.B., and B.F.E. designed experiments; J.E.J. provided reagents; E.A.M., L.E.S., C.L.D., and N.P.B. collected data; E.A.M., L.E.S., N.P.B., M.D.O., W.J.C., and B.F.E. analyzed data; E.A.M., L.E.S., N.P.B., 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 TC 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 AI complex. The points at which the complex were split and then reassembled are indicated with arrowheads. **E**. Schematic overlay aligned on the central rigid body 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 TC subcomplex, or one of the distinct configurations of the DI and DE complexes (DI-1, DI-2, DE-1, or DE-2) is indicated below the class averages. ~3% 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

1023 the primers, respectively. The positions of the interdomain linkers are indicated with
1024 arrowheads. **D,E**. Distinct configurations of the DI and DE complexes. Models were constructed
1025 from four rigid bodies, comprised of the $POLA1_{cat}$/$PRIM2_C$ models shown in panel C and the
1026 PRIM1/$PRIM2_N$/$POLA1_C$/$POLA2_C$ models shown in **Fig. 2**, and docked into the low-resolution
1027 3D reconstructions shown in **Supplemental Figs. S19 and S20**. **D**. DI and DE complexes
1028 oriented as in the 2D class averages. **E**. DI and DE complexes aligned on the tetramer core.
1029
1030 **Figure 4. Primer/template binding and elongation. A**. Separation and detection of RNA and
1031 RNA-DNA elongation products by denaturing gel electrophoresis and autoradiography.
1032 Intermediate and doublet bands result from a mixture of RNA and DNA incorporation at
1033 positions 7–10 (**Supplemental Fig. S22**). **B**. Substrate schematic. Triphosphate moieties are
1034 depicted as brown circles. [α-^{32}P]-labelled ATP and [α-^{32}P]-labelled dATP are only incorporated
1035 at position 7 in the primer. **C**. Quantitation of the 30-min time points from the gel in panel A. The
1036 intensities of the elongation products are normalized to the intensity of the 30-mer. **D**. Kinetics of
1037 30-mer formation. **E**. Binding of DNA/DNA, RNA/DNA, and RNA-DNA/DNA duplexes by
1038 $POLA1_{cat}$. Data in panels D and E are presented as the mean ± SEM from three replicate
1039 experiments.
1040
1041 **Figure 5. Substrate recognition and duplex remodeling.** In each panel, the DI and DE
1042 structures are positioned on the left and the right, respectively. **A**. Schematic of substrate-
1043 binding contacts in $POLA1_{cat}$ (blue) and $PRIM2_C$ (green). **B**. Recognition of an incoming dGTP
1044 by $POLA1_{cat}$. Phosphodiester bond formation is prevented by a 2′,3′-dideoxycytidine (ddC) at the
1045 end of the primers. **C**. Recognition of RNA/DNA and DNA/DNA duplexes through substrate-
1046 adaptive conformational changes in $POLA1_{cat}$. **D**. Recognition of distinct primer/template
1047 conformations by $PRIM2_C$. **E**. RNA/DNA and RNA-DNA/DNA duplexes extracted from the DI
1048 and DE subcomplexes. Major (M) and minor (m) grooves are indicated. Numbers to the left of
1049 the duplexes indicate the base step starting from the 5′-end of the primers. **F**. Analysis of duplex
1050 conformation. Z_P is the interstrand phosphate-phosphate distance after projection onto the
1051 helical (Z) axis. The regions of the duplexes bound by $POLA1_{cat}$ (blue) and $PRIM2_C$ (green) are
1052 shown below the plots.
1053
1054 **Figure 6. Configurational rearrangements upon DNA termination. A**. Substrate schematic.
1055 Triphosphate moieties are depicted as brown circles. The primer terminates in a 2′,3′-
1056 dideoxycytidine to prevent incorporation of dGTP. **B**. Representative 2D class averages and

1057 corresponding subunit schematics. The percentage of particles belonging to the DT
1058 subcomplex, the TC subcomplex, or the AI complex is indicated below the class averages. No
1059 2D class averages were observed with clear density for the complete DT complex. All final 2D
1060 class averages are shown in **Supplemental Fig. S28**. **C**. Masked 3D reconstruction of $POLA1_{cat}$
1061 bound to a full-length primer/template. **D**. 3D reconstructions and molecular models of $POLA1_{cat}$
1062 bound to a full-length primer/template with (DT subcomplex 1, left) and without (DT subcomplex
1063 2, right) simultaneous association of $PRIM2_C$. The positions of the interdomain linkers are
1064 indicated with arrowheads. **E**. Analysis of duplex conformation. Z_P is the interstrand phosphate-
1065 phosphate distance after projection onto the helical (Z) axis. The regions of the duplexes bound
1066 by $POLA1_{cat}$ (blue) and $PRIM2_C$ (green) are shown below the plots. **F**. RNA-DNA/DNA duplexes
1067 extracted from DT subcomplex 1 (left) and DT subcomplex 2 (right). Major (M) and minor (m)
1068 grooves are indicated. Numbers to the left of the duplexes indicate the base step starting from
1069 the 5′-end of the primers.
1070
1071 **Figure 7. Structural rearrangements during chimeric primer synthesis. A**. Stages of RNA
1072 and DNA synthesis. $PRIM2_C$ remains bound to the 5′-end of the primer during all stages of
1073 synthesis. The transition from RNA elongation to DNA initiation occurs when the 3′-end of the
1074 primer is transferred from PRIM1 to $POLA1_{cat}$. After generation of a full-length primer, DNA
1075 elongation is terminated by dissociation of polα–primase from the primer/template. Protein-
1076 protein contacts between $PRIM2_C$ and $POLA1_{cat}$ are represented by a dashed line. **B**. Subunit
1077 schematics showing configurational rearrangements throughout the catalytic cycle of primer
1078 synthesis. Schematics are aligned on the tetramer core (PRIM1/$PRIM2_N$/$POLA1_C$/$POLA2_C$). The
1079 RNA initiation and RNA elongation configurations were derived from hypothetical models in
1080 which the positions of $POLA1_{cat}$ are indeterminate and the positions of $PRIM2_C$ are dictated by
1081 the geometric requirements of phosphodiester bond formation in the active site of PRIM1. The
1082 DNA initiation and DNA elongation configurations correspond to DI-2 and DE-2, respectively
1083 (**Fig. 3**). The DNA termination configuration was derived from a hypothetical model in which the
1084 DE subcomplex in DE-2 was replaced by DT subcomplex 1 while retaining the position of
1085 $PRIM2_C$. Loose association of $POLA1_{cat}$ and $PRIM2_C$ with the full-length primer/template is
1086 indicated with wobble lines. At each stage of primer synthesis, the configuration depicted is
1087 likely representative of a broad configurational/conformational distribution.

Figure 1

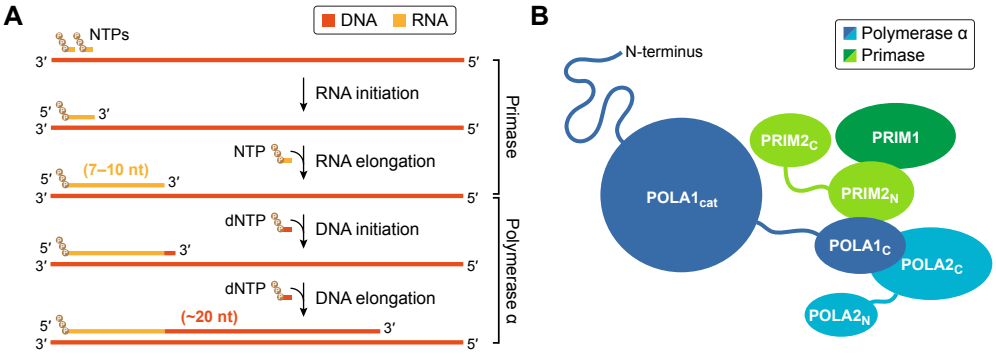

A
DNA
RNA
NTPs
3′
5′
RNA initiation
5′
3′
3′
5′
NTP
RNA elongation
5′
(7–10 nt)
3′
3′
5′
dNTP
DNA initiation
5′
3′
3′
5′
dNTP
DNA elongation
5′
(~20 nt)
3′
3′
5′
Primase
Polymerase α
B
Polymerase α
Primase
N-terminus
POLA1cat
PRIM2C
PRIM1
PRIM2N
POLA1C
POLA2C
POLA2N

Figure 2

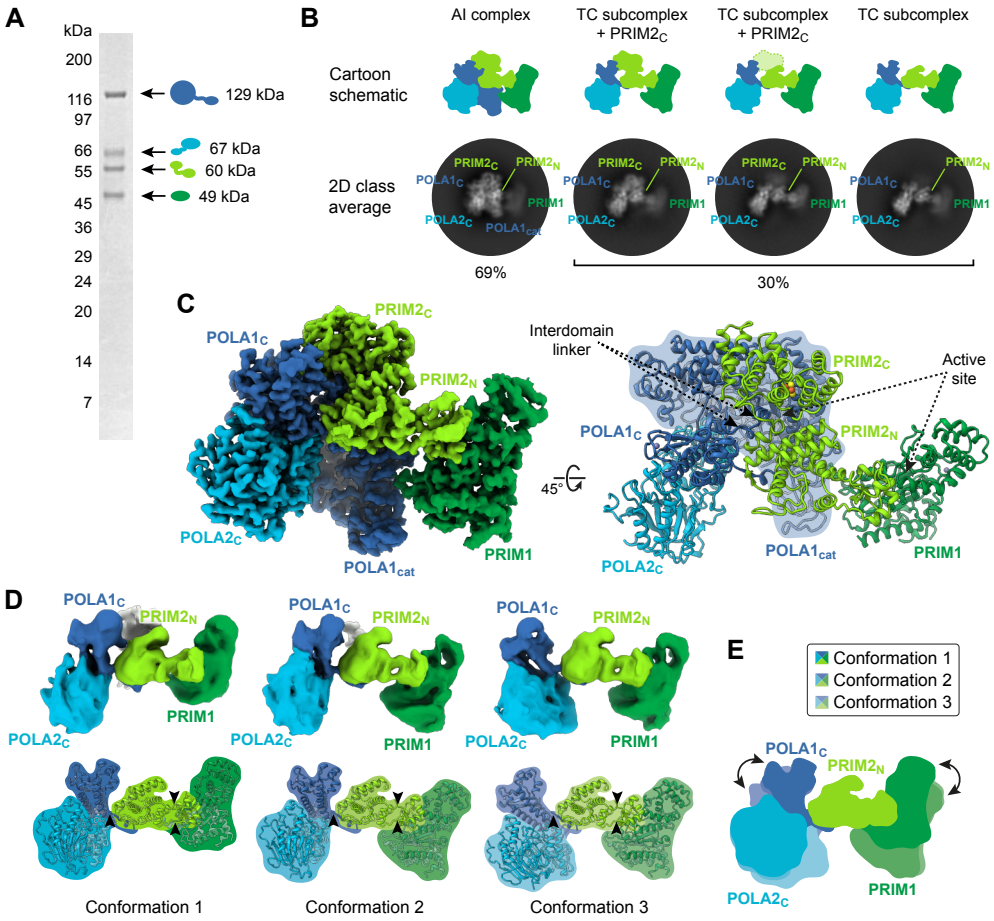

A
kDa
200
116
97
66
55
45
36
29
24
20
14
7
129 kDa
67 kDa
60 kDa
49 kDa
B
AI complex
TC subcomplex + PRIM2C
TC subcomplex + PRIM2C
TC subcomplex
Cartoon schematic
2D class average
PRIM2C
PRIM2N
POLA1C
PRIM1
POLA2C
POLA1cat
69%
30%
C
PRIM2C
POLA1C
PRIM2N
POLA2C
POLA1cat
PRIM1
Interdomain linker
Active site
45°
D
POLA1C
PRIM2N
PRIM1
POLA2C
Conformation 1
Conformation 2
Conformation 3
E
Conformation 1
Conformation 2
Conformation 3

Figure 3

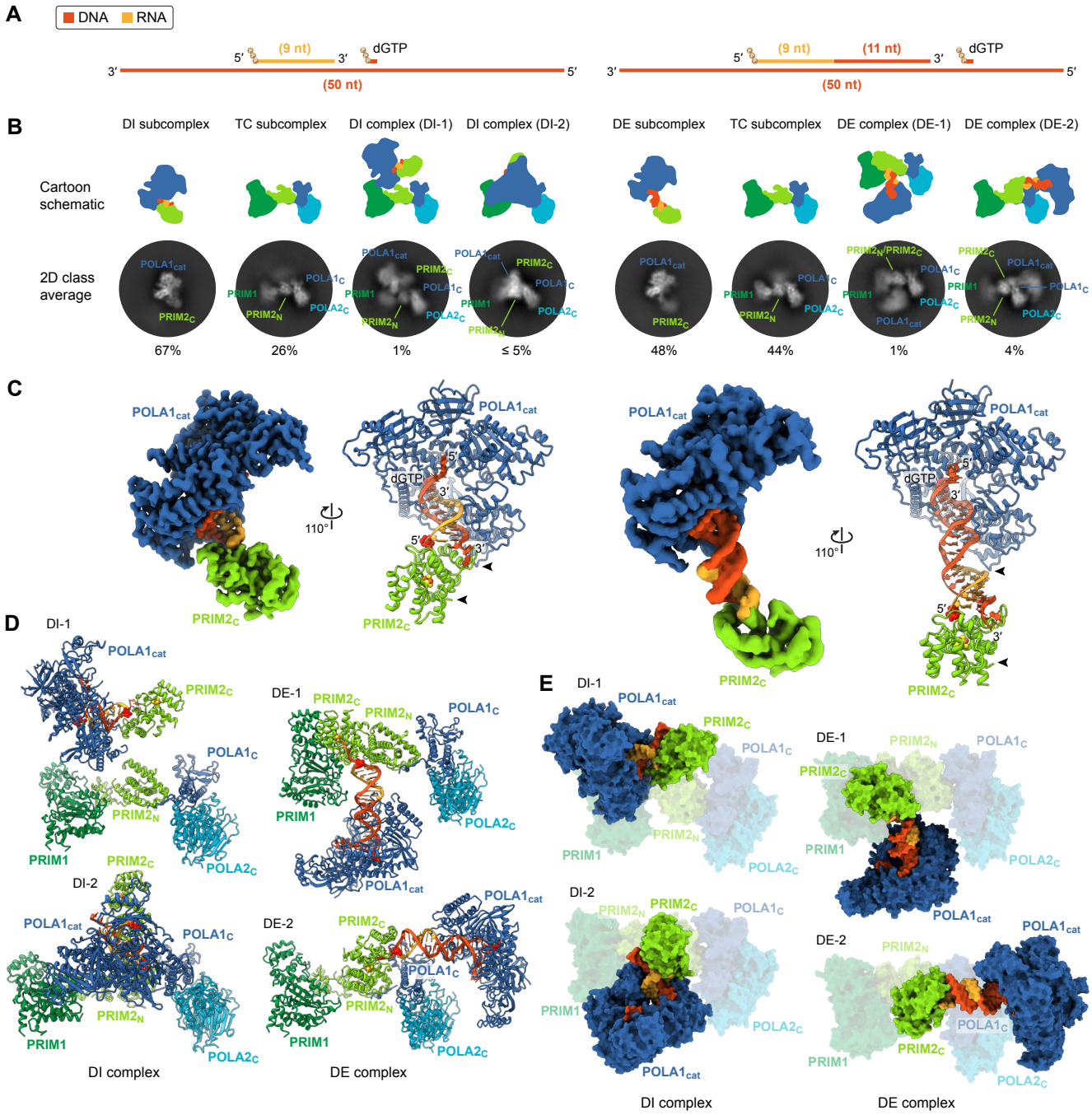

A
DNA
RNA
(9 nt)
dGTP
(50 nt)
(11 nt)
B
DI subcomplex
TC subcomplex
DI complex (DI-1)
DI complex (DI-2)
DE subcomplex
TC subcomplex
DE complex (DE-1)
DE complex (DE-2)
Cartoon schematic
2D class average
67%
26%
1%
≤ 5%
48%
44%
1%
4%
C
110°
D
DI-1
DE-1
DI-2
DE-2
DI complex
DE complex
E
DI complex
DE complex

Figure 4

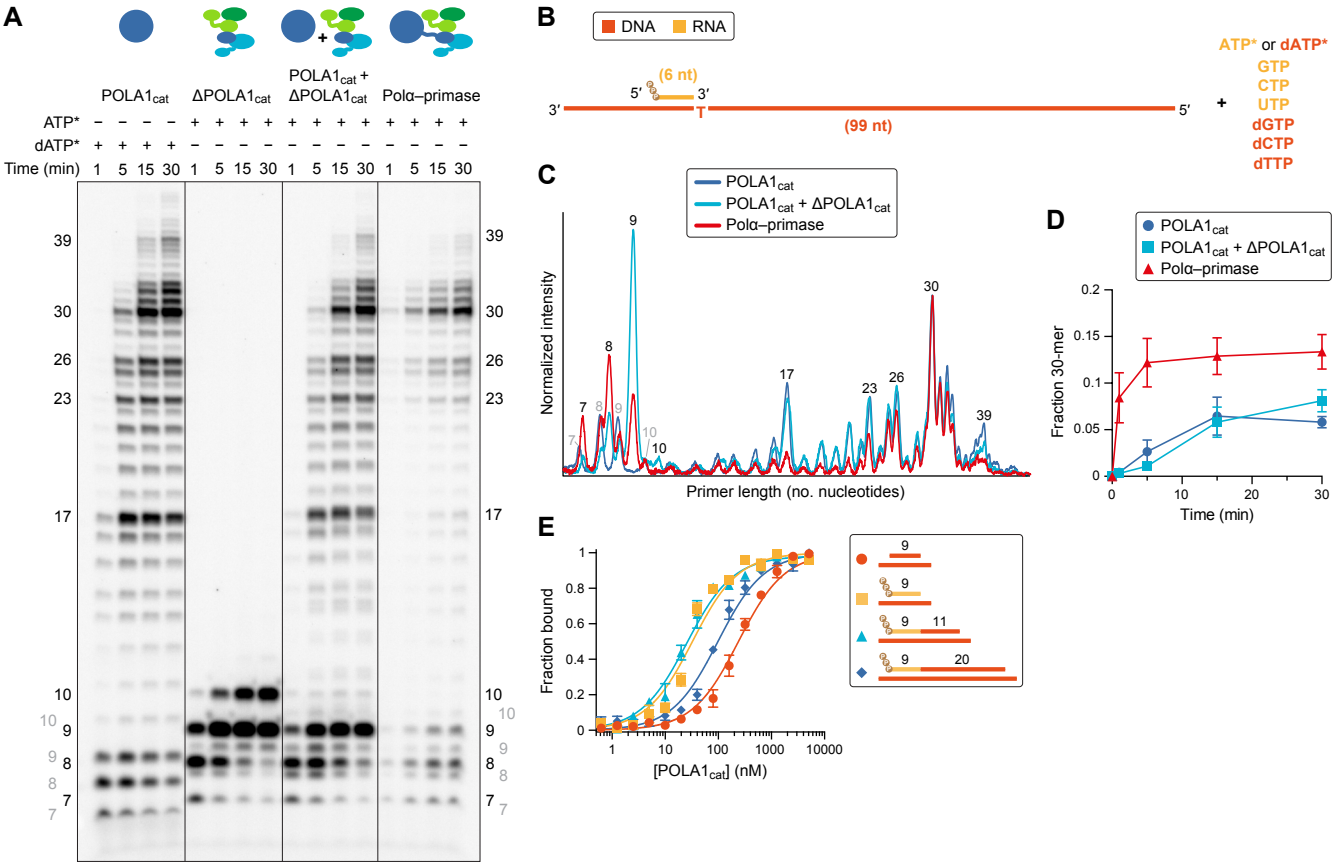

A
POLA1cat
ΔPOLA1cat
POLA1cat + ΔPOLA1cat
Polα–primase
ATP*
dATP*
Time (min)
B
DNA
RNA
(6 nt)
(99 nt)
ATP* or dATP*
GTP
CTP
UTP
dGTP
dCTP
dTTP
C
POLA1cat
POLA1cat + ΔPOLA1cat
Polα–primase
Normalized intensity
Primer length (no. nucleotides)
D
Fraction 30-mer
Time (min)
E
Fraction bound
[POLA1cat] (nM)

Figure 5

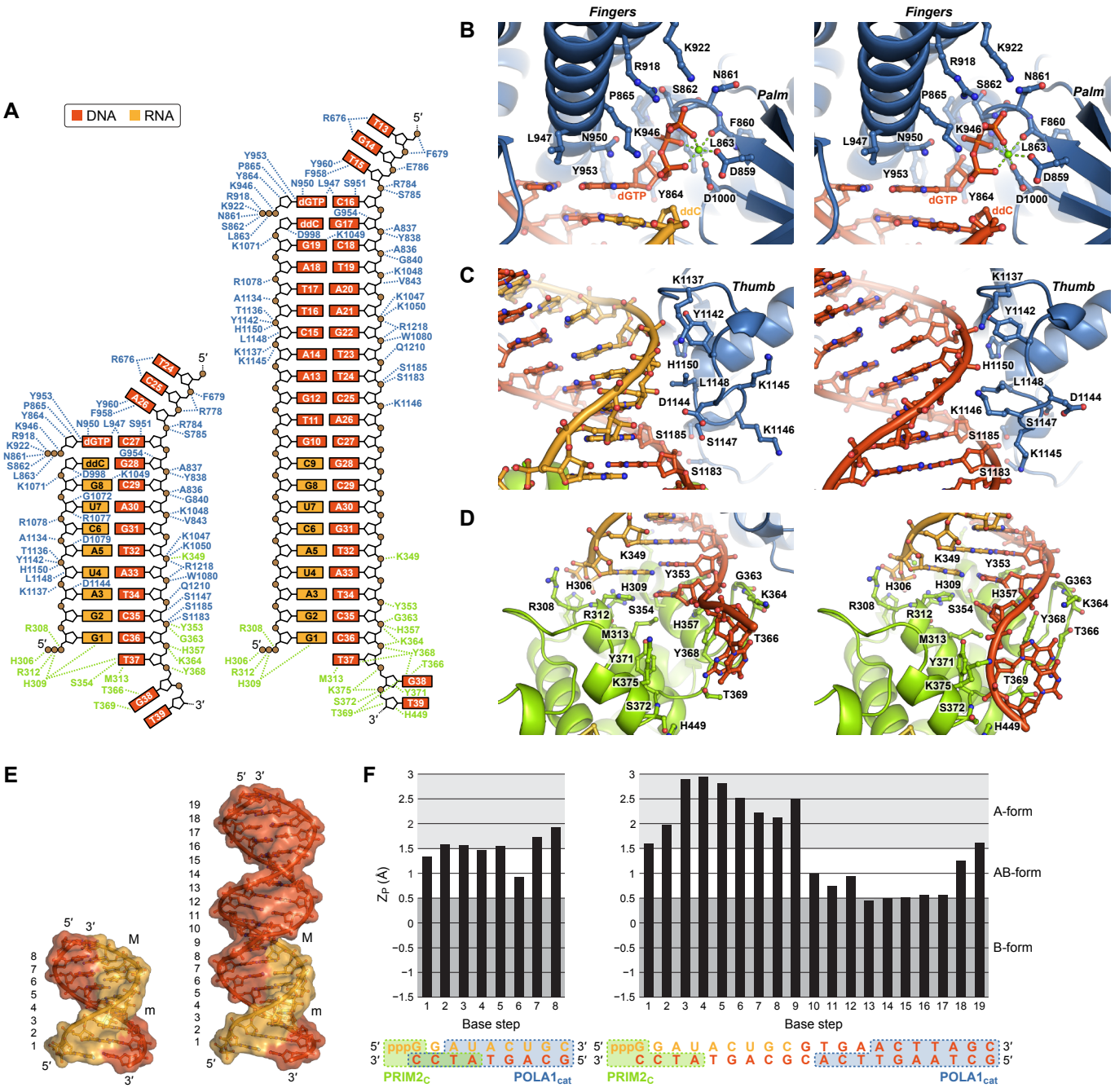

A
DNA
RNA
B
Fingers
Palm
dGTP
ddC
C
Thumb
D
E
M
m
F
A-form
AB-form
B-form
Base step
PRIM2C
POLA1cat

Figure 6

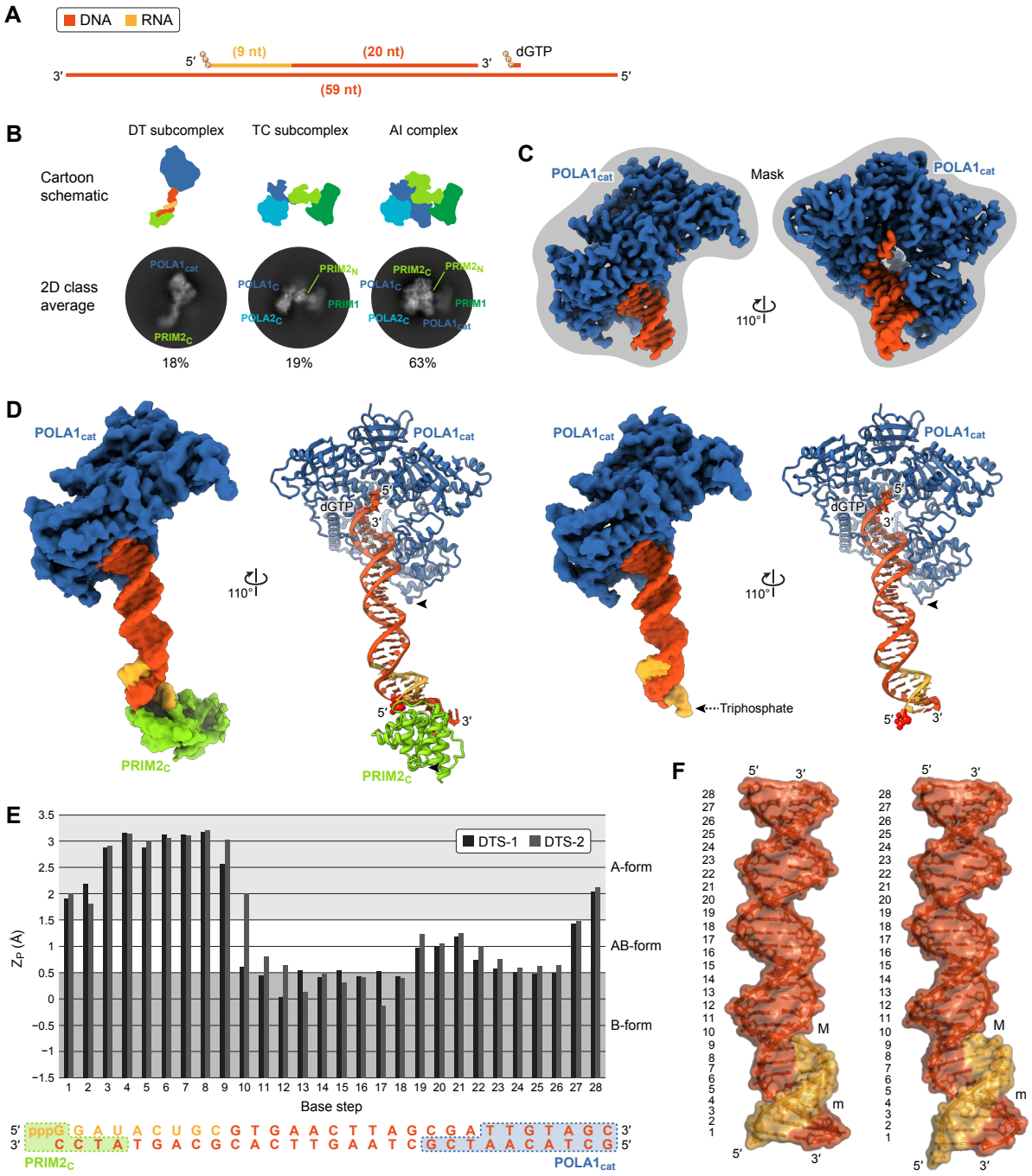
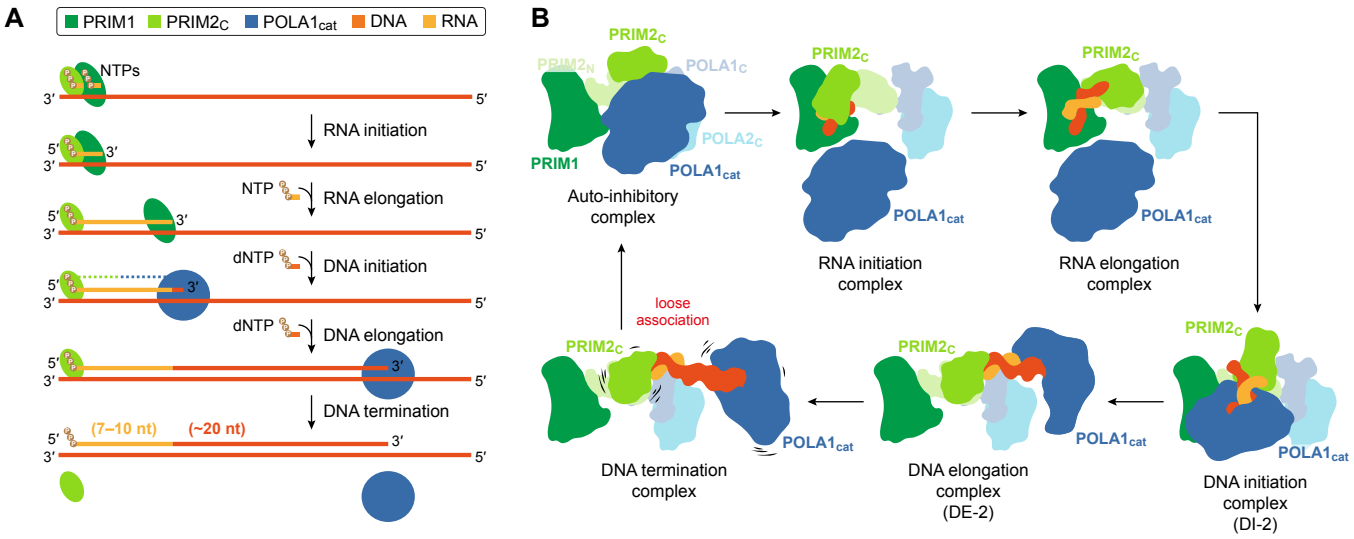

A
DNA
RNA
(9 nt)
(20 nt)
dGTP
(59 nt)
B
DT subcomplex
TC subcomplex
AI complex
Cartoon schematic
2D class average
18%
19%
63%
C
Mask
110°
D
dGTP
Triphosphate
PRIM2C
E
DTS-1
DTS-2
A-form
AB-form
B-form
Base step
F

Figure 7


A
PRIM1
PRIM2C
POLA1cat
DNA
RNA
NTPs
3′
5′
RNA initiation
NTP
RNA elongation
dNTP
DNA initiation
dNTP
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 (DI-2)
DNA elongation complex (DE-2)
loose association
DNA termination complex