Published in final edited form as:
Nature. 2022 June 01; 606(7912): 204–210. doi:10.1038/s41586-022-04759-1.

Fast and efficient DNA replication with purified human proteins

Yasemin Baris[1], **Martin R. G. Taylor**[1], **Valentina Aria**[1], **Joseph T. P. Yeeles**[1,✉]

[1]MRC Laboratory of Molecular Biology, Cambridge, UK

Abstract

Chromosome replication is performed by a complex and intricate ensemble of proteins termed the replisome, where the DNA polymerases Polδ and Polε, DNA polymerase α-primase (Polα) and accessory proteins including AND-1, CLASPIN and TIMELESS–TIPIN (respectively known as Ctf4, Mrc1 and Tof1–Csm3 in *Saccharomyces cerevisiae)* are organized around the CDC45–MCM–GINS (CMG) replicative helicase[1–7]. Because a functional human replisome has not been reconstituted from purified proteins, how these factors contribute to human DNA replication and whether additional proteins are required for optimal DNA synthesis are poorly understood. Here we report the biochemical reconstitution of human replisomes that perform fast and efficient DNA replication using 11 purified human replication factors made from 43 polypeptides. Polε, but not Polδ, is crucial for optimal leading-strand synthesis. Unexpectedly, Polε-mediated leading-strand replication is highly dependent on the sliding-clamp processivity factor PCNA and the alternative clamp loader complex CTF18–RFC. We show how CLASPIN and TIMELESS–TIPIN contribute to replisome progression and demonstrate that, in contrast to the budding yeast replisome[8], AND-1 directly augments leading-strand replication. Moreover, although AND-1 binds to Polα[9,10], the interaction is dispensable for lagging-strand replication, indicating that Polα is functionally recruited via an AND-1-independent mechanism for priming in the human replisome. Collectively, our work reveals how the human replisome achieves fast and efficient leading-strand and lagging-strand DNA replication, and provides a powerful system for future studies of the human replisome and its interactions with other DNA metabolic processes.

At the onset of eukaryotic chromosome replication, the CMG replicative helicase is assembled and activated at DNA replication origins. CMG then unwinds the parental DNA duplex to generate single-stranded DNA templates for leading-strand and lagging-strand replication, and many additional factors associate with the advancing replication forks to form replisomes. In human cells, replisomes progress along chromosomes at approximately

✉**Correspondence and requests for materials** should be addressed to Joseph T. P. Yeeles. jyeeles@mrc-lmb.cam.ac.uk.

Author contributions Y.B. performed all experiments, generated the expression vectors, prepared DNA templates and purified proteins, wrote the methods, and reviewed and edited the manuscript. M.R.G.T. conceptualized the study, acquired funding, performed preliminary DNA replication assays, generated the expression vectors, prepared DNA templates and purified proteins, and reviewed the manuscript. V.A. identified the CMG–Polα interaction. J.T.P.Y. conceptualized and supervised the study, acquired funding, prepared the DNA template and performed protein purification, wrote the original draft of the manuscript, and reviewed and edited the manuscript.

Competing interests The authors declare no competing interests.

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

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Europe PMC Funders Author Manuscripts

1–2 kb min^{-1} (refs.[11,12]). Analysis of DNA fibres from cultured cells has identified various replisome components that are required for optimal replication fork progression including TIMELESS–TIPIN (TIM–TIPIN)[13], CLASPIN[11] and AND-1 (ref. [14]), which all associate directly with CMG[2,5]. However, because the replisome coordinates many additional processes–including parental histone transfer and sister chromatid cohesion–and frequently encounters obstacles such as DNA damage and RNA polymerase, all of which can affect fork progression, how these replisome factors influence DNA synthesis, and whether they do so directly, is unknown. Therefore, to directly address how the human replisome accomplishes fast and efficient DNA replication, we reconstituted functional human replisomes from individual purified proteins.

A minimal leading-strand replisome

To assemble human replisomes, we purified the 11-subunit CMG helicase and 10 additional replication factors including the replicative DNA polymerases Polδ and Polε, the primase Polα, the single-stranded DNA-binding protein RPA, the sliding clamp processivity factor PCNA and its loaders RFC and CTF18–RFC, and the replisome adaptor proteins AND-1, TIM–TIPIN and CLASPIN (Fig. 1a, Extended Data Fig. 1). We began by attempting to reconstitute CMG-dependent leading-strand replication by Polε, because Polε performs the bulk of leading-strand replication in yeast[15,16], and human Polε interacts with CMG in a comparable manner to the *S. cerevisiae* proteins[2,3]. As outlined in Fig. 1b, CMG was first loaded onto a 9.7-kbp forked DNA template containing a primer for leading-strand replication. After a 5-min incubation with Polε, template unwinding and DNA synthesis were initiated by addition of ATP and RPA. Figure 1c shows that Polε synthesized long leading-strand products that were entirely dependent on CMG. However, although some leading strands were extended to the end of the template (9.7 kb) by later time points, the majority were not, even after 120 min (Fig. 1c, lanes 5–8). Therefore, although Polε can perform leading-strand replication when functioning only with CMG, synthesis is slow and non-processive. This finding is notable because *S. cerevisiae* Polε can perform fast and efficient leading-strand replication without PCNA[8,17]. However, PCNA is still required for optimal leading-strand replication by reconstituted budding yeast replisomes[8,17]. We therefore reasoned that human Polε might have a greater dependence on PCNA for leading-strand replication than its yeast counterpart. The addition of PCNA and its loader RFC markedly stimulated leading-strand replication with full-length duplex replication products (Extended Data Fig. 2a, native) and complete 9.7-kb leading strands synthesized within 10 min (Fig. 1d). Full-length duplex products were largely absent when RFC and PCNA were omitted and replication products migrated as a smear of replication intermediates above the position of full-length products (Extended Data Fig. 2a, native), indicating that replisomes had failed to fully unwind the template under these conditions. Quantification of leading-strand replication products from pulse-chase experiments containing RFC and PCNA revealed a maximal synthesis rate of 1.67 kb min^{-1} (Fig. 1e, Extended Data Fig. 2b), demonstrating that a minimal human replisome comprising only CMG, Polε and PCNA can execute leading-strand replication at rates matching those measured in cells. Moreover, in contrast to *S. cerevisiae*[8,17], leading-strand replication by human Polε appears to be highly dependent on PCNA.

Fast leading-strand replication

We next examined how AND-1, TIM–TIPIN and CLASPIN contribute to leading-strand DNA replication. Figure 2a shows that addition of AND-1, T1M–T1P1N and CLASPIN stimulated the rate of leading-strand synthesis, with fully replicated 9.7-kb products synthesized within 4.5 min (lane 9). Given this marked enhancement in rate, we asked whether PCNA was still required for optimal leading-strand synthesis. Omission of PCNA and RFC markedly reduced the rate and efficiency of leading-strand replication despite the presence of AND-1, TIM–TIPIN and CLASPIN, and products were not extended beyond a few kilobases after 5.5 min (Extended Data Fig. 2c). This pronounced dependence on PCNA for robust leading-strand replication led us to consider that the alternative clamp loader complex CTF18–RFC, which interacts with Polε[18–20] and enhances primer extension reactions[21], might be important for leading-strand DNA replication. Despite the presence of RFC in the reaction, addition of CTF18–RFC further accelerated leading-strand replication (Fig. 2b). Quantification of the maximal replication rate in the presence of AND-1, TIM–TIPIN, CLASPIN and CTF18–RFC revealed that replisomes were moving at up to 4.41 kb min^{-1} (Fig. 2c, Extended Data Fig. 2d), which is over twice the mean rate and towards the upper limit of observed fork rate distributions measured in human cells[11–13].

Polε-coupled PCNA loading by CTF18–RFC

To investigate how CTF18–RFC augments leading-strand synthesis, we performed reactions in the presence or absence of PCNA with different combinations of clamp loader complex. Lanes 1–4 in Fig. 3a show that both CTF18–RFC and RFC failed to stimulate replication in reactions that lacked PCNA. In reactions containing PCNA (lanes 5–8), synthesis rates were comparable when CTF18–RFC was the only clamp loader and when both clamp loaders were present. Moreover, higher concentrations of RFC failed to compensate for the omission of CTF18–RFC (Extended Data Fig. 3a). CTF18–RFC also accelerated minimal replisomes that comprised only CMG–Polε, demonstrating that it functions independently of AND-1, TIM–TIPIN and CLASPIN. Here, leading strands were extended at up to 2.34 kb min^{-1} (Fig. 3b, Extended Data Fig. 3b), which is approximately 40% faster than the equivalent reaction with RFC (Fig. 1e). Loading of PCNA by CTF18–RFC is therefore sufficient for fast and robust leading-strand DNA replication, whereas RFC is dispensable.

Work with *S. cerevisiae* and *Homo sapiens* proteins has demonstrated that CTF18–RFC interacts with the Polε catalytic subunit[18–20]. Consistent with these findings, glycerol gradient sedimentation analysis revealed that CTF18–RFC associates with replisomes reconstituted on forked DNA substrates (Fig. 3c, Extended Data Fig. 3c). To investigate whether the interaction between Polε and CTF18–RFC was required for optimal leading-strand replication, we generated mutants in Polε ($Pol\varepsilon^{5A}$) and CTF18–RFC ($CTF18\text{-}RFC^{RAA}$) that disrupt complex formation[19] (Extended Data Fig. 3d, e). In the presence of RFC, mutation of either Polε or CTF18–RFC markedly reduced the length of leading strands in a 3-min reaction (Fig. 3d). When mutants were combined, replication products were similar to those synthesized in the absence of CTF18–RFC. Similar results were obtained in reactions that lacked RFC, except leading-strand products were longer when mutants were combined than when CTF18–RFC was omitted (Extended Data Fig. 3f, lanes

1 and 6), which was probably the result of PCNA loading occurring uncoupled from Polε. These data indicate that the interaction between Pol ε and CTF18–RFC is critical for leading-strand replication in the human replisome.

Because CTF18–RFC supported faster rates of PCNA-dependent replication than RFC (Fig. 3a), we considered that Polε might lose association with PCNA before completing replication. This would necessitate PCNA reloading at the replication fork, which we hypothesized might be performed more efficiently by CTF18–RFC than by RFC due to the interaction between CTF18–RFC and Polε[18,19]. To test this, we first performed a pulse-chase experiment in the presence of RFC on a longer 15.8-kbp template that facilitates visualization of rate differences at fast-moving replication forks. CTF18–RFC was either omitted, included during the pulse or added with the chase. The addition of CTF18–RFC in the chase accelerated almost the entire population of leading strands (Fig. 3e, Extended Data Fig. 4a) and this acceleration was dependent on the interaction between Polε and CTF18–RFC (Extended Data Fig. 4b, c). Next, we purified a truncated CTF18 complex (CTF18-1-8 module) that retains the ability to interact with Polε[18,19](Extended Data Fig. 5a, b) but lacks RFC2–5 and cannot therefore load PCNA. The CTF18-1-8 module failed to accelerate leading-strand replication in a reaction containing RFC and PCNA (Fig. 3f), strongly suggesting that the binding of CTF18–RFC to Polε accelerates replication by modulating PCNA loading, rather than by directly influencing the Polε catalytic domain. To further evaluate the requirement for PCNA loading by CTF18–RFC, we purified a complex with a single point mutation in the Walker A motif of CTF18 (CTF18^{K380E}–RFC) (Extended Data Fig. 5c). This complex still associated with Polε but failed to promote PCNA-dependent synthesis by both Polε and Polδ in primer extension assays, consistent with a pronounced defect in PCNA loading (Extended Data Fig. 5d–f). The addition of CTF18^{K380E}–RFC during the chase phase of a pulse-chase experiment failed to enhance leading-strand replication and, at longer time points, appeared to inhibit replication, perhaps by competing with Polε for the 3’ end of the leading strand (Extended Data Fig. 5g). Collectively, our experiments support a model in which repetitive PCNA loading by CTF18–RFC contributes to fast and efficient leading-strand replication in the human replisome; however, they do not exclude the possibility that CTF18–RFC also influences replication through structural or allosteric effects.

Roles of AND-1, TIM–TIPIN and CLASPIN

Collectively, AND-1, TIM–TIPIN and CLASPIN accelerate replication forks by approximately 1.9-fold (compare rates in Figs. 2c and 3b). To characterize the individual contributions of AND-1, TIM–TIPIN and CLASPIN, we performed the drop-out experiment in Fig. 4a. The ionic strength of the reaction buffer was increased (Extended Data Fig. 6a) because budding yeast replisomes have a greater dependence on Tof1–Csm3 and Mrc1–orthologues of TIM–TIPIN and CLASPIN, respectively–under these conditions[8]. Omission of any of the three proteins reduced the length of leading-strand products compared to the complete reaction, with omission of CLASPIN having the greatest effect (Fig. 4a, Extended Data Fig. 6b). Hence, AND-1, TIM–TIPIN and CLASPIN are all required for optimal leading-strand DNA replication by reconstituted human replisomes. Figure 4b and Extended Data Fig. 6c, d show that AND-1 and CLASPIN stimulated leading-strand

synthesis independently of other adaptor proteins. By contrast, TIM–TIPIN failed to enhance replication in the absence of AND-1 and CLASPIN, suggesting that it functions with one or both factors. Accordingly, addition of TIM–TIPIN to a reaction containing CLASPIN but lacking AND-1 stimulated replication (Fig. 4c, Extended Data Fig. 6e), whereas TIM–TIPIN did not stimulate a reaction containing AND-1 but lacking CLASPIN (Extended Data Fig. 6f). Furthermore, increasing the concentration of CLASPIN in reactions lacking TIM–TIPIN failed to accelerate replication (Extended Data Fig. 6g). These results show that acceleration of leading-strand replication by TIM–TIPIN depends on CLASPIN. They also show that, although CLASPIN can modestly influence replication without TIM–TIPIN, CLASPIN requires TIM–TIPIN for optimal activity.

Interaction between CLASPIN and TIM–TIPIN

How CLASPIN and TIM–TIPIN cooperate to enhance replication is unknown. Because DNA synthesis is stimulated by CLASPIN without TIM–TIPIN, but not vice versa (Fig. 4b, c, Extended Data Fig. 6c, e), we hypothesized that TIM–TIPIN acts by augmenting CLASPIN function. Although CLASPIN has been predicted to be largely disordered[22], a recent structure of a human replisome revealed an interface between the leading edge of the TIMELESS α-solenoid and CLASPIN residues 284–319 (ref. [2]) (Fig. 4d, Extended Data Fig. 7a), suggesting that TIM–TIPIN might promote CLASPIN function by facilitating its positioning at the replication fork. To test this, we purified two N-terminal truncation mutants: CLASPIN$^{\Delta 2–283}$ and CLASPIN$^{\Delta 2–319}$ (Fig. 4d, Extended Data Fig. 7b). In the absence of TIM–TIPIN, both mutants retained the ability to modestly stimulate leading-strand replication (Fig. 4e, lanes 1–4, Extended Data Fig. 7c). The addition of TIM–TIPIN stimulated replication with CLASPIN$^{\Delta 2–283}$ to a comparable extent to full-length CLASPIN (Fig. 4e, Extended Data Fig. 7c, d), demonstrating that the N-terminal 283 amino acids of CLASPIN are dispensable for rate enhancement. By contrast, acceleration of leading-strand synthesis by TIM–TIPIN was markedly reduced with CLASPIN$^{\Delta 2–319}$ (Fig. 4e, Extended Data Fig. 7c, d), although not completely abolished, probably due to a small interface between CLASPIN and the N-terminal region of the TIMELESS α-solenoid[2] (Fig. 4d, Extended Data Fig. 7a). These data reveal that the interaction between CLASPIN (amino acids 284–319) and TIMELESS is crucial for leading-strand replication, indicating that TIM–TIPIN controls replication fork rate primarily by coordinating CLASPIN in the replisome.

AND-1 and leading-strand synthesis

Our observation that AND-1 directly stimulates leading-strand replication (Fig. 4a, b) was notable because its budding yeast orthologue, Ctf4, does not display this behaviour[8]. AND-1 has a C-terminal extension containing an HMG box and a DNA-binding domain[9] (Fig. 4f), which are both absent from *S. cerevisiae* Ctf4. Moreover, removal of the AND-1 HMG box slowed replication forks in DT40 cells[14]. To identify the region of human AND-1 that is responsible for stimulating leading-strand replication, we purified three C-terminal truncation mutants (Fig. 4f, Extended Data Fig. 7e). Figure 4g shows that removal of the AND-1 HMG box (AND-1$^{\Delta 1017}$) or the HMG box and an additional 30 amino acids (AND-1$^{\Delta 987}$) did not compromise the ability of AND-1 to promote leading-strand

replication (see also Extended Data Fig. 7f). However, removal of the AND-1 DNA-binding domain (AND-1$^{\Delta 895}$) impaired leading-strand synthesis (Fig. 4g, compare lanes 1 and 5). Thus, AND-1 directly enhances leading-strand replication, probably via binding DNA, independently of its HMG box.

Lagging-strand synthesis

As illustrated in Fig. 5a, discontinuous lagging-strand replication will generate labelled lagging-strand daughter molecules (full length) and distinct replication intermediates (intermediates). Accordingly, addition of Polα to replication reactions also supplemented with Polδ–the main lagging-strand polymerase in yeast[15]–altered the profile of replication intermediates (Fig. 5b, native, compare lanes 3 and 8). Inspec-tion of the denaturing gel revealed the presence of a population of Polα-dependent products displaying a broad length distribution from approximately 0.15 to 2 kb (Fig. 5b, denaturing, lagging strands). Omission of Polδ reduced the length of these products to approximately 0.6 kb or less (Extended Data Fig. 8a, b), as did omission of RFC, despite the presence of CTF18–RFC (Fig. 5c). Two-dimensional electrophoresis confirmed that the Polα-dependent products, synthesized in the presence or absence of Polδ, were constituents of both replication intermediates and full-length daughter molecules (Extended Data Fig. 8c–e). These observations are indicative of replisomes performing lagging-strand replication.

To determine the proportion of replication forks performing concomitant leading-strand and lagging-strand replication, we performed reactions on a template containing a site-specific leading-strand DNA lesion (cyclobutane pyrimidine dimer) placed approximately 3.5 kb from the forked end of the template (Extended Data Fig. 9a). Inhibition of leading-strand synthesis at the cyclobutane pyrimidine dimer generates stalled replication forks[23] (Extended Data Fig. 9b) and their migration should be influenced by lagging-strand replication (Extended Data Fig. 9a). In the absence of Polα, stalled replication forks migrated only a short distance above full-length products (Extended Data Fig. 9c, d). Upon addition of Polα, migration of the entire population of stalled forks was retarded (Extended Data Fig. 9c, e), demonstrating that leading-strand and lagging-strand replication was occurring concomitantly at most, if not all, replication forks. Collectively, these data demonstrate the reconstitution of robust lagging-strand replication, where Polα synthesizes primers that are extended by Polδ in conjunction with PCNA that is loaded by RFC.

An AND-1-independent priming mechanism

AND-1 interacts with Polα primarily via its HMG box[9], providing an obvious mechanism for the recruitment of Polα to the replisome for priming. However, although AND-1 is essential for proliferation of DT40 cells, bulk replication is still completed upon depletion of AND-1, and an AND-1 truncation lacking the HMG box supported cell proliferation, albeit at reduced rates[14]. These observations are inconsistent with an essential function for AND-1 in nascent-strand priming. To evaluate the role of AND-1 in nascent-strand priming, we compared replication in the presence and absence of AND-1 and with AND-1$^{\Delta 1017}$, in which the HMG box but not the DNA-binding domain is deleted (Fig. 4f). In the absence of AND-1, lagging-strand products were still synthesized. However, there was a subtle

increase in their length, indicative of less frequent priming (Fig. 5d, compare lanes 4 and 5). Consistent with the role of AND-1 in leading-strand replication (Fig. 4a, b, e), the intensity of leading-strand products was reduced, in both the presence and the absence of Polα. AND-1$^{\Delta 1017}$ supported comparable levels of leading-strand replication to wild-type AND-1. However, as seen with omission of AND-1, there was a subtle increase in the length of lagging-strand products, although robust synthesis still occurred (Fig. 5d, compare lanes 5 and 6, Extended Data Fig. 10a).

We reasoned that AND-1-mediated Polα recruitment for priming might become apparent when Polα was limiting. To test this, we performed a Polα titration in the presence of wild-type AND-1 or AND-1$^{\Delta 1017}$. Figure 5e and Extended Data Fig. 10b show that, although 5 nM Polα supported robust lagging-strand synthesis with both AND-1 proteins, lagging strands were significantly longer in the reaction containing AND-1$^{\Delta 1017}$, whereas leading-strand replication was unaffected. Conversely, at 40 nM Polα, lagging strands were almost indistinguishable. To further define the requirements for lagging-strand priming, we asked whether Polα could function with a minimal replisome composed only of CMG, Polε, CTF18–RFC, RFC and PCNA. Again, robust Polα-dependent lagging-strand synthesis was observed (Extended Data Fig. 10c), indicating that Polα might be recruited for priming directly by CMG. Consistent with this idea, Fig. 5f and Extended Data Fig. 10d show that Polα co-migrated with CMG in a glycerol gradient performed in the absence of DNA, demonstrating a direct interaction between the two complexes. Collectively, these data reveal that the human replisome targets Polα to the lagging strand for priming via an AND-1-independent mechanism, probably involving direct binding to the CMG helicase. In addition, they show that, although it is not essential, the interaction between AND-1 and Polα can modulate priming, potentially by increasing the local concentration of Polα at replication forks.

Discussion

By reconstituting human replisomes that recapitulate cellular rates of DNA replication, we have identified the minimal requirements for fast and efficient leading-strand and lagging-strand DNA synthesis. Moreover, our reconstituted system has enabled us to study human DNA replication without the additional complexity of the myriad of replication-coupled process, many of which involve core replisome proteins. This has revealed numerous important insights into some of the most basic and fundamental aspects of human DNA replication, including the discovery that Polε-coupled PCNA loading by CTF18–RFC contributes directly to leading-strand synthesis. This, together with the dispensability of Polδ for optimal leading-strand synthesis, provides strong evidence that Polε is the principal leading-strand polymerase in human cells. Conversely, on the lagging strand, Polδ is more active than Polε at extending primers made by Polα. This clear division of labour between Polε and Polδ has also been observed in yeast[8,15,16,24], indicating that it is a conserved feature of eukaryotic DNA replication. Moreover, our results also reveal a clear division of labour for PCNA loading: RFC is essential for Polδ function on the lagging strand, and CTF18–RFC is necessary and sufficient for maximal rates of leading-strand replication by Polε. The requirement for Polε-coupled CTF18-dependent PCNA loading for maximal leading-strand replication rates indicates that human Polε frequently loses association with

PCNA. Although this might seem counterproductive, we propose that it functions to ensure that the leading-strand daughter is sufficiently populated with PCNA for downstream processes including sister chromatid cohesion and checkpoint activation, both of which involve CTF18–RFC[25–29].

Our discovery that TIM–TIPIN, CLASPIN and AND-1 all modulate replication fork rate in reconstituted reactions lacking the replication-coupled pathways in which they participate, indicates that they accelerate replisomes by directly influencing template unwinding and/or DNA synthesis. These properties of TIM–TIPIN and CLASPIN, and their mechanism (or mechanisms) of action, including our discovery that TIM–TIPIN promotes CLASPIN activity through direct binding, are probably conserved features of the eukaryotic replisome given the notably similar structural organization of the core yeast[1] and human[2] replisomes. By contrast, the capacity of AND-1 to enhance leading-strand replication appears to be specific to eukaryotic species where AND-1 contains a C-terminal DNA-binding domain[9]. How DNA binding by AND-1 promotes fork progression is an interesting subject for future exploration. Although TIM–TIPIN, CLASPIN and AND-1 are required for maximal fork rates, replisomes lacking all three proteins still progress at over 2 kb min^{-1}. We suggest that the presence of three adaptor proteins that can each modify replication rate, together with the ability of replisomes to sustain reasonably fast DNA replication without these proteins, enables replisome progression rates to be fine-tuned in response to environmental cues; for example, in response to redox changes, TIM–TIPIN is evicted from the replisome to slow fork progression[13]. On the lagging strand, our experiments show that Polα is functionally recruited to the human replisome for priming independently of its interactions with AND-1. Because we observed a direct interaction between CMG and Polα and robust priming by replisomes composed of CMG, Polε, RFC, CTF18–RFC and PCNA, and because yeast Polα can synthesize primers when functioning only with CMG[30], we consider it likely that priming is facilitated by direct interactions between Polα and CMG and this is an important area for future investigation.

The reconstitution of functional human replisomes from purified proteins represents an important milestone in the study of eukaryotic DNA replication that has taken many years to accomplish. Using these reconstituted replisomes, we can now explore, in molecular and atomic detail, how the human replisome navigates chromatinized templates, responds to obstacles such as DNA damage and protein roadblocks, and coordinates important replication-coupled nuclear processes.

Methods

Plasmid construction

Sequences for the expression of Polα (composed of POLA1, POLA2, PRIM1 and PRIM2) and Polδ (composed of POLD1, POLD2, POLD3 and POLD4) were codon optimized for overexpression in insect cells and synthesized by GeneArt Gene synthesis. For CTF18–RFC, CTF18, CTF8 and DCC1 were codon optimized for expression in insect cells and synthesized by Epoch Life Science Gene Synthesis. POLA1 was encoded with an N-terminal twin strep tag with a 3C cleavage site. CTF18 and POLD4 contained an N-terminal twin strep tag along with a TEV cleavage site (see Supplementary Table 1 for the affinity

tag sequences). The codon-optimized sequences were cloned into the pACEBac1 vector (see Supplementary Table 2 for plasmid details). For expression of CTF18–RFC, genes encoding RFC2, RFC3, RFC4 and RFC5 were amplified by PCR and cloned into pBIG1a.

Protein expression

For CTF18–RFC expression, separate viruses expressing RFC2–5, CTF18, CTF8 and DCC1 were amplified in SF9 cells and used to co-infect 31 of Hi5 cells at a density of 1.5×10^6 cells per ml. Viruses expressing individual subunits of Polδ and Polα were used to co-infect 2 l of Hi5 cells at a density of 1×10^6 cells per ml. Cells were collected at approximately 60 h after infection.

Protein purification

CMG, RFC, PCNA, TIM–TIPIN, CLASPIN and AND-1 were expressed and purified as previously described[2]. RPA was purified as previously described[2] following expression in *Escherichia coli* Rosetta 2(DE3) (Merck) from the pLK966 plasmid[31,32]. AND-1 and CLASPIN mutants were purified as reported for wild-type proteins[2]. Polε was purified as previously described[2] and using the method described in detail below. No differences were observed between the different preparations of Polε.

CTF18–RFC purification

Cell pellet was resuspended in lysis buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10% glycerol, 0.5 mM TCEP, 0.02% NP-40-S (Nonidet P-40 substitute (Roche)) and protease inhibitors (cOmplete, EDTA-free, Roche, one tablet per 50 ml buffer). Cells were lysed by dounce homogenization, and insoluble material was removed by centrifugation (235,000*g* at 4 °C for 45 min). Strep-Tactin XT superflow high-capacity resin (iba; 1.5 ml) was added to the lysate and incubated for 45 min at 4 °C. Resin was collected in a 20-ml column (Bio-Rad) and washed with 75 ml lysis buffer. Protein was eluted with 10 column volumes (CV) (1.5 ml each fraction) lysis buffer + 30 mM biotin. Fractions were pooled and applied to 1 ml MonoQ column (cytiva) equilibrated in 25 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.01% NP-40-S. The protein was eluted with a 30 CV gradient from 100 to 1,000 mM NaCl. Peak fractions were pooled, concentrated to approximately 500 µl in an Amicon Ultra-15 30 kDa MWCO concentrator and applied to a Superdex200 Increase 10/300 gel-filtration column (cytiva) equilibrated in 25 mM Tris-HCl (pH 7.2), 10% glycerol, 0.005% TWEEN 20, 0.5 mM TCEP and 150 mM NaCl. Peak fractions were pooled, frozen in liquid nitrogen and stored at –80 °C. Approximately 0.5 mg total protein was obtained from 3 l of cells. CTF18–RFC mutants and the CTF18-1-8 module were purified as described for the wild-type protein.

Polε purification

Wild-type and mutant Polε were purified using the following protocol that is based on a previously published method for yeast Polε[33]. Cells from a 1.5-l culture, grown as previously described[2], were resuspended in 50 ml lysis buffer (100 mM Tris-OAc (pH 7.5), 100 mM NaOAc, 1 mM EDTA and 0.5 mM TCEP) + protease inhibitors (cOmplete, EDTA-free, Roche; one tablet per 50 ml buffer). Cells were lysed by dounce homogenization and

insoluble material was cleared by centrifugation (235,000*g* at 4 °C for 45 min). Ammonium sulfate was added to the supernatant to a concentration of 150 mM from a 3 M stock. Then, polyethylen-imine (03880, Sigma) was added to 0.4% by slow dropwise addition of a 9% stock (pH 8). Following a 10-min incubation with stirring at 4 °C, insoluble material was removed by centrifugation (48,000*g* at 4 °C for 20 min). Solid ammonium sulfate (0.28 g ml^{-1}) was slowly added to the supernatant followed by incubation overnight with stirring at 4 °C. Precipitated protein was collected by centrifugation (48,000*g* at 4 °C for 20 min) and resuspended in 40 ml 25 mM HEPES-NaOH (pH 7.6), 10% glycerol, 1 mM EDTA, 0.005% NP-40-S, 1 mM EDTA and 50 mM NaOAC (buffer B + 50 mM NaOAC). Solid ammonium sulfate (0.16 g ml^{-1}) was slowly added followed by incubation for 10 min with stirring at 4 °C. Insoluble material was removed by centrifugation (48,000*g* at 4 °C for 20 min) and a further 0.11 g ml^{-1} solid ammonium sulfate was added to the sample. Following a 10-min incubation at 4 °C with stirring, precipitated protein was pelleted by centrifugation (48,000*g* at 4 °C for 20 min) and resuspended in approximately 50 ml buffer B + 50 mM NaOAC. The sample was applied to 3 × 1 ml HiTrap SP FF columns (cytiva) arranged in series and equilibrated in buffer B + 50 mM NaOAC. The columns were washed with buffer B + 200 mM NaOAC and proteins were eluted in buffer B + 750 mM NaOAC. The eluate was diluted threefold in buffer B + 50 mM NaOAC and applied to a 5-ml HiTrap Q FF column (cytiva) equilibrated in buffer B + 50 mM NaOAC. Bound proteins were eluted with a 20 CV gradient to buffer B + 900 mM NaOAC. Peak fractions were dialysed overnight against buffer B + 75 mM NaOAC and applied to a 5-ml HiTrap heparin column (cytiva) equilibrated in buffer B + 100 mM NaOAC. Protein was eluted with a 20 CV gradient to buffer B + 1.2 M NaOAC. Peak fractions were pooled, diluted 2.5-fold in buffer B + 50 mM NaOAC and applied to a 1-ml MonoQ column (cytiva) equilibrated in buffer B + 300 mM NaOAc. Bound proteins were eluted with a 30 CV gradient to buffer B + 1.2 M NaOAc. Peak fractions were concentrated to approximately 600 µl and applied to a Superdex 200 increase 10/300 column (cytiva) equilibrated in 25 mM HEPES-NaOH (pH7.6), 10% glycerol, 1 mM EDTA, 0.005% NP-40-S, 1 mM EDTA, 1 mM DTT and 500 mM NaOAC. Peak fractions were pooled, frozen in liquid nitrogen and stored at –80 °C. Protein yield was approximately 1 mg from 1 l of cell culture. Polε mutants were purified as described for the wild-type protein.

Polα purification

Cell pellet was resuspended in lysis buffer (25 mM Tris-HCl (pH 7.2), 300 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.02% NP-40-S) + protease inhibitors (cOmplete, EDTA-free, Roche; one tablet per 50 ml buffer). Cells were lysed by dounce homogenization and insoluble material was removed by centrifugation (235,000*g* at 4 °C for 45 min). Strep-Tactin XT superflow high-capacity resin (iba; 1 ml) was added to the lysate and incubated for 30 min at 4 °C. Resin was collected in 20-ml column and was washed with 50 ml lysis buffer. Protein was eluted with 10 CV (1 ml each fraction) buffer (25 mM Tris-HCl (pH 7.2), 100 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.02% NP-40-S) + 30 mM biotin. Fractions were pooled and applied to a 1-ml MonoQ column (cytiva) equilibrated in 25 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.01% NP-40-S. The protein was eluted with a 20 CV gradient from 100 to 1,000 mM NaCl. Peak fractions were pooled, concentrated to approximately 500 µl in an Amicon Ultra-15 30 kDa MWCO

concentrator and applied to a Superdex 200 Increase 10/300 gel-filtration column (cytiva) equilibrated in 25 mM Tris-HCl (pH 7.2), 10% glycerol, 0.005% TWEEN 20, 0.5 mM TCEP and 150 mM NaCl. Peak fractions were pooled, frozen in liquid nitrogen and stored at –80°C. The total amount of protein purified from 2 l cells was approximately 0.3 mg.

Polδ purification

Cell pellet was resuspended in lysis buffer (50 mM Tris-HCl (pH 8), 300 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.02% NP-40-S) + protease inhibitors (cOmplete, EDTA-free, Roche; one tablet per 50 ml buffer). Cells were lysed by dounce homogenization and insoluble material was removed by centrifugation (235,000*g* at 4 °C for 45 min). Strep-Tactin XT superflow high-capacity resin (iba; 1.5 ml) was added to the lysate and incubated for 30 min at 4°C. Resin was collected in a 20-ml column and was washed with 75 ml lysis buffer. Resin was further washed with 15 ml lysis buffer + 5 mM $Mg(OAc)_2$ + 0.5 mM ATP followed by 15 ml wash without ATP and $Mg(OAc)_2$. Protein was eluted with 10 CV (1.5 ml each fraction) buffer (50 mM Tris-HCl (pH 8), 100 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.02% NP-40-S) + 30 mM biotin. Fractions were pooled and applied to a 1-ml MonoQ column (cytiva) equilibrated in 50 mM Tris-HCl (pH 8), 100 mM NaCl, 10% glycerol, 0.5 mM TCEP and 0.01% NP-40-S. The protein was eluted with a 30 CV gradient from 100 to 1,000 mM NaCl. Peak fractions were pooled, concentrated to approximately 500 µl in an Amicon Ultra-15 30 kDa MWCO concentrator and applied to a Superdex200 Increase 10/300 gel-filtration column (cytiva) equilibrated in 25 mM Tris-HCl (pH 7.2), 10% glycerol, 0.005% TWEEN 20, 0.5 mM TCEP and 150 mM NaCl. Peak fractions were pooled, frozen in liquid nitrogen and stored at –80 °C. Approximately 0.06 mg of protein was obtained from 2 l cells.

Generation of vJY170

To generate a 15.8-kbp plasmid for DNA template preparation, an approximately 5.7-kbp fragment was excised from vJY22 (ref. [8]) with restriction enzymes NarI and SpeI (New England Biolabs) and cloned into vVA20 (ref. [17]) linearized with the same enzymes to give plasmid vJY170.

Fork annealing for linear fork DNA template preparation

The sequence of the replication fork was adapted from a previously published study[24]. MT096 (10 µM; 5′phos/GCTATGTGGTAGGAAGT GAGAATTGGAGAGTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTGAGGAAAGAATGTTGGTGAGGGTTGGGAAGTGGAAGGA TGGGCTCGAGAGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT) was mixed with JY197 (20 µM; 5′-TTTTTTTTTTTTTTTTTTTCA CACTCTCCAATTCTCACTTCCTACCACAT) andJY195 (20 µM; 5′CCTCTC GAGCCCATCCTTCCACTTCCCAACCCTCACC) in a 25 µl reaction containing 200 mM NaCl and 25 mM EDTA. The mixture was heated to 75 °C and cooled to room temperature. Annealing leaves a 5′-phosphorylated GCT overhang for ligation to one end of the SapI-linearized plasmid (ZN3 or vJY170).

Preparation of linear forked DNA template

CsCl gradient purified ZN3 (50 μg) (ref. [23]) or vJY170 plasmid was digested with 15 μl SapI (R0569, New England Biolabs) in a 200 μl reaction containing 1X CutSmart Buffer (B7204, New England Biolabs) for 4 h at 37 °C. The DNA was incubated with SDS (0.2%), 30 mM EDTA and 2 μl proteinase K (20 mg ml^{-1}; New England Biolabs) for 30 min at 37 °C and then extracted with phenol:chloroform:isoamyl alcohol 25:24:1 saturated with TE (10 mM Tris (pH 8), 1 mM EDTA) (P2069, Sigma-Aldrich). The aqueous phase was collected, and DNA precipitated with 0.3 M NaCl + 2.8 volumes ice-cold 100% ethanol on dry ice. DNA was washed with 70% ethanol, centrifuged and the pellet was air-dried and resuspended in 50 μl TE (10 mM Tris-HCL (pH 7.2) and 1 mM EDTA). To ligate the fork, 25-fold molar excess of pre-annealed fork was mixed with 15 nM SapI-linearized DNA and incubated in a 300 μl reaction containing 60,000 units T4 DNA ligase (New England Biolabs), 5 mM $MgCl_2$, 30 μl T4 ligase buffer at 16 °C for 17 h. The ligation mix was then incubated with 30 mM EDTA, 0.3% SDS and 3.5 μl proteinase K (20 mg ml^{-1}; New England Biolabs) for 30 min at 37 °C. Excess unligated fork was removed by gel filtration over a 0.7 × 50 cm Sepharose-4B (Sigma) column equilibrated in 5 mM Tris-HCL (pH 7.5), 0.1 mM EDTA. Peak fractions containing the linear forked DNA were pooled, frozen, lyophilized in a vacuum concentrator and precipitated by adding 0.3 M NaCl + 2.8 volumes ice-cold 100% ethanol on dry ice. DNA was pelleted and washed with 70% ethanol. The pellet was air dried and resuspended in 10 mM Tris-HCL (pH 7.2) and 1 mM EDTA to approximately 350–500 ng $μl^{-1}$. The site-specific cyclobutane pyrimidine dimer was introduced into the ZN3 plasmid as previously described[23]. Linear forked DNA from this plasmid was prepared as described above for undamaged templates. To reduce template dimers, after fork ligation and purification, the cyclobutane pyrimidine dimer template was digested with Pci1–which cuts approximately 100 bp from the distal end of the template–and re-purified over a Sepharose-4B column as described above.

Standard DNA replication assays

Replication reactions were conducted at 37 °C in a replication buffer consisting of 25 mM HEPES-KOH (pH 7.6), 100 mM potassium glutamate, 0.01% NP-40-S, 1 mM DTT, 10 mM $Mg(OAc)_2$ and 0.1 $mgml^{-1}$ BSA. Unless stated in the figure legends, the DNA, protein and nucleotide concentrations in the final replication reaction were: 1.2 nM DNA, 25 nM CMG, 20 nM Polε, 20 nM RFC, 20 nM PCNA, 20 nM AND-1, 10 nM Polα, 100 nM RPA, 20 nM CLASPIN, 20 nM TIM–TIPIN, 20 nM CTF18–RFC, 4 mM ATP, 30 μM dC/dT/dG/dATP, 200 μM C/G/UTP and 33 nM α-[^{32}P]-dCTP. Reactions were set up as follows: CMG (50 nM) was pre-incubated with the DNA (2.4 nM) for 10 min in replication buffer supplemented with 0.1 mM AMP-PNP. The reaction was diluted twofold by the addition of replication buffer containing 60 μM dA/dCTP and, where indicated in the figure legends, PCNA, RFC, Polε, Polα, CLASPIN, TIM–TIPIN, AND-1, CTF18–RFC and Polδ. Replication was initiated by addition of a 10X solution containing ATP, dTTP, dGTP, GTP, CTP, UTP, α-[^{32}P]-dCTP and RPA. In reactions containing Polδ, RFC was added when replication was initiated. Where indicated in the figure legends, the final concentration of potassium glutamate in reactions was increased to 250 mM when replication was initiated. Reactions were quenched by addition of 50 mM EDTA. Unincorporated α-[^{32}P]-dCTP was removed with illustra MicroSpin G-50 columns (GE Healthcare). Reactions were run on

0.7% alkaline agarose gels in 30 mM NaOH and 2 mM EDTA for 16 h at 24 V. Gels were fixed with 5% cold trichloroacetic acid, dried onto Whatman paper and imaged using an Amersham Typhoon laser-scanner or autoradiographed with Amersham Hyperfilm-MP (GE Healthcare).

For native agarose gel analysis, SDS (0.1%) and proteinase K (0.2 mg ml^{-1}; New England Biolabs) were added to quenched reactions and incubated for 30 min at 37 °C. DNA was extracted with phenol:chloroform:isoamyl alcohol 25:24:1 saturated with TE (P2069, Sigma-Aldrich) and unincorporated nucleotides were removed with illustra MicroSpin G-50 columns (GE Healthcare). Samples were analysed in 0.8% agarose gels run at 20 V overnight in modified 1X TAE (50 mM Tris-HCl (pH 7.9), 40 mM NaOAc and 1 mM EDTA (pH 8.0)). Gels were dried onto Whatman paper and imaged as described for alkaline gels.

For two-dimensional gels, the sample was split, and approximately 5% was loaded in one lane and the remaining approximately 95% was loaded in another lane on the same native gel. The latter lane was excised from the gel and laid horizontally along the top of a 0.7% denaturing gel and run at 26 V for 17 h. The gels were processed as described above.

Pulse-chase experiments

Pulse-chase experiments were performed using the conditions for standard replication reactions, except that the concentration of dCTP in the pulse was reduced to 3 μM. The concentrations of dGTP, dCTP, dATP and dTTP were increased to 600 μM in the chase.

Replisome assembly and glycerol gradient sedimentation

Replisome assembly reactions were set up to yield a final volume of 10 μl containing: 150 nM CMG with a 1.5-fold molar excess of fork DNA[1], TIM–TIPIN, CLASPIN, AND-1, Polε, CTF18–RFC in reconstitution buffer (25 mM HEPES-NaOH (pH 7.6), 150 mM NaOAc, 0.5 mM TCEP, 500 μM AMP-PNP and 10 mM $Mg(OAc)_2$). CMG (or CMG storage buffer in the –CMG reaction) was first incubated with fork DNA[1] for 15 min on ice.

Next, the remaining proteins were added and the final volume was adjusted to 10 μl. The reaction was loaded onto a 230 μl 10–30% glycerol gradient in a buffer containing 40 mM HEPES-NaOH (pH 7.5), 150 mM NaOAc, 0.5 mM TCEP, 500 μM AMP-PNP and 3 mM $Mg(OAc)_2$. Samples were separated by centrifugation (Beckman TLS-55 rotor; 200,000*g* at 4 °C for 1 h). Fractions (10 μl) were collected manually and analysed by 4–12% SDS–PAGE with silver staining (Invitrogen). For the Polα–CMG gradient, 100 nM CMG was mixed with 80 nM Polα in reconstitution buffer for a final volume of 10 μl and incubated on ice for 30 min. The reaction was loaded onto the gradient and fractions were analysed as described above.

Pulldown experiments

All reaction steps were performed in a modified replication buffer consisting of 25 mM HEPES-KOH (pH 7.6), 150 mM potassium glutamate, 0.05% NP-40-S, 1 mM DTT and 10 mM $Mg(OAc)_2$. Of bait protein (wild-type CTF18–RFC complex or its variants), 5 pmol

was incubated with 1 µl MagStrep type3 XT beads (iba) in 100 µl buffer for 1 h at 4°C with gentle rotation. The beads were washed three times with 100 µl buffer and 10 pmol prey protein (wild-type or mutant Polε) in 100 µl buffer was added to the washed beads. Following 1-h incubation at 4°C with gentle rotation, the beads were washed with 3 × 100 µl buffer and bound proteins were eluted with 20 µl buffer + 50 mM biotin. Eluted proteins were analysed by 4–12% SDS–PAGE with silver staining (Thermo Fisher).

Primer extension assay

Oligonucleotide (500 nM; sequence: 5′-GAATAATGGAAGGGTT AGAACCTACCAT) was annealed to 50 nM M13mp18 single-stranded DNA (New England Biolabs) in 10 mM Tris-HCl (pH 7.6), 100 mM NaCl and 5 mM EDTA. The mixture was heated to 75 °C and gradually cooled to room temperature. Unannealed oligonucleotide was removed using an S400 column (cytiva). The reaction was performed at 37 °C in a buffer containing 25 mM HEPES-KOH (pH 7.6), 100 mM potassium glutamate, 0.01% NP-40-S, 1 mM DTT, 10 mM $Mg(OAc)_2$, 0.1 mg ml^{-1} BSA, 5 mM ATP, 200 µM CTP, GTP, UTP, 30 µM dATP, dCTP, dGTP, dTTP, and 33 nM α-[^{32}P]-dCTP. Primed template (1 nM) was first incubated with 400 nM RPA, 20 nM PCNA and 20 nM each clamp loader complex where indicated in the figure legends for 10 min. The reaction was initiated by the addition of 20 nM Polδ or Polε where indicated in the figure legends. Reactions were stopped at the indicated time point by addition of EDTA to 50 mM. Samples were processed as described for standard DNA replication assays.

Rate measurements

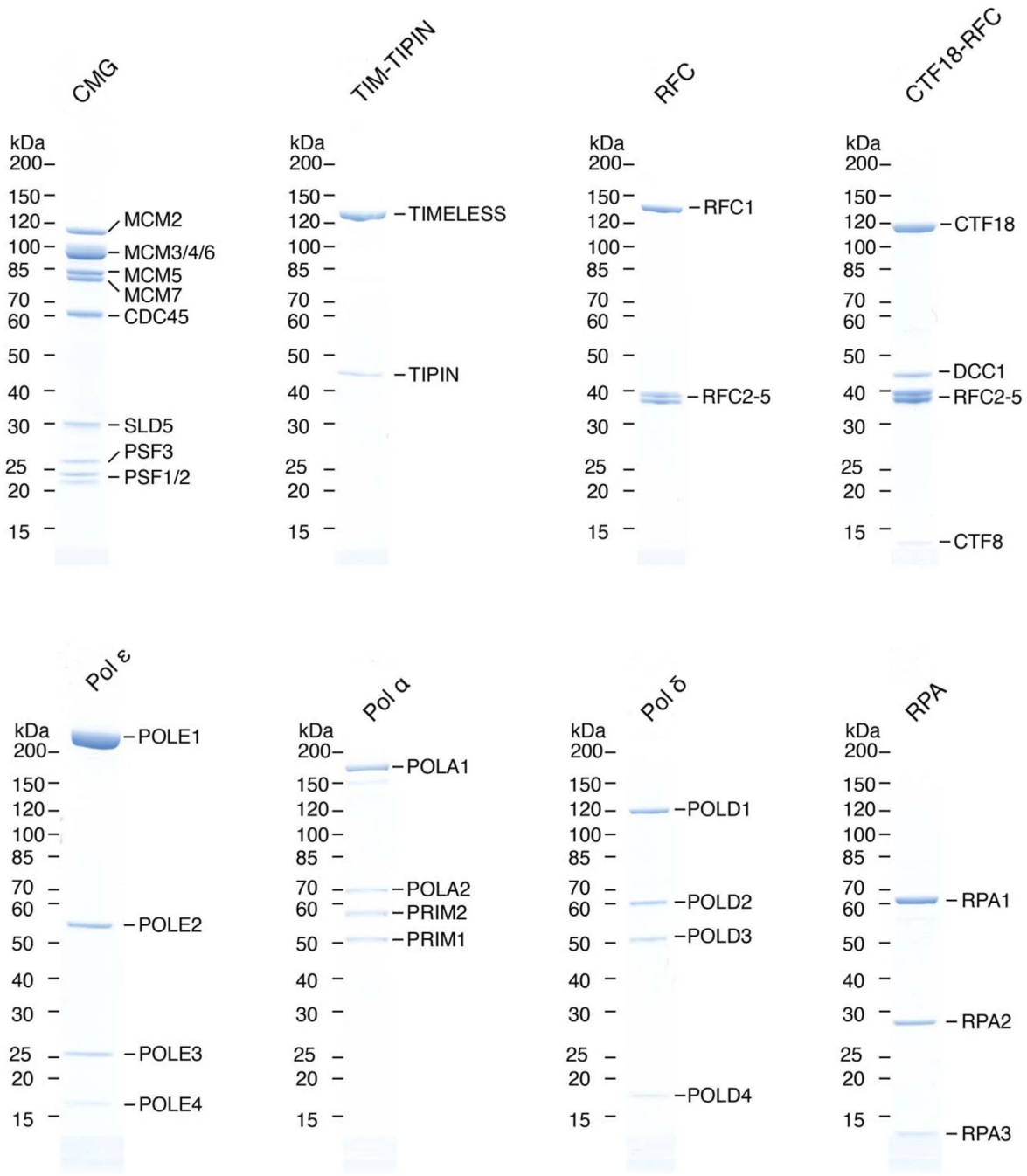
Maximal leading-strand synthesis rates were determined as previously described[8,17]. In brief, to determine the maximal product length from pulse-chase experiments, lane profiles were first generated in ImageJ. Straight lines were then manually fit to the steepest phase of the leading-strand peak. For all experiments, maximal product length was taken to be the point at which the line intercepted the lane background signal. Data were fit to linear regressions in GraphPad Prism and the slope of the regression was used to calculate the maximal rate of leading-strand synthesis.

Statistics and reproducibility

All leading-strand synthesis rates (Figs. 1e, 2c, 3b) were derived from three independent experiments. The experiments in Figs. 1c, d, 2a, b, 3a, d, f, 4a–c, e, g, 5b, c were performed three times. The experiments in Figs. 3c, e, 5d–f were performed twice. The experiments in Extended Data Figs. 2c, 6d, f, 8a were performed three times. The experiments in Extended Data Figs. 2a, 3a, c, e, f, 4b, 5b, d–g, 6a, g, 8d, e, 9b–e, 10c, d were performed twice.

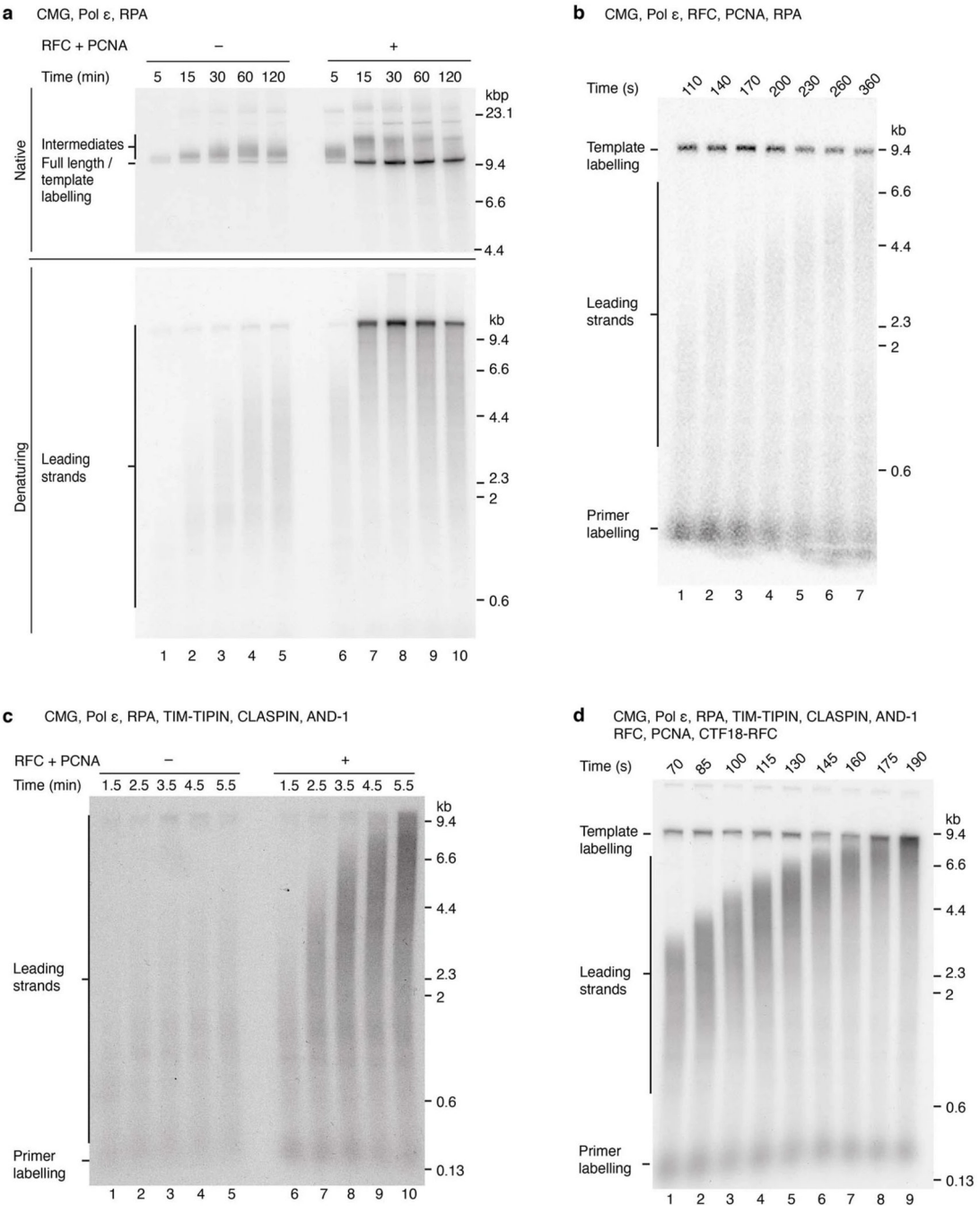
Europe PMC Funders Author Manuscripts

Extended Data



Extended Data Fig. 1. Purified human DNA replication proteins.

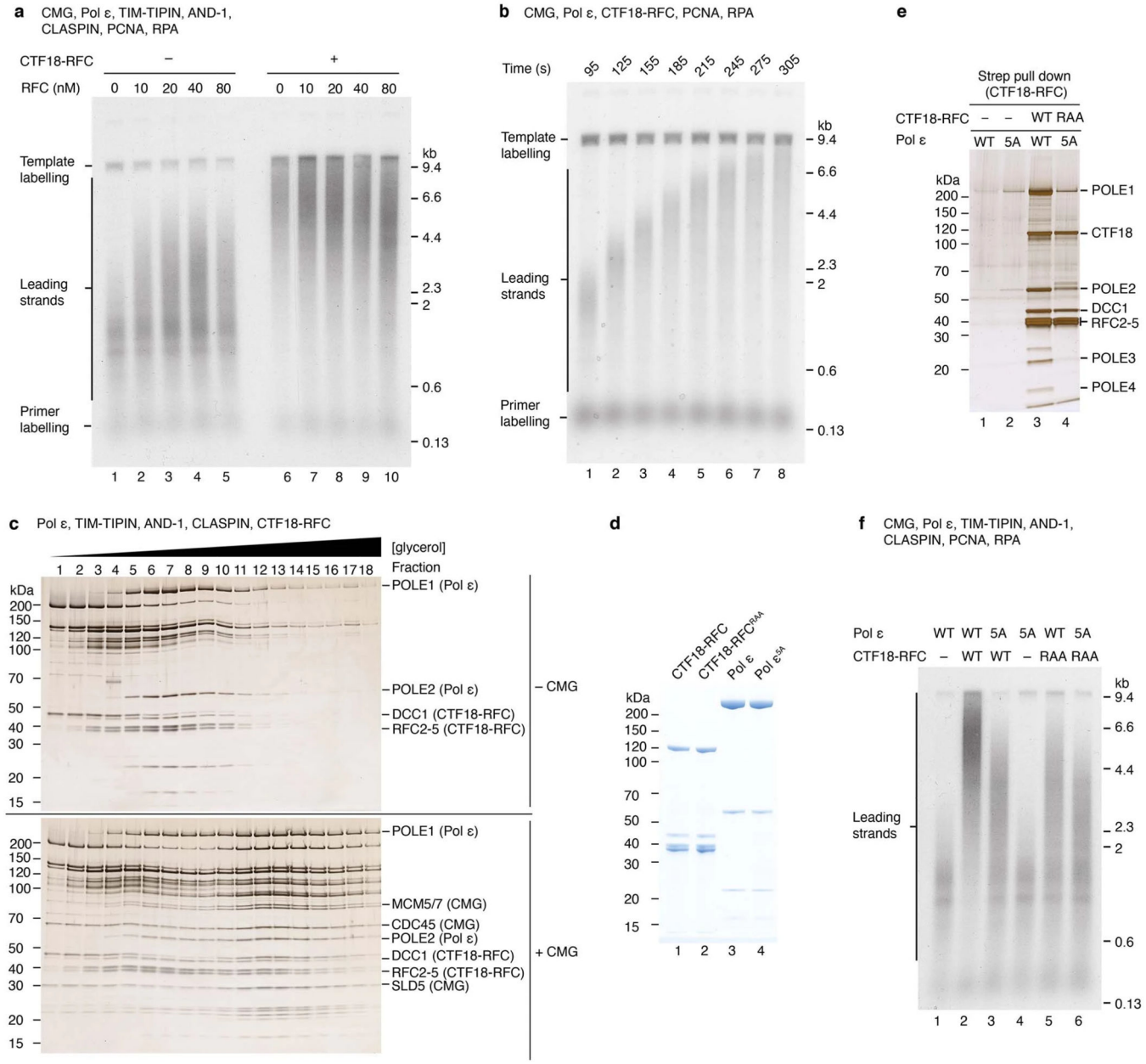
Coomassie stained SDS-PAGE of human DNA replication proteins. Individual lanes from the gel in Fig. 1a are shown with each subunit labelled.



Extended Data Fig. 2. Leading-strand synthesis.

a, Standard replication reaction on the 9.7 kbp template performed with the indicated proteins and analysed by native and denaturing agarose gel electrophoresis as indicated. In the absence of RFC and PCNA the predominant replication products (intermediates) migrate above the position of full length in the native gel. As indicated, in the native gel template labelling products and complete full-length replication products migrate in the same position. **b**, Denaturing agarose gel analysis of a pulse chase experiment on the 9.7 kbp template with the indicated proteins. Unless otherwise stated, in this and all pulse chase

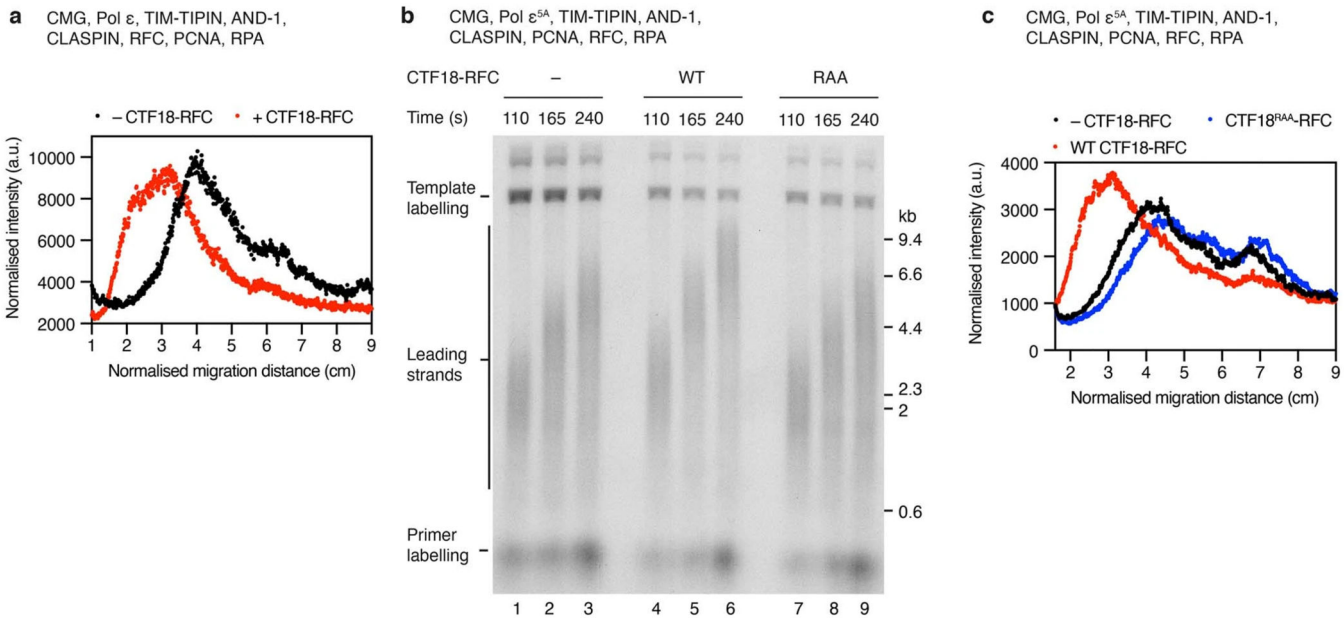
experiments, the chase was added at 50 s. **c**, Denaturing agarose gel analysis of a replication reaction on the 9.7 kbp template with the indicated proteins. **d**, Denaturing agarose gel analysis of a pulse chase experiment on the 9.7 kbp template with the indicated proteins.



Extended Data Fig. 3. CTF18-RFC is required for optimal leading-strand synthesis.

a, Denaturing agarose gel analysis of a replication reaction on the 9.7 kbp template with the indicated proteins. **b**, Denaturing agarose gel analysis of a pulse chase experiment on the 9.7 kbp template with the indicated proteins. **c**, Silver-stained SDS-PAGE analysis of glycerol gradients performed with the indicated proteins. For clarity, only CMG, CTF18-RFC and Polε subunits are annotated. **d**, Coomassie stained SDS-PAGE of CTF18-RFC and Pol ε interaction mutants. **e**, Silver-stained SDS-PAGE analysis of a pull-down experiment with

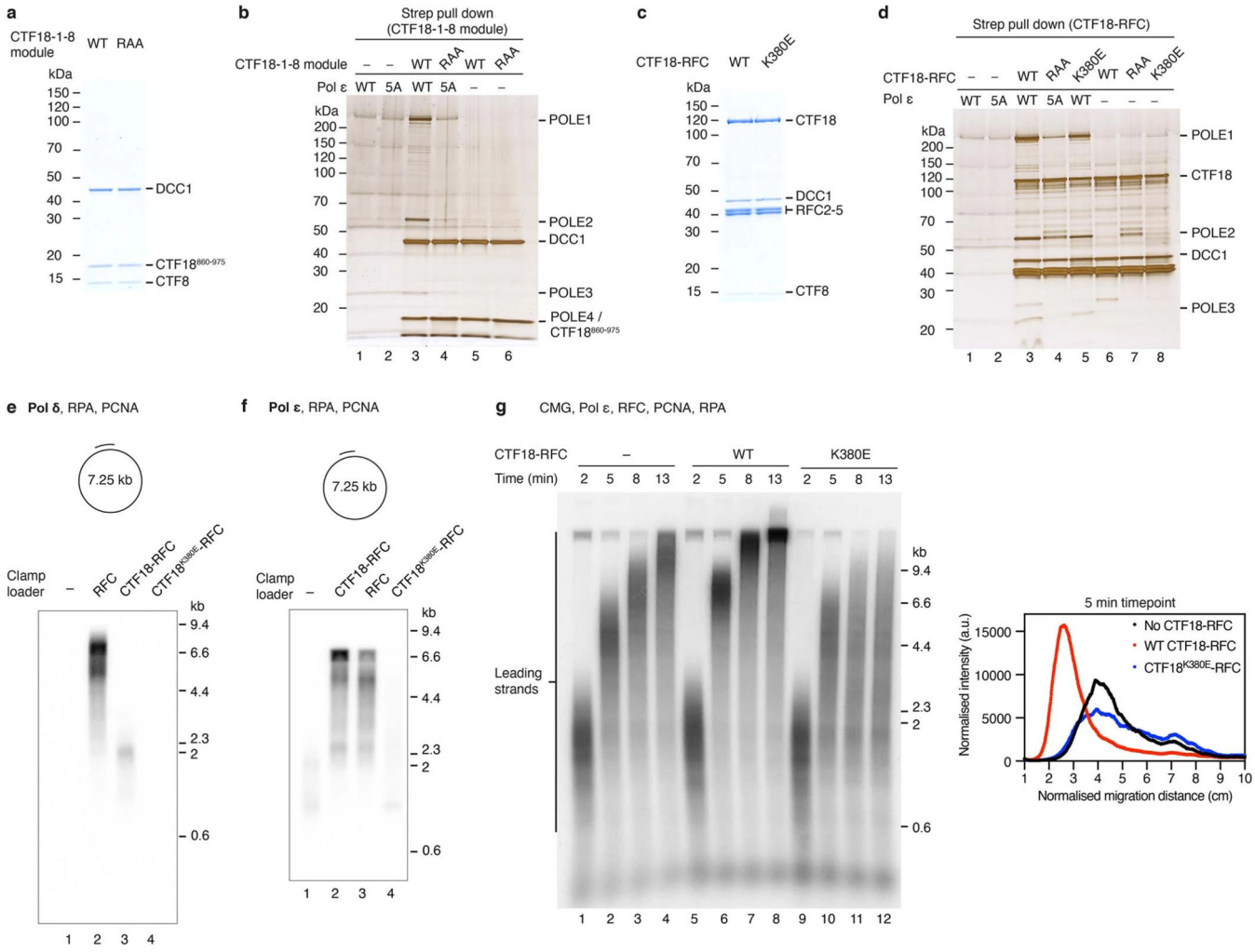
the indicated proteins showing that mutation of CTF18 and POLE1 disrupt the interaction between the two proteins. **f**, Denaturing agarose gel analysis of a replication reaction performed for 3 min on the 9.7 kbp template with the indicated proteins.



Extended Data Fig. 4. CTF18-RFC accelerates established replication forks.

a, Lane profiles of the 165 s timepoints in Fig. 3e where CTF18-RFC was absent or added in the chase (lanes 2 and 8 respectively). **b**, Denaturing agarose gel analysis of a pulse chase experiment on the 15.8 kbp template with the indicated proteins. Where indicated CTF18-RFC or CTF18-RFC^{RAA} were added with the chase. **c**, Lane profiles of the 240 s timepoints in (**b**). For lane profiles (**a**, **c**), product intensities were normalised by dividing each value by the relative intensity of the total signal in a given lane.

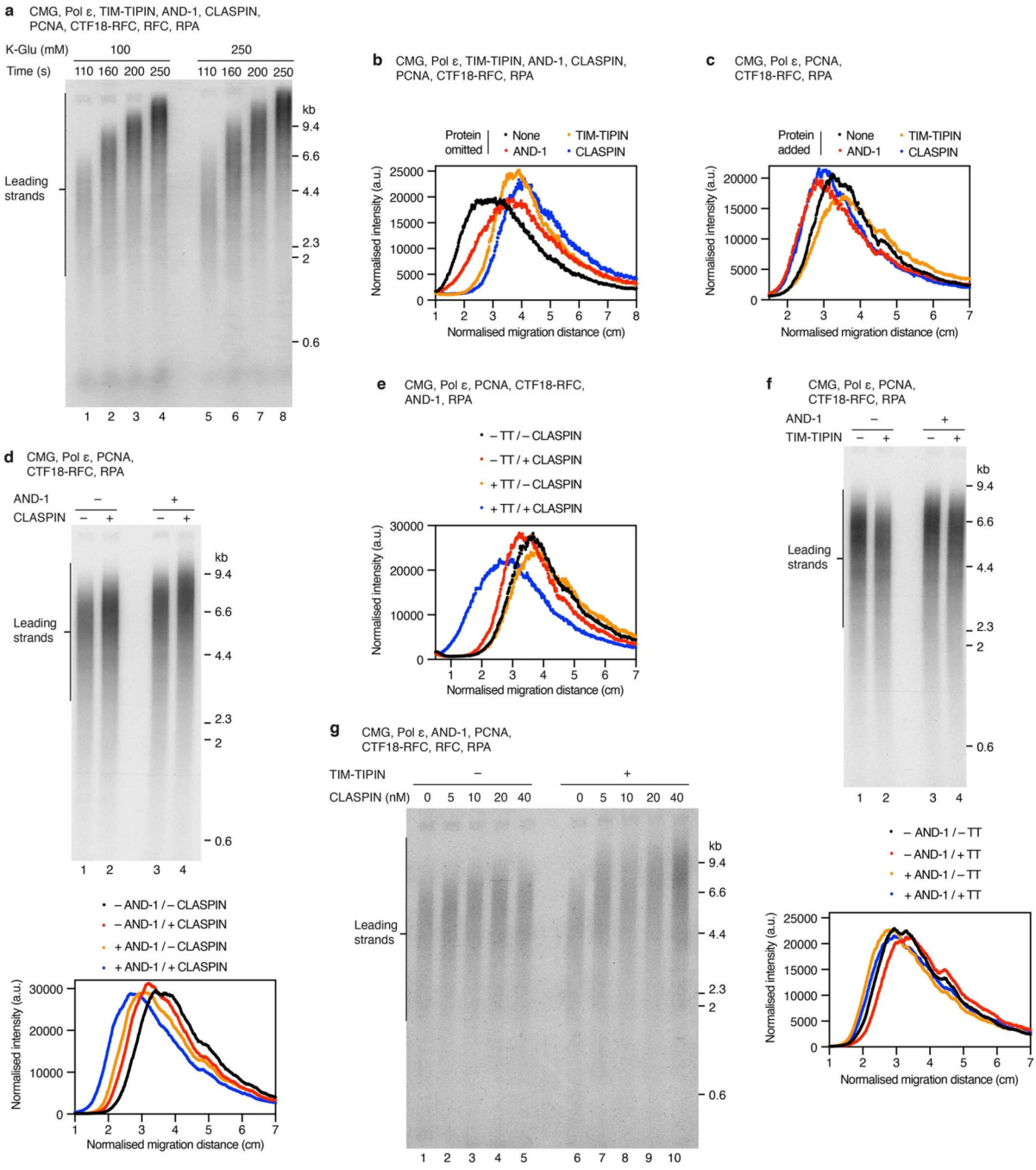
Europe PMC Funders Author Manuscripts



Extended Data Fig. 5. PCNA loading by CTF18-RFC is required for optimal leading-strand synthesis.

a, Coomassie stained SDS-PAGE of CTF18-1-8 module complexes. **b**, Silver-stained SDS-PAGE analysis of a pull-down experiment with the indicated proteins showing that the CTF18-1-8 module interacts specifically with Pol ε. **c**, Coomassie stained SDS-PAGE of WT and CTF18^{K380E} complexes. **d**, Silver-stained SDS-PAGE analysis of a pull-down experiment with the indicated proteins showing that CTF18^{K380E}-RFC retains the capacity to interact with Pol ε. **e**, **f**, Primer extension reactions on M13mp18 single-strand DNA with Pol δ and Pol ε showing that CTF18^{K380E}-RFC has a severe defect in supporting PCNA-dependent DNA synthesis by both polymerases. **g**, (left) Denaturing agarose gel analysis of a pulse-chase experiment on the 15.8 kbp template with the indicated proteins. The chase was added at 1 min 45 s. Where indicated CTF18-RFC or CTF18^{K380E}-RFC were added with the chase. (right) Lane profiles for the 5 min timepoint. Data were normalised by dividing each intensity value by the relative total signal at the 2 min timepoint.

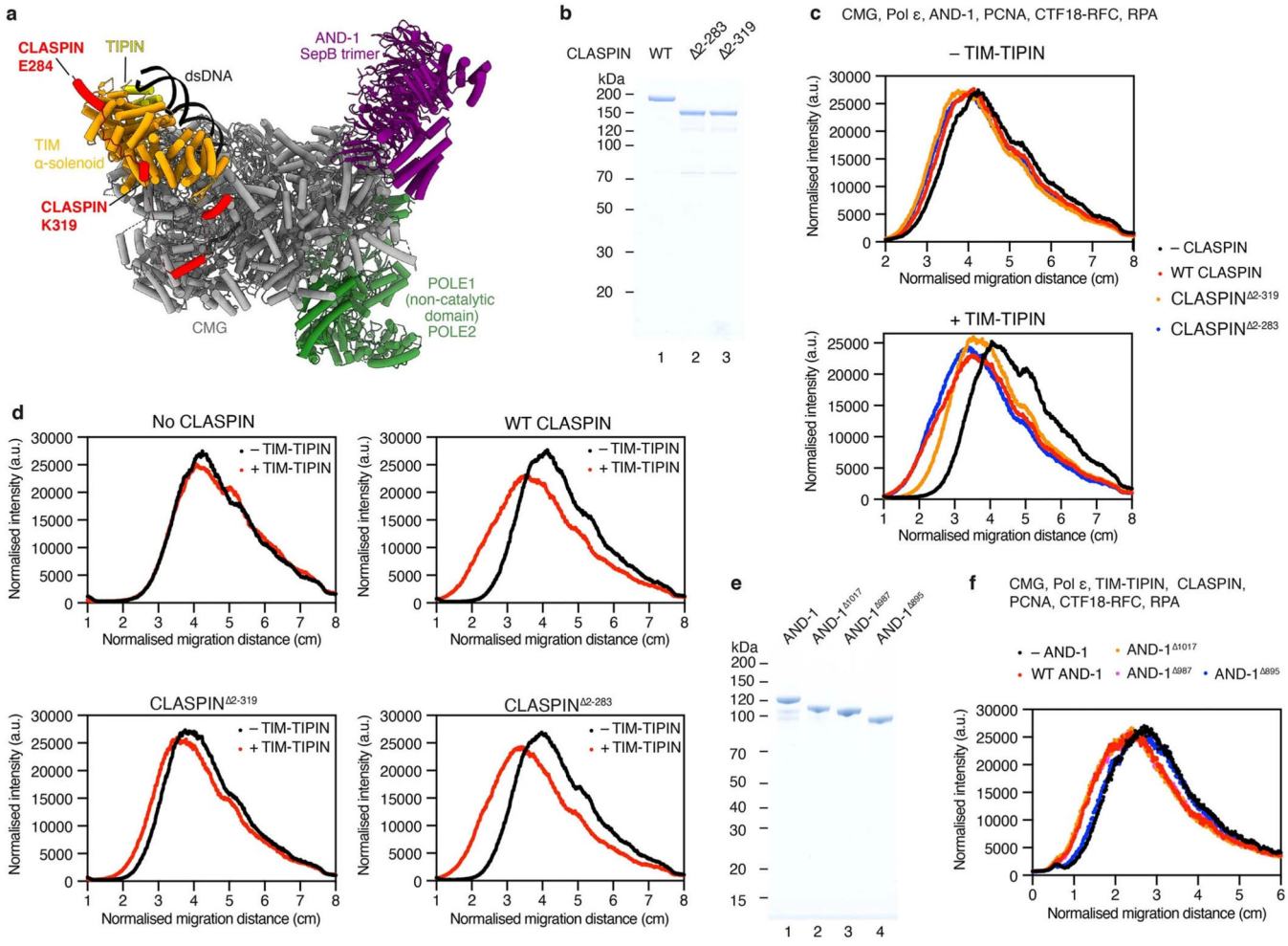
Europe PMC Funders Author Manuscripts



Extended Data Fig. 6. TIM-TIPIN, CLASPIN and AND-1 enhance leading-strand replication.

a, Denaturing agarose gel analysis of a time course experiment on the 15.8 kbp template with the indicated proteins at two concentrations of potassium glutamate (K-Glu). **b**, **c**, Lane profiles of the data in Fig. 4a, b respectively. **d**, Denaturing agarose gel analysis (top) and lane profiles (bottom) of a 3 min 45 s replication reaction on the 15.8 kbp template with the indicated proteins. **e**, Lane profiles of the data in Fig. 4c. TT, TIM-TIPIN. **f**, Denaturing agarose gel analysis (top) and lane profile (bottom) of a 3 min 45 s replication reaction on the 15.8 kbp template with the indicated proteins. TT, TIM-TIPIN. **g**, Denaturing agarose

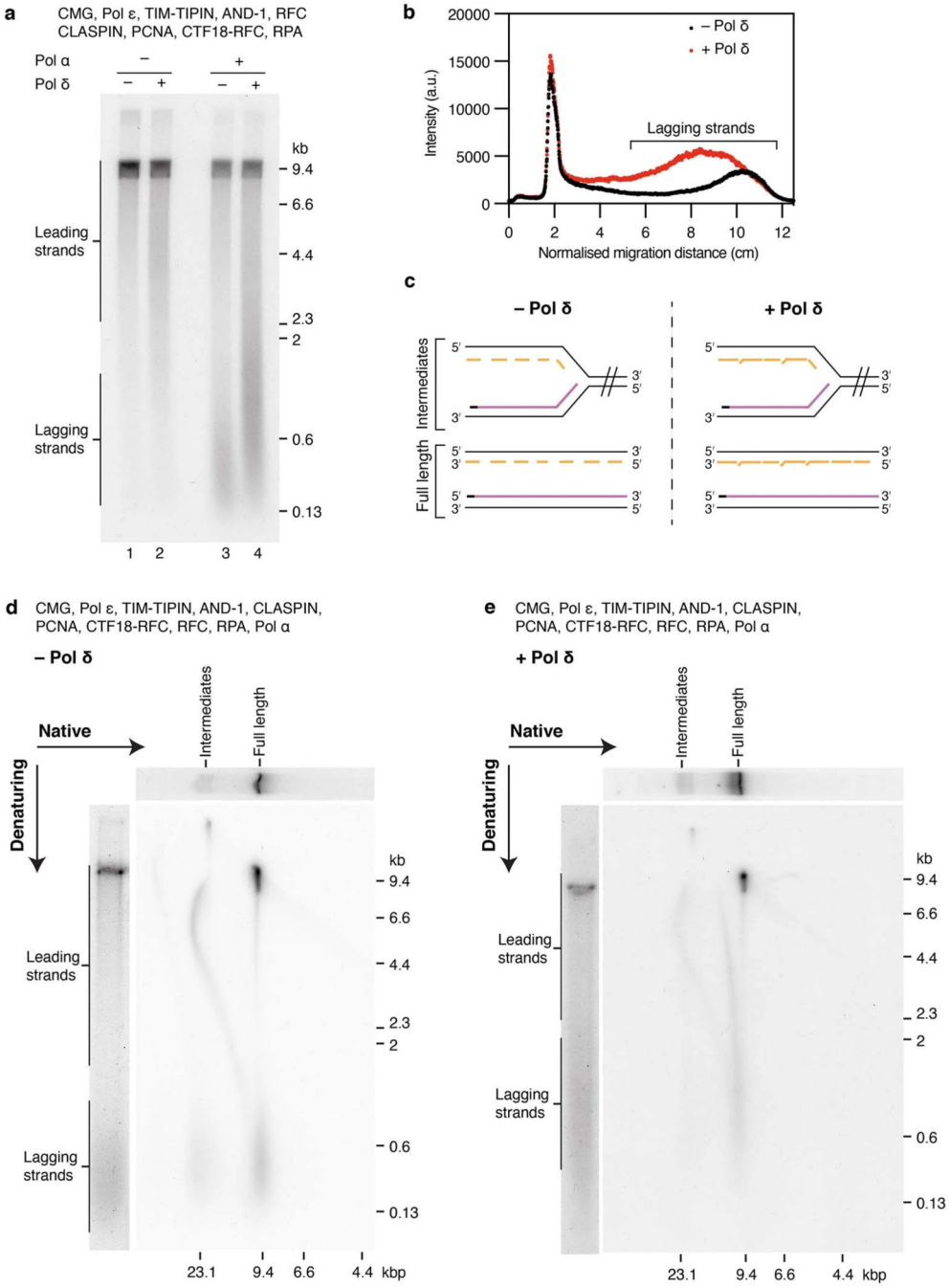
gel analysis of a 3.5 min replication reaction on the 15.8 kbp template with the indicated proteins. In **d**, **f**, **g**, the potassium glutamate concentration was 250 mM. For lane profiles (**b**–**f**), product intensities were normalised by dividing each value by the relative intensity of the total signal in a given lane.



Extended Data Fig. 7. CLASPIN and AND-1 truncations.

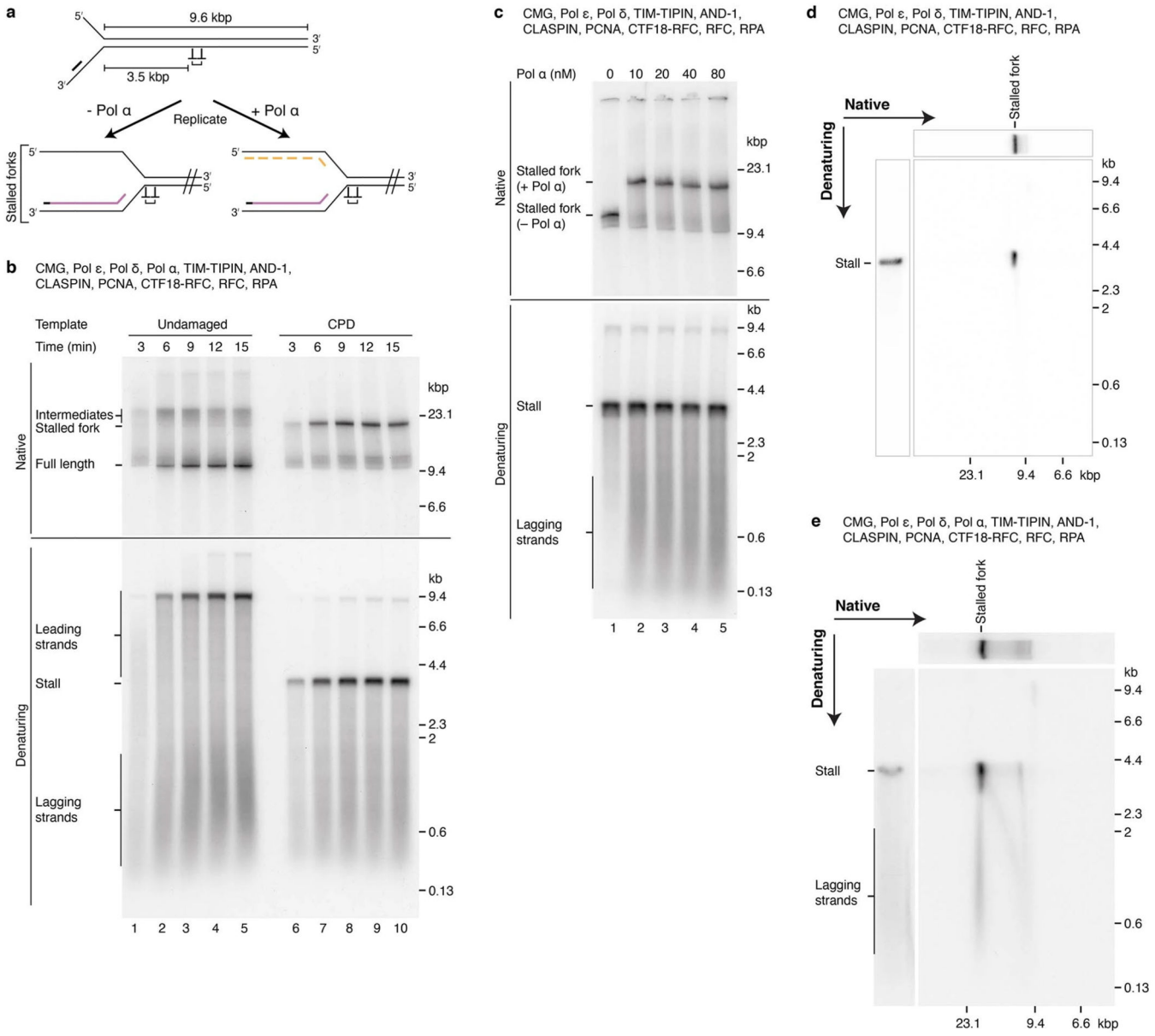
a, Cartoon representation of the core human replisome (PDB:7PFO)[2] showing the region of CLASPIN (E284–K319) that interacts with the TIM α-solenoid. **b**, Coomassie stained SDS-PAGE of CLASPIN truncation mutants. **c**, **d** Lane profiles of the data in Fig. 4e. **e**, Coomassie stained SDS-PAGE of AND-1 truncation mutants. **f**, Lane profiles of the data in Fig. 4g. For lane profiles (**c**, **d**, **f**), product intensities were normalised by dividing each value by the relative intensity of the total signal in a given lane.

Europe PMC Funders Author Manuscripts



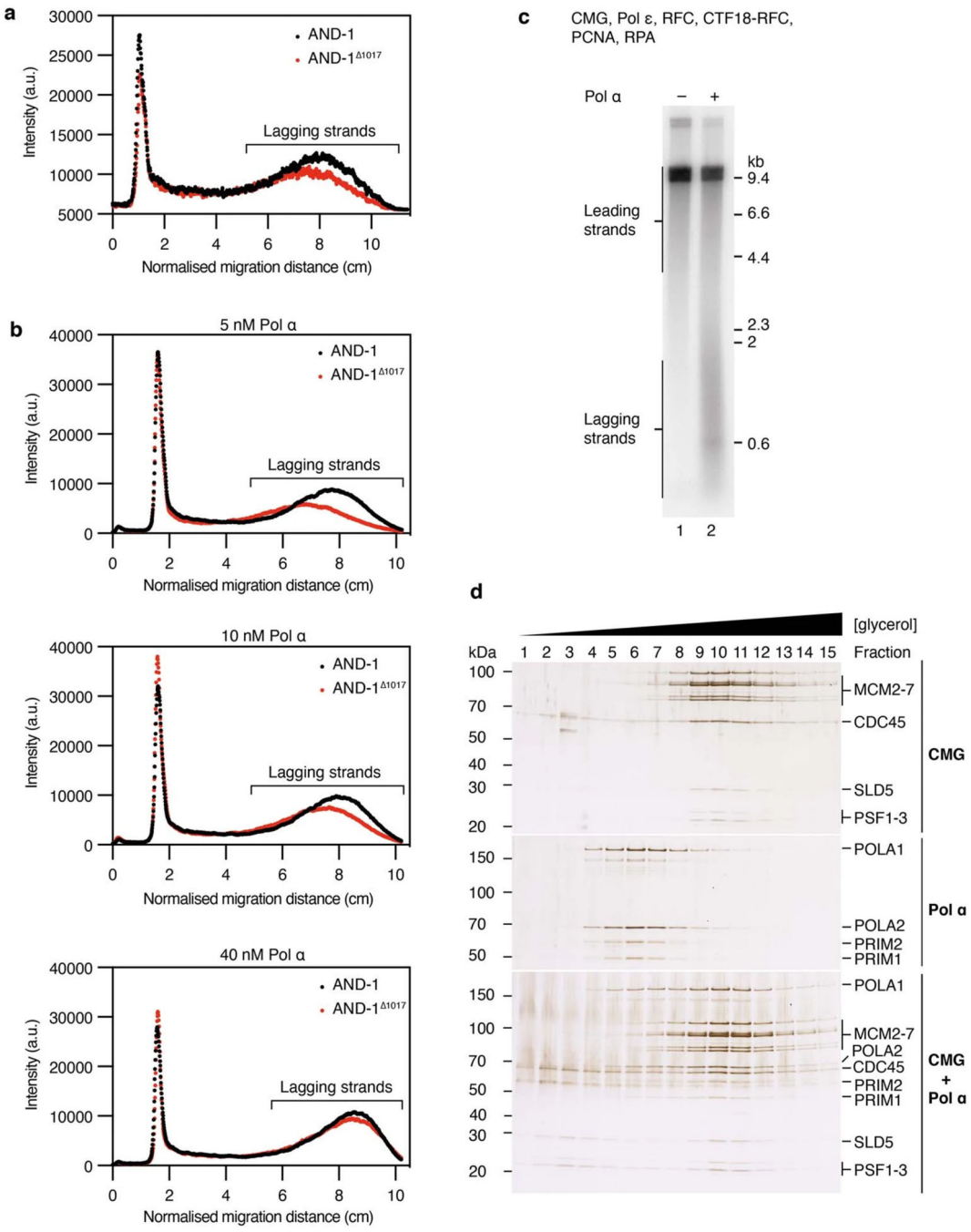
Extended Data Fig. 8. Reconstitution of lagging-strand replication.

a, Denaturing agarose gel analysis of a 20 min replication on the 9.7 kbp template with the indicated proteins. **b**, Lane profiles from lanes 3 and 4 in (**a**). **c**, Schematic showing the possible replication products (–/+ Pol δ) if lagging strands are extended by Pol δ and are constituents of both replication intermediates and full-length products. **d**, **e**, Two-dimensional agarose gel analysis of 20 min replication reactions performed with the indicated proteins on the 9.7 kbp template in the absence (**d**) and presence (**e**) of Pol δ. In all reactions, the concentration of potassium glutamate was 250 mM.



Extended Data Fig. 9. Lagging-strand replication occurs at all replication forks.

a, Schematic of a replication reaction on the cyclobutane pyrimidine dimer (CPD) template. **b**, Native and denaturing gel analysis of a time course experiment on undamaged and CPD templates with the indicated proteins. **c**, Native and denaturing gel analysis of a 60 min reaction on the CPD template with different concentrations of Pol α as indicated. **d**, **e**, Two-dimensional agarose gel analysis of 30 min replication reactions performed with the indicated proteins on the CPD template in the absence (**d**) and presence (**e**) of Pol α. In all reactions, the concentration of potassium glutamate was 250 mM.



Extended Data Fig. 10. Role of AND-1 in lagging-strand replication.

a, Lane profiles from Fig. 5d, lanes 5 and 6. **b**, Lane profiles from the experiment in Fig. 5e, lanes 2, 3, 5, 7, 8 and 10. **c**, Denaturing agarose gel analysis of a 30 min reaction on the 9.7 kbp template with the indicated proteins. **d**, Silver-stained SDS-PAGE analysis of glycerol gradients performed with the indicated proteins demonstrating complex formation between Pol α and CMG in the absence of replication fork DNA.

Europe PMC Funders Author Manuscripts

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank J. Shi for operation of the LMB baculovirus facility, L. Passmore for protein expression vectors, M. Jones for advice on glycerol gradients, and L. Krejci and M. Wold for the RPA expression plasmid. This work was supported by the MRC, as part of UK Research and Innovation (MRC grant MC_UP_1201/12 to J.T.P.Y.), the Wellcome Trust (reference 110014/Z/15/Z) for a Sir Henry Wellcome Postdoctoral Fellowship to M.R.G.T and an LMB Cambridge Trust International Scholarship to Y.B.

Data availability

The data supporting the findings of this study are available within the paper and its Supplementary Information files.

References

1. Baretić D, et al. Cryo-EM structure of the fork protection complex bound to CMG at a replication fork. Mol Cell. 2020; 78: 926–940. e13 [PubMed: 32369734]
2. Jones ML, Baris Y, Taylor MRG, Yeeles JTP. Structure of a human replisome shows the organisation and interactions of a DNA replication machine. EMBO J. 2021; 40 e108819 [PubMed: 34694004]
3. Goswami P, et al. Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. Nat Commun. 2018; 9 5061 [PubMed: 30498216]
4. Yuan Z, et al. Ctf4 organizes sister replisomes and Pol alpha into a replication factory. eLife. 2019; 8 e47405 [PubMed: 31589141]
5. Rzechorzek NJ, et al. CryoEM structures of human CMG–ATPγS–DNA and CMG–AND-1 complexes. Nucleic Acids Res. 2020; 48: 6980–6995. [PubMed: 32453425]
6. Kapadia N, et al. Processive activity of replicative DNA polymerases in the replisome of live eukaryotic cells. Mol Cell. 2020; 80: 114–126. e8 [PubMed: 32916094]
7. Lewis JS, et al. Tunability of DNA polymerase stability during eukaryotic DNA replication. Mol Cell. 2020; 77: 17–25. e5 [PubMed: 31704183]
8. Yeeles JTP, Janska A, Early A, Diffley JFX. How the eukaryotic replisome achieves rapid and efficient DNA replication. Mol Cell. 2017; 65: 105–116. [PubMed: 27989442]
9. Kilkenny ML, et al. The human CTF4-orthologue AND-1 interacts with DNA polymerase alpha/primase via its unique C-terminal HMG box. Open Biol. 2017; 7 170217 [PubMed: 29167311]
10. Guan C, Li J, Sun D, Liu Y, Liang H. The structure and polymerase-recognition mechanism of the crucial adaptor protein AND-1 in the human replisome. J Biol Chem. 2017; 292: 9627–9636. [PubMed: 28381552]
11. Petermann E, Helleday T, Caldecott KW. Claspin promotes normal replication fork rates in human cells. Mol Biol Cell. 2008; 19: 2373–2378. [PubMed: 18353973]
12. Conti C, et al. Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. Mol Biol Cell. 2007; 18: 3059–3067. [PubMed: 17522385]
13. Somyajit K, et al. Redox-sensitive alteration of replisome architecture safeguards genome integrity. Science. 2017; 358: 797–802. [PubMed: 29123070]
14. Abe T, et al. AND-1 fork protection function prevents fork resection and is essential for proliferation. Nat Commun. 2018; 9 3091 [PubMed: 30082684]
15. Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PM, Kunkel TA. Division of labor at the eukaryotic replication fork. Mol Cell. 2008; 30: 137–144. [PubMed: 18439893]

16. Pursell ZF, Isoz I, Lundstrom EB, Johansson E, Kunkel TA. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science. 2007; 317: 127–130. [PubMed: 17615360]
17. Aria V, Yeeles JTP. Mechanism of bidirectional leading-strand synthesis establishment at eukaryotic DNA replication origins. Mol Cell. 2019; 73: 199–211. e10
18. Grabarczyk DB, Silkenat S, Kisker C. Structural basis for the recruitment of Ctf18-RFC to the replisome. Structure. 2018; 26: 137–144. e3 [PubMed: 29225079]
19. Stokes K, Winczura A, Song B, Piccoli G, Grabarczyk DB. Ctf18-RFC and DNA Pol form a stable leading strand polymerase/clamp loader complex required for normal and perturbed DNA replication. Nucleic Acids Res. 2020; 48: 8128–8145. [PubMed: 32585006]
20. Murakami T, et al. Stable interaction between the human proliferating cell nuclear antigen loader complex Ctf18-replication factor C (RFC) and DNA polymerase ε is mediated by the cohesion-specific subunits, Ctf18, Dcc1, and Ctf8*. J Biol Chem. 2010; 285: 34608–34615. [PubMed: 20826785]
21. Fujisawa R, Ohashi E, Hirota K, Tsurimoto T. Human CTF18-RFC clamp-loader complexed with non-synthesising DNA polymerase ε efficiently loads the PCNA sliding clamp. Nucleic Acids Res. 2017; 45: 4550–4563. [PubMed: 28199690]
22. Tunyasuvunakool K, et al. Highly accurate protein structure prediction for the human proteome. Nature. 2021; 596: 590–596. [PubMed: 34293799]
23. Taylor MRG, Yeeles JTP. The initial response of a eukaryotic replisome to DNA damage. Mol Cell. 2018; 70: 1067–1080. e12 [PubMed: 29944888]
24. Georgescu RE, et al. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. Nat Struct Mol Biol. 2014; 21: 664–670. [PubMed: 24997598]
25. Terret ME, Sherwood R, Rahman S, Qin J, Jallepalli PV. Cohesin acetylation speeds the replication fork. Nature. 2009; 462: 231–234. [PubMed: 19907496]
26. Crabbe L, et al. Analysis of replication profiles reveals key role of RFC-Ctf18 in yeast replication stress response. Nat Struct Mol Biol. 2010; 17: 1391–1397. [PubMed: 20972444]
27. Hanna JS, Kroll ES, Lundblad V, Spencer FA. Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. Mol Cell Biol. 2001; 21: 3144–3158. [PubMed: 11287619]
28. Mayer ML, Gygi SP, Aebersold R, Hieter P. Identification of RFC(Ctf18p, Ctf8p, Dcc1p): an alternative RFC complex required for sister chromatid cohesion in S. cerevisiae. Mol Cell. 2001; 7: 959–970. [PubMed: 11389843]
29. Kawasumi R, et al. Vertebrate CTF18 and DDX11 essential function in cohesion is bypassed by preventing WAPL-mediated cohesin release. Genes Dev. 2021; 35: 1368–1382. [PubMed: 34503989]
30. Georgescu RE, et al. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. eLife. 2015; 4 e04988 [PubMed: 25871847]
31. Henricksen LA, Umbricht CB, Wold MS. Recombinant replication protein A: expression, complex formation, and functional characterization. J Biol Chem. 1994; 269: 11121–11132. [PubMed: 8157639]
32. Sebesta M, et al. Role of PCNA and TLS polymerases in D-loop extension during homologous recombination in humans. DNA Repair. 2013; 12: 691–698. [PubMed: 23731732]
33. Xing X, et al. A recurrent cancer-associated substitution in DNA polymerase ε produces a hyperactive enzyme. Nat Commun. 2019; 10: 374. [PubMed: 30670691]

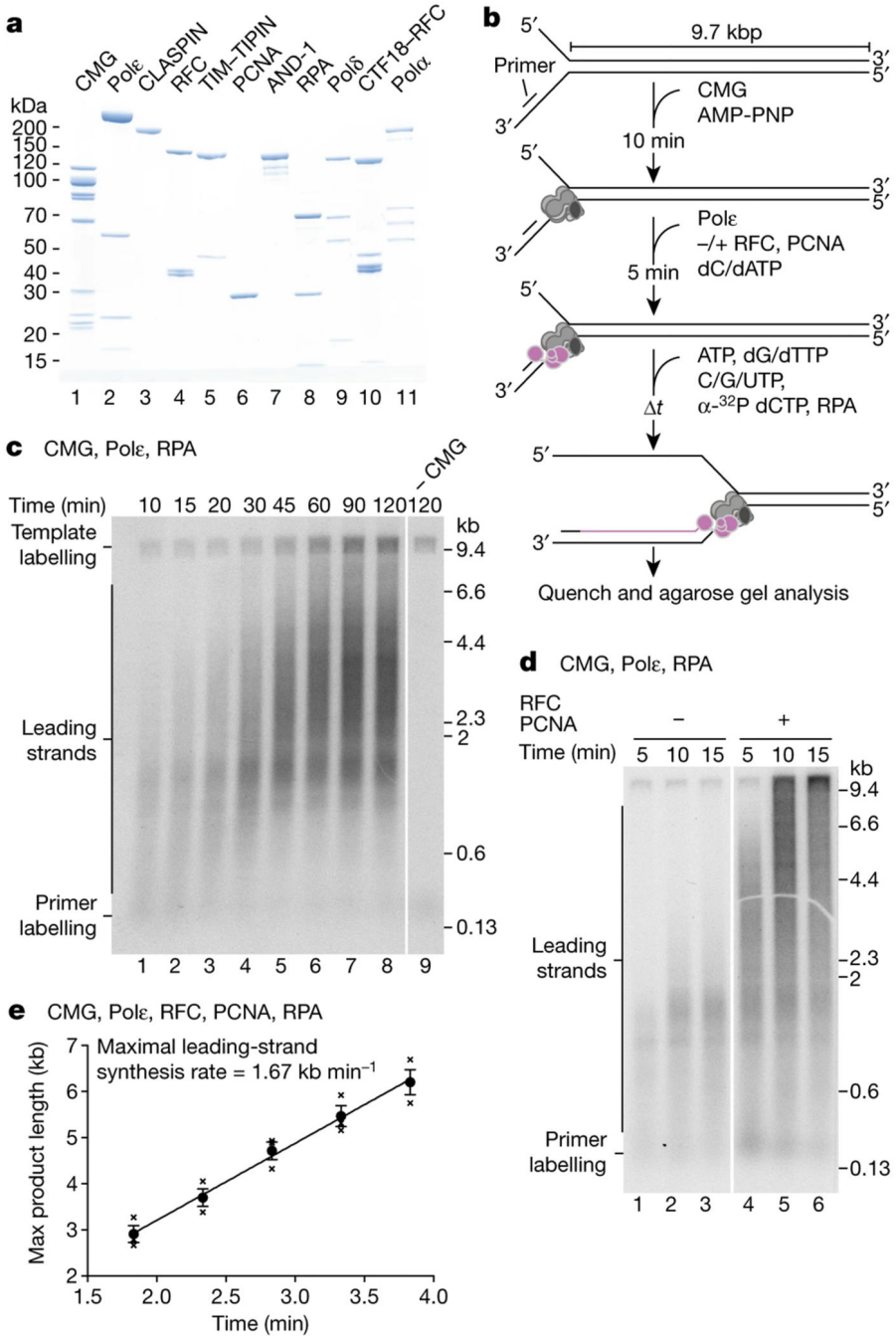


Fig. 1. PCNA is required for efficient leading-strand synthesis by CMG-Polε.

a, Coomassie-stained SDS–PAGE analysis of purified human DNA replication proteins. **b**, Schematic illustrating the reaction scheme for CMG-based DNA replication assays using forked DNA templates. The CMG complex is coloured grey and Polε is coloured purple. **c**, Denaturing agarose gel analysis of a time course experiment performed as in **b** in the absence of RFC and PCNA on a 9.7-kbp forked DNA template. The CMG-independent primer and template labelling products are indicated. **d**, Experiment performed as in **c** with RFC and PCNA included where indicated. **e**, Quantification of the maximal leading-strand

synthesis rate from pulse-chase experiments (Extended Data Fig. 2b). Linear regression is fit to the mean of three experiments. The error bars represent the s.e.m., the mean is indicated by filled circles and individual data points are represented as crosses. In all figures, the proteins present in each reaction are shown above the gel or graph. For gel source data in this and all subsequent figures, see Supplementary Fig. 1.

Europe PMC Funders Author Manuscripts

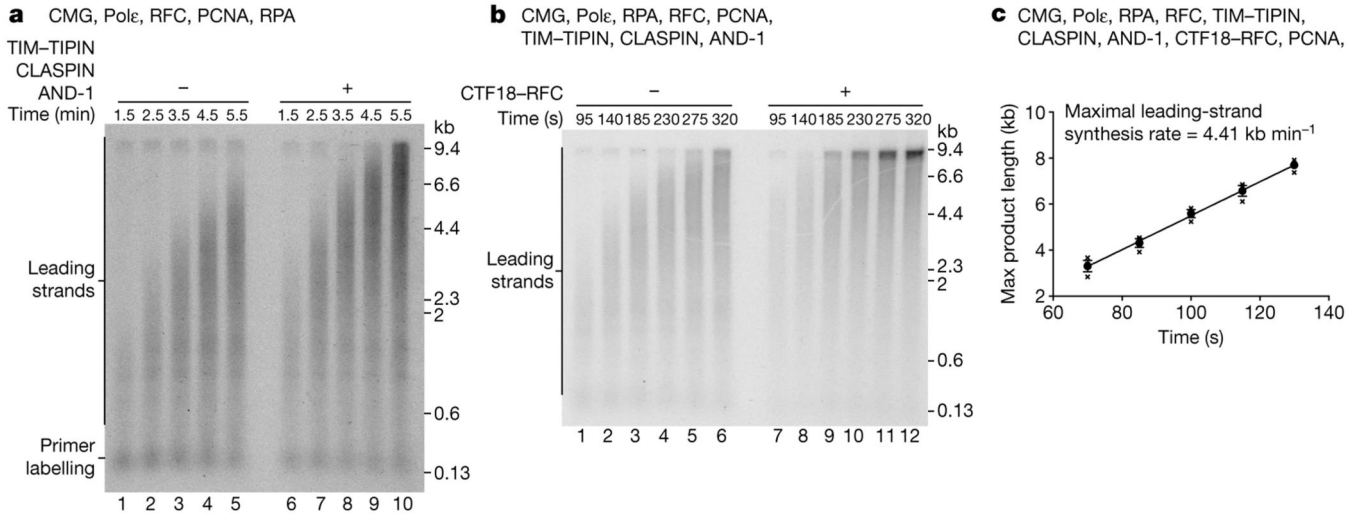


Fig. 2. Reconstitution of cellular DNA replication rates with purified proteins.

a, Replication reaction performed as in Fig. 1b, c but including RFC and PCNA. TIM–TIPIN, CLASPIN and AND-1 were included where indicated. The additional proteins were added after the CMG-binding step (Fig. 1b). **b**, Replication reaction performed in the presence of TIM–TIPIN, CLASPIN and AND-1, with CTF18–RFC added where indicated. **c**, Quantification of the maximal leading-strand synthesis rate from pulse-chase experiments (Extended Data Fig. 2d). Linear regression is fit to the mean of three experiments. The error bars represent the s.e.m., the mean is indicated by filled circles and the individual data points are represented as crosses.

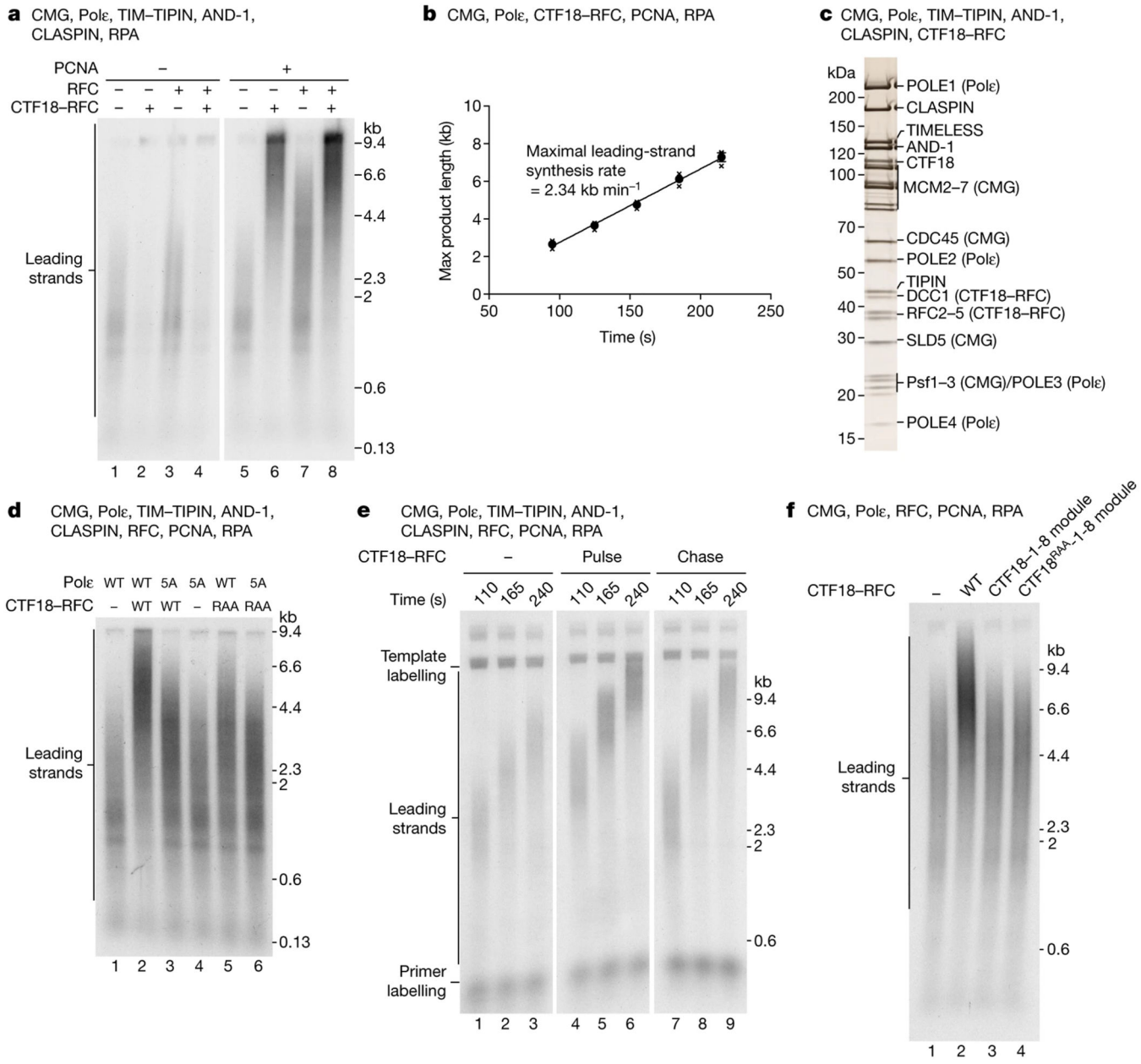


Fig. 3. PCNA loading by CTF18–RFC coupled to Polε is required for maximal replication rates.

a, Replication reaction performed on the 9.7-kbp template for 3 min 30 s with PCNA, RFC and CTF18–RFC added or omitted where indicated. **b**, Quantification of the maximal leading-strand synthesis rate from pulse-chase experiments (Extended Data Fig. 3b). Linear regression is fit to the mean of three experiments. The error bars represent the s.e.m., the mean is indicated by filled circles and the individual data points are represented as crosses. **c**, Silver-stained SDS-PAGE of a peak fraction from a glycerol sedimentation gradient performed with the indicated proteins (Extended Data Fig. 3c; +CMG, fraction 13). **d**, Denaturing agarose gel of a 3-min replication reaction comparing Polε and CTF18–RFC interaction mutants. WT, wild type. **e**, Pulse-chase experiment on the 15.8-kbp template where CTF18–RFC was either added in the pulse or with the chase. Unless stated otherwise,

in this and all pulse-chase experiments, the chase was added after 50 s. **f**, Denaturing agarose gel analysis of a replication reaction performed for 3 min 40 s on the 15.8-kbp template with the indicated proteins.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

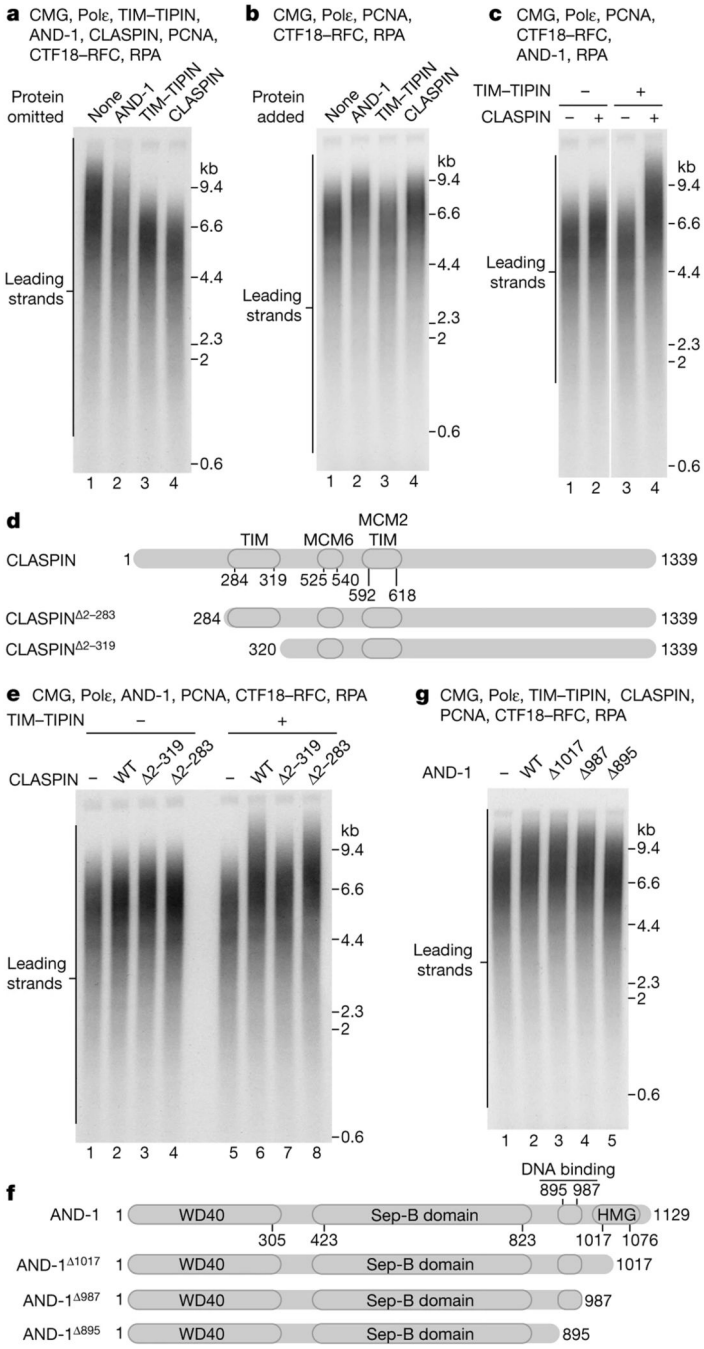


Fig. 4. How TIM–TIPIN, CLASPIN and AND-1 facilitate leading-strand replication.

a-c, Denaturing agarose gel analysis of replication reactions performed on the 15.8-kbp forked DNA template for 3 min 30 s with the indicated proteins. **d**, Diagrams of the primary structure for full-length CLASPIN and CLASPIN truncation mutants. Regions of CLASPIN that bind to MCM and TIMELESS are indicated[2]. **e**, Denaturing agarose gel analysis of a replication reaction performed on the 15.8-kbp forked DNA template for 3 min 40 s with the indicated proteins. **f**, Diagrams of the primary structure for full-length AND-1 and AND-1 truncation mutants. **g**, Denaturing agarose gel analysis of a replication reaction performed

on the 15.8-kbp forked DNA template for 3 min 40 s with the indicated proteins. In all reactions, the concentration of potassium glutamate was 250 mM.

Europe PMC Funders Author Manuscripts

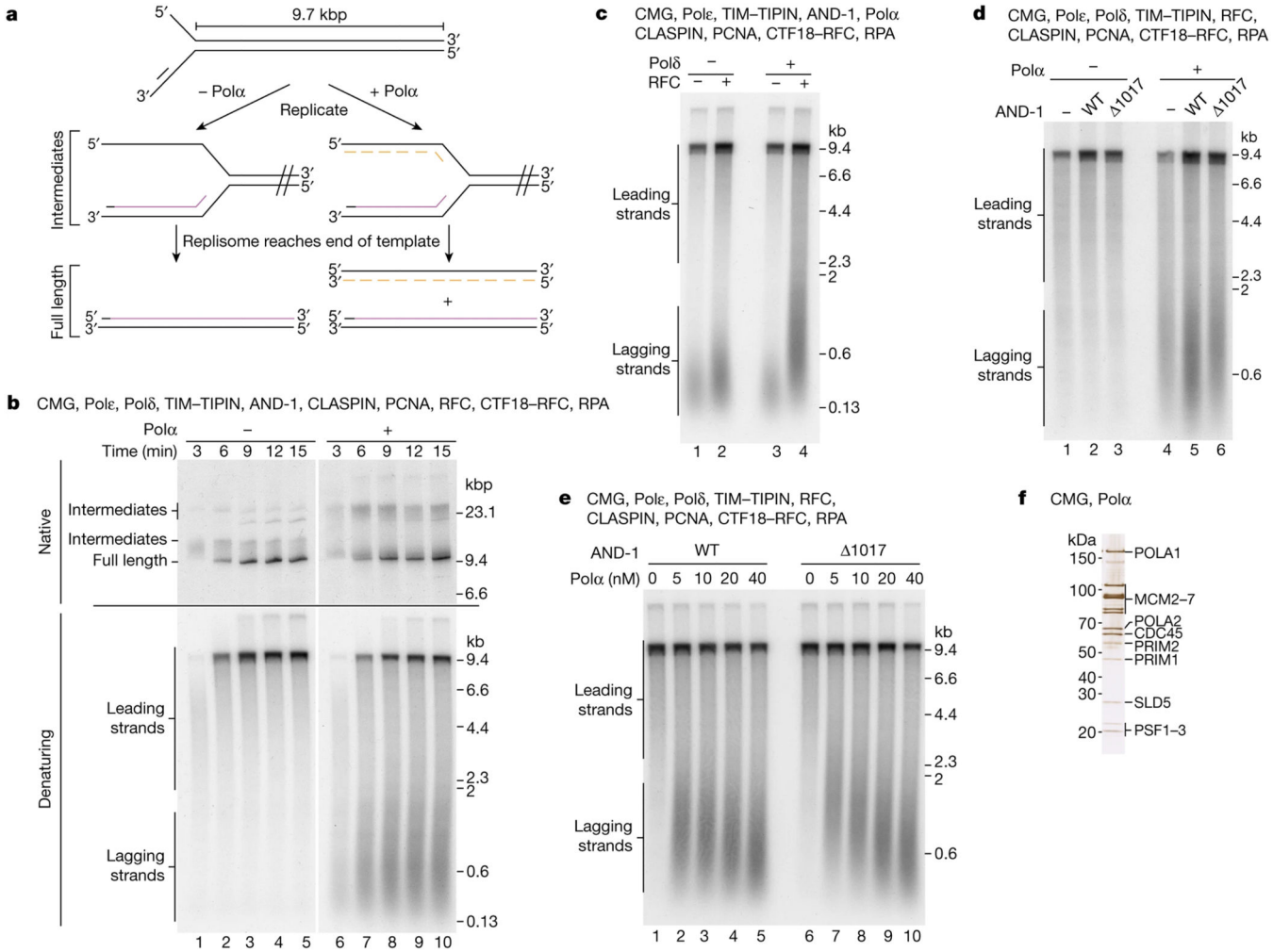


Fig. 5. Reconstitution of lagging-strand DNA replication with purified human proteins.

a, Schematic illustrating the expected DNA replication products in reactions with or without the primase Polα. **b**, Time course experiment on the 9.7-kbp template with or without Polα as indicated. Products were analysed by native and denaturing agarose gel electrophoresis. **c-e**, Denaturing agarose gel analysis of replication reactions performed on the 9.7-kbp template for 20 min with the indicated proteins. In all reactions, the concentration of potassium glutamate was 250 mM. f, Silver-stained SDS–PAGE of a peak fraction from a glycerol sedimentation gradient performed with the indicated proteins (Extended Data Fig. 10d; CMG + Polα, fraction 10).