HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2019 August 05.

Published in final edited form as:

Science. 2019 February 22; 363(6429): . doi:10.1126/science.aav7003.

Structures and operating principles of the replisome

Yang Gao[1], **Yanxiang Cui**[2], **Tara Fox**[3,4], **Shiqiang Lin**[1,5], **Huaibin Wang**[1], **Natalia de Val**[3,4], **Z. Hong Zhou**[2], **Wei Yang**[1,*]

[1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[2]The California NanoSystems Institute, UCLA, Los Angeles, CA 90095, USA

[3]Center for Molecular Microscopy, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20892, USA

[4]Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research Inc., Frederick, MD 21701, USA

[5]Current address: College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, P. R. China

Abstract

INTRODUCTION—DNA replication has been studied since the 1950s. It is well established that double helical DNA needs to be separated for replication by a helicase. Each strand is then copied by a DNA polymerase, continuously on the leading and discontinuously (via Okazaki fragments) on the lagging strand, where each DNA synthesis initiates from an RNA primer provided by primase. After six decades, how DNA polymerases, helicase, primase, and their accessory factors form a replisome and perform concerted leading and lagging strand synthesis at a replication fork had never been visualized in atomic detail.

RATIONALE—Bacteriophage T7 presents the simplest known DNA replication system, consisting of only three proteins. Helicase and primase reside in one polypeptide chain that forms a hexamer in the presence of DNA and ATP or dTTP. T7 DNA polymerase, aided by *E. coli* thioredoxin as its processivity factor, carries out both leading and lagging strand DNA synthesis. Based on published biochemical data, we designed a minimal DNA fork to trap these essential proteins in replication competent states.

RESULTS—We determined cryogenic-electron microscopy (cryo-EM) structures of the T7 replisome and showed how its essential enzymatic functions are coordinated in three dimensions. The hexameric helicase adopts a spiral “lock washer” form that encircles the coil-like lagging

*Correspondence to wei.yang@nih.gov.

Author contributions: Y.G. prepared all samples and carried out cryoEM data analysis and structure determination; S. L. cloned gp4 and gp5; Y.C. and Z. H. Z. collected cryoEM data of gp4-DNA complexes and helped with data processing; T. F. and N. de V. helped with data collection of gp4-gp5 complexes; W.Y. conceived and led the research project; Y.G. and W.Y. wrote the manuscript.

Competing interests: Authors declare having no competing interests.

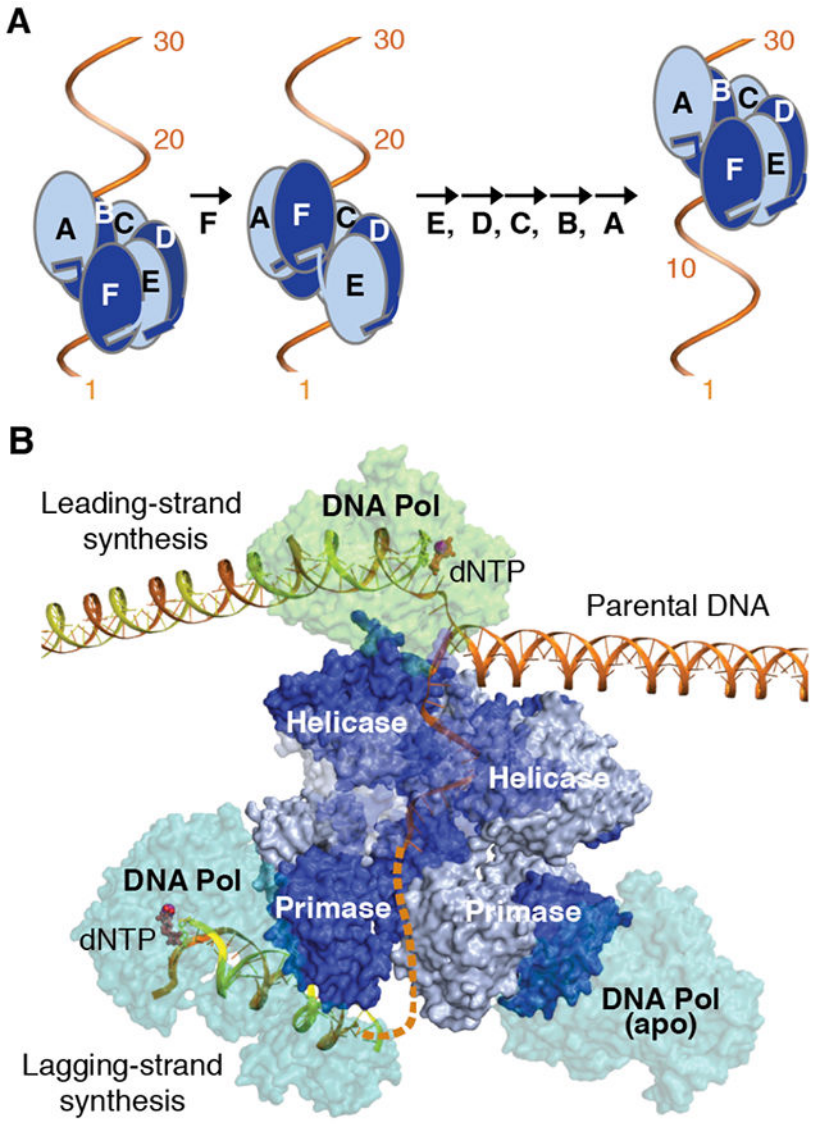
Data and materials availability: All data are available in the manuscript or the supplementary material.

Accession numbers: 10 in PDB, which are 6N7I, 6N7N, 6N7S, 6N7T, 6N7V,6N7W, 6N9U-X; 20 in EMDB, which are 0357, 0359, 0362 to 0365, 0379 to 0382, and 0386 to 0395.

DNA strand, with two nucleotides (nt) bound to each protein subunit and adjacent helicase subunits linked by domain swapping. ATP hydrolysis propels each helicase domain to translocate sequentially and coaxially along DNA in a hand-over-hand fashion, advancing 2 nt per step in the 5′ to 3′ direction (Fig. A). Instead of all enzymes moving in the same direction parallel to the downstream parental DNA, a β-hairpin from the leading-strand polymerase separates the two parental DNA strands into a T-shaped fork that enables the closely coupled helicase to unspool the downstream DNA tangentially (Fig. B). By protein-protein and DNA-mediated interactions, the leading-strand DNA polymerase and helicase cooperate to determine the rate of replication. For every ATP hydrolyzed and 2 nt advanced on DNA by the helicase, the DNA polymerase incorporates two deoxyribonucelotides. T7 primase, separated from the leading-strand polymerase by the helicase domain, synthesizes the RNA primers needed to initiate lagging-strand DNA synthesis. Transfer of a short RNA primer from the primase to DNA polymerase is facilitated by a zinc-binding-domain at the N-terminus of the T7 primase-helicase protein. Two lagging strand polymerases can be attached to the hexameric primases with one actively synthesizing DNA and the other waiting for a primer (Fig. B). Such a relay system may allow the discontinuous lagging-strand synthesis to keep pace with the leading-strand synthesis.

CONCLUSION—We note the similarity between hexameric DNA helicases and AAA+ protein chaperones and unfoldases, which form spiral-shaped hexamers around protein substrates, bind two amino-acid residues with each subunit and move proteins by a hand-over-hand subunit translocation mechanism. The operating principles of the bacteriophage replisome observed here rationalize many well-known features of bacterial and eukaryotic replication. Despite moving in opposite directions (3′ to 5′ on the leading strand in eukaryotes vs. 5′ to 3′ on the lagging strand in bacteria) and association with divergent polymerases, primase and accessory factors, in each replisome a helicase is the central organizer with leading- and lagging-strand DNA synthesis occurring on opposite sides of the helicase. The tight association of leading-strand DNA synthesis with movement of the replication fork sheds light on the effects of DNA lesions on replication and also the direct transfer of histones from parental DNA to the newly synthesized daughter strands in eukaryotes.

Graphical Abstract



Structure of T7 replisomes

(**A**) A Hand-over-hand mechanism for hexameric helicase translocating along ssDNA. (**B**) In the T7 replisome the leading-strand DNA polymerase and helicase establish a T-shaped DNA replication fork. Primase supplies RNA primers for lagging-strand DNA polymerases to make Okazaki fragments in a relay mechanism.

Abstract

The replisome that performs concerted leading and lagging DNA strand synthesis at a replication fork has never been visualized in atomic detail. Using bacteriophage T7 as a model system, we determined cryo-EM structures up to 3.2 Å of helicase translocating along DNA, and of helicase-polymerase-primase complexes engaging in synthesis of both DNA strands. Each domain of the spiral-shaped hexameric helicase translocates hand-over-hand sequentially along a ssDNA coil, akin to the way AAA+ ATPases unfold peptides. Two lagging-strand polymerases are attached to

the primase ready for Okazaki-fragment synthesis in tandem. A β-hairpin from the leading-strand polymerase separates two parental DNA strands into a T-shaped fork, thus enabling the closely coupled helicase to advance perpendicular to the downstream DNA duplex. These structures reveal the molecular organization and operating principles of a replisome.

One Sentence Summary:

Cryo-EM structures of T7 helicase translocating along DNA and T7 replisome carrying out concerted leading-and lagging-strand DNA synthesis.

DNA replication involves continuous copying of the leading strand and discontinuous synthesis of Okazaki fragments of the lagging strand (1, 2). The DNA replisome is a multi-protein machine that performs concerted parental strand separation and RNA-primed DNA synthesis on both strands. A helicase is central to all replisomes by coupling DNA unwinding to catalysis of leading-strand polymerase and lagging-strand primase (3–5). Lagging strand polymerases are often attached to the replisome at the beginning of Okazaki fragment synthesis (2). Leading- and lagging-strand synthesis, although organized differently, proceed at a similar speed during replication (6, 7). The replisome inevitably encounters a variety of obstacles, and repair and homologous recombination are required to rescue impeded replication forks (8, 9). Despite its importance, a replisome has never been visualized at atomic resolution.

Within a replisome, how the helicase separates a DNA double helix is also uncertain. All replicative helicases are homo- or hetero-hexameric ATPases that belong to RecA-like or AAA+ families and translocate either in the 5′ to 3′ or 3′ to 5′ direction (10). Although the ATPase sites are always at subunit interfaces, and single-stranded (ss) DNA-binding occurs in the central channel, the overall structure and DNA-binding mode differs among these helicases. Papillomavirus E1 helicase adopts a planar ring shape and encircles a narrow 6-nt coil in a fashion similar to *E. coli* Rho transcription terminator wrapping around RNA (11, 12). Although movement of DNA-binding loops lining the central channel of E1 helicase is hypothesized to drive DNA translocation, with or without DNA the E1 helicase structures including the spirally arranged DNA-binding loops are similar (11, 13). In contrast, bacterial DnaB helicase adopts a lock-washer shape surrounding a 14-nt coil, and each subunit may sequentially translocate from one end of the lock-washer to the other while unwinding DNA (14). Eukaryotic CMG helicase binds DNA similarly to DnaB but only contacts 6–8 nt in a half coil with a subset of six subunits, and how it operates is unclear (15, 16).

We chose to examine the simplest DNA replisome from bacteriophage T7 (1), in which DNA polymerase (gene product 5, gp5), together with a host processivity factor thioredoxin (Trx), carries out both leading- and lagging-strand synthesis, and the essential helicase and primase reside in one polypeptide (gene product 4, gp4). Like DnaB, T7 helicase is a RecA-like ATPase and translocates on the lagging strand, but it can use dTTP with improved processivity (17). Structures of T7 DNA polymerase-Trx-DNA-dGTP complexes and fragments of gp4 (primase including the zinc-binding-domain (ZBD), helicase alone, and primase with helicase) have been determined (18–21). Structures of gp4-gp5 complexes

have also been reported, but the DNA substrate was either absent (22) or invisible (23). Here we report snapshots of the T7 DNA replisome in action and its similarity to *E. coli* and yeast replisomes.

Cryo-EM structure of gp4-DNA complex

A 25-nt ssDNA with primase recognition site (PRS, 3′-CTGG-5′) (24) near its 5′-end was used as the substrate for both helicase and primase of gp4 (Fig. 1A). The E343Q mutant gp4, which is known to lack the NTP hydrolytic activity and have increased DNA affinity (25), readily formed a hexamer on the ssDNA in the presence of dTTP (Fig. 1B). After purification, the gp4-DNA- dTTP complex synthesized RNA primers when ATP, CTP and Mg^{2+} were supplied (Fig. 1C). Removal of the 3′-OH by using 3′-deoxy-CTP resulted in a terminated $pppAC_d$ dinucleotide. Subsequent addition of CTP did not yield longer product but stabilized the gp4-DNA/$pppAC_d$- CTP/dTTP complex or gp4-DNA complex in short (Fig. 1, C and D).

A cryo-EM structure of such gp4-DNA complex was determined at 3.2 Å resolution, in which 14 nt and five helicase domains (HelA to HelE) were well ordered, and thus called $gp4_5$-DNA (fig. S1). The primase (aa 65–250) and helicase (aa 264–549) form two tiers, the latter of which forms a lock washer wrapping around the right-handed ssDNA coil (Fig. 1E) with two nucleotides bound to each helicase domain as in the crystal structure of the DnaB-DNA complex (fig. S2 and S3) (14). Upon further 3D classification, the sixth helicase subunit (HelF) was found in three distinct conformations (gp4-DNA I-III, fig. S1A). In the gp4-DNA-I structure HelF bound 2-nt at the 5′ end trailing behind HelE, and in the gp4-DNA-III structure HelF was positioned forward by ~24 Å and bound to the 3′ end of DNA ahead of HelA (Fig. 1E and Fig. 2, A and B). The gp4-DNA-II structure with HelF in between HelA and HelE likely represents an intermediate state of translocation. The moving HelF illustrates a “hand-over-hand” translocation mechanism (Fig. 2C).

Individual helicase domains are superimposable with those in the ring-shaped hexameric and heptameric gp4 crystal structures (fig. S2C) (19, 21). The linker helix (αX, X for exchange) between the helicase and primase domains (Pri) is cyclically domain-swapped and links PriF to HelA, PriA to HelB, etc. (Fig. 1E and 2A), as observed in the apo structures (19, 21). Each primase domain is flexible relative to the helicase tier, resulting in poorly resolved structures (fig. S1, D and E). We docked six primase catalytic domains (20) into separate regions of density. The cryo- EM density corresponding to PriF, which binds the 5′-end of DNA, is significantly larger than for the other five regions (Fig. 1E). PRS/$pppAC_d$/CTP and an engaged ZBD of gp4 (see below) may account for the increased volume of PriF.

Mechanism for DNA unwinding by gp4

DNA-binding LoopD1 and LoopD2 from each helicase subunit sandwich the backbone of two nucleotides (Fig. 2, A and D). Residues R487, G488, and G490 from LoopD1 and K467 and N468 from LoopD2 contact the phosphosugar groups. Mutation of any of these residues reduces DNA binding (25–28). Each helicase domain interface harbors a dTTP·Mg^{2+}. Sidechains of Y535 and R504 sandwich the base of dTTP, as they do with adenine of ATP

(19), while the Walker A motif interacts with the triphosphates. The arginine finger (R522) supplied by the 3′ neighboring subunit contacts the γ-phosphate of dTTP. The Mg^{2+} is coordinated by the β- and γ-phosphate of dTTP and sidechains of S319 and E343Q. D424 from the Walker B motif possibly has water- mediated interactions with Mg^{2+}, as observed in UvrD and Rho helicase (12, 29). Interestingly, a nucleophilic water molecule is observed near the γ-phosphate of dTTP at the HelDE interface in $gp4_5$-DNA (Fig. 2D and fig. S2D) which is the next subunit to dissociate and propel HelE to translocate. The water is coordinated by the Q494 mainchain and H465 sidechain, which are adjacent to DNA-binding residues on Loop D1 and LoopD2, respectively. Mutation of H465 eliminates DNA stimulation of dTTPase activity (28). The gp4-DNA structure thus reveals a direct connection between DNA binding and NTP hydrolysis.

Gradual conformational changes are evident in the lock-washer gp4 hexamer. When six pairs of neighboring subunits (including HelEF and HelFA in gp4-DNA-I and III, respectively) are superimposed using the 3′ Hel domain, the 5′ subunits increasingly rotate anti-clockwise from the A to F subunits, thus making the NTPase active site tighter (Fig. 2E). Similar to the gradual closing of ATPase sites in an AAA+ protein translocase (30), the gradient of conformational change may control the sequential dTTP hydrolysis from 5′ to the 3′ end. Concurrently, HelA at the 3′ end has more extended DNA interface than HelF and HelE at the 5′ end (fig. S2F), thus facilitating helicase translocation.

The ssDNA complexed with gp4 helicase assumes a helical structure of 12-nt per turn, reminiscent of B-DNA. Distinct from B-DNA, however, the diameter of the gp4-bound ssDNA expands from 20 Å to 23 Å (Fig. 2F) and the average rise per base shrinks from 3.4 Å to 2.9 Å. The ssDNAs bound to DnaB and CMG helicase adopt a similar helical conformation (fig. S3) (14, 16). In contrast, the ssDNA in complex with E1, and ssRNA with Rho helicases, have a reduced diameter (less than 15 Å) and a short helical turn of 6 nt instead of 12 nt (11, 12). The similar helical structures of B-DNA and the ssDNAs bound by gp4, DnaB and CMG imply that these helicases can separate two intertwined DNA strands by translocating along one and excluding the other with little rotational movement of either protein or DNA. The B-like ssDNA structure also explains why T7 gp4 and bacterial DnaB can bind and translocate along dsDNA when loaded via a 5′ overhang (31, 32). These helicases likely translocate along one strand of the duplex in the preferred orientation (5′ to 3′) and pull the strand slightly away from the helical axis as they do with ssDNA, thus causing local strand separation and partial duplex melting (fig. S2, G to I).

Our structures suggest a sequential “hand-over-hand” translocation of each helicase domain from 5′ to 3′ along the ssDNA coil (Fig. 2c). Consistent with published data (17, 33), each domain movement is driven by one dTTP hydrolysis, which allows the Hel domain at the 5′ end to reposition to the 3′ end and acquire a new dTTP at the newly formed domain interface. While the cyclically domain-swapped linkers (αX) hold the hexamer intact, gp4 advances by 2 nt for each dTTP hydrolyzed. Interestingly, both 1- and 2-nt step sizes have been reported for T7 helicase (34–36). For a complete translocation cycle, a gp4 hexamer translocates over 12 nt in six steps (movie 1). The proposed translocation mechanism for DnaB helicase is rather similar to the “hand- over-hand” model depicted here, although the gradient of conformational changes and movement of HelF were not observed in the DnaB

crystal structure (14). Eukaryotic CMG, which binds ssDNA similarly to gp4 and DnaB (fig. 3D), contacts only 6–8 nt with 3–4 subunits (MCM6, 2, 5, 3 (16) or MCM6, 4, 7 (15)) and translocates in the reverse direction, 3′ to 5′. The direction of helicase movement can be readily switched by reversing the order of sequential ATP hydrolysis and the direction of helicase domain movement (37). In the case of CMG, we suspect that instead of one subunit translocating 2 nt per step, two or three MCM subunits may translocate *en bloc*, and a full turn of translocation over 12 nt may take 3 or 2 steps, respectively (38). Thereby not all six ATPase sites of CMG need to hydrolyze ATP for the translocation (39).

Attachment of lagging-strand polymerase to gp4 primase

We designed a 44-nt ssDNA and short DNA/RNA hybrid fusion substrate to capture initiation of lagging-strand DNA synthesis (Fig. 3A). Four cryo-EM structures of gp4-gp5-DNA/RNA complexes (LagS1–4) with resolution mostly at 4 Å were obtained, and using focused refinement the gp5 polymerase-primase-DNA/RNA portion (LagS1-gp5-DNA) reached 3.7 Å resolution (fig. S4). The gp4 hexamer surrounded the ssDNA coil as in the gp4-DNA-I structure, so the nomenclature of A-F subunit from the 3′ to 5′ end remains applicable. In all four structures, a polymerase-DNA/RNA-dTTP ternary complex was attached to two adjacent primase domains, EF, AB or BC, in a similar fashion (fig. S5), thus a maximum of three lagging-strand polymerases may be attached to gp4 simultaneously (40). Indeed, two DNA polymerases were found in the LagS4 structure. One attached to PriAB contained the DNA/RNA substrate, and the other attached to PriCD had no substrate (fig. S5E), as if waiting for the next primer to arrive. Two DNA polymerases are also observed adjacent to the primase domains in the 13.8 Å gp4-gp5 structure without DNA (23), but they were proposed to engage in leading and lagging strand synthesis, respectively.

In the 3.7 Å LagS1-gp5-DNA structure (Fig. 3, B and C), the gp5 polymerase interacts with PriE and PriF, the latter of which was engaged in primer synthesis in the absence of gp5 polymerase (Fig. 1E). PriE formed extensive polar and hydrophobic contacts with Helix H on the Thumb domain of polymerase (Fig. 3D), burying 969 $Å^2$ at each interface. The interactions between polymerase and PriF were less extensive and involve the tip of Helix E′ of polymerase (Exo domain) and R77 and T82 from PriF (Fig. 3E). The gp4-gp5 interactions stabilized both primase domains (fig. S4 and S5B), and the flexible hinges between the primase and helicase (aa 255–265) became fully traceable. Similar association of polymerase with two contiguous primase domains (EF, AB and BC) and the flexible hinge in gp4 (fig. S5G) indicate that the lagging-strand polymerase may attach to the same pair of primase domains while helicase domains continuously translocate with F subunit becoming the new A, then the new B, and so forth, which corroborates the single molecule observations (7, 40).

While the 6-bp DNA/RNA hybrid and an incoming dTTP engage the polymerase active site (Fig. 3F), a ZBD at the N-terminus of gp4 (aa 1–44), which is linked to the primase domain by twenty disordered residues over 40 Å (fig. S5B), buttresses the 5′ end of RNA primer and is fully embraced by the polymerase. The sidechain of Y37 of the ZBD stacks against the first DNA/RNA base pair (dT-rA), and the downstream ssDNA is bent by ~40°. The H14 sidechain is also within ~3.5 Å of the dT-rA pair. The first nucleotide in PRS, dC, is

sequestered in a pocket and contacted by H33 (Fig. 3F). Mutations of H14 or Y37 severely hinder primer synthesis by gp4 and extension by gp5 polymerase (41). An H33A mutation alters the PRS preference from 3′-CTGG-5′ to 3′- A/GTGG-5′ (42). The structures of gp4-gp5-DNA/RNA complexes reveal not only how the ZBD binds PRS and stabilizes the short DNA/RNA hybrid, but how the flexible ZBD may act as a DNA mimic to facilitate primer synthesis, handoff, and extension without impeding helicase movement (41, 42). In parallel, the flexible FeS domain in the eukaryotic primase also binds to the 5′-end of RNA primer and enables its synthesis and delivery to DNA polymerase (43). Similar primase-polymerase attachment may also occur at the beginning and re-priming of leading-strand synthesis.

Concerted DNA unwinding and leading-strand synthesis

A DNA fork with a dideoxy 3′-end at the 25-nt leading-strand primer was made to form a replisome with gp4 helicase-primase and gp5 polymerase for cryo-EM analysis (Fig. 4A). Several different complexes were generated: gp4 or gp5 alone on DNA, and gp4-gp5-DNA complexes with different stoichiometries and configurations (fig. S6). After 3D classification, we obtained five structures of leading-strand complexes (Lead1–5) at resolutions ranging from 9 to 14 Å. These structures contain one each of gp4 hexamer, gp5 polymerase and DNA fork, but polymerase contacts helicase domains in five different orientations (fig. S7). Separate local refinement of gp4 and gp5 yielded a 4.5 Å gp5-Trx-DNA-dTTP·Mg^{2+} structure, which resembles crystal structures of the polymerase quaternary complexes (18), and a 3.8 Å Lead-gp4-DNA structure, which is identical to gp4-DNA-I except for two additional unpaired nucleotides beyond HelA and before the parental duplex (fig. S6, S8, A and B, and Fig. 4, B and C).

Eight base pairs of the parental DNA were readily traceable along with the leading-strand polymerase (Fig. 4D and fig. S8C). The first base pair at the fork (A-T) is partially separated and stacked with the sidechain of W579 on a b-hairpin from the polymerase (aa 575–585). The β-hairpin, which is conserved among many bacteriophages (fig. S8D), is disordered without the DNA fork (18) and functions like the separation pin in SF1 helicases (Fig. 4D) (29). At the fork, the leading-strand template enters DNA polymerase leaving only one unpaired nucleotide, as depicted by previous analysis of T7 replisome (4). Some differences between the cryo-EM and crystal structures of Pol-DNA complexes (18) are observed, particularly in regions near crystal lattice contacts (fig. S8, E to G). In the cryo-EM structure, LoopDE′ (residues 100–132) in the Exo domain is detached from DNA, and the distal end of dsDNA was shifted as a rigid body together with Trx by ~10 Å towards gp4 helicase.

The leading-strand complexes (Fig. 4, B and C) were readily modeled by docking the well-resolved gp4 and gp5 structures with minor adjustments around the DNA fork (fig. S8H). The parental DNA sits between the helicase and the polymerase. At the T-shaped fork, the two newly separated strands extend in opposite directions, one threading through the helicase central channel and the other entering the polymerase (Fig. 4C). Because DNA polymerase advances 1-nt at a time while helicase takes a 2-nt step, separation between the top of the helicase and parental DNA may oscillate between 1 and 3 nt while the polymerase

remains 1 nt away from the downstream duplex. The two proteins are connected through direct contacts. In Lead1 complex (fig. S7) the acidic C-terminal tails (C-tails) from HelA and HelE (spanning two ends of the lock washer as in gp4-DNA-I) contact two basic patches on polymerase, one near the Trx-binding domain (known as TBDbp (44)) and another in the front adjacent to the parental DNA and separation pin (known as Fbp (45)) (Fig. 4, E to G). Although not visualized without DNA (22, 23), these two basic patches and the acidic tail have been well established to be involved in gp4-gp5 interactions and the efficiency and processivity of the T7 replisome (44, 45). Such charge-charge interactions appear dynamic and allow partner switch (HelE in Lead1–3 vs. HelA or HelB in Lead4–5) and up to 180° rotation of helicase relative to the polymerase and parental DNA with the interface of TBDbp detached in Lead2–5 structures (fig. S7 and movie 2). However, the Fbp of polymerase remains associated with an acidic tail of helicase in all five complexes, which may stabilize the fork (45).

The cryo-EM structures reported here reveal how helicase and polymerase work in concert at a DNA fork. The two cooperatively move the fork forward by translocating on opposite strands in opposite directions tangential to the parental DNA helix. DNA synthesis accelerates DNA unwinding by the helicase (46), and helicase increases dNTP binding by the polymerase (4). Moreover, DNA polymerase provides the separation pin to physically divide DNA duplex (Fig. 4) and orients the helicase for DNA unspooling. In contrast to the popular speculation that helicase translocates along the parental helical axis and forcibly separates two intertwined strands, pulling one strand perpendicular to the helical axis, which has also been observed with monomeric DNA helicases (29) and RNA transcriptome (47–49), may be the most efficient way of unspooling the double helix. Together they unwind DNA 8-fold more efficiently than gp4 alone, making unwinding processive and independent of CG-content (4, 17), which also explains the coupling between leading-strand DNA polymerase and lagging-strand helicase observed in E. coli replisome (7). In the presence of a DNA lesion or roadblock, the two proteins may temporarily dissociate and thus result in reduced fork progression. The extended ssDNA generated by helicase and slowing of the replisome serve as signals for DNA repair and also allow re-priming (8, 9). After lesion bypass and re-priming, helicase and polymerase can reattach and form the replisome again.

Replisome architecture

Four lagging-strand complexes (LagL1–4) were also obtained using the same DNA fork substrate (Fig. 4A and fig. S6 and S9). Like the complexes formed on the short RNA/DNA hybrid (LagS structures, Fig. 3 and fig. S5), the polymerase complexed with 25-bp DNA and dTTP·Mg^{2+} was attached to adjacent primase domains AB, BC, CD or EF. With the long DNA substrate, the processivity factor Trx became ordered, and the Helix H-primase interaction was strengthened (fig. S9, B to D). The LagL complexes were probably formed by the two proteins each bound to a different DNA fork substrate (fig. S9A). In fact, within the same sample preparation, complete replisome particles with the leading-strand polymerase contacting the helicase domains and one or two lagging-strand polymerases on the primase side were found in cryo-EM micrographs (Fig. 5A).

Because of limited particle number and an almost unlimited variety of combinations of helicase-leading and primase-lagging strand associations, we have not obtained a well-resolved *ab initio* replisome structure. However, guided by the 2D classification results (Fig. 5A), a replisome model can readily be built by superposition of the conserved gp4 helicase portions of Lead1, LagL1 and LagS4 structures (Fig. 5B and movie 3). In this replisome model, gp4 is sandwiched between leading- and lagging-strand polymerases. The replication fork is held between the leading-strand polymerase and helicase, and cooperatively pushed forward by the two proteins. Lagging-strand DNA that emerges from the helicase central channel is captured first by primase and then transferred to lagging-strand polymerases. Binding of more than one polymerase to the primase side ensures parallel synthesis of Okazaki fragments (Fig. 5C) and similar speed of leading and lagging strand synthesis.

The organizing principles governing the T7 replisome is consistent with the details ascribed to bacterial DNA replication (7) even though in bacteria the helicase-primase association is intermolecular and mediated by the Tau protein (3). In eukaryotic replisomes, CMG, a heterohexameric helicase in the 11-subunit complex, translocates along the leading strand instead of the lagging strand. As a result, the DNA fork must be positioned between helicase and primase / lagging-strand polymerases (Fig. 5D) (2). Nevertheless, the CMG helicase remains the central organizer and contacts the leading-strand polymerase (Pol ε) on one side and primase on the other (5). As both CMG and Pol ε move along the leading strand DNA, the replisome may ignore and skip over lesions on the lagging strand. Recent data suggest that CMG helps DNA polymerase to traverse lesions and roadblocks, and together they keep the replication-fork moving forward (50).

Conclusion

Here we present the atomic structures of a primordial DNA replisome, which illustrate how a replicative helicase translocates in a direction perpendicular to the DNA fork and how leading- and lagging-strand DNA synthesis occur in the front and back side of the helicase. flexible connections observed among replisome components depict a highly dynamic organization yet with operating principles conserved in bacterial and eukaryotic replisomes, which not only replicate genomic DNA, but also detect DNA lesions and coordinate stress responses. The detailed replisome structure explains many previously published results and provides a basis understanding coordination among DNA replication, recombination and repair. In closing, we note the similarity between hexameric DNA helicases (movie 1) and AAA+ protein chaperones unfoldase, including Hsp101, Hsp104, p97, ClpB, Vps4, Yme1 and 26S proteasome (fig. S10), which form spiral lock-washer-like hexamers around protein substrates, bind two amino-acid residues with each subunit and move proteins by a hand-over-hand subunit translocation mechanism (30, 51–56).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank M. Gellert. and D. Leahy for critical reading of the manuscript, C. Lima for the help and protocol of local refinement, T. Grant and N. Grigorieff for helping with initial cryoEM data analysis, and J. Jiang for helping with FSC calculation.

Funding: This research was supported by the National Institute of Diabetes and Digestive and Kidney Diseases to W. Y. (DK036146) and National Cancer Institute to N. d. V. The authors acknowledge the use of instruments at the Electron Imaging Center for NanoMachines supported by NIH (1S10RR23057, 1S10OD018111 and GM071940), NSF (DBI-1338135) and CNSI at UCLA.

References

1. Hamdan SM, Richardson CC, Motors, switches, and contacts in the replisome. Annu Rev Biochem 78, 205–243 (2009). [PubMed: 19298182]
2. O'Donnell M, Langston L, Stillman B, Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring Harb Perspect Biol 5, (2013).
3. Kim S, Dallmann HG, McHenry CS, Marians KJ, Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement. Cell 84, 643–650 (1996). [PubMed: 8598050]
4. Nandakumar D, Pandey M, Patel SS, Cooperative base pair melting by helicase and polymerase positioned one nucleotide from each other. Elife 4, (2015).
5. Sun J et al., The architecture of a eukaryotic replisome. Nat Struct Mol Biol 22, 976–982 (2015). [PubMed: 26524492]
6. Lee J, Chastain PD, 2nd, T. Kusakabe, J. D. Griffith, C. C. Richardson, Coordinated leading and lagging strand DNA synthesis on a minicircular template. Mol Cell 1, 1001–1010 (1998). [PubMed: 9651583]
7. Graham JE, Marians KJ, Kowalczykowski SC, Independent and Stochastic Action of DNA Polymerases in the Replisome. Cell 169, 1201–1213 e1217 (2017). [PubMed: 28622507]
8. Marians KJ, Lesion Bypass and the Reactivation of Stalled Replication Forks. Annu Rev Biochem 87, 217–238 (2018). [PubMed: 29298091]
9. Branzei D, Foiani M, Maintaining genome stability at the replication fork. Nat Rev Mol Cell Biol 11, 208–219 (2010). [PubMed: 20177396]
10. O'Donnell ME, Li H, The ring-shaped hexameric helicases that function at DNA replication forks. Nat Struct Mol Biol 25, 122–130 (2018). [PubMed: 29379175]
11. Enemark EJ, Joshua-Tor L, Mechanism of DNA translocation in a replicative hexameric helicase. Nature 442, 270–275 (2006). [PubMed: 16855583]
12. Thomsen ND, Berger JM, Running in reverse: the structural basis for translocation polarity in hexameric helicases. Cell 139, 523–534 (2009). [PubMed: 19879839]
13. Sanders CM et al., Papillomavirus E1 helicase assembly maintains an asymmetric state in the absence of DNA and nucleotide cofactors. Nucleic Acids Res 35, 6451–6457 (2007). [PubMed: 17881379]
14. Itsathitphaisarn O, Wing RA, Eliason WK, Wang J, Steitz TA, The hexameric helicase DnaB adopts a nonplanar conformation during translocation. Cell 151, 267–277 (2012). [PubMed: 23022319]
15. Abid Ali F et al., Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. Nat Commun 7, 10708 (2016). [PubMed: 26888060]
16. Georgescu R et al., Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. Proc Natl Acad Sci U S A 114, E697–E706 (2017). [PubMed: 28096349]
17. Sun B et al., ATP-induced helicase slippage reveals highly coordinated subunits. Nature 478, 132–135 (2011). [PubMed: 21927003]
18. Doublie S, Tabor S, Long AM, Richardson CC, Ellenberger T, Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. Nature 391, 251–258 (1998). [PubMed: 9440688]

19. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB, Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. Cell 101, 589–600 (2000). [PubMed: 10892646]
20. Kato M, Ito T, Wagner G, Richardson CC, Ellenberger T, Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. Mol Cell 11, 1349–1360 (2003). [PubMed: 12769857]
21. Toth EA, Li Y, Sawaya MR, Cheng Y, Ellenberger T, The crystal structure of the bifunctional primase-helicase of bacteriophage T7. Mol Cell 12, 1113–1123 (2003). [PubMed: 14636571]
22. Wallen JR et al., Hybrid Methods Reveal Multiple Flexibly Linked DNA Polymerases within the Bacteriophage T7 Replisome. Structure 25, 157–166 (2017). [PubMed: 28052235]
23. Kulczyk AW, Moeller A, Meyer P, Sliz P, Richardson CC, Cryo-EM structure of the replisome reveals multiple interactions coordinating DNA synthesis. Proc Natl Acad Sci U S A 114, E1848–E1856 (2017). [PubMed: 28223502]
24. Mendelman LV, Richardson CC, Requirements for primer synthesis by bacteriophage T7 63-kDa gene 4 protein. Roles of template sequence and T7 56-kDa gene 4 protein. J Biol Chem 266, 23240–23250 (1991). [PubMed: 1744119]
25. Crampton DJ, Mukherjee S, Richardson CC, DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase. Mol Cell 21, 165–174 (2006). [PubMed: 16427007]
26. Washington MT, Rosenberg AH, Griffin K, Studier FW, Patel SS, Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. J Biol Chem 271, 26825–26834 (1996). [PubMed: 8900164]
27. Lee SJ, Qimron U, Richardson CC, Communication between subunits critical to DNA binding by hexameric helicase of bacteriophage T7. Proc Natl Acad Sci U S A 105, 8908–8913 (2008). [PubMed: 18574147]
28. Satapathy AK et al., Residues in the central beta-hairpin of the DNA helicase of bacteriophage T7 are important in DNA unwinding. Proc Natl Acad Sci U S A 107, 6782–6787 (2010). [PubMed: 20351255]
29. Lee JY, Yang W, UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. Cell 127, 1349–1360 (2006). [PubMed: 17190599]
30. Han H, Monroe N, Sundquist WI, Shen PS, Hill CP, The AAA ATPase Vps4 binds ESCRT-III substrates through a repeating array of dipeptide-binding pockets. Elife 6, (2017).
31. Kaplan DL, The 3'-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase. J Mol Biol 301, 285–299 (2000). [PubMed: 10926510]
32. Rasnik I et al., Branch migration enzyme as a Brownian ratchet. EMBO J 27, 1727–1735 (2008). [PubMed: 18511910]
33. Liao JC, Jeong YJ, Kim DE, Patel SS, Oster G, Mechanochemistry of t7 DNA helicase. J Mol Biol 350, 452–475 (2005). [PubMed: 15950239]
34. Kim DE, Narayan M, Patel SS, T7 DNA helicase: a molecular motor that processively and unidirectionally translocates along single-stranded DNA. J Mol Biol 321, 807–819 (2002). [PubMed: 12206763]
35. Johnson DS, Bai L, Smith BY, Patel SS, Wang MD, Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. Cell 129, 1299–1309 (2007). [PubMed: 17604719]
36. Pandey M, Patel SS, Helicase and polymerase move together close to the fork junction and copy DNA in one-nucleotide steps. Cell Rep 6, 1129–1138 (2014). [PubMed: 24630996]
37. Saikrishnan K, Powell B, Cook NJ, Webb MR, Wigley DB, Mechanistic basis of 5'−3' translocation in SF1B helicases. Cell 137, 849–859 (2009). [PubMed: 19490894]
38. Li H, O'Donnell ME, The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. Bioessays 40, (2018).
39. Ilves I, Petojevic T, Pesavento JJ, Botchan MR, Activation of the MCM2–7 helicase by association with Cdc45 and GINS proteins. Mol Cell 37, 247–258 (2010). [PubMed: 20122406]

40. Duderstadt KE et al., Simultaneous Real-Time Imaging of Leading and Lagging Strand Synthesis Reveals the Coordination Dynamics of Single Replisomes. Mol Cell 64, 1035–1047 (2016). [PubMed: 27889453]
41. Kato M, Ito T, Wagner G, Ellenberger T, A molecular handoff between bacteriophage T7 DNA primase and T7 DNA polymerase initiates DNA synthesis. J Biol Chem 279, 30554–30562 (2004). [PubMed: 15133047]
42. Kusakabe T, Hine AV, Hyberts SG, Richardson CC, The Cys4 zinc finger of bacteriophage T7 primase in sequence-specific single-stranded DNA recognition. Proc Natl Acad Sci U S A 96, 4295–4300 (1999). [PubMed: 10200256]
43. Baranovskiy AG et al., Mechanism of Concerted RNA-DNA Primer Synthesis by the Human Primosome. J Biol Chem 291, 10006–10020 (2016). [PubMed: 26975377]
44. Hamdan SM et al., A unique loop in T7 DNA polymerase mediates the binding of helicase-primase, DNA binding protein, and processivity factor. Proc Natl Acad Sci U S A 102, 5096–5101 (2005). [PubMed: 15795374]
45. Zhang H et al., Helicase-DNA polymerase interaction is critical to initiate leading-strand DNA synthesis. Proc Natl Acad Sci U S A 108, 9372–9377 (2011). [PubMed: 21606333]
46. Stano NM et al., DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase. Nature 435, 370–373 (2005). [PubMed: 15902262]
47. Yin YW, Steitz TA, Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science 298, 1387–1395 (2002). [PubMed: 12242451]
48. Liu B, Hong C, Huang RK, Yu Z, Steitz TA, Structural basis of bacterial transcription activation. Science 358, 947–951 (2017). [PubMed: 29146813]
49. Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD, Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell 127, 941–954 (2006). [PubMed: 17129781]
50. Sparks JL et al., The CMG helicase bypasses DNA protein cross-links to facilitate their repair. bioRxiv, (2018).
51. Ho CM et al., Malaria parasite translocon structure and mechanism of effector export. Nature 561, 70–75 (2018). [PubMed: 30150771]
52. Gates SN et al., Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science 357, 273–279 (2017). [PubMed: 28619716]
53. Ripstein ZA, Huang R, Augustyniak R, Kay LE, Rubinstein JL, Structure of a AAA+ unfoldase in the process of unfolding substrate. Elife 6, (2017).
54. Yu H et al., ATP hydrolysis-coupled peptide translocation mechanism of Mycobacterium tuberculosis ClpB. Proc Natl Acad Sci U S A 115, E9560–E9569 (2018). [PubMed: 30257943]
55. de la Pena AH, Goodall EA, Gates SN, Lander GC, Martin A, Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. Science 362, (2018).
56. Puchades C et al., Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. Science 358, (2017).
57. Zheng SQ et al., MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods 14, 331–332 (2017). [PubMed: 28250466]
58. Zhang K, Gctf: Real-time CTF determination and correction. J Struct Biol 193, 1–12 (2016). [PubMed: 26592709]
59. Heymann JB, Single particle reconstruction and validation using Bsoft for the map challenge. J Struct Biol, (2018).
60. Kimanius D, Forsberg BO, Scheres SH, Lindahl E, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife 5, (2016).
61. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods 14, 290–296 (2017). [PubMed: 28165473]
62. Kucukelbir A, Sigworth FJ, Tagare HD, Quantifying the local resolution of cryo-EM density maps. Nat Methods 11, 63–65 (2014). [PubMed: 24213166]
63. Brieba LG et al., Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. EMBO J 23, 3452–3461 (2004). [PubMed: 15297882]

64. Emsley P, Lohkamp B, Scott WG, Cowtan K, Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66, 486–501 (2010). [PubMed: 20383002]

65. Goddard TD, Huang CC, Ferrin TE, Visualizing density maps with UCSF Chimera. J Struct Biol 157, 281–287 (2007). [PubMed: 16963278]

66. Adams PD et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213–221 (2010). [PubMed: 20124702]

67. Thomsen ND, Lawson MR, Witkowsky LB, Qu S, Berger JM, Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-dependent hexameric motor. Proc Natl Acad Sci U S A 113, E7691–E7700 (2016). [PubMed: 27856760]

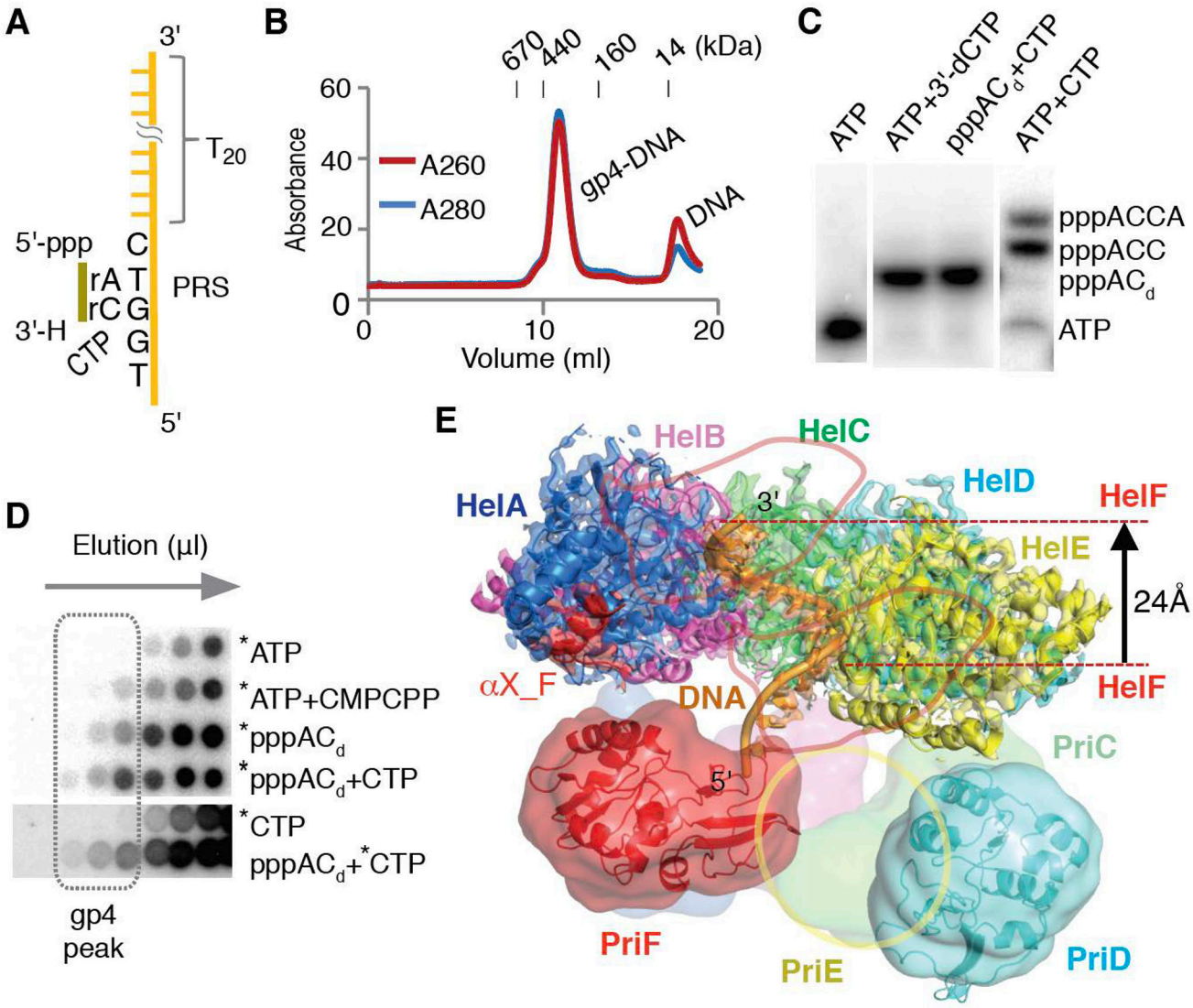


Figure 1. Assembly and structure of the gp4-DNA complex.

(**A**) Substrate design. (**B**) Gel filtration profile of the gp4-DNA complex. (**C**) The gp4 primase activity was assayed with ^{32}P-ATP and 3′-dCTP and chased with CTP, or with ^{32}P-ATP and CTP. The products are labeled on the side of the gel. (**D**) Analysis of gp4-DNA complexes by size-exclusion and dot-blot assays. The gp4 fractions are boxed. Different ^{32}P-labeled components are indicated on the side. (**E**) Cryo-EM structure of the gp4$_5$-DNA complex. A, B, C, D, E and F subunits and DNA are color-coded. The semi-transparent cryo-EM densities are shown without any filter for helicase domains, and low-pass-filtered to 8 Å for primase domains. The translocation of HelF is indicated by the semi-transparent red outlines of its beginning (gp4-DNA-I) and end (gp4-DNA-III) position. PriE is diagramed as a yellow outline.

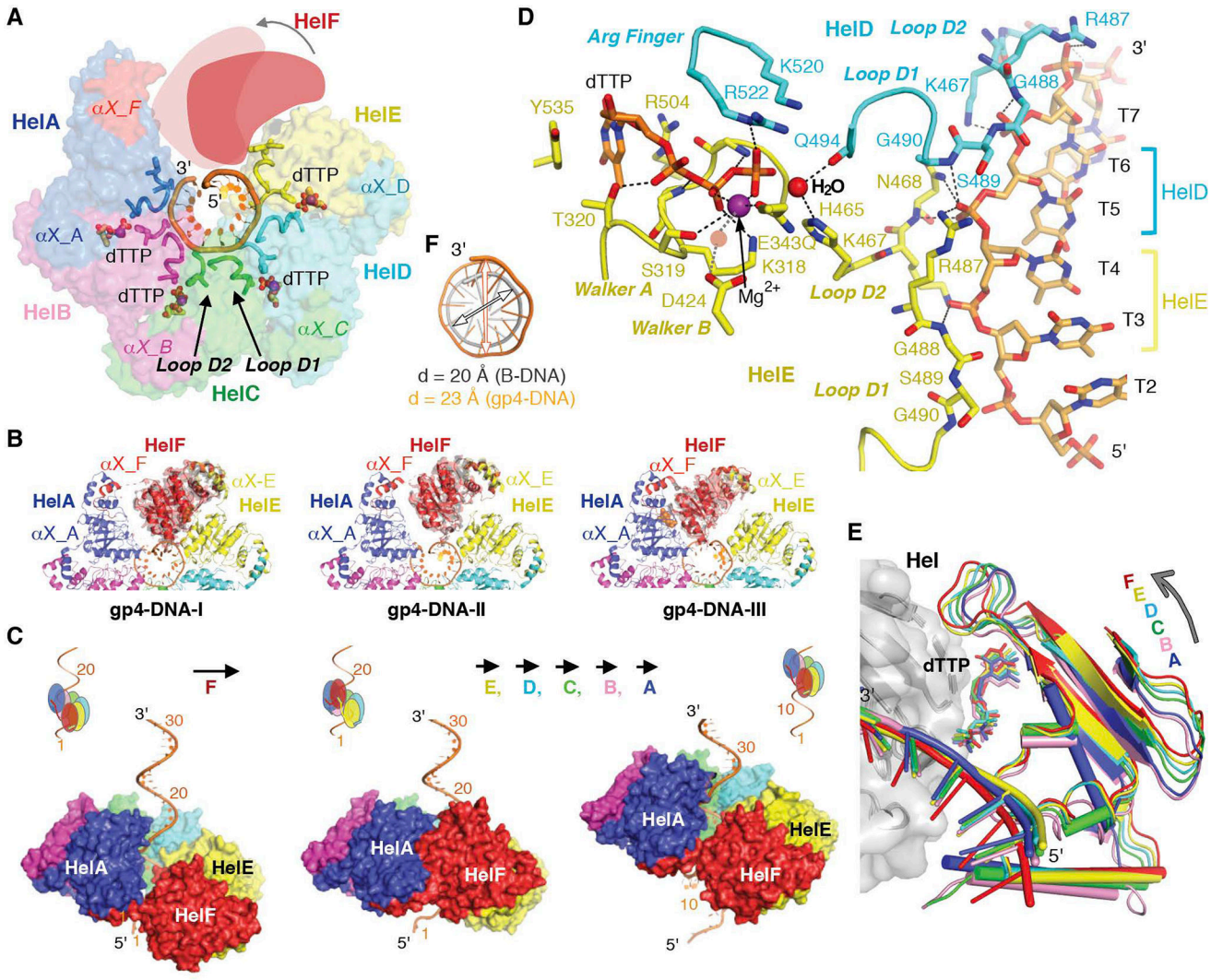


Figure 2. Mechanism of helicase translocation.

(**A**) Gp4-DNA complex (semi-transparent surface) is viewed from the 5′-end of DNA. The beginning (dark red) and end position of HelF (pale red) are shown. (**B**) Translocation of HelF. Structures of gp4-DNA-I to III are shown from left to right. The cryo-EM densities for HelF and the linker of HelE are shown as semi-transparent surface. (**C**) Diagrams of helicase translocation. dTTP hydrolysis and release from HelEF interface releases HelF at the 5′ end. After HelF translocation toward the 3′ end, the newly formed HelFA interface acquires a fresh dTTP. Sequential translocations of E to A subunits each toward the 3′ end complete a helical turn of 12 nucleotides. (**D**) NTPase active site at the HelDE interface and its connection to DNA binding by LoopD1 and LoopD2 are diagrammed. (**E**) Gradual changes at the helicase subunit interface. Hel dimers are superimposed by their 3¢ subunits (shown as grey surface and ribbons), and the 5′ subunits are color-coded as in (A). (**F**) Superposition of ssDNA complexed with gp4 (orange) and B-form DNA (cyan).

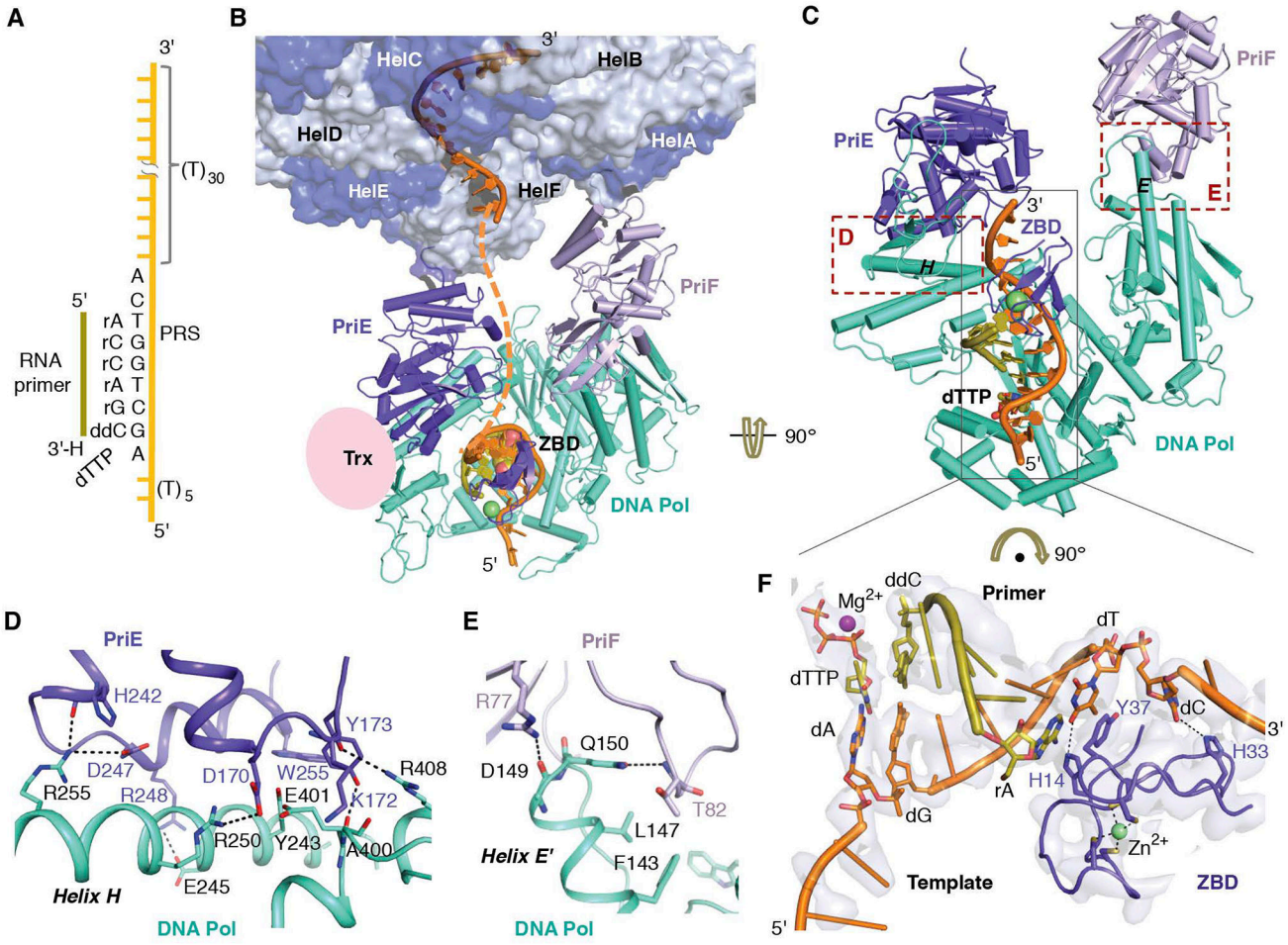


Figure 3. Structure of the lagging-strand gp4-gp5 complex.

(**A**) The DNA/RNA substrate. (**B**) Structure of gp4 (alternate dark and light molecular surface for helicase domains and cartoon for primases), gp5 polymerase (green cartoon) and DNA (orange and yellow). (**C**) A zoom-in view of the gp4 primase, gp5 polymerase, DNA/RNA hybrid and dTTP complex. (**D** and **E**) The zoom-in views of polymerase and primase interfaces that are boxed in (C). The possible interactions are shown as dashed lines. (**F**) The DNA template (PRS) and the 5′-end of RNA primer are stabilized by the gp4 ZBD. Cryo-EM density is superimposed as grey semi-transparent surface. Interactions between ZBD and DNA/RNA are indicated by dotted lines.

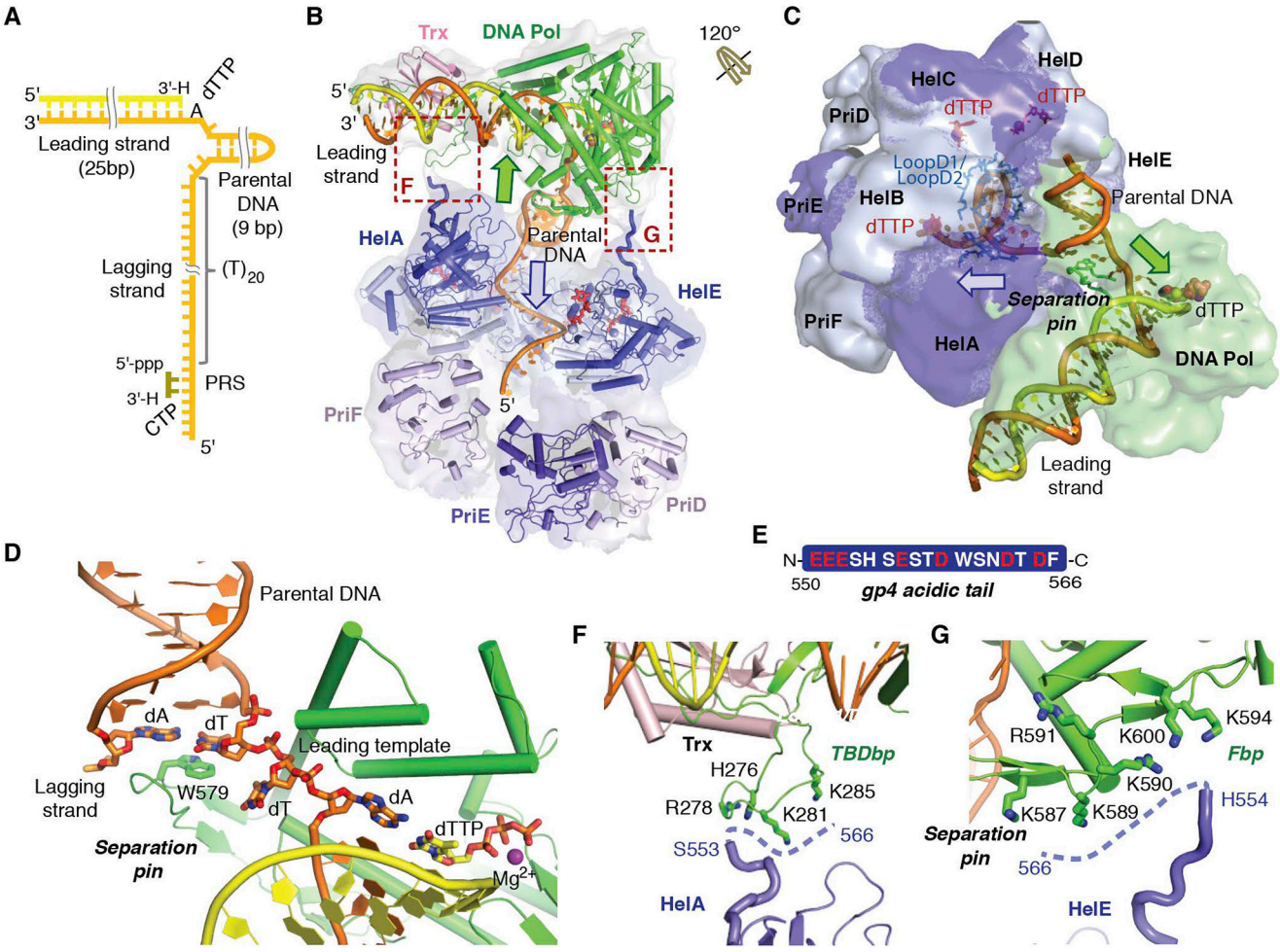


Figure 4. Structure of the leading-strand gp4-gp5 complex.

(**A**) The DNA fork substrate. (**B** and **C**) The gp4-gp5-DNA fork structure. Cartoon of polymerase (green), Trx (light pink), and helicase (alternate dark and light blue) are shown with cryo-EM density (semi-transparent). The cartoons of HelF in panel B and of all proteins in panel C are omitted for clarity. dTTP and DNA-binding loops are shown as red and blue sticks, respectively in C. Green and blue arrows indicate the directions of DNA unspooling by polymerase and helicase. (**D**) A zoom-in view of the DNA fork bound by the leading-strand polymerase. Nucleotides at the fork are shown as sticks. The separation pin and W579 stack with the first base-pair of the parental DNA. (**E**) Sequence of the gp4 acidic tail. (**F** and **G**) Zoom-in views of the gp4 primase and gp5 polymerase interfaces as boxed in B.

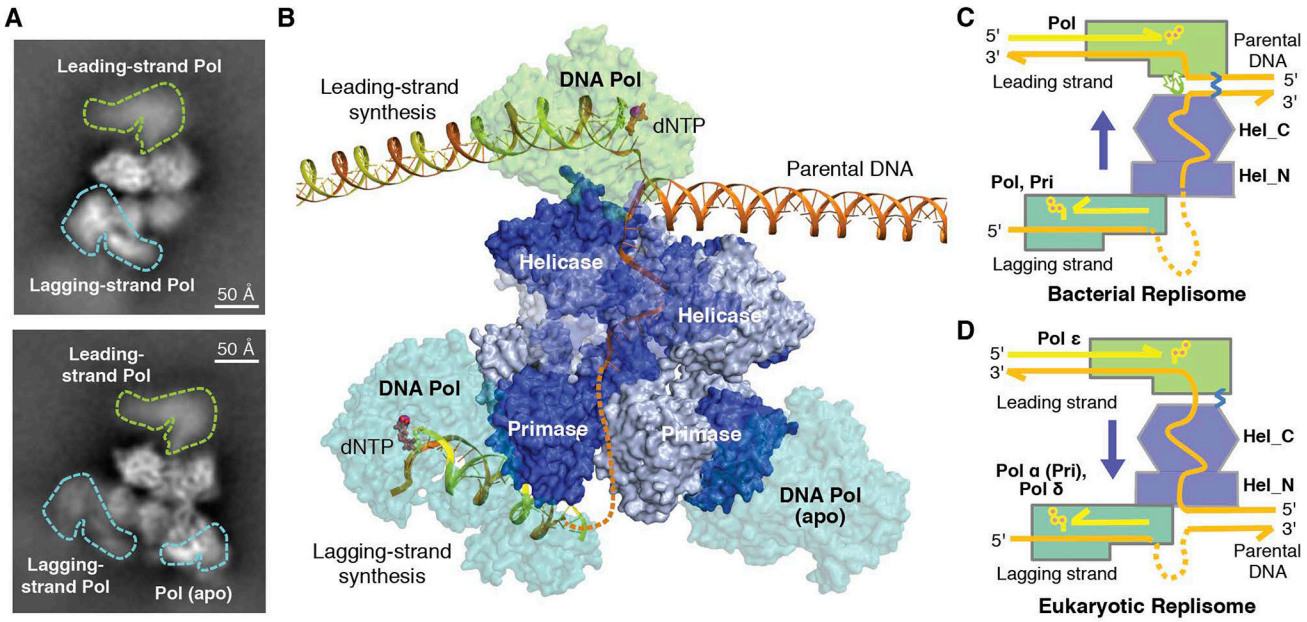


Figure 5. Architecture of T7 replisome.

(**A**) Representative 2D-classified replisome structures. Polymerases are outlined in green and cyan. (**B**) A replisome model composed of leading- (Lead1) and lagging-strand complex structures (LagL1 and LagS4) is obtained by superimposition of the conserved gp4 helicase domains. Proteins are shown as semi-transparent surface, and DNA as cartoon. Parental DNA and leading-strand product are extended artificially with B-DNA. (**C** and **D**) Diagrams of bacterial and eukaryotic replisome with helicase on the lagging strand (C) and leading strand (D), respectively. In each case the direction of helicase translocation is marked by a blue arrow.