LETTERS

DNA primase acts as a molecular brake in DNA replication

Jong-Bong Lee[1], Richard K. Hite[2], Samir M. Hamdan[1], X. Sunney Xie[3], Charles C. Richardson[1] & Antoine M. van Oijen[1]

A hallmark feature of DNA replication is the coordination between the continuous polymerization of nucleotides on the leading strand and the discontinuous synthesis of DNA on the lagging strand[1]. This synchronization requires a precisely timed series of enzymatic steps that control the synthesis of an RNA primer, the recycling of the lagging-strand DNA polymerase, and the production of an Okazaki fragment. Primases synthesize RNA primers at a rate that is orders of magnitude lower[2–4] than the rate of DNA synthesis by the DNA polymerases at the fork. Furthermore, the recycling of the lagging-strand DNA polymerase from a finished Okazaki fragment to a new primer is inherently slower than the rate of nucleotide polymerization[5]. Different models have been put forward to explain how these slow enzymatic steps can take place at the lagging strand without losing coordination with the continuous and fast leading-strand synthesis[6–8]. Nonetheless, a clear picture remains elusive. Here we use single-molecule techniques to study the kinetics of a multiprotein replication complex from bacteriophage T7 and to characterize the effect of primase activity on fork progression. We observe the synthesis of primers on the lagging strand to cause transient pausing of the highly processive leading-strand synthesis. In the presence of both leading- and lagging-strand synthesis, we observe the formation and release of a replication loop on the lagging strand. Before loop formation, the primase acts as a molecular brake and transiently halts progression of the replication fork. This observation suggests a mechanism that prevents leading-strand synthesis from outpacing lagging-strand synthesis during the slow enzymatic steps on the lagging strand.

An attractive model system for studying the orchestration at the replication fork is the replication machinery of bacteriophage T7. It can be reconstituted *in vitro* using a small number of purified proteins (Fig. 1a) and its organization closely mimics that of *Escherichia coli* and more complex organisms[9]. The T7 DNA polymerase consists of a 1:1 complex of the T7 gene 5 protein (gp5), encoded by the phage, and the thioredoxin (trx) processivity factor, encoded by the *E. coli* host[10]. The T7 gene 4 protein (gp4) assembles as a hexamer and provides both helicase and primase activities[11,12]. The helicase activity, required for unwinding the parental DNA strands, is located in the carboxy-terminal half, and the primase activity, capable of synthesizing the tetraribonucleotide primers that are required to initiate lagging-strand DNA synthesis, is located in the amino-terminal half (see Fig. 1a, b). The position of the primase active site on the outside of the hexamer[13] would, in principle, allow the helicase to continue unwinding while a primer is synthesized on the lagging strand. A replication loop formed on the lagging strand of the replication fork reorients the lagging-strand DNA polymerase so that it advances in parallel with the leading-strand polymerase.

Here we characterize the kinetics of leading- and lagging-strand synthesis at the single-molecule level by monitoring the length of individual DNA molecules during the replication reaction. The 5′ end of one strand of a 48.5 kilobase (kb)-long duplex λ phage DNA molecule is attached to the bottom surface of a glass flow cell using a biotin–streptavidin linker. The opposite 3′ end bears a digoxigenin moiety and is linked to a bead that is 2.8 μm in diameter and coated with anti-digoxigenin antibody (Fig. 2a, b). A constant laminar flow is applied above the surface such that the resultant drag on the beads stretches the DNA molecules with a force of 3 pN (ref. 14). This force is well within the regime in which the force reduces the configurational freedom of the DNA but does not influence protein–DNA interactions[15]. We measure changes in the lengths of the individual DNA molecules by imaging the beads and tracking their positions.

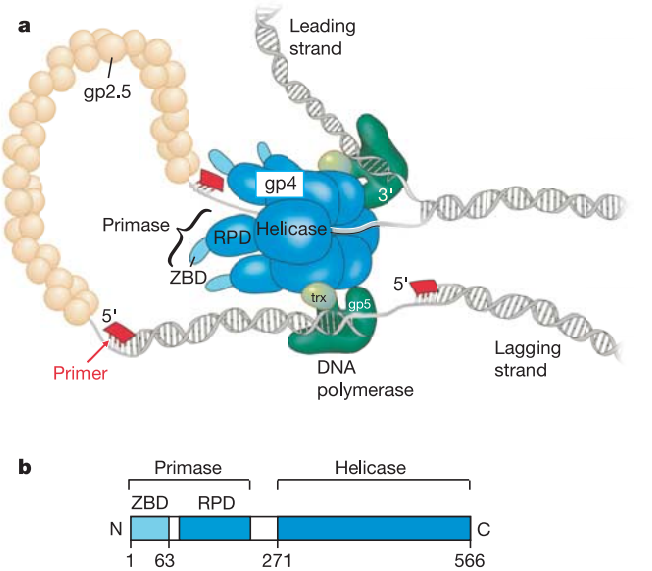

Figure 1 | The T7 replication fork. a, The bacteriophage T7 replisome consists of the hexameric T7 gene 4 protein (gp4) and two copies of the T7 DNA polymerase (T7 gene 5 protein (gp5) complexed with *E. coli* thioredoxin (trx)). T7 gene 2.5 protein (gp2.5), the ssDNA-binding protein, coats the transiently exposed ssDNA in the replication loop. **b**, Gp4 consists of a primase and a helicase domain, connected by a linker region. The primase domain consists of two subdomains: a zinc-binding domain (ZBD) and the RNA polymerase domain (RPD).

[1]Harvard Medical School, Department of Biological Chemistry and Molecular Pharmacology, 250 Longwood Avenue, Boston, Massachusetts 02115, USA. [2]Harvard Medical School, Graduate Program of Biological and Biomedical Sciences, 220 Longwood Avenue, Boston, Massachusetts 02115, USA. [3]Harvard University, Department of Chemistry and Chemical Biology, 12 Oxford Street, Cambridge, Massachusetts 02138, USA.

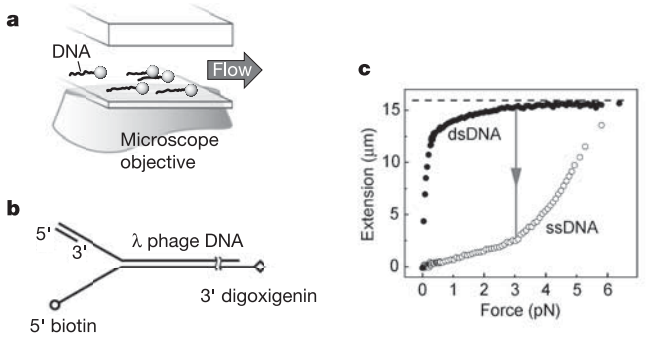

Figure 2 | Experimental design. a, Schematic representation of the experimental arrangement (not drawn to scale). **b**, Pre-made forked DNA substrate with single-stranded regions exposed at both 3′ and 5′ ends to allow for assembly of T7 DNA polymerase and gp4, respectively. **c**, Extension of λ-phage ssDNA (open circles) and dsDNA (filled circles) as a function of the applied stretching force. The dashed horizontal line at 16.2 μm represents the crystallographic length of B-form λ-phage dsDNA. The vertical arrow illustrates the change in length as a result of enzymatic conversion from dsDNA to ssDNA.

The coiling of single-stranded (ss)DNA causes it to be shorter than double-stranded (ds)DNA at low stretching forces (<6 pN; Fig. 2c). Consequently, conversion from dsDNA to ssDNA can be monitored through a decrease in total DNA length[14,16,17].

We preassemble a complex consisting of one T7 DNA polymerase (gp5–trx) and helicase/primase (gp4) on a forked λ phage DNA. This primer-template has an exposed 5′ ssDNA tail to facilitate loading of the hexameric helicase[18] and a DNA primer annealed to the other strand to initiate leading-strand synthesis (Fig. 2b). In the presence of deoxynucleoside 5′ triphosphates (dNTPs), but in the absence of Mg^{2+} ions, only the gp4 and leading-strand DNA polymerase will remain stably bound to the DNA fork, without initiating synthesis[19]. Stringent washing of the flow cell after pre-assembly of the leading-strand complex removes all proteins not stably bound, including the lagging-strand DNA polymerase. Unwinding and nucleotide incorporation are started by addition of Mg^{2+}.

Leading-strand synthesis catalysed by T7 DNA polymerase converts one DNA strand arising from gp4 helicase activity into dsDNA[20]. In the absence of the lagging-strand DNA polymerase, the lagging strand will remain in the single-stranded form. By attaching the DNA to the surface of the flow cell by the 5′ biotin-labelled tail, leading-strand synthesis can be detected by an effective shortening of the DNA (Fig. 3a, b). The number of nucleotides polymerized can be obtained by using the difference between the lengths of ssDNA and dsDNA[14,16,17].

Averaged over multiple single-molecule traces ($n = 13$ traces), we measure the processivity of T7 leading-strand synthesis to be 17 ± 3 kb (mean ± s.e.m.), a dramatic increase compared to that of unwinding by the gp4 alone (60 base pairs (bp))[21] or synthesis by the T7 DNA polymerase alone (~1 kb, see Supplementary Figs S1, S2 and ref. 10). The origin of this high processivity is likely to be found in a stabilization of the complex by a specific interaction of the acidic C-terminal domain of gp4 with basic regions on the T7 DNA polymerase[19]. The rate of leading-strand synthesis (164 ± 20 bp s^{-1}, $n = 13$) is identical to that measured in ensemble experiments[22], confirming that the force exerted on the DNA has no influence on the enzymatic activities.

We enabled the primase activity of gp4 by adding the required ribonucleotides to the reaction mixture. The sequence recognized by the T7 gp4 primase is 3′CTG(G/T)(G/T)5′, where the 3′C is essential for recognition but is not copied into the RNA primer[23]. Consequently, the primase requires only the ribonucleotides rATP and rCTP. In the presence of both rATP and rCTP, we observe the appearance of short pauses in the single-molecule leading-strand synthesis traces (Fig. 3a, trace 2; pauses are indicated by grey arrows).

To confirm that the pauses originate from primase activity, we measured leading-strand synthesis catalysed by the T7 DNA polymerase and a variant of gp4 that is lacking the 63 N-terminal amino acids that form the zinc-binding domain (ZBD). This truncated gp4 cannot catalyse template-directed RNA synthesis[24], although it does form hexamers and show normal helicase activity[25]. We performed leading-strand synthesis with this gp4 isoform both in the presence and absence of rATP and rCTP. The absence of pausing in either case confirms that the pausing is dependent on primase activity (Fig. 3a, traces 3 and 4). Pausing by wild-type gp4 is abolished when only a

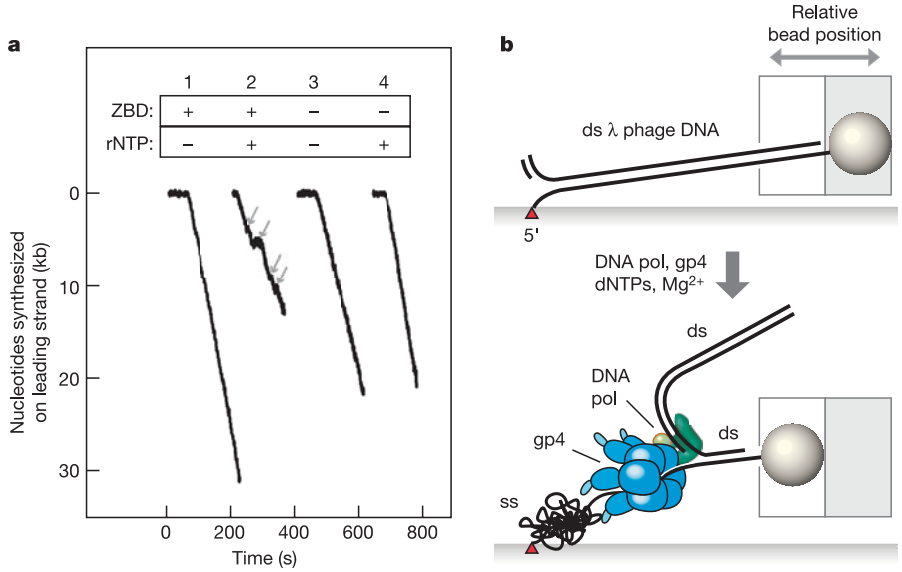

Figure 3 | Single-molecule observation of T7 leading-strand synthesis. a, Single-molecule trajectories. Above the traces is indicated whether the gp4 contained the zinc-binding domain (ZBD), and whether rATP and rCTP were present in the reaction mixture. Grey arrows denote pauses. **b**, Schematic depiction of events observed in **a**. In the absence of lagging-strand synthesis, leading-strand synthesis causes the 5′ tail of the DNA to be converted to the single-stranded form. Attachment of the 5′ end to the surface allows us to monitor this conversion as a change in total length of the DNA.

single ribonucleotide is present (Supplementary Information). Furthermore, the pause sites in the presence of rATP and rCTP correspond to the positions of known primase recognition sites in the λ phage genome (Supplementary Information).

To exclude the possibility that pauses in leading-strand synthesis are simply caused by an inhibition of the primase to transfer its primer, we repeated the single-molecule replication experiments in the continuous presence of excess T7 DNA polymerase. The ensuing lagging-strand synthesis leads to the formation and release of a replication loop on the lagging strand, expressed in the single-molecule traces as a gradual shortening of the DNA followed by a prompt lengthening (Fig. 4). A quantitative analysis of the various length changes is presented in the Supplementary Information. The presence of replication loops confirms that correctly functioning replication forks are formed in our replication experiments. Loop formation is not observed when the ribonucleotides are withheld from the reaction. The low concentrations of DNA used in the single-molecule experiment result in a DNA–protein stoichiometry different from previously published bulk experiments[8], and may explain the reduced processivity of the replisome.

The pauses before the initiation of replication loop formation represent a transient halting of the whole replication fork during primer synthesis and delivery. The average pause duration of 6 ± 2 s ($n = 5$) is similar to that measured in the leading-strand synthesis experiments (5.6 ± 0.7 s, $n = 62$ at 3 pN; 5 ± 2 s, $n = 8$ at 1 pN), indicating that it is primarily determined by primer synthesis as opposed to delivery. This observation also confirms that the absence of the T7 single-stranded DNA-binding protein, gp2.5, is not likely to have an influence on the pausing behaviour of the fork. After all, gp2.5 mediates primer site recognition, but does not influence the RNA polymerization rate[26].

The several seconds needed to synthesize a tetraribonucleotide correspond well with the low rate of RNA synthesis by the primase as measured in ensemble experiments[23]. In the time required by the primase to synthesize a primer, the leading-strand synthesis would have synthesized close to 1,000 bases. The fact that an Okazaki fragment in T7 replication has an average length of 3,000 nucleotides[8] would mean that the lagging-strand DNA polymerase has to synthesize appreciably faster than the leading-strand polymerase to make up for this delay. However, closer analysis of replication loop formation as observed in our experiments shows that the rate of synthesis of the two DNA polymerases is equal (Supplementary Information). An alternative explanation could be that a primer for the next Okazaki fragment is already synthesized during the synthesis of the nascent fragment[7], although the delay arising from the transfer of the lagging-strand DNA polymerase to the new primer remains. Our observation that leading-strand synthesis momentarily stops during primer synthesis provides a simpler solution. A transient halting of the whole replication complex during the slow enzymatic events on the lagging strand would allow the necessary transactions to be completed without leading-strand synthesis progressing too far ahead of the lagging-strand synthesis.

At this point, it is unclear what the molecular mechanism is that underlies the halting of leading-strand synthesis by the primase activity. The low affinity of the primase for the primer-template[23] makes it unlikely that stalling is caused by a direct primase–RNA–DNA interaction. This leaves the possibility that ribonucleotide condensation by the primase results in an allosteric regulation of helicase activity through protein–protein interactions. The replication machinery of other organisms, such as T4 and *E. coli*, differs from that of the T7 bacteriophage in that the primase is not covalently linked to the helicase. Nonetheless, oligomerization of the primase, with a similar organization as that observed for the T7 gp4, is achieved in these systems through the interaction between the primase and helicase, and has been demonstrated to be critical for primase synthesis[27,28]. Albeit noncovalent in nature, this interaction could facilitate a similar regulation of helicase activity to that we have observed in the T7 replication machinery.

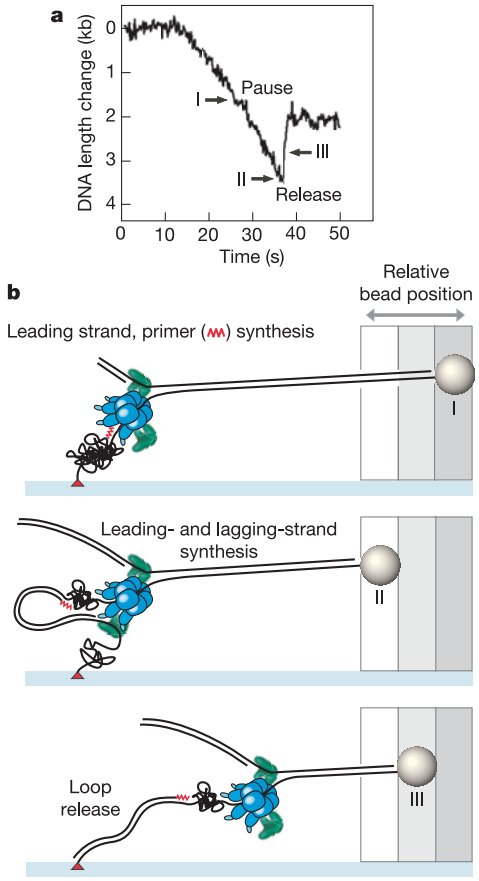

Figure 4 | Lagging-strand synthesis and loop formation. a, Single-molecule trajectory of replication loop formation. In the presence of additional T7 DNA polymerase, lagging-strand synthesis is initiated after primer synthesis (indicated by the pause), and a replication loop is formed in the lagging strand. Replication loop release is clearly visible as instantaneous lengthening of DNA. A quantitative interpretation of the different rates and length changes is presented in the Supplementary Information. **b**, Schematic depiction of the events observed in **a**.

METHODS

Tethering and observation of single DNA molecules. Forked DNA molecules (see Supplementary Information for DNA substrate preparation) are attached by one end to the glass surface of a flow cell and by the other end to a bead (see ref. 14 and Supplementary Information). To prevent nonspecific interactions between the beads and the surface, we used paramagnetic beads with a diameter of 2.8 μm (Dynal) and applied a small (1 pN) magnetic force upwards by positioning a permanent magnet above the flow cell. Combined with the horizontal drag force created by laminar flow, this magnetic force held the centre of the beads ~4 μm above the surface of the flow cell. The flow cells were placed on an inverted microscope and the beads were imaged using a digital CCD camera. Typically, an area of 850 μm × 850 μm was imaged with a time-resolution of 200 ms. Particle-tracking software (Semasopht) was then used to determine the centre position of every bead for every frame with a precision of 1 nm (ref. 29). About 50 individual, tethered DNA molecules can be observed simultaneously, allowing for a high data throughput.

Replication reactions. T7 gp4 with and without the ZBD, and T7 gp5–trx were purified as previously described[10,30]. The flow cell containing tethered, forked DNA molecules was incubated for 15 min with 18 nM T7 gp4 (hexameric) and 18 nM T7 gp5 (DNA polymerase) complexed 1:1 with *E. coli* thioredoxin in replication buffer (40 mM Tris pH 7.5, 50 mM potassium glutamate, 10 mM dithiothreitol, 100 μg ml^{-1} BSA), in the presence of 700 μM each of dATP, dTTP, dCTP and dGTP. After washing with 15 flow cell volumes (300 μl) of replication buffer and nucleotides, replication was started by introducing 10 mM $MgCl_2$,

700 μM dNTPs and 300 μM each of rATP and rCTP in replication buffer. Lagging-strand synthesis was initiated by adding 18 nM T7 gp5 to the latter mixture. All experiments were performed at a temperature of 22 ± 1 °C.

Data analysis. After particle tracking, traces were corrected for residual instabilities in the flow by subtracting traces corresponding to tethers that were not enzymatically altered. Bead displacements were converted into numbers of nucleotides synthesized, using the known length difference between ssDNA and dsDNA at our experimental conditions (Fig. 2c). The resultant traces were smoothed and fitted with series of line segments to detect pauses and determine rates. See Supplementary Information for details regarding data analysis.

Received 14 June; accepted 13 October 2005.

1. Alberts, B. DNA replication and recombination. *Nature* **421,** 431–435 (2003).
2. Sheaff, R. J. & Kuchta, R. D. Mechanism of calf thymus DNA primase: slow initiation, rapid polymerization, and intelligent termination. *Biochemistry* **32,** 3027–3037 (1993).
3. Swart, J. R. & Griep, M. A. Primer synthesis kinetics by *Escherichia coli* primase on single-stranded DNA templates. *Biochemistry* **34,** 16097–16106 (1995).
4. Frick, D. N., Kumar, S. & Richardson, C. C. Interaction of ribonucleoside triphosphates with the gene 4 primase of bacteriophage T7. *J. Biol. Chem.* **274,** 35899–35907 (1999).
5. Stukenberg, P. T., Turner, J. & O'Donnell, M. An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. *Cell* **78,** 877–887 (1994).
6. Salinas, F. & Benkovic, S. J. Characterization of bacteriophage T4-coordinated leading- and lagging-strand synthesis on a minicircle substrate. *Proc. Natl Acad. Sci. USA* **97,** 7196–7201 (2000).
7. Tougu, K. & Marians, K. J. The interaction between helicase and primase sets the replication fork clock. *J. Biol. Chem.* **271,** 21398–21405 (1996).
8. Lee, J., Chastain, P. D. II, Griffith, J. D. & Richardson, C. C. Lagging strand synthesis in coordinated DNA synthesis by bacteriophage T7 replication proteins. *J. Mol. Biol.* **316,** 19–34 (2002).
9. Benkovic, S. J., Valentine, A. M. & Salinas, F. Replisome-mediated DNA replication. *Annu. Rev. Biochem.* **70,** 181–208 (2001).
10. Tabor, S., Huber, H. E. & Richardson, C. C. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* **262,** 16212–16223 (1987).
11. Guo, S., Tabor, S. & Richardson, C. C. The linker region between the helicase and primase domains of the bacteriophage T7 gene 4 protein is critical for hexamer formation. *J. Biol. Chem.* **274,** 30303–30309 (1999).
12. Frick, D. N., Baradaran, K. & Richardson, C. C. An N-terminal fragment of the gene 4 helicase/primase of bacteriophage T7 retains primase activity in the absence of helicase activity. *Proc. Natl Acad. Sci. USA* **95,** 7957–7962 (1998).
13. Kusakabe, T., Baradaran, K., Lee, J. & Richardson, C. C. R. Roles of the helicase and primase domain of the gene 4 protein of bacteriophage T7 in accessing the primase recognition site. *EMBO J.* **17,** 1542–1552 (1998).
14. van Oijen, A. M. *et al.* Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* **301,** 1235–1238 (2003).
15. Bustamante, C., Smith, S. B., Liphardt, J. & Smith, D. Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.* **10,** 279–285 (2000).
16. Wuite, G. J., Smith, S. B., Young, M., Keller, D. & Bustamante, C. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* **404,** 103–106 (2000).
17. Maier, B., Bensimon, D. & Croquette, V. Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proc. Natl Acad. Sci. USA* **97,** 12002–12007 (2000).
18. Ahnert, P., Picha, K. M. & Patel, S. S. A ring-opening mechanism for DNA binding in the central channel of the T7 helicase-primase protein. *EMBO J.* **19,** 3418–3427 (2000).
19. Hamdan, S. M. *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. USA* **102,** 5096–5101 (2005).
20. Kolodner, R. & Richardson, C. C. Gene 4 protein of bacteriophage T7. Characterization of the product synthesized by the T7 DNA polymerase and gene 4 protein in the absence of ribonucleoside 5′-triphosphates. *J. Biol. Chem.* **253,** 574–584 (1978).
21. Jeong, Y. J., Levin, M. K. & Patel, S. S. The DNA-unwinding mechanism of the ring helicase of bacteriophage T7. *Proc. Natl Acad. Sci. USA* **101,** 7264–7269 (2004).
22. Stano, N. M. *et al.* DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase. *Nature* **435,** 370–373 (2005).
23. Frick, D. N. & Richardson, C. C. Interaction of bacteriophage T7 gene 4 primase with its template recognition site. *J. Biol. Chem.* **274,** 35889–35898 (1999).
24. Bernstein, J. A. & Richardson, C. C. A 7-kDa region of the bacteriophage T7 gene 4 protein is required for primase but not for helicase activity. *Proc. Natl Acad. Sci. USA* **85,** 396–400 (1988).
25. Bernstein, J. A. & Richardson, C. C. Purification of the 56-kDa component of the bacteriophage T7 primase/helicase and characterization of its nucleoside 5′-triphosphatase activity. *J. Biol. Chem.* **263,** 14891–14899 (1988).
26. He, Z.-G. & Richardson, C. C. Effect of single-stranded DNA-binding proteins on the helicase and primase activities of the bacteriophage T7 gene 4 protein. *J. Biol. Chem.* **279,** 22190–22197 (2004).
27. Yang, J., Xi, J., Zhuang, Z. & Benkovic, S. J. The oligomeric T4 primase is the functional form during replication. *J. Biol. Chem.* **280,** 25416–25423 (2005).
28. Mitkova, A. V., Khopde, S. M. & Biswas, S. B. Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of *Escherichia coli* in RNA primer synthesis. *J. Biol. Chem.* **278,** 52253–52261 (2005).
29. Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82,** 2775–2783 (2002).
30. Notarnicola, S. M., Mulcahy, H. L., Lee, J. & Richardson, C. C. The acidic carboxyl terminus of the bacteriophage T7 gene 4 helicase/primase interacts with T7 DNA polymerase. *J. Biol. Chem.* **272,** 18425–18433 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We wish to thank T. Ellenberger, D. Crampton and P. Blainey for discussions and comments, and S. Buratowski for critically reading the manuscript. We are grateful to S.-J. Lee for providing the purified ZBD-less gp4 variant. We thank S. Moskowitz for preparation of the figures. This work was supported by grants from the NIH to C.C.R. and X.S.X.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.M.v.O. (antoine_van_oijen@hms.harvard.edu).