

Dynamic look at DNA unwinding by a replicative helicase

Seung-Jae Lee^{a,b,c}, Salman Syed^d, Eric J. Enemark^{a,c,1}, Stephen Schuck^c, Arne Stenlund^c, Taekjip Ha^{d,e,2}, and Leemor Joshua-Tor^{a,b,c,2}

^aW. M. Keck Structural Biology Laboratory, ^bHoward Hughes Medical Institute, and ^cCold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; and ^dCenter for Biophysics and Computational Biology and ^eDepartment of Physics and Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Edited by Stephen J. Benkovic, The Pennsylvania State University, University Park, PA, and approved January 28, 2014 (received for review December 2, 2013)

A prerequisite for DNA replication is the unwinding of duplex DNA catalyzed by a replicative hexameric helicase. Despite a growing body of research, key elements of helicase mechanism remain under substantial debate. In particular, the number of DNA strands encircled by the helicase ring during unwinding and the ring orientation at the replication fork completely contrast in contemporary mechanistic models. Here we use single-molecule and ensemble assays to address these questions for the papillomavirus E1 helicase. We find that E1 unwinds DNA with a strand-exclusion mechanism, with the N-terminal side of the helicase ring facing the replication fork. We show that E1 generates strikingly heterogeneous unwinding patterns stemming from varying degrees of repetitive movements, which is modulated by the DNA-binding domain. Together, our studies reveal previously unrecognized dynamic facets of replicative helicase unwinding mechanisms.

ATPase | molecular motors

DNA replication is the most fundamental of all of life's processes. One of the key requisites in the initiation of replication is the separation of the two strands of the double helix, which is carried out by a hexameric helicase. Despite their prominent roles in biology, some of the basic aspects of these helicases, whether they use a strand-exclusion mechanism or whether they translocate along double-stranded DNA, for example, have been subjects of considerable debate (1, 2). Viral replicative helicases, such as SV40 Large-T antigen (LTag) and papillomavirus E1, have provided the opportunity to study some of these basic features largely owing to their homohexameric architecture. These viral helicases recognize their respective origin of DNA replication (ori) through their dsDNA-binding domains (DBDs) and assemble in a stepwise fashion, ultimately forming double-hexameric (DH) structures on their ori and unwind the DNA bidirectionally.

E1 consists of an N-terminal domain, a DBD, an oligomerization domain (OD), a helicase/ATPase domain (HD), and a C-terminal acidic tail (Fig. 1A). Biochemical and structural data have demonstrated that the DBDs bound to the pseudopalindromic E1 binding site are at the center of the double hexamer and that the helicase domains that bind to the flanks of the ori are on either end in a head-to-head arrangement (3, 4). This arrangement is supported by EM studies of LTag that show a dumbbell-shaped structure for the DH (5, 6), with each half of the dumbbell containing two lobes: a larger HD outer lobe and a smaller DBD inner lobe. The assembled DH appears to place the HD of E1 proximal to the dsDNA to be unwound.

However, in the structure of the E1 helicase with ssDNA and Mg^{2+} -ADP, the ssDNA is oriented such that the N-terminal part of the polypeptide (the OD) is closest to the 5' end of the DNA (7). The 3'→5' polarity of E1 helicase indicates that the translocating helicase moves with the N-terminal OD leading and the C-terminal helicase domain trailing, or alternatively, that the DNA is pumped through the helicase from the OD side to the HD side. Does this apparent discrepancy imply that the DH functions fundamentally differently from the single hexameric helicase? It has been sug-

gested that a DH might function as a helicase by encircling and pumping dsDNA for unwinding (2). Early studies of LTag on DNA seemed to support such a model showing "rabbit-ear" structures where the two hexamers remain in close proximity, and the DNA emanates roughly from the center between the hexamers (8). However, other unwinding intermediates in the same study showed isolated hexamers, making it difficult to unequivocally assign a mode of action. Indeed, more recent EM studies on forked DNA showed "back-to-back" arrangement of LTag where the larger lobes are adjacent to each other. In addition, whereas the lobes often seem to be collinear, in some cases they are clearly offset (9). In fact, a similar back-to-back interaction between hexamers, which are offset rather than collinear, was observed in the crystal of E1-ssDNA-ADP. We suggested previously (10) that the two strands of the DNA duplex would be separated during hexamer assembly, where each hexamer assembles around different ssDNA strands and the helicase then acts as a translocase. This view is reinforced both by the E1-ssDNA-MgADP crystal structure as well as by recent work by Yardimci et al. (11) on LTag.

Several mechanisms for helicase unwinding by hexameric helicases have been proposed (1, 2). In one mechanism, described for LTag, termed the "squeeze-pumping" model, the dsDNA enters the channel formed by the hexameric helicase from the ATPase domain side, and the narrowing of the channel induced by "concerted" ATP hydrolysis melts the dsDNA by a "squeeze/crush" or "squeeze-to-open" mechanism.

In contrast, the structure of the E1 helicase with ssDNA and Mg^{2+} -ADP (7) suggested that each hexamer encircles and

Significance

Precise replication of the genome is essential for maintaining the integrity of genetic information in all forms of life. A key step in this process is the unwinding of the DNA double helix, a subject of intensive research and debate. We took a multifaceted approach to study fundamental mechanisms of helicase function, using the E1 helicase from papillomavirus. Our findings reveal that E1 employs a strand exclusion mechanism to unwind DNA with the N-terminal side leading at the replication fork. Intriguingly, DNA unwinding by E1 is modulated by the origin-recognition domain, suggesting a previously unsuspected role for this domain in regulating helicase activity.

Author contributions: S.-J.L., E.J.E., A.S., T.H., and L.J. designed research; S.-J.L., S. Syed, E.J.E., S. Schuck, and A.S. performed research; S.-J.L. contributed new reagents/analytic tools; S.-J.L., S. Syed, E.J.E., S. Schuck, A.S., T.H., and L.J. analyzed data; and S.-J.L., E.J.E., A.S., T.H., and L.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹Present address: Department of Structural Biology, St Jude Children's Research Hospital, Memphis, TN 38105.

²To whom correspondence may be addressed. E-mail: tjha@illinois.edu or leemor@cshl.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1322254111/-DCSupplemental.

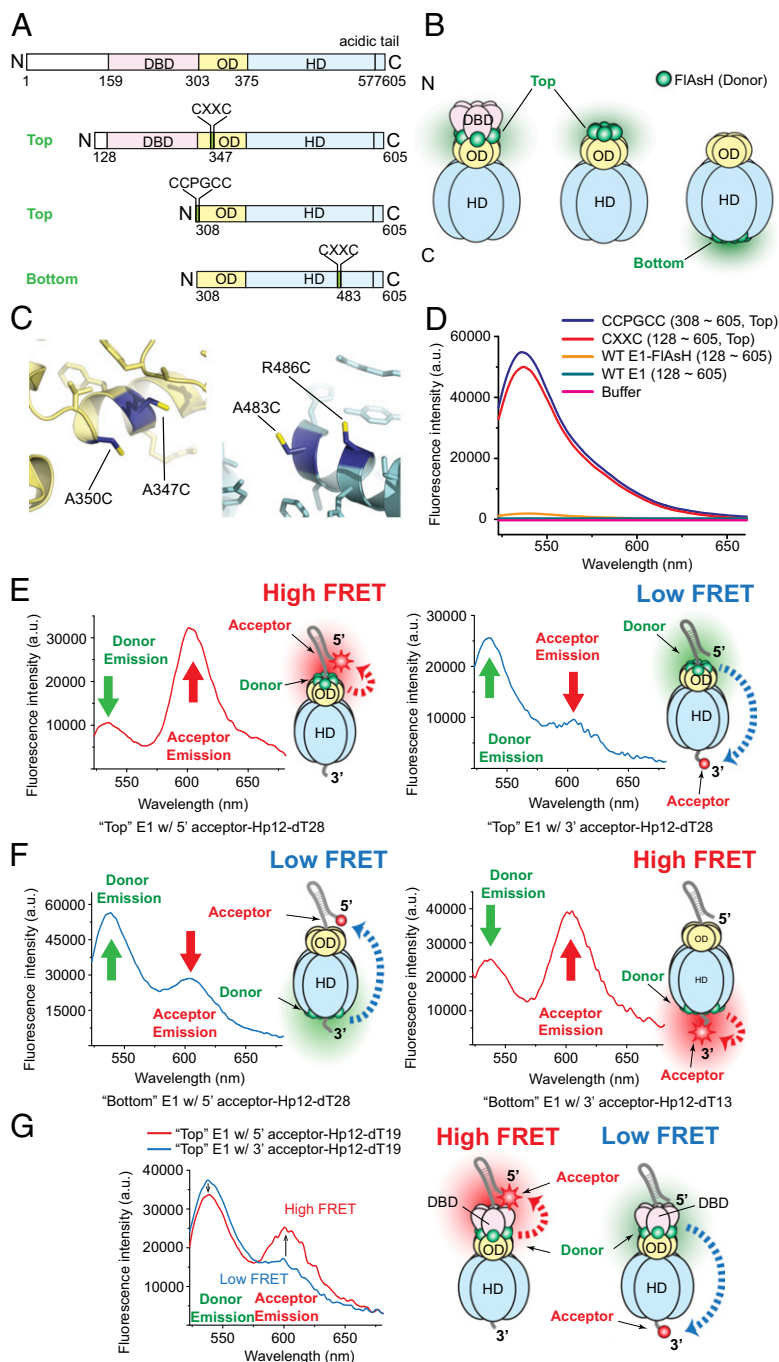


Fig. 1. Orientation of E1 helicase on a “replication fork” substrate. (A) E1 constructs used in this study. FIAsh binding sites (CCPGCC or CXXC) were introduced into the constructs as indicated. (B) Schematic diagram illustrating the location of the labels on the helicase. (C) (Left) Di-cysteine mutations (A347C and A350C) introduced into the OD (“top” position, 128–605). (Right) Di-cysteine mutations (A483C and R486C) incorporated into the HD (“bottom” position, 308–605). These di-cysteine substitutions were introduced to the noncatalytic region, where the target residues in the helical region are spaced by two to three residues apart (4–6 Å) and show low sequence conservation. In addition, they are solvent-exposed and not part of a network of hydrogen bonds. (D) Emission spectra were scanned using 2 μM of E1 helicases labeled with FIAsh via CCPGCC or CXXC. Wild-type E1(128–605) without a FIAsh binding site (green curve) did not bind FIAsh after washes. The labeling efficiencies of FIAsh on all three sites including the CCPGCC and CXXC were all above 95% (based on absorbance measurements using the extinction coefficient $\epsilon_{\text{FIAsh},528} = 70,000 \text{ cm}^{-1} \cdot \text{M}^{-1}$). (E) (Left) High FRET interaction between the “top”-labeled E1 (OD+HD, residues 308–605) helicase and the 5'-labeled DNA. (Right) Low FRET interaction between “top”-labeled E1 (308–605) helicase and the 3'-labeled DNA. (F) (Left) Low FRET interaction between “bottom”-labeled E1 (308–605) helicase and 5'-labeled DNA. (Right) High FRET interaction between “bottom”-labeled E1 (308–605) helicase and 3'-labeled DNA. In this experiment, we used a shorter 3'-ssDNA tail (dT13) that resulted in a stronger FRET signal than using the longer 3'-ssDNA (dT28) tail. (G) Orientation of E1 (DBD+OD+HD, residues 128–605) helicase on a “replication fork” substrate. A higher FRET interaction occurs between “top”-labeled E1 (128–605) helicase and the 5'-labeled DNA than with the 3'-labeled DNA. For the FRET experiment, E1 was purified as a stable monomer. The protein was then hexamerized using either ssDNA (dT₃₀) or hairpin DNA substrates, which were prelabeled with Alexa 568 on their 5' or 3' end, in the presence of Mg²⁺ and ADP, a condition that does not support substrate unwinding. The hairpin DNA substrates were designed to resemble a DNA replication fork with a duplex region of 12 bp and a long, 28-base 3' single-stranded extension. The protein–DNA complex was separated from the monomeric species by gel filtration and labeled with FIAsh as described in *Materials and Methods*. The final concentration of monomeric E1 was 4–7 M, and the ratio between donor and acceptor is ~6:1, as expected. All FRET experiments were performed in the presence of Mg²⁺ and ADP to trap the ssDNA-bound hexameric form of the helicase. The FRET efficiency or distance between donor and acceptor was not determined quantitatively owing to the complex nature of multidonor single-acceptor system in E1 hexamer–DNA complex. See also Fig. S1.

translocates on ssDNA during unwinding by using a sequential ATP hydrolysis mechanism. A similar mechanism, albeit with an opposite polarity, was described for an ssRNA translocase: the *Escherichia coli* transcription termination factor Rho (12) (also reviewed in ref. 13) and for the bacterial hexameric helicase DnaB, although in this case the authors propose a two-nucleotide translocation step (14). In addition, the sequential hydrolysis model requires subunits to switch conformations at the nucleotide-binding site during the hydrolysis cycle, as described in the ClpX AAA+ unfoldase (15).

In seeking to determine whether the E1 double hexamer uses a fundamentally different mechanism for helicase activity than what is observed for the hexamer, we find that E1 uses a strand-exclusion mechanism for DNA unwinding in the double hexamer

as well as in the hexamer, and that the helicase is oriented such that the N terminus with its DBD is near the replication fork. We further show that rather than continuous, monotonic unwinding, implied by the sequential ATP hydrolysis mechanism, E1 harbors heterogeneous unwinding patterns we have not recognized previously. The helicase devoid of its DBD displays very repetitive unwinding characteristics, showing extensive slippage and reinitiation, which are significantly diminished when the DBD is present. In addition, the DBD reduces the assembly time of the helicase on a forked-DNA substrate.

Results

Orientation of the E1 Helicase on DNA in Solution. The E1–ssDNA–MgADP crystal structure implies that the ds/ss fork junction is

located above the ring formed by the ODs (Fig. 1B). To determine the relative orientation of the helicase with respect to substrate DNA we performed a series of equilibrium FRET measurements between a labeled helicase and a labeled DNA substrate. We placed the donor fluorophore, fluorescein arsenical hairpin binder (FAsH) (16), on either side of the helicase—on what we denote as the “top” of the helicase, which is at the N-terminal OD, or on the “bottom” of the helicase (Fig. 1B). Single-stranded or hairpin DNA was labeled with the acceptor, Alexa 568, on either their 5′ or 3′ ends, so that we could compare two different pairs of measurements. To simplify our analysis, we used two versions of the helicase: a shorter form (residues 308–605) and a longer form (residues 128–605) that also includes the DBD (Fig. 1A). Inclusion of the acidic tail was important to facilitate hexamerization of the helicase (17). The distance between the “top” and the “bottom” of E1 in the E1(308–577)–ssDNA–MgADP crystal structure is ~ 58 Å. With an R_0 value of 60 Å for the FAsH–Alexa 568 FRET pair, we expect to clearly detect changes in the 30- to 90-Å range (18). To label the “top” of the shorter form of the helicase, we incorporated the canonical tetra-cysteine motif (CCPGCC) at the N-terminal end of the E1 construct, which according to the crystal structure points directly up from the OD (Fig. 1B). Owing to the potential variability of the positions of DBDs in the hexamer, we decided to keep the label on the more rigid OD in the longer form of the helicase. To retain structural integrity, we chose to use a modified form of a di-cysteine motif (CX_nC; X is any amino acid; $n = 2$ –3) shown previously to be as efficient for FAsH binding as the canonical tetra-cysteine motif (19). This design principle was also used to label the “bottom” of the helicase, again with minimal to no disruption to the structure of the helicase (Fig. 1C).

The emission spectra of FAsH-labeled CCPGCC–E1 or CXXC–E1 were practically identical (Fig. 1D) and control FAsH labeling reactions with “wild-type” E1 constructs lacking tetra- or di-cysteine motifs did not result in detectable fluorescence intensity. These results indicate that CCPGCC–E1 and CXXC–E1 are selectively labeled with the FAsH donor fluorophore. The N-terminally “top”-labeled E1 (308–605) displayed a strong FRET signal with the 5′-labeled hairpin DNA, with highly depressed donor signal around 530 nm and distinctively sensitized acceptor emission around 610 nm. In contrast, the same labeled protein displayed a low FRET signal with a 3′-labeled hairpin DNA (Fig. 1E). The sample preparation and data correction for bulk-phase FRET assay were performed as presented in *Materials and Methods*. Next, we performed similar FRET experiments with the “bottom”-labeled E1 with the same set of labeled 3′-end extended hairpin DNA substrates. Here, the “bottom”-labeled E1 hexamer produced a low FRET signal around the acceptor emission wavelength with a 5′-labeled substrate DNA (Fig. 1F). These results corroborate our model that the ds/ss junction representing the replication fork is positioned on the OD side of the helicase.

A similar set of experiments was performed on a longer form of E1 that includes the DBD spanning residues 128–605. The same mutants were used to label the helicase with FAsH at the “top” of the helicase on the OD. Using two different substrates—a single-strand dT₂₆ (Fig. S1) as well as a 3′-end extended hairpin DNA labeled on either end (Fig. 1G)—a high FRET signal was obtained with both substrates when they were labeled at their 5′ end, whereas a lower FRET signal was obtained for both when they were labeled at their 3′ end. Thus, all FRET experiments performed were consistent with the ds/ss junction located nearer the DBD and the N terminus of the helicase and further from the HD.

E1 Uses a Steric Exclusion Mechanism for DNA Unwinding. The FRET experiments described above are consistent with the crystal structure of the E1 helicase domain in complex with ssDNA, which demonstrates that a hexamer of the E1 helicase and

oligomerization domains binds to ssDNA such that the ssDNA passes through the central channel of the hexamer. Curiously, early EM data from a close relative of the E1 protein, SV40 T-antigen, seemed to show that dsDNA enters both ends of the DH complex and that ssDNA is then extruded from the DH (8).

A possible explanation for this apparent discrepancy could be that the path of the DNA in a hexamer formed on ssDNA is radically different from that in a DH complex formed on a dsDNA template. Specifically, we wanted to distinguish between models where one or two strands pass through the hexameric ring. Prior observations have demonstrated that some hexameric helicases, in the process of unwinding, can displace streptavidin bound to a biotinylated template (20). It occurred to us that such displacement could be used to determine whether one or both DNA strands of a template pass through the hexameric ring. We have shown that E1 binds to a dsDNA ori probe, forming a double trimer (DT); this DT is converted into a DH, which unwinds the template in the presence of *E. coli* single-strand DNA-binding protein (SSB), resulting in ssDNA+SSB complexes that can be detected by EMSA. This reaction is dependent on the formation of the E1 DH, which in turn is dependent on the centrally located E1 binding sites (BSs) and sequences on the flanks of the E1 BS (3, 21, 22). If both strands of the DNA pass through the central channel of the helicase, we expect that a streptavidin attached to either the 5′ or 3′ end of the DNA would be displaced, whereas if only one strand passes through the central channel, only streptavidin attached to that strand would be displaced. Because E1 is a 3′ to 5′ helicase we would expect that the streptavidin bound to the 5′ end of the DNA would be displaced.

We generated three probes corresponding to the minimal 84-bp ori probe that we have previously shown is unwound by a DH of E1 (Fig. 2). In one probe no biotin was incorporated (no biotin). In another probe, we attached a biotin to the 3′ end of one strand and labeled the other end of that same strand with ³²P (3′ biotin). In the third probe, we attached biotin to the 5′ end of one strand and labeled the 3′ end of the same strand with ³²P (5′ biotin). For the no-biotin template (Fig. 2B, lanes 1–7), addition of streptavidin had no effect on the mobility of the untreated probe as expected (compare lanes 1 and 2). The mobility of the boiled probe was reduced greatly owing to the complex formation with *E. coli* SSB (lane 3), but the mobility of the ssDNA+SSB complex was not changed by the presence of streptavidin (lane 4). E1 can generate the same ssDNA+SSB complex generated by boiling (lanes 5–7), demonstrating that E1 unwinds the template and that streptavidin has no effect on this process (lanes 5 and 6).

To test the 3′ and 5′ biotin templates for unwinding, we first bound streptavidin to the biotinylated probes. As shown in Fig. 2B, lanes 9 and 16, the majority of the probe was complexed with streptavidin. We next tested these streptavidin-associated probes for unwinding by E1 (lanes 13, 14, 20, and 21). Both probes were unwound by E1 as indicated by the formation of ssDNA+SSB complexes (lanes 13, 14, 20, and 21). In the absence of free biotin (lanes 13 and 20), both of the unwound probes were bound to streptavidin. In the presence of free biotin, however (lanes 14 and 21), where rebinding to the biotinylated probe is prevented, the unwound 5′-biotin probe lacked bound streptavidin. In contrast, the 3′-biotin probe had retained bound streptavidin. These results demonstrate that streptavidin is displaced from the 5′-biotin probe but not from the 3′-biotin probe during unwinding. As markers we generated ssDNA by boiling the template and complexed this ssDNA with *E. coli* SSB (lanes 3, 10, and 17) or with *E. coli* SSB and streptavidin (lanes 4, 11, and 18).

These results very clearly demonstrate that the two DNA strands are treated differently by the E1 helicase that melts and unwinds the viral ori, effectively ruling out a mechanism where both DNA strands pass through the central channel in the DH. Indeed, the most likely trajectory for the DNA strands is that during unwinding by the E1 DH helicase one strand passes

unwinding pattern. To simplify our analysis, molecules were grouped into three different subpopulations according to their unwinding patterns that were defined as monotonic (11% of traces), partially repetitive (46%), and repetitive (43%), based on increasing levels of slippage events that result in repetitive movements along the substrate (Fig. 3B). Unwinding traces were categorized with decreasing monotonicity for each unwinding experiment containing ~ 300 traces. The traces in which no slippage was observed during unwinding were classified as monotonic. If one or more slippage/unwinding event with a FRET increase during slippage larger than 0.4 were observed, they were considered repetitive. All other traces showing an intermediate degree of nonmonotonicity were classified as partially repetitive. Repeated experiments showed similar distribution of unwinding patterns (monotonic $8.3 \pm 1.4\%$, partially repetitive $45.3 \pm 2.3\%$, and repetitive $46.3 \pm 2.4\%$). To exclude the possibility that repetitive unwinding results from the specific substrate geometry or the positions of the fluorophores, we also tested a different forked DNA substrate where the FRET value would increase rather than decrease as unwinding progresses (Fig. S2), displaying similar heterogeneous unwinding patterns (27, 28).

The DBD of E1 Promotes Monotonic Unwinding. To understand what role the DBD might play in unwinding, we performed the single-molecule FRET unwinding experiment using the longer construct of the E1 helicase that includes the DBD (residues 128–605). In this case, the percentage of the population characterized by repetitive movements dramatically decreased from 43 to 6%, promoting more monotonous unwinding with less slippage (Fig. 4). The DBD seems to prevent slippage, possibly with an extra “grip” on the ds/ss junction, consistent with the well-established role of the DBD in binding dsDNA during initial origin recognition (10, 29, 30). Alternatively, the DBD may have an allosteric effect on nucleotide binding or hydrolysis. Our data demonstrate that the DBD plays an additional role of enhancing the unwinding activity by preventing extensive slippage, consistent with the observation that mutations on the surface of the E1 DBD affect the helicase activity of full-length E1, indicating that the DBD plays a role in the helicase activity of E1 (31), even though the DBD itself is not required for helicase activity (32). We note that the addition of single-stranded DNA-binding protein does not alter the unwinding behavior of E1 helicase either with or without the DBD (Fig. S3).

ADP Inhibits Monotonic Unwinding by E1. In the sequential hydrolysis mechanism described above, forward translocation, and thus progression of unwinding, occurs with binding and subsequent hydrolysis of ATP. Therefore, increasing the concentration of ADP should interfere with this cycle and thus interfere with unwinding. Using the longer form of E1, which includes the DBD (residues 128–605) that displays less slippage, we added ADP to the reaction and performed a single-molecule FRET unwinding assay as described above. The portion of traces that displayed repetitive unwinding behavior increased from 6% with no ADP to 55% with a ratio of 1:2 and to 65% with a ratio of 1:1 ADP:ATP at 1.5 mM each (Fig. 5). Increasing the ratio of ADP to ATP thus increases the percentage of molecules that show extensive slippage behavior, resulting in an inhibitory effect on the helicase.

The E1 DBD Promotes Assembly of the Functional Helicase on a Forked Substrate. In the single-molecule FRET unwinding assays described here, there is a noticeable lag time between the introduction of the helicase and nucleotide into the reaction chamber and commencement of measureable unwinding activity. We refer to this as the assembly time of the functional E1 helicase on the substrate. During assembly, E1, loaded with nucleotides, would recognize and assemble on the DNA substrate. It should be

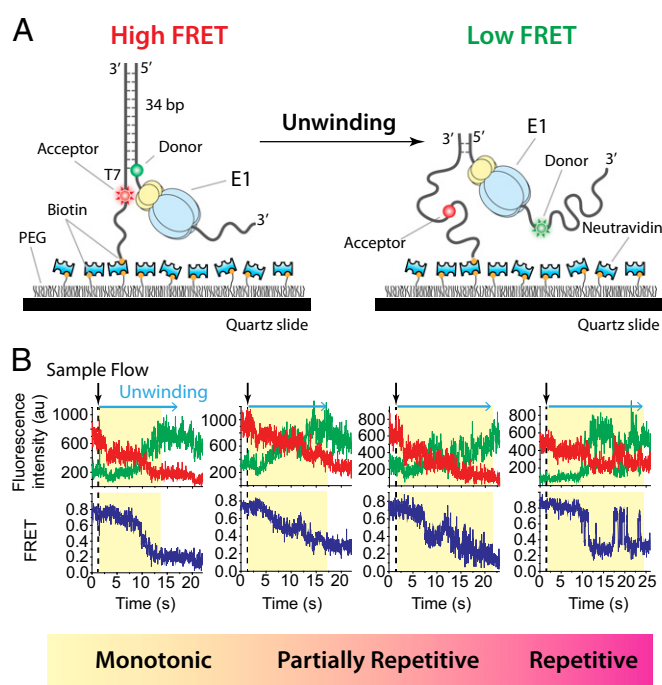


Fig. 3. Heterogeneous unwinding patterns of E1 helicase (residues 308–605) on a forked DNA substrate revealed by smFRET (high to low FRET). (A) A schematic diagram showing the smFRET unwinding assay for E1 helicase (high to low FRET). To probe the unwinding by E1 helicase on a single DNA molecule, we prepared a fluorescently labeled forked substrate, which contains a 34-bp duplex DNA with a 3' dT₆₀ single-stranded tail. Donor fluorophore (Cy3, green) was located at the junction on the 3' tracking strand and acceptor fluorophore (Cy5, red) was introduced onto the 5' nontracking single-stranded tail, seven nucleotides away from the junction. The forked DNA substrate is tethered to the PEG surface through a biotin at the end of the 5' tail. Before unwinding, FRET between the two fluorophores is high owing to close proximity and is expected to decrease as unwinding progresses. (B) A high degree of heterogeneity in the unwinding patterns of E1 helicase. smFRET unwinding assay revealed that E1 helicase (residues 308–605) has heterogeneous unwinding behaviors ranging from monotonic to repetitive movements. See the main text for criteria unwinding pattern classification; see also Fig. S2.

noted that in this experiment we could not rule out that a potential conformational change from a nonfunctional hexameric helicase to a functional one might contribute to the assembly time. The reaction was initiated in all cases by flowing mixtures of E1 (700 nM), MgCl₂ (10 mM), and ATP (2 mM).

Three E1 constructs were tested. The first consists of the OD and HD alone (OD+HD) and showed essentially no unwinding activity ($<1\%$) and thus did not permit calculation of assembly times. Indeed, as noted previously, the acidic tail has been reported to help maintain the oligomeric state of the helicase (17). The second construct includes the acidic tail, OD+HD+acidic tail (residues 308–605) and displayed a mean assembly time of 10.1 ± 0.25 s (Fig. 6). Strikingly, the more complete E1 helicase that also includes the DBD (residues 128–605) showed a considerably shorter mean assembly time of 4.9 ± 0.25 s (Fig. 6). Therefore, it seems that the E1 DBD serves to significantly promote assembly of the helicase on DNA. This is consistent with a role for the DBD where the E1 hexamer is assembled in place around a single strand of DNA at the ori.

Discussion

Our understanding of the mechanism by which replicative helicases unwind dsDNA has benefited tremendously from a combination of biochemical, structural, and biophysical studies. Here, we address

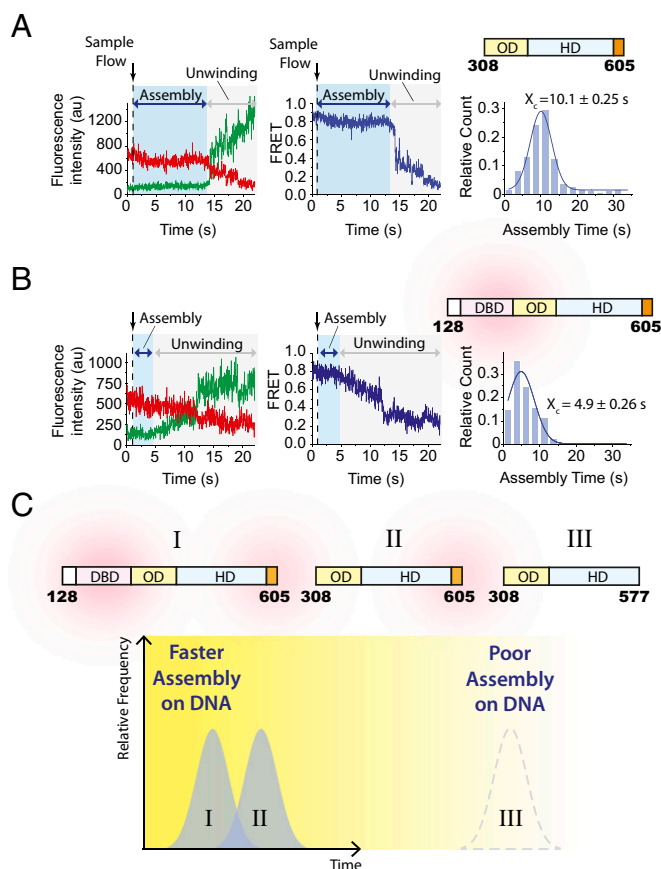


Fig. 6. The E1 DBD promotes functional helicase assembly on a forked substrate. (A) E1 containing OD+HD+acidic tail (residues 308–605) displayed assembly time with a mean of 10.1 ± 0.25 s. (B) The longer E1 helicase (residues 128–605), which includes the DBD, showed a considerably shorter assembly time with a mean of 4.9 ± 0.26 s. For the experiments shown in A and B, the reaction was initiated by flowing mixtures of E1 (700 nM), $MgCl_2$ (5 mM), ATP (2 mM), and an oxygen-scavenging buffer system. The lag time between the introduction of sample and initiation of unwinding was used as a parameter to assess assembly time for each construct. (C) A schematic diagram illustrating the population shift in the assembly time of various E1 constructs. E1 containing only OD+HD has practically no enzymatic activity in the smFRET unwinding assay. Once the 30-residues-long acidic tail was included, there was a sharp increase in molecule assembly (decrease in assembly time) followed by unwinding activity. Incorporation of the DBD further decreases the assembly time of E1 helicase.

As presented in our studies, smFRET analysis reveals that E1 harbors strikingly heterogeneous population derived from varying degree of repetitive unwinding events. The DBD of E1 greatly reduces this repetitive unwinding and promotes a much smoother, monotonic process. This is an unexpected role for the DBD that we did not anticipate. However, Schuck and Stenlund (31) had previously found several mutations in the DBD that affected helicase activity, some quite dramatically. The result is an increase in processivity in a sense, although not strictly speaking, because the enzyme does not dissociate in this case. In addition to promoting smooth monotonic unwinding, the DBD also decreases the assembly time the helicase needs to get on the DNA and start unwinding by about twofold.

The fact that the DBD does not completely mitigate slippage suggests that either the N-terminal domain, which is missing in all of these constructs, contributes to the helicase activity of the E1 protein or that limited slippage is a normal part of the helicase function and possibly serves a regulatory role for the E1 helicase. We believe that the former possibility is unlikely; our

data demonstrate that the helicase activity of full-length E1 on very long substrates (>1 kb) is reduced compared with E1_{308–605}, indicating that the presence of the N-terminal domain reduces processivity of E1 (Fig. S4). Although no direct evidence exists for the second possibility, it is conceivable that the slippage represents a mechanism to prevent extensive unwinding until all required cofactors are present and a functional complex is assembled on the origin of DNA. It is possible that the presence of polymerase would affect the nonmonotonic unwinding. It was shown that T7 polymerase enhances the activity of T7 helicase (46), although it is unclear at present whether this effect is through the prevention of back-slipping.

Here we present a comprehensive mechanism by which E1, a replicative helicase, unwinds duplex DNA with its specific orientation toward ds/ss junction. We also demonstrate remarkable heterogeneity of E1 unwinding activity. Taken together, our studies reveal previously unrecognized dynamic aspects of replicative helicase as well as its detailed functional mechanisms. These dynamic elements including the heterogeneity add yet another layer of complexity to replicative helicases that remains to be investigated further. In addition, an intriguing question to be answered will be how the dynamic elements of replicative helicases are related to their diverse cellular functions in the context of DNA replication.

Materials and Methods

Protein Expression and Purification. All bovine papillomavirus (BPV) E1 wild-type and mutant constructs were expressed in *E. coli* strain BL21DE3 as N-terminal GST fusions with a thrombin cleavage site. The proteins were purified as previously described (33, 47).

Site-Directed Mutagenesis. Multiple amino acid substitutions or insertions were introduced into vectors encoding BPV E1 fragments of varying lengths using the Phusion site-directed mutagenesis kit (Finnzymes) according to the manufacturer's instructions.

Site-Specific Labeling of E1 and DNA Substrates for Ensemble Steady-State FRET Assays.

For bulk FRET experiments with hexameric E1–DNA complexes, we chose to use FIAsh labeling as opposed to maleimide coupling because E1 is rich in cysteines, which would complicate specific labeling. We placed the donor fluorophore (FIAsh) on either side of the helicase—on what we denote as the “top” of the helicase, which is at the N-terminal OD, or on the “bottom” of the helicase. Single-stranded or hairpin DNA was labeled with the acceptor Alexa 568 on either their 5′ or 3′ ends, so that we could compare two different pairs of measurements. For each E1 hexamer–DNA complex, one sample labeled with acceptor only and the other sample doubly labeled with donor and acceptor were prepared in parallel. To generate the doubly labeled samples with donor and acceptor, E1 hexamer complexed with prelabeled Alexa 568–DNA oligo were incubated for >1 h at room temperature in the dark with more than two equivalents of FIAsh (commercially known as Lumio Green). The unbound free dye was removed by gel filtration on a desalting column equilibrated with buffer. To label FIAsh onto E1, a canonical CCGGCC motif was used together with a modified form of a di-cysteine motif (CX_nC ; X is any amino acid; $n = 2–3$). Therefore, residues A347 and A350 were mutated to cysteines for the “top” OD mutations, and residues A483 and R486 mutated to cysteines for the “bottom” label. In both cases, the pair of residues are positioned on a helix ~ 5 Å apart, are solvent-exposed, are not part of hydrogen bonding network, and show low sequence conservation, resulting in a CXXC FIAsh-binding motif incorporated directly into the structure of the helicase. The labeling efficiencies of FIAsh on all three sites including the CCGGCC and CXXC were all above 95% (based on absorbance measurements using the extinction coefficient $\epsilon_{FIAsh,528} = 70,000$ $cm^{-1} \cdot M^{-1}$).

Single-strand or 3′-extended hairpin–DNA oligos labeled with Alexa Fluor 568 were purchased from Eurofins MWG Operon. Single-strand or 3′-extended hairpin–DNA oligos labeled with Alexa Fluor 568 were purchased from Eurofins MWG Operon. The hairpin substrates included a double-stranded region at the 5′ end of an extended single-stranded tail to be at the leading edge of the 3′→5′ helicase. Sequences of DNA substrates are available in [Supporting Information](#).

For the FRET experiment, E1 was purified as a stable monomer. The protein was then hexamerized using either ssDNA (dT₃₀) or hairpin DNA substrates,

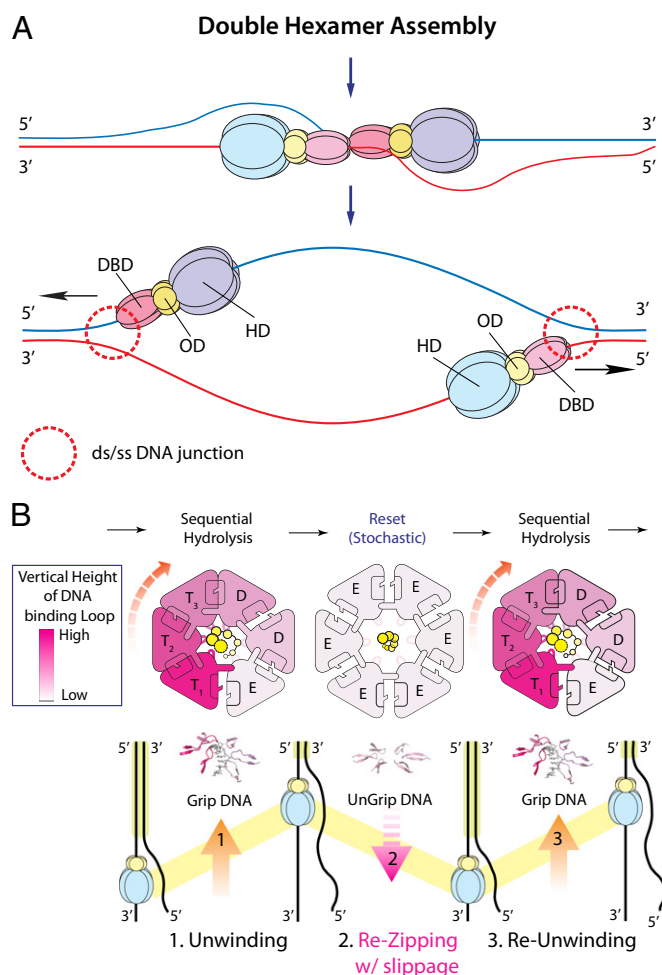


Fig. 7. DNA unwinding by the E1 helicase. (A) Model illustrating the strand-exclusion mechanism for unwinding in which single E1 hexamers translocate along ssDNA, with the ds/ss DNA junction located on the DBD side of the helicase. (B) Repetitive unwinding in the E1 helicase. In our coordinated escort model for E1 translocation on DNA, the helicase uses sequential ATP hydrolysis for forward movement along ssDNA, stripping away the other strand, thus unwinding dsDNA. The vertical positions of the DNA-binding hairpins in the central channel of the helicase are associated with different nucleotide states between the subunits such that ATP binding is associated with the top-most position, ADP with a lower position, and an empty binding site with the lowest position in the channel. Other states, such as ADP+phosphate, are inferred to exist and would occupy intermediate positions. ATP binding is also associated with higher affinity to DNA. The slippage events that we observe here most likely correspond to disassociation of nucleotides from most, if not all, binding sites around the hexameric ring, leading to ungridding of the DNA, slippage, and re-zipping of the dsDNA. Rebinding of nucleotides would restart the hydrolysis cycle to restart unwinding in the forward direction. The DBD reduces the occurrence of these stochastic slippage events.

which were prelabeled with Alexa 568 on their 5' or 3' end, in the presence of Mg^{2+} and ADP, a condition that does not support substrate unwinding. The hairpin DNA substrates were designed to resemble a DNA replication fork with a duplex region of 12 bp and a long, 28-base 3' single-stranded extension. The protein–DNA complex was separated from the monomeric species by gel filtration and labeled with FIAsh. The final concentration of monomeric E1 was 4–7 μM , and the ratio between donor and acceptor is ~6:1, as expected. All FRET experiments were performed in the presence of Mg^{2+} and ADP to trap the ssDNA-bound hexameric form of the helicase.

For the experiment in which the fluorophore was placed at the 3' end of the substrate DNA, we shortened the 3' extension to 13 nucleotides to

reduce the distance between the donor on the bottom of the helicase and the 3' end of the DNA substrate to obtain a clearer FRET signal.

Ensemble Fluorescence Spectroscopy. For each E1 hexamer–DNA complex investigated, we prepared two types of samples, one sample (DA) labeled with donor and acceptor fluorophores and the other one (A) labeled only with acceptor fluorophore. For each DA sample, we measured two emission spectra obtained with the excitation wavelengths for the donor (FIAsh) and for the acceptor (Alexa Fluor 568), which are 500 nm and 578 nm, respectively. Likewise, two emission spectra from each A sample were collected with the excitation wavelengths for the donor and the acceptor. In the donor–acceptor system used in this study, donor (FIAsh) and acceptor (Alexa Fluor 568) emissions are peaked at around 530 nm and 600 nm, respectively. When the E1 hexamer–DNA complex labeled with both probes is excited at the excitation wavelength for FIAsh (500 nm), a combination of three signals is obtained including donor (FIAsh) emission, acceptor (Alexa Fluor 568) emission arising from FRET interaction, and some acceptor emission derived from direct excitation of acceptor at 500 nm, which is the excitation wavelength for donor FIAsh. The latter bleed-through fluorescence resulting from direct excitation of acceptor at the donor excitation wavelength was corrected and subtracted as described below. After spectrum correction, we were able to observe relative changes in fluorescence intensity for the donor and acceptor emission peaks and compare FRET signals. For each placement of donor in a given E1 hexamer–DNA complex, all experimental parameters and conditions were kept identical, except for the relative location of the acceptor fluorophore in the DNA substrates. The FRET efficiency or distance between donor and acceptor was not determined quantitatively owing to the complex nature of multidonor single-acceptor system in E1 hexamer–DNA complex.

Ensemble FRET measurements were performed on a GEN5 spectrofluorimeter thermostated to 25 °C. For each E1 hexamer–DNA complex, emission spectra were taken from two types of samples: DA, which is labeled with a FIAsh donor on the protein and an Alexa 568 acceptor on the DNA, and A, which is labeled only with the Alexa 568 acceptor on the DNA. For each type of sample (DA or A), two types of emission scans were obtained. The first emission scan (donor scan) used the excitation maximum for the donor fluorophore (FIAsh). The second emission scan (acceptor scan) used the excitation maximum for the acceptor fluorophore (Alexa Fluor 568). The excitation wavelength for the donor and acceptor was 500 nm and 578 nm, respectively. The emission spectra were scanned from 500 nm to 700 nm. Background fluorescence of the buffer and unlabeled E1–DNA complex was negligible. The procedures for data correction are available in [Supporting Information](#).

Origin Unwinding Assays with Biotinylated Probes. The 84-bp ori probe was generated by PCR using the BPV-1 viral ori as a template. For the 5'-labeled probes the top strand primer was 5'-labeled with [γ - ^{32}P]ATP before the PCR. The bottom-strand primer contained a restriction site for the enzyme Mlu I. After PCR, the template was digested with Mlu I and filled in with biotin-dCTP to place a biotin at the 3' end. To place a biotin at the 5' end we used a bottom-strand primer containing a biotin for PCR. The top-strand primer contained a restriction site for Mlu I. After PCR, the template was digested with Mlu I and filled in with [α - ^{32}P]dCTP generating a template with the 3' end of one strand labeled with ^{32}P and the 5' end of the same strand labeled with biotin. Binding of streptavidin was carried out by incubating 0.2 pmol biotin-substituted probe with 2 pmol of streptavidin overnight at room temperature. Unwinding assays were performed by incubating 0.25 pmol of E1 with 2 fmol of ori probe in 10 μL of binding buffer [20 mM Hepes (pH 7.5), 70 mM NaCl, 0.7 mg/mL BSA, 0.1% Nonidet P-40, 5% (vol/vol) glycerol, 5 mM DTT, and 5 mM $MgCl_2$] in the presence of 10 ng/ μL *E. coli* SSB and 2 mM ATP. The samples were incubated for 30 min at 32 °C and Sarkosyl was added to 0.1% immediately before loading of the samples onto 4% poly acrylamide gels. In samples containing free biotin it was added to a 1,000-fold molar excess over the probe.

smFRET Assays. smFRET experiments to measure unwinding activities of E1 were performed on a wide-field total internal-reflection fluorescence microscope with 30-ms time resolution and imaged by means of a charged-coupled-device camera (iXon DV 887-BI; Andor Technology) (48). Oligonucleotides were purchased from Integrated DNA Technology. The Cy3 and Cy5 fluorophores were internally labeled on dT through a C_6 amino linker. Forked DNA substrates were prepared by mixing the appropriate biotinylated and nonbiotinylated oligonucleotides in a 1:1.3 molar ratio at 5 μM in T50 buffer [10 mM Tris (pH 8.0) and 50 mM NaCl]. Annealing reactions were in-

cubated at 95 °C for 3 min followed by slow cooling to room temperature for 3 h. Quartz slides and glass coverslips were surface-passivated with PEG containing 1% biotin-PEG (Laysan Bio, Inc.), as previously described (24, 49). A typical E1 unwinding buffer for smFRET assays includes an oxygen scavenger system and a triplet-quenching agent (50). Additional information about DNA substrates is available in [Supporting Information](#).

- Enemark EJ, Joshua-Tor L (2008) On helicases and other motor proteins. *Curr Opin Struct Biol* 18(2):243–257.
- Gai D, Chang YP, Chen XS (2010) Origin DNA melting and unwinding in DNA replication. *Curr Opin Struct Biol* 20(6):756–762.
- Schuck S, Stenlund A (2005) Assembly of a double hexameric helicase. *Mol Cell* 20(3):377–389.
- Stenlund A (2003) Initiation of DNA replication: Lessons from viral initiator proteins. *Nat Rev Mol Cell Biol* 4(10):777–785.
- Valle M, Gruss C, Halmer L, Carazo JM, Donate LE (2000) Large T-antigen double hexamers imaged at the simian virus 40 origin of replication. *Mol Cell Biol* 20(1):34–41.
- Gomez-Lorenzo MG, et al. (2003) Large T antigen on the simian virus 40 origin of replication: A 3D snapshot prior to DNA replication. *EMBO J* 22(23):6205–6213.
- Enemark EJ, Joshua-Tor L (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* 442(7100):270–275.
- Wessel R, Schweizer J, Stahl H (1992) Simian virus 40 T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. *J Virol* 66(2):804–815.
- VanLook MS, Alexandrov A, Yu X, Cozzarelli NR, Egelman EH (2002) SV40 large T antigen hexamer structure: Domain organization and DNA-induced conformational changes. *Curr Biol* 12(6):472–476.
- Enemark EJ, Stenlund A, Joshua-Tor L (2002) Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. *EMBO J* 21(6):1487–1496.
- Yardimci H, et al. (2012) Bypass of a protein barrier by a replicative DNA helicase. *Nature* 492(7428):205–209.
- Thomsen ND, Berger JM (2009) Running in reverse: The structural basis for translocation polarity in hexameric helicases. *Cell* 139(3):523–534.
- Lyubimov AY, Strycharzka M, Berger JM (2011) The nuts and bolts of ring-translocase structure and mechanism. *Curr Opin Struct Biol* 21(2):240–248.
- Itsathitphaisarn O, Wing RA, Eliason WK, Wang J, Steitz TA (2012) The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* 151(2):267–277.
- Stinson BM, et al. (2013) Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* 153(3):628–639.
- Griffin BA, Adams SR, Tsien RY (1998) Specific covalent labeling of recombinant protein molecules inside live cells. *Science* 281(5374):269–272.
- Whelan F, et al. (2012) A flexible brace maintains the assembly of a hexameric replicative helicase during DNA unwinding. *Nucleic Acids Res* 40(5):2271–2283.
- Granier S, et al. (2007) Structure and conformational changes in the C-terminal domain of the beta2-adrenoceptor: Insights from fluorescence resonance energy transfer studies. *J Biol Chem* 282(18):13895–13905.
- Zhang X-Y, Bishop AC (2008) Engineered inhibitor sensitivity in the WPD loop of a protein tyrosine phosphatase. *Biochemistry* 47(15):4491–4500.
- Morris PD, Raney KD (1999) DNA helicases displace streptavidin from biotin-labeled oligonucleotides. *Biochemistry* 38(16):5164–5171.
- Schuck S, Stenlund A (2011) Mechanistic analysis of local ori melting and helicase assembly by the papillomavirus E1 protein. *Mol Cell* 43(5):776–787.
- Schuck S, Stenlund A (2007) ATP-dependent minor groove recognition of TA base pairs is required for template melting by the E1 initiator protein. *J Virol* 81(7):3293–3302.
- Ha T, et al. (1996) Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci USA* 93(13):6264–6268.
- Ha T, et al. (2002) Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* 419(6907):638–641.
- Graham BW, Schauer GD, Leuba SH, Trakselis MA (2011) Steric exclusion and wrapping of the excluded DNA strand occurs along discrete external binding paths during MCM helicase unwinding. *Nucleic Acids Res* 39(15):6585–6595.
- Rothenberg E, Trakselis MA, Bell SD, Ha T (2007) MCM forked substrate specificity involves dynamic interaction with the 5'-tail. *J Biol Chem* 282(47):34229–34234.
- Lee G, Yoo J, Leslie BJ, Ha T (2011) Single-molecule analysis reveals three phases of DNA degradation by an exonuclease. *Nat Chem Biol* 7(6):367–374.
- Lee G, Bratkowski MA, Ding F, Ke A, Ha T (2012) Elastic coupling between RNA degradation and unwinding by an exoribonuclease. *Science* 336(6089):1726–1729.
- Chen G, Stenlund A (1998) Characterization of the DNA-binding domain of the bovine papillomavirus replication initiator E1. *J Virol* 72(4):2567–2576.
- Enemark EJ, Chen G, Vaughn DE, Stenlund A, Joshua-Tor L (2000) Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Mol Cell* 6(1):149–158.
- Schuck S, Stenlund A (2006) Surface mutagenesis of the bovine papillomavirus E1 DNA binding domain reveals residues required for multiple functions related to DNA replication. *J Virol* 80(15):7491–7499.
- Castella S, Burgin D, Sanders CM (2006) Role of ATP hydrolysis in the DNA translocase activity of the bovine papillomavirus (BPV-1) E1 helicase. *Nucleic Acids Res* 34(13):3731–3741.
- Sedman J, Stenlund A (1998) The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *J Virol* 72(8):6893–6897.
- McGeoch AT, Trakselis MA, Laskey RA, Bell SD (2005) Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat Struct Mol Biol* 12(9):756–762.
- Hohng S, Joo C, Ha T (2004) Single-molecule three-color FRET. *Biophys J* 87(2):1328–1337.
- Sun B, et al. (2011) ATP-induced helicase slippage reveals highly coordinated subunits. *Nature* 478(7367):132–135.
- Myong S, Rasnik I, Joo C, Lohman TM, Ha T (2005) Repetitive shuttling of a motor protein on DNA. *Nature* 437(7063):1321–1325.
- Myong S, Bruno MM, Pyle AM, Ha T (2007) Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase. *Science* 317(5837):513–516.
- Myong S, et al. (2009) Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA. *Science* 323(5917):1070–1074.
- Yodh JG, Stevens BC, Kanagaraj R, Janscak P, Ha T (2009) BLM helicase measures DNA unwound before switching strands and hRPA promotes unwinding reinitiation. *EMBO J* 28(4):405–416.
- Park J, et al. (2010) PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. *Cell* 142(4):544–555.
- Sun Y, et al. (2012) The eukaryotic initiation factor eIF4H facilitates loop-binding, repetitive RNA unwinding by the eIF4A DEAD-box helicase. *Nucleic Acids Res* 40(13):6199–6207.
- Klaue D, et al. (2013) Fork sensing and strand switching control antagonistic activities of RecQ helicases. *Nat Commun* 4:2024.
- Qi Z, Pugh RA, Spies M, Chemla YR (2013) Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. *Elife* 2:e00334.
- Masuda-Ozawa T, Hoang T, Seo YS, Chen LF, Spies M (2013) Single-molecule sorting reveals how ubiquitylation affects substrate recognition and activities of FBH1 helicase. *Nucleic Acids Res* 41(6):3576–3587.
- Stano NM, et al. (2005) DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase. *Nature* 435(7040):370–373.
- Sedman J, Sedman J, Stenlund A (1997) Binding of the E1 and E2 proteins to the origin of replication of bovine papillomavirus. *J Virol* 71(4):2887–2896.
- Ha T (2001) Single-molecule fluorescence resonance energy transfer. *Methods* 25(1):78–86.
- Rasnik I, Myong S, Cheng W, Lohman TM, Ha T (2004) DNA-binding orientation and domain conformation of the E. coli rep helicase monomer bound to a partial duplex junction: single-molecule studies of fluorescently labeled enzymes. *J Mol Biol* 336(2):395–408.
- Rasnik I, McKinney SA, Ha T (2006) Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* 3(11):891–893.