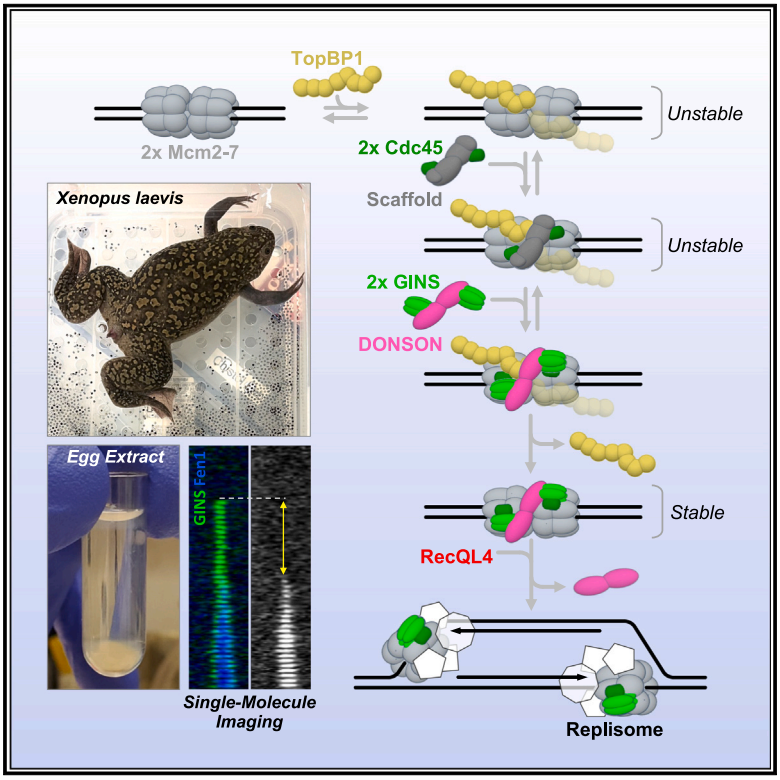
Cell

Article

Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa

Graphical abstract



Authors

Riki Terui, Scott E. Berger, Larissa A. Sambel, Dan Song, Gheorghe Chistol

Correspondence

chistol@stanford.edu

In brief

Single-molecule imaging was used to visualize how DNA replication initiates bidirectionally from individual origins, providing critical insights into the function and dynamics of several key replication initiation factors.

Highlights

- Single-molecule experiments shed light on how DNA replication is initiated
- Reveals the order in which replication initiation factors are recruited to origins
- 2-fold symmetry is established by dimerization scaffolds that recruit GINS and Cdc45
- The ATPase activity of RecQL4 plays an important role in efficient origin firing

Terui et al., 2024, Cell *187*, 1–18
July 25, 2024
https://doi.org/10.1016/j.cell.2024.05.024

CellPress

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

Cell

CellPress

Article

Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa

Riki Terui,[1,7] Scott E. Berger,[2,7] Larissa A. Sambel,[1] Dan Song,[1,6] and Gheorghe Chistol[1,2,3,4,5,8,*]
[1]Chemical and Systems Biology Department, Stanford School of Medicine, Stanford, CA 94305, USA
[2]Biophysics Program, Stanford School of Medicine, Stanford, CA 94305, USA
[3]Cancer Biology Program, Stanford School of Medicine, Stanford, CA 94305, USA
[4]Stanford Cancer Institute, Stanford School of Medicine, Stanford, CA 94305, USA
[5]BioX Interdisciplinary Institute, Stanford School of Medicine, Stanford, CA 94305, USA
[6]Present address: Eikon Therapeutics Inc., Hayward, CA 94545, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: chistol@stanford.edu
https://doi.org/10.1016/j.cell.2024.05.024

SUMMARY

Metazoan genomes are copied bidirectionally from thousands of replication origins. Replication initiation entails the assembly and activation of two CMG helicases (Cdc45•Mcm2-7•GINS) at each origin. This requires several replication firing factors (including TopBP1, RecQL4, and DONSON) whose exact roles are still under debate. How two helicases are correctly assembled and activated at each origin is a long-standing question. By visualizing the recruitment of GINS, Cdc45, TopBP1, RecQL4, and DONSON in real time, we uncovered that replication initiation is surprisingly dynamic. First, TopBP1 transiently binds to the origin and dissociates before the start of DNA synthesis. Second, two Cdc45 are recruited together, even though Cdc45 alone cannot dimerize. Next, two copies of DONSON and two GINS simultaneously arrive at the origin, completing the assembly of two CMG helicases. Finally, RecQL4 is recruited to the CMG•DONSON•DONSON•CMG complex and promotes DONSON dissociation and CMG activation via its ATPase activity.

INTRODUCTION

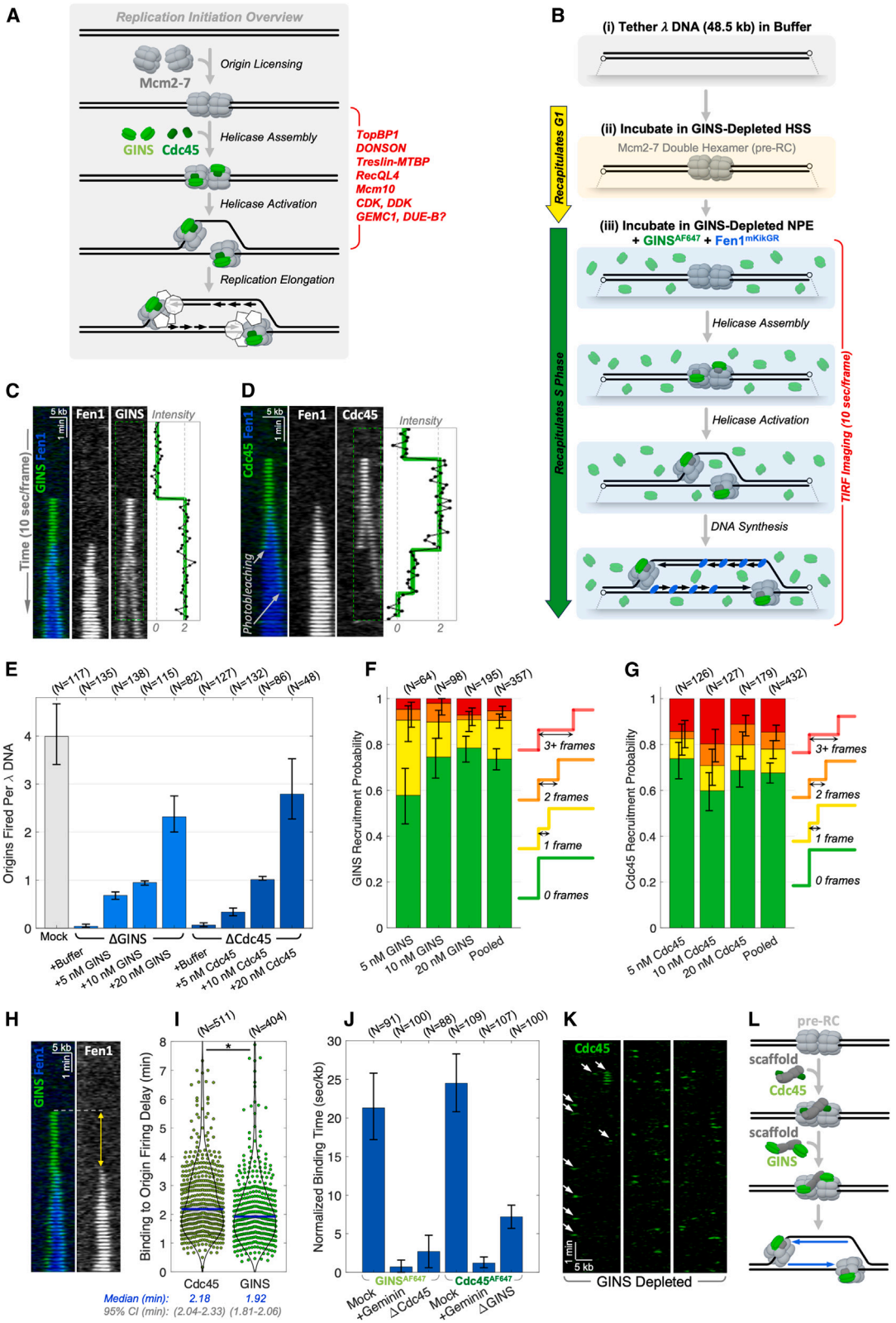
To efficiently duplicate their large genomes, eukaryotes initiate DNA replication in parallel from thousands of sites called origins. Importantly, DNA is replicated bidirectionally from each origin, ensuring that chromosomes are fully replicated regardless of the exact number and location of origins. How bidirectional replication is faithfully established at each origin is a long-standing question in the field.

Eukaryotic replication occurs in a few highly regulated stages (Figure 1A). During G1, chromatin is "licensed" for replication, wherein two Mcm2-7 (mini-chromosome maintenance proteins 2–7) complexes are loaded onto double-stranded DNA (dsDNA) at each origin. Each dimer of Mcm2-7 forms a stable pre-replication complex (pre-RC).[1,2] During S phase, GINS[3] (Go-Ichi-Ni-San, the Japanese acronym for 5-1-2-3, short for its subunits Sld5, Psf1, Psf2, and Psf3) and Cdc45[4,5] (cell division cycle 45) are recruited to the pre-RC to assemble a pair of CMG (Cdc45•Mcm2-7•GINS) helicases.[6–8] Next, the two helicases are activated and each CMG begins to unwind DNA.[9,10] Finally, several other replication proteins (polymerases, processivity clamps, nucleases, structural scaffolds, etc.) are recruited to each helicase, forming a mature replisome.[11–13] Helicase assembly and activation involves several "firing factors" whose exact functions are under debate: TopBP1 (topoisomerase II binding protein 1), RecQL4 (RecQ-like helicase 4), DONSON (downstream neighbor of SON), TRESLIN-MTBP (TopBP1-interacting, replication-stimulating protein; Mdm2-binding protein), Mcm10 (mini-chromosome maintenance protein 10), GEMC1 (Geminin coiled-coil containing protein 1), DUE-B (DNA unwinding element binding protein), etc.[7,8]

TopBP1 is a large protein consisting of several BRCA1 C-terminal (BRCT) phosphopeptide-binding domains interspersed with disordered regions. TopBP1 is required for recruiting Cdc45 and GINS to chromatin, but its mechanism of action is unclear.[14,15] In addition to its role in initiation, TopBP1 is involved in DNA damage response as an allosteric activator of ATR (ataxia telangiectasia and Rad53 related).[16] TopBP1 oligomerization is critical for its ATR activator function,[17] but it is unclear how TopBP1 oligomerization regulates origin firing.

RecQL4 contains a C-terminal ATPase domain and an N-terminal domain distantly related to Sld2—an essential firing factor in yeast.[18] Mutations in RecQL4 are associated with Rothmund-Thomson syndrome (RTS), Baller-Gerold syndrome (BGS), and RAPADILINO syndrome.[19,20] RecQL4 was implicated in replication initiation, fork progression, DNA damage response, and telomere maintenance, making it especially challenging to interrogate its role in origin firing.[21] First, there are conflicting reports about RecQL4's function during initiation: RecQL4 was

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024



(legend on next page)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

implicated in CMG assembly,[22,23] as well as CMG activation.[24,25] Second, the function of RecQL4's ATPase domain is highly debated.[24–28] Finally, it was reported that RecQL4 facilitates fork progression on damaged DNA,[29] but it is unknown if RecQL4 travels with the replisome.

DONSON was recently identified as a regulator of initiation in metazoa with no direct homolog in yeast. The emerging consensus is that DONSON facilitates GINS recruitment to the pre-RC.[30–33] It was also proposed that DONSON travels with the replisome, acting as a fork protection factor.[34,35] How DONSON accomplishes these two distinct functions remains unclear.

Replication initiation is also regulated by CDK (cyclin-dependent kinase) and DDK (Dbf4-dependent kinase), which phosphorylate the pre-RC and several other replication factors.[8] In yeast, CDK is thought to promote the formation of a pre-loading complex (pre-LC)[36] consisting of Dpb11 (the yeast homolog of TopBP1), Sld2, GINS, and Polε (the leading strand polymerase). It was hypothesized that the pre-LC is recruited to Mcm2-7 during initiation.[36] A similar model was recently proposed for metazoa, with the putative pre-LC consisting of either TopBP1•DONSON•GINS•Polε[32] or DONSON•GINS•Cdc45•Polε.[33] However, there is no direct evidence showing that these proteins are recruited to DNA as a complex.

Origin firing remains the most poorly understood aspect of DNA replication. Specifically, how are two CMG helicases correctly assembled and activated at thousands of origins in each dividing cell? What is the exact function of each firing factor, and how are they coordinated? To systematically dissect the mechanism of metazoan replication initiation, we employed single-molecule imaging to directly monitor the recruitment of GINS, Cdc45, TopBP1, RecQL4, and DONSON to origins in *Xenopus* egg extracts.

Our data indicate that replication initiation is a highly dynamic process that entails the stepwise assembly of several short-lived protein complexes, many of which represent non-productive intermediates that do not result in origin firing. First, TopBP1 docks to the pre-RC and most likely serves as an adapter for subsequent GINS recruitment. Next, two copies of Cdc45 simultaneously bind to the pre-RC, suggesting that another replication factor acts as a dimerization scaffold for Cdc45. Shortly after Cdc45 arrival, a DONSON dimer loaded with two copies of GINS docks to the origin, completing the assembly of two inactive replicative helicases. DONSON remains stably bound until RecQL4 is recruited to the origin. We show that RecQL4's ATPase activity is critical for efficient CMG activation and DONSON release from CMG. Importantly, TopBP1, RecQL4, and DONSON all dissociate from the origin before the start of DNA synthesis, and none of them travel with the replication fork. We propose a model of replication initiation that helps explain the high fidelity of bidirectional origin firing in metazoa.

RESULTS

A single-molecule assay to visualize replication initiation in real time

To directly visualize replication initiation, we adapted a single-molecule imaging approach called KEHRMIT (kinetics of the eukaryotic helicase by real-time molecular imaging and tracking).[37] First, λ DNA was flow-stretched and immobilized in a microfluidic flow cell (Figure 1Bi). Next, DNA was incubated with high-speed supernatant (HSS) of cytoplasmic extract from *Xenopus laevis* eggs (Figure 1Bii). HSS supports efficient loading of Mcm2-7 onto DNA without initiating replication. Since metazoan origins are not defined by conserved DNA sequences, pre-RCs are loaded at random locations on λ DNA. Licensed DNA was incubated with nucleo-plasmic extract (NPE), which recapitulates S phase and supports efficient DNA replication (Figure 1Biii).[38] GINS-depleted NPE was supplemented with $GINS^{AF647}$ (GINS labeled with Alexa Fluor [AF] 647) and $Fen1^{mKikGR}$ (a marker of DNA synthesis[39]). Replication was monitored using a total internal reflection fluorescence (TIRF) microscope.[40]

Figure 1. A dimer of GINS and a dimer of Cdc45 are recruited to each Mcm2-7 double hexamer

(A) Overview of replication initiation in metazoa.

(B) Single-molecule workflow to visualize $GINS^{AF647}$ (green) recruitment during replication initiation. $Fen1^{mKikGR}$ (blue) binds nascent lagging strands. Other replication proteins are not shown for clarity.

(C) Representative kymogram showing $GINS^{AF647}$ (green) recruitment during initiation. $Fen1^{mKikGR}$ (blue) serves as a marker of DNA synthesis. Inset: integrated $GINS^{AF647}$ signal (black) and a changepoint fit to the raw signal (green).

(D) Representative kymogram showing $Cdc45^{AF647}$ (green) recruitment during initiation. Inset: integrated $Cdc45^{AF647}$ signal (black) and the changepoint fit to the raw signal (green). A biochemical assay was used to verify that fluorescent labeling did not impair Cdc45's function (Figures S1E–S1G).

(E) Origin firing efficiency for $GINS^{AF647}$ (light blue) and $Cdc45^{AF647}$ (dark blue) titration experiments.

(F) Breakdown of how two copies of GINS were recruited to the origin. At most origins, two GINS molecules appear simultaneously (0 frames, green). At the remaining origins, two copies of GINS show up sequentially, where the second molecule appears 1 frame (yellow), 2 frames (orange), or 3+ frames (red) after the first GINS.

(G) Breakdown of how two copies of Cdc45 were recruited to the origin.

(H) Kymogram illustrating the time delay between $GINS^{AF647}$ binding and origin firing.

(I) Time delay between protein binding and origin firing for $Cdc45^{AF647}$ and $GINS^{AF647}$ (blue bar marks the median; gray box: 95% confidence interval [CI] for the median; p values were computed using the two-sample Kolmogorov-Smirnov test; * denotes $p \leq 0.05$).

(J) Quantification of $GINS^{AF647}$ and $Cdc45^{AF647}$ binding to DNA in indicated reactions (see "Global protein binding analysis" in STAR Methods).

(K) Representative kymograms illustrating short-lived docking of $Cdc45^{AF647}$ to DNA in the GINS-depleted reaction.

(L) Model illustrating how two copies of GINS or Cdc45 could be simultaneously recruited to the pre-RC.

In all panels, error bars denote the 95% CI estimated via bootstrapping. In panels (E) and (J), N denotes the number of DNA molecules. In panels (F), (G), and (I), N marks the number of origins.

See also Figure S1.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

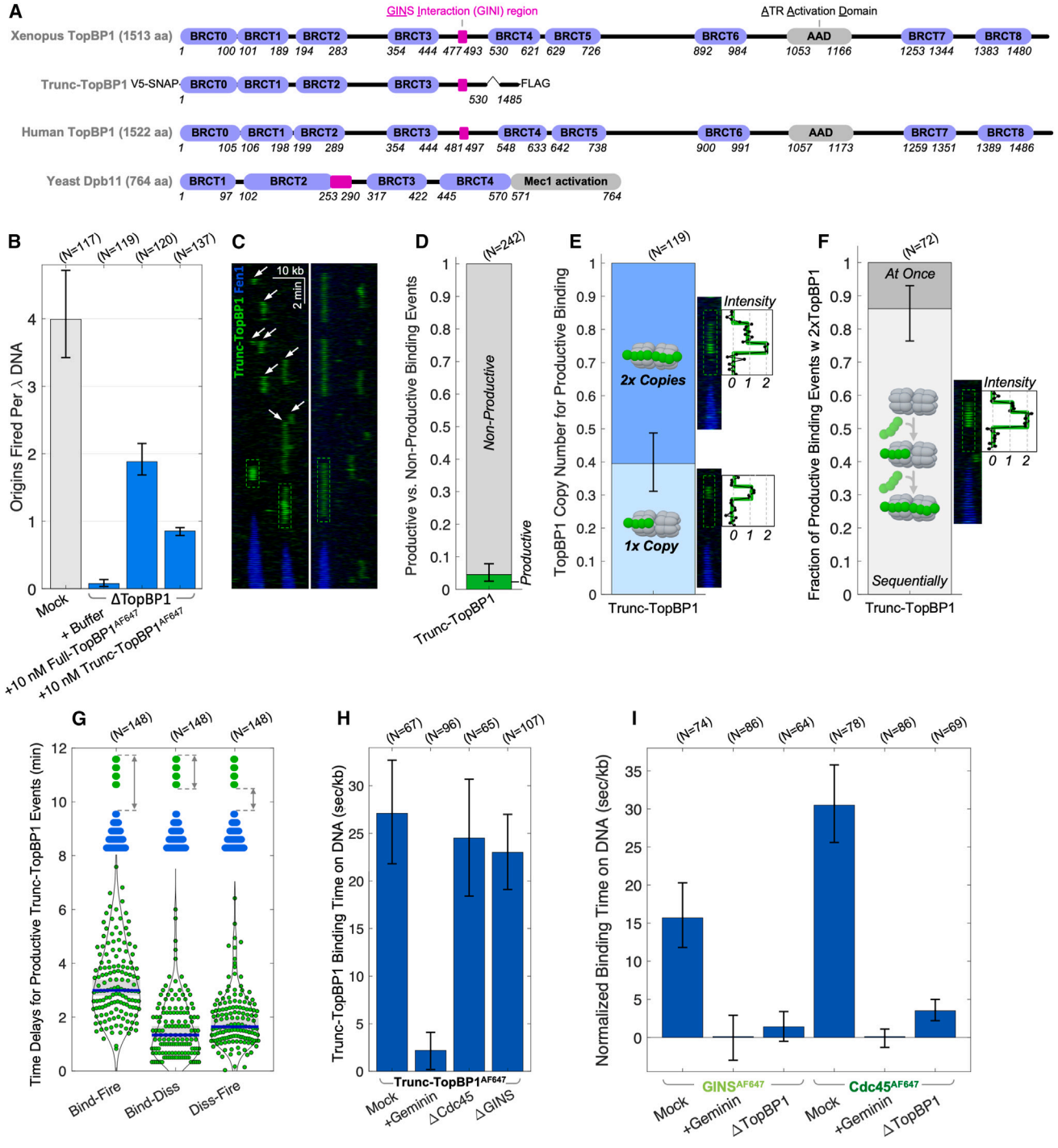


Figure 2. TopBP1 is recruited to the pre-RC before GINS and Cdc45

(A) Domain map of full-length and truncated *Xenopus* TopBP1 versus human TopBP1 and yeast Dpb11.

(B) Origin firing efficiency for full-length TopBP1^{AF647} and truncated TopBP1^{AF647} compared to mock-depleted and TopBP1-depleted reactions.

(C) Kymograms showing truncated TopBP1^{AF647} (green) binding during initiation. Dashed boxes mark productive binding events, white arrows mark non-productive binding. The SNAP tag used for fluorescent labeling did not impair trunc-TopBP1's function (Figures S2F–S2H), and >95% of trunc-TopBP1 was fluorescently labeled (Figure S2E).

(D) Probability of productive (green) versus non-productive (gray) trunc-TopBP1^{AF647} binding events.

(E) Fraction of productive binding events that contain one or two copies of trunc-TopBP1^{AF647}. Insets show example kymograms with integrated signal.

(legend continued on next page)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

A dimer of GINS and a dimer of Cdc45 are recruited to each Mcm2-7 double hexamer

It has long been known that DNA is replicated bidirectionally from each origin of replication.[41] How bidirectionality is established and enforced during origin firing is not known. To address this question, we prepared active GINS^{AF647} and Cdc45^{AF647} (Figures S1A–S1G) and visualized their recruitment to DNA.

Figure 1C shows a representative kymogram of GINS^{AF647} recruitment to the origin. Shortly after binding to the pre-RC, two GINS^{AF647} molecules began moving bidirectionally away from the origin, indicating that they were successfully incorporated into active helicases (Figure 1C, green channel). Each initiation event was followed by a growing replication bubble (Figure 1C, blue channel). Intriguingly, two GINS molecules were recruited at once (Figure 1C, inset). Next, we monitored Cdc45^{AF647} recruitment to the pre-RC using the same workflow. Two copies of Cdc45^{AF647} were simultaneously recruited to the origin (Figure 1D).

We initially hypothesized that two GINS are recruited sequentially but in rapid succession, making their binding appear simultaneous. If so, the recruitment of the second molecule should be resolvable when GINS binding is rate-limiting. Lowering GINS^{AF647} concentration from 20 nM to 5 nM dramatically reduced the probability of origin firing (Figure 1E, light blue bars), suggesting that in these experiments GINS was rate-limiting for initiation. However, even at the lowest concentration sampled, two GINS copies were recruited within the same movie frame (Figure 1F, see 5 nM GINS). Similarly, making Cdc45^{AF647} rate-limiting for origin firing did not significantly change the probability that two Cdc45 were recruited simultaneously (Figure 1G). Although we cannot strictly rule out that Cdc45/GINS recruitment is sequential and highly cooperative, our data suggest that Cdc45/GINS are loaded as pre-formed dimers.

To determine their recruitment order, we made extensive efforts to simultaneously visualize GINS^{AF568} and Cdc45^{AF647}. However, at the GINS/Cdc45 concentrations compatible with single-molecule imaging, replication initiated very inefficiently. Instead, we inferred recruitment order by comparing the time delay between protein binding and origin firing (Figure 1H). Although the distribution of binding-firing delays was broad, Cdc45 arrived at the origin ~15 s before GINS (Figure 1I). This difference is highly reproducible between different extract preparations and different Cdc45 and GINS concentrations (Figures S1H–S1J). Consistent with this observation, GINS was not recruited to licensed DNA in Cdc45-depleted extract (Figure 1J). Interestingly, although Cdc45^{AF647} recruitment to DNA was suppressed in GINS-depleted extract, it was not completely eliminated (Figure 1J). Kymograms of individual DNA molecules reveal that in the absence of GINS, Cdc45^{AF647} transiently docked to DNA and typically resided on DNA for only ~10 s (Figure 1K). These data suggest that, in metazoa, Cdc45 transiently docks to the pre-RC first but requires subsequent GINS recruitment to remain stably bound to Mcm2-7, consistent with previous reports in yeast.[12,42]

Since neither GINS nor Cdc45 alone can dimerize *in vitro* (Figures S1A and S1B), we hypothesized the existence of distinct dimerization scaffolds for GINS and Cdc45 (Figure 1L). These putative scaffolds should satisfy a few key requirements: they must dimerize, they must directly bind their client protein (GINS or Cdc45), and they must bind Mcm2-7 and/or another protein that docks to the pre-RC. Several replication firing factors fit these criteria, including TopBP1, RecQL4, and DONSON.

TopBP1 is recruited to the pre-RC before GINS or Cdc45

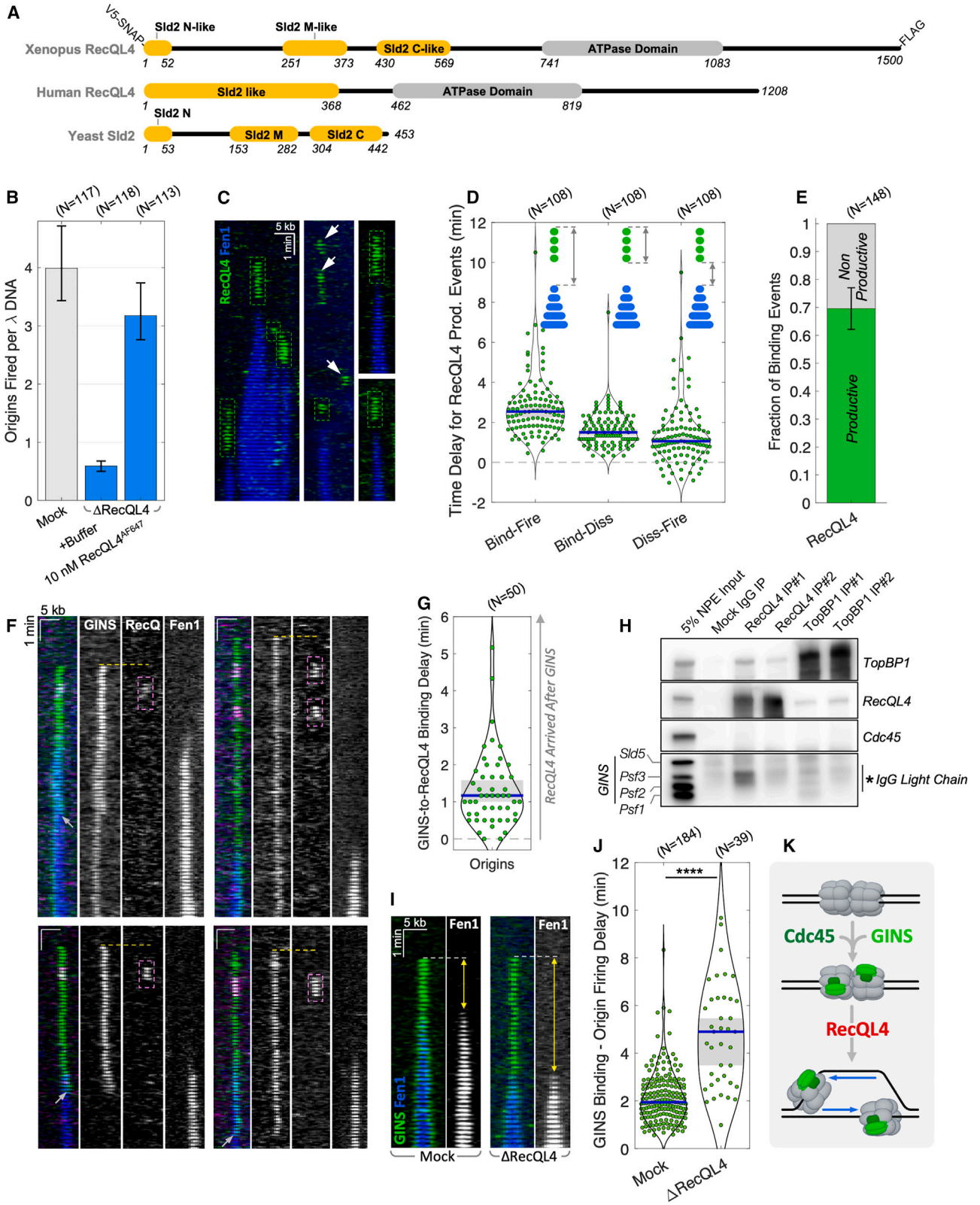
We initially regarded TopBP1 as the most promising dimerization scaffold candidate because it can bind both GINS and Cdc45.[43–45] Moreover, the yeast ortholog of TopBP1 (Dpb11) is thought to be a subunit of the pre-LC proposed to recruit GINS to the pre-RC.[36] To test this idea, we visualized TopBP1 binding to licensed DNA using the same workflow as for GINS. Since full length TopBP1^{AF647} formed large assemblies that hindered our ability to analyze and interpret the data (Figures S2A–S2C), we prepared a truncated construct (Figures S2D and S2E) that lacks BRCT4-8 and the ATR activation domain[16,46] (Figure 2A, trunc-TopBP1). The truncated protein retained its origin firing activity (Figures 2B and S2F–S2H), suggesting it was a good surrogate for full-length TopBP1.

TopBP1 recruitment was surprisingly dynamic—it repeatedly bound to and dissociated from DNA, sometimes at the same location (Figure 2C). These transient binding events lasting 1–2 min were observed only on licensed DNA (Figure S2I), indicating that TopBP1 is specifically recruited to pre-RCs, consistent with a previous report.[46] Importantly, origin firing occurred shortly after trunc-TopBP1^{AF647} dissociation (Figure 2C, green boxes), suggesting that TopBP1 binding and dissociation are important steps in replication initiation. Interestingly, most trunc-TopBP1^{AF647} binding events were non-productive—i.e., did not trigger origin firing (Figure 2C white arrows; Figure 2D). Finally, trunc-TopBP1^{AF647} always dissociated before the start of DNA synthesis and did not travel with the fork, indicating that TopBP1 is not a constitutive replisome subunit.

If TopBP1 did serve as the dimerization scaffold for GINS (or Cdc45), it should be recruited to the pre-RC as a pre-formed dimer in complex with GINS (or Cdc45). To determine TopBP1 stoichiometry during initiation, we measured the intensity of productive binding events in the trunc-TopBP1^{AF647} experiment (>90% of the trunc-TopBP1 was fluorescently labeled [Figure S2E], with exactly one AF647 per protein). A significant fraction (~40%) of origin firing events was facilitated by a single copy

(F) For all productive binding events with two copies of trunc-TopBP1^{AF647}, breakdown of how the two molecules were recruited.
(G) Distributions of the measured time delay between protein binding and origin firing for trunc-TopBP1^{AF647} (blue bar: median, gray box: 95% CI).
(H) Quantification of trunc-TopBP1^{AF647} binding time on DNA.
(I) Quantification of GINS^{AF647} and Cdc45^{AF647} binding time on DNA.
In all panels, error bars denote the 95% CI. In panels (B), (H), and (I), N denotes the number of DNA molecules. In (D)–(F) and (G), N marks the number of binding events.
See also Figure S2.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024



(legend on next page)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

of trunc-TopBP1^{AF647} (Figure 2E). When two trunc-TopBP1^{AF647} copies were detected, their recruitment was sequential in most cases (Figure 2F), suggesting that trunc-TopBP1 did not form stable dimers in our experiments. The photobleaching assay also revealed that the vast majority of full-length TopBP1 and truncated TopBP1 are monomeric *in vitro* (Figure S2J). These results indicate that a single copy of trunc-TopBP1 is sufficient to support bidirectional origin firing and suggest that TopBP1 does not enforce 2-fold symmetry during initiation.

Simultaneous imaging of TopBP1^{AF546} and GINS^{AF647} (or Cdc45^{AF647}) was not feasible due to poor origin firing efficiency at the protein concentrations required for single-molecule detection. Instead, we resorted to measuring the time delay between protein binding and origin firing. Both trunc-TopBP1^{AF647} and full-length TopBP1^{AF647} bound to DNA ~3 min before origin firing (Figures 2G and S2K), i.e., ~1 min before Cdc45/GINS recruitment (see Figure 1I), suggesting that TopBP1 acts upstream of GINS/Cdc45 binding. Consistent with this finding, depleting GINS or Cdc45 did not impair trunc-TopBP1^{AF647} binding to licensed DNA (Figure 2H). Conversely, depleting TopBP1 abolished GINS^{AF647} binding to licensed DNA, and strongly suppressed Cdc45^{AF647} binding (Figure 2I). Like the GINS-depleted extract (Figure 1K), TopBP1-depleted extract supported only short-lived docking of Cdc45^{AF647} (most events lasted 10 s or less), suggesting that TopBP1 and GINS are required to stabilize Cdc45's weak binding to Mcm2-7. Taken together, our data indicate that TopBP1 is recruited to the pre-RC before GINS or Cdc45 and therefore cannot serve as a dimerization scaffold for either protein.

RecQL4 arrives at the origin after GINS recruitment and facilitates CMG activation

RecQL4 has several properties predicted for the GINS dimerization scaffold: RecQL4 forms stable dimers *in vitro* (Figures S3A–S3B); RecQL4 is specifically recruited to licensed chromatin[24]; and RecQL4 is a distant homolog of Sld2 (Figure 3A), which has been implicated in recruiting GINS to origins in yeast.[36]

To interrogate RecQL4's function during initiation, we monitored origin firing in a RecQL4-depleted reaction supplemented with 10 nM of RecQL4^{AF647} (Figures 3B and S3C–S3F). Like TopBP1, RecQL4's behavior was very dynamic: it transiently bound to DNA shortly before origin firing (Figures 3C and 3D), but unlike TopBP1, most RecQL4 binding events were productive (Figure 3E). Importantly, RecQL4 did not travel with replication forks, indicating that it is not a constitutive replisome subunit.

To directly test the hypothesis that RecQL4 is a dimerization scaffold for GINS, we simultaneously visualized GINS^{AF647} and RecQL4^{AF546} recruitment to origins during initiation. Both GINS^{AF647} and RecQL4^{AF546} were clearly detected in 50 origin firing events, and in nearly all cases RecQL4 was recruited after GINS (Figures 3F and 3G). This finding indicates that RecQL4 is not the GINS dimerization scaffold. Consistent with this conclusion, neither GINS nor Cdc45 coimmunoprecipitated efficiently with RecQL4 (or TopBP1) as would be expected for a dimerization scaffold (Figure 3H).

Since RecQL4 is recruited after GINS but before the start of DNA synthesis, we hypothesized that RecQL4 plays a role in CMG activation. If so, making RecQL4 rate-limiting for initiation should prolong the time delay between GINS recruitment and the start of DNA replication. Depleting RecQL4 led to a ~2.5-fold increase in the delay between GINS^{AF647} recruitment and origin firing (Figures 3I and 3J), strongly suggesting that RecQL4 facilitates CMG activation (Figure 3K).

DONSON and GINS simultaneously arrive at the origin

Several recent studies have shown that DONSON is an essential factor for replication initiation in metazoa, that it forms dimers, and that it mediates GINS recruitment to origins.[30–33] We independently verified that recombinant DONSON forms dimers (Figures S4A–S4C) and that DONSON interacts with GINS (Figures 4A and 4B).

To directly interrogate DONSON's function, we visualized origin firing in a DONSON-depleted reaction supplemented with 10 nM of DONSON^{AF647} (Figures 4C and S4D–S4G). Like TopBP1 and RecQL4, DONSON^{AF647} transiently bound to DNA shortly before replication initiation (Figure 4D green boxes; Figure 4E). ~35% of DONSON binding events were non-productive and failed to trigger origin firing (Figure 4D white arrows; Figure 4F). Importantly, the delay between DONSON binding and origin firing was nearly identical to that measured for GINS

Figure 3. RecQL4 arrives at the origin after GINS recruitment and facilitates CMG activation

(A) Domain maps of *Xenopus* and human RecQL4 versus yeast Sld2.

(B) Origin firing efficiency for RecQL4^{AF647} compared to mock-depleted and RecQL4-depleted reactions.

(C) Kymograms showing the recruitment of RecQL4^{AF647} (green) during replication initiation. Green boxes mark productive binding, white arrows mark non-productive binding. The SNAP tag did not impair RecQL4's function (Figures S3C–S3E), and >90% of RecQL4 was fluorescently labeled (Figure S3F).

(D) Measured time delays for productive RecQL4 binding events.

(E) Probability of productive (green) versus non-productive (gray) RecQL4^{AF647} binding events.

(F) Representative kymograms illustrating the recruitment order of GINS^{AF647} (green) and RecQL4^{AF546} (magenta). At some origins, only one GINS molecule is fluorescently labeled; gray arrows mark GINS^{AF647} photobleaching.

(G) Measured time delays between GINS recruitment and RecQL4 binding for origins where both proteins were detected (positive values correspond to RecQL4 arriving after GINS).

(H) Immunoblots of proteins immunoprecipitated from NPE using RecQL4 and TopBP1 antibodies raised against two non-overlapping fragments of each protein.

(I) Representative kymograms illustrating the delay between GINS binding and origin firing in mock versus RecQL4-depleted reactions.

(J) Time delay between GINS^{AF647} binding and origin firing in mock versus RecQL4-depleted extract (**** denotes $p \leq 0.0001$).

(K) Model depicting the timing of RecQL4 arrival during replication initiation.

In panels (D), (G), and (J), blue bars denote medians, gray boxes: 95% CI. In panels (B) and (E), error bars denote 95% CI. In panel (B), N marks the number of DNA molecules. In panels (D) and (E), N denotes the number of binding events. In panels (G) and (J), N represents the number of origins.

See also Figure S3.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

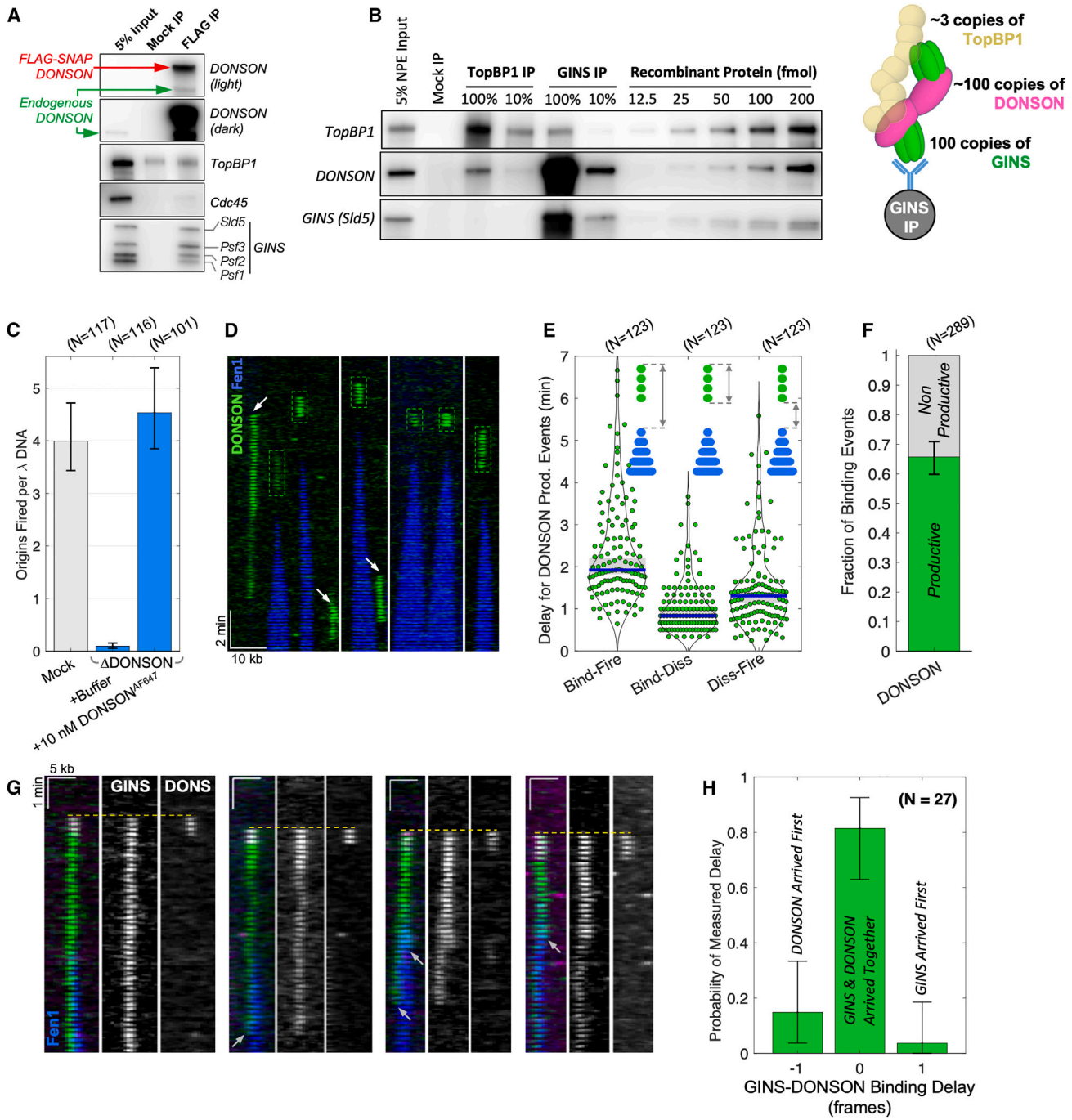


Figure 4. DONSON and GINS are simultaneously recruited to the pre-RC

(A) Immunoblots of proteins immunoprecipitated with FLAG-SNAP-DONSON from NPE.

(B) Quantitative immunoblots of proteins immunoprecipitated from NPE with TopBP1 or GINS versus known amounts of recombinant protein. Cartoon illustrates the approximate GINS coIP stoichiometry.

(C) Origin firing efficiency for DONSON^{AF647} compared to mock-depleted and DONSON-depleted reactions.

(D) Representative kymograms showing transient DONSON^{AF647} binding (green) during replication initiation. Dashed boxes mark productive binding, white arrows mark non-productive binding. Recombinant DONSON efficiently supported DNA replication in bulk (Figures S4E–S4G).

(E) Time delays for productive DONSON binding events (blue bars denote medians, gray boxes: 95% CI).

(F) Probability of productive (green) versus non-productive (gray) DONSON^{AF647} binding events.

(legend continued on next page)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

(Figure 4E; compare to Figure 1I), suggesting that DONSON and GINS are recruited to the origin together.

If DONSON is the GINS dimerization scaffold, two copies of GINS and two copies of DONSON should be simultaneously recruited to the origin. To test this hypothesis, we depleted GINS and DONSON from extract, supplemented the reaction with 10 nM DONSON^{AF546} and 30 nM GINS^{AF647}, and visualized their recruitment in the same experiment. Both GINS^{AF647} and DONSON^{AF546} signals were clearly detectable in 27 origin firing events (Figure 4G), and in ~80% of them DONSON and GINS were recruited to DNA at the same time (Figure 4H). Our data support a model where a DONSON dimer ensures the simultaneous delivery of two GINS molecules to the origin (Figure 1L).

A recent study proposed that DONSON also mediates Cdc45 recruitment to origins.[33] Unlike GINS, Cdc45 did not co-precipitate with DONSON (Figure 4A), suggesting that a different dimerization scaffold is responsible for recruiting two copies of Cdc45.

Long-lived non-productive GINS/Cdc45 binding events represent assembled but inactive helicases

The recruitment of two GINS^{AF647} molecules to DNA resulted in origin firing only ~55% of the time, and we call these events "productive GINS binding" (Figure 5A, green bars). In the remaining ~45% of events, two GINS^{AF647} were stably recruited to DNA without initiating replication, and we call them "non-productive" (Figure 5A, gray bars). These non-productive 2xGINS^{AF647} binding events were very long-lived (median ~7 min, Figures 5B and 5C), and were detected only when Cdc45 was present in extract (Figure S5A). Importantly, similar long-lived non-productive 2xCdc45^{AF647} binding events were observed only when GINS was present in extract (Figures S5B–S5D).

Such long-lived non-productive 2xGINS (or 2xCdc45) binding events may represent pairs of assembled but inactive CMG helicases. Consistent with this idea, the fraction of non-productive 2xGINS^{AF647} binding events increased to >90% when replication initiation was supressed by depleting RecQL4 (Figure 5A). Taken together, these observations suggest that long-lived non-productive 2xGINS (or 2xCdc45) binding events represent pairs of assembled helicases whose activation requires RecQL4 (Figure 5D).

In addition to the long-lived non-productive binding events, origin firing was often preceded by one or more very short-lived (median ~10 s) non-productive GINS^{AF647} docking events (Figure 5E, red arrows; Figures 5F and S5E). Interestingly, most of these events contained a single copy of GINS^{AF647} (Figure 5G). We observed similar short-lived docking events containing a single copy of Cdc45^{AF647} (Figures S5E–S5H). These observations indicate that although GINS/Cdc45 monomers can bind to the origins, they dissociate quickly and do not lead to unidirectional initiation (Figure 5H).

RecQL4 promotes DONSON dissociation from CMG

When visualizing DONSON recruitment during replication initiation, we frequently observed long-lived non-productive binding events that contained two DONSON^{AF647} molecules (Figure 6A, white arrows). In GINS+DONSON imaging experiments, non-productive DONSON^{AF546} recruitment was accompanied by long-lived binding of two GINS^{AF647} molecules (Figures 6B and S6A). This suggests that after delivering GINS to the origin, DONSON remains stably bound to inactive CMGs that await activation. Consistent with this idea, depleting RecQL4 sharply increased the fraction of non-productive DONSON binding (Figure 6C white arrows; compare to Figure 6A) to the same extent as non-productive GINS binding (Figure 6D; compare to Figure 5A). This defect was robustly rescued by supplementing RecQL4-depleted extract with 100 nM of wild-type recombinant RecQL4 (Figures 6D and S6B). Importantly, depleting RecQL4 led to a large increase in the time delay between DONSON binding and origin firing (Figures 6E and S6C). Most of this effect is due to the slow dissociation of DONSON from the origin in the RecQL4 depletion (Figures 6F, 6G, S6D, and S6E). Taken together, these results indicate that RecQL4 promotes DONSON dissociation from CMG during origin firing (Figure 6H).

The ATPase activity of RecQL4 is required to efficiently evict DONSON from CMG

RecQL4 shares some sequence similarity with the yeast Sld2, but unlike Sld2, RecQL4 has an ATPase domain that is conserved in metazoa[24,47] (Figure 3A). Due to conflicting reports about the exact function of this domain, we sought to interrogate the role of RecQL4's ATPase activity in replication initiation.

First, we purified recombinant *Xenopus laevis* RecQL4 bearing a Walker B motif mutation (RecQL4-D856A) that blocks its ability to hydrolyze ATP. In an ensemble assay, RecQL4-D856A rescued plasmid replication in RecQL4-depleted extract (Figure 6I), albeit slower than wild-type recombinant RecQL4 (Figure 6J; compare the time needed to reach 50% replication). Surprisingly, the replication activity at 20 nM RecQL4 was indistinguishable from that at 100 nM, suggesting that even low RecQL4 concentrations are not rate-limiting in the biochemical assay (Figures 6I–6K).

Next, we visualized DONSON^{AF647} recruitment and origin firing in RecQL4-depleted extract supplemented with 100 nM recombinant RecQL4-D856A. In the single-molecule assay, RecQL4-D856A supported a lower origin firing efficiency than the wild-type protein, but higher than the RecQL4-depleted extract (Figures 6L and S6F). The fraction of productive DONSON binding events was only ~11% in the RecQL4-D856A reaction compared to ~46% in the wild-type RecQL4 experiment and only ~4% in the depleted extract (Figure 6D), suggesting that

(G) Representative kymograms illustrating the relative timing of GINS^{AF647} and DONSON^{AF546} recruitment. At some origins, only one GINS molecule is fluorescently labeled; gray arrows mark GINS^{AF647} photobleaching.

(H) Measured time delay between GINS recruitment and DONSON binding (time is shown in movie frames, 1 frame = 10 s). N denotes the number of origins where both GINS and DONSON were detected.

In panels (C), (F), and (H), error bars denote 95% CI. In panel (C), N marks the number of DNA molecules. In panels (E) and (H), N denotes the number of origins. In panel (F), N denotes the number of binding events.

See also Figure S4.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

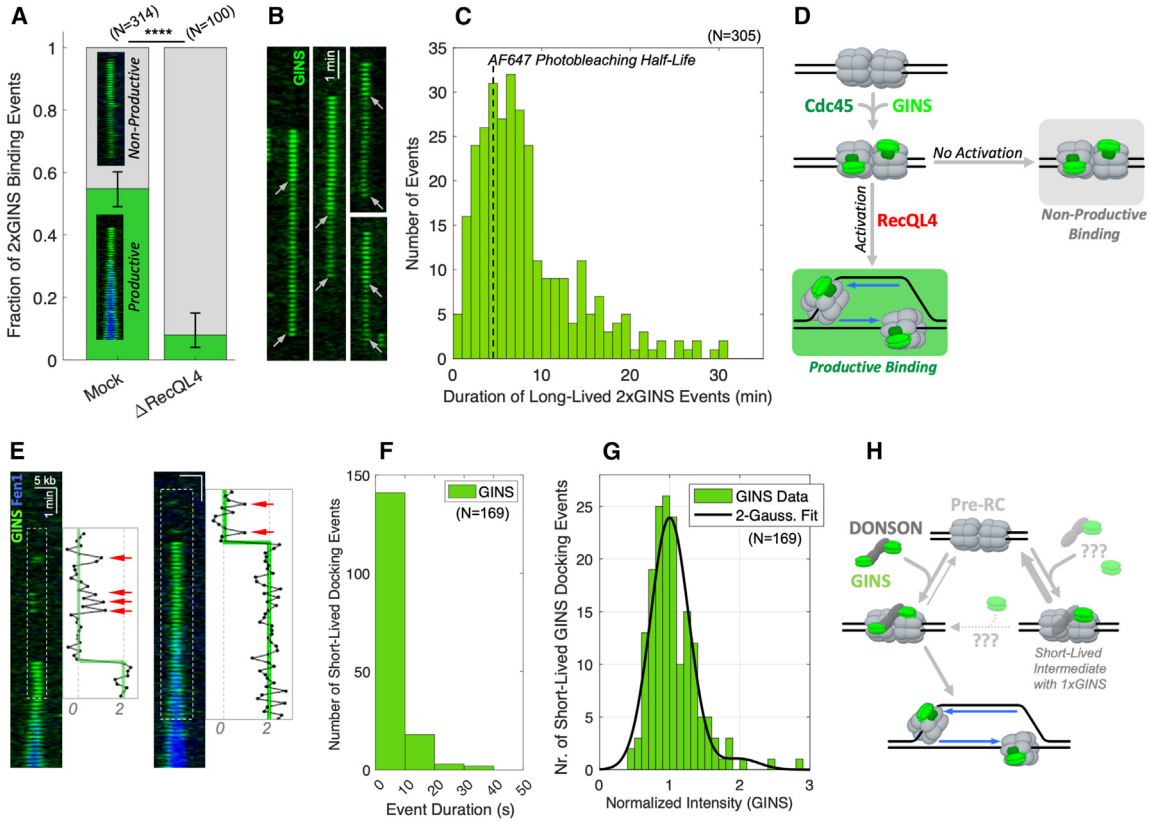


Figure 5. Long-lived non-productive GINS binding events represent assembled but inactive helicases

(A) Fraction of productive (green) and non-productive (gray) dual-GINS recruitment events in mock versus RecQL4-depleted extract (**** denotes $p \leq 0.0001$, error bars denote 95% CI).

(B) Representative kymograms showing non-productive dual-GINS^{AF647} recruitment (gray arrows mark GINS^{AF647} photobleaching).

(C) The duration of long-lived non-productive dual-GINS^{AF647} binding events.

(D) Model illustrating the difference between productive and non-productive long-lived GINS/Cdc45 binding events.

(E) Kymograms showing short-lived docking of GINS^{AF647} before origin firing (inset, red arrows).

(F) The duration of short-lived GINS docking events.

(G) The normalized intensity of short-lived GINS docking events. Black curve shows the fit to two Gaussians centered at 1.0 and 2.0, respectively.

(H) Model depicting the short-lived non-productive recruitment of a single GINS molecule versus the productive recruitment of two GINS copies (thicker arrows depict faster rate constants). It is unclear if short-lived docking events are mediated by DONSON or whether GINS alone can bind to the origin.

In panels (A), (C), (F), and (G), N denotes the number of binding events.

See also Figure S5.

RecQL4's ATPase activity plays a role in replication initiation. Importantly, the RecQL4-D856A mutant did not rescue the defect in DONSON dissociation before origin firing (Figure 6E), indicating that RecQL4's ATPase activity is required to efficiently evict DONSON from CMG. These data clearly implicate RecQL4's ATPase activity in replication initiation.

DISCUSSION

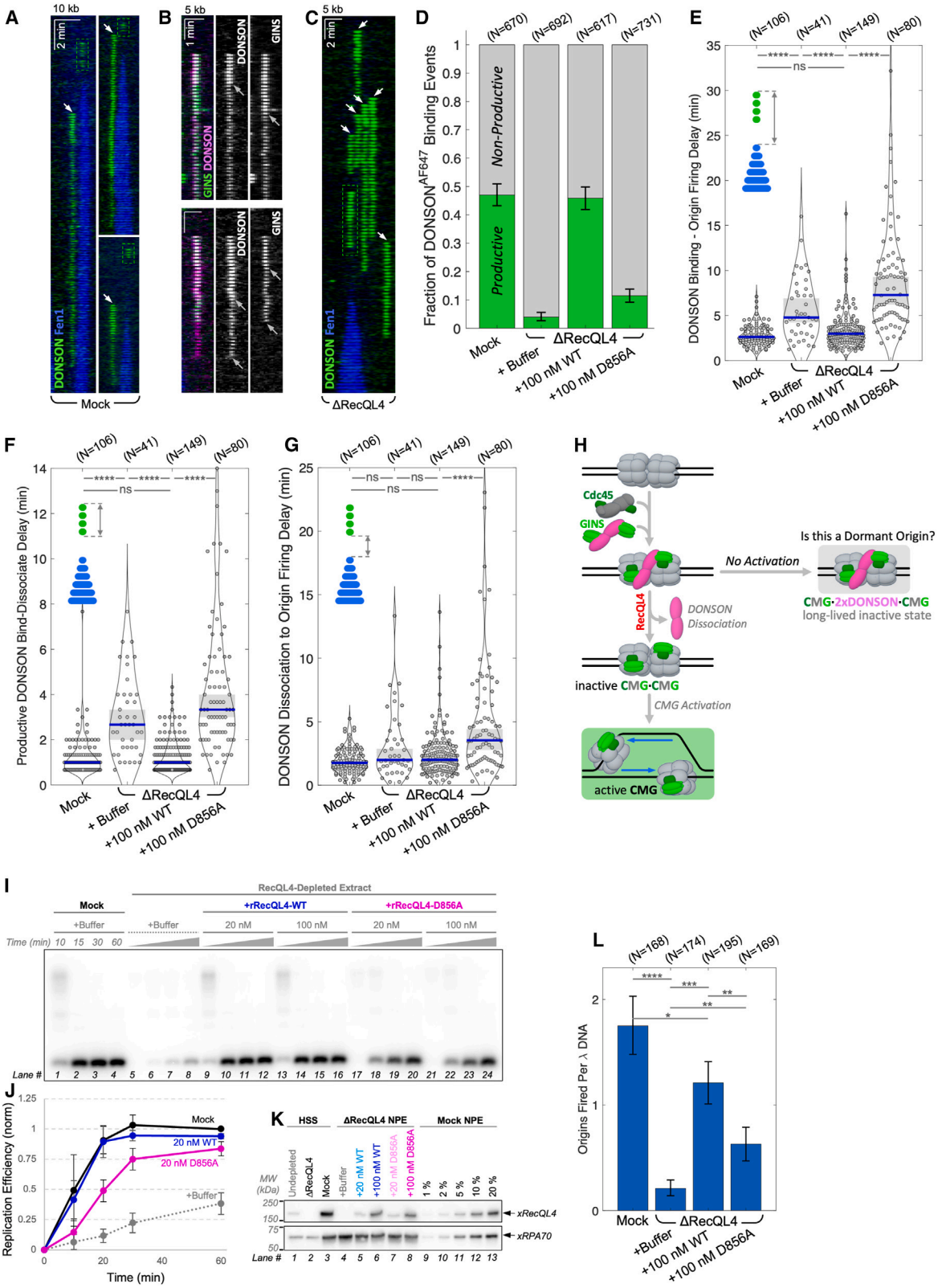
Model of bidirectional replication initiation in metazoa

In this study, we leveraged single-molecule imaging to visualize replication initiation in real time. Our experiments revealed a surprisingly dynamic process governed by a network of transient protein-protein interactions that previously confounded the interpretation of ensemble experiments. Our observations suggest a model of replication initiation where several short-lived intermediates are needed to correctly assemble and activate two CMG helicases at each origin (Figure 7A).

Proper assembly of two CMGs at each origin is achieved by simultaneously recruiting two Cdc45 molecules, and briefly afterward—two GINS molecules. The delivery of these proteins to the origin and their incorporation into replicative helicases is orchestrated by several firing factors. By directly visualizing three key firing factors (TopBP1, DONSON, and RecQL4), we revealed that these proteins transiently dock to the origin but do not travel with the replication fork. This paints a clearer picture of their function and reconciles previous conflicting reports.

Our experiments strongly suggest that TopBP1 is recruited to Mcm2-7 before Cdc45 or GINS. Consistent with this conclusion, careful quantification of protein co-immunoprecipitation (coIP) from egg extract suggests that TopBP1 does not form a stable stoichiometric complex with DONSON and GINS (Figure 4B).

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024



(legend on next page)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

Our findings are inconsistent with the currently accepted pre-LC model of initiation.[32,33,36] The pre-LC model originated from a report that Dpb11, GINS, Polε, and Sld2 weakly co-immunoprecipitated from yeast lysate and that chemical cross-linking was required to efficiently co-immunoprecipitate these proteins.[36] An alternative interpretation is that the coIP of Dpb11, GINS, Polε, and Sld2 in the presence of a cross-linker reflects a network of pairwise protein interactions (GINS-Dpb11, Dpb11-Sld2, etc.) rather than a specific stoichiometric complex.

Once bound to the pre-RC, TopBP1 may act as a "docking adapter" to recruit the GINS•DONSON•DONSON•GINS complex from solution. Indeed, TopBP1 is required for GINS binding to pre-RCs (Figure 2I), and the interaction between TopBP1 and DONSON is important for recruiting GINS and DONSON to chromatin.[32] Since TopBP1 contains several BRCT phosphopeptide-binding modules, we speculate that TopBP1 acts as a "reader" of Mcm2-7 phosphorylation (for example, TopBP1 may bind phosphorylated pre-RCs more stably, increasing the likelihood of origin firing). We found that TopBP1 dimerization is not required for replication initiation (Figure 2E). It is possible that TopBP1 dimerizes only at high concentrations but remains monomeric at low concentrations, which are nevertheless sufficient to support origin firing. This may explain why knocking down TopBP1 in human cells did not impair DNA replication but fully suppressed ATR signaling,[22] which requires TopBP1 oligomerization.[16,17,48]

We showed that DONSON acts as a dimerization scaffold for GINS, ensuring the simultaneous delivery of two copies of GINS to the pre-RC. The recruitment of GINS•DONSON•DONSON•GINS to the origin is the last step in CMG assembly. Our data indicates that DONSON remains stably bound to the inactive CMG dimer until RecQL4 promotes DONSON dissociation, with origin firing occurring shortly thereafter. It is unclear exactly how RecQL4 promotes origin firing: it may directly facilitate DONSON dissociation from CMG, or DONSON dissociation may be a byproduct of CMG activation by RecQL4.

DONSON does not travel with the replication fork

DONSON was initially proposed to travel with replication forks and stabilize them in response to DNA damage.[34,35] Knocking down DONSON led to a significant increase in replication fork asymmetry due to spontaneous fork stalling.[34] It is plausible that DONSON prevents the fork asymmetry observed in fiber assays by promoting efficient origin firing near stalled replication forks (Figure 7B).

More recently, chromatin binding assays were used to infer that DONSON is a replisome subunit.[31,33] By comparing structures of CMG•DONSON[30] and CMG•Polα[49] (the polymerase that primes DNA synthesis), the Labib group recently proposed that DONSON binding to CMG would prevent Polα recruitment to the replisome. The Walter group arrived at a similar conclusion by examining AlphaFold-generated models of DONSON bound to Mcm2-7.[32] Our data clearly show that DONSON dissociates from the origin before the start of DNA synthesis. Our single-molecule experiments combined with recent structural studies[30,32,50] point to DONSON dissociation being a pre-requisite for origin firing.

Our data suggest that the CMG•DONSON•DONSON•CMG intermediate complex is very stable on chromatin. It was hypothesized that dormant origins[51] may contain partially assembled pre-initiation intermediates that remain inactive until they are activated in response to DNA damage (via ATR signaling).[52] It is tempting to speculate that the CMG•DONSON•DONSON•CMG intermediates detected in our experiments represent dormant origins. This may explain why a significant fraction of long-lived 2xGINS (or 2xDONSON) binding events do not lead to origin firing (Figures 5A and 6D, gray bars), even though egg extracts contain a large excess of replication initiation factors. Interestingly, in rare cases, we observed that active replication forks "pushed" long-lived intermediates containing DONSON (Figure 7C). Finally, it is unclear how the inactive CMG•DONSON•DONSON•CMG complex is removed from DNA—via passive dissociation or active extraction (perhaps through a ubiquitin-dependent pathway).

RecQL4's ATPase activity regulates replication initiation

The role of RecQL4's ATPase domain in genome maintenance has been highly debated. RecQL4's ATPase activity was reported to be essential for viability in *Drosophila*,[26] but it was not required for the proliferation of chicken DT40 cells.[28] Mice with ATPase-dead RecQL4 exhibited no detectable defects in development, survival, or DNA damage sensitivity,[27] whereas

Figure 6. RecQL4 facilitates DONSON dissociation from CMG

(A) Representative kymograms showing non-productive (white arrows) and productive (green boxes) DONSON^{AF647} recruitment.

(B) Representative kymograms showing non-productive dual-DONSON^{AF546} (magenta) binding accompanied dual-GINS^{AF647} (green) recruitment (gray arrows mark photobleaching).

(C) Representative kymogram illustrating DONSON binding to licensed DNA in RecQL4-depleted extract (green box: productive binding; white arrows: non-productive binding).

(D) Fraction of productive (green) and non-productive (gray) DONSON binding events (N: number of binding events; error bars: 95% CI).

(E–G) The distribution of measured time delays for productive DONSON binding events (N represents the number of origins; blue lines: median values; gray boxes: 95% CI; ns denotes $p > 0.05$, **** denotes $p \leq 0.0001$).

(H) Model illustrating that RecQL4 promotes DONSON release prior to origin firing.

(I) Plasmid replication assay comparing the biochemical activity of recombinant wild-type RecQL4 (WT) and ATPase dead RecQL4 (D856A).

(J) Integrated signal intensities from panel (I) with matching line styles and colors. For clarity, only 20 nM recombinant RecQL4 curves are shown (error bars denote the standard deviation from three technical repeats).

(K) Immunoblots of samples corresponding to reactions from panel (I). RPA70 was used as a loading control.

(L) Origin firing efficiency for experiments containing 20 nM DONSON^{AF647} (N represents the number of DNA molecules; error bars mark 95% CI; * denotes $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$).

See also Figure S6.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

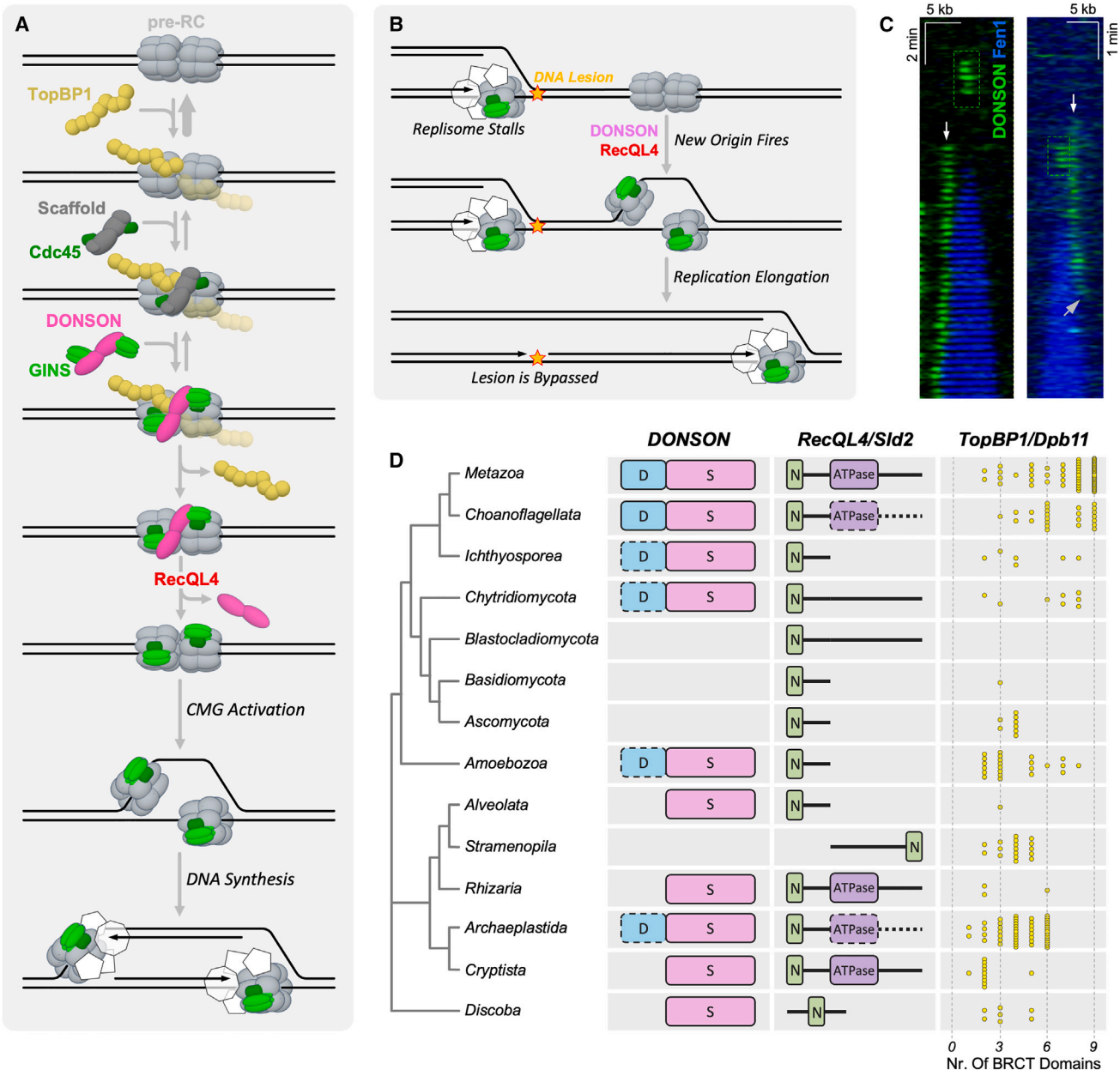


Figure 7. Model of replication initiation in metazoa

(A) Proposed model of replication initiation that is consistent with our data. The identity of the Cdc45 dimerization scaffold remains to be determined, and the timing of its dissociation is unknown (not shown).

(B) Model illustrating how a replication fork stalled at DNA lesion (orange) can be rescued by firing a nearby dormant origin (which would require DONSON, RecQL4, and other replication factors).

(C) Kymograms showing that long-lived non-productive DONSON-containing complexes (white arrows) are pushed by an active replication fork (productive DONSON binding events are marked with green boxes). The disappearance of the DONSON signal (gray arrow) may be due to bona fide unloading or photobleaching.

(D) Phylogenetic analysis of firing factor domain architecture (see STAR Methods, "Bioinformatic analysis"). For DONSON, the N-terminal disordered region (D) and the structured region involved in self-dimerization (S) are shown. For RecQL4, the conserved N-terminal region (~50 residues) and the ATPase domain are shown. For TopBP1, the number of BRCT domains is shown (each dot represents one species). A solid border indicates that the domain was found present in all putative orthologs in the taxa; a dashed border indicates that the domain could not be identified in some orthologs (lengths are not to scale).

See also Figure S7.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

mice with RecQL4 lacking the ATPase domain were viable but exhibited some phenotypes observed in human patients with RTS, BGS, and RAPADILINO syndromes.[53,54] Finally, RecQL4 truncations lacking the ATPase domain supported DNA replication *in vitro*, but only when used at very high concentrations.[24,25]

Our single-molecule experiments clearly indicate that RecQL4's ATPase promotes the efficient eviction of DONSON from CMG (Figures 6D and 6E). In contrast, the ensemble replication assay was somewhat insensitive to defects in origin firing activity (Figure 6I), because a single origin is sufficient to fully replicate a plasmid. Replication initiation defects caused by RecQL4 ATPase mutations may have gone unnoticed in some cell-based assays because these mutants retain low origin firing activity.

Although we depleted >99% of endogenous RecQL4 (Figure S3E), the depleted extract retained some origin firing activity (Figure 3B). This may be due to trace amounts of RecQL4 remaining after the depletion. Alternatively, replication origins may be able to fire in the absence of RecQL4, albeit inefficiently. Consistent with this idea, a recent study found that U2OS cells remained viable after RecQL4 was knocked out by introducing premature stop codons via CRISPR-Cas9.[55] Although truncated RecQL4 was not detected in these cells via immunoblotting, it is possible that small but undetectable amounts of protein may be sufficient to support origin firing.[55]

It is unclear exactly how RecQL4's ATPase activity contributes to replication initiation. It has been reported that RecQL4 can unwind short stretches of dsDNA in an ATP-dependent manner,[26,56,57] so it is plausible that RecQL4 may facilitate initiation by melting origin DNA. Alternatively, RecQL4 may use its ATPase activity to remodel the inactive CMG•DONSON•DONSON•CMG awaiting activation at the origin. Further studies will be needed to elucidate how CMG is activated in metazoa.

A potential mechanism to explain the high fidelity of bidirectional initiation

As far as we know, all eukaryotic origins fire bidirectionally. Human cells fire 10^4–10^5 origins during each cell cycle. This implies a very low rate of accidental unidirectional firing—at most $\sim 10^{-5}$ in human cells. Other biological processes that require such high fidelity include tRNA aminoacylation,[58] T cell receptor signaling,[59] and homologous recombination.[60] They achieve error rates of 10^{-4} or better via Hopfield kinetic proofreading[61–63]—a mechanism wherein a multi-step biochemical reaction has one or more reversible steps and at least one proof-reading step driven by an energy input such as phosphate hydrolysis. Incorrect intermediates prematurely exit this pathway, while correct intermediates successfully complete the reaction.

Although we have insufficient evidence, it is tempting to speculate that Hopfield kinetic proofreading underlies the high fidelity of bidirectional replication initiation. First, we identified at least three reversible steps: TopBP1 binding, Cdc45 recruitment, and GINS recruitment (Figure 7A). Second, we identified at least one irreversible step: the activation of two CMGs assembled at the origin (with the accompanying DONSON eviction from the helicase). Finally, replication initiation is driven by an energy input: we showed that ATP hydrolysis by RecQL4 contributes to efficient origin firing, and CMG helicase activation requires ATP hydrolysis by Mcm2-7.[10,64]

Differences between replication initiation mechanisms in yeast versus metazoa

We found that two Cdc45 molecules simultaneously bind to the pre-RC. To explain this observation, we postulated the existence of a Cdc45 dimerization scaffold. There are a few Cdc45 scaffold candidates: Treslin-MTBP[65–67] (the metazoan ortholog of yeast Sld3-Sld7), GemC1[68,69] (only found in metazoa), and potentially DUE-B.[70,71] These proteins have been implicated in replication initiation, have been shown to dimerize, and are known to interact with Cdc45 and other replication factors. We speculate that the Cdc45 dimerization scaffold helps ensure a high fidelity of bidirectional replication initiation in metazoa. In contrast to our findings, a recent single-molecule study reported that in budding yeast, Cdc45 is sequentially recruited to the pre-RC one molecule at a time.[42] This suggests that yeast may not employ a Cdc45 dimerization scaffold.

Origin firing was previously reconstituted with purified budding yeast proteins[12] but not with human proteins,[13] suggesting the existence of additional regulators of initiation in metazoa. One of them was recently identified in the form of DONSON.[30–33] Using InterPro,[72] thousands of putative DONSON homologs can be found broadly across eukaryotic supergroups[73]: metazoa, fungi, amoeba, brown algae, red algae, green algae, and plants. This suggests DONSON was present in the last eukaryotic common ancestor (LECA). BLAST-searching EukProt,[74] a curated database of ~1,000 eukaryotic species distributed between the major supergroups, suggests that LECA had a RecQL4 homolog with an Sld2-like N-terminus and an ATPase domain (consistent with a previous bioinformatic screen of Sld2 orthologues[75]) and a TopBP1 homolog containing several BRCT domains (Figure 7D). Budding yeast Sld2 lacks an ATPase domain (Figure 3A) and yeast lacks DONSON altogether, suggesting that yeast has undergone reductive evolution of some replication initiation factors.

Limitations of the study

Xenopus egg extracts contain high concentrations of replication factors that are needed to support several fast cell divisions following egg fertilization.[76] Our mock-depleted replication reactions contain ~200 nM of GINS, ~200 nM of Cdc45, ~150 nM of TopBP1, ~250 nM of RecQL4, and ~350 nM of DONSON. To maintain an acceptable signal-to-noise ratio in our single-molecule experiments, the concentration of fluorescent proteins was capped at 30 nM. Our data indicate that the replication initiation mechanism is robust across a wide range of GINS, Cdc45, TopBP1, RecQL4, and DONSON concentrations. It is unclear how physiological conditions compare to concentrations used in this study, as there are no accurate measurements of replication protein concentrations in cells. Estimates from the Protein Abundance Database[77] vary widely among cell types: 50–500 nM for GINS, 20–100 nM for Cdc45, 5–50 nM for TopBP1, 10–50 nM for DONSON, and 1–10 nM for RecQL4.

To track CMG movement after origin firing and to resolve protein binding to different locations on DNA, the DNA was stretched to ~80% of its contour length; therefore, it could not be fully chromatinized. Our data show that TopBP1, DONSON, and RecQL4 do not undergo long-range diffusion along DNA during initiation. However, given the spatial resolution of our assay (~300–500 base pairs [bp]), we cannot rule out that

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

some proteins (for example, RecQL4) may travel short distances near the origin.

To date we have not found an oxygen scavenging system that does not inhibit replication in extract. In most experiments, the image acquisition rate was limited to 10 s/frame to minimize photobleaching (20 s/frame for experiments in Figure 6). Due to photobleaching (Figures S7A–S7G), we under-estimated the duration of long-lived non-productive binding events for 2xDONSON, 2xGINS, or 2xCdc45, which lasted several minutes (Figures 5C and S5C). The effect of photobleaching was minimal on productive binding events for TopBP1, DONSON, and RecQL4, which lasted only 1–2 min (Figures 2G, 3D, and 4E).

We spent considerable effort optimizing fluorescent labeling of recombinant proteins to maximize labeling efficiency and minimize any adverse effect on protein function. TopBP1 and RecQL4 labeling efficiency exceeded 90%, whereas GINS and Cdc45 were labeled less efficiently (~50% and ~70%, respectively). When examining how two copies of GINS (or Cdc45) are recruited to the origin (Figures 1F and 1G), we analyzed only initiation events where both copies of GINS or Cdc45 were labeled and visible. When measuring the time delay between GINS (or Cdc45) binding and origin firing (Figure 1I), we analyzed events where one or two copies of GINS (or Cdc45) were fluorescently labeled (Figures S7H–S7J). Due to incomplete labeling (~46%), some DONSON binding events were invisible, while others contained only one copy of labeled DONSON. Unless explicitly stated, we analyzed events that included one or two fluorescently labeled DONSON.

In a small but significant fraction of origin firing events (Figures 1F and 1G, yellow, orange, and red), our automated code scored that GINS/Cdc45 were recruited sequentially. These findings could be explained by transient docking of 1xGINS/Cdc45, overfitting of noisy signal, and fluorophore blinking as discussed in the STAR Methods section (see Figures S7K and S7L for examples).

When analyzing the duration of TopBP1, RecQL4, and DONSON binding events, we counted only events that lasted two or more frames (Figure S7M) because we were not confident that single-frame events represented bona-fide binding (see STAR Methods for details).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2024.05.024.

ACKNOWLEDGMENTS

We thank members of the Chistol, Cimprich, and Ferrell labs for suggestions and feedback; Karlene Cimprich, James Ferrell, and Agnel Sfeir for critical reading of the manuscript; William Dunphy and Akiko Kumagai for *Xenopus* TopBP1 cDNA; and Johannes Walter for sharing unpublished results. R.T. was supported in part by the Stanford School of Medicine Dean's fellowship. S.E.B. was supported by the Stanford Graduate Fellowship, the New Science fellowship, and the ARCS Northern California Scholarship. L.A.S. is supported in part by a Blavatnik Family Fellowship Fund. G.C. is supported by an NSF CAREER Award (2144481), an NIGMS R35 award (GM147060), and an American Cancer Society seed grant (228425).

AUTHOR CONTRIBUTIONS

G.C., R.T., and S.E.B. wrote the manuscript with input from all authors. G.C. conceived and directed the project. G.C., R.T., and S.E.B. conceptualized and designed most experiments. S.E.B. prepared the initial batch of recombinant Cdc45. D.S. prepared recombinant Fen1^{mKikGR}. R.T. prepared all other recombinant proteins. R.T. and S.E.B. performed biochemical assays in egg extract and single-molecule experiments. R.T. and L.A.S. performed immunoprecipitation assays. S.E.B. raised and validated the Cdc45 custom antibody. D.S. raised and validated TopBP1 custom antibodies. R.T. raised and validated DONSON and RPA (replication protein A) custom antibodies. S.E.B. and L.A.S. raised and validated RecQL4 custom antibodies. G.C. wrote MATLAB code to analyze single-molecule experiments. G.C., R.T., and S.E.B. analyzed single-molecule data. S.E.B. performed bioinformatic analysis of replication initiation factors.

DECLARATION OF INTERESTS

Authors declare no conflicts of interest.

Received: October 19, 2023
Revised: March 28, 2024
Accepted: May 13, 2024
Published: June 11, 2024

REFERENCES

1. Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc. Natl. Acad. Sci. USA *106*, 20240–20245. https://doi.org/10.1073/pnas.0911500106.
2. Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F.X. (2009). Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. Cell *139*, 719–730. https://doi.org/10.1016/j.cell.2009.10.015.
3. Kubota, Y., Takase, Y., Komori, Y., Hashimoto, Y., Arata, T., Kamimura, Y., Araki, H., and Takisawa, H. (2003). A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. Genes Dev. *17*, 1141–1152. https://doi.org/10.1101/gad.1070003.
4. Aparicio, O.M., Stout, A.M., and Bell, S.P. (1999). Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. Proc. Natl. Acad. Sci. USA *96*, 9130–9135. https://doi.org/10.1073/pnas.96.16.9130.
5. Tercero, J.A., Labib, K., and Diffley, J.F. (2000). DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. EMBO J *19*, 2082–2093. https://doi.org/10.1093/emboj/19.9.2082.
6. Ilves, I., Petojevic, T., Pesavento, J.J., and Botchan, M.R. (2010). Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins. Mol. Cell *37*, 247–258. https://doi.org/10.1016/j.molcel.2009.12.030.
7. Parker, M.W., Botchan, M.R., and Berger, J.M. (2017). Mechanisms and regulation of DNA replication initiation in eukaryotes. Crit. Rev. Biochem. Mol. Biol. *52*, 107–144. https://doi.org/10.1080/10409238.2016.1274717.
8. Costa, A., and Diffley, J.F.X. (2022). The Initiation of Eukaryotic DNA Replication. Annu. Rev. Biochem. *91*, 107–131. https://doi.org/10.1146/annurev-biochem-072321-110228.
9. Tognetti, S., Riera, A., and Speck, C. (2015). Switch on the engine: how the eukaryotic replicative helicase MCM2–7 becomes activated. Chromosoma *124*, 13–26. https://doi.org/10.1007/s00412-014-0489-2.
10. Douglas, M.E., Ali, F.A., Costa, A., and Diffley, J.F.X. (2018). The mechanism of eukaryotic CMG helicase activation. Nature *555*, 265–268. https://doi.org/10.1038/nature25787.
11. Yeeles, J.T.P., Janska, A., Early, A., and Diffley, J.F.X. (2017). How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. Mol. Cell *65*, 105–116. https://doi.org/10.1016/j.molcel.2016.11.017.
12. Yeeles, J.T.P., Deegan, T.D., Janska, A., Early, A., and Diffley, J.F.X. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. Nature *519*, 431–435. https://doi.org/10.1038/nature14285.
13. Baris, Y., Taylor, M.R.G., Aria, V., and Yeeles, J.T.P. (2022). Fast and efficient DNA replication with purified human proteins. Nature *606*, 204–210. https://doi.org/10.1038/s41586-022-04759-1.
14. Hashimoto, Y., and Takisawa, H. (2003). Xenopus Cut5 is essential for a CDK-dependent process in the initiation of DNA replication. EMBO J *22*, 2526–2535. https://doi.org/10.1093/emboj/cdg238.
15. Van Hatten, R.A., Tutter, A.V., Holway, A.H., Khederian, A.M., Walter, J.C., and Michael, W.M. (2002). The Xenopus Xmus101 protein is required for the recruitment of Cdc45 to origins of DNA replication. J. Cell Biol. *159*, 541–547. https://doi.org/10.1083/jcb.200207090.
16. Kumagai, A., Lee, J., Yoo, H.Y., and Dunphy, W.G. (2006). TopBP1 activates the ATR-ATRIP complex. Cell *124*, 943–955. https://doi.org/10.1016/j.cell.2005.12.041.
17. Kim, A., Montales, K., Ruis, K., Senebandith, H., Gasparyan, H., Cowan, Q., and Michael, W.M. (2020). Biochemical analysis of TOPBP1 oligomerization. DNA Repair (Amst) *96*, 102973. https://doi.org/10.1016/j.dnarep.2020.102973.
18. Luong, T.T., and Bernstein, K.A. (2021). Role and regulation of the recql4 family during genomic integrity maintenance. Genes (Basel) *12*, 1919. https://doi.org/10.3390/genes12121919.
19. Siitonen, H.A., Sotkasiira, J., Biervliet, M., Benmansour, A., Capri, Y., Cormier-Daire, V., Crandall, B., Hannula-Jouppi, K., Hennekam, R., Herzog, D., et al. (2009). The mutation spectrum in RECQL4 diseases. Eur. J. Hum. Genet. *17*, 151–158. https://doi.org/10.1038/ejhg.2008.154.
20. Mo, D., Zhao, Y., and Balajee, A.S. (2018). Human RecQL4 helicase plays multifaceted roles in the genomic stability of normal and cancer cells. Cancer Lett. *413*, 1–10. https://doi.org/10.1016/j.canlet.2017.10.021.
21. Croteau, D.L., Popuri, V., Opresko, P.L., and Bohr, V.A. (2014). Human RecQ helicases in DNA repair, recombination, and replication. Annu. Rev. Biochem. *83*, 519–552. https://doi.org/10.1146/annurev-biochem-060713-035428.
22. Im, J.S., Ki, S.H., Farina, A., Jung, D.S., Hurwitz, J., and Lee, J.K. (2009). Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. Proc. Natl. Acad. Sci. USA *106*, 15628–15632. https://doi.org/10.1073/pnas.0908039106.
23. Xu, X., Rochette, P.J., Feyissa, E.A., Su, T.V., and Liu, Y. (2009). MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. EMBO J *28*, 3005–3014. https://doi.org/10.1038/emboj.2009.235.
24. Sangrithi, M.N., Bernal, J.A., Madine, M., Philpott, A., Lee, J., Dunphy, W.G., and Venkitaraman, A.R. (2005). Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. Cell *121*, 887–898. https://doi.org/10.1016/j.cell.2005.05.015.
25. Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y., and Takisawa, H. (2006). The N-Terminal Noncatalytic Region of Xenopus RecQ4 Is Required for Chromatin Binding of DNA Polymerase α in the Initiation of DNA Replication. Mol. Cell. Biol. *26*, 4843–4852. https://doi.org/10.1128/mcb.02267-05.
26. Capp, C., Wu, J., and Hsieh, T.S. (2009). Drosophila RecQ4 has a 3′-5′ DNA helicase activity that is essential for viability. J. Biol. Chem. *284*, 30845–30852. https://doi.org/10.1074/jbc.M109.008052.
27. Castillo-Tandazo, W., Smeets, M.F., Murphy, V., Liu, R., Hodson, C., Heierhorst, J., Deans, A.J., and Walkley, C.R. (2019). ATP-dependent helicase activity is dispensable for the physiological functions of Recql4. PLoS Genet. *15*, e1008319. https://doi.org/10.1371/journal.pgen.1008266.
28. Abe, T., Yoshimura, A., Hosono, Y., Tada, S., Seki, M., and Enomoto, T. (2011). The N-terminal region of RECQL4 lacking the helicase domain is both essential and sufficient for the viability of vertebrate cells. Role of the N-terminal region of RECQL4 in cells. Biochim. Biophys. Acta *1813*, 473–479. https://doi.org/10.1016/j.bbamcr.2011.01.001.
29. Kohzaki, M., Chiourea, M., Versini, G., Adachi, N., Takeda, S., Gagos, S., and Halazonetis, T.D. (2012). The helicase domain and C-terminus of human RecQL4 facilitate replication elongation on DNA templates damaged by ionizing radiation. Carcinogenesis *33*, 1203–1210. https://doi.org/10.1093/carcin/bgs149.
30. Xia, Y., Sonneville, R., Jenkyn-Bedford, M., Ji, L., Alabert, C., Hong, Y., Yeeles, J.T.P., and Labib, K.P.M. (2023). DNSN-1 recruits GINS for CMG helicase assembly during DNA replication initiation in Caenorhabditis elegans. Science (80-.) *381*, eadi4932–100. https://doi.org/10.1126/science.adi4932.
31. Kingsley, G., Skagia, A., Passaretti, P., Fernandez-Cuesta, C., Reynolds-Winczura, A., Koscielniak, K., and Gambus, A. (2023). DONSON facilitates Cdc45 and GINS chromatin association and is essential for DNA replication initiation. Nucleic Acids Res. *51*, 9748–9763. https://doi.org/10.1093/nar/gkad694.
32. Lim, Y., Tamayo-Orrego, L., Schmid, E., Tarnauskaite, Z., Kochenova, O.V., Gruar, R., Muramatsu, S., Lynch, L., Schlie, A.V., Carroll, P.L., et al. (2023). In silico protein interaction screening uncovers DONSON's role in replication initiation. Science *381*, 3448. https://doi.org/10.1126/science.adi3448.

33. Hashimoto, Y., Sadano, K., Miyata, N., Ito, H., and Tanaka, H. (2023). Novel role of DONSON in CMG helicase assembly during vertebrate DNA replication initiation. EMBO J *42*, e114217. https://doi.org/10.15252/embj.2023114131.

34. Reynolds, J.J., Bicknell, L.S., Carroll, P., Higgs, M.R., Shaheen, R., Murray, J.E., Papadopoulos, D.K., Leitch, A., Murina, O., Tarnauskaitė, Ž., et al. (2017). Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism. Nat. Genet. *49*, 537–549. https://doi.org/10.1038/ng.3790.

35. Zhang, J., Bellani, M.A., James, R.C., Pokharel, D., Zhang, Y., Reynolds, J.J., McNee, G.S., Jackson, A.P., Stewart, G.S., and Seidman, M.M. (2020). DONSON and FANCM associate with different replisomes distinguished by replication timing and chromatin domain. Nat. Commun. *11*, 3951. https://doi.org/10.1038/s41467-020-17449-1.

36. Muramatsu, S., Hirai, K., Tak, Y.S., Kamimura, Y., and Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon), and GINS in budding yeast. Genes Dev. *24*, 602–612. https://doi.org/10.1101/gad.1883410.

37. Sparks, J.L., Chistol, G., Gao, A.O., Räschle, M., Larsen, N.B., Mann, M., Duxin, J.P., and Walter, J.C. (2019). The CMG Helicase Bypasses DNA-Protein Cross-Links to Facilitate Their Repair. Cell *176*, 167–181.e21. https://doi.org/10.1016/j.cell.2018.10.053.

38. Walter, J., Sun, L., and Newport, J. (1998). Regulated chromosomal DNA replication in the absence of a nucleus. Mol. Cell *1*, 519–529. https://doi.org/10.1016/S1097-2765(00)80052-0.

39. Loveland, A.B., Habuchi, S., Walter, J.C., and Van Oijen, A.M. (2012). A general approach to break the concentration barrier in single-molecule imaging. Nat. Methods *9*, 987–992. https://doi.org/10.1038/nmeth.2174.

40. Berger, S., and Chistol, G. (2023). Visualizing the Dynamics of DNA Replication and Repair at the Single-Molecule Molecule Level, 1st ed. (Elsevier Inc.) https://doi.org/10.1016/bs.mcb.2023.07.001.

41. Burhans, W.C., Vassilev, L.T., Caddle, M.S., Heintz, N.H., and DePamphilis, M.L. (1990). Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell *62*, 955–965. https://doi.org/10.1016/0092-8674(90)90270-O.

42. De Jesús-Kim, L., Friedman, L.J., Lõoke, M., Ramsoomair, C.K., Gelles, J., and Bell, S.P. (2021). DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2-7. Elife *10*, 1–30. https://doi.org/10.7554/eLife.65471.

43. Tanaka, S., Komeda, Y., Umemori, T., Kubota, Y., Takisawa, H., and Araki, H. (2013). Efficient Initiation of DNA Replication in Eukaryotes Requires Dpb11/TopBP1-GINS Interaction. Mol. Cell. Biol. *33*, 2614–2622. https://doi.org/10.1128/mcb.00431-13.

44. Schmidt, U., Wollmann, Y., Franke, C., Grosse, F., Saluz, H.P., and Hänel, F. (2008). Characterization of the interaction between the human DNA topoisomerase IIbeta-binding protein 1 (TopBP1) and the cell division cycle 45 (Cdc45) protein. Biochem. J. *409*, 169–177. https://doi.org/10.1042/BJ20070872.

45. Day, M., Tetik, B., Parlak, M., Almeida-Hernández, Y., Räschle, M., Kaschani, F., Siegert, H., Marko, A., Sanchez-Garcia, E., Kaiser, M., et al. (2024). TopBP1 utilises a bipartite GINS binding mode to support genome replication. Nat. Commun. *15*, 1797. https://doi.org/10.1038/s41467-024-45946-0.

46. Kumagai, A., Shevchenko, A., Shevchenko, A., and Dunphy, W.G. (2010). Treslin Collaborates with TopBP1 in Triggering the Initiation of DNA Replication. Cell *140*, 349–359. https://doi.org/10.1016/j.cell.2009.12.049.

47. Marino, F., Vindigni, A., and Onesti, S. (2013). Bioinformatic analysis of RecQ4 helicases reveals the presence of a RQC domain and a Zn knuckle. Biophys. Chem. *177–178*, 34–39. https://doi.org/10.1016/j.bpc.2013.02.009.

48. Zhou, Z.W., Liu, C., Li, T.L., Bruhn, C., Krueger, A., Min, W., Wang, Z.Q., and Carr, A.M. (2013). An Essential Function for the ATR-Activation-Domain (AAD) of TopBP1 in Mouse Development and Cellular Senescence. PLoS Genet. *9*, e1003702. https://doi.org/10.1371/journal.pgen.1003702.

49. Jones, M.L., Aria, V., Baris, Y., and Yeeles, J.T.P. (2023). How Pol α-primase is targeted to replisomes to prime eukaryotic DNA replication. Mol. Cell *83*, 2911–2924.e16. https://doi.org/10.1016/j.molcel.2023.06.035.

50. Cvetkovic, M.A., Passaretti, P., Butryn, A., Reynolds-Winczura, A., Kingsley, G., Skagia, A., Fernandez-Cuesta, C., Poovathumkadavil, D., George, R., Chauhan, A.S., et al. (2023). The structural mechanism of dimeric DONSON in replicative helicase activation. Mol. Cell *83*, 4017–4031.e9. https://doi.org/10.1016/j.molcel.2023.09.029.

51. Blow, J.J., Ge, X.Q., and Jackson, D.A. (2011). How dormant origins promote complete genome replication. Trends Biochem. Sci. *36*, 405–414. https://doi.org/10.1016/j.tibs.2011.05.002.

52. Koundrioukoff, S., Kim, S.-J., Alary, N., Toffano, A., Melendez-Garcia, R., Wu, X., Liu, Y., Gnan, S., El-Hilali, S., Brison, O., et al. (2023). The balance between ATR and DDK activities controls TopBP1-mediated locking of dormant origins at the pre-IC stage. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.29.569233.

53. Hoki, Y., Araki, R., Fujimori, A., Ohhata, T., Koseki, H., Fukumura, R., Nakamura, M., Takahashi, H., Noda, Y., Kito, S., and Abe, M. (2003). Growth retardation and skin abnormalities of the Recql4-deficient mouse. Hum. Mol. Genet. *12*, 2293–2299. https://doi.org/10.1093/hmg/ddg254.

54. Mann, M.B., Hodges, C.A., Barnes, E., Vogel, H., Hassold, T.J., and Luo, G. (2005). Defective sister-chromatid cohesion, aneuploidy and cancer predisposition in a mouse model of type II Rothmund-Thomson syndrome. Hum. Mol. Genet. *14*, 813–825. https://doi.org/10.1093/hmg/ddi075.

55. Padayachy, L., Ntallis, S.G., and Halazonetis, T.D. (2024). RECQL4 is not critical for firing of human DNA replication origins. Sci. Rep. *14*, 7708–7712. https://doi.org/10.1038/s41598-024-58404-0.

56. Macris, M.A., Krejci, L., Bussen, W., Shimamoto, A., and Sung, P. (2006). Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. DNA Repair (Amst) *5*, 172–180. https://doi.org/10.1016/j.dnarep.2005.09.005.

57. Suzuki, T., Kohno, T., and Ishimi, Y. (2009). DNA helicase activity in purified human RECQL4 protein. J. Biochem. *146*, 327–335. https://doi.org/10.1093/jb/mvp074.

58. Hopfield, J.J., Yamane, T., Yue, V., and Coutts, S.M. (1976). Direct experimental evidence for kinetic proofreading in amino acylation of tRNA(Ile). Proc. Natl. Acad. Sci. USA *73*, 1164–1168. https://doi.org/10.1073/pnas.73.4.1164.

59. Mckeithan, T.W. (1995). Kinetic proofreading in T-cell receptor signal transduction. Proc. Natl. Acad. Sci. USA *92*, 5042–5046. https://doi.org/10.1073/pnas.92.11.5042.

60. Sagi, D., Tlusty, T., and Stavans, J. (2006). High fidelity of RecA-catalyzed recombination: A watchdog of genetic diversity. Nucleic Acids Res. *34*, 5021–5031. https://doi.org/10.1093/nar/gkl586.

61. Ninio, J. (1975). Kinetic amplification of enzyme discrimination. Biochimie *57*, 587–595. https://doi.org/10.1016/S0300-9084(75)80139-8.

62. Hopfield, J.J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc. Natl. Acad. Sci. USA *71*, 4135–4139. https://doi.org/10.1073/pnas.71.10.4135.

63. Boeger, H. (2022). Kinetic Proofreading. Annu. Rev. Biochem. *91*, 423–447. https://doi.org/10.1146/annurev-biochem-040320-103630.

64. Kang, S., Warner, M.D., and Bell, S.P. (2014). Multiple Functions for Mcm2-7 ATPase Motifs during Replication Initiation. Mol. Cell *55*, 655–665. https://doi.org/10.1016/j.molcel.2014.06.033.

65. Kumagai, A., and Dunphy, W.G. (2017). MTBP, the partner of Treslin, contains a novel DNA-binding domain that is essential for proper initiation of DNA replication. Mol. Biol. Cell *28*, 2998–3012. https://doi.org/10.1091/mbc.E17-07-0448.

66. Volpi, I., Gillespie, P.J., Chadha, G.S., and Blow, J.J. (2021). The role of DDK and Treslin-MTBP in coordinating replication licensing and

pre-initiation complex formation. Open Biol. *11*, 210121. https://doi.org/10.1098/rsob.210121.

67. Kumagai, A., Shevchenko, A., Shevchenko, A., and Dunphy, W.G. (2011). Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. J. Cell Biol. *193*, 995–1007. https://doi.org/10.1083/jcb.201102003.
68. Balestrini, A., Cosentino, C., Errico, A., Garner, E., and Costanzo, V. (2010). GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication. Nat. Cell Biol. *12*, 484–491. https://doi.org/10.1038/ncb2050.
69. Caillat, C., Fish, A., Pefani, D.E., Taraviras, S., Lygerou, Z., and Perrakis, A. (2015). The structure of the GemC1 coiled coil and its interaction with the Geminin family of coiled-coil proteins. Acta Crystallogr. D Biol. Crystallogr. *71*, 2278–2286. https://doi.org/10.1107/S1399004715016892.
70. Chowdhury, A., Liu, G., Kemp, M., Chen, X., Katrangi, N., Myers, S., Ghosh, M., Yao, J., Gao, Y., Bubulya, P., and Leffak, M. (2010). The DNA Unwinding Element Binding Protein DUE-B Interacts with Cdc45 in Preinitiation Complex Formation. Mol. Cell. Biol. *30*, 1495–1507. https://doi.org/10.1128/mcb.00710-09.
71. Poudel, S., Yao, J., Kemp, M.G., and Leffak, M. (2018). Interaction between DUE-B and Treslin is required to load Cdc45 on chromatin in human cells. J. Biol. Chem. *293*, 14497–14506. https://doi.org/10.1074/jbc.RA118.004519.
72. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., et al. (2023). InterPro in 2022. Nucleic Acids Res. *51*, D418–D427. https://doi.org/10.1093/nar/gkac993.
73. Burki, F., Roger, A.J., Brown, M.W., and Simpson, A.G.B. (2020). The New Tree of Eukaryotes. Trends Ecol. Evol. *35*, 43–55. https://doi.org/10.1016/j.tree.2019.08.008.
74. Richter, D.J., Berney, C., Strassert, J.F.H., Poh, Y.P., Herman, E.K., Muñoz-Gómez, S.A., Wideman, J.G., Burki, F., and de Vargas, C. (2022). EukProt: A database of genome-scale predicted proteins across the diversity of eukaryotes. Peer Community J. *2*, e56. https://doi.org/10.24072/pcjournal.173.
75. Gaggioli, V., Zeiser, E., Rivers, D., Bradshaw, C.R., Ahringer, J., and Zegerman, P. (2014). CDK phosphorylation of SLD-2 is required for replication initiation and germline development in C. elegans. J. Cell Biol. *204*, 507–522. https://doi.org/10.1083/jcb.201310083.
76. Hoogenboom, W.S., Klein Douwel, D., and Knipscheer, P. (2017). Xenopus egg extract: A powerful tool to study genome maintenance mechanisms. Dev. Biol. *428*, 300–309. https://doi.org/10.1016/j.ydbio.2017.03.033.
77. Wang, M., Weiss, M., Simonovic, M., Haertinger, G., Schrimpf, S.P., Hengartner, M.O., and Von Mering, C. (2012). PaxDb, a database of protein abundance averages across all three domains of life. Mol. Cell. Proteomics. *11*, 492–500. https://doi.org/10.1074/mcp.O111.014704.
78. Arias, E.E., and Walter, J.C. (2005). Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in Xenopus egg extracts. Genes Dev. *19*, 114–126. https://doi.org/10.1101/gad.1255805.
79. Yoo, H.Y., Kumagai, A., Shevchenko, A., Shevchenko, A., and Dunphy, W.G. (2007). Ataxia-telangiectasia Mutated (ATM) -dependent Activation of ATR Occurs through Phosphorylation of TopBP1 by ATM. J. Biol. Chem. *282*, 17501–17506. https://doi.org/10.1074/jbc.M701770200.
80. Lebofsky, R., Takahashi, T., and Walter, J.C. (2009). DNA replication in nucleus-free Xenopus egg extracts. Methods Mol. Biol. *521*, 229–252. https://doi.org/10.1007/978-1-60327-815-7_13.
81. Sparks, J., and Walter, J.C. (2019). Extracts for analysis of DNA replication in a nucleus-free system. Cold Spring Harb. Protoc. *2019*, 194–206. https://doi.org/10.1101/pdb.prot097154.
82. Modesti, M. (2011). Fluorescent Labeling of Proteins. Methods Mol. Biol. *783*, 281–296. https://doi.org/10.1007/978-1-61779-282-3.
83. Sergé, A., Bertaux, N., Rigneault, H., and Marguet, D. (2008). Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. Nat. Methods *5*, 687–694. https://doi.org/10.1038/nmeth.1233.
84. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: Protein domains identifier. Nucleic Acids Res. *33*, 116–120. https://doi.org/10.1093/nar/gki442.
85. Necci, M., Piovesan, D., Dosztányi, Z., and Tosatto, S.C.E. (2017). MobiDB-lite: Fast and highly specific consensus prediction of intrinsic disorder in proteins. Bioinformatics *33*, 1402–1404. https://doi.org/10.1093/bioinformatics/btx015.
86. Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci. *31*, 8–22. https://doi.org/10.1002/pro.4218.
87. Schoch, C.L., Ciufo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Leipe, D., McVeigh, R., O'Neill, K., Robbertse, B., et al. (2020). NCBI Taxonomy: A comprehensive update on curation, resources and tools. Database. *2020*, baaa062–21. https://doi.org/10.1093/database/baaa062.
88. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., et al. (2020). CDD/SPARCLE: The conserved domain database in 2020. Nucleic Acids Res. *48*, D265–D268. https://doi.org/10.1093/nar/gkz991.
89. Lewis, T.E., Sillitoe, I., Dawson, N., Lam, S.D., Clarke, T., Lee, D., Orengo, C., and Lees, J. (2018). Gene3D: Extensive prediction of globular domains in proteins. Nucleic Acids Res. *46*, D435–D439. https://doi.org/10.1093/nar/gkx1069.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Anti-xTopBP1-SU276	this paper	N/A
Rabbit Anti-xTopBP1-SU277	this paper	N/A
Rabbit Anti-xRecQL4-SU307	this paper	N/A
Rabbit Anti-xRecQL4-SU305	this paper	N/A
Rabbit Anti-xDONSON-SU349	this paper	N/A
Rabbit Anti-xDONSON-SU351	this paper	N/A
Rabbit Anti-xRPA-SU299	this paper	N/A
Rabbit Anti-xCdc45-SU294	this paper	N/A
Rabbit Anti-xGINS-SU289	this paper	N/A
Rabbit Anti-xPsf3-SU279	this paper	N/A
Rabbit Anti-xSld5-SU283	this paper	N/A
Peroxidase AffiniPure Rabbit Anti-Mouse IgG (H + L)	Jackson Immuno Research Labs	Cat# 315-035-003; RRID:AB_2340061
Bacterial and virus strains		
DH5α	NEB	Cat# C2987H
DH10Bac	Thermo Fisher Scientific	Cat# 10361012
DH10EMBacY	Geneva Biotech	Cat# DH10EMBacY
Rosetta (DE3) pLysS	Novagen	Cat# 70956-4
Chemicals, peptides, and recombinant proteins		
ApaI	NEB	Cat# R0114S
BamHI-HF	NEB	Cat# R3136S
BmgBI	NEB	Cat# R0628S
BstXI	NEB	Cat# R0113S
HindIII-HF	NEB	Cat# R3104S
NcoI-HF	NEB	Cat# R3193S
NdeI	NEB	Cat# R0111S
NotI-HF	NEB	Cat# R3189S
SalI-HF	NEB	Cat# R3138S
SmaI	NEB	Cat# R0141S
SpeI-HF	NEB	Cat# R3133S
SphI-HF	NEB	Cat# R3182S
XbaI	NEB	Cat# R0145S
XhoI	NEB	Cat# R0146S
Klenow Fragment (3′–5′ exo-) polymerase	NEB	Cat# M0212S
Proteinase K	NEB	Cat# P8107S
[α-32P] dCTP	PerkinElmer	Cat# BLU513A250UC
biotin-dCTP	Thermo Fisher Scientific	Cat# 19518018
biotin-dGTP	PerkinElmer	Cat# NEL541001EA
L-Cysteine hydrochloride monohydrate	Fisher Scientific	Cat# ICN10144601
Cycloheximide	Sigma Aldrich	Cat# 01810
Aprotinin	Sigma Aldrich	Cat# 11583794001
Leupeptin	Sigma Aldrich	Cat# 11529048001
Cytochalasin B from *Drechslera dematioidea*	Sigma Aldrich	Cat# C6762
Dithiothreitol (DTT)	Bio-Rad	Cat# 1610611

(*Continued on next page*)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Nocodazole	Sigma Aldrich	Cat# M1404
Dimethyl sulfoxide (DMSO)	Sigma Aldrich	Cat# D8418
Adenosine 5′-triphosphate disodium salt hydrate (ATP)	Sigma Aldrich	Cat# A7699
Phosphocreatine disodium salt hydrate (PC)	Sigma Aldrich	Cat# P7956
Creatine Phosphokinase from rabbit muscle (CPK)	Sigma Aldrich	C3755
Geminin	Arias et al.[78]	N/A
GINS expressed from pGC128	Sparks et al.[37]	N/A
Cdc45 expressed from pGC045	this paper	N/A
Fen1^{mkikGR}	Loveland et al.[39]	N/A
TopBP1 expressed from pRT005	this paper	N/A
RecQL4 expressed from pRT027	this paper	N/A
RecQL4-D856A expressed from pRT076	this paper	N/A
DONSON expressed from pRT066	this paper	N/A
His6-Human Rhinovirus (HRV) 3C Protease	this paper	N/A
Sortase	Sparks et al.[37]	N/A
SNAP-Surface Alexa Fluor 546	NEB	Cat# S9132S
SNAP-Surface Alexa Fluor 647	NEB	Cat# S9136S
Alexa Fluor 546 C_5 Maleimide	Thermo Fisher Scientific	Cat# A10258
Alexa Fluor 647 C_2 Maleimide	Thermo Fisher Scientific	Cat# A20347
AZDye 647 Maleimide	Fluoroprobes	Cat# 1122-1
GGGGYKCK Peptide	Biosynth	N/A
ANTI-FLAG M2 Affinity Gel	Sigma Aldrich	Cat# A2220
Anti-FLAG M2 Magnetic Beads	Millipore	Cat# M8823
Anti-V5 Agarose Affinity Gel antibody produced in mouse	Sigma Aldrich	Cat# A7345-1ML
3X FLAG Peptide	APExBio	Cat# A6001
Nuvia HP-Q Strong Anion Exchange Media	Bio-Rad	Cat# 12006693
Ni-NTA Agarose	Qiagen	Cat# 30230
Imidazole	Fisher Scientific	Cat# 03196-500
Dynabeads Protein A	Invitrogen	Cat# 10002D
rProtein A Sepharose Fast Flow	GE Healthcare	Cat# 171279
AminoLink Plus Coupling Resin	Fisher Scientific	Cat# PI20501
Cellfectin II Reagent	Fisher Scientific	Cat# 10-362-100
LB Broth	Thermo Fisher Scientific	Cat# BP9723
LB Agar	Thermo Fisher Scientific	Cat# BP9724
ESF 921 Insect Cell Culture Medium, Protein Free	Expression Systems	Cat# 96-001-01
Superdex 200 Increase 10/300 GL	Cytiva	Cat# 28990944
Mono Q 5/50 GL	Cytiva	Cat# 17516601
PD-10 Desalting Column	Cytiva	Cat# 1785101
Amicon Ultra-4-10K	Merck Millipore	Cat# UFC801024
Spectrum Spectra/Por 1 RC Dialysis Membrane Tubing 6–8 kD	Spectrum Labs	Cat# 132645
Spectrum Spectra/Por 4 RC Dialysis Membrane Tubing 12–14 kD	Spectrum Labs	Cat# 132700
Pierce Protease Inhibitor Tablets, EDTA-free	ThermoFisher Scientific	Cat# A32965
4-15% Mini-PROTEAN TGX Precast Protein Gels	Bio-Rad	Cat# 4561086
Tris-Glycine-SDS Running Buffer	Boston BioProducts	Cat# BP-150
EZ-Run Prestained Rec Protein Ladder	Fisher Scientific	Cat# BP36031

(*Continued on next page*)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Precision Plus Protein Dual Color Standard	Bio-Rad	Cat# 1610374EDU
SuperSignal West Pico PLUS Chemiluminescent Substrate	Fisher Scientific	Cat# PI34577
Optiblot Blue InstantBlue Coomassie Protein Stain	Abcam	Cat# ab119211
PVDF membranes	VWR	Cat# 10061-492
Transfer Buffer	Fisher Scientific	Cat# NC9297917
Phosphate Buffered Saline (10x, pH 7.4)	Fisher Scientific	Cat# NC9140736
CHORULON Hormone	Merck	Cat# 133754
75 × 25 × 1 mm Glass Slides	VWR	Cat# 48300-026
High Precision Glass CoverSlip, No 1.5, 24 × 60 mm	Bioscience Tools	Cat# CSHP-No1.5-24x60
Double-sided tape	Grace Bio-Labs	SKU 62001
mPEG-Succinimidyl Valerate, MW 5,000	Laysan Bio	Cat# MPEG-SVA-5000-1g
Biotin-PEG-SVA, MW 5,000	Laysan Bio	Cat# Biotin-PEG-SVA-5000-100MG
Streptavidin from *Streptomyces avidinii*	Sigma Aldrich	Cat# S4762-5MG
Streptavidin, Alexa Fluor 647 conjugate	Thermo Fisher Scientific	Cat# S21374
Critical commercial assays		
NEBuilder HiFi DNA Assembly Master Kit	NEB	Cat# E2621S
ZR BAC DNA Miniprep kit	Zymo Research	Cat# D4049
TnT SP6 High-Yield Wheat Germ Protein Expression System	Promega	Cat# L3261
Experimental models: Cell lines		
Sf9 insect cells	Expression Systems	Cat# 94-001S
Tni insect cells	Expression Systems	Cat# 94-002S
Experimental models: Organisms/strains		
Xenopus laevis, adult female	Nasco Healthcare	Cat# LM00531; RRID:XEP_Xla
Xenopus laevis, adult male	Nasco Healthcare	Cat# LM00715MX
Oligonucleotides		
See Table S1 for oligonucleotides and gene blocks	–	N/A
Recombinant DNA		
pFastBac1	Thermo Fisher Scientific	Cat# 10360014
pET28b(+)	Novagen	Cat# 69865
pSP64 Poly(A) Vector	Promega	Cat# P1241
pACE2	Geneva Biotech	N/A
pFastBacHAHT-Xtopbp1-FLAG	Yoo et al.[79]	N/A
λ DNA	NEB	Cat# N3011S
pBlueScript	Agilent	Cat# 212205
pGC045	this paper	N/A
pGC127	Sparks et al.[37]	N/A
pGC283	this paper	N/A
pGC314	gift from Tycho Mevissen	N/A
pLAS025	this paper	N/A
pLS019	this paper	N/A
pLS020	this paper	N/A
pLS029	this paper	N/A
pLS032	this paper	N/A
pLS033	this paper	N/A
pSEB054	this paper	N/A
pSEB057	this paper	N/A

(*Continued on next page*)

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pRT005	this paper	N/A
pRT009	this paper	N/A
pRT018	this paper	N/A
pRT019	this paper	N/A
pRT023	this paper	N/A
pRT025	this paper	N/A
pRT026	this paper	N/A
pRT027	this paper	N/A
pRT029	this paper	N/A
pRT030	this paper	N/A
pRT031	this paper	N/A
pRT032	this paper	N/A
pRT033	this paper	N/A
pRT034	this paper	N/A
pRT035	this paper	N/A
pRT036	this paper	N/A
pRT037	this paper	N/A
pRT055	this paper	N/A
pRT056	this paper	N/A
pRT057	this paper	N/A
pRT063	this paper	N/A
pRT064	this paper	N/A
pRT065	this paper	N/A
pRT066	this paper	N/A
pRT076	this paper	N/A
Software and algorithms		
MATLAB	Mathworks	matlab; RRID:SCR_001622
ImageJ	NIH	imagej; RRID:SCR_003070
Image Stabilizer Plugin for ImageJ	Kang Li	https://www.cs.cmu.edu/~kangli/code/Image_Stabilizer.html
Jupyter Notebook	Project Jupyter	https://jupyter.org/; RRID:SCR_018315
Anaconda	Anaconda	https://www.anaconda.com/products/individual
Pandas	PyData	https://pandas.pydata.org/; RRID:SCR_018214
Biopython	Biopython	https://biopython.org/; RRID:SCR_007173

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Gheorghe Chistol (chistol@stanford.edu).

Materials availability

All newly generated reagents generated in this study are available upon request from the lead contact.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report any original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Xenopus laevis

Eggs from adult, female *Xenopus laevis* frogs (NASCO Healthcare, Cat# LM00531) were used to prepare egg extracts and the testes of adult, male *Xenopus laevis* frogs (NASCO Healthcare, Cat# LM00715MX) were used to prepare sperm chromatin. All animals were healthy, not subjected to previous procedures, and no animal husbandry was performed. Frogs were housed at the Boswell Small Amphibians Research Facility at Stanford School of Medicine in compliance with Institute Animal Care and Use Committee (IACUC) regulations and all experiments involving frogs were approved by the Stanford University APLAC (Administrative Panel on Laboratory Animal Care).

Insect cell lines

Sf9 cells (Expression Systems, Cat# 94-001S) and Tni cells (Expression Systems, Cat# 94-002S) were cultured in ESF 921 insect cell culture medium (Expression Systems, Cat# 96-001-01) at 27 °C for baculovirus production and protein overexpression.

Bacterial cell lines

DH10Bac (Thermo Fisher Scientific, Cat# 10361012), DH10EMBacY (Geneva Biotech, Cat# DH10EMBacY), DH5α (New England Biorabs (NEB), Cat# C2987H), and Rosetta (DE3) pLysS (Novagen, Cat# 70956-4) bacteria were cultured in LB broth (Thermo Fisher Scientific, Cat# BP9723) or Terrific Broth (TB) at 37 C for plasmid production and protein overexpression.

IVTT system

IVTT reactions were performed using TnT SP6 High-Yield Wheat Germ Protein Expression System (Promega, Cat# L3261) according to the manufacturer's instructions. Proteins were not purified after their production by IVTT. The IVTT reaction was directly added replication reactions as the source of a given protein. To avoid liquid-liquid phase separation of the target protein, its final concentration was kept under 1 μM. The optimal concentration of the template plasmid was titrated for each protein.

METHOD DETAILS

Preparation of DNA substrates

All single molecule experiments used doubly tethered λ phage DNA. The 48.5 kb DNA substrate as purchased from NEB (#N3011S) has 12-nt 5′ ssDNA overhangs on both ends. We used the Klenow Fragment (3′–5′ exo-) polymerase (NEB, Cat# M0212S) to fill in the overhangs with biotin-dCTP (Thermo Fisher Scientific, Cat# 19518018) and biotin-dGTP (PerkinElmer, Cat# NEL541001EA). The product was separated via electrophoresis and was electro-eluted from the gel before use in tethering. A step-by-step detailed protocol for preparing biotinylated λ DNA is available in our recent methods article.[40]

Xenopus egg extract preparation

Xenopus laevis egg extracts were prepared following published protocols.[80,81] Frogs were primed with 75 IU of human chorionic gonadotropin (CHORULON) (Merck, Cat# 133754) 5–7 days before extract preparation. To induce ovulation, 660 IU of CHORULON was injected 20–21 h before extract preparation. High-speed supernatant (HSS) was prepared from 5 to 6 adult female frogs, and nucleoplasmic extract (NPE) was prepared from 15 to 20 frogs. Sperm chromatin was purified from adult male frog testes and used for NPE preparation.

Ensemble replication assay

To validate antibodies and recombinant proteins, ensemble DNA replication experiments were performed following a detailed protocol described in a recent methods article.[40] Since all proteins of interest (POIs) in this study are dispensable for Mcm2-7 double-hexamer loading, all licensing reactions were performed with HSS depleted of POI.

Protein co-immunoprecipitation assay

To immunoprecipitate an endogenous protein of interest from extract, 10 μg of affinity-purified antibody was incubated with 1.25 mg of Dynabeads Protein A (Invitrogen, Cat# 10002D) for 1 h at 25 °C with gentle rotation. Beads were washed twice with IP wash buffer (10 mM HEPES-KOH pH 7.7, 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose, 0.05% NP-40) and incubated with 30 μL NPE for 1 h at 25 °C with gentle rotation. Next, beads were pelleted using a magnetic rack and washed three times with cold IP wash buffer. Bound proteins were eluted by boiling beads in 30 μL 1x Laemmli buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 0.1% bromophenol blue, 5% β-mercaptoethanol).

To immunoprecipitate FLAG-tagged DONSON from NPE, Anti-FLAG M2 Magnetic Beads (Millipore #M8823) were pelleted and washed three times with IP wash buffer and aliquots with 2 μL packed beads were made. Then, C-terminally FLAG-tagged DONSON (FLAG-DONSON) expressed in IVTT was immobilized onto Anti-FLAG M2 Magnetic Beads. To this end, 2 μL of packed anti-FLAG M2 Magnetic Beads was incubated with 20 μL of wheat germ extract containing FLAG-DONSON for 1 h at 4 °C with gentle rotation. Beads were washed twice with cold IP wash buffer. Beads coated with FLAG-DONSON were incubated with either 30 μL of

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

30% NPE (diluted in IP wash buffer), or 100 μL of undiluted HSS for 1 h at 4 °C with gentle rotation. Finally, beads were pelleted and washed three times with cold IP wash buffer. Bound proteins were eluted by boiling beads in 30 μL 1x Laemmli buffer. 4 μL of each eluted sample was loaded onto an SDS-PAGE gel and analyzed by immunoblotting. An equivalent amount of extract was loaded to verify immunoprecipitation efficiency. Optionally, to carefully quantify the coIP, a known amount of recombinant protein standard (200, 100, 50, 25, 12.5 fmol) was loaded alongside eluted samples.

For the IP of V5-RecQL4 and FLAG-RecQL4 expressed in IVTT, N-terminally V5 tagged RecQL4 and N-terminally FLAG tagged RecQL4 were co-expressed in IVTT from [pRT035]pSP64-V5-xRECQL4 and [pRT036]pSP64-FLAG-xRECQL4 at 25 °C for 2 h. 5 μL of the IVTT reaction was diluted with 5 μL of ELB-Sucrose. 2 μL of ANTI-FLAG M2 Affinity Gel (Sigma Aldrich, Cat# A2220) or 2 μL of Anti-V5 Agarose Affinity Gel (Sigma Aldrich, Cat# A7345-1ML) was incubated with 10 μL of diluted IVTT reaction for 1 h at 25 °C with gentle rotation. Beads were pelleted and washed three times with cold IP wash buffer. Bound proteins were eluted by boiling beads in 20 μL of 1x Laemmli buffer. When V5-RecQL4 and FLAG-RecQL4 were expressed separately, their coIP efficiency significantly decreased, suggesting that RecQL4 forms stable oligomers that exchange subunits at 25 °C on a timescale much longer than 1 h.

For the IP of DONSON-SNAP-FLAG and untagged DONSON expressed in IVTT, untagged DONSON and C-terminally SNAP-FLAG tagged DONSON were co-expressed in IVTT from [pRT063]pSP64-DONSON and [pRT064]pSP64-FLAG-SNAP-DONSON at 25 °C for 2 h. 5 μL of the IVTT reaction was diluted with 5 μL of ELB-Sucrose. 10 μL of diluted IVTT reaction was incubated with 2 μL of Anti-FLAG M2 Affinity Gel for 1 h at 25 °C with gentle rotation. Beads were pellets and washed three times with cold IP wash buffer. Bound proteins were eluted by boiling beads in 20 μL of 1x Laemmli buffer.

Western blotting and coomasie staining

All samples were boiled in 1x Laemmli buffer and run on 4–15% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad #4561086) using Tris-Glycine-SDS Running Buffer (Boston BioProducts #BP-150). EZ-Run Prestained Rec Protein Ladder (Fisher Scientific #BP36031) or Precision Plus Protein Dual Color Standard (Bio-Rad #1610374EDU) was run alongside samples to infer protein band sizes. For Coomasie staining, protein gels were incubated with Optiblot Blue InstantBlue Coomassie Protein Stain (Abcam #ab119211) for at least 1 h at room temperature with gentle shaking. For immunoblotting, proteins were transferred from the gel onto a PVDF membrane (VWR #10061-492) in transfer buffer (Fisher Scientific #NC9297917). Membranes were blocked in 1x PBST (1x Phosphate Buffered Saline (PBS) (Fisher Scientific, Cat# NC9140736), 0.05% Tween 20) containing 5% (w/v) nonfat milk for 1 h at room temperature with gentle shaking. Following a wash in 1xPBST, the membrane was incubated with primary antibodies diluted in 1x PBST containing 1% (w/v) BSA overnight at 4 °C with gentle shaking. Membranes were washed three times with 1x PBST and incubated with secondary antibody diluted in 1x PBST +5% nonfat milk for 1 h at room temperature with gentle shaking. Membranes were washed three times with 1x PBST, developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (Fisher Scientific, Cat# PI34577), and imaged using Azure c600 (Azure Biosystems).

Single-molecule replication assay

All single-molecule experiments were conducted following a detailed protocol,[40] except for the following modifications. In the λ DNA tethering step, 20 pM of streptavidin^{AF647} (Thermo Fisher Scientific, Cat# S21374) was added to the streptavidin solution. Streptavidin^{AF647} spots were used as a fiducial marker for focusing immediately before starting imaging. After λ DNA tethering, DNA was licensed in HSS for 10 min. In the original KEHRMIT protocol, after a brief pulse of replication initiation at high GINS^{AF647} concentration (~200 nM), the flow cell was flushed with GINS-depleted extract to prevent further origin firing and to minimize background.[37,40] In this study, Replication Mix (10 μL of NPE, 10 μL of HSS, 6 μL of ELB-Sucrose, 1 μL of 600 ng/μL pBlueScript carrier plasmid DNA, 1 μL of ATP regeneration Mix (67 mM ATP, 0.17 mg/mL creatine phosphokinase, 0.67 M phosphocreatine; same as in our recent methods article[40]), 1 μL of 60 μM Fen1^{mKikGR}, and 1 μL of 150–900 nM fluorescently labeled protein of interest) was flowed into the microfluidic chamber. The reaction was then imaged continuously (10 s/frame) without washing out the free fluorescently labeled protein. Thus, fluorescent protein (5–30 nM) was present in the replication reaction during the entire duration of the experiment. Optimal replication efficiency was achieved when the single-molecule replication reaction contained equal volumes of NPE, HSS, and buffer.[40] The concentrations of proteins examined in this study were estimated using quantitative western blots (see table below).

Protein of Interest	Concentration in HSS (nM)	Concentration in NPE (nM)	Concentration in Replication Reaction (nM)
GINS	~30 nM	~600 nM	~200 nM
Cdc45	~20 nM	~600 nM	~200 nM
TopBP1	~50 nM	~400 nM	~150 nM
DONSON	~100 nM	~1000 nM	~350 nM
RecQL4	~80 nM	~700 nM	~250 nM

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

For two-color imaging, Fen1^{mKikGR} (488 nm laser excitation) and POI labeled with Alexa Fluor 647 (640 nm excitation) were used. The following settings were used for imaging Alexa Fluor 647: TIRF arm set to 9000 (~67°), 500 ms exposure, 10 s/frame, 18% 640 nm laser power (~250 μW power out of the objective, ~2.5 W/cm^2 power density), 150 EM Gain. Fen1^{mKikGR} was imaged using the following settings: TIRF arm set to 9000 (~67°), 200 ms exposure, 10 s/frame, 11% 488 nm laser power (~70 μW power out of the objective, ~0.7 W/cm^2 power density), 150 EM Gain. Given these exposure durations, 6 fields of view (FOVs) could be sequentially imaged during the 10 s/frame interval. For three-color imaging we used Fen1^{mKikGR} (488 nm laser excitation), Alexa Fluor 546 (561 nm excitation), and Alexa Fluor 647 (640 nm excitation). To image Alexa Fluor 546 the following settings were used: TIRF arm set to 9000 (~67°), 500 ms exposure, 10 s/frame, 35% 561 nm laser power (~0.7 μW power out of the objective, ~7.0 W/cm^2 power density), 150 EM Gain. To maintain the time resolution of 10 s/frame time, only 4 FOVs could be sequentially imaged during a three-color experiment.

Cloning and plasmids

To prepare an antigen consisting of *Xenopus laevis* TopBP1 amino acids 718–857, the gene block gGC614 and the vector pET28b(+) linearized by NcoI-HF and SalI-HF were assembled using the NEBuilder HiFi mix (NEB, Cat# E2621S), yielding the plasmid [pLS019].

To prepare an antigen consisting of *Xenopus laevis* TopBP1 amino acids 985–1189, the gene block gGC615 and the vector pET28b(+) linearized by NcoI-HF and SalI-HF were assembled using the NEBuilder HiFi mix, yielding the plasmid [pLS020]. Plasmids [pLS019] and [pLS020] were used to express the respective antigens in bacteria.

Cloning of the full-length frog *topbp1* gene was performed as follows. The plasmid pFastBacHF-xTopBP1 encoding the *Xenopus laevis topbp1* gene was a kind gift from William G. Dunphy and Akiko Kumagai.[79] To add a GGGGSGGGGS linker (10aa), an LPETG tag for sortase labeling (sort), an FLAG tag, and a His6 tag to the C-terminus, the *topbp1* gene was PCR amplified from pFastBacHF-xTopBP1 using primers oRT008 and oRT020. The resulting PCR product, the gene block gRT21, and the pFastBac1 plasmid linealized by XhoI and NotI-HF were assembled using NEBuilder HiFi mix, yielding the plasmid [pRT005]pFastBac-xTopBP1-10aa-sort-FLAG-His6. To add a V5 tag to the N-terminus of TopBP1, oRT068 and oRT069 were annealed and the resulting DNA fragment was ligated into the [pRT005] plasmid digested at NcoI, yielding the plasmid [pRT023]pFastBac-V5-xTopBP1-10aa-sort-FLAG-His6.

To express TopBP1 in IVTT, the *topbp1* gene was PCR amplified from pFastBacHF-xTopBP1 using oRT008 and phosphorylated oRT009. The resulting PCR product was then digested with SpeI-HF and cloned into the XbaI-SmaI region of pSP64 Poly(A), resulting in the plasmid [pRT029]pSP64-xTopBP1. To add a V5 tag, an SNAP tag, and a 10aa linker to the N-terminus and an FLAG tag to the C-terminus of TopBP1, the *topbp1* gene was PCR amplified from pFastBacHF-xTopBP1 using primers oRT091 and oRT092, while the tandem V5-SNAP-10aa linker sequence was PCR amplified from the gene block gRT88 using primers oRT089 and oRT090. The two resulting PCR products were fused by overlap extension PCR using primers oRT089 and oRT092. The V5-SNAP-10aa-xTopBP1 fragment was cloned into the XbaI-SmaI site of pSP64 Poly(A), yielding the plasmid [pRT031]pSP64-V5-SNAP-10aa-xTopBP1-FLAG.

To add a tandem 10aa-SNAP-FLAG tag to the C-terminus of TopBP1, the *topbp1* gene was PCR amplified from pRT023 using primers oRT090 and oRT095, while the 10aa-SNAP-FLAG sequence was PCR amplified from the gene block gRT88 using primers oRT093 and oRT094. The two resulting PCR products were fused by overlap extension PCR using primers oRT095 and phosphorylated oRT094. The resulting V5-xTopBP1-10aa-SNAP-FLAG fragment was digested with SpeI-HF and cloned into the XbaI-SmaI site of pSP64 Poly(A), yielding the plasmid [pRT032]pSP64-V5-xTopBP1-10aa-SNAP-FLAG.

To clone the plasmid needed for expressing truncated TopBP1 in IVTT, a fragment of the *topbp1* gene was PCR amplified from pRT031 using phosphorylated primers oRT72 and oRT73. The resulting PCR product was circularized by ligation, yielding the plasmid [pRT037]pSP64-V5-SNAP-10aa-xTopBP1(Δ531-1485)-FLAG.

The *Xenopus laevis recql4* gene was cloned as follows. The *recql4* gene was PCR amplified from *Xenopus* egg cDNA from using primers oSEB005 and oSEB006. The resulting PCR product was then cloned into the NotI-HF digested pFastBac1 vector using the NEBuilder HiFi mix, resulting in the plasmid [pRT025]pFastBac-xRecQL4. The *recql4* gene cloned from cDNA prepared in our lab was 99% identical to the *recql4* sequence NM-001095713 from the NCBI database. We sequenced several *recql4* clones and most of them had the same sequence, suggesting that the differences are due to SNPs present in our frog colony. The *recql4* gene from our cDNA was used for all subsequent cloning and experiments.

To express an antigen consisting of *Xenopus laevis* RecQL4 amino acids 1145–1319, the *recql4* gene fragment was PCR amplified from pRT025 using primers oSEB036 and oSEB037. The backbone was generated by PCR amplifying pET-28b(+) using primers oLS019 and oLS023. The two PCR products were assembled using NEBuilder HiFi mix, yielding the plasmid [pSEB011]pET28-xRecQL4(1145–1319)-His6.

To express an antigen consisting of *Xenopus laevis* RecQL4 amino acids 56–165, the *recql4* gene fragment was PCR amplified from pRT025 using primers oSEB042 and oSEB043. The backbone was generated by PCR amplifying pET-28b(+) using primers oLS019 and oLS023. The two PCR products were assembled using NEBuilder HiFi mix, yielding the plasmid [pSEB014]pET28-xRecQL4(56–165)-His6.

To add an FLAG tag and a Human Rhinovirus (HRV) 3C protease cleavage sequence to the N-terminus and a tandem 10aa-sort-His_6 tag to the C-terminus of RecQL4, the *recql4* gene was PCR amplified from pRT025 using primers oRT078 and oRT079. The resulting DNA fragment was PCR amplified using primers oRT080 and oRT081, yielding a fragment that encoded FLAG-HRV3C-xRecQL4-10aa. This fragment was then PCR amplified using primers oRT082 and oRT084, yielding a fragment that encoded

FLAG-HRV3C-xRecQL4-10aa-sort-His6. This DNA fragment was inserted into the plasmid pFastBac1 linearized with BamHI-HF and XhoI using the NEBuilder HiFi mix, resulting in the plasmid [pRT027]pFastBac-FLAG-HRV3C-xRecQL4-10aa-sort-His6.

To introduce the aspartic acid 856 to alanine substitution in the Walker B motif, the gene fragments were amplified by PCR using primer pairs, oSEB011 and oRT108, and oRT107 and oRT002, using pRT025 as a template. The two PCR fragments were fused by overlap-extension PCR with primers oSEB011 and oRT002. The resulting RecQL4-D856A fragment was digested with ApaI and HindIII-HF, and cloned into the same sites in pRT027 resulting in [pRT076]pFastBac-FLAG-HRV3C-xRecQL4-D856A.

To express RecQL4 in bacteria, the FLAG-xRecQL4-10aa-sort-His6 sequence was PCR amplified from pRT027 using primers oRT083 and oRT085 and cloned into the NcoI-XhoI region of pET28b(+), resulting in [pRT026]pET28-FLAG-xRecQL4-10aa-sort-His6. However, RecQL4 expressed very poorly in bacteria.

To express RecQL4 in IVTT, the *recql4* sequence was PCR amplified from pRT027 using oRT087 and phosphorylated oRT086. The resulting PCR product was digested with XbaI and ligated into the XbaI-SmaI region of pSP64 Poly(A), yielding the plasmid [pRT030] pSP64-xRecQL4. To add a V5-SNAP-10aa tag to the N-terminus of RecQL4 and an FLAG tag to its C-terminus, the *recql4* gene was PCR amplified from pRT027 using primers oRT101 and oRT102, while the backbone was prepared by PCR amplifying [pRT031] pSP64-V5-SNAP-10aa-xTopBP1-FLAG using primers oRT090 and oRT100. These two PCR products were then assembled using NEBuilder HiFi mix, yielding the plasmid [pRT033]pSP64-V5-SNAP-10aa-xRecQL4-FLAG.

To generate a construct with a V5 tag at the N-terminus of RecQL4 and a 10aa-SNAP-FLAG tag at the C-terminus, the *recql4* gene was PCR amplified from pRT027 using primers oRT096 and oRT097, while the backbone was PCR amplified from [pRT032]pSP64-V5-xTopBP1-10aa-SNAP-FLAG using primers oRT098 and oRT099. The resulting two PCR products were assembled using NEBuilder HiFi mix, resulting in the plasmid [pRT034]pSP64-V5-xRecQL4-10aa-SNAP-FLAG.

To express RecQL4 with an N-terminal V5 tag in IVTT, the SalI-BstXI region of pRT030 was substituted with the SalI-BstXI region of pRT034, yielding plasmid [pRT035]pSP64-V5-xRecQL4.

To express RecQL4 with an N-terminal FLAG tag in IVTT, the XbaI-NdeI region of pRT030 was replaced with the XbaI-NdeI region of pRT026, yielding [pRT036]pSP64-FLAG-xRecQL4.

The *donson* gene was PCR amplified from *Xenopus* egg cDNA using primers oRT126 and oRT129. The resulting PCR fragment was digested with NotI-HF and SphI-HF, then cloned into pFastBac1, yielding the plasmid [pRT055]pFastBac-xDONSON. The *donson* gene cloned from our cDNA was an exact match to the sequence NM_001095087 in the NCBI database.

To express an antigen consisting of *Xenopus laevis* DONSON amino acids 1–175, the *donson* gene fragment was PCR amplified from pRT055 using primers oRT131 and oRT132. The backbone was generated by PCR amplifying pET-28b(+) using primers oRT119 and oRT125. The two PCR products were assembled using NEBuilder HiFi mix, yielding the plasmid [pRT056]pET28-xDONSON (1–175)-His6.

To express an antigen consisting of *Xenopus laevis* DONSON amino acids 155–579, the *donson* gene fragment was PCR amplified from pRT055 using oRT133 and oRT134. The backbone was generated by PCR amplifying pET-28b(+) using primers oRT119 and oRT125. The two PCR products were assembled using NEBuilder HiFi mix, yielding the plasmid [pRT057]pET28-His6-xDONSON(155–579).

To express DONSON in IVTT, the *donson* gene was PCR amplified from pRT055 using primers oRT146 and phosphorylated oRT145. The resulting PCR product was digested with XbaI and cloned into the XbaI-SmaI site of pSP64 Poly(A), yielding [pRT063]pSP64-xDONSON.

To add an FLAG-SNAP-10aa tag to the N-terminus of DONSON, the *donson* gene was PCR amplified from pRT055 using primers oRT144 and oRT145. Separately, the FLAG-SNAP-10aa sequence was PCR amplified from gRT88 using primers oRT090 and oRT143. The two PCR products were then fused by overlapping extension PCR using primers oRT143 and phosphorylated oRT145. The resulting DNA fragment was digested with XbaI and cloned into the XbaI-SmaI site of pSP64 Poly(A), yielding plasmid [pRT064]pSP64-FLAG-SNAP-10aa-xDONSON.

To add a 10aa-SNAP-FLAG tag to the C-terminus of DONSON, the *donson* gene was PCR amplified from pRT055 using primers oRT146 and oRT147. Separately, the 10aa-SNAP-FLAG sequence was PCR amplified from pRT034 using primers oRT093 and oRT094. The two PCR products were then fused by overlapping extension PCR using primers oRT146 and phosphorylated oRT94. The resulting DNA fragment was digested with XbaI and cloned into the XbaI-SmaI site of pSP64 Poly(A), yielding plasmid [pRT065]pSP64-xDONSON-10aa-SNAP-FLAG.

To express DONSON with an N-terminal V5 tag in IVTT, the SalI-BmgBI region of pRT064 was substituted with the SalI-BmgBI region of pRT033, yielding plasmid [pLAS025]pSP64-V5-SNAP-xDONSON.

To express DONSON in insect cells, the *donson* gene was PCR amplified from pRT065 using primers oRT126 and oRT148. The resulting product was PCR amplified again using primers oRT126 and oRT051 (this added a tandem 10aa-sort-His6 tag to the C-terminus of DONSON). The resulting PCR product was then digested with NotI-HF and SphI-HF, then cloned into the NotI-SphI region of pRT055, yielding the plasmid [pRT066]pFastBac-xDONSON-10aa-sort-His6.

To express an antigen consisting of *Xenopus laevis* Psf3 amino acids 82–240, the gene block gLS003 and pET28b(+) vector linearized by NcoI-HF and SacI-HF were assembled using NEBuilder HiFi mix, yielding plasmid [pLS032].

To clone the full length *Xenopus laevis sld5* gene the gene block gGC625 was assembled with the pET28b(+) vector linearized by NcoI-HF and SacI-HF using the NEBuilder HiFi mix, yielding plasmid [pLS029].

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

To express an antigen consisting of Sld5 amino acids 1–82, the *sld5* gene fragment and the backbone were PCR amplified from [pLS029] using phosphorylated primers oLS019 and oLS020. The resulting PCR product was circularized by ligation, yielding plasmid [pLS033].

To prevent leaky expression of target proteins from the MultiColi system (Geneva Biotech), we inserted a LacI expression cassette into the pACE2 vector, which by default lacks one. The lacI sequence was PCR amplified from pET28b(+) using primers oRT052 and oRT053. The backbone was PCR amplified from pACE2 using oRT054 and oRT055. The two PCR products fragments were assembled using NEBuilder HiFi mix, yielding plasmid [pRT018]pACE2-LacI.

To express the Xenopus laevis RPA complex in bacteria, the genes rpa1, rpa2, and rpa3 were cloned into a polycistronic expression vector as follows. The pACE2 vector backbone was linearized via PCR amplification using primers oRT024 and oRT025. To place a His6 tag at the C-terminus of RPA1, the rpa1 gene was PCR amplified from cDNA using primers oRT026 and oRT028. The resulting PCR product was cloned into the linearized pACE2 vector using the NEBuilder HiFi mix, resulting in plasmid [pRT009]pACE2-xRPA1-His6. The amino acid sequence of our *rpa1* gene exactly matches NP_001081585 from the NCBI database. The rpa1 gene (with a C-terminal His6 tag) was PCR amplified from pRT009 using primers oRT056 and oRT057. The rpa2 gene was PCR amplified from the gene block gRT022 using primers oRT058 and oRT059. The rpa3 gene was PCR amplified from the gene block gRT023 using primers oRT060 and oRT067. The backbone was generated by digesting [pRT018]pACE2-LacI with NdeI and XhoI. All four fragments were assembled by NEBuilder HiFi mix, yielding [pRT019]pACE2-LacI-xRPA.

To express *Xenopus laevis* Cdc45 in insect cells, a GGGGSGGGGS (10aa) linker, an LPETG tag (for sortase labeling), and a His6 tag were added to the C-terminus of Cdc45 as follows. The backbone was generated by linearizing pFastBac1 with Xho1, while the insert was PCR amplified from the gene block gCdc45 using primers oGC099 and oGC100. The two resulting DNA fragments were assembled using NEBuilder HiFi mix, yielding the plasmid [pGC045]pFastBac1-xCdc45-10aa-sort-His6.

To express an antigen in bacteria corresponding to full length *Xenopus laevis* Cdc45, the plasmid [pGC278]pET28b(+)-xCdc45-His6 was cloned as follows. The backbone was PCR amplified from pET28b(+) using primers oGC518 and oGC519. The insert was codon optimized and synthesized as the gene block gGC520. The two DNA fragments were then assembled into the final plasmid pGC278 using NEBuilder HiFi mix. Full-length Cdc45-His6 expressed very poorly in bacteria.

To express an antigen in bacteria corresponding to *Xenopus laevis* Cdc45 amino acids 1–361, the plasmid [pGC283]pET28b(+)-xCdc45(1-361aa)-His6 was cloned as follows. The plasmid pGC278 was PCR amplified using phosphorylated primers oGC526 and oGC527. The resulting PCR product was circularized via ligation, yielding plasmid pGC283.

Antibody preparation

Custom polyclonal antibodies were generated as previously described.[40] Briefly, the His6 tagged antigen sequence was cloned into pET28b(+) (see Cloning and Plasmids). The antigen was expressed in Rosetta (DE3) pLysS (Novagen, Cat# 70956-4) upon induction with 0.5–1.0 mM IPTG for 3 h at 37 °C or overnight at 16 °C. Soluble antigens were purified in native buffer, whereas insoluble antigens were purified in denaturing buffer (see table and 'Recombinant protein expression,purification' section below). In most cases, the Ni-NTA eluate (still containing 400 mM imidazole) was used to immunize rabbits. For some soluble antigens, additional chromatography was used to polish the protein preparation. All antibodies used in this study were raised by Cocalico Biologicals.

All antibodies used in this study were affinity purified. 1–5 mg of purified antigen was cross-linked to 1 mL of AminoLink Plus Coupling Resin (Fisher Scientific, Cat# PI20501) according to the manufacturer's protocol. 10–30 mL of serum was incubated with the affinity resin for 30 min at room temperature. The resin was washed twice with 10 mL of 1x PBS + 500 mM NaCl, and twice again with 10 mL of 1x PBS. Next, antibodies bound to the affinity resin were eluted 5–6 times with 1.0 mL of 200 mM glycine-HCl pH 2.5. Each elution was collected into an tube containing 0.1 mL of 1 M Tris-HCl pH 9.0, which immediately neutralized the glycine. Eluted fractions were pooled and dialyzed against 1 L of 1x TBS (50 mM Tris-HCl pH 8.0, 150 mM NaCl) + 10% (w/v) sucrose. The IgG concentration was measured using a Nanodrop, normalized to 1 mg/mL via dilution or spin-concentration, aliquots were flash frozen in liquid nitrogen, and stored at −80 °C.

Target Protein (Rabbit #)	Antigen	Plasmid	Antigen Purification	Depletion Condition μg IgG per μL extract	Primary Antibody Concentration for WB
Anti-xTopBP1 (SU276)	(985–1189 aa)-His6	pLS020 (bacteria)	Native	N/A	0.2 μg/mL
Anti-xTopBP1 (SU277)	(718–857 aa)-His6	pLS051 (bacteria)	Native	2:1	N/A
Anti-xPsf3 (SU279)	His6-(82–210 aa)	pLS032 (bacteria)	Denaturing	N/A	0.2 μg/mL
Anti-xSld5 (SU283)	(1–82 aa)-His6	pLS033 (bacteria)	Native	N/A	0.2 μg/mL
Anti-xGINS (SU289)	whole GINS complex, Psf2-10aa-sort-His6	pGC127 (insect)	Native + Mono Q Col.	2:1	0.2 μg/mL

(*Continued on next page*)

Continued

Target Protein (Rabbit #)	Antigen	Plasmid	Antigen Purification	Depletion Condition μg IgG per μL extract	Primary Antibody Concentration for WB
Anti-Cdc45 (SU294)	(1–361 aa)-His6	pGC283 (bacteria)	Denaturing	1:1	0.2 μg/mL
Anti-xRecQL4 (SU307)	(56–165 aa)-His6	pSEB014 (bacteria)	Denaturing	1:1	0.2 μg/mL
Anti-xRecQL4 (SU305)	(1145–1319 aa)-His6	pSEB011 (bacteria)	Denaturing	N/A	0.2 μg/mL
Anti-xDONSON (SU349)	(1–175 aa)-His6	pRT056 (bacteria)	Denaturing	N/A	0.2 μg/mL
Anti-xDONSON (SU351)	His6-(155–579 aa)	pRT057 (bacteria)	Native	1:1	0.2 μg/mL
Anti-xRPA (SU299)	whole RPA complex, RPA70-His6	pRT019 (bacteria)	Native + Heparin Col.	N/A	0.2 μg/mL

Recombinant protein expression, purification

Geminin, GINS, RPA, and Fen1^{mKikGR} were purified following previously published protocols.[37,39,78,82]

Cdc45 was purified as follows. A bacmid was generated by transforming DH10Bac cells with [pGC045]pFastBac1-xCdc45-10aa-sort-His6, and the resulting bacmid was purified with the ZR BAC DNA Miniprep Kit (Zymo Research, Cat# D4049). Baculovirus was amplified in three stages (P1, P2, P3) in Sf9 cells. xCdc45-10aa-sort-His6 was expressed in 500 mL of Tni suspension culture by infection with the P3 baculovirus for 48 h. Cells were harvested, washed with PBS, pelleted by centrifugation, and flash frozen in liquid nitrogen. Cell pellets were thawed and resuspended in four volumes of buffer N (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 20 mM imidazole, 5% glycerol) containing 1X Pierce Protease Inhibitor EDTA-free (ThermoFisher Scientific, Cat# A32965). Lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at 4 °C in a Beckman Type 70 Ti rotor (Beckman Coulter). Cleared lysate was incubated with 500 μL of Ni-NTA resin (Qiagen, Cat# 30230) for 1 h at 4 °C. The protein was eluted from the Ni-NTA resin with 300 mM imidazole in buffer N containing 1X Pierce Protease Inhibitor EDTA-free. Peak fractions were pooled, diluted with 9 volumes of buffer A (20 mM Tris-HCl pH 7.5, 5% glycerol, 1 mM DTT), and loaded on a Mono Q 5/50 GL column (Cytiva, Cat# 17516601). The protein was eluted off the column with a linear gradient of 50 mM–1000 mM NaCl in buffer A. Peak fractions were pooled and dialyzed against storage buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT). Dialyzed protein was concentrated to 1–2 mg/mL using the Amicon Ultra-4-10K column (Merck Millipore, Cat# UFC801024), snap frozen in liquid nitrogen as small aliquots, and stored at −80 °C.

TopBP1 was purified as follows. A bacmid was generated by transforming DH10Bac cells with [pRT005]pFastBac1-xTopBP1-10aa-sort-FLAG-His6, and the resulting bacmid was purified with the ZR BAC DNA Miniprep Kit. The baculovirus was amplified in three stages (P1, P2, P3) in Sf9 cells. xTopBP1-10aa-sort-FLAG-His6 was expressed in 500 mL of Sf9 suspension culture by infection with the P3 baculovirus for 48 h. Cells were harvested, washed with PBS, pelleted by centrifugation, and flash frozen in liquid nitrogen. Cell pellets were thawed and resuspended in four volumes of buffer F (25 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 1 mM DTT) containing 1X Pierce Protease Inhibitor EDTA-free. Lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at 4 °C in a Beckman Type 70 Ti rotor. Cleared lysate was incubated with 500 μL of ANTI-FLAG M2 Affinity Gel for 1 h at 4 °C. The protein was eluted from the FLAG-M2 resin with 100 μg/mL 3X FLAG peptide (APExBio, Cat# A6001) in buffer F containing 1X Pierce Protease Inhibitor EDTA-free. Peak fractions were pooled, diluted with 2 volumes of buffer Q (20 mM HEPES pH 8.0, 10% glycerol, 5 mM DTT), and incubated with 200 μL of Nuvia HP-Q (Bio-Rad, Cat# 12006693) at 4 °C for 1 h. The protein was eluted from the Nuvia HP-Q resin with buffer Q containing 300 mM NaCl, snap frozen in liquid nitrogen as small aliquots, and stored at −80 °C. We initially attempted to fluorescently label recombinant TopBP1 via sortase mediated transpeptidation. However, the fluorescent labeling efficiency of TopBP1 was low (<40%).

RecQL4-WT and RecQL4-D856A was purified as follows: A bacmid was generated by transforming DH10Bac cells with [pRT027] pFastBac1-FLAG-HRV3C-xRecQL4-10aa-sort-His6 or [pRT076]pFastBac1-FLAG-HRV3C-xRecQL4, and the resulting bacmid was purified with the ZR BAC DNA Miniprep Kit. The baculovirus was amplified in three stages (P1, P2, P3) in Sf9 cells. The protein was expressed in 250 mL of Sf9 suspension culture by infection with the P3 baculovirus for 48 h. Cells were harvested, washed with PBS, pelleted by centrifugation, and flash frozen in liquid nitrogen. Cell pellets were thawed and resuspended in four volumes of buffer F containing 1X Pierce Protease Inhibitor EDTA-free. Lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at 4 °C in a Beckman Type 70 Ti rotor. Cleared lysate was incubated with 200 μL of ANTI-FLAG M2 Affinity Gel for 1 h at 4 °C. The protein was eluted from the resin by cutting the HRV3C site with 50 μg/mL HRV3C protease in buffer F for 2 h at 4 °C. The eluted RecQL4 was separated from the HRV3C protease via size exclusion chromatography on a Superdex 200 Increase 10/300 GL column (Cytiva, Cat# 28990944). The peak fractions were pooled and concentrated, snap frozen in liquid nitrogen as small aliquots, and stored at −80 °C.

DONSON was purified as follows. A bacmid was generated by transforming DH10EMBacY cells with [pRT066]pFastBac1-xDONSON-10aa-sort-His6, and the resulting bacmid was purified with the ZR BAC DNA Miniprep Kit. The baculovirus was amplified in

three stages (P1, P2, P3) in Sf9 cells. The protein was expressed in 100 mL of Sf9 suspension culture by infection with the P3 baculovirus for 48 h. Cells were harvested, washed with PBS, pelleted by centrifugation, and flash frozen in liquid nitrogen. Cell pellets were thawed and resuspended in four volumes of buffer N containing 1X Pierce Protease Inhibitor EDTA-free. Lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at 4 °C in a Beckman Type 70 Ti rotor. Cleared lysate was incubated with 500 μL of Ni-NTA resin for 1 h at 4 °C. The protein was eluted from the Ni-NTA resin with 300 mM imidazole in buffer N containing 1X Pierce Protease Inhibitor EDTA-free. Peak fractions were pooled and dialyzed twice against 500 mL of High-salt buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 5% glycerol). After dialysis, the protein was snap frozen in liquid nitrogen as small aliquots, and stored at −80 °C. For analytical gel filtration, 100 μg of purified xDONSON-10aa-sort-His6 was loaded onto Superdex 200 Increase 10/300 GL and eluted with High-salt buffer.

His6-HRV 3C protease was purified as follows. His-HRV 3C protease was expressed in Rosetta (DE3) pLysS (Novagen) upon induction with 0.5 mM IPTG overnight at 16 °C. The Rosetta (DE3) pLysS cell pellet overexpressing the His6-HRV 3C protease was thawed in Lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% glycerol) containing 1 mM PMSF, and sonicated. The lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at 4 °C in a Beckman Type 70 Ti rotor. Cleared lysate was incubated with 0.5 mL of Ni-NTA resin for 1 h at 4 °C. The protein was eluted with 250 mM imidazole Elution buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 250 mM imidazole, 10% glycerol). Peak fractions were pooled and dialyzed twice against 1 L of Dialysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT). After dialysis, the protein was snap frozen in liquid nitrogen as small aliquots, and stored at −80 °C.

The native purification of antigens was performed as follows. The Rosetta (DE3) pLysS cell pellet overexpressing the His6-tagged antigen was thawed in Lysis buffer containing 1 mM PMSF, and sonicated. The lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at 4 °C in a Beckman Type 70 Ti rotor. Cleared lysate was incubated with 1 mL of Ni-NTA resin for 1 h at 4 °C. The protein was eluted with 400 mM imidazole in Lysis buffer containing 1 mM PMSF.

The denaturing purification of antigens was performed as follows. The Rosetta (DE3) pLysS cell pellet overexpressing His6-tagged antigen was thawed in Denaturing Lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole, 8 M Urea) containing 1 mM PMSF, and sonicated. The lysate was cleared by ultracentrifugation at 35,000 rpm for 40 min at room temperature in a Beckman Type 70 Ti rotor. Cleared lysate was incubated with 1 mL of Ni-NTA resin for 1 h at room temperature. The protein was eluted with 400 mM imidazole in denaturing elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 6 M Urea, 1 mM PMSF).

Quantification of proteins expressed in IVTT

The concentration of proteins expressed in IVTT reactions was determined by semi-quantitative western blotting. Typically, 0.05 μL of the IVTT reaction expressing the protein of interest (POI) was loaded on an SDS-PAGE gel, and a dilution series (5–200 fmol) of purified recombinant POI was loaded alongside. After transfer onto a PVDF membrane, the membrane was blotted with specific antibody against the POI. The concentration of POI in the IVTT reaction was estimated by comparing its band intensity to those from the dilution series of purified proteins. To quantify V5-SNAP-xTopBP1(Δ531-1485)-FLAG, FLAG-tagged recombinant TopBP1 was used as a standard and the membrane was blotted with FLAG antibody because the truncated TopBP1 lacks the epitope region for our TopBP1 antibody.

Fluorescent labeling using sortase

First, the peptide GGGGYKCK was reacted with Alexa Fluor 647 maleimide or Alexa Fluor 546 maleimide. The resulting peptide-dye product was separated from unreacted dye and unreacted peptide via HPLC. Next, 1–2 nmol of purified recombinant protein (GINS, Cdc45, or DONSON) containing the LPETG tag was incubated with a 10-fold molar excess of labeled peptide and a 0.5-fold molar excess of SrtA-His6 enzyme in sortase buffer (5 mM $CaCl_2$, 1 mM DTT, 20 mM Tris pH 7.5, and 150 mM NaCl). The reaction was performed overnight at 4 °C with gentle end-over-end rotation. Upon sortase-mediated conjugation of the fluorescent peptide to the recombinant protein, the His6 tag was removed from the labeled protein. The unreacted protein as well as the Srt-His6 enzyme were captured by incubating the reaction for 1 h with 50 μL of Ni-NTA resin. The labeled protein and excess peptide remained in the flowthrough. The labeled protein was separated from the excess peptide via size exclusion chromatography on Superdex 200 Increase 10/300. The peak fractions were pooled and concentrated. Protein labeling efficiency was estimated based on absorbance measurements at 280 nm and the absorption peak for the respective Alexa Fluor dye. The biochemical activity of the labeled protein was tested following a published detailed protocol.[40] Alexa Fluor 647 did not significantly affect the biochemical activity of GINS or Cdc45 expressed in insect cells.

Fluorescent labeling using SNAP

After expressing TopBP1 or RecQL4 in the IVTT lysate, the reaction was kept on ice, and the concentration of the target protein was determined via quantitative western blotting. The SNAP labeling reagent was used in a 2-fold molar excess relative to the target protein. SNAP-Surface Alexa Fluor 546 or SNAP-Surface Alexa Fluor 647 was added to the IVTT reaction containing the target protein and reacted overnight at 4 °C with gentle end-over-end rotation. The final reaction was aliquoted, snap frozen in liquid nitrogen, and stored at −80 °C. Fluorescent labeling efficiency was estimated by loading the IVTT reaction on a PAGE-Gel alongside a titration series of purified fluorescent protein standards ($GINS^{AF546}$ or $GINS^{AF647}$). The labeled proteins were not purified after the SNAP-labeling reaction. The IVTT reaction was directly added replication reactions as the source of a given protein.

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

QUANTIFICATION AND STATISTICAL ANALYSIS

Data reporting

Details of the statistical analyses performed are indicated in the figure legends. Most of the quantities measured in this study did not have symmetric distributions, hence we reported the median of the distribution. Unless otherwise specified in the figure captions, error bars represent the 95% confidence interval for the median. This was calculated using MATLAB's built in bootstrapping function. The exact meaning of the number 'N' reported in each figure is indicated in the figure captions, it usually refers to the number of DNA molecules, the number of replication origins, or the number of binding events. We typically used N∼100 events to report the duration of an intrinsically stochastic event (for example the duration of a productive TopBP1 binding event). P-values were computed using the two-sample Kolmogorov-Smirnov test function built into MATLAB; 'n.s.' denotes not significant, * = p value < 0.05, ** = p value < 0.01, *** = p value < 0.001, **** = p value $\leq$ 0.0001.

Reproducibility and blinded analysis

All experiments were repeated at least twice, with independent extract preparations. The data presented in the main figures comes from one of these preparations. In the supplemental figures we compare the results from two biological repeats. Titration series (for example Figure 1E) and treatment series (for example Figures 1J, 2H and 2I) were performed with the same extract for consistency. The data presented in Figures 6D–6G was analyzed in a blinded manner. Specifically, raw data from Mock, RecQL4-depletion, and WT or D856A add-back experiments were placed in a common folder, the order of individual DNA molecules was scrambled, and the exact experiment corresponding to each DNA molecule was unknown to the user performing the analysis. The results were then unscrambled using automated code in MATLAB.

Single-molecule data analysis

Each replication initiation experiment imaged 4–6 fields of view (FOVs) over the course of 45 min with a resolution of 10 s/frame. Each FOV typically contained 50–200 DNA molecules stretched to 80–90% of their contour length. A suite of MATLAB scripts written in house were used to select and analyze data as described in our recent methods article.[40] Below we summarize the data analysis pipeline specific for replication initiation experiments.

Selecting DNA molecules

Regions of interest (ROIs) corresponding to individual DNA molecules were selected manually based on the previews of DNA stained with SYTOX Green (these were acquired before the start of the experiment). Only well-separated DNA molecules were used in the analysis.

Quantifying the Fen1 signal

Custom MATLAB scripts written in house were used to generate kymograms corresponding to each selected DNA molecule. Origin firing was monitored by visualizing the binding of Fen1^{mKikGR} to nascent DNA. Importantly, the replication fork speed (measured from the growth rate of the Fen1^{mKikGR}-labeled bubble), was the same at low concentrations (5–20 nM) of GINS^{AF647}/Cdc45^{AF647} as in mock-depleted extract, indicating that fluorescent labeling did not impair GINS/Cdc45 function. The raw Fen1 signal was automatically thresholded using OTSU's method via a built-in MATLAB function. The OTSU threshold binary image was used to fit straight lines to the edges of the replication bubble, and the fork speed value was extracted from the slope of the line (taking into account the pixel size, the time resolution of the experiment, and the extent to which DNA was stretched). Finally, the raw Fen1 signal was integrated over time, and the integrated signal was fit by two straight line segments (see Figure S7H): the first segment represents the background noise before origin firing, while the second segment corresponds to a linear increase in Fen1 signal after the start of DNA synthesis. The intersection of the two segments corresponds to the time of origin firing. This estimate agrees well with the origin firing time prediction from the OTSU threshold analysis.

Quantifying GINS/Cdc45 recruitment

The raw GINS/Cdc45 signal was integrated and analyzed using MATLAB's 'findchangepts' function, which identified the moment in time when the signal changed from "low" to "high" (see Figure S7H). This change point marked the time when GINS bound to the pre-RC. Given that in most cases GINS/Cdc45 arrive to the origin as a pre-formed dimer, we could determine the time of GINS/Cdc45 recruitment even when only one GINS/Cdc45 copy was fluorescently labeled (see Figure S7H, right side).

Limitations of GINS/Cdc45 data analysis

In 20–40% of origin firing events (depending on GINS/Cdc45 concentration), our changepoint analysis code scored that GINS/Cdc45 were recruited sequentially, one copy at a time. To assist in reading this section, several representative examples are shown in Figure S7K. It is possible that sometimes a single molecule of GINS/Cdc45 is recruited to the pre-RC and a second molecule is quickly recruited from solution before the scaffold+[GINS/Cdc45] dissociate. However, there are several technical aspects of data acquisition and analysis that could lead to GINS/Cdc45 recruitment being scored as sequential, even if the two GINS/Cdc45 molecules were recruited simultaneously.

(i) The integrated GINS/Cdc45 signal is noisy, and the changepoint analysis code may sometimes overfit the noisy data, thus scoring simultaneous GINS/Cdc45 binding as sequential recruitment.

(ii) The GINS/Cdc45 signal is acquired during a 0.5 s exposure, while new images are taken every 10 s. This means that sometimes, two GINS/Cdc45 molecules are recruited halfway through the exposure, resulting in a lower signal. This signal is then scored as "sequential recruitment" by our changepoint analysis.

(iii) The Alexa Fluor 647 dye used to label GINS/Cdc45 sometimes blinks (see Figure S7L). If two GINS/Cdc45 molecules are recruited simultaneously but one fluorophore is in the dark state, binding will be scored as sequential.

(iv) We observed very short-lived non-productive docking of a single GINS/Cdc45 molecule to pre-RC (Figure S7K, red arrows). If such an event occurs immediately before the stable binding of two GINS/Cdc45 molecules, we may be unable to observe the dissociation of the single GINS/Cdc45 molecule. Therefore, the entire initiation event will be scored as sequential recruitment of GINS/Cdc45.

Analyzing firing factor binding/dissociation

To analyze the binding and dissociation of firing factors (TopBP1, RecQL4, and DONSON) we employed the Generalized Likelihood Ratio Test (GLRT). GLRT was used to convert noisy single-molecule images into binary images that were then segmented into binding events, as shown in Figure S7M. The GLRT algorithm uses a sliding window (we used a window with a 5-pixel radius) to determine whether the signal at a given pixel is consistent with noise or is significantly different from noise. This approach is superior to simple image thresholding because it accounts for the intrinsic properties of background noise. We adapted GLRT code written in MATLAB from a previous publication.[83] We did not score any events that lasted a single frame because we could not be confident that they represented bona-fide binding.

Global protein binding analysis

To quantify the binding of TopBP1, Cdc45, and GINS to the pre-RC in mock versus depleted extracts, we devised a global binding analysis. First, ROIs with and without DNA were selected based on the SYTOX Green images. Next, kymograms were generated for both regions and the GLRT algorithm was employed to identify binding events (including single-frame events). The total duration of all binding events was added and normalized to DNA length (s/kb). Matched selections without DNA were used to subtract non-specific background binding (which may vary between experiments).

Analyzing in vitro photobleaching data

Along with gel filtration, the *in vitro* photobleaching assay was used to determine the stoichiometry of recombinant proteins used in this study. First, a flow cell was constructed as previously described.[40] The fluorescently labeled recombinant protein of interest was flown into the flow cell at a concentration of 10–100 pM (diluted in 1x ELB-sucrose buffer) and incubated for 5 min. During this incubation the proteins non-specifically attached to the coverslip surface. The number of proteins immobilized on the coverslip were estimated first. If the density of immobilized proteins was too low (<100 molecules per FOV), a higher concentration of protein was incubated in the flow cell. If the density of immobilized proteins was too high (>1000 molecules per FOV), a new flow cell was prepared and incubated with a lower concentration of recombinant protein. In an optimal experiment 200–400 molecules were present in a field of view of 81 microns by 81 microns (this is captured by our 100x objective and EMCCD camera). Once optimal fluorescent protein density was achieved, the field of view was imaged continuously until all fluorophores photobleached (usually several hundred frames). Imaging was conducted using TIRF illumination conditions identical to those in a typical replication experiment. Individual molecules were automatically segmented using a MATLAB script, and the integrated intensity of individual spots was analyzed using the changepoint algorithm.

Bioinformatic analysis

The conservation of domain architecture for TopBP1, DONSON, and RecQL4 was determined through the identification and characterization of putative sequence homologs via BLASTp (v2.15.0) searches against the EukProt V3 database[74] using Geneious Prime 2024.0.3 (https://www.geneious.com).

For DONSON, the human protein sequence (Uniprot Q9NYP3) was used to identify 250 hits which were annotated using InterProScan[84] to obtain the MobiDB-lite disorder predictions[85] and PANTHER[86] annotations. Submitting the hits to these databases identified putative disordered regions and the DONSON homology region (PTHR12972). Hits were organized by the taxonomic classification of the species with Python code written to query the NCBI Taxonomy database.[87] Taxa were selected to capture the breadth of eukaryotes and provide extra detail within fungi. The presence or absence of domains within each gene was summarized for each of the selected taxa.

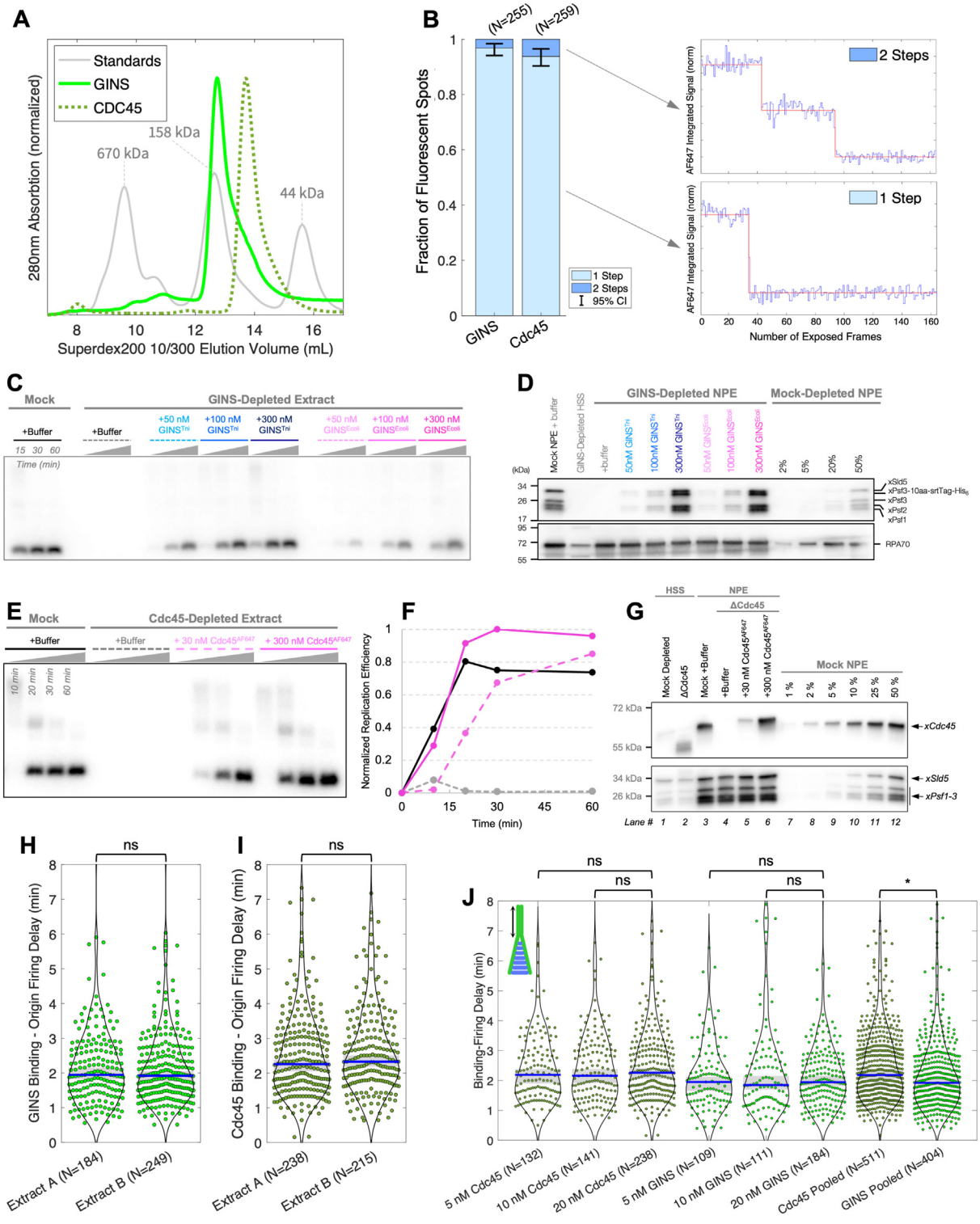
For RecQL4, 19 annotated Sld2/RecQL4 sequence orthologs were aligned (Uniprot O14216, O94761, P34252, Q0UZV8, Q2UQ97, Q4HW93, Q4WXD3, Q5B8U0, Q6BXN9, Q6C7D7, Q6CK38, Q6FME9, Q7S438, Q59UH5, Q75DR2, Q75NR7, D4A5W5, Q9VSE6, Q33DM4). The consensus sequence was extracted for the Sld2 N-like region, and any uncertainties were replaced with the human residues resulting in a 49 AA sequence: ERLAQLRAELKEWERAFAAQNGRKPSKDDIKAAPEIAAKYKEYSKLKKK. Searching with this sequence obtained 100 hits which were annotated using InterProScan[84] to obtain the CDD[88] annotations. Submitting the hits to this database identified the Sld2 homology domain (cd22289), and the ATPase domain of RecQL4 (cd18018). Hits were organized by the

Please cite this article in press as: Terui et al., Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa, Cell (2024), https://doi.org/10.1016/j.cell.2024.05.024

taxonomic classification of the species with Python code written to query the NCBI Taxonomy database.[87] Taxa were selected to capture the breadth of eukaryotes and provide extra detail within fungi. Finally, the presence or absence of domains within each gene was summarized for each of the selected taxa.

For TopBP1, the human protein sequence (Uniprot Q92547) was used to identify 1000 hits, which were then annotated using InterProScan[84] to obtain the Gene3D[89] annotations to identify BRCT domains (G3DSA:3.40.50.10190). Unlike PANTHER and CDD, only Gene3D accurately identified the accepted BRCT domains for the human (BRCT0-8, Uniprot Q92547), *Mus musculus* (BRCT0-8, Uniprot Q6ZQF0), *Xenopus laevis* (BRCT0-8, Uniprot Q800K6), *Saccharomyces cerevisiae* (BRCT1-4, Uniprot P47027), *Schizosaccharomyces pombe* (BRCT1-4, Uniprot P32372), and *Arabidopsis thaliana* (BRCT0-5, Uniprot Q70X85) orthologs of TopBP1. Using code written in Python, the domain annotations of all 1000 hits were used to calculate how many BRCT domains are present in the TopBP1 ortholog in each species. Species were then organized by their taxonomic classification and grouped in the same way as DONSON and RecQL4, and the distribution of BRCT domains per ortholog was plotted in MATLAB.

Supplemental figures



(legend on next page)

Figure S1. Validation and characterization of recombinant GINS and Cdc45, related to Figure 1

(A) Gel filtration chromatogram of recombinant GINS (MW = 98 kDa), Cdc45 (MW = 66 kDa), and gel filtration standards.

(B) The oligomeric state of GINS and Cdc45 was measured using an *in vitro* photobleaching assay. Left: bar plots showing the probability that spots corresponding to labeled GINS/Cdc45 photobleached in one step (light blue) or two steps (dark blue). Right: representative traces illustrating 1-step and 2-step photobleaching. N represents the number of fluorescent spots analyzed; error bars represent 95% CI estimated via bootstrapping.

(C) Plasmid replication assay comparing the biochemical activity of mock-depleted extract versus GINS-depleted extract supplemented with buffer, 50–300 nM of recombinant GINS purified from insect cells (Tni) or bacteria (*E. coli).* GINS purified from bacteria was roughly half as active as GINS purified from insect cells.

(D) Immunoblots of samples corresponding to reactions from panel (C). RPA70 was used as a loading control. The replication reaction consists of equal parts HSS, NPE, and buffer. Based on this blot, the estimated concentration of GINS is ~200 nM in the replication reaction and ~600 nM in undiluted NPE.

(E) Plasmid replication assay comparing the biochemical activity of mock-depleted extract versus Cdc45-depleted extract supplemented with buffer or with 30–300 nM of fluorescently labeled recombinant Cdc45.

(F) Plot of integrated signal intensities from panel (E) with matching line styles and colors.

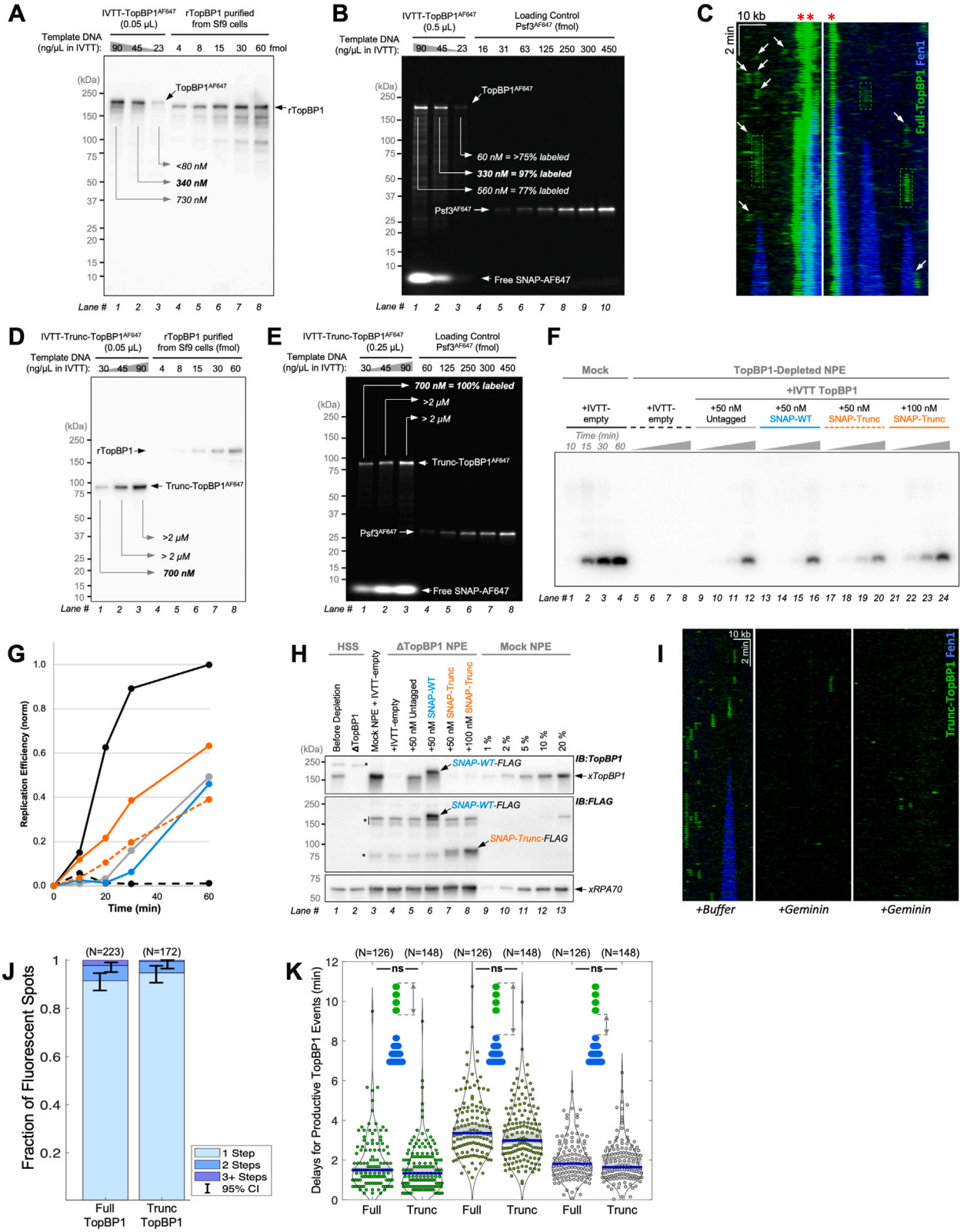
(G) Immunoblots of samples corresponding to reactions from panel (E). GINS was used as a loading control.

(H) Time delay between GINS binding and origin firing measured using two different egg extract preparations.

(I) Time delay between Cdc45 binding and origin firing measured using two different egg extract preparations.

(J) Time delay between Cdc45 or GINS binding and origin firing for 5 nM, 10 nM, and 20 nM concentrations of each protein. Experiments were conducted with the same extract preparation.

For panels (H)–(J), p values were calculated using the two-sample Kolmogorov-Smirnov test, "ns" corresponds to $p > 0.05$, * denotes $p \leq 0.05$.



(legend on next page)

Figure S2. Validation and characterization of recombinant TopBP1, related to Figure 2

(A) SNAP-tagged TopBP1 was expressed via *In Vitro* Transcription-Translation (IVTT) by using the indicated concentration of template plasmid in IVTT reactions. SNAP-tagged TopBP1 expressed in IVTT was labeled with Alexa Fluor 647 (AF647) by incubating the IVTT reaction with SNAP-Surface AF647 at 4 °C overnight. IVTT-TopBP1^{AF647} was separated by SDS-PAGE, and TopBP1 concentration was estimated via semi-quantitative western blotting.

(B) IVTT-TopBP1^{AF647} fluorescent labeling efficiency was estimated by comparing the AF647 signal with known amounts of GINS^{AF647} where the fluorophore is on Psf3. The 330 nM TopBP1^{AF647} preparation (lane 2) was used for single-molecule assays.

(C) Representative kymograms from a single-molecule experiment with 10 nM of full-length TopBP1^{AF647}. Boxes indicate productive TopBP1 binding events, arrows show non-productive events, red asterisks mark large TopBP1 assemblies that may represent liquid-liquid phase separation (LLPS).

(D) Semi-quantitative immunoblot of truncated TopBP1 expressed via IVTT.

(E) Estimation of trunc-TopBP1 fluorescent labeling efficiency. The 700 nM Trunc-TopBP1^{AF647} preparation (lane 1) was used for single-molecule assays.

(F) Plasmid replication assay comparing the biochemical activity of mock-depleted reaction (lanes 1–4) versus TopBP1-depleted reaction supplemented with buffer (lanes 5–8), or TopBP1 constructs expressed in IVTT: 50 nM of untagged TopBP1 (lanes 9–12), 50 nM of SNAP-tagged full-length wild-type TopBP1 (lanes 13–16), and 50 nM or 100 nM of SNAP-tagged truncated TopBP1 (lanes 17–20 and lanes 21–24, respectively).

(G) Plot of integrated signal intensities from panel (F). 50 nM (~20% of endogenous TopBP1 concentration) of untagged (solid gray), SNAP-tagged full length (solid blue), and SNAP-tagged truncated TopBP1 (dashed orange) rescue DNA replication to ~50% level of the mock reaction.

(H) Immunoblots of samples corresponding to reactions from panel (F).

(I) Licensing DNA in HSS supplemented with 300 nM geminin (which inhibits pre-RC loading) greatly reduced trunc-TopBP1 binding in the single-molecule assay.

(J) *In vitro* photobleaching analysis of full-length TopBP1^{AF647} and truncated TopBP1^{AF647} (N represents the number of fluorescent spots; error bars: 95% CI).

(K) Measured time delays for productive TopBP1 binding events, comparing full-length and truncated constructs.

N represents the number of origins; blue bars mark the median, gray boxes denote the 95% CI; "ns" corresponds to $p > 0.05$.

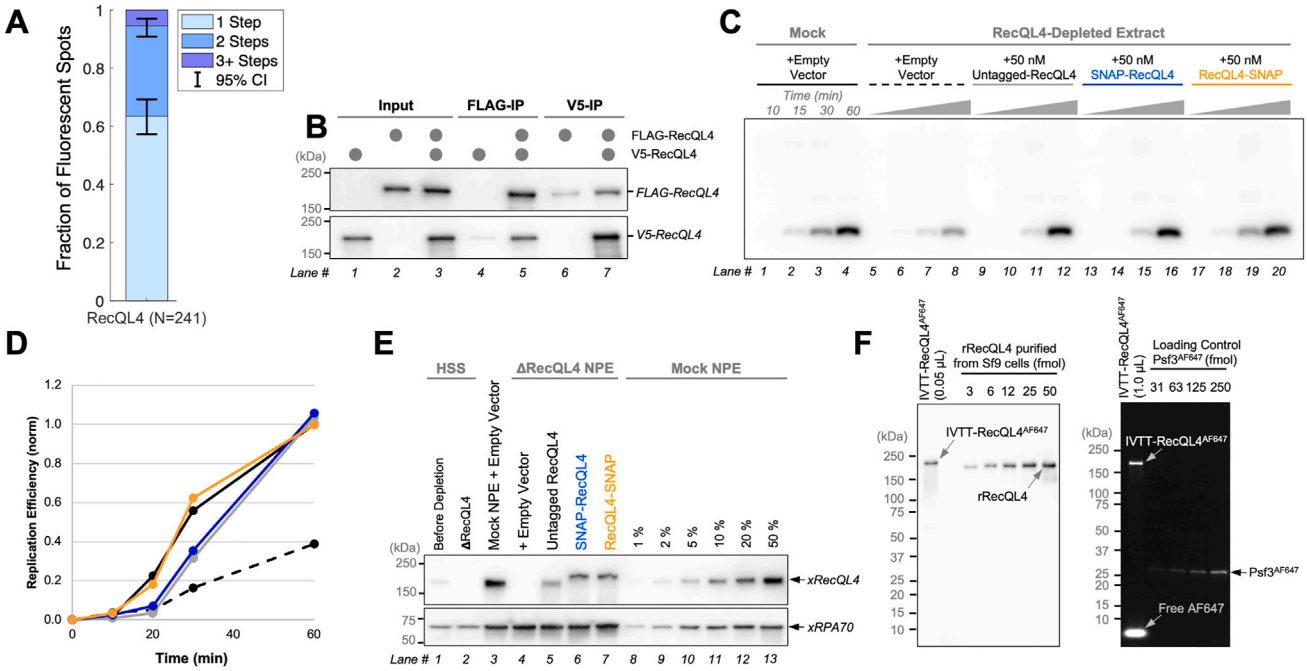


Figure S3. Validation and characterization of recombinant RecQL4, related to Figure 3

(A) The oligomeric state of IVTT-RecQL4^{AF647} was measured using an *in vitro* photobleaching assay. Bar plot depicts the probability that spots corresponding to labeled RecQL4 photobleached in one step (light blue), two steps (medium blue), or more steps (dark blue). N denotes the number of fluorescent spots analyzed, error bars represent the 95% CI.

(B) Co-IP experiment of FLAG-RecQL4 and V5-RecQL4. N-terminally V5-tagged RecQL4 (lanes 1, 3–5, 7) and N-terminally FLAG-tagged RecQL4 (lanes 2, 3, 5–7) were expressed via IVTT, and immunoprecipitated using anti-FLAG beads (lanes 4 and 5) or anti-V5 beads (lanes 6 and 7). Samples were separated by SDS-PAGE, transferred to a PVDF membrane, and blotted with anti-FLAG (upper panel) or anti-V5 antibody (lower panel). V5-RecQL4 co-precipitated with FLAG-RecQL4 (lanes 4 and 5), and vice versa (lanes 6 and 7), suggesting that RecQL4 forms oligomers.

(C) Plasmid replication assay comparing the biochemical activity of mock-depleted extract (lanes 1–4) versus RecQL4-depleted extract supplemented with buffer (lanes 5–8), 50 nM of untagged RecQL4 (lanes 9–12), 50 nM of N-term SNAP-tagged RecQL4 (lanes 13–16), or 50 nM of C-term SNAP-tagged RecQL4 (lanes 17–20).

(D) Plot of integrated signal intensities from panel (C) with matching line styles and colors.

(E) Immunoblots of samples corresponding to reactions from panel (C). 50 nM RecQL4 corresponds to ~15% of endogenous RecQL4 concentration. RPA70 was used as a loading control.

(F) Estimation of RecQL4^{AF647} fluorescent labeling efficiency. C-terminally SNAP tagged RecQL4 was expressed via IVTT and incubated with SNAP-Surface AF647 at 4 °C overnight. Left: RecQL4 concentration was estimated by semi-quantitative western blotting. Right: RecQL4 fluorescent labeling efficiency was estimated by comparing the AF647 signal with known amounts of GINS^{AF647} where the fluorophore is on Psf3. This preparation yielded ~210 nM of RecQL4, of which ~200 nM was fluorescently labeled (~95% labeling efficiency).

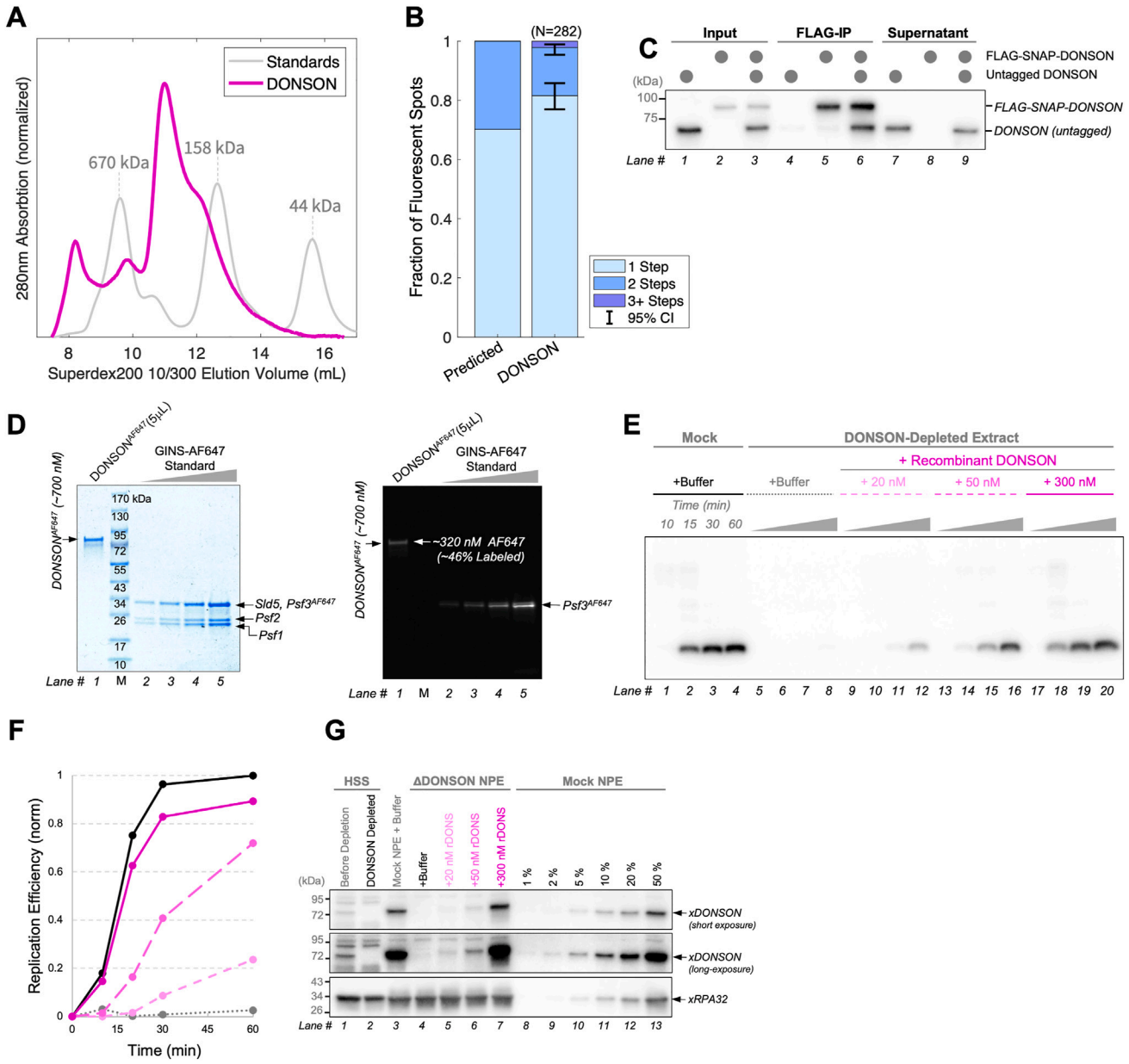


Figure S4. Validation and characterization of recombinant DONSON, related to Figure 4

(A) Gel filtration chromatogram of recombinant DONSON (MW = 64 kDa) with standards.

(B) *In vitro* photobleaching analysis of recombinant DONSON, of which ~46% was labeled with AF647. Left bar: the predicted probability of 1-step and 2-step photobleaching events given the ~46% labeling efficiency under the assumption that DONSON forms stable dimers *in vitro*. Right bar: the measured probability that DONSON^{AF647} spots photobleached in one, two, or more steps. N represents the number of fluorescent spots analyzed, error bars denote the 95% CI.

(C) CoIP experiment of untagged DONSON and N-terminally FLAG-SNAP-tagged DONSON (FLAG-SNAP-DONSON). Untagged DONSON (lanes 1, 3, 4, 6, 7, and 9) and FLAG-SNAP-DONSON (lanes 2, 3, 5, 6, 8, and 9) were expressed by IVTT, and were immunoprecipitated using anti-FLAG beads (lanes 4–6). Samples were separated by SDS-PAGE, transferred to a PVDF membrane, and blotted with anti-DONSON antibody. Untagged DONSON co-precipitated with FLAG-SNAP-DONSON, suggesting that DONSON forms oligomers.

(D) Left: DONSON^{AF647} was separated by SDS-PAGE and stained with Coomassie brilliant blue. Right: DONSON fluorescent labeling efficiency was estimated by comparing the AF647 signal with known amounts of GINS^{AF647} where the fluorophore is on Psf3. This preparation yielded ~700 nM of DONSON, of which ~320 nM was fluorescently labeled (~46% labeling efficiency).

(E) Plasmid replication assay comparing the biochemical activity of mock-depleted extract (lanes 1–4) versus DONSON-depleted extract supplemented with buffer (lanes 5–8), 20 nM (lanes 9–12), 50 nM (lanes 13–16), or 300 nM (lanes 17–20) of recombinant DONSON purified from Sf9 insect cells.

(F) Plot of integrated signal intensities from panel (E) with matching line styles and colors.

(G) Immunoblots of samples corresponding to reactions from panel (E). 50 nM DONSON corresponds to ~5% of endogenous DONSON concentration in NPE. RPA32 was used as a loading control.

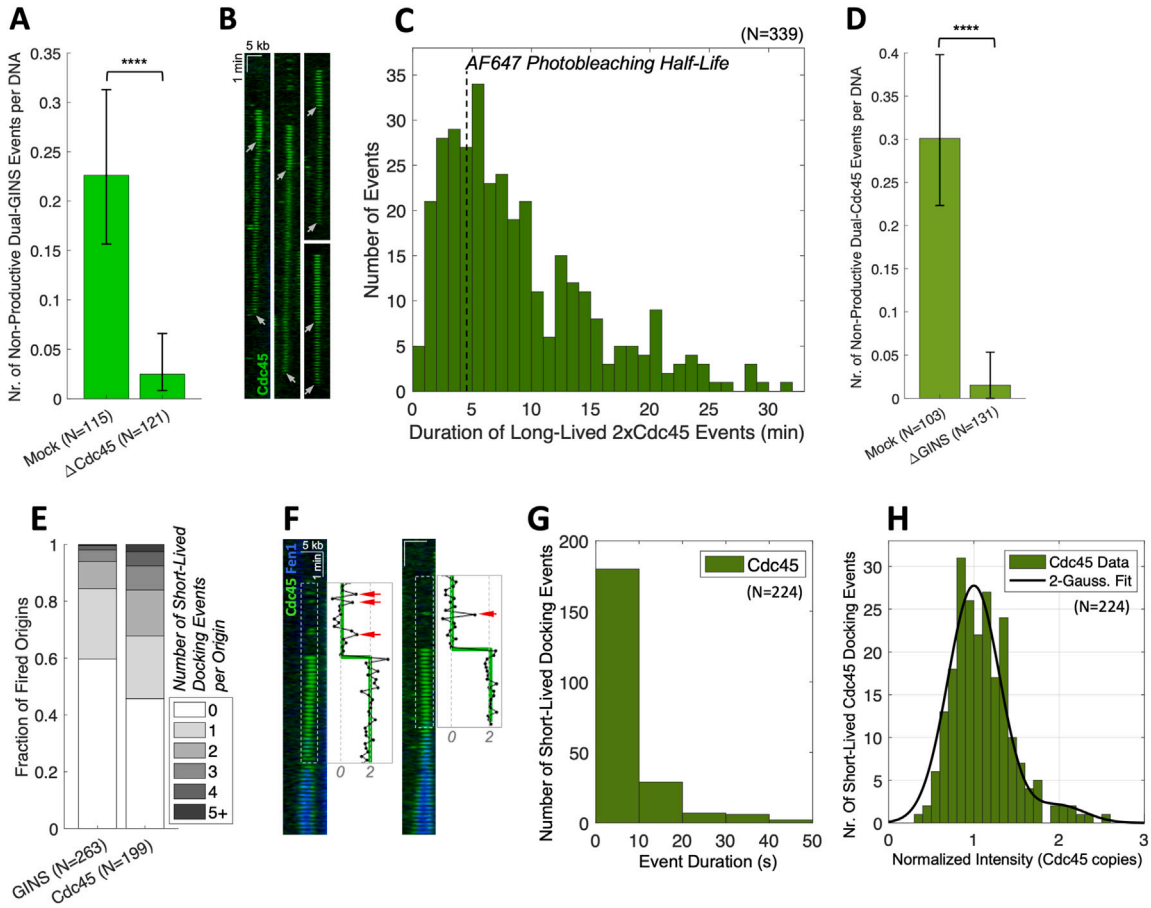


Figure S5. Long-lived non-productive Cdc45 binding events, related to Figure 5

(A) The frequency of non-productive long-lived binding events that contain two copies of GINS^{AF647} (i.e., they photobleach in two steps) in mock-depleted versus Cdc45-depleted extract (N denotes the number of DNA molecules analyzed; error bars represent the 95% CI; **** indicates $p \leq 0.0001$).

(B) Representative kymograms showing non-productive dual-Cdc45^{AF647} recruitment (gray arrows mark photobleaching).

(C) The duration of long-lived non-productive dual-Cdc45^{AF647} binding events (N represents the number of binding events).

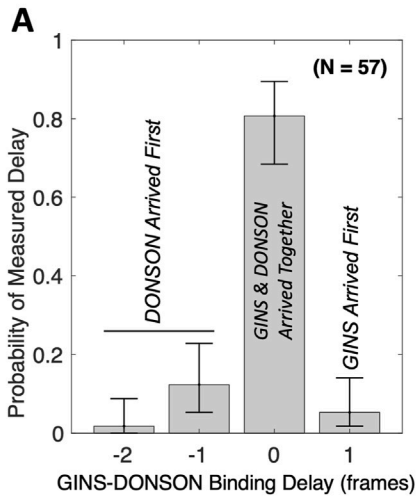
(D) The frequency of non-productive long-lived dual-Cdc45^{AF647} binding events in mock-depleted versus GINS-depleted extract (N denotes the number of DNA molecules analyzed; error-bars represent the 95% CI; **** indicates $p \leq 0.0001$).

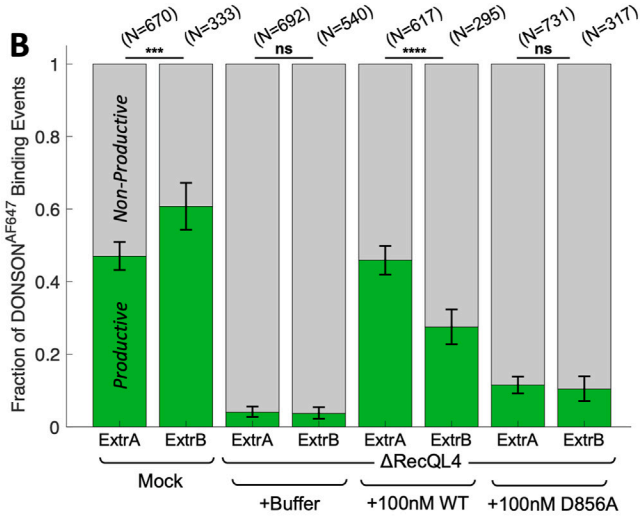
(E) The fraction of fired origins that were preceded by 0, 1, 2, 3, 4, or more short-lived GINS or Cdc45 docking events (N represents the number of origins analyzed).

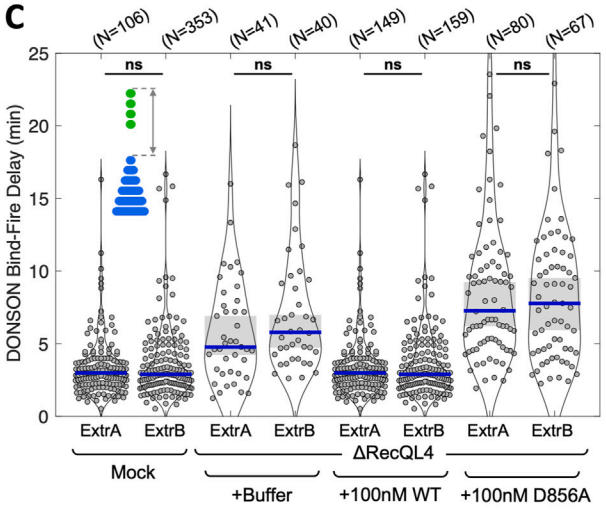
(F) Kymograms showing short-lived docking of Cdc45^{AF647} before origin firing (inset, red arrows).

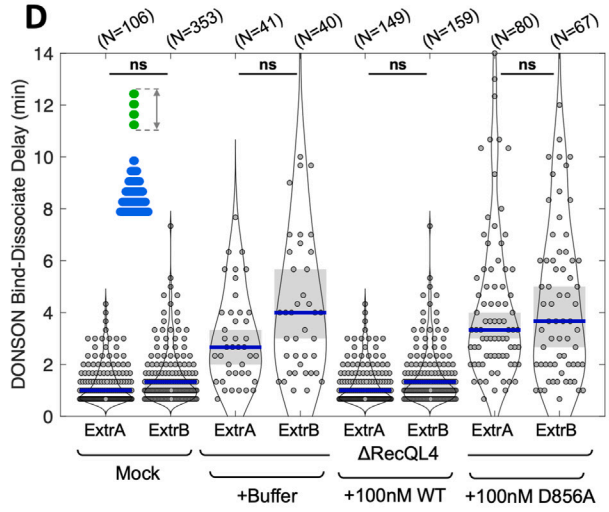
(G) The duration of short-lived Cdc45 docking events (N represents the number of binding events).

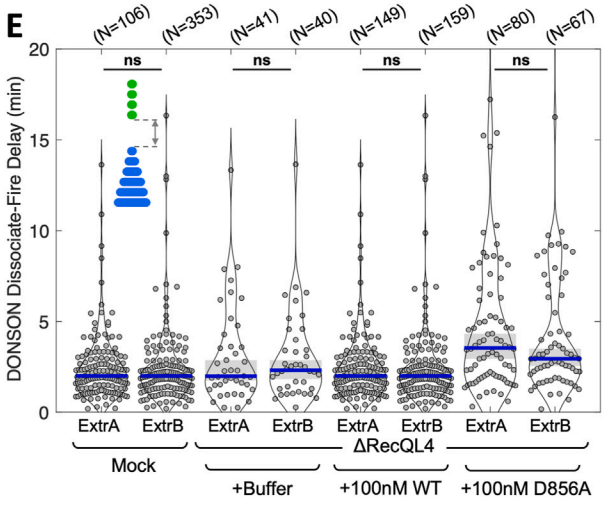
(H) The normalized intensity of short-lived Cdc45 docking events. Black curve shows the fit to two Gaussians centered at 1.0 and 2.0 respectively.

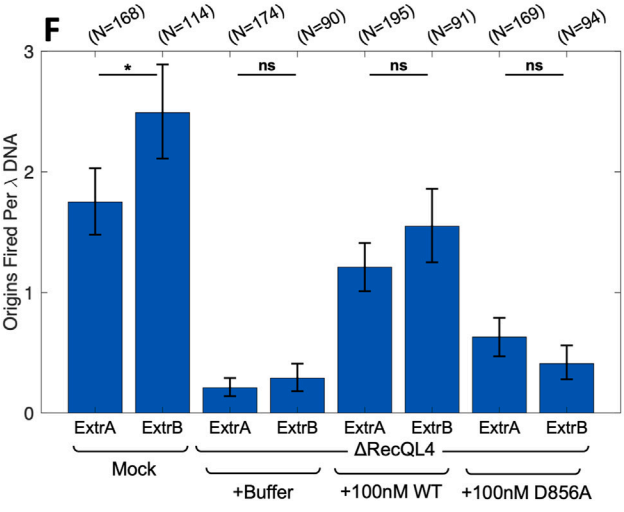












(legend on next page)

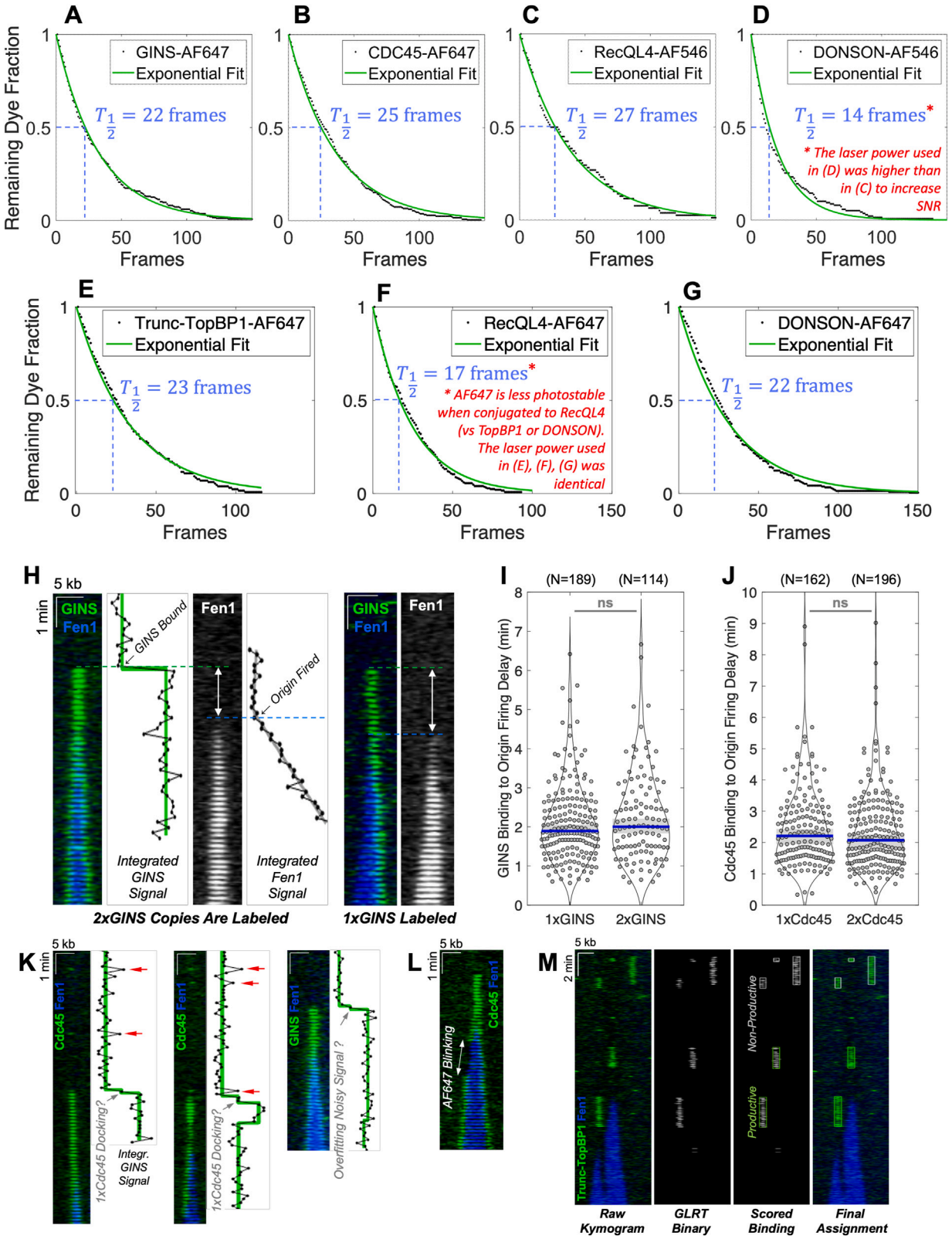
Figure S6. RecQL4 facilitates DONSON dissociation from CMG, related to Figure 6

(A) The distribution of measured time delays between GINS recruitment and DONSON binding for non-productive events (time is shown in movie frames (1 frame = 10 s); N represents the number of non-productive events where both GINS and DONSON were detected; error bars denote the 95% CI).

(B) Fraction of productive (green) and non-productive (gray) DONSON binding events measured with two different extract preparations (N represents the number of binding events; error bars denote the 95% CI; "ns" corresponds to $p > 0.05$, *** corresponds to $p \leq 0.001$, **** corresponds to $p \leq 0.0001$). Note that in the experiment corresponding to ExtrB: RecQL4-depletion +100 nM WT we observed a higher than usual non-specific binding of $DONSON^{AF647}$, leading us to over-estimate the fraction of non-productive events.

(C–E) The distribution of measured time delays for productive DONSON binding events. Each experiment was repeated with two different extract preparations. N represents the number of origins, blue lines mark median values, gray boxes denote the 95% CI; "ns" denotes $p > 0.05$.

(F) Origin firing efficiency for the single-molecule replication reaction containing 20 nM $DONSON^{AF647}$. Each experiment was repeated with two different extract preparations. N represents the number of DNA molecules; error bars represent the 95% CI; "ns" denotes $p > 0.05$; * corresponds to $p \leq 0.05$.



(legend on next page)

Figure S7. Reproducibility and technical limitations, related to Figure 7

(A–G) Photobleaching decay curves for GINS^{AF647}, Cdc45^{AF647}, RecQL4^{AF546}, DONSON^{AF546}, trunc-TopBP1^{AF647}, RecQL4^{AF647}, DONSON^{AF647}. In each case, the data was well fit by a single-exponential decay, and the estimated half-life $T_{1/2}$ was extracted from the fit (measured in exposure frames). Imaging conditions (exposure, laser power, TIRF angle) were identical to KEHRMIT experiments in this study. In panel (D) and the corresponding KEHRMIT experiments, a higher laser power was used than in panel (C) to increase the signal-to-noise (SNR) ratio. Even though the same laser power was used in panels (A), (B), and (E)–(G) and the corresponding KEHRMIT experiments, AF647 was less photostable when conjugated to RecQL4 (presumably due to a different local chemical environment).

(H) Representative kymograms showing origin firing events where both GINS molecules were fluorescently labeled (left) or only one GINS molecule was labeled (right).

(I) Time delay between GINS binding and origin firing for origins where both GINS were labeled (2xGINS), and origins where only one GINS was labeled (1xGINS). Blue bars denote the median; gray boxes mark the 95% CI; N represents the number of origins analyzed; "ns" denotes $p > 0.05$.

(J) Time delay between Cdc45 binding and origin firing for origins where both Cdc45 were labeled (2xCdc45), and origins where only one Cdc45 was labeled (1xCdc45). Blue bars denote the median; gray boxes mark the 95% CI; N represents the number of origins analyzed; "ns" denotes $p > 0.05$.

(K) Representative kymograms depicting events where two copies of GINS/Cdc45 appear to be recruited sequentially. Potential interpretations are discussed in the STAR Methods section.

(L) Kymogram illustrating AF647 blinking. Implications are discussed in STAR Methods.

(M) Kymogram illustrating how the generalized likelihood ratio test (GLRT) was used to measure the duration of transient binding events for replication initiation factors (TopBP1 in this case).