Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6

Jan Marten Schmidt1,2 & Franziska Bleichert3✉

Eukaryotic DNA replication initiation relies on the origin recognition complex (ORC), a DNA-binding ATPase that loads the Mcm2–7 replicative helicase onto replication origins. Here, we report cryo-electron microscopy (cryo-EM) structures of DNA-bound Drosophila ORC with and without the co-loader Cdc6. These structures reveal that Orc1 and Orc4 constitute the primary DNA binding site in the ORC ring and cooperate with the winged-helix domains to stabilize DNA bending. A loop region near the catalytic Walker B motif of Orc1 directly contacts DNA, allosterically coupling DNA binding to ORC’s ATPase site. Correlating structural and biochemical data show that DNA sequence modulates DNA binding and remodeling by ORC, and that DNA bending promotes Mcm2–7 loading in vitro. Together, these findings explain the distinct DNA sequence-dependencies of metazoan and S. cerevisiae initiators in origin recognition and support a model in which DNA geometry and bendability contribute to Mcm2–7 loading site selection in metazoans.
The loading of ring-shaped, replicative helicases onto DNA constitutes a key event during the initiation of chromosomal DNA replication. Across all domains of life, replicative helicases are recruited and deposited onto DNA at or near specialized chromosomal regions, termed origins, by the action of dedicated initiator and loader proteins. A critical task of initiators during this process is to recognize helicase loading sites and to remodel origins prior to helicase recruitment. The structural choreography that ultimately leads to helicase loading is only beginning to be understood in the different initiation systems.

Cellular replication initiators share a modular domain organization that optimally supports origin recognition and remodeling. A helix-turn-helix (HTH) fold serves as the primary origin DNA-binding and origin-targeting region in bacterial and archaeal initiators. This HTH domain is appended to an ATP hydrolyase associated with various cellular activities (AAAT+) fold, a domain prevalent in a large group of proteins involved in diverse biological processes beyond DNA replication. A common theme among members of this superfamily is that they use molecular motors or switches to remodel client molecules bound in the central pore of ring-shaped or helical AAA+ assemblies in response to ATP binding and hydrolysis. In initiators, the AAA+ regions also engage DNA, and these interactions directly remodel origin DNA in the prokaryotic proteins. However, the extent to which the initiator HTH and AAA+ modules contribute to DNA binding and remodeling across the eukaryotic domain is not fully understood. How origin DNA binding is sensed by initiators and coupled to the ATPase sites in the eukaryotic domain is also unclear.

In many eukaryotes, including fungi, plants, and metazoa, the initiator is a heterohexameric assembly, termed origin recognition complex (ORC), that acts in concert with the co-loaders Cdc6 and Cdt1 to deposit the hexameric Mcm2–7 helicase core onto DNA. ORC is composed of five evolutionary related AAA+ proteins (Orc1–5) and a sixth subunit (Orc6) of distinct origin. In Orc1–5, the AAA+ modules are augmented by a C-terminal HTH domain of the winged-helix (WH) type and co-assemble into a pentameric ring that can exist in both open (active) or closed (autoinhibited) configurations. Conversely, Orc6 is recruited to the periphery of the ORC ring through its interaction with Orc3,16,19. Structural studies have revealed that the initiator binds DNA in the center of the ORC ring, bending DNA to prepare origins for Mcm2–7 recruitment and loading. A recent cryo-EM structure of S. cerevisiae ORC has further detailed how the AAA+ and WH domains, in conjunction with several basic patch regions, recognize specific autonomous replication sequence (ARS) DNA elements at budding yeast origins. Notably, sequence-specific DNA binding as exemplified by S. cerevisiae ORC is a highly specialized mode of origin recognition outside prokaryotic lineages. Consensus sequences are not present at origins in more distantly related fungi or in multicellular eukaryotes; instead, contextual DNA sequence and chromatin cues appear to target ORC to preferred chromosomal regions in these systems. Yet, the physical basis for how ORC engages DNA in the absence of specific consensus DNA motifs remains poorly defined. Whether DNA remodeling is required for Mcm2–7 loading is likewise uncertain. Lastly, whether and how the recruitment of co-loaders such as Cdc6 (a AAA+ ATPase itself) elicits conformational changes in ORC and/or DNA prior to Mcm2–7 recruitment is also not understood.

To resolve these outstanding questions, we determined cryo-EM structures of DNA-bound Drosophila ORC (DmORC) assemblies in the presence and absence of the co-loader Cdc6. These structures, in conjunction with biochemical assays examining DNA binding and ATP hydrolysis activities of mutant DmORC assemblies, reveal unexpected differences to the S. cerevisiae ORC-DNA cryo-EM structure and indicate that the roles of the AAA+ and WH domains in DNA binding have diverged not only among the eukaryotic Orc1–5 subunits, but also between metazoan and related archaean initiators. We further reconstitute Mcm2–7 loading in vitro using recombinant Drosophila proteins and find that DNA binding facilitates Mcm2–7 loading. Together our results suggest mechanistic models for the coupling of DNA binding to the ATPase sites within ORC and ORC-Cdc6, and for the role of DNA sequence context both in targeting of metazoan ORC to origins and in Mcm2–7 loading.

Results

Structure of the DmORC-DNA-Cdc6 complex. Previous attempts to determine 3D structures of Drosophila ORC-DNA or ORC-DNA–Cdc6 by cryo-EM were unsuccessful because the particles adopt a strong preferred orientation on continuous carbon support EM grids and aggregate at the higher protein concentrations required for imaging without carbon layer. Through testing of various deletion constructs of ORC and Cdc6, we found that removal of the N-terminal regions of Orc1 (Orc1ΔN) and Cdc6 (Cdc6ΔN) improves solubility in the presence of DNA and monodispersity of the sample after freezing. These regions precede the AAA+ domains, are predicted to be mostly disordered, and have been reported to drive liquid-liquid phase separation. Deletion of Orc1 and Cdc6 N-termini did not substantially weaken DmORC’s ATP-dependent affinity for DNA, nor did it impair the ability of DmORC to hydrolyze ATP and of DmORC and DmCdc6 to co-associate in the presence of ATP and DNA (Supplementary Fig. 1). We also established an in vitro helicase loading assay that supported the ATP-dependent, salt-resistant association of Drosophila Mcm2–7 with DNA and Mcm2–7 double hexamer formation in the presence of full-length loading factors (Fig. 1). Importantly, DmOrc1, DmCdc6, but also DmCdt1 proteins lacking N-terminal, intrinsically disordered regions (IDRs) loaded Mcm2–7 onto DNA to similar extents as full-length proteins (Fig. 1a, b). Together, these findings indicate that the truncated DmORC and DmCdc6 constructs are functional and that the IDRs in Orc1, Cdc6, and Cdt1 are not essential for Mcm2–7 loading in vitro.

For structural studies, we reconstituted the DmORC-DNA–Cdc6 complex on a 60 bp AT-rich DNA duplex in the presence of ATP using N-terminally trimmed Orc1 and Cdc6 and purified the assembly by size exclusion chromatography (Fig. 2a, b). Cryo-EM analysis of the purified ternary complex revealed uniform particles in different orientations, with secondary structure elements clearly visible in 2D class averages (Supplementary Fig. 2a, b). The majority of complexes retained all components and allowed us to determine the 3D structure of the ternary DmORC-DNA–Cdc6 assembly at an overall resolution of 3.4 Å (Supplementary Figs. 2c, d, 3a, Supplementary Table 1). The 3D EM map showed sufficient detail to generate atomic models for protein components, with Cdc6 being built completely de novo (Supplementary Figs. 3c–f, Supplementary Table 1). The Orc2 N-terminus and the Orc6 TFB domain, both of which were disordered or flexible in the cryo-EM map, constitute notable exceptions. The DNA backbone is traceable for ~30 bp of the duplex; however, we were not able to unambiguously assign a DNA sequence to the density, most likely because DNA is bound to ORC in different registers (Supplementary Fig. 3g).

A near-planar DmORC-Cdc6 ring encircles and bends DNA. In the DNA-bound initiator-co-loader complex, DmORC and
DmCdc6 form a two-tiered hexameric ring, with each layer comprising the Orc1–5/Cdc6 AAA+ ATPase and WH domains, respectively (Fig. 2c, d, Supplementary Movie 1). The DNA passes through the center of the complex and is bent at the winged-helix surface toward the Orc3 insert (Fig. 2c–e). While the ATPase ring is closed, the WH tier contains a visible gap between Orc2 and Orc3 (Fig. 2e). In the *Drosophila* complex, both tiers are planar and oriented perpendicular to the DNA axis within ORC’s central channel (Fig. 2d). Comparison of the ternary complex structure with the previously determined crystal structure of apo-DmORC16 reveals that Orc1 has transitioned from the auto-inhibited conformation, in which the Orc1 AAA+ module is disengaged from its partner subunit Orc4 and laterally blocks ORC’s central channel, into an active conformation (Fig. 2f). The establishment of canonical AAA+/AAA+ interactions between Orc1 and Orc4, which are similar to those seen in the crystal structure of a subcomplex of human ORC17, is facilitated by two distinct conformational changes in Orc1: (a) rotation of the entire Orc1 AAA+ domain and (b) an ATP-induced reorientation of the helical lid subdomain with respect to the AAA+ -base (Fig. 2f, g). The latter rearrangement closes the Orc1/4 AAA+/AAA+ interface by enabling interactions between the Orc1 AAA+-lid and the Orc4 AAA+-base (increasing the buried solvent accessible area from 846 Å² to 1412 Å²), rationalizing why ATP binding to ORC stabilizes the active state (Fig. 2g). Besides Orc1, the WH domain of Orc2 has also repositioned in the ORC-DNA-Cdc6 complex and packs against the AAA+ module of Cdc6, propagating the domain-swap observed between AAA+ and WH domains of adjacent protomers around the ORC ring (Fig. 2e, f).

The co-loader Cdc6 is phylogenetically most closely related to ORC’s Orc1 subunit15 and adopts a similar 3D structure as Orc1 in the ternary initiator-co-loader complex (Supplementary Fig. 4a). In DmORC-DNA-Cdc6, the co-loader occupies the position between Orc1 and Orc2 and engages in canonical and non-canonical AAA+/AAA+ interactions with Orc1 and Orc2, respectively, to close the ORC ring (Supplementary Fig. 4c). The N-terminus of Orc3, part of which folds into a β-sheet of the Orc2 AAA+ core, is sandwiched in between the Orc2 and Cdc6 AAA+ domains, sealing the ATPase ring and forming a tripartite interface (Supplementary Fig. 4b, c). Two conserved Cdc6 arginines positioned in a helix that typically harbor the arginine finger in AAA+ ATPases—a residue involved in stimulating ATP hydrolysis of the adjacent protomer—contribute to the interaction network, indicating that this region in Cdc6 has been co-opted to mediate protein-protein interactions at the non-catalytic Orc2/Cdc6 site (Supplementary Fig. 4c, d).

**DNA recognition by DmORC-Cdc6.** A distinctive feature of the ternary DmORC-DNA-Cdc6 assembly is a large channel in the center of the complex that is occupied by DNA (Fig. 2e). Inspection of the cryo-EM density reveals that the DNA is positioned off-center within the DmORC-Cdc6 channel, arguing that the various subunits differentially contribute to DNA binding in the *Drosophila* complex (Fig. 2e, Supplementary Movie 1). In the AAA+ layer, the DNA closely abuts the ATPase domains of Orc1, Orc4, and Cdc6. This placement is stabilized by three different types of elements that directly engage the duplex (Fig. 3a).
Fig. 2 Structure of a licensing-competent DmORC·DNA·Cdc6 complex. a Domain architecture of DmORC subunits and Cdc6. Dashed lines demarcate regions that are flexible and structurally not resolved. Gray regions in Orc1 and Cdc6 were removed to improve sample behavior for cryo-EM. The color scheme is maintained throughout the figures unless noted otherwise. b Purification of DmORC·DNA·Cdc6 for cryo-EM. Chromatograms are shown for the ternary complex (solid lines) and for isolated DNA (dashed lines). Peak fractions retain all DmORC subunits and DmCdc6 when analyzed by SDS-PAGE (inset). The asterisk corresponds to a degradation product of Orc2 and/or Orc3. c Side view of the unsharpened cryo-EM map with AAA+- and WH domains colored differentially, highlighting the two-tiered organization of the complex. d Near-planar arrangement of the AAA+- and WH modules with respect to the main DNA axis (gray line) in the ORC·Cdc6 ring. The centers of mass of the AAA+- base and WH domains are depicted as spheres. e AAA+- and WH view of the cryo-EM density (unsharpened) with ORC and Cdc6 subunits colored differentially. The gap between Orc3 and Orc2 in the WH layer is indicated. f, g Conformational rearrangements in DmORC during the transition from the autoinhibited to the activated state. In f the positions of the Orc1 AAA+- and Orc2 WH domains in the autoinhibited apo-DmORC state (shown as cartoon; PDB 4xgc16) and the ORC·DNA·Cdc6 complex (ODC; depicted as molecular surface) are compared after structural alignment of both assemblies. Arrows indicate the domain movements during activation. Side views are shown, and other subunits/domains are displayed as light gray cartoon. g ATP binding to Orc1 reorients the Orc1 AAA+-lid and closes the Orc1/4 interface. The Orc1 AAA+-base subdomain of Orc1 in apo-DmORC (PDB 4xgc16) was superposed on the same region in the ternary complex. BAH, bromo-adjacent homology domain; IDR, intrinsically disordered region; BP, basic patch; WH, winged-helix; NTD, N-terminal domain; CTD, C-terminal domain; TFIIB, transcription factor IIb-like domain. Source data are provided as a Source data file.
The initiator-specific motifs (ISMs) of DmOrc1–5 and Cdc6, a helical insertion in the AAA+ fold characteristic for members of the initiator clade of AAA+ ATPases, form a circular collar around DNA. The ISMs of Orc1, Orc4, and Cdc6 form several hydrogen bonds and van der Waals contacts with the sugar-phosphate backbone of one of the DNA strands (Fig. 3b, c, Supplementary Fig. 5). Conversely, the ISMs of Orc2 and Orc3 make more limited contacts with DNA, while the ISM of Orc5 does not directly engage the DNA duplex (Supplementary Fig. 5). Nonetheless, these ISMs are still expected to contribute to DNA binding through the positive N-terminal dipoles of their parallel ISM helices and basic amino acid side chains, which point toward the central channel and add to its electropositive surface potential (Fig. 3d). Unexpectedly, the ISM collar is flanked on either side by additional DNA binding regions. On the outer surface of the AAA+ tier, a basic patch element in Orc1 (preceding the Orc1 AAA+ domain, Fig. 2a) folds into a short α-helix and binds the DNA backbone (Fig. 3a, b, Supplementary Fig. 5). Behind the ISM collar (when viewed from the Orc1 basic patch), Orc1 and Cdc6 each use an additional loop region (referred to as B-loop hereafter) to engage DNA (Fig. 3a–c, Supplementary Fig. 5). In Orc1, the side chain of B-loop arginine 692 inserts into the minor DNA groove. In Cdc6, this B-loop replaces the ISM collar. In Orc2 and Orc6, the side chains of B-loop arginines 97 and 141, respectively, insert into the minor groove. In Orc3 and Orc5, the side chains of these basic residues make limited contacts with DNA (Fig. 3c, Supplementary Fig. 5).
DNA (Supplementary Fig. 6), these binding assays report on measured′s channel, whereby the AAA contacts position the duplex in the central substrate-binding ORC together, these structures reveal that multivalent protein/DNA-dependent manner18,34. To understand how the different DNA-dipoles of the ISM helices are marked.

groove where it is stabilized by interactions with the deoxyribose and by the negative electrostatic potential of the backbone phosphate groups but makes no contacts with the nucleobases (Fig. 3b, Supplementary Figs. 3d, 5). Together with the ISMs and the Orc1 basic patch, these B-loops form an extended DNA binding site that promotes the off-center position of the DNA duplex in ORC·Cdc6′s channel.

In the ORC·Cdc6 WH tier, several subunits are also seen to bind DNA (Fig. 3e, Supplementary Fig. 5). Three of them—Orc2, Orc3, and Cdc6—engage the duplex using a β-hairpin element, a signature motif of winged-helix domains that often inserts into the minor groove as exemplified by the related archaeal initiator Orc1/Cdc6.6,7 However, in the ternary DmORC·DNA·Cdc6 complex, only the β-hairpin of Orc3 is poised to participate in minor groove interactions (Supplementary Fig. 3e). By contrast, the β-hairpins in Orc1 and Orc4 do not directly engage DNA, while Orc5 contacts the DNA backbone in a non-canonical manner using basic side chains in helix H2 of the WH fold (Fig. 3f, Supplementary Fig. 5).

Binding of the ATPase and WH domains in DmORC·Cdc6 to DNA leads to significant bending of the DNA duplex by approximately 46 degrees, with a bending point between the AAA+ and WH rings (Figs. 2d, 3f). The curved DNA segment is flanked on either side by the minor-groove-binding Orc1 B-loop and Orc3 β-hairpin, interactions that likely stabilize DNA binding. Contacts between the Orc5 WH-H2 helix and DNA may additionally contribute to duplex remodeling. Importantly, we observe a similar DNA configuration and DNA interactions in a ~3.9 Å resolution cryo-EM structure of a binary DmORC·DNA complex reconstituted in the absence of Cdc6, indicating that Cdc6 binding does not substantially remodel ORC·DNA contacts (Supplementary Fig. 6). The main difference between both structures pertains to the Orc2 WH domain, which is not visible in DmORC·DNA (Supplementary Fig. 6e). By contrast, more extensive remodeling is observed in the ORC ring during DNA binding to the initiator when comparing structures of DNA-free, active DmORC and DmORC·DNA (Supplementary Fig. 6). Taken together, these structures reveal that multivalent protein/DNA contacts position the duplex in the central substrate-binding ORC channel, whereby the AAA+ domains exert a firm grip on the DNA, and both AAA+ and WH domains cooperate to induce DNA bending (Fig. 3f).

DmORC binds DNA with low nanomolar affinity in an ATP-dependent manner18,34. To understand how the different DNA-binding sites in DmORC·DNA and DmORC·DNA·Cdc6 contribute to ORC′s affinity for DNA, we substituted single or multiple conserved residues in these regions with alanine and measured DmORC′s ability to associate with DNA using a fluorescence anisotropy-based binding assay. As Cdc6 binding to DmORC·DNA does not alter how the initiator engages DNA (Supplementary Fig. 6), these binding assays report on ORC′s DNA contacts in both the binary and ternary complexes. All mutant DmORC assemblies purified as a stable hexamer, indicating that the substitutions did not interfere with complex assembly (Supplementary Fig. 7a). As seen previously18, wild type DmORC bound a fluorescein-labeled 40 bp AT-rich DNA duplex with an apparent dissociation constant (Kd app) of ~16 nM in the presence of ATP (Fig. 3g). Alanine substitutions of conserved basic amino acids in the ISMs of Orc3 (K141) or Orc4 (R97) slightly reduced ORC′s affinity for DNA by approximately 3-fold (Fig. 3g). Surprisingly, a double mutation of Orc1-ISM residues S657 and Q660, which engage in van der Waals and hydrogen bond interactions with the DNA backbone, did not impede DmORCs ability to bind DNA. Unlike Orc3-K141 and Orc4-R97, these Orc1-ISM side chains cannot form salt bridges with the DNA phosphates, rationalizing the disparate findings (Fig. 3g, Supplementary Fig. 5). However, alanine substitutions in the Orc1 B-loop motif had a more striking effect. A triple alanine mutant (Orc1N691A/R692A/R693A) completely abolished ATP-dependent, high-affinity DNA binding by DmORC, which could not be restored by the addition of Cdc6 (Fig. 3g, Supplementary Fig. 5, Supplementary Fig. 7b). By comparison, the β-hairpin of Orc3 and the WH-helix 2 of Orc5 only moderately contribute to stabilizing DmORC on DNA, with triple mutations of basic residues (Orc3K710A/R711A/K712A or Orc5K341A/K344A/R345A) decreasing DmORC′s affinity for DNA by only 2–3 fold (Fig. 3g). These biochemical data are congruent with the multivalent protein/DNA interactions observed in both DmORC·DNA and DmORC·DNA·Cdc6 cryo-EM structures, and also uncover the B-loop in Orc1 as a DNA binding element that is essential for high-affinity interactions of DmORC with DNA in the absence and presence of Cdc6.

DNA-induced rearrangements in Orc1 regulate ATP hydrolysis. Several ORC subunits and Cdc6 are known to form bipartite ATP binding sites at AAA+/AAA+ interfaces in the ORC and ORC·Cdc6 rings16,17,20,21,35–37. In agreement with these studies, we observe clear densities for ATP bound to Orc1/4, Orc4/5, Orc5/3, and Cdc6/Orc1 in our ternary DmORC·DNA·Cdc6 complex structure (Supplementary Fig. 8). Of these sites, only those formed by Orc1 and Orc4, as well as by Cdc6 and Orc1, retain catalytic activity in yeast and metazoans17,34,35,37–39. Notably, ATP hydrolysis by ORC (i.e., at the Orc1/4 interface) is inhibited by DNA34,35, but how duplex binding is sensed by ORC and transmitted to the Orc1/4 ATPase site is unknown. To better understand how both events are coupled, we measured the ATP hydrolysis rates of wild type DmORC and our various DNA binding site mutants in the absence and presence of DNA (Fig. 4a). For wild type DmORC, we observe a 5-fold decrease in ATP turnover upon DNA addition, a reduction comparable to that reported for S. cerevisiae ORC35. Amino acid substitutions in the ISMs and WH domains of DmORC did not drastically alter
the basal ATPase rate and still supported DNA-mediated inhibition. Strikingly, the ATP turnover rate of DmORC carrying a triple alanine mutation in the Orc1 B-loop motif was reduced by ~3-fold and was not attenuated further by DNA, indicating the B-loop may be involved in regulating ATP hydrolysis at the Orc1/4 ATPase site.

Insights into a possible relay mechanism become apparent by comparing the Orc1 B-loop conformation in the ternary DmORC·DNA·Cdc6 complex with that in the DNA-free apo-DmORC crystal structure and our DNA-free ATP-DmORC cryo-EM structure. In the presence of DNA, the Orc1 B-loop adopts a more extended configuration, which leads to partial unwrapping of a short α-helix that connects the B-loop to the catalytic Walker B motif. Interestingly, the B-loop residues and the preceding α-helix are highly conserved in metazoan Orc1 proteins. An invariant aspartate (D687) in this helix is part of a bonding network that positions the sensor 1 (a residue required for ATP hydrolysis in AAA+ ATPases) and likely also the catalytic glutamate (a conserved amino acid that helps activate the lytic water for catalysis) in the Orc1/4 ATPase site. Importantly, an Orc1D687A mutation abolishes ATP hydrolysis by DmORC, supporting a critical role for the pre-B-loop helix in organizing this ATPase center.
conformation in a manner that impedes both DNA binding and ATP-hydrolysis (Figs. 3g, 4a). Although higher resolution structures will be required to resolve the exact rotamer conformations of residues involved in the allosteric coupling, these experiments nonetheless uncover a crucial role for the Orc1 B-loop not only in DNA binding, but also in regulating ATP hydrolysis by ORC. A comparable relay mechanism might also extend to Cdc6, where the B-loop likewise contacts DNA and connects to the Cdc6/Orc1 ATPase center via a similar bonding network as seen for Orc1 (Fig. 4e, f). Interactions between the Orc1 and Cdc6 B-loops and neighboring residues could additionally help regulate the temporal order of ATP hydrolysis at both ATPase sites (Fig. 4f, g).

Comparison of DNA binding in DmORC and ScORC. Among eukaryotic initiators, the sequence-specific recognition of DNA consensus motifs is a unique property of ORC in S. cerevisiae and a few related Saccharomycetes spp., while metazoan ORC assemblies bind DNA in a sequence-independent manner. Our structure of the ternary DmORC-DNA-Cdc6 complex afforded comparative analysis with a recent cryo-EM model of DNA-complexed S. cerevisiae ORC (ScORC) to unveil the physical basis for these divergent DNA binding modes. Common features of both initiators include DNA bending and the use of the ISM collar and the Orc3 WH domain (colored by subunit). Source data are provided as a Source data file.

DNA geometry and malleability modulate DmORC activities. Although DmORC engages DNA predominantly through DNA-sequence-independent backbone contacts in our structures, human and Drosophila ORC have been reported to prefer to bind AT-rich DNA, but the physical basis for these observations is unknown. Given the paucity of DNA base-contacts by DmORC, we hypothesized that the affinity differences for dissimilar DNA duplexes may result from sequence-dependent differences in DNA duplex geometry. Therefore, we systematically determined the dissociation constants for DmORC and DNA duplexes of varying GC-content and correlated the measured affinities with predicted DNA shape parameters. We found that DmORC binds more tightly to AT-rich sequences, with an overall 20-fold difference in apparent $K_d$ between homopolymeric poly(dA-dT) duplexes and duplexes of 100% GC-content, a range larger than observed previously for DmORC. 

<ref>ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.10538/s41467-020-18067-7</ref>
Orchestrating DNA remodeling through thermodynamics and electrostatics

**Fig. 5 Comparison of S. cerevisiae and Drosophila ORC–DNA contacts.** a The B-loop of Orc1 engages DNA in Drosophila but not in budding yeast ORC, while interactions between the Orc1-BP and DNA differ in both systems. DNA binding elements shared with S. cerevisiae ORC are not shown. b S. cerevisiae ORC uses numerous additional regions to associate with DNA in the cryo-EM structure of the DNA-bound budding yeast initiator (PDB 5zr120). A subset of these are yeast-specific insertions in ORC subunits (e.g., Orc2ISM-insert, Orc4IH), while others are flexible in DmORC·DNA and/or DmORC·DNA·Cdc6 (i.e., Orc2BP, Orc5BP, Orc175BP). Only duplex-binding regions unique to the S. cerevisiae structure are shown. Note also that the basic patch of S. cerevisiae Orc1 tracks in the DNA minor groove. c In Drosophila Orc1, the basic patch packs against the ISM of Orc4 and N-terminally folds into a helix that contacts the phosphodiester backbone but does not engage the minor groove. Unsharpened EM map density is shown. d Docking of the Orc1-basic patch reorients the second Orc4-ISM helix, priming it for DNA binding. The Orc4-ISMs from DmORC structures determined in this study (all containing the Orc1-basic patch) are superposed onto the Orc4-ISMs from the apo-DmORC and human Orc1/4/5 crystal structures (PDBs 4gxc16 and 5uj717; both determined in the absence of the Orc1-basic patch). The repositioning of the positive ISM helix 2 dipole is indicated. BP, basic patch; IH, insertion helix; ISM, initiator-specific motif.

Fig. 10a). Importantly, the apparent Kd values of DmORC binding to DNA did not only correlate with GC-content, but also with minor groove width and, more pronouncedly, with minor groove electrostatic potential, suggesting that thermodynamic duplex properties modulate ORC–DNA interactions (Fig. 6b). A similar relationship is observed when using endogenous Drosophila DNA sequences; known ORC binding sites at the well-studied Droso- phila chorion locus, such as ACE3, are characterized by increased negativity of the minor groove electrostatic potential (Fig. 6a, c)43. The robust nucleotide-dependence of DNA binding by DmORC in our assay conditions (>100-fold tighter DNA association in the presence of ATP)18 argues that the varying affinities for different DNA sequences are mediated by binding elements in the ORC core rather than secondary DNA binding sites. Concordantly, deletion of secondary DNA binding sites, which include the Orc6 TFIIB domain or the long Orc1 N-terminus23,32, enhances DmORC’s ability to discriminate between duplexes of high and low GC content (Fig. 6d, Supplementary Fig. 10a–c).

To understand the structural mechanisms for DmORC’s preferential binding to AT-rich DNA and its consequences for ORC-dependent DNA remodeling, we reconstituted two additional DmORC·DNA·Cdc6 complexes, using poly(dA·dT) dsDNA and a 70% GC duplex (referred to as GC-rich hereafter), respectively, and determined their structures by cryo-EM. The overall resolution of these structures was with 3.2 Å and 3.4 Å comparable to that of the DmORC·DNA·Cdc6 structure containing AT-rich dsDNA (GC content of 31.7%) (Supplementary Fig. 10d–f). Unexpectedly, we identified two different DNA conformations in our poly(dA·dT) dataset, which are characterized by different bending angles (~22° in state 1 vs. ~30° in state 2), and both are bent by DmORC to a lesser extent than the AT-rich duplex (46° bend) (Fig. 6c, f). The reduced bending of poly(dA·dT) DNA in the ternary complex probably results from an increased rigidity of the duplex that prevents the Orc3 WH hairpin to engage the minor groove (Fig. 6g). By contrast, the ORC subunits and Cdc6 adopt very similar configurations when bound to poly(dA-dT) or mixed nucleotide duplexes (Fig. 6e). One exception is the WH domain of Orc2, for which we observe weaker density when the DNA is bent to only 22° (state 1 of the poly(dA-dT)-containing complex), suggesting that a reduced curvature of DNA interferes with docking of this module onto the ORC-Cdc6 AAA+ ring (Supplementary Fig. 10d). The GC-rich DNA, on the other hand, assumes a conformation that globally resembled the AT-rich duplex (Fig. 6c, f). Nonetheless, closer inspection of the different DNA densities revealed variations in the position of the phosphodiester backbone,
especially within the AAA+ ring where the DNA densities are best resolved. While the phosphates of the different DNAs superimpose well near ISMs of Orc1 and Orc4, their position is more variable in other regions (Fig. 6h). Thus, the AAA+ ring is the principal ORC region that stabilizes the Drosophila initiator on DNA, with the ISMs of Orc1 and Orc4 serving as an anchor point to position the DNA duplex, while adjacent regions modulate ORC’s affinity for distinct DNA duplexes.

Our structure of DNA-bound Drosophila initiator assemblies, together with that of ScORC-DNA determined previously20, suggests that DNA bending is a conserved activity of eukaryotic initiators; yet, whether DNA remodeling is required for Mcm2–7
**Fig. 6 DNA sequence influences DNA binding and remodeling activities of DmORC.** a Increasing GC-content alleviates DmORC’s ATP-dependent affinity for DNA. The apparent dissociation constants (Kd, app) and standard errors (S.E.) of fit for ATP-dependent DmORC binding to different DNA duplexes is summarized in a bar graph (see Supplementary Fig. 10a for corresponding binding curves and number of replicates). The ARS1-like duplex used in Fig. 3g is shown for comparison. ACE3 and S18 are endogenous DNA sequences of the Drosophila chorion locus. b Predicted minor groove width and electrostatic potential of DNA duplexes used in (a) are plotted for each base pair to highlight the correlation with measured Kd, app for DmORC binding. c Minor groove electrostatic potential across the Drosophila chorion locus. The ORC-binding ACE3 and ori-β regions are highlighted by black bars. The positions of the endogenous DNA sequences used in (a) are indicated by triangles. d Deletion of secondary DNA binding sites in DmORC enhances DmORC’s preference for associating with AT-rich duplexes. The ratios of Kd, app-100% GC over Kd, app-poly(dA·dT) are plotted for wild type, full-length DmORC, and complexes lacking respective secondary DNA binding regions. See Supplementary Fig. 10a–c for binding curves. e Superposition of cryo-EM maps (unsharpened) of DmORC and DmCdc6 (all colored in gray tones) complexed with different DNA duplexes. The structure of the AT-rich (31.7% GC) assembly corresponds to the one shown in previous figures. f DNA axes of structures in e illustrate different degrees of DNA bending by DmORC. g Diminished DNA bending precludes minor groove engagement by the Orc3 WH β-hairpin. h Variability of DNA backbone positions within the AAA+ ring. EM densities (locally sharpened with LocScale) for DNA duplexes in the different DmORC·DNA·Cdc6 structures are shown (except for state 1 of the poly(dA·dT) complex). i DNA sequence modulates Mcm2-7 loading efficiency. DmMcm2-7 is recruited (low-salt wash) similarly to AT-rich and poly(dA·dT) duplexes but loaded (high-salt wash) more efficiently onto AT-rich DNA compared to poly(dA·dT) duplexes. Note that Cdt1 and Cdc6 nonspecifically associate with beads. Source data are provided as a Source data file.

loading is unknown. To address this question, we examined whether a poly(dA·dT) duplex, which is bound by ORC with high affinity but bent to a lesser degree as the mixed AT-rich duplex (Fig. 6a, e, f), is a substrate for Mcm2–7 loading. Strikingly, using our in vitro reconstitution system for *Drosophila* Mcm2–7 loading we find that Mcm2–7 is loaded onto poly(dA·dT) duplexes much less efficiently than onto the AT-rich substrate (Fig. 6i, high-salt wash). By contrast, Mcm2–7 is recruited to both DNA substrates to similar extents in low-salt conditions (Fig. 6i). Collectively, these biochemical and structural findings argue that sequence-dependent thermodynamic properties of DNA, including duplex geometry and deformability, tune ORC’s ability to productively engage nucleic acid and to load Mcm2–7 onto DNA.

**Discussion**

The various cryo-EM structures of DNA- and Cdc6-bound DmORC presented here, in combination with the previous crystal structure of the DNA-free apo-complex 

[Further discussion and analysis related to the image and text content.]

in ORC’s central channel. Undocking of Orc1 from Orc4 after ATP hydrolysis and/or transitioning into an autoinhibited state could also provide a means to release ORC from DNA or to disengage the complex from Mcm2–7 after loading.

Activation of ORC and DNA binding to the initiator allows the association of the co-loader Cdc6 (Fig. 7a, vi). Surprisingly, Cdc6 binding does not remodel ORC–DNA contacts or alter DmORC’s ATPase activity, nor does it induce major conformational changes in the ORC ring apart from providing an interaction surface for the Orc2 WH domain, which is flexible in Cdc6’s absence. These findings indicate that Cdc6 predominantly has a structural role during Mcm2–7 loading. In agreement with this model, Cdc6 and the Orc2 WH module form a combined binding platform for the Mcm3 WH domain in the OCCM complex, and biochemical studies have found that this Mcm3 region is essential for DNA binding and may, therefore, provide a means to regulate Mcm2–7 recruitment and loading.

Structural comparison of DNA-bound *S. cerevisiae* and DmORC assemblies reveals that both initiators have evolved a subset of specialized DNA binding elements that reflect the distinct modes of origin recognition in these systems. In budding yeast ORC, DNA interactions facilitate recognition of the ARS consensus sequence at replication origins. By contrast, DmORC and probably all other metazoan initiators are optimized to bind DNA in a sequence-independent manner through DNA backbone interactions with regions of the positively charged ORC central channel. Nonetheless, these contacts, likely in conjunction with the DNA minor groove engagement by a conserved Orc1 B-loop residue (either arginine, lysine or histidine in metazoan Orc1 orthologs), also provide a means for recognition of DNA-sequence-dependent variations in DNA duplex geometry. These associations can account for *Drosophila* and human ORC’s preference for binding AT-rich or poly(dA·dT) DNA segments in vitro and are also consistent with an enrichment of these sequence features at chromosomal ORC binding sites and near replication start sites in vivo. Thus, AT-rich sequences and poly(dA·dT) stretches at metazoan replication origins may not only promote initiation by maintaining nucleosome-free regions, but also by stabilizing metazoan ORC on DNA through high-affinity interactions in ORC’s central channel. These mechanisms are expected to work in concert with chromatin cues (e.g. interactions of the Orc1-BAH domain with H4K20me2 on nucleosomes and other chromatin-associated ORC partner proteins) and secondary DNA binding sites.
(including the Orc6-TFIIB module and the phase separating IDR in Orc1) to recruit ORC to origins on metazoan chromosomes\textsuperscript{22,23,25,27,32,48,49,51–56}. The extent to which these distinct strategies are used in combination at individual origins or may change during development is an important question for future studies.

Both \textit{S. cerevisiae} and \textit{Dm} ORC bend DNA, which has been proposed to help align the DNA duplex with the gate in the Mcm2–7 ring during helicase loading (Fig. 7a, vii)\textsuperscript{18,20}. Unexpectedly, we observed varying degrees of \textit{Dm} ORC-induced DNA bending of duplexes containing different DNA sequences, demonstrating that DNA binding and bending can be uncoupled. Although ORC’s inability to bend certain DNA sequences, in particular poly(dA·dT) did not prevent high-affinity DNA binding or co-association of Cdc6, it inhibited Mcm2–7 loading, likely by impeding DNA insertion into the helicase ring (Fig. 7a). Given the importance of DNA bending for efficient Mcm2–7 loading, it is somewhat surprising that DNA bending in our DNA-complexed \textit{Dm} ORC structures is not as extensive as in the \textit{Sc} ORC·DNA assembly, in which the bent duplex is stabilized by yeast-specific ORC/DNA contacts mediated by Orc6 and basic patches in Orc2 and Orc5\textsuperscript{20}. However, it is possible that similar
interactions are more transient in the Drosophila initiator-DNA complexes, or that Mcm2–7 recruitment is necessary to promote more extensive DNA remodeling prior to helicase loading. Collectively, our findings suggest a model in which the location of metazoan Mcm2–7 loading sites, and thereby licensed origins, is in part modulated by DNA-sequence induced variations in DNA geometry and malleability. Binding of metazoan ORC to DNA regions that are more resistant to deformation (such as polyA tracts) may require the initiator to passively slide on DNA until it encounters a malleable DNA segment30,50,57. Alternatively, metazoan ORC may recruit specific protein cofactors that help productively bend DNA to increase helicase loading efficiency. In this regard, reconstituting metazoan Mcm2–7 loading in vitro, which we have done here with recombinant Drosophila proteins, will provide a useful system to investigate these different possibilities in the future.

Binding of substrates in the central pore of oligomeric AAA+ ATPases often regulates the enzymatic activity of these assemblies11,58. In ORC, DNA binding inhibits ATP hydrolysis at the Orc1/4 site, the principal ATPase center, in yeast and the Drosophila complex34,35 (Fig. 4a). Our structural and biochemical studies identify the B-loop region in Drosophila Orc1 to be involved in sensing and communicating the DNA binding status in ORC to the Orc1/4 ATPase site. Although in the DNA-bound S. cerevisiae initiator (unlike in DmORC-DNA assemblies), the Orc1 B-loop element does not directly contact the DNA duplex, conformational changes in this region may be achieved indirectly upon DNA engagement of yeast-specific DNA binding element20; high-resolution structural information of DNA-free S. cerevisiae ORC will be necessary to test this hypothesis. Notably, the corresponding regions in Drosophila and budding yeast Cdc6 (in the OCCM complex21) directly interact with DNA, suggesting that Orc1 and Cdc6 may use similar strategies to regulate ATP hydrolysis at the Orc1/4 and Cdc6/Orc1 sites, respectively. Inspection of structures of related initiator and loader AAA+ ATPases further revealed that the B-loop region is universally located at the nucleic-acid binding face of these proteins and, in addition to the ISMs, often directly contacts DNA (Fig. 7b). In the bacterial helicase loader DnaC, the B-loop also undergoes a conformational change upon DNA association and is required for ssDNA binding and ATP hydrolysis59. This structural conguency suggests that nucleic-acid engagement by the B-loop element may constitute a more general strategy of AAA+ ATPases in the initiator and helicase loader clade to stabilize these proteins on DNA. Future studies will help determine how these interactions are coupled to the ATPase cycle in the different systems to support the diverse biological functions of these molecular switches.

Methods

Protein constructs and baculovirus generation. The various Drosophila ORC (DmORC) assemblies used in this study were reconstituted in insect cells by co-expression of baculoviruses expressing either individual or multiple ORC subunits. To generate ORC-containing vectors for baculovirus production, full-length or N-terminally truncated ORC subunits were first cloned into a ligation-independent cloning (LIC)-compatible pFastBac vector (Macrolab, University of California Berkeley, USA). For purification, a hex-histidine (6xHis) tag and a maltose binding protein (MBP) tag, both followed by a tobacco etch virus (TEV) protease cleavage site, were added to the N-termini of Orc1 and Orc4, respectively. These constructs served as templates for site-directed mutagenesis to introduce point mutations into Orc1, Orc3, Orc4, and Orc5. All individual wild type and mutant ORC subunit constructs were verified by DNA sequencing. Subsequently, several wild type and mutant ORC subunits were combined into a pFastBac-derived BioBrics MultiSac expression vector (Macrolab, University of California Berkeley, USA) by subcloning different combinations of DmORC genes. Bamcids of single and MultiSac expression constructs were generated in DH10Bac cells and used for transfections of S9 cells with Colefactin II (Thermo Scientific Fischer). Baculoviruses were amplified for two rounds in S9 cells to obtain high-titer viruses for infection of Hi5 cells for large-scale DmORC expression.

Drosophila Cdc6 was also expressed in insect cells. The coding sequences of full-length or N-terminally truncated DmCdc6 (2421 amino acids, 2422-662) were cloned into the LIC-compatible pFastBac vector (Macrolab, University of California Berkeley, USA). The full-length DmCdc6 construct contained an N-terminal MBP tag followed by a TEV cleavage site, while that of DmCdc6AN214 was fused to an N-terminal 6xHis-MBP tag and a TEV cleavage site. Bacmid generation and baculovirus amplification were performed as described for DmORC.

Drosophila Cdt1 and Mcm7 were expressed in E. coli and Hi5 insect cells, respectively. Full-length or N-terminally trimmed (amino acids 298-743) DmCdt1 were cloned as N-terminal 6xHis-TEV fusions into a LIC-compatible pFastBac vector (Macrolab, University of California Berkeley, USA), while Mcm2–7 subunits were cloned into two pFastBac-derived BioBrics MultiSac expression vectors (Macrolab, University of California Berkeley, USA), one construct containing Mcm2, Mcm4, and Mcm6, and the other Mcm3, Mcm5, and Mcm7. For affinity purification, Mcm4 and Mcm7 were tagged N-terminally with 6xHis and MBP, respectively, each followed by a TEV protease cleavage site.

Expression and purification of recombinant DmORC and DmCdc6. The different DmORC assemblies were purified at 4 °C from 4 liters Hi5 cells infected with combinations of baculoviruses expressing single or multiple DmORC subunits as previously16,19 with minor modifications. 48 h post-infection, Hi5 cells were harvested, resuspended in ~130 mL lysing buffer (50 mM Tris-HCl, pH 7.8, 300 mM KCl, 50 mM imidazole, pH 7.8, 10% glycerol, 1 mM BME, and lysed by sonication. The crude cell lysate was clarified by two ultracentrifugation runs in a Beckman Coulter Optima L-80 XP ultracentrifuge, each at 42,414 × g for 45 min in a 45Ti rotor. After the first centrifugation step, the supernatant was added to the lysate supernatant and dialyzed against 20% (v/v) supernatant of the second ultracentrifugation run was applied to a 5 mL HisTrap HP nickel-affinity chromatography column (GE Healthcare). After extensive washing with lysis buffer, bound DmORC was eluted with a 50–250 mM imidazole gradient in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 1 mM BME and further purified using a 10 mL amylase column (New England Biolabs). 6xHis and MBP tags were cleaved by overnight incubation with 6xHis-tagged TEV protease, which was removed by subsequent nickel-affinity chromatography using a 5 mL HisTrap HP column (GE Healthcare). After extensive washing with lysis buffer, bound DmCdc6 was eluted with a 50–250 mM imidazole gradient in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 1 mM BME and further purified using a 10 mL amylose column (New England Biolabs). Proteins were then aliquoted, flash frozen in liquid nitrogen, and stored at −80 °C.

Full-length MBP-tagged DmCdc6 and 6xHis-MBP-tagged DmCdc6AN214 were expressed in Hi5 cells for 48 h. For purification of full-length DmCdc6, cells were resuspended in 100 mL buffer containing 50 mM Tris-HCl, pH 7.8, 600 mM KCl, 50 mM imidazole, 1 μg/mL leupeptin, 1 μM aprotinin, and 10% glycerol, purified by nickel-af

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-18067-7 © 2020 Macmillan Media Limited, part of Springer Nature. All rights reserved.
Expression and purification of DmCdt1. *Drosophila* Cdt1 was expressed in BL21 RIL E. coli (cultured in Terrific Broth (TB)) upon induction with 0.5 mM Isopropyl β-D-thiogalactopyranoside (IPTG) overnight in 1 L M9 medium supplemented with 50 μg/mL ampicillin. Cells were harvested, resuspended in 120 mL lysis buffer (50 mM Tris-HCl, pH 7.8, 1 M NaCl, 10% glycerol, 50 mM imidazole, pH 7.8, 5 mM BME, 200μM PMSF, 1 μg/mL leupeptin) and lysed by sonication. The lysate was clarified by 30 min centrifugation at 30,996 x g in a Beckman Coulter Avanti J-26 XP centrifuge using a JA-17 rotor. Solubilized Cdt1 in the supernatant was bound to a 5 mL HisTrap HP nickel-affinity chromatography column (GE Healthcare) and, after washing with 50 mL lysis buffer and 50 mL low salt buffer (50 mM Tris-HCl pH 7.8, 150 mM KCl, 10% glycerol, 50 mM imidazole, pH 7.8, 5 mM BME), eluted onto a 5 mL HiTrap Q HP anion exchange column (GE Healthcare) with 250 mM imidazole in 50 mM Tris-HCl, pH 7.8, 150 mM KCl, 10% glycerol, 5 mM BME. The flow-through was pooled and digested with 6xHis-tagged TEV protease while being dialyzed into 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 5 mM BME overnight. TEV protease and uncleaved Cdt1 were removed nickel-affinity chromatography. Cleaved Cdt1 was concentrated in a 30K Amicon Ultra-15 concentrator (Millipore) and further purified by gel filtration on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 1 mM DTT. Peak fractions containing DmCdt1 were pooled, concentrated, flash frozen, and stored at −80 °C. All purification steps were performed at 4 °C.

Expression and purification of DmMcm2−7. Four liters of Hi5 cells were co-infected with high-titer baculoviruses encoding for *Drosophila* Mcm2−7 subunits for two days. Cells were harvested by centrifugation and resuspended in 130 mL lysis buffer (50 mM HEPES-KOH, pH 7.5, 200 mM potassium acetate, 10% glycerol, 50 mM imidazole, pH 7.5, 1 mM BME, 200 μM PMFSF). After sonication, the lysis was clarified by two rounds of ultracentrifugation and ammonium sulfate precipitation as described for *DmORC*. Mcm2−7 were isolated using a 5 mL HisTrap HP nickel-affinity chromatography column (GE Healthcare) and eluted using a 50–250 mM imidazole gradient in 50 mM HEPES-KOH, pH 7.5, 200 mM potassium acetate, 10% glycerol, 1 mM BME. Eluted proteins were loaded onto a MonoQ HR 5/5 column (GE Healthcare) washed in 50 mM HEPES-KOH, pH 7.5, 200 mM potassium glutamate, 10% glycerol, 1 mM DTT. DmMcm2−7 peak fractions were pooled, concentrated in a 30K Amicon Ultra-15 concentrator (Millipore), flash frozen and stored at −80 °C. All purification steps were performed at 4 °C. The presence of all six DmMcm2−7 subunits was verified SDS-PAGE and mass spectrometry, and the integrity of the hexameric complex was validated by negative-stain electron microscopy (4 μL of 80 mM Mcm2−7 were applied to continuous carbon EM grids and stained with 2% uranyl acetate), showing the expected assembly of subunits into open, hexameric rings (Fig. 1c).

Assembly and purification of complexes for cryo-EM. Attempts to reconstitute *Drosophila* ORC-DNA-Cdc6 and ORC-DNA complexes with full-length proteins at the same stoichiometry as necessary for cryo-EM studies using open holobodies of cryo-EM grids resulted in visible sample opaqueness when mixing ORC and Cdc6 with DNA, and substantial protein aggregation was visible on cryo-EM grids. Removal of the N-terminal 439 amino acid residues of DmORC1 (referred to as ORC1ΔN) and the N-terminal 241 amino acid residues of DmCdc6 (ΔCdc6), which have been reported to drive liquid-liquid phase separation, improved the behavior of these proteins in the presence of DNA without impeding DmORC’s ATPase activity, the ATP-dependent co-association of ORC, DNA, and Cdc6, and Mcm2−7 loading (Fig. 1, Supplementary Figure 1). Therefore, these trimmed but Mcm2−7 loading-competent constructs were used to assess DmORC-DNA and DmORC-DNA-Cdc6 complexes for cryo-EM experiments.

To reconstitute *Drosophila* ORC-DNA-Cdc6 assemblies, *DmORC*ΔN,D, *DmDNA*, and *DmCdc6*ΔN,D, which have been reported to drive liquid-liquid phase separation, were expressed in *Drosophila* Kc cells by using the 25 μM HEPES-KOH, pH 7.6, 250 mM potassium acetate, 10 μM magnesium acetate, 1 mM DTT, and 1 mM ATP. Assembly reactions were incubated for 5 min at 25 °C after mixing ORC with DNA, and again after adding Cdc6. Subsequently, 300 μL of the reconstitution reaction were loaded onto a Superose 6 Increase 10/300 GL column (GE Healthcare) equilibrated in 25 mM HEPES-KOH, pH 7.6, 250 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, and 0.2 mM ATP to purify the ternary DmORC-DNA-Cdc6 complex by gel filtration chromatography. Peak fractions containing DNA (judged by absorbance at 254 nm), *DmORC*ΔN,D, and *DmCdc6*ΔN,D were pooled and concentrated in a 30K Amicon Ultra-4 concentrator (Millipore) and then applied to a 0.7 μm cryo-EM grid (200 × 200 mm Cytodex3) and stained with 10 μL of a 2% uranyl acetate plate of a target defocus of 0.4–0.6 μm (using a phase thickness of 20–30 degrees) and 1–2.2 μm (without phase plate), respectively. For each dataset, 40 (with Volta phase plate) or 50 (without Volta phase plate) frames were recorded per movie at a dose rate of 5–7 e−/Å2 per second, yielding a total electron dose of 40 e−/Å2 and 50 e−/Å2, respectively (see also Supplementary Table 1 for additional information on data collection settings). Data of the DmORC-DNA-Cdc6 complex was assembled with poly(DA DTT) DNA was collected as described above with the exception that a Falcon 3EC direct electron detector. Dose-fractionated images were collected in an automated manner with EPU (Thermo Fisher Scientific) in EFTEM mode (slit width 20 eV) with or without Volta phase plate (two rounds of 0.4–0.6 μm (using a phase thickness of 20–30 degrees) and 1–2.2 μm (without phase plate), respectively. Therefore, these trimmed but Mcm2−7 loading-competent constructs were used to assess DmORC-DNA-Cdc6 assemblies, low-pass filtered to 30 Å, and served as templates for reference-based particle selection with GATOUMATCH for all datasets. Further EM image processing steps, including particle extraction, 2D and 3D classification, 3D refinement, particle polishing, and postprocessing were performed using RELION 3.0 and RELION 2.1 and RELION 3.0-25. Picked particles were extracted using a box size of 300 × 300 pixels, normalized, and subjected to 2D and 3D classification. To remove damaged or falsely picked particles, we performed one to two rounds of 3D classification of the data into four to eight classes, which yielded better results than 2D classification in preserving particles corresponding to underdispersed views of the complex. A low-resolution 3D reconstruction (low-pass filtered to 50 Å) from negatively stained particles of the ternary complex served as initial reference for 3D classification of the data. Particles of 3D class volumes with clearly visible secondary structure elements (cleaned particles) were subjected to 3D auto refinement (with a soft-edged, global mask low-pass filtered to 15 Å) and particle polishing. B-factors were automatically estimated in RELION 3.0 and used for map sharpening as implemented in the RELION postprocessing procedure. Global resolution of entire datasets and local resolutions were estimated using gold standard Fourier shell correlations between half-maps in RELION 3.0.

Model building and refinement. Model building was initiated by individually docking the AAA+ and winged-helix domains of the Orc1−5 subunits of the DmORC crystal structure (PDB 4e7x) into the cryo-EM density map of DmORC-DNA-Cdc6 (with 60 bp AT rich dsDNA, dataset 1) using UCSF Chimera. The DmORC model was then manually rebuilt in COOT to accommodate for structural changes in DmORC between the apo-ORC state and the ternary complex structure, and to build additional segments that were poorly or not resolved in the previously unsharpened, globally sharpened, and locally sharpened (with Locs) cryo-EM density maps were used. DmCdc6 was built de novo in COOT into the sharpened EM density, which was of sufficient resolution to almost completely trace the main chain and to assign the amino acid register to the density. The close evolutionary kinship between DmCdc6 and Orc1 provided additional information on the topological fold of DmCdc6 to confirm our assignment. For initial DNA placement, short segments of idealized B-form DNA were docked into the cryo-EM map using UCSF Chimera and locally rebuilt in COOT. The resulting DmORC-DNA-Cdc6 model was improved by iterative rounds of real-space refinement against the sharpened cryo-EM map in PHENIX or Coot. For mCdc6, we used annotations for both protein and DNA, as well Ramachandran and rotamer restraints for protein chains, and subsequent rebuilding in COOT. In addition, a morphing step was included during the first refinement round in Phenix. The final model has excellent geometry ( MolProbity score 1.52; Supplementary Table 1) and includes the AAA+ and winged-helix domains of Orc1−5 and Cdc6, the basic patch region in Orc1, the C-terminal helix in Orc6, and 34 base pair duplex DNA bound by DmCdc6.
Models for all other DmORC-DNA-Cdc6 assemblies, for the binary DmORC-DNA complex, and for DNA-free DmORC were built using the ternary complex with the 60 bp AT-rich duplex as a starting model. Briefly, the pdb model was docked into the respective cryo-EM density in UCSF Chimera, locally rebuilt in COOT if necessary, and real-space-refined in PHENIX using a combination of rigid body, XYZ, and ADP refinement. For cryo-EM maps of a resolution worse than 3.5 Å, we refrained from extensive rebuilding the corresponding coordinate models in COOT. We note that the assignment of DNA sequences in our models is consistent with the higher model and map B-factors in this region of the complex. Moreover, although full-length DmORC was used to reconstitute the binary and ternary complexes, no density was visible for the Orc6 TFIB domain in 2D class averages or any of the 3D reconstructions, indicating flexibility of this region.

Structure analysis. Structural alignments and superpositions were performed with PyMOL (The PyMOL Molecular Graphics System, Version 1.8.2.0 Schrödinger, LLC). Buried solvent accessible surface areas were determined using the PDB Epstein server22 implemented in COOT. Protein–DNA contacts were explored using PDBePISA and DnaproDB23. Multiple protein sequence alignments were performed with MAFFT24,25 and visualized in JALVIEW26 by or generating sequence logos with WEBLOGO27. DNA shape parameters were determined using CLDVD.S. All figures were rendered using PyMOL, UCSF Chimera, and UCSF ChimeraX88.

Annealing of DNA duplexes. DNA duplexes used for reconstituting DNA-bound DmORC assemblies for structural studies and for biochemical assays were obtained by annealing single-stranded, complementary DNA oligonucleotides at 10 or 50 μM, respectively, each in 10 mM Tris- HCl, pH 8 and 5 mM MgCl2. For annealing, oligonucleotides were heated to 95 °C for 5 min in a 1 L water bath and slowly cooled to room temperature overnight. Successful annealing was verified by separating 0.25 pmol dsDNA on a 12% native-PAGE gel in 1× Tris-borate buffer pH 8.3 and subsequent staining with SYBR Safe (Thermo Fisher Scientific). DNA duplexes used throughout this study are summarized in Supplementary Table 2.

Fluorescence anisotropy DNA binding assays. Binding of full-length wild type or mutant DmORC, or of truncated DmORC
to DNA was measured by fluorescence anisotropy using 40 bp dsDNAs (Supplementary Table 2) with one strand 5′-labeled with fluoroscein similarly as done previously16. 1 μM dsDNA was mixed with increasing concentrations of DmORC (from 0.04 nM to 2.5 μM) in 25 mM HEPES-KOH, pH 7.6, 300 mM potassium glutamate, 10% glycerol, 10 mM magnesium acetate, 0.1% NP40, 1 mM DTT, 0.2 mg/mL BSA, and 0 or 1 mM ATP. After 40 min incubation at 25 °C, fluorescence anisotropy was measured using excitation and emission filters of 485 nm and 520 nm, respectively, in a PHERASer Mx PS90 plate reader (BMG Labtech). To obtain apparent dissociation constant (Kapp) for DmORC binding to the respective DNA duplexes, the means and standard deviation (SD) of data points of n independent experiments (as listed in each figure) were plotted and fit to the Hill equation using the GraphPad Prism software:

\[ A = A_1 + (A_0 - A_1) \times \left[ \frac{[R]}{K_{\text{app}} + [R]} \right]^{K_{\text{app}}} \]

(1)

where A0 and A1 are the measured anisotropy of free and bound, fluorescently labeled DNA duplex, respectively, [R] is the concentration of DmORC, Kapp is the apparent binding constant, and n is the Hill coefficient. For the (poly(dA-dT)) duplex, we included additionally binding data to the quadratic binding equation to account for potential DmORC depletion at low concentrations due to its high affinity for this DNA:

\[ A = A_1 + (A_0 - A_1) \times \left[ \frac{L}{K_{\text{app}} + [R]} \right] + \sqrt{\left[ \frac{L}{K_{\text{app}} + [R]} \right]^2 - 4 \times \left[ \frac{L}{2} \right]} \times \left[ \frac{L}{2} \right] \]

(2)

where A0 and A1 are the measured anisotropy of free and bound, fluorescently labeled poly(dA-dT) duplex, respectively, [L] is the concentration of poly(dA-dT) used in binding assays, [R] is the concentration of DmORC, and Kapp is the apparent binding constant. Kapp values obtained for both methods were with 2.1 mM (HEPES-KOH) or 2.4 mM (quadratic equation) comparable. For simplicity, only Hill curves are shown in Supplementary Fig. 10a-c. For bar graphs, Kapp and standard error (SE) of fits are plotted. DNA binding of DmORC protein in the presence of DmcCd6 was measured as described above with the exception that wild type or truncated DmcCd6 was added at equimolar concentrations to DmORC in the reaction mix.

Pull-down assays. To verify that N-terminal truncations of DmORC1 and DmCd6 do not interfere with ternary complex formation, we performed pull-down assays using MBP-tagged DmCd6c or 6xHis-MBP-tagged DmcCd6ΔN as baits. 400 μL reactions were assembled with 60 bp AT-rich dsDNA (Supplementary Table 2) and different combinations of full-length DmORC or DmORCΔCdc6, 6xHis-MBP-tagged DmCd6c, and full-length MBP-6xHis-MBP-tagged N-terminal 100 nM, 200 nM, respectively, in a PHERAstar FSX plate reader (BMG Labtech). To obtain NATURE COMMUNICATIONS | (2020) 11:4263 | https://doi.org/10.1038/s41467-020-18067-7 | www.nature.com/naturecommunications
2. O. Iyer, L. M., Leipe, D. D., Koonin, E. V. & Aravind, L. Evolutionary history and 9. Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. Structure and function of the 11. Riera, A. et al. From structure to mechanism-understanding initiation of DNA 13. Chuang, R. Y. & Kelly, T. J. The fission yeast homologue of Orc6p binds to 16. On, K. F., Jaremko, M., Stillman, B. & Joshua-Tor, L. A structural view of the
50. Segal, E. & Widom, J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* 19, 65–71 (2009).

51. Eaton, M. L. et al. Chromatin signatures of the *Drosophila* replication program. *Genome Res.* 21, 164–174 (2011).

52. Cayrou, C. et al. Genome-scale analysis of nucleosome replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Res.* 21, 1438–1449 (2011).

53. Aladjem, M. I. & Fanning, E. The replicon revisited: an old model learns new tricks in metazoan chromosomes. *EMBO Rep.* 5, 686–691 (2004).

54. Shen, Z. et al. A WD-repeat protein stabilizes ORC binding to chromatin. *Mol. Cell* 40, 99–111 (2010).

55. Bartke, T. et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143, 470–484 (2010).

56. Vermeulen, M. et al. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967–980 (2010).

57. Dudzhevich, D. et al. The dynamics of eukaryotic replication initiation: origin specificity, licensing, and firing at the single-molecule level. *Mol. Cell* 58, 483–494 (2015).

58. Miller, J. M. & Enemark, E. J. Fundamental characteristics of AAA+ protein family structure and function. *Archaea* 2016, 9293407 (2016).

59. Arias-Palomino, E., Puri, N., O’Shea Murray, V. L., Yan, Q. & Berger, J. M. Physical basis for the loading of a bacterial replicative helicase onto DNA. *Mol. Cell* 74, 173–184 e4 (2019).

60. Zheng, S. Q. et al. MotonCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332 (2017).

61. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* 193, 1–12 (2016).

62. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-electron microscopy with parallelisation using GPUs in RELION-2. *eLife* 5, e18722 (2016).

63. Zhang, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* 6, 5–17 (2019).

64. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* 157, 281–287 (2007).

65. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612 (2004).

66. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 486–501 (2010).

67. Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *eLife* 6, e27131 (2017).

68. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 213–221 (2010).

69. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D. Struct. Biol.* 74, 531–544 (2018).

70. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 12–21 (2010).

71. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797 (2007).

72. Sagedorf, J. M., Berman, H. M. & Ruh, R. DNAProDR: an interactive tool for structural analysis of DNA-protein complexes. *Nucleic Acids Res.* 45, W89–W97 (2017).

73. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33, 518–521 (2005).

74. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* 9, 286–298 (2008).

75. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191 (2009).

76. Crooks, G. E., Hon, G., Chandria, I. M. & Brenner, S. E. WebLogos: a sequence logo generator. *Genome Res.* 14, 1188–1190 (2004).

77. Lavery, R., Moakher, M., Maddocks, J. H., Petkeviciute, D. & Zakrzewska, K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* 37, 5917–5929 (2009).

78. Blanchet, C., Pasi, M., Zakrzewska, K. & Lavery, R. CURVES+: web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res.* 39, W68–W73 (2011).

79. Colasanit, A. V., Lu, X. J. & Olson, W. K. Analyzing and building nucleic acid structures with 3DNA. *J. Vis. Exp.* e4401 (2013).

80. Lu, X. J. & Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 31, 5108–5121 (2003).

81. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* 27, 14–25 (2018).

82. Kianatka, K., Solinger, J. A. & Heyer, W. D. NADH-coupled microplate photometric assay for kinetic studies of ATP-lyase enzymes with low and high specific activities. *Anal. Biochem.* 321, 266–271 (2003).

83. Coster, G., Frigola, J., Beuron, F., Morris, E. P. & Difley, J. F. Origin licensing requires ATP binding and hydrolysis by the MCM replicative helicase. *Mol. Cell* 55, 666–677 (2014).

84. Keel, B. A., Makino, D. L., O’Donnell, M. & Kuriyan, J. How a DNA polymerase clamp loader opens a sliding clamp. *Science* 334, 1675–1680 (2011).

Acknowledgements
The authors are grateful to Simone Cavadini and Andreas Schenk for assistance with cryo-EM data collection, to Federica Galliano for helping clone one of the Cd6 expression constructs, and to Sandra Muehlhaeusser for culturing and maintaining insect cell stocks. We also thank members of the Bleichert laboratory and Susanne Kassube for helpful discussions and feedback on the manuscript. This work was supported by the Novartis Research Foundation and the European Research Council under the European Union’s Horizon 2020 research and innovation program (ERC-STG-757909 to F.B.).

Author contributions
Conceptualization, F.B.; methodology, J.M.S. and F.B.; formal analysis, J.M.S. and F.B.; visualization and analysis, J.M.S. and F.B.; writing—original draft, F.B.; writing—review & editing, J.M.S. and F.B.; visualization, F.B.; supervision, F.B.; funding acquisition, F.B.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-18067-7.

Correspondence and requests for materials should be addressed to F.B.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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

© The Author(s) 2020