A structural perspective on origin recognition and remodeling by the Origin Recognition Complex and its co-loader Cdc6 in eukaryotes

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Jan Marten Schmidt

aus

Deutschland

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Susan M. Gasser

Assistenz-Prof. Dr. Franziska Bleichert

Prof. Dr. Marc Bühler

Prof. Dr. Beat Fierz

Basel, den 17.11.2020

Prof. Dr. Martin Spiess

Dekan

Table of Contents

List of Figures

Summary

Prior to each cell division, genomic information must be duplicated precisely once to support cell proliferation and to preserve normal cell function. Hence, over- or under-replication of the genome can have detrimental consequences such as carcinogenesis and developmental disorders[1–7]. Across the tree of life, DNA replication initiates from discrete sites (also known as origins) to achieve timely genome duplication[8,9]. Origins are recognized by dedicated initiator proteins that collaborate with co-factor proteins to load replicative helicases onto DNA[8]. Replicative helicases are essential for unwinding the DNA and constitute the core of the DNA replication machinery[8,10,11].

In bacteria, archaea and a few eukaryotes, like *S. cerevisiae*, origins are defined by specific DNA sequence motifs that direct initiator binding[8,9]. In contrast, metazoan origins are not characterized by a DNA consensus sequence but by contextual factors in cis (like DNA structural properties and chromatin environment) and trans (protein factors), and the metazoan initiator Origin Recognition Complex (ORC) is known to bind DNA independent of a consensus motif[9,12,13]. However, it is a long-standing question how metazoan ORC recognizes origin DNA. In my Ph.D., I used single-particle cryo-EM to determine high-resolution structures of *Drosophila* ORC bound to different DNAs in presence or absence of the co-loader Cdc6. These structural studies provided insight into how the metazoan initiator and co-loader recognize DNA and help explain how DNA binding by ORC regulates initiator activity.

Upon binding by the initiator, origins are remodeled to support efficient helicase loading[8,9]. Origin remodeling by the initiator is likely influenced by intrinsic DNA sequence properties and local chromatin structure, but it is unclear how these factors contribute to initiator recruitment and helicase loading at origins in metazoans. We established a metazoan helicase loading assay *in vitro* and combined it with biochemical assays and cryo-EM to investigate how DNA sequence affects initiator binding, DNA remodeling and replicative helicase loading. We found that DNA sequence properties, such as minor groove width and negative electrostatic potential, contribute to metazoan ORC binding affinity and that metazoan ORC binding to a DNA substrate can be uncoupled from ORC's ability to bend DNA. However, we show that DNA bending by ORC is crucial for efficient

helicase loading. Hence, we suggest that DNA shape is an important factor in origin selection and licensing in metazoans.

Initiation of DNA replication is integrated into the cell cycle to ensure that all genomic information is duplicated exactly once before cells divide[14–16]. Cell-cycle dependent kinases (CDK) phosphorylate initiation and loading factors to regulate their activities, thereby coordinating the initiation of DNA replication with diverse cellular processes. Although phosphorylation-dependent regulation of initiation factors is well established, the underlying structural mechanisms are not well defined. Furthermore, origin licensing requires that the initiator, co-loader and helicase proteins undergo large conformational changes that are tightly coupled to ATP binding and hydrolysis[8,11,17]. We have just begun to understand how these dynamic structural rearrangements are regulated, and the physical basis for coordinated nucleotide binding and hydrolysis is largely unclear. We determined high-resolution cryo-EM structures of the ternary budding yeast ORC-DNA-Cdc6 complex to understand better the physical basis by which CDK-dependent phosphorylation regulates initiator activity and to investigate how the ternary ORC-DNA-Cdc6 assembly coordinates ATP hydrolysis. We propose a structural mechanism for coordinated ORC/Cdc6 ATPase activity and report a potentially autoinhibited conformation of the budding yeast ORC-DNA-Cdc6 complex that appears to be incompatible with helicase loading. Biochemical analysis suggest that ORC is phosphorylated in our ORC-DNA-Cdc6 structure, suggesting that we have uncovered a structural explanation of how CDK-dependent phosphorylation of ORC inhibits helicase loading.

Metazoan initiation factors are delicate multi-protein assemblies that have evolved highly-functionalized domains to support timely DNA replication[11,17]. Amino acid mutations in these specialized initiator domains have been shown to severely impair human development leading to the Meier-Gorlin Syndrome (MGS), but our molecular understanding of how these mutations derail initiation factor function is limited[6,7]. We have purified metazoan ORC harboring MGS-mutations and characterized their functional consequences using biochemical tools *in vitro*. We find that MGS-associated mutations

in *Drosophila* ORC compromise different initiator functions to varying degrees, which helps define the molecular basis for MGS.

Taken together, the studies presented here advance our knowledge of how metazoan origins are defined on a molecular level and how these characteristics might be read out by the initiator to facilitate origin selection and efficient helicase loading. Furthermore, our comprehensive structural characterization of *Drosophila* and budding yeast ORC in complex with DNA and Cdc6 reveals dynamic conformational rearrangements that are necessary for helicase loading and suggest mechanisms for the controlled hydrolysis of ATP by the initiator and its co-loader. Moreover, we unveiled an unknown conformational state of the budding yeast ORC-DNA-Cdc6 complex that provides a structural explanation for how phosphorylation of ORC regulates origin licensing. Finally, our functional characterization of MGS mutations in ORC aids our understanding of the molecular mechanism underlying this developmental disease. The *in vitro* metazoan helicase loading assay established by my work will facilitate future research efforts to explore how origin licensing is regulated in higher eukaryotes.

Abbreviations

AAA+	ATPases Associated with various cellular Activities
ACS	Autonomously replicating sequence Consensus Sequence
ARS	Autonomously Replicating Sequence
ATP	Adenosine TriPhosphate
AT-rich	Adenosine-Thymidine-rich
BAH	Bromo-Adjacent homology
BP	Basic Patch
Budding yeast	Saccharomyces cerevisiae
CBF	Cyclin-Box Fold
Cdc6	Cell division cycle 6
Cdt1	Chromatin licensing and DNA replication factor 1
CE	Catalytic Glutamate
CpG	Cytosine-Guanine dinucleotide
Cryo-EM	(single particle) cryo-electron microscopy
DCAF14	DDB1- and CUL4-associated factor
DNA	DeoxyriboNucleic Acid
DnaA	Chromosomal replication initiator protein DnaA
DnaB	Replicative DNA helicase DnaB
DnaC	DNA replication protein DnaC
Drosophila	*Drosophila melanogaster*
ds	double-stranded
G4	Guanine-quadruplex
GC-rich	Guanine-Cytosine-rich
HMGA1a	High Mobility Group protein A1a
HP1	Heterochromatin Protein 1
IDR	Intrinsically Disordered Region
IS	Initiation Site
ISM	Initiator Specific Motif
LRWD1	Leucine Rich Repeats And WD Repeat Domain Containing 1
Mcm2-7	Minichromosome maintenance complex2-7
MGS	Meier-Gorlin Syndrome
OD	ORC-DNA
ODC	ORC-DNA-Cdc6
OGRE	Origin G-rich Repeated Element
ORC	Origin Recognition Complex
RF	Arginine Finger
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Saccharomyces pombe*

S1	Sensor 1
S2	Sensor 2
ss	single-stranded
TFIIB	Transcription Factor II B
WA	Walker A
WB	Walker B
WH	Winged Helix

Introduction

Replication of DNA

Organisms in all domains of life use double-helical DeoxyriboNucleic Acid (DNA) to encode and store genetic information essential for cellular life[18–20]. The genetic information encoded in DNA must be correctly duplicated in a semi-conservative manner to support cell proliferation[21]. Erroneous replication of DNA can result in gene amplification, genome instability and genetic disorders that interfere with cell functions and development[1,3–7]. Consequently, DNA synthesis must be tightly controlled and integrated with various cellular processes to ensure precise and timely replication of genomic information[8].

DNA is replicated by highly functional replisomes, which are assembled on chromatin in several discrete steps[11,22–25]. First, dedicated initiator proteins recognize defined genomic loci, also known as origins, to which they recruit and deposit replicative helicases in a process termed origin licensing[9]. Second, the loaded helicases are activated leading to origin DNA unwinding[11]. Third, the replisome is assembled on the activated helicase core and DNA synthesis begins[10,26].

Although DNA replication is functionally highly conserved across domains of life, the enzymes executing it have evolved independently in bacteria and archaea/eukaryotes resulting in a significant degree of divergence[8,27,28]. In bacteria, origin of replications are defined by DNA consensus motifs that are recognized by the initiator DnaA, which oligomerizes on origin DNA to deposit the replicative helicase DnaB in collaboration with the loader protein DnaC onto single-stranded (ss) DNA[29–31]. Archaea are phylogenetically closer to eukaryotes than to bacteria[32], which is reflected in the high degree of conservation between eukaryotic and archaeal initiator proteins[33]. The archaeal initiator Orc1/Cdc6 binds to origins in a sequence-dependent manner to recruit and load the homo-hexameric replicative helicase minichromosome maintenance complex (MCM) on double-stranded (ds) DNA[34,35]. Notably, loading of archaeal MCM does not rely on additional co-loading factors in contrast to the bacterial and eukaryotic systems. Eukaryotic genomes are significantly larger than those of prokaryotes, which required that

the eukaryotic DNA replication system co-evolved in its complexity. Eukaryotic origins are recognized by the initiator Origin Recognition Complex (ORC), which is a hetero-hexameric ATPase that, in concert with the co-loaders Cdc6 and Cdt1, deposits replicative helicases onto origin DNA[11,17]. In contrast to bacteria and archaea, origins are not defined by a DNA consensus motif in the majority of eukaryotes, with the exception of budding yeast and closely related fungi, and it is not well understood how the eukaryotic initiator selects those replicator regions[9,13].

Origin specification in eukaryotes

Eukaryotes rely on several hundred (*S. cerevisiae*: approx. 400) to thousands (human: approx. 50000) of origins of replication to duplicate their genomes faithfully during one cell cycle[36–39]. The location of origins is not random and DNA replication must be tightly coordinated with transcription to avoid replication-transcription conflicts, which can result in incomplete genome duplication and promote genome instability[40–42]. However, the factors that determine genomic loci as origins have diverged among eukaryotes. They can be broadly divided into two classes, one that is guided by a DNA consensus sequence, and a second that is context-dependent[9,13].

DNA-sequence-guided specification of replication origins is rare in eukaryotes and has only been observed in budding yeast and a few closely related fungi. In *S. cerevisiae*, origins of replication are defined by an autonomously replicating sequence (ARS)[43–45], which is composed of up to four functional regions: the A, B1, B2 and B3 element (Figure 1)[46,47]. The budding yeast initiator Origin Recognition Complex (ORC) recognizes a 11bp-long ARS consensus sequence (ACS) of the AT-rich A element together with the B1 element in an ATP-dependent manner[48–52]. The B2 region may represent a low-affinity binding site for ORC[53–57], while the B3 element has been shown to recruit the transcription factor Abf1[46,58,59]. However, not all origins contain a B3 region and Abf1 recruitment is not required for origin activity [46,58,59].

In most other eukaryotic systems, replication initiation sites (IS) are context-dependent and are specified by a combination of epigenetic marks, chromatin state, DNA topology and DNA structural properties[9,13,60–63]. Although metazoan ORC has a slight preference for AT-rich DNA *in vitro*[64–66], no consensus sequence for ORC DNA recognition has been identified *in vivo*[37]. Instead ORC seems to bind accessible chromatin sites and regions that contain active chromatin marks[37,67–72]. However, it is largely unclear what DNA sequences and chromatin features actually determine initiator binding or the efficiency of initiation sites in different eukaryotic systems.

Several sequence characteristics have been found to correlate with proximal replication initiation sites (IS) in metazoans although no causal relationship between DNA sequence and genomic metazoan ORC binding sites has been established (Figure 1)[9,13].

It has been suggested that metazoan replication initiation sites are associated with CpG islands and G-rich motifs, so called Origin G-rich Repeated Element (OGRE), which may fold into G-quadruplex structures[36,73–81]. In accordance with these observation, human Orc1 and Orc2 genomic binding sites correlate with GC-rich sequences and it has been suggested that human ORC preferentially binds to G-quadruplex forming ssDNA *in vitro*[37,82,83]. Interestingly, it has been proposed that asymmetric distribution of A/T and G/C nucleotides and flanking poly(dA:dT) tracts, both of which being prone to deviate from B-DNA structure, correlate with origin activity in mammals[84,85]. Notably, polyA and polyG tracts are characterized by high intrinsic stiffness that is incompatible with nucleosome binding and therefore might promote the formation of nucleosome-free regions[69,79,85–88]. The nucleosome-excluding properties of polyA and polyG tracts may help explain the enrichment of polyA and polyG sites close to replication origins.

Mounting evidence suggest that DNA geometry and topology rather than specific sequence might be of importance for regulating origin recognition and helicase loading in higher eukaryotes. Negative supercoiling of DNA has been shown to promote *Drosophila* ORC binding to origins and budding yeast and *Drosophila* ORC bend DNA, which may aid helicase loading[52,66,89–91]. The role of DNA geometry in defining metazoan replication origins is supported by work from Comoglio et al., who identified that DNA at active *Drosophila* and human origins has a relatively low helix twist while propeller twist, minor groove width and roll are increased, rendering the origin DNA sequence conformational flexible[79]. Nonetheless, how these contextual cues are integrated to guide origin recognition in eukaryotes remains enigmatic. Even less clear is how these factors ultimately influence origin activity throughout the genome under various conditions.

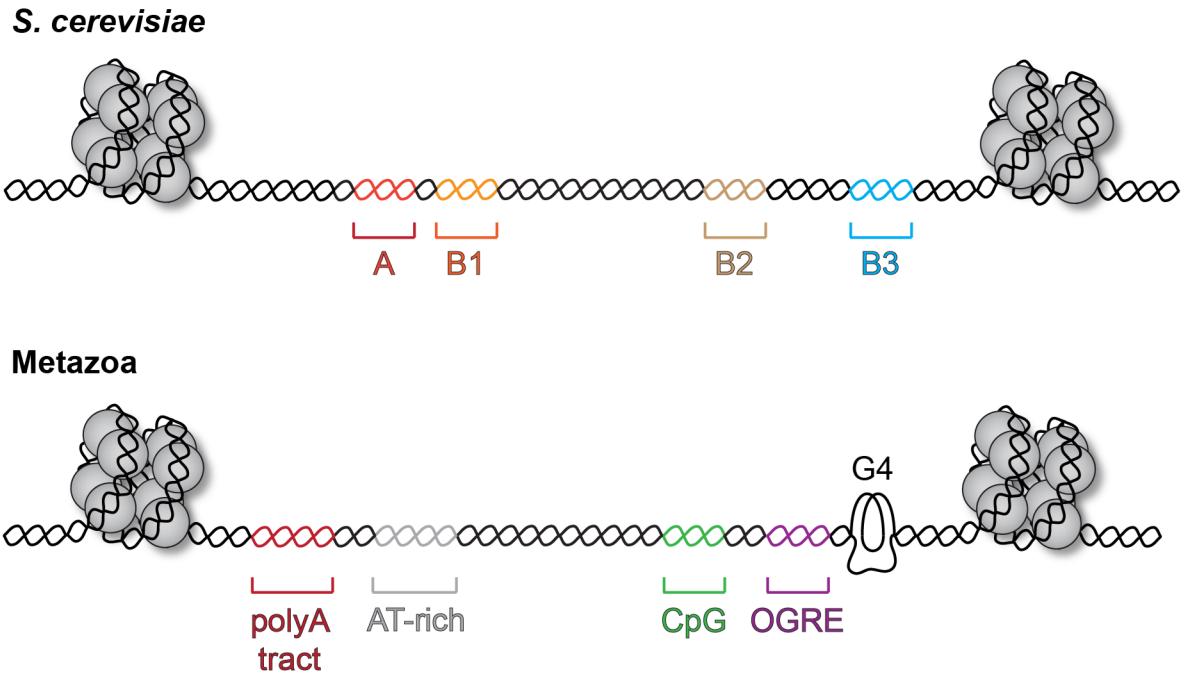


Figure 1: Origin specification in eukaryotes. Illustration of key characteristics of origins in budding yeast and metazoans. In *S. cerevisiae*, origins harbor an A and B1 element, the primary binding site of the initiator, and a B2 and B3/Abf motif. Metazoan origins are defined by skewed A/T and G/C nucleotide distribution, CpG islands, flanking polyA or polyG stretches, OGRE and potentially G-quadruplexes. For simplicity, histone marks that have been found to correlate with metazoan replication start sites have been omitted. This figure has been adapted from Ref. 9.

Structure and function of the Origin Recognition Complex and its co-loader

The eukaryotic Origin Recognition Complex (ORC) is organized as a hetero-hexamer with subunits Orc1-6 (Figure 2). Orc1-5 are characterized by an N-terminal ATPases associated with various cellular activities (AAA+) domain and a C-terminal winged helix (WH) fold[50,92–95], as is the co-factor Cell division cycle 6 (Cdc6), which is closely related to Orc1[96–99]. By contrast, Orc6 consists of two cyclin-box motifs that share similarities with the transcription factor IIB (TFIIB), and a C-terminal helix[100–102]. The AAA+ and WH domains of Orc1-5 assemble into an open two-tiered ring (arranged as Orc1–Orc4–Orc5–Orc3–Orc2), in which the WH motifs dock against the AAA+ domains of adjacent subunits[52,103,104]. Orc1-5 encircle a central conduit, which can engage double-stranded (ds) DNA[52,90,103–105]. Orc6 does not participate in Orc1-5 ring assembly but is bound outside of the Orc1-5 ring to an insertion fold in Orc3 through the Orc6 C-terminal domain[52,102,103,106]. Oligomerization of the AAA+ domains of Orc1-5 leads to the formation of composite ATPase sites between neighboring protomers[52,103,104]. However, only ORC subunits 1, 4 and 5 bind nucleotides and it has been suggested that the composite site formed by Orc1 and Orc4 is the primary catalytic center of the initiator, albeit it is unclear whether the Orc4/Orc5 ATPase site has residual activity[17,51,103,104,107,108]. Binding of Cdc6 to ORC on DNA results in an additional, functional composite ATPase sites between Cdc6 and Orc1[94,99,105,109,110].

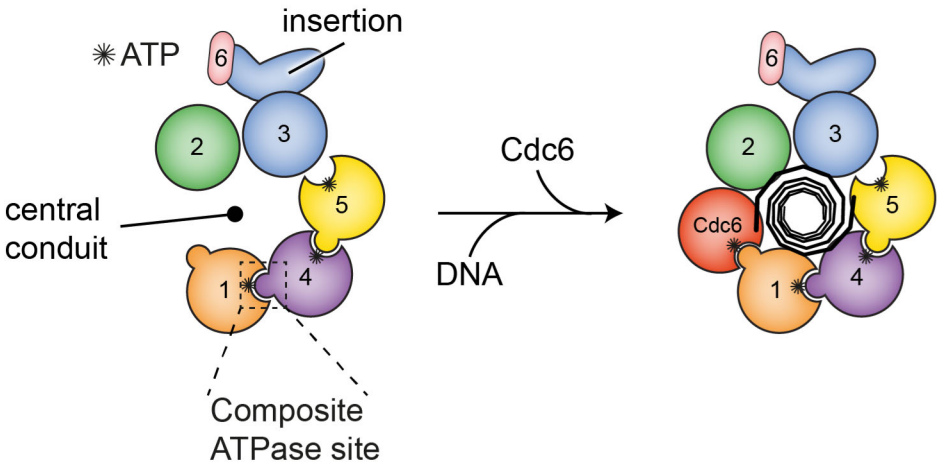


Figure 2: Organization of the AAA+ tier of the Origin Recognition Complex with or without its co-factor Cdc6. In the presence of ATP, Orc1-5 oligomerize into an open-ring assembly. The AAA+ domain of Orc3 includes an insertion fold that links Orc6 to the Orc1-5 ring. The AAA+ domains of Orc1, Orc4 and Orc5 are nucleotide-bound and the composite ATPase site of Orc1/4 has hydrolase activity. Upon DNA binding, Cdc6 joins the complex, forming a composite ATPase center with Orc1 resulting in a closed ORC-Cdc6 ring. This figure has been adapted from Ref. 17.

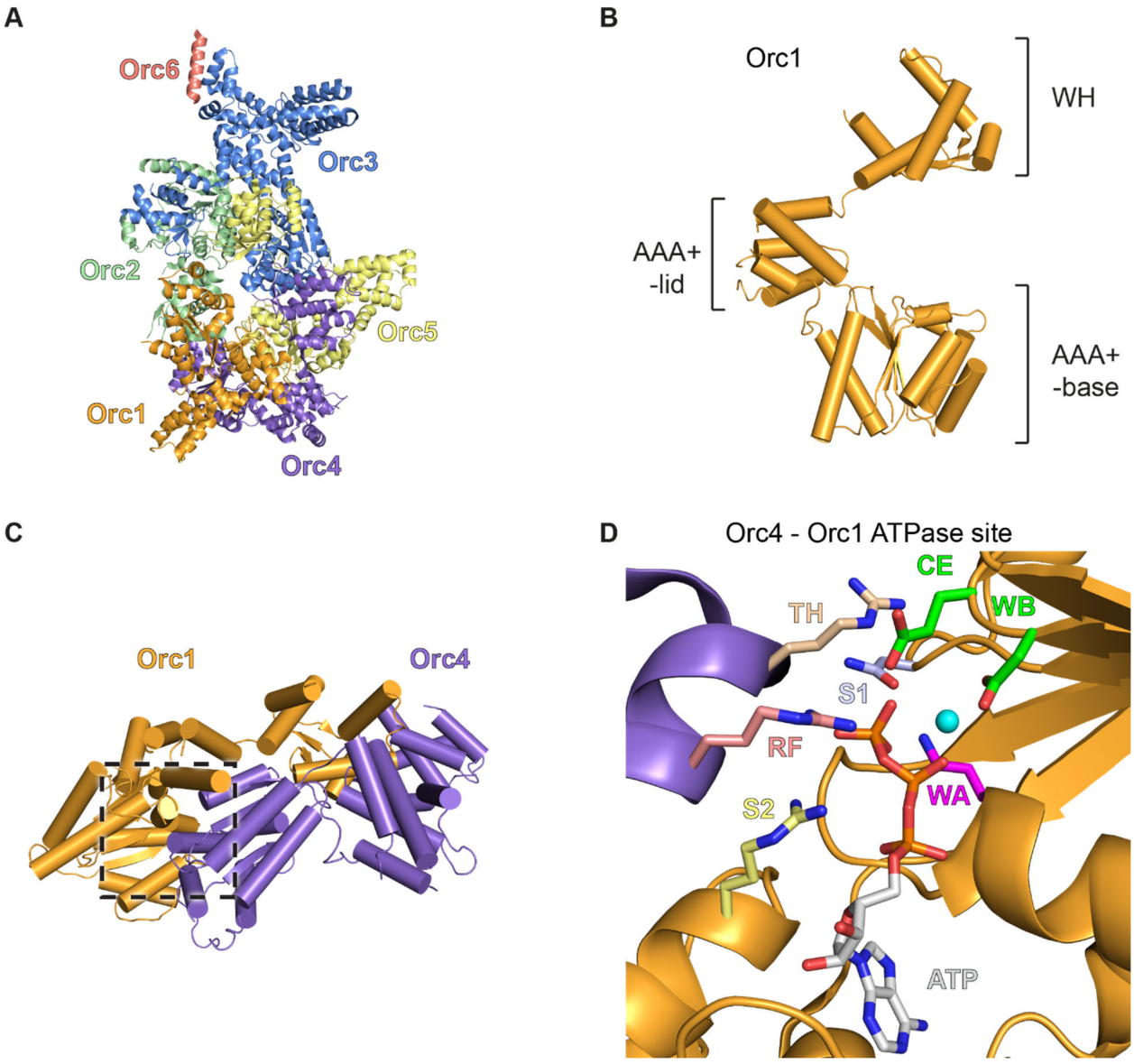


Figure 3: ORC subunit organization and Orc1/4 composite ATPase site. **A**) Crystal structure of *Drosophila* ORC (PDB: 4XGC[103]). **B**) Domain architecture of the human Orc1 subunit. Orc1, like Orc2-5, has a AAA+ domain and a C-terminal winged-helix domain. The AAA+ fold can be further divided into a AAA+-base and a AAA+-lid subdomain. **C**) Overview of the Orc1/Orc4 assembly within the human Orc1-5 complex. **D**) Zoom into the human Orc1/4 composite ATPase center. Orc4 and Orc1 are shown in purple and orange, respectively. Key residues of the composite ATPase site are shown as sticks. **B-D**) are based on PDB: 5UJ7[104]. ATP- Adenosine Triphosphate, WA – Walker A, WB - Walker B, CE – Catalytic Glutamate, S1 – Sensor 1, TH – Tether, RF – Arginine Finger, S2 – Sensor 2.

The AAA+ domains of ORC and Cdc6 can be further divided into a N-terminal AAA+-base and a C-terminal AAA+-lid sub-domain, which feature several residues essential for ATP binding and hydrolysis (Figure 3)[92,93,111–113]. In detail, the AAA+-base subdomain encompasses a Walker A (GXXGXGKT/S) and B (D(D/E)) motif that mediate ATP binding and catalysis, respectively[51,113,114]. Furthermore, the AAA+-base domains of ORC and Cdc6 have a polar Sensor 1 residue, which assists the WB motif in nucleotide hydrolysis[110,115–117]. ATP hydrolysis by the composite ATPase centers within the ORC/Cdc6 complex also requires a highly conserved arginine residue, known as the R-finger[108,109,117,118]. The Walker A and B, Sensor 1 and R-finger are complemented by a second conserved basic residue, referred to as a Tether, that contributes to the hydrogen-bond network within the ATPase center[52,103,108,113]. The AAA+-lid subdomain harbors a conserved arginine or lysine residue, known as Sensor 2, which has been proposed to be involved in sensing the status of the bound nucleotide and in the transmission of conformational changes[111,113,119].

Notably, the functionalities of the AAA+ domains of Orc1-5 have diverged. The Cdc6/Orc1 and Orc1/Orc4 ATPase centers contain all functional motifs (WA, WB, R-finger, Sensor 1 and 2, Tether) and have hydrolase activity, while the Orc4/Orc5 interface possesses a R-finger (Orc5) and a WA and WB motif (Orc4), but lacks the Sensor 2 arginine and shows no detectable ATP hydrolysis in *Drosophila* and *S. cerevisiae*[51,52,94,103,105,107,109,110]. The interface between Orc5 and Orc3 has deviated even more from the canonical AAA+ composite ATPase site. *Drosophila* and budding yeast Orc5 contain a canonical Walker A and a non-canonical Walker B motif (DN instead of D(D/E)) that facilitate nucleotide binding, but lack a Sensor 2 element, and Orc3 misses the arginine finger residue[52,103]. As a consequence, the Orc5/Orc3 interface binds ATP but is more open than the Orc1/Orc4 and Orc4/Orc5 sites and shows no hydrolase activity[51,52,103,104,107,120]. The composite AAA+ site formed by Orc3 and Orc2 has no nucleotide binding and hydrolysis residues and likely serves a structural role in the ORC assembly[103].

In addition to the AAA+ and WH domains, Orc1-5 have ancillary domains that provide further functionalities to the complex and mediate its regulation (Figure 4).

Accordingly, Orc1 possesses a N-terminal Bromo-Adjacent-Homology (BAH) domain that has been implicated in ORC-chromatin binding and specifically recognizes di-methylated lysine-20 of histone 4 in mammals[121–126]. The BAH domain of Orc1 is appended by an intrinsically disordered region (IDR), which allows ORC-droplet formation in the presence of DNA[127]. A region just N-terminal of Orc1's AAA+ is characterized by multiple basic amino acid residues (so called basic patch (BP)), which are crucial for DNA binding[90,128].

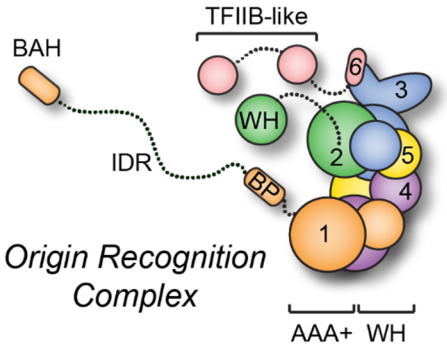


Figure 4: The eukaryotic Origin Recognition complex. Cartoon representation of eukaryotic ORC highlighting the AAA+ and WH tier organization and ancillary domains of the complex. BAH – Bromo-Adjacent-Homology domain, IDR – Intrinsically Disordered Region, BP – Basic Patch, WH – Winged Helix domain, TFIIB – Transcription Factor II B. This figure has been adapted from Ref. 17.

Origin DNA recognition by the eukaryotic Origin Recognition Complex

How DNA is recognized by ORC is different among eukaryotes and can be either sequence-specific, as in *S. cerevisiae*, or sequence-independent, as in almost all other eukaryotes[17]. A recently published high-resolution structure of the *S. cerevisiae* ORC bound to origin DNA addressed the long-asked question of how the budding yeast initiator achieves sequence-specific DNA recognition (Figure 5)[52]. All subunits of the budding yeast initiator contribute to DNA recognition, albeit by using different motifs that either interact with the sugar-phosphate backbone of the DNA or make sequence-specific base contacts[52]. Accordingly, the AAA+ domains of Orc1, Orc2, Orc3 and Orc5 contact the DNA backbone by an initiator specific motifs (ISM) in the AAA+ domain[52,93]. The beta-hairpin elements, which are part of the winged-helix domain, of Orc2, Orc4 and Orc5 extent into the DNA major grooves, while the beta-hairpin motif of Orc3 interacts with the DNA backbone[52]. In addition, budding yeast ORC uses basic regions in Orc2 and Orc5 as well as the C-terminal cyclin-box-fold of Orc6 to bind to the B1 element[52]. Sequence-specific recognition of the 11bp-long ARS consensus sequence (ACS) is mediated by the basic patch of Orc1, the ISM of Orc2 and a yeast-specific insertion helix in the WHD of Orc4, all of which bind thymine bases within the ACS[52]. Although all three elements make base-specific contacts, it has been proposed that the insertion helix in Orc4 is mainly responsible for origin recognition in budding yeast[52,129]. However, further investigations are required to clearly establish this finding.

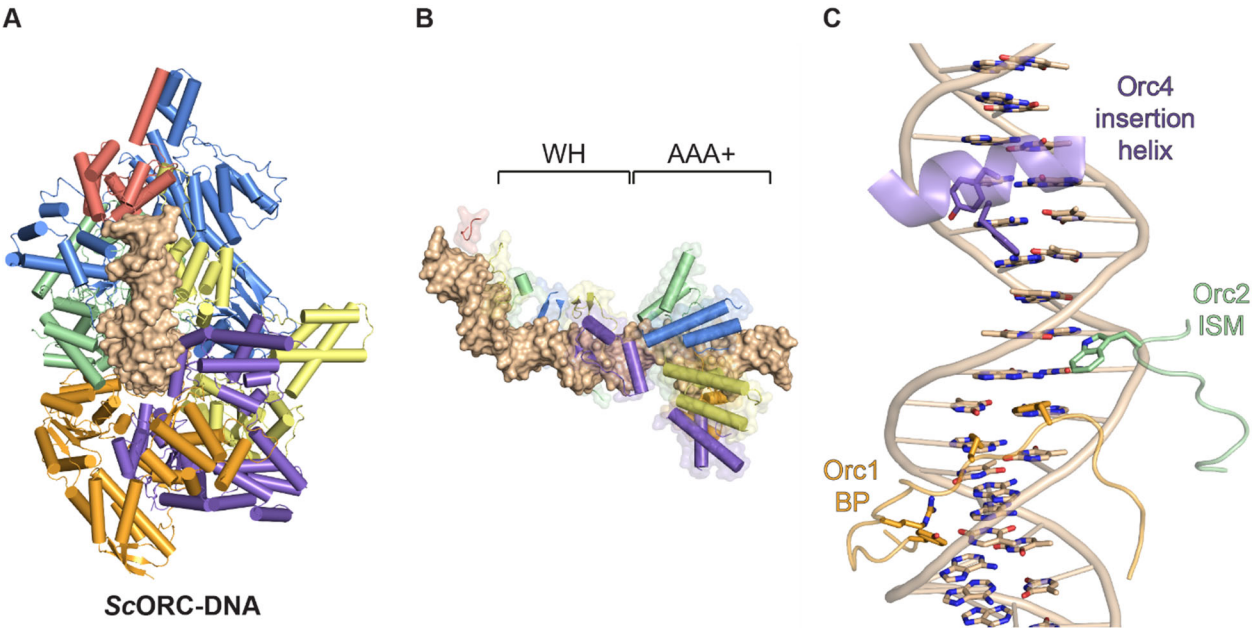


Figure 5: Sequence-specific origin recognition by S. cerevisiae ORC. **A**) Structure of DNA-bound budding yeast ORC (PDB: 5ZR1)[52]. **B**) Overview of region contacting DNA in the *Sc*ORC-DNA complex. The location of the AAA+ and WH tier is indicated. **C**) Close-up on Orc1-BP, Orc2-ISM and Orc4-insertion-helix, which mediate base-specific DNA recognition in *S. cerevisiae* ORC. Residues in *Sc*ORC that make base-specific contacts are shown as sticks. This figure has been adapted from Ref. 17.

In contrast to the yeast initiator, we know very little about how the AAA+ and WH domains of the different subunits of metazoan ORC cooperate to facilitate sequence-independent origin recognition. Low purification yields and poor behavior of the full length *Drosophila* and human initiator have impeded structural efforts to obtain high resolution structures of the DNA-bound metazoan ORC complex that would help us to understand how metazoan ORC engages DNA in a sequence-independent manner. To this end, our structural knowledge of the metazoan initiator is limited to a high-resolution structure of *Drosophila* ORC in an auto-inhibited conformation[103], a medium of active human ORC without DNA[104] and low resolution 2D class averages of the DNA-bound *Drosophila* ORC/Cdc6 complex[90], none of which could report on precise protein-DNA contacts. Therefore, we still await high resolution structures of DNA-bound metazoan ORC that will help us establish how the metazoan initiator binds DNA independent of a consensus sequence.

It is important to note, that *S. cerevisiae* ORC induces a bend into the DNA upon binding, as observed for the *Drosophila* initiator[52,90], which might aid replicative helicase loading [52,90]. However, the functional consequence of ORC-induced DNA bending on helicase loading *in vitro* and *in vivo* is largely unclear. Likewise, it is unknown how DNA bending by ORC might influence origin selection.

Origin licensing and activation in eukaryotes

Initiation of DNA replication can be described as an orchestrated sequence of initiator-cofactor intermediates that undergo conformational changes to achieve coordinated replicative helicase recruitment and loading[11,17]. Once ORC has bound DNA in its central channel, it can recruit and load the replicative helicase minichromosome maintenance protein 2–7 (Mcm2-7) complex onto origins with the help of the co-loading factors Cdc6 and Chromatin licensing and DNA replication factor (Cdt1)[130–139]. In this process, two replicative helicases are sequentially deposited onto origins in a head-to-head orientation resulting in the formation double-hexamer (DH) complexes[11,140–142].

The molecular mechanisms by which ORC and Cdc6 recruit and load two Mcm2-7 hexamers leading to the assembly of the DH complex remain an area of active research (Figure 6 and 7). Prior to helicase recruitment, Cdc6 associates with DNA-bound ORC, which closes the initiator ring locking it on DNA and resulting in ORC-DNA-Cdc6 (ODC) complex formation[90,105,143,144]. The closed ORC/Cdc6 ring can then bind the replicative helicase by interacting with the C-termini of Mcm3 and Mcm7[91,105,145,146]. Like ORC, the Mcm2-7 complex is a heterohexameric AAA+ ATPase and is organized in a N-terminal and C-terminal tiered arrangement[92,93,147]. The MCM ring is not closed but possesses an opening (gate) between Mcm2 and Mcm5, which allows substrate binding in the channel of the MCM ring[105,148–150]. Binding of the co-factor Cdt1 to Mcm2-7 is crucial for loading by stabilizing the replicative helicase ring in an open conformation and by releasing an autoinhibitory mechanism of Mcm6[137–139,146,151–153]. Upon initial Mcm2-7 recruitment to ORC/Cdc6, the bent DNA is step-wise aligned with the gap in the MCM ring and then released from ORC to enter the central channel of the helicase resulting in the transitory ORC-Cdc6-Cdt1-MCM (OCCM) complex[91,105,154].

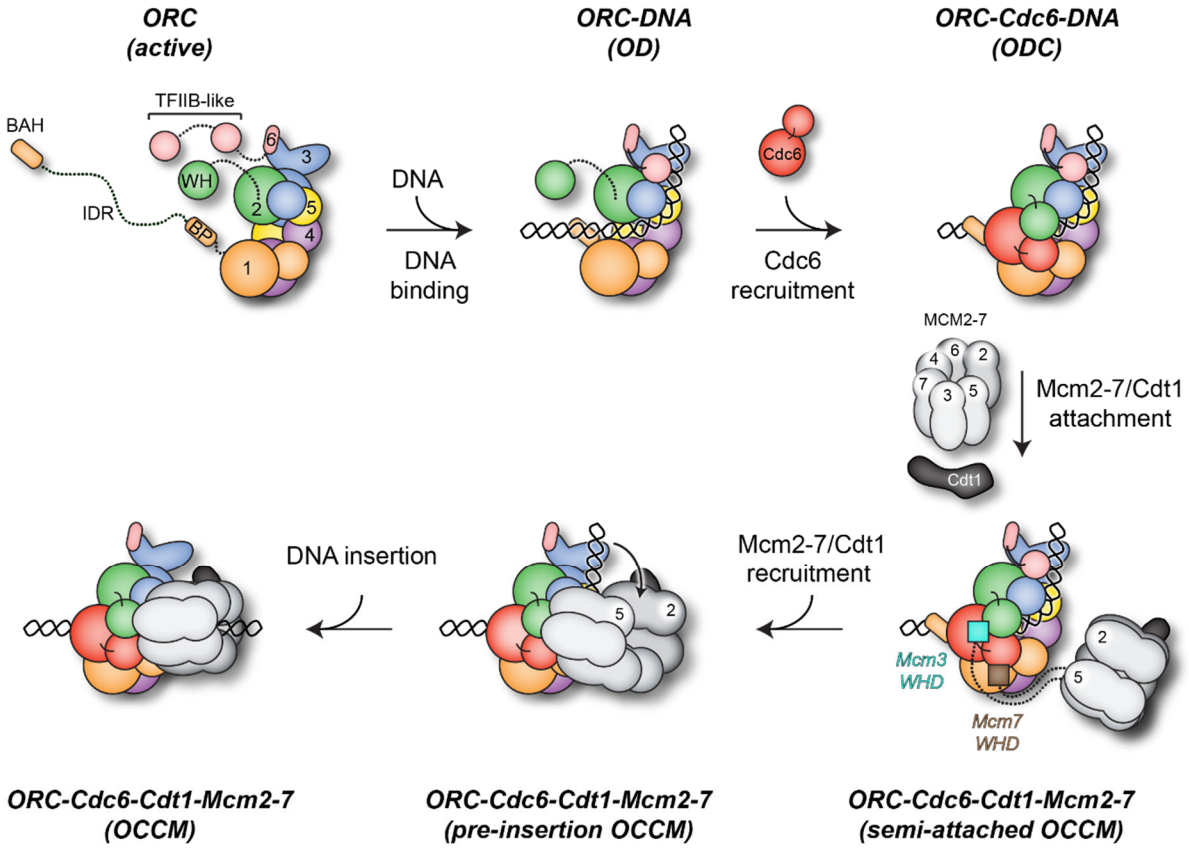


Figure 6: Origin DNA remodeling and loading of the first Mcm2-7 hexamer by eukaryotic ORC. ATP-bound ORC binds DNA in its central channel and bends DNA (OD complex), resulting in stabile docking of the C-terminal cyclin-box fold of Orc6 in budding yeast. DNA-bound ORC recruits Cdc6 (ODC complex), which locks the complex on DNA. Next, the first Mcm2-7 hexamer dynamically associates with the ODC complex, mediated by the WH domains of Mcm3 and Mcm7 (semi-attached OCCM). Stable binding of Mcm2-7 to ORC/Cdc6 aligns the bent DNA with the Mcm5/Mcm2 gate (pre-insertion OCCM). Subsequently, the DNA is inserted into the central channel of Mcm2-7 (OCCM). Dashed lines indicate flexible regions. This figure has been adapted from Ref. 90.

It has been long-debated how the second Mcm2-7 hexamer is recruited to origins (Figure 7)[11,145,155]. A structural study by Miller et al. provides an explanation how ORC orchestrates the loading of the two Mcm2-7 hexamers by the OCCM pathway in the budding yeast system[154]. Accordingly, engagement of the high-affinity ARS consensus sequence (ACS) DNA element by ORC leads to the recruitment of the first Mcm2-7 hexamer and formation of the OCCM intermediate[154]. Once Cdc6 and Cdt1 have been

released and the Mcm2-7 gate has closed, ORC detaches from the loaded Mcm2-7 hexamer[154]. Next, ORC binds at the inverted B2 element and associates with the N-terminus of the loaded Mcm2-7 hexamer resulting in the Mcm2-7-ORC (MO) intermediate[154]. In the MO structure, the interface between ORC and the MCM hexamer is primarily formed by the TFIIB domain of Orc6 and the N-termini of Mcm2 and Mcm5[154]. This binding mode secures the locked Mcm2/5 gate while allowing the binding of the second Mcm2-7 hexamer to the C-terminal face of the ORC/Cdc6 ring[154]. These structural findings argue that ORC binds at two inverted DNA sites to load both Mcm2-7 hexamers via the same interface[145,154]. This idea is supported by the observations that double hexamer assembly is facilitated by binding of ORC to a high and an inverted and up to 400bp distant low affinity DNA site in *S. cerevisiae* suggesting a symmetric mechanism for helicase loading[57]. However, the relevance of the MO pathway is unknown as the finding leaves the possibility that Mcm2-7 hexamers are loaded individually at two inverted DNA sites without going through the MO intermediate, and then detach from the initiator and passively slide along the DNA until they encounter each other[57]. Furthermore, it is unclear whether Mcm2-7 loading involves one or two ORC molecule. How one ORC can disengage from the high-affinity ACS site to subsequently bind to the low-affinity B2 element during MO formation is unknown. Moreover, if either one or two ORC molecules can load the double hexamer, are both mechanisms or only one relevant *in vivo*?

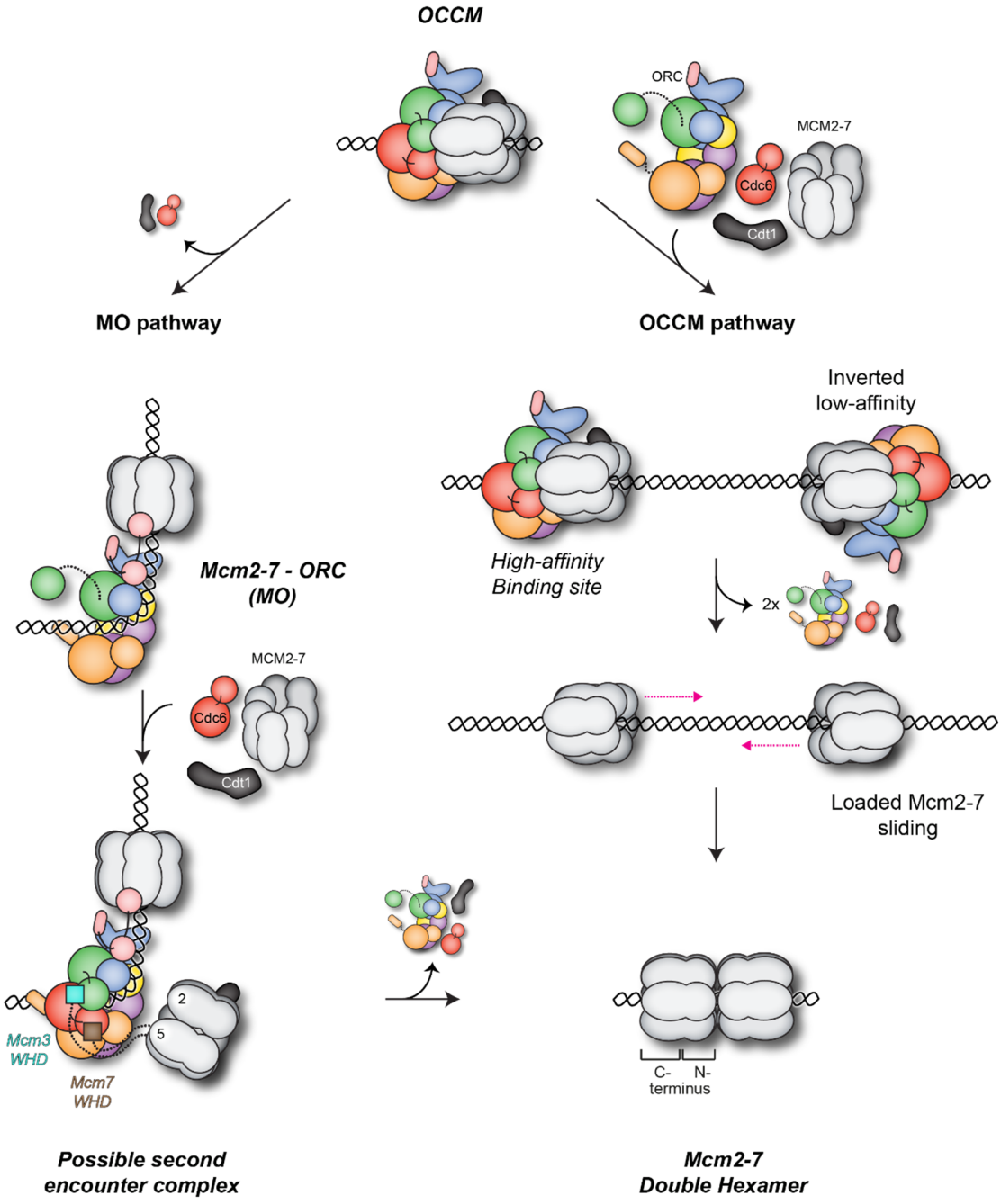


Figure 7: Loading of the second Mcm2-7 hexamer. Comparison of two proposed mechanisms for second helicase loading. In case of the Mcm2-7-ORC (MO) pathway, Cdc6 and Cdt1 disengage from OCCM and the gate of the Mcm2-7 closes. ORC is released from the ACS-B1 DNA site and binds the low-affinity B2 element. The cyclin-box folds of Orc6 play a critical role in this mechanisms as they represent the only contact point between ORC and Mcm2-7 in the MO complex and help orient ORC for recruitment of the second Mcm2-7. In the MO complex the winged-helix interface of ORC is free for binding of the second helicase hexamer through the OCCM mechanism. Loading of the

second hexamer and DH formation likely occurs through formation of another OCCM intermediate and requires the release of ORC, Cdc6 and Cdt1. In contrast, second Mcm2-7 hexamer loading could be achieved without the MO intermediate by assembly of a second OCCM complex at an inverted low-affinity binding site. This would result in loading of two individual Mcm2-7 hexamers, which may passively slide along DNA until encounter leading to DH formation. Both pathways could be executed by one or two ORC molecules. This figure has been adapted from Ref. 17 and 90.

After loading, the Mcm2-7 double hexamer is activated by the combined action of cell-cycle-regulated kinases and firing factors (Figure 8). To become activated, the loaded Mcm2-7 DHs are phosphorylated at the N-termini of Mcm4 and Mcm6 by the Dbf4-dependent Cdc7 kinase (DDK) to recruit the Cell division control protein 45 (Cdc45) and the co-factor Synthetic lethal with Dpb11 (Sld) 3 in complex with Sld7[156–161]. Subsequently, the firing factors Sld2 and Sld3 are phosphorylated by S phase cyclin-dependent kinase (S-CDK), which allows Sld2 and Sld3 to bind to DNA polymerase B 11 (Dpb11)[162–164]. Dpb11 itself interacts with the coactivator Go-Ichi-Ni-San (GINS) complex, which is conserved in eukaryotes[165–168]. Dpb11 together with Sld2, GINS and the leading strand DNA polymerase epsilon can assemble into a pre-loading complex suggesting that these components are recruited as one entity to the loaded Mcm2-7 through interaction of Dpb11 with phosphorylated Sld3[168,169]. Association of Cdc45 and GINS to the Mcm2-7 double hexamer results in the formation of two, individual Cdc45-Mcm2-7-GINS (CMG) entities, each with DNA unwinding activity[11,167,170–176].

The CMG complex constitutes the motor module of the moving replisome and is assisted by co-factors to support efficient DNA replication. CMG assembly, initial DNA unwinding and subsequent replisome processivity is promoted by binding of another initiation factor, Mcm10, to the CMG[177–181]. The CMGs translocate with the N-terminal face of the helicase ring first 3’ to 5’ on the leading strand, which passes through the Mcm2-7 central pore whilst the lagging-strand is excluded[176,182–184]. DNA unwinding by CMG+Mcm10 facilitates the recruitment of polymerase alpha/primase and further replisome processivity factors, cumulating in replisome assembly, which in turn initiates leading and lagging strand DNA synthesis by DNA polymerase epsilon and DNA polymerase delta, respectively[11,24,176,185–188].

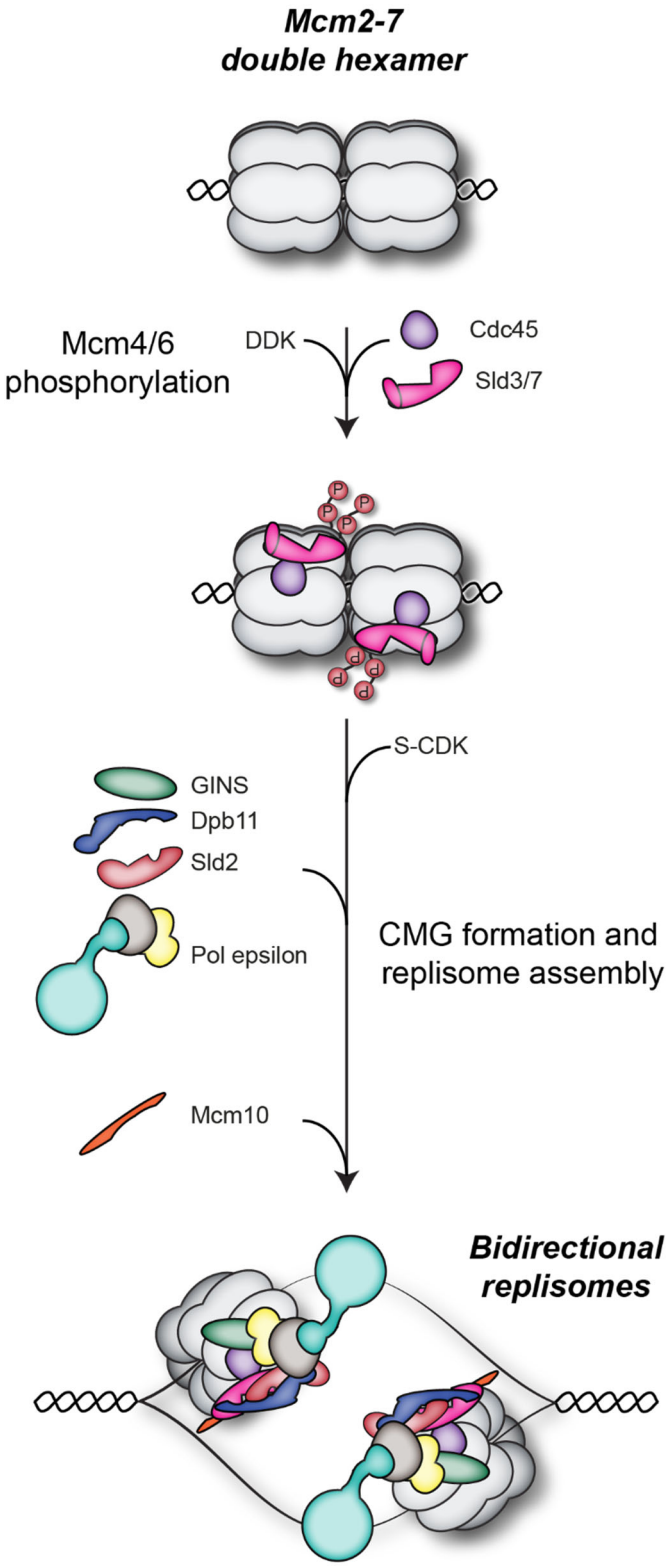


Figure 8: Bidirectional replisome assembly. First, loaded Mcm2-7 DHs are activated by DDK-mediated phosphorylation of the Mcm4 and Mcm6 N-termini and Cdc45 and Sld3/7 are recruited. Second, S-CDK phosphorylates the co-factors Sld2 and Sld3 triggering the formation of two CMG entities and the assembly of bidirectional, leading-strand replisomes. This figure has been adapted from Ref. 11.

ATP binding and hydrolysis is a central mechanism of regulating the initiation of DNA replication

Nucleotide binding and hydrolysis by ORC, Cdc6 and Mcm2-7 are important regulators for origin licensing and are essential factors *in vivo*[11,17]. ORC requires ATP to transition from an autoinhibited to an active conformation, in which it can engage DNA (Figure 9)[90,103]. Cdc6 is recruited to DNA-bound ORC in a nucleotide-dependent manner to load Mcm2-7, thereby forming a composite ATPase site with Orc1[94,105,109,110,132–134,136,189]. Subsequently, the ORC-Cdc6 complex can bind the Mcm2-7 hexamer, with the first point of contact being between Cdc6 and the WH domain of Mcm3 and Mcm7[91]. ATP hydrolysis by Cdc6 precedes ORC's hydrolase activity[109] and likely destabilizes the Cdc6/Orc1 interface resulting in the release of the co-loader and not yet loaded or non-functional Mcm2-7 hexamers[118,145,146,150,190]. By contrast, ORC hydrolase activity is likely required for iterative rounds of helicase loading by promoting initiator release from DNA[50,51,108,155]. Interestingly, DNA binding to ORC's central conduit reduces ORC's ATP activity, but how ATP hydrolysis is coupled to substrate binding, is not understood[51,66,107]. Moreover, how the sequential nucleotide hydrolysis events (first Cdc6 then Orc1) are coordinated by the ternary ORC-DNA-Cdc6 is an open question. Despite the fact that ATPase activity by ORC and Cdc6 is crucial for cell survival, only ATP hydrolysis by Mcm2-7 is required for double hexamer (DH) formation *in vitro*[98,108,116,118,190–192]. The mechanisms controlling ORC and Cdc6 nucleotide binding and hydrolysis and their effect on downstream initiation events are poorly defined. Further research is needed to decipher the precise role of ORC and Cdc6 ATPase activity during origin licensing *in vivo*.

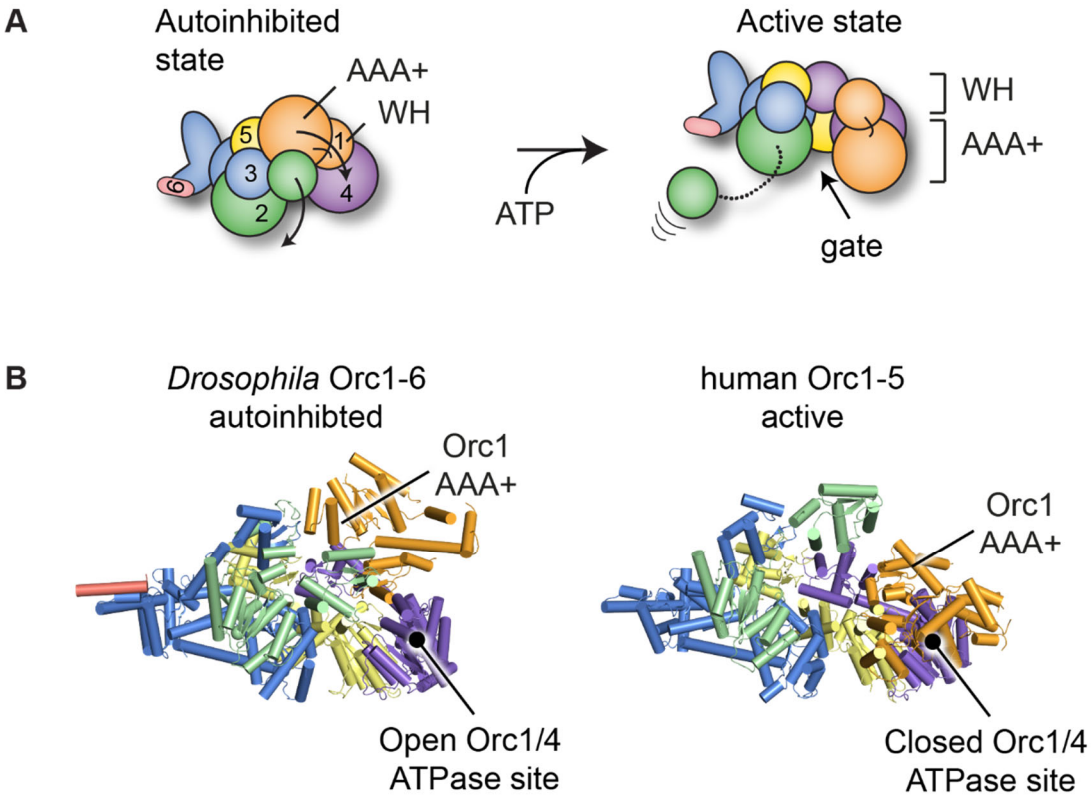


Figure 9: Eukaryotic ORC can assume an autoinhibited and an active conformation. Illustration of the autoinhibited and active conformational state of eukaryotic ORC. **A)** In the absence of nucleotides, the composite ATPase center of Orc1/4 is disengaged due to an out-of-plane movement of the Orc1-AAA+ domain while the Orc2-WH domain obstructs the gate of the ORC ring. Upon ATP binding, the Orc1-AAA+ fold aligns with the AAA+ tier forming a closed ATPase site with the Orc4-AAA+ domain and the Orc2-WH fold becomes flexible. **B)** Structural representation of autoinhibited *Drosophila* Orc1-6 and of active human Orc1-5 is based on PDB: 4XCG[103] and PDB: 5UJM[104], respectively. This figure has been adapted from Ref. 17.

Coupling DNA replication initiation to the cell cycle

Cells go through iterative phases of genome duplication and segregation in the cell cycle. Transition from one cell cycle phase to the other is tightly controlled by the activity of serine/threonine cyclin-dependent kinases (CDKs) that are specified via accessory cyclin (Cl) co-factors and transfer phosphate groups to target proteins to modulate their functionalities[193]. The activity of different CDKs oscillates with the cell cycle and is mainly determined by an interplay of cell-cycle-dependent cyclin expression and degradation[15]. All cyclins are marked for degradation by the proteasome by ubiquitin-targeting, either by the Skp, Cullin, F-box containing complex (SCF) ubiquitin ligase[194] or by the E3 ubiquitin ligase anaphase-promoting complex/cyclosome (APC/C)[195]. The APC/C is activated in prophase of mitosis by M-CDK-mediated phosphorylation[196–198] and continues to operate until G1/S-transition to maintain degradation of S- and M-phase cyclins[196,199–202]. However, in G1 phase, G-CDK activity rises resulting in the inhibition of APC/C[199,200]. This allows the accumulation of S-phase cyclins, which activate licensed origins and further promote the shutdown of APC/C activity. The interdependence of CDK and APC/C results in discrete time spans for high CDK and APC/C activity during the cell cycle. This oscillation is enforced by additional layers of regulation that are highly interconnected and include cyclin inhibitors, phosphatases and transcription factors[14].

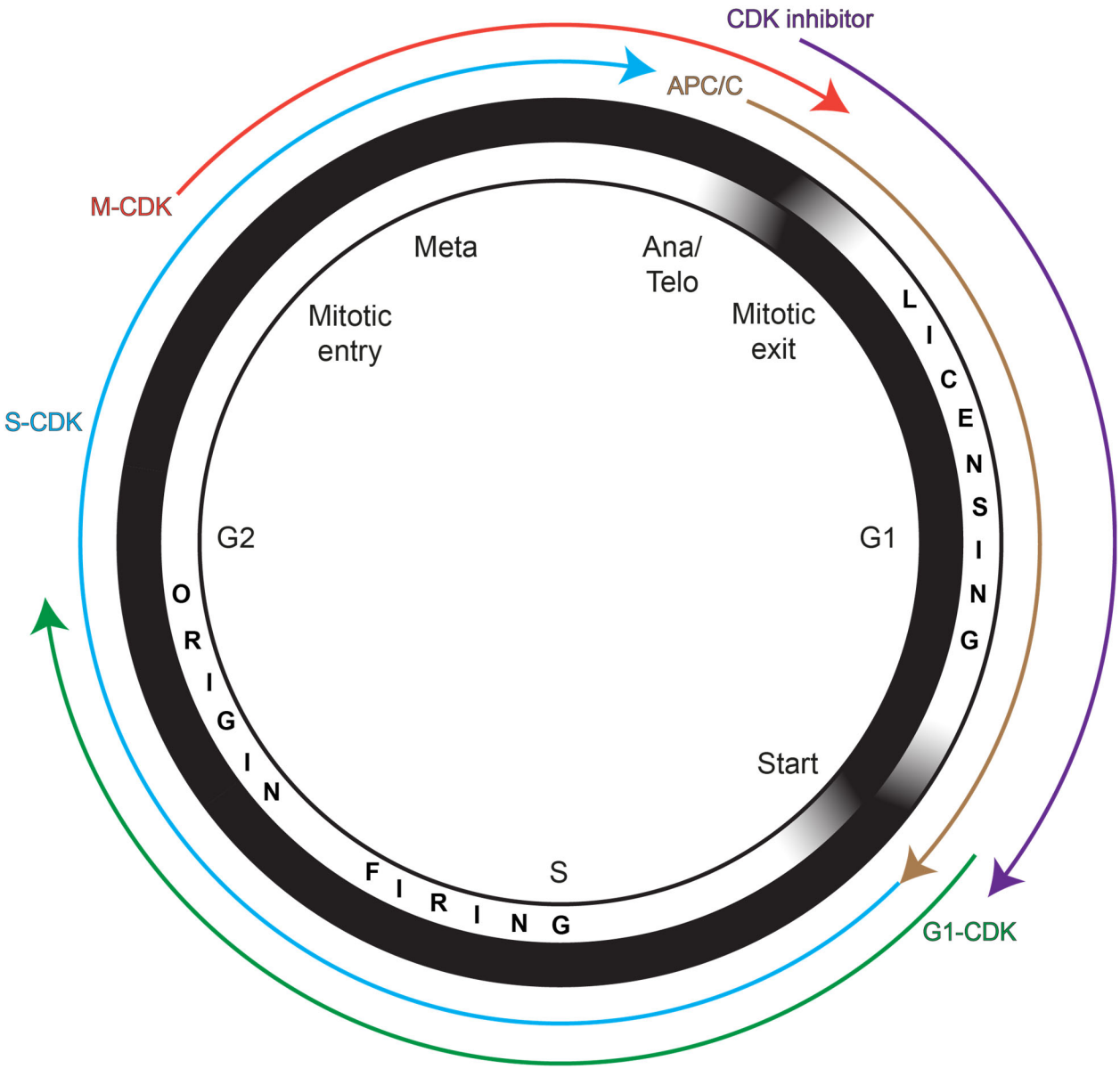


Figure 10: CDK activity during the cell cycle regulates the timing of origin licensing and activation. Illustration of relative CDK, CDK inhibitor and APC/C activity over the cell cycle. In mitosis, S-CDK and M-CDK levels are high. In order for origin licensing to occur at mitotic exit and G1 phase, CDK activity is downregulated by high levels of CDK inhibitor and APC/C. In late G1 and S phase, G1- and S-CDK activity rise, which prevents further helicase loading and activates the licensed origins. Origin firing is stopped by increasing activity of M-CDK, which mark the mitotic entry. This figure has been adapted from Ref. 15.

To ensure that DNA is replicated precisely once per cell cycle origin licensing must be restricted to late M and G1 phase, during which S-CDK activity is low and APC/C activity is high, while origin firing can only occur in S phase, when S-CDK activity is high and APC/C activity is low (Figure 10)[15,16]. Notably, S-CDK fulfills a dual role as inhibitor for origin licensing and activator for origin firing in S phase, which helps to establish the strict separation of both processes. In budding yeast, origin licensing is limited to late M and G1 phase by the simultaneous inhibition of the initiation factors ORC, Cdc6 and Cdt1 by S-CDK creating multiple, redundant pathways that block re-licensing and re-replication[203–205].

Budding yeast Cdc6 is a key target of CDK-dependent regulation while Cdc6 itself inhibits M-CDK function. In detail, Cdc6 possesses an N- and C-terminal phospho-degron and is labeled for degradation in late G1 / early S phase[206–209]. Interestingly, the N-terminus of Cdc6 has been shown to associate with M-CDK, which leads to reduced kinase activity at the end of mitosis[207,210,211]. In addition, binding of the M-phase cyclin Clb2 to the N-terminus of Cdc6 has been demonstrated to obstruct ORC-Cdc6 complex formation and might inhibit origin licensing *in vivo*[211,212]. Consequently, Cdc6 regulated inhibition of M-CDK during mitotic exit and its own degradation at G1/S transition by S-CDK defines a period for when helicase loading can occur[208,212,213].

In budding yeast, Orc1, Orc2 and Orc6 are also subject to CDK-dependent regulation[203,204,214,215]. S-phase CDK promotes hyper-phosphorylation of the N-terminus of budding yeast Orc2 and a linker region in Orc6 in S phase[203,204,210]. Although phosphorylation of *S. cerevisiae* ORC does not seem to affect helicase recruitment to the initiator[145], it has been shown to inhibit helicase loading *in vitro* and helps prevent re-replication *in vivo*[203,204,215]. Notably, Orc6 is the primary CDK-target in the budding yeast initiator complex and is known to cooperate with the co-loader Cdt1 to license origins[106,215]. Interestingly, S-CDK seems to inhibit Orc6 function during helicase loading by two distinct mechanism: a) binding of the S-phase cyclin Clb5 to Orc6 sterically blocks Orc6 interaction with Cdt1-Mcm2-7 and b) phosphorylation of Orc6 hinders the association with Cdt1[215]. In contrast to Orc6, phosphorylation of Orc2 has little effect on helicase loading on its own, but enhances re-replication when Orc6 is deregulated[215].

However, how phosphorylation of Orc2 interferes with helicase loading is still not well understood. In *S. cerevisiae*, Orc1 is only weakly phosphorylated by S-CDK *in vitro* and the role of Orc1 phosphorylation is largely unclear[215].

CDK-mediated regulation of ORC activity is functionally conserved from budding yeast to metazoan, albeit through different mechanisms[14,15]. In *Drosophila*, it has been shown that the N-terminal disordered region in Orc1 is phosphorylated *in vitro* and *in vivo*[214]. Phosphorylation of *Drosophila* Orc1 and Orc2 reduces nucleotide-dependent ORC-DNA binding and inhibits ORC's ATPase activity suggesting that CDK-mediated phosphorylation of Orc1 is an important regulator of *Drosophila* ORC function[214,216]. Similarly, CDK-dependent phosphorylation of human Orc1 inhibits Orc1 binding to chromatin[217]. It is worth noting, that, in contrast to budding yeast, *Drosophila* Orc6 does not contain CDK consensus sites suggesting that, although budding yeast and *Drosophila* ORC are both phosphorylated to modulate their functions, the exact regulatory mechanisms might have diverged significantly.

Origin licensing is further controlled by restricting Mcm2-7 and Cdt1 activity to G1 phase. In budding yeast, the co-loader Cdt1 is stably bound to Mcm2-7 and hence both can be regarded as one entity with respect to regulation[138]. Cellular localization of Mcm2-7 oscillates with the cell cycle and is controlled by a transport module formed by two partial nuclear localization sequences (NLS) in Mcm2 and Mcm3, which function in cis, and a nuclear export signal (NES) in Mcm3[218]. In G1 phase, Mcm2-7 protein levels concentrate in the nucleus and decrease in late G1 due to net nuclear export[219,220]. The balance between nuclear import and export by the Mcm2-3 transport unit is skewed towards nuclear export when Mcm3 is phosphorylated by M-CDK leading to the exclusion of soluble Mcm2-7/Cdt1 from the nucleus, thereby preventing origin licensing[218]. In contrast to the nuclear export mechanism of Mcm2-7/Cdt1 in budding yeast, fission yeast and metazoan Cdt1 is degraded by the E3 ubiquitin ligase containing Cul4 pathway (Crl4) in S phase[115,221–224]. In addition, metazoans have evolved a Cdt1-inhibitor protein termed Geminin that blocks Cdt1-Mcm2-7 interaction and is itself controlled by cell-cycle dependent degradation through APC/C[225–230]. Although inactivation of Geminin-mediated inhibition of Cdt1 results in re-replication in *Drosophila* and some human cell lines, Cdt1

degradation seems to be the primary pathway to control Cdt1 activity[222]. In metazoans, both mechanism act as redundant pathways to prevent genome over-replication[222].

In summary, eukaryotic DNA replication initiation is controlled by an intricate network of redundant cell-cycle dependent pathways to ensure robustness of the system. Genome size and complexity, along with number of origins, all are expected to correlate with regulatory mechanisms[14]. Consequently, the overall principle of CDK-mediated regulation of DNA replication initiation is conserved in eukaryotes, though the specific network factor and pathways differ considerably[14].

The role of DNA replication proteins in human development – Meier-Gorlin Syndrome

Proper function of initiation and DNA replication factors is pivotal for human development. Mutations in initiation factors can cause severe genetic diseases like dwarfism[2]. Dwarfism disorders are characterized by impaired growth and tissue development, which among other factors may be caused by defects in DNA replication and deregulated cell cycle progression resulting in an overall reduced cell number[2]. One such dwarfism disorder is the Meier-Gorlin Syndrome (MGS); a rare, autosomal and recessive disease that has been first described by Meier and Gorlin in 1959 and 1975, respectively[231,232]. Individuals with MGS show a broad spectrum of phenotypic features that are also observed in other primordial dwarfism diseases, but small stature in combination with microtia and small or absent patellae are the most prevalent characteristics suggesting that specifically bone and cartilage tissue development is compromised[232,233].

In the last decade it has become increasingly clear that the Meier-Gorlin Syndrome is linked to mutations in initiation and DNA replication factors. Approximately 67-78% of patients with MGS have mutations in either Orc1, Orc4, Orc6, Cdt1, or Cdc6[6]. MGS missense mutations in ORC and Cdc6 are primarily found in the functionally important Cdc6/Orc1 and Orc1/Orc4 ATPase site as well as in the Orc1-BAH domain (Figure 11)[234–236]. These mutations are hypomorphs and hinder cell cycle progression resulting in impaired development[7]. However, detailed characterization of MGS-associated mutations through comprehensive *in vitro* and *in vivo* studies are needed to understand fully how specific mutations in licensing and firing factors result in the Meier-Gorlin Syndrome phenotype. Further scientific efforts are needed to elucidate why cartilage and bone tissue development seems to be mainly affected in MGS patients.

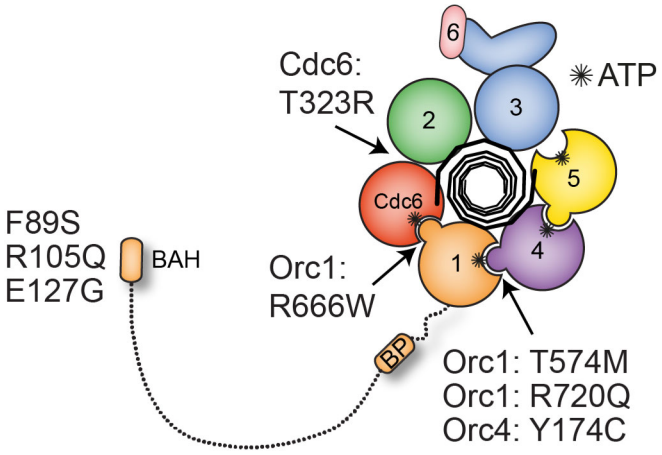


Figure 11: MGS-associated mutations in the human ORC-DNA-Cdc6 complex. Cartoon representation of the AAA+ domain tier of the ORC-DNA-Cdc6 complex and the Orc1-BAH domain. The location of MGS mutations in the ODC complex are indicated. Note that mutations cluster in the Orc1/4 composite ATPase site and the Orc1-BAH domain.

The goals of my thesis

DNA replication is largely controlled at the level of initiation. Thus, to understand how DNA replication is controlled in space and time it is pivotal to investigate how ORC and Cdc6 function.

The goals of my thesis were to:

Goal 1: Elucidate how metazoan ORC recognizes DNA in a sequence-independent manner and investigate how DNA shape contributes to ORC binding and replicative helicase loading.

Although the location of origins in metazoan is not random, metazoan origins are not defined by a DNA consensus sequence and metazoan ORC is known to bind DNA in a sequence-independent manner[9,13]. However, it is unclear how the metazoan initiator recognizes DNA without a guiding DNA consensus motif. To investigate how metazoan ORC recognizes DNA, we employed single-particle cryo-EM to determine high-resolution structures of *Drosophila* ORC bound to DNA in presence or absence of Cdc6.

Furthermore, eukaryotic ORC has been shown to bend the DNA substrate upon binding and it has been speculated that DNA shape may affect initiator binding and subsequent helicase loading[52,79,90]. To test these hypotheses, we probed ORC binding to DNA sequences with different AT-content *in vitro*. We also established an *in vitro* metazoan helicase loading assay to measure helicase loading efficiency by ORC on different DNA substrates, and we determined cryo-EM density maps of ORC/Cdc6 bound to DNAs with varying AT-content.

Goal 2: Explore the physical basis by which the budding yeast ORC-DNA-Cdc6 complex is regulated.

CDK-dependent phosphorylation of budding yeast ORC is an important regulator of initiator function and inhibits re-replication *in vivo*[203,204,215]. Furthermore, ORC activity is tightly controlled by nucleotide binding and hydrolysis[17]. Despite their importance for regulating origin licensing, the structural mechanisms responsible for CDK-dependent inhibition of ORC function and for coordinated ATPase hydrolysis by ORC/Cdc6 are not well understood. We used cryo-EM to obtain high-resolution structural data on the ternary *S. cerevisiae* ORC-DNA-Cdc6 complex. Our structural work unveiled a potentially inactive conformational state of the ODC complex that is likely linked to ORC's phosphorylation status, suggesting a physical basis for how phosphorylation of ORC inhibits replicative helicase loading. Moreover, we propose a physical explanation for how ORC/Cdc6 activity is regulated.

Goal 3: Define how mutations in ORC and Cdc6 that have been associated with MGS affect protein function in vitro.

ORC and Cdc6 are frequently mutated in MGS patients, but little is known how these mutations affect initiator and co-loader function[6,7]. To learn more about the consequences of MGS-associated mutations in ORC and Cdc6, we used biochemical tools to probe MGS mutant initiator and co-loader function *in vitro*.

Together, my results have advanced our understanding of how origins are selected and remodeled by metazoan ORC to support efficient helicase loading. Furthermore, our findings, provided insight into the structural mechanisms underlying conformational rearrangements of ORC/Cdc6 and how they help regulate initiator activity. Our understanding of initiator and co-loader function was further aided by the biochemical interrogation of ORC and Cdc6 harboring MGS-associated mutations. Finally, the established ATPase activity assay for metazoan ORC/Cdc6 and the *in vitro* metazoan helicase loading assay will enable the research community to address how origin licensing is controlled in metazoans.

Results

Goal 1: Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex.

Jan Marten Schmidt and Franziska Bleichert

Nat. Commun. 11, 4263 (2020). https://doi.org/10.1038/s41467-020-18067-7

The published article including supplementary materials can be found in the appendix.

Summary

The metazoan origin recognition complex selects replication origins in the absence of a DNA consensus sequence and has been shown to bend DNA, which likely contributes to replicative helicase loading in eukaryotes[9]. In our study, we determined cryo-EM structures of *Drosophila* ORC bound to DNA with or without Cdc6 (Figure 12) and probed binding of *Dm*ORC to different DNA sequences *in vitro*. Our work revealed that the AAA+ domains of *Drosophila* ORC use multiple motifs to contact DNA and organize in a collar around the DNA, which likely helps to stabilize the initiator and co-factor on duplex DNA. Notably, Orc1 and Orc4 contact the DNA more extensively than the other subunits in the ORC-DNA-Cdc6 complex and the WH domain of Orc3 might contribute to DNA bending. Furthermore our structures showed that a loop element in proximity to the Walker B motif of Orc1 binds the DNA minor groove. This arrangement functionally links the Orc1/Orc4 ATPase center to DNA binding, which we confirmed by biochemical assays. We found that *Dm*ORC preferentially binds to AT-rich DNA sequences and associates tightly with a poly(dA:dT) homomer *in vitro*. We suggest that *Dm*ORC binding affinity to a DNA sequence correlates with minor groove widths and negative electrostatic potential. In addition, we provided structural evidence that ORC DNA binding can be separated from ORC-induced DNA bending. In securing these

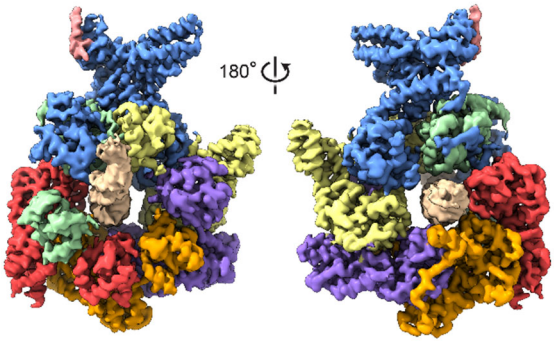


Figure 12: Cryo-EM density map of the ternary *Drosophila* ORC-DNA-Cdc6 complex colored by subunits. Figure has been adapted from Ref. 244

results, we established an *in vitro* replicative helicase loading assay using *Drosophila* proteins and demonstrated that DNA bending by ORC promotes replicative helicase loading.

My contribution

This work is a shared effort by myself and my thesis supervisor Franziska Bleichert. For Franziska Bleichert's contribution to this study I refer to the author information section of the article below. We received technical support for cryo-EM data collection from Simone Cavadini and Andreas Schenk. Sandra Muehlhaeusser maintained our cell culture stocks and Federica Galliano contributed to cloning a Cdc6 expression construct. Throughout my Ph.D., my contributions included development and execution of biochemical assays, design and cloning of protein expression constructs, protein expression and purification, negative-stain and cryo-EM sample preparation and data collection, EM data processing, data analysis and interpretation and editing of the manuscript.

Goal 2: Structure of a loading competent S. cerevisiae ORC-Cdc6 complex bound to origin DNA.

Jan Marten Schmidt, Jan Seebacher and Franziska Bleichert

The preliminary manuscript can be found in the appendix.

Summary

The initiator ORC associates with the co-factor Cdc6 in an ATP-dependent manner to recruit and load the replicative helicase Mcm2-7 onto origins in eukaryotes[17]. CDK-dependent phosphorylation of ORC is an important mechanism of inhibiting Mcm2-7 loading and helps prevent re-replication in budding yeast[106,203,204,215]. However, the structural mechanism underlying CDK-dependent inhibition of ORC activity are not understood. In addition, the physical basis for sequential ATP hydrolysis of the initiator and its co-loader, which is essential *in vivo*, remains elusive [98,108–110,116,191,192]. In our study, we determined cryo-EM structures of the binary and ternary budding yeast ORC-DNA(-Cdc6) complex that extend our understanding of how budding yeast ORC recognizes different origin DNA sequences. We showed that the ODC complex can adopt two distinct conformations. Furthermore, we proposed a structural mechanism for sequential ATPase activity of the Cdc6/Orc1 and Orc1/Orc4 ATPase centers. Our structural data

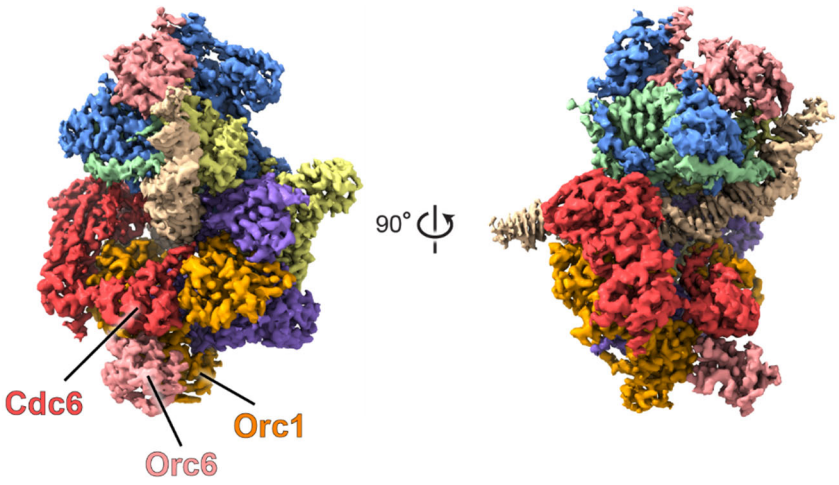


Figure 13: Cryo-EM density map of the budding yeast ORC-DNA-Cdc6 complex revealing an Orc6/Cdc6/Orc1 interface. The cryo-EM density is colored by subunit and the densities corresponding to Cdc6, Orc6 and Orc1 are indicated. This figure has been adapted from Ref. 262.

revealed binding of the Orc6 N-terminus to Cdc6 and Orc1, which may depend on phosphorylation of ORC and may aid the regulation of helicase loading in *S. cerevisiae* (Figure 13). We confirmed our discovery of the Orc6/Cdc6/Orc1 interface by cross-linking mass spectrometry.

My contribution

This project is work of myself, Jan Seebacher and Franziska Bleichert. For details on the contributions of Jan Seebacher and Franziska Bleichert, please see the author contribution section in the attached manuscript. Andreas Schenk and Simone Cavadini provided technical assistance for cryo-EM data collection. Sandra Muehlhaeusser managed and maintained our cell culture. I contributed as follows: I expressed and purified proteins, performed cross-linking mass spectrometry experiments, prepared samples for cryo-EM and collected cryo-EM data. Together with FB, I processed cryo-EM data, analyzed, interpreted and visualized data and wrote the manuscript.

Goal 3: Probing the functional consequences of Meier-Gorlin Syndrome mutations in a metazoan initiator and its co-loader.

Jan Marten Schmidt, Federica Galliano and Franziska Bleichert

This project is ongoing. A detailed presentation of the current status can be found in the appendix.

Summary

Meier-Gorlin Syndrome is a rare, genetic disorder that mainly affects the development of bone and cartilage tissue and leads to a dwarfism phenotype in humans[7]. Modern whole-genome sequence techniques have unveiled that hypomorphic mutations in ORC and Cdc6 play an important role in the majority of MGS cases. Notably, MGS-associated missense mutations frequently map to the composite ATPase sites within the ORC-DNA-Cdc6 complex[6] (Figure 14). In our study we sought to understand how these mutations impact replicative helicase loading and how nucleotide binding and hydrolysis by ORC is coupled to downstream initiation events. Accordingly, we probed mutant *Drosophila* ORC DNA binding, ATP hydrolysis and ORC-DNA-Cdc6 complex assembly by biochemical assays *in vitro*.

My contribution

This study is a joint effort of myself (JS), Federica Galliano (FG) and Franziska Bleichert (FB). FB conceptualized the study and supervised the project. JS, FB and FG designed and cloned protein expression constructs. JS, FB, and FG purified proteins. Sandra Muehlhaeusser managed and maintained our cell culture stocks. JS performed biochemical (DNA binding, ATPase activity, pull-down) assays. With the help of FB, JS analyzed, interpreted and visualized the data.

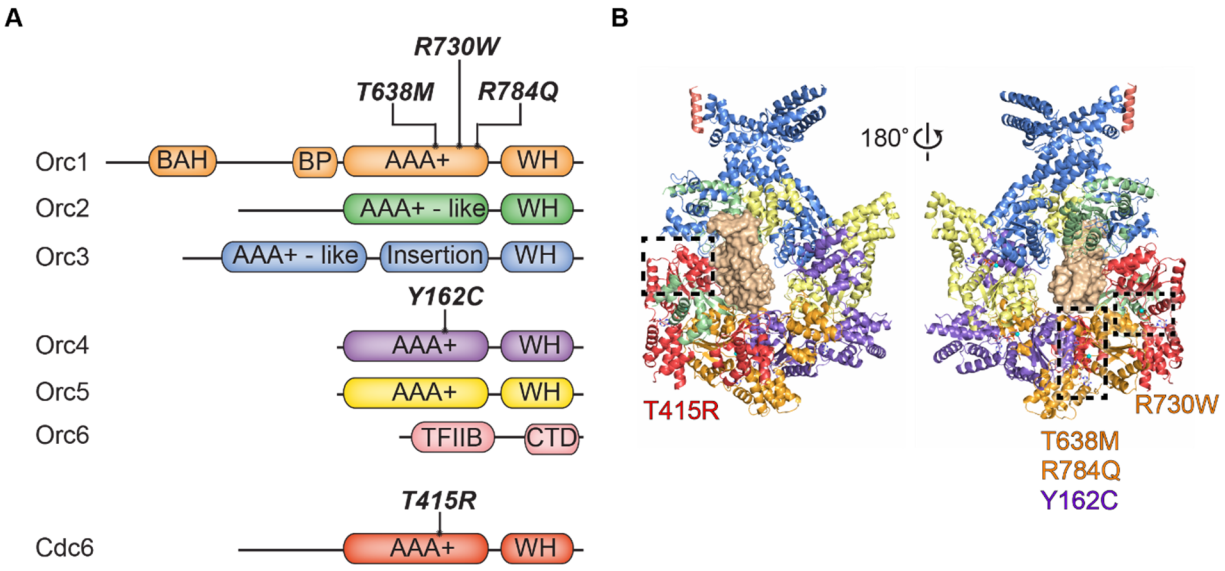


Figure 14: Localization of MGS-associated missense mutations in ORC and Cdc6 in the context of the ternary *Drosophila* ORC-DNA-Cdc6 assembly. **A**) Domain organization of ORC and Cdc6. Position of MGS-associated mutations is shown. Mutations correspond to residues in *Drosophila* ORC and Cdc6. **B**) View at the WH-tier and AAA+-tier of *Drosophila* ODC (PDB: 7JK4)[244]. Boxed regions indicate location of MGS mutation(s) within in the ternary complex. This figure has been adapted from Ref. 269.

Discussion and Outlook

Origin licensing is mediated by numerous initiation factor assemblies that undergo diverse conformational changes[11,17]. To understand how these events are coordinated and controlled, it is essential to learn more about how physical features shape initiation factor function and how these intrinsic properties determine the conformational dynamics of protein complexes during origin licensing. Our biochemical and structural studies on *Drosophila* and *S. cerevisiae* ORC and its co-loader Cdc6 provide comprehensive mechanistic descriptions of initiator function, conformational dynamics and regulation. Moreover, our work contributes to our understanding of how missense mutations identified in Meier-Gorlin Syndrome patients may compromise DNA replication initiation in humans. The *in vitro* replicative helicase loading assay presented here is a powerful tool that will enable the research community to address further how origin selection and licensing is controlled in metazoans.

The autoinhibited state of the Origin Recognition Complex

From yeast to man, ORC is known to adopt an autoinhibited (closed) conformation in the absence of nucleotide *in vitro*[52,90,103,237]. In the autoinhibited form, the AAA+ domain of Orc1 and the Orc2-WH domain sterically occlude ORC's central channel, preventing DNA to enter the ORC ring[103]. Therefore, it has been suggested that the autoinhibited state of ORC might be important for regulating ORC activity *in vivo*[103]. Upon ATP binding, eukaryotic ORC can transition to an active (open) conformation that can associate with origin DNA[50–52,65,66,90,103,104,107]. However, the presence of nucleotide alone is insufficient to force all ORC molecules into an active conformation *in vitro* suggesting that unknown co-factors may aid the conformational switch *in vivo*[103]. The role and prevalence of the autoinhibited state *in vivo* is unclear, but it may serve as a dormant state that is rapidly available if the demand for DNA replication increases. For instance, large quantities of ORC are stored in oocytes to facilitate fast cell divisions during early embryogenesis[238]. However, DNA replication must not occur before fertilization. Consequently, ORC might be stockpiled in oocytes in the autoinhibited state[103]. Furthermore, it can be speculated

that the autoinhibited state might be an intermediate of a potential reactivation cycle of ORC, which upon ATP hydrolysis likely detaches from DNA and needs to exchange the bound ADP with ATP before it can participate in another round of helicase loading[17,108,155]. Development of imaging techniques that allow tracking of different conformational states of ORC with high temporal-resolution *in vivo* are required to address further the role of the autoinhibited state of ORC during development and repeated rounds of origin licensing. Alternatively, investigating the functional consequence of ORC mutants that are locked in the active form - for instance by artificially stabilizing the Orc1/Orc4 interface - might help to shed light on the functional relevance of the autoinhibited state during helicase loading.

The role of ATP in ORC assembly

It is worth noting that the role of ATP binding for the structural integrity of ORC may have evolved from yeast to humans. In contrast to *S. cerevisiae* and *Drosophila* ORC, it has been reported that ATP binding is required for human ORC assembly suggesting that human ORC is intrinsically unstable, which may be important for complex disintegration in S phase[239,240]. However, human Orc1-5 can also be purified in the absence of nucleotides, indicating that ATP is not required for complex integrity[90]. On the other hand, it has been proposed that human Orc1 is the first ORC subunit that associates with DNA independent of Orc2-5, suggesting that human Orc1 dynamically associates with Orc2-5[241]. In support of this finding, recent structural work on the human Orc1-5 complex has revealed that Orc2-5 form a stable sub-complex that adopts a closed, likely inactive, conformation, which requires Orc1 to become activated[237]. Interestingly, protein levels of human Orc2-5 remain constant over the cell cycle while Orc1 levels are elevated in G1 phase and decline in S phase due to proteasome-dependent degradation[242,243]. Consequently, Jaremko et al.'s finding[237] that Orc2-5 form a stable, yet likely inactive sub-complex, provides a structural explanation why only Orc1 protein levels oscillate during the cell cycle and Orc2-5 protein abundance remains constant. However, it is unclear whether stable Orc1-5 complex formation is solely controlled by protein abundance or whether other mechanisms such as ancillary proteins help regulate human Orc1-5

assembly and thereby activity *in vivo*. Future studies are needed to better understand the dynamic assembly and disassembly of human ORC over the cell cycle and how potential co-factors may regulate human ORC integrity.

Modes of origin recognition: sequence-specific vs. context-dependent

It is well-established that DNA recognition by ORC is different among eukaryotes, being primarily sequence-specific in *S. cerevisiae*, and context-dependent in almost all other eukaryotes[9]. The physical basis for the sequence-specific origin recognition by budding yeast ORC has been revealed by a high-resolution structure of *S. cerevisiae* ORC bound to origin DNA[52]. Our structural studies addressed the long-standing question how the metazoan initiator engages DNA in a sequence-independent manner[244]. Despite the contrasting modes of DNA recognition, our studies unveiled basic features of ORC DNA binding that are conserved in eukaryotes. In budding yeast and *Drosophila*, the ISM motifs in the AAA+ domains of ORC form a ring that encircles DNA, holding the initiator on DNA[52,103,244]. Moreover, both eukaryotes use the WH domain tier as a DNA binding unit, which likely is required for ORC induced DNA bending[52,103,244].

The shared characteristics of DNA binding among eukaryotic ORCs stand in contrast to the multiple differences in structural elements that contact DNA and mediate the species-specific DNA recognition. The basic-patch in Orc1, which is crucial for DNA binding in budding yeast and *Drosophila*[90,128], contacts DNA in our *Drosophila* ORC-DNA structure, but binds DNA much more extensively in budding yeast, where it contributes to the sequence-specific DNA read out[52,244]. By contrast, the basic-patch of *Drosophila* Orc1 re-positions the ISM of Orc4 forming a structural platform that may promote tight DNA binding[244]. In addition to the Orc1-BP, a loop in the Orc2-ISM and a yeast-specific helix-insertion in the WHD of Orc4 mediate the recognition of the ACS by *S. cerevisiae* ORC (Figure 5)[52,129,244]. Instead of an insertion helix, *Schizosaccharomyces pombe* Orc4 possesses N-terminal AT-hook elements that specifically bind to AT-rich DNA and are essential for origin recognition by *Sp*ORC[245–248]. *Drosophila* and human ORC lack the helix-insertion in the WHD of Orc4 and the ISM of *Drosophila* Orc2 makes only minimal

contacts with DNA explaining the different sequence-specificities for DNA recognition of budding yeast, *Drosophila* and human ORC[103,104,244].

In contrast to budding yeast, *Drosophila* ORC binds DNA primarily through Orc1 and Orc4, which results in an off-center position of the bound duplex in the channel of the ORC ring[244]. *Drosophila* ORC binding DNA mainly through Orc1 and Orc4 may strengthen the link between substrate recognition and ATPase hydrolysis in the Orc1/Orc4 interface. Like *Drosophila* ORC, a recently published structure of human Orc1-5 bound to endogenous DNA suggest that human ORC binds DNA primarily through Orc1, which is in agreement with the observation that human Orc1 associates with chromatin independent of Orc2-5[237,241]. However, the density assigned to DNA in the human OD complex is not well resolved, rendering the comparison with the DNA binding mode of *Drosophila* ORC difficult. High-resolution structural studies of DNA-bound human ORC are required to elucidate the exact mechanisms by which human ORC recognizes DNA.

The Orc1 B-loop motif couples DNA binding to ATPase activity in Drosophila ORC

Timely ATPase hydrolysis by ORC is essential for cell survival and it has been long known that DNA binding to ORC reduces ORC's ATPase activity, likely helping to stabilize the complex on DNA[51,66,107,108]. However, it remained enigmatic how the DNA binding status is sensed and transmitted to down-regulate ATP hydrolysis by ORC. Intriguingly, we discovered that a conserved loop element (B-loop) in physical proximity to the Walker B motif of *Drosophila* Orc1 inserts into the DNA minor groove and couples DNA substrate binding to ORC's ATPase activity[244]. The B-loop of human Orc1 is also in proximity to DNA in the cryo-EM structure of human Orc1-5 bound to endogenous DNA[237]. Although the Orc1 B-loop does not directly contact the DNA duplex in the *Sc*ORC-DNA complex, DNA binding to budding yeast ORC may nevertheless indirectly reposition the Orc1 B-loop, which may affect ATPase hydrolysis[244]. High-resolution structural studies of active *Sc*ORC without DNA and DNA-bound human ORC will be required to better understand the dynamics of the B-loop and its role in regulating DNA binding and ATPase activity of the initiator. Remarkably, we found that a similar loop region is also used by

bacteriophage T4 clamp loader and *Escherichia coli* DnaA and DnaC to contact DNA, suggesting that the B-loop might be a conserved DNA binding element in diverse initiator and loader proteins[244].

DNA geometry is an important regulator of helicase loading efficiency in Drosophila

S. cerevisiae and *Drosophila* ORC have been shown to induce a bend into the DNA substrate and it has been postulated that DNA bending is required for helicase loading[52,90]. In *Drosophila*, DNA bending by ORC is likely mediated by coordinated binding of the Orc1 B-loop and the WH domain of Orc3 that together stabilize the DNA curvature[52,90,244]. Although the Orc3-WH domain engages the bent DNA in budding yeast, *Sc*ORC also uses basic regions in Orc2 and Orc5 and the C-terminal cyclin-box fold of Orc6 to bend the DNA much more extensively than it is seen for metazoan ORC[52]. We could not observe these contacts in our structure of the *Dm*ORC-DNA complex, and they may be transient[244].

Our *in vitro* DNA binding assays confirmed previous observations that metazoan ORC preferentially binds to AT-rich DNA and suggest that metazoan ORC may favor DNA sequences with certain thermodynamic properties[64–66,79,100,244]. In accord with these findings, replication start sites are AT-rich in *Drosophila* and it has been proposed that geometric DNA characteristics are important determinants for origin activity in *Drosophila* and human[79,249,250]. Analysis of DNA sequence properties tested in our DNA binding assay unveiled that binding affinity of *Drosophila* ORC to a DNA sequence correlates with minor groove width and in particular negative electrostatic potential *in vitro*[244]. In *Drosophila*, the Orc1 B-loop tightly inserts into the DNA minor groove, therefore we propose that it may function as a reader for DNA shape[244]. Comprehensive mutational analysis of the Orc1 B-loop will be needed to see if the Orc1 B-loop is involved in modulating ORC's affinity for different DNA shapes in *Drosophila*. Furthermore, it will be important to explore whether such a potential read out mechanism for DNA shape is also employed by other metazoans.

The role of polyA tracts in defining the location and activity of replication origins

Intriguingly, our *in vitro* DNA binding studies revealed a particularly high affinity of ORC for polyA duplexes[244]. Stretches of polyA homomers are numerous in eukaryotic genomes and can be found in proximity of replication origins in yeast, *Drosophila* and human, where they correlate with origin activity[79,85,250–260]. The correlation of polyA tracts with the location and activity of replication origins might be a consequence of the nucleosome-disfavoring properties of polyA stretches as nucleosome-depleted regions are the main commonality of origins in eukaryotes[13,62,252]. In agreement with previous studies[68,100,250,257], our work argues that polyA tracts not only create nucleosome-free genomic regions, but also serve as high affinity binding sites for *Drosophila* ORC, and may facilitate ORC recruitment to origins[244].

The high intrinsic stiffness of polyA tracts may be unfavorable for DNA bending by ORC, which has been suggested to be important for replicative helicase loading[52,90,252]. Strikingly, our ORC/Cdc6-polyA structure showed that polyA DNA largely resists deformation by ORC and helicase loading efficiency by ORC was markedly reduced on a polyA DNA compared to an AT-rich, bendable substrate *in vitro*[244]. Hence, we suggest that polyA tracts are likely not direct sites of replicative helicase loading *in vivo*.

Taken together, our studies support the idea that metazoan ORC reads out intrinsic geometric DNA features and that DNA malleability is crucial for efficient replicative helicase loading[244]. In accordance with Comoglio et al., we propose that in addition to being nucleosome-free regions metazoan origins are likely defined by high intrinsic deformability instead of a specific DNA sequence or nucleotide-content[79]. This would reconcile observations that origins in *Drosophila* are AT-rich while origins in mouse and human are generally GC-rich[69,72,77,79,249,250].

Currently available data on genomic ORC binding and replication initiation sites is limited to a resolution of approx. 300bp while 40bp of DNA are sufficient for high-affinity, nucleotide-dependent metazoan ORC binding *in vitro*[13,82,90]. Thus, the resolution of mapped origins and ORC binding sites in the genome does not permit direct conclusions about how mapped DNA sequences may determine ORC binding. In the future, high-

resolution mapping of ORC genomic binding sites will be required to further dissect how DNA geometries may influence origin selection in higher eukaryotes.

Structural insight into ORC-DNA-Cdc6 complex formation and function

Upon DNA-binding ORC can recruit Cdc6, which stabilizes the initiator on DNA and results in the formation of a docking platform for Mcm2-7 recruitment and loading[17]. Despite being essential for origin licensing a structure of the isolated ORC-DNA-Cdc6 evaded structural investigations for a long time. Moreover, the physical architecture of the eukaryotic co-loader itself had only been assumed based on an archaeal homologue[105,261]. Our structural work, provides the first high-resolution view of the isolated ternary budding yeast and *Drosophila* ORC/Cdc6 complex bound to different DNA substrates[244,262]. Cdc6 recruitment to DNA-bound ORC did not induce significant structural rearrangements and ORC-DNA contacts remained constant in *Drosophila* and budding yeast[244,262]. Contrary, the basic patches of Orc2 and Orc5 as well as the C-terminal cyclin-box fold (CBF_C) of Orc6 are positioned farther away from DNA in the ORC-DNA-Cdc6 assembly in context of the OCCM intermediate suggesting that these elements do not contact the DNA anymore[91,105]. In addition, the Orc1 basic patch that meanders through the DNA minor groove in the budding yeast OD complex remains DNA-bound in our structures of the *Sc*ODC assembly, but is flexible in the OCCM intermediate[52,91,105,262]. Hence, association of Mcm2-7 with ODC seems to trigger the repositioning of the Orc2-BP, Orc5-BP and Orc6- CBF_C as well as the detachment of the Orc1-BP, which together may loosen ORC's grip on DNA allowing more conformational freedom. This greater conformational flexibility might be required for stable association of Mcm2-7 to the ORC/Cdc6 ring and the step-wise insertion of the DNA into the core of the helicase. However, how association of Mcm2-7 to ODC induces these changes and whether the underlying mechanisms can be translated to metazoans is unknown. Future studies are required to further address how Mcm2-7 recruitment leads to structural rearrangements in the ODC complex and how these alterations support efficient helicase loading.

The winged helix domain of Orc2 is flexible in the active and DNA-bound budding yeast, *Drosophila* and human ORC complex[52,237,244]. Upon Cdc6 recruitment, the Orc2-WHD anchors onto the AAA+ domain of Cdc6 forming a binding site for the Mcm3-WHD, which mediates the initial association of Mcm2-7 to ORC-DNA-Cdc6[91,145]. Thus, Orc2-WHD binding to Cdc6 likely helps to regulate helicase recruitment. Strikingly, we found that stable association of the Orc2-WHD with Cdc6 requires pronounced DNA bending by *Drosophila* ORC suggesting a possible read-out mechanisms for DNA curvature during metazoan helicase recruitment[244]. However, in the budding yeast ODC, the WHD of Orc2 remains dynamic despite extensive DNA bending suggesting that the DNA curvature read-out by the WHD of Orc2 might be specific to metazoans[262]. This divergence may be explained by the fact that budding yeast ORC uses additional elements to stabilize the bound DNA in a sharply bent configuration[52]. In addition, the sequence-specific origin recognition by *S. cerevisiae* ensures binding to an AT-rich and flexible substrate. Accordingly, budding yeast ORC likely bends origin DNA more stably and consistently, which makes a mechanism for sensing the DNA configuration unnecessary.

A structural explanation for the coordinated ATP hydrolysis by ORC and Cdc6

Sequential ATP hydrolysis by the Cdc6/Orc1 and Orc1/Orc4 ATPase centers is pivotal for cell viability[94,98,108,109,116,191,192]. However, the mechanisms that control and coordinate nucleotide hydrolysis by both ATPase sites are largely unclear. Our sub-3Å cryo-EM structures of budding yeast OD and ODC provide unprecedented insight into the organization of key catalytic residues in the ATPase sites of the OD and ODC loading intermediates[262]. While we did not observe major differences in side chain orientation of most catalytic residues when comparing the Cdc6/Orc1 and Orc1/Orc4 ATPase centers in the budding yeast ODC complex, the Sensor 1 side chain assumes a nucleotide-facing position at the Cdc6/Orc1 interface but is flipped away from the nucleotide in the Orc1/Orc4 ATPase site[262]. Notably, the Sensor 1 residue is known to play an important role in regulating ATPase activity and has been suggested to help couple ATP hydrolysis to conformational changes in AAA+ proteins[113,263,264]. In the ternary budding yeast ODC, the position of the Sensor 1 seems to be determined by a hydrogen bond network

(referred to as S-TH-D) formed by a Tether residue and a conserved aspartate that is located between the B-loop and Walker B motif[244,262]. The aspartate between the B-loop and the Walker B motif is positioned to communicate the status of the bound DNA to the ATPase sites through the Sensor 1 asparagine[262]. Consequently, we postulate that the orientation of the Sensor 1 residue, and thereby the S-TH-D hydrogen bond network, may help coordinate ORC/Cdc6 ATPase activity[262]. High-resolution structures of DNA-bound human ORC and active yeast and *Drosophila* ORC will be required to further address the physical basis by which DNA binding and Cdc6 recruitment regulate sequential ATPase activity by the initiator and its co-loader across eukaryotes.

Discovery of a potentially phosphorylation-dependent state of the budding yeast ODC complex

Budding yeast Orc6 has four CDK consensus sites and is known to be the primary target of CDK-mediated phosphorylation in *Sc*ORC[203,204,215]. Phosphorylation of *Sc*Orc6 inhibits Orc6-Cdt1 binding and impairs helicase loading[106,215]. However, the underlying molecular mechanisms are unknown. Furthermore, phosphorylation of Orc6 by S-CDK is an important inhibitor of re-replication in budding yeast suggesting that phosphorylated Orc6 interferes with Mcm2-7 binding to ORC-DNA-Cdc6 and/or helicase loading *in vivo*[215]. Yet, the physical basis for how phosphorylation of *Sc*Orc6 leads to inhibition of ORC function is unclear. Our cryo-EM studies on the budding yeast ODC complex revealed that the poorly characterized N-terminal cyclin-box fold (CBF_N) of Orc6 binds an interface formed by the WHD of Cdc6 and the AAA+ domain of Orc1, which is occupied by the WHD of Mcm7 in the OCCM intermediate[262]. Consequently, we suggest that binding of the Orc6-CBF_N to Cdc6/Orc1 blocks stable Mcm2-7 recruitment rendering the ODC complex unproductive for replicative helicase loading[262]. Given the biological context, it is likely that Orc6-CBF_N binding to Cdc6/Orc1 is linked to Orc6 phosphorylation and that this association is incompatible with helicase loading. Consistent with this hypothesis, approximately 50% of the purified ORC protein used for our structural studies contained phosphorylated Orc6[262]. Future studies are required to test whether the Orc6/Cdc6/Orc1 interface is regulated by phosphorylation status of Orc6 and prevents re-

replication *in vivo*. Moreover, it is unclear which and how many of the CDK sites in Orc6 have to be phosphorylated to stabilize the Orc6/Cdc6/Orc1 interface. It is important to note that the binding sites of the C-terminal and N-terminal cyclin-box fold of Orc6 are approx. 120Å apart in our structure of the budding yeast ODC complex[262]. Consequently, this organization can only occur through a long linker that connects both CBFs in budding yeast and a few related fungi, but is likely not relevant for other eukaryotes that lack such an extended spacer[262].

Our discovery of the Orc6/Cdc6/Orc1 interface that may render the complex inactive is supported by the recently published structure of the semi-attached OCCM intermediate, in which the helicase is only dynamically attached to the ORC/Cdc6 ring[91]. Yuan et al. report that the WHD of Mcm3 and Mcm7 tether the helicase hexamer to ORC/Cdc6 in the semi-attached OCCM structure[91]. However, comparative structural analysis strongly suggest that the electron density assigned to Mcm7-WHD in the semi-attached OCCM corresponds to the N-terminal cyclin-box fold of Orc6 instead[262]. Hence, the semi-attached OCCM complex likely represents a loading-incompetent conformational state. Furthermore, it suggests that the binding of the CBF_N of Orc6 to Cdc6/Orc1 allows helicase recruitment, but not loading, resulting in the formation of unproductive loading intermediates on chromatin, which may be resolved by Cdc6 ATP hydrolysis[118,145,146,190]. These structural observations argue that the Mcm3-WHD is the first (stable) point of contact between Mcm2-7 and the ODC complex and support the idea that the WHD of Mcm3 is essential for helicase recruitment[91,145]. Future studies will be needed to further dissect how the WHD of Mcm3, in concert with Cdc6 and the Orc2-WHD, regulates initial helicase recruitment and possibly helps to resolve unproductive loading intermediates.

Recent technological advances in cryo-electron microscopy hardware combined with improved algorithms for data processing that address heterogeneity in multi-component protein complexes, have allowed us to get unprecedented insight into the details of the protein assemblies that mediate the initiation of DNA replication[11,154,265,266]. However, resolving the complex compositions and kinetics of conformational changes at high resolution during origin licensing and activation remains challenging. In the future,

optimized cryo-EM sample preparation work flows promise to dissect molecular processes at a millisecond time scale[267]. Such time-resolved cryo-EM techniques together with real-time single molecule studies may allow us explore the still largely elusive dynamics that must occur during the initiation of DNA replication. Furthermore, identifications of mutants or chemical entities that slow down or even stall the helicase loading process might aid in capturing more transient intermediates, which will help us better understand how origin licensing is orchestrated. Advances in versatile cryo-EM affinity-grids that allow to enrich specifically for delicate target-proteins may pave the way towards on-grid protein purification and might open new avenues for capturing intrinsically unstable loading intermediates that have avoided structural investigation to date[268].

Missense mutations in ORC and Cdc6 have been identified as potential cause for the Meier-Gorlin Syndrome

Initiator and loader protein function is crucial for timely genome duplication and development. Missense and truncation mutations in ORC and Cdc6 have been found to impair normal human development resulting in the Meier-Gorlin Syndrome[6,7]. However, the molecular mechanisms by which the identified missense mutations in ORC and Cdc6 hinder initiator and co-loader function are not well understood. Our biochemical studies revealed that ORC and Cdc6 harboring MGS-associated mutations in their AAA+ domains show divers defects in ATPase activity, DNA binding and ternary complex assembly, respectively, suggesting partial loss of protein function[269]. Notably, we found that mutations in Orc1 (K730W) and Orc4 (Y174C) reduce ternary ODC complex formation and therefore are expected to compromise helicase loading severely, in agreement with the observation that mutations in Orc1 and Orc4 are associated with a particularly serious MGS phenotype[6]. In the future, more comprehensive characterization of the functional consequences of MGS-mutations in ORC and other initiation factors will be needed to understand how these mutations deregulate origin licensing and firing in MGS patients. We firmly believe that our newly developed *in vitro* metazoan replicative helicase loading assay will greatly facilitate this endeavor.

Protein factors and chromatin cues may help guide ORC binding to chromatin

In addition to contextual cis-acting factors such as DNA sequence properties, specification of origin in higher eukaryotes may be aided by other proteins such as HP1, LRWD1, HMGA1a and DCAF14, which may assist initiator recruitment to distinct chromatin loci[9,270–279]. These ORC-interactors are expected to help integrate various chromatin cues with DNA properties and ORC function to establish origins *in vivo*. However, how these factors contribute to origin selection in different tissues and how they might modulate initiator activity during cell development are exciting questions for the future. In addition, it is likely that more transient, cell-type specific or cell-cycle-stage-specific interactors of ORC are yet to be revealed.

Eukaryotic ORC has been shown to directly interact with chromatin through the Orc1-BAH domain[125,280]. However, the mechanisms by which ORC interacts with nucleosomes have diverged during evolution and the role of epigenetic marks in origin selection across eukaryotes is not well defined[9]. Accordingly, the budding yeast Orc1-BAH domain directly interacts with the nucleosome core independent of histone tail modification[280]. This is important for the selection of origins in *S. cerevisiae*[124,280]. In contrast, mouse and human ORC have been shown to specifically recognize the H4K20me2 histone mark[125]. This interaction may be crucial for licensing and activation of early replication origins in human as the mutation of H4K20 in mammalian cells inhibits progression through S phase[126,281]. On the other hand, H4K20me2 is dispensable for origin activity and DNA replication in *Drosophila*[282,283]. Further research efforts are required to address the significant differences in chromatin recognition by ORC to understand how chromatin cues help shape origin selection and activity across eukaryotes.

Most of our knowledge of the initiation of DNA replication has been derived from the budding yeast system and relatively little is known about how metazoans select origins of replication, control replicative helicase loading and its activation in space and time. My work provides important insights into how metazoan ORC recognizes origin DNA without the help of a consensus sequence. Furthermore, we show that DNA geometrical

features contribute to metazoan ORC binding site selection. We have broaden our understanding how these thermodynamic DNA properties influence origin remodeling and ultimately helicase loading by metazoan ORC *in vitro*. These findings may have important implications for how the location of metazoan origins is establish *in vivo*. High-resolution mapping of ORC binding sites *in vivo* will be required to confirm the role of DNA shape and structural features in defining origins in metazoans.

In addition, our high-resolution structural studies of eukaryotic ORC in complex with DNA and Cdc6 together with comprehensive mutational analysis of the initiator provide a physical explanation for how DNA binding by ORC is coupled to its ATPase activity and how sequential ATP hydrolysis by ORC and Cdc6 might be coordinated. We identified a probable autoinhibited conformation of the budding yeast ODC, which potentially resolves the long-standing question of how CDK-dependent phosphorylation of *Sc*Orc6 inhibits excessive helicase loading outside of late M and G1 phase.

Lastly, it is important to note that our knowledge of DNA replication initiation in eukaryotes is largely limited to a few species, namely budding yeast, *Drosophila* and human[9]. All of these organisms belong to the opisthokonts clade, which represents only a small portion of the eukaryotic domain[9,284]. Although it is likely that the process of DNA replication initiation will be functionally conserved, it is unclear to what extent the underlying mechanisms have diverged during the evolution of eukaryotes[9]. Future efforts are needed to uncover the true diversity of eukaryotic DNA replication initiation.

Acknowledgements

I am grateful to Franziska Bleichert for allowing me to conduct my Ph.D. in her laboratory, for being my mentor over the last 4 years, for all her support and scientific guidance. I would like to extend my gratitude to Susan M. Gasser for hosting me in her laboratory during the last year of my Ph.D., for being my co-first thesis supervisor and for her continuous support.

I would like to thank my thesis committee (Franziska Bleichert, Susan M. Gasser, Marc Bühler, Beat Fierz and Stefanie Jonas) for helpful discussions, their scientific advice and guidance.

I would like thank Mehdi Goudarzi, Stefanie Jonas, Annemarthe van der Veen, Sven Thamm and Sofia Quinodoz for being my supervisors on various different research projects over the course of my Bachelor's and Master's studies. It was a great experience learning from you and working with you greatly shaped the way I do research today.

I would like to thank all current and past members of the Bleichert lab and would like to use this opportunity to thank Sandra Muehlhaeusser, Babatunde Ekundayo, Maxim Kolesnikov, Asim Akbani, Federica Galliano and Stephane Lefevre for their support, for helping to establish the lab, for training me, for stimulating discussions and for being great colleagues.

I would like to thank the Gasser lab for their warm welcome, for teaching me how to work with yeast and for creating a nice atmosphere in the lab. I am especially grateful to Kenji Shimada for working with me on the yeast project.

I am grateful to all the FMI facilities for their work. I am particularly thankful for the support I received from the Structural Biology platform (Simone Cavadini, Andreas Daniel Schenk, Georg Kempf, Alexandra Graff and Christel Genoud) and the Protein Analysis facility (Jan Seebacher, Daniel Hess and Vytautas Iesmantavicius). On this occasion, I would like to send many thanks to Simone and Andreas for teaching me a lot about electron microscopy. Thanks for always being approachable, for answering all of the many questions I had and your help. I would also like to thank very much Jan for introducing me to (cross-linking) mass spectrometry.

I had the pleasure to interact with many great people at the FMI during my Ph.D. and I would like to thank the whole FMI community for creating a welcoming and supportive environment. I thank especially all the members of the 3rd floor labs and the Thomä lab for the great companionship and scientific exchange.

I am very fortunate to have Danilo Maddalo, Ulf Neumann and Anamaria Rauh as my mentors and would like to thank them very much for taking the time to listen to my thoughts, for all the inspiring discussions and for helping me to take the next steps in my career.

My deepest gratitude to Viki, my family and friends for their infinite support and love. I cannot thank you enough for all what you have done for me and feel truly privileged to have you in my life.

References

1. Green, B. M., Finn, K. J. & Li, J. J. Loss of DNA replication control is a potent inducer of gene amplification. *Science* **329**, 943–946 (2010).

2. Klingseisen, A. & Jackson, A. P. Mechanisms and pathways of growth failure in primordial dwarfism. *Genes Dev.* **25**, 2011–2024 (2011).

3. Abbas, T., Keaton, M. A. & Dutta, A. Genomic instability in cancer. *Cold Spring Harb. Perspect. Biol.* **5**, a012914 (2013).

4. Barlow, J. H. & Nussenzweig, A. Replication initiation and genome instability: a crossroads for DNA and RNA synthesis. *Cell. Mol. Life Sci.* **71**, 4545–4559 (2014).

5. Hills, S. A. & Diffley, J. F. X. DNA replication and oncogene-induced replicative stress. *Curr. Biol.* **24**, R435-44 (2014).

6. de Munnik, S. A. *et al.* Meier-Gorlin syndrome. *Orphanet J. Rare Dis.* **10**, 114 (2015).

7. Hossain, M. & Stillman, B. Meier-Gorlin Syndrome. in *The Initiation of DNA Replication in Eukaryotes* (ed. Kaplan, D. L.) 503–524 (Springer International Publishing, 2016). doi:10.1007/978-3-319-24696-3_25

8. Bleichert, F., Botchan, M. R. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Science* **355**, eaah6317 (2017).

9. Ekundayo, B. & Bleichert, F. Origins of DNA replication. *PLoS Genet.* **15**, e1008320 (2019).

10. Bai, L. *et al.* Architecture of the Saccharomyces cerevisiae Replisome. *Adv. Exp. Med. Biol.* **1042**, 207–228 (2017).

11. Lewis, J. S. & Costa, A. Caught in the act: structural dynamics of replication origin activation and fork progression. *Biochem. Soc. Trans.* **48**, 1057–1066 (2020).

12. Parker, M. W., Botchan, M. R. & Berger, J. M. Mechanisms and regulation of DNA

replication initiation in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **52**, 107–144 (2017).

13. Ganier, O., Prorok, P., Akerman, I. & Méchali, M. Metazoan DNA replication origins. *Curr. Opin. Cell Biol.* **58**, 134–141 (2019).

14. Diffley, J. F. X. Quality control in the initiation of eukaryotic DNA replication. *Philos. Trans. R. Soc. London. Ser. B, Biol. Sci.* **366**, 3545–3553 (2011).

15. Siddiqui, K., On, K. F. & Diffley, J. F. X. Regulating DNA replication in eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).

16. Reusswig, K.-U. & Pfander, B. Control of Eukaryotic DNA Replication Initiation-Mechanisms to Ensure Smooth Transitions. *Genes (Basel).* **10**, (2019).

17. Bleichert, F. Mechanisms of replication origin licensing: a structural perspective. *Curr. Opin. Struct. Biol.* **59**, 195–204 (2019).

18. Avery, O. T., Macleod, C. M. & McCarty, M. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. *J. Exp. Med.* **79**, 137–158 (1944).

19. WATSON, J. D. & CRICK, F. H. The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **18**, 123–131 (1953).

20. Dahm, R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.* **122**, 565–581 (2008).

21. Meselson, M. & Stahl, F. W. The replication of DNA in Escherichia coli. *Proc. Natl. Acad. Sci.* **44**, 671–682 (1958).

22. Pellegrini, L. & Costa, A. New Insights into the Mechanism of DNA Duplication by the Eukaryotic Replisome. *Trends Biochem. Sci.* **41**, 859–871 (2016).

23. Miller, T. C. & Costa, A. The architecture and function of the chromatin replication machinery. *Curr. Opin. Struct. Biol.* **47**, 9–16 (2017).

24. Gao, Y. *et al.* Structures and operating principles of the replisome. *Science (80-.).* **363**, eaav7003 (2019).

25. Yuan, Z., Georgescu, R., Schauer, G. D., O'Donnell, M. E. & Li, H. Structure of the polymerase ε holoenzyme and atomic model of the leading strand replisome. *Nat. Commun.* **11**, 3156 (2020).

26. Yang, W., Seidman, M. M., Rupp, W. D. & Gao, Y. Replisome structure suggests mechanism for continuous fork progression and post-replication repair. *DNA Repair (Amst).* **81**, 102658 (2019).

27. Leipe, D. D., Aravind, L. & Koonin, E. V. Did DNA replication evolve twice independently? *Nucleic Acids Res.* **27**, 3389–3401 (1999).

28. Yao, N. Y. & O'Donnell, M. E. Evolution of replication machines. *Crit. Rev. Biochem. Mol. Biol.* **51**, 135–149 (2016).

29. Wolański, M., Donczew, R., Zawilak-Pawlik, A. & Zakrzewska-Czerwińska, J. oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* **5**, 735 (2015).

30. Chodavarapu, S. & Kaguni, J. M. Replication Initiation in Bacteria. *Enzym.* **39**, 1–30 (2016).

31. Katayama, T. Initiation of DNA Replication at the Chromosomal Origin of E. coli, oriC. *Adv. Exp. Med. Biol.* **1042**, 79–98 (2017).

32. Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5088–5090 (1977).

33. Bell, S. D. Archaeal orc1/cdc6 proteins. *Subcell. Biochem.* **62**, 59–69 (2012).

34. Samson, R. Y., Abeyrathne, P. D. & Bell, S. D. Mechanism of Archaeal MCM Helicase Recruitment to DNA Replication Origins. *Mol. Cell* **61**, 287–296 (2016).

35. Bell, S. D. Initiation of DNA Replication in the Archaea. *Adv. Exp. Med. Biol.* **1042**, 99–115 (2017).

36. Cayrou, C. *et al.* Genome-scale analysis of metazoan replication origins reveals

their organization in specific but flexible sites defined by conserved features. *Genome Res.* **21**, 1438–1449 (2011).

37. Miotto, B., Ji, Z. & Struhl, K. Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. *Proc. Natl. Acad. Sci.* **113**, E4810 LP-E4819 (2016).

38. Chagin, V. O. *et al.* 4D Visualization of replication foci in mammalian cells corresponding to individual replicons. *Nat. Commun.* **7**, 11231 (2016).

39. Wang, D. & Gao, F. Comprehensive Analysis of Replication Origins in Saccharomyces cerevisiae Genomes . *Frontiers in Microbiology* **10**, 2122 (2019).

40. Chen, Y.-H. *et al.* Transcription shapes DNA replication initiation and termination in human cells. *Nat. Struct. Mol. Biol.* **26**, 67–77 (2019).

41. Gómez-González, B. & Aguilera, A. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes Dev.* **33**, 1008–1026 (2019).

42. Kelly, T. & Callegari, A. J. Dynamics of DNA replication in a eukaryotic cell. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 4973–4982 (2019).

43. Stinchcomb, D. T., Struhl, K. & Davis, R. W. Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**, 39–43 (1979).

44. Brewer, B. J. & Fangman, W. L. The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell* **51**, 463–471 (1987).

45. Huberman, J. A., Spotila, L. D., Nawotka, K. A., el-Assouli, S. M. & Davis, L. R. The in vivo replication origin of the yeast 2 microns plasmid. *Cell* **51**, 473–481 (1987).

46. Marahrens, Y. & Stillman, B. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**, 817–823 (1992).

47. Rao, H., Marahrens, Y. & Stillman, B. Functional conservation of multiple elements in yeast chromosomal replicators. *Mol. Cell. Biol.* **14**, 7643–7651 (1994).

48. Broach, J. R. *et al.* Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harb. Symp. Quant. Biol.* **47 Pt 2**, 1165–1173 (1983).

49. Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E. & Campbell, J. L. Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae. *Mol. Cell. Biol.* **4**, 2455–2466 (1984).

50. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).

51. Klemm, R. D., Austin, R. J. & Bell, S. P. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell* **88**, 493–502 (1997).

52. Li, N. *et al.* Structure of the origin recognition complex bound to DNA replication origin. *Nature* **559**, 217–222 (2018).

53. Palzkill, T. G. & Newlon, C. S. A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell* **53**, 441–450 (1988).

54. Zou, L. & Stillman, B. Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. *Mol. Cell. Biol.* **20**, 3086–3096 (2000).

55. Lipford, J. R. & Bell, S. P. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell* **7**, 21–30 (2001).

56. Wilmes, G. M. & Bell, S. P. The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 101–106 (2002).

57. Coster, G. & Diffley, J. F. X. Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science* **357**, 314–318 (2017).

58. Diffley, J. F. & Stillman, B. Purification of a yeast protein that binds to origins of

DNA replication and a transcriptional silencer. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2120–2124 (1988).

59. Diffley, J. F. X. & Cocker, J. H. Protein-DNA interactions at a yeast replication origin. *Nature* **357**, 169–172 (1992).

60. Gilbert, D. M. In search of the holy replicator. *Nat. Rev. Mol. Cell Biol.* **5**, 848–855 (2004).

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

62. Prioleau, M.-N. & MacAlpine, D. M. DNA replication origins-where do we begin? *Genes Dev.* **30**, 1683–1697 (2016).

63. Aladjem, M. I. & Redon, C. E. Order from clutter: selective interactions at mammalian replication origins. *Nat. Rev. Genet.* **18**, 101–116 (2017).

64. Austin, R. J., Orr-Weaver, T. L. & Bell, S. P. Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev.* **13**, 2639–2649 (1999).

65. Vashee, S. *et al.* Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* **17**, 1894–1908 (2003).

66. Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *EMBO J.* **23**, 897–907 (2004).

67. Eaton, M. L., Galani, K., Kang, S., Bell, S. P. & MacAlpine, D. M. Conserved nucleosome positioning defines replication origins. *Genes Dev.* **24**, 748–753 (2010).

68. MacAlpine, H. K., Gordân, R., Powell, S. K., Hartemink, A. J. & MacAlpine, D. M. Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Res.* **20**, 201–211 (2010).

69. Xu, J. *et al.* Genome-wide identification and characterization of replication origins

by deep sequencing. *Genome Biol.* **13**, R27 (2012).

70. Gindin, Y., Valenzuela, M. S., Aladjem, M. I., Meltzer, P. S. & Bilke, S. A chromatin structure-based model accurately predicts DNA replication timing in human cells. *Mol. Syst. Biol.* **10**, 722 (2014).

71. Lubelsky, Y. *et al.* DNA replication and transcription programs respond to the same chromatin cues. *Genome Res.* **24**, 1102–1114 (2014).

72. Wang, W. *et al.* Genome-Wide Mapping of Human DNA Replication by Optical Replication Mapping Supports a Stochastic Model of Eukaryotic Replication Timing. *bioRxiv* 2020.08.24.263459 (2020). doi:10.1101/2020.08.24.263459

73. Delgado, S., Gómez, M., Bird, A. & Antequera, F. Initiation of DNA replication at CpG islands in mammalian chromosomes. *EMBO J.* **17**, 2426–2435 (1998).

74. Cayrou, C. *et al.* New insights into replication origin characteristics in metazoans. *Cell Cycle* **11**, 658–667 (2012).

75. Picard, F. *et al.* The spatiotemporal program of DNA replication is associated with specific combinations of chromatin marks in human cells. *PLoS Genet.* **10**, e1004282–e1004282 (2014).

76. Cadoret, J.-C. *et al.* Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15837–15842 (2008).

77. Cayrou, C. *et al.* The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome Res.* **25**, 1873–1885 (2015).

78. Valton, A.-L. *et al.* G4 motifs affect origin positioning and efficiency in two vertebrate replicators. *EMBO J.* **33**, 732–746 (2014).

79. Comoglio, F. *et al.* High-resolution profiling of Drosophila replication start sites reveals a DNA shape and chromatin signature of metazoan origins. *Cell Rep.* **11**, 821–834 (2015).

80. Langley, A. R., Gräf, S., Smith, J. C. & Krude, T. Genome-wide identification and

characterisation of human DNA replication origins by initiation site sequencing (ini-seq). *Nucleic Acids Res.* **44**, 10230–10247 (2016).

81. Prorok, P. *et al.* Involvement of G-quadruplex regions in mammalian replication origin activity. *Nat. Commun.* **10**, 3274 (2019).

82. Dellino, G. I. *et al.* Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. *Genome Res.* **23**, 1–11 (2013).

83. Hoshina, S. *et al.* Human origin recognition complex binds preferentially to G-quadruplex-preferable RNA and single-stranded DNA. *J. Biol. Chem.* **288**, 30161–30171 (2013).

84. Bartholdy, B., Mukhopadhyay, R., Lajugie, J., Aladjem, M. I. & Bouhassira, E. E. Allele-specific analysis of DNA replication origins in mammalian cells. *Nat. Commun.* **6**, 7051 (2015).

85. Tubbs, A. *et al.* Dual Roles of Poly(dA:dT) Tracts in Replication Initiation and Fork Collapse. *Cell* **174**, 1127-1142.e19 (2018).

86. Drew, H. R. & Travers, A. A. DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **186**, 773–790 (1985).

87. Nelson, H. C. M., Finch, J. T., Luisi, B. F. & Klug, A. The structure of an oligo(dA)·oligo(dT) tract and its biological implications. *Nature* **330**, 221–226 (1987).

88. Aymami, J., Coll, M., Frederick, C. A., Wang, A. H. & Rich, A. The propeller DNA conformation of poly(dA).poly(dT). *Nucleic Acids Res.* **17**, 3229–3245 (1989).

89. Lee, D. G. & Bell, S. P. Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol. Cell. Biol.* **17**, 7159 LP – 7168 (1997).

90. Bleichert, F., Leitner, A., Aebersold, R., Botchan, M. R. & Berger, J. M. Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5906–E5915 (2018).

91. Yuan, Z. *et al.* Structural mechanism of helicase loading onto replication origin DNA by ORC-Cdc6. *Proc. Natl. Acad. Sci.* **117**, 17747 LP – 17756 (2020).

92. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**, 27–43 (1999).

93. Iyer, L. M., Leipe, D. D., Koonin, E. V & Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. in *Journal of Structural Biology* **146**, 11–31 (2004).

94. Speck, C., Chen, Z., Li, H. & Stillman, B. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat. Struct. Mol. Biol.* **12**, 965–971 (2005).

95. Duncker, B. P., Chesnokov, I. N. & McConkey, B. J. The origin recognition complex protein family. *Genome Biol.* **10**, 214 (2009).

96. Muzi-Falconi, M. & Kelly, T. J. Orp1, a member of the Cdc18/Cdc6 family of S-phase regulators, is homologous to a component of the origin recognition complex. *Proc. Natl. Acad. Sci.* **92**, 12475 LP – 12479 (1995).

97. Bell, S. P., Mitchell, J., Leber, J., Kobayashi, R. & Stillman, B. The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell* **83**, 563–568 (1995).

98. Perkins, G. & Diffley, J. F. X. Nucleotide-Dependent Prereplicative Complex Assembly by Cdc6p, a Homolog of Eukaryotic and Prokaryotic Clamp-Loaders. *Mol. Cell* **2**, 23–32 (1998).

99. Liu, J. *et al.* Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol. Cell* **6**, 637–648 (2000).

100. Balasov, M., Huijbregts, R. P. H. & Chesnokov, I. Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in Drosophila melanogaster. *Mol. Cell. Biol.* **27**, 3143–3153 (2007).

101. Liu, S. *et al.* Structural analysis of human Orc6 protein reveals a homology with

transcription factor TFIIB. *Proc. Natl. Acad. Sci.* **108**, 7373 LP – 7378 (2011).

102. Bleichert, F. *et al.* A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *Elife* **2**, e00882 (2013).

103. Bleichert, F., Botchan, M. R. & Berger, J. M. Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321–326 (2015).

104. Tocilj, A. *et al.* Structure of the active form of human origin recognition complex and its ATPase motor module. *Elife* **6**, (2017).

105. Yuan, Z. *et al.* Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nat. Struct. Mol. Biol.* **24**, 316–324 (2017).

106. Chen, S., de Vries, M. A. & Bell, S. P. Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. *Genes Dev.* **21**, 2897–2907 (2007).

107. Chesnokov, I., Remus, D. & Botchan, M. Functional analysis of mutant and wild-type Drosophila origin recognition complex. *Proc. Natl. Acad. Sci.* **98**, 11997 LP – 12002 (2001).

108. Bowers, J. L., Randell, J. C. W., Chen, S. & Bell, S. P. ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol. Cell* **16**, 967–978 (2004).

109. Randell, J. C. W., Bowers, J. L., Rodríguez, H. K. & Bell, S. P. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol. Cell* **21**, 29–39 (2006).

110. Speck, C. & Stillman, B. Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *J. Biol. Chem.* **282**, 11705–11714 (2007).

111. Ammelburg, M., Frickey, T. & Lupas, A. N. Classification of AAA+ proteins. *J. Struct. Biol.* **156**, 2–11 (2006).

112. Erzberger, J. & Berger, J. Evolutionary relationships and structural mechanisms of

AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).

113. Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. Structure and function of the AAA+ nucleotide binding pocket. *Biochim. Biophys. Acta* **1823**, 2–14 (2012).

114. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951 (1982).

115. Guarino, E. *et al.* Cdt1 proteolysis is promoted by dual PIP degrons and is modulated by PCNA ubiquitylation. *Nucleic Acids Res.* **39**, 5978–5990 (2011).

116. Schepers, A. & Diffley, J. F. X. Mutational analysis of conserved sequence motifs in the budding yeast cdc6 protein11Edited by M. Yaniv. *J. Mol. Biol.* **308**, 597–608 (2001).

117. Evrin, C. *et al.* In the absence of ATPase activity, pre-RC formation is blocked prior to MCM2–7 hexamer dimerization. *Nucleic Acids Res.* **41**, 3162–3172 (2013).

118. Coster, G., Frigola, J., Beuron, F., Morris, E. P. & Diffley, J. F. X. Origin Licensing Requires ATP Binding and Hydrolysis by the MCM Replicative Helicase. *Mol. Cell* **55**, 666–677 (2014).

119. Ogura, T., Whiteheart, S. W. & Wilkinson, A. J. Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. *J. Struct. Biol.* **146**, 106–112 (2004).

120. Giordano-Coltart, J., Ying, C. Y., Gautier, J. & Hurwitz, J. Studies of the properties of human origin recognition complex and its Walker A motif mutants. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 69 LP – 74 (2005).

121. Callebaut, I., Courvalin, J.-C. & Mornon, J.-P. The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation. *FEBS Lett.* **446**, 189–193 (1999).

122. Zhang, Z., Hayashi, M. K., Merkel, O., Stillman, B. & Xu, R.-M. Structure and function of the BAH-containing domain of Orc1p in epigenetic silencing. *EMBO J.* **21**, 4600–4611 (2002).

123. Noguchi, K., Vassilev, A., Ghosh, S., Yates, J. L. & DePamphilis, M. L. The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. *EMBO J.* **25**, 5372–5382 (2006).

124. Müller, P. *et al.* The conserved bromo-adjacent homology domain of yeast Orc1 functions in the selection of DNA replication origins within chromatin. *Genes Dev.* **24**, 1418–1433 (2010).

125. Kuo, A. J. *et al.* The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484**, 115–119 (2012).

126. Long, H. *et al.* H2A.Z facilitates licensing and activation of early replication origins. *Nature* **577**, 576–581 (2020).

127. Parker, M. W. *et al.* A new class of disordered elements controls DNA replication through initiator self-assembly. *Elife* **8**, e48562 (2019).

128. Kawakami, H., Ohashi, E., Kanamoto, S., Tsurimoto, T. & Katayama, T. Specific binding of eukaryotic ORC to DNA replication origins depends on highly conserved basic residues. *Sci. Rep.* **5**, 14929 (2015).

129. Hu, Y. *et al.* Evolution of DNA Replication Origin Specification and Gene Silencing Mechanisms. *bioRxiv* 2020.07.04.187286 (2020). doi:10.1101/2020.07.04.187286

130. Hartwell, L. H. Sequential function of gene products relative to DNA synthesis in the yeast cell cycle. *J. Mol. Biol.* **104**, 803–817 (1976).

131. Maine, G. T., Sinha, P. & Tye, B. K. Mutants of S. cerevisiae defective in the maintenance of minichromosomes. *Genetics* **106**, 365–385 (1984).

132. Coleman, T. R., Carpenter, P. B. & Dunphy, W. G. The Xenopus Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts. *Cell* **87**, 53–63 (1996).

133. Donovan, S., Harwood, J., Drury, L. S. & Diffley, J. F. Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5611–5616 (1997).

134. Wang, B. *et al.* The essential role of Saccharomyces cerevisiae CDC6 nucleotide-binding site in cell growth, DNA synthesis, and Orc1 association. *J. Biol. Chem.* **274**, 8291–8298 (1999).

135. Weinreich, M., Liang, C. & Stillman, B. The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 441–446 (1999).

136. Mizushima, T., Takahashi, N. & Stillman, B. Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes Dev.* **14**, 1631–1641 (2000).

137. Nishitani, H., Lygerou, Z., Nishimoto, T. & Nurse, P. The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature* **404**, 625–628 (2000).

138. Tanaka, S. & Diffley, J. F. X. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2–7 during G1 phase. *Nat. Cell Biol.* **4**, 198–207 (2002).

139. Maiorano, D., Moreau, J. & Méchali, M. XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis. *Nature* **404**, 622–625 (2000).

140. Remus, D. *et al.* Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).

141. Evrin, C. *et al.* A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20240–20245 (2009).

142. Gambus, A., Khoudoli, G. A., Jones, R. C. & Blow, J. J. MCM2-7 form double hexamers at licensed origins in Xenopus egg extract. *J. Biol. Chem.* **286**, 11855–11864 (2011).

143. Sun, J. *et al.* Cdc6-Induced Conformational Changes in ORC Bound to Origin

DNA Revealed by Cryo-Electron Microscopy. *Structure* **20**, 534–544 (2012).

144. Sun, J. *et al.* Cryo-EM structure of a helicase loading intermediate containing ORC–Cdc6–Cdt1–MCM2-7 bound to DNA. *Nat. Struct. Mol. Biol.* **20**, 944–951 (2013).

145. Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. X. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339–343 (2013).

146. Fernández-Cid, A. *et al.* An ORC/Cdc6/MCM2-7 Complex Is Formed in a Multistep Reaction to Serve as a Platform for MCM Double-Hexamer Assembly. *Mol. Cell* **50**, 577–588 (2013).

147. Li, N. *et al.* Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).

148. Bochman, M. L. & Schwacha, A. The Mcm2-7 Complex Has In Vitro Helicase Activity. *Mol. Cell* **31**, 287–293 (2008).

149. Costa, A. *et al.* The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.* **18**, 471–477 (2011).

150. Samel, S. A. *et al.* A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA. *Genes Dev.* **28**, 1653–1666 (2014).

151. Kawasaki, Y., Kim, H.-D., Kojima, A., Seki, T. & Sugino, A. Reconstitution of Saccharomyces cerevisiae prereplicative complex assembly in vitro. *Genes to Cells* **11**, 745–756 (2006).

152. Frigola, J. *et al.* Cdt1 stabilizes an open MCM ring for helicase loading. *Nat. Commun.* **8**, 15720 (2017).

153. Zhai, Y. *et al.* Open-ringed structure of the Cdt1-Mcm2-7 complex as a precursor of the MCM double hexamer. *Nat. Struct. Mol. Biol.* **24**, 300–308 (2017).

154. Miller, T. C. R., Locke, J., Greiwe, J. F., Diffley, J. F. X. & Costa, A. Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature* **575**,

704–710 (2019).

155. Ticau, S., Friedman, L. J., Ivica, N. A., Gelles, J. & Bell, S. P. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell* **161**, 513–525 (2015).

156. Randell, J. C. W. *et al.* Mec1 Is One of Multiple Kinases that Prime the Mcm2-7 Helicase for Phosphorylation by Cdc7. *Mol. Cell* **40**, 353–363 (2010).

157. Sheu, Y.-J. & Stillman, B. The Dbf4–Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* **463**, 113–117 (2010).

158. Heller, R. C. *et al.* Eukaryotic Origin-Dependent DNA Replication In Vitro Reveals Sequential Action of DDK and S-CDK Kinases. *Cell* **146**, 80–91 (2011).

159. Tanaka, S., Nakato, R., Katou, Y., Shirahige, K. & Araki, H. Origin Association of Sld3, Sld7, and Cdc45 Proteins Is a Key Step for Determination of Origin-Firing Timing. *Curr. Biol.* **21**, 2055–2063 (2011).

160. Tanaka, T. *et al.* Sld7, an Sld3-associated protein required for efficient chromosomal DNA replication in budding yeast. *EMBO J.* **30**, 2019–2030 (2011).

161. Deegan, T. D., Yeeles, J. T. & Diffley, J. F. Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. *EMBO J.* **35**, 961–973 (2016).

162. Kamimura, Y., Masumoto, H., Sugino, A. & Araki, H. Sld2, which interacts with Dpb11 in Saccharomyces cerevisiae, is required for chromosomal DNA replication. *Mol. Cell. Biol.* **18**, 6102–6109 (1998).

163. Tanaka, S. *et al.* CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* **445**, 328–332 (2007).

164. Zegerman, P. & Diffley, J. F. X. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* **445**, 281–285 (2007).

165. Kubota, Y. *et al.* A novel ring-like complex of Xenopus proteins essential for the

initiation of DNA replication. *Genes Dev.* **17**, 1141–1152 (2003).

166. Takayama, Y. *et al.* GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev.* **17**, 1153–1165 (2003).

167. Kanemaki, M., Sanchez-Diaz, A., Gambus, A. & Labib, K. Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature* **423**, 720–724 (2003).

168. Tanaka, S. *et al.* Efficient initiation of DNA replication in eukaryotes requires Dpb11/TopBP1-GINS interaction. *Mol. Cell. Biol.* **33**, 2614–2622 (2013).

169. Muramatsu, S., Hirai, K., Tak, Y.-S., Kamimura, Y. & Araki, H. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. *Genes Dev.* **24**, 602–612 (2010).

170. Labib, K., Tercero, J. A. & Diffley, J. F. Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science* **288**, 1643–1647 (2000).

171. Tercero, J. A., Labib, K. & Diffley, J. F. DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. *EMBO J.* **19**, 2082–2093 (2000).

172. Pacek, M. & Walter, J. C. A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *EMBO J.* **23**, 3667–3676 (2004).

173. Shechter, D., Ying, C. Y. & Gautier, J. DNA unwinding is an Mcm complex-dependent and ATP hydrolysis-dependent process. *J. Biol. Chem.* **279**, 45586–45593 (2004).

174. Moyer, S. E., Lewis, P. W. & Botchan, M. R. Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl. Acad. Sci.* **103**, 10236 LP – 10241 (2006).

175. Pacek, M., Tutter, A. V, Kubota, Y., Takisawa, H. & Walter, J. C. Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell* **21**, 581–587 (2006).

176. Douglas, M. E., Ali, F. A., Costa, A. & Diffley, J. F. X. The mechanism of eukaryotic CMG helicase activation. *Nature* **555**, 265–268 (2018).

177. Solomon, N. A. *et al.* Genetic and molecular analysis of DNA43 and DNA52: Two new cell-cycle genes in Saccharomyces cerevisiae. *Yeast* **8**, 273–289 (1992).

178. Merchant, A. M., Kawasaki, Y., Chen, Y., Lei, M. & Tye, B. K. A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **17**, 3261 LP – 3271 (1997).

179. Lõoke, M., Maloney, M. F. & Bell, S. P. Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes Dev.* **31**, 291–305 (2017).

180. Mayle, R. *et al.* Mcm10 has potent strand-annealing activity and limits translocase-mediated fork regression. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 798–803 (2019).

181. Jr., R. M. B. & Trakselis, M. A. Fine-tuning of the replisome: Mcm10 regulates fork progression and regression. *Cell Cycle* **18**, 1047–1055 (2019).

182. Fu, Y. V *et al.* Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* **146**, 931–941 (2011).

183. Georgescu, R. *et al.* Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E697–E706 (2017).

184. Yuan, Z. *et al.* DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat. Commun.* **11**, 688 (2020).

185. Villa, F. *et al.* Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. *Mol. Cell* **63**, 385–396 (2016).

186. Yuan, Z. *et al.* Ctf4 organizes sister replisomes and Pol α into a replication factory. *Elife* **8**, (2019).

187. Wasserman, M. R., Schauer, G. D., O'Donnell, M. E. & Liu, S. Replication Fork Activation Is Enabled by a Single-Stranded DNA Gate in CMG Helicase. *Cell* **178**, 600-611.e16 (2019).

188. Baretić, D. *et al.* Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. *Mol. Cell* **78**, 926-940.e13 (2020).

189. Seki, T. & Diffley, J. F. X. Stepwise assembly of initiation proteins at budding yeast replication origins in vitro *Proc. Natl. Acad. Sci.* **97**, 14115 LP – 14120 (2000).

190. Kang, S., Warner, M. D. & Bell, S. P. Multiple Functions for Mcm2–7 ATPase Motifs during Replication Initiation. *Mol. Cell* **55**, 655–665 (2014).

191. Takahashi, N., Tsutsumi, S., Tsuchiya, T., Stillman, B. & Mizushima, T. Functions of sensor 1 and sensor 2 regions of Saccharomyces cerevisiae Cdc6p in vivo and in vitro. *J. Biol. Chem.* **277**, 16033–16040 (2002).

192. Chang, F. *et al.* Cdc6 ATPase activity disengages Cdc6 from the pre-replicative complex to promote DNA replication. *Elife* **4**, (2015).

193. Poon, R. Y. C. Cell Cycle Control: A System of Interlinking Oscillators. *Methods Mol. Biol.* **1342**, 3–19 (2016).

194. Ang, X. L. & Wade Harper, J. SCF-mediated protein degradation and cell cycle control. *Oncogene* **24**, 2860–2870 (2005).

195. Peters, J.-M. The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat. Rev. Mol. Cell Biol.* **7**, 644–656 (2006).

196. Visintin, R. *et al.* CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis. *Cell* **278**, 683–693 (1997).

197. Rudner, A. D. & Murray, A. W. Phosphorylation by Cdc28 activates the Cdc20-dependent activity of the anaphase-promoting complex. *J. Cell Biol.* **149**, 1377–1390 (2000).

198. Kraft, C. *et al.* Mitotic regulation of the human anaphase-promoting complex by

phosphorylation. *EMBO J.* **22**, 6598–6609 (2003).

199. Zachariae, W., Schwab, M., Nasmyth, K. & Seufert, W. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science* **282**, 1721–1724 (1998).

200. Jaspersen, S. L., Charles, J. F. & Morgan, D. O. Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Curr. Biol.* **9**, 227–236 (1999).

201. Schwab, M., Lutum, A. S. & Seufert, W. Yeast Hct1 is a regulator of Clb2 cyclin proteolysis. *Cell* **90**, 683–693 (1997).

202. Schwab, M., Neutzner, M., Möcker, D. & Seufert, W. Yeast Hct1 recognizes the mitotic cyclin Clb2 and other substrates of the ubiquitin ligase APC. *EMBO J.* **20**, 5165—5175 (2001).

203. Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411**, 1068–1073 (2001).

204. Wilmes, G. M. *et al.* Interaction of the S-phase cyclin Clb5 with an 'RXL' docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev.* **18**, 981–991 (2004).

205. Green, B. M., Morreale, R. J., Ozaydin, B., Derisi, J. L. & Li, J. J. Genome-wide mapping of DNA synthesis in Saccharomyces cerevisiae reveals that mechanisms preventing reinitiation of DNA replication are not redundant. *Mol. Biol. Cell* **17**, 2401–2414 (2006).

206. Drury, L. S., Perkins, G. & Diffley, J. F. The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J.* **16**, 5966–5976 (1997).

207. Perkins, G., Drury, L. S. & Diffley, J. F. Separate SCF(CDC4) recognition elements target Cdc6 for proteolysis in S phase and mitosis. *EMBO J.* **20**, 4836–4845 (2001).

208. Calzada, A., Sánchez, M., Sánchez, E. & Bueno, A. The stability of the Cdc6

protein is regulated by cyclin-dependent kinase/cyclin B complexes in Saccharomyces cerevisiae. *J. Biol. Chem.* **275**, 9734–9741 (2000).

209. Drury, L. S., Perkins, G. & Diffley, J. F. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Curr. Biol.* **10**, 231–240 (2000).

210. Weinreich, M., Liang, C., Chen, H.-H. & Stillman, B. Binding of cyclin-dependent kinases to ORC and Cdc6p regulates the chromosome replication cycle. *Proc. Natl. Acad. Sci.* **98**, 11211 LP – 11217 (2001).

211. Mimura, S., Seki, T., Tanaka, S. & Diffley, J. F. X. Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature* **431**, 1118–1123 (2004).

212. Örd, M., Venta, R., Möll, K., Valk, E. & Loog, M. Cyclin-Specific Docking Mechanisms Reveal the Complexity of M-CDK Function in the Cell Cycle. *Mol. Cell* **75**, 76-89.e3 (2019).

213. Calzada, A., Sacristán, M., Sánchez, E. & Bueno, A. Cdc6 cooperates with Sic1 and Hct1 to inactivate mitotic cyclin-dependent kinases. *Nature* **412**, 355–358 (2001).

214. Remus, D., Blanchette, M., Rio, D. C. & Botchan, M. R. CDK phosphorylation inhibits the DNA-binding and ATP-hydrolysis activities of the Drosophila origin recognition complex. *J. Biol. Chem.* **280**, 39740–39751 (2005).

215. Chen, S. & Bell, S. P. CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev.* **25**, 363–372 (2011).

216. Baldinger, T. & Gossen, M. Binding of Drosophila ORC proteins to anaphase chromosomes requires cessation of mitotic cyclin-dependent kinase activity. *Mol. Cell. Biol.* **29**, 140–149 (2009).

217. Li, C., Vassilev, A. & DePamphilis, M. L. Role for Cdk1 (Cdc2)/Cyclin A in Preventing the Mammalian Origin Recognition Complex's Largest Subunit (Orc1) from Binding to Chromatin during Mitosis. *Mol. Cell. Biol.* **24**, 5875 LP – 5886

(2004).

218. Liku, M. E., Nguyen, V. Q., Rosales, A. W., Irie, K. & Li, J. J. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. *Mol. Biol. Cell* **16**, 5026–5039 (2005).

219. Labib, K., Diffley, J. F. & Kearsey, S. E. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nat. Cell Biol.* **1**, 415–422 (1999).

220. Nguyen, V. Q., Co, C., Irie, K. & Li, J. J. Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. *Curr. Biol.* **10**, 195–205 (2000).

221. Ralph, E., Boye, E. & Kearsey, S. E. DNA damage induces Cdt1 proteolysis in fission yeast through a pathway dependent on Cdt2 and Ddb1. *EMBO Rep.* **7**, 1134–1139 (2006).

222. Kim, Y. & Kipreos, E. T. Cdt1 degradation to prevent DNA re-replication: conserved and non-conserved pathways. *Cell Div.* **2**, 18 (2007).

223. Abbas, T. & Dutta, A. CRL4Cdt2: master coordinator of cell cycle progression and genome stability. *Cell Cycle* **10**, 241–249 (2011).

224. Havens, C. G. & Walter, J. C. Mechanism of CRL4(Cdt2), a PCNA-dependent E3 ubiquitin ligase. *Genes Dev.* **25**, 1568–1582 (2011).

225. McGarry, T. J. & Kirschner, M. W. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93**, 1043–1053 (1998).

226. Wohlschlegel, J. A. *et al.* Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* **290**, 2309–2312 (2000).

227. Tada, S., Li, A., Maiorano, D., Méchali, M. & Blow, J. J. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nat. Cell Biol.* **3**, 107–113 (2001).

228. Nishitani, H., Taraviras, S., Lygerou, Z. & Nishimoto, T. The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after

initiation of S-phase. *J. Biol. Chem.* **276**, 44905–44911 (2001).

229. Yanagi, K., Mizuno, T., You, Z. & Hanaoka, F. Mouse geminin inhibits not only Cdt1-MCM6 interactions but also a novel intrinsic Cdt1 DNA binding activity. *J. Biol. Chem.* **277**, 40871–40880 (2002).

230. Lee, C. *et al.* Structural basis for inhibition of the replication licensing factor Cdt1 by geminin. *Nature* **430**, 913–917 (2004).

231. MEIER, Z., POSCHIAVO & ROTHSCHILD, M. [Case of arthrogryposis multiplex congenita with mandibulofacial dysostosis (Franceschetti syndrome)]. *Helv. Paediatr. Acta* **14**, 213–216 (1959).

232. Gorlin, R. J., Cervenka, J., Moller, K., Horrobin, M. & Witkop, C. J. J. Malformation syndromes. A selected miscellany. *Birth Defects Orig. Artic. Ser.* **11**, 39–50 (1975).

233. Boles, R. G., Teebi, A. S., Schwartz, D. & Harper, J. F. Further delineation of the ear, patella, short stature syndrome (Meier-Gorlin syndrome). *Clin. Dysmorphol.* **3**, 207–214 (1994).

234. Bicknell, L. S. *et al.* Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome. *Nat. Genet.* **43**, 350–356 (2011).

235. Bicknell, L. S. *et al.* Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat. Genet.* **43**, 350–356 (2011).

236. Guernsey, D. L. *et al.* Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nat. Genet.* **43**, 360–365 (2011).

237. Jaremko, M. J., On, K. F., Thomas, D. R., Stillman, B. & Joshua-Tor, L. The dynamic nature of the human origin recognition complex revealed through five cryoEM structures. *Elife* **9**, e58622 (2020).

238. Kronja, I. *et al.* Quantitative proteomics reveals the dynamics of protein changes during Drosophila oocyte maturation and the oocyte-to-embryo transition. *Proc.*

Natl. Acad. Sci. U. S. A. **111**, 16023–16028 (2014).

239. Ranjan, A. & Gossen, M. A structural role for ATP in the formation and stability of the human origin recognition complex. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4864 LP – 4869 (2006).

240. Siddiqui, K. & Stillman, B. ATP-dependent assembly of the human origin recognition complex. *J. Biol. Chem.* **282**, 32370–32383 (2007).

241. Kara, N., Hossain, M., Prasanth, S. G. & Stillman, B. Orc1 Binding to Mitotic Chromosomes Precedes Spatial Patterning during G1 Phase and Assembly of the Origin Recognition Complex in Human Cells. *J. Biol. Chem.* **290**, 12355–12369 (2015).

242. Méndez, J. *et al.* Human Origin Recognition Complex Large Subunit Is Degraded by Ubiquitin-Mediated Proteolysis after Initiation of DNA Replication. *Mol. Cell* **9**, 481–491 (2002).

243. Tatsumi, Y., Ohta, S., Kimura, H., Tsurimoto, T. & Obuse, C. The ORC1 cycle in human cells: I. cell cycle-regulated oscillation of human ORC1. *J. Biol. Chem.* **278**, 41528–41534 (2003).

244. Schmidt, J. M. & Bleichert, F. Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nat. Commun.* **11**, 4263 (2020).

245. Chuang, R. Y. & Kelly, T. J. The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 2656–2661 (1999).

246. Lee, J. K., Moon, K. Y., Jiang, Y. & Hurwitz, J. The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 13589–13594 (2001).

247. Chuang, R.-Y., Chretien, L., Dai, J. & Kelly, T. J. Purification and characterization of the Schizosaccharomyces pombe origin recognition complex: interaction with

origin DNA and Cdc18 protein. *J. Biol. Chem.* **277**, 16920–16927 (2002).

248. Gaczynska, M. *et al.* Atomic force microscopic analysis of the binding of the Schizosaccharomyces pombe origin recognition complex and the spOrc4 protein with origin DNA. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17952–17957 (2004).

249. MacAlpine, D. M., Rodríguez, H. K. & Bell, S. P. Coordination of replication and transcription along a Drosophila chromosome. *Genes Dev.* **18**, 3094–3105 (2004).

250. Liu, J. *et al.* DNA sequence templates adjacent nucleosome and ORC sites at gene amplification origins in Drosophila. *Nucleic Acids Res.* **43**, 8746–8761 (2015).

251. Field, Y. *et al.* Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* **4**, e1000216 (2008).

252. Segal, E. & Widom, J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **19**, 65–71 (2009).

253. Mojardín, L., Vázquez, E. & Antequera, F. Specification of DNA replication origins and genomic base composition in fission yeasts. *J. Mol. Biol.* **425**, 4706–4713 (2013).

254. Levine, J. & Spradling, A. DNA sequence of a 3.8 kilobase pair region controlling Drosophila chorion gene amplification. *Chromosoma* **92**, 136–142 (1985).

255. Heck, M. M. & Spradling, A. C. Multiple replication origins are used during Drosophila chorion gene amplification. *J. Cell Biol.* **110**, 903–914 (1990).

256. Van Houten, J. V & Newlon, C. S. Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol. Cell. Biol.* **10**, 3917 LP – 3925 (1990).

257. Amati, B., Pick, L., Laroche, T. & Gasser, S. M. Nuclear scaffold attachment stimulates, but is not essential for ARS activity in Saccharomyces cerevisiae: analysis of the Drosophila ftz SAR. *EMBO J.* **9**, 4007–4016 (1990).

258. Amati, B. & Gasser, S. M. Drosophila scaffold-attached regions bind nuclear scaffolds and can function as ARS elements in both budding and fission yeasts. *Mol. Cell. Biol.* **10**, 5442 LP – 5454 (1990).

259. Dechering, K. J., Cuelenaere, K., Konings, R. N. & Leunissen, J. A. Distinct frequency-distributions of homopolymeric DNA tracts in different genomes. *Nucleic Acids Res.* **26**, 4056–4062 (1998).

260. Okuno, Y., Satoh, H., Sekiguchi, M. & Masukata, H. Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin. *Mol. Cell. Biol.* **19**, 6699–6709 (1999).

261. Gaudier, M., Schuwirth, B. S., Westcott, S. L. & Wigley, D. B. Structural basis of DNA replication origin recognition by an ORC protein. *Science* **317**, 1213–1216 (2007).

262. Schmidt, J. M., Seebacher, J. & Bleichert, F. *No Title*. (2020).

263. Hanson, P. I. & Whiteheart, S. W. AAA+ proteins: have engine, will work. *Nat. Rev. Mol. Cell Biol.* **6**, 519–529 (2005).

264. Miller, J. M. & Enemark, E. J. Fundamental Characteristics of AAA+ Protein Family Structure and Function. *Archaea* **2016**, 9294307 (2016).

265. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).

266. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Breaking the next Cryo-EM resolution barrier – Atomic resolution determination of proteins! (2020). doi:10.1101/2020.05.21.106740

267. Kaledhonkar, S., Fu, Z., White, H. & Frank, J. Time-Resolved Cryo-electron Microscopy Using a Microfluidic Chip. *Methods Mol. Biol.* **1764**, 59–71 (2018).

268. Wang, F. *et al.* General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. *Proc. Natl. Acad. Sci.* **117**, 24269 LP – 24273 (2020).

269. Schmidt, J. M., Galliano, F. & Bleichert, F. *No Title*. (2020).

270. Pak, D. T. S. *et al.* Association of the Origin Recognition Complex with Heterochromatin and HP1 in Higher Eukaryotes. *Cell* **91**, 311–323 (1997).

271. Wang, Y. *et al.* Temporal association of ORCA/LRWD1 to late-firing origins during G1 dictates heterochromatin replication and organization. *Nucleic Acids Res.* **45**, 2490–2502 (2017).

272. Thomae, A. W. *et al.* Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 1692–1697 (2008).

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

274. Prasanth, S. G., Shen, Z., Prasanth, K. V & Stillman, B. Human origin recognition complex is essential for HP1 binding to chromatin and heterochromatin organization. *Proc. Natl. Acad. Sci.* **107**, 15093 LP – 15098 (2010).

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

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

277. Shen, Z. *et al.* Dynamic association of ORCA with prereplicative complex components regulates DNA replication initiation. *Mol. Cell. Biol.* **32**, 3107–3120 (2012).

278. Hein, M. Y. *et al.* A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).

279. Zhang, Y. *et al.* A replicator-specific binding protein essential for site-specific initiation of DNA replication in mammalian cells. *Nat. Commun.* **7**, 11748 (2016).

280. De Ioannes, P. *et al.* Structure and function of the Orc1 BAH-nucleosome complex. *Nat. Commun.* **10**, 2894 (2019).

281. Brustel, J. *et al.* Histone H4K20 tri-methylation at late-firing origins ensures timely

heterochromatin replication. *EMBO J.* **36**, 2726–2741 (2017).

282. McKay, D. J. *et al.* Interrogating the function of metazoan histones using engineered gene clusters. *Dev. Cell* **32**, 373–386 (2015).

283. Li, Y., Armstrong, R. L., Duronio, R. J. & MacAlpine, D. M. Methylation of histone H4 lysine 20 by PR-Set7 ensures the integrity of late replicating sequence domains in Drosophila. *Nucleic Acids Res.* **44**, 7204–7218 (2016).

284. Burki, F. The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harb. Perspect. Biol.* **6**, a016147 (2014).

Appendix

Goal 1: Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex.

Jan Marten Schmidt and Franziska Bleichert

Nat. Commun. 11, 4263 (2020). https://doi.org/10.1038/s41467-020-18067-7

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-18067-7 OPEN

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

Jan Marten Schmidt[1,2] & Franziska Bleichert[3]✉

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.

[1] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, Basel 4058, Switzerland. [2] University of Basel, Petersplatz 1, Basel 4051, Switzerland. [3] Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06520, USA. ✉email: franziska.bleichert@yale.edu

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[1,2]. 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[3,4]. A helix-turn-helix (HTH) fold serves as the primary origin DNA-binding and origin-targeting region in bacterial and archaeal initiators[5–7]. This HTH domain is appended to an ATP hydrolase associated with various cellular activities (AAA+) fold, a domain prevalent in a large group of proteins involved in diverse biological processes beyond DNA replication[8,9]. A common theme among members of this superfamily is that they act as 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[10,11]. In initiators, the AAA+ regions also engage DNA, and these interactions directly remodel origin DNA in the prokaryotic proteins[6,7,12]. 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 AAA+ oligomer and how this mechanism relates to other AAA+ systems 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[1,13]. ORC is composed of five evolutionary related AAA+ proteins (Orc1–5) and a sixth subunit (Orc6) of distinct origin[14,15]. 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[16–18]. 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[18,20,21]. 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[20]. 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[22–31]. 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 (*Dm*ORC) 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 *Dm*ORC assemblies, reveal unexpected differences to the *S. cerevisiae* ORC·DNA cryo-EM structure[20] 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 archaeal initiators. We further reconstitute Mcm2–7 loading in vitro using recombinant *Drosophila* proteins and find that DNA bending 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 *Dm*ORC·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[18]. 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[32]. Deletion of Orc1 and Cdc6 N-termini did not substantially weaken *Dm*ORC's ATP-dependent affinity for DNA[18], nor did it impair the ability of *Dm*ORC to hydrolyze ATP and of *Dm*ORC and *Dm*Cdc6 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, *Dm*Orc1, *Dm*Cdc6, but also *Dm*Cdt1 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 *Dm*ORC and *Dm*Cdc6 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 *Dm*ORC·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 *Dm*ORC·DNA·Cdc6 assembly at an overall resolution of 3.4 Å (Supplementary Figs. 2c, d, 3a, b, 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 TFIIB 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 *Dm*ORC·Cdc6 ring encircles and bends DNA. In the DNA-bound initiator·co-loader complex, *Dm*ORC and

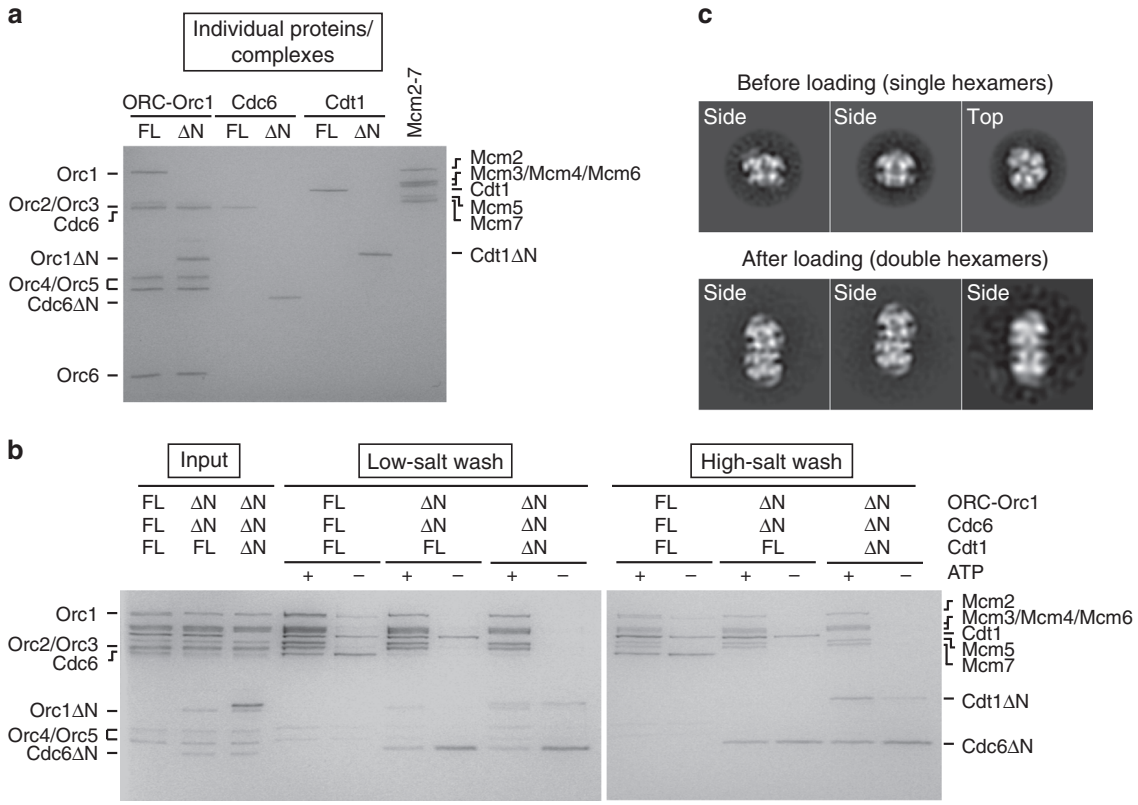


Fig. 1 Mcm2–7 loading does not require intrinsically disordered N-terminal regions (IDRs) of loading factors. a SDS-PAGE gel of purified *Drosophila* loading factors, either full-length proteins or constructs lacking IDRs (ΔN). **b** *Dm*Mcm2–7 is recruited (low-salt wash) and loaded (high-salt wash) in an ATP-dependent manner by full-length and IDR-less *Dm*Orc1, *Dm*Cdc6, and *Dm*Cdt1 with similar efficiency. Note that FL-Cdt1, Cdt1ΔN, Cdc6, Cdc6ΔN, and to some extent FL-Orc1 associate with beads nonspecifically. **c** Loaded *Dm*Mcm2–7 are double-hexamers. *Dm*Mcm2–7 before and after loading was analyzed by negative-stain EM. 2D class averages show Mcm2–7 single-hexamers before loading but double-hexamers after loading. Source data are provided as a Source data file.

*Dm*Cdc6 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-Dm*ORC[16] 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 ORC[17], 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 Å^2 to 1412 Å^2), 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 subunit[15] and adopts a similar 3D structure as Orc1 in the ternary initiator·co-loader complex (Supplementary Fig. 4a). In *Dm*ORC·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-strand that abuts the central 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 harbors 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 *Dm*ORC·Cdc6. A distinctive feature of the ternary *Dm*ORC·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 *Dm*ORC·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).

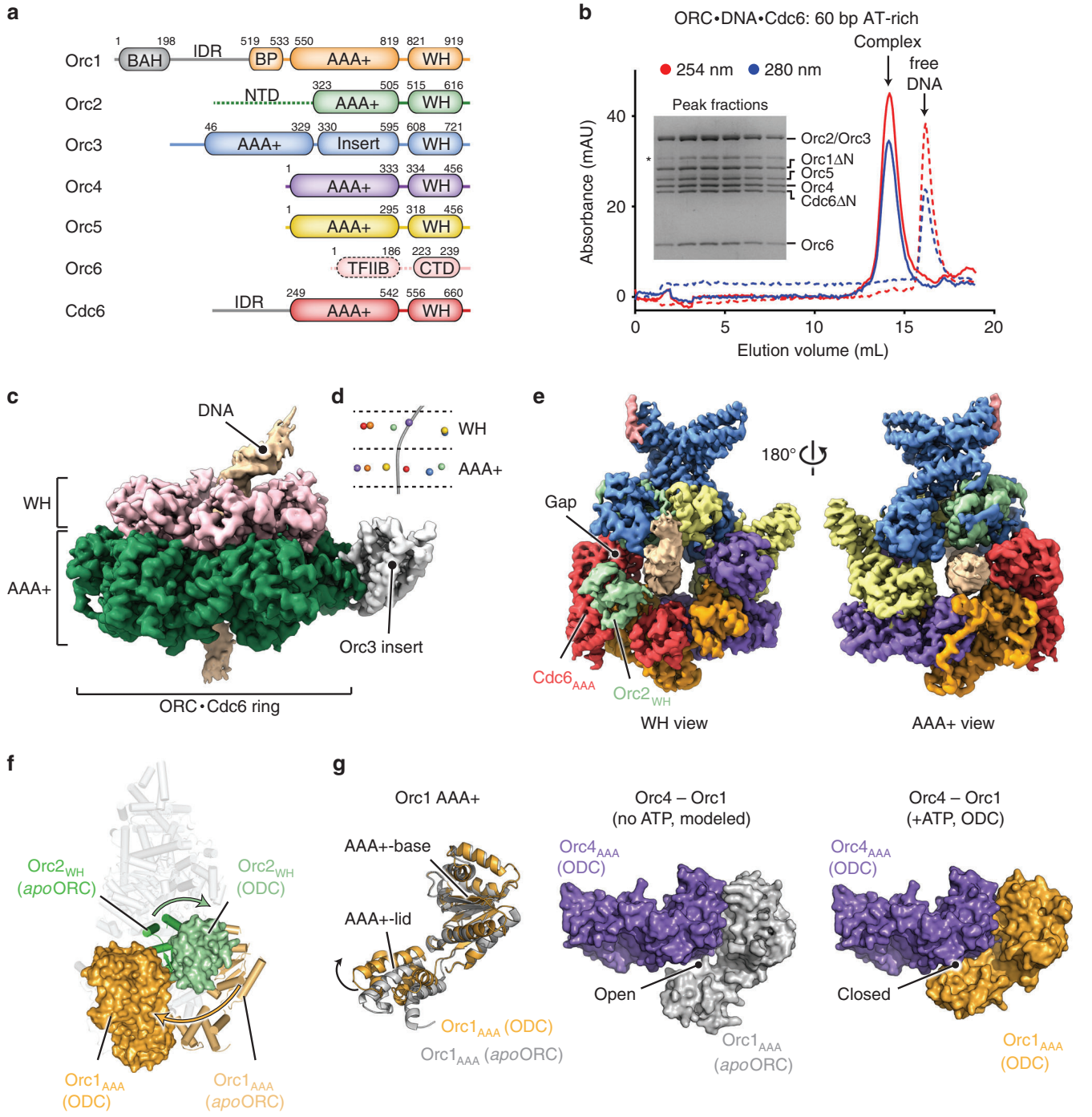
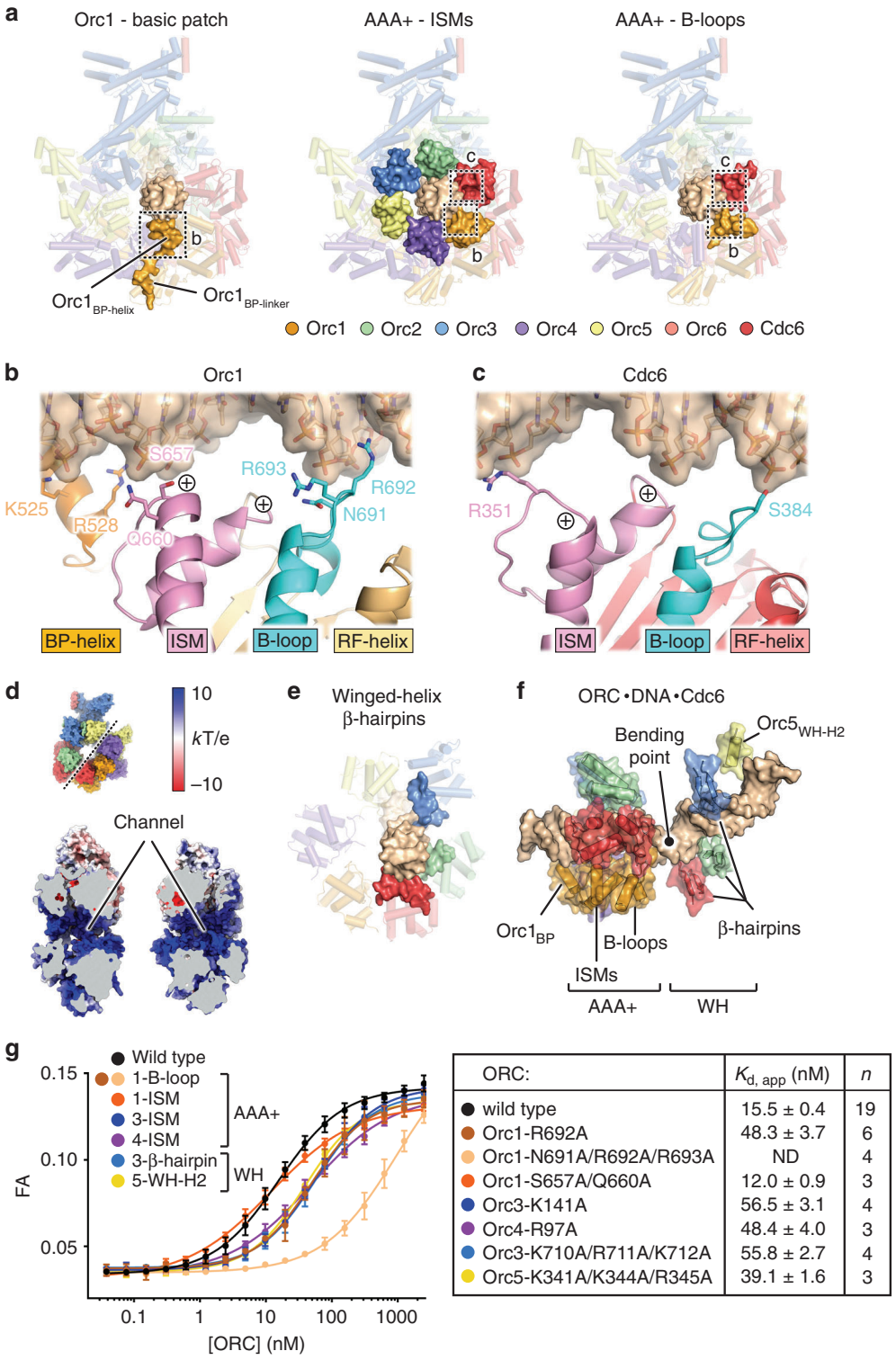


Fig. 2 Structure of a licensing-competent *Dm*ORC·DNA·Cdc6 complex. a Domain architecture of *Dm*ORC subunits and Cdc6. Dashed lines demarcate regions that are flexible and structurally not resolved. Gray regions in Orc1 and Cdc6 were removed to improve sample behavior for cryo-EM. The color scheme is maintained throughout the figures unless noted otherwise. **b** Purification of *Dm*ORC·DNA·Cdc6 for cryo-EM. Chromatograms are shown for the ternary complex (solid lines) and for isolated DNA (dashed lines). Peak fractions retain all *Dm*ORC subunits and *Dm*Cdc6 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 *Dm*ORC 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-Dm*ORC state (shown as cartoon; PDB 4xgc[16]) 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-Dm*ORC (PDB 4xgc[16]) 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.



ORC:	$K_{d,\,app}$ (nM)	*n*
wild type	15.5 ± 0.4	19
Orc1-R692A	48.3 ± 3.7	6
Orc1-N691A/R692A/R693A	ND	4
Orc1-S657A/Q660A	12.0 ± 0.9	3
Orc3-K141A	56.5 ± 3.1	4
Orc4-R97A	48.4 ± 4.0	3
Orc3-K710A/R711A/K712A	55.8 ± 2.7	4
Orc5-K341A/K344A/R345A	39.1 ± 1.6	3

The initiator-specific motifs (ISMs) of *Dm*Orc1–5 and Cdc6, a helical insertion in the AAA+ fold characteristic for members of the initiator clade of AAA+ ATPases[10,33], 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

Fig. 3 Protein–DNA contacts in the ternary *Dm*ORC·DNA·Cdc6 assembly. a The Orc1-basic batch, the ISM ring, and the B-loops of Orc1 and Cdc6 (all shown as molecular surfaces) form three layers within the AAA+ tier that interact with DNA. Close-up views of the Orc1·DNA and Cdc6·DNA interfaces corresponding to boxed areas in (**a**) are displayed in (**b**, **c**), respectively. Amino acid side chains in contact with DNA are shown as sticks, and the positive dipoles of the ISM helices are marked. **d** Open-book view of the *Dm*ORC·Cdc6 channel colored by electrostatic surface potential. The opening plane is noted as a dashed line in the upper overview image of the complex. **e** The β-hairpins of Orc2, Orc3, and Cdc6 (all shown as molecular surface) bind DNA in the WH-tier. The other WH regions are depicted as transparent cartoon. AAA+ domains are omitted for clarity. **f** Summary of protein–DNA contacts in *Dm*ORC·DNA·Cdc6. Only regions that engage DNA are shown. In addition to a subset of the WH β-hairpins, helix 2 of the Orc5 WH domain ($Orc5_{WH\text{-}H2}$) is also positioned close to the duplex. Note that in the binary *Dm*ORC·DNA complex, the WH domain of Orc2 does not contact DNA (see Supplementary Fig. 6e). **g** Mutations in DNA binding regions decrease *Dm*ORC's affinity for DNA as measured by fluorescence anisotropy in the presence of ATP. The means and standard deviations, as well as fitted binding curves, for wild type and mutant ORC assemblies are shown. Amino acid substitutions, the apparent dissociation constants ($K_{d,\ app}$ ±S.E.) for DNA binding, and numbers (*n*) of independent replicates are summarized. ND – $K_{d,\ app}$ could not be determined since binding curves did not reach saturation. Source data are provided as a Source data file.

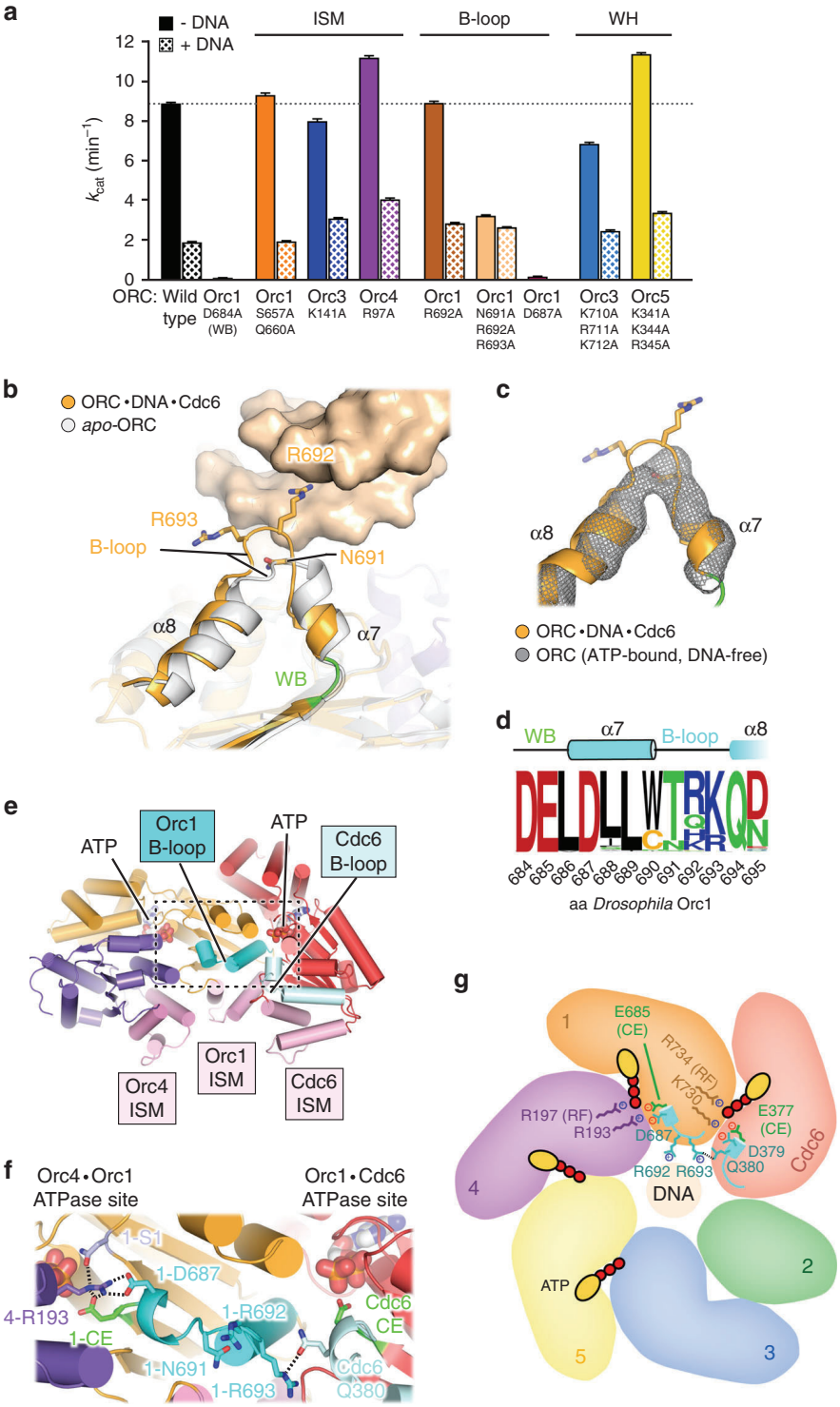
groove where it is stabilized by interactions with the deoxyribose and by the negative electrostatic potential of the backbone phosphate groups but makes no contacts with the nucleobases (Fig. 3b, Supplementary Figs. 3d, 5). Together with the ISMs and the Orc1 basic patch, these B-loops form an extended 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 *Dm*ORC·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 *Dm*ORC·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 bending. 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 *Dm*ORC·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 *Dm*ORC·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 *Dm*ORC and *Dm*ORC·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).

*Dm*ORC binds DNA with low nanomolar affinity in an ATP-dependent manner[18,34]. To understand how the different DNA-binding sites in *Dm*ORC·DNA and *Dm*ORC·DNA·Cdc6 contribute to ORC's affinity for DNA, we substituted single or multiple conserved residues in these regions with alanine and measured *Dm*ORC's ability to associate with DNA using a fluorescence anisotropy-based binding assay. As Cdc6 binding to *Dm*ORC·DNA does not alter how the initiator engages DNA (Supplementary Fig. 6), these binding assays report on ORC–DNA contacts in both the binary and ternary complexes. All mutant *Dm*ORC assemblies purified as a stable hexamer, indicating that the substitutions did not interfere with complex assembly (Supplementary Fig. 7a). As seen previously[18], wild type *Dm*ORC bound a fluorescein-labeled 40 bp AT-rich DNA duplex with an apparent dissociation constant ($K_{d,\ app}$) 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 *Dm*ORCs 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 ($Orc1^{N691A/R692A/R693A}$) completely abolished ATP-dependent, high-affinity DNA binding by *Dm*ORC, 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 *Dm*ORC on DNA, with triple mutations of basic residues ($Orc3^{K710A/R711A/K712A}$ or $Orc5^{K341A/K344A/R345A}$) decreasing *Dm*ORC'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 *Dm*ORC·DNA and *Dm*ORC·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 *Dm*ORC 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 rings[16,17,20,21,35–37]. 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 *Dm*ORC·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 metazoans[17,34,35,37–39]. Notably, ATP hydrolysis by ORC (i.e., at the Orc1/4 interface) is inhibited by DNA[34,35], 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 *Dm*ORC and our various DNA binding site mutants in the absence and presence of DNA (Fig. 4a). For wild type *Dm*ORC, we observe a 5-fold decrease in ATP turnover upon DNA addition, a reduction comparable to that reported for *S. cerevisiae* ORC[35]. Amino acid substitutions in the ISMs and WH domains of *Dm*ORC did not drastically alter



the basal ATPase rate and still supported DNA-mediated inhibition. Strikingly, the ATP turnover rate of *Dm*ORC 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 *Dm*ORC·DNA·Cdc6 complex with that in the DNA-free *apo*-*Dm*ORC crystal structure and our DNA-free ATP-*Dm*ORC cryo-EM structure[16]. 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 (Fig. 4b, c). Interestingly, the B-loop residues and the preceding α-helix are highly conserved in metazoan Orc1 proteins (Fig. 4d). 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 (Fig. 4e–g). Importantly, an Orc1^{D687A} mutation abolishes ATP hydrolysis by *Dm*ORC, supporting a critical role for the pre-B-loop helix in organizing this ATPase center (Fig. 4a). We propose that the Orc1$^{N691A/R692A/R693A}$ mutant alters the B-loop

Fig. 4 The Orc1 B-loop helps couple DNA binding and ATPase activities in *Dm*ORC. a Basal (−DNA) and DNA-inhibited (+DNA) ATPase activities of wild type *Dm*ORC and complexes harboring alanine substitutions of DNA-interacting residues. Mutations in the B-loop of Orc1 substantially reduce *Dm*ORC's basal ATPase rate and abrogate further inhibition by DNA. Substitutions of Orc1-D684 in the Walker B motif and of Orc1-D687, which links the Orc1 B-loop to active site residues, render the complex catalytically inactive. ATP hydrolysis rates from three independent ATP titrations (except for wild type ORC ($n = 10$ −DNA; $n = 4$ +DNA) and Orc1^{R692A} ($n = 4$ −DNA, $n = 4$ +DNA)) were fit to the Michaelis–Menten equation, and k_{cat} and S.E. of fits are plotted as bar graph. **b**, **c** Conformational rearrangement of the Orc1 B-loop upon DNA binding. **b** The Orc1 B-loops of DNA-free *apo-Dm*ORC (gray; PDB 4xgc[16]) and the ternary *Dm*ORC·DNA·Cdc6 complex (orange) adopt different configurations. Orc1-R692 inserts into the DNA minor groove. A short α-helix (α7) connects the DNA-binding B-loop to the Walker B motif (green). **c** A comparison of Orc1 B-loop conformations in the ternary complex (orange cartoon) and in ATP-bound, DNA-free ORC (unsharpened EM map depicted as gray mesh). **d** Sequence frequency logo of metazoan Orc1 shows the conservation of B-loop residues and identifies an invariant aspartate (Orc1-D687) between the Walker B motif and the helix preceding the B-loop. Secondary structure features are illustrated above the sequence logo. **e** The B-loops and adjacent helices of both Orc1 and Cdc6, but not their ISMs, are located near the Orc1/4 and Cdc6/Orc1 ATPase sites. AAA+ domains of Orc1, Orc4, and Cdc6 are shown, viewed from ORC's central channel. **f** Zoomed view of dashed region in (**e**) illustrates the bonding network between the Orc1 B-loop region, adjacent helices, and Orc1/4 active site residues. Dashed lines indicate hydrogen bonds and salt bridges. **g** Summary of interactions from (**f**). The Orc1-sensor 1 residue, which also bonds with Orc4-R193, is omitted for clarity. WB, Walker B (green); CE, catalytic glutamate (green); S1, sensor 1 (light blue); RF, arginine finger (colored by subunit). Source data are provided as a Source data file.

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 critical 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 *Sc*ORC. 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[14,25,30,40,41]. Our structure of the ternary *Dm*ORC·DNA·Cdc6 complex afforded comparative analysis with a recent cryo-EM model of DNA-complexed *S. cerevisiae* ORC (*Sc*ORC)[20] 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 β-hairpin for DNA binding, albeit how these elements interact with nucleic acid varies slightly. Nonetheless, numerous differences in ORC/DNA contacts are also apparent (Fig. 5a, b, Supplementary Movie 1). In the *Sc*ORC·DNA complex, the Orc1 B-loop is pushed away from DNA by the Orc2 WH β-hairpin, suggesting it is a DNA binding element unique to metazoans. Conversely, many additional regions of *Sc*ORC, but not *Dm*ORC, interact with DNA, which results in a substantially larger solvent-accessible surface area that is buried upon DNA binding of the budding yeast initiator (~3900 Å^2 for *Sc*ORC vs. ~1200 Å^2 for *Dm*ORC) (Fig. 5b). Notably, the TFIIB domain of Orc6, the C-terminal part of which binds DNA in *Sc*ORC[20], is not visible in the binary or ternary DNA-bound *Dm*ORC assembly, even when the length of the DNA duplex is increased (Fig. 5b, Supplementary Fig. 9a–c). This disparity can be explained by the lack of a yeast-specific loop insertion in the WH domain of Orc5, which in *Sc*ORC likely helps position the C-terminal TFIIB module near the DNA duplex (Supplementary Fig. 9d–f). Basic patches in Orc2 and Orc5 that cooperate with Orc6 to achieve DNA bending in *Sc*ORC·DNA are likewise not ordered in DNA-bound *Dm*ORC assemblies (Fig. 5a, b, Supplementary Fig. 9d, e); these observations suggest that the corresponding *Dm*ORC regions are not strictly required for inducing DNA bending, although they could become important during later steps of Mcm2–7 loading and help stabilize the duplex segment distal of the bend.

Another important difference relates to the role of the Orc1 basic patch (Orc1-BP) in DNA binding. Mutation of this element in *Dm*ORC and *Sc*ORC attenuates the affinity of both initiators for DNA[18,42], but the underlying structural mechanism for these concordant observations is strikingly different. Instead of snaking along the minor groove as seen in *Sc*ORC·DNA[20], the N-terminal part of *Dm*Orc1's basic patch folds into a helix that only locally interacts with the sugar-phosphate backbone (Figs. 3b, 5a–c, Supplementary Fig. 5). In addition, this *Dm*Orc1-BP helix and the adjacent linker also pack against the ISM of Orc4 (Fig. 5c). Comparing *Drosophila* and human ORC structures obtained in the absence or presence of the Orc1 basic patch reveals that these interactions trigger a rotation of the second Orc4-ISM helix, moving the positive, N-terminal Orc4 helix dipole closer to DNA (Fig. 5d). Therefore, the strong DNA binding defect of *Dm*ORC seen previously upon deleting the Orc1 basic patch is likely caused by the combined effect of abrogating Orc1-BP contacts with DNA and the allosteric influence on the Orc4-ISM[18]. At the same time, the paucity of interactions with the DNA grooves, which in *Sc*ORC are involved in base-pair recognition of ARS consensus sequences[20], rationalizes the reported sequence promiscuity of *Dm*ORC and orthologous initiators in metazoans, and most likely also within the broader eukaryotic domain. Taken together, our results show that budding yeast and metazoan initiators share a subset of common DNA binding regions but have also evolved specialized elements that suit the specific needs of these initiators in origin recognition.

DNA geometry and malleability modulate *Dm*ORC activities. Although *Dm*ORC 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[25,30], but the physical basis for these observations is unknown. Given the paucity of DNA base-contacts by *Dm*ORC, 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 *Dm*ORC 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 *Dm*ORC[25] (Fig. 6a, Supplementary

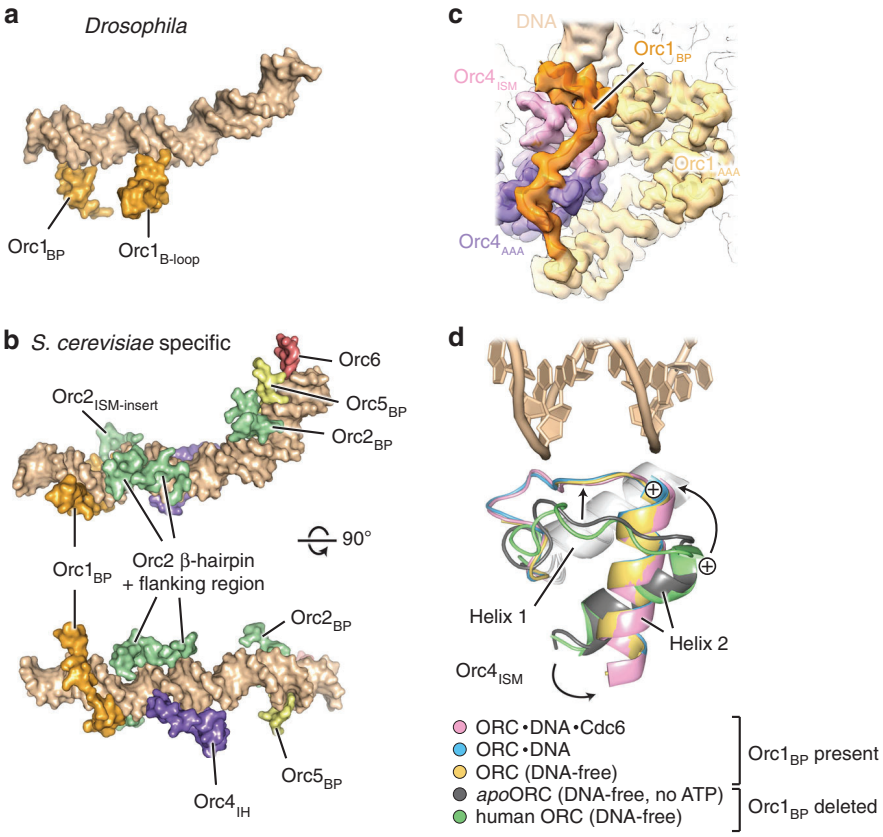
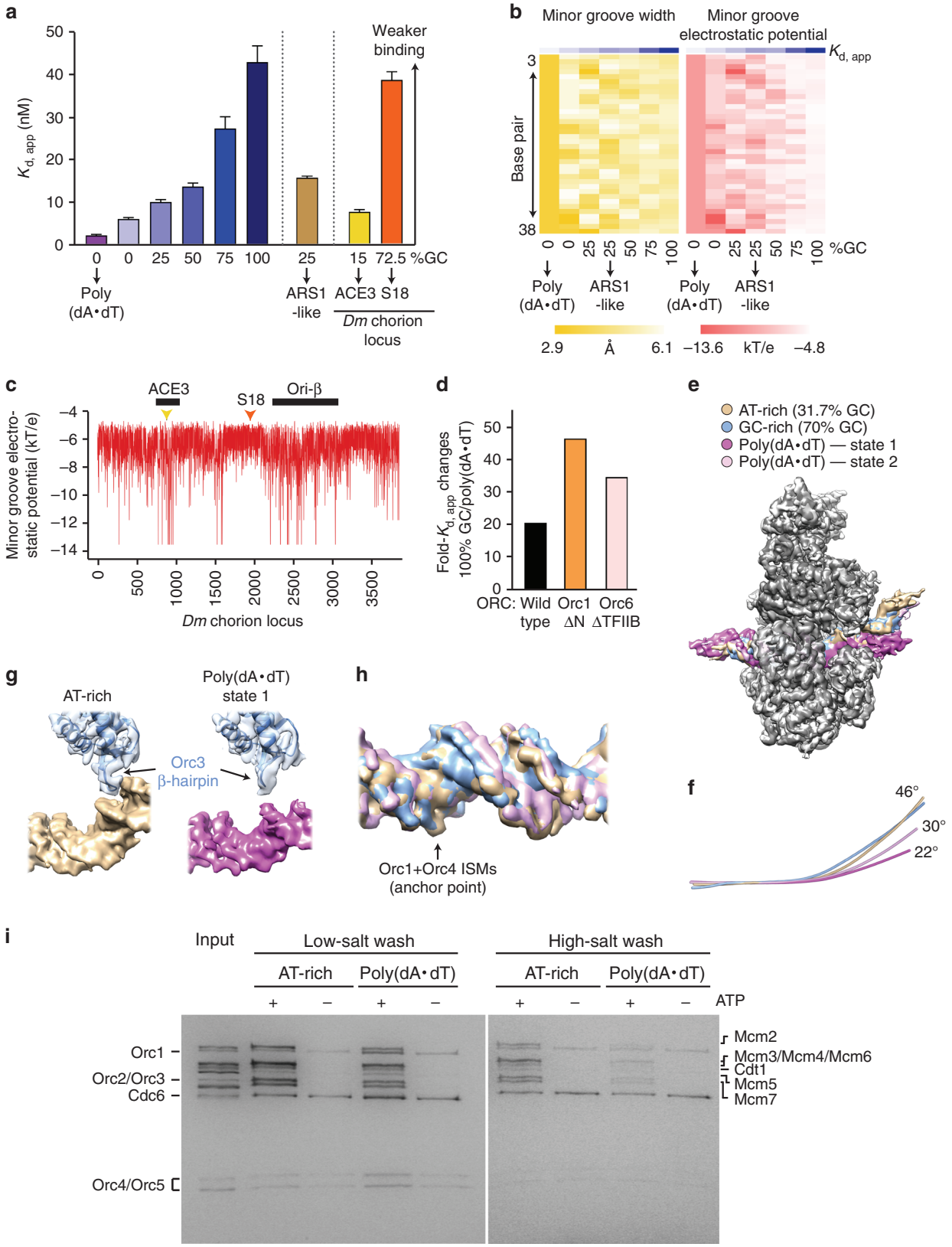


Fig. 5 Comparison of *S. cerevisiae* and *Drosophila* ORC–DNA contacts. a The B-loop of Orc1 engages DNA in *Drosophila* but not in budding yeast ORC, while interactions between the Orc1-BP and DNA differ in both systems. DNA binding elements shared with *S. cerevisiae* ORC are not shown. **b** *S. cerevisiae* ORC uses numerous additional regions to associate with DNA in the cryo-EM structure of the DNA-bound budding yeast initiator (PDB 5zr1[20]). A subset of these are yeast-specific insertions in ORC subunits (e.g., $Orc2_{ISM\text{-}insert}$, $Orc4_{IH}$), while others are flexible in *Dm*ORC·DNA and/or *Dm*ORC·DNA·Cdc6 (i.e., $Orc2_{BP}$, $Orc5_{BP}$, $Orc6_{TFIIB}$). 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 *Dm*ORC structures determined in this study (all containing the Orc1-basic patch) are superposed onto the Orc4-ISMs from the *apo-Dm*ORC and human Orc1/4/5 crystal structures (PDBs 4gxc[16] and 5uj7[17]; 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 K_d values of *Dm*ORC 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 *Drosophila* 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 *Dm*ORC 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-terminus[23,32], enhances *Dm*ORC's ability to discriminate between duplexes of high and low GC content (Fig. 6d, Supplementary Fig. 10a–c).

To understand the structural mechanisms for *Dm*ORC's preferential binding to AT-rich DNA and its consequences for ORC-dependent DNA remodeling, we reconstituted two additional *Dm*ORC·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 *Dm*ORC·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 *Dm*ORC to a lesser extent than the AT-rich duplex (46° bend) (Fig. 6e, 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 β-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. 6e, 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 *Sc*ORC·DNA determined previously[20], 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 *Dm*ORC. a Increasing GC-content alleviates *Dm*ORC's ATP-dependent affinity for DNA. The apparent dissociation constants ($K_{d,\ app}$) and standard errors (S.E.) of fit for ATP-dependent *Dm*ORC 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 $K_{d,\ app}$ for *Dm*ORC 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 *Dm*ORC enhances *Dm*ORC's preference for associating with AT-rich duplexes. The ratios of $K_{d,\ app}$-100% GC over $K_{d,\ app}$-poly(dA·dT) are plotted for wild type, full-length *Dm*ORC, and complexes lacking respective secondary DNA binding regions. See Supplementary Fig. 10a–c for binding curves. **e** Superposition of cryo-EM maps (unsharpened) of *Dm*ORC and *Dm*Cdc6 (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 *Dm*ORC. **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 *Dm*ORC·DNA·Cdc6 structures are shown (except for state 1 of the poly(dA·dT) complex). **i** DNA sequence modulates Mcm2–7 loading efficiency. *Dm*Mcm2–7 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 *Dm*ORC presented here, in combination with the previous crystal structure of the DNA-free *apo*-complex[16], afford the direct comparison of different ORC functional states in a single species to define structural transitions in the initiator during early stages of replication initiation. In nucleotide-free *Dm*ORC, the Orc1 AAA+ domain is disengaged from its Orc4 partner, precluding productive DNA binding in ORC's channel (Fig. 7a, *i*–*ii*). The first major restructuring event of ORC occurs during ATP binding, which alters the conformation of Orc1′s AAA+ domain to stabilize interactions of this module with the AAA+ domain of Orc4 during ORC activation (Fig. 7a, *iii*). Juxtaposition of the Orc1 and Orc4 ATPase modules facilitates docking of the Orc1 basic patch helix and linker onto Orc4, which in turn triggers a rotation of the Orc4 ISM helix and primes ORC for DNA binding in its central channel. Mutations or deletions of the Orc1-BP reduce *Dm*ORC's affinity for DNA, supporting the notion that the Orc4-ISM/Orc1-BP platform is critical for DNA engagement[18]. Interestingly, the Orc1-BP harbors two cyclin-dependent kinase (CDK) phosphorylation sites; phosphorylation of *Dm*ORC has been reported to inhibit ATP-dependent DNA binding[44], an outcome that may be caused by destabilization of the Orc1-BP/Orc4-ISM DNA binding site. Lastly, conformational flexibility of the ORC ring, mediated by motions at the Orc3/Orc5 interface, likely promotes DNA entry into ORC's channel by transiently widening of the gap between Orc1 and Orc2 (Fig. 7a, *iv*–*v*). Thus, multiple conformational rearrangements in the initiator, some of which are conserved among eukaryotic ORC assemblies, control its association with DNA. For example, instability of the Orc1/4 AAA+/AAA+ interface in the absence of ATP has also been observed in human and *S. cerevisiae* ORC, offering a structural explanation for the ATP-dependence of ORC/DNA interactions in ORC's central channel[16,18,20]. 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 *Dm*ORC'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 recruitment of both Mcm2–7 hexamers[21,45–47]. In this regard, it is interesting to note that efficient docking of the Orc2 WH domain against the ORC·Cdc6 ring also requires a certain degree of DNA bending and may, therefore, provide a means to regulate Mcm2–7 recruitment and loading.

Structural comparison of DNA-bound *S. cerevisiae* and *Dm*ORC 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[20,21]. By contrast, *Dm*ORC 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 association 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[25,26,30,43,48,49]. Thus, AT-rich sequences and poly(dA·dT) stretches at metazoan replication origins may not only promote initiation by maintaining nucleosome-free regions[50], 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

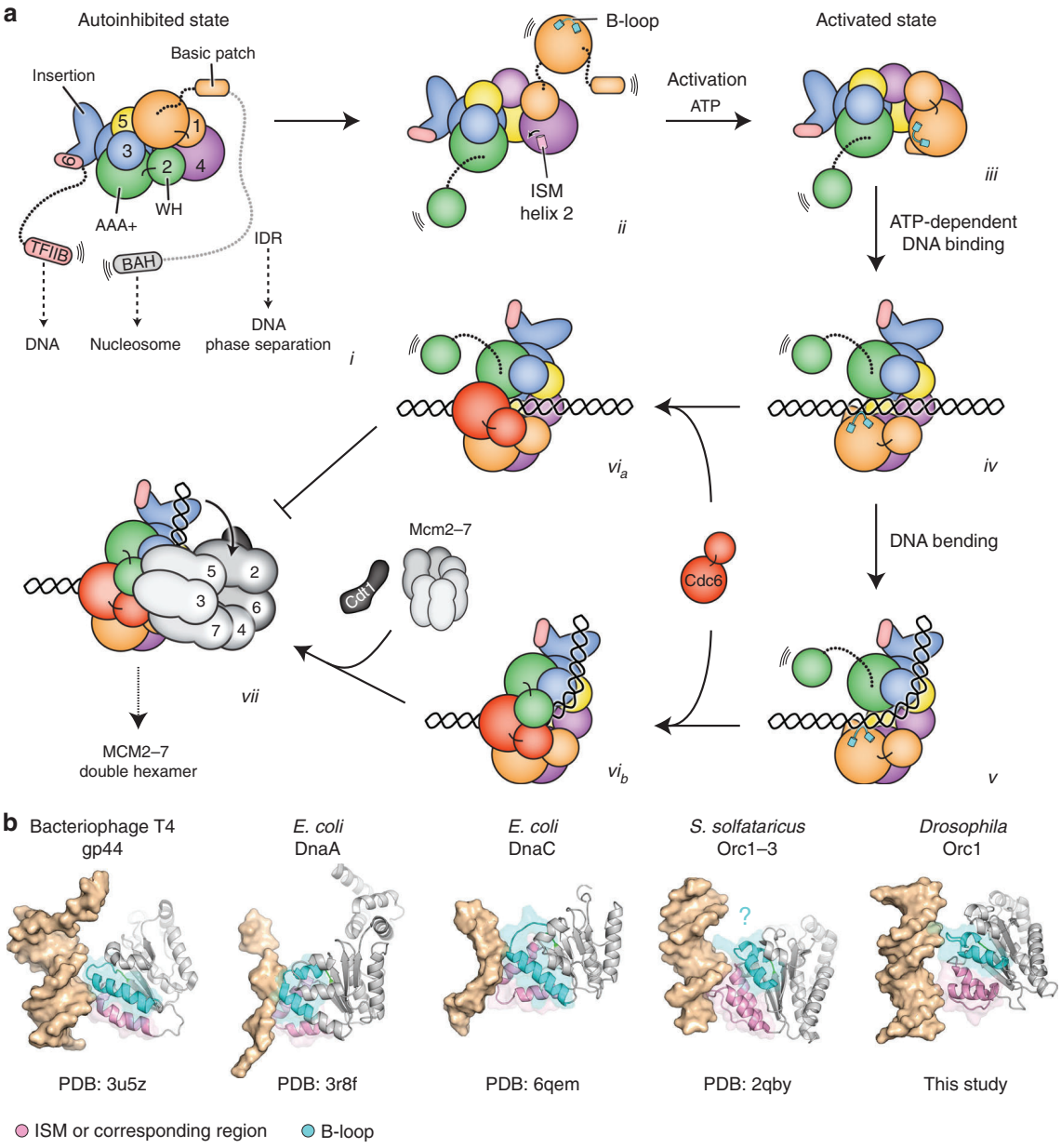


Fig. 7 Metazoan ORC activities during Mcm2–7 loading. a Summary of conformational transitions in metazoan ORC and of DNA binding and remodeling mechanisms (panels *i-vii*, see text for details). ORC regions that have been implicated in recruiting ORC to origins by serving as secondary DNA binding sites (Orc1-IDR, Orc6-TFIIB) or nucleosome-interacting domains (Orc1-BAH) are indicated. The N-terminal region in Orc1 deleted in our construct used for structure determination is shown in gray. TFIIB, transcription factor IIB-like domain in Orc6; BAH, bromo-adjacent homology domain in Orc1; IDR, intrinsically disordered region. **b** Comparison of DNA contacts in diverse initiator and loader proteins highlights the B-loop element as a common DNA binding region. Archaeal Orc1/Cdc6 constitutes an exception since its B-loop is more distant from DNA compared to the other AAA+ ATPase modules. Nucleic-acid bound AAA+ domain structures of bacteriophage T4 clamp loader (PDB 3u5z[85]), *E. coli* DnaA (PDB 3r8f[12]) and DnaC (PDB 6qem[59]), *S. solfataricus* Orc1-3 (PDB 2qby[6]), and *Drosophila* Orc1 (this study) are shown. The initiator-specific motifs (ISM) or corresponding region in the clamp loader, and B-loop regions are depicted as transparent pink and cyan surfaces, respectively.

(including the Orc6-TFIIB module and the phase separating IDR in Orc1) to recruit ORC to origins on metazoan chromosomes[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 *S. cerevisiae* and *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*)[18,20]. Unexpectedly, we observed varying degrees of *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 particularly 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 *Dm*ORC structures is not as extensive as in the *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[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 poly (dA·dT) stretches), may require the initiator to passively slide on DNA until it encounters a malleable DNA segment[30,50,57]. Alternatively, metazoan ORC may recruit specific protein co-factors 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 assemblies[11,58]. In ORC, DNA binding inhibits ATP hydrolysis at the Orc1/4 site, the principal ATPase center, in the yeast and the *Drosophila* complex[34,35] (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 *Dm*ORC·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 elements[20]; 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 complex[21]) 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 hydrolysis[59]. This structural congruency 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 (*Dm*ORC) assemblies used in this study were reconstituted in insect cells by co-infection 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 hexa-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 BioBricks MultiBac expression vector (Macrolab, University of California Berkeley, USA) by subcloning different combinations of *Dm*ORC genes. Bacmids of single and MultiBac expression constructs were generated in DH10Bac cells and used for transfections of Sf9 cells with Cellfectin II (Thermo Scientific Fisher). Baculoviruses were amplified for two rounds in Sf9 cells to obtain high-titer viruses for infection of Hi5 cells for large-scale *Dm*ORC expression.

Drosophila Cdc6 was also expressed in insect cells. The coding sequences of full-length or N-terminally truncated *Dm*Cdc6 (ΔN241, amino acids 242-662) were cloned into the LIC-compatible pFastBac vector (Macrolab, University of California Berkeley, USA). The full-length *Dm*Cdc6 construct contained an N-terminal MBP tag followed by a TEV cleavage site, while that of *Dm*Cdc6ΔN241 was fused to an N-terminal 6xHis-MBP tag and a TEV cleavage site. Bacmid generation and baculovirus amplification were performed as described for *Dm*ORC.

Drosophila Cdt1 and Mcm2–7 were expressed in *E. coli* and Hi5 insect cells, respectively. Full-length or N-terminally trimmed (amino acids 298-743) *Dm*Cdct1 were cloned as N-terminal 6xHis-TEV fusions into a LIC-converted, pET-derived *E. coli* expression vector (Macrolab, University of California Berkeley, USA), while Mcm2–7 subunits were cloned into two pFastBac-derived BioBricks MultiBac 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 MBP and 6xHis, respectively, each followed by a TEV protease cleavage site.

Expression and purification of recombinant *Dm*ORC and *Dm*Cdc6. The different *Dm*ORC assemblies were purified at 4 °C from 4 liters Hi5 cells infected with combinations of baculoviruses expressing single or multiple *Dm*ORC subunits as previously[16,19] with minor modifications. 48 h post-infection, Hi5 cells were harvested, resuspended in ~130 mL lysis buffer (50 mM Tris-HCl, pH 7.8, 300 mM KCl, 50 mM imidazole, pH 7.8, 10% glycerol, 200 µM PMSF, 1 µg/mL leupeptin, 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 142,414 × *g* for 45 min in a 45Ti rotor. After the first centrifugation step, ammonium sulfate was added to the lysate supernatant to a final concentration of 20% (v/v). The 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 *Dm*ORC 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). 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). *Dm*ORC was concentrated in a 30K Amicon Ultra-15 concentrator (Millipore) and subjected to gel filtration chromatography on a HiPrep 16/60 Sephacryl S-400 HR column (GE Healthcare) equilibrated in 25 mM HEPES-KOH pH 7.6, 10% glycerol, 1 mM DTT and either 500 mM potassium glutamate (for most biochemical assays), or 300 mM potassium acetate and 10 mM magnesium acetate (for reconstituting DNA·ORC and DNA·ORC·Cdc6 assemblies for cryo-EM). Peak fractions containing *Dm*ORC were pooled and concentrated in 30K Amicon Ultra-15 concentrators (Millipore). Proteins were then aliquoted, flash frozen in liquid nitrogen, and stored at −80 °C.

Full-length MBP-tagged *Dm*Cdc6 and 6xHis-MBP-tagged *Dm*Cdc6ΔN241 were expressed in Hi5 cells for 48 h. For purification of full-length *Dm*Cdc6, cells were resuspended in 100 mL buffer containing 50 mM Tris-HCl, pH 7.8, 600 mM KCl, 10% glycerol, 1 mM DTT, 200 µM PMSF, 1 µg/mL leupeptin. Cell lysis and extract clarification were performed as described for *Dm*ORC. Full-length MBP-tagged *Dm*Cdc6 was purified by amylose affinity chromatography on a 10 mL column (New England Biolabs). After extensive washing with lysis buffer, the salt concentration was lowered to 300 mM KCl and the protein eluted with 20 mM maltose in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 1 mM DTT. *Dm*Cdc6-containing fractions were pooled, concentrated in 30K Amicon Ultra-15 concentrators (Millipore), and the protein was further purified by gel filtration chromatography on a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 25 mM HEPES-KOH, pH 7.6, 500 mM potassium glutamate, 10% glycerol, 1 mM DTT.

To purify N-terminally trimmed *Dm*Cdc6, Hi5 cells expressing 6xHis-MBP-tagged *Dm*Cdc6ΔN241 were resuspended in 150 mL lysis buffer containing 50 mM Tris-HCl, pH 7.8, 600 mM KCl, 50 mM imidazole, pH 7.8, 10% glycerol, 5 mM magnesium acetate, 1 mM BME, 200 µM PMSF, and 1 µg/mL leupeptin. After clarifying the lysate by ultracentrifugation and 20% (v/v) ammonium sulfate precipitation, *Dm*Cdc6ΔN241 was purified by affinity chromatography using a 5 mL HisTrap HP column (GE Healthcare) and a 10 mL amylose column (New England Biolabs). The protein was eluted from these columns with a 50–250 mM imidazole gradient and 20 mM maltose, respectively, both in 50 mM Tris-HCl, pH 7.8, 600 mM KCl, 10% glycerol, 5 mM magnesium acetate, 1 mM BME. The affinity tags were removed by overnight cleavage with 6xHis-tagged TEV protease and a subsequent nickel-affinity step to remove the 6xHis-MBP tag, uncleaved protein, and TEV protease. The flow-through was concentrated in a 10K Amicon Ultra-15 concentrator (Millipore) and loaded onto a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 25 mM HEPES-KOH, pH 7.6, 300 mM potassium acetate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT. Peak *Dm*Cdc6ΔN241 fractions eluting from the gel filtration column were pooled, concentrated in a 10K Amicon Ultra-15 concentrators (Millipore), and flash frozen in liquid nitrogen for subsequent storage at −80 °C. A 6xHis-MBP-tagged version of *Dm*Cdc6ΔN241 was purified in a similar manner except that TEV cleavage and the second nickel-affinity chromatography step were omitted. All purification steps were performed at 4 °C.

Expression and purification of *Dm*Cdt1. *Drosophila* Cdt1 was expressed in BL21 RIL *E. coli* (cultured in Terrific Broth (TB)) upon induction with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 16 °C. Eight liters of culture 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, 200 µM PMSF, 1 µg/mL leupeptin) and lysed by sonication. The lysate was clarified by 30 min centrifugation at 30,996 × *g* in a Beckman Coulter Avanti J-26 XP centrifuge using a JA-17 rotor. Soluble Cdt1 in the supernatant was bound to a 5 mL HisTrap HP nickel-affinity chromatography column (GE Healthcare) and, after washing with 500 mL lysis buffer and 50 mL low salt buffer (50 mM Tris-HCl 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 *Dm*Cdt1 were pooled, concentrated, flash frozen, and stored at −80 °C. All purification steps were performed at 4 °C.

Expression and purification of *Dm*Mcm2–7. 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, 1 µg/mL leupeptin, 200 µM PMSF). After sonication, the lysate was clarified by two rounds of ultracentrifugation and ammonium sulfate precipitation as described for *Dm*ORC. 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 an 8 mL amylose column (New England Biolabs), washed in 50 mM HEPES-KOH, pH 7.5, 200 mM potassium acetate, 10% glycerol, 1 mM BME, and Mcm2–7 was eluted by adding 20 mM maltose. Affinity tags were removed by overnight digestion with 6xHis-tagged TEV protease, followed by nickel-affinity and gel filtration chromatography using a Superose 6 Increase 10/300 GL column (GE Healthcare) in 25 mM HEPES-KOH, pH 7.5, 300 mM potassium glutamate, 10% glycerol, 1 mM DTT. *Dm*Mcm2–7 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 *Dm*Mcm2–7 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 nM 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 micromolar concentrations necessary for cryo-EM studies using open hole 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 *Dm*Orc1 (referred to as Orc1ΔN) and the N-terminal 241 amino acid residues of *Dm*Cdc6 (Cdc6ΔN), which have been reported to drive liquid-liquid phase separation[32], improved the behavior of these proteins in the presence of DNA without impeding *Dm*ORC's ATPase activity, the ATP-dependent co-association of ORC, DNA, and Cdc6, and Mcm2–7 loading (Fig. 1, Supplementary Fig. 1). Therefore, these trimmed but Mcm2–7 loading-competent constructs were used to assemble *Dm*ORC·DNA and *Dm*ORC·DNA·Cdc6 complexes for cryo-EM experiments.

To reconstitute *Drosophila* ORC·DNA·Cdc6 assemblies, $Dm\text{ORC}^{\text{Orc1}\Delta\text{N}}$, duplex DNA (60 bp AT-rich, 84 bp AT-rich, 60 bp poly(dA·dT), or 60 bp GC-rich; Supplementary Table 2), and *Dm*Cdc6ΔN were combined at a final concentration of 2.5 µM, 3 µM, and 6.8 µM, respectively, in buffer containing 25 mM HEPES-KOH, pH 7.6, 250 mM potassium acetate, 10 mM 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 10/300GL 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 *Dm*ORC·DNA·Cdc6 complex by gel filtration chromatography. Peak fractions containing DNA (judged by absorbance at 254 nm), $Dm\text{ORC}^{\text{Orc1}\Delta\text{N}}$, and *Dm*Cdc6ΔN were pooled and concentrated in a 30K Amicon Ultra-4 concentrator (Millipore) to an absorbance of 0.7–1.4 at 280 nm for cryo-EM grid preparation. The ORC·DNA complex was assembled and purified in an analogous manner with the exception that Cdc6 was omitted in the reconstitution reaction.

Cryo-EM data collection and image processing. 3.5 µL of purified, concentrated ternary *Dm*ORC·DNA·Cdc6 or the binary *Dm*ORC·DNA complexes were applied to plasma-cleaned (using a 18.9%/81.1% H_2/O_2 gas mixture for 30 s at 5 W) 300-mesh R1.2/1.3 UltrAuFoil grids (Quantifoil Micro Tools GmbH), adsorbed for 10 s, and then frozen in liquid ethane using a Vitrobot Mark IV plunge freezer (Thermo Fisher Scientific). Cryo-EM grids were imaged in a Titan Krios transmission electron microscope (Thermo Fisher Scientific) operated at an acceleration voltage of 300 kV and equipped with a spherical aberration (Cs) corrector (CEOS GmbH), a Quantum LS energy filter, and a post-GIF Gatan K2 summit 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 a Volta phase plate using a target defocus of 0.4–0.6 µm (with phase plate, aiming for a phase shift of 20–130 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 $e^-/Å^2$ per second, yielding a total electron dose of 40 $e^-/Å^2$ and 50 $e^-/Å^2$, respectively (see also Supplementary Table 1 for additional information on data collection settings). Data of the *Dm*ORC·DNA·Cdc6 complex assembled with poly(dA·dT) DNA was collected as described above with the exception that a Falcon 3EC direct electron detector (Thermo Fisher Scientific) in electron counting mode was used to record dose-fractionated movies at a target defocus of 0.8–1.5 µm. Movies were recorded as 50 frames over 41.47 s at a dose rate of 1.21 $e^-/Å^2$ per second, resulting in a total dose of 50.2 $e^-/Å^2$.

Recorded movie frames were motion-corrected using MotionCor2[60], and contrast transfer function parameters, including phase shifts, were determined from averaged movie frames with GCTF[61]. Subsequently, averaged movie images and power spectra were manually inspected and micrographs containing aggregated protein, severe ice contamination, or having GCTF resolution estimates of worse than 4–6 Å were excluded from further processing. For the first dataset, the ternary *Dm*ORC·DNA·Cdc6 complex assembled on 60 bp AT-rich DNA (dataset 1), particles were initially automatically picked in a reference-free manner using GAUTOMATCH (developed by Kai Zhang, https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) and subjected to 2D classification. Class averages representing different views of the ternary complex were selected, low-pass filtered to 30 Å, and served as templates for reference-based particle selection with GAUTOMATCH for all datasets. All further EM image processing steps, including particle extraction, 2D and 3D classification, 3D refinement, particle polishing, and postprocessing were performed using RELION 2.1 and RELION 3.0[62–64]. 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 retaining particles corresponding to underrepresented 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 *Dm*ORC·DNA·Cdc6 dataset 1. For all other datasets, the cryo-EM volume of the ternary complex, low-pass filtered to 50 Å, provided the starting reference for 3D classification. Particles of 3D class volumes with clearly visible secondary structure elements (cleaned particles) were subjected to 3D autorefinement (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 maps and local resolutions of masked sub-regions were determined 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 *Dm*ORC crystal structure (PDB 4xgc[16]) into the cryo-EM density map of *Dm*ORC·DNA·Cdc6 (with 60 bp AT-rich dsDNA, dataset 1) using UCSF Chimera[65,66]. The *Dm*ORC model was then manually rebuilt in COOT[67] to accommodate for structural changes in *Dm*ORC between the *apo*-ORC state and the ternary complex structure, and to build additional segments that were poorly or not resolved in the *Dm*ORC crystal structure. For model building, a combination of unsharpened, globally sharpened, and locally sharpened (with LocScale[68]) cryo-EM density maps were used. *Dm*Cdc6 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 *Dm*Cdc6 and *Dm*Orc1 provided additional information on the topological fold of *Dm*Cdc6 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 *Dm*ORC·DNA·Cdc6 model was improved by iterative rounds of real-space refinement against the sharpened cryo-EM map in PHENIX[69,70] using secondary structure restraints 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[71]; 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 *Dm*ORC·Cdc6.

Models for all other *Dm*ORC·DNA·Cdc6 assemblies, for the binary *Dm*ORC·DNA complex, and for DNA-free *Dm*ORC 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 (except in those containing poly(dA·dT) duplexes) is speculative due to the sequence-independent DNA binding activity of *Dm*ORC. Consequently, individual particles in the 3D reconstructions likely bind DNA in different registers, which is consistent with the higher model and map B-factors in this region of the complex. Moreover, although full-length *Dm*Orc6 was used to reconstitute the binary and ternary complexes, no density was visible for the Orc6 TFIIB 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 PDBePISA server[72] implemented in COOT. Protein–DNA contacts were explored using PDBePISA and DNAproDB[73]. Multiple protein sequence alignments were performed with MAFFT[74,75] and visualized in JALVIEW[76] or by generating sequence logos with WEBLOGO[77]. DNA shape parameters were determined using CURVES+[78,79] and 3DNA[80,81]. Figures were rendered using PyMOL, UCSF Chimera, and UCSF ChimeraX[82].

Annealing of DNA duplexes. DNA duplexes used for reconstituting DNA-bound *Dm*ORC 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 $MgCl_2$. 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 *Dm*ORC, or of truncated *Dm*ORC$^{Orc1\Delta N}$ and *Dm*ORC$^{Orc6\Delta TFIIB}$, to DNA was measured by fluorescence anisotropy using 40 bp dsDNAs (Supplementary Table 2) with one strand 5′-labeled with fluorescein similarly as done previously[18]. 1 nM dsDNA was mixed with increasing concentrations of *Dm*ORC (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, 0.1 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 PHERAstar FSX plate reader (BMG Labtech). To obtain apparent dissociation constant ($K_{d,\,app}$) for *Dm*ORC 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:

$$A = A_f + \frac{(A_b - A_f) * [R]^h}{(K_{d,app}^h + [R]^h)} \quad (1)$$

where A_f and A_b are the measured anisotropy of free and bound, fluorescently labeled DNA duplex, respectively, [R] is the concentration of *Dm*ORC, $K_{d,\,app}$ is the apparent binding constant, and h is the Hill coefficient. For the poly(dA·dT) duplex, we additionally fit the binding data points to the quadratic binding equation to account for potential *Dm*ORC depletion at low concentrations due to its high affinity for this DNA:

$$A = A_f + (A_b - A_f) * \frac{[L] + K_{d,app} + [R] - \sqrt{([L] + K_{d,app} + [R])^2 - 4 * [L] * [R]}}{2 * [L]} \quad (2)$$

where A_f and A_b 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 *Dm*ORC, and $K_{d,\,app}$ is the apparent binding constant. $K_{d,\,app}$ values obtained for both methods were with 2.1 nM (Hill equation) and 1.7 nM (quadratic equation) comparable. For simplicity, only Hill curves are shown in Supplementary Fig. 10a–c. For bar graphs, $K_{d,\,app}$ and standard error (SE) of fits are plotted. DNA binding of *Dm*ORC protein in the presence of *Dm*Cdc6 was measured as described above with the exception that wild type or truncated *Dm*Cdc6 was added at equimolar concentrations to *Dm*ORC in the reaction mix.

Pull-down assays. To verify that N-terminal truncations of *Dm*Orc1 and *Dm*Cdc6 do not interfere with ternary complex formation, we performed pull-down assays using MBP-tagged *Dm*Cdc6 or 6xHis-MBP-tagged *Dm*Cdc6ΔN as baits. 400 µL reactions were assembled with 60 bp AT-rich dsDNA (Supplementary Table 2) and different combinations of full-length *Dm*ORC or *Dm*ORC$^{Orc1\Delta N}$, and full-length MBP-*Dm*Cdc6 or 6xHis-MBP-*Dm*Cdc6ΔN at final concentrations of 100 nM, 100 nM, and 200 nM, respectively, in binding buffer containing 25 mM HEPES-KOH, pH 7.6, 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT, and 0 or 1 mM ATP. *Dm*ORC and DNA were incubated for 15 min at 25 °C prior to the addition of *Dm*Cdc6. After an additional 15 min at 25 °C, reactions were added to 25 µL amylose beads (80% slurry, New England Biolabs) and incubated for 30 min at 25 °C. Beads were then washed twice with 1 mL binding buffer, followed by elution of bound proteins in 25 µL binding buffer supplemented with 20 mM maltose. 0.5% input and 32% eluted proteins were analyzed by 10% SDS-PAGE and silver staining. Pull-down assays were performed three times as independent experiments.

ATPase activity assays. Steady-state ATP hydrolysis activity of *Dm*ORC was measured using the enzymatic NADH-coupled assay[83]. In this system, ATP hydrolyzed by *Dm*ORC is regenerated by pyruvate kinase, a reaction that is in turn coupled to the lactate dehydrogenase-mediated oxidation of NADH to NAD^+, which is measured spectrophotometrically. Full-length wild type or mutant *Dm*ORC, or *Dm*ORC$^{Orc1\Delta N}$, were diluted to 2 µM and mixed with 2.4–4 units/mL pyruvate kinase/3.6–5.6 units/mL lactate dehydrogenase (PK/LDH from rabbit muscle, Sigma Aldrich), 8 mM phosphoenolpyruvate (Sigma Aldrich), 0.6 mM NADH (Sigma Aldrich) in buffer containing 25 mM HEPES-KOH, pH 7.6, 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 0.2% NP40, 2 mM DTT, 0.2 mg/mL BSA. 25 µL of this 2x *Dm*ORC/PK/LDH mix were transferred to a 96-well, half area microplate (Greiner) and reactions were initiated by adding 25 µL of 2x-ATP titrations (in two-fold steps, either in the absence or presence of 4 µM DNA duplex) in 25 mM HEPES-KOH pH 7.6, 300 mM potassium glutamate, 10% glycerol, 10 mM magnesium acetate. Final reactions (50 µL) therefore contained 1 µM *Dm*ORC assemblies, 1.2–2 units/mL PK/1.8–2.8 units/mL LDH, 4 mM phosphoenolpyruvate, and 0.3 mM NADH in 25 mM HEPES-KOH, pH 7.6, 300 mM potassium glutamate, 10% glycerol, 10 mM magnesium acetate, 0.1% NP40, 1 mM DTT, 0.1 mg/mL BSA. Final ATP concentrations in titration reactions ranged from 4 mM to 2 µM, serially 2-fold diluted. For DNA inhibition experiments, the 60 bp AT-rich DNA duplex was included in reactions at a final concentration of 2 µM. The decrease in NADH concentration resulting from ATP hydrolysis was monitored by measuring absorbance at 340 nm in a PHERAstar FSX plate reader (BMG Labtech) at 27 °C for 2 h at 1 min intervals. Raw absorbance values were converted to NADH concentrations using a standard curve of NADH titrations (in 2-fold steps, as well as 0 mM) in assay buffer conditions included on each 96-well plate measured. ATP hydrolysis rates were extracted for each ATP titration from the linear portions of the NADH consumption curves. The means and standard deviations of hydrolysis rates from three or more independent experiments were plotted as a function of ATP concentration and fit to the standard Michaelis–Menten equation using GraphPad Prism:

$$v = \frac{V_{max} * [ATP]}{K_M + [ATP]} \quad (3)$$

where v is the rate of ATP hydrolysis at a given substrate concentration (ATP), V_{max} is the maximum velocity of ATP hydrolysis, and K_M is the Michaelis–Menten constant. For bar graphs, k_{cat} and the standard error of the corresponding fits are plotted. Steady-state ATP hydrolysis activity of *Dm*Cdc6 and *Dm*Cd6ΔN were determined as described for *Dm*ORC. To investigate steady-state ATP hydrolysis of full-length and truncated ternary *Dm*ORC·DNA·*Dm*Cdc6 complexes, assays were set up in the presence of DNA as described above with the exception that 2 µM *Dm*Cdc6 or *Dm*Cd6ΔN was added to the 2x *Dm*ORC/PK/LDH mix.

Mcm2–7 loading assay. In vitro Mcm2–7 loading assays using *Drosophila* proteins were adapted from those established previously for the *S. cerevisiae* system[46,84]. 0.6 pmoles of biotinylated duplex DNA (Supplementary Table 2) were coupled to 15 µL streptavidin sepharose high performance beads (GE Healthcare) in 20 µL coupling buffer (5 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 1 M NaCl, 0.1% NP40) for 30 min at room temperature. After washing with coupling buffer to remove free DNA, 0.6 pmoles streptavidin (Sigma Aldrich) were added for 15 min to block any free DNA ends and prevent Mcm2–7 sliding off after loading. Unbound streptavidin was removed and beads equilibrated by washing beads thrice in low salt buffer (25 mM HEPES-KOH, pH 7.6, 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 0.1% NP40, 1 mM DTT) in the absence or presence of 1 mM ATP. *Dm*ORC, *Dm*Cdc6, *Dm*Cdt1, and *Dm*Mcm2–7 were mixed at 250 nM, 250 nM, 500 nM, and 500 nM final concentration in low salt buffer with or without 1 mM ATP. 40 µL of the loading protein mix were added to DNA-coupled beads and incubated for 30 min at 27 °C. To assess Mcm2–7 loading, beads were washed once with high salt buffer (25 mM HEPES-KOH, pH 7.6, 1 M KCl, 10 mM magnesium acetate, 10% glycerol, 0.1% NP40, 1 mM DTT, 0 or 1 mM ATP), followed by two washes with low salt buffer (25 mM HEPES-KOH, pH 7.6, 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 0.1% NP40, 1 mM DTT, 0 or 1 mM ATP). To analyze Mcm2–7 recruitment, beads were washed with low salt buffer instead of high salt buffer. Bound proteins were eluted by digestion with 1000 units micrococcal nuclease (New England Biolabs) for 10 min at 37 °C in 25 mM HEPES-KOH, pH 7.6, 300 mM KCl, 10 mM magnesium acetate,

10% glycerol, 0.1% NP40, 5 mM $CaCl_2$, 1 mM DTT, 0 or 1 mM ATP, and analyzed by SDS-PAGE and silver staining. All loading assays were performed at least twice as independent experiments. 4 µL of eluted proteins were also applied to glow-discharged continuous carbon EM grids and stained with 2% uranyl formate prior to imaging in a FEI Tecnai T12 Spirit transmission electron microscope to verify double hexamer formation (Fig. 1c, Supplementary Fig. 1e).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support this study is available from the corresponding author upon reasonable request. The cryo-EM maps and model coordinates have been deposited in the Electron Microscopy Databank (EMDB) and Protein Databank (PDB), respectively, under the following accession numbers: EMD-22361 and PDB 7JK4 (ORC·DNA·Cdc6, 60 bp AT-rich), EMD-22362 and PDB 7JK5 (ORC·DNA, 60 bp AT-rich), EMD-22363 and PDB 7JK6 (ORC, active conformation), EMD-22329 and PDB 7JGR (ORC·DNA·Cdc6, 84 bp AT-rich), EMD-22360 and PDB 7JK3 (ORC·DNA·Cdc6, 60 bp GC-rich), EMD-22359 and PDB 7JK2 (ORC·DNA·Cdc6, 60 bp poly(dA·dT) state 1), and EMD-22330 and PDB 7JGS (ORC·DNA·Cdc6, 60 bp poly(dA·dT) state 2). Plasmids generated in this study are available upon request from the corresponding author. Source data are provided with this paper.

Received: 16 June 2020; Accepted: 3 August 2020;
Published online: 26 August 2020

References

1. Bleichert, F., Botchan, M. R. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Science* **355**, eaah6317 (2017).
2. O'Donnell, M., Langston, L. & Stillman, B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, a010108 (2013).
3. Duderstadt, K. E. & Berger, J. M. A structural framework for replication origin opening by AAA+ initiation factors. *Curr. Opin. Struct. Biol.* **23**, 144–153 (2013).
4. On, K. F., Jaremko, M., Stillman, B. & Joshua-Tor, L. A structural view of the initiators for chromosome replication. *Curr. Opin. Struct. Biol.* **53**, 131–139 (2018).
5. Fujikawa, N. et al. Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res.* **31**, 2077–2086 (2003).
6. Dueber, E. L., Corn, J. E., Bell, S. D. & Berger, J. M. Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science* **317**, 1210–1213 (2007).
7. Gaudier, M., Schuwirth, B. S., Westcott, S. L. & Wigley, D. B. Structural basis of DNA replication origin recognition by an ORC protein. *Science* **317**, 1213–1216 (2007).
8. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**, 27–43 (1999).
9. Iyer, L. M., Leipe, D. D., Koonin, E. V. & Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* **146**, 11–31 (2004).
10. Erzberger, J. P. & Berger, J. M. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).
11. Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. Structure and function of the AAA+ nucleotide binding pocket. *Biochim Biophys. Acta* **1823**, 2–14 (2012).
12. Duderstadt, K. E., Chuang, K. & Berger, J. M. DNA stretching by bacterial initiators promotes replication origin opening. *Nature* **478**, 209–213 (2011).
13. Riera, A. et al. From structure to mechanism-understanding initiation of DNA replication. *Genes Dev.* **31**, 1073–1088 (2017).
14. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).
15. Duncker, B. P., Chesnokov, I. N. & McConkey, B. J. The origin recognition complex protein family. *Genome Biol.* **10**, 214 (2009).
16. Bleichert, F., Botchan, M. R. & Berger, J. M. Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321–326 (2015).
17. Tocilj, A. et al. Structure of the active form of human origin recognition complex and its ATPase motor module. *eLife* **6**, e20818 (2017).
18. Bleichert, F., Leitner, A., Aebersold, R., Botchan, M. R. & Berger, J. M. Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc. Natl Acad. Sci. USA* **115**, E5906–E5915 (2018).
19. Bleichert, F. et al. A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *eLife* **2**, e00882 (2013).
20. Li, N. et al. Structure of the origin recognition complex bound to DNA replication origin. *Nature* **559**, 217–222 (2018).
21. Yuan, Z. et al. Structural basis of Mcm2–7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nat. Struct. Mol. Biol.* **24**, 316–324 (2017).
22. Kuo, A. J. et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484**, 115–119 (2012).
23. Balasov, M., Huijbregts, R. P. & Chesnokov, I. Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in *Drosophila melanogaster*. *Mol. Cell Biol.* **27**, 3143–3153 (2007).
24. Chuang, R. Y. & Kelly, T. J. The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc. Natl Acad. Sci. USA* **96**, 2656–2661 (1999).
25. Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *EMBO J.* **23**, 897–907 (2004).
26. MacAlpine, H. K., Gordan, R., Powell, S. K., Hartemink, A. J. & MacAlpine, D. M. *Drosophila* ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Res.* **20**, 201–211 (2010).
27. Miotto, B., Ji, Z. & Struhl, K. Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. *Proc. Natl Acad. Sci. USA* **113**, E4810–E4819 (2016).
28. Cayrou, C. et al. The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome Res.* **25**, 1873–1885 (2015).
29. Dellino, G. I. et al. Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. *Genome Res.* **23**, 1–11 (2013).
30. Vashee, S. et al. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* **17**, 1894–1908 (2003).
31. Ekundayo, B. & Bleichert, F. Origins of DNA replication. *PLoS Genet.* **15**, e1008320 (2019).
32. Parker, M. W. et al. A new class of disordered elements controls DNA replication through initiator self-assembly. *eLife* **8**, e48562 (2019).
33. Erzberger, J. P., Mott, M. L. & Berger, J. M. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* **13**, 676–683 (2006).
34. Chesnokov, I., Remus, D. & Botchan, M. Functional analysis of mutant and wild-type Drosophila origin recognition complex. *Proc. Natl Acad. Sci. USA* **98**, 11997–12002 (2001).
35. Klemm, R. D., Austin, R. J. & Bell, S. P. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell* **88**, 493–502 (1997).
36. Perkins, G. & Diffley, J. F. Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders. *Mol. Cell* **2**, 23–32 (1998).
37. Bowers, J. L., Randell, J. C., Chen, S. & Bell, S. P. ATP hydrolysis by ORC catalyzes reiterative Mcm2–7 assembly at a defined origin of replication. *Mol. Cell* **16**, 967–978 (2004).
38. Randell, J. C., Bowers, J. L., Rodriguez, H. K. & Bell, S. P. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2–7 helicase. *Mol. Cell* **21**, 29–39 (2006).
39. Speck, C. & Stillman, B. Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *J. Biol. Chem.* **282**, 11705–11714 (2007).
40. Rao, H. & Stillman, B. The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proc. Natl Acad. Sci. USA* **92**, 2224–2228 (1995).
41. Rowley, A., Cocker, J. H., Harwood, J. & Diffley, J. F. Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO J.* **14**, 2631–2641 (1995).
42. Kawakami, H., Ohashi, E., Kanamoto, S., Tsurimoto, T. & Katayama, T. Specific binding of eukaryotic ORC to DNA replication origins depends on highly conserved basic residues. *Sci. Rep.* **5**, 14929 (2015).
43. Austin, R. J., Orr-Weaver, T. L. & Bell, S. P. Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev.* **13**, 2639–2649 (1999).
44. Remus, D., Blanchette, M., Rio, D. C. & Botchan, M. R. CDK phosphorylation inhibits the DNA-binding and ATP-hydrolysis activities of the *Drosophila* origin recognition complex. *J. Biol. Chem.* **280**, 39740–39751 (2005).
45. Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339–343 (2013).
46. Coster, G. & Diffley, J. F. X. Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science* **357**, 314–318 (2017).
47. Miller, T. C. R., Locke, J., Greiwe, J. F., Diffley, J. F. X. & Costa, A. Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature* **575**, 704–710 (2019).

48. Tubbs, A. et al. Dual roles of poly(dA:dT) tracts in replication initiation and fork collapse. *Cell* **174**, 1127–1142 e19 (2018).
49. Comoglio, F. et al. High-resolution profiling of *Drosophila* replication start sites reveals a DNA shape and chromatin signature of metazoan origins. *Cell Rep.* **11**, 821–834 (2015).
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 metazoan 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. Duzdevich, 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**, 9294307 (2016).
59. Arias-Palomo, 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. MotionCor2: 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-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).
63. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
64. Zivanov, 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).
65. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157**, 281–287 (2007).
66. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
67. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
68. Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *eLife* **6**, e27131 (2017).
69. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
70. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D. Struct. Biol.* **74**, 531–544 (2018).
71. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
72. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
73. Sagendorf, J. M., Berman, H. M. & Rohs, R. DNAproDB: an interactive tool for structural analysis of DNA-protein complexes. *Nucleic Acids Res.* **45**, W89–W97 (2017).
74. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
75. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* **9**, 286–298 (2008).
76. 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).
77. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
78. 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).
79. 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).
80. Colasanti, A. V., Lu, X. J. & Olson, W. K. Analyzing and building nucleic acid structures with 3DNA. *J. Vis. Exp.* e4401 (2013).
81. 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).
82. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
83. Kiianitsa, K., Solinger, J. A. & Heyer, W. D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Anal. Biochem.* **321**, 266–271 (2003).
84. Coster, G., Frigola, J., Beuron, F., Morris, E. P. & Diffley, J. F. Origin licensing requires ATP binding and hydrolysis by the MCM replicative helicase. *Mol. Cell* **55**, 666–677 (2014).
85. Kelch, 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 Cdc6 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.; investigation, 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.

Structural Mechanism for Replication Origin Binding and Remodeling by a Metazoan Origin Recognition Complex and its Co-loader Cdc6

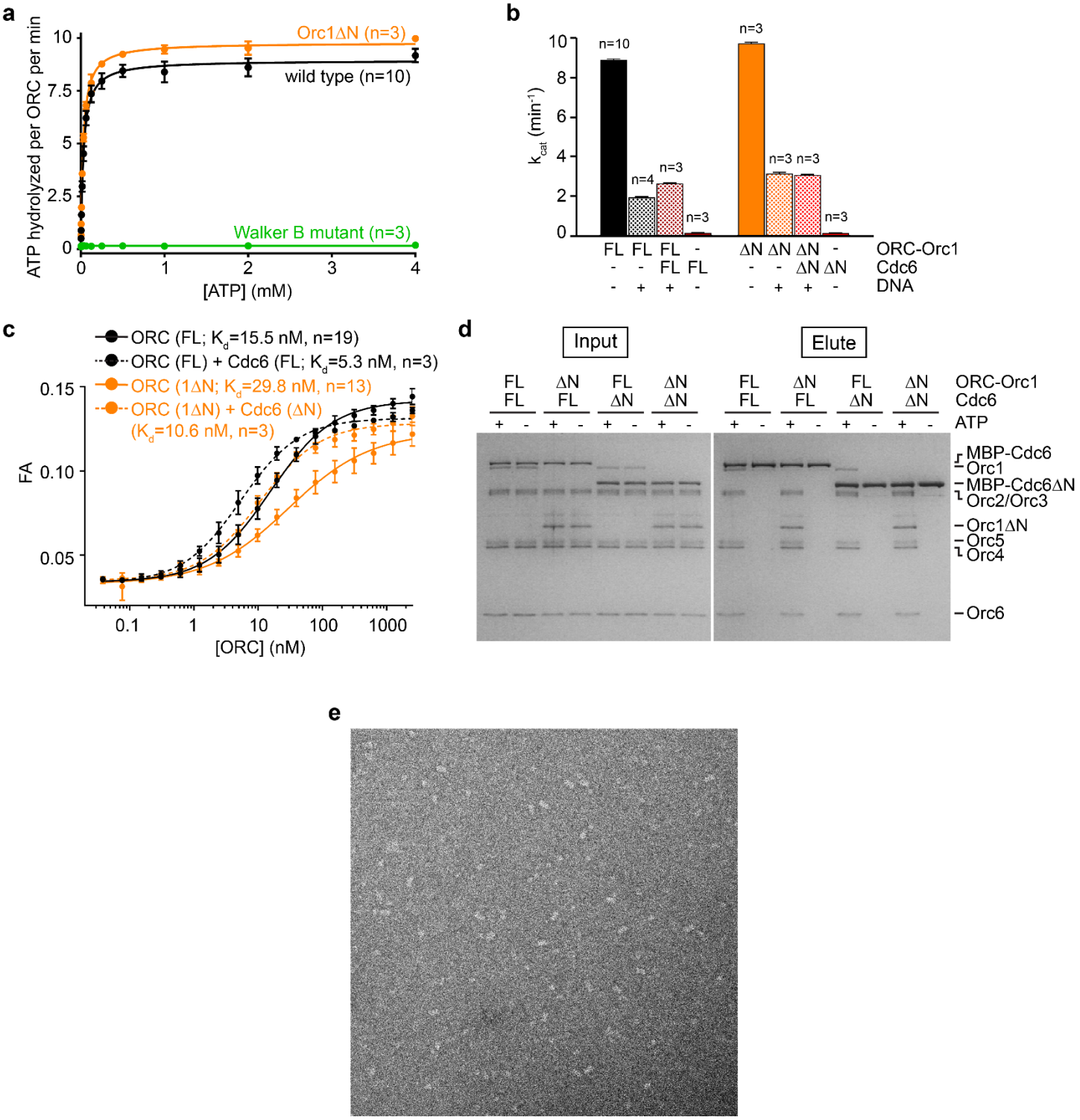
Schmidt and Bleichert

Supplementary Information

Supplementary Figures 1-10

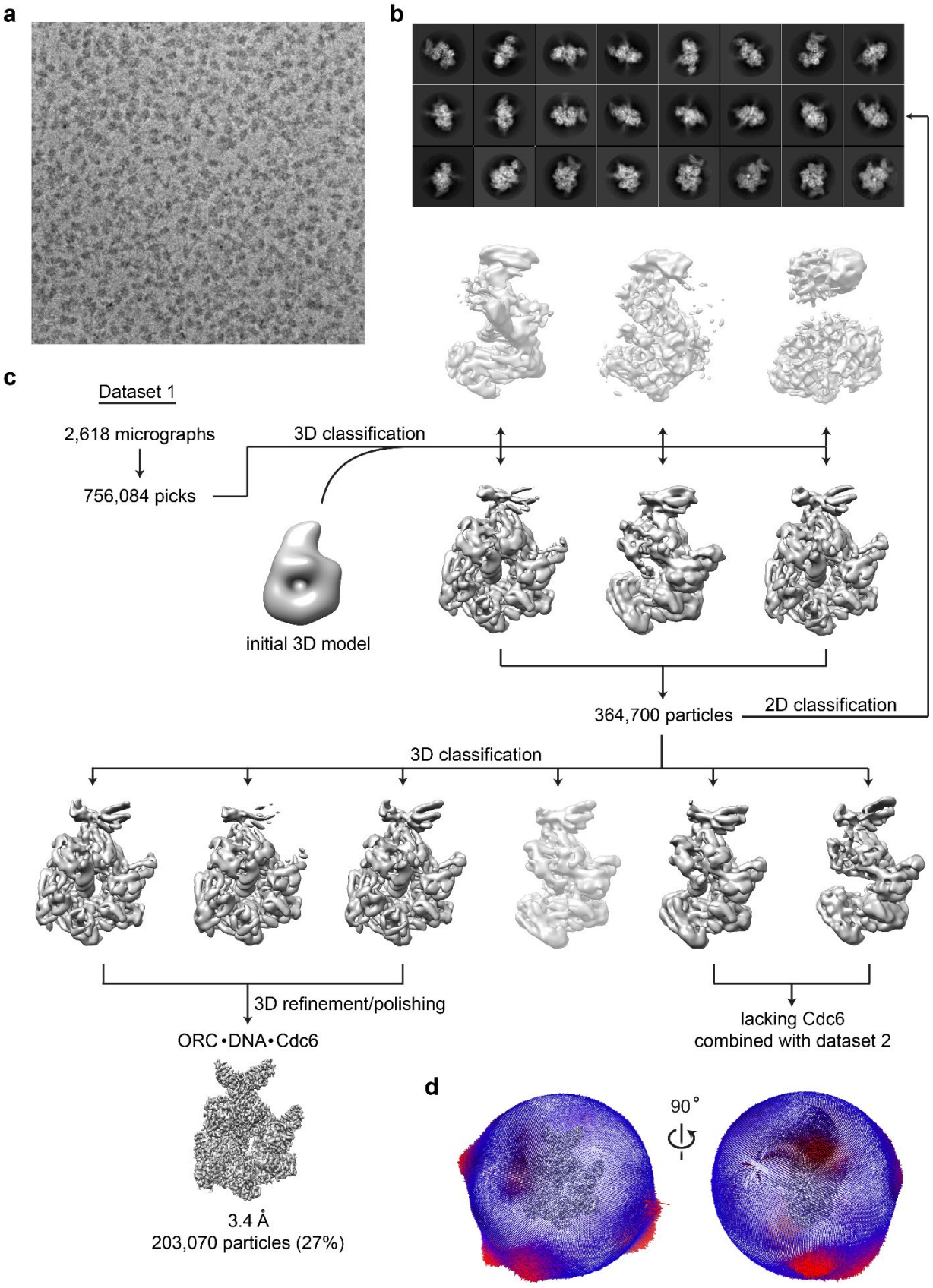
Supplementary Tables 1-2

Suplementary References



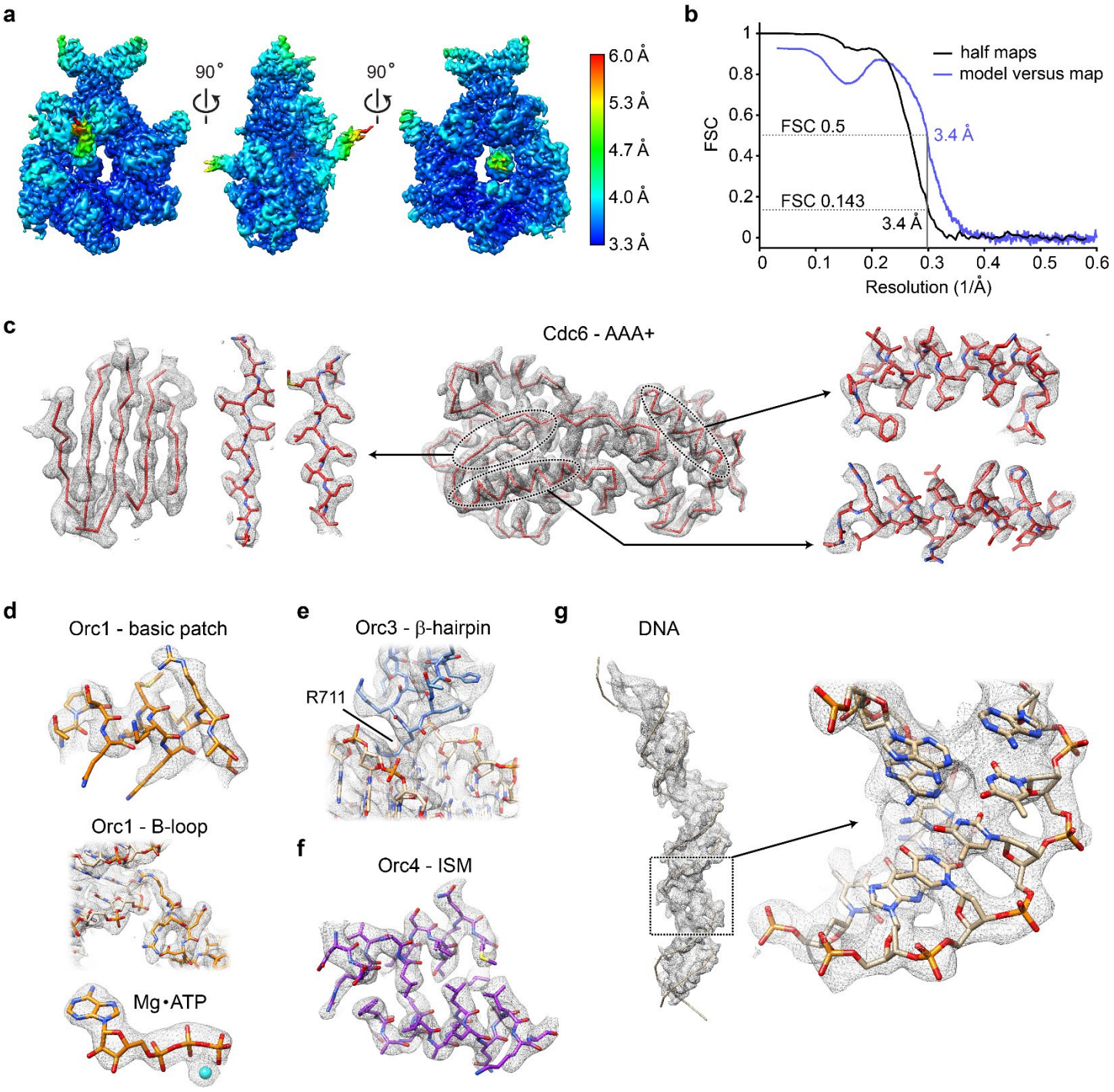
Supplementary Figure 1. **Reconstitution of an Mcm2-7 loading-competent, ternary *Dm*ORC· DNA·Cdc6 complex. a**) Full-length (wild type) *Dm*ORC and *Dm*ORC lacking the N-terminal 439 amino acids of Orc1 (Orc1ΔN) hydrolyze ATP with similar rates (~9-10 ATP hydrolyzed per min per ORC at saturating ATP concentrations) following Michaelis-Menten kinetics. *Dm*ORC bearing a mutation in the conserved Walker B motif of Orc1 (Orc1^{D684A}) is catalytically inactive. The means and standard deviations of ATP turnover rates of independent replicates (n is indicated) are plotted as a function of ATP concentration and fit to the Michaelis-Menten equation. **b**) ATPase activity of ORC is altered by DNA but not Cdc6. k_{cat} and S.E. of Michaelis-Menten equation fits to ATP titrations from independent experiments (n is indicated) are plotted. **c** and **d**) Removal of

the *Dm*Orc1 and *Dm*Cdc6 N-termini does not interfere with ternary complex formation. In **c**, ATP-dependent DNA binding by *Dm*ORC in the absence or presence of Cdc6 was measured by fluorescence anisotropy. Removal of the Orc1 N-terminal region has a <2-fold effect on *Dm*ORC's affinity for DNA, while addition of *Dm*Cdc6 or of N-terminally truncated *Dm*Cdc6 (*Dm*Cdc6ΔN) slightly stabilizes ORC on DNA (~3-fold decrease in $K_{d,\,app}$). Mean and standard deviations of n independent replicates (see figure for n) are plotted, and apparent dissociation constants ($K_{d,\,app}$) are listed. In **d**, Pull-down assays using MBP-tagged Cdc6 as bait show ATP-dependent co-purification of *Dm*ORC in the presence of 60 bp AT-rich dsDNA to similar extents when full-length (FL) or truncated (ΔN) Cdc6 or Orc1 (in the context of hexameric ORC) are used. Input (0.5%) and eluted proteins were separated by SDS-PAGE and visualized by silver staining. **e**) Negative-stain electron micrograph of loaded Mcm2-7. Source data are provided as a Source Data file.

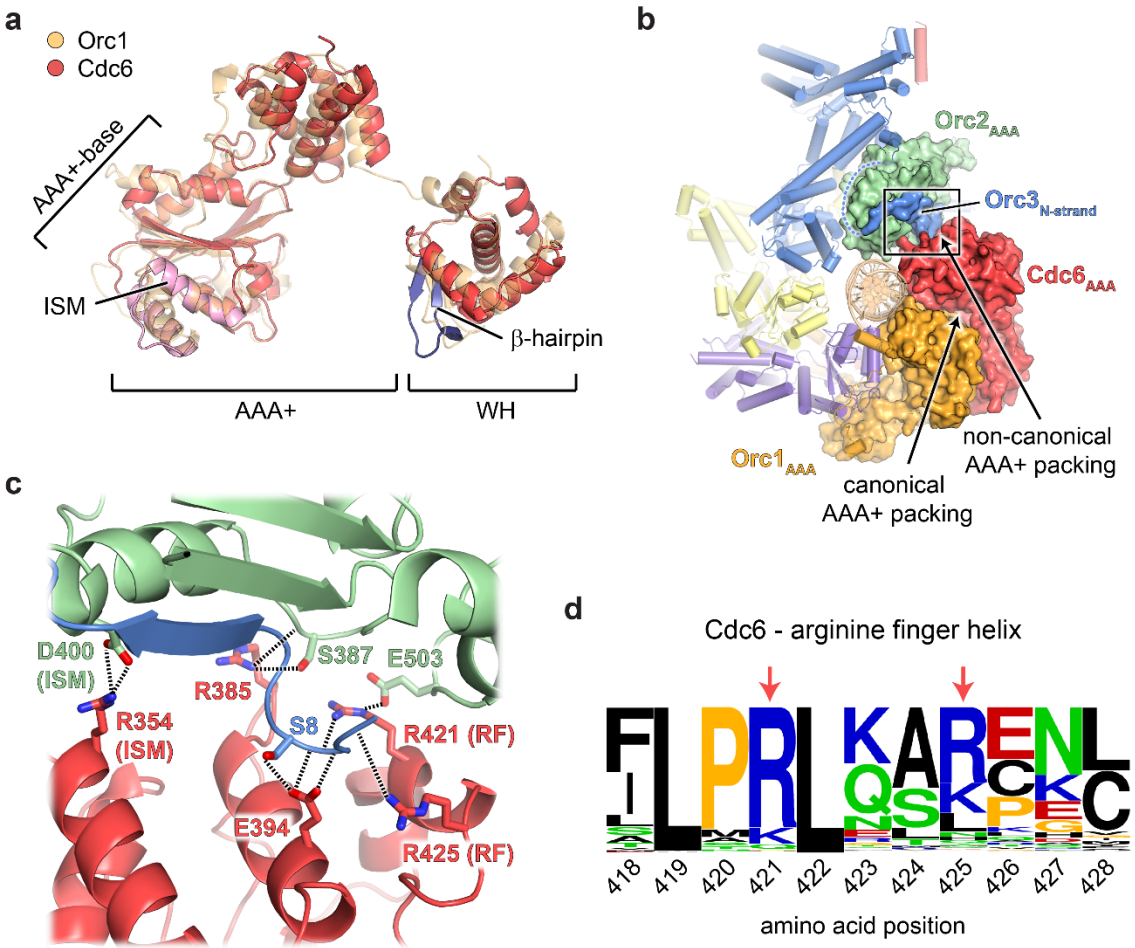


Supplementary Figure 2. **Cryo-EM data collection and processing workflow for the ternary *Dm*ORC·DNA·Cdc6 complex assembled on an AT-rich 60 bp duplex (dataset 1).** Representative **a**) electron micrograph and **b**) 2D class averages. **c**) Two rounds of 3D classification were performed to sort particles, and

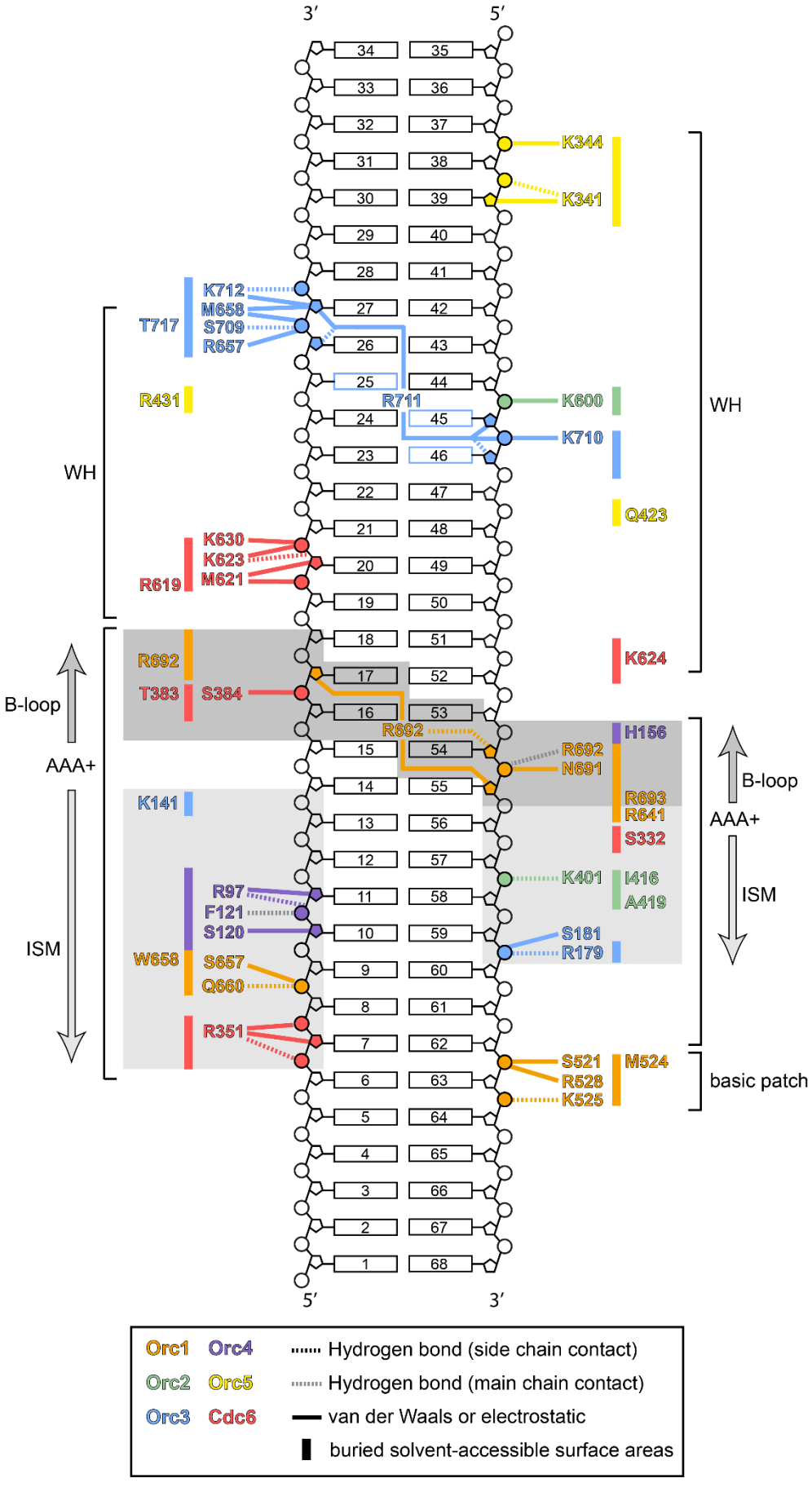
those in well-resolved 3D classes (dark grey) corresponding to DNA-free *Dm*ORC, the binary *Dm*ORC·DNA assembly, and the ternary *Dm*ORC·DNA·Cdc6 complex were kept for further processing. Particles from other classes (light grey) were discarded. *Dm*ORC·DNA·Cdc6 particles were subjected to 3D refinement and polishing, yielding a 3D map with an overall resolution of 3.4 Å. *Dm*ORC and *Dm*ORC·DNA particles were combined prior to 3D refinement with respective classes from a second dataset that was collected using the binary *Dm*ORC·DNA complex (see also **Supplementary Fig. 6**). **d**) Angular distribution of particles contributing to the final 3D reconstruction of *Dm*ORC·DNA·Cdc6. Similar processing schemes as outlined here were used for all other datasets.



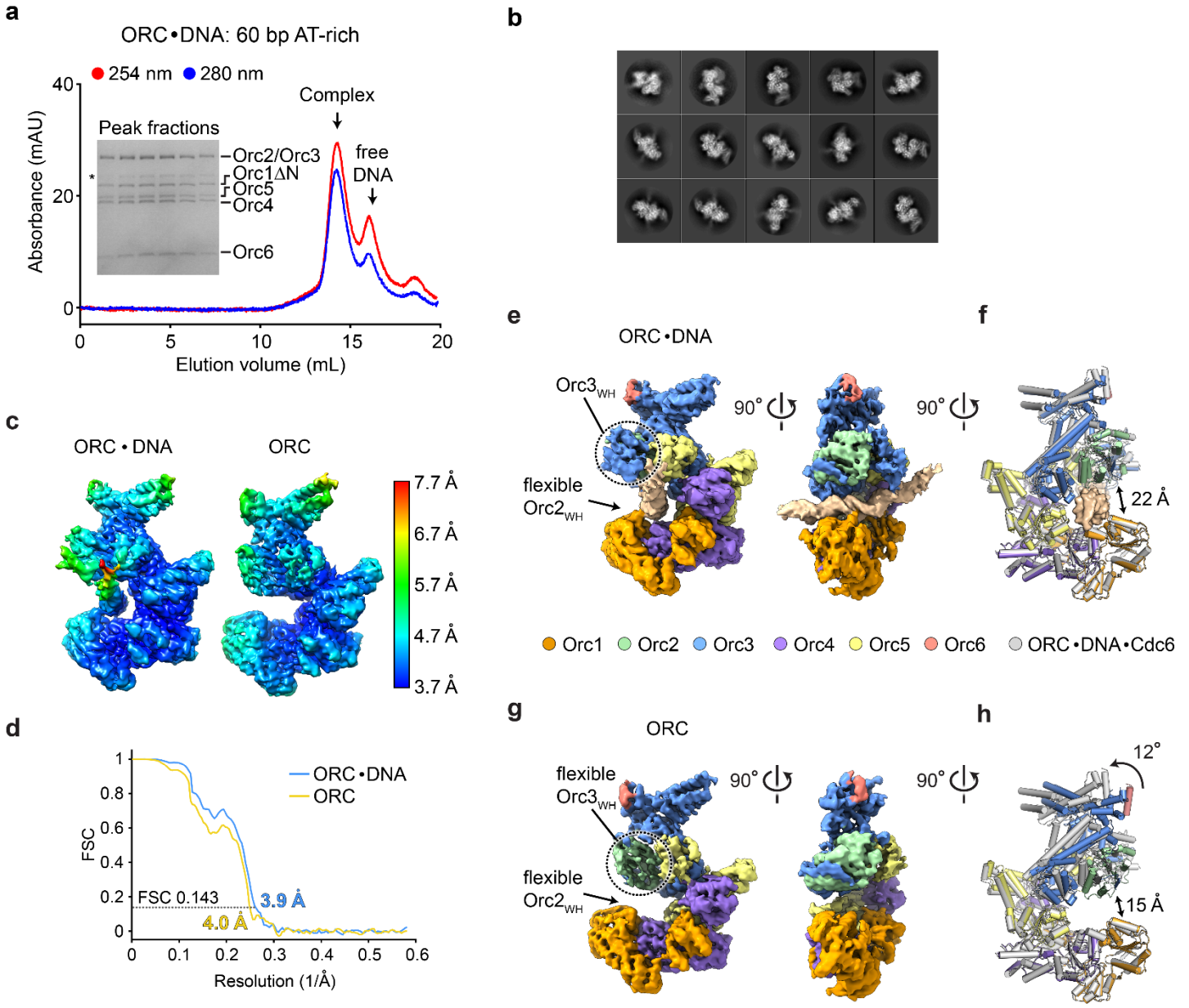
Supplementary Figure 3. **Local resolution, model building, and refinement of *Dm*ORC·DNA·Cdc6. a**) Unsharpened EM map of the ternary complex colored by local resolution. **b**) Fourier shell correlation (FSC) curves calculated using EM half-maps, or the full EM map and a pdb-derived model map. The resolutions at $FSC_{0.143}$ (for EM half-maps) and $FSC_{0.5}$ (for model versus EM map) are indicated. **c** to **g**) EM density map (sharpened and filtered to local resolution) and model is shown for various regions of the *Dm*ORC·DNA·Cdc6 structure.



Supplementary Figure 4. **Molecular interactions between *Dm*Cdc6 and *Dm*ORC. a**) Superposition of Orc1 and Cdc6 by aligning their AAA+ base regions emphasizes the structural similarity of both proteins. The AAA+ initiator specific motif (ISM) and WH β-hairpin in Cdc6, two signature motifs in related archaeal and metazoan initiators, are colored pink and deep blue, respectively. **b**) *Dm*Cdc6 is recruited to *Dm*ORC by extensive, canonical AAA+/AAA+ interactions with Orc1, as well as by non-canonical AAA+/AAA+ packing with the Orc2 AAA+ domain and an additional β-strand in the Orc3 N-terminus. Modules involved in interactions are rendered as molecular surface, while other subunits and DNA are depicted as cartoon. The linker between the N-terminal Orc3 β-strand and the AAA+ core module is flexible (dashed line). **c**) Zoomed view of boxed region in **b** reveals numerous contacts between Cdc6 and the composite binding site formed by Orc2 and the Orc3 N-terminus. Residues in the ISM helix and arginine finger helix of Cdc6 contribute to the bonding network. **d**) Sequence frequency logo of a Cdc6 multiple sequence alignment illustrates the conservation of the arginines (red arrows) in the RF-helix that participate in interactions with Orc2 and Orc3. ISM – initiator specific motif, RF – arginine finger.

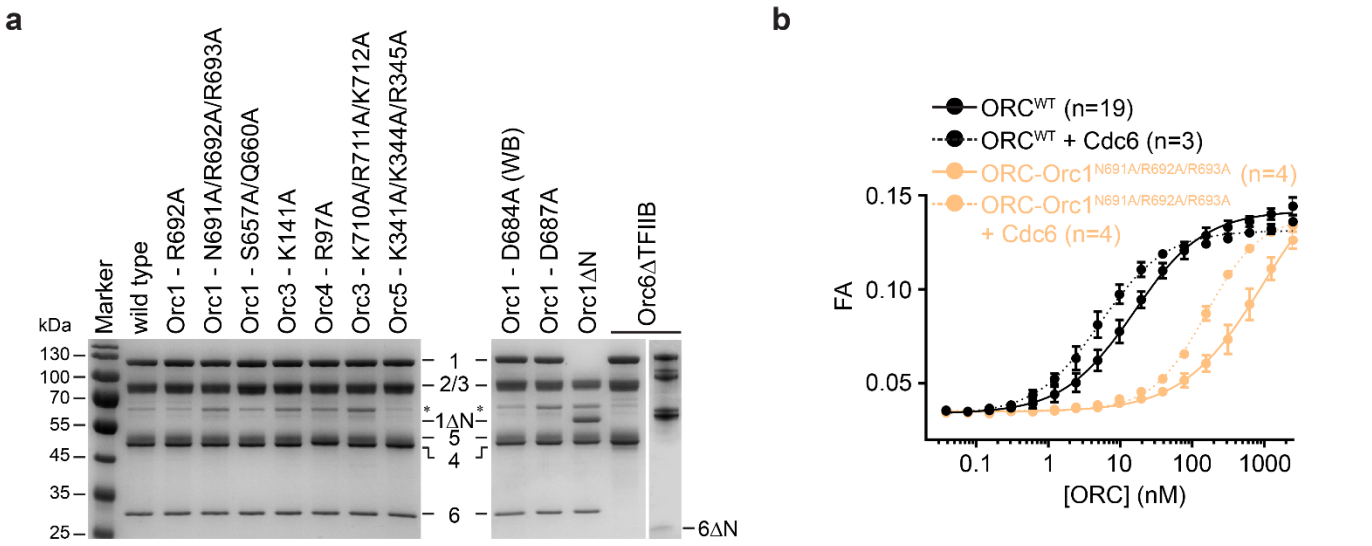

3′
5′
34
35
33
36
32
37
31
38
30
39
29
40
28
41
27
42
26
43
25
44
24
45
23
46
22
47
21
48
20
49
19
50
18
51
17
52
16
53
15
54
14
55
13
56
12
57
11
58
10
59
9
60
8
61
7
62
6
63
5
64
4
65
3
66
2
67
1
68
5′
3′
K712
M658
S709
R657
T717
R431
WH
K630
K623
M621
R619
R692
T383
S384
B-loop
AAA+
ISM
K141
R97
F121
S120
W658
S657
Q660
R351
R711
K344
K341
K600
K710
Q423
K624
WH
R692
N691
H156
R693
R641
S332
K401
I416
A419
S181
R179
B-loop
AAA+
ISM
S521
R528
K525
M524
basic patch
Orc1
Orc4
Orc2
Orc5
Orc3
Cdc6
Hydrogen bond (side chain contact)
Hydrogen bond (main chain contact)
van der Waals or electrostatic
buried solvent-accessible surface areas

Supplementary Figure 5. **Summary of DNA contacts by *Dm*ORC and Cdc6.** Residue contact map for the *Dm*ORC·DNA·Cdc6 complex (reconstituted with the 60 bp AT-rich DNA duplex). Amino acid residues within hydrogen bonding distance of DNA (3.6 Å cut-off), and those engaged in van der Waals or electrostatic interactions, are indicated by dashed and solid lines, respectively. Colored bars represent buried solvent-accessible surface areas. Interactions involving ISM and B-loop residues are highlighted by light grey and dark grey boxes.

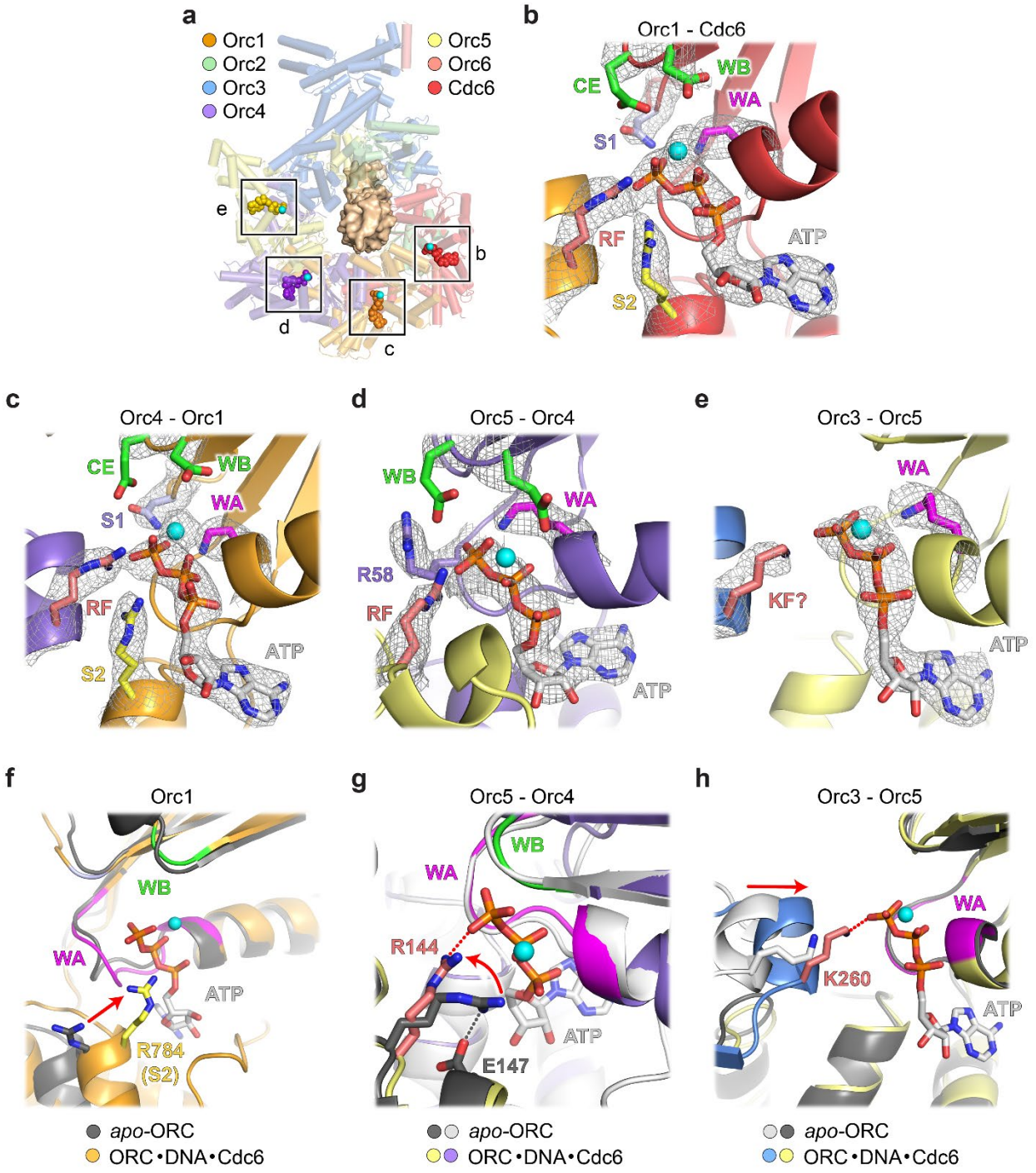


Supplementary Figure 6. **Cryo-EM structures of the binary *Dm*ORC·DNA complex and DNA-free *Dm*ORC. a**) Reconstitution and purification of *Dm*ORC·DNA. Gel filtration chromatogram and SDS-PAGE gel of complex peak fractions are shown. **b**) Representative 2D class averages obtained for the *Dm*ORC·DNA sample. Classification revealed that the sample contained a mixture of DNA-free and DNA-bound *Dm*ORC. **c**) Unsharpened cryo-EM maps for *Dm*ORC·DNA and DNA-free *Dm*ORC (ATP-bound) colored by local resolution. For 3D refinement of the final maps, DNA-bound and DNA-free *Dm*ORC 3D classes from dataset 2 were merged with the corresponding ones from dataset 1 (see **Supplementary Fig. 2**). Both EM volumes show an active ORC state with a large gap in the ORC ring that is otherwise occupied by Cdc6 in *Dm*ORC·DNA·Cdc6. **d**) Gold-standard FSC curves for *Dm*ORC and *Dm*ORC·DNA calculated using EM half-maps. The resolutions at $FSC_{0.143}$ are 3.9 Å and 4 Å (albeit slightly anisotropic due to particle orientation bias). **e** and **f**) Cdc6 binds *Dm*ORC·DNA by a lock-and-key-like mechanism and does not substantially remodel ORC·DNA contacts (except for Orc2 WH). **e**) WH and side views of the *Dm*ORC·DNA cryo-EM map (unsharpened) colored by subunit. No cryo-EM density is observed for the Orc2 WH domain which is detached from the AAA+ layer and flexible in the absence of Cdc6. **f**) Superposition of the *Dm*ORC model in the binary (cartoon, colored by subunit) and ternary (grey cartoon) complex. Cdc6 binding does not extensively remodel

the Orc1-5 ring (apart from the Orc2 WH module). **g** and **h**) DNA binding by ORC stabilizes the WH domain of Orc3 and induces a slight opening of the ORC ring. **g**) WH and side views of the DNA-free *Dm*ORC cryo-EM map (unsharpened) colored by subunit. The Orc3 WH domain is flexible. **h**) Structural comparison of the DNA-free, active *Dm*ORC model (cartoon, colored by subunit) with *Dm*ORC in the ORC·DNA·Cdc6 complex (grey cartoon). DNA binding widens the Orc1-5 ring and the gap between Orc1 and Orc2, and is accompanied by a ~12° rotation of the Orc2/Orc3 module with respect to Orc1/Orc4/Orc5. Source data are provided as a Source Data file.

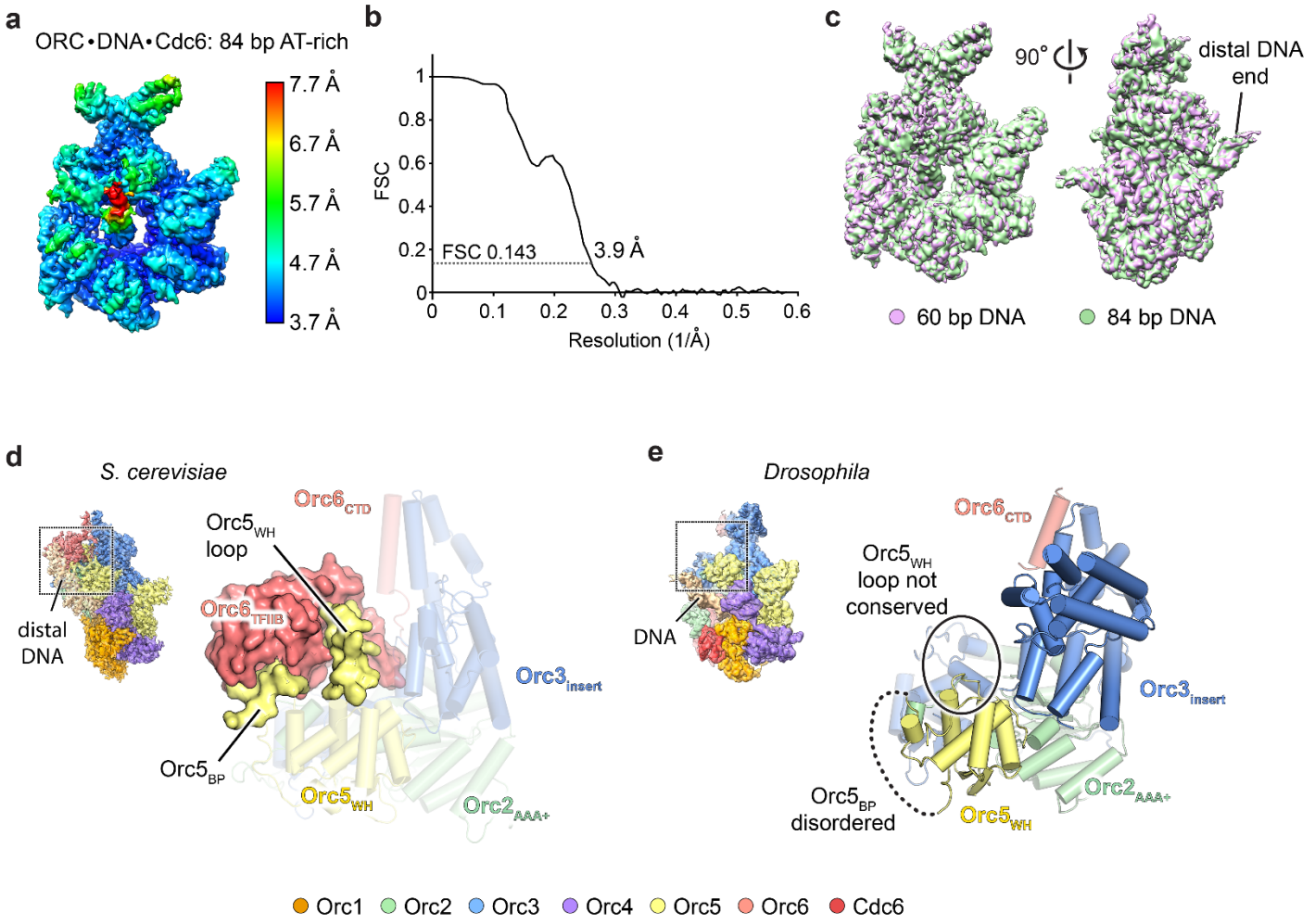


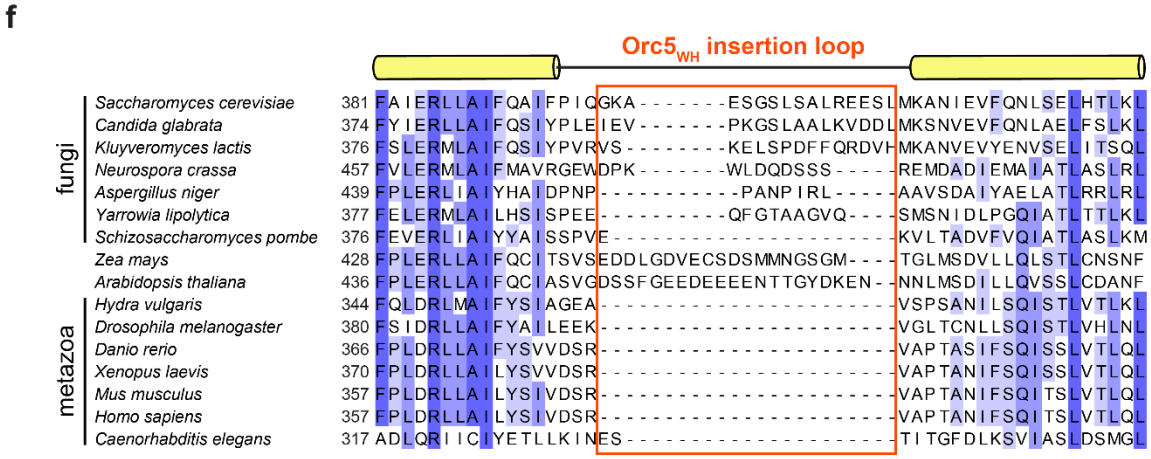
Supplementary Figure 7. **Purification and DNA binding analysis of mutant *Dm*ORC assemblies. a**) SDS-PAGE gel of purified wild type and mutant ORC assemblies used in this study. For Orc6ΔTFIIB, an additional higher percentage gel was run (right lane) to resolve the 8.1 kDa protein. The asterisk marks a degradation product of Orc2 and/or Orc3. **b**) *Dm*Cdc6 does not rescue the DNA binding defect resulting from mutations in the Orc1-B-loop. ATP-dependent DNA binding curves (determined by fluorescent anisotropy, means +/-SD of data points from independent replicates (n) are shown) for *Dm*ORC containing Orc1$^{N691A,\ R692A,\ R693A}$ in the absence and presence of *Dm*Cdc6. Source data are provided as a Source Data file.



Supplementary Figure 8. **Configuration of ATP-binding centers in *Dm*ORC·DNA·Cdc6. a**) ATP (spheres, colored by subunit) and Mg^{2+} (cyan spheres) are bound at four AAA+/AAA+ interfaces: Cdc6/Orc1, Orc1/4, Orc4/5, and Orc5/3. **b** to **e**) Detailed view of each bipartite ATP-binding site with ATP and side chains of conserved active site residues (Walker A, Walker B, sensors 1 and 2, and arginine finger) shown as sticks. Compared to **a**, the ATPase sites are rotated to position the Walker B motif or corresponding regions at the top of each image. The grey mesh outlines the cryo-EM map density (sharpened) for ATP, Mg^{2+}, and displayed side chains. Note that only the Cdc6/Orc1 and Orc1/4 ATPase sites retain all active site residues and possess catalytic activity in yeast and metazoans. Consequently, both ATPase centers adopt a closed configuration

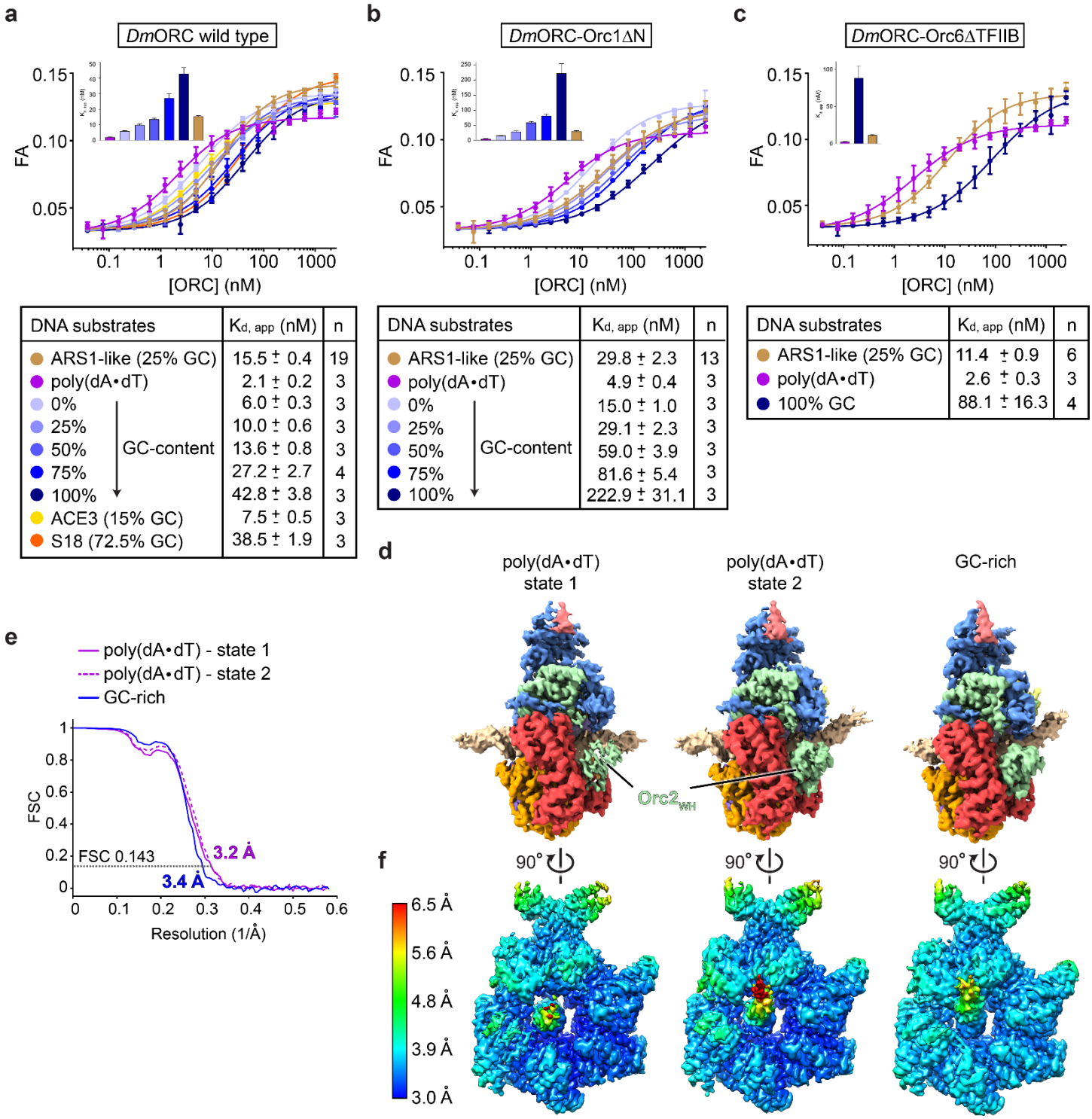
with conserved Walker A, sensor 1, sensor 2, and trans-acting arginine finger residues engaging the nucleotide triphosphate. The Walker B aspartate and catalytic glutamate side chains are not well resolved in the EM map. **f** to **h**) Comparison of ATPase centers in *Dm*ORC·DNA·Cdc6 and *apo-Dm*ORC crystallized previously (PDB 4xcg[1]). Except for the Orc1/4 ATPase center, which is not formed in *apo-Dm*ORC, the active sites are organized similarly in both structures with minor conformational changes. For example, the α-helical lid subdomain of Orc1 rotates towards the AAA+ base to position the conserved sensor 2 near the nucleotide (in **f**), while putative arginine and lysine fingers of Orc5 (in **g**) and Orc3 (in **h**) reorient or move into the ATP binding site to engage the γ-phosphate. These changes likely contribute to the ATP-mediated stabilization of hexameric ORC assemblies and/or the active ORC conformation[2-5]. Movements are highlighted by red arrows.





Supplementary Figure 9. Increasing duplex DNA length does not facilitate additional interactions between *Dm*ORC and DNA. a) Cryo-EM structure of a ternary *Dm*ORC·DNA·Cdc6 complex reconstituted using an 84 bp AT-rich duplex. The cryo-EM map (unsharpened) is colored by local resolution. **b**) Gold-standard FSC curve comparing EM half-maps of the structure in **a**. The overall map resolution is 3.9 Å at $FSC_{0.143}$. **c**) Cryo-EM maps (unsharpened) of *Dm*ORC·DNA·Cdc6 assembled on the 60 bp and 84 bp AT-rich duplexes are superposed. No additional DNA density is resolved distal of the DNA bend with the longer duplex. **d** and **e**) Orc6 contacts DNA and docks against the Orc5 WH domain in *S. cerevisiae* but not *Drosophila* ORC. **d**) In the cryo-EM density of *S. cerevisiae* ORC bound to 72 bp origin DNA[6] (left part of panel), the distal DNA

segment is well resolved and binds the TFIIB domain of Orc6. **e**) Contrariwise, the distal DNA segment is flexible and the Orc6 TFIIB domain is not seen in the EM density of *Drosophila* ORC·Cdc6 bound to an 84 bp duplex (compare boxed regions in **d** and **e**). Close-up views in **d** and **e** unveil that the Orc5-basic patch ($Orc5_{BP}$) and a nearby loop ($Orc5_{WH}$-loop), both of which facilitate docking of the Orc6 TFIIB domain in budding yeast ORC, are disordered ($Orc5_{BP}$) or absent ($Orc5_{WH}$-loop) in *Dm*Orc5. Orc6-TFIIB and Orc5 elements important for Orc6 docking are rendered as molecular surface in the zoomed view in **d**. **f**) Multiple sequence alignment of Orc5 protein sequences (colored by % identity) reveals that the $Orc5_{WH}$-loop is an insertion specific to fungi and is not observed in metazoan species.



DNA substrates	$K_{d,\,app}$ (nM)	n
ARS1-like (25% GC)	15.5 ± 0.4	19
poly(dA•dT)	2.1 ± 0.2	3
0%	6.0 ± 0.3	3
25%	10.0 ± 0.6	3
50%	13.6 ± 0.8	3
75%	27.2 ± 2.7	4
100%	42.8 ± 3.8	3
ACE3 (15% GC)	7.5 ± 0.5	3
S18 (72.5% GC)	38.5 ± 1.9	3

DNA substrates	$K_{d,\,app}$ (nM)	n
ARS1-like (25% GC)	29.8 ± 2.3	13
poly(dA•dT)	4.9 ± 0.4	3
0%	15.0 ± 1.0	3
25%	29.1 ± 2.3	3
50%	59.0 ± 3.9	3
75%	81.6 ± 5.4	3
100%	222.9 ± 31.1	3

DNA substrates	$K_{d,\,app}$ (nM)	n
ARS1-like (25% GC)	11.4 ± 0.9	6
poly(dA•dT)	2.6 ± 0.3	3
100% GC	88.1 ± 16.3	4

Supplementary Figure 10. **DNA binding and remodeling of different DNA substrates by *Dm*ORC. a**) *Dm*ORC has a higher affinity for AT-rich DNA substrates than for GC-rich ones. DNA binding of wild type, full-length *Dm*ORC to fluorescein-labeled dsDNA was measured by fluorescence anisotropy. Mean and standard deviations of independent experiments (n is listed) are shown, and apparent dissociation constants ($K_{d,\,app}$) and standard errors of fits are summarized and also plotted as a bar graph in the inset. **b** and **c**) Removal of the N-terminal 439 amino acid residues in Orc1 (Orc1ΔN, in **b**) or the Orc6 TFIIB domain (Orc6ΔTFIIB, in **c**) does not mitigate *Dm*ORC's ability to associate more strongly with AT-rich duplexes as compared to GC-rich dsDNA. Fluorescence anisotropy experiments were performed as in **a**. **d** to **f**) Cryo-EM reconstructions of

*Dm*ORC·DNA·Cdc6 complexes reconstituted with a 60 bp poly(dA·dT) dsDNA and a 60 bp GC-rich duplex reveal different extents of DNA remodeling by *Dm*ORC. **d**) Side views of cryo-EM maps (unsharpened) of ternary complexes reconstituted with poly(dA·dT) and GC-rich dsDNA. In the case of the poly(dA·dT)-containing ternary complex, 3D classification revealed two distinct states characterized by different DNA bending angles. The DNA in state 1 exhibits a small DNA bend, while the degree of bending in state 2 is intermediate to state 1 and that observed for the GC- and AT-rich duplexes (see also **Figs. 6e-f**). The density for the Orc2 WH domain in state 1 is weak, suggesting that some DNA bending is necessary to allow docking of Orc2 WH onto the ORC·Cdc6 ring. **e**) Gold-standard FSC curves calculated using half-maps of each 3D reconstruction. The resolution at $FSC_{0.143}$ is 3.2 Å for both poly(dA·dT) states, and 3.4 Å for the reconstruction containing the GC-rich duplex. **f**) Unsharpened cryo-EM map for each reconstruction colored by local resolution. Source data are provided as a Source Data file.

Table 1. Summary of cryo-EM data collection, refinement, and validation statistics.

Sample	**ORC·DNA·Cdc6 60 bp AT-rich (31.7% GC)**	**ORC·DNA 60 bp AT-rich (31.7% GC)**	**ORC·DNA·Cdc6 84 bp AT-rich (32.1% GC)**	**ORC·DNA·Cdc6 60 bp GC-rich (70% GC)**		**ORC·DNA·Cdc6 60 bp poly(dA·dT) (0% GC)**	
EM data collection and processing:							
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	
Microscope	Titan Krios (Volta phase plate, EFTEM)	Titan Krios (Volta phase plate, EFTEM)	Titan Krios (Volta phase plate, EFTEM)	Titan Krios (Volta phase plate, EFTEM)	Titan Krios (EFTEM)	Titan Krios	
Camera	K2 Summit	K2 Summit	K2 Summit	K2 Summit	K2 Summit	Falcon 3EC (electron counting)	
Voltage (kV)	300	300	300	300	300	300	
Magnification	x130,000	x130,000	x130,000	x130,000	x130,000	x75,000	
Frames (no.)	40	40	40	40	50	50	
Total electron dose ($e^-/Å^2$)	40	40	40	40	50	50	
Electron dose rate (e^-/pixel/s)	4.9	4.9	4.2	4.2	3.7	0.91	
Calibrated pixel size (Å)	0.86	0.86	0.86	0.86	0.86	0.867	
Defocus range (μm)	-0.4 to -0.6	-0.4 to -0.6	-0.4 to -0.6	-0.4 to -0.6	-1 to -2.2	-0.8 to -1.5	
Micrographs	2,618	1,786	545	1,267	2,533	2,236	
Initial picks (no.)	756,084	596,067	204,063	360,046	843,811	1,030,990	
		114,247					
	ORC·DNA·Cdc6	ORC·DNA	ORC	ORC·DNA·Cdc6	ORC·DNA·Cdc6	ORC·DNA·Cdc6 state 1	ORC·DNA·Cdc6 state 2
Refined particles (no.)	203,070	131,750	126,560	44,520	174,891	82,964	81,248
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1
Global resolution (Å)							
FSC 0.5 (unmasked/masked)	4.2/3.7	7.0/4.3	7.6/4.4	7.6/4.5	4.3/3.9	4.3/3.8	4.3/3.7
FSC 0.143 (unmasked/masked)	3.7/3.4	4.2/3.9	4.4/4.0	4.4/3.9	3.9/3.4	3.7/3.2	3.7/3.2
Local resolution range (Å)	3.2 - 7.5	3.7 - 8.5	3.7 - 7.0	3.7 - 11.3	3.3 - 6.4	3.0 - 9.6	3.0 - 11.1
Map sharpening B factor ($Å^2$)	-95	-138	-97	-94	-100	-82	-73
EMDB accession number	EMD-22361	EMD-22362	EMD-22363	EMD-22329	EMD-22360	EMD-22359	EMD-22330
Model refinement and validation:							
PDB accession number	7JK4	7JK5	7JK6	7JGR	7JK3	7JK2	7JGS
Model composition							
Non-hydrogen atoms	21717	17408	15161	21717	21661	21619	21743
Protein residues	2531	2003	1887	2531	2528	2523	2528
DNA residues	68	64	0	68	66	66	70
Ligands (ATP, Mg)	4, 4	3, 3	3, 3	4, 4	4, 4	4, 4	4, 4
Root mean square deviation							
Bond lengths (Å)	0.006	0.007	0.015	0.006	0.004	0.005	0.005
Bond angles (°)	0.878	1.044	1.207	0.878	0.784	0.846	0.875

	ORC·DNA·Cdc6 AT-rich	ORC·DNA AT-rich	ORC	ORC·DNA·Cdc6 AT-rich (84 bp)	ORC·DNA·Cdc6 GC-rich	ORC·DNA·Cdc6 poly(dA·dT) state 1	ORC·DNA·Cdc6 poly(dA·dT) state 2
B factors (Å^2)							
Protein	113.06	193.96	238.72	167.65	129.38	133.81	122.88
DNA	216.36	323.45	-	295.57	250.15	206.49	219.38
Ligands	92.2	152.81	195.84	140.86	105.63	101.69	96.42
Ramachandran plot							
% favored	95.89	95.1	95.6	95.89	96.85	96.48	97.3
% allowed	4.11	4.9	4.4	4.11	3.15	3.52	2.7
% outliers	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Rotamer outliers (%)	0.32	0.29	0.24	0.36	0.18	0.23	0.18
MolProbity							
Clashscore	4.61	5.55	5.9	5.63	3.67	4.03	4.05
MolProbity score	1.52	1.64	1.63	1.59	1.35	1.42	1.32
Model-map comparison							
EM Ringer score*	3.11	1.52	0.69	1.84	2.28	2.71	2.92
CC_{mask}	0.88	0.84	0.78	0.84	0.87	0.86	0.87
$FSC_{model/map}$ 0.5	3.4	3.9	4.1	3.9	3.4	3.2	3.2

* The EM Ringer score was calculated using the B factor-sharpened cryo-EM maps.

Table 2. Summary of DNA oligonucleotides used in this study

Name	Sequence (5'-3')	used for
40 bp ARS1-like (25% GC)		
Top	/5FluorT/TTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTG	DNA binding
Bottom	CAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAA	
40 bp poly(dA·dT)		
Top	/5FluorT/TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	DNA binding
Bottom	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
40 bp 0% GC		
Top	/5FluorT/TATATTATATATATAAAATAAAATATATTATAAAATTTTT	DNA binding
Bottom	AAAAATTTTATAATATATTTTATTTTATATATATAATATA	
40 bp 25% GC		
Top	/5FluorT/TAGAATTTGAATGTATCCTTAGTCGATTCAAATCTAAATT	DNA binding
Bottom	AATTTAGATTTGAATCGACTAAGGATACATTCAAATTCTA	
40 bp 50% GC		
Top	/5FluorT/TGCCAGCTCTTTCAGTATCATGGAGCCCATGGTTGAGTGA	DNA binding
Bottom	TCACTCAACCATGGGCTCCATGATACTGAAAGAGCTGGCA	
40 bp 75% GC		
Top	/5FluorT/TCCCGGCGCCTGGGTGGTCGTACTGCCAGCTGAGCCCGCG	DNA binding
Bottom	CGCGGGCTCAGCTGGCAGTACGACCACCCAGGCGCCGGGA	
40 bp 100% GC		
Top	/5FluorT/CGCCGGGGCCGGGGGCGCCCCGCCGCGCGGGGCCGGGCCC	DNA binding
Bottom	GGGCCCGGCCCCGCGCGGCGGGGCGCCCCCGGCCCCGGCG	
40 bp ACE3 (15% GC)		
Top	/5FluorT/ GTTTATAATTTTATTGTAATTTTATCTCAATTTTTTTTGC	DNA binding
Bottom	GCAAAAAAAATTGAGATAAAATTACAATAAAATTATAAAC	
40 bp S18 (72.5% GC)		
Top	/5FluorT/ TCCAGACCAGAGGAGCTCCAGCGCTGGGAGCGGCCAGGGC	DNA binding
Bottom	GCCCTGGCCGCTCCCAGCGCTGGAGCTCCTCTGGTCTGGA	
60 bp AT-rich (31.7% GC)		
Top	CCTGCAGGCCTTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTGTAAAATAACA	cryo-EM ATPase pull-down
Bottom	TGTTATTTTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAAGGCCTGCAGG	
60 bp GC-rich (70% GC)		
Top	GGGGCTCGACTCCAAGTGCGGGGCACTCGTCTGTGAGGGGGCTCAGTTCGCACTGCGGGG	cryo-EM
Bottom	CCCCGCAGTGCGAACTGAGCCCCCTCACAGACGAGTGCCCCGCACTTGGAGTCGAGCCCC	
60 bp poly(dA·dT)		
Top	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	cryo-EM
Bottom	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	

84 bp AT-rich (32.1% GC)		
Top	TTTGTGCACTTGCCTGCAGGCCTTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTGT AAAATAACAAGATGTAAAGAT	cryo-EM
Bottom	ATCTTTACATCTTGTTATTTTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAAGGC CTGCAGGCAAGTGCACAAA	
178 bp AT-rich (30.9% GC)		
Top	/5Biosg/GCCGGCATTTAAATCAAAATAGCAAATTTCGTCAAAAATGCTAAGAAATAGGTTATTA CTGAGTAGTATTTATTTAAGTATTGTTTGTGCACTTGCCTGCAGGCCTTTTGAAAAGCAAGCA TAAAAGATCTAAACATAAAATCTGTAAAATAACAAGATGTAAAATTTAAATCGCCGG	Mcm2-7 loading
Bottom	/5Biosg/CCGGCGATTTAAATTTTACATCTTGTTATTTTACAGATTTTATGTTTAGATCTTTTATG CTTGCTTTTCAAAAGGCCTGCAGGCAAGTGCACAAACAATACTTAAATAAATACTACTCAGTA ATAACCTATTTCTTAGCATTTTTGACGAAATTTGCTATTTTGATTTAAATGCCGGC	
178 bp poly(dA·dT)		
Top	/5Biosg/GCCGGCATTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTAAATCGCCGG	Mcm2-7 loading
Bottom	/5Biosg/GCCGGCATTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTAAATCGCCGG	

5FluorT – 5’ Fluorescein dT label; 5Biosg – 5’ Biotin label

Supplementary References

1. Bleichert, F., Botchan, M.R. & Berger, J.M. Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321-6 (2015).

2. Bleichert, F., Leitner, A., Aebersold, R., Botchan, M.R. & Berger, J.M. Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc Natl Acad Sci U S A* **115**, E5906-E5915 (2018).

3. Ranjan, A. & Gossen, M. A structural role for ATP in the formation and stability of the human origin recognition complex. *Proc Natl Acad Sci U S A* **103**, 4864-9 (2006).

4. Tocilj, A. et al. Structure of the active form of human origin recognition complex and its ATPase motor module. *eLife* **6**, e20818 (2017).

5. Siddiqui, K. & Stillman, B. ATP-dependent assembly of the human origin recognition complex. *J Biol Chem* **282**, 32370-83 (2007).

6. Li, N. et al. Structure of the origin recognition complex bound to DNA replication origin. *Nature* **559**, 217-222 (2018).

Goal 2: Structure of a loading competent S. cerevisiae ORC-Cdc6 complex bound to origin DNA.

Jan Marten Schmidt, Jan Seebacher and Franziska Bleichert

Structure of a loading competent *S. cerevisiae* ORC-Cdc6 complex bound to origin DNA

Jan Marten Schmidt[1,2], Jan Seebacher[1] and Franziska Bleichert[3*]

[1] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, Basel, 4058, Switzerland

[2] University of Basel, Faculty of Natural Sciences, Klingelbergstr. 50, Basel, 4056, Switzerland

[3] Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06520, USA

*Corresponding author: Franziska Bleichert
franziska.bleichert@yale.edu

Abstract

Initiation of DNA replication is mediated by the Origin Recognition Complex (ORC) initiator, which recruits and loads replicative helicases at origins of replication in eukaryotes. Deposition of replicative helicases onto DNA by ORC requires that the initiator associates with the co-loader Cdc6 and undergoes an intricate sequence of conformational changes driven by nucleotide binding and hydrolysis. How the structural rearrangements of ORC are regulated and orchestrated with co-loader binding to support helicase recruitment is unknown. Here we report near-atomic resolution cryo-EM structures of the *S. cerevisiae* Origin Recognition Complex (ORC) bound to origin DNA with or without Cdc6. Our structural studies reveal two discrete conformations of the ORC-Cdc6 ring and provide insight into the physical basis for coordinated activity of the ATPase centers in the ternary ORC-DNA-Cdc6 (ODC) complex. Further, we describe a tripartite binding site formed by Orc1, Cdc6 and the N-terminal cyclin-box fold (CBF_N) of Orc6 in the ternary complex and suggest a role for the N-terminus of Orc6 in regulating replicative helicase recruitment in budding yeast.

Introduction

In eukaryotes, the replication of genomic information starts with the binding of the six-subunit initiator origin recognition complex (ORC) to chromosomal sites throughout the genome in late M and the beginning of G1 phase[1,2]. Recruitment of the co-loader Cdc6 to DNA-bound ORC renders the initiator competent to load the hetero-hexametric Mini-chromosome maintenance (Mcm2-7) complex onto DNA, with the help of the co-factor Cdt1, in a process termed origin licensing[3,4]. The Mcm2-7 complex constitutes the core of the replicative helicase, which functions as a motor for unwinding chromosomal DNA during DNA replication. Mcm2-7 helicases are loaded onto DNA as inactive double-hexamers that are oriented in a head-to-head fashion[5–7]. In S phase, a subset of the loaded helicases is activated to assemble into bi-directional replisomes[3,8]. Replicative helicase loading onto chromosomes and helicase activation are strictly separated in time and are integrated with cell-cycle signaling to guarantee that all genomic information is replicated precisely once per cell cycle[9,10].

ORC subunits 1-5 and Cdc6, which is closely related to Orc1, belong to the family of ATPases associated with various cellular activities (AAA+)[11,12]. AAA+ proteins, which consists of a base and lid domain fold, form composite ATPase sites between neighboring subunits and undergo significant conformational changes upon binding and hydrolysis of ATP to modulate bound proteins or nucleic acids[11,12]. The AAA+ domains of Orc1-5 adopt an open pentameric-ring conformation and cooperate with adjacent winged-helix motifs to bind duplex DNA in an ATP-dependent manner in a channel in the center of the complex[13–17]. Recent structural studies of DNA-bound *S. cerevisiae* ORC[15], ORC-Cdc6-Cdt1-Mcm2-7(OCCM)[18,19] and ORC-Mcm2-7 (MO)[20] have expanded our understanding how budding yeast ORC specifically recognizes origin DNA and how the first Mcm2-7 hexamer is tethered to the ORC-DNA-Cdc6 complex during helicase loading. However, it is not well understood how the recruitment of Cdc6 switches the ORC-DNA complex into a replicative helicase loading-competent state. Understanding the physical basis of this highly regulated event requires the knowledge of all loading-intermediate states at high resolution. These provide mechanistic insight into how intra- and inter-molecular

conformational changes are orchestrated through co-factors, post-translational modifications and ATPase hydrolysis.

In contrast to Orc1-5, Orc6 does not belong to the AAA+ protein superfamily and does not align with the Orc1-5 ring. Instead, Orc6 associates with ORC1-5 by interacting with an insertion-domain in Orc3 outside the Orc1-5 ring through its C-terminal domain (CTD)[13,15,21]. N-terminal to the CTD, Orc6 possesses a TFIIB-like domain comprising two cyclin-box folds (CBF)[20,22,23]. The C-terminal cyclin-box fold is tethered to *Sc*Orc1-5 by a yeast-specific loop-insertion in Orc5 and has been shown to contribute to DNA-bending in *S. cerevisiae*[15]. In contrast to the well-characterized C-terminal region, the structure and function of the N-terminus of Orc6 is not well understood. This is of particular relevance as the N-terminus of Orc6 is well-known to be a target of cell-cycle regulated kinases (CDKs) and has been proposed to contribute to replicative helicase loading by interacting with Cdt1 and Mcm2 and Mcm5[20,24–29].

Although Orc1-5 belong to the AAA+ superfamily, only the AAA+ domains of Orc1, Orc4 and Orc5 have preserved functional motifs that bind ATP[12,13,15,18,30]. ATPase hydrolysis is has only been observed for the composite ATPase site formed at the Orc1/4 interface, and is an important regulator of origin licensing *in vivo*[12,31–35]. It is known that DNA binding to ORC's central channel reduces its ATPase velocity[31,32,36], yet, how ORC binding to DNA is coupled mechanistically to ORC's ATP turnover is still not understood. Cdc6 recruitment to DNA-bound ORC results in the formation of a Mcm2-7 loading complex with an additional functional ATPase site at the Orc1/Cdc6 interface[18,37–45]. Cdc6/Orc1 ATPase activity is essential for cell viability while being dispensable for Mcm2-7 loading *in vitro* and *in vivo*[38,46–50]. Instead it has been suggested that ATP hydrolysis by Cdc6/Orc1 may evict erroneous Mcm2-7 loading intermediates to ensure that only correctly assembled Mcm2-7 double hexamers become activated[48–50]. Together, ORC and Cdc6 ATPase activity allows the timely transition between different functional intermediates of the loading complex, whereby ATPase rate may serve as temporal regulator of the orchestrated assembly and disassembly of replicative helicase loading intermediates. The underlying structural basis of how ATP-binding and hydrolysis fuels and coordinates origin licensing has yet to be revealed.

To understand how Cdc6 recruitment to DNA-bound ORC transitions the complex into a loading-competent state and how the functional ORC/Cdc6 ATPase sites are mechanistically coupled to the assembly and disassembly of loading intermediates, we determined the cryo-EM structures of *S. cerevisiae* ORC bound to an 84bp fragment of the ARS1 origin in the absence and presence of Cdc6 at 2.6Å and 2.7Å resolutions, respectively. Cdc6 docking to ORC results in structural alterations in the Orc1 AAA+ domain that are coupled to the formation of the Orc1/Cdc6 ATPase site. Our structural analyses further identifies two distinct conformational states of the ORC-Cdc6 ring in the ternary complex that differ in helical pitch and specific DNA contacts formed by the AAA+ regions of Cdc6. In addition, we find that key catalytic residues in the functional ATPase site of ORC and Cdc6 adopt distinct side-chain conformations that help explain how sequential ATP hydrolysis of the Orc1/Cdc6 and Orc4/Orc1 site may be achieved. Strikingly, the N-terminal cyclin-box fold of Orc6 can be seen to bind to the winged-helix domain (WHD) of Cdc6 and the AAA+-lid of Orc1, a region that overlaps with the binding site of Mcm7 in the OCCM loading intermediate[18,19]. Taken together, these structural findings provide models for how DNA-bound ORC switches into a loading-competent state upon Cdc6 recruitment and how the N-terminus of Orc6 helps regulate replicative helicase loading in budding yeast.

Results

Structural insight in the *S. cerevisiae* ORC complex bound to ARS1-DNA

To understand how the DNA-bound Origin Recognition Complex can switch into a loading competent state upon Cdc6 recruitment, we assembled and purified a ternary, wild-type *S. cerevisae* ORC-Cdc6 complex bound to 84bp dsDNA of the ARS1 locus in presence of ATP and performed structural studies using single-particle cryo-EM (**Supplementary Figure 1**). We were able to determine the structure of a DNA-bound ORC complex (OD) and structures of the ORC-DNA-Cdc6 complex (ODC) at 2.6Å and 2.7Å resolution, respectively (**Supplementary Figure 2-4**). For all structures, we could assign 41bp of the ARS1 DNA to the density in the central channel of the complex (**Supplementary Figure 4B**). As previously reported[15], the bound DNA is strongly bent upwards when exiting the central channel at the WHD tier towards Orc3 and shows a second inflection point at the rear end of the AAA+ domain ring of ORC (**Supplementary Figure 5A**).

In our OD structure, the overall domain architecture of ORC confirmed previous studies of the DNA-bound complex[15] (**Supplementary Figure 5A**). However, we observed only strongly fragmented density for the Orc2-WHD in our density map suggesting high flexibility of this region (**Supplementary Figure 5A and B, Supplementary Figure 2, Supplementary Figure 5A and B**). This is in contrast to the structure of the cross-linked DNA-bound initiator, in which the Orc2-WHD is stably bound between the AAA+ domains of Orc1 and 2[15]. Notably, we found that arginine-254 (R254) of the Orc2 basic-patch (BP) makes base-specific contacts with Guanine-32 (G32) in our structure (**Supplementary Figure 5C**). These contacts seem to be specific to the ARS1-ORC complex as the R254 in the ARS305-ORC complex[15], which has a adenine at position 32, is flipped away from the base revealing a physical basis for differential substrate recognition by *S. cerevisiae* ORC (**Supplementary Figure 5C**).

Cdc6 adopts distinct conformational states in the ORC-DNA-Cdc6 complex

The resolution of our density map of the ternary initiator complex allowed us to build Cdc6 *de novo* and analyze structural changes that are induced upon Cdc6 recruitment to DNA-bound ORC (**Supplementary Figure 4C**). The AAA+ domain of Cdc6 seals the gap

between Orc1 and Orc2, thereby locking the bound duplex in ORC's central channel, while the winged-helix domain of Cdc6 docks on top of the Orc1 AAA+-lid domain complementing the domain-swapped ORC-WH tier (**Figure 1A**). Contrary to the ORC-DNA-Cdc6 sub-complex in the OCCM intermediate[18,19], we did not observe density for the Orc2-WHD suggesting that Mcm2-7 recruitment to ORC-DNA-Cdc6 stabilizes Orc2-WHD binding to Cdc6 (**Figure 1A**). In addition, the Orc1-BP is tightly bound to DNA in our ODC structure, but not resolved in the OCCM[18] and OCCM intermediates[19] suggesting that replicative helicase recruitment to ORC-Cdc6 restructures ORC-DNA contacts. Future high-resolution structural studies will be required to resolve how DNA recognition by ORC is modulated during Mcm2-7 recruitment and loading.

Focused 3D classification on the density assigned to Cdc6 revealed that Cdc6 adopts two distinct states (state 1 and 2) in context of the ternary ODC complex (**Figure 1B**). In state 1, the AAA+-base fold of Cdc6 is positioned out of plane in respect to the AAA+ domains of Orc1-5 resulting in a more spiral arrangement compared to state 2 (**Figure 1B and C**). This spiral arrangement of ODC state 1 contrasts the planar AAA+ domain organization of ORC-Cdc6 in the OCCM intermediate (**Figure 1C**). Further, the relative arrangement of the AAA+-lid and base domains of Cdc6 differs between both states, while the overall fold remains unchanged (**Figure 1D**).

The functional consequences of the different Cdc6 states are unclear, but it seems that the AAA+-base domains of ORC and Cdc6 transition to a planar organization in order to support Mcm2-7 recruitment.

Cdc6 states feature distinct Cdc6-DNA contacts

In our ODC structures, Cdc6 contacts DNA through three motifs: the initiator specific motif (ISM) and B-loop element, which are part of the AAA+-base fold, and a beta-hairpin motif in the WHD. All three regions in Cdc6 have been previously reported to contact DNA in the OCCM intermediate, although some of them through different amino acid side chain residues. In ODC (state 1), Cdc6 interacts with the DNA backbone by Q198 (ISM), N232 and T233 (B-loop) as well as K479, K481 and K483 (WHD) (**Supplementary Figure 6**).

In contrast, Cdc6 (OCCM) has been described to interact with DNA by T196 (ISM), N232 and T233 (B-loop) and WHD residues K477, K479 and K483[18]. These differences are difficult to interpret as the OCCM density map is insufficient to reliably assign side-chains in the ISM, B-loop and WHD of Cdc6. However, it is important to note that the ISM and WHD of Cdc6 in the OCCM may contact DNA differently than in our ODC structures.

Interestingly, transition from ODC state 1 to state 2 is accompanied by changes in how Cdc6 engages DNA. The ISM of Cdc6 shifts away from the duplex breaking the contact between Q198 and the DNA backbone (**Figure 1E**). The B-loop instead moves towards the DNA minor groove closer to the deoxyribose of thymidine-66 (**Figure 1E**). The interactions between the Cdc6 beta-hairpin motif and DNA do not change between ODC states (**Supplementary Figure 6**). In conclusion, transition from Cdc6 state 1 to 2 is accompanied (or may even be triggered) by the release of ISM-DNA contacts and a repositioning of the B-loop towards the DNA minor groove, which may represent a means of DNA shape read-out.

Physical basis for ATPase site regulation in the ORC-DNA-Cdc6 complex

The AAA+ domains of Orc4/Orc1 assemble into a functional ATPase site, which is essential for repeated Mcm2-7 loading *in vivo*[32,34]. Cdc6 recruitment to DNA-bound ORC results in the formation of an additional ATPase site between the AAA+ domains of Orc1 and Cdc6. However, it is unclear how Cdc6 recruitment to ORC-DNA modulates the Orc1 AAA+ domain to form a composite Orc1/Cdc6 ATPase site. Orc1/Cdc6 ATPase activity is essential *in vivo* and has been proposed to ensure productive helicase loading[25,38,46–50]. Nucleotide hydrolysis by Cd6/Orc1 and Orc4/Orc1 occurs in a sequential manner with Cdc6/Orc1 preceding Orc4/Orc1[45]. Despite its regulatory importance, it is unknown how sequential ATP hydrolysis by ORC and Cdc6 is accomplished.

To analyze Cdc6-induced structural alteration in the Orc1 AAA+ domain, we align the ODC structure (state 1) on the AAA+ domain of our OD model. Upon Cdc6 recruitment to DNA-bound ORC, the helix α4 of the Orc1 AAA+ base domain, which harbors the catalytic arginine finger and a tether residue, is pushed back and tilted (**Figure 2A**). This movement re-positions the arginine finger and tether of Orc1 and is therefore likely required for hydrolysis to occur.

When comparing key motifs (Walker A (WA) and B (WB), Arginine Finger (RF), Sensor 1 (S1) and 2 (S2), Tether (TH))[51–53] in the ATPase interface of Orc4/Orc1 and Orc1/Cdc6, we found that the Sensor 1 asparagine located on strand beta-4 occupies different positions in the ATPase sites (**Figure 2B-D**). The Sensor 1 residue plays an important role in modulating ATPase activity by putatively helping to align a water molecule for nucleophilic attack and is essential *in vivo*[47,53–57]. In Orc4/Orc1, the Sensor 1 (Orc1: N600) is flipped away from the client nucleotide, putatively rendering the ATPase site inactive (**Figure 2B**). In contrast, the Sensor 1 asparagine (Cdc6: N262) points towards the γ phosphate of ATP in the Orc1/Cdc6 interface, which likely primes the ATPase site for hydrolysis (**Figure 2C**). The orientation of the Sensor 1 motif seems to be regulated by a hydrogen-bond network (hereafter referred to as S1-TH-D) between Sensor 1, a tether and an invariant aspartate adjacent to the Walker B motif. Notably, mutations in the invariant aspartate of Orc1 (D569) have been shown to reduce ORC's ATPase activity *in vitro* and are lethal when overexpressed *in vivo*[33,44]. Interestingly, the tether is an arginine in Orc4, instead of a lysine as in Orc1, and seems to position the Sensor 1 away from ATP in the Orc4/Orc1 interface (**Figure 2E**). In addition, the S1-TH-D network in the Orc1/Cdc6 site is opened up likely allowing the Sensor 1 asparagine to orient towards the bound nucleotide (**Figure 2F**). Taken together, the orientation of the Sensor 1 residue, and thereby the S1-TH-D network, may represent an important regulator for ORC-Cdc6 ATPase activity and may provide a physical basis for sequential ATP hydrolysis of the ORC-Cdc6 complex. It is worth noting that the invariant aspartate of the S1-TH-D network links the Walker B motif to the B-loop, which contacts DNA and thereby may couple DNA-binding to ATPase hydrolysis in the budding yeast ORC-Cdc6 complex.

We also compared the distances of the Orc4/Orc1 and Orc1/Cdc6 S1-TH-D networks, respectively, in our OD and ODC structures to investigate whether the ATPase sites change with Cdc6 recruitment and complex status. No significant differences in Orc4/Orc1 and Orc1/Cdc6 ATPase site organization could be observed between OD, ODC (state 1 +/- Orc6) and ODC (state 2 +/- Orc6) suggesting that these intermediates do not differ in ATPase activity.

The N-terminal cyclin-box fold of Orc6 forms a tripartite interface with Orc1 and Cdc6

Intriguingly, we observed an additional density located near the Orc1 AAA+-lid and the Cdc6-WH domain in our ODC consensus map, which led us to perform focused 3D classification using a masked around the density (**Supplementary Figure 2**). The 3D classification yielded two complex states (with and without the additional density), which were further classified into ODC state 1 and 2 (**Supplementary Figure 2 and 3**). We were able to assign and *de novo* build the N-terminal cyclin-box fold (CBF_N) of Orc6 into the density in ODC state 1 and 2 (**Figure 3A, Supplementary Figure 3F and 4A**). Alignment of the *de novo* built N-terminal cyclin-box fold (CBF_N) of Orc6 with the known Orc6-CBF_N of the Mcm2-7-ORC (MO) intermediate[20] showed good agreement of both structural motifs further supporting our model (**Figure 3B**). The CBF_N of Orc6 binds an interface formed by the Orc1 AAA+-lid and the Cdc6-WH domain in our structure (**Figure 3A**). Further, we performed cross-linking mass spectrometry on the *in vitro* reconstituted ODC complex to validate the tripartite Orc6/Orc1/Cdc6 interface. In line with our structural studies, we observed specific DiSuccinimidyl SulfOxide (DSSO) cross-links between lysine 92 of the Orc6-CBF_N and lysine 761 of the Orc1 AAA+-lid domain, as well as lysine 50 of the Orc6-CBF_N and lysine 409 of the Cdc6-WHD, confirming the tripartite interface (**Figure 3C-E**).

We performed multiple sequence alignments of Orc1, Cdc6 and Orc6-CBF_N and mapped the conservation scores onto the Orc6/Orc1/Cdc6 interface (**Supplementary Figure 7A**). The interface shows an overall moderate degree of conservation with only a few highly conserved residues, most of which are likely important for the integrity of the domain fold. The interaction of Orc6-CBF_N with Cdc6-WHD seems to be mediated by the first helix-loop-helix motif of the Orc6-CBF_N and the second helix plus a loop succeeding the third helix of the Cdc6-WHD. Orc6-CBF_N binding to Cdc6-WHD relies on polar (Cdc6: N411 and Orc6: E54, R57) and non-polar (Cdc6: I412, L450 and Orc6: L44) residues. The Orc1 AAA+-lid domain mainly contacts the Orc6-CBF_N through the highly conserved second helix, which provides polar residues D687, E690 and R694 to the interface. These

sidechains are in hydrogen-bond distance to residues H93 and E53, respectively, which are located on the second and third helix of Orc6-CBF$_N$.

In summary, the Orc6/Cdc6/Orc1 interface relies on polar and non-polar interactions, which likely only support the dynamic binding of Orc6-CBF$_N$ to the ODC complex. In addition, the Orc6/Cdc6/Orc1 interface is only moderately conserved in eukaryotes suggesting that Orc6-CBF$_N$ binding to Cdc6/Orc1 may be a yeast-specific phenomenon. This is supported by the yeast-specific 157 amino-acid long insertion between the Orc6 cyclin-box folds, which likely is required for the CBF$_N$ of Orc6 to reach the Cdc6/Orc1 interface (**Supplementary Figure 7B**).

The Orc6 N-terminal cyclin-box fold sterically hinders Mcm2-7 recruitment

To learn more about structural alterations in ORC-Cdc6 that may occur in order to recruit Mcm2-7, or after its recruitment, we docked the ORC-DNA-Cdc6 (including the WHD domains of Mcm3 and Mcm7) structure isolated from the OCCM intermediate into our ODC-Orc6-CBF$_N$ density map. We found that the Orc2-WHD, which is stably bound to the AAA+ domain of Cdc6 in the OCCM intermediate, is absent in all our structures (**Supplementary Figure 2 and 8A**). Interestingly, the Orc2-WHD and the AAA+ domain of Cdc6 form a binding site for the Mcm3-WHD in the OCCM complex[18,19]. The WHD of Orc2 makes up approx. 25% of the interface area, which indicates a potential role of Orc2-WHD in Mcm2-7 recruitment by facilitating Mcm3-WHD binding to ORC-Cdc6 (**Supplementary Figure 8B**). Further, we found that the Mcm7-WHD bound to the ORC-DNA-Cdc6 complex in the OCCM sterically clashes with the Orc6-CBF$_N$ in our ODC structures suggesting that the Orc6-CBF$_N$ has to be released from Cdc6/Orc1 in order to recruit Mcm2-7 (**Supplementary Figure 8A**). This is of particular interest as a recent study has captured an ORC-DNA-Cdc6-Mcm2-7 intermediate (termed “semi-attached” OCCM), in which Mcm2-7 is dynamically attached to the ORC-DNA-Cdc6 ring by contacts between Mcm3-WHD/Orc2-WHD/Cdc6-AAA+-lid and Mcm7-WHD/Cdc6-WHD/Orc1-AAA+-lid[19]. Strikingly, the density assigned to the Mcm7-WHD in the “semi-attached” OCCM cannot be fully accounted for by the Mcm7-WH fold, but is remarkable similar and perfectly overlaps with the corresponding density of the Orc6-CBF$_N$ in our ODC electron

map (**Supplementary Figure 8C**). Indeed, docking our ODC-Orc6-CBF$_N$ model in the electron density map of the "semi-attached" OCCM revealed a much better fit between model and density map suggesting that the density has been falsely assigned to Mcm7-WHD and most likely corresponds to Orc6-CBF$_N$ (**Supplementary Figure 8C**). We therefore suggest that the "semi-attached" OCCM represents a stalled loading intermediate, which allows loose Mcm2-7 attachment by means of the Mcm3-WHD/Orc2-WHD/Cdc6-AAA+-lid interaction, but blocks stable Mcm2-7 recruitment and loading. The "semi-attached" OCCM is probably a reasonably long-lived intermediate as it could be captured by cryo-EM. This further supports the idea that it inhibits the recruitment process.

In conclusion, our ODC structures represent distinct origin licensing intermediates that contribute to our understanding of how structural re-arrangements in ORC and Cdc6 regulate replicative helicase loading and reveal a putative role of the N-terminal cyclin-box fold of Orc6 in inhibiting Mcm2-7 recruitment and loading.

Discussion

Here, we present near-atomic resolution cryo-EM structures of the *S. cerevisiae* DNA-bound ORC complex with or without Cdc6. The structures provide insight into the flexibility of the initiator complex flexibility and provide a physical basis for regulation of ORC-Cdc6 ATPase activity. We reveal a conformational state of the ternary ODC complex that may be incompatible with replicative helicase loading.

Our structure of budding yeast ORC bound to ARS1 DNA reveals that the basic patch of Orc2, which reaches into the major groove of the B1 element, is involved in base-specific recognition of origin DNA by contacting the guanine base 32 via R254. Interestingly, this contact is not observed in the previously reported structure of ARS305-DNA bound *Sc*ORC, in which the R254 of Orc2 is flipped away from the DNA base and the Orc2-BP engages DNA solely via backbone interactions[15](**Supplementary Figure 5C**). ARS305 has an adenine instead of guanine base at position 32 suggesting that the Orc2-BP may allow for differential recognition of origin DNA. It is tempting to speculate that this observation may apply to other DNA-contacting regions in *Sc*ORC, for instance the basic patch of Orc5, adding plasticity to *Sc*ORC-DNA recognition. This would help fine-tune origin binding depending on sequence variation in the consensus. Comprehensive structural and functional studies of budding yeast ORC bound to different ARS substrates will be required to confirm this hypothesis.

Recruitment of the co-loader Cdc6 to DNA-bound ORC closes the gap in the AAA+ tier of the initiator ring leading to the formation of a docking platform required for Mcm2-7 binding (**Figure 1A**). Notably, we observed that Cdc6 can acquire two discrete states in context of the ODC complex, which resemble either a spiral (state 1) or a close to planar (state 2) organization of the AAA+-base folds of ORC and Cdc6 (**Figure 1B**). The transition between state 1 and 2 is accompanied by changes in Cdc6-DNA contacts (**Figure 1E and Supplementary Figure 6**). In ODC (state 1), Cdc6 engages the DNA backbone by Q198 (ISM), N232 and T233 (B-loop), and K479, K481 and K483 (WHD). In the second ODC state, the ISM of Cdc6 is no longer interacting with the DNA backbone and the B-loop is re-positioned towards the DNA minor groove in proximity to the

deoxyribose of thymidine-66. However, DNA interactions of the Cdc6 beta-hairpin remain unchanged.

In contrast to ODC (state 1), the AAA+ domain in the semi-attached OCCM adopt a planar arrangement, which is a third variant, different from state 2 of the ODC complex[19] (**Figure 1C**). It is noteworthy that the Cdc6-DNA contacts reported for the OCCM intermediate differ from ODC state 1 and 2[18]. The relevance of the different AAA+ arrangements in the ODC is unclear, but the transition from a more spiral to planar organization of the complex seems to be required for and/or induced by Mcm2-7 recruitment. It is equally unknown whether the observed changes in Cdc6-DNA contacts are causal for the transition to a near-planar organization of the ODC ring, or a consequence thereof. We hypothesize that the ODC complex can dynamically transition between a spiral and more planar AAA+ domain organization. This latter might be stabilized upon Mcm2-7 recruitment. The differences between our isolated ODC structures and the ODC sub-complex of the OCCM intermediates further extent to the Orc2-WHD and Orc1-BP. In contrast to the OCCM intermediates[18,19], the Orc2-WHD remains flexible in our ODC structures, suggesting that binding of the Mcm3-WHD is required for stable association of the Orc2-WHD to the AAA+-lid domain of Cdc6 (**Supplementary Figure 8A and B**). Furthermore, the basic patch of Orc1, which closely follows the ACS minor groove and makes base-specific contacts with the thymine bases 7, 9 and 11[15], is present in our ODC structures but not in the semi-attached OCCM intermediate[19], suggesting that Mcm2-7 recruitment triggers detachment of the Orc1-BP. The release of ORC-DNA contacts may increase complex flexibility, which in turn may promote conformational transitions necessary for loading Mcm2-7.

ATP binding and hydrolysis by ORC and Cdc6 are important regulators for initiator function and ORC ATPase activity is essential for origin function *in vivo*[25,34,38,44–50,56]. However, how the hydrolysis of client nucleotides in the Cdc6/Orc1 and Orc4/Orc1 composite ATPase sites is coordinated in time to fuel conformational changes of loading intermediates is not understood. Comparative analysis of our OD and ODC structures show that the formation of the Cdc6/Orc1 ATPase center induces a repositioning of the Orc1 arginine finger and the adjacent tether residue, which is likely to be required to

promote ATP hydrolysis (**Figure 2A**). ATP hydrolysis by Cdc6/Orc1 has been proposed to be an important check-point for accurate origin licensing[48–50,58]. The role of Orc4/Orc1 ATPase function is controversial, but it may help to regulate the residence time of ORC on DNA. However, Cdc6/Orc1 ATPase hydrolysis must precede Orc4/Orc1 ATPase activity to support helicase loading[45]. The physical basis for the temporal regulation of these sequential events is unclear. Our data shed light on this. Strikingly, we observed different orientations of the Sensor 1 asparagine residue, which is part of a hydrogen-bond network coordinating the water molecule for nucleophilic attack on the γ phosphate[53], in the Cdc6/Orc1 and Orc4/1 ATPase interfaces. Accordingly, the Sensor 1 (Orc1: N600) in the Orc4/Orc1 ATPase site points away from the client nucleotide potentially rendering it inactive (**Figure 2B**). The outward orientation of the Sensor 1 asparagine of Orc1 might be directed by the basic tether residue in Orc4, which is an arginine instead of a lysine as in Orc1 (**Figure 2E**). In contrast, the Sensor 1 asparagine (Cdc6: N262) in the Orc1/Cdc6 ATPase site reaches towards the γ phosphate of ATP, which may prime the ATPase center for hydrolysis (**Figure 2C**). Moreover, orientation of the Cdc6 Sensor 1 residue towards the bound nucleotide may be aided by a widened S-TH-D hydrogen bond network (**Figure 2F**). We propose that the orientation of the Sensor 1 residue in the Cdc6/Orc1 and Orc4/Orc1 ATPase site may regulate the sequential ATP hydrolysis within the ORC-Cdc6 complex.

Unexpectedly, a subset of particles in our high-resolution cryo-EM data set of the ORC-DNA-Cdc6 complex showed additional density that we could assign to the N-terminal cyclin-box fold of Orc6 (aa 1-114) (**Supplementary Figure 2, Figure 3A and B**). This assignment was validated by cross-linking mass spectrometry (**Figure 3C-E**). Contrary to the C-terminal cyclin-box fold of Orc6, the N-terminal CBF has only been poorly structurally characterized and was just recently captured by cryo-EM in the Mcm2-7-ORC (MO) loading intermediate[20]. In our structures, the Orc6-CBF_N is located at the bottom of the ORC-Cdc6 ring, approx. 120Å away from the C-terminal CBF. It forms a tripartite interface with the winged-helix domain of Cdc6 and the AAA+-lid domain of Orc1 (**Figure 3A**). The Orc6/Cdc6/Orc1 interface is formed by polar and hydrophobic contacts, but is overall not well conserved among eukaryotes; it may be specific to budding yeast and closely related species (**Supplementary Figure 7A**). One explanation for this might be

the long linker insertion that separates the two Orc6-CBFs in budding yeast, as it is required for the N-terminal CBF to reach down to Cdc6 and Orc1 (**Supplementary Figure 7B**). Intriguingly, the linker region in budding yeast Orc6 contains multiple CDK consensus sites[26]. Phosphorylation of those sites has been shown to inhibit Orc6 interaction with Cdt1 and to prevent re-replication[26–29]. In addition, Frigola et al. showed that phosphorylated ORC can recruit but not load Mcm2-7 *in vitro*[58]. Nonetheless, the physical basis of this inhibitory mechanism is unknown.

We docked the ODC substructure of OCCM intermediate into the density of the ODC-Orc6-CBF$_N$ and found that the winged-helix domain of Mcm7 (OCCM) occupies the same binding interface as the Orc6-CBF$_N$. This suggests that the N-terminus of Orc6 sterically blocks OCCM formation, potentially creating an auto-inhibited conformation of the ODC complex (**Supplementary Figure 8A**). In addition, the N-terminus of Orc6 has been proposed to interact with Cdt1 and some regions of the Orc6-CBF$_N$ that mediate binding to Orc1 and Cdc6 also contact Mcm2 and Mcm5 in the MO intermediate[20,28]. Thus, formation of the Orc6-CBF$_N$/Orc1/Cdc6 interface would be able to sterically block replicative helicase recruitment by preventing Mcm7-WHD binding and may sequester the Orc6-CBF$_N$ away from other interaction partners such as Cdt1, Mcm2, and Mcm6. This is likely to inhibit Mcm2-7 loading. Interestingly, our purification protocol of *Sc*ORC yielded partially phosphorylated Orc6 protein, which could be removed by lambda phosphatase treatment (**Supplementary Figure 8D**). We suggest that the ODC-Orc6$_{CBFN}$ conformational state identified here constitutes the phosphorylated form of the complex, although the conformational state of the fully dephosphorylated complex remains to be determined.

A recent study by Yuan et al. identified a loading intermediate (semi-attached OCCM), in which only the WHD of Mcm3 and Mcm7 contact ORC-Cdc6, resulting in a loose attachment of the helicase core with the initiator ring. The Mcm3 and Mcm7 WHDs occupy the same interfaces on the ODC complex, as has been reported for the OCCM intermediate[18,19]. However, the density assigned to Mcm7-WHD differs significantly between the semi-attached OCCM and OCCM and is not fully accounted for by the Mcm7-WHD[19] (**Supplementary Figure 8C**). Since the fold of the Mcm7-WHD and Orc6-CBF$_N$

are similar, we speculated that the density assigned to Mcm7-WHD in the semi-attached OCCM corresponds to Orc6-CBF$_N$ instead. Indeed, docking of our ODC-Orc6$_{CBFN}$ model into the semi-attached OCCM density showed that the Orc6-CBF$_N$ fold accounted much better for the density originally assigned to Mcm7-WHD (**Supplementary Figure 8C**). Further, the semi-attached OCCM was captured by cryo-EM, suggesting that it represents a relatively long-lived intermediate that may not be able to proceed with helicase loading. Interestingly, cross-linking mass spectrometry performed on the OCCM structure revealed a cross-link between the N-terminus of Orc6 and the Cdc6-WHD supporting our model.

Taken together, our high resolution structural studies of the budding yeast ORC complex bound to ARS1 DNA in the presence or absence of the co-loader Cdc6 have advanced our understanding of how ORC recognizes different origin DNA through the Orc2-BP element. We provide evidence that Cdc6 can exist in at least two distinct conformations in context of the ORC-DNA complex and speculate that the ORC-Cdc6 ring transitions from a spiral to planar organization in order to support helicase recruitment. In addition, we identify different orientations of the Sensor 1 residue in the Cdc6/Orc1 and Orc4/Orc1 composite ATPase sites, which may provide a physical basis for sequential ATP hydrolysis. Finally, we propose a role for the Orc6-CBF$_N$ in inhibiting Mcm2-7 recruitment and loading, and speculate that this may be dependent on ORC's phosphorylation status.

Methods

Expression and purification of *S. cerevisiae* ORC and Cdc6

Wild-type *Saccharomyces cerevisiae* (*Sc*) Orc1-6 subunits were cloned into pFastBac-derived baculovirus expression vectors (series 4 vectors, QB3 MacroLab, UC Berkeley) via ligation independent cloning (LIC). Of the six ORC subunits, *Sc*Orc1 and *Sc*Orc4 were N-terminally tagged with hexa-histidine (6xHis) and maltose binding protein (MBP), respectively, each followed by a Tobacco Etch Virus (TEV) protease cleavage site, while the remaining subunits were cloned without affinity tags. The coding sequence of wild-type *Sc*Cdc6 was cloned into a series 4 baculovirus expression vector (QB3 MacroLab, UC Berkeley) in frame with an N-terminal 6xHis-MBP affinity tag followed by a TEV protease site.

Full-length *Sc*ORC was expressed in High5 insect cells via the Bac-to-Bac Baculovirus Expression System (Thermo Fisher Scientific). In brief, baculovirus expression vectors of individual *Sc*Orc1-6 subunits were transformed into DH10bac competent cells for bacmid generation, which were subsequently transfected into Sf9 cells to obtain baculoviruses using Cellfectin II (Thermo Fisher Scientific). The viruses of individual *Sc*ORC subunits were amplified in Sf9 cells to produce high-titer viruses, which were freshly used for co-infection of 4L High5 cell cultures for large-scale *Sc*ORC expression. Two days after infection, High5 cells were harvested by centrifugation and the cell pellet was resuspended in 140 mL lysis buffer (50 mM Tris-HCl, 300 mM KCl, 10% glycerol, 50 mM Imidazole, 1 mM ß-ME, 200 µM PMSF, 1 µg/mL leupeptin, final pH 7.8). Cells were disrupted by sonication and lysates clarified by ultracentrifugation at 142,414 x *g* for 45 min in a 45Ti rotor (Beckman Coulter Optima L-80 XP ultracentrifuge). The soluble fraction was subjected to salt-precipitation by the addition of ammonium sulfate to a final concentration of 20% (v/v) and subsequently cleared once more by ultracentrifugation. The supernatant was filtered through a 1.2 µm Acrodisc syringe filter (VWR) and loaded onto a 5 mL HisTrap HP Nickel-affinity chromatography column (GE Healthcare). After loading, the column was washed with 60 mL lysis buffer prior to *Sc*ORC elution with a 50-250 mM imidazole gradient in lysis buffer. Peak fractions were pooled and further purified on a 10 mL amylose column (New England Biolabs) in 50 mM Tris-HCl (final pH 7.8), 300

mM KCl, 10% glycerol, 1 mM ß-ME and eluted with 20 mM maltose. The affinity-tags on *Sc*Orc1 and *Sc*Orc4 were cleaved by overnight digestion with 6xHis-tagged TEV protease and subsequent Nickel-affinity chromatography using a 5 mL HisTrap HP column (GE Healthcare, equilibrated in 50 mM Tris-HCl, 300 mM KCl, 10% glycerol, 50 mM Imidazole, 1 mM ß-ME, final pH 7.8) to remove TEV protease and uncleaved *Sc*ORC. The flow-through was concentrated in a 30K Amicon Ultra-15 concentrator (Millipore) and further purified by gel filtration chromatography on a HiPrep 16/60 Sephacryl S-400 HR column (GE Healthcare) equilibrated in 25 mM HEPES-KOH (final pH 7.6), 300 mM KCl, 10% glycerol, 1 mM DTT. *Sc*ORC peak fractions were pooled and concentrated in a 30K Amicon Ultra-15 concentrator (Millipore). Purified *Sc*ORC protein was aliquoted and flash-frozen in liquid nitrogen for storage at -80°C. All purification steps were done at 4°C. Full-length ScCdc6 was expressed and purified as described for *Sc*ORC except that all buffers were supplemented with 10 mM magnesium acetate and a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) was used for final gel filtration chromatography.

Reconstitution and purification of *Saccharomyces cerevisiae* ORC•DNA•Cdc6

For complex assembly, *Sc*ORC and *Sc*Cdc6 were first dialyzed overnight at 4°C into assembly buffer containing 25 mM HEPES-KOH, 250 mM KOAc, 10 mM MgOAc, 1 mM DTT at a final pH of 7.6. An 84bp DNA duplex harboring the ARS1 sequence (5'-TTTGTGCACTTGCCTGCAGGCCTTTTGAAAAGCAAGCATAAAAGATCTAAACATAAA ATCTGTAAAATAACAAGATGTAAAGAT-3', annealed at 50 µM in 10 mM Tris-HCl pH 8, 5 mM $MgCl_2$) was then mixed with *Sc*ORC in assembly buffer supplemented with 1 mM ATP, and after 5 min incubation at room temperature, *Sc*Cdc6 was added. The final reaction (300 µL total volume) contained 3 µM DNA, 2.5 µM *Sc*ORC and 6.8 µM ScCdc6. After 5 min incubation at ambient temperature, the assembly mix was loaded onto a Superose 6 Increase 10/300GL column (GE Healthcare) equilibrated in 25 mM HEPES-KOH (final pH 7.6), 250 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, 0.2 mM ATP, to isolate a ternary *Sc*ORC•DNA•Cdc6 complex. Peak fraction containing *Sc*ORC, DNA and *Sc*Cdc6 were pooled and concentrated in a 30K Amicon Ultra-4 concentrator (Millipore) to an absorbance (at 280 nm) of 1.8.

Cryo-electron microscopy data collection and image processing

Cryo-EM grids were prepared by applying 3.5 µL of purified and concentrated *Sc*ORC•DNA•Cdc6 complex (in 25 mM HEPES-KOH pH 7.6, 250 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, 0.2 mM ATP, 0.001% TWEEN20) to 300-mesh R1.2/1.3 UltrAuFoil grids (Quantifoil Micro Tools GmbH), which had been freshly plasma-cleaned for 2 min at 5W (18.9% H_2 /81.1% O_2 gas mixture). The sample was allowed to adsorb to the grid for 10 sec, blotted for 2 sec and subsequently vitrified in liquid ethane using a Vitrobot Mark IV plunge freezer (Thermo Fisher Scientific). Cryo-EM data were recorded as dose-fractionated movies in a Titan Krios G2 cryo-electron microscope (Thermo Fisher Scientific) at an acceleration voltage of 300 kV using a post-GIF (Quantum LS imaging filter with a slit width of 20 eV) Gatan K2 summit direct electron detector. The microscope was also equipped with a spherical aberration corrector (CEOS GmbH Heidelberg, Germany). Automated data collection was performed with EPU software (Thermo Fisher Scientific) using fringe-free illumination with a beam-size of 750 nm and 3 exposures per 1.2 µm hole. The targeting defocus was set to -0.8 to -1.6 µm and movies were recorded as 50 frames for a total of 7.5 s with a dose rate of 6.6 $e^-/Å^2$ per second, yielding a total electron dose of 49.3 $e^-/Å^2$.

For image processing, dose-fractionated movies were first motion-corrected with MotionCor2 (Zheng, 2017), and contrast transfer function parameters were determined with GCTF using non-dose-weighted, motion-corrected sums of movie frames (Kai Zhang, 2016). Particles were picked automatically with GAUTOMATCH (Kia Zhang) from the dose-weighted movie-frame sums using low-pass filtered (to 30Å) class averages obtained from a smaller prior test dataset as template. Picked particles were extracted and normalized in RELION 2.0 (Sheres, 2012) with a box size of 300 x 300 pixels at a pixel size of 0.86Å.

Extracted particles were subjected to one round of 3D classification with six classes using a reference map of a previously recorded *Sc*ORC•DNA•ScCdc6 dataset with an initial low-pass filter of 60Å in Relion2.0 to filter out ice and damaged particles. The 3D classification resulted in one meaningful class showing clearly resolved secondary structure, which was refined via 3D auto-refinement in Relion2.0. The refined subset was

then sorted into four classes by masked 3D classification in Relion3.0. The applied mask was generated by aligning the PDB model of *Sc*Cdc6 (OCCM) and the N-terminus of a homology model of *Sc*Orc6 (PHYRE2) to the refined *Sc*ORC•84bp_ARS1•ScCdc6 electron density map in UCSF Chimera. A density map was created from the aligned *Sc*Cdc6 and ScOrc6 models, low-pass filtered to 30Å and resampled to the refined *Sc*ORC•84bp_ARS1•ScCdc6 map. The *Sc*Cdc6-ScOrc6 map was then taken as input to create a soft-edged mask in Relion3.0. Size and position of the resulting mask relative to the refined *Sc*ORC•84bp_ARS1•ScCdc6 density map was confirmed in UCSF Chimera. Individual classes from masked 3D classification were selected and subjected to 3D Refinement, which resulted in density maps representing three different states (ORC-DNA; ORC-DNA-Cdc6 (contracted); ORC-DNA-Cdc6 (relaxed)) of the complex and a junk class. The density maps were sharpened and B-factors automatically estimated via Relion's post-processing function using a global volume mask, low-passed filtered to 15Å. Particles of 3D volumes representing the three different complex states were polished and shiny particles were used for another round of 3D autorefinement. After CTF refinement, the particles were once more polished and refined. The final maps were sharpened and local resolution was determined within Relion3.0 using automatically estimated B-factors.

Model building and refinement

Structural analysis

Multiple sequence alignments were performed via MAFFT version 7[59,60] and visualized in JALVIEW[61] or used to map amino acid conservations on model builds via the ConSurf[62] python plugin in PyMOL(TM) (version 2.3.2 - Incentive Product Copyright (C) Schrodinger, LLC). Cdc6-DNA contacts were determined via the DNAproDB server[63]. Alignments of structural models was performed in PyMOL and RMSD of C-Alpha atoms were projected onto models by the python module ColorByRMSD [https://pymolwiki.org/index.php/ColorByRMSD]. UCSF Chimera and UCSF ChimeraX[64]

were used for docking of structural models into cryo-EM density maps. Topology diagrams of Mcm7-WHD (semi-attached OCCM) and Orc6-CBF$_N$ were generated by PDBsum[65]. All figures were generated using PyMOL, UCSF Chimera and UCSF ChimeraX.

Cross-linking mass spectrometry

*Sc*ORC and *Sc*Cdc6 protein were dialyzed against 25 mM HEPES-KOH pH 7.6, 250 mM KOAc, 10 mM MgOAc, 1 mM DTT at 4°C overnight. After dialysis, proteins were spun for 10 min at 21000 x *g* at 4°C to remove any precipitated protein and protein concentration was measured based on absorption at 280 nm. The ternary ORC-DNA-Cdc6 complex was reconstituted as described for our structural studies, with the exception that the assembled ternary complex was not further purified via gel filtration chromatography. For cross-linking, approx. 35 µg of reconstituted ternary complex was applied to a 30 kDA Amicon spin concentrator, that had been pre-equilibrated with 25 mM HEPES-KOH pH 7.6, 250 mM KOAc, 10 mM MgOAc, 1 mM ATP, 1 mM DTT (reconstitution buffer). The concentrator was topped up with 74.5, 73 and 71 µL reconstitution buffer, respectively, and the sample was carefully mixed by pipetting up and down. In the meantime, CID-cleavable DSSO cross-linking reagent (disuccinimidyl sulfoxide, MW = 388.35, Spacer Arm = 10.3Å, A33545, Thermo Scientific) was allowed to warm up to room temperature and subsequently was dissolved in anhydrous DMSO to a final concentration of 100 mM. The cross-linker was added to the protein solution (in the Amicon concentrator) to a final concentration of 0.5, 2 and 4 mM, respectively. The cross-linking was performed for 1 hour at 10°C while shaking at 400 rpm in the dark. Excess of unreacted cross-linker was quenched by the addition of 50 mM Tris-HCl pH 6.8 (final conc.) and incubation for 1 hour at room temperature. Cross-linking reagent was removed by centrifugation for 5 min at 14k x *g*. The protein was denatured and washed by adding 400 µL of freshly prepared 8 M urea dissolved in 50 mM HEPES-KOH pH 8.5 to the Amicon concentrator, which was then spun for 5 min at 14k x *g*. This step was repeated twice. At the end of each spin, the sample was concentrated to approx. 50 µL total volume. After denaturation and washing, the sample was reduced and alkylated by adding TCEP and 2-chloroacetamide (final

concentration 5 mM and 10 mM, respectively) for 30 min in the dark, while shaking at 400 rpm. The samples were centrifuged for 5 min at 14k x *g* and washed three times with 8M urea in 50 mM HEPES-KOH pH 8.5. The sample volume was reduced to approx. 30 µL by centrifugation and digested with Lys-C enzyme (Wako Chemicals, 1:100 enzyme to protein ratio) for approx. 4 hrs at room temperature, while shaking at 400 rpm. The Lys-C digest was diluted 4-fold by addition of 50 mM HEPES-KOH, pH 8.5 to a final concentration of 2 M urea, followed by addition of trypsin (Sequencing Grade Modified, Promega) 1:100 enzyme to protein ratio) and overnight digest at 37°C, while shaking at 400 rpm. Next, fresh trypsin (1:100 enzyme to protein ratio) and acetonitrile (5% final concentration) were added and the sample was incubated for further 4 hours at 37°C, while shaking at 400 rpm. The Amicon concentrator with trypsin digest was inverted, then spun for 1 min at 1000 x *g* and collected into a fresh Eppendorf tube. The digest was acidified with 50% TFA to give a 1% final concentration, then sonicated for 10 sec in a water bath, and centrifuged for 5 min at 20k x *g*. The supernatant was transferred to a new Eppendorf tube and the sample was stored at -80°C.

LC-MS data acquisition and analysis (this section has been written by Jan Seebacher)

Note: the MS data acquisition was essentially identical, except two different LC column setups were used in 1802 (two column setup: EASY trapping and analytical column, EASY-spray source) and 1925 (Pharmafluidics analyitical column, DPV source, drop-desalting).

We are omitting here the results from the SCX fractions in Exp 1925 as this essentially did not result in any additional relevant results.

20 µl of each sample were transferred to MS autosampler vials for single-shot analyses. The remaining samples were off-line fractionated using SCX on stage-tips.

5 µl of digest was injected, loaded, online-desalted, and then separated on a 50 cm uPAC C18 HPLC column (Pharmafluidics) connected to a modified Digital PicoView nano-source (New Objective), or 15 µl of digest was injected and trapped onto a PepMap 100

C18 2 cm trap using an EASY nLC-1000 system (Thermo Fisher), followed by on-line peptide separation on a 15 cm EASY-Spray C18 column (ES801) connected to an EASY-Spray source (all Thermo Scientific). In either case, a linear gradient of increasing acetonitrile in 0.1% formic acid and water was used for the duration of 240 min, and MS spectra were collected on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) in "MS2_MS3" mode, essentially according to Liu et al 2017[66]. Peptide MS1 precursor ions were measured in the Orbitrap at 120k resolution, with advanced peak determination (APD) feature enabled, and those with assigned charge states between 3-8 were subjected to CID–MS2 fragmentation (25% CID collision energy), and their fragments were detected in the Orbitrap at 30k resolution. Data-dependent HCD-MS3 scans were performed if a unique mass difference (Δm) of 31.9721 Da was found in the CID–MS2 scans with detection in the ion trap (35% HCD collision energy).

MS raw data was analyzed in Proteome Discoverer 2.4 (Thermo Scientific) using Sequest[67] search for linear peptides, including crosslinker-modifications, and XlinkX search to identify crosslinked peptides.

MS2 fragment ion spectra not indicative of the DSSO crosslink delta mass were searched with the Sequest search engine against a custom protein database containing the expected protein components, a database of proteins previously identified in the Sf9 insect cell expression system[68], as well as a database of contaminants taken from MaxQuant, cRAP and those commonly identified in our analyses at the FMI, using the target-decoy strategy[69]. The following variable crosslinker modifications were considered: DSSO Hydrolyzed / +176.014 Da (K); DSSO Tris / +279.078 Da (K), DSSO alkene fragment / +54.011 Da (K); DSSO sulfenic acid fragment / +103.993 Da (K), as well as Oxidation / +15.995 Da (M). Carbamidomethyl / +57.021 Da (C) was set as a static modification. Trypsin was selected as the cleavage reagent, allowing max. 2 missed cleavage sites, peptide lengths between 6-150, 10 ppm precursor mass tolerance and 0.02 Da fragment mass tolerance. PSM validation was performed using the Percolator node in PD 2.4 and a target FDR of 1%.

XlinkX v2.0[66] was used to perform a database search against custom protein database containing the expected complex components to identify crosslinked peptides. Crosslink-

to-spectrum match (CSM) were accepted above an XlinkX score cutoff of 20, requiring a min. of 3 MS3 spectra used for crosslink identification (both peptides must be identified by MS3, at least one peptide as thiol and alkene fragment for DSSO cleavage), MS2-only crosslink identifications were not accepted.

Crosslinks were exported to a csv table in xiNET format. Together with the fasta database of the complex components used in the database search, a network graph was generated in the xiNET and xVis crosslinkviewer[70,71]. Crosslinks were mapped onto the *Sc*ODC structure in PyMOL using the PyXlinkViewer plugin[72].

Data Availability

Acknowledgements

We like to thank Simone Cavadini and Andreas Schenk for their technical support during cryo-EM data collection. We express our gratitude to Sandra Muehlhaeusser for maintaining insect cell culture stocks. We also would like to thank all current and past members of the Bleichert laboratory for their support and helpful discussions.

Author Contributions

Conceptualization, F.B.; methodology, J.M.S. and F.B.; formal analysis, J.M.S., F.B. and J.S.; investigation, J.M.S. and F.B.; writing—original draft, J.M.S and F.B.; visualization, J.M.S. and F.B.; supervision, F.B.; funding acquisition, F.B.

Competing Interests

The authors declare no competing interests.

Figure legends

Figure 1. Architecture and conformational states of the ternary *Sc*ORC-Cdc6-DNA complex. **a**) Overview of the cryo-EM density map of the ternary *Sc*ORC-Cdc6-DNA complex. Cdc6 closes the gap between Orc1 and Orc2/3 by completing the AAA+ and the domain-swapped WHD-tier of Orc1-5. The overall complex architecture remains unchanged. **b**) Conformational states of Cdc6 in context of the ODC complex. Cdc6 adopts two discrete states in the ODC complex, which differ in the relative position of the AAA+-base to the AAA+-lid domain of Cdc6. In state 1 (blue), the AAA+-base of Cdc6 assumes a position lower than the AAA+-lid domain. In contrast, in state 2 the Cdc6 AAA+-base fold moves up towards the plane of the AAA+-lid domain. **c**) Centers of mass of the WHD and AAA+-base domains of Orc1-5 and Cdc6 of ODC state1 (blue) and semi-attached OCCM[19] (light-orange) are shown as spheres. In ODC state 1, the AAA+-base domains of Cdc6 and Orc1-5 adopt a more spiral. The organization of the AAA+-base domains in the OCCM is near planar and contrast the arrangement of ODC state 1. The center of masses of the WHD of Orc1-5 and Cdc6 of ODC state 1 and semi-attached OCCM are similar with the exception that the OCCM Orc1-WHD and Cdc6-WHD are positioned higher. Note, the Orc2-WHD has been omitted from the figure. The spheres are encircled to indicate the corresponding subunit of the ODC complex. Color code as in **a**. **d**) Alignment of the Cdc6-fold in the semi-attached OCCM with ODC state 1 (left) and ODC state 2 with ODC state 1 (right). Structural models are colored by the Root-Mean-Square-deviation of aligned C-alpha atoms. **e**) Cdc6-DNA contacts in ODC state 1 (blue) and 2 (yellow). Glutamine 198 in the ISM of Cdc6 interacts with the DNA backbone in ODC state 1 but not in ODC state 2. Asparagine 232 in the Cdc6 B-loop move towards the DNA minor groove from state 1 to state 2.

Figure 2. Organization of composite ATPase sites in the OD and ODC complex. **a**) Binding of Cdc6 to the OD complex results in the formation of a composite ATPase site between Orc1 and Cdc6. Zoom into the Orc1/Cdc6 site shows that Cdc6 binding tilts the helix α4 of Orc1, which contains the catalytic arginine finger, backwards. **b**) and **c**) Close up view on the composite ATPase sites of Orc4/Orc1 and Orc1/Cd6, respectively, of the ODC complex. All side chain orientations and bound nucleotide are supported by well-

defined density. **d**) Overlay of the Orc4/Orc1 and Orc1/Cd6 ATPase site reveals that the Sensor 1 asparagine points towards the client nucleotide in the Orc1/Cd6 interface, but is flipped outwards in the Orc4/Orc1 ATPase site. **e**) and **f**) Hydrogen-bond network surrounding the Sensor 1 residue in the Orc4/Orc1 and Orc1/Cd6 ATPase interface. The Sensor 1 residue of Orc1 forms hydrogen bonds with an invariant aspartate in Orc1 and the arginine tether of Orc4, which stabilizes the Sensor 1 an outward facing orientation. In the Orc1/Cdc6 ATPase site, the hydrogen bond network between Sensor 1, Tether and invariant aspartate is released and the Sensor 1 orients to the bound ATP. S1 – Sensor 1; WB – Walker B; WA – Walker A; ATP – Adenosine Triphosphate; S2 – Sensor 2; RF – Arginine Finger; TH – Tether; CE – Catalytic Glutamate; D – Invariant Aspartate.

Figure 3. The N-terminal cyclin-box fold of Orc6 forms a tripartite interface with Orc1 and Cdc6. **a**) Overview of the ODC-Orc6-CBFN complex. The cryo-EM density corresponding to the Cdc6-WHD, Orc1-AAA+-lid domain and Orc6-CBFN are indicated. **b**) Alignment of the N-terminal cyclin-box fold (CBF_N) of Orc6 our ODC-CBF_N structure (salmon) and the MO intermediate[20] (light cyan) shows high similarity. RMSD of C-alpha atoms of the alignment is given in Angstrom (Å). Grey color indicates unaligned regions. **c**) Cross-linking mass spectrometry performed on the ODC complex validates that the CBF_N of Orc6 interacts with the WHD of Cdc6 (red dashed line) and the AAA+-lid domain of Orc1 (orange dashed line). For simplicity, only inter-protein cross-links between Orc6/Cdc6 and Orc6/Orc1 are shown. **d**) Inter-protein crosslinks of the ODC-CBF_N assembly. Cross-links are mapped onto the structural model of the ODC-CBF_N complex. Only cross-links between atoms that are part of the structural models are depicted. In black are crosslinks that are within the distance of the spacer (10.3Å) of the used crosslinking agent. In magenta are crosslinks that do not satisfy the threshold of 10.3Å. The Orc6-50aa/Cdc6-409aa crosslink is highlighted by a white circle. **e**) Overview of all observed inter- (blue) and intra-(red) crosslinks of the ODC-CBF_N complex**.**

Supplementary Figure 1. Gel filtration chromatography of *Sc*ORC, Cdc6 and the reconstituted ODC complex. **a**) Final gel filtration run of isolated *Sc*Cdc6 and *Sc*ORC protein. **b**) Sizing trace of the reconstituted ternary ORC-DNA-Cdc6 complex (solid line).

Inlet shows SDS-PAGE gel of peak fractions confirming the presence of ORC and Cdc6. As a control, the sizing trace of isolated DNA is included (dashed lines).

Supplementary Figure 2. Processing scheme for ODC cryo-EM data set. **a**) Representative cryo-electron micrograph showing monodispersed particles. **b**) 2D class averages representing the OC and ODC complex. **c**) Processing scheme for the ODC cryo-EM data set leading to the determination of ORC-DNA, ORC-DNA-Cdc6 (state 1 and 2) and ORC-DNA-Cdc6-Orc6-CBF_N (state 1 and 2) electron density maps. Areas of the complex that were used as masked for focused 3D classification are demarcated with dotted lines.

Supplementary Figure 3. **a**) Fourier shell correlations (FSC) between half-maps of the OD and ODC complexes. Global resolution limits for the corresponding maps are given as FSC 0.143. **b**) Angular distribution of particles of the ORC-DNA and ORC-DNA Cdc6 3D reconstructions. **d**), **e**) and **f**) Cryo-EM density map of ORC-DNA, ORC-DNA-Cdc6 and Orc6-CBF_N colored by local resolution.

Supplementary Figure 4. Model building of Orc6-CBFN, DNA and Cdc6. Model build and sharpened and local-resolution filtered electron density is shown for **a**) the N-terminal cyclin-box fold of Orc6, **b**) DNA and **c**) Cdc6.

Supplementary Figure 5. *Sc*ORC bound to ARS1 origin DNA. **a**) Overview of the cryo-EM density map of *Sc*ORC-DNA. No defined density for the Orc2-WHD was observed. Instead, a gap between Orc1 and Orc2/3 is visible suggesting that the Orc2-WHD is flexible in our structure. **b**) Side-view of the *Sc*ORC-DNA complex. The model of cross-linked *Sc*ORC-DNA (PDB: 5ZR1[15]) was docked into our density map of the uncross-linked ORC-DNA assembly to illustrate that the Orc2-WHD occupies the gap between Orc1 and Orc2/3, when the complex is cross-linked. The position of the Orc2 basic patch is indicated. **c**) Zoom into the Orc2-BP region that is contacting DNA in the ScORC-ARS1 complex. Arginine 254 within the Orc2-BP makes base-specific contact with G32 of the ARS1 DNA (green density map and cartoon). In *Sc*ORC bound to ARS305 DNA (PDB: 5ZR1[15]) the Orc2-Arg254 is flipped away from the adenine base at position 32 (grey density map and cartoon).

Supplementary Figure 6. Comparison of Cdc6-DNA contacts in the ORC-DNA-Cdc6 complex state 1 and state 2. Transition from state 1 to state 2 is accompanied by changes in Cdc6-ISM and B-loop DNA interactions.

Supplementary Figure 7. The budding yeast linker insertion in Orc6, which likely enables the Orc6/Orc1/Cdc6 interface, is not conserved in eukaryotes. **a**) Degree of conservation of residues forming the tripartite Orc6-CBFN/Cdc6-WHD/Orc1-AAA+-lid interface. Residues that are likely important for Orc6-CBFN binding to Cdc6-WHD and Orc1-AAA+-lid domain are shown as sticks. **b**) Multiple sequence alignment of eukaryotic Orc6 protein sequences unveils that the linker region separating the N- and C-terminal cyclin-box fold of Orc6 in budding yeast is specific to some fungi and not present in metazoan. Amino acid conservation is displayed as % identity.

Supplementary Figure 8. The ODC-Orc6-CBF_N complex may be an inhibitory loading intermediate that is incompatible with OCCM formation. **a**) Docking of the ODC sub-complex of the OCCM intermediate (PDB: 5V8F[18]) into our ODC-Orc6-CBF_N electron density map. **i.** Zoom on the binding site of the Orc2 and Mcm3 WH domains. **ii.** Zoom on the electron density corresponding to the CBF_N of Orc6 in the ODC-Orc6-CBF_N complex. The WHD of Mcm7 in the OCCM (grey cartoon) occupies the same binding site at the ORC-Cdc6 ring as the N-terminal CBF of Orc6 (shown as density) in the ODC-Orc6-CBFN complex suggesting that simultaneously binding cannot occur. **b**) Tripartite interface of Mcm3-WHD/Orc2-WHD/Cdc6-AAA+ domain is illustrated as cartoon. The size of the interface between Mcm3-WHD/Orc2-WHD and Mcm3-WHD/Cdc6-AAA+-lid are indicated. **c**) Fit of the Orc6-CBF_N model of the ODC-Orc6-CBF_N complex into the electron density map of the semi-attached OCCM (PDB: 6WGC[19]) suggest that the density assigned to the Mcm7-WHD likely corresponds to the Orc6-CBF_N instead. Goodness of model to map fit is given as Correlation Above Mean (CAM) and Average Map Value (AMV) generated by Fit in map tool of ChimeraX[64]. Topology diagrams of Mcm7-WHD (semi-attached OCCM) and Orc6-CBF_N are displayed. **d**) SDS-PAGE gel of untreated or with Lambda Phosphatase treated *Sc*ORC protein shows that the isolated *Sc*ORC protein contains phosphorylated Orc6. Control refers to protein that was incubated at 30°C in the absence of phosphatase for the indicated time.

References

1. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).

2. Ekundayo, B. & Bleichert, F. Origins of DNA replication. *PLoS Genet.* **15**, e1008320 (2019).

3. Lewis, J. S. & Costa, A. Caught in the act: structural dynamics of replication origin activation and fork progression. *Biochem. Soc. Trans.* **48**, 1057–1066 (2020).

4. Bleichert, F. Mechanisms of replication origin licensing: a structural perspective. *Curr. Opin. Struct. Biol.* **59**, 195–204 (2019).

5. Remus, D. *et al.* Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).

6. Evrin, C. *et al.* A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20240–20245 (2009).

7. Gambus, A., Khoudoli, G. A., Jones, R. C. & Blow, J. J. MCM2-7 form double hexamers at licensed origins in Xenopus egg extract. *J. Biol. Chem.* **286**, 11855–11864 (2011).

8. Riera, A. *et al.* From structure to mechanism— understanding initiation of DNA replication. *Genes and Development* **31**, 1073–1088 (2017).

9. Fragkos, M., Ganier, O., Coulombe, P. & Méchali, M. DNA replication origin activation in space and time. *Nat. Rev. Mol. Cell Biol.* **16**, 360–374 (2015).

10. Reusswig, K.-U. & Pfander, B. Control of Eukaryotic DNA Replication Initiation-Mechanisms to Ensure Smooth Transitions. *Genes (Basel).* **10**, (2019).

11. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**, 27–43 (1999).

12. Iyer, L. M., Leipe, D. D., Koonin, E. V & Aravind, L. Evolutionary history and

higher order classification of AAA+ ATPases. in *Journal of Structural Biology* **146**, 11–31 (2004).

13. Bleichert, F., Botchan, M. R. & Berger, J. M. Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321–326 (2015).

14. Bleichert, F., Leitner, A., Aebersold, R., Botchan, M. R. & Berger, J. M. Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5906–E5915 (2018).

15. Li, N. *et al.* Structure of the origin recognition complex bound to DNA replication origin. *Nature* **559**, 217–222 (2018).

16. Schmidt, J. M. & Bleichert, F. Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nat. Commun.* **11**, 4263 (2020).

17. Jaremko, M. J., On, K. F., Thomas, D. R., Stillman, B. & Joshua-Tor, L. The dynamic nature of the human origin recognition complex revealed through five cryoEM structures. *Elife* **9**, e58622 (2020).

18. Yuan, Z. *et al.* Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nat. Struct. Mol. Biol.* **24**, 316–324 (2017).

19. Yuan, Z. *et al.* Structural mechanism of helicase loading onto replication origin DNA by ORC-Cdc6. *Proc. Natl. Acad. Sci.* **117**, 17747 LP – 17756 (2020).

20. Miller, T. C. R., Locke, J., Greiwe, J. F., Diffley, J. F. X. & Costa, A. Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature* **575**, 704–710 (2019).

21. Bleichert, F. *et al.* A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *Elife* **2**, e00882 (2013).

22. Balasov, M., Huijbregts, R. P. H. & Chesnokov, I. Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in Drosophila

melanogaster. *Mol. Cell. Biol.* **27**, 3143–3153 (2007).

23. Liu, S. *et al.* Structural analysis of human Orc6 protein reveals a homology with transcription factor TFIIB. *Proc. Natl. Acad. Sci.* **108**, 7373 LP – 7378 (2011).

24. Liang, C. & Stillman, B. Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in cdc6 mutants. *Genes Dev.* **11**, 3375–3386 (1997).

25. Weinreich, M., Liang, C., Chen, H.-H. & Stillman, B. Binding of cyclin-dependent kinases to ORC and Cdc6p regulates the chromosome replication cycle. *Proc. Natl. Acad. Sci.* **98**, 11211 LP – 11217 (2001).

26. Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411**, 1068–1073 (2001).

27. Wilmes, G. M. *et al.* Interaction of the S-phase cyclin Clb5 with an 'RXL' docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev.* **18**, 981–991 (2004).

28. Chen, S., de Vries, M. A. & Bell, S. P. Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. *Genes Dev.* **21**, 2897–2907 (2007).

29. Chen, S. & Bell, S. P. CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev.* **25**, 363–372 (2011).

30. Tocilj, A. *et al.* Structure of the active form of human origin recognition complex and its ATPase motor module. *Elife* **6**, (2017).

31. Klemm, R. D., Austin, R. J. & Bell, S. P. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell* **88**, 493–502 (1997).

32. Chesnokov, I., Remus, D. & Botchan, M. Functional analysis of mutant and wild-type Drosophila origin recognition complex. *Proc. Natl. Acad. Sci.* **98**, 11997 LP – 12002 (2001).

33. Klemm, R. D. & Bell, S. P. ATP bound to the origin recognition complex is

important for preRC formation. *Proc. Natl. Acad. Sci.* **98**, 8361 LP – 8367 (2001).

34. Bowers, J. L., Randell, J. C. W., Chen, S. & Bell, S. P. ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol. Cell* **16**, 967–978 (2004).

35. Giordano-Coltart, J., Ying, C. Y., Gautier, J. & Hurwitz, J. Studies of the properties of human origin recognition complex and its Walker A motif mutants. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 69 LP – 74 (2005).

36. Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *EMBO J.* **23**, 897–907 (2004).

37. Donovan, S., Harwood, J., Drury, L. S. & Diffley, J. F. Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5611–5616 (1997).

38. Perkins, G. & Diffley, J. F. X. Nucleotide-Dependent Prereplicative Complex Assembly by Cdc6p, a Homolog of Eukaryotic and Prokaryotic Clamp-Loaders. *Mol. Cell* **2**, 23–32 (1998).

39. Wang, B. *et al.* The essential role of Saccharomyces cerevisiae CDC6 nucleotide-binding site in cell growth, DNA synthesis, and Orc1 association. *J. Biol. Chem.* **274**, 8291–8298 (1999).

40. Weinreich, M., Liang, C. & Stillman, B. The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 441–446 (1999).

41. Seki, T. & Diffley, J. F. X. Stepwise assembly of initiation proteins at budding yeast replication origins in vitro *Proc. Natl. Acad. Sci.* **97**, 14115 LP – 14120 (2000).

42. Mizushima, T., Takahashi, N. & Stillman, B. Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes Dev.* **14**, 1631–1641 (2000).

43. Oehlmann, M., Score, A. J. & Blow, J. J. The role of Cdc6 in ensuring complete genome licensing and S phase checkpoint activation. *J. Cell Biol.* **165**, 181–190 (2004).

44. Speck, C., Chen, Z., Li, H. & Stillman, B. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat. Struct. Mol. Biol.* **12**, 965–971 (2005).

45. Randell, J. C. W., Bowers, J. L., Rodríguez, H. K. & Bell, S. P. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol. Cell* **21**, 29–39 (2006).

46. Schepers, A. & Diffley, J. F. X. Mutational analysis of conserved sequence motifs in the budding yeast cdc6 protein11Edited by M. Yaniv. *J. Mol. Biol.* **308**, 597–608 (2001).

47. Takahashi, N., Tsutsumi, S., Tsuchiya, T., Stillman, B. & Mizushima, T. Functions of sensor 1 and sensor 2 regions of Saccharomyces cerevisiae Cdc6p in vivo and in vitro. *J. Biol. Chem.* **277**, 16033–16040 (2002).

48. Coster, G., Frigola, J., Beuron, F., Morris, E. P. & Diffley, J. F. X. Origin Licensing Requires ATP Binding and Hydrolysis by the MCM Replicative Helicase. *Mol. Cell* **55**, 666–677 (2014).

49. Kang, S., Warner, M. D. & Bell, S. P. Multiple Functions for Mcm2–7 ATPase Motifs during Replication Initiation. *Mol. Cell* **55**, 655–665 (2014).

50. Chang, F. *et al.* Cdc6 ATPase activity disengages Cdc6 from the pre-replicative complex to promote DNA replication. *Elife* **4**, (2015).

51. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951 (1982).

52. Guenther, B., Onrust, R., Sali, A., O'Donnell, M. & Kuriyan, J. Crystal Structure of the δ′ Subunit of the Clamp-Loader Complex of E. coli DNA Polymerase III. *Cell* **91**, 335–345 (1997).

53. Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. Structure and function of the AAA+ nucleotide binding pocket. *Biochim. Biophys. Acta* **1823**, 2–14 (2012).

54. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. Crystal Structure of the Hexamerization Domain of N-ethylmaleimide–Sensitive Fusion Protein. *Cell* **94**, 525–536 (1998).

55. Gai, D., Zhao, R., Li, D., Finkielstein, C. V & Chen, X. S. Mechanisms of Conformational Change for a Replicative Hexameric Helicase of SV40 Large Tumor Antigen. *Cell* **119**, 47–60 (2004).

56. Takehara, M., Makise, M., Takenaka, H., Asano, T. & Mizushima, T. Analysis of mutant origin recognition complex with reduced ATPase activity in vivo and in vitro. *Biochem. J.* **413**, 535–543 (2008).

57. Chen, B. *et al.* Engagement of Arginine Finger to ATP Triggers Large Conformational Changes in NtrC1 AAA+ ATPase for Remodeling Bacterial RNA Polymerase. *Structure* **18**, 1420–1430 (2010).

58. Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. X. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339–343 (2013).

59. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).

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

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

62. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).

63. Sagendorf, J. M., Berman, H. M. & Rohs, R. DNAproDB: an interactive tool for

structural analysis of DNA-protein complexes. *Nucleic Acids Res.* **45**, W89–W97 (2017).

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

65. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **27**, 129–134 (2018).

66. Liu, F., Lössl, P., Scheltema, R., Viner, R. & Heck, A. J. R. Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nat. Commun.* **8**, 15473 (2017).

67. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).

68. Carinhas, N. *et al.* Quantitative proteomics of Spodoptera frugiperda cells during growth and baculovirus infection. *PLoS One* **6**, e26444–e26444 (2011).

69. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol. Biol.* **604**, 55–71 (2010).

70. Combe, C. W., Fischer, L. & Rappsilber, J. xiNET: cross-link network maps with residue resolution. *Mol. Cell. Proteomics* **14**, 1137–1147 (2015).

71. Grimm, M., Zimniak, T., Kahraman, A. & Herzog, F. xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. *Nucleic Acids Res.* **43**, W362–W369 (2015).

72. Schiffrin, B., Radford, S. E., Brockwell, D. J. & Calabrese, A. N. PyXlinkViewer: A flexible tool for visualization of protein chemical crosslinking data within the PyMOL molecular graphics system. *Protein Sci.* **29**, 1851–1857 (2020).

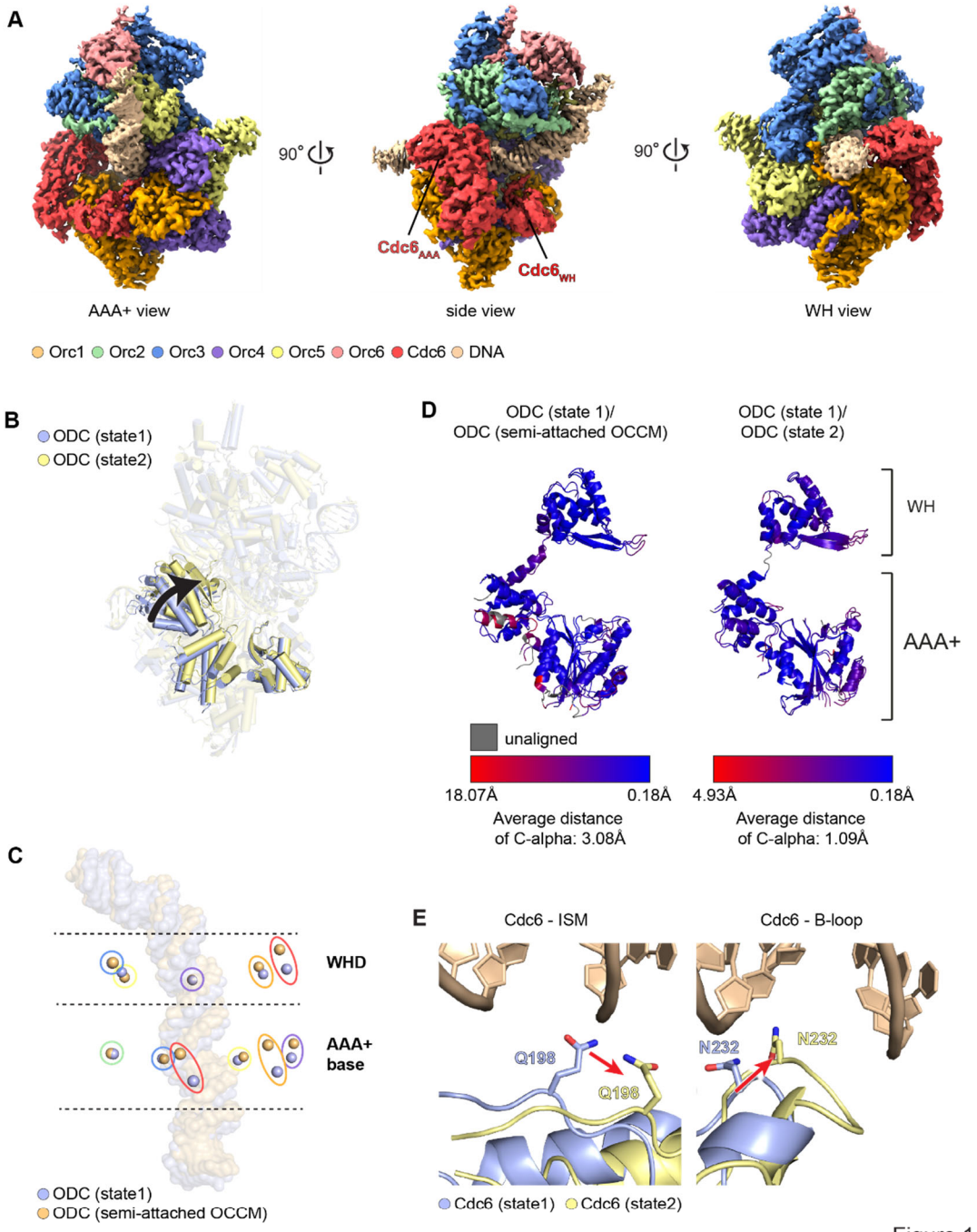

A
90°
90°
Cdc6AAA
Cdc6WH
AAA+ view
side view
WH view
Orc1 Orc2 Orc3 Orc4 Orc5 Orc6 Cdc6 DNA
B
ODC (state1)
ODC (state2)
C
WHD
AAA+ base
ODC (state1)
ODC (semi-attached OCCM)
D
ODC (state 1)/
ODC (semi-attached OCCM)
ODC (state 1)/
ODC (state 2)
WH
AAA+
unaligned
18.07Å
0.18Å
4.93Å
0.18Å
Average distance of C-alpha: 3.08Å
Average distance of C-alpha: 1.09Å
E
Cdc6 - ISM
Cdc6 - B-loop
Q198
Q198
N232
N232
Cdc6 (state1)
Cdc6 (state2)

Figure 1

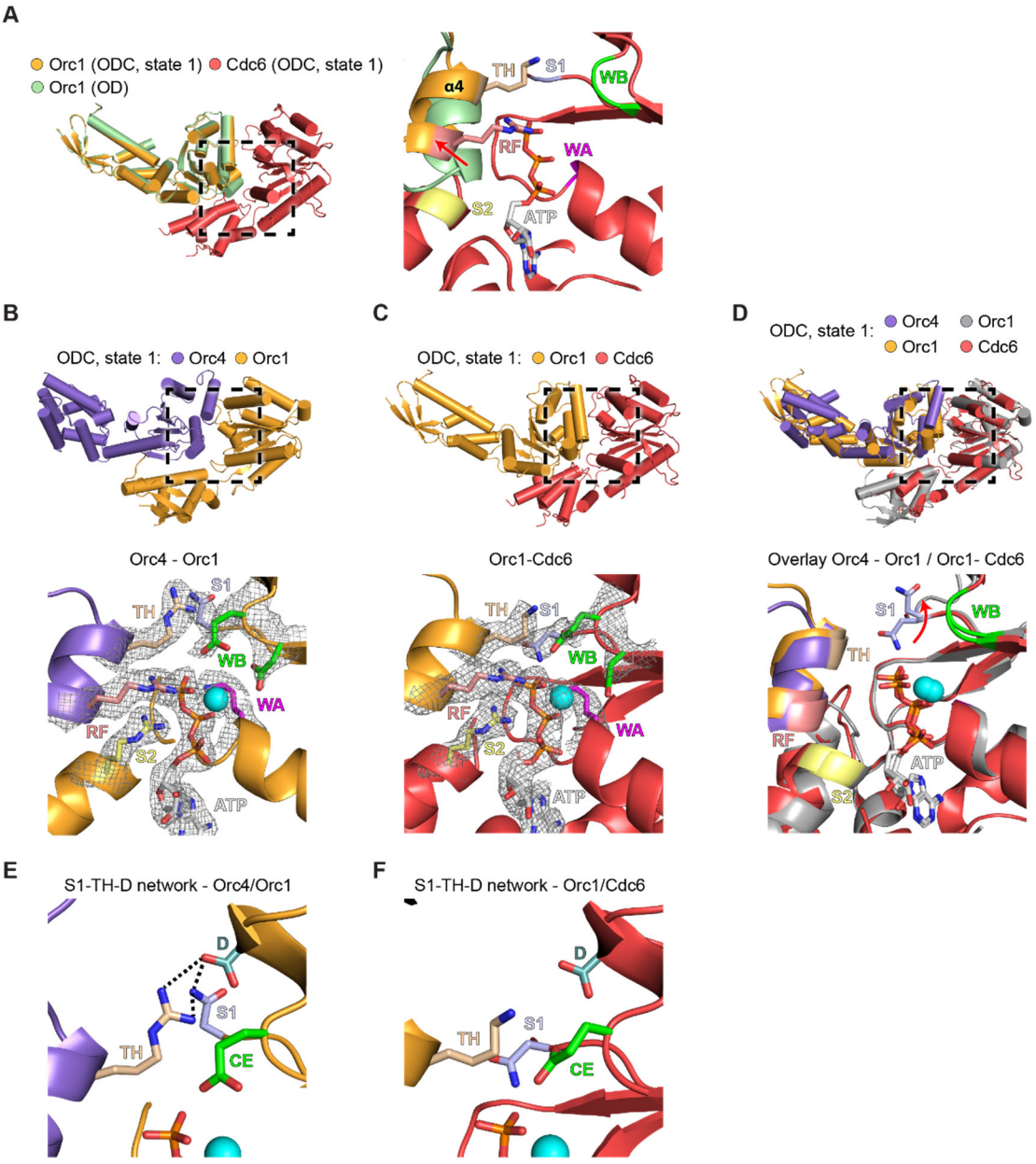

A
Orc1 (ODC, state 1)
Cdc6 (ODC, state 1)
Orc1 (OD)
α4
TH
S1
WB
RF
WA
S2
ATP
B
ODC, state 1:
Orc4
Orc1
C
ODC, state 1:
Orc1
Cdc6
D
ODC, state 1:
Orc4
Orc1
Orc1
Cdc6
Orc4 - Orc1
Orc1-Cdc6
Overlay Orc4 - Orc1 / Orc1- Cdc6
E
S1-TH-D network - Orc4/Orc1
D
S1
TH
CE
F
S1-TH-D network - Orc1/Cdc6
D
S1
TH
CE

Figure 2

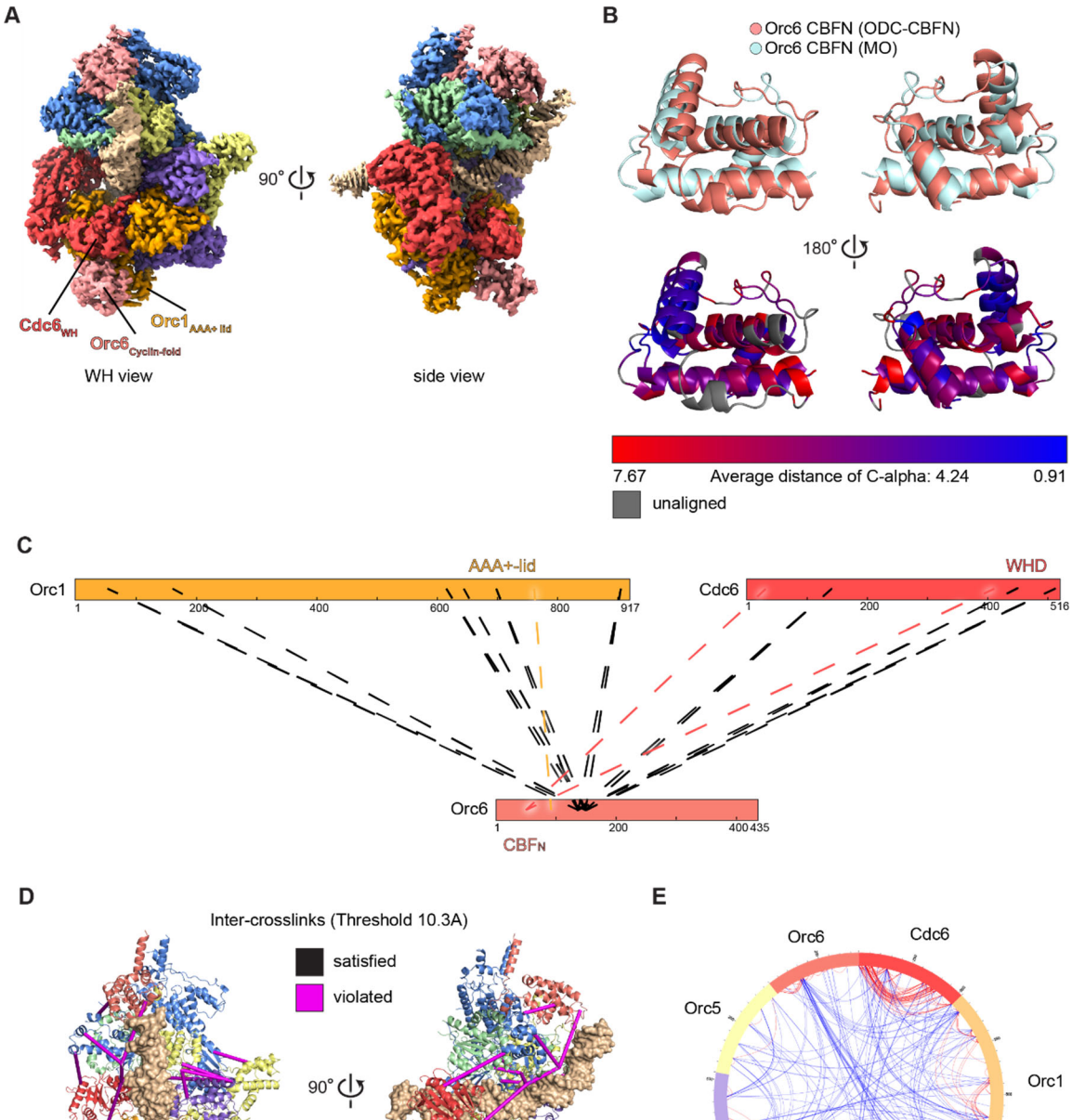
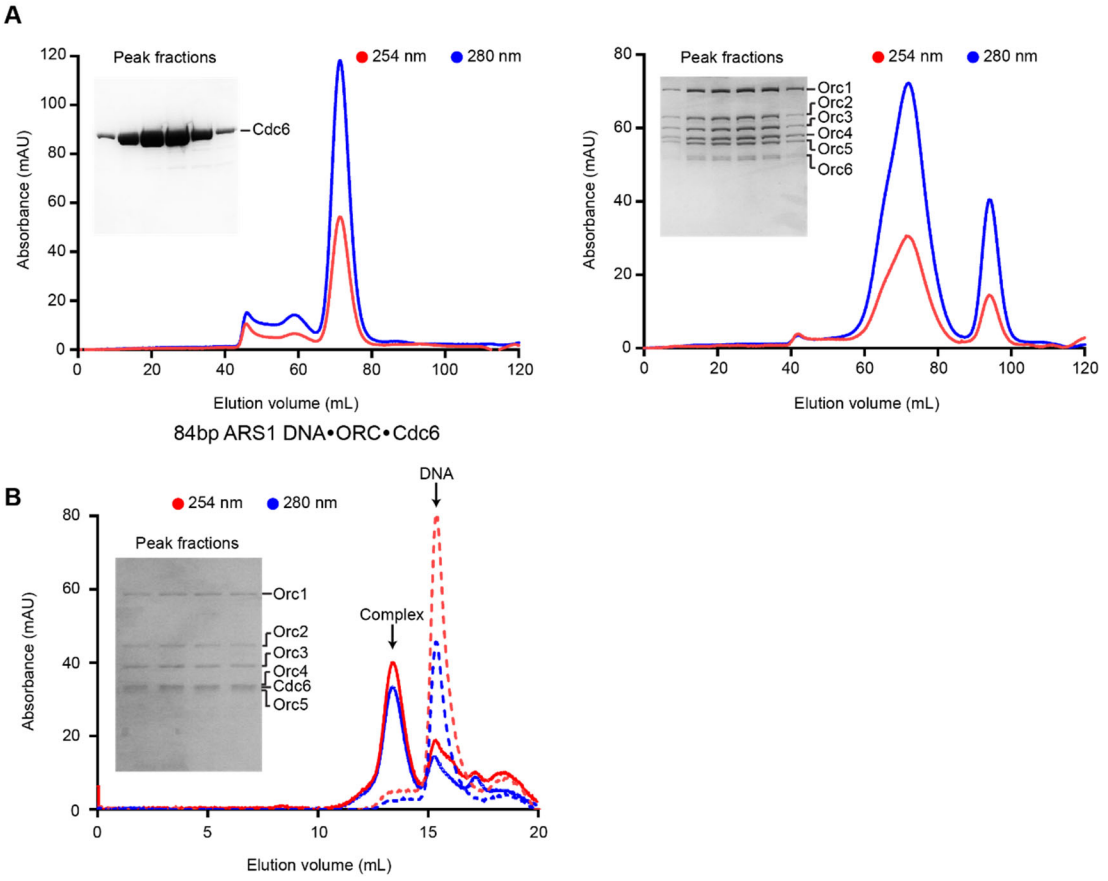
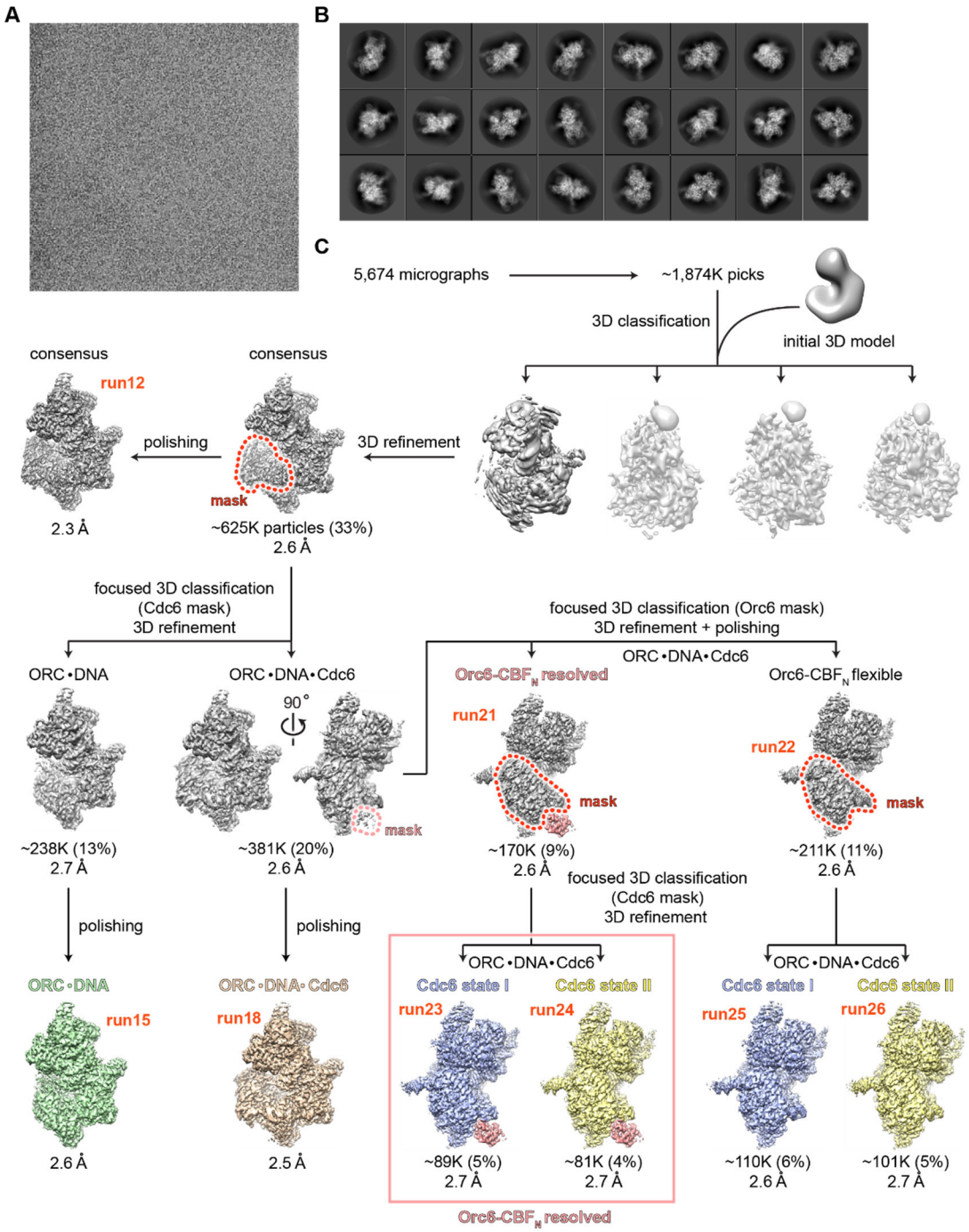

A
90°
Cdc6WH
Orc6Cyclin-fold
Orc1AAA+ lid
WH view
side view
B
Orc6 CBFN (ODC-CBFN)
Orc6 CBFN (MO)
180°
7.67
Average distance of C-alpha: 4.24
0.91
unaligned
C
Orc1
AAA+-lid
1
200
400
600
800
917
Cdc6
WHD
1
200
400
516
Orc6
1
200
400 435
CBFN
D
Inter-crosslinks (Threshold 10.3A)
satisfied
violated
90°
Cdc6WH
Orc1AAA+ lid
Orc6Cyclin-fold
E
Orc6
Cdc6
Orc5
Orc1
Orc4
Orc3
Orc2

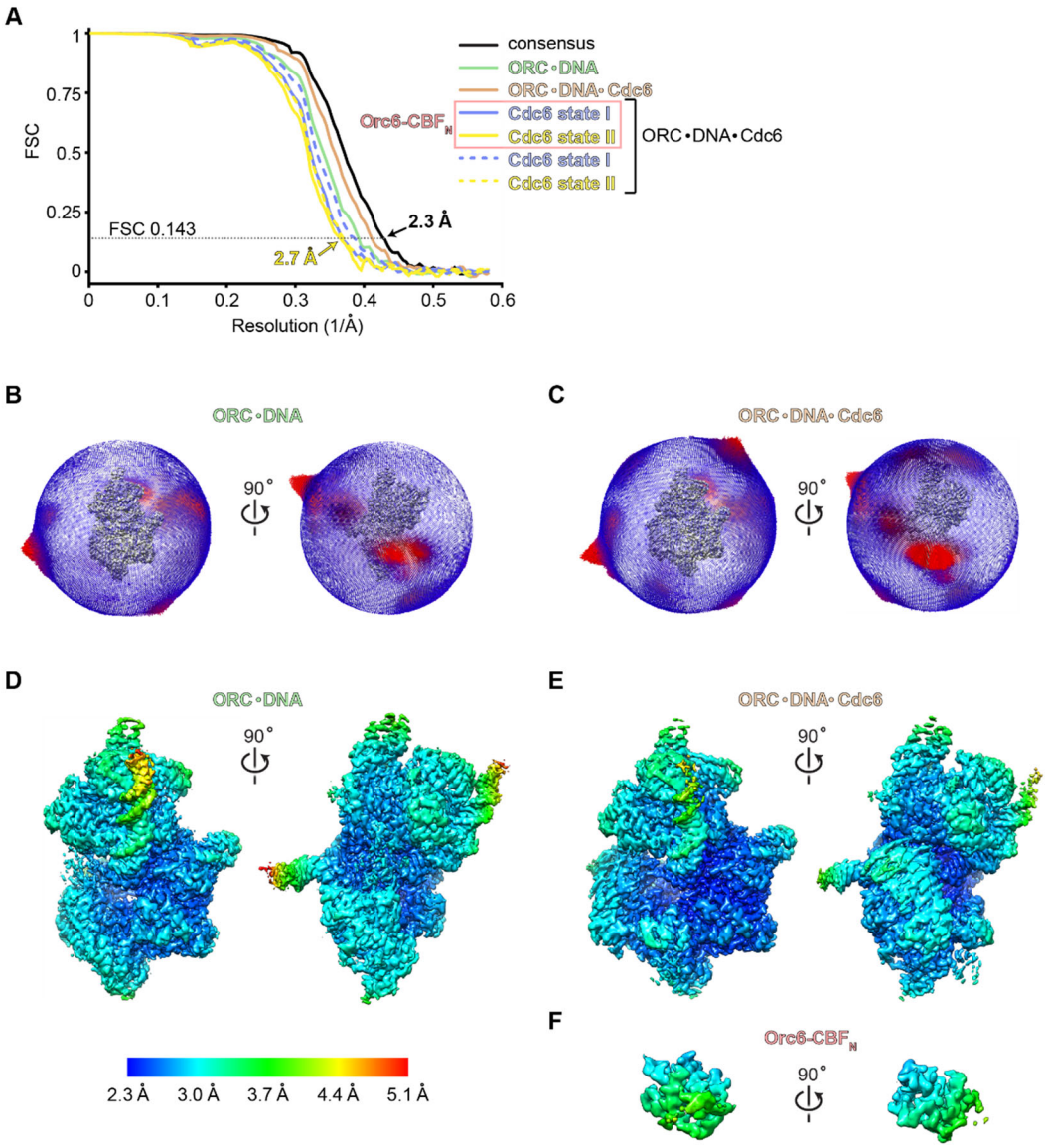
Figure 3



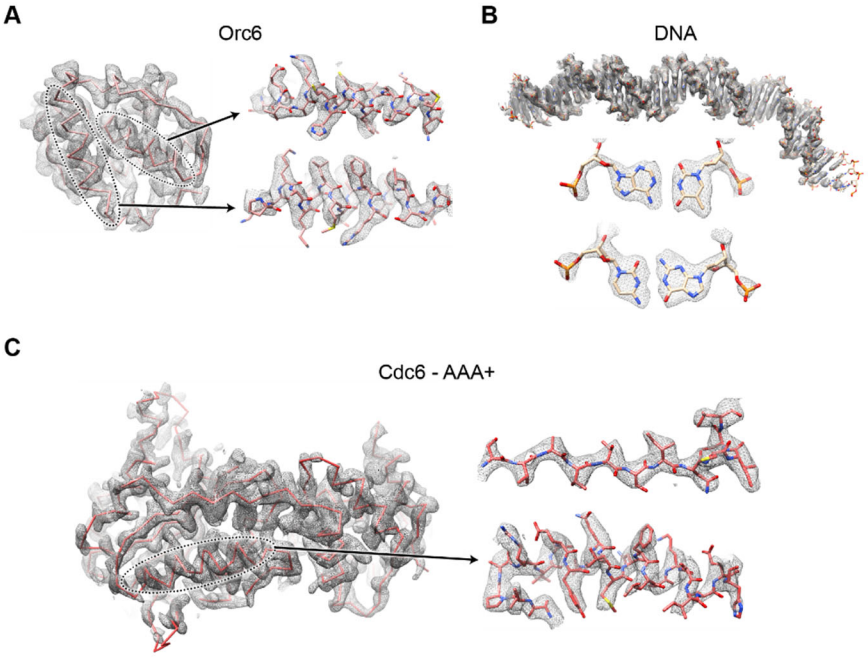
Supplementary Figure 1



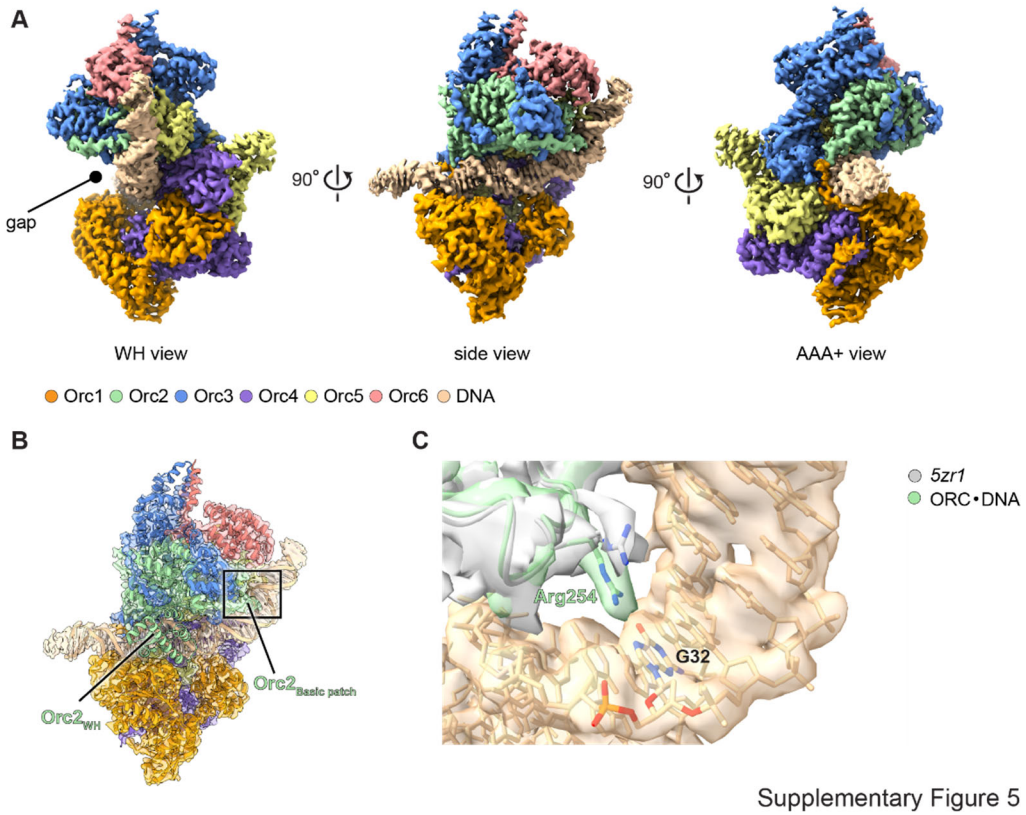
Supplementary Figure 2



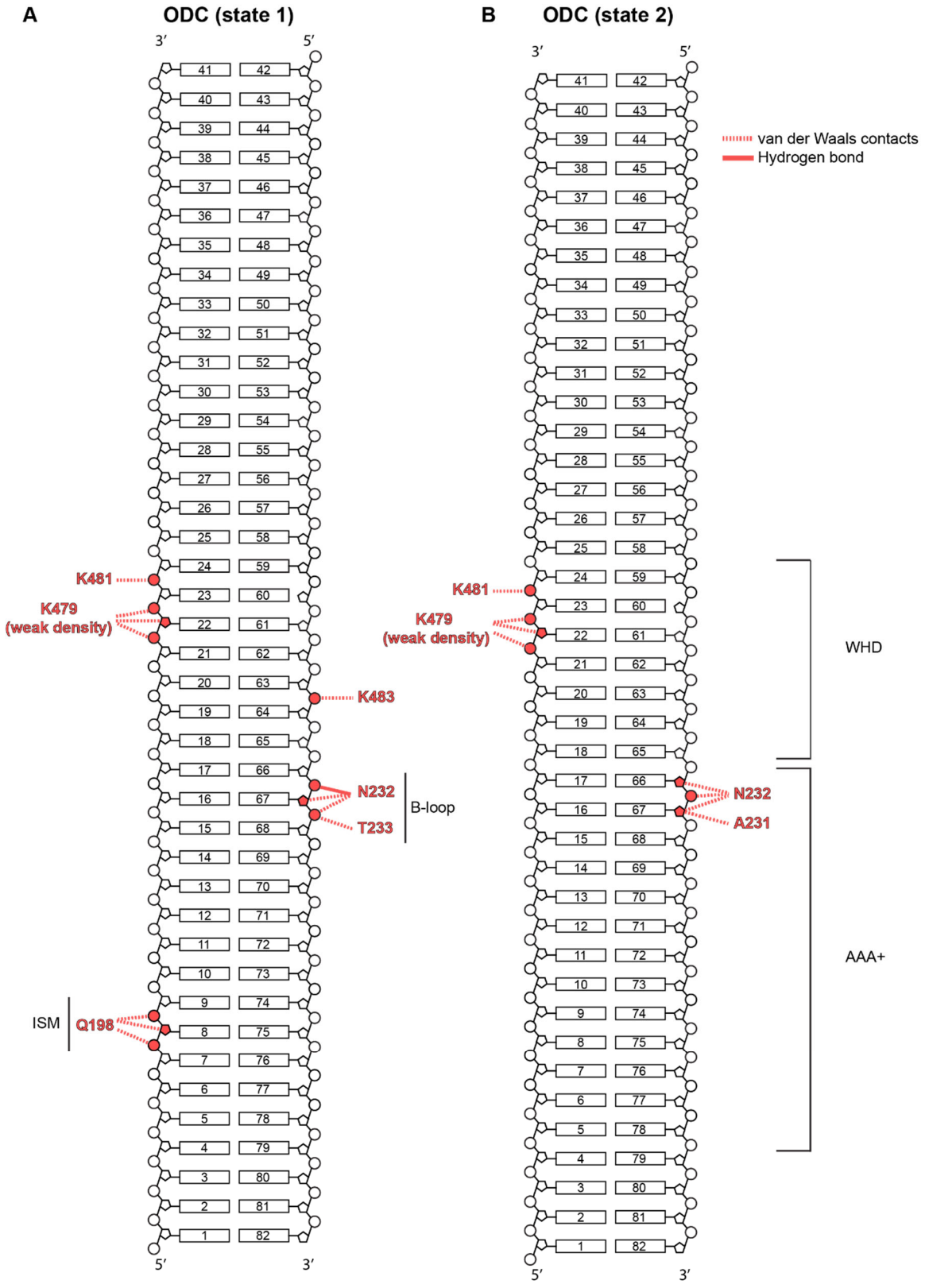
Supplementary Figure 3



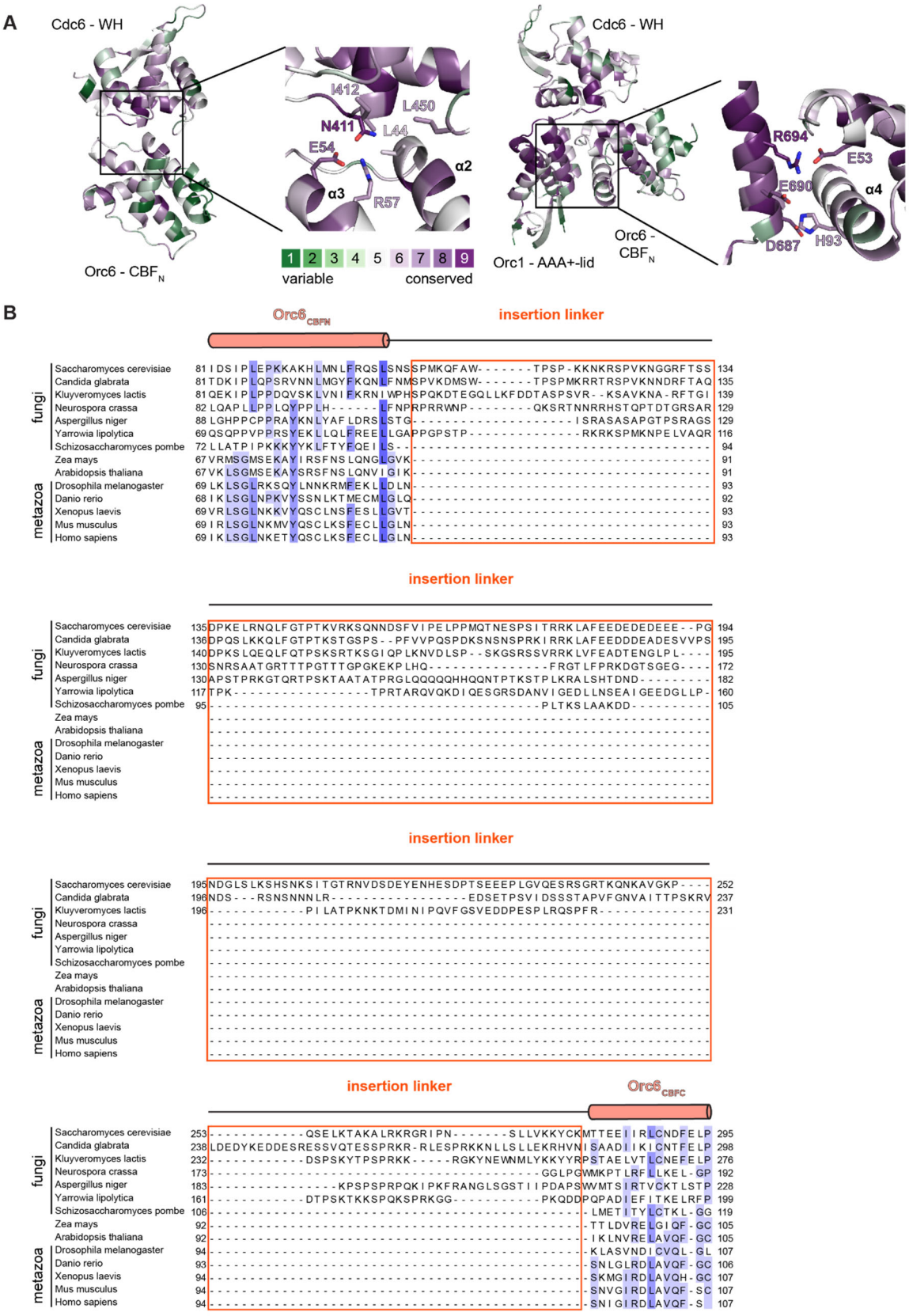
Supplementary Figure 4



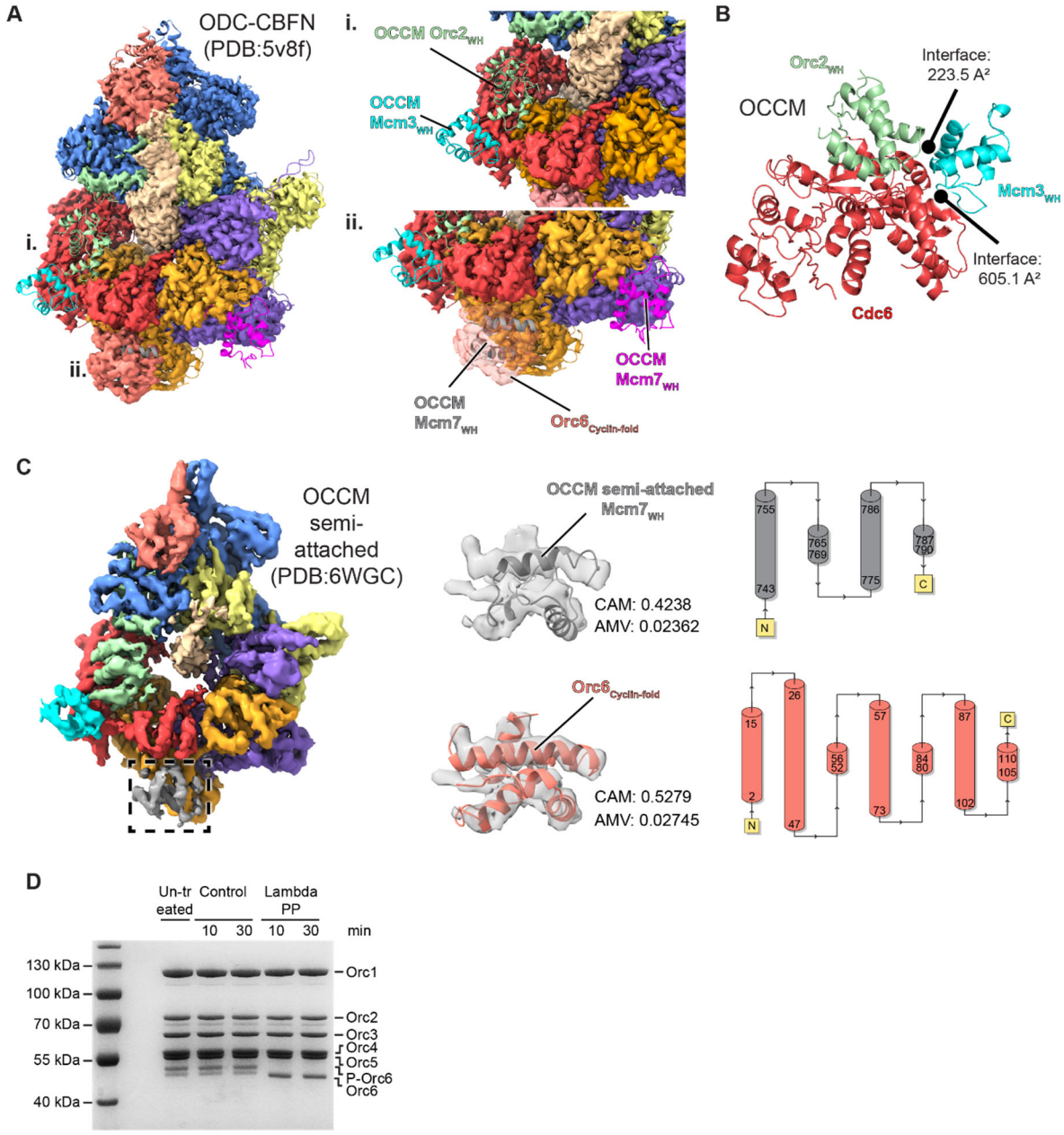
Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure 7



Supplementary Figure 8

Goal 3: Probing the functional consequences of Meier-Gorlin Syndrome mutations in a metazoan initiator and its co-loader.

Jan Marten Schmidt, Federica Galliano and Franziska Bleichert

Introduction

Mutations in ORC and its co-loader Cdc6 have been correlated with dwarfism disorders in humans, such as the Meier-Gorlin Syndrome[1–3]. The Meier-Gorlin Syndrome is a rare, autosomal recessive developmental disease that is characterized by a small stature, microtia and hypoplasia[4–6]. Notably, the majority of examined MGS patients have truncation or missense mutation in the DNA replication initiation factors Orc1, Orc4, Orc6, Cdc6, or Cdt1, which reduce protein function either in context of origin licensing or in regard to chromosome segregation[7,8]. The AAA+ domains of Orc1, Orc4 and Cdc6 are frequently mutated in individuals with MGS, but little is known about how these mutations affect different aspects of origin licensing[9].

Orc1 has been found to be mutated either at T574M, R666W or R720Q (the respective residues in *Drosophila* are T638, K730 and R784) while Orc4 is mutated at Y174C (*Drosophila*: Y162) and Cdc6 has the missense mutation T323R (*Drosophila*: T415) (Figure 1A, Table 1)[1–3]. These mutations are either monoallelic, homozygous or compound heterozygous[9]. Interestingly, all identified mutations map to or are in close proximity to protomer interfaces within the ORC/Cdc6 ring (Figure 1B). Orc1-R720 and Orc4-Y174 are the Sensor 2 and Tether residue, respectively, of the composite Orc1/Orc4 ATPase site (Figure 1C). Orc1-R666 (corresponds to K730 in *Drosophila*) functions as Tether in the Cdc6/Orc1 ATPase center (Figure 1C). The composite ATPase site formed by Cdc6/Orc1 and Orc1/Orc4 are the only ATPase centers that exhibit detectable activity in the eukaryotic initiator/co-loader assembly and are essential in budding yeast[10–19]. Hence, mutations in these interfaces, namely Orc1-R666W in Cdc6/Orc1, and Orc1-R720Q and Orc4-Y174C in Orc1/Orc4, are likely to impair nucleotide binding and/or ATP hydrolysis by ORC/Cdc6. Accordingly, the Orc1-R720Q and Orc4-Y174C mutant abolish and inhibit, respectively, ORC's hydrolase activity in context of human Orc1-5 *in vitro*[19].

Notably, Cdc6-T323 resides in an alpha-helix connected to the Sensor 1 residue of the Cdc6/Orc1 ATPase site (Figure 1C). Thus, mutation T323R in Cdc6 might influence Cdc6/Orc1 ATPase activity. T574 of Orc1 is located in proximity to the DNA-contacting initiator specific motif (ISM) of Orc4 and outside the Orc1/Orc4 ATPase site suggesting that it may inhibit DNA binding (Figure 1C).

To investigate the functional consequences of mutations in the AAA+ domains of ORC and Cdc6 identified in MGS patients, we probed mutant ORC DNA binding, ATP hydrolysis and ORC-DNA-Cdc6 (ODC) complex assembly *in vitro* using purified *Drosophila* proteins. Our *in vitro* studies thereby may contribute to our understanding of how nucleotide binding and hydrolysis by ORC controls initiator activity and regulates downstream initiation events.

Table 1: Summary of MGS in the AAA+ domains of human ORC and Cdc6 and their equivalents in *Drosophila*.

Protein	**Human**	***Drosophila***	**Function**
Orc1	T547M	T638M	Unknown
	R666W	K730W	Tether
	R720Q	R784Q	Sensor 2
Orc4	Y174C	Y162C	Tether
Cdc6	T323R	T415R	Unknown

Results

To better understand the functional consequences of the MGS-mutations in the AAA+ domains of ORC and Cdc6 on origin licensing, we purified mutant ORC and Cdc6 proteins and characterized them biochemically. As isolation of meaningful quantities of highly pure human ORC and Cdc6 is challenging and as the AAA+ domains are structurally and functionally highly conserved across eukaryotes, we used *Drosophila* proteins harboring the respective MGS-associated mutations for our study (Figure 2A). From here on, all mutations referred to will be the human-equivalent residues in *Drosophila*.

ATP hydrolysis by ORC is an essential function *in vivo* and has been suggested to be required for iterative cycles of helicase loading[15]. Consequently, we probed the ATPase activity of *Drosophila* MGS mutant ORC proteins by performing an enzyme-coupled spectrophotometric assay (Figure 2B). As a control, we used wild-type *Dm*ORC, which has a catalysis rate of 8.6 ATP per ORC per min, and a Orc1-WB (D684A, E685A) and Orc4-RF (R197A), both render the complex catalytically dead. As expected, an alanine substitution of the Orc1-RF (R734A), which mediates hydrolysis in the Cdc6/Orc1, has only a minor inhibitory effect on the ATPase rate of ORC in our assay. Interestingly, substitution of threonine by methionine at position 638 in Orc1 accelerated ORC's ATPase activity by approximately 2-fold (Figure 2B). In contrast, MGS mutants Orc1-R784Q and Orc4-Y162C showed no detectable ATPase activity suggesting that either nucleotide binding and/or hydrolysis are severely inhibited by these mutations (Figure 2B). Substitution of lysine 730 to tryptophan in Orc1 resulted in a 2-fold reduction of ATPase hydrolysis to 4.7 ATP per ORC per min (Figure 2C). Taken together, our data shows that MGS mutants in the AAA+ domain of Orc1 (T638M, K730W, R784Q) and Orc4 (Y162C) deregulated ORC's ATPase activity suggesting that these disease-associated mutants may interfere with iterative rounds of helicase loading *in vivo*.

Recognition of origin DNA by ORC precedes co-loader recruitment and is fundamental for helicase loading. Therefore, we used a fluorescence anisotropy DNA binding assay to test mutant ORC binding to fluorophore-labeled DNA at steady-state (Figure 2B, Table 2). In presence of ATP, wild-type *Dm*ORC binds DNA with an apparent (app.) K_D of approx. 14 nM, which is only slightly different for the mutants Orc1-R734 ($K_{D,app.}$: 11.6

nM) and Orc1-K730W ($K_{D,app.}$: 8.3 nM). Notably, alanine substitutions in the Walker B motif (D684A, E685A) of Orc1, the Orc4 R-finger (R197A) and MGS mutation Orc1-R784Q and Orc4-Y162C abolished ORC-DNA binding. The Orc1-T638M, which is hyperactive in terms of ATPase hydrolysis, showed reduced but not impaired DNA binding ($K_{D,app.}$: 104.7 nM). Accordingly, we suggest that MGS mutations Orc1-R784Q and Orc4-Y162C severely compromise ORC's association with chromatin and consequently helicase loading *in vivo*, while mutations Orc1-K730W and Orc1-T638M likely have no or only a mild effect on recruitment of ORC to chromatin.

ATP-dependent recruitment of Cdc6 to DNA-bound ORC is imperative for origin licensing. To investigate whether the described MGS mutations in ORC and Cdc6 have an effect on ternary ORC-DNA-Cdc6 complex formation, we incubated ORC and Cdc6 presence of DNA and ATP and probed complex formation by affinity-tag pulldown using an MBP-tag on the N-terminus of Cdc6 as a bait (Figure 3, Table 2). When pulling on wild-type MBP-Cdc6 protein, we observed markedly reduced ternary complex assembly in case of the Orc1-K730W, RF (R734A) and WB mutants (Figure 3A). Mutant protein Orc1-R784Q, Orc4-Y162C and Orc4-RF (R197A) were not able to stably associate with Cdc6 under the tested conditions. Surprisingly, we found that Orc1-T638M, which shows reduced DNA binding and upregulated ATPase activity, assembles into a complex with DNA and Cdc6 at comparable levels to wild-type protein. We then also tested complex assembly with Cdc6-T415R and E377Q (WB) mutants and wild type ORC protein (Figure 3B). MBP-Cdc6-T415R efficiently pulled down ORC suggesting that ternary complex formation is not inhibited by Cdc6-T415R in MGS. As a control we used a mild Cdc6-WB mutant (E377Q), which showed reduced complex assembly. In summary, our pull-down experiments suggest that MGS mutations in the ATPase centers of Cdc6/Orc1 (Orc1-K730W) and Orc1/Orc4 (Orc1-R784Q and Orc4-Y162C), which may affect nucleotide binding, impair ternary ORC-DNA-Cdc6 complex assembly and consequently are likely to hinder origin licensing.

Taken together, our *in vitro* data shows that MGS mutation in the AAA+ domain of Orc1 and Orc4 deregulate multiple initiator functions. In agreement with previous studies[8], we suggest that these mutations are hypomorphs and likely do not support efficient origin

licensing. Notably, substitution of the Sensor 2 residue in Orc1 (R784) to glutamine and mutation of the Orc4-Tether (Y162) to cysteine, respectively, impairs all three tested complex activities suggesting that they severely reduce ORC function *in vivo*.

Discussion

Our biochemical interrogation of monoallelic, homozygous or compound heterozygous mutations in the AAA+ domains of ORC and Cdc6 that are found in MGS patients suggest that compromised ATP hydrolysis, DNA binding and Cdc6 recruitment by the initiator may be an underlying cause for MGS. All tested MGS mutants exhibited reduced activity in one or more of the tested functionalities suggesting that all of them are hypomorph mutants. However, MGS mutations located in the Cdc6/Orc1 (K730W) and Orc1/Orc4 (R784, Y162C) ATPase sites are likely more severe than Orc1-T638M and Cdc6-T415R as they prevent stable Cdc6 recruitment to DNA-bound ORC, which is essential for replicative helicase loading (Figure 3A).

ATPase activity of Orc1-R784Q (Sensor 2) and Orc4-Y162C (Tether) have been studied before in context of the human Orc1-5 complex[19]. The authors reported that R720Q (human equivalent to *Drosophila* R784Q) abolishes ORC's ATPase activity while Orc4-Y174C (human equivalent to *Drosophila* Y162C) showed a 2-fold reduction[19]. In contrast, our findings suggest that both (Orc1-R784Q and Orc4-Y162) render the complex catalytically inactive. The reason for this discrepancy is unclear, but could be due to the different assays used for measuring ATPase activity. Tocilj et al. performed endpoint measurements based on the quantification of γ-^{32}P labeled nucleotide by PEI-cellulose TLC plate and PhosphoImager[19]. The authors used a concentration of 10 µM ORC protein in their assay and covered a range of 12.5-200 µM ATP. We determined steady-state ATPase hydrolysis by an NADH-coupled enzymatic ATPase assay over 2 hours at 1 min intervals, using 10-fold less ORC protein for our assay and measuring ATPase activity at 2 µM to 4 mM ATP. Thus, we may be below the threshold for detecting very low levels of activity at high ORC concentrations.

Despite the fact that the Orc4-Y162C mutant is catalytically inactive, does not bind DNA and shows defective ODC complex formation, it seems to retain some function *in vivo,* as homozygous MGS patients have been found. The corresponding mutation in yeast (Y232C) cannot fully rescue an haploid Orc4 deletion strain when expressed on a plasmid, but results in slow growth and defective S phase dynamics[3]. Moreover, haploid Orc4-Y232C yeast cells do not initiate replication at rDNA sites as frequently as wild-type

cells resulting in chromosome breakage and lower ribosomal DNA copy number[20]. Notably, *Drosophila melanogaster* homozygous for the Orc4-Y162C mutation were viable, but exhibited tissue-specific defects and females were sterile[21] [McDaniel 2020]. The observation that homozygous Orc4-Y162C *Drosophila* mutants are viable is likely due to maternally stored wild-type protein as fly embryos with only maternal Orc4-Y162C do not survive[21].

The molecular mechanisms by which the monoallelic Orc1-T638M mutation may result in the MGS phenotype are unclear. Orc1-T638 is located outside the Orc1/Orc4 ATPase site, but reaches towards the ISM of Orc4, an element important for DNA binding by metazoan ORC (Figure 1C). Consequently, we anticipated that the Orc1-T638M mutant might be defective in DNA recognition, but has normal ATP activity. Surprisingly, we found that substitution of threonine-638 to methionine in *Dm*ORC increased ATPase activity by 2-fold while inhibiting DNA binding (Figure 2B and C). The physical basis for the former remains to be shown, but since ATPase hydrolysis is thought to destabilize ORC on DNA we suggest that the observed reduction of DNA binding is a consequence of the accelerated ATPase rate[10,12,15,22–24]. Notably, ODC complex assembly remained at wild-type levels suggesting that Cdc6 recruitment compensates for the observed DNA binding defect of Orc1-T638M (Figure 3A).

Both copies of Cdc6 have been found to be mutated at position 323 from threonine to arginine in a MGS patient, which has been proposed to result in a partial loss of protein function[1,25]. Here we show that *Drosophila* Cdc6-T415R binds to ORC-DNA with wild-type efficiency suggesting that other protein functions might be affected. In *Drosophila*, T415 is part of an alpha-helix directly downstream of the Sensor 1 residue (N410) of the Cdc6/Orc1 interface. Mutations in the Sensor 1 of budding yeast Cdc6 have been shown to inhibit ATPase activity and impair helicase loading potentially due to reduced Cdt1 release *in vitro*[26,27]. In yeast, Cdc6 ATPase activity is not required for helicase loading *in vitro* but is critical for cell survival and has been suggested to help resolve non-functional loading intermediates important for the recycling of origin licensing factors[11,13,14,18,28–31]. Hence, we suggest that MGS mutation Cdc6-T323R may inhibit Cdc6 ATPase activity

and thereby deregulate the ability of Cdc6 to release non-functional or incomplete loading intermediates *in vivo*.

Our findings presented here extend our knowledge about how MGS-associated mutations in the AAA+ domains of Orc1, Orc4 and Cdc6 alter initiator function. These defects are likely to deregulate origin licensing *in vivo*. In the future, it will be necessary to extend our *in vitro* studies to address more comprehensively how MGS mutations influence the dynamics of replicative helicase recruitment and loading. For instance, the use of isotope-labeled ATP to investigate nucleotide binding by ORC mutants would allow one to test how MGS mutations impact ORC's ability to transition from an autoinhibited to an active conformation by negative-stain EM. DNA binding and ATPase hydrolysis assays in presence of Cdc6 would allow one to evaluate ternary complex behavior and ATPase activity by Cdc6. Finally, the use of our recently established *in vitro* metazoan helicase loading assay[32], would allow one to probe Mcm2-7 recruitment and loading by mutant ORC and Cdc6. Such studies can be extended to further MGS-associated mutations, found in other initiation factors, such as Cdt1 and Mcm5 and will confirm predictions of function based on high resolution structural studies.

Methods

Cloning of *Drosophila* ORC and Cdc6 expression constructs

Individual wild-type *Dm*ORC subunits and *Dm*Cdc6 were cloned into pFastBac vectors adapted for Ligation-independent cloning (University of California, Berkeley MacroLab). Orc1 and Orc4 were N-terminally tagged with 6xHis and MBP, respectively. Both affinity-tags were separated by a TEV protease cleavage site from the open reading frame. DmCdc6 expression vectors contained either an N-terminal 6xHis-MBP-10xN-TEV or MBP-TEV affinity-tag. MGS and control ORC and Cdc6 mutants were generated by site-directed mutagenesis. In some cases, individual ORC subunits were combined into MultiBac expression vectors by BioBrick cloning (University of California, Berkeley MacroLab) to boost protein expression.

Generation of baculoviruses for protein expression

Individual pFastBac and MulitBac expression vectors were transformed into DH10Bac cells for generation of Bacmids. Bacmids were transfected with Cellfectin II (Thermo Fisher Scientific) into Sf9 cells and high-titer viruses were obtained by two rounds of baculoviruses amplification in Sf9 cells.

Purification of recombinant *Drosophila* ORC and Cdc6 protein from insect cells

Purification of wild-type and mutant *Dm*ORC was performed as described in Ref. 24,32,33. In brief, four liters of Hi5 cells were co-infected with high-titer viruses encoding for Orc1-6. Infected cells were grown for 48h and harvested by centrifugation. The cell pellet was resuspended in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 50 mM imidazole, pH 7.8, 10% glycerol, 200 µM PMSF, 1 µg/mL leupeptin, 1 mM BME (lysis buffer) prior to cell lysis by sonication. The soluble proteins were separated from the whole cell lysate by centrifugation at 142,414 × *g* for 45 min. The lysate was further clarified by the addition of ammonium sulfate (20% (v/v) final concentration) followed by centrifugation at 142,414 × *g* for 45 min. The clarified lysate was applied to a 5 mL HisTrap HP nickel-affinity chromatography column (GE Healthcare) and washed with lysis buffer. Subsequently, a 50–250 mM imidazole gradient in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 1 mM BME was applied to the column to elute bound *Dm*ORC. The eluate

was run over a 10 mL amylose resin (New England Biolabs), washed with 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10% glycerol, 1 mM BME and captured protein was eluted with 20 mM maltose. Affinity-tags were removed from Orc1 and Orc4 by the addition of TEV protease and overnight incubation. Digested ORC protein was run over a 5 mL HisTrap HP nickel-affinity chromatography column (GE Healthcare) and the flow-through was further purified by applying it to a HiPrep 16/60 Sephacryl S-400 HR gel filtration chromatography column (GE Healthcare) equilibrated in 25 mM HEPES-KOH pH 7.6, 500 mM potassium glutamate, 10% glycerol, 1 mM DTT. Concentrated peak fractions were flash frozen and stored at -80°C. All purification steps were performed at 4°C.

MBP-tagged and 6xHis-MBP-tagged *Dm*Cdc6 were purified as described in Ref. 244 and similarly as outlined for *Dm*ORC. In case of MBP-tagged *Dm*Cdc6, gel filtration chromatography was performed on a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 25 mM HEPES-KOH, pH 7.6, 500 mM Kglut, 10% glycerol, 1 mM DTT. Wild-type and mutant 6xHis-MBP-DmCdc6 was purified by gel filtration chromatography on a a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) or HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 25 mM HEPES-KOH, pH 7.6, 300 mM KCl, 10% glycerol, 10 MgOAc, 1 mM DTT. Purified proteins were flash frozen and stored at -80°C before use.

ATPase activity assay

The spectrophotometric enzyme-coupled ATPase activity assay was performed precisely as in Ref. 32.

Fluorescence anisotropy assay for ORC DNA binding

The fluorescence anisotropy assay was performed as described in Ref. 32,34 with the exception that fluorescence anisotropy was measured at 17 (instead of 18) ORC protein concentrations with the lowest concentration being 0.08 nM *Dm*ORC.

DNA annealed

Single-stranded complementary oligos were obtained from IDT (LubioScience). Oligos were resuspended in 200 µL of 10 mM Tris-HCl pH 8.0, 5 mM MgCl2 at 10 µM (fluorophore-labeled) and 25 µM (unlabeled), respectively, and annealed by heating them to 95°C for 5 min in a water bath, which subsequently was allowed to cool down to room temperature. Formation of double-stranded DNA was confirmed by native-PAGE, which was run in 1xTris-borate buffer pH 8.3 and stained with SYBR Safe (Thermo Fisher Scientific) as described in Ref. 32.

Table 2: List of oligonucleotides used in this study.

Name	Sequence (5'-3')
FL_ARS1_40bp.S	/5FluorT/TTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTG
ARS1_40bp.AS	CAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAA
ARS1_60bp.S	CCTGCAGGCCTTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTGTAAAATAACA
ARS1_60bp.AS	TGTTATTTTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAAGGCCTGCAGG
ARS1_84bp.S	TTTGTGCACTTGCCTGCAGGCCTTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTGTAAAATAACAAGATGTAAAGAT
ARS1_84bp.AS	ATCTTTACATCTTGTTATTTTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAAGGCCTGCAGGCAAGTGCACAAA

/5FluorT/: 5' Fluorescein dT

Affinity-tag pull down

Pull down assays using 6xHis-MBP- or MBP-tagged wild-type or mutant *Dm*Cdc6 as a bait were performed by incubating wild-type or mutant *Dm*ORC with 60bp (Figure 3B) and 84bp (Figure 3A) ARS1 dsDNA, respectively, at an equimolar concentration of 100 nM in binding buffer (25 mM HEPES 7.6, 300 mM Kglut, 10% glycerol, 10 mM MgOAc, 1 mM DTT, +/- 1 mM ATP) for 15 min at room temperature. Subsequently, wild-type MBP-DmCdc6 (Figure 3A) and wild-type or mutant 6xHis-MBP-DmCdc6 (Figure 3B), respectively, were added to the reaction at a final concentration of 200 nM. Samples were incubated for further 15 min at room temperature before 5 µL of each sample was transferred to a new Eppendorf tube containing 20 µL 1x SDS-loading dye (Input sample).

Subsequently, 25 µL of Amylose-resin beads (NEB) equilibrated in binding buffer were added to the pull down reaction. The samples were incubated for 30 min at room temperature and frequently mixed by careful pipetting. From here on the sample was kept on ice. Protein-bound beads were washed twice with 1 mL ice-cold binding buffer. Afterwards, all residual supernatant was removed and bound proteins were eluted by adding 25 µL elution buffer (25 mM HEPES 7.6, 300 mM Kglut, 10% glycerol, 10 mM MgOAc, 20 mM Maltose, 1 mM DTT, +/- 1 mM ATP) to the beads. The samples were incubated for 5 min at room temperature before 20 µL of eluate was transferred to a fresh Eppendorf tube and topped up with 5 µL 5xSDS-loading dye (Eluate sample). 10 µL of Input and Eluate samples were separated on a 10%-SDS-PAGE gel run in 1x Tris-Glycine-SDS buffer. SDS-PAGE gels were developed by silver-staining.

Figure legend

Figure 1: Meier-Gorlin Syndrome mutations in the AAA+ domain of *Dm*ORC and Cdc6. A) Domain organization of eukaryotic Orc1-6 and Cdc6. MGS-associated mutations in the AAA+ domain of Orc1, Orc4 and Cdc6 are indicated as the equivalent residues in *Drosophila*. **B)** Structure of the *Drosophila* ORC-DNA-Cdc6 (ODC) complex colored by subunit (color scheme as in **a**) (PDB: 7JK4[32]). The location of the indicated MGS mutations is boxed. **C)** Zoom on MGS mutations in context of *Dm*ODC. ORC subunits and Cdc6 are colored as in Figure 1. Residues mutated in MGS and key catalytic amino acid side chains are displayed as sticks. ISM – Initiator Specific Motif, TH – Tether, RF – Arginine Finger, S2 - Sensor 2, ATP – Adenosine TriPhosphate, WA – Walker A, WB – Walker B, S1 – Sensor 1.

Figure 2: ATPase activity and DNA binding by *Drosophila* ORC harboring MGS mutations. A) SDS-PAGE gel of purified wild-type and mutant ORC and Cdc6 protein. **B)** ATPase activity of wild-type and mutant *Dm*ORC measured by a enzyme-coupled spectrophotometric assay. ATPase rates are plotted as k_{cat} per min, n = number of independent experiments. **C)** Fluorescence-anisotropy DNA binding assay. Steady-state binding affinities of *Dm*ORC proteins to fluorophore-labeled DNA are given as apparent K_D, n = number of independent experiments. ND – binding curve did not saturate, therefore binding affinity could not be resolved.

Figure 3: Ternary ORC-DNA-Cdc6 complex formation is reduced by MGS mutations. A) Silver-stained SDS-PAGE gel of affinity-tag pulldown of wild type and mutant *Dm*ORC in presence of DNA and ATP using MBP-*Dm*Cdc6 as a bait. **B)** Same as in **a** but with wild-type *Dm*ORC and wild-type and mutant MBP-*Dm*Cdc6 protein. **C)** Zoom on Threonine-415 (T415) of Cdc6 in the *Drosophila* ODC complex (PDB: 7JK4[32]). T415 is part of a short alpha-helix that is directly connected to the Sensor 1 of the Cdc6/Orc1 ATPase site.

References

1. Bicknell, L. S. *et al.* Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat. Genet.* **43**, 350–356 (2011).

2. Bicknell, L. S. *et al.* Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome. *Nat. Genet.* **43**, 350–356 (2011).

3. Guernsey, D. L. *et al.* Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nat. Genet.* **43**, 360–365 (2011).

4. MEIER, Z., POSCHIAVO & ROTHSCHILD, M. [Case of arthrogryposis multiplex congenita with mandibulofacial dysostosis (Franceschetti syndrome)]. *Helv. Paediatr. Acta* **14**, 213–216 (1959).

5. Gorlin, R. J., Cervenka, J., Moller, K., Horrobin, M. & Witkop, C. J. J. Malformation syndromes. A selected miscellany. *Birth Defects Orig. Artic. Ser.* **11**, 39–50 (1975).

6. Boles, R. G., Teebi, A. S., Schwartz, D. & Harper, J. F. Further delineation of the ear, patella, short stature syndrome (Meier-Gorlin syndrome). *Clin. Dysmorphol.* **3**, 207–214 (1994).

7. de Munnik, S. A. *et al.* Meier-Gorlin syndrome. *Orphanet J. Rare Dis.* **10**, 114 (2015).

8. Hossain, M. & Stillman, B. Meier-Gorlin Syndrome. in *The Initiation of DNA Replication in Eukaryotes* (ed. Kaplan, D. L.) 503–524 (Springer International Publishing, 2016). doi:10.1007/978-3-319-24696-3_25

9. De Munnik, S. A. *et al.* Meier-Gorlin syndrome genotype-phenotype studies: 35 individuals with pre-replication complex gene mutations and 10 without molecular diagnosis. *Eur. J. Hum. Genet.* **20**, 598–606 (2012).

10. Klemm, R. D., Austin, R. J. & Bell, S. P. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell* **88**,

493–502 (1997).

11. Perkins, G. & Diffley, J. F. X. Nucleotide-Dependent Prereplicative Complex Assembly by Cdc6p, a Homolog of Eukaryotic and Prokaryotic Clamp-Loaders. *Mol. Cell* **2**, 23–32 (1998).

12. Chesnokov, I., Remus, D. & Botchan, M. Functional analysis of mutant and wild-type Drosophila origin recognition complex. *Proc. Natl. Acad. Sci.* **98**, 11997 LP – 12002 (2001).

13. Schepers, A. & Diffley, J. F. X. Mutational analysis of conserved sequence motifs in the budding yeast cdc6 protein11Edited by M. Yaniv. *J. Mol. Biol.* **308**, 597–608 (2001).

14. Takahashi, N., Tsutsumi, S., Tsuchiya, T., Stillman, B. & Mizushima, T. Functions of sensor 1 and sensor 2 regions of Saccharomyces cerevisiae Cdc6p in vivo and in vitro. *J. Biol. Chem.* **277**, 16033–16040 (2002).

15. Bowers, J. L., Randell, J. C. W., Chen, S. & Bell, S. P. ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol. Cell* **16**, 967–978 (2004).

16. Speck, C., Chen, Z., Li, H. & Stillman, B. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat. Struct. Mol. Biol.* **12**, 965–971 (2005).

17. Randell, J. C. W., Bowers, J. L., Rodríguez, H. K. & Bell, S. P. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol. Cell* **21**, 29–39 (2006).

18. Chang, F. *et al.* Cdc6 ATPase activity disengages Cdc6 from the pre-replicative complex to promote DNA replication. *Elife* **4**, (2015).

19. Tocilj, A. *et al.* Structure of the active form of human origin recognition complex and its ATPase motor module. *Elife* **6**, (2017).

20. Sanchez, J. C. *et al.* Defective replication initiation results in locus specific chromosome breakage and a ribosomal RNA deficiency in yeast. *PLoS Genet.*

13, e1007041 (2017).

21. McDaniel, S. L. *et al.* Tissue-Specific DNA Replication Defects in Drosophila melanogaster Caused by a Meier-Gorlin Syndrome Mutation in Orc4. *Genetics* **214**, 355 LP – 367 (2020).

22. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).

23. Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *EMBO J.* **23**, 897–907 (2004).

24. Bleichert, F., Botchan, M. R. & Berger, J. M. Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321–326 (2015).

25. Yao, L., Chen, J., Wu, X., Jia, S. & Meng, A. Zebrafish cdc6 hypomorphic mutation causes Meier-Gorlin syndrome-like phenotype. *Hum. Mol. Genet.* **26**, 4168–4180 (2017).

26. Speck, C. & Stillman, B. Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *J. Biol. Chem.* **282**, 11705–11714 (2007).

27. Evrin, C. *et al.* In the absence of ATPase activity, pre-RC formation is blocked prior to MCM2–7 hexamer dimerization. *Nucleic Acids Res.* **41**, 3162–3172 (2013).

28. Fernández-Cid, A. *et al.* An ORC/Cdc6/MCM2-7 Complex Is Formed in a Multistep Reaction to Serve as a Platform for MCM Double-Hexamer Assembly. *Mol. Cell* **50**, 577–588 (2013).

29. Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. X. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339–343 (2013).

30. Coster, G., Frigola, J., Beuron, F., Morris, E. P. & Diffley, J. F. X. Origin Licensing Requires ATP Binding and Hydrolysis by the MCM Replicative Helicase. *Mol. Cell*

55, 666–677 (2014).

31. Kang, S., Warner, M. D. & Bell, S. P. Multiple Functions for Mcm2–7 ATPase Motifs during Replication Initiation. *Mol. Cell* **55**, 655–665 (2014).

32. Schmidt, J. M. & Bleichert, F. Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nat. Commun.* **11**, 4263 (2020).

33. Bleichert, F. *et al.* A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *Elife* **2**, e00882 (2013).

34. Bleichert, F., Leitner, A., Aebersold, R., Botchan, M. R. & Berger, J. M. Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5906–E5915 (2018).

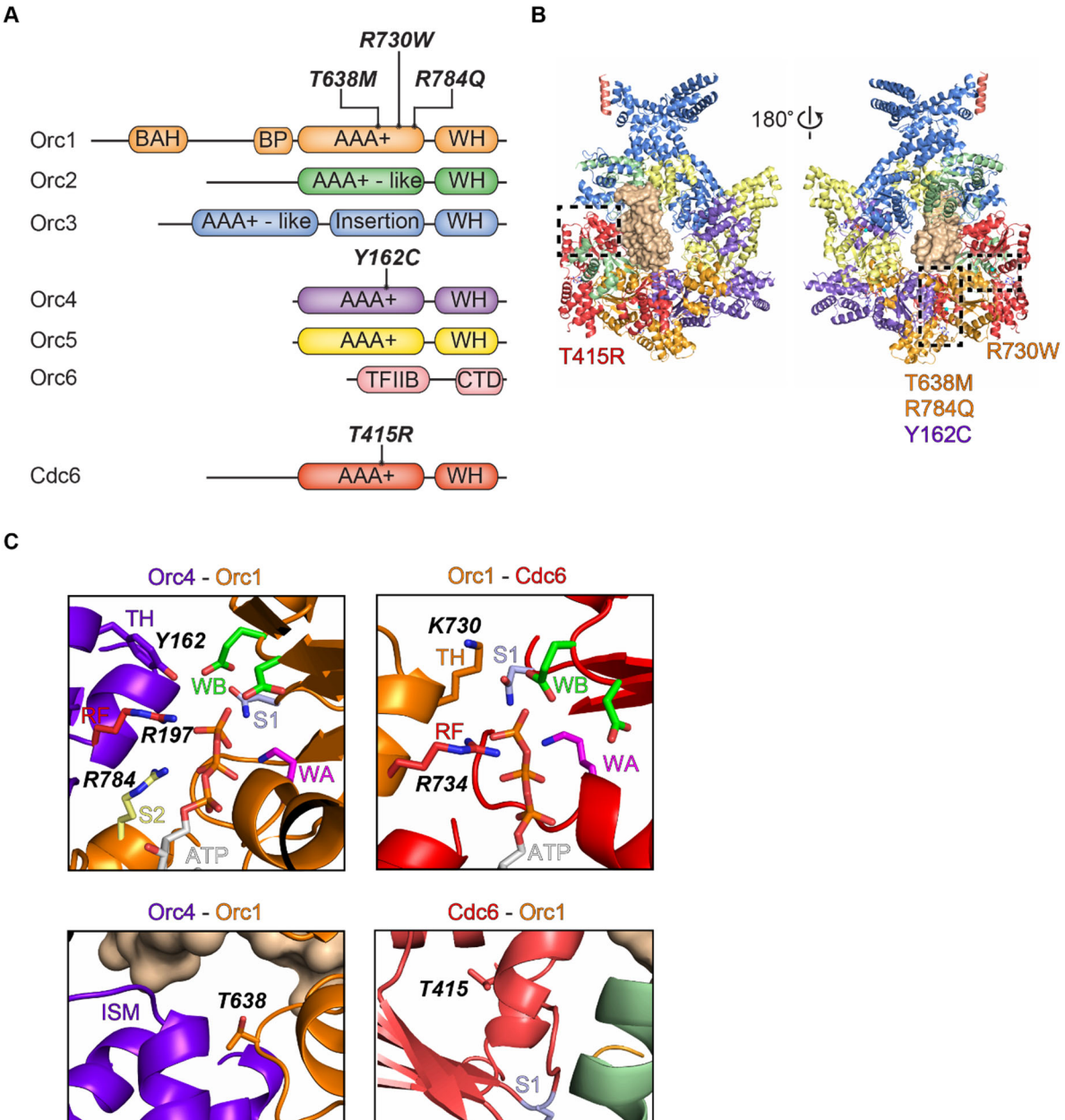
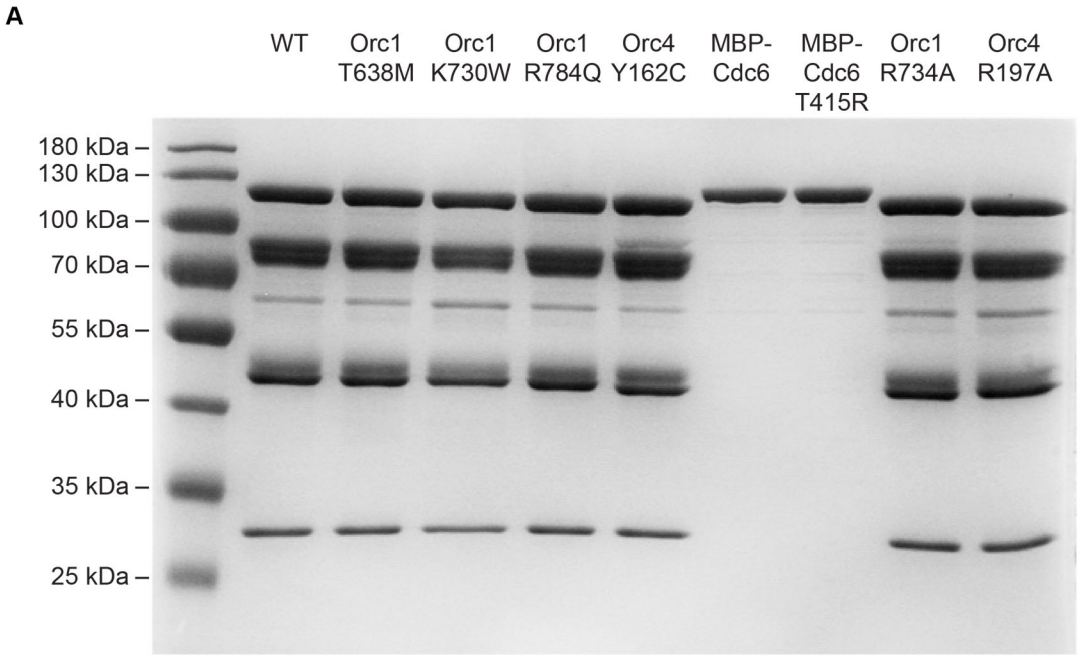
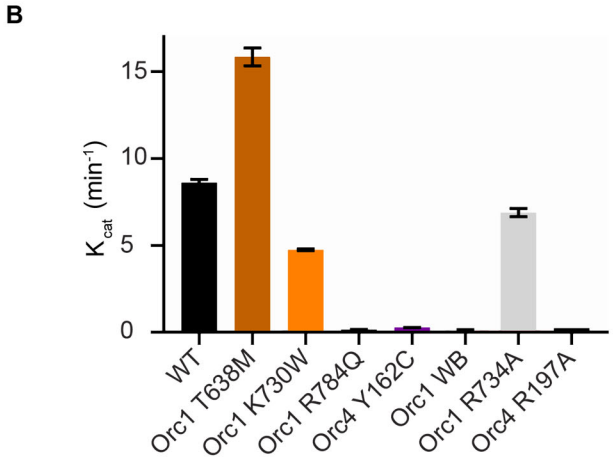
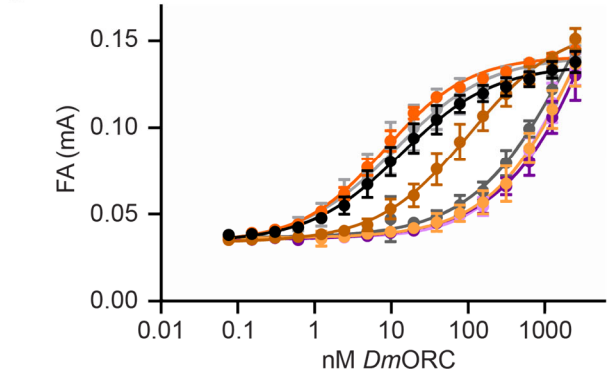

A
R730W
T638M
R784Q
Orc1
BAH
BP
AAA+
WH
Orc2
AAA+ - like
WH
Orc3
AAA+ - like
Insertion
WH
Y162C
Orc4
AAA+
WH
Orc5
AAA+
WH
Orc6
TFIIB
CTD
T415R
Cdc6
AAA+
WH
B
180°
T415R
R730W
T638M
R784Q
Y162C
C
Orc4 - Orc1
TH
Y162
WB
RF
S1
R197
WA
R784
S2
ATP
Orc1 - Cdc6
K730
TH
S1
WB
RF
WA
R734
ATP
Orc4 - Orc1
ISM
T638
Cdc6 - Orc1
T415
S1

Figure 1



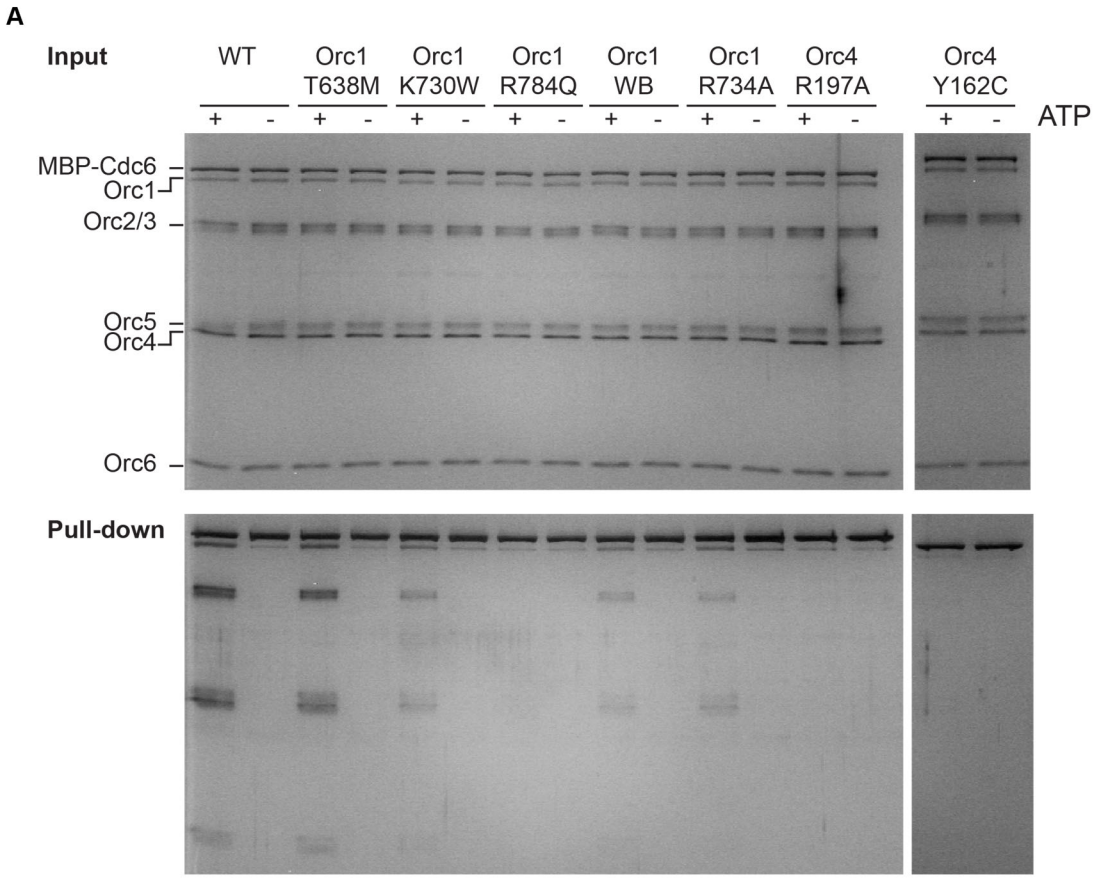


ORC:	K_{cat} (min^{-1})	n
wild type	8.6 ± 0.06	3
Orc1 T638M	15.8 ± 0.18	3
Orc1 K730W	4.7 ± 0.04	3
Orc1 R784Q	0.15 ± 0.004	3
Orc4 Y162C	0.27 ± 0.009	3
Orc1 WB	0.13 ± 0.005	3
Orc1 R734A	6.9 ± 0.06	3
Orc4 R197A	0.23 ± 0.01	1



ORC:	$K_{d, app}$ (nM)	n
wild type	13.5 ± 0.9	11
Orc1 T638M	104.7 ± 11.6	6
Orc1 K730W	8.3 ± 0.5	3
Orc1 R784Q	ND	6
Orc4 Y162C	ND	5
Orc1 WB	ND	6
Orc1 R734A	11.6 ± 1.4	6
Orc4 R197A	ND	6

Figure 2


A
Input
WT
Orc1 T638M
Orc1 K730W
Orc1 R784Q
Orc1 WB
Orc1 R734A
Orc4 R197A
Orc4 Y162C
ATP
MBP-Cdc6
Orc1
Orc2/3
Orc5
Orc4
Orc6
Pull-down

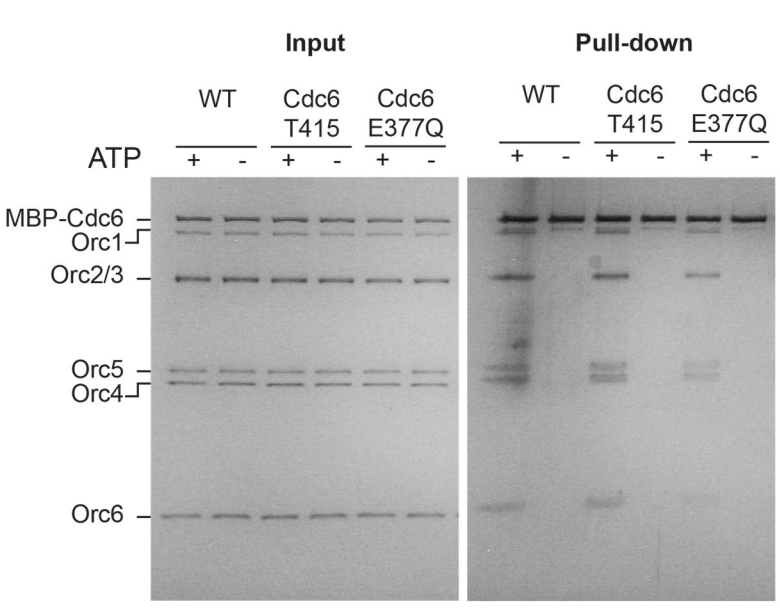

B
Input
Pull-down
WT
Cdc6 T415
Cdc6 E377Q
ATP
MBP-Cdc6
Orc1
Orc2/3
Orc5
Orc4
Orc6

Figure 3