PNAS
www.pnas.org

Supplementary Information for

Conformational control and DNA binding mechanism of the metazoan origin recognition complex

Franziska Bleichert[1*], Alexander Leitner[2], Ruedi Aebersold[2,3], Michael R. Botchan[4*], James M. Berger[5*]

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

[2]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland

[3]Faculty of Science, University of Zurich, 8057 Zurich, Switzerland

[4]Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA 94720, USA

[5]Department of Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

www.pnas.org/cgi/doi/10.1073/pnas.1806315115

*Corresponding authors: Franziska Bleichert, franziska.bleichert@fmi.ch

James M. Berger, jmberger@jhmi.edu

Michael R. Botchan, mbotchan@berkeley.edu

This PDF file includes:

SI Materials and Methods

Figs. S1 to S6

Legends for Dataset S1 and S2

SI Materials and Methods

Cloning, expression, and purification of ORC

Hexameric, full-length *Drosophila* ORC and the *Dm*ORC core complex were reconstituted in insect cells by co-infection with a MultiBac virus expressing subunits Orc1-5 and a single virus expressing Orc6. To create the MultiBac vector, full-length Orc1-5 or N-terminally truncated Orc1-3 (Orc1ΔN: amino acid (aa) residues 533-924, Orc2ΔN: aa 266-618, Orc3ΔN: aa 47-721), in combination with full-length Orc4 and Orc5, were cloned into a pFastBac-derived BioBricks MultiBac expression vector (Macrolab, University of California Berkeley, USA). Orc6 or Orc6ΔN (amino acid residues 187-257) were cloned into pFastBac or a ligation-independent-cloning (LIC)-compatible pFastBac vector. All other versions of *Dm*ORC lacking various N-terminal regions of certain subunits were obtained by co-infecting insect cells with a full-length Orc4/Orc5 MultiBac virus and individual viruses for the remaining subunits (cloned individually into pFastBac or a LIC-compatible pFastBac vector). The following regions of Orc1-3 and Orc6 were placed into such expression constructs: Orc1ΔN194 – aa 195-924, Orc1ΔN343 – aa 344-924, Orc1ΔN439 – aa 440-924, Orc1ΔN514 – aa 515-924, Orc1ΔN528 – aa 529-924, Orc1ΔN – aa 533-924, Orc2ΔN – aa 266-618, Orc3ΔN – aa 47-721, Orc6ΔN – aa 187-257. The Orc1 R492E/K523E/R528E point mutant was generated by site-directed mutagenesis of the pFastBac-Orc1 vector and verified by DNA sequencing. To aid purification, a hexa-histidine (6xHis) tag was added to the N-terminus of all Orc1 constructs and a maltose binding protein (MBP) tag to the N-terminus of Orc4, both followed by a tobacco etch virus (TEV)

protease cleavage site. All viruses were generated in Sf9 cells by transfecting bacmid DNA (obtained from DH10Bac cells) with Cellfectin II (Thermo Scientific Fisher) according to the manufacturer's instruction. Baculoviruses were amplified for two rounds to yield high-titer virus stocks for large-scale protein expression.

*Dm*ORC purification was performed as described previously with minor modifications (1, 2). Four liters of High Five cells (Thermo Scientific Fisher) in spinner flasks were co-infected for 2 days with a combination of baculoviruses to express wild-type or mutant full-length or "trimmed" *Dm*ORC assemblies. Subsequently, cells were harvested by centrifugation, resuspended in 100-150 mL 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 50 mM imidazole (pH 7.8), 10% glycerol, 200 µM PMSF, 1 µg/L leupeptin, 1 mM BME, and lysed by sonication. The cell lysate was clarified by centrifugation and a 20% ammonium sulfate precipitation for 30 min on ice, followed by another centrifugation step. The supernatant was passed over a 5 mL HisTrap HP Nickel-affinity chromatography column (GE Healthcare), and after extensive washing, ORC was eluted using a 50 – 250 mM imidazole gradient in 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 50 mM imidazole (pH 7.8), 10% glycerol, 1 mM BME. Fractions containing ORC were pooled and further purified by amylose-affinity chromatography using a 10 mL column (New England Biolabs). 6xHis and MBP tags were cleaved off by digestion using 6xHis-tagged TEV protease overnight. A subsequent nickel-affinity chromatography step was used to remove TEV and any uncleaved His-tagged Orc1-containing *Dm*ORC. The flow-through was concentrated in 30K

Amicon Ultra-15 concentrators (Millipore) and further purified by gel filtration chromatography using a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 25 mM HEPES (pH 7.6), 500 mM potassium glutamate, 10% glycerol, and 1 mM DTT. Peak fractions were pooled, concentrated using 30K Amicon Ultra-15 concentrators (Millipore), and flash frozen in liquid nitrogen for storage.

Human ORC (subunits Orc1-5) was co-expressed in High Five cells using individual baculoviruses encoding a single subunit each. The human complex was purified by Nickel-affinity and size-exclusion chromatography similarly to the *Dm*ORC complexes, with the exception that gel filtration chromatography was performed in 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 10% glycerol, 1 mM DTT. Human Orc6 only weakly binds to human ORC1-5 and does not co-purify with the other subunits (1), hence it was omitted from the expression and purification procedure.

Cloning, expression, and purification of *Drosophila* Cdc6

Full-length *Drosophila* Cdc6 was cloned into a LIC-modified pFastBac vector with an N-terminal MBP tag and a TEV protease site. Bacmids were generated in DH10Bac cells, purified according to the Bac-to-Bac protocol (Thermo Scientific Fisher), and used for baculovirus generation by transfecting bacmid DNA into Sf9 cells using Cellfectin II (Thermo Scientific Fisher). After two rounds of amplification in Sf9 cells, the virus was used to infect 4 L of cultured

High Five cells. 48 h post-infection, cells were harvested and resuspended in 100 mL lysis buffer (50 mM Tris-HCl (pH 7.8), 300 mM KCl, 10% glycerol, 200 μM PMSF, 1 μg/L leupeptin). Cells were lysed by sonication, followed by a centrifugation at 38,724 x g for 1 h at 4°C, a 30 min precipitation with 20% ammonium sulfate, and an additional centrifugation step at 38,724 x g for 1 h at 4°C. The clarified lysate was passed over a 25 mL amylose column (NEB), which was then washed with 100 mL lysis buffer. Cdc6 was eluted with 20 mM maltose in 50 mM Tris-HCl pH 7.8, 300 mM KCl, 10% glycerol. Cdc6 containing fractions were concentrated in a Centriprep concentrator (molecular weight cut-off 30K) to 2 mL and then further purified by gel filtration chromatography using a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 10% glycerol. Peak fractions were pooled, concentrated in a 30K Centriprep concentrator, and flash frozen in liquid nitrogen for storage.

Fluorescence anisotropy DNA binding assays

A fluorescence anisotropy-based DNA binding assay was developed to measure the dissociation constants of ORC for fluorescein-labeled DNA substrates. Increasing concentrations of ORC (38 pM to 2.5 μM) were incubated with 1 nM fluorescein-labeled duplex DNA in 25 mM HEPES-KOH (pH 7.6), 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT, 0.01% NP40, 0.1 mg/mL BSA for 30 min at 23°C. Twenty μL of each reaction were transferred into a 384-well plate and fluorescence polarization measured in

a Clariostar or a Pherastar FSX plate reader (BMG) using excitation and emission filters for fluorescein. Polarization values were converted to anisotropy and the averages and standard deviation of at least three independent experiments were plotted as a function of ORC concentration. Data points were fitted to the Hill equation to obtain dissociation constants. All binding reactions were performed using a 40 bp DNA duplex (**Table 1**) derived from the yeast ARS1 sequence in the presence of 1 mM ATP unless otherwise indicated. To investigate the DNA length-dependence of ORC-DNA interactions, duplex lengths were varied from 20 bp to 84 bp (**Table 1**). Double-stranded DNA fragments were generated by annealing 5'-fluorescein labeled oligonucleotides (obtained from IDT) with complementary, unlabeled oligonucleotides in 20 mM Tris-HCl (pH 8.0) and 10 mM $MgCl_2$.

Competition DNA-binding experiments were carried out in 20 μL reactions containing 100 nM full-length *Dm*ORC, 1 nM fluorescein-labeled 40 bp duplex DNA, and 300 pM to 2.5 μM competitor DNA in 25 mM HEPES-KOH (pH 7.6), 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT, 1 mM ATP, 0.01% NP40, and 0.2 mg/mL BSA. *Dm*ORC and fluorescein-labeled DNA were first mixed together at 200 nM and 2 nM concentration, respectively, and incubated 20 min at 23°C. Reactions were then diluted 2-fold by adding competitor DNAs. An unlabeled 40 bp DNA duplex, as well as supercoiled and EcoRV-linearized pRS313 plasmid DNA were used as competitors. Fluorescence polarization was measured in 384-well plates in a

Pherastar FSX plate reader (BMG), converted to anisotropy, and the averages and standard deviations of three or more experiments were plotted.

ORC•DNA and ORC•DNA•Cdc6 pull-down assays

To confirm ORC-DNA interactions using an orthogonal approach to fluorescence anisotropy, we performed pull-down assays using an 84 bp ARS1-derived DNA duplex (see **Table 1**) labeled with a photocleavable 5'-biotin (IDT) as a bait. For each reaction, 100 nM of DNA and 10 μL of streptavidin-coupled Dynabeads (Thermo Fisher Scientific) were mixed in 100 μL binding buffer (25 mM HEPES-KOH (pH 7.6), 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT, 0.01% NP40) and incubated for 30 min at 23°C. Subsequently, beads were washed twice with binding buffer and incubated with 100 nM ORC in a total reaction volume of 100 μL, either in the absence or presence of 1 mM ATP. After a 30 min incubation at 23°C, beads were washed twice with 1 mL of binding buffer, and ORC was eluted by UV cleavage. Input and eluted proteins were separated on a 10% SDS-PAGE gel and visualized by silver staining.

To investigate the ability of different ORC assemblies to associate with Cdc6, we performed pull-down assays using MBP-tagged *Dm*Cdc6 as bait. Unless otherwise noted, 100 nM full-length wild-type *Dm*ORC or ORC assemblies lacking specific N-terminal regions (see main text) were incubated with 100 nM 84 bp duplex DNA for 15 min at 23°C. MBP-tagged *Dm*Cdc6 was

added to a final concentration of 200 nM, followed by another 15 min incubation at room temperature. Binding reactions were performed in 25 mM HEPES-KOH (pH 7.5), 300 mM potassium glutamate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT, 0.01% NP40 in a total volume of 200 μL, either in the absence or presence of 1 mM nucleotide. The entire reaction was then added to 25 μL amylose beads (NEB) for 15 min, and beads were subsequently washed twice with 1 mL binding buffer. Bound proteins were eluted with 20 mM maltose, and both input and eluted proteins were separated by SDS-PAGE electrophoresis and visualized by silver staining.

Chemical crosslinking and mass spectrometry

Full-length wild-type *Dm*ORC or the *Dm*ORC core assembly were first dialyzed overnight into 25 mM PIPES buffer (pH 7.5), 300 mM KCl, 10% glycerol, 1 mM DTT to remove primary amines from the ORC storage buffer prior to crosslinking. Crosslinking reactions were performed with disuccinimidyl suberate (DSS) at a final concentration of 100 μM and 1 mg/mL of the respective *Dm*ORC assembly. After incubation for 45 min at 25 °C, the crosslinking reaction was quenched by adding 1 M ammonium bicarbonate solution to a final concentration of 50 mM. After reduction and alkylation of cysteines, proteins were digested with endoproteinase Lys-C (Wako) and trypsin protease (Promega). The digested sample was fractionated by size exclusion chromatography using a Superdex Peptide PC 3.2/300 column (GE Healthcare) (3), and fractions were analyzed by LC-MS/MS on an Orbitrap Fusion Lumos mass spectrometer, as described

previously (4). The resulting data were searched against the database containing sequences for full-length or “trimmed” *Dm*ORC subunits with xQuest (3, 5). Score threshold were set to achieve a false discovery rate of <5%. Inter- and intra-protein crosslinks were visualized as a network plot with xVis (6) and mapped onto the *Drosophila* ORC crystal structure or a model of an active *Dm*ORC complex with a closed Orc1/4 ATPase site with PyMOL (The PyMOL Molecular Graphics System, Version 1.8.2.0 Schrödinger, LLC). The active *Dm*Orc1/4 ATPase site was modeled by superpositioning the *Dm*Orc1 and *Dm*Orc4 AAA+ domains onto the DmOrc4 and DmOrc5 AAA+ domains, respectively, as described in (2). To differentiate crosslinks formed between the activated and autoinhibited conformation, only crosslinks between Orc1, Orc4, and Orc3 were considered but not crosslinks involving the winged-helix domain of Orc2, since the position and orientation of this domain is unknown in the active ORC conformation.

Negative-stain electron microscopy and image processing

For negative-stain EM, full-length *Drosophila* ORC, *Dm*ORC lacking the N-terminal 532 amino acid residues in Orc1, or human ORC1-5 were diluted to a concentration of 80 nM using a buffer containing 25 mM HEPES-KOH (pH 7.6), 300 mM potassium glutamate, 10 mM magnesium acetate, 5% glycerol, 1 mM DTT, and 1 mM ATP or ATPγS. 4 μL of diluted protein samples were spotted onto glow-discharged, continuous-carbon film EM grids and stained with four drops of 2% uranyl formate. Grids were imaged either in a Tecnai T12 TWIN

transmission electron microscope operated at 100 keV, or a Tecnai G2 Spirit BIOTWIN electron microscope operated at 120 kV, both equipped with a LaB_6 cathode as electron source and an FEI Eagle CCD camera. Approximately 150-250 micrographs were automatically acquired using the EPU software package at a nominal magnification of 52,000 or 49,000, respectively. Micrographs were visually inspected and those with overly thin or thick staining discarded from further processing. Contrast transfer function parameters were estimated with CTFFIND4 (7) and corrected by phase flipping using SPIDER (8). Automatic particle picking was performed with DOGPICKER (9) in a reference-free manner. Particle images were normalized with XMIPP (10, 11) and subjected to 2D classification using the RELION software package (12-14). The final image stacks contained 70,000 to 120,000 particles.

Cryo-electron microscopy and image processing

The ORC•DNA•Cdc6 complex was assembled by mixing full-length *Dm*ORC and a 84 bp DNA duplex (**Table 1**) at a final concentration of 80 and 100 nM, respectively, in 25 mM HEPES pH 7.6, 300 mM potassium glutamate, 10 mM magnesium acetate, 1% glycerol, 1 mM DTT, 1 mM ATP, 0.02% NP40. After a 25 minute-incubation at room temperature, full-length *Dm*Cdc6 was added to a final concentration of 100 nM. 4 μL of sample were applied for 22 s at 22°C and 100% humidity to a glow-discharged 400-mesh C-flat 1.2/1.3 EM grid coated with a thin film of continuous carbon in a Vitrobot, blotted for 5s, and plunge-frozen in liquid ethane. Cryo-EM grids were imaged using a Titan Krios

electron microscope (at Janelia Research Campus) operated at 300 kV and equipped with a Cs corrector and an energy filter (slit width 20 eV). Dose-fractionated images were recorded automatically using Serial EM (15) on a post-GIF Gatan K2 Summit direct electron detector in super-resolution counting mode with a pixel size of 0.675 Å (1.35 Å after 2x binning) and a dose rate on the specimen of 5.49 electrons per $Å^2$ per second. 50 frames of 0.3 s each were recorded per movie, resulting in a total dose of ~82 electrons per $Å^2$. In total, 3859 movies were collected using a defocus range of 2-5 μm.

For processing, recorded movie frames were down-sampled by a factor of 2 to a pixel size of 1.35Å, whole-frame motion-corrected using UNBLUR, and frames were summed with SUMMOVIE (16). CTF parameters were then determined using CTFFIND4 (7). An initial set of particles was selected automatically from a subset of summed movies in a reference-free manner using GAUTOMATCH (http://www.mrc-lmb.cam.ac.uk/kzhang/) and subjected to 2D classification with RELION 2.0 (14). Several class averages resembling the ORC•DNA•Cdc6 complex were selected, low-pass filtered, and used as templates for particle autopicking from all summed movies using GAUTOMATCH. This process yielded a total of ~850 K particles, which were used for 2D classification in RELION 2.0. Class averages showed that the sample contained a mixture of autoinhibited and active *Dm*ORC, as well as the *Dm*ORC•DNA•Cdc6 complex (**Figures 5C** and **S6B**). Unfortunately, the strong

preferred orientation of the particles on the thin carbon layer prevented the reconstruction of any reliable 3D volume.

Multiple sequence alignment

Orc1 protein sequences of 210 different metazoan species were aligned using MAFFT (17, 18). The sequence logo used to represent the basic patch identified in the N-terminus of Orc1 was generated with the WebLogo server (19).

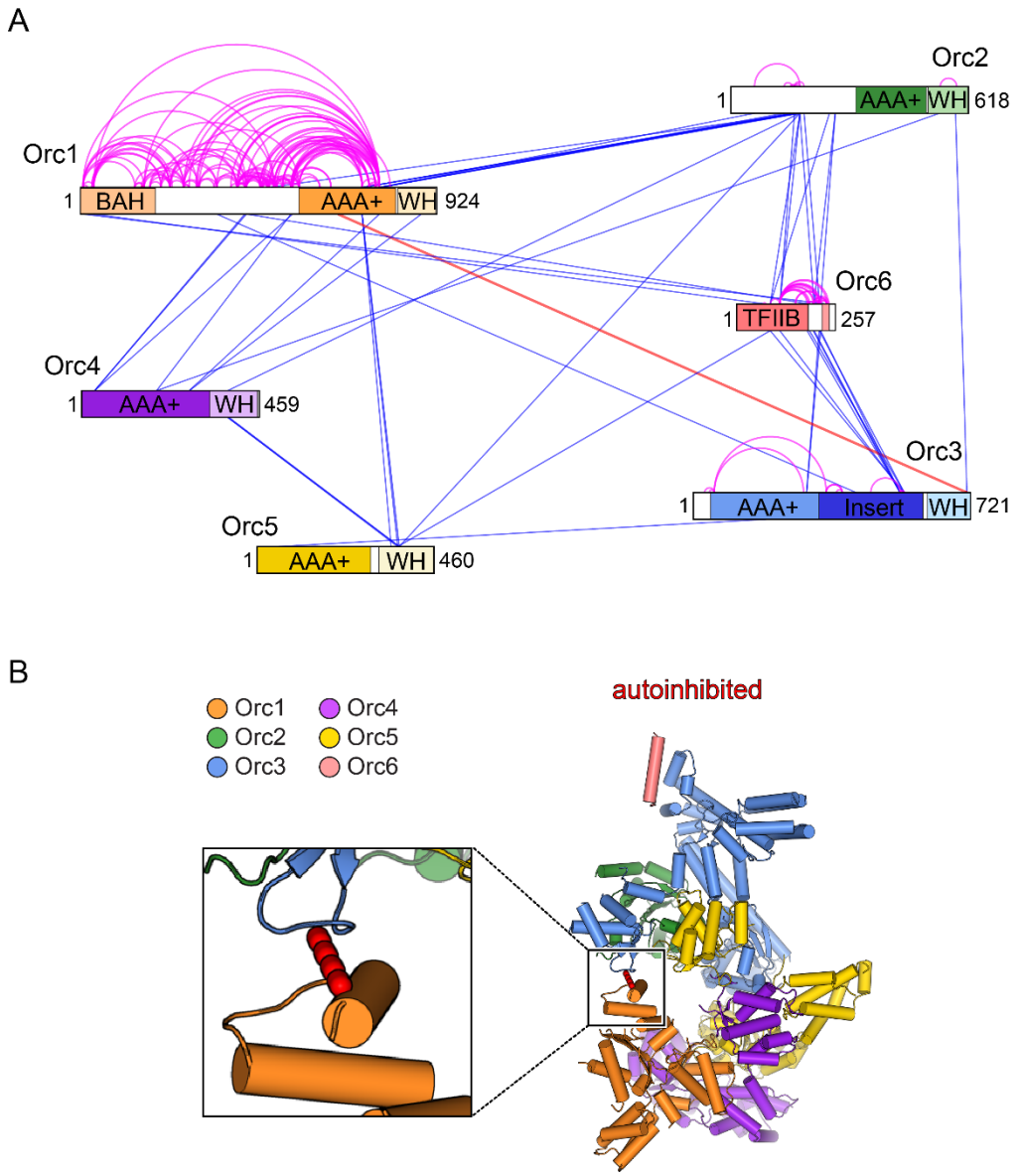

Figure S1. Crosslinking mass spectrometry analysis of full-length *Dm*ORC. **A**) Network plot of inter-protein (blue lines) and intra-protein (magenta lines) crosslinks. The individual domains in ORC subunits are indicated. WH – winged helix, BAH – bromo-adjacent homology. The presence of the largely unstructured Orc1 N-terminal domain acts as a sink for the vast majority of crosslinking interactions; as a consequence, only one crosslink was identified that matches the autoinhibited conformation exclusively (red

dashed line; no unique crosslink corresponding to the active state was detected). The position of the first and last amino acid of ORC subunits are indicated. **B**) The crosslink consistent with the autoinhibited conformation was mapped onto the crystal structure of *Drosophila* ORC (2).

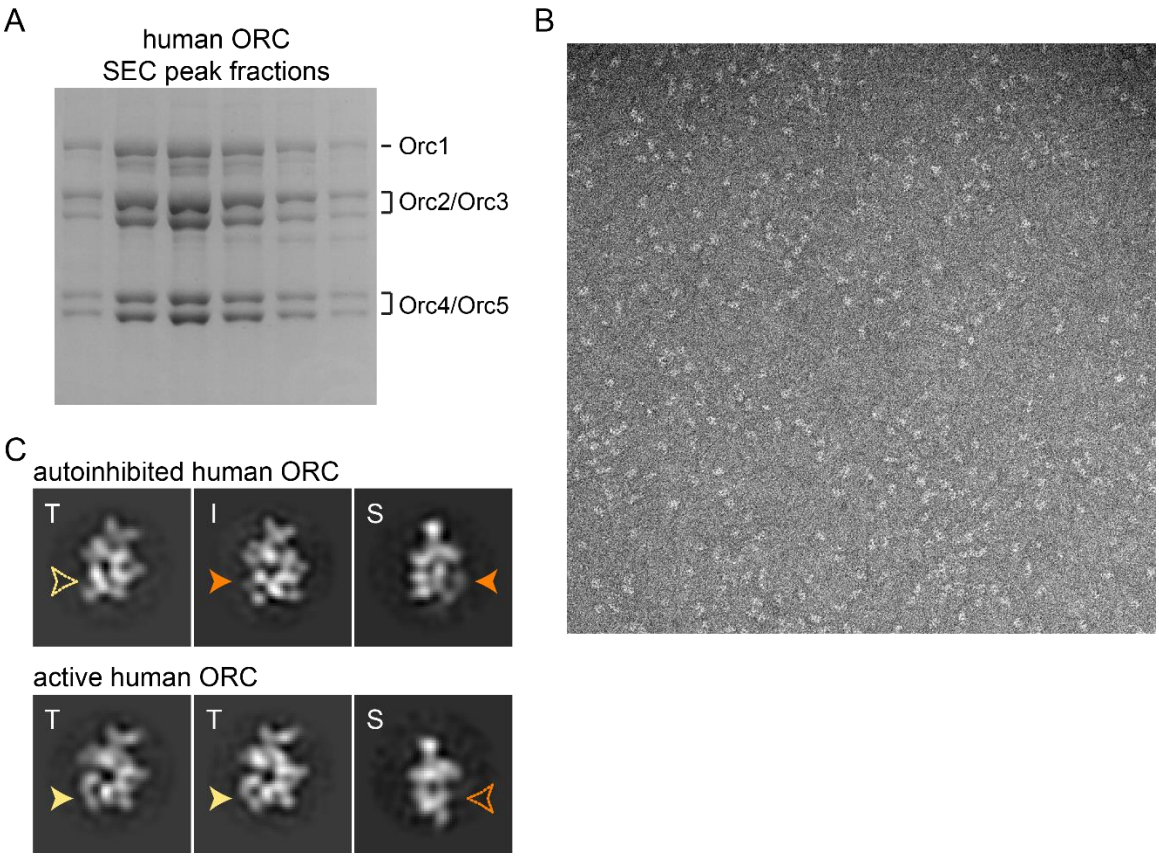

Figure S2. Purification and EM analysis of human ORC1-5. **A**) Coomassie-stained SDS-PAGE gel of peak human ORC1-5 fractions eluted from the final size-exclusion column. **B**) Electron micrograph of negatively-stained human ORC1-5 in the presence of ATPγS. **C**) Additional class averages (T – top view, I – intermediate view, and S – side view) of human ORC1-5 are shown, both in the autoinhibited and active configurations. Full and dotted arrowheads point towards the Orc1 density that is repositioned between both states. Overall, the class averages of human ORC in both conformations are similar to those of *Dm*ORC, albeit more particles are observed in the active state than with *Drosophila* ORC in the presence of ATPγS.

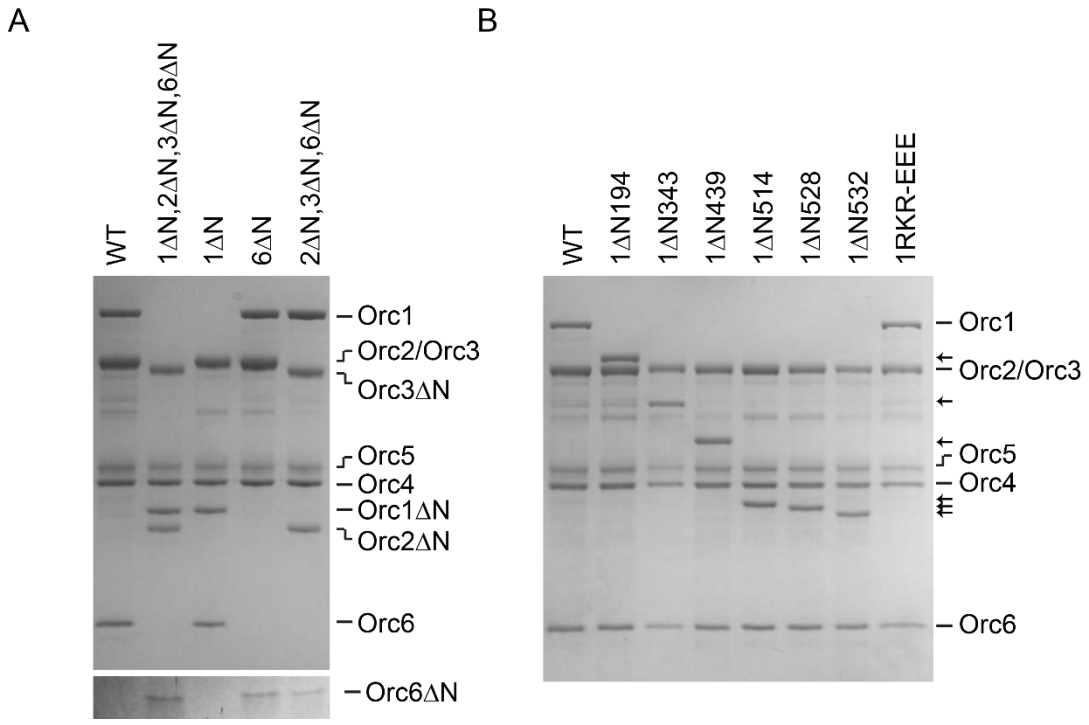

Figure S3. Purification of different *Dm*ORC complexes. **A**) Wild-type full-length *Dm*ORC and several *Dm*ORC assemblies lacking N-terminal regions of single (Orc1ΔN, Orc6ΔN) or multiple subunits (Orc1ΔN+Orc2ΔN+Orc3ΔN+Orc6ΔN, Orc2ΔN+Orc3ΔN+Orc6ΔN) were purified. A Coomassie-stained SDS-PAGE gel of purified complexes is shown. **B**) Purification of *Dm*ORC assemblies with different N-terminal truncations of the Orc1 subunit or point mutations in the Orc1 N-terminus (Orc1$^{R492E/K523E/R528E}$: abbreviated as 1RKR-EEE) were purified and analyzed by SDS-PAGE electrophoresis and Coomassie staining. Wild-type, full-length *Dm*ORC is shown as a reference. Arrows mark the bands of Orc1 subunits containing different N-terminal deletions.

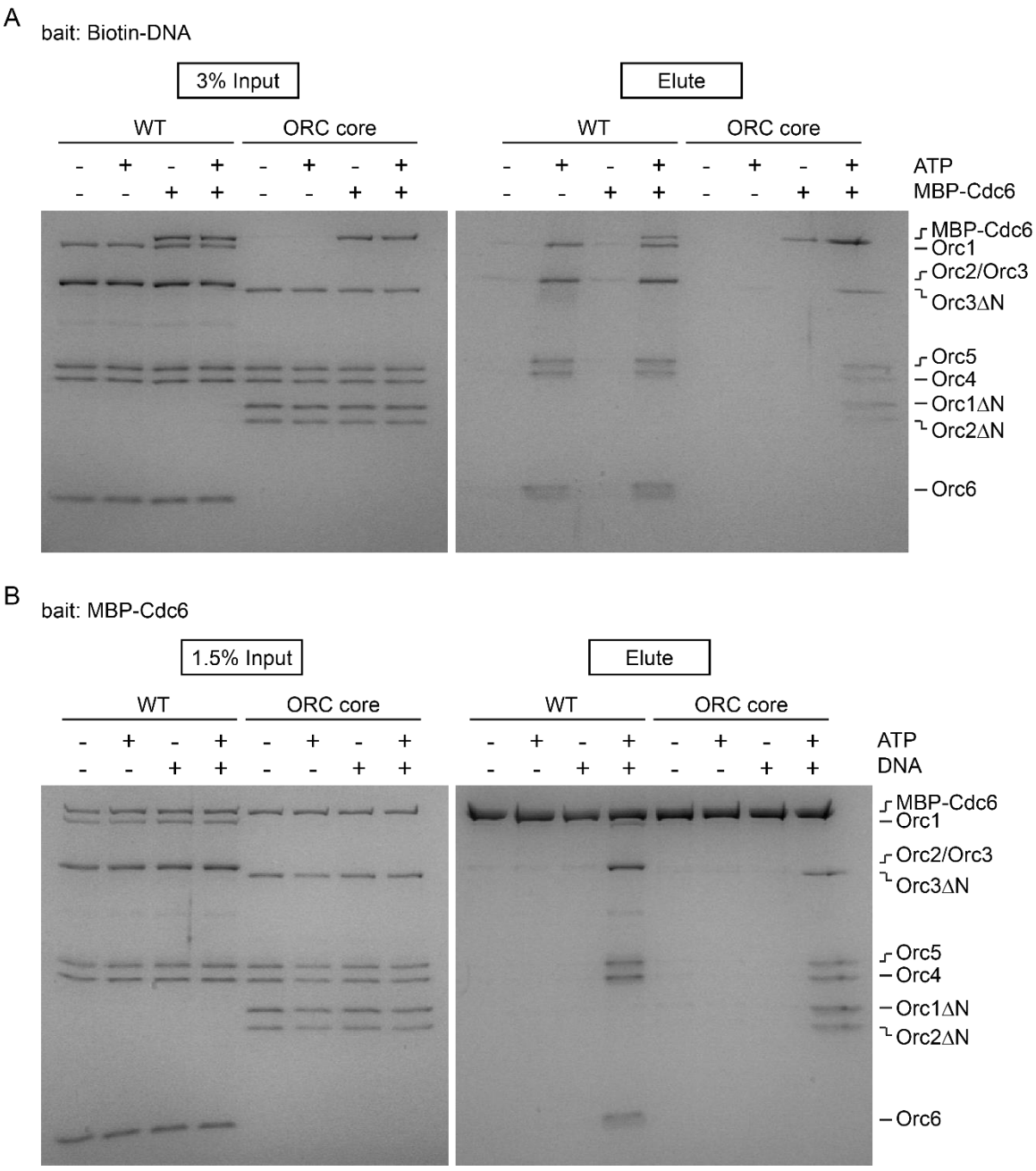

Figure S4. *Drosophila* Cdc6 stabilizes *Dm*ORC on DNA and ternary complex formation requires both DNA and ATP. **A**) The ability of both full-length, wild type *Dm*ORC and the *Dm*ORC core complex (lacking the N-terminal regions in Orc1, Orc2, Orc3, and Orc6) to associate with DNA in the presence or absence of *Dm*Cdc6 was compared in pull-down

assays using biotinylated DNA as bait. **B**) Wild-type full-length *Dm*ORC or the truncated *Dm*ORC core complex were mixed with MBP-tagged *Dm*Cdc6 either in the absence or presence of DNA and ATP, respectively, and subjected to pull-downs using amylose beads. For **A**) and **B**), input and eluted proteins were separated by SDS-PAGE and visualized by silver staining. Note that the Orc6 C-terminal peptide (Orc6ΔN) is not resolved and visible on the gels shown due to its small size.

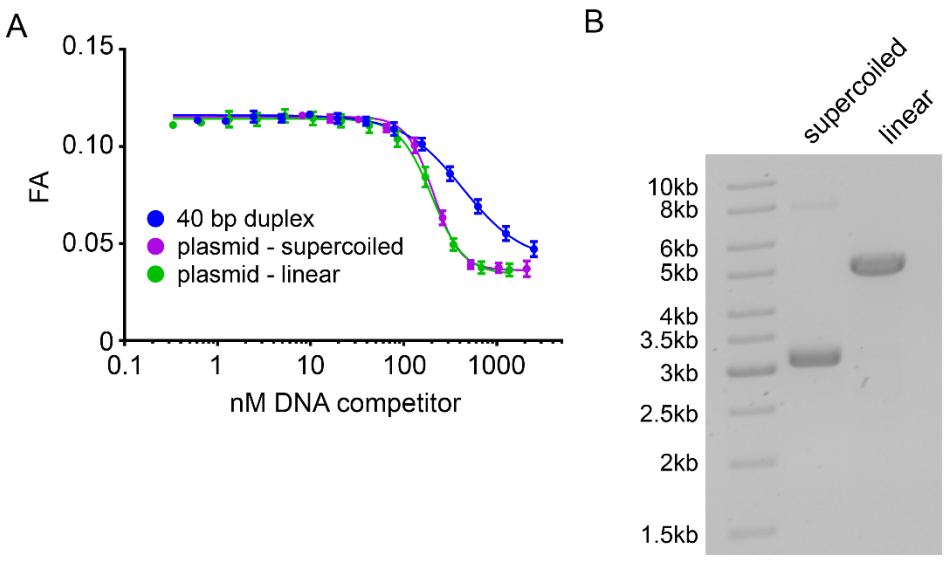

Figure S5. ORC has a higher affinity for linearized and supercoiled plasmid DNA than for short DNA duplexes. **A**) Fluorescence polarization was used to monitor DNA binding in competition experiments. Increasing amounts of 40 bp duplex DNA, linearized pRS313 plasmid DNA, or supercoiled pRS313 were titrated as competitors in reactions containing constant amounts of *Dm*ORC and fluorescently labeled 40 bp duplex DNA. All experiments were performed in the presence of 1 mM ATP. The concentration of plasmid DNA is corrected for length and expressed as a concentration of 40 bp DNA fragments. **B**) Agarose gel electrophoresis of linearized and supercoiled plasmid DNA.

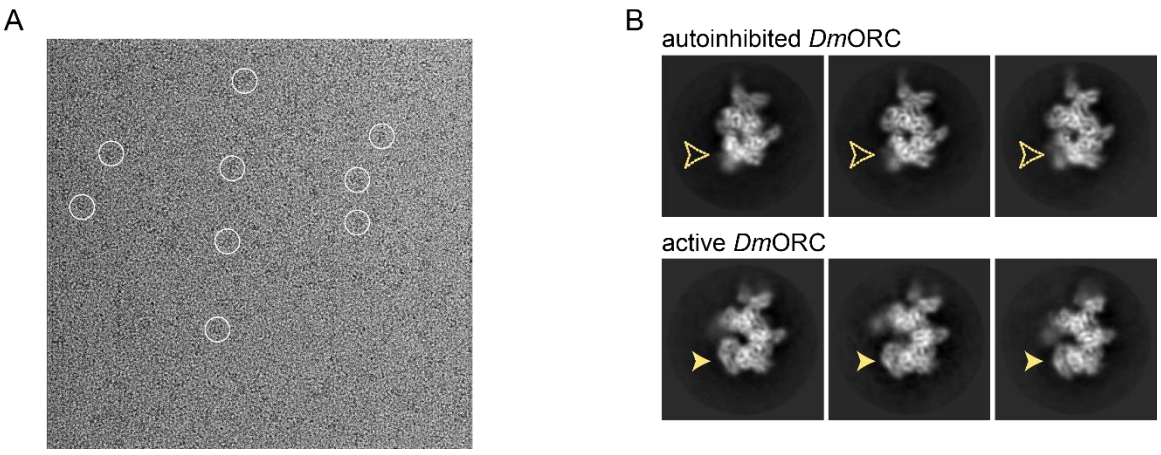

Figure S6. Cryo-EM analysis of a *Drosophila* ORC•DNA•Cdc6 complex. **A**) Example cryo-EM micrograph with some particles highlighted by white circles. **B**) 2D cryo-EM class averages of autoinhibited and activated *Dm*ORC present in the *Drosophila* ORC•DNA•Cdc6 sample are shown. Arrowheads demarcate the Orc1-AAA+ density, which is repositioned between the two structural configurations. Class averages of the ternary complex are shown in **Figure 5C**.

Dataset S1. List of crosslinked peptides observed with the *Dm*ORC core complex.

Dataset S2. List of crosslinked peptides observed with full-length *Dm*ORC.

Supporting References

1. Bleichert F*, et al.* (2013) A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *eLlife* **2**:e00882.
2. Bleichert F, Botchan MR, & Berger JM (2015) Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**(7543):321-326.
3. Leitner A, Walzthoeni T, & Aebersold R (2014) Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc* **9**(1):120-137.
4. Eliseev B*, et al.* (2018) Structure of a human cap-dependent 48S translation pre-initiation complex. *Nucleic Acids Res* **46**(5):2678-2689.
5. Walzthoeni T*, et al.* (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Methods* **9**(9):901-903.
6. Grimm M, Zimniak T, Kahraman A, & Herzog F (2015) xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. *Nucleic Acids Res* **43**(W1):W362-369.
7. Rohou A & Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**(2):216-221.
8. Frank J*, et al.* (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**(1):190-199.
9. Voss NR, Yoshioka CK, Radermacher M, Potter CS, & Carragher B (2009) DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J Struct Biol* **166**(2):205-213.
10. Scheres SH, Nunez-Ramirez R, Sorzano CO, Carazo JM, & Marabini R (2008) Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc* **3**(6):977-990.
11. Sorzano CO*, et al.* (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148**(2):194-204.

12. Scheres SH (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**(3):519-530.
13. Scheres SH (2016) Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods Enzymol* **579**:125-157.
14. Kimanius D, Forsberg BO, Scheres SH, & Lindahl E (2016) Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**.
15. Mastronarde DN (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**(1):36-51.
16. Grant T & Grigorieff N (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6. *eLife* **4**:e06980.
17. Katoh K, Kuma K, Toh H, & Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**(2):511-518.
18. Katoh K & Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* **9**(4):286-298.
19. Crooks GE, Hon G, Chandonia JM, & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* **14**(6):1188-1190.