ARTICLE

https://doi.org/10.1038/s41586-018-0293-x

Structure of the origin recognition complex bound to DNA replication origin

Ningning Li[1,7], Wai Hei Lam[2,7], Yuanliang Zhai[2,3,6,7]*, Jiaxuan Cheng[4,7], Erchao Cheng[4], Yongqian Zhao[2,3], Ning Gao[1]* & Bik-Kwoon Tye[2,5]*

The six-subunit origin recognition complex (ORC) binds to DNA to mark the site for the initiation of replication in eukaryotes. Here we report a 3 Å cryo-electron microscopy structure of the *Saccharomyces cerevisiae* ORC bound to a 72-base-pair origin DNA sequence that contains the ARS consensus sequence (ACS) and the B1 element. The ORC encircles DNA through extensive interactions with both phosphate backbone and bases, and bends DNA at the ACS and B1 sites. Specific recognition of thymine residues in the ACS is carried out by a conserved basic amino acid motif of Orc1 in the minor groove, and by a species-specific helical insertion motif of Orc4 in the major groove. Moreover, similar insertions into major and minor grooves are also embedded in the B1 site by basic patch motifs from Orc2 and Orc5, respectively, to contact bases and to bend DNA. This work pinpoints a conserved role of ORC in modulating DNA structure to facilitate origin selection and helicase loading in eukaryotes.

Initiation of DNA replication begins with the binding of an initiator at a replicator followed by the recruitment of a replicative helicase that unwinds DNA[1,2]. Bacterial and archaeal initiators contain an AAA+ domain and a helix–turn–helix (HTH) motif that are implicated in specific DNA interactions[3–7]. In eukaryotes, the initiator is a highly conserved six-subunit origin recognition complex (ORC)[8–10]. The five subunits Orc1–Orc5 each bear an AAA+ domain and a winged-helix domain (WHD)[1,9], whereas Orc6 bears little resemblance to the other ORC members and its role in the assembly of ORC around origin DNA seems to differ among organisms[9,10]. Despite the high conservation of ORC in protein sequence among eukaryotes, its selectivity for replicators varies from species to species, in which specific DNA sequence has a determinant role in certain yeasts[11,12] and chromatin structure has a predominant role in humans[9,13]. Both of these modes are used in all eukaryotes but the prevalence of these selectivity modes is influenced by the interplay of several factors[14–23].

To understand how the ORC selects and binds DNA, much effort has been devoted to the study of the ORC architecture either alone or with Cdc6, which ORC recruits to origin DNA in G1 phase as a prelude to the assembly of the CMG (Cdc45–MCM–GINS) helicase[24–29]. More recently, several high-resolution structures that include the ORC, such as the OCCM (ORC–Cdc6–Cdt1–Mcm2–Mcm7)[30], or that contain trimmed portions of ORC in fly[31] and human[32] have been determined. Together, these structures provide a clear picture of how ORC subunits interact with each other. Unfortunately, because DNA was either omitted or the resolution does not provide sufficient detail, we can only speculate on how the ORC may interact with DNA to perform its functions.

In this study, we have determined a series of structures of *S. cerevisiae* ORC, either alone or with ARS305 DNA (36 or 72 base pairs (bp) in length), using full-length proteins by single-particle cryo-electron microscopy (cryo-EM) (Extended Data Figs. 1, 2, Extended Data Table 1). The first ORC–DNA complex (36 bp, containing both ACS and B1 elements) was obtained at 3.6 Å, which allows a nearly complete identification of base sequence from end to end (Fig. 1a). However, the interaction between the ORC and the B1 element is relatively unstable. Therefore, we extended the original DNA to 72 bp from the B1 end and determined a second structure of the ORC–DNA complex at 3.0 Å (Fig. 1b–d). These two structures are basically the same except that atomic interactions at the B1 proximal site in the second one are clearly resolved.

Subunit organization around the ACS DNA

In the ORC–DNA (72 bp) structure, over 50 bp of the 72-bp duplex, from one end to beyond the B1 element, can be visualized in the density map (Fig. 1b and Supplementary Video 1). The DNA beyond that point appears highly flexible. There is no melting of the duplex DNA, but a continuous curving of the backbone is clearly seen with two major bending points (Fig. 1d). One resides at position A8, right in the middle of ACS, and the other is in the middle of the B1 element at position T28.

Similar to previous structures[30–32], Orc1–Orc5 subunits oligomerise to encircle the ACS DNA through both canonical interactions between inter AAA+ domains (Fig. 1a–c and Extended Data Fig. 3a) and interdigitated domain-swapping interactions between the WHDs and the AAA+ domains of adjacent ORC subunits (Extended Data Fig. 3d). In addition to these conventional interactions, many flexible extensions and linkers of ORC subunits also contribute largely to the inter-subunit stabilization (Supplementary Video 1, Extended Data Fig. 3e–g). ORC recognizes and binds to origin DNA in an ATP-dependent manner[8]. Orc1–Orc5 subunits each contain AAA+ or AAA+-like domains, but only Orc1, Orc4 and Orc5 contain functional ATPase-related motifs[9,33]. As expected, we identified three ATPγS molecules and their associated Mg^{2+} at three corresponding interfaces (Extended Data Fig. 4). However, each of the bound ATPγS molecules is uniquely coordinated, probably reflecting distinct roles for these centres during replication initiation (Supplementary Discussion).

Two regions of Orc6, the transcription factor II B (TFIIB) domain B (residues 271–386) and the carboxy-terminal domain (CTD, residues 387–430), were solved in the structure (Extended Data Fig. 1i). Orc6 interacts with three subunits, Orc3, Orc2 and Orc5, and their interfaces

[1]State Key Laboratory of Membrane Biology, Peking-Tsinghua Center for Life Sciences, School of Life Sciences, Peking University, Beijing, China. [2]Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, China. [3]Institute of Advanced Study, The Hong Kong University of Science and Technology, Hong Kong, China. [4]Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China. [5]Department of Molecular Biology & Genetics, Cornell University, Ithaca, NY, USA. [6]Present address: School of Biological Sciences, The University of Hong Kong, Hong Kong, China. [7]These authors contributed equally: Ningning Li, Wai Hei Lam, Yuanliang Zhai, Jiaxuan Cheng. *e-mail: zhai@hku.hk; gaon@pku.edu.cn; biktye@ust.hk

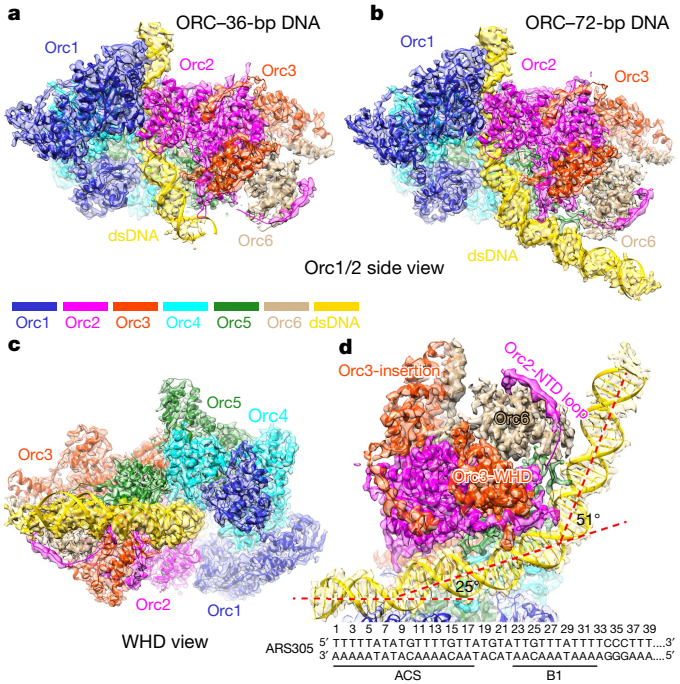


Fig. 1 | Overall structure of ORC bound to origin DNA. a, Side view of the cryo-EM density map of the ORC–DNA complex (36 bp) with the atomic model superimposed. dsDNA, double-stranded DNA. **b**, **c**, Side (Orc1/Orc2) (**b**) and bottom (WHD) (**c**) views of the ORC–DNA complex (72 bp). **d**, Cut-away view of the ORC–72-bp DNA complex, highlighting the interaction of Orc6 with Orc2 and Orc3. ARS305 sequence is shown, with successive bending of the DNA illustrated by the dashed red lines. Other Orc subunits are omitted for clarity. NTD, N-terminal domain.

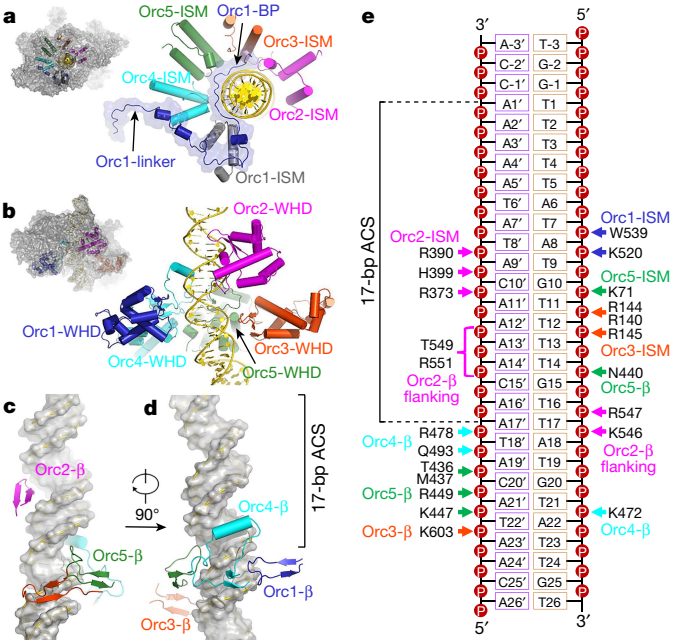


Fig. 2 | Extensive interactions between ORC and DNA around the ACS region. a, Distribution of the ISMs of ORC subunits around origin DNA. The long N-terminal loop upstream of Orc1-AAA+ is highlighted in blue. Orc1-BP is inserted between the ISMs of Orc1, Orc4, Orc5, Orc3 and DNA. Orientations of the ISMs in the ORC–DNA structure are shown in top-left thumbnail. **b**, Same as in **a**, but for the five WHDs of the ORC subunits. **c**, **d**, Distribution of the β-hairpin motifs from the WHDs of ORC subunits around the DNA (surface representation). **e**, Summary of non-specific ORC–DNA backbone interactions around the ACS region. The residues of ORC subunits, mostly basic, that contribute to DNA backbone phosphate interactions are labelled.

are dominated by hydrophobic residues (Extended Data Fig. 5a–d). Orc6 is situated distal to the ORC central channel and is not involved in the ACS DNA recognition but interacts with the B1 element through its TFIIB domain (Fig. 1d, Extended Data Fig. 5e).

The AAA+ modules of the Orc1–Orc5 subunits form a tilted and breached ring around DNA with a gap between Orc1 and Orc2 (Extended Data Fig. 3a). Notably, the Orc2 WHD is sandwiched between the AAA+ domains of Orc1 and Orc2, deviating from the WHD tier (Extended Data Fig. 3a, b). This is in contrast to the OCCM, in which the six WHDs occupy nearly symmetric positions[30]. 3D classification of the ORC–DNA particles showed that the Orc2 WHD is indeed flexible and takes up distinct positions in different subpopulations (Extended Data Fig. 6a–c). In fact, before DNA binding, as seen in our structure of ORC alone, both the entire Orc1-AAA+ and Orc2-WHD appear highly flexible and a very large opening can be observed between Orc1 and Orc2 (Extended Data Fig. 6h). This observation is consistent with the low-resolution electron microscopy structure of human ORC1–ORC5, in which a gap was also observed between ORC1 and ORC2[32]. Also consistently, the crystal structure of *Drosophila* ORC[31] in an auto-inhibitive state showed that Orc1-AAA+ and Orc2-WHD occupy completely different positions (Extended Data Fig. 7). The function of ORC in helicase loading requires Cdc6 to be inserted transiently into the gap between Orc1 and Orc2[30,34]. Therefore, these structures indicate that the interface between Orc1 and Orc2 is intrinsically dynamic, serving for both DNA entry and Cdc6 dock.

Non-specific gripping of origin DNA by ORC

The archaeal initiator Orc1 recognizes origin DNA using three major modules, an initiation-specific motif (ISM) from the AAA+ domain, a β-hairpin 'wing' and one HTH motif from the WHD[5,6]. These DNA recognition modules are largely preserved in each of the Orc1–Orc5 subunits in eukaryotes; however, their roles in contacting DNA have diverged. In the ORC–DNA structure, the five ISMs are arranged in a helical spiral around DNA, but Orc4-ISM has no direct contact with origin DNA (Fig. 2a, e). In sharp contrast to the archaeal ORC, all of the HTH motifs of Orc1–Orc5 are positioned away from DNA (Extended Data Fig. 3c). Three β-hairpin motifs of the WHDs (Fig. 2b) from Orc2, Orc4 and Orc5 contact the major groove in a helical manner, spanning nearly a full turn of the ACS DNA (Fig. 2c, d). Particularly, β-hairpins of Orc2 and Orc5 are both inserted into the major groove, with either the loop or flanking residues interacting with DNA backbone (Fig. 2c–e). The β-hairpin of Orc3, although not inserted into the major groove, is also in close contact with DNA phosphate backbone (Fig. 2c–e). Notably, some of the DNA contacts through these conserved modules of ORC are dynamic. As seen in the OCCM structure[30], ISMs of Orc1 and Orc5 and the β-hairpin of Orc2 are no longer in contact with DNA. Together, these features suggest that these conserved elements of ORC act combinatorially, probably through a universal mechanism, to provide a versatile grip for holding origin DNA with fairly high affinity but low sequence-specificity in different functional states of ORC.

Specific recognition of ACS by Orc1, Orc2 and Orc4

For specific origin DNA binding, eukaryotic ORC has acquired new elements to augment the cross-kingdom conserved DNA binding modules. Through structural analysis, we identified three critical motifs from Orc1, Orc2 and Orc4 that are important for sequence-specific recognition of ACS DNA. Notably, a basic patch of Orc1 (residues 358–371, Orc1-BP), which is located in a highly disordered region between the BAH and the AAA+ domain, is fully inserted into the minor groove of the ACS DNA for a half turn atop of all the other DNA-binding modules (Figs. 2a, 3a, b). This insertion is facilitated by Orc2-ISM, which places a bulky Trp396 deep into the minor groove to form a hydrogen bond between the N1 of the indole group and the O2 of the thymine ring from the invariant T11 of the ACS (Fig. 3c). This tryptophan insertion seems to serve as a steric roadblock, forcing the upstream sequence of Orc1-BP to make a 90° turn away from DNA (Fig. 3a).

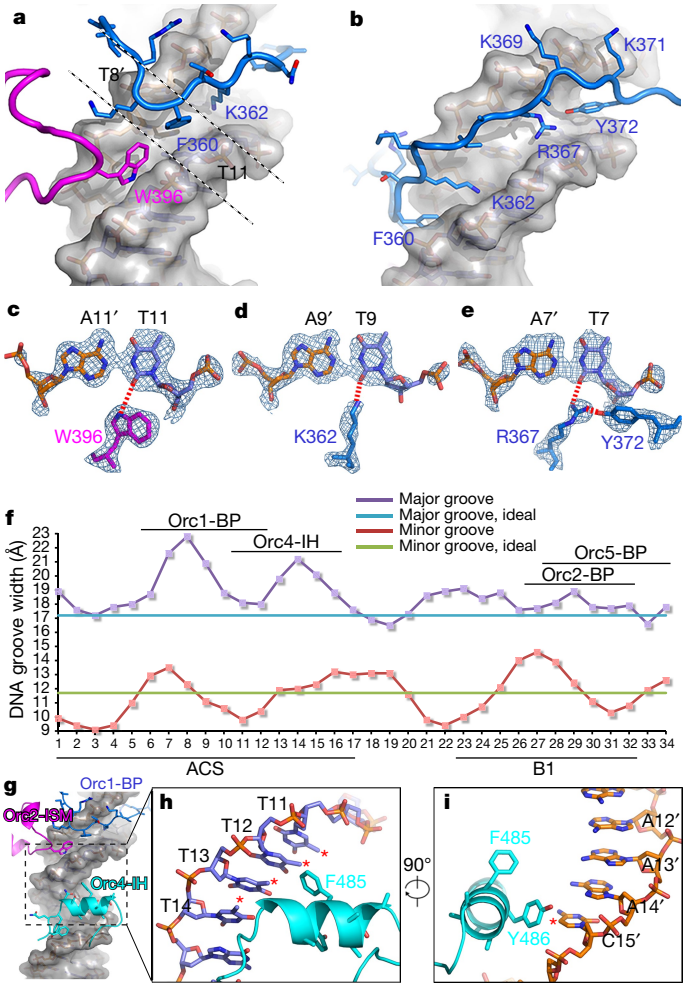


Fig. 3 | Base-specific recognition of ACS DNA by ORC. a, Zoomed-in view showing insertion of Orc1-BP into the minor groove of the ACS. Orc2-ISM and Orc1-BP are coloured magenta and blue, respectively. DNA is highlighted in grey surface representation. Selected residues from Orc1 and Orc2 and relevant nucleotides are labelled. **b**, Same as in **a**, but in a 90°-rotated view. **c**, Base recognition of T11 from the T-rich strand by W396 of Orc2. The DNA base pair is shown in stick model with cryo-EM density superimposed. The hydrogen bond between W396 and O2 of T11 is indicated by dashed lines (approximately 3.0 Å). Prime symbols denote bases on the opposite strand. **d**, Same as **c**, but for the hydrogen bond (approximately 2.5 Å) between K362 and the O2 of T9 (T-rich strand). **e**, Same as in **c**, but for the hydrogen bonds (approximately 2.5 Å) between R367 and the O2 of T7 (T-rich strand) as well as the OH of Y372. **f**, Plots of DNA groove widths of the ORC-bound DNA. Widths were measured as the distances between opposite phosphates of major and minor grooves. The major and minor groove widths of ideal B-form DNA are shown as two constant values, 17 and 12 Å, respectively. Contact regions for groove-inserting motifs are marked. Measurements were done with the program 3DNA[49]. **g**, The major groove insertion of the ACS by Orc4-IH. Minor groove insertion by Orc1-BP is also shown. **h**, Magnified view of the boxed region in **g** showing the hydrophobic interaction between F485 and the T-stretch (T11–T14) of the T-rich strand. Red asterisks denote the methyl group of thymine. **i**, Same as in **h**, but in a 90°-rotated view for the hydrophobic interaction between Y486 of Orc4-IH and C15′ of the A-rich strand. The red asterisk denotes the Hoogsteen edge of C15′.

Upon binding of Orc1-BP, the major groove is widened by more than 4 Å while the minor groove is widened by 2.5 Å at the C-terminal half but compressed by more than 1 Å at the N-terminal end of the basic patch, relative to ideal B-form DNA (Fig. 3f). By embedding in the minor groove, the side chains of Lys362 and Arg367 establish specific interactions with the bases of T9 and T7, respectively, on the T-rich strand via hydrogen bonds (Fig. 3d, e). These two specific thymine recognitions are facilitated by flanking residues near Lys362 and Arg367. The guanidinium group of Arg367 coordinates with the hydroxyl group of Tyr372 (also inserted in the minor groove) and O2 of T7 in a perfect triangular configuration (Fig. 3e). Phe360 (also inserted into the minor groove) is sandwiched between two opposing deoxyribose moieties of opposite strands by hydrophobic interactions (Fig. 3a), probably contributing to the positioning of the side chain of Lys362 towards the O2 of T9 (Fig. 3d).

Importantly, this basic patch was proposed as the eukaryotic origin sensor, as substitution mutations of Lys362 and Arg367 showed that they are crucial in both in vitro ACS binding and cell viability[35]. Basic patches are common motifs among DNA-binding proteins such as TATA box binding proteins and high mobility group (HMG) proteins that show a preference for A–T bases[36]. Comparative sequence analyses show that similar basic patches are found in Orc1 across species from yeast to humans[35] (Extended Data Fig. 8a), suggesting that this motif may also help to anchor ORC in T-rich sequences in all eukaryotes. In support of this hypothesis, metazoan ORCs from frog, fly and human also prefer AT-rich DNA substrates to some extent[37–41].

Another important motif contributing to the sequence-specific interaction is a special insertion α-helix (IH) in the β-hairpin of Orc4-WHD. As also seen in OCCM[30], Orc4-IH is deeply inserted into the major groove of the ACS (Fig. 3g). In particular, this region contains four consecutive thymines (T11–T14) in ARS305, and their methyl groups create a local hydrophobic environment (Fig. 3h). Orc4-IH, on the other hand, contains mostly hydrophobic residues with four aromatic side chains, in which Phe485 is especially close to the methyl groups of T12 and T13 (Fig. 3h). This observation suggests that Orc4-IH is most likely involved in sequence-specific

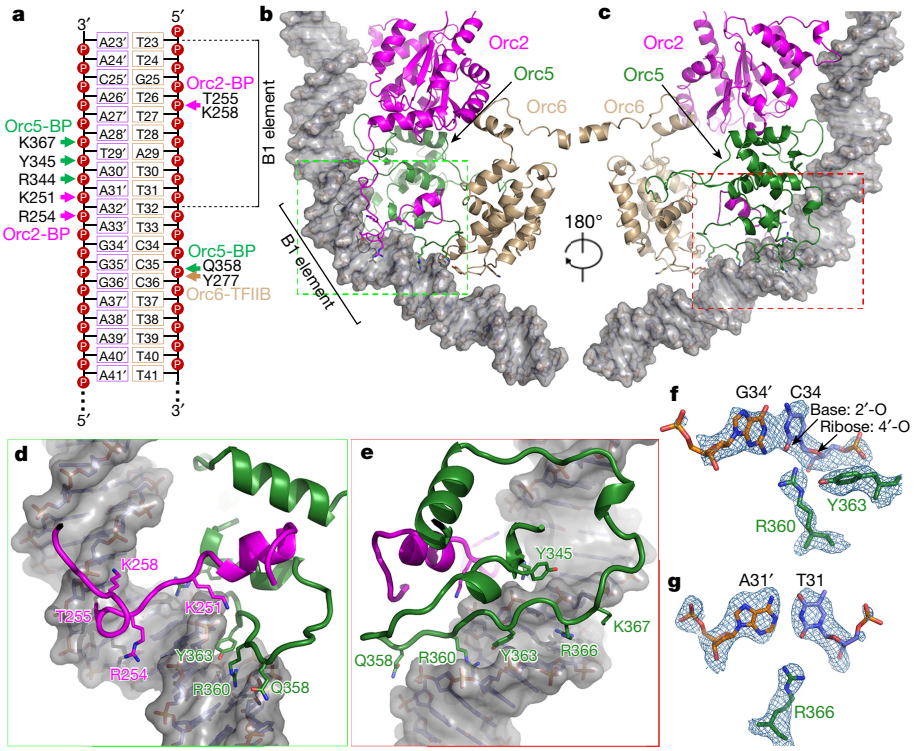


Fig. 4 | Interactions between ORC and DNA around the B1 element. **a**, Summary of non-specific ORC–DNA backbone interactions around B1 region. **b**, **c** Orc2, Orc5, and Orc6 interact with DNA around the B1 site. **d**, Magnified view of the boxed region in **b** showing insertion of the helical turn motif of Orc2 in the major groove. **e**, Magnified view of the boxed region in **c** showing insertion of Orc5-BP in the minor groove. **f**, Arg360 of Orc5-BP is orientated towards the O2 of C34 and O4 of the sugar ring (both less than 3 Å distance). **g**, Arg366 of Orc5-BP points to the O2 of T31 of the T-rich strand (approximately 3.5 Å).

interaction with the thymine stretch in the ACS sequence. Upon Orc4-IH binding, the major groove is widened by 4 Å compared to ideal B-form DNA (Fig. 3f). In addition, the hydrophobic ring of Tyr486 stacks upon the hydrophobic Hoogsteen edge of the conserved C15′ from the A-rich strand (Fig. 3i), also contributing to sequence specificity. Notably, this special insertion α-helix is completely absent in metazoans (Extended Data Fig. 8b), which may partially explain the divergent specificity of ORC for origin DNA among different species[30].

Throughout these analyses, we have discovered several base-specific interactions from both the conserved and species-specific elements of the ORC subunits. Notably, most of them involve thymine nucleotides through specific recognition of the methyl group from the major groove (Hoogsteen edge) or the free O2 from the minor groove (sugar edge) of the thymine ring (Fig. 3). These thymine-specific interactions perfectly explain the 'uniqueness' in structure and sequence of the signature asymmetric T-rich sequence of the *S. cerevisiae* ACS.

B1 interaction with Orc2, Orc5 and Orc6

Biochemical and genetic evidence suggest that ORC also binds to the B1 element of origins[24,42,43]. Notably, in the ORC–DNA structure, two motifs from Orc2 and Orc5 were found to be inserted in the major and minor grooves of the B1 DNA, respectively (Fig. 4). A flexible loop (residues 251–258, also a basic patch) upstream Orc2-AAA+ forms a helical turn in the major groove (Fig. 4b, d) contacting DNA backbone through basic residues such as Lys251, Arg254 and Lys258 (Fig. 4a, d). In analogy to the ACS recognition by Orc1-BP, a basic patch (residues 358–367) from Orc5-WHD is inserted into the adjacent minor groove (Fig. 4c, e), which is also conserved to a certain extent across species (Extended Data Fig. 8c). However, by contrast, these embedded side chains of Orc5-BP display relatively weak interactions with the bases. Arg360 of Orc5-BP, coordinated by Tyr363, points to the O2 of C34 and the O4 of the sugar ring (both approximately 3 Å); Arg366 is orientated towards the O2 of T31, at a distance of around 3.5 Å (Fig. 4f, g). As to Orc6, the N-terminal sequence of its TFIIB domain B (TFIIB-B), including Tyr277, interacts with the backbone of DNA (Fig. 4a, Extended Data Fig. 5e). Importantly, as seen in the structures from 3D classification, the extent of DNA curving positively correlates with the stability of Orc6 in the ORC–DNA complex (Extended Data Fig. 6f, g), suggesting the bending at the B1 site comes from cooperative effects from all three interacting ORC subunits.

The lack of a strong discrimination of bases at the B1 interaction site echoes the diversified B1 elements in yeast origins[44]. The A/T-rich nature of B1 element suggests that Orc5-BP could potentially contribute to specific base recognition as well, because the DNA bending is a dynamic process and hydrogen bonds similar to Orc1-BP could be established readily with the O2 of thymines in the minor groove of the B1 element.

Implications in origin DNA selection and MCM loading

There are three salient lessons learned from this study. First is the importance of Orc1-BP in binding to the minor groove of the ACS to recognize three conserved bases, T7, T9 and T11 (Fig. 3). We note that there is some degeneracy in the T7 and T9 positions for either A or T in the ACS (Fig. 5a). An explanation for this degeneracy is that hydrogen donors such as Arg367 are in the same proximity to hydrogen receptors of O2 of thymine or N3 of adenine in the minor groove such that A–T or T–A base pairs are indistinguishable for hydrogen bonding[45,46]. Another explanation is that induced-fit by the flexible Orc1-BP allows contact with the O2 of thymine on either strand at these positions. Further examination of Orc1 sequence indicates that multiple basic patches are present in the flexible linker region (Extended Data Fig. 8e). Notably, these additional basic patches in Orc1 are also found in other species (Extended Data Fig. 8e, f).

Second is the role of Orc4-IH in binding to consecutive thymines (11–14, mostly invariant) and a conserved C15′ (A-rich strand) of the ACS in the major groove (Fig. 3g–i). Given the fact that this Orc4-IH is absent in metazoans (Extended Data Fig. 8b), Orc1-BP may be the

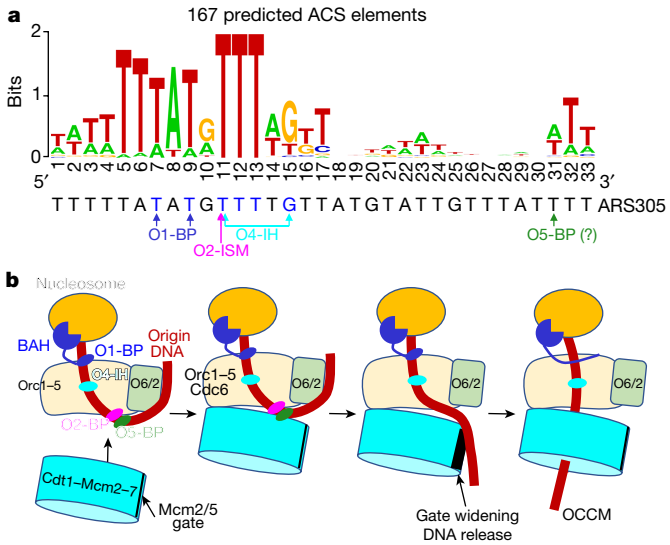


Fig. 5 | Model of origin recognition and MCM loading by ORC. **a**, Position weight matrix of the ORC binding consensus built with a predicted set including 167 ACS elements. All ACS used in these analyses were previously published[16] (Supplementary Table 1). The sequence logo was generated using http://weblogo.berkeley.edu. Conserved bases of ARS305 recognized by ORC motifs (O1-BP, O2-ISM, O4-IH and probably O5-BP) are highlighted. **b**, Schematic model showing the positioning of ORC next to a nucleosome and the subsequent loading of MCM. Docking of ORC at the initiation site is facilitated by the binding of Orc1-BAH to the nucleosome and the searching of ACS by Orc1-BPs. Establishment of ACS and B1 interactions by ORC base-specific recognition motifs curves the DNA to facilitate docking of Cdt1–MCM. Regulated DNA liberation and widening of the Mcm2/Mcm5 gate ensure precise insertion of origin DNA into the MCM ring.

major or sole determinant for DNA recognition and hence the more relaxed sequence specificity for ORC in metazoans.

Third is the bending of the origin DNA at successive points at the ACS and B1 elements (Fig. 1d) induced by deformation of major and minor grooves through interactions with various ORC subunits (Fig. 3f). As a consequence of this bending, the TFIIB-B of Orc6 conceivably could come into contact with the origin DNA beyond the B1 element (Extended Data Fig. 5g). This particular feature, if true, will shed light on the special nature of ORC–origin DNA complex in organizing chromatin structure. The formation of a nucleosome-like structure between ORC and origin DNA has been postulated based on biochemical studies[24,25] and a low-resolution electron microscopy study of *Drosophila* ORC bound to DNA[29]. This nucleoprotein structure has been implicated to provide a nucleosome-free region for replication initiation by phasing adjacent nucleosomes[15,16,21].

In brief, our ORC–DNA structure provides important insights into origin recognition in all eukaryotes. Our working hypothesis is that ORC-binding sites in eukaryotes are determined by a combination of factors at least including the BAH domain and Orc1-BP. The BAH domain of yeast, fly, plant and human Orc1 are known to interact with nucleosomes juxtaposed to replication origins[9,14,20,23]. We believe that the BAH domain and Orc1 basic patches may be involved in the initial searching and anchoring of ORC at DNA sites close to the designated initiation site (Fig. 5b). The precise binding to specific sites for replication initiation is likely to depend on a variety of means such as recruitment by other protein factors[14,47], modified histones, and specific DNA structures (G4 DNA)[48]. In *S. cerevisiae*, the positioning of ORC at the ACS is achieved by specific interactions between the essential Orc1-BP and Orc4-IH with invariant bases of the ACS (Fig. 5b).

Although the specificity of ORC in DNA binding is highly diverged, DNA bending induced by ORC could be shared by all eukaryotes to serve as a mechanism for the loading of MCM complex by ORC onto origin DNA. Notably, superimposition of the ORC–DNA and OCCM structures[30] indicates that the bent DNA is positioned at the gate between the Mcm2 and Mcm5 subunits of the MCM ring in the OCCM (Extended Data Fig. 9). This perfect alignment suggests that DNA bending by ORC has a unique function in MCM loading. We envision that by bending, DNA is tugged away from the interacting surfaces of ORC and the MCM ring to maximize their contacts while aligned with the Mcm2/Mcm5 gate of the MCM ring (Fig. 5b). The coordination of DNA straightening and gate opening in the formation of OCCM is a testable hypothesis.

In summary, information derived from the ORC–DNA structure not only provides a structural framework for understanding how ORC recognizes and binds yeast origin DNA, but also provides insight into the origin selection mechanism in metazoans. The induced bending of origin DNA paves the way for studying the potentially conserved roles of ORC in regulating both MCM loading and chromatin organization in all eukaryotes.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0293-x.

Received: 26 January 2018; Accepted: 8 May 2018;
Published online 4 July 2018.

1. Bleichert, F., Botchan, M. R. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Science* **355**, eaah6317 (2017).
2. Costa, A., Hood, I. V. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Annu. Rev. Biochem.* **82**, 25–54 (2013).
3. 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).
4. Duderstadt, K. E., Chuang, K. & Berger, J. M. DNA stretching by bacterial initiators promotes replication origin opening. *Nature* **478**, 209–213 (2011).
5. 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).
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. Miller, J. M. & Enemark, E. J. Fundamental Characteristics of AAA+ Protein Family Structure and Function. *Archaea* **2016**, 9294307 (2016).
8. Bell, S. P. The origin recognition complex: from simple origins to complex functions. *Genes Dev.* **16**, 659–672 (2002).
9. Hoggard, T. & Fox, C. A. in *The Initiation of DNA Replication in Eukaryotes* (ed. Kaplan, D. L.) 159–188 (Springer International Publishing, 2016).
10. Duncker, B. P., Chesnokov, I. N. & McConkey, B. J. The origin recognition complex protein family. *Genome Biol.* **10**, 214 (2009).
11. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).
12. Nieduszynski, C. A., Knox, Y. & Donaldson, A. D. Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev.* **20**, 1874–1879 (2006).
13. 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).
14. 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).
15. Lipford, J. R. & Bell, S. P. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell* **7**, 21–30 (2001).
16. 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).
17. Deshpande, A. M. & Newlon, C. S. The ARS consensus sequence is required for chromosomal origin function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**, 4305–4313 (1992).
18. Rao, H., Marahrens, Y. & Stillman, B. Functional conservation of multiple elements in yeast chromosomal replicators. *Mol. Cell. Biol.* **14**, 7643–7651 (1994).
19. Royzman, I., Austin, R. J., Bosco, G., Bell, S. P. & Orr-Weaver, T. L. ORC localization in *Drosophila* follicle cells and the effects of mutations in dE2F and dDP. *Genes Dev.* **13**, 827–840 (1999).
20. 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).
21. Berbenetz, N. M., Nislow, C. & Brown, G. W. Diversity of eukaryotic DNA replication origins revealed by genome-wide analysis of chromatin structure. *PLoS Genet.* **6**, e1001092 (2010).

22. Simpson, R. T. Nucleosome positioning can affect the function of a cis-acting DNA element *in vivo*. *Nature* **343**, 387–389 (1990).
23. Li, S. et al. Structural basis for the unique multivalent readout of unmodified H3 tail by *Arabidopsis* ORC1b BAH-PHD cassette. *Structure* **24**, 486–494 (2016).
24. Lee, D. G. & Bell, S. P. Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol. Cell. Biol.* **17**, 7159–7168 (1997).
25. 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).
26. Chen, Z. et al. The architecture of the DNA replication origin recognition complex in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **105**, 10326–10331 (2008).
27. Sun, J. et al. Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. *Structure* **20**, 534–544 (2012).
28. 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).
29. Clarey, M. G., Botchan, M. & Nogales, E. Single particle EM studies of the *Drosophila melanogaster* origin recognition complex and evidence for DNA wrapping. *J. Struct. Biol.* **164**, 241–249 (2008).
30. 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).
31. Bleichert, F., Botchan, M. R. & Berger, J. M. Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321–326 (2015).
32. Tocilj, A. et al. Structure of the active form of human origin recognition complex and its ATPase motor module. *eLife* **6**, e20818 (2017).
33. 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).
34. Zhai, Y. et al. Unique roles of the non-identical MCM subunits in DNA replication licensing. *Mol. Cell* **67**, 168–179 (2017).
35. 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).
36. Bewley, C. A., Gronenborn, A. M. & Clore, G. M. Minor groove-binding architectural proteins: structure, function, and DNA recognition. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 105–131 (1998).
37. 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).
38. 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).
39. Kong, D., Coleman, T. R. & DePamphilis, M. L. *Xenopus* origin recognition complex (ORC) initiates DNA replication preferentially at sequences targeted by *Schizosaccharomyces pombe* ORC. *EMBO J.* **22**, 3441–3450 (2003).
40. Vashee, S. et al. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* **17**, 1894–1908 (2003).
41. 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).
42. 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).
43. 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).
44. Lucas, I. A. & Raghuraman, M. K. The dynamics of chromosome replication in yeast. *Curr. Top. Dev. Biol.* **55**, 1–73 (2003).
45. Hélène, C. DNA recognition. Reading the minor groove. *Nature* **391**, 436–438 (1998).
46. Seeman, N. C., Rosenberg, J. M. & Rich, A. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA* **73**, 804–808 (1976).
47. Thomae, A. W. et al. Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins. *Proc. Natl Acad. Sci. USA* **105**, 1692–1697 (2008).
48. 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).
49. Lu, X. J. & Olson, W. K. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* **3**, 1213–1227 (2008).

Acknowledgements We thank the Electron Microscopy Laboratory of Peking University (cryo-EM platform) and the Tsinghua University Branch of the China National Center for Protein Sciences (Beijing) for the data collection of the ORC–DNA and apoORC samples, respectively. The computation was supported by High-performance Computing Platform of Peking University. This work was supported by the Ministry of Science and Technology of China (2016YFA0500700 to N.G.), the National Natural Science Foundation of China (NSFC) (31761163004, 31725007 and 31630087 to N.G.; 31700655 to N.L.), the Research Grants Council (RGC) of Hong Kong (GRF16138716 to B.K.T.; GRF16104115, GRF16143016 and GRF16104617 to Y.L.Z. and B.K.T.), NSFC/RGC Joint Research Scheme (N_HKUST614/17 to N.G., B.K.T., Y.L.Z. and N.L.), and the China Postdoctoral Science Foundation (2017M610013 to N.L.). N.L. is supported by Young Elite Scientists Sponsorship Program by CAST and a postdoctoral fellowship from the Peking-Tsinghua Centre for Life Sciences.

Reviewer information *Nature* thanks C. Fox and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions Y.L.Z., N.G. and B.K.T. conceived the study; W.H.L., Y.L.Z. and Y.Q.Z. purified ORC; J.C., N.L., E.C., W.H.L. and Y.L.Z. prepared cryo grids; N.L. and J.C. collected data; N.L., J.C., Y.L.Z. and N.G. processed data; and N.L., W.H.L., Y.L.Z., N.G. and B.K.T. prepared the manuscript. N.L., W.H.L., Y.L.Z. and J.C. contributed equally to the study.

Competing interests The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0293-x.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0293-x.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to Y.L.Z. or N.G. or B.-K.T.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

ORC purification. Orc1–Orc6 was purified with a yeast strain ySD-ORC (a gift from J. Diffley) as previously described[50] with the following modifications. In brief, 9 l of yeast cells were cultured in YPA medium containing 2% (w/v) raffinose at 30 °C until OD_{600} reached 3.0. Alpha factor was then added to a final concentration of 100 ng ml^{-1} to synchronize cells at G1 phase before overexpression induction of Orc1–Orc6 by addition of 2% (w/v) D-galactose for 3 h. Cells were collected and the pellet was washed with ice-cold water twice, and resuspended in 50 ml lysis buffer (25 mM HEPES-KOH pH 7.6, 0.5 M KCl, 10% glycerol, 0.05% (v/v) NP-40, 1× PI cocktail (Roche) and 1 mM PMSF). The cell suspension was frozen drop-wise in liquid nitrogen. Popcorns of the frozen cells were then crushed using a freezer mill (SPEX CertiPrep 6850 Freezer/Mill) with a setting of 15 cps, 2 min on, 2 min off, for 6 cycles. The powder was thawed on ice with addition of an equal volume of lysis buffer. The suspension was centrifuged twice for clarification at 38,900*g* for 20 min at 4 °C with a R20A2 rotor. The supernatant was incubated with 5 ml of 50% slurry of calmodulin affinity resin (Agilent Technologies) in the presence of 2 mM $CaCl_2$ for 4 h at 4 °C. After removal of flow-through and extensive washing of the resin with 50 column volumes of ice-cold washing buffer (25 mM HEPES-KOH pH 7.6, 0.3 M KCl, 10% glycerol, 2 mM $CaCl_2$ and 0.05% (v/v) NP-40), elution was performed by incubating the resins with 4 ml of elution buffer (25 mM HEPES-KOH pH 7.6, 0.3 M KCl, 10% glycerol, 1 mM EDTA, 2 mM EGTA and 0.05% (v/v) NP-40) for 30 min at 4 °C. The eluent was concentrated and further fractionated with size-exclusion chromatography (SEC) with Superose 6 10/300 GL column (GE Healthcare) in SEC buffer (25 mM HEPES-KOH pH 7.6, 0.15 M KCl, 10% glycerol, 2 mM β-mercaptoethanol, 0.05% (v/v) NP-40). Peak fractions containing the target complex were pooled and frozen in liquid nitrogen for storage.

Preparation of the ORC–DNA sample. DNA oligonucleotides used in this study (36 bp forward: 5′-TGGTTTTTATATGTTTTGTTATGTATTGTTTATTTT-3′ and 36 bp reverse: 5′-AAAATAAACAATACATAACAAAACATATAAAAACCA-3′, 72 bp forward: 5′-TGGTTTTTATATGTTTTGTTATGTATTGTTTATTTTCCCTTTA ATTTTAGGATATGAAAACAAGAATTTATC-3′ and 72 bp reverse: 5′-GATAA ATTCTTGTTTTCATATCCTAAAATTAAAGGGAAAATAAACAATACATAAC AAAACATATAAAAACCA-3′) were first dissolved in a buffer containing 10 mM HEPES-KOH (pH 7.5), 1 mM EDTA and 50 mM NaCl to a final concentration of 100 μM. Equal volumes of the pairs of dissolved DNA oligonucleotides were mixed and incubated at 94 °C for 5 min. The reaction was left in the heat block to cool naturally to room temperature for DNA annealing. To assemble ORC onto the annealed origin DNA, 370 μl of 1.6 μM ORC was first mixed with an equal volumes of a buffer containing 25 mM HEPES-KOH pH 7.6, 0.1 M potassium acetate, 8 mM magnesium acetate, 0.02% (v/v) NP-40 and 2 mM ATPγS, and incubated on ice for 30 min. The purified ORC was unstable and prone to dissociate into sub-complexes (Extended Data Fig. 1b). To preserve the integrity of the assembled ORC–DNA, a mild fixation was applied to the complexes before cryo-EM grid preparation. Specifically, after a further incubation for 30 min at room temperature, the mixture was concentrated to 100 μl and then applied on the top of a 2-ml glycerol (10–30%) gradient containing glutaraldehyde (0–0.025%) in a buffer (25 mM HEPES-KOH pH 7.6, 0.1 M potassium acetate, 8 mM magnesium acetate, 0.5 mM β-mercaptoethanol) for GraFix[51]. The gradient was centrifuged in a Beckman TLS55 rotor (Beckman Optima TLX ultracentrifuge) for 13.5 h at a speed of 82,000*g* at 4 °C. Fractions were collected and the crosslinking reaction was quenched by addition of ice-cold Tris-HCl (pH 8.0) to a final concentration of 40 mM. For apoORC sample, the ORC protein was prepared in the same way but without incubation with ATPγS and dsDNA oligonucleotides. The GraFix treatment dramatically stabilized the ORC–DNA sample on cryo-EM grids (Extended Data Fig. 1d).

Electron microscopy. Fractions containing intact ORC–DNA (or ORC) complexes (Extended Data Fig. 1a) were pooled. Ultrafiltration for removal of glycerol and buffer exchange (25 mM HEPES-KOH (pH 7.6), 100 mM potassium acetate, 5 mM magnesium acetate, 0.5 mM β-mercaptoethanol) was performed with a centrifugal filter (Amicon Ultra-0.5ml 50 k) at 6,000*g* at 4 °C. For negative staining, 4-μl aliquots of samples in serial dilutions were applied onto copper grids, washed and stained with 2% uranyl acetate. The negative-stained grids were examined using an FEI Tecnai T12 electron microscope operated at 120 kV, equipped with a 4 k × 4 k CCD camera (Gatan). The relative concentration of the samples used for cryo-grid preparation (about 10 times higher) were estimated from negative-staining electron microscopy. Aliquots (4 μl) of samples were applied onto the glow-discharged holey-carbon gold grids (Quantifoil, R1.2/1.3, 300 mesh) and blotted using an FEI Vitrobot Mark IV. Cryo-grids were screened using an FEI Talos Arctica operated at 200 kV with an FEI Ceta camera. Good cryo-grids were recovered and stored in liquid nitrogen. For the ORC–DNA sample, data were collected using a Titan Krios microscope (FEI) operated at 300 kV and images were collected using Leginon[52] with a nominal magnification of 130,000× and a defocus range of 1.5–2.5 μm. The images were recorded using a K2 summit camera equipped with GIF Quantum energy filter (Gatan) in the super-resolution counting and movie mode, with a dose rate of 4.4e^- s^- $Å^{-2}$ and a total exposure time of 12 s. Each movie stack contains 40 frames with a calibrated pixel size of 0.526 Å at object scale (super-resolution). For the sample of apoORC, data acquisition was performed on a Titan Krios operated at 300 kV with defocus ranging from 1.5 to 2.5 μm. Images were recorded using a K2 summit camera (Gatan) in a super-resolution counting mode at a nominal magnification of 22,500 × corresponding to a calibrated pixel size of 0.66 Å at object scale. Data were acquired semi-automatically using UCSF-Image4[53] in a movie mode with a dose rate of 10.2 e^- s^- $Å^{-2}$ and a total exposure time of 8 s, rendering a movie stack of 32 frames for each micrograph.

Image processing. A total of 1,765 movie stacks were acquired for the ORC–DNA (36 bp) sample (Extended Data Fig. 2a). Drift-correction, electron-dose weighting and two-fold binning were applied to the movie stacks using MotionCor2[54], which generate summed images with or without dose weighting. Program of CTFFIND4[55] was used to evaluate the parameters of contrast transfer function (CTF) of each micrograph based on images without dose weighting. Summed images and CTF power spectra were screened manually and high quality images were kept for further processing. Around 2,000 particles were manually picked and subjected to two-dimensional (2D) classification using RELION2.0[56] to provide templates for large-scale particle auto-picking. With the templates, 478,000 particles were auto-picked from images without dose weighting using RELION2.0. After two rounds of 2D classification, 464,000 particles were kept for further processing (Extended Data Fig. 2a). The initial three-dimensional (3D) model was calculated using CryoSPARC[57]. To improve the performance of 3D classification, one round of 3D refinement was first performed on all particles selected after 2D classification using RELION2.0. Particles were re-centred according to the parameters from the last iteration of 3D refinement and re-extracted from dose-weighted images (particles that are too close to or exceed image edges were discarded). The first round of 3D classification was applied to exclude particles with obvious compositional or conformational heterogeneity. Around 42% good particles (196,000) were selected for further fine 3D classification with different parameters tested to separate particles corresponding to different conformational states as completely as possible (Extended Data Fig. 2a). Three different conformational states (I, II and III) were finally separated. Among these three, state I contained approximately 52% of particles and displayed more high-resolution features; states II and III each had ~12% of particles. States II and III were further refined to final resolutions of 4.7 Å and 5.6 Å, respectively. In the map of state II, Orc2-WHD moves towards the WHD tier of the ORC and the gap between Orc1 and Orc2 AAA+ domain is slightly smaller than State I (Extended Data Fig. 6b, d). While in the map of state III, Orc2-WHD is in the same position as that of state I but very unstable and the Orc1/Orc2 gap becomes larger compared to state I (Extended Data Fig. 6c, e). For state I, 3D refinement with a global mask applied improved the resolution to 3.8 Å. Application of particle-level local defocus evaluated using Gctf[58] based on images without dose weighting further improved the resolution to 3.6 Å. The resolution was estimated by gold-standard FSC 0.143 criteria, after correction of the mask effect. The map was sharpened by an auto-evaluated *B*-factors using RELION2.0. The local resolution map was created using ResMap[59] and displayed using UCSF Chimera[60].

For the ORC–DNA (72 bp) dataset, 3,603 micrographs were similarly processed (Extended Data Fig. 1d–f). A total of 427,000 particles were selected after 2D classification. After 3D classification, three classes (state I) were combined (approximately 164,000 particles) and subjected to high-resolution refinement, resulting in a final density map at a resolution of 3.0 Å. A major structural difference in other classes (states II and IV) from the ORC–DNA (72 bp) dataset is the extent of DNA curving (Extended Data Fig. 6f, g).

Micrographs (1,041 in total) of apoORC complexes were also similarly processed (Extended Data Fig. 2b). In total, 262,000 particles were auto-picked, and 189,000 particles were selected after 2D classification. After two rounds of 3D classification, 45,000 good particles were selected for final refinement, resulting in a final density map at a resolution of 8.2 Å. The overall subunit arrangement of this apoORC is in general the same as that of the origin-bound complex, but ORC alone displayed more tendency to dissociate during sample preparation even when GraFix was applied, indicating that DNA binding stabilizes the whole complex.

Model building. Modelling was first performed using the map of state I from the 36-bp dataset. To facilitate the modelling, multiple versions of density maps were obtained by applying additional rounds of mask-based refinement using RELION2.0. For example, to improve the resolution of the core region, a mask excluding Orc6, the Orc3 insertion domain and unstable portions of dsDNA was applied. These additional procedures have improved the quality of local densities, although the reported resolutions had no marked gain (Extended Data Fig. 2a, c).

The initial template used for the modelling was from the OCCM structure (PDB code 5UDB)[30]. Each domain of the Orc1–Orc5 subunits was manually fitted into the density map in Chimera, followed by manual rebuilding using Coot[61].

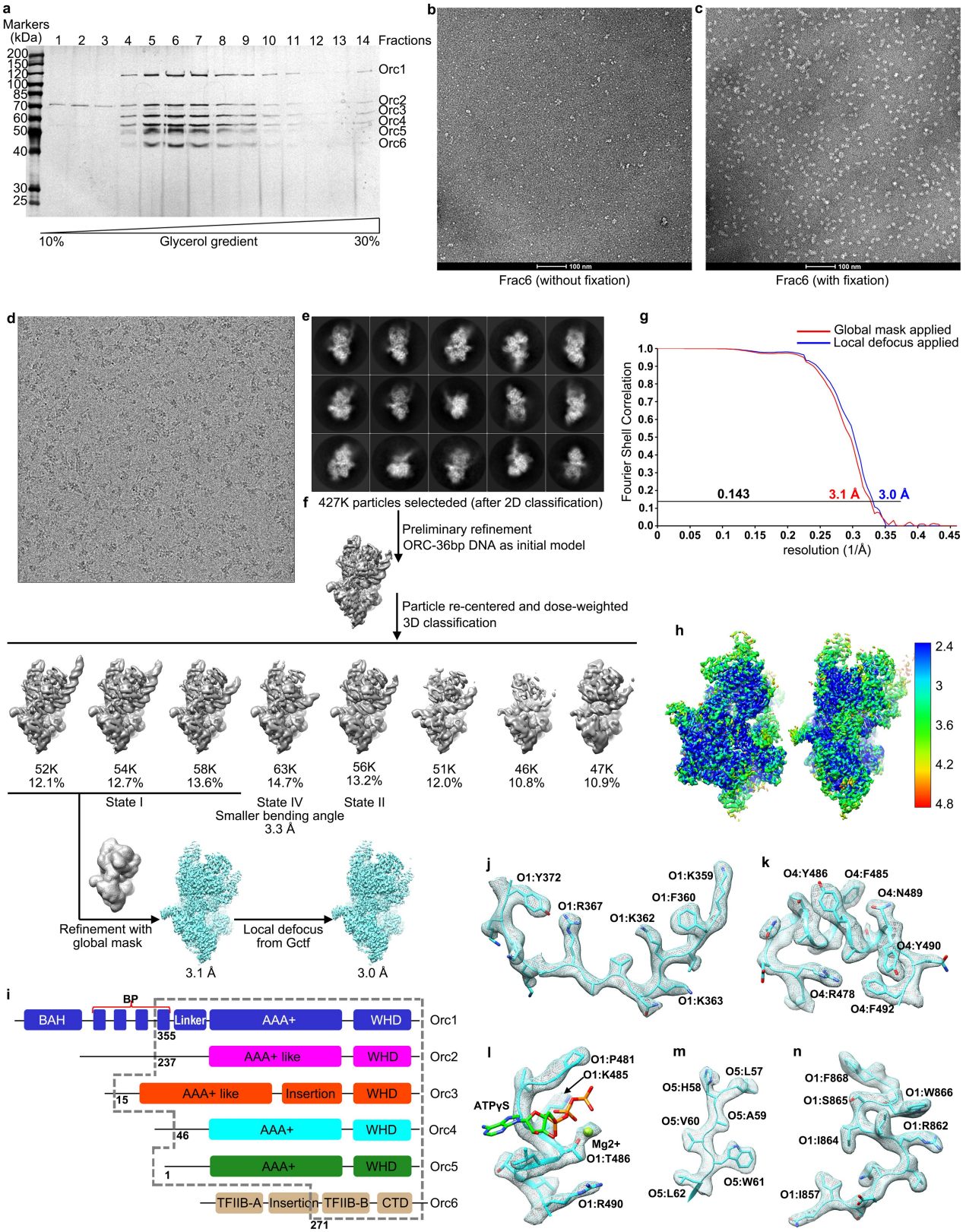
Linkers, flexible extensions and TFIIB-B domain of Orc6 were modelled de novo. Secondary structure prediction of the Orc1–Orc6 subunits was performed using PSIPRED[62]. For modelling the origin DNA, a 36-bp ideal B-form dsDNA (with ARS305) was generated, fitted in the density map and manually adjusted using Coot. The model was refined against the 3.6-Å density map using Phenix.real_space_refine[63] with secondary structure restraints, geometry restraints and DNA-specific restraints applied. Chains of the model were then extended with the 3.0-Å map from the 72-bp dataset, refined and optimized similarly. The final models were evaluated by MolProbity[64], and statistics are presented in Extended Data Table 1.

For the 8.2 Å map of the apoORC, each domain of the Orc1–Orc6 subunits from the atom model of the ORC–DNA structure was fitted into the density map in Chimera, generating a pseudo-atomic model for the apoORC complex. Chimera and Pymol (http://pymol.org) were used for figure preparation.

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

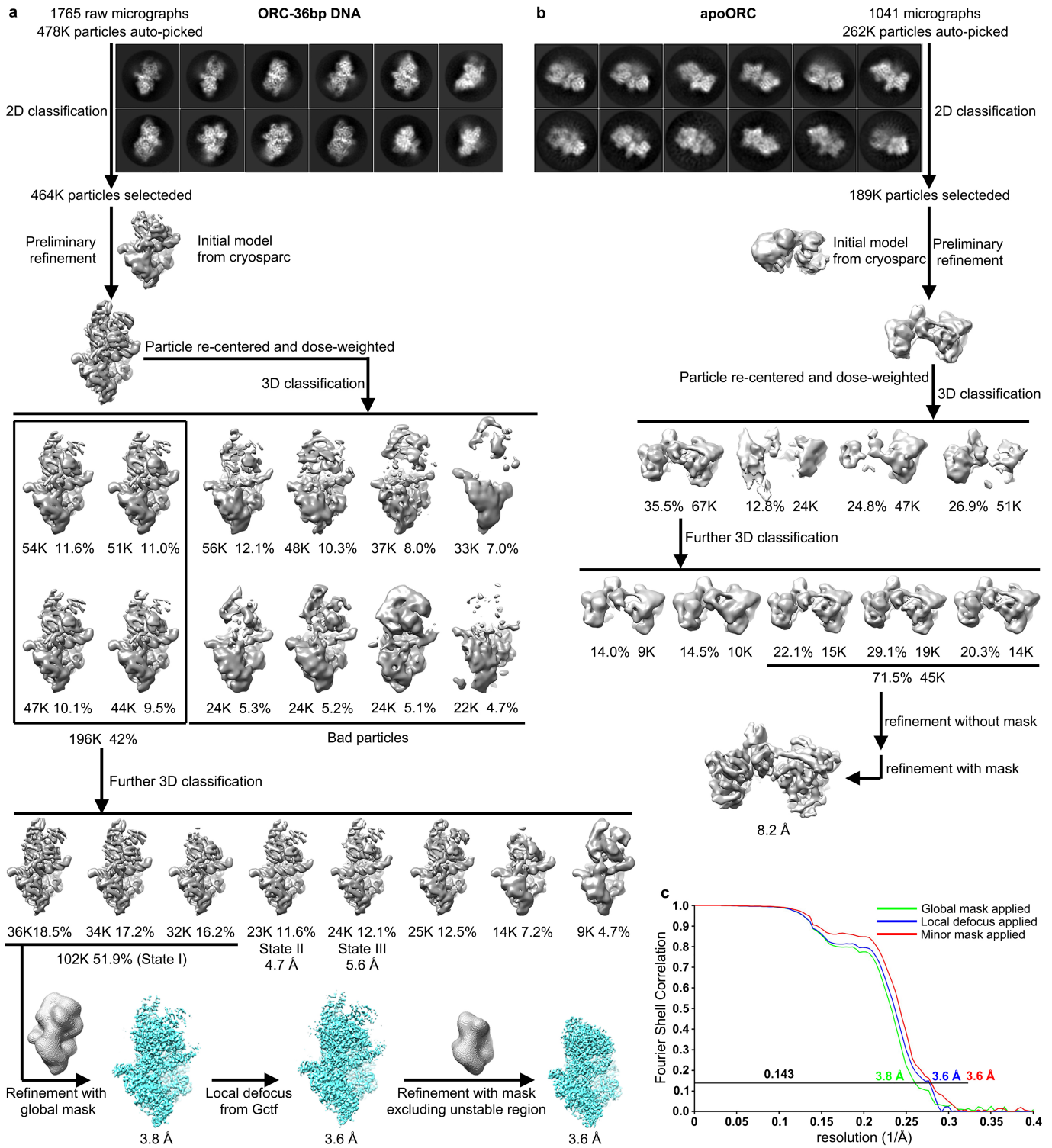
Data availability. Cryo-EM maps of the apoORC, ORC–36-bp and ORC–72-bp complexes have been deposited in the Electron Microscopy Data Bank (EMDB) with accession codes EMD-6943, EMD-6942 and EMD-6941, respectively. Atomic coordinates of the ORC–72-bp complex have been deposited in the Protein Data Bank (PDB) with accession code 5ZR1. All other data are available from the corresponding author upon reasonable request.

50. Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339–343 (2013).
51. Kastner, B. et al. GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008).
52. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol*. **151**, 41–60 (2005).
53. Li, X., Zheng, S., Agard, D. A. & Cheng, Y. Asynchronous data acquisition and on-the-fly analysis of dose fractionated cryoEM images by UCSFImage. *J. Struct. Biol*. **192**, 174–178 (2015).
54. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
55. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol*. **192**, 216–221 (2015).
56. 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).
57. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
58. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol*. **193**, 1–12 (2016).
59. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
60. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem*. **25**, 1605–1612 (2004).
61. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
62. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res*. **41**, W349-57 (2013).
63. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
64. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
65. Liu, S. et al. Structural analysis of human Orc6 protein reveals a homology with transcription factor TFIIB. *Proc. Natl Acad. Sci. USA* **108**, 7373–7378 (2011).
66. Nikolov, D. B. et al. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**, 119–128 (1995).

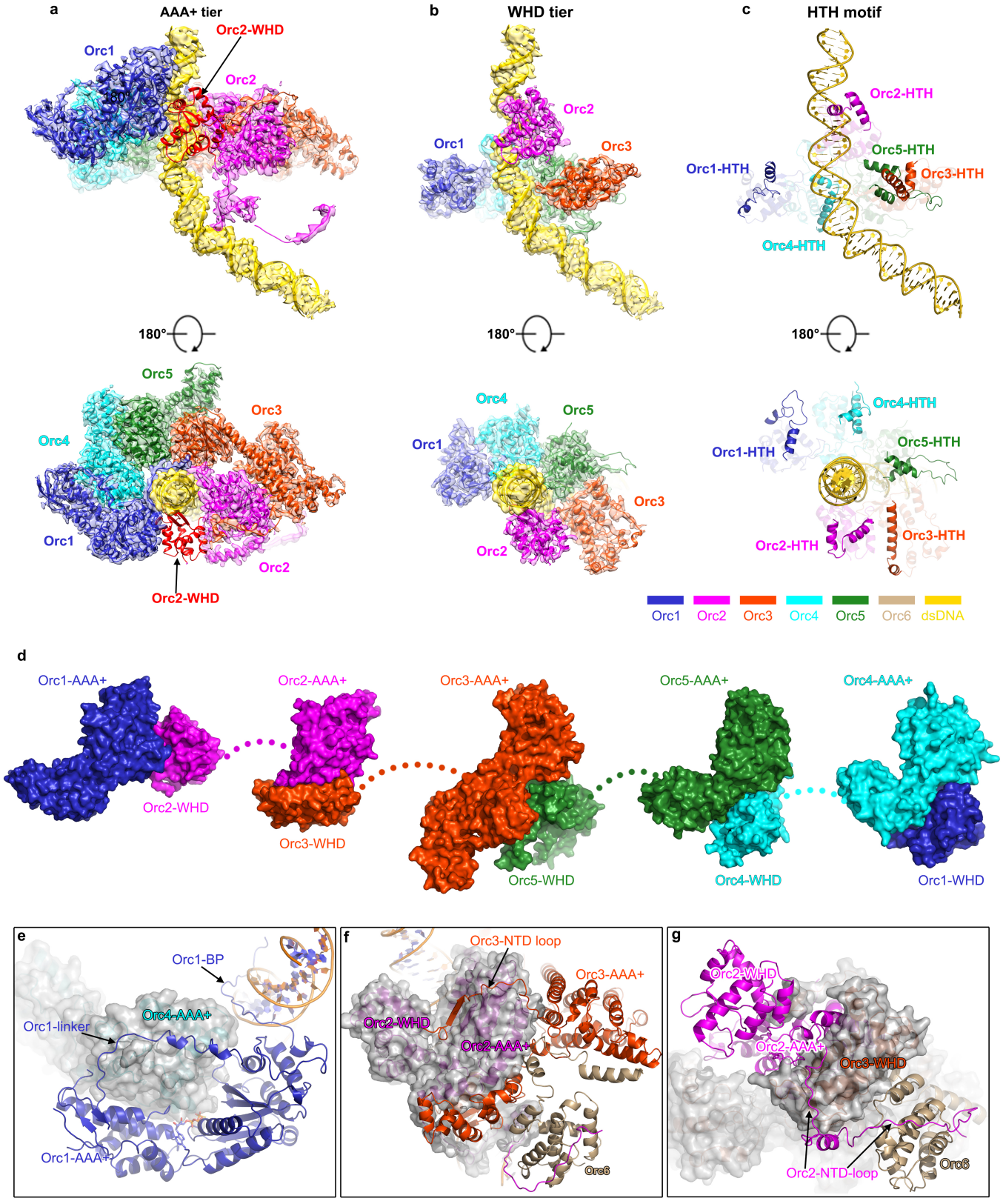


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Sample preparation and image processing of the ORC–DNA complex. a, SDS–PAGE analysis of the glycerol gradient fractions. ORC–DNA complexes (no fixation) were subjected to 10–30% glycerol gradient centrifugation. Fractions were collected and resolved on SDS–PAGE. Peak fractions (5–7) containing intact ORC complexes were processed for further electron microscopy analysis. Experiments were repeated multiple times ($n > 10$), and similar results were obtained. **b**, Negative-staining electron microscopy of the ORC–DNA complex. Samples from fraction 6 were subjected to negative staining. A severe dissociation of complexes was observed. **c**, Negative-staining electron microscopy of the ORC–DNA complex prepared with GraFix method using a gradient of glutaraldehyde (0–0.025%). **d**, A representative raw cryo-EM image of the ORC–DNA (72 bp) complex. **e**, 2D class averages of the ORC–DNA (72 bp) particles. **f**, Workflow of image processing of the ORC–DNA (72 bp) particles. The processing includes rounds of 2D classification, 3D classification, structural refinement and masked-based refinement procedures. **g**, FSC curves of the final density map of the ORC–DNA (72 bp) complex. **h**, The local resolution map of the final density map. **i**, Schematic domain organization of Orc1–Orc6 subunits. Regions that were built in the final atomic model were boxed in dashed grey lines. **j**–**n**, Local density of representative regions of the final cryo-EM density map, for Orc1-BP (**j**), Orc4-IH (**k**), the Orc1 ATP-binding pocket (**l**) and two other regions (**m**, **n**). For clarity, density of ATPγS is omitted in **j** to highlight the Walker A motif of Orc1.

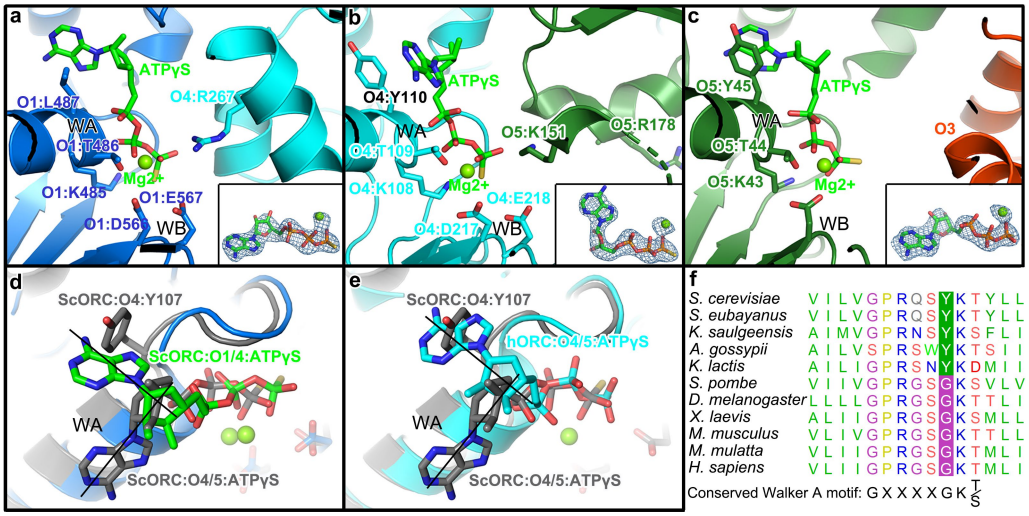


Extended Data Fig. 2 | Workflow of the image processing of the ORC–DNA (36 bp) and apoORC particles. a, Image processing workflow of the ORC–DNA (36 bp) particles. Processing includes rounds of 2D classification, 3D classification, structural refinement and masked-based refinement procedures. **b**, Image processing workflow of the apoORC particles. **c**, FSC curves of the density maps of the ORC–DNA (36 bp) complex.

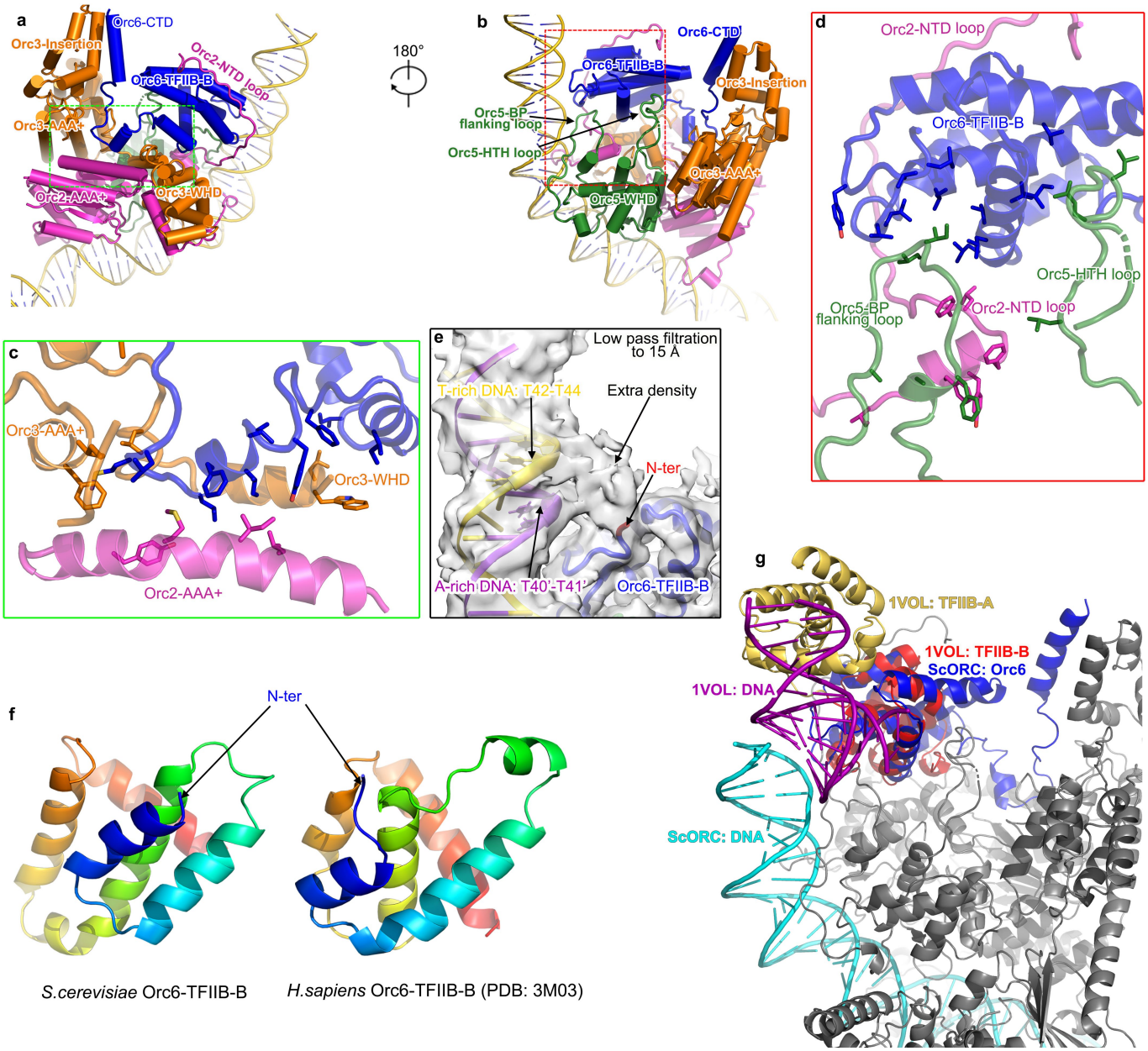


Extended Data Fig. 3 | See next page for caption.

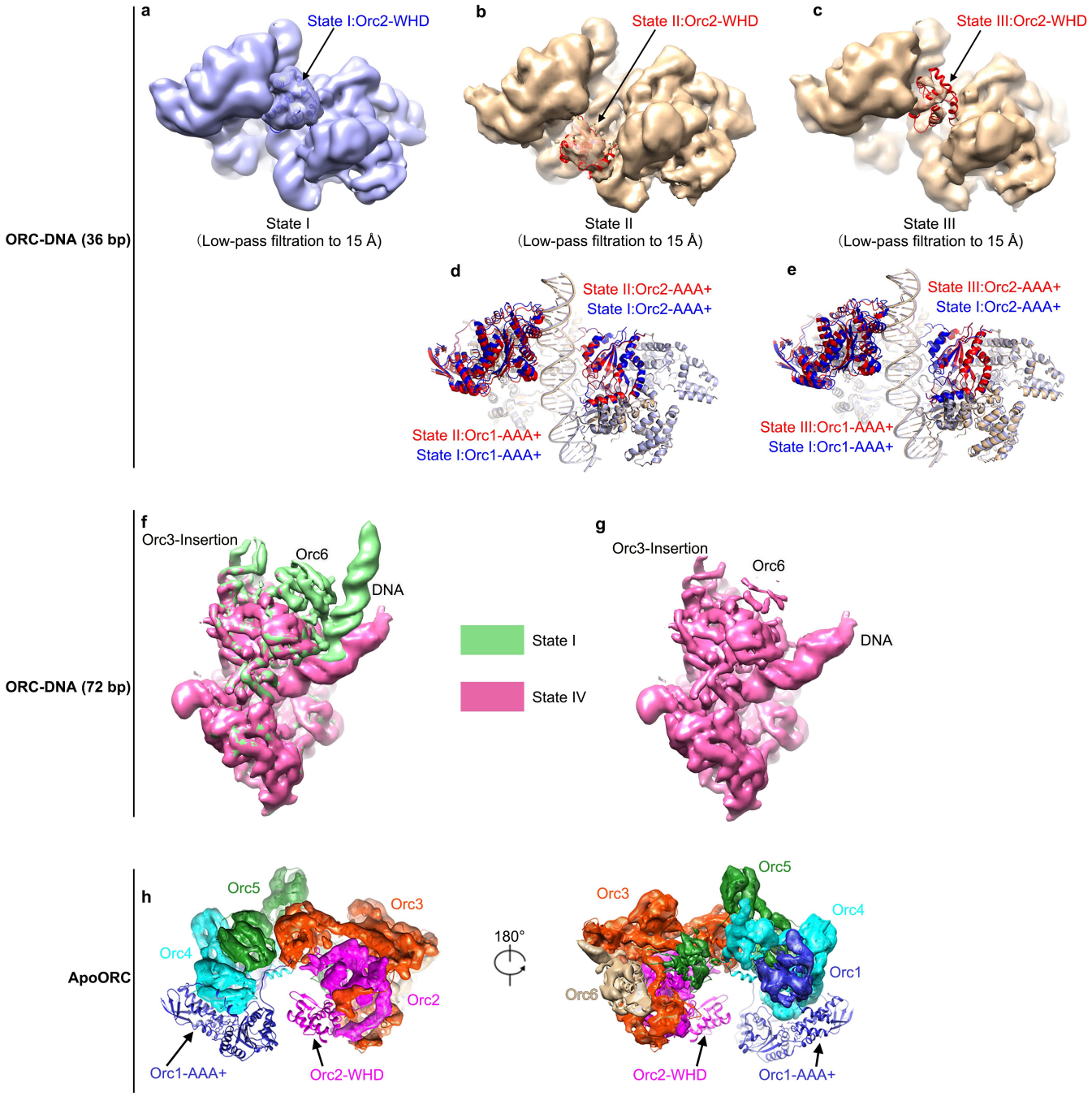
Extended Data Fig. 3 | Organization of AAA+ and WHD domains around the origin DNA. a, Organization of AAA+ domains of Orc1–Orc5 subunits around origin DNA. Cryo-EM maps of the AAA+ domains and DNA are shown in solid surface representation and colour-coded. The WHD of Orc2, which blocks the gap between the AAA+ domains of Orc1 and Orc2 is shown in cartoon representation. **b**, Same as in **a**, but for the WHDs of the Orc1–Orc5 subunits. **c**, Distribution of the HTH motifs of WHDs around the origin DNA. The WHDs of the Orc1–Orc5 subunits are shown in cartoon representation with the HTH motifs highlighted. **d**, Domain swapping between the AAA+ and WHD tiers. As shown, Orc2-WHD is in a different position from the rest. **e**, A flexible linker (residues 375–436) upstream of Orc1 AAA+ domains extends on the surface of Orc4 AAA+ domain (surface representation), with the further upstream basic patch sequences inserted into the minor groove of the ACS. The bound ATPγS is shown in stick model (orange). **f**, The very N-terminal extension (residues 15–50) of Orc3 wraps around the AAA+ domain of Orc2 (surface representation) and ends in the interface between the Orc2-WHD and Orc2-AAA+ domain. **g**, A very long N-terminal linker (NTD loop) upstream AAA+ domain of Orc2 extends on the surface of Orc3-WHD and TFIIB-B domain of Orc6. Note that the linker of Orc2 wrapping around Orc6 is traceable in the cryo-EM density map but the model could not be built at atomic level.



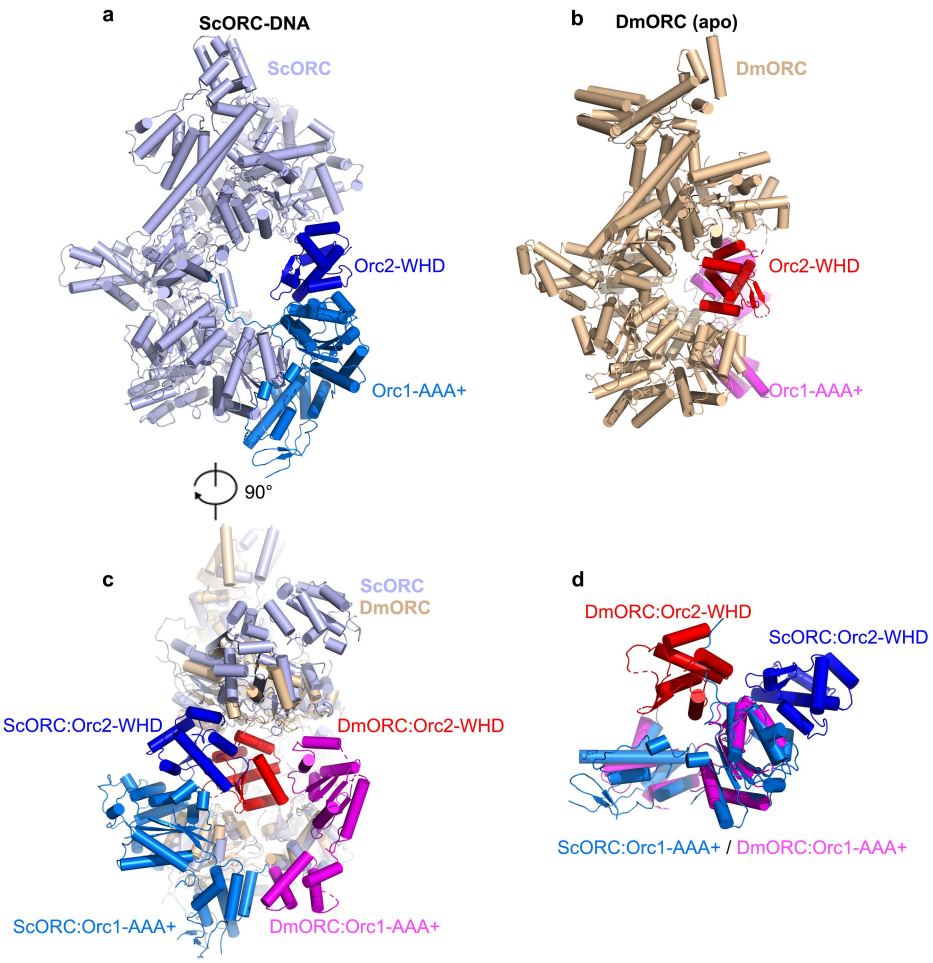
Extended Data Fig. 4 | Configuration of the three ATPase centres in the ORC–DNA complex. a, Zoomed-in view of the ATPase centre formed between Orc1 (O1) and Orc4 (O4). Orc1, Orc4 and ATPγS-Mg^{2+} are coloured blue, cyan and green, respectively. The Walker A and B motifs (WA and WB) of Orc1 and the arginine finger of Orc4 (R267) are highlighted in stick models. Inset, the stick model of ATPγS-Mg^{2+} superimposed with the cryo-EM density. **b**, Zoomed-in view of the ATPase centre formed between Orc4 and Orc5 (O5). Orc4, Orc5 and ATPγS-Mg^{2+} are coloured cyan, dark green and green, respectively. The Walker A and B motifs (WA and WB) of Orc4 and the equivalent arginine finger of Orc5 (R178) are highlighted in stick models. K151 of Orc5 within 4 Å distance from the γ-phosphate is shown. Inset, the stick model of ATPγS-Mg^{2+} superimposed with the cryo-EM density. **c**, Zoomed-in view of the ATPase centre formed between Orc5 and Orc3 (O3). Orc5, Orc3 and ATPγS-Mg^{2+} are coloured dark green, orange and green, respectively. The Walker A and B motifs of Orc5 are highlighted in stick models. Inset, the stick model of ATPγS-Mg^{2+} superimposed with the cryo-EM density. **d**, Comparison between the ATPase centres of O1:O4 and O4:O5, highlighting the flip of the base moiety of the bound ATPγS within the O4:O5 centre. The flip is forced by the replacement of a conserved glycine by a bulky tyrosine residue (Y107) of the Walker A motif of Orc4. The Walker A motifs were used as reference in the alignment. **e**, Comparison between the ATPase centres of the yeast O4:O5 and human O4:O5 (PDB code 5UJ7)[32], highlighting the flip of the base moiety of the bound ATPγS within the yeast O4:O5 centre. The Walker A motifs were used as reference in the alignment. **f**, Sequence alignment of the Walker-A motif of Orc4 from different species.



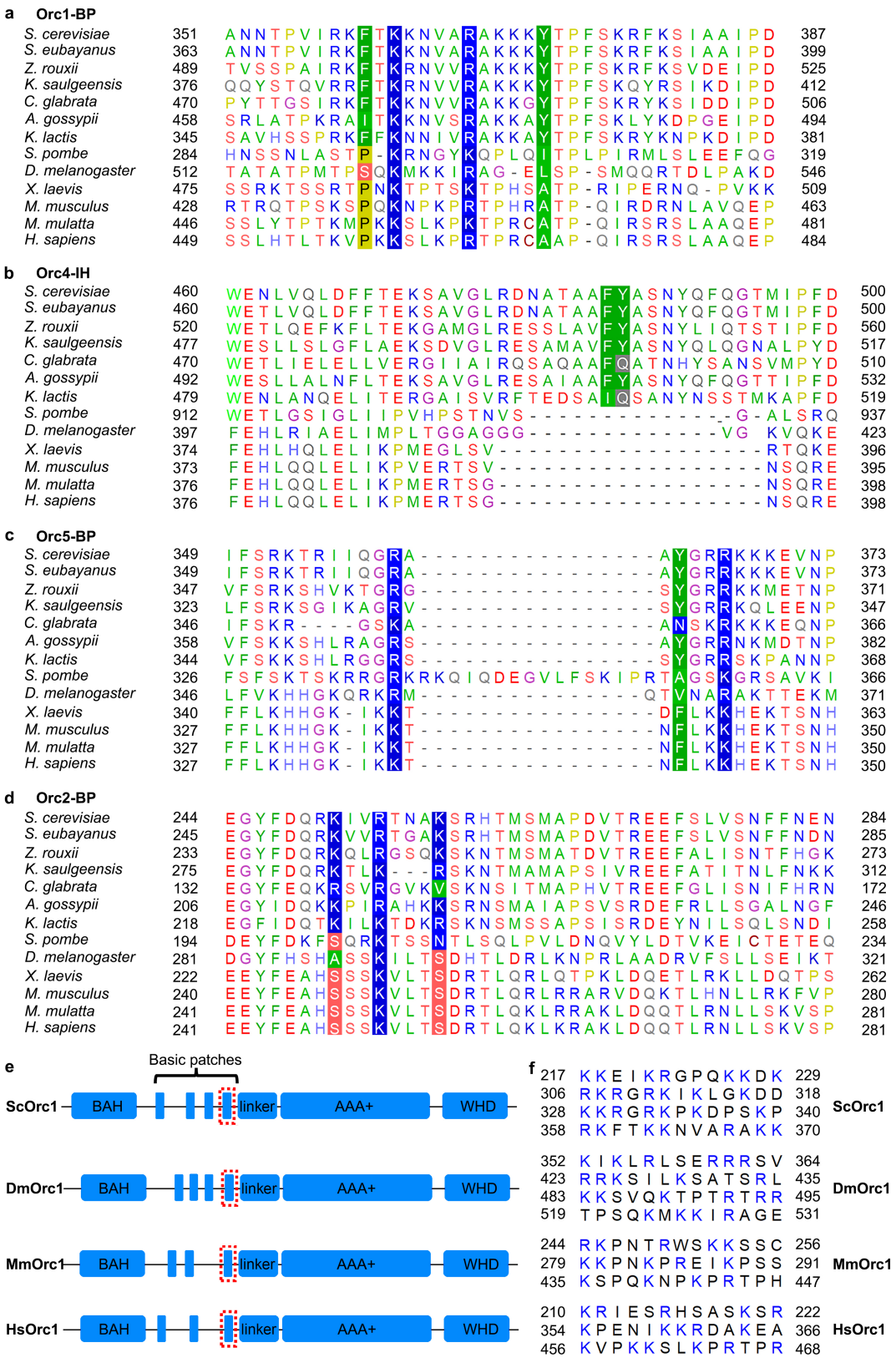
Extended Data Fig. 5 | Orc6 interacts with Orc2, Orc3 and Orc5. **a**, **b**, Overview of the interactions between Orc3, Orc2, Orc5 and Orc6. **c**, **d**, Zoomed-in views of the boxed regions in **a** and **b** to highlight their relatively hydrophobic interfaces. Selected hydrophobic residues at the interface are displayed in stick model. A short helix in the linker between Orc6-CTD and Orc6-TFIIB-B packs with two helices from Orc2-AAA+ and Orc3-WHD (**c**). A long N-terminal linker of Orc2 (upstream the AAA+ domain) wraps the TFIIB-B domain of Orc6. Note that the linker of Orc2 (Orc2-NTD loop) is traceable in the cryo-EM density map but the model could not be built at atomic level. **e**, Low-pass filtered map of the ORC–DNA complex, highlighting the interactions (indicated by the presence of extra density) between the linker sequence of Orc6 (between TFIIB domains A and B) and DNA. The N-terminal (N-ter) end of the model built for Orc6 in our map is S217. **f**, Comparison between the yeast Orc6–TFIIB-B and human ORC6–TFIIB-B. The structure of human ORC6 is from a crystallography study (PDB code 3M03)[65]. The overall structure of the yeast ORC6–TFIIB-B is quite similar to its human counterpart. **g**, Superimposition of the structure of TFIIB-DNA onto the ORC–DNA complex. The crystal structure of a human TFIIB-TBP-DNA (PDB code 1VOL)[66] was aligned using ORC6–TFIIB-B as reference. As shown, ORC6–TFIIB-B has not established extensive interactions with DNA. It is possible that further conformational change in Orc6 is required to form extensive interactions with DNA as the TFIIB does, probably at a later stage of replication licensing.



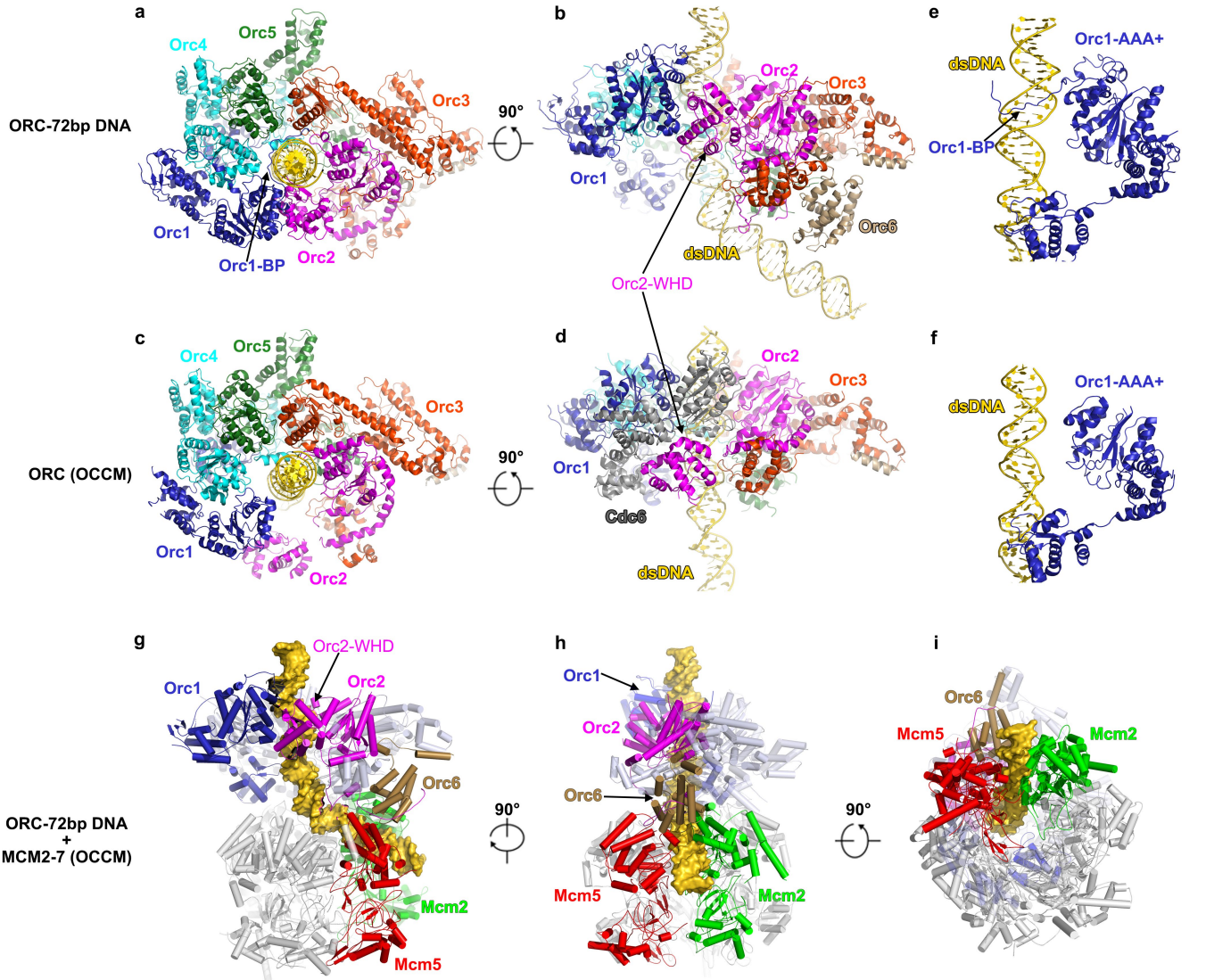
Extended Data Fig. 6 | Flexibility of the ORC complexes. **a**–**c**, Comparison of states I, II and III of the ORC–DNA complex (36 bp). Density maps of the three states are displayed in surface representation and in the Orc1–Orc2 side-view. The model of Orc2-WHD is highlighted in red cartoon. As shown, Orc2-WHD occupies different positions in the three maps. In the map of state II, the density of Orc2-WHD is relatively weak and it takes a position similar to that of the OCCM structure[30]. In the map of state III, Orc2-WHD is in a similar position as in state I, but its density is highly fragmented. Together, these indicate a floppy nature of the Orc2-WHD. **d**, Superimposition of the models of states I and II. The atomic model of state II was derived by flexible fitting of the state I model into the density map of state II. The alignment was done using Orc2 and Orc3. Compared with state I, the opening of the gap in the structure of state II is narrower. For clarity, the WHDs of Orc2 in the two states are omitted. **e**, Superimposition of the models of states I and III. The atomic model of state III was derived by flexible fitting of the state I model into the density map of state III. The alignment was done using Orc2 and Orc3. Compared with state I, the opening of the gap in the structure of state III is slightly larger. For clarity, the WHDs of Orc2 in the two states are omitted. **f**, **g**, Comparison of the density maps of states I and IV from the ORC–DNA (72 bp) dataset. A major difference between the two maps is the bending angle of the DNA. The extent of DNA bending correlates with the stability of Orc6 and Orc3 (Insertion domain of the AAA+ module). **h**, Top (left) and bottom (right) views of the cryo-EM map of the apoORC complex with the atomic model superimposed, which was derived by flexible fitting of the ORC–DNA model into the density map. ORC subunits are colour-coded. The AAA+ domain of Orc1 and the WHD of Orc2 are highly flexible, resulting in a large opening between Orc1 and Orc2, as indicated by the reduced EM densities of the corresponding regions.



Extended Data Fig. 7 | Structural comparison between the *S. cerevisiae* ORC–DNA and the *Drosophila* apoORC complexes. **a**, **b**, Side-by-side comparison of the yeast ORC–DNA and the *Drosophila* apoORC (PDB code 4XGC)[31] complexes. **a**, The yeast ORC–DNA structure is shown in cartoon representation, with Orc1-AAA+ and Orc2-WHD highlighted in marine and blue, respectively. **b**, The *Drosophila* apoORC structure is shown in cartoon representation, with Orc1-AAA+ and Orc2-WHD highlighted in magenta and red, respectively. **c**, Superimposition of the yeast ORC–DNA and the *Drosophila* apoORC structures using Orc3–Orc5 as reference. Note that the positions and orientations of Orc1-AAA+ and Orc2-WHD are markedly different in the two structures. **d**, Superimposition of Orc1 from the structures of the yeast ORC–DNA and the *Drosophila* apoORC complexes using Orc1-AAA+ as a reference. Note that the Orc2-WHDs in two structures are in totally different positions relative to Orc1-AAA+, highlighting the distinct interfaces between Orc1-AAA+ and Orc2-WHD in the two structures.



Extended Data Fig. 8 | Multiple sequence alignment of Orc1-BP, Orc4-IH, Orc5-BP and Orc2-BP. a–d, Multi-sequence alignment of Orc1 N-terminal basic patches (**a**), Orc4 insertion helixes (**b**), Orc5 WHD basic patches (**c**) and Orc2 N-terminal basic patches (**d**) from various species as indicated. **e**, Multiple basic patches found between the BAH and AAA+ domains of Orc1 from yeast to human. The criteria for basic patches are a stretch of 10 to 14 amino acids flanked by either lysine or arginine with at least three basic (K or R) residues in between and a pair of them are spaced 3–4 residues apart as found in Orc1 (R367 and K362). **f**, Sequence information of the Orc1 basic patches in **d** from various species are listed as indicated.



Extended Data Fig. 9 | Structural comparison between the yeast ORC–DNA and OCCM complexes. a, **b**, AAA+ (**a**) and side (**b**) views of the ORC–DNA complex. ORC subunits and DNA are shown in cartoon representation and colour-coded. **c**, **d**, AAA+ (**c**) and side (**d**) views of the ORC in the context of the OCCM complex. ORC subunits and DNA are shown in cartoon representation and colour-coded. The OCCM structure (PDB code 5UDB) is from previous cryo-EM work[30]. Compared with the OCCM structure, ORC subunits of Orc1 and Orc2 in the structure of ORC–DNA are more compact around the DNA. Cdc6 (grey) is included in the side view. **e**, Relative orientation of the origin DNA with Orc1 in the ORC–DNA complex. Orc1-BP is inserted into the minor groove of ACS DNA. **f**, Same as in **e**, but for the DNA and Orc1 in OCCM complex. The distance between the AAA+ domain and DNA is considerably larger, resulting in the loss of DNA contact. **g**–**i**, Superimposition of the ORC–DNA (72 bp) and OCCM (PDB code 5UDB)[30] structures. For clarity, ORC subunits from the OCCM is not shown. The Mcm2–Mcm7 subunits from the OCCM are shown in grey. Only Mcm2 and Mcm5 are labelled and colour-coded as indicated.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	ORC-72bp DNA (EMDB-6941) (PDB 5ZR1)	ORC-36bp DNA (EMDB-6942)	apoORC (EMDB-6943)
Data collection and processing			
Magnification	130,000	130,000	22,500
Voltage (kV)	300	300	300
Electron exposure (e–/Å^2)	52.3/dose weighting	52.3/dose weighting	50/dose weighting
Defocus range (μm)	1.3-2.3	1.3-2.3	1.5-2.5
Pixel size (Å)	1.052	1.052	1.32
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	427K	464K	189K
Final particle images (no.)	164K	102K	45K
Map resolution (Å)	3.0	3.6	8.2
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	2.4-4.9	3.4-4.9	5.8-10.8
Refinement			
Initial model used (PDB code)	Built based on atom model of OCCM (5UDB) or *de novo*	Built based on atom model of OCCM (5UDB) or *de novo*	
Map sharpening *B* factor (Å^2)	-94	-142	-167
Model composition			
Non-hydrogen atoms	22260	21850	
Protein and DNA residues	2479	2448	
Ligands (ATPγS and Mg^{2+})	6	6	
R.m.s. deviations			
Bond lengths (Å)	0.0069	0.0098	
Bond angles (°)	1.29	1.50	
Validation			
MolProbity score	1.49	1.79	
Clashscore	4.10	4.46	
Poor rotamers (%)	0.65	1.49	
Ramachandran plot			
Favored (%)	95.68	92.65	
Allowed (%)	4.32	7.36	
Disallowed (%)	0.00	0.29	

nature research

Corresponding author(s): Yuanliang Zhai, Ning Gao & Bik Tye

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☒	☐ Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Leginon and UCSF-Image4 were used for data collection.
Data analysis	MotionCor2, CTFFIND4, Relion2.0.3, CryoSPARC, Pymol, UCSF Chimera, COOT, PSIPRED, Phenix, MolProbity, ResMap.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Cryo-EM maps of the ApoORC, ORC-36bp, ORC-72bp complexes have been deposited in the EM databank with accession codes EMD-6943, EMD-6942, and EMD-6941, respectively. Coordinates of ORC-72bp have been deposited in the Protein Data Bank with accession code 5ZR1.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Biochemical experiments, including protein purification, GraFix and SDS-PAGE were done in biological replicates (N>10)
Data exclusions	Regarding the cryo-EM raw micrograph screening, exclusion was done based on the quality of the images and the presence of ice contamination. Regarding the particle selection, 2D and 3D classification were used and criterion is based on the quality of resulting 2D class average and 3D maps.
Replication	all attempts of replication was succesful.
Randomization	No randomization.
Blinding	No blinding.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Yeast strain ySD-ORC (a gift from John Diffley, Francis Crick Institute, UK) was used for protein purification.
Authentication	The ySD-ORC was used only for protein purification, and not further authenticated.
Mycoplasma contamination	Cell lines for purification were not tested.
Commonly misidentified lines (See ICLAC register)	No cell lines used in this study were commonly misidentified lines.