Crystal structure of the eukaryotic origin recognition complex
Franziska Bleichert¹, Michael R. Botchan² & James M. Berger¹

Initiation of cellular DNA replication is tightly controlled to sustain genomic integrity. In eukaryotes, the heterohexameric origin recognition complex (ORC) is essential for coordinating replication onset. Here we describe the crystal structure of Drosophila ORC at 3.5 Å resolution, showing that the 270 kilodalton initiator core complex comprises a two-tiered notched ring in which a collar of winged-helix domains from the Orc1–5 subunits sits atop a layer of AAA+ (ATPases associated with a variety of cellular activities) folds. Although canonical inter-AAA+ domain interactions exist between four of the six ORC subunits, unanticipated features are also evident. These include highly interdigitated domain-swapping interactions between the winged-helix folds and AAA+ modules of neighbouring protomers, and a quasi-spiral arrangement of DNA binding elements that circumnavigate an approximately 20 Å wide channel in the centre of the complex. Comparative analyses indicate that ORC encircles DNA, using its winged-helix domain face to engage the mini-chromosome maintenance 2–7 (MCM2–7) complex during replicative helicase loading; however, an observed out-of-plane rotation of more than 90° for the Orc1 AAA+ domain disrupts interactions with catalytic amino acids in Orc4, narrowing and sealing off entry into the central channel. Prima facie, our data indicate that Drosophila ORC can switch between active and autoinhibited conformations, suggesting a novel means for cell cycle and/or developmental control of ORC functions.

The faithful replication of chromosomes relies on evolutionarily conserved initiator proteins to recruit ring-shaped helicases to DNA in a cell-cycle-regulated manner (reviewed in refs 1–3). Replication initiators belong to the AAA+ protein superfamily, a large group of multi-subunit nucleotide hydrolases that function as motors or molecular switches in many cellular processes. AAA+ NTPases assemble into homo- or hetero-oligomeric complexes that actively alter the conformation or position of client macromolecules in response to ATP binding and hydrolysis.

In eukaryotes, replication initiation is promoted by the six-subunit ORC assembly⁴⁵. Five of the six subunits of ORC (Orc1–5) retain AAA+ modules⁶⁷⁷, while the sixth (Orc6) is composed of tandem cyclin-box folds similar to transcription factor IIB (TFIIB)⁸⁹. During initiation, ORC binds replication origins, recruiting another AAA+ ATPase, Cdc6, to DNA in a nucleotide-dependent manner (reviewed in refs 3, 12). The DNA-bound ORC•Cdc6 complex in turn recruits the MCM2–7 replicative helicase and its associated Cdt1 chaperone to origins, promoting the loading of MCM2–7 complexes onto DNA (reviewed in refs 3, 13).

So far, many of the molecular mechanisms by which ORC assemblies and operates have remained enigmatic. To better understand ORC function, we therefore determined the crystal structure of the Drosophila complex to 3.5 Å resolution. The structure highlights a domain-swapped organization for ORC and captures the complex in an unanticipated, autoinhibited conformation. Analysis of the structure leads to a revised model for DNA binding and proposed ORC•MCM2–7 contacts, and adds to a stepwise series of assembly and conformational intermediates that help account for how the complex acts during the early stages of replication initiation.

Crystal structure of the ORC hexamer

Sequence analyses had indicated that the Orc1–5 subunits would share a domain architecture similar to that of archaeal Orc proteins, with an AAA+–type ATPase fold fused to at least one carboxy (C)-terminal winged-helix (WH) DNA-binding domain (Fig. 1a)⁷⁸. For its part, the Orc6 C terminus has been reported to bind to ORC1–5 through a domain insertion in Orc3, leaving its TFIIB-like domain conformationally independent of the ORC core (Fig. 1a)⁴¹⁴. For crystallizing Drosophila ORC, we designed a ‘trimmed’ construct lacking the flexible amino (N)-terminal extensions of Orc1, Orc2 and Orc3 (ref. 14), and the Orc6 TFIIB region (Fig. 1a). Neither modification interfered with ORC assembly, nor did they affect the overall architecture of ORC (Extended Data Fig. 1a–c). This ORC core (referred to as ORC hereafter) crystallized in space group P2221 with one Orc1–6 heterohexamer per asymmetric unit.

The structure was phased by single-wavelength anomalous dispersion and refined to 3.5 Å with Rwork/Rfree values of 0.22/0.26 (Extended Data Fig. 2a–c and Extended Data Table 1).

AAA+ and WH domains interlock within the ORC body

The ORC structure shows that the complex forms a lopsided, two-tiered ring with a cashew-shaped protuberance off of one edge (Fig. 1b, c and Supplementary Video 1). Orc1–5 comprise the ring body, which bears a prominent central channel, while a large domain insertion in Orc3 forms a bi-lobed, β-helical extension that engages a short β-helix formed at the Orc6 C terminus. In contrast to models based on prior electron microscopy (EM) reconstructions¹⁴¹⁵, the AAA+ subunits are arranged in the order of Orc1–Orc4–Orc5–Orc3–Orc2, thus revising the placement for Orc2 and Orc3 within the pentameric ORC ring (Fig. 1b).

In the structure, Orc1–5 each comprises one AAA+ fold, followed by a single C-terminal WH domain (Extended Data Fig. 3). The AAA+ and WH regions coassociate, but are segregated between the two ring tiers (Fig. 1b, c and Supplementary Video 1). Interestingly, the collar of WH domains is rotationally offset from the AAA+ domains, leading to a domain-swapped organization wherein, apart from Orc2, the WH domain of one subunit packs against the AAA+ domain of its adjoining partner (Fig. 1c). Domain swapping is facilitated by long linkages between the AAA+ and WH modules, a region known to be conformationally

¹Department of Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, USA. ²Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, California 94720, USA.

©2015 Macmillan Publishers Limited. All rights reserved
**DNA binding elements in the WH domain collar**

In archaean Orc homologues, the WH element recognizes origin sequences by using a helix–turn–helix motif and a β-hairpin ‘wing’ to interact with the adjacent major and minor grooves of double-stranded DNA\(^{16,17,20,21}\) (Fig. 2a). The conservation between archaean Orcs and eukaryotic Orc1–5 proteins, we anticipated that the WH domains of ORC would bind DNA in a similar manner. However, the second \(\alpha\)-helix of the helix–turn–helix in the WH domain (corresponding to the DNA ‘recognition helix’) is buried against the AAA + tier in all subunits but Orc2 (Fig. 2b). This arrangement leaves the β-hairpin wings of Orc1, Orc4, Orc5 and Orc3 solvent exposed, which in turn co-localize to form a portion of the interior surface within the central channel in the ORC body (Fig. 2b, c). Given the extensive contacts between the WH and AAA + tiers, sequestration of the recognition helices seems necessary to maintain ORC integrity. Thus, certain aspects of DNA recognition by the WH domains of ORC probably differ from the approach used by archaean Orcs.

**Quasi-canonical AAA + domain packing within ORC**

AAA + ATPases tend to oligomerize into closed-ring, open-ring or helical assemblies\(^{6,22}\). On the basis of phylogenetic, biochemical and EM data, ORC has been proposed to follow a similar trend\(^{7,8,14,23,24}\). The ORC structure shows that, of the various AAA + domain interactions, the one between Orc4 and Orc5 is most like that seen in typical AAA + systems, whereby the two subunits associate to form a bipartite nucleotide-binding site at the inter-protomer interface (Fig. 3a and Extended Data Fig. 5a). Metazoan Orc4 manifests conserved Walker A and B ATPase motifs (GKT and D(D/E), respectively), which in related AAA + ATPases typically contact both \(\beta\)- and \(\gamma\)-phosphates of bound nucleotide and a catalytically important Mg\(^{2+}\) ion. Interestingly, Orc5 donates a well-conserved ‘arginine-finger’ (Arg144) into the Orc4 active site (Extended Data Fig. 5a, b). Moreover, although Orc4 does not possess a typical ‘Sensor II’ arginine (an amino-acid motif in AAA + ATPases that often aids hydrolysis), its Walker A region does retain a conserved basic amino acid (Arg58) (Extended Data Fig. 5c) that occupies an analogous position (Extended Data Fig. 5a). Overall, the observed structural organization and conservation of catalytically important residues in the Orc4•Orc5 interface raises the possibility that, in some organisms, Orc4 might be able to support some level of ATP turnover, conformationally modulate ORC function in response to ATP, or require nucleotide to promote ORC stability (as has been observed for the human complex\(^{25–27}\)). Any such activity probably varies between species, however, as *Saccharomyces cerevisiae* Orc4 neither contains a conserved Walker A motif nor has been found to bind/hydrolyse ATP\(^{37}\).

Other AAA +/AAA + interactions within ORC deviate to a greater or lesser extent from the canonical packing arrangements exemplified by Orc4•Orc5. For example, the Orc5•Orc3 interface is relatively open, with fewer contacts between AAA + modules (Fig. 3a and Extended Data Fig. 5d). Although Orc5 possesses a canonical Walker A motif in most eukaryotes, the same is not true for its Walker B motif, nor is there a candidate Orc5 Sensor II residue or an Orc3 arginine finger. These structural features are consistent with biochemical observations showing that Orc5 binds ATP but lacks ATP hydrolysis activity\(^{26–30}\). By comparison, residues for nucleotide binding and hydrolysis are not conserved at the Orc3•Orc2 AAA +/AAA + interface; hence, the Orc2...
and Orc3 AAA+ domains and their interactions appear to play a predominantly structural part in ORC assembly.

Consistent with its blend of typical and atypical local AAA+/AAA+ interactions, the global organization between the ORC AAA+ domains is also mixed. Of AAA+ proteins, ORC is most closely related phylogenetically to other replication initiators, including bacterial DnaA, and to DNA polymerase clamp loaders. In comparing the ORC ATPase assembly with these systems, it can be seen that the Orc4-Orc5 and Orc3-Orc2 AAA+ folds are organized similarly to protomers in ATP-assembled DnaA (Fig. 3a). By contrast, the AAA+/AAA+ organization at the Orc5•Orc3 interface is relatively open. Consequently, the clean helical symmetry observed in ATP-assembled DnaA oligomers is broken within the Orc5•Orc3 junction of ORC, a configuration that creates a cracked ring-like architecture reminiscent of clamp loaders (Fig. 3b–d). Despite the somewhat more planar arrangement of ATPase folds in ORC, the Orc2–5 initiator specific motifs (or 'ISMs', an α-helical modification that both distinguishes AAA+ family replication initiators and binds DNA directly) nonetheless cluster together, forming a shallow, quasi-spiral shaped set of 'threads' that line the interior of the ~20 Å wide central ORC channel (Fig. 3b). Thus, the structural features of ORC are a hybrid of both clamp loader and prokaryotic initiator systems.

An unanticipated Orc1 conformation

In addition to Orc4 and Orc5, one other ORC subunit also known to bind ATP is Orc1 (refs 27, 30). Indeed, Orc1 serves as the major source of ATPase activity in ORC and requires a conserved arginine residue from Orc4 for catalytic activity. Given this activity, we expected, as with Orc2–Orc5, to see relatively canonical AAA+ interactions between Orc1 and Orc4 in the structure. Surprisingly, the AAA+ domain of Orc1 is completely disengaged from Orc4, owing to a buckling of the linker at a single region between the Orc1 ATPase and WH folds (residues 819–821) that gives rise to a ~105° out-of-plane rotation from the Orc2–5 AAA+ domains (Fig. 4a). Although this Orc1 AAA+/WH domain juxtaposition is unique compared with other subunits in the structure (Extended Data Fig. 3), it is similar to one of the conformations adopted by an ADP-bound archaeal Orc homologue, Aeropyrum pernix Orc2 (ref. 18) (Extended Data Fig. 6a, b). Notably, the movement of Orc1 does not considerably affect the relative arrangement of its two AAA+ subdomains, which is maintained as in Orc3–5 and archaeal Orcs (Extended Data Fig. 6c–e). The en bloc reorientation of Orc1 appears stabilized by contacts between the Orc1 AAA+ domain and the WH domains of Orc1, Orc2 and Orc3, together burying a total of 4,256 Å² at the interface (Fig. 4b and Extended Data Fig. 6f–h).

A consequence of the disposition of Orc1 within the complex is that its nucleotide binding cleft resides ~40 Å away from the arginine finger of Orc4. Importantly, comparison of the crystal structure with a prior three-dimensional EM reconstruction of ATPγS-bound Drosophila ORC shows excellent agreement between the two models (Fig. 4c and Supplementary Video 2), indicating that the Orc1 conformation in the crystal corresponds to the predominant state of the complex in solution. Moreover, co-crystallization of ORC with the ATP analogue ATPγS, while showing clear density for nucleotide binding to the Orc1, Orc4 and Orc5 AAA+ folds (Extended Data Fig. 7a–c), recapitulates the configuration seen in apo-ORC. Together, these data indicate that Drosophila Orc1 must undergo a large structural change to support ATPase activity, but that ATP binding is itself insufficient to drive such a rearrangement in most ORC particles.

Implications of the structure for DNA binding by ORC

In the ORC structure, the central channel within the body of the complex is formed by both the Orc2–5 ISMs and the β-hairpin wings of

Figure 3 | AAA+/AAA+ domain interactions in ORC. a, Pairwise AAA+ interactions in the Orc2–Orc5 oligomer compared with canonical AAA+ interactions in an ATP-bound DnaA dimer (PDB accession number 2HCB). Superpositions were performed using the AAA+ domain of the left-most protomer of a given ORC subunit pair and DnaA homodimer. b–d, Comparison of AAA+ domain assemblies (to scale) for Orc2–5 (b), ATP-assembled DnaA (PDB accession number 2HCB) (c) and replication factor C (RFC, PDB accession number 1SXJ) (d). Upper panels depict side views, with the purple AAA+ domains of DnaA and RFC co-oriented as per the Orc4 AAA+ region. Lower panels show views down the central channel of ORC and RFC, or slightly offset from the helical axis of DnaA. The initiator specific motifs (ISMs, red) of Orc2–5 and DnaA are highlighted (RFC lacks this element). Only the AAA+ domains of ORC, DnaA and RFC are shown; the RFC-A subunit is omitted. Colouring for ORC is maintained as in a, with sequential subunits in DnaA and RFC coloured accordingly.
Orc1 and Orc3–5. In archael Orcs, these two elements both bind to duplex DNA\textsuperscript{16,17}. To investigate if the ORC central channel could accommodate any of the DNA interactions typified by archael Orcs, we superposed the DNA-bound crystal structure of Sulfolobus solfataricus Orc1–1 (ref. 16) onto Orc4 (after Orc1, Orc4 is most closely related to archael Orcs). Notably, superpositioning of the AAA\textsuperscript{+} domains of the two proteins (Fig. 5a) resulted in a placement for DNA that aligns the duplex coaxially with the central ORC channel (Fig. 5b). Inspection of the resultant ORC•DNA model not only reveals that the quasi-spiral formed by the Orc2–5 ISMs approximates that of the docked duplex, but that the β-hairpin wings of the Orc1 and Orc3–5 WH domains also reside in a position where they can access the nucleic-acid segment (Fig. 5b).

The superpositioning between DNA-bound archael Orc1–1 and the DNA-free Drosophila ORC imaged here has important implications for understanding how the eukaryotic initiator engages origin regions. One is that ORC probably binds DNA using a mechanism similar to that of sliding clamp loaders, which encircle primer-template junctions\textsuperscript{34,35}. This binding mode is congruent with a recent proposal based on EM analysis of an ORC•Gdc6•Cd1•MCM2–7 complex, which posits that ORC helps align the ring of an MCM2–7 complex around DNA\textsuperscript{24}. The apparent sequence-specificity of DNA binding by \textit{S. cerevisiae} ORC\textsuperscript{3} (in contrast to metazoan ORC\textsuperscript{36,37}) probably results from specific interactions between amino-acid side chains in the ORC channel and nucleotide bases in specific autonomously replicating sequence elements.

At present it is unclear from the structure why \textit{Drosophila} ORC prefers to bind negatively supercoiled DNA over linear segments\textsuperscript{36}, a preference also reported for \textit{Schizosaccharomyces pombe} ORC\textsuperscript{38}. Archael Orcs are known to underwind DNA upon binding\textsuperscript{16,17}, suggesting that the eukaryotic ORC AAA\textsuperscript{+} and WH domains may cooperate to do likewise; topological changes in DNA structure induced by \textit{S. pombe} and human ORC are consistent with such an interpretation\textsuperscript{39}. Alternatively, if ORC were to bind B-form DNA, conformational transitions within the ORC body and its associated DNA-binding elements would be required to accommodate the DNA duplex. In this instance, specificity for negatively supercoiled substrates could arise from the relative positioning of the TFIIB-like DNA-binding elements in Orc6 (refs 10, 11, 40) and the DNA-recognition elements in the central ORC channel.

**ORC activation and implications for MCM2–7 loading**

Modelling indicates that the DNA binding elements in the central ORC channel can encircle a DNA duplex; however, certain features of the complex would appear to preclude \textit{Drosophila} ORC from doing so in the state seen both here and by EM. For example, passage of the DNA through the entirety of the central channel is prevented by a constriction formed by both the Orc2 WH domain and the Orc1 AAA\textsuperscript{+} fold (Fig. 1b). Similarly, although the ATPase region of the ORC ring is cracked open, the Orc2 WH domain and the Orc1 AAA\textsuperscript{+} fold occlude...
this crack, thereby preventing the lateral entry of DNA into the central channel from the side of the complex (Figs 1b and 5c). This observation probably helps explain the weak effect that nucleotide has on DNA binding by *Drosophila* ORC purified from fly embryos or recombinant sources. It is interesting to note that by preventing DNA binding, the placement of the Orc1 ATPase region and the Orc2 WH fold also blocks the known interaction site for another critical component of replication initiation, Cdc6 (refs 23, 24). Overall, the simplest interpretation of the ORC conformation imaged here is that it corresponds to a naturally autoinhibited form of the complex, and that in some organisms, only a fraction of the total ORC pool that can be obtained from asynchronously dividing cells may be capable of productively altering its interactions with DNA in response to ATP.

If *Drosophila* ORC first assembles into an inactive form, then what manner of transition might push the complex into a new state in which its ATPase region is now competent to bind DNA (or Cdc6)? Insights into a simple structural rearrangement that could support such a switch can be gleaned from what is known about archetypal AAA+ ATPase organization. Using the Orc4•Orc5 interaction seen in the crystal structure as a template, we generated a model for the expected arrangement of a functional Orc1•Orc4 AAA+ ATPase centre by swivelling the Orc1 AAA+ fold around a single hinge point in the linker region just before its WH domain (Supplementary Video 1). The resultant model not only restores expected AAA+ interactions between the Orc1 active site and the Orc4 arginine finger (Fig. 6A and Extended Data Fig. 7d), but also both removes the Orc1-mediated blockage of the putative path for DNA in the central channel and co-aligns the Orc1 ISM with the ISM helix formed by Orc2–Orc5 (Fig. 6b and Supplementary Video 1). Docking of the rearranged model into the cryo-EM densities of *S. cerevisiae* ORC shows a reasonable fit for *Drosophila* ORC containing the repositioned Orc1 AAA+ domain, and further reveals that a region of EM density, which extends from the centre of the ORC body, actually corresponds to DNA (Extended Data Fig. 8a, b). In accord with a two-state model, the EM density for *S. cerevisiae* ORC cannot accommodate the AAA+ domain of Orc1 in the state imaged crystallographically (Extended Data Fig. 8a), nor can the *Drosophila* ATP-S-ORC EM volume account for a remodelled Orc1 conformation for ORC (Extended Data Fig. 8c and Supplementary Video 2).

Collectively, the structure and analysis presented here helps to define a framework for understanding how *Drosophila* ORC interfaces with its partner proteins and DNA during the initial stages of MCM2–7 loading (Fig. 6c). In this scheme, ORC would start off in an ATP-bound but autoinhibited form—either dissociated from chromatin or in a chromatin-bound state via secondary binding sites/partners—that is restricted in its ability to either bind DNA within its central channel or bind Cdc6 to its ring. Conversion of this state into an activated configuration would involve the *en bloc* movement of the Orc1 ATPase domain (Supplementary Video 1), allowing Orc1 to engage the arginine finger of Orc4, and unlatching the Orc2 WH domain to open a gap in the Orc1–5 ring. Once open, DNA would bind to the ISM and β-hairpin elements in the central ORC channel, after which Cdc6 would dock into the Orc1/Orc2 gap, trapping DNA within the centre of the complex. One prediction of this model is that Cdc6 should bind to ORC using its ATPase centre to engage an arginine finger in Orc1 (*Drosophila* residue Arg734). After formation of a ternary ORC•DNA•Cdc6 complex, the WH domains of Cdc6 and Orc2 would be expected to engage the AAA+ folds of Orc1 and Cdc6, respectively, creating a circuit of WH domains (and their associated β-hairpin elements) to help lock the complex into place. Interestingly, these structural findings and analyses indicate that, rather than using its AAA+ domains to bind MCM2–7 (ref. 24), which requires an inverted order of ATPase site-arginine finger interactions around the ORC ring (Extended Data Fig. 9a, b), ORC probably uses its WH domain collar instead (Extended Data Fig. 9c).

A major question still remaining is what event might trigger ORC rearrangement, or why ORC should exist in an autoinhibited state. On the basis of our EM and co-crystallization data, ATP binding alone is incapable of efficiently driving such a transition for most of the particles.
in a purified ORC preparation. Interestingly, phosphatase treatment of *Drosophila* ORC stimulates DNA binding⁴¹, suggesting that removal of one or more post-translational marks might help convert ORC to an active form. Moreover, metazoan ORC associates with chromatin in a cell-cycle-dependent manner (reviewed in refs 1, 42), and the targeting of ORC to chromosomes in metazoans (and in fission yeast) is known to require protein–DNA contacts distinct from those in budding yeast, such as the TFIIH-homology domain in metazoan Orc6 (refs 10, 40) or the AT-hook in *S. pombe* Orc4 (ref. 43). The action of these elements, together with the recognition of nucleosomes by the N-terminal BAH domain of Orc1 (refs 44–46), suggest that the formation of nucleotide-dependent contacts between the ATPase region of ORC and DNA may take place after the formation of initial ORC/chromatin encounters. Moreover, metazoans need to stockpile ORC in oocytes, yet keep this pool of ORC from prematurely initiating DNA replication before fertilization; hence, ORC may be maternally stored in an inactive form during oogenesis. In these contexts, autoinhibition could provide a novel mechanism for regulating ORC’s productive association with DNA in a cell-cycle-dependent and/or developmental stage-specific manner.

**Online Content** Methods, along with any additional Extended Data display items, can be found in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 October 2014; accepted 16 January 2015.

Published online 11 March 2015.

1. Siddiqui, K., On, K. F. & Diffley, J. F. Regulating DNA replication in eukarya. Cold Spring Harb. Perspect. Biol. 5, a012930 (2013).
2. O’Donnell, M., Langston, L. & Stillman, B. ATP-dependent cooperative binding of ORC to origin DNA. Nature Struct. Mol. Biol. 10, 965–971 (2003).
3. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA helicases in prokaryotes and eukaryotes. J. Mol. Evol. 56, 357–435 (2003).
4. Balasov, M., Huijbregts, R. P. & Chesnokov, I. Role of the Orc6 protein in origin recognition and replication in *Drosophila melanogaster*. Mol. Cell. Biol. 34, 8414–8423 (2014).
5. Schouen, C. R. et al. Multiple mechanisms contribute to Schizosaccharomyces pombe origin recognition complex DNA interactions. J. Biol. Chem. 283, 30216–30224 (2008).
6. Makino, D. L., Kuriyama, K., Botchan, M. & O'Donnell, M. A structural role for Orc1p in DNA replication in archaea and eukarya. Cold Spring Harb. Perspect. Biol. 5, a010108 (2013).
7. Iyer, L. M., Leipe, D. D., Koonin, E. V. & Aravind, L. Evolutionary history and higher order classification of AAA + ATPases. J. Struct. Biol. 146, 11–34 (2004).
8. Chernev, I. N., Chernev, N. O. & Botchan, M. A. Cytochrome function of *Drosophila* ORC6 protein resides in a domain distinct from its replication activity. Proc. Natl Acad. Sci. USA 100, 9150–9155 (2003).
9. Liu, S. et al. Structural analysis of human ORC protein reveals a homology with transcription factor TFIIB. Proc. Natl Acad. Sci. USA 108, 7373–7378 (2011).
10. Costa, A., Hood, I. V. & Berger, J. M. Mechanisms for initiating cellular DNA replication. Annu. Rev. Biochem. 82, 25–54 (2013).
11. Yardimci, H. & Walker, J. C. Pre-replication complex formation: a molecular double take? Nature Struct. Mol. Biol. 21, 20–25 (2014).
12. Bleichert, F. et al. A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc1 impedes origin recognition complex formation. eLife 2, e00882 (2013).
13. Shen, Z. et al. The architecture of the DNA replication origin recognition complex in *Saccharomyces cerevisiae*. Proc. Natl Acad. Sci. USA 105, 10326–10331 (2008).
14. Siegerer, 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).
15. Gaudier, M., Schwurth, B. S., Westcott, S. L. & Wigley, D. B. Structural basis of DNA replication origin recognition by an Orc protein. Science 317, 1213–1216 (2007).
16. Singleton, M. R. et al. Conformational changes induced by nucleic acid binding in *Saccharomyces cerevisiae* Orc4p. J. Mol. Biol. 343, 547–557 (2004).
17. Liu, J. et al. Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. Mol. Cell 6, 637–648 (2000).
18. Dueber, E. C. et al. Molecular determinants of origin discrimination by Orc1 initiator protein. Nucleic Acids Res. 39, 3621–3631 (2011).
19. Robinson, N. P. et al. Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. Cell 116, 25–38 (2004).
20. Zeidler, J. P., Mott, M. L. & Berger, J. M. Structural basis for ATP-dependent DNA assembly and replication-origin remodeling. Nature Struct. Mol. Biol. 13, 676–683 (2006).
21. Sun, J. et al. Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. Structure 20, 534–544 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the beamline scientists at beamlines 8.3.1 of the Advanced Light Source (Berkeley Lawrence National Laboratory), 23-ID-B of the Advanced Photon Source (Argonne National Laboratory) and X25 of the National Synchrotron Light Source (Brookhaven National Laboratory) for technical support with X-ray diffraction data collection, and the Nogales laboratory (University of California, Berkeley) for EM support. We also thank A. Fisher and Y. Li for assistance in insect cell culture. This work was supported by the National Institutes of Health (GM071747 to J.M.B. and CA37-30490 to M.R.B.) and by a fellowship from the UC Berkeley Miller Institute for Basic Research in Science (to F.B.).

**Author Contributions** F.B. performed all biochemical and crystallization experiments, collected X-ray diffraction data and determined the structure with guidance from J.M.B. All authors interpreted and discussed results, and wrote the manuscript.

**Funding Information** Atomic coordinates and structure factors have been deposited in the RSCB Protein Data Bank under accession code 4XGC. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M.B. (jberge29@jhmi.edu) or M.R.B. (mbotchan@berkeley.edu).
MULTISPECIES ATOM-derivative crystals were obtained by incubating crystals in cryo-protecting solution containing 100 μM ethyl mercapturic phosphate for 3 h; data from these soaks were used to identify metal-binding sites for confirming amino-acid registers but were not used for phasing.

Crystallization and crystal harvesting for ATPS-bound ORC was performed as described for apo-ORC but with the following modifications: (1) ORC was dialyzed into crystallization buffer containing MgCl₂ (50 mM Tris-HCl (pH 7.8), 200 mM KCl, 5% glycerol, 0.5 mM TCEP, 5 mM MgCl₂); (2) before crystallization, 1 mM ATPS was added to ORC to a final concentration of 1 mM; and (3) for harvesting, 0.5 mM ATPS was added to the cryo-protectant to prevent dissociation of the nucleotide.

Data collection and structure determination. The diffraction quality of individual crystals varied greatly, necessitating the screening of many hundreds of different crystals to identify acceptable targets for data collection. Crystal screening was performed at beamlines 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory, X25 at the National Synchrotron Light Source at Brookhaven National Laboratory and 23-ID-B at the Advanced Photon Source at Argonne National Laboratory. Native diffraction data sets (λ = 1.0332 Å) as well as single-wavelength anomalous dispersion (SAD) data sets for gadolinium (λ = 1.71083 Å) and sulphur (λ = 2.7121 Å, see below) were collected at the Advanced Photon Source 23-ID-B equipped with a MAR CCD detector. Although complete data sets were typically obtained by exposing multiple regions within a single crystal using the ‘vector data collection’ option to minimize radiation damage, the best native data set was collected at a single site from a crystal harvested at 4°C. Data sets of ethyl mercapturic phosphate-derivatized crystals were collected at the National Synchrotron Light Source X25 (λ = 1.060 Å) on a Pilatus 6M detector.

Diffraction data were processed with XDS [50, 51]. Merging with the program AIMLESS [52] (Extended Data Table 1) indicated that the crystals belonged to space group P2₁2₁2₁, with unit cell dimension of a = 145.5 Å, b = 259.0 Å and c = 257.0 Å for the best native crystal. The Gd-SAD data set was obtained by merging data from four different crystals. Despite slight non-isomorphism between crystals, merging data from multiple crystals substantially improved the anomalous signal, phases and electron-density interpretability compared with data sets collected from single crystals.

For initial phasing by SAD, gadolinium sites were identified with SHELXD [53]. The strongest sites were then used as input into Phaser [54] as implemented in PHENIX [55] to find additional sites and to obtain initial phases to ~4 Å. Maximum likelihood density modification with RESOLVE [56] was used to break the phase ambiguity and to improve electron-density maps. At this point, experimental phases were next applied (again using PHENIX) to the native data set, which was of better quality than the gadolinium derivative. The resulting electron-density maps (at 4 Å resolution) allowed identification of all five AAA+ domains and four of the five WH domains, and revealed density for bulky side chains as expected for this resolution. In parallel to data processing with SHELXD/PHASER/RESOLVE, experimental phases and density maps were also calculated with SHARP and improved by solvent flattening in SOLOMON [57, 58]. Although slightly less featured, the SOLOMON electron-density maps were overall very similar to those obtained from RESOLVE but additionally revealed clear protein density in some regions that were poorly defined in RESOLVE density maps. Thus, while model building was performed predominantly into RESOLVE density maps, SOLOMON density maps were used as an additional guide to trace the main chain of the model. Phases were gradually improved by iterative cycles of model building, density modification with phase combination of experimental and model phases in RESOLVE, and phase extension to 3.7 Å.

Model building was initiated by automated searches using MOLREP [59, 60] and by manual docking using UCSF Chimera [61, 62] to place the AAA+ and WH domains of archaeal Orc1/Cdc6 (PDB accession numbers: 2QBY [63]; IFNN [64]) and/or homology models for Drosophila ORC subunits (as generated by Phyre2 (ref. 62)) into cry-protectant density maps. These structures were valuable reference points and facilitated tracing of most of the main chain; insights into likely subunit positions, allowed assignment of specific subunits to map density regions. Using COOT [65], a nearly complete model of ORC was manually built de novo into phase-combined and B-factor-sharpened RESOLVE density maps, guided by the topology of archaeal Orc1/Cdc6 AAA+ and WH domains, as well as by secondary structure prediction and multiple sequence alignments. The initial model was improved by iterative rounds of refinement in PHENIX (real-space, individual sxy, individual atomic displacement parameter refinement using secondary structure and (in early stages of building) experimental phase restraints, as well as stereochemistry and atomic displacement parameter weight optimization; subsequent rounds of model rebuilding were performed using COOT. During the course of refinement, a slightly higher-resolution and more complete native data set (to 3.5 Å) was obtained and used for
the final rounds of refinement, which also included refinement of TLS parameters. The final model contains the AAA+ and WH domains of Orc1 to Orc5, the Orc3 domain insertion and the conserved C-terminal helix of Orc6; an N-terminal region of Orc2 (preceeding the AAA+ domain) was built as a poly-alanine model, since the amino-acid register for this region could not be assigned unambiguously. The final model was validated with MOLPROBITY\textsuperscript{44} and has excellent geometry (MolProbity score 1.88), with no Ramachandran outliers and only a small fraction (1.9\%) of rotamer outliers (Extended Data Table 1).

During the course of model building, several approaches were used to validate the sequence register of the various ORC chains. These included (1) using Hg-binding sites in ethyl mercury phosphate-derivatized crystals as fiducials for cysteines and histidines, (2) using sulphur sites in native S-SAD data sets to verify the location of a subset of cysteines and methionines (the weak signal present in these data precluded the use of this information for phasing) and (3) conducting Orc3–Or6 crossing experiments to confirm the register of the Orc6 C-terminal helix (Extended Data Fig. 4h). Hg and S sites were identified from log-likelihood-gradient maps calculated using the MR-SAD option in Phaser\textsuperscript{45}. Additionally, MR-SAD for sulphur sites also revealed the position of two ions with anomalous scattering properties at the wavelength of data collection ($\lambda = 1.7712$ Å). These ions showed clear density in experimental and $2F_{o} - F_{c}$ maps and were interpreted as a chloride ion in the P-loop of Orc5 and as a potassium ion bound to Orc2.

Once a satisfactory apo-ORC model was obtained, it was used as a search model for molecular replacement (using PHENIX–PHASER) to phase data collected from ORC co-crystallized with ATP5. The resulting solution (Z score = 86.7, log-likelihood-gain = 7747) revealed clear (more than 2-to-4\%) positive difference density in the nucleotide binding clefts of Orc1, Orc4 and Orc5 that could accommodate ATP5 (Extended Data Fig. 7a–c). Since the resolution of these crystals was limited to 4 Å, and since only small structural changes were observed throughout the remainder of ORC, we refrained from building and refining a model against this data set.

**Structure analysis.** Structural superpositions and docking into EM maps were performed using UCSF Chimera\textsuperscript{64,65}. Buried surface area at domain/subunit interfaces was calculated with PyMOL (PyMOL Molecular Graphics System, version 1.7.0.0, Schrödinger). Multiple protein sequence alignments were performed with MAFFT\textsuperscript{66,67} and conservation scores were calculated and mapped onto the structure with Consurf\textsuperscript{68}. Sequence logos were generated with WEBLOGO\textsuperscript{69}. Figures were rendered using both PyMOL (The PyMOL Molecular Graphics System, version 1.7.0.0, Schrödinger) and UCSF Chimera\textsuperscript{64,65}.

**Expression and purification of ORC-5 and Orc6 for binding and crosslinking studies.** ORC containing subunits Orc1–5 was expressed in High5 cells using a His and MBP tags were removed from ORC1–5, whereas for Orc6 proteins, as well as untagged ORC1–5, were dialysed overnight into 50 mM PIPES (pH 7.2), 300 mM KCl, 5% glycerol, 1 mM DTT and 0.1 mg ml\textsuperscript{-1} CTDA236E, and different concentrations of ORC1–5 (ranging from 122 pM to 300 mM ATPases were performed with either wild-type or an A236C Orc6-CTD mutant and untagged Orc1–5 using the homobifunctional maleimide crosslinker BMOE (bis(maleimido)ethane; Thermo Scientific) under non-reducing conditions. Ala236 of Orc6 was chosen for mutation because inspection of the structure revealed that the insert of Orc3 contained a nearby native cysteine (Cys372), which we reasoned could form a crosslink with an A236C Orc6 mutant if our build register was correct (Extended Data Fig. 4e). Before crosslinking, both wild-type Orc6-CTD and Orc6-CTD A236C proteins, as well as untagged ORC1–5, were dialysed overnight into 50 mM PIPES (pH 7.5), 300 mM KCl and 10% glycerol to remove reducing agents. Binding of the mutant-type Orc6 to ORC1–5 was performed in 50 μl reactions containing 4 μM of each protein in 50 mM PIPES (pH 7.5), 300 mM KCl and 10% glycerol for 30 min. BMOE was then added to reactions to a final concentration of 0.2 mM for 5 min, after which crosslinking reactions were stopped by adding SDS–polyacrylamide gel electrophoresis (SDS–PAGE) sample-loading dye containing β-mercaptoethanol (100 mM final concentration). Stopped reactions were analysed by SDS–PAGE and Coomassie staining.

To ensure that Orc6 still bound to ORC1–5 under non-reducing conditions, we performed binding reactions using ORC1–5 containing MBP-tagged Orc4 and either wild-type or A236C Orc6-CTD under non-reducing conditions as described for the crosslinking experiments, but instead subjected the reactions to pull-downs using amylase beads (New England Biolabs). Beads were washed three times with 1 ml of 2% (w/v) ammonium acetate before crosslinking experiments, but instead subjected the reactions to pull-downs using amylase beads (New England Biolabs). Beads were washed three times with 1 ml of 2% (w/v) ammonium acetate.

**Electron microscopy.** Four microticles of a 30 nM Drosophila ORC solution containing Tertiary crystals of Orc1, Orc2 and Orc3 subunits (in 20 mM Tris (pH 7.8), 125 mM potassium glutamate, 5 mM MgCl\textsubscript{2}, 1 mM ATP5) were spotted onto glow-discharged, continuous-carbon film EM grids, stained with four drops of 2% uranyl formate for 10 s each and blotted. Grids were imaged in a Tecnai T12 BIOWIN transmission electron microscope operated at 120 kV with a LaB\textsubscript{6} cathode as electron source. Data collection, image processing and two-dimensional classification were performed as described previously\textsuperscript{42,43}, and two-dimensional class averages were compared with class averages of the full-length ATP8–ORC data set reported in ref. 14 (Extended Data Fig. 1c).

49. Parks, T. D. et al. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. Anal. Biochem. 216, 413–417 (1994).

50. Kabsch, W. Xds. Acta Crystallogr. D 66, 125–132 (2010).

51. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D 66, 133–144 (2010).

52. Winn, M. D. Overview of the CCP4 suite and current developments. Acta Crystallogr. D 67, 235–242 (2011).

53. Sheldrick, G. M. Experimental phasing with SHELXLC/D/E: combining chain tracing with density modification. Acta Crystallogr. D 66, 479–488 (2010).

54. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).

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

56. Terwilliger, T. C. Maximum-likelihood density modification. Acta Crystallogr. D 56, 965–972 (2000).

57. Vornhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. Methods Mol. Biol. 364, 215–230 (2007).

58. Bragin, G. et al. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr. D 59, 2023–2030 (2003).

59. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. Acta Crystallogr. D 66, 22–25 (2010).

©2015 Macmillan Publishers Limited. All rights reserved.
60. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157**, 281–287 (2007).
61. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
62. Kelley, L. A. & Sternberg, M. J. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363–371 (2009).
63. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
64. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
65. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
66. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* **9**, 286–298 (2008).
67. Ashkenazy, H. et al. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010).
68. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
Extended Data Figure 1 | Deletion of variable N-terminal extensions in Orc1, Orc2 and Orc3 alters neither ORC stability nor overall ORC architecture. a, Gel-filtration chromatography trace of the ORC core used for crystallography, with (b) SDS–PAGE of respective ORC peak fractions from a, indicates the formation of a stable hexameric complex. c, Full-length ORC and ORC containing N-terminal truncations display similar structural features in two-dimensional EM class averages. Both complexes were imaged by negative-stain EM in the presence of ATPγS. Note that although class averages from ORC with truncated Orc1–3 subunits contain full-length Orc6, Orc6 is not visible owing to its flexible nature. Class averages for full-length ORC are derived from a data set used in ref. 14.
Extended Data Figure 2 | Experimental electron density contoured at 1σ for different regions of the ORC structure. The Orc1 WH domain is shown in a, the Orc3 insertion in b and the Orc4 AAA+ domain in c.
Extended Data Figure 3 | Comparison of individual ORC subunits and archaeal Orc structures. Structures of individual ORC subunits are compared to *S. solfataricus* Orc1-1 (PDB accession number 2QBY chain A) and *A. pernix* Orc2 (PDB accession numbers 1W5S chain A (left) and 1W5T chain C (right)). Different structural elements are coloured as indicated. The initiator-specific motif (ISM) of the AAA+ ATPase fold is shown in the inset. No electron density was observed for the region linking the AAA+ and WH domains of Orc5 (indicated by a dashed line). The very N-terminal region of Orc2, which could only be built as stretches of polyalanine, is not shown.
Extended Data Figure 4 | The Orc3 domain insertion forms a conserved, hydrophobic binding platform for Orc6.  

a, Surface representation of ORC. The Orc3 insertion, which extends from the Orc3 AAA+ lid subdomain and interacts with the C-terminal helix of Orc6, is boxed.  
b, Secondary structure representation of the boxed region shown in a. The Orc3 insertion forms a bi-lobed, α-helical fold, three helices of which create a binding site for Orc6.  
c, Surface conservation of the Orc3 insertion. Conserved Orc3 residues cluster in the region that interacts with the Orc3 lid and in the Orc6 binding pocket. The latter region contacts highly conserved residues in Orc6 (Y225 and W228).  
d, Close-up view of Orc3•Orc6 interactions, showing a primarily hydrophobic binding site in Orc3 for Orc6 residues (Y225, W228, M232, A236). Y225, which in Drosophila Orc6 is equivalent to an amino acid altered in a subset of patients with Meier–Gorlin syndrome, appears positioned within a hydrogen-bonding distance of E354 in Orc3 (dashed line). Colours are as in b.  
e–h, Biochemical validation of the binding register for Drosophila Orc6.  
e, Close-up of the Orc6•Orc3 interface. Orc6–Ala236 faces a hydrophobic surface formed by Orc3 residues and is also in close proximity to a natural cysteine in Orc3 (Cys372). To validate the register of the short C-terminal Orc6 helix and the Orc6•Orc3 interface, we mutated Orc6–Ala236 to either glutamate, which we hypothesized would impede binding to ORC1–5 because of clashes with hydrophobic residues in Orc3, or to cysteine, which we presumed would not affect Orc3 binding but would allow site-specific crosslinking to Orc3-Cys372.  
f, Orc6A236E has a reduced affinity for the ORC1–5 complex. The C-terminal domains (CTDs) of wild-type (WT) Orc6, Orc6A236E or the Meier–Gorlin syndrome equivalent Orc6 Y225S were each N-terminally labelled with Alexa Fluor 488 and tested for ORC1–5 binding using fluorescence anisotropy. As shown previously4, the C-terminal domain of Orc6 binds ORC1–5 with low nanomolar affinity, whereas the Y225S mutation strongly reduces binding. As predicted on the basis of the structure of the Orc6•Orc3 interface, the A236E mutation also reduces the affinity of the Orc6-CTD for ORC1–5. Mean and standard deviations from three (for Orc6Y225S and Orc6A236E) or six (for wild-type Orc6) independent experiments are shown.  
g, Orc6A236C is able to bind to the ORC1–5 complex. Orc6-CTDWT or Orc6-CTDA236C were incubated with ORC1–5 (containing MBP-tagged Orc4) and subjected to pull-down experiments using amylose resin. Both Orc6-CTDWT and Orc6-CTDA236C co-purified with ORC1–5. The pull-down experiment was performed under non-reducing experimental conditions similar to the crosslinking experiment in h. Asterisks mark two likely proteolytic fragments of Orc3.  
h, The Orc6-CTD A236C mutant, but not the wild-type Orc6-CTD, specifically crosslinks to Orc3 within the ORC1–5 complex. Orc6-CTDWT or Orc6-CTDA236C, either alone or in the presence of ORC1–5, was incubated with a bifunctional maleimide crosslinker and the proteins subsequently analysed by SDS–PAGE. In reactions containing ORC1–5 and Orc6-CTDA236C, crosslinking gives rise to a novel band with higher molecular mass than Orc3; the appearance of this band correlates with a decrease in the amount of uncrosslinked Orc3 and Orc6-CTD, and does not appear with reactions containing ORC1–5 and wild-type Orc6-CTD, indicating that this species corresponds to an Orc3–Orc6 crosslink (a moderately strong higher molecular-mass band that appears in the absence of Orc6 probably corresponds to homotypic adducts between exposed cysteines in Orc3). These results are consistent with the structure, which places Orc6-Ala236 in close proximity to Orc3-Cys372. Note that ORC1–5 contained MBP-tagged Orc4 in g but that the tag was removed in h.
Extended Data Figure 5 | ATP-binding site configuration at the Orc4•Orc5 and Orc5•Orc3 interfaces. a, Inter-AAA\(^+\) interactions between Orc4 and Orc5 are similar to canonical AAA\(^+\) interactions between DnaA protomers (top panel, only Orc4 is used for superpositioning onto the left (light grey) AAA\(^+\) domain of an ATP-bound DnaA dimer, PDB accession number 2HCB\(^22\)). Close-up views of the nucleotide-binding site are shown for Orc4 (bottom panel) and for DnaA for comparison (middle panel). The resemblance of the Orc4 nucleotide-binding pocket to the active site of functional AAA\(^+\) ATPases is somewhat surprising considering that mutations in the active site of Drosophila and human Orc4 have no reported effect on the ATPase activity of ORC as measured \textit{in vitro}\(^{28,29}\), but may help explain why a \textit{Drosophila} ORC mutant bearing a Walker A or B substitution in Orc4 exhibits modest DNA replication defects in extracts\(^{28}\). b, The putative arginine finger in Orc5 is well conserved across homologues. A sequence logo of a multiple sequence alignment of the region containing the putative arginine finger (marked with an arrow) in eukaryotic Orc5 protein sequences is shown. Amino-acid numberings correspond to the \textit{Drosophila} Orc5 sequence. c, A potential Sensor II equivalent arginine (marked with an arrow) in the Orc4 Walker A motif is conserved in eukaryotic Orc4 homologues. A sequence logo of the Walker A motif from a multiple sequence alignment of eukaryotic Orc4 protein sequences is shown. Amino-acid positions are numbered as in \textit{Drosophila} Orc4. d, Inter-AAA\(^+\) interactions between Orc5 and Orc3. The top panel shows a superposition derived from placing the AAA\(^+\) domain of Orc5 atop the AAA\(^+\) domain of the left (dark grey) protomer of an ATP-bound DnaA dimer; the bottom panel shows a close-up view of the nucleotide-binding site at the Orc5•Orc3 interface. Side chains of conserved residues known to be involved in nucleotide binding and hydrolysis in AAA\(^+\) ATPases are represented as sticks in both a and d. WA, Walker A; WB, Walker B; SI, Sensor I; SII, Sensor II; RF, arginine finger.
Extended Data Figure 6 | The conformation of Orc1 arises from a reorientation between its AAA+ and WH domains, not from changes within the AAA+ ATPase domain itself. a, Superpositioning of the WH domains of Orc1 and S. solfataricus Orc1-1 (PDB accession number 2QBY chain A) reveals different conformations for both proteins, resulting from a large domain rotation of the Orc1 AAA+ domain around a pivot point in the linker preceding its WH domain. b, The Orc1 conformation is most similar to a state seen for A. pernix Orc2 (PDB accession number 1W5T chain C). The WH domains of both proteins were superposed as in a. c–e, Superposing the AAA+ base subdomains of Orc1 and S. solfataricus Orc1-1 (c, PDB accession number 2QBY chain A), A. pernix Orc2 (d, PDB accession number 1W5T chain C) and Orc3, Orc4 or Orc5 (e) shows that the typical AAA+ configuration between the base and lid subdomains are maintained in Orc1. Only a slight opening of the nucleotide-binding cleft is observed in Orc1, which is probably caused by the absence of bound nucleotide. f, g, The most C-terminal α-helix of the Orc1 WH domain mediates interactions with the Orc1 lid subdomain. An overview of the interaction is shown in f, with a close-up view of contacts between a conserved tyrosine (Tyr915) in the C-terminal Orc1 helix and a hydrophobic pocket of the Orc1 lid depicted in g. h, The tyrosine in the C-terminal helix of Orc1 is well conserved across metazoan but not fungal Orc1 homologues. Alignments are shown as sequence logos. The numbering of amino acids is based on Drosophila Orc1, and the tyrosine is marked by an arrow.
Extended Data Figure 7  | Nucleotide binding by Orc1, Orc4 and Orc5. For a–c, molecular replacement with the apo-ORC model was used to phase diffraction data collected from an ORC–ATP$_7$S co-crystal. Positive $F_o - F_c$ difference density contoured at different sigma levels reveals clear features for nucleotide binding to the AAA$^1$ domains of (a) Orc1, (b) Orc4 and (c) Orc5. ATP$_7$S is docked into the difference density for reference; owing to the moderate (4.0 Å) resolution of the data, this structure was not refined.

d. Modelling of canonical AAA$^+$ interactions between Orc1 and Orc4, generated using the Orc4•Orc5 interaction as a reference. Upper panel: structural overview of modelled AAA$^+$ domain positioning between Orc1 and Orc4. Lower panel: close-up of the modelled Orc1•Orc4 ATPase site. Side chains (taken from their place in the apo-ORC model as a reference) are shown for conserved catalytically important residues.
Extended Data Figure 8 | Comparison of crystallographic and EM models.

Docking of the observed and remodelled ORC structures into the cryo-EM density of *S. cerevisiae* and *Drosophila* ORC indicates that the ATPase domain of Orc1 is repositioned into a canonical AAA$^+$/AAA$^+$ interaction with Orc4 when Cdc6 is present, and supports a model where DNA passes through the central channel in ORC. a, The three-dimensional EM volume for *S. cerevisiae* ORC (as present in a complex with Cdc6, Cdt1 and Mcm2–7 and assembled in the presence of DNA, EMD-5625 (ref. 24)) contains Orc1 in the activated conformation. ORC with Orc1 in the autoinhibited conformation (left panel, as observed in the crystal structure) and remodelled conformation (right panel, remodelled) were docked into the ORC cryo-EM map (only the density for ORC$^+$Cdc6 is shown). The ORC$^+$Cdc6 EM density readily accommodates Orc1 in the activated conformation, but not in its autoinhibited state. The EM density corresponding to Cdc6 is indicated in the right panel, as observed in the crystal structure. DNA passes through the central ORC channel in the DNA•ORC•Cdc6 complex. ORC (with Orc1 in the remodelled conformation) was first docked into the cryo-EM map derived from a DNA•ORC•Cdc6 complex (EMD-5381 (ref. 23)). The AAA$^+$ domain and its associated DNA from either *S. solfataricus* Orc1-1 or Orc1-3 (PDB accession number 2QBY$^+$) were then superposed using the AAA$^+$ domain of Orc4 as a guide. Both dockings indicate that a region of density previously assigned to the Orc6 subunit$^+$ actually corresponds to the DNA duplex. Although superpositioning of the AAA$^+$ domain of *S. solfataricus* Orc1-3 onto Orc4 better positions duplex DNA in the observed EM density (than does the comparable exercise using the Orc1-1•DNA complex), the curvature of the DNA (as present in the Orc1-3•DNA co-crystal structure) results in a greater number of clashes between DNA and ORC subunits. Nevertheless, both docking scenarios are consistent with a DNA binding mode of ORC where DNA runs through the central channel. Note that the handedness of the EM map (EMD-5381 (ref. 23)) has been corrected in this figure because it has been reported that the original handedness was inverted$^{24}$. For clarity, the WH domain of Orc2 is omitted from the remodelled ORC structure in c. b, DNA passes through the central ORC channel in the DNA•ORC•Cdc6 complex. ORC (with Orc1 in the remodelled conformation) was first docked into a prior three-dimensional EM reconstruction of *Drosophila* ORC (EMD-2479$^+$) reveals excellent agreement between EM and crystal structures, but not between EM and remodelled activated ORC structures. The poor fit of the remodelled Orc1 conformation into the EM density suggests that the EM structure represents the autoinhibited state of ORC as seen in the crystal, indicating it is the predominant state in solution. See also Supplementary Video 2.
Extended Data Figure 9 | Docking of the ORC structure into the cryo-EM structure of an S. cerevisiae replication initiation intermediate indicates that ORC recruits the MCM2–7 complex by binding to the ORC WH domains. A prior model for ORC–MCM2–7 engagement24, proposed from an ORC–Cdc6–Cdt1–MCM2–7 cryo-EM structure generated in the presence of DNA (shown in a, EMD-5625 (ref. 24)), used the crystal structure of replication factor C (RFC) bound to the sliding clamp PCNA (shown in b, PDB accession number 1SXJ31) to suggest that the AAA+ domains of ORC engage the MCM2–7–Cdt1 complex. However, using the handedness of the EM volume as reported24, this organization of ORC subunits leads to an inverted ATPase site assembly, requiring that the Orc4 arginine finger (which is known to stimulate Orc1 ATP hydrolysis33) points towards the Orc5 nucleotide-binding site rather than the appropriate Orc1 active site. Schematics for the ATP site assemblies of ORC and RFC derived from these structures are shown in the lower panels in a and b. The location of the WH domain collar of ORC and the C-terminal collar of RFC is indicated by a grey circle. c, Docking of the ORC crystal structure (with Orc1 in its remodelled or 'activated' conformation) into the cryo-EM map shown in a reveals that the WH domains of ORC face an MCM2–7 complex. This switched polarity of WH domains and AAA+ domains in the EM map corrects the ATPase site assembly and is schematized in the right panel.
### Extended Data Table 1 | Summary of data collection, phasing and refinement statistics

|                         | Native       | GdCl₃ soak | Native - ATPγS |
|-------------------------|--------------|------------|----------------|
| **Data collection:**    |              |            |                |
| Beamline                | APS 23ID-B   | APS 23ID-B | NSLS X25       |
| Wavelength (Å)          | 1.0332       | 1.71083    | 1.1            |
| Resolution range (Å)    | 48.8-3.5 (3.59-3.5) | 50.23-3.96 (4.13-3.96) | 58.11-4.0 (4.16-4.0) |
| Space group             | I222         | I222       | I222           |
| Unit cell dimensions    |              |            |                |
| a, b, c (Å)             | 145.5, 259.0, 257.0 | 145.3, 257.7, 256.1 | 144.2, 255.3, 261.1 |
| α, β, γ (°)             | 90, 90, 90   | 90, 90, 90 | 90, 90, 90     |
| **Reflections**         |              |            |                |
| Total                   | 518,695 (38,421) | 2,359,387 (209,066) | 554,356 (62,459) |
| Unique                  | 61,418 (4,491) | 40,955 (3,602) | 41,025 (4,568) |
| Completeness (%)        | 100.0 (100.0) | 97.3 (76.2) | 100.0 (100.0)  |
| Multiplicity            | 8.4 (8.6)    | 57.6 (58.0) | 13.5 (13.7)    |
| Mean I/σ                | 7.7 (1.1)    | 23.4 (1.9) | 16.6 (1.3)     |
| CC₁/₂                   | 0.995 (0.427) | 0.999 (0.745) | 0.999 (0.678) |
| Rmerge                  | 0.214 (2.181) | 0.355 (3.857) | 0.095 (2.174) |
| Rfree                   | 0.078 (0.788) | 0.047 (0.509) | 0.027 (0.606) |
| Wilson B factor         | 118.6        | 165        | 197.8          |
| **Phasing:**            |              |            |                |
| FOM before density modification | 0.27 |            |                |
| FOM after density modification | 0.64 |            |                |
| **Refinement:**         |              |            |                |
| Rfree/RRfree            | 0.2243 / 0.2581 |            |                |
| Number of atoms         |              |            |                |
| Protein                 | 16336        |            |                |
| Ligands                 | 2            |            |                |
| **B factors**           |              |            |                |
| Mean                    | 118.1        |            |                |
| Macromolecule           | 118.1        |            |                |
| Ligands                 | 109.2        |            |                |
| Root mean square deviation |          |            |                |
| Bond lengths (Å)        | 0.006        |            |                |
| Bond angles (°)         | 0.705        |            |                |
| **Ramachandran plot**   |              |            |                |
| % favored               | 92.9         |            |                |
| % allowed               | 7.1          |            |                |
| % outliers              | 0.0          |            |                |
| Molprobity              |              |            |                |
| Clashscore              | 4.53         |            |                |

Highest resolution shell is shown in parenthesis.