

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

Chew Theng Lim^{1,†}, Thomas C.R. Miller^{2,5,†,*}, Kang Wei Tan¹, Saurabh Talele⁴, Anne Early¹, Philip East³, Humberto Sánchez⁴, Nynke H. Dekker⁴, Alessandro Costa² and John F.X. Diffley^{1,*}

¹ Chromosome Replication Laboratory

² Macromolecular Machines Laboratory

³ Bioinformatics & Biostatistics

The Francis Crick Institute, 1 Midland Road, London, NW1 1AT

⁴ Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands

⁵ DNRF Center for Chromosome Stability, Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark

[†] Equal contribution

*Corresponding Authors

john.diffley@crick.ac.uk

Tel: +44 (0) 203 796 1833

tmiller@sund.ku.dk

Tel: +45 (0) 35 32 70 74

Running title: Multiple pathways for MCM loading

Summary

Eukaryotic DNA replication initiates from multiple genomic loci known as origins. At budding yeast origins like ARS1, a double hexamer (DH) of the MCM replicative helicase is assembled by Origin Recognition Complex (ORC), Cdc6 and Cdt1 via sequential hexamer loading from two opposed ORC binding sites. Cyclin Dependent Kinase (CDK) inhibits DH assembly, which prevents re-replication by restricting helicase loading to G1 phase. Here we show that an intrinsically disordered region (IDR) in the Orc2 subunit promotes interaction between ORC and the first loaded, closed-ring MCM hexamer (the MO intermediate); CDK phosphorylation of this IDR blocks MO formation and DH assembly. We show that MO functions by stabilising ORC at the lower affinity binding sites required for second hexamer loading. Origins comprising two high affinity ORC sites can assemble DH efficiently without MO by independently loading single hexamers; these origins escape CDK inhibition *in vitro* and *in vivo*. Our work reveals mechanistic plasticity in MCM loading with implications for

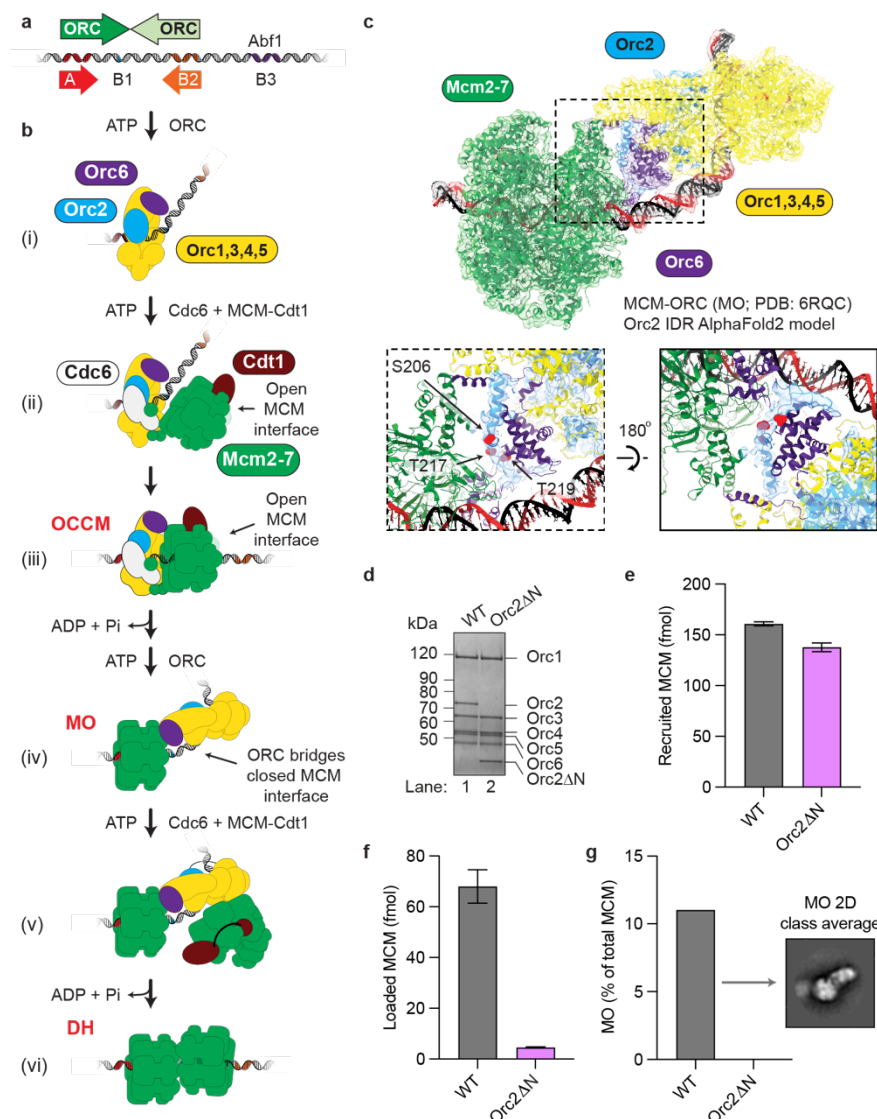
understanding how CDK regulation has shaped yeast origin evolution and how natural origins might escape cell cycle regulation. We also identify key steps common to loading pathways, with implications for understanding how MCM is loaded in other eukaryotes.

Introduction

The initiation of eukaryotic DNA replication is separated into two temporally distinct steps that ensures no region of the genome is replicated more than once¹⁻³. During G1 phase, a pair of hexameric MCM helicases is loaded at each replication origin around double stranded DNA in the form of a head-to-head MCM DH. During S-phase, each DH is converted into two active CMG (Cdc45-MCM-GINS) helicases, thereby initiating bidirectional DNA replication.

Budding yeast DNA replication origins generally contain a high affinity ORC binding site flanked by one or more low affinity sites in the opposite orientation⁴⁻⁶. At ARS1 (Fig. 1a), the bipartite high affinity site comprises the A and B1 elements (A/B1) and a lower affinity site corresponds to the B2 element⁶⁻⁸. Both the A and B2 elements contain the extended ARS consensus sequence (EACS): the A element is a closer match to the EACS than B2. DH assembly at ARS1 occurs via sequential loading of two MCM hexamers (outlined in Fig. 1b)^{5,9,10}. After loading of the first hexamer (steps i-iii), ORC releases from A/B1. ORC then binds both B2 and the N-terminal domain of the first loaded MCM, generating the MCM-ORC or 'MO' complex (iv)¹¹. The same ORC that was bound to A/B1 can 'flip' and bind B2 (Ref ¹²); alternatively, a second ORC molecule can bind B2 and engage the first recruited MCM (Ref ⁵). In MO, the MCM ring is closed and MCM subunits are primarily in a post-hydrolysis (ADP-bound) state, very similar to the DH¹¹. Because this second ORC binding event also induces a DNA bend, the face of ORC that promotes initial MCM recruitment and OCCM formation is available in MO to receive the second MCM-Cdt1 without steric interference (v). MO is thus poised to be a key intermediate in DH assembly (vi); nonetheless, the exact mechanism of MO formation, its function and its importance in DH assembly remain unclear. Moreover, whether all DH assembly proceeds via MO is unknown.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Results

MO is crucial for MCM loading at ARS1

We first sought to understand the importance and role of MO in DH assembly by identifying an ORC mutant that cannot form MO. We have previously shown that an N-terminal deletion of the Orc6 subunit of ORC (Orc6Δ119) reduced both MO and DH formation by approximately 50% consistent with the idea that MO is a precursor of DH¹¹; however, this was a relatively modest effect and we wanted a mutant with more complete loss of MO formation. Orc2 contains a long (235aa) N-terminal extension that is predicted by AlphaFold¹³ to be intrinsically disordered. The Orc2 IDR has not been resolved in any ORC structure determined experimentally, including MO. However, a region of the Orc2 IDR was tentatively assigned as wrapping around Orc6 in a cryo-EM density map of DNA-

bound ORC¹⁴, albeit at a resolution that prevented assignment of the amino acid sequence and inclusion of this region in the deposited PDB (5ZR1-Ref¹⁴). Using AlphaFold2 multimer¹⁵, we identified a region of the Orc2 IDR docking on the C-terminal domain of Orc6 with high confidence (Fig. 1c; Extended Data Fig. 1), consistent with the density seen by Li *et al.* Notably, overlaying our AlphaFold prediction with our previously determined structure of the MO intermediate places the region of the Orc2 N-terminal IDR (aas 190-231) into a contiguous stretch of density at the core of the MCM-ORC interface, suggesting that it may play a key role for MO formation (Fig. 1c). Deletion of this IDR from Orc2 did not affect the ability of Orc2 to form a stable complex with the other ORC subunits (Fig. 1d). ORC complexes containing this deletion (Orc2ΔN) recruited MCM normally in ATP_γS (Fig. 1e). However, Orc2ΔN loaded MCM at a level only 7% that of

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

wild type ORC (Fig. 1f). To visualise the effect of Orc2 IDR truncation on DH formation directly, we performed MCM loading assays on a 454bp linear DNA template containing ARS1 flanked by nucleosomes and imaged the reactions by negative stain EM. Whilst MO was readily detectable with wild type ORC in ATP, consistent with previous work¹¹, MO was not detected in reactions with Orc2ΔN (Fig. 1g). Thus Orc2ΔN can efficiently recruit MCM to DNA but does not form the MO intermediate, and thus cannot form DHs efficiently. These experiments define a role for Orc2 IDR as a key element of the MO intermediate and provide stronger evidence that MO is a critical intermediate in the assembly of the DH at ARS1.

MO is Required to Stabilise ORC at Weak Binding Sites

The two key ORC binding sites at ARS1 (A/B1 and B2) are so close together that ORC cannot bind both sites simultaneously⁶. Thus, MCM loading must occur in two sequential ORC binding steps. In Fig. 2a we arranged budding yeast origins according to the distance between the best match to the EACS (A domain) and the best predicted secondary ORC site in the opposite orientation (B2-like domain). In roughly 1/3 of yeast origins, the two ORC binding sites are as close or closer together than A/B1 and B2 in ARS1 (Fig. 2a). Using synthetic origins with two high affinity ORC sites (perfect EACS/B1) in which distances are measured from the inside edges of the two B1 elements, 10bp corresponds to the distance between A/B1 and B2 in ARS1, and this construct, like ARS1, can only bind a single ORC molecule (Extended Data Fig. 2a lanes 5 and 6). Moving binding sites just two bp further apart, however, allows two ORC molecules to bind simultaneously (Extended Data Fig. 2a lanes 7 and 8). Thus, roughly 2/3 of yeast origins are predicted to allow simultaneous binding of two ORC molecules (Fig. 2a). In approximately 20% of yeast origins, the two predicted ORC binding sites are far enough apart to fit a single MCM

hexamer between them and in 23% they are far enough apart to fit a DH between them. However, in only 15% of origins, the weaker ORC binding site is a better EACS match than ARS1 B2 (Extended Data Fig. 2b). Thus, asymmetry in ORC binding affinities appears to be a general property of yeast origins regardless of whether or not the two sites overlap. The MCM loading mechanism has primarily been studied using ARS1; it remains unclear whether the mechanism of MCM loading is the same at all origins or whether other mechanisms may be important at origins with different ORC binding affinities or wider spacing between ORC binding sites.

To begin to address this, we tested the ability of Orc2ΔN, which cannot form MO, to load MCM on a series of synthetic origins with two high affinity ORC binding sites in the opposite orientation and placed at distances from 10 to 300 base pairs apart. As shown previously⁵, wild type ORC can load MCM efficiently on this series of origins, with a peak of MCM loading at 70-90bp but with efficient loading even at 300bp. Orc2ΔN cannot load MCM efficiently at ARS1, but surprisingly, loads MCM at all distances from 10 to 300bp as efficiently as wild type ORC (Fig. 2b). This suggests that MO is not required when MCM is loaded from two high affinity sites, regardless of whether those sites effectively overlap (10bp) or are separated by a relatively large distance (300bp). To test this directly, we changed one high affinity EACS/B1 site in the most efficient synthetic template (90bp) to the B2 element from ARS1, a lower affinity ORC site (Fig. 2c). As shown in Fig. 2d, wild type ORC loads MCM as efficiently on this template as it does on the parent template with two high affinity sites, demonstrating that weak ORC sites can be effectively used even when the two ORC sites do not overlap. MCM loading with Orc2ΔN on the 90bp template was at least as efficient as wild type, but MCM loading on the 90bp-B2 construct was greatly reduced.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

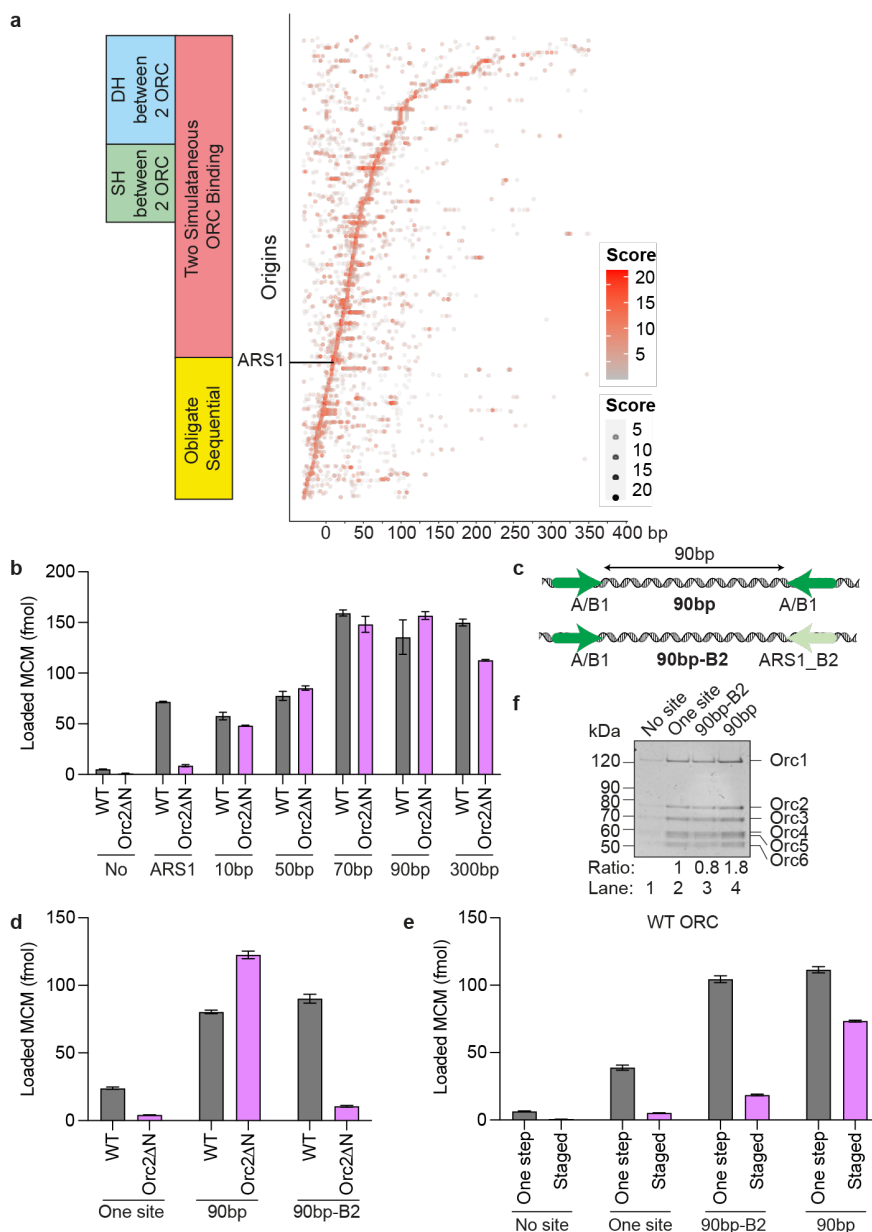


Figure 2. The MO promotes MCM loading by stabilizing ORC at low-affinity binding sites.

a. Matches to the EACS were identified using a position weight matrix (PWM) as previously described⁵. Origins are arranged according to the distance between the centre of the best match to the EACS and the centre of the best B2-like EACS by PWM in the opposite orientation. PWM scores indicated by size/colour of dots as indicated in the figure. **b.** Orc2ΔN loads MCMs onto origins containing two high-affinity inverted ORC binding sites spaced ≥10 bp apart with comparable affinity to WT ORC. **c.** Schematic of synthetic origins containing inverted high affinity ORC binding sites spaced 90bp apart (top), or with one high affinity site replaced by the ARS1 B2 element (bottom). **d.** MCM loading on 90bp origins by Orc2ΔN is dependent on two high affinity ORC binding sites. **e.** MCM loading by WT ORC shows two DNA-bound ORCs load MCM more efficiently than one ORC in staged reaction. **f.** ORC is retained on 90bp and 90bp-B2 origins after washing. The binding ratio is normalised to the one site template, showing 90bp symmetrical origins retain approximately twice as much ORC as 90bp-B2 origins.

From these experiments we conclude that the MO intermediate is critical for MCM loading when one of the ORC binding sites is lower affinity, and, therefore, that the essential role of MO in DH assembly is to stabilise ORC binding at weak sites after first hexamer loading. Furthermore, MO is still essential when the two ORC sites are 90bp apart indicating that MO is not just required when ORC sites overlap, as in ARS1. However, MO is not required when MCM is loaded from two high affinity ORC binding sites.

Because the EACS/B1 sequences in both 90bp and 90bp-B2 are high affinity ORC binding sites, ORC binding to these sequences is stable to stringent (80 mM

NaCl) washing whilst binding to B2 is not⁵, which gives us an opportunity to examine the efficiency of the ‘one ORC’ loading via ORC flipping vs the ‘two ORC’ mechanisms where a second ORC molecule loads the second hexamer. To test this, reactions were ‘staged’: near-saturating amount of wild type ORC was pre-bound to either the 90bp or the 90bp-B2 origin (Extended Data Fig. 2c), then free ORC along with B2-bound ORC was washed off (Extended Data Fig. 2d). We then added Cdc6 and MCM-Cdt1 and examined MCM loading. As shown in Fig. 2e, in staged reactions the 90bp template loaded approximately 4 times more MCM than the 90bp-B2 template (pink bars) despite the fact that both templates work equally well in a one-step reaction (grey bars) and ORC was retained on both

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

templates after washing (Fig. 2f). This result is consistent with the idea that a single ORC can load a double hexamer from a single high affinity site via ORC flipping and MO formation but also shows that this pathway is less efficient than the pathway in which each MCM single hexamer is loaded by separate ORC molecules at two high affinity sites. Moreover, the difference between the one-step and staged reaction with 90bp-B2 also argues that the single ORC pathway is less efficient than the two ORC pathway even when loading works through MO. Therefore, we can distinguish three separate ways MCM can be loaded: one ORC via MO, two ORC via MO and two ORC MO-independent.

CDK phosphorylation of Orc2 blocks MO formation

CDK promotes Cdc6 proteolysis¹⁶⁻¹⁹ and MCM-Cdt1 nuclear export²⁰⁻²⁴ and also blocks ORC's ability to assemble DHs *in vitro*²⁵⁻²⁷. The region in the Orc2 IDR predicted to form part of the MCM-ORC interface in MO (Fig. 1c and yellow box in Extended Data Fig. 3a) is moderately conserved amongst closely related budding yeasts, which all contain multiple potential CDK phosphorylation sites near the predicted interface with MCM in MO (Fig. 1c red dots). We therefore hypothesised that CDK phosphorylation of Orc2 might inhibit MCM loading by interfering with MO formation. Phosphorylation of wild type ORC with S phase CDK (Clb5-Cdc28-Cks1) caused a reduction in mobility of both the Orc2 and Orc6 subunits in SDS PAGE (Fig. 3a); this CDK phosphorylated ORC was very inefficient in DH formation at ARS1 (Fig. 3b), consistent with previous work²⁵⁻²⁷. Similar to Orc2ΔN, however, phosphorylated ORC loaded MCM almost as well as unphosphorylated ORC on the 90bp synthetic origin (Fig. 3b). To ensure this loaded MCM was functional, we tested its ability to support DNA replication *in vitro* with purified proteins. Whilst pre-phosphorylated ORC did not support replication from an ARS1-containing template (compare lanes 5 and 6, Fig. 3c), it supported replication from a template containing the 90bp origin almost as well as unphosphorylated ORC (lanes 3 and 4, Fig. 3c and Extended Data Fig. 4a) indicating that the DH formed is functional. Consistent with this, a 2.8 Å resolution cryo-EM structure of DH assembled with phosphorylated ORC is virtually identical to DH assembled with unphosphorylated

ORC (Extended Data Fig. 5a). To test whether this 90bp origin could also bypass CDK regulation of ORC *in vivo*, we used a yeast strain in which CDK regulation of MCM was constitutively eliminated by fusing an unregulated nuclear localisation sequence (SV40-TAg NLS) to Mcm7, and CDK regulation of Cdc6 was conditionally deregulated with a copy of a stable version of Cdc6 (Cdc6ΔNT) under a galactose-inducible promoter. When both Cdc6 and MCM are deregulated, cells rely entirely on wild type ORC phosphorylation to prevent re-replication²⁸. We integrated a series of origins in place of ARS419, then arrested cells in G2/M, expressed stable Cdc6 and examined copy number after 3 hrs. As shown in Fig. 3d, deregulation of Cdc6 and MCM did not lead to any increase in DNA copy number of ARS1; ARS317, which had been previously shown to re-replicate under similar conditions²⁹, showed elevated copy number, but the 90bp origin showed the largest increase in copy number after Cdc6 deregulation consistent with the idea that this synthetic origin bypasses CDK regulation of ORC both *in vitro* and *in vivo*.

We next sought to determine which step in DH assembly was blocked by Orc2 phosphorylation. To examine the effects of Orc2 phosphorylation alone, we mutated four Ser/Thr residues in the CDK consensus sequences (^{S/T}PX^{K/R}) in Orc6 to alanine to generate Orc6-4A. As shown in Fig. 3a, the phosphorylation-dependent shift of Orc6 was greatly reduced in Orc6-4A, but Orc2 was still shifted by CDK phosphorylation. On ARS1, CDK phosphorylation of Orc2 did not affect MCM recruitment in ATPγS (Fig. 3e). However, Orc2 phosphorylation led to a substantial reduction in DH formation in both DNA-pulldown (Fig. 3fi) and EM-based (Fig. 3fii) assays and a complete inhibition of MO formation (Fig. 3g), indicating that phosphorylation of Orc2 alone is sufficient to inhibit DH formation. Thus, CDK phosphorylation of Orc2, like deletion of the Orc2 IDR (Fig. 1), blocks DH assembly before MO formation. Whilst CDK phosphorylation of Orc2 blocks MO formation and DH formation on the ARS1 origin, it does not block DH formation on the 90bp origin (Fig 3b), consistent with the results in the previous section showing that MO is not required when origins contain two high affinity binding sites (Fig. 2b and d).

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

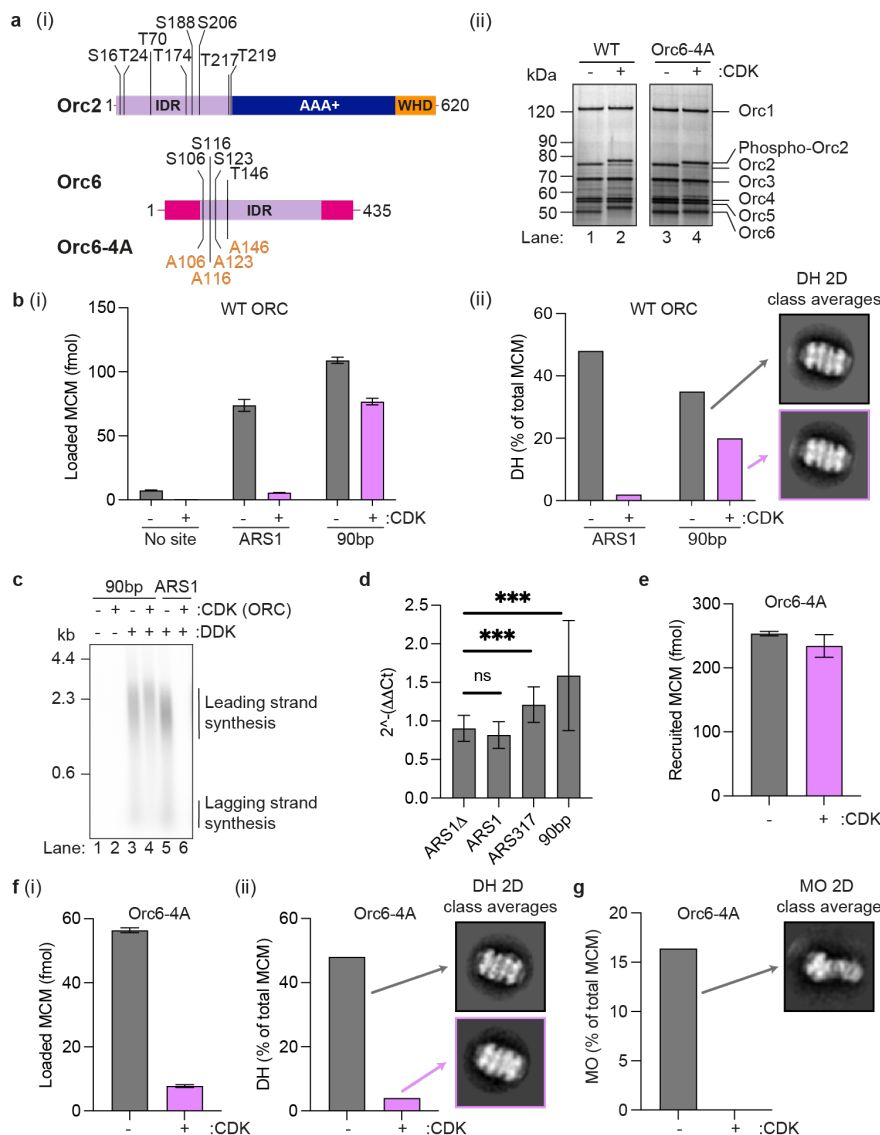


Figure 3. CDK phosphorylation of Orc2 inhibits the formation of MO.

a. (i) Potential CDK phosphorylation sites (consensus sequences (S/T PK R)) in Orc2 and Orc6. **(ii)** An Orc6 mutant with four CDK target sites mutated to alanine (Orc6-4A) can be stably incorporated into the ORC complex but not phosphorylated. CDK phosphorylation of Orc2 is comparable to WT. **b.** MCM loading on ARS1, but not 90bp origins, is highly sensitive to CDK phosphorylation ORC. **c.** MCMs loaded onto 90bp origins by phosphorylated ORC are replication competent. **d.** In yeast cells dependent on ORC phosphorylation to prevent re-replication, 90bp spaced origins inserted into the yeast genome re-replicate causing an increase in copy number, as detected by qPCR. ORC phosphorylated on Orc2 by CDK (Orc6-4A mutant) efficiently recruits MCMs to ARS1 origins **(e)** but is inhibited in MCM loading **(f)**. **g.** ORC phosphorylated on Orc2 by CDK (Orc6-4A mutant) fails to form MO.

RECONSIL experiments enabled us to identify SHs apparently trapped on DNA between nucleosomes (Fig. 4a). Thus, the Orc2 IDR seems not to be required for OCCM assembly or disassembly to release SH onto DNA but is required for MO-formation, which can be inhibited by CDK phosphorylation of the Orc2 IDR.

To confirm this, we used

Structure of the MCM single hexamer

Formation of MO requires Orc2 and Orc6 to bridge the N-terminal interface between the Mcm2 and Mcm5 subunits that form MCM's DNA entry 'gate'. However, it is currently unknown whether closure of the Mcm2-5 interface precedes ORC binding in the MO or whether ORC binding closes the Mcm2-5 gate to load the first MCM hexamer. Phosphorylated Orc2 (Orc6-4A) recruits MCM normally but does not form MO suggesting that Orc2 phosphorylation either blocks a step in the loading of the SH, for example OCCM disassembly, and/or directly prevents ORC binding to a loaded SH during the formation of MO. To investigate this, we performed single particle reconstitution in silico (RECONSIL)¹¹ experiments to visualise entire chromatinised origins of replication during MCM loading reactions containing phosphorylated or truncated Orc2.

cryo-EM to visualise the helicase loading reaction with CDK-phosphorylated Orc2. The 3.4 Å resolution structure (3.1 Å after density modification) shows a single MCM ring encircling duplex DNA (Fig. 4b, Extended Data Fig. 5). Unlike the MCM in the OCCM structure, the SH contains a completely closed Mcm2-5 interface identical to that seen in the DH (Fig. 4b, Extended Data Fig. 5b). Thus, phosphorylation of ORC2 blocks DH formation after a SH has been fully closed around DNA, implying that ring closure precedes and is not a consequence of N-terminal ORC binding that results in MO formation. Comparison with MCM in the OCCM complex reveals that the DNA grip in the helicase ring changes after closure of the Mcm2-5 gate (from Mcm7-4-6-2 to Mcm3-7-4-6 duplex engagement), likely as a consequence of ATP hydrolysis (Fig. 4b). Nucleotide occupancy changes from ATPγS in

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

the 3-7, 7-4, 4-6 and 6-2 interfaces in OCCM, to ATP in the 4-6 and 6-2 interfaces and ADP in the 2-5, 5-3, 3-7 and 7-4 interfaces in SH. This change in nucleotide occupancy indicates a minimum of 4 ATP hydrolysis events that may drive the re-orientation of the ATPase pore loops, with pre-sensor 1 (PS1)–arranged in a staircase configuration that follows the helical rise of the leading-strand template (Fig. 4c), as observed in the active CMG helicase (Extended Data Fig. 5c).

Mechanism of MO-independent DH assembly

To examine the properties of the loaded SH, we assembled MCM with Orc2 Δ N on either a one site template, where only SH assembles, or a template containing two ORC binding sites separated by 90bp where DH efficiently forms; we then challenged the loaded MCM products with buffers containing different salt concentrations. As shown in Fig. 5a, MCM loaded on both templates was resistant to salt concentrations up to 250 mM NaCl; virtually all of the MCM was removed from the one site templates at 500 and 1000 mM NaCl whilst about 40% of the MCM remained on the 90bp template after 500 and 1000 mM NaCl washes. These experiments indicate that SH is stable up to ~ 250 mM NaCl, but is efficiently removed at salt concentrations of 500 mM NaCl and above. The DH is stable up to 2 M NaCl³⁰, so the fraction of MCM removed by 500 mM NaCl from the 90bp origin likely represents SH, with the MCM remaining at 500 and 1000 mM NaCl being DH. This suggests that a substantial fraction of the MCM loaded even on this very efficient origin are SHs that have not matured into DHs. This is consistent with single molecule experiments showing that SHs are readily detected during normal MCM loading reactions³¹.

To determine the off rate of SH under MCM loading conditions (low salt), MCM was loaded as above using Orc2 Δ N on a single site sequence, excess protein was removed by low salt wash, and the amount of SH remaining on DNA was then assessed after incubation for the indicated times followed by low salt wash. Fig. 5b shows that SHs were lost from DNA with a half-life of approximately 3.5min. To compare this to the DH, we loaded MCM onto the 90bp template then washed free

proteins and SHs off with high salt (1 M NaCl) before taking time points. As shown in Fig. 5b, DHs are stable for the entire duration of the experiment. Thus, the SH is much less stable than the DH on DNA.

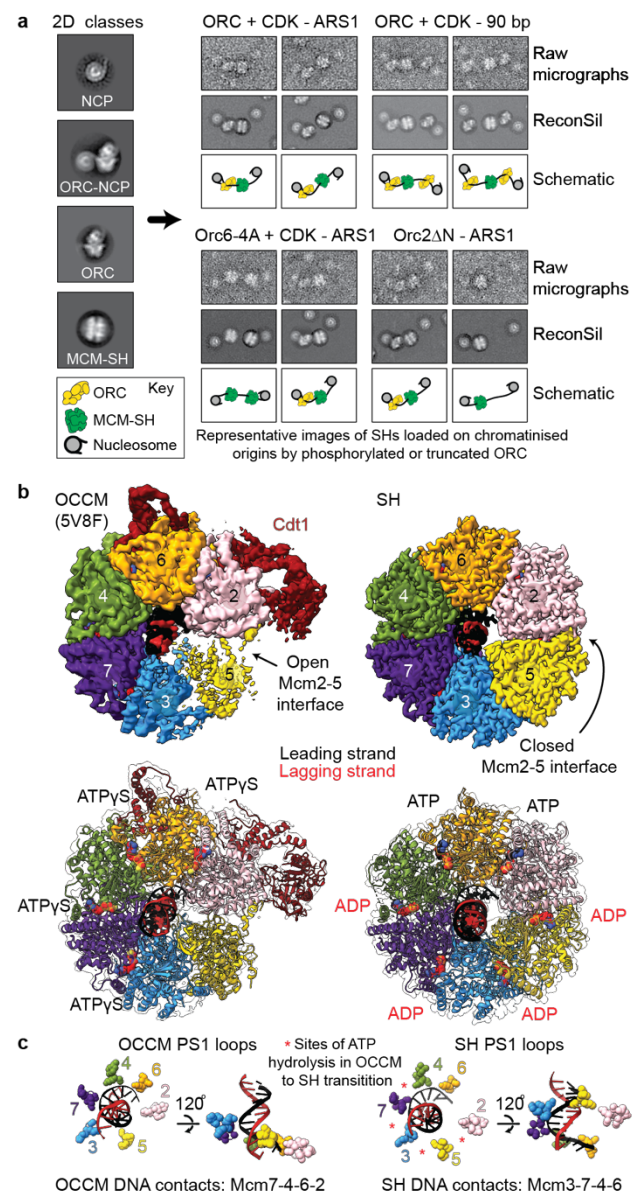


Figure 4. Structure of the MCM SH

a. Negative stain EM in silico reconstitution (ReconSil) of MCM loading reactions performed with phosphorylated or truncated Orc2. Left, conventional EM 2D class averages. Right, representative images of chromatinised origins bound by single MCM helicases and/or DNA-bound ORCs flanked by nucleosomes, shown as raw micrographs (top rows), after ReconSil (middle rows) or as schematics (bottom rows). **b.** Cryo-EM structures and atomic models for OCCM (left; 5V8F [Ref⁶⁵]) and a SH (right; this study). The EM density map displayed for the SH has been density modified using EMReady⁶². **c.** ATP hydrolysis in the Mcm4-7, 7-3, 3-5 and 5-2 ATPase sites during SH loading is associated with a reconfiguration of the MCM PS1 loops into a staircase configuration that engages the leading strand template DNA backbone.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

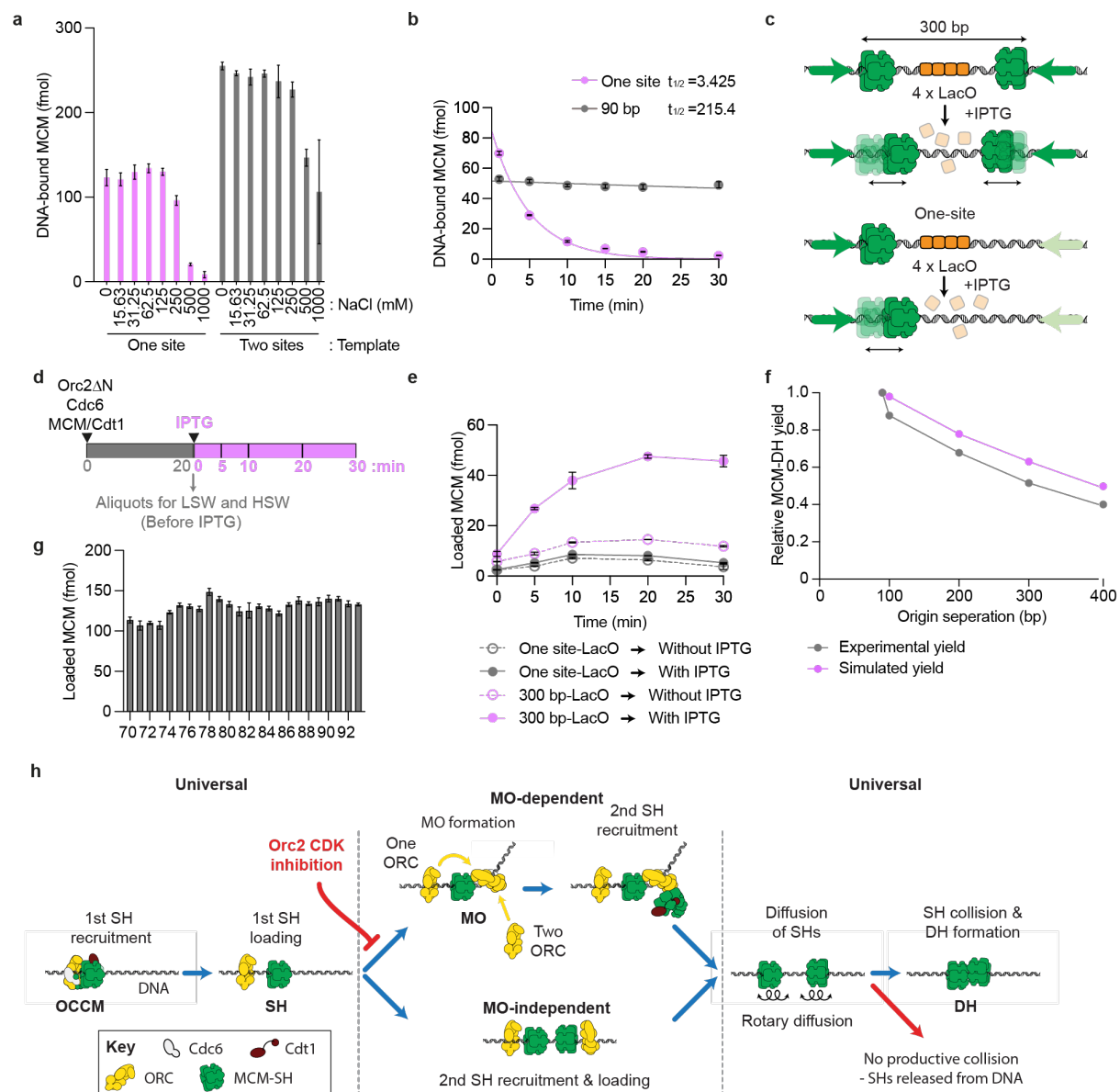


Figure 5. Mechanism of MO-independent DH formation.

SHs have a lower salt stability (**a**) and shorter half-life (**b**) on DNA than DHs. Schematics illustrating origins (**c**) and assay (**d**) used for characterising MO-independent DH formation via collision of two independently loaded SHs (**e**). **f**. Efficiency of DH formation decreases with increase in separation between ORC binding sites and can be modelled as the meeting probability of two diffusing SHs loaded at ORC binding sites at a given separation in the experimental time (20 minutes). **g**. MCM loading is insensitive to the separation distance between two ORC binding sites (between 70-93 bp) and relative orientation of recruited ORCs with respect to each other along the DNA helix. **h**. Proposed model for MCM loading via MO-dependent and independent pathways. Both pathways converge on a 'universal' mechanism that requires loading of two SHs that must collide in a head-to-head orientation to form a DH.

To show more directly that two independently loaded SHs can assemble into a DH, we generated a template with two high affinity ORC sites separated by 300bp and 4 lac operators between the sites (Fig. 5c). As a control we also generated a template with a single ORC site and 4 lac operators. We pre-assembled SHs on these templates with Orc2 Δ N for 20 minutes (Fig. 5d). As shown in Extended Data Fig. 6a, there is considerable MCM

loaded on both templates at this point; however, this MCM is all removed by high salt wash indicating that there are no DHs. DH formation (high salt resistant MCM) was detected at significant levels only on the two ORC binding site template and only after IPTG addition (Fig. 5e) in a time-dependent manner. This strongly supports the idea that two SHs can generate a DH after interaction.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

The efficiency of DH assembly decreases as two ORC sites are moved further apart. We reasoned that the further apart two SHs are loaded, the higher the probability that one or the other will fall off before meeting. To model this, we used the observed one dimensional diffusion constant for the DH³², together with the off rate determined in Fig. 5b (Extended Data Fig. 9 and 10), and as shown in Fig. 5f, the predicted loading efficiencies are remarkably close to those seen in the experimental data. Diffusion could occur with rotation following the path of the DNA helix, or by diffusion without rotation. For several reasons we favour the idea that movement occurs coupled with rotation tracking the DNA helix. First, the structure of the SH shows a single mode of DNA engagement within the MCM central channel, with individual nucleotides clearly resolved (Extended Data Fig. 5d). Despite the large number of particles (~400k), further 3D classification failed to detect SHs with an alternate DNA engagement, suggesting that the DNA register remains the same in all imaged SHs, regardless of their relative position on the origin DNA. If SHs diffused without rotation, they would be unlikely to have a highly preferred mode of DNA engagement. Second, without rotation, two SHs would meet in different registers depending on how far apart the loading sites were. One might expect to see evidence of this in loading efficiency; however Fig. 5g shows that loading efficiency is very similar at every distance from 70 to 93bp. Rotation along the DNA helix ensures that the two SHs always meet with the same register; using the first steric clash between hexamers as the end point of rotational diffusion along B-form DNA, this register is very close to the register seen in the DH (Extended Data Movie 1). Taken together, we propose DH assembly occurs by random rotational diffusion and meeting of two SH along the DNA followed by a presently uncharacterised ‘locking’ of the two into a stable DH.

Discussion

Our results, summarised in Fig. 5h, indicate that MCM can be loaded by at least three distinct pathways in budding yeast. In each pathway, the unifying mechanism is the loading of two SHs with inverted orientations. DH assembly can be coordinated by the MO, which facilitates ORC

recruitment to a weak DNA site. Alternatively, DH assembly can be done in the absence of the MO, in which case two independently loaded SHs diffuse and rotate until meeting to form a stable DH. The loading of the SH defines a precise register of DNA engagement that ensures two loaded SHs with inverted orientations will always meet in the same subunit register, regardless of the distance between their loading sites. Our data (Fig. 2b) shows that DH formation is at least as efficient without MO, as with MO, indicating that the SH to DH ‘locking’ mechanism is intrinsic to two correctly oriented and DNA engaged SHs and is independent of the MO.

SHs are considerably less stable on DNA than DHs. We suggest that this is because the Mcm2-5 interaction is intrinsically weak, providing a gate for DNA to exit the central channel, consistent with recent FRET experiments³³. The Mcm2-5 gates are not aligned in the DH but are located approximately 180° apart, so, even if both gates open at the same time, the DH would remain topologically bound around DNA³⁴ which may prevent dissociation of the DH. Alternatively, contacts between the hexamers may stabilise the Mcm2-5 gate in the DH, preventing opening³³.

Previous work in yeast has shown that a substantial fraction of MCM is released from G1 chromatin with a moderate salt concentration (350 mM NaCl)³⁵ suggesting the presence of SH on chromatin. From yeast to humans, the total amount of MCM bound to chromatin far exceeds the number of active replication origins. This ‘MCM paradox’ has been explained at least in part by the existence of dormant origins which are activated under stress conditions³⁶⁻³⁹; our results suggest that, in addition, some of the excess MCM may be in the form of SH, which cannot support replication (e.g. Fig. 3c). Though SHs are relatively short lived, there may be active mechanisms to remove them from chromatin. Further work is required to assess what fraction of chromatin-bound MCM is SH.

Synthetic origins comprising two high affinity ORC binding sites load MCM even more efficiently than ARS1 (e.g. Fig. 2b) and function as autonomously replicating sequences *in vivo*⁵. However, yeast origins generally have a single high affinity site flanked by low affinity sites,

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

indicating that they likely assemble DH primarily via MO. Our work suggests an explanation for this: origins with two high affinity sites cannot be inhibited by CDK phosphorylation of Orc2 (Fig.3 c,d). We propose this arrangement has been selected against during evolution: co-evolution of sequence-specific DNA binding by ORC^{40,41} and CDK regulation of ORC function has resulted in the asymmetry in modern yeast replication origins.

Human Orc2 also contains a long N-terminal IDR⁴² but has no detectable sequence similarity to the yeast IDR. Whether human ORC loads MCM via an MO mechanism remains to be seen. Orc2 from multicellular plants entirely lacks an N-terminal IDR and therefore plants are unlikely to use the same MO-dependent mechanism. Whether plants and metazoans use a mechanism related to mechanism 3 or some other, novel mechanism requires further work. Finally, the recent sequencing of several metamonad genomes has shown that, while all six MCM genes are present, organisms in the genus *Carpodomonas* lack all ORC subunits, Cdc6 and Cdt1 (ref⁴³) indicating additional MCM loading mechanisms remain to be discovered.

Materials and Methods

Strains

A full list of primers, plasmids and strains used for strain construction are listed in Supplementary Materials. The YJL3239 strain was a gift from Joachim Li⁴⁴. The ARS317, ARS1, ARS1Δ and synthetic 90bp origins were inserted in ARS419 using the method described in Richardson and Li, 2014⁴⁵ with some modifications. The endogenous ARS317 was deleted and replaced with kanMX6 from pFA6a-3HAKanMX6 with amplification using CTL61 and CTL62 (yCTL5). URA3 was replaced with hphNT1 from pFA6a-hphNT1 with amplification using CTL98 and CTL99 (yCTL18). URA3 was amplified using pRS306 with primers CTL86 and CTL87 and additionally inserted in ARS419 at position 567 kb in Chromosome IV (yCTL19). The ARS317, ARS1, ARS1Δ and synthetic 90bp origins were amplified from plasmid ARS317-Abf1(pCTL69), ARS1-Abf1 (pCTL68), ARS1Δ (pGC404 in Coster and Diffley, 2017⁵) and 90bp-Abf1 (pCTL67) respectively using CTL160 and CTL161 and counter-selected with 5-FOA, resulting in strains yCTL30, yCTL29, yCTL28 and yCTL33. 1 mg/ml of 5-FOA was added to minimal media containing 1X YNB, 2% glucose, 1X CSM (Complete Supplement Mixture, Formedium, UK) and 80 μg/ml uracil. The ARS317-RIP sequence was identified according to the L4-L12 region in Richardson and Li, 2014⁴⁵.

The sequences for Orc6 phosphorylation mutant Orc6-4A was obtained from Invitrogen GeneArt Synthesis

(Thermo Fisher Scientific). This sequence was then used to replace the wild type Orc6 in the plasmid pJF18 as in Frigola *et al.*, 2013⁴⁶ to create strain yCTL6 (Orc6-4A). Endogenous Orc6 was Flag-tagged to allow removal of the endogenous wild-type protein.

Amino acids 2-235 in Orc2 were deleted in the Orc2-N-terminal truncation mutant (Orc2ΔN). This mutant was introduced into pJF19 as in Frigola *et al.* (2013)⁴⁶ to replace the wild type Orc2. The wild-type Orc2 was Flag-tagged so that the endogenous protein could be removed. The resulting strain was named yCTL17.

Protein expression and purification

In this study, ORC, Cdc6, Luc-MCM/Cdt1, CDK and Sic1 were expressed and purified according to previously described methods in the references Frigola *et al.*, 2013⁴⁶, Coster and Diffley, 2017⁵, Yeeles *et al.*, 2015⁴⁷. The Orc2ΔN mutant was purified identical to wild-type ORC, but with the additional step of using Flag beads after CBP pull-down to remove endogenous wild-type proteins. The same was done for the Orc6 phosphorylation mutants.

ORC binding assay

ORC binding assays were performed as described in Coster and Diffley, 2017⁵ with some modifications. Briefly, 0.15 pmol DNA of 3 kb PCR substrates with a single biotin (GC227 and GC228) were coupled with 5 μl of Dynabeads M-280 streptavidin (Invitrogen 11205D) in 10 μl of binding buffer (5 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 1 M NaCl), for 30 mins at 30°C with mixing at 1250 rpm in a 1.5 ml microcentrifuge tube. Beads were washed twice and resuspended in buffer containing 10 mM HEPES-KOH pH7.6 and 1 mM EDTA. The bead-bound DNA substrate was then incubated with 12 nM ORC in 40 μl loading buffer (25 mM HEPES-KOH pH7.6, 100 mM NaOAc, 10 mM MgOAc, 0.02% NP-40, 5% glycerol, 1 mM DTT, 5 mM ATP and 80 mM KCl). The reaction was incubated for 15 mins at 30°C with mixing at 1600 rpm on 96-well plate. Reactions were then washed with low salt wash (LSW) additionally containing 80 mM NaCl twice (45 mM HEPES-KOH pH7.6, 5 mM MgOAc, 300 mM NaOAc, 0.02% NP-40, 10% glycerol, 2 mM CaCl₂, 80 mM NaCl) for 2 mins, 30 °C at 1600 rpm. The washed beads were then washed once with LSW once before resuspending in 12 μl LSW with 600 units of micrococcal nuclease (MNase, NEB M0247S) and incubated for 2 min, 30°C at 1600 rpm. The supernatant from two reactions were collected and separated by Criterion XT Tris Acetate 3-8% (Bio-rad) SDS-PAGE and visualized by silver staining.

MCM loading

The MCM loading assay was performed as described in Coster and Diffley, 2017⁵. Biotinylated DNA substrate was attached to magnetic beads and then incubated with 12 nM ORC, 20 nM Cdc6, 80 nM Luc-MCM/Cdt1 in 40 μl loading buffer for 20 mins at 30°C with mixing at 1600 rpm. Beads were washed twice with high salt wash (HSW, 45 mM HEPES-KOH pH7.6, 5 mM MgOAc, 0.02% NP-40, 10% glycerol, and 1 M NaCl) for 2 mins and twice with LSW (45 mM HEPES-KOH pH7.6, 5 mM MgOAc, 300 mM NaOAc, 0.02% NP-40, 10% glycerol and 2 mM CaCl₂) 20 secs. Proteins were released by 30 μl of low salt wash with MNase treatment for 2 mins and transferred to white flat bottom, half area 96-well plates (Corning, 3642). Nano-Glo Luciferase assay substrate (Promega N1120) were diluted

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

1:50 and 30 μ l of diluted substrate was added to the sample. A standard curve was also prepared with 2.5-fold dilution of Luc-MCM (ranging from 0.027 to 256.33 fmol). The luminescence was measured on a PHERAstar plate reader using "LUM Plus" optic module (BMG Labtech) with the read time of one second per well at 12.1 mm focal height. The background (low salt wash only) was subtracted and the amount of Luc-MCM was calculated using a linear regression of standard curve of Luc-MCM. All experiments were done in triplicates.

In the one step reaction in Fig. 2e, 40 nM ORC, 40 nM Cdc6 and 80 nM MCM/Cdt1 were used to saturate the MCM loading activity.

In the staged reaction in Fig. 2e, 40 nM ORC was preincubated with DNA beads for 15 min at 30°C with mixing at 1600 rpm. The reaction was washed twice with LSW+80 mM NaCl buffer described above. Then, 40 nM Cdc6 and 80 nM MCM/Cdt1 were added to the reaction and incubate for another 10 min.

In the inducible roadblock assay, beads were resuspended in 20 μ l binding buffer (25 mM HEPES-KOH pH7.6, 100 mM NaOAc, 10 mM MgOAc, 0.02% NP-40, 5% glycerol, 5 mM DTT and 80 mM KCl). Then, 2.4 pmol of LacI (a gift from George Cameron and Hasan Yardimci) was added and incubated for 30 mins at 30°C with mixing at 1250 rpm. The beads were washed twice and resuspended in 5 μ l of binding buffer. This mixture was subsequently added to the MCM loading reaction and incubated for 20 min. Following this, 10 mM of IPTG and a 50X molar excess of competitor DNA containing one ORC binding site (EACS.B1 F and EACS.B1 R) were added. The reaction was aliquoted at indicated time points for HSW, followed by LSW, as described above.

MCM recruitment

MCM recruitment assay was performed identically to MCM loading assay, with the exception that 5 mM ATP was replaced with 5 mM ATP γ S. After the incubation, the beads were washed twice with an ice-cold low salt wash for 20 secs.

CDK phosphorylation of ORC

For CDK phosphorylated ORC, 12 nM of ORC was incubated with 12 nM CDK for 10 mins at 30°C in loading buffer containing 0.5 mM ATP. To stop the reaction, 60 nM Sic1 was added and incubated for an additional 5 mins at 30°C. For non-phosphorylated ORC, 12 nM ORC was first incubated with 60 nM Sic1 for 5 mins at 30°C and then with 12 nM CDK for 10 mins at 30°C. The subsequent loading assay was done as described above. For Fig. 1b, proteins were separated by Criterion XT Tris Acetate 3-8% (Bio-rad) SDS-PAGE and visualized by silver staining.

Salt sensitivity assay

The salt sensitivity assay was performed identically to the MCM loading assay. After 20 min incubation, the samples were washed twice with cold LSW containing the indicated NaCl concentration, as shown in Fig. 6a, for 30 sec, followed by two washes with cold LSW.

Retention assay

Retention assay was performed identically to the MCM loading assay. After 20 min of incubation, the beads were washed with HSW (for 90bp origin only) and then cold

loading buffer (for both one site and 90bp origins) for 40 sec at 30°C with mixing at 1600 rpm. 50X molar excess of competitor DNA in loading buffer (preheat at 30°C) was then added and the mixture was aliquoted for an additional wash with HSW (for 90bp origin only), followed by cold LSW (for both one site and 90bp origins) at the indicated times.

Replication assay

The replication assays were performed according to a protocol described by Yeeles *et al.*, 2015⁴⁷. 40 nM ORC or phosphorylated ORC, 40 nM Cdc6, 60 nM MCM/Cdt1 and 4 nM DNA were incubated for 20 mins at 30°C, 1250 rpm in reaction buffer containing 25 mM HEPES-KOH pH7.6, 10 mM MgOAc, 2 mM DTT, 0.02% NP-40, 100 mM KCl, 5 mM ATP. The ORC was phosphorylated after CBP pull-down during purification and excess CDK was removed through gel filtration. 50 nM DDK was added with further incubation for 10 mins. The reaction was then incubated with 40 nM Cdc45, 30 nM Dpb11, 20 nM Pol ϵ , 20 nM GINS, 20 nM CDK, 100 nM RPA, 20 nM Ctf4, 10 nM Topo1, 30 nM Pol α , 25 nM Sld3/7, 20 nM Mcm10, 10 nM Sld2, 33 μ M ³²P-dCTP, 200 μ M NTPs and 80 μ M dNTPs for 30 mins, 30°C, 1250 rpm. The reactions were then stopped by adding 50 mM EDTA (final concentration) and filtered through an Illustra Microspin G-50 column. The replication products were resuspended in loading buffer and separated through 0.8% alkaline agarose gels in 30 mM NaOH, 2 mM EDTA for 16 hrs at 25 volts. Gels were fixed with 5% cold trichloroacetic acid and then dried onto chromatography paper (Whatman) and autoradiographed with Amersham Hyperfilm-MP (GE Healthcare). Gel images were scanned using a Typhoon phosphorimager (GE Healthcare) and were quantified using ImageJ.

EMSA

Double-stranded DNA probes for electrophoretic mobility shift assays (EMSA) were generated by PCR using primers (50bp up 209 and 50bp down 209) to give 265bp probes. A full list of primers and plasmids used for DNA templates are listed in Supplementary Materials.

62.5ng gel-purified PCR product (Nucleospin Gel and PCR clean-up, 740609.250 Macherey Nagel) was 5' end labelled in a final volume of 10 μ l using 0.6 μ l T4 PNK (10U/ μ l, M0201 New England Biolabs) with 2 μ l gamma 32P ATP 3000Ci/mmol (SRP-301 Hartmann). After 1 h on ice, 40 μ l 10 mM HEPES-KOH pH7.6, 1 mM EDTA was added and unincorporated nucleotides were removed by passing through 2x microspin G-50 columns (GE27-5330-02 Cytiva), pre-spun for 1' at 735g followed by elution for 2' at 735g.

For each EMSA reaction, 12 fmols DNA probe was used (assuming 80% elution losses per column). Reaction components were added on ice to give a final concentration of 25 mM HEPES-KOH pH7.6, 10 mM MgOAc, 100 mM NaOAc, 0.02% NP40, 5 mM DTT, 80 mM KCl, 5% glycerol, 2 μ g/ml poly dI-dC and 5 mM ATP. Lastly, 100 fmol of purified ORC was added as required and the reactions were incubated for 30 mins at 30°C in a thermomixer at 800 rpm. After 30 mins the samples were placed on ice and 2 μ l loading dye (Purple gel loading dye, no SDS, B7025 New England Biolabs) was added prior to loading 12 μ l onto a 1.5% 15 x 15 cm 100 ml 0.5X TB agarose gel. The gel was run for 1h at 200 V (4°C) before

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

fixation for 30 mins in 5% TCA and drying on a vacuum gel dryer at 55°C for 1h (Hoefer GD2000). After overnight exposure to a storage phosphor screen (BAS-IP MS 2025 E GE Healthcare) the signal was detected using a Typhoon FLA7000 phosphorimager.

Induction of re-replication

Cells were grown overnight in medium lacking methionine and containing 2% raffinose (1X YNB, 1X CSM-Met (Formedium, DCS0111)). At a density of about 1×10^7 cells/ml, 50 ng/ml alpha factor was added to arrest the cell growth for 3 hrs. Cells were spun down and switched to YPRaff medium with alpha factor for 1 hr. Cells were washed 2X with YPRaff media (1X with YPRaff and 1X with YPRaff + 50 μ g/ml Pronase E (Sigma-Aldrich, 1074330001)). Cells were resuspended in YPRaff medium containing 100 μ g/ml Pronase E with 5 μ g/ml nocodazole for 3 hours. Cells were then divided into two and 2% of glucose or galactose was added, respectively, to induce Cdc6 Δ NT for 3 hrs in the presence of galactose only. Cells were collected and lysed with lysis buffer (2% Triton, 1% SDS, 100 mM NaCl, 10 mM Tris, pH8.0, 1 mM EDTA and protease inhibitor (AEBSF, Leupeptin, Pepstatin A)). Glass beads and phenol/chloroform were added together and tubes were vortexed for 1 min. The supernatant was treated with RNases at RT for 30 min and then extracted with phenol/chloroform and ethanol precipitated. The DNA was used in qPCR. 9 μ l of qPCR reaction was set up in 384 well plates with 4.5 μ l of 2X FastStart Universal SYBR Green Master mix ROX (Roche), ~0.28 μ M of primers (Set 1: CTL141+CTL142 for re-replication region; Set 2: CTL143+CTL144 for internal control) for 10 min at 95°C and then 40 cycles of 10s at 95°C and 30s at 60°C. Each DNA sample was first evaluated for the linearity of PCR by performing serial dilutions. The fold change of threshold cycle (Ct) was determined using $2^{\Delta\Delta Ct}$ method⁴⁸, where $\Delta\Delta Ct = (C_{t\text{re-replication region}} - C_{t\text{internal control}})_{\text{galactose}} - (C_{t\text{re-replication region}} - C_{t\text{internal control}})_{\text{glucose}}$

Nucleosome flanked origins

Nucleosome-flanked origins of replication were prepared with purified yeast histone octamers as described¹¹ on the DNA substrates of N-ARS1-N or N-90bp-N. Plasmids containing N-ARS1-N and N-90bp-N were synthesised by Eurofins and used as a template for PCR to produce linear DNA substrates using the primers NCP F and NCP R. Origin substrates were amplified using standard PCR protocols and were purified by anion exchange chromatography using a 1 ml RESOURCE Q column (GE healthcare). Peak fractions were ethanol precipitated and resuspended in TE buffer.

Nucleosome assembly

Purified origin DNA substrates were combined with soluble histone octamers to form nucleosomes by salt-deposition^{11,49}. Nucleosome reconstitution was optimized in small-scale titrations and the products checked by 4% native PAGE.

Negative stain EM MCM loading assays

MCM loading experiments, negative stain sample preparation, imaging and analysis were carried out as described¹¹, with minor modifications:

Sample preparation

Nucleosome flanked ARS1/90bp origin substrates (7.5 nM) were incubated with wild-type or mutant ORC (20 nM), Cdc6 (20 nM) and MCM-Cdt1 (40 nM) with mixing (1250 RPM) in EM buffer (25 mM HEPES pH 7.6, 10 mM MgOAc, 100 mM NaOAc, 0.02% NP-40, 80 mM KCl, 1 mM DTT and 5% glycerol) and 2 mM ATP in a final reaction volume of 20 μ l. In experiments testing CDK-phosphorylated ORC, ORC was pre-phosphorylated for 10 minutes before Sic1 was added for 5 minutes to inhibit further CDK activity. Negative stain grids were prepared after 2-minutes (MO detection) and 10-minutes (DH detection) incubation at 24°C. In all cases, samples were diluted 5-fold prior to making grids.

Grid preparation

Samples were applied to glow-discharged 300-mesh copper grids with carbon film (EM Resolutions). 3 μ l of sample was applied to each grid and incubated for 1 minute. Staining was performed with four drops of 40 μ l 2% uranyl acetate, before grids were blotted to remove excess stain.

Data collection

Micrographs were collected on a Tecnai LaB6 G² Spirit transmission electron microscope (FEI) operating at 120 keV, with a 2K x 2K GATAN Ultrascan 100 camera. Images were recorded at a nominal magnification of 30K (3.45 Å pixel size), and a defocus range of ~0.5 – 2 μ m.

Image processing

All datasets were processed in Relion3.1⁵⁰ using a common processing pipeline. The contrast transfer function (CTF) of each micrograph was estimated using Gctf v1.06⁵¹. Particles were picked with Topaz⁵² using a model that had been pre-trained to pick MCM-Cdt1, ORCs, MOs and MCM-DHs in negative stain EM data. Particles were initially extracted with a 128-pixel box, rescaled to 64 pixels and were subjected to two rounds of reference-free 2D classification. All MCM containing particles were re-extracted with a 128-pixel box without rescaling. The particles were subjected to a further round of 2D classification. From this classification, particles that contributed to classes containing the MO and DH intermediates were selected and further classified, as required. Finally, particles contributing to classes that contained an MCM (total MCMs, MOs and DHs) were quantified from the appropriate 2D classification. The proportion of MCM-containing particles contributing to each intermediate state was plotted using Prism.

ReconSil

Datasets for ReconSil images were processed together in Relion3.1⁵⁰. CTF estimation was performed using Gctf v1.06⁵¹. Particles were picked with Topaz⁵² using models pre-trained to pick MCM-containing particles (as above), ORCs and nucleosomes. Particle picks were combined, duplicates removed and extracted particles were subjected to multiple rounds of reference-free 2D classification. ReconSil micrographs were generated using the command-line tool 'relion_particle_reposition' in Relion3.1 to overlay particles in the raw micrographs with the 2D averages that those particles contributed to. Individual origins were extracted from raw and ReconSil micrographs (box size: 320 pixels). Representative examples were selected from fully reconstituted origins where confident

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

assignment of co-localisation to the same origin could be made (as described in¹¹).

Cryo-EM

Image acquisition, refinement and validation statistics for the experiments below can be found in **Extended Data Table 1**.

Sample preparation

Samples were prepared as described above for the negative stain EM experiments with the following modifications:

Mutant ORC (Orc6-4A; 40 nM) was phosphorylated by CDK (10 nM) for 10 minutes in CDK buffer (25 mM HEPES pH 7.6, 10 mM MgOAc, 100 mM NaOAc, 80 mM KCl, 1 mM DTT, 2 mM ATP), before Sic1 (30 nM) was added to inhibit further CDK activity. Nucleosome flanked ARS1 origin substrates (15 nM) were incubated with pre-phosphorylated Orc6-4A mix (as above) and MCM-Cdt1 (80 nM), with mixing (1250 RPM) in cryo-EM buffer (25 mM HEPES pH 7.6, 10 mM MgOAc, 100 mM NaOAc, 80 mM KCl, 1 mM DTT) and 2.2 mM ATP in a final reaction volume of 50 μ l. The reaction was incubated at 30°C for 30 minutes before being used to prepare cryo-EM grids. 4 μ l of sample was applied to fresh graphene-oxide coated 300-mesh UltrAuFoil R1.2/1.3 grids⁵³ and incubated for 30 seconds before vitrification using a Vitrobot Mark IV (Thermo Fisher) cooled to 10°C with 100% humidity. Grids were blotted for 5 seconds and plunged into liquid ethane.

Cryo-EM data collection

Data were collected on a Titan Krios EM equipped with a K2 Summit direct electron

detector (Gatan Inc.) at the Francis Crick Institute (Structural Biology STP).

Cryo-EM image processing: SH and DH

Image processing was performed in RELION-3.1⁵⁰ (Extended Data Figs. 7 & 8). Movie stacks were aligned and motion-corrected using MotionCor2⁵⁴, retaining all frames. The contrast transfer function (CTF) of each micrograph was estimated using Gctf v1.06⁵¹. Particles were automatically picked using Topaz⁵². First, particles were picked using a general Topaz model (scale factor = 4; model = resnet8_u32; radius = 20). These particles were extracted (320-pixel box, rescaled to 80 pixels) and subjected to multiple rounds of reference-free 2D classification to isolate SH and DH particles for subsequent processing.

High-quality DH particles were subjected to iterative rounds of 3D refinement, particle polishing (incl. increasing the box size to 420 pixels), CTF refinement, and a final round of 2D classification, yielding the final EM map at an average resolution of 2.8 Å, reconstructed from 135,143 MCM-DH particles. The final RELION map was subject to density modification using resolve_cryo_em in Phenix 1.19.2⁵⁵ to generate a 2.6 Å map used for model building (Extended Data Fig. 5).

SH particles obtained from 2D classification were used for Topaz training (scale factor = 4; cnn_model = resnet8; radius = 3). SH particles were picked using the new Topaz model (selection threshold = -2), extracted (280-pixel box, rescaled to 70 pixels) and subjected to two-rounds of reference-free 2D classification. SH particles

were re-extracted (280-pixel box, unbinned) and classified into 3 classes using a 30-Å filtered reference from an ab-initio reconstruction in CryoSPARC⁵⁶. High-quality SH particles were subjected to iterative rounds of particle polishing (incl. increasing the box size to 420 pixels), CTF refinement, masked 3D refinement, and a final round of 2D classification (without alignment), yielding the final EM map at an average resolution of 3.4 Å, reconstructed from 396,366 SH particles. The final RELION map was subject to density modification using resolve_cryo_em in Phenix 1.19.2⁵⁵ to generate a 3.1 Å map used for model building (Extended Data Fig. 5).

Molecular modelling:

DH

Molecular modelling of the DH complex was performed using a density modified map generated using resolve_cryo_em. A published DH model (PDB 7P30⁵⁷) was used as an initial model and rigid-body docked into the new DH EM map. The resulting model was subject to manual modification in Coot⁵⁸ and real space refinement in Phenix 1.19.2⁵⁹. All figures were generated using UCSF ChimeraX⁶⁰.

SH

Molecular modelling of the SH complex was performed using a density modified map generated using resolve_cryo_em. An initial model of a loaded SH was extracted from the higher-resolution DH model (this study) and rigid-body docked into the SH EM map. The resulting model was subject to iterative rounds of manual modification in Coot⁵⁸ and real space refinement in Phenix 1.19.2⁵⁹. All figures were generated using UCSF ChimeraX⁶⁰.

Orc2 IDR AlphaFold2 Modelling

The MCM-ORC (MO) interaction interface, consisting of *S. cerevisiae* Orc2 (P32833), Orc3 (p54790), Orc5 (P50874, aa 319-479), Orc6 (P38826), Mcm2 (P29469, aa 178-457), Mcm5 (P29496 aa 1-336) and Mcm6 (P53091, aa 88-463), was modelled using AlphaFold2 (2.3.1¹³) implemented in AlphaFold2 (0.30.0⁶¹). The five resulting models positioned the Orc2 IDR amino acids 190-231 in the same location, wrapping around the Orc6 TFIIB-B domain. Orc3 (aa 270-419) in the top ranked model was aligned with the corresponding sequence in the previously determined MO structure (6RQC¹¹) using the matchmaker tool in ChimeraX (v1.6⁵⁹). The Orc2 IDR (aa 190-231) was fit as a rigid body into the previously unassigned density in the MO map using the 'Fit in Map' tool in ChimeraX. Fig. 1c and Extended Data Fig. 1 (AlphaFold2 figure) display EMready⁶² density modified EM maps obtained from the multibody refinement of the MO intermediate. Similar AlphaFold2 predictions of the ORC subunits alone, or the ORC subunits with the complete N-terminal domains of Mcm2-7, yielded almost identical predictions for the interaction of the Orc2 IDR with Orc6.

Simulating DH formation efficiency from two diffusing SHs

We model the separation dependent DH formation efficiency using stochastic simulations of 1-dimensional diffusion of SHs sliding along the DNA. DH formation efficiency is computed from the probability of collision of two diffusing SHs given by intersection of their diffusion

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

trajectories. The simulation consists of four key steps: SH loading, SH diffusion, SH dissociation, and DH formation,

1. SH loading:

The simulation takes place on a DNA substrate consisting of two ORC binding sequences in a head-to-head orientation with the B1 sequences of the EACS/B1 (ORC binding site) pointed towards each other. The separation between the two origins is defined as the distance between end of the B1 sequences in ORC binding sequences as shown in Extended Data Fig. 9a. ORC binds to an EACS/B1 sequence with an overhang of 6 base pairs beyond the B1^{14,63} (Extended Data Fig. 9b). This ORC complex can then load an SH ahead of itself which occupies 35 base pairs on the DNA substrate⁶⁴. This implies that the loading of one SH complex onto the DNA occupies at least 41 base pairs (6 bp ORC overhang + 35 bp SH) beyond the origin on the DNA substrate. The simulation is initialized with both ORC sites loaded with SHs with N-termini pointing towards each other.

2. SH diffusion

Since SH sliding is independent of ATP hydrolysis, we assume that its diffusion occurs at thermal equilibrium and is therefore not driven by input of external energy. In such conditions, the displacement of SH per unit time follows a gaussian distribution given by:

$$P(x, \Delta t) = \frac{1}{\sqrt{4\pi D \Delta t}} e^{-\frac{(x-x_0)^2}{4D \Delta t}} \quad \text{Eq 1}$$

where, x is the position of the SH, D is the diffusion coefficient, Δt is the duration of the time step and x_0 is the initial position. After initialization of the simulation, the displacement of SH in the following time step is determined by sampling this probability distribution using the value of SH diffusion constants measured from single molecule fluorescence tracking experiments^{31,32}.

$$x_{n+1} = x_n + P(x, \Delta t) \quad \text{Eq 2}$$

$$x_{n+1} = x_n + \sqrt{2D \Delta t} * \text{randn}() \quad \text{Eq 3}$$

where 'randn()' is the MATLAB function to generate normally distributed random numbers. The time step of simulation is chosen such that the mean squared displacement (MSD) of SH for the given value of diffusion constant is less than 1 base pair of the DNA.

$$MSD = \langle (x(\Delta t) - x_0)^2 \rangle = 2D \Delta t \quad \text{Eq 4}$$

Thus, the time step of the simulation was chosen to be 10 ms.

3. SH dissociation from the DNA.

Once loaded onto the DNA substrate, SHs dissociate from the DNA with an experimentally measured half-life of 3 minutes (Fig. 5b). We implement this phenomenon in our simulations by assigning a dissociation probability to each loaded SH onto the DNA. The dissociation probability is computed using the exponential distribution that satisfies the experimentally measured half-life:

$$\tau_{MCM SH} = -\frac{\tau_{MCM SH}^{1/2}}{\log_e(0.5)} \quad \text{Eq 5}$$

$$P_{dissociation} = \frac{1}{\tau_{MCM SH}} e^{-\frac{t}{\tau_{MCM SH}}} \quad \text{Eq 6}$$

where $\tau_{MCM SH}^{1/2}$ is the experimentally measured SH half-life and t is the total time spent by SH on the DNA since its loading. At every time step, a random number, R is sampled from a uniform distribution using MATLAB function rand(). If R is smaller than $P_{dissociation}$, the corresponding SH is said to be dissociated from the DNA, and the simulation is stopped as DH formation is no longer possible.

4. DH formation

To successfully form a DH, the two loaded SHs must collide with each other while diffusing along the DNA. We define a collision event as the first instance when the distance between the N-termini of both SHs is less than or equal to zero base pairs during the simulation. The time duration from the start of the simulation to the time step corresponding to the collision event is referred to as the collision time which also corresponds to the first-passage time of the collision process. The probability of DH formation is thus the probability of collision of two SHs loaded onto the DNA which depends on the separation between the ORC binding sites, the availability of binding site onto the DNA for the second SH, the diffusion constants of the SHs, the half-life of SH onto the DNA and the total time for the simulation.

Simulation Results

The diffusion coefficient of MCM onto the DNA has been observed to be $800 \pm 200 \text{ bp}^2/\text{s}$ from single-molecule fluorescence tracking experiments³². The half-life of SH on the DNA is measured to be 3 minutes (Fig. 5b). Using these values as our simulation parameters, we quantify the dynamics of DH formation as a function of the separation between the ORC loading sites as follows. For each case of separation, we first quantify the number of simulated trajectories that contain a collision event between two SHs and measure the time between the start of the simulation to the collision event, hereafter referred to as collision time (t_c) Extended Data Fig. 10a. The probability distribution of collision times, $P(t_c)$, is generated from all the iterations (Extended Data Fig. 10b). The cumulative sum of this distribution given by $C(t) = \sum_{t_c=0}^t P(t_c)$, refers to the probability of collisions occurring before time t , which is by definition, the probability of DH formation within time t (Extended Data Fig. 10c). Computing $C(t)$ for various separations between ORC binding sites at a fixed time (20 minutes) determined by the incubation time in bulk experiments (Fig. 5d) allows us to compare the relative DH formation efficiencies on the corresponding DNA substrates (Extended Data Fig. 10d).

We observe that the relative DH formation efficiency has a maximum at separation of 82 base pairs between the ORC binding sites, which corresponds to the case when SHs are loaded closest to each other onto the DNA (ORC overhang + SH + SH + ORC overhang = $6+35+35+6 = 82 \text{ bp}$). At separations greater than 82 bp, the DH formation efficiency decreases with increasing separation (and shows a good agreement with exponential trend) as both SHs loaded onto the DNA must diffuse along the DNA and collide with each other before either or both dissociate from the DNA.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

References

- 1 Lin, Y. C. & Prasanth, S. G. Replication initiation: Implications in genome integrity. *DNA Repair (Amst)* **103**, 103131, doi:10.1016/j.dnarep.2021.103131 (2021).
- 2 Hu, Y. & Stillman, B. Origins of DNA replication in eukaryotes. *Mol Cell* **83**, 352-372, doi:10.1016/j.molcel.2022.12.024 (2023).
- 3 Costa, A. & Diffley, J. F. X. The Initiation of Eukaryotic DNA Replication. *Annu Rev Biochem* **91**, 107-131, doi:10.1146/annurev-biochem-072321-110228 (2022).
- 4 Palzkill, T. G. & Newlon, C. S. A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell* **53**, 441-450 (1988).
- 5 Coster, G. & Diffley, J. F. X. Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science* **357**, 314-318, doi:10.1126/science.aan0063 (2017).
- 6 Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128-134 (1992).
- 7 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).
- 8 Rowley, A., Cocker, J. H., Harwood, J. & Diffley, J. F. X. 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).
- 9 Ticau, S. *et al.* Mechanism and timing of Mcm2-7 ring closure during DNA replication origin licensing. *Nat Struct Mol Biol* **24**, 309-315, doi:10.1038/nsmb.3375 (2017).
- 10 Ticau, S., Friedman, L. J., Ivica, N. A., Gelles, J. & Bell, S. P. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell* **161**, 513-525, doi:10.1016/j.cell.2015.03.012 (2015).
- 11 Miller, T. C. R., Locke, J., Greiwe, J. F., Diffley, J. F. X. & Costa, A. Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature* **575**, 704-710, doi:10.1038/s41586-019-1768-0 (2019).
- 12 Gupta, S., Friedman, L. J., Gelles, J. & Bell, S. P. A helicase-tethered ORC flip enables bidirectional helicase loading. *eLife* **10**, doi:10.7554/eLife.74282 (2021).
- 13 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
- 14 Li, N. *et al.* Structure of the origin recognition complex bound to DNA replication origin. *Nature* **559**, 217-222, doi:10.1038/s41586-018-0293-x (2018).
- 15 Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034, doi:10.1101/2021.10.04.463034 (2022).
- 16 Perkins, G., Drury, L. S. & Diffley, J. F. X. Separate SCF(CDC4) recognition elements target Cdc6 for proteolysis in S phase and mitosis. *Embo J* **20**, 4836-4845 (2001).
- 17 Drury, L. S., Perkins, G. & Diffley, J. F. X. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Curr Biol* **10**, 231-240 (2000).
- 18 Drury, L. S., Perkins, G. & Diffley, J. F. X. The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J* **16**, 5966-5976, doi:10.1093/emboj/16.19.5966 (1997).
- 19 Elsasser, S., Chi, Y., Yang, P. & Campbell, J. L. Phosphorylation controls timing of Cdc6p destruction: A biochemical analysis. *Mol Biol Cell* **10**, 3263-3277 (1999).
- 20 Moses, A. M., Liku, M. E., Li, J. J. & Durbin, R. Regulatory evolution in proteins by turnover and lineage-specific changes of cyclin-dependent kinase consensus sites. *Proc Natl Acad Sci U S A* **104**, 17713-17718, doi:0700997104 [pii] 10.1073/pnas.0700997104 (2007).
- 21 Liku, M. E., Nguyen, V. Q., Rosales, A. W., Irie, K. & Li, J. J. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. *Mol Biol Cell* **16**, 5026-5039, doi:E05-05-0412 [pii] 10.1091/mbc.E05-05-0412 (2005).
- 22 Nguyen, V. Q., Co, C., Irie, K. & Li, J. J. Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. *Curr. Biol.* **10**, 195-205 (2000).
- 23 Labib, K., Diffley, J. F. X. & Kearsey, S. E. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nat Cell Biol* **1**, 415-422 (1999).
- 24 Tanaka, S. & Diffley, J. F. X. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nat Cell Biol* **4**, 198-207, doi:10.1038/ncb757 (2002).
- 25 Yeeles, J. T., Deegan, T. D., Janska, A., Early, A. & Diffley, J. F. X. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431-435, doi:10.1038/nature14285 (2015).
- 26 Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. X. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339-343, doi:10.1038/nature11920 (2013).
- 27 Chen, S. & Bell, S. P. CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev* **25**, 363-372, doi:gad.2011511 [pii] 10.1101/gad.2011511 (2011).
- 28 Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411**, 1068-1073. (2001).
- 29 Richardson, C. D. & Li, J. J. Regulatory mechanisms that prevent re-initiation of DNA replication can be locally modulated at origins by nearby sequence elements. *PLoS genetics* **10**,

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

- e1004358, doi:10.1371/journal.pgen.1004358 (2014).
- 30 Frigola, J. *et al.* Cdt1 stabilizes an open MCM ring for helicase loading. *Nature communications* **8**, 15720, doi:10.1038/ncomms15720 (2017).
- 31 Sanchez, H. *et al.* DNA replication origins retain mobile licensing proteins. *Nature communications* **12**, 1908, doi:10.1038/s41467-021-22216-x (2021).
- 32 Sanchez, H. *et al.* A chromatinized origin reduces the mobility of ORC and MCM through interactions and spatial constraint. *Nature communications* **14**, 6735, doi:10.1038/s41467-023-42524-8 (2023).
- 33 Amasino, A. L., Gupta, S., Friedman, L. J., Gelles, J. & Bell, S. P. Regulation of replication origin licensing by ORC phosphorylation reveals a two-step mechanism for Mcm2-7 ring closing. *Proc Natl Acad Sci U S A* **120**, e2221484120, doi:10.1073/pnas.2221484120 (2023).
- 34 Hizume, K., Kominami, H., Kobayashi, K., Yamada, H. & Araki, H. Flexible DNA Path in the MCM Double Hexamer Loaded on DNA. *Biochemistry* **56**, 2435-2445, doi:10.1021/acs.biochem.6b00922 (2017).
- 35 Donovan, S., Harwood, J., Drury, L. S. & Diffley, J. F. X. Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc Natl Acad Sci U S A* **94**, 5611-5616 (1997).
- 36 Santocanale, C., Sharma, K. & Diffley, J. F. X. Activation of dormant origins of DNA replication in budding yeast. *Genes Dev* **13**, 2360-2364 (1999).
- 37 Vujcic, M., Miller, C. A. & Kowalski, D. Activation of silent replication origins at autonomously replicating sequence elements near the HML locus in budding yeast. *Mol. Cell. Biol.* **19**, 6098-6109 (1999).
- 38 Ibarra, A., Schwob, E. & Mendez, J. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc Natl Acad Sci U S A* **105**, 8956-8961 (2008).
- 39 Woodward, A. M. *et al.* Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J Cell Biol* **173**, 673-683 (2006).
- 40 Hu, Y. *et al.* Evolution of DNA replication origin specification and gene silencing mechanisms. *Nature communications* **11**, 5175, doi:10.1038/s41467-020-18964-x (2020).
- 41 Lee, C. S. K. *et al.* Humanizing the yeast origin recognition complex. *Nature communications* **12**, 33, doi:10.1038/s41467-020-20277-y (2021).
- 42 Parker, M. W. *et al.* A new class of disordered elements controls DNA replication through initiator self-assembly. *eLife* **8**, doi:10.7554/eLife.48562 (2019).
- 43 Salas-Leiva, D. E. *et al.* Genomic analysis finds no evidence of canonical eukaryotic DNA processing complexes in a free-living protist. *Nature communications* **12**, 6003, doi:10.1038/s41467-021-26077-2 (2021).
- 44 Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411**, 1068-1073, doi:10.1038/35082600 (2001).
- 45 Richardson, C. D. & Li, J. J. Regulatory Mechanisms That Prevent Re-initiation of DNA Replication Can Be Locally Modulated at Origins by Nearby Sequence Elements. *PLOS Genetics* **10**, e1004358, doi:10.1371/journal.pgen.1004358 (2014).
- 46 Frigola, J., Remus, D., Mehanna, A. & Diffley, J. F. X. ATPase-dependent quality control of DNA replication origin licensing. *Nature* **495**, 339-343, doi:10.1038/nature11920 (2013).
- 47 Yeeles, J. T. P., Deegan, T. D., Janska, A., Early, A. & Diffley, J. F. X. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431-435, doi:10.1038/nature14285 (2015).
- 48 Schmittgen, T. D. *et al.* Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. *Anal Biochem* **285**, 194-204, doi:10.1006/abio.2000.4753 (2000).
- 49 Luger, K., Rechsteiner, T. J. & Richmond, T. J. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* **304**, 3-19 (1999).
- 50 Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCr J* **7**, 253-267, doi:10.1107/S2052252520000081 (2020).
- 51 Zhang, K. Gctf: Real-time CTF determination and correction. *Journal of structural biology* **193**, 1-12, doi:10.1016/j.jsb.2015.11.003 (2016).
- 52 Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160, doi:10.1038/s41592-019-0575-8 (2019).
- 53 Martin, T. G. B., A.; Fitzpatrick, A. W. P. & Scheres, S. H. W. Graphene Oxide Grid Preparation. doi:10.6084/m9.figshare.3178669.v1 (2016).
- 54 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).
- 55 Terwilliger, T. C., Ludtke, S. J., Read, R. J., Adams, P. D. & Afonine, P. V. Improvement of cryo-EM maps by density modification. *Nat Methods* **17**, 923-927, doi:10.1038/s41592-020-0914-9 (2020).
- 56 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, doi:10.1038/nmeth.4169 (2017).
- 57 Greiwe, J. F. *et al.* Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase. *Nat Struct Mol Biol* **29**, 10-20, doi:10.1038/s41594-021-00698-z (2022).

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

- 58 Casanal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci* **29**, 1069-1078, doi:10.1002/pro.3791 (2020).
- 59 Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, doi:10.1107/s2059798318006551 (2018).
- 60 Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein science : a publication of the Protein Society* **30**, 70-82, doi:10.1002/pro.3943 (2021).
- 61 Yu, D., Chojnowski, G., Rosenthal, M. & Kosinski, J. AlphaPulldown-a python package for protein-protein interaction screens using AlphaFold-Multimer. *Bioinformatics* **39**, doi:10.1093/bioinformatics/btac749 (2023).
- 62 He, J., Li, T. & Huang, S. Y. Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nat Commun* **14**, 3217, doi:10.1038/s41467-023-39031-1 (2023).
- 63 Feng, X. *et al.* The structure of ORC-Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6. *Nature communications* **12**, 3883, doi:10.1038/s41467-021-24199-1 (2021).
- 64 Noguchi, Y. *et al.* Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1712537114 (2017).
- 65 Yuan, Z. *et al.* Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nat Struct Mol Biol* **24**, 316-324, doi:10.1038/nsmb.3372 (2017).
- 66 Baretic, D. *et al.* Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. *Mol Cell* **78**, 926-940 e913, doi:10.1016/j.molcel.2020.04.012 (2020).

Acknowledgements

We thank A. Alidoust, N. Patel, and D. Patel in the Structural Biology STP for yeast protein expression. We thank A. Nans, A. Purkiss and P. Walker from the Structural Biology STP for support with cryo-EM and computing. We thank G. Cameron and H. Yardimci for the gift of Lac repressor protein and plasmid. We thank F. Palmero Moya of the NHD lab for discussions of simulations. This work

was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001065 and FC001066), the UK Medical Research Council (FC001065 and FC001066), and the Wellcome Trust (FC001065 and FC001066). This work was also funded by Wellcome Trust Senior Investigator Awards (106252/Z/14/Z and 219527/Z/19/Z) to JFXD, and European Research Council Advanced Grants (669424-CHROMOREP and 101020432-MeChroRep) to JFXD. AC receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 820102). NHD acknowledges funding from the Netherlands Organization for Scientific Research (NWO) through grant OCENW.M.21.173 and from a European Research Council Advanced Grant (789267-REPLICHRONA). T.C.R.M. was supported by a Novo Nordisk Fonden Hallas-Møller Emerging Investigator Grant (NNF22OC0073571), the Danish National Research Foundation (DNRF115) and the Carlsberg Foundation (CF21-0571). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

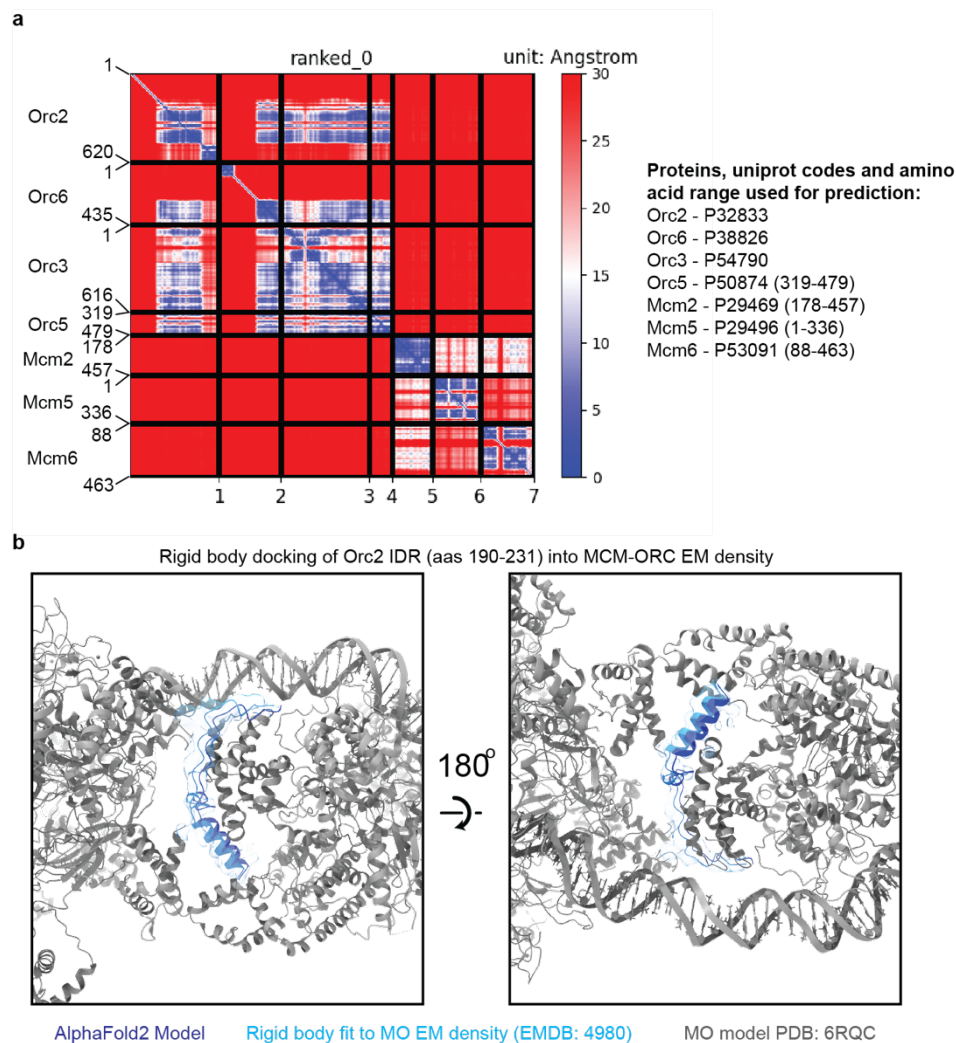
CTL and JFXD conceived the initial study. CTL performed all biochemical MCM loading experiments, analysis of re-replication *in vivo* and helped with EM experiments. TCRM performed all EM experiments, image processing and atomic model building. AC helped with AlphaFold-Multimer, cryo-EM grid preparation and screening. KWT performed *in vitro* replication experiments. AE performed the EMSA experiment. PE performed the PWM analysis of yeast origins. ST, HS, CTL, JFXD and NHD conceived the simulation experiment which was developed and executed by ST with advice from HS and NHD. CTL, TM and JFXD wrote the manuscript with input from other authors.

Data availability

Atomic model coordinates and cryo-EM maps have been deposited in the Protein Data Bank (PDB) and Electron Microscopy Data Bank (EMDB), under the accession codes 8RIF/EMD-19186 (DH), and 8RIG/EMD-19187 (SH).

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

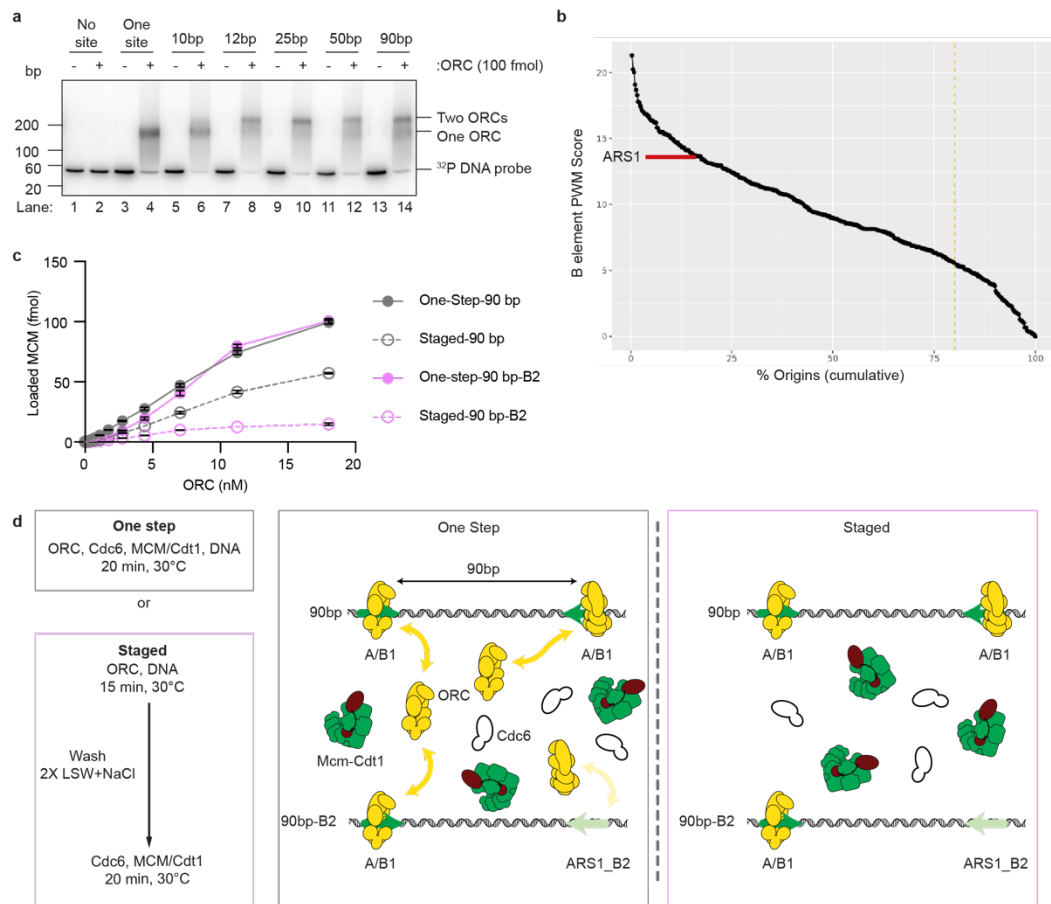
Extended Data



Extended Data Figure 1. AlphaFold2 places the Orc2 IDR at the MCM-ORC (MO) interface.

a. Predicted aligned error (PAE) plot for the AlphaFold2 model of the MO interface. The structure prediction was based on proteins present at the MO interface in the published structure (PDB: 6RQC¹¹). **b.** Following alignment of Orc3 (aa 270-419) in the top ranked AlphaFold2 model with the respective region of Orc3 in the MO structure, the predicted Orc2 IDR (aas 190-231) was fit as a rigid body into previously unassigned EM density in the MO map. Here, the density displayed has been isolated from a density modified (EMready⁶²) map of ORC obtained from the multibody refinement of the MO (EMDB: 4980¹¹). The prediction places the Orc2 IDR as wrapping around the Orc6 TFIIIB domain at the center of the MO interface.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 2. The formation of MO stabilizes ORC binding at weak binding sites

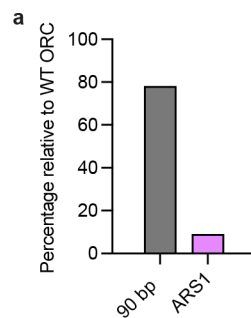
a. EMSA of ORC binding at different distances on synthetic templates. The 10bp distance is similar to ARS1, which can only bind a single ORC. Two ORCs can bind simultaneously on origins with a spacing of ≥ 12 bp. **b.** Approximately 15% of origins have a better match to the EACS at the weaker ORC binding site than ARS1 B2. **c.** Near-saturating amounts of MCM are loaded onto both 90bp and 90bp-B2 origins in one-step reactions using 18 nM of ORC. **d.** Schematic figure of 'One Step' and 'Staged' reactions, where excess ORC is always available for the one step reaction, but ORC only remains bound to high affinity sites after stringent washing with 80 mM NaCl in the staged reactions.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

<i>Saccharomyces cerevisiae</i> /1-620	1-----MLNGDFEHNIDLSKAKSRNVTPKRVDPHGE-----RQLRRIHSSKKNLLERISLVGNERKNTSPDALKP	68
<i>Kluyveromyces lactis</i> /1-598	1MTKKERSDNEEEEEEDIDINNDQIFISKSVRYQSPKKLKSDFLSPS-KTVRPKSPKRSQTLKSVSF-----ETLVKDEG---	77
<i>Eremothecium gossypii</i> /1-579	1-----MRAREIFSA-----RNCPAWSCAERMDGQDAEAGAPGDIWAHA-----DLCVABAGAGAP	51
<i>Zygosaccharomyces rouxii</i> /1-603	1-----MSDDGIVKNDILSPKARKGVTCRVVSKRRSTDPVDPGLDRKRSPKKNLLERIE-----ERSPTCGGIWQE	66
<i>Torulaspora delbrueckii</i> /1-588	1MFGYIVREIYEVVMNNDIEDIVKNDILSPKARKGVTCRVVSKRRSTDPVDPGLDRKRSPKKNLLERIE-----ETEEVVEE---	70
<i>Naumovozyma castellii</i> /1-730	1-----MSEGPDDIVTHIDILCSAVKSSNKAIRNPALAHIDSPNKK---SRVSTKRPALETVKLL-----ETTETIGKFGVGP	68
<i>Saccharomyces cerevisiae</i> /1-620	69-----KTPSKAPRRKGRPRKIQEELTDRIKKDEKDTISSKK-----	104
<i>Kluyveromyces lactis</i> /1-598	78-----NRIHSSDTDTATRRRQRKTTTRVKVLEDDAIK-----	109
<i>Eremothecium gossypii</i> /1-579	52-----SSKFLRFYRAATAARRASRRGRASPRRGAA-----	84
<i>Zygosaccharomyces rouxii</i> /1-603	67-----SSDLSSSESPGKRRKRSERRTKSDNDKKEEE-----	99
<i>Torulaspora delbrueckii</i> /1-588	71-----VADEVADEQADAEVVRGELQ-----	88
<i>Naumovozyma castellii</i> /1-730	69GEP I I K R K R G R P R K I V T S E P P K L K R K R G R P R K E E Q H Y K E E D E S E G T M K E R S M S P L S A K I S S D N I L T S A K R T R G I K I T V Q L S S A S	154
<i>Saccharomyces cerevisiae</i> /1-620	105-----KRKLDKDTSGNVNEESKTSNNKQVMETGICEKREKERIQVATTITYEDNVTPQTDNDFVSNSEPEPEPATPSK	177
<i>Kluyveromyces lactis</i> /1-598	110-----YVSDTNELIHDDDEYE-----DLKIGIVSKESSPNKSPRRKLTFFNAEPPSPQRKTT	159
<i>Eremothecium gossypii</i> /1-579	85-----DSSVDEGATAVPSRRDAAE-----EESPRRKQPVVRP-QTPAKLAAPALAEADSDGFVTARSTP--LRSPR	147
<i>Zygosaccharomyces rouxii</i> /1-603	100-----DNNDGGC-DAYVSNNGDEEDKLDRENKRHRKLP SRSETPIPVLASPVKTTESENEVSIITTPQ--KKTPK	165
<i>Torulaspora delbrueckii</i> /1-588	89-----EETEGG-----LETPRKRK-----RTKIEVKLEPVEPVELEQEVKVLDEVPSPSKKPI	137
<i>Naumovozyma castellii</i> /1-730	155EADSSDDDTFTSAKEEVEESTEEENDGCFEENEVKEEIEFTKRRKGRPRKYP IPETPKIKRRKGRPRKEPVI SENNSECKSKRP	240
<i>Saccharomyces cerevisiae</i> /1-620	178K-----SLTTNHDFTSPLKQIMNNLKEYKDS--SPGKLTLSRNFPTTPVPKNN--KLY-QTSE	232
<i>Kluyveromyces lactis</i> /1-598	160R-----NTVQRNKIIDDDLAYDDIQYKHLKLNLPNFKP KLPKSDNND--QTSD	208
<i>Eremothecium gossypii</i> /1-579	148R-----QLTFSSPSKS--PRKTRKQDVLQAQVPSFKLPDPFVPAAPVPGY-----RLPE	194
<i>Zygosaccharomyces rouxii</i> /1-603	166K-----TLSTEHDFTSPLKQVLSNLDKEYKNTA--SENKLSRNFPVTPPLTAK--EKYVPTTT	221
<i>Torulaspora delbrueckii</i> /1-588	138K-----STSEVHDFTSPLKSVLSNLDKEYKNTA--SSQSLKLSRNFPVTPPLPNN--DRF-KFQP	192
<i>Naumovozyma castellii</i> /1-730	241KKGESSDIESDRESLAMDQTPSPKSIAPLTHEHDFTSPLKQVIMNSLRQYQEDSLLSSKKLVLRNREVPVQVPTDKYEKRLEASE	326
<i>Saccharomyces cerevisiae</i> /1-620	233TKASSFLDTFEGYFDQR--KIVRTNAKSRHTMSMAPDVTRFEESLVSNFFNENFQKRPRQKLFIEIQKMFPPQYWFELTQGFSLF	316
<i>Kluyveromyces lactis</i> /1-598	209EQSQIYFFDGFEGFIDQT--KILKTDKRSKNSMSSAPSRSDREYNVLSQSLNDIFHRSTRVQKIHETLFFQFTFELLOGFTLLF	292
<i>Eremothecium gossypii</i> /1-579	195DRNLTYFFDGFEGYIDQK--KPIRAHKSSRNSMAIAPSVSRDRELLSGALNGFMHQTPRAVLTGHQRLFPQYWFELMOCFTLLF	278
<i>Zygosaccharomyces rouxii</i> /1-603	222DKSMSSFFDTFEGYFDQR--KQLRGSQKSKNTMSMATDTRREEFALISNTFHGKLHRAARDGLYEITQTRLFPQYWFELTQGFSLF	305
<i>Torulaspora delbrueckii</i> /1-588	193EKAVSSFFDTFEGYFDQR--KNTRGKLSKNTMSMATDTRREEFALISDI FHKKLHRTARDGLYEITQTRLFPQYWFELTQGFSLF	276
<i>Naumovozyma castellii</i> /1-730	327NKSSKTFDTFEGYFDQRSLSKYTHSKNTMSMCPVSKSODEFAMFTNVFNKHLYLKEPRAKLQALQRKMFPPQYWFELTQGFSLF	412
<i>Saccharomyces cerevisiae</i> /1-620	317YGVGSKRNFLFEFAIDYLSPKIAYSQ-----LAYENELQKNKPVNSTPCLILNGYNPSCNRYDVFKEITDLVPAEL--TRSET	393
<i>Kluyveromyces lactis</i> /1-598	293YIGGSKKFLFESFAFNLSMKIALLG-----NPDIQKDDPCIPVFMNGVTTLS--KNIFLDIFKILMEGETIDDOGGQS	364
<i>Eremothecium gossypii</i> /1-579	279YIGGSKRRFLETLVFEYLSPKLALSG-----APEGADVEGVPCVVIINGFNPAENYRDCQSIAAYMLPDEL--RHAET	349
<i>Zygosaccharomyces rouxii</i> /1-603	306YGVGSKREFLERYVFEYLAPNLVS-----DESSVGVPVGVVNGYNPTCNRYDVFKEISRALLPDEL--SRNET	371
<i>Torulaspora delbrueckii</i> /1-588	277YIGGSKREFMEKYAFEYLVNPLAYSQ--MHKQVAVPKKQLDINDYNTPCIVINGYNPTCNRYDVFKEISAAALLPEL--TRNET	357
<i>Naumovozyma castellii</i> /1-730	413YIGGSKRNFLDLAIKIMISPKLTFGKRIALQKLELKKVKNKSTGSIQGPCLILINGYNPTCYREIFQDITDIMEFDGL--TRNET	496
<i>Saccharomyces cerevisiae</i> /1-620	394KYWGNHVLQIQKMIIDFYKNDPLDIKLLVHNLDCPSIRKNTFOTMLSFLSVIRQIAIVASTDHIYAPILWDMKKAQNYNFVFDH	479
<i>Kluyveromyces lactis</i> /1-598	365NYWDMRIDLQMDRLSKYFKRPPFTVKMILLIHNLDGCPFRREHFOTRMSILAQIKQICVVASVDHIQAPFLWDFHRAQSTNFVYFDH	450
<i>Eremothecium gossypii</i> /1-579	350KYWHNHVHLQIQKMIIDFYKNDPLDIKLLVHNLDCPMLRKPFOQMLSSARIRQIAIVASVDHIYAPILWDSLRQILNFVFDH	435
<i>Zygosaccharomyces rouxii</i> /1-603	372KYWGNHVLQIQKMIIDFYKNDPLDIKLLVHNLDCPSLRKDNFOTMLCYLALIRQVALVASTDHIYAPILWDMKKAQNYNFVFDH	457
<i>Torulaspora delbrueckii</i> /1-588	358KYWGNHVLQIQKMIIDFYKNDPLDIKLLVHNLDCPSLRKDNFOTMLCYLALIRQVALVASTDHIYAPILWDMKKAQNYNFVFDH	443
<i>Naumovozyma castellii</i> /1-730	497KFWGNLAILQVQKMIIDFYKNDPLDIKLLVHNLDCPALRKEIFQTIFFSLAQIKQIVLASVDHIYAPILFDMKKAQNYNFVFDH	582
<i>Saccharomyces cerevisiae</i> /1-620	480ISNFEPTVESTFOD-VMKMGKSDTSFG-AEGAKYVLSLTVNSKKMYKLLIETQMOMNMGNSANTG-----PKRGRTORTGVEL	556
<i>Kluyveromyces lactis</i> /1-598	451ITNFEPTVETVQTDNAIDFSKDSGTLFNANGAKYVLSLTVNSKKMYKLLIETQMOMNMGNSANTG-----SSLNKIPGIEF	533
<i>Eremothecium gossypii</i> /1-579	436VTNYPEPAIEAAEQE-SYRLNRSLEQAISDAARYVLSLTVNSKKMYKLLIETQMOMNMGNSANTG-----SAKRKILTN--SRAGISFGVPI	515
<i>Zygosaccharomyces rouxii</i> /1-603	458VSNYPEPAIEAAEQE-SYRLNRSLEQAISDAARYVLSLTVNSKKMYKLLIETQMOMNMGNSANTG-----SAKRKILTN--SRAGISFGVPI	539
<i>Torulaspora delbrueckii</i> /1-588	444VTNYPEFTVESSFRD-VMKMKSSDASG-AEGAKYVLSLTVNSKKMYKLLIETQIRNME-KSTNSKGVAP--TKRGSPVTCVEL	524
<i>Naumovozyma castellii</i> /1-730	583VTNYPEFTVESSFRD-KMNLGKSSSSG-AEGAKYVLSLTVNSKKMYKLLIETQIRNME-LASGSKNKSASSVSKRGGLSTGVEF	665
<i>Saccharomyces cerevisiae</i> /1-620	557KLFNHLCAADFIASNEIALSRMTEFIEHKMANITKNNSGMEIIVWPPTYAELEKLLKTVLNTL-	620
<i>Kluyveromyces lactis</i> /1-598	534SELFKTCTEQFVVSSEMLRTLISEFIDHKMANVNTKNMGKEIIVWKYSYGDMEKMITIDSMLQ	598
<i>Eremothecium gossypii</i> /1-579	516SAFYQACAAQFVASNEMSLRSMLREFVEHKMAHLADKAGQEIIVYVNYSGFEMQKLLSDALSSV-	579
<i>Zygosaccharomyces rouxii</i> /1-603	540KQLLHLCAAEFIASNEIALSRMTEFIEHKMAAISKNPAGTEHVWIPYNYSEINKLLSTVMQSI-	603
<i>Torulaspora delbrueckii</i> /1-588	525KQLLHLCAAEFIASNEIALSRMTEFIEHKMANLKNPAGTEHVWIPYNYSEMTKLYSTVLKEL-	588
<i>Naumovozyma castellii</i> /1-730	666KSFVTMCASEFIASNEIALSRMTEFIEHKMALLSRHSSGETEYLYVYPRYSEMEKILSTVLNNIE	730

Extended Data Figure 3. Closely related yeast species contain multiple potential CDK phosphorylation sites in the Orc2 IDR.

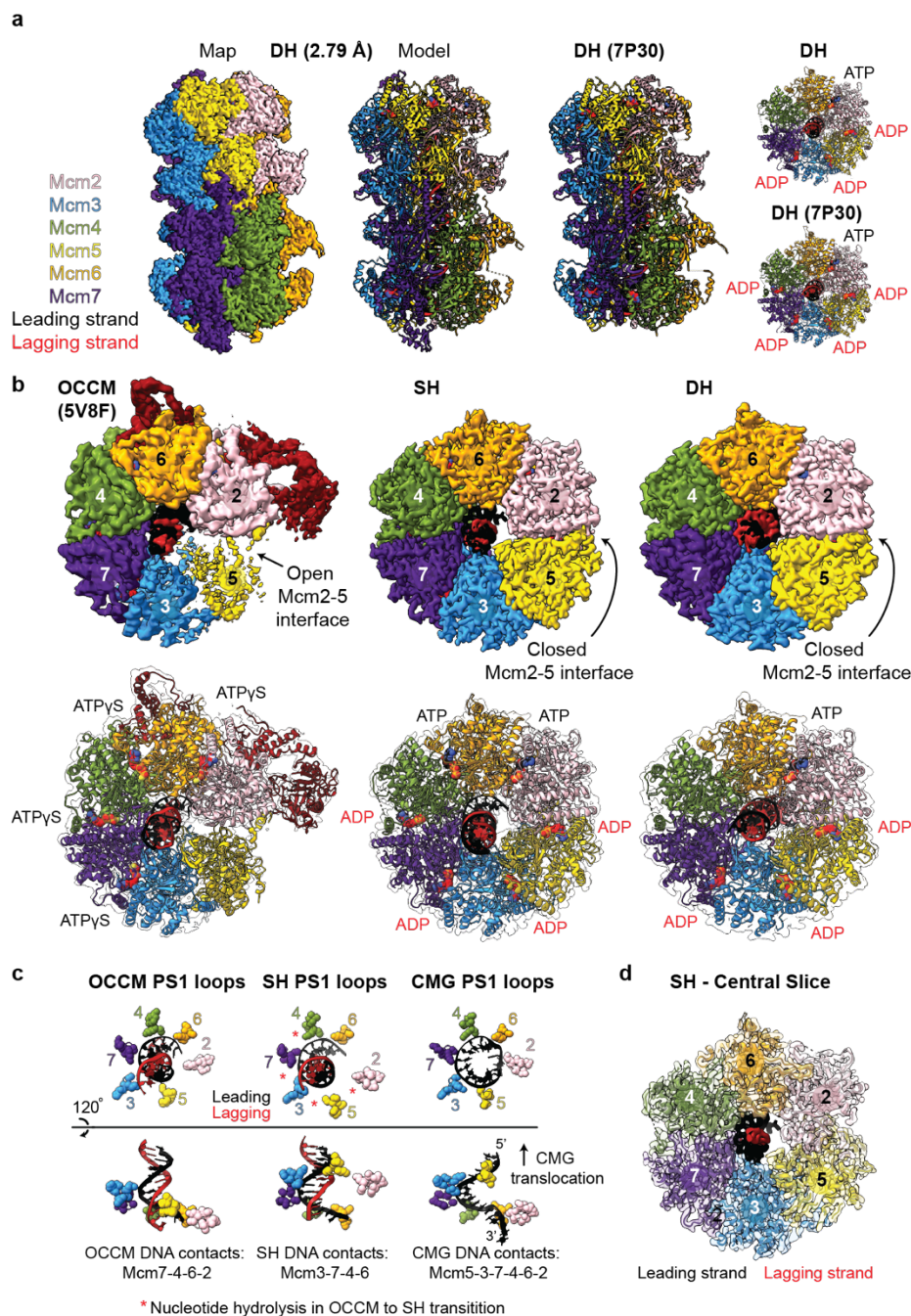
The region of the Orc2 IDR that is predicted to form part of the MCM-ORC interface (yellow box) is moderately conserved and harbors multiple potential CDK phosphorylation sites.



Extended Data Figure 4. CDK phosphorylated ORC is replication competent in vitro

a. The percentage of DNA synthesis for CDK-phosphorylated ORC relative to WT ORC as shown in Fig. 3c.

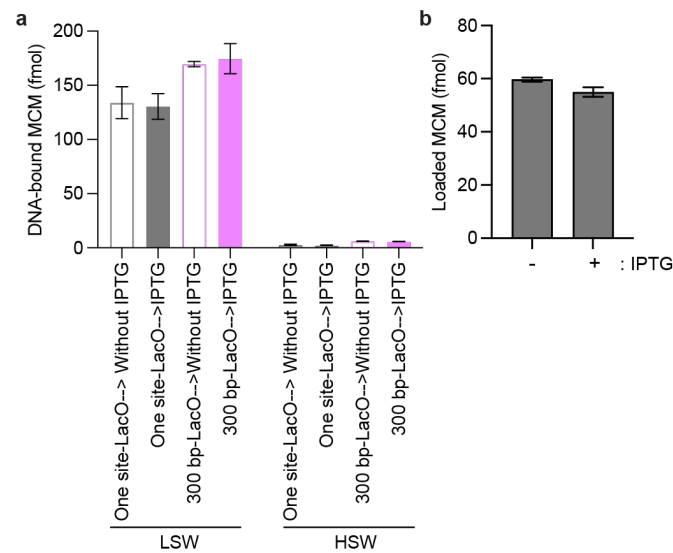
Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 5. Comparison of DHs assembled by ORC and ORC phosphorylated on Orc2.

a. Cryo-EM structure and atomic model for DHs assembled by ORC phosphorylated on Orc2 (left; this study) and atomic model for DHs assembled by unphosphorylated ORC (right; PDB: 7P30⁵⁷). The DH EM density map displayed has been density modified using EMReady⁶². **b.** Comparison of the MCM ring in the OCCM (left), SH (middle) and DH (right), showing that the SH and DH have fully closed Mcm2-5 interfaces and both appear to be in post-ATP hydrolysis states relative to the OCCM. **c.** Comparison of the PS1 loops of the OCCM (5V8F⁶⁵; left), SH (this study; middle) and CMG (6SKO⁶⁶; right), showing that the SH PS1 loops adopt a staircase configuration and engage the backbone of the leading strand DNA template like the CMG. **d.** Central slice through the SH structure, highlighting the well resolved EM density for the bases in the leading strand DNA template.

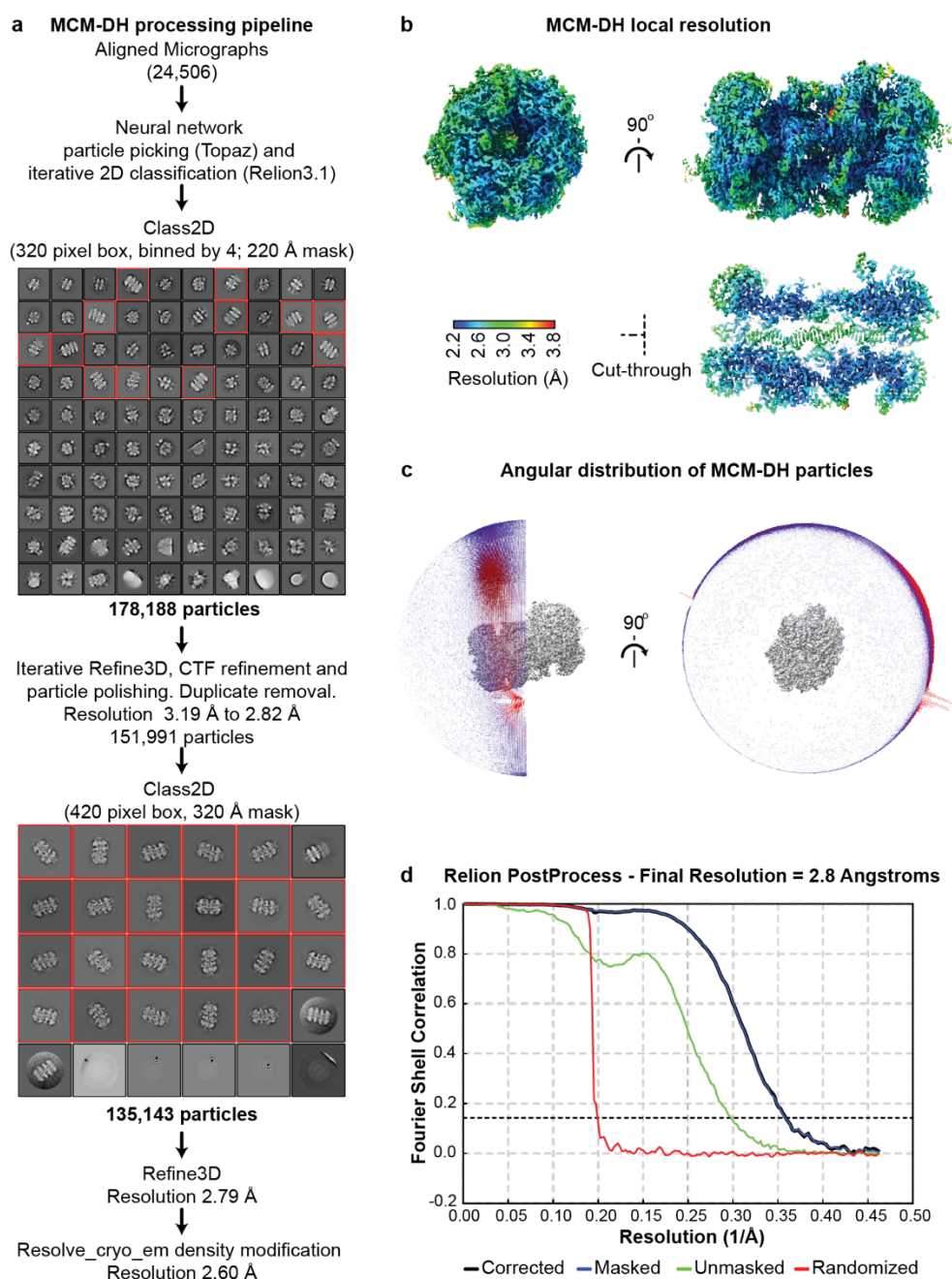
Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 6. DH formation by two single hexamers

a. MCM loading assay performed with 20 minutes incubation of Orc2ΔN, Cdc6 and MCM-Cdt1 with ATP and a DNA template containing two high affinity ORC sites separated by 300bp and 4 lac operators. Aliquots taken before IPTG addition show MCM loaded onto DNA (detectable after LSW), but no HSW stable DH formation. **b.** The addition of IPTG in the MCM loading assay does not affect MCM loading.

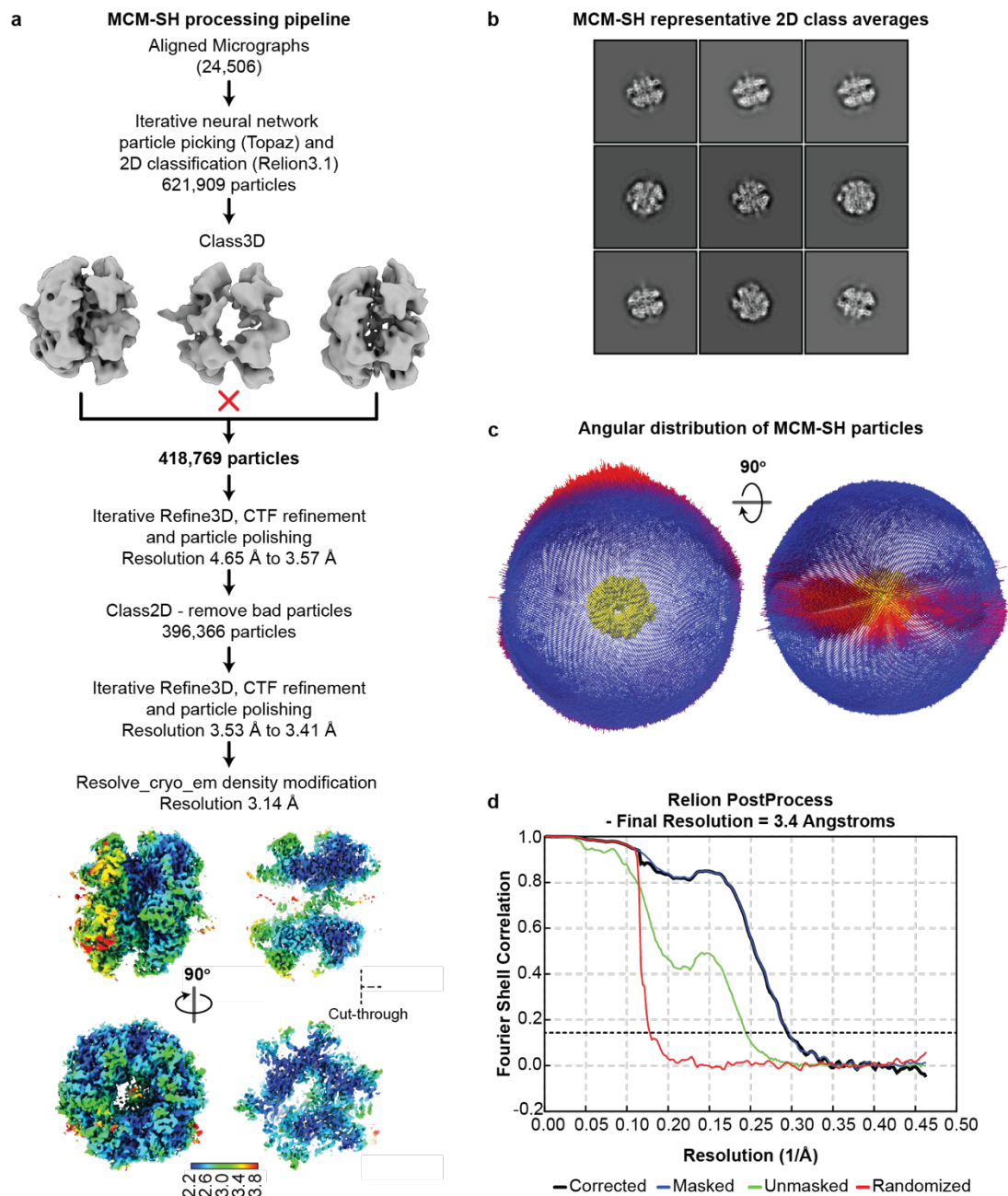
Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 7. DH cryo-EM data processing.

a. Cryo-EM data processing pipeline for the DH. **b.** Final DH EM density map following density modification with resolve_cryo_em⁵⁵, coloured according to the local resolution. **c.** Angular distribution plot. **d.** Fourier shell correlation (FSC) plot obtained from Relion PostProcess.

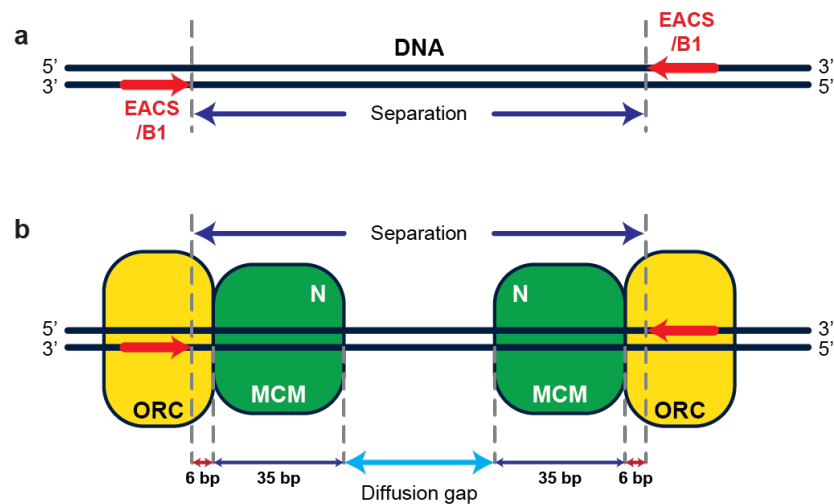
Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 8. SH data cryo-EM processing.

a. Cryo-EM data processing pipeline for the SH, including views of the resulting 3D density modified map coloured according to local resolution. **b.** Representative 2D classes for the SH. **c.** Angular distribution plot. **d.** Fourier shell correlation (FSC) plot obtained from Relion PostProcess.

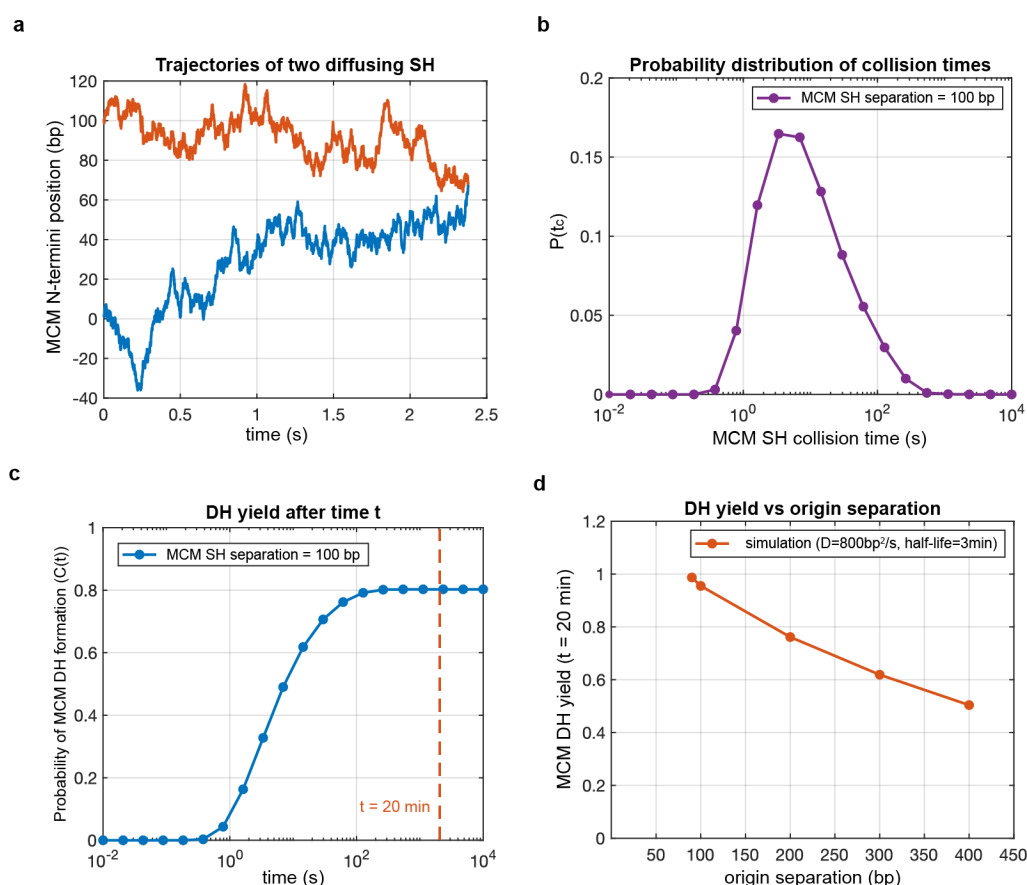
Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 9. Schematic of the simulated system and definition of relevant distances

a. The separation between two ORC binding sites oriented towards each other on DNA is defined as the distance in base pairs between the end of the B1 sequences in the EACS/B1 ('ORC binding sites'). **b.** ORC binds to the EACS/B1 sequences with an overhang of 6 bp beyond the end of the B1 sequence. A SH binds in front of the ORC and occupies 35 bp with the distance between the origin and the N-termini of SH being $6+35 = 41$ bp. The distance between the N-termini of SHs is the diffusion gap that must be overcome to result in DH formation and is determined by the separation between origin sequences and the space occupied by ORC and MCMs on the DNA.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function



Extended Data Figure 10. Simulating efficiency of DH formation

a. an example of a simulated traces showing two MCM-SHs loaded at a separation of 100 base pairs which then diffuse along the DNA substrate and ultimately collide with each other resulting in MCM-DH formation. **b.** Probability distribution of MCM-SH collision times. **c.** Probability of MCM-DH formation with respect to time which is the Cumulative sum of the MCM-SH collision time distribution shown in (b). MCM-DH yield is defined as the MCM-DH formation probability at a cutoff time determined by the incubation time in vitro experiments (in this case $t_{\text{cutoff}} = 20$ min). **d.** Simulated MCM-DH yield at various values of separation between origin sequences.

Extended Data Movie 1. Model for DH formation via rotation-coupled diffusion of SHs

Two MCM single hexamers diffusing via rotation along B-form DNA. By using the first steric clash between hexamers as the end point of rotational diffusion, an inter-ring register is obtained which appears close to the register observed in the MCM double hexamer.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

Extended Data Table 1.

	MCM-SH (EMDB-19187) (PDB-8RIG)	MCM-DH (EMDB-19186) (PDB-8RIF)
Data collection and processing		
Magnification	130,000	130,000
Voltage (kV)	300	300
Electron exposure (e-/Å ²)	51.4	51.4
Defocus range (μm)	1.0 – 5.0	1.0 – 5.0
Pixel size (Å)	1.08	1.08
Symmetry imposed	C1	C2
Initial particle images (no.)	621,909	178,188
Final particle images (no.)	396,366	135,143
Map resolution (Å)		
RELION (PostProcess)	3.41	2.8
FSC threshold	0.143	0.143
Resolve Cryo-EM	3.14	2.60
FSCref threshold	0.5	0.5
Map resolution range (Å)	2.2-4.0	2.2-3.8
Refinement		
Initial model used (PDB code)	8RIF	7P30
Map resolution (Å)	2.6	3.0
FSCref threshold	0.5	0.5
Map sharpening <i>B</i> factor (Å ²)	Resolve Cryo-EM (27.3)	Resolve Cryo-EM (0)
Model composition		
Non-hydrogen atoms	30,221	64,145
Protein residues	3,721	7,820
Nucleotides	38	106
Ligands	2 ATP, 4 ADP,	2 ATP, 6 ADP,
(ATP, ADP, Mg ²⁺ , Zn ²⁺)	6 Mg ²⁺ , 5 Zn ²⁺	8 Mg ²⁺ , 10 Zn ²⁺
<i>B</i> factors (Å²)		
Protein	65.37	61.85
Nucleotide	20.00	102.70
Ligand	57.35	63.70
R.m.s. deviations		
Bond lengths (Å ²)	0.007	0.007
Bond angles (°)	1.117	1.232
Validation		
Molprobity score	1.26	1.30
Clashscore	2.04	1.81
Poor rotamers (%)	0.97	0.98
Ramachandran plot		
Favored (%)	95.88	94.87
Allowed (%)	3.96	4.98
Disallowed (%)	0.16	0.16

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

Supplementary Materials

EcoRI site

XmaI site

Matches to the ORC consensus sequence on the top or bottom strand

LacO

Nucleosome binding site

Abf1

These sequences were cloned in pUC19 at EcoRI and XmaI sites

pCTL40 10bp

gaattcCTCGA TTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGTGGCGAAAAGCGTCGAAGCGG AACT
AAACATAAAAAAAGCGGTGTTTCGCCGAGATCGGCTCGCGCATGGCCGAGTTGAGCGGTTCCAGGCTGGCCACGCAG
CATCAGATGGTAGGCCTCCTGGCGCCGACCGGCCTCAGCATCCGGTACCTCAGCTGGCCACCGTCGGCGTCTCGC
ACGACCACCAAGTGAAGGGTCTGAGCAGCGCGCTCGTGCTCCTCGGAGTGGAGGCAGCCGAGCGCGACGGTGTGC
CCGCTTCCTGGAGACCTCCGCGCTCCGCAACCTCCACTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGT
CGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCTCGCCACACGACCCGCA
GCGCCCGACCGAAAAGGAGCGCACGA cccggg

pCTL9 12bp

gaattcCTCGA TTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGTGGCGGTAAAAGCGTCGAAGCGG AA
CTAAACATAAAAAAAGGTGTTTCGCCGAGATCGGCTCGCGCATGGCCGAGTTGAGCGGTTCCAGGCTGGCCACGCAGC
ATCAGATGGTAGGCCTCCTGGCGCCGACCGGCCTCAGCATCCGGTACCTCAGCTGGCCACCGTCGGCGTCTCGCA
CGACCACCAAGTGAAGGGTCTGAGCAGCGCGCTCGTGCTCCTCGGAGTGGAGGCAGCCGAGCGCGACGGTGTGCC
CGCCTTCCTGGAGACCTCCGCGCTCCGCAACCTCCACTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTC
GAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCTCGCCACACGACCCGCGAG
CGCCCGACCGAAAAGGAGCGCACGA cccggg

pCTL86 ARS317

gaattcACTAGTACTTAAAAAACTGTAGTTTCAGTGCAAAAAAGTTTTAACATTACGTATCTTGTACCCTTTTTAT
TGCATATAGAAAAGGTCAAATAATCCTTCACATCATGAAATATAAGCTAAATCGCATTTCTTTTCGTCCACATTTGCAAAC
AAAACTTTTCAATAATAATTTTATAAATAGTATCAATATATATATATATATATATTTATTTGTTTACTTTTCTATCAGTG
TTTTCAATTTTTTATTAACAATGTTTGATTTTTTAAATCGCAATTTAAT ACCTAAATATAAAAAAT GTTATTATATTGCAA
AAACCCATCAACCTTGAAAAAAGTAGAAACGTTTTATTTAATTCATCAATACATCATAAAATACGAACGATCCCCGTC
CAAGTTATGAGCTTAATCTTCATAAAAAATTTTGAAGCAATAGATCATGTACTAACTAAATCAGGGAAATT cccggg

pCTL91 90bp-B2

gaattcCTCGA TTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGTGGCGGTCTGGACCACGCCGGAGA
GCGTCGAAGCGGAGGCGGTGTTTCGCCGAGATCGGCTCGCGCATGGCCGAGTTCAAAGGCCTGCAGGCAAGTGCA
CAAACAATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTTTGACGAAATTCGGTCGGCGTCTCGCAC
GACCACCAAGTGAAGGGTCTGAGCAGCGCCGTCGTGCTCCTCGGAGTGGAGGCAGCCGAGCGCGACGGTGTGCC
GCCTTCCTGGAGACCTCCGCGCTCCGCAACCTCCACTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCG
AGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCTCGCCACACGACCCGCGAGC
GCCCGACCGAAAAGGAGCGCACGA cccggg

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

pCTL93 300bp+LacO

gaattcCTCGAATTTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGTGGCGGTCTGGACCACGCCGGAGA
GCGTCGAAGCGGAGGCGGTGTTTCGCCGAGATCGGGAGCTCTCACACCTACAAGGGATGTACATCAATTGTGAGCGG
ATAACAATTGTTAGGGAGGAATTGTGAGCGGATAACAATTTGGAGTTGATAATTGTGAGCGGATAACAATTGGCTTCA
ACGTAATTGTGAGCGGATAACAATTTCCGTACGAATGTGCCGAACCTTATGGTACCGCCGAGCGCGACGGTGTGCCCG
CCTTCCTGGAGACCTCCGCGCTCCGCAACCTCCACTTCTAAAAGCGGCTCGGCTTCAACTAAACATAAAAAATACGAGG
TGCCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCTCGCCACACGACCCCGCAGCGCC
CGACCGAAAGGAGCGCACGAaccggg

pCTL94 One site+LacO

gaattcCTCGAATTTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGTGGCGGTCTGGACCACGCCGGAGA
GCGTCGAAGCGGAGGCGGTGTTTCGCCGAGATCGGGAGCTCTCACACCTACAAGGGATGTACATCAATTGTGAGCGG
ATAACAATTGTTAGGGAGGAATTGTGAGCGGATAACAATTTGGAGTTGATAATTGTGAGCGGATAACAATTGGCTTCA
ACGTAATTGTGAGCGGATAACAATTTCCGTACGAATGTGCCGAACCTTATGGTACCGCCGAGCGCGACGGTGTGCCCG
CCTTCCTGGAGACCTCCGCGCTCCGCAACCTCCACTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTGCA
GGTGCCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCTCGCCACACGACCCCGCAGCG
CCCGACCGAAAGGAGCGCACGAaccggg

pTM0005 N-ARS1-N

ccgccctggagaatcGcggtgcgagggccgctcaattggtcgtagacagcttagcaccgcttaaacgcacgtacgcgctgtcccccgcgttttaaccgcaa
ggggattactccctagctccaggcacgtgtcagatatatacatcctgATTTTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAG
GCCTGCAGGCAAGTGCACAAACAATACTTAAATAAATACTACTCAGTAATAACCTATTTCTAGCATTTTGTACGAAATT
TGCTATTTTCCAGTTTCGCGCGCCACCTACCGTGTGAAGTCGTCACTCGGGCTTCTAAGTACGCTTAGCGCACGGT
AGAGCGCAATCCAAGGCTAACCACCGTGCATCGATGTTGAAAGAGGCCCTCCGTCCTTATTACTTCAAGTCCCTGGG
GTgcagg

pTM0023 N-90bp-N

ccgccctggagaatcGcggtgcgagggccgctcaattggtcgtagacagcttagcaccgcttaaacgcacgtacgcgctgtcccccgcgttttaaccgcaa
ggggattactccctagctccaggcacgtgtcagatatatacatcctgATCTCGAATTTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGT
GGCGGTCTGGACCACGCCGAGAGCGTGTTCAGGCTGGCCACGCAGCGAAGCGGAGGCGGTGTTTCGCCGAGATC
GGCTCGCGCAAAGCCGAGTTGAGCGAACTAAACATAAAAAATACAGCAACCCAGTTTCGCGCGCCACCTACCGTGTGAA
GTCTCACTCGGGCTTCTAAGTACGCTTAGCGCACGGTAGAGCGCAATCCAAGGCTAACCACCGTGCATCGATGTTG
AAAGAGGCCCTCCGTCCTTATTACTTCAAGTCCCTGGGGTgcagg

pCTL68 ARS1-Abf1

catatgcggtgtgaaataccgcacagatcgtaaggagaaaataccgcacatcaggcgccattcgccattcaggctgcgcaactgttggaaggcgatcggtg
cgggcctcttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagtgggtaacTATTTCCTAGCATTTTGACGAAATTgtaaaacg
acggccagtgaattcATTTTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAGGCCCTGCAGGCAAGTGCACAAAC
AATACTTAAATAAATACTACTCAGTAATAAACCATTTCCTAGCATTTTGACGAAATTGTCTATTTTGTAGAGTCTTTTA
CACCATTGTCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCATCACCACATTTTCTGG
CGTCAGTCCACAGCTAATCAATAAATGTAAGCTTTTCGGGGCTCTCTTGCTTCCAACCCAGTCAGAAATCGAGTTCCA
ATCCAAAAGTTACCTATCCCACTGCTTCTGAATCAAACAAGGGAATAAACGAATGAGGTTTCTGTGAAGCTGCACT
GAGTAGTATGTTGCAaccggggatcc

pCTL67 90bp-Abf1

catatgcggtgtgaaataccgcacagatcgtaaggagaaaataccgcacatcaggcgccattcgccattcaggctgcgcaactgttggaaggcgatcggtg
cgggcctcttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagtgggtaacccagggtTATTTCCTAGCATTTTGACGAAATT
acggccagtgaattcCTCGAATTTTTATGTTTAGTTTCGCGGACGACGGTTTCGAGGTGGCGGTCTGGACCACGCCGGAGA
GCGTCGAAGCGGAGGCGGTGTTTCGCCGAGATCGGCTCGCGCATGGCCGAGTTGAGCGGTTCAAACTGGCCACGC
AGAACTAAACATAAAAAATCCTGGCGCCGCTATTTCCTAGCATTTTGACGAAATTCTGGCCACCGTCTGGCGTCTCGCA
CGACCACAGTGCAAGGCTGTGAGCAGCGCCGTCGTGCTCCTCGGAGTGAGGCAGCCGAGCGCGACGGTGTGCC
CGCCTTCTGGAGACCTCCGCGCTCCGCAACCTCCACTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTC
GAGGTGCCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCTCGCCACACGACCCCGCAG
CGCCGACCGAAAGGAGCGCACGAaccggggatcc

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

pCTL69 ARS317-Abf1

catatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgcacaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtg
cgggcctcttcgctattacgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaac **TATTCTTAGCATTTTGACGAAATT**gtaaaacg
acggccagtgaattcACTAGTACTTAAAAAACTGTAGTTTCAGTGCAAAAAAGTTTTAACATTACGTATCTTGTACCCCTTTTAA
TTGCATATAGAAAGGTCAAATAATCCTTCACATCATGAAATATAAGCTAAATCGCATTTCTTTTCGTCCACATTTGCAAA
CAAACTTTTCAATAATAATTTTATAAATAGTATCAATATATATATATATATATATATATTTATTTGTTTACTTTTTCTATCAGT
GTTTTCAATTTTTTATTAAACAATGTTTGATTTTTTAAATCGCAATTTAAT **ACCTAAATATAAAAAAT**GTTATTATATTGCA
AAAACCCATCAACCTTGAAAAAAGTAGAAACGTTTTTATTTAATTCTATCAATAC **ATCATAAAATACG**AACGATCCCCGT
CCAAGTTATGAGCTTAATCTTCCATAAAAAATATTTGAAAGCAATAGATCATGTACTAAACTAAAATCAGGGAAATTcccgg
gatatcc

Orc6-4A (S106A, S116A, S123A, T146A)

ATGCAACAAGTTCAACACTGTGTTGCTGAAGTTTTGAGATTGGACCCACAAGAAAAGCCAGACTGGTCTTCT
GGTTACTTGAAGAAGTTGACTAACGCTACTTCTATTTTGTACAACACTTCTTTGAACAAGGTTATGTTGAAGCAAGACG
AAGAAGTTGCTAGATGTCACATTTGTGCTTACATTGCTTCTCAAAAGATGAACGAAAAGCACATGCCAGACTTGTGTTA
CTACATTGACTCTATTCCATTGGAACCAAGAAGGCTAAGCACTTGATGAACCTGTTTCAGACAATCTTTGTCTAACTCT
GCTCCAATGAAGCAATTCGCTTGGACTCCAGCTCCAAAGAAGAACAAGAGAGCTCCAGTTAAGAACGGTGGTAGATT
CACTTCTTCTGACCCAAAGGAATTGAGAAACCAATTGTTTCGGTGCTCCAACCTAAGGTTAGAAAGTCTCAAAACAACGA
CTCTTTCGTTATTCCAGAATTGCCACCAATGCAAACTAACGAATCTCCATCTATTACTAGAAGAAAGTTGGCTTTGAA
GAAGACGAAGACGAAGACGAAGAAGAACAGGTAACGACGGTTTGTCTTTGAAGTCTCACTCTAACAAAGTCTATTACT
GGTACTAGAAACGTTGACTCTGACGAATACGAAAACCACGAATCTGACCCAACTTCTGAAGAAGAACCATTGGGTGTT
CAAGAATCTAGATCTGGTAGAATAAGCAAAACAAGGCTGTTGGTAAGCCACAATCTGAATTGAAGACTGCTAAGGCT
TTGAGAAAGAGAGGTAGAATTCCAAACCTCTTTGTTGGTTAAGAAAGTACTGTAAGATGACTACTGAAGAAATTTATTAGAT
TGTGTAACGACTTCGAATTGCCAAGAGAAGTTGCTTACAAGATTGTTGACGAATACAACATTAACGCTTCTAGATTGGT
TTGTCCATGGCAATTGTTTGTGGTTTGGTTTGAAGTGTACTTTTCACTGTTTCAACGAAAAGAAGAAGAAAGGACCCA
AGAATTGACCACTTCATTGTTTCTAAGATGTGTTCTTTGATGTTGACTTCTAAGGTTGACGACGTTATTGAATGTGTTAA
GTTGGTTAAGGAATTGATTATTGGTGAAAAGTGGTTCAGAGACTTGCAAATTAGATACGACGACTTCGACGGTATTAG
ATACGACGAAATTTATTTTCAGAAAGTTGGGTTCTATGTTGCAAACCTACTAACATTTTGGTTACTGACGACCAATACAAC
ATTTGGAAGAAGAGAATTGAAATGGACTTGGCTTTGACTGAACCATTGTAA

Orc2-N-terminal truncation (delete a.a. 2-235)

ATGGCTTCTTCTTTCTTGGACACTTTCGAAGGTTACTTCGACCAAAGAAAAGATTGTTAGAACTAACGCTAAGT
CTAGACACACTATGTCTATGGCTCCAGACGTTACTAGAGAAGAATTCTCTTTGGTTTCTAAGCTTCTTCAACGAAAACCTT
CCAAAAGAGACCAAGACAAAAGTTGTTTCGAAATTCAAAAGAAGATGTTCCACAATACTGGTTCGAATTGACTCAAGG
TTTCTCTTTGTTGTTCTACGGTGTGTTGTTCTAAGAGAACTTCTTGAAGAATTGCTATTGACTACTTGTCTCCAAAGA
TTGCTTACTCTCAATTGGCTTACGAAAACGAATTGCAACAAAACAAGCCAGTTAACTCTATTCCATGTTTGATTTTGAAC
GGTTACAACCCATCTTGTAACTACAGAGACGTTTTCAAGGAAATTAAGTACTGACTTGTGTTCCAGCTGAATTGACTAGAT
CTGAAACTAAGTACTGGGGTAACCAAGTTATTTTGCAAATTCAAAAGATGATTGACTTCTACAAGAACCAACCATTGGA
CATTAAGTTGATTTTGGTTGTTTCACAACTTGGACGGTCCATCTATTAGAAAGAACAATTTCCAAACTATGTTGTCTTTCT
TGTCTGTTATTAGACAAATTGCTATTGTTGCTTCTACTGACCACATTTACGCTCCATTGTTGTGGGACAACATGAAGGC
TCAAACTACAACCTTCGTTTTCCACGACATTTCTAACTTCAACCATCTACTGTTGAATCTACTTTCAAGACGTTATGA
AGATGGGTAAGTCTGACACTTCTTCTGGTGCTGAAGGTGCTAAGTACGTTTTGCAATCTTTGACTGTTAACTCTAAGAA
GATGTACAAGTTGTTGATTGAAACTCAAATGCAAAACATGGGTAACCTGTCTGCTAACACTGGTCCAAAGAGAGGTAC
TCAAGAAGTGGTGTGATTGAAGTTGTTCAACCACTTGTGTGCTGCTGACTTCATTGCTTCTAACGAAATTGCTTTG
AGATCTATGTTGAGAGAATTCATTGAACACAAGATGGCTAACATTACTAAGAACAACCTCTGGTATGGAAATTTATTTGGG
TTCCATACACTTACGCTGAATTGGAAGGTTGTTGAAGACTGTTTTGAACACTTTGTAA

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

Supplementary Table 1: Primers

Primers	Sequence	Note	Source
GC227	Biotin-gctgcaatgataccgagagaccac	Amplify 3 kb template	Coster and Diffley, 2017
GC228	Biotin-gggccagatggaagccctc	Amplify 3 kb template	Coster and Diffley, 2017
EACS.B1 F	attcCTCGATTTTTTATGTTTAGTTTCGCGACGACGGTTTCGAGGTGG	Competitor DNA	Coster and Diffley, 2017
EACS.B1 R	CCACCTCGAAACCGTCGTCGCCGAACTAAACATAAAAAATCGAGgaat	Competitor DNA	Coster and Diffley, 2017
NCP F	Desthiobiotin – CGAtagaaCTCGGgcccctggagaatcGcgggtccg	Amplify nucleosome template	Miller et al., 2020
NCP R	CctgcACCCAGGGACTTGAAGTAATAAGGAC	Amplify nucleosome template	Miller et al., 2020
CTL38	ATCTTTTGAAATTTGC AAAATAACGTGGTTACCCAGTACTTAG	Insertion for ORC2ΔN	This study
CTL39	GCAAAATCAAAGATGATTGACTTCTACAAGAAC	Vector for ORC2ΔN	This study
CTL40	TATAGTTTTTCTCCTTGACGTAAAGTATAGAGG	Vector for ORC2ΔN	This study
CTL61	ATCAGTGTTCATTTTATTTTATTAACAATGTTTGATTTTTTAAATCGCAATTTAATAC CTggtacccggccagcg	Replace endogenous ARS317::kanMX6	This study
CTL62	ATGATCTATTGCTTTCAAATATTTTATGGAAGATTAGCTCATAACTTGACGGGA TCGTTTAAACTGGATGGCGGCGTTAG	Replace endogenous ARS317::kanMX6	This study
CTL98	TGCTACTCATCTAGTCTGTGCTGCCAAGCTATTTAATATCATGCACGAAAAGCA AACGgcatggaggccag	Replace URA3::hphNT1	This study
CTL99	GGCCGCATCTTCTCAAATATGCTTCCAGCCTGCTTTCTGTAACTTACCCTCTA CTTcttgagcggtcccaaac	Replace URA3::hphNT1	This study
CTL86	GCCTTTTATATACCGATTAAATATTTATTTCTGACTTTTCTTACCACCAACa ggcclgttlaaacGCTTTTCAATTCATCTTTTTTTTTTTG	Insertion of URA3 in ARS419	This study
CTL87	TTCTATTGTAAGAAGCTGTTTATAGAATACTCTTTTTTTATGTTTGGCCTACACC TtctagaagctcatatgTTAGTTTTGCTGGCCGCATC	Insertion of URA3 in ARS419	This study
CTL125	GGAGAAAAAATACACCGGTGTATAAAACAATGGCTTCTTCTTCTGGACACTT TCGAAGGTTACTTC	Insertion for ORC2ΔN	This study
CTL141	ggtgtgaataccgcacagatg	qPCR detection site	This study
CTL142	gctggcgtaatagogaagag	qPCR detection site	This study
CTL143	GTGCATGGCCTCTCTTTCTAATTG	Internal control for qPCR	This study
CTL144	GCAGGATTAGGTAGCGTTTAATTTATCC	Internal control for qPCR	This study
CTL147	TATTTCTTAGCATTTTTGACGAAATTacggccagtgaattcCTCG	Insert for pCTL67	This study
CTL148	AATTTCTGCAAAAATGCTAAGAAATAGCGGCCAGGTaTTTTATG	Insert for pCTL67	This study
CTL149	ATTTTCTGCAAAATCTGCGCCACCGTCGGC	Vector for pCTL67	This study
CTL150	AAATGCTAAGAAATAacctggcggtacccaac	Vector for pCTL67	This study
CTL156	TATTTCTTAGCATTTTTGACGAAATTacggccagtgaattcACTAGTAC	Insert for pCTL68 and pCTL69	This study
CTL157	AATTTCTGCAAAAATGCTAAGAAATActtagaggatccccggg	Insert for pCTL68 and pCTL69	This study
CTL158	AAATGCTAAGAAATAgttaccaactaatgccttgc	Vector for pCTL68 and pCTL69	This study
CTL159	ATTTTCTGCAAAATgtaatcatggtcatagctgttctcg	Vector for pCTL68 and pCTL69	This study
CTL160	GCAGCTGGTGTTCATCGAACATGATGAGCTTTAAACAAGGAGGAGTATAAACTTT TAacatagcggtgtgaaataccgc	Amplify sequence for ARS419 insertion	This study
CTL161	ATAACCGATGACTTGGCAAGAAAAGGATTTTAAGAGTTGTTCTATTCTGTAAGAACTG TTTcaggaaacagctatgaccatgattac	Amplify sequence for ARS419 insertion	This study
CTL172	TAGGTGTGAGAGCTCCGATCTCGGCGAACAC	Amplify sequence for LacO plasmid	This study
CTL173	CGAACTTATGGTACCGCCGAGCGCGACG	Amplify sequence for LacO plasmid	This study
CTL174	GAGCTCTCACACTACAAGGGATGTACATCAATTGTGAGCGGATAACAATTGTTAG GGAGGAATTGTGAGCGGATAACAATTTGGAGTTGATAATTGTGAGCGGATAACAATT GGCTTCAACGTAATTGTGAGCGGATAACAATTTCCGTACGAATGTGCCGAACCTATG GTACC	Duplex DNA of LacO	This study
CTL208	TCAAAAGGCCTGCAGGCAAGTGCACAAACAATCTTAATAAATACTACTCAGTAAT AACCTATTTCTTAGCATTTTTGACGAAATT	oligo for ARS1_B2	This study

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

CTL209	AATTTGTCAGGCTGCTAAGAAATAGGTTATTACTGAGTAGTATTATTTAAGTATT GTTTGTGCACTTGCCTGCAGGCTTTTGA	oligo for ARS1_B2	This study
CTL214	CTTGCTGCAGGCTTTTGAATCGGCCATGCGC	Amplify sequence for 90bp-B2	This study
CTL215	TTAGCATTTTGTACGAAATTCGTCGGCGTCTCG	Amplify sequence for 90bp-B2	This study
50bp up 209	TTAAGTTGGGTAACGCCAGGG	Amplify sequence for EMSA	This study
50bp down 209	CCAGCTGAGGTACCGGATGCT	Amplify sequence for EMSA	This study

Supplementary Table 2: Plasmids

Plasmid number	Plasmid	Vector	Insert source	Insert sites	Source
pGC199	No site	pUC19			Coster <i>et al.</i> , Science 2017.
pGC200	One site	pUC19			Coster <i>et al.</i> , Science 2017.
pGC201	25bp	pUC19			Coster <i>et al.</i> , Science 2017.
pGC202	50bp	pUC19			Coster <i>et al.</i> , Science 2017.
pGC211	90bp	pUC19			Coster <i>et al.</i> , Science 2017.
pCTL8	pRS304-Orc5/Orc6-4A	pJF18	synthetic GeneArt strings	BspI/EcoRI	pJF18 from Frigola <i>et al.</i> , Nature 2013.
pCTL9	12bp	pUC19	synthetic GeneArt strings	BspI/EcoRI	This study
pCTL40	10bp	pUC19	synthetic GeneArt strings	BspI/EcoRI	This study
pCTL64	pRS306-Orc1/Orc2ΔN	PCR from pJF19 with CTL39 and CTL40	PCR from pJF19 with CTL125 and CTL38	In-fusion	pJF19 from Frigola <i>et al.</i> , Nature 2013.
pCTL67	pUC19-90bp-Abf1	PCR from pGC211 with CTL149 and CTL150	PCR from pGC211 with CTL147 and CTL148	In-fusion	pGC211 from Coster <i>et al.</i> , Science 2017.
pCTL68	pUC19-ARS1-Abf1	PCR from pGC403 with CTL158 and CTL159	PCR from pGC403 with CTL156 and CTL157	In-fusion	pGC403 from Coster <i>et al.</i> , Science 2017.
pCTL69	pUC19-ARS317-Abf1	PCR from pCTL45 with CTL158 and CTL159	PCR from pCTL45 with CTL156 and CTL157	In-fusion	This study
pCTL86	pUC19-ARS317	pUC19	synthetic GeneArt strings 495bp	EcoRI/BspI	This study
pCTL91	pUC19-90bp-B2	pGC211	PCR pGC211 with primer CTL214 and CTL215; ARS1_B2 sequence generated from IDT as oligo CTL208 and CTL209	In-fusion	pGC211 from Coster <i>et al.</i> , Science 2017.
pCTL93	pUC19-300bp+LacO	pGC205	PCR pGC200 with primer CTL172 and CTL173; LacO sequence generated from IDT as duplex DNA CTL174	In-fusion	pGC 205 from Coster <i>et al.</i> , Science 2017.
pCTL94	pUC-One site+LacO	pGC200	PCR pGC200 with primer CTL172 and CTL173; LacO sequence generated from IDT as duplex DNA CTL174	In-fusion	pGC200 from Coster <i>et al.</i> , Science 2017.

Cell Cycle Regulation has Shaped Budding Yeast Replication Origin Structure and Function

Supplementary Table 3: Yeast strains

Yeast strain	Background	Genotype	Source	
ySDORC	W303-1a	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> <i>bar1::Hyg</i> <i>pep4::KanMx</i> <i>his3::pRS303-Gal1,10-ORC3/4 (HIS3)</i> <i>ura3::pRS306-Gal1,10-ORC1/2 CBP-Orc1 (URA3)</i> <i>trp1::pRS304-Gal1,10-ORC5/6 (TRP1)</i> <i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> <i>bar1::Hyg</i> <i>pep4::KanMx</i>	Frigola <i>et al.</i> , Nature 2013.	
YGC229	W303-1a	<i>his3::pRS303-Gal1,10-Cdt1/Gal4 (HIS3)</i> <i>ura3::pRS306-Gal1,10-Mcm2/3 (URA3) CBP-TEV-NanoLuc-Mcm3</i> <i>trp1::pRS304-Gal1,10-Mcm4/5 (TRP1)</i> <i>leu2::pRS305-Gal1,10-Mcm6/7 (LEU2)</i> <i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> <i>bar1::Hyg</i> <i>pep4::KanMx</i>	Coster <i>et al.</i> , Science 2017.	
yCTL6 (Orc6-4A)	W303-1a	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> <i>bar1::Hyg</i> <i>pep4::KanMX</i> Flag at WT ORC6 (Leu) c-terminal <i>ura::URApRS306/SDORC1,2</i> <i>his::HISpRS303/SDORC3,4</i> <i>trp::TRPpRS304/SDORC5,ORC6-4S/T-A</i>	This study	
yCTL17 (Orc2_deltaN)	W303-1a	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> <i>bar1::Hyg</i> <i>pep4::KanMX</i> Flag at WT ORC2 (Nat) C-terminal <i>ura::URApRS306/SDORC1,ORC2_del2-235</i> <i>his::HISpRS303/SDORC3,4</i> <i>trp::TRPpRS304/SDORC5,ORC6</i>	This study	RE digestion with NcoI on pCTL64
yCTL28 (ARS1Δ)	A364a	<i>MATa ade2 ade3 ura3-52 trp1-289 leu2-3,112</i> <i>bar1::LEU2</i> <i>ORC2</i> <i>ORC6</i> <i>ura3-52::GAL1p-Δntcdc6, URA3 ΔURA3::hph</i> <i>cdc20::MET3p-HA3-CDC20, TRP1</i> <i>MCM7-2NLS</i> <i>ARS317::KanMX6 ΔARS419::ARS1Δ</i>	This study	PCR on pGC404 with CTL160 and CTL161. pGC404 from Coster <i>et al.</i> , Science 2017.
yCTL29 (ARS1)	A364a	<i>MATa ade2 ade3 ura3-52 trp1-289 leu2-3,112</i> <i>bar1::LEU2</i> <i>ORC2</i> <i>ORC6</i> <i>ura3-52::GAL1p-Δntcdc6, URA3 ΔURA3::hph</i> <i>cdc20::MET3p-HA3-CDC20, TRP1</i> <i>MCM7-2NLS</i> <i>ARS317::KanMX6 ΔARS419::ARS1</i>	This study	PCR on pCTL68 with CTL160 and CTL161
yCTL30 (ARS317)	A364a	<i>MATa ade2 ade3 ura3-52 trp1-289 leu2-3,112</i> <i>bar1::LEU2</i> <i>ORC2</i> <i>ORC6</i> <i>ura3-52::GAL1p-Δntcdc6, URA3 ΔURA3::hph</i> <i>cdc20::MET3p-HA3-CDC20, TRP1</i> <i>MCM7-2NLS</i> <i>ARS317::KanMX6 ΔARS419::ARS317</i>	This study	PCR on pCTL69 with CTL160 and CTL161
yCTL33 (90bp)	A364a	<i>MATa ade2 ade3 ura3-52 trp1-289 leu2-3,112</i> <i>bar1::LEU2</i> <i>ORC2</i> <i>ORC6</i> <i>ura3-52::GAL1p-Δntcdc6, URA3 ΔURA3::hph</i> <i>cdc20::MET3p-HA3-CDC20, TRP1</i> <i>MCM7-2NLS</i> <i>ARS317::KanMX6 ΔARS419::90bp</i>	This study	PCR on pCTL67 with CTL160 and CTL161