Published in final edited form as:
Science. 2017 July 21; 357(6348): 314–318. doi:10.1126/science.aan0063.

Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading

Gideon Coster and **John F.X Diffley***
The Francis Crick Institute, 1 Midland Road, London NW1 1AT, U.K.

Abstract

Bidirectional replication from eukaryotic DNA replication origins requires the loading of two ring-shaped Minichromosome Maintenance (MCM) helicases around DNA in opposite orientations. MCM loading is orchestrated by binding of the Origin Recognition Complex (ORC) to DNA, but how ORC coordinates symmetrical MCM loading is unclear. We used natural budding yeast origins and synthetic sequences to show that efficient MCM loading requires binding of two ORC molecules to two ORC binding sites. The relative orientation of these sites, but not the distance between them, was critical for MCM loading in vitro and origin function in vivo. We propose that quasi-symmetrical loading of individual MCM hexamers by ORC and directed MCM translocation into double hexamers acts as a unifying mechanism for the establishment of bidirectional replication in Archaea and Eukarya.

The motor of the eukaryotic replicative helicase – the heterohexameric MCM complex – is loaded onto replication origins as an inactive, head-to-head double hexamer (DH) during the G1 phase of the cell cycle (1–3). In S phase, the DH is converted into two active CMG (Cdc45-MCM-GINS) helicases (4–6), that nucleate assembly of the two bidirectional replisomes. MCM loading begins with the binding of ORC to DNA. ORC, together with Cdc6 and Cdt1, then recruits the first MCM hexamer (7, 8). ATP binding by ORC, Cdc6 and MCM is required for this recruitment, and ATP hydrolysis by MCM then drives DH assembly (9, 10). But how the second hexamer is recruited and loaded in the correct position and orientation is unclear (Fig. S1A).

The recruitment of the first MCM hexamer to ORC/Cdc6 is mediated by the C-terminus of Mcm3 (11). Based on single molecule approaches, it has been proposed that the second MCM hexamer is directly recruited by interaction with the first hexamer, aided by a second Cdc6 but not a second ORC molecule (8, 12). If this were the case, point mutants in the C-terminus of Mcm3 (e.g. Mcm3-13) that cannot be recruited as the first hexamer, should still be able to load as a second hexamer in the presence of wild type MCM. However, such mutants fail to load even when mixed with wild type MCM, suggesting that both hexamers must be able to bind ORC/Cdc6 for loading (11). It remains possible that Mcm3-13 loading still occurs, but at a level below detection limits. We therefore developed a more sensitive and quantitative assay for MCM loading based on a fusion of Mcm3 to luciferase (Fig. 1A and Fig. S2). The Mcm3-13 mutation had no effect on MCM complex stability (Fig. S1B) or

*Correspondence: John.Diffley@crick.ac.uk.

Europe PMC Funders Author Manuscripts

subunit stoichiometry (Fig. S1C). However, despite the improved sensitivity, no loading of Mcm3-13 was observed in the absence or presence of wild type MCM over a range of ratios (Fig. 1B and Fig. S1D, E).

Although this result suggested that both MCM hexamers must be able to bind ORC/Cdc6, budding yeast origins generally contain a single high affinity ORC binding site located at one end of a Nucleosome Free Region (NFR) (13). To test whether these features are sufficient for origin activity in vivo, we constructed synthetic sequences bearing a single T-rich ORC binding site in a GC-rich background with or without nucleosome excluding sequences associated with origins – either binding sites for the transcription factor Abf1 or different lengths of poly(dA) sequences (13, 14). Yeast cells transformed with these constructs failed to sustain growth on selective media, indicating that they are not functional origins (Fig. 1C-E). The synthetic ORC binding site supported ATP-dependent (Fig. 1F) and salt-stable (Fig. S1F) ORC binding in vitro comparable to or better than several natural origins (Fig. 1F, G). However, MCM loading on this synthetic sequence was poor relative to natural origins, even those like ARS606 with weaker ORC binding (Fig 1H).

Budding yeast origins also often contain multiple, degenerate ORC binding sites in the opposite orientation to the high affinity site (15). We constructed additional templates that contained two ORC sites separated by 100bp in all possible orientations (Fig. S3A). Whilst ORC binding and MCM recruitment in ATPγS were unaffected by the relative orientation of the sites (Fig. 2A, B), MCM loading was greatly stimulated by two sites (> 7-fold) only when the two-sites were in a head-to-head orientation (Fig. 2C). This synergy required an intact second site (Fig. S3B) and resulted in loading efficiencies equivalent to natural origins (see below). Two-sites, only in the head-to-head orientation, also generated functional origins in vivo, with or without Abf1 sites (Fig. 2D, E). Furthermore, changing only three bases within the poly(dA) stretch created a close match to the ORC binding sequence that supported ORC binding in vitro (Fig. S3C) and converted all of the inactive poly(dA) templates (Fig. 1E) into active origins (Fig. 2F). Thus, two ORC binding sites display orientation-dependent synergy that is necessary and sufficient for maximal MCM loading in vitro and origin function in vivo.

Templates with single high affinity ORC binding sites supported low levels of MCM loading in vitro (Fig. 2C), indicating either that a single ORC binding event is sufficient for MCM loading, or that loading on single-site templates involves two ORC binding events: one at the high affinity site, and one non-specifically in adjacent sequences. Several lines of evidence support the latter possibility. First, ORC binding to a specific site (Fig. S1F) and MCM loading onto two-site templates (Fig. 3A) were resistant to salt concentrations of up to 100 mM NaCl whereas non-specific binding of ORC (Fig. S1F) and loading onto templates lacking ORC sites (Fig. 3A) were both sensitive to salt in this range. MCM loading on one-site templates exhibited intermediate salt sensitivity suggesting that both specific and non-specific ORC binding contribute to loading onto these templates. Consistent with this, MCM loading on the one-site template was almost completely eliminated by an oligonucleotide competitor containing an ORC binding site (Fig. 3B), whilst the two-site template was largely resistant to competition (~23 fold difference). Finally, MCM loading on the one-site template was also inhibited by excess template (Fig. S4), suggesting that excess high affinity

binding sites sequester ORC, thereby preventing non-specific binding needed for MCM loading.

If occupancy of two ORC sites in opposite orientations is required for loading, it is likely to be especially inefficient at low ORC concentrations. Indeed, whilst mathematical modelling predicts a simple linear relationship between ORC concentration and MCM loading in a one-ORC mechanism, a more complex, sigmoidal relationship is predicted if two ORC molecules are required for loading (Fig. S5). As expected, in the presence of ATPγS under non-specific (no NaCl) conditions, MCM recruitment increased linearly with increasing ORC concentration and fitted well to a one-ORC model (Fig. 3C and Fig. S5), consistent with the fact that there is a single MCM in the recruited complex. However, in the presence of ATP, MCM loading did not increase linearly with increasing ORC concentration but instead generated a sigmoidal curve that fit well to the two-ORC model, but not the one-ORC model (Fig. 3C and Fig. S5). Thus, efficient MCM loading requires two ORC molecules.

We next examined the role of secondary ORC binding sites in natural origins. We selected several well-studied origins, all of which contained multiple near matches to the ORC binding site 3’ and on the opposite strand relative to the essential site (Fig. S6). In each case, the essential site was the closest match to the consensus sequence (Fig. S6) and was the site with the highest affinity, since robust ORC binding was lost upon deletion (Fig. S7A). However, under milder wash conditions, ORC bound above background levels to the remaining sequences (Fig. 3D). This binding was stabilized by ATP (Fig. S7B), a hallmark of sequence-specific ORC binding (16). Moreover, MCM was loaded on these sequences at levels greater than the no site control and comparable to the synthetic one-site construct (Fig. 3E). ARS606, which had the weakest high affinity site (Fig. 1F,G), had the strongest secondary sites (Fig. S7B) and loaded MCM better than other origins with stronger high affinity sites (Fig. 3E). To test the importance of secondary ORC sites, we identified two origins, ARS600.1 and ARS1216, with small numbers of secondary ORC binding sites (Fig. S7C). Mutation of these secondary sites greatly reduced MCM loading in vitro and impaired origin activity in vivo (Fig. 3F and Fig. S7D, E). Thus, natural origins use one high affinity ORC binding site with one or more lower affinity secondary sites in the opposite orientation.

The fact that two ORC binding events promote MCM loading, but the distance between high affinity and secondary sites within natural origins varies, prompted us to test the effects of distance between two specific sites. Neither ORC binding nor MCM recruitment were affected by altering the spacing between sites (Fig. S8). Synergistic MCM loading occurred over a wide range of distances from 25 to 400bp, with a peak at 70bp (Fig. 4A). Moreover, origin activity in vivo was seen with all two-ORC site templates at distances up to 400bp apart (Fig. 4B). This suggested that MCM hexamers may translocate along DNA before forming a stable DH. Such long-range directed translocation is used by type III restriction enzymes and mismatch repair proteins (17, 18), and can be blocked by an intervening obstacle. A covalent protein roadblock (19) impaired MCM loading when positioned between two ORC sites, but not when positioned 3’ of either one or two-sites (Fig. 4C, D and Fig. S9).

We propose (Fig. 4E) that each MCM hexamer in the DH is loaded by the same mechanism in a reaction requiring ORC, Cdc6 and Cdt1. Our results do not address whether these are independent or concerted events. Budding yeast origins comprise a single high affinity ORC binding site and multiple lower affinity sites in the opposite orientation; we describe this arrangement as “quasi-symmetrical”. This role for multiple ORC binding sites is supported by the bimodal distribution of ORC bound at origins in vivo (20). It helps explain why there is a switch in sequence bias from T-rich to A-rich between the high affinity ORC site and flanking sequence and how the positioning of nucleosomes at both ends of the NFR at yeast origins can be affected by ORC (13). We note that Archaeal origins have a related architecture with head-to-head high affinity ORC binding sites, separated by 65-70bp of AT rich DNA (21) – close to the distance that yielded optimal synergistic loading in budding yeast and the length of DNA covered by the DH (~68 bp) (1). Deletion analysis of a fission yeast replication origin in vivo also revealed a requirement for two ORC binding sites in MCM loading (22). Origins in higher eukaryotes show little sequence specificity (23), but we suggest act in a similar two ORC mechanism as the yeast system under non-specific conditions.

Why the second ORC was not detected in single molecule experiments (8) is unclear; perhaps transient ORC binding at weak sites was missed for technical reasons or maybe unlabeled ORC was responsible for loading on the secondary site. Alternatively, there may be a second, inefficient, pathway in which a single ORC molecule hops between two sites to load each hexamer, which might be more prevalent at the low ORC concentrations used in these experiments and with the specific natural origin employed. Such a “hopping” model still preserves the inherent symmetry of our proposed model, but based on our results cannot be a major pathway.

Stable double hexamer formation from imprecisely spaced ORC binding sites requires translocation of single MCM hexamers. Although this may occur by passive sliding, we note that MCM loading requires ATP hydrolysis by both MCM hexamers (9, 10) and the predicted direction of movement of MCM loaded at distant ORC sites is the same as CMG during replication (24), leading us to suggest translocation is an active process. Consistent with this, the eukaryotic MCM complex has DNA translocase activity without Cdc45 or GINS (25) and archaeal MCM can translocate over duplex DNA (26).

Supplementary Materials

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank H. Yardimci (Francis Crick Institute, UK) for advice and reagents, H. Patel and G. Kelly (Francis Crick Institute, UK) for bioinformatic analyses and A. McClure for critical reading of the manuscript. This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001066), the UK Medical Research Council (FC001066), and the Wellcome Trust (FC001066). This work was also funded by a Wellcome Trust Senior Investigator Award (106252/Z/14/Z) and a European Research Council Advanced Grant (669424-CHROMOREP) to J.F.X.D. The authors declare no conflict of interest. All data is contained within the paper and the accompanying supplementary materials. J.F.X.D. and G.C. designed the experiments, analysed data and wrote the manuscript. G.C. performed all experiments.

References and Notes

1. Remus D, et al. Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. Cell. 2009; 139:719–730. [PubMed: 19896182]
2. Evrin C, et al. A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:20240–20245. [PubMed: 19910535]
3. Bell SP, Labib K. Chromosome Duplication in Saccharomyces cerevisiae. Genetics. 2016; 203:1027–1067. [PubMed: 27384026]
4. Costa A, et al. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nature Structural & Molecular Biology. 2011; 18:471–7.
5. Moyer SE, Lewis PW, Botchan MR. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proceedings of the National Academy of Sciences. 2006; 103:10236–10241.
6. Ilves I, Petojevic T, Pesavento JJ, Botchan MR. Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins. Molecular Cell. 2010; 37:247–258. [PubMed: 20122406]
7. Sun J, et al. Cryo-EM structure of a helicase loading intermediate containing ORC-Cdc6-Cdt1-MCM2-7 bound to DNA. Nature Structural & Molecular Biology. 2013; 20:944–951.
8. Ticau S, Friedman LJ, Ivica NA, Gelles J, Bell SP. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. Cell. 2015; 161:513–525. [PubMed: 25892223]
9. Coster G, Frigola J, Beuron F, Morris EP, Diffley JFX. Origin Licensing Requires ATP Binding and Hydrolysis by the MCM Replicative Helicase. Molecular Cell. 2014; 55:666–677. [PubMed: 25087873]
10. Kang S, Warner MD, Bell SP. Multiple functions for Mcm2-7 ATPase motifs during replication initiation. Molecular Cell. 2014; 55:655–665. [PubMed: 25087876]
11. Frigola J, Remus D, Mehanna A, Diffley JFX. ATPase-dependent quality control of DNA replication origin licensing. Nature. 2013; 495:339–343. [PubMed: 23474987]
12. Sun J, et al. Structural and mechanistic insights into Mcm2-7 double-hexamer assembly and function. Genes & Development. 2014; 28:2291–2303. [PubMed: 25319829]
13. Eaton ML, Galani K, Kang S, Bell SP, MacAlpine DM. Conserved nucleosome positioning defines replication origins. Genes & Development. 2010; 4:748–753.
14. Breier AM, Chatterji S, Cozzarelli NR. Prediction of Saccharomyces cerevisiae replication origins. Genome Biol. 2004; 5:R22. [PubMed: 15059255]
15. Palzkill TG, Newlon CS. A yeast replication origin consists of multiple copies of a small conserved sequence. Cell. 1988; 53:441–450. [PubMed: 3284655]
16. Bell SP, Stillman B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature. 1992; 357:128–134. [PubMed: 1579162]
17. Kolodner RD, Mendillo ML, Putnam CD. Coupling distant sites in DNA during DNA mismatch repair. Proceedings of the National Academy of Sciences. 2007; 104:12953–12954.
18. Szczelkun MD, Friedhoff P, Seidel R. Maintaining a sense of direction during long-range communication on DNA. Biochem Soc Trans. 2010; 38:404–409. [PubMed: 20298192]
19. Yardimci H, et al. Bypass of a protein barrier by a replicative DNA helicase. Nature. 2012; 492:205–209. [PubMed: 23201686]
20. Belsky JA, MacAlpine HK, Lubelsky Y, Hartemink AJ, MacAlpine DM. Genome-wide chromatin footprinting reveals changes in replication origin architecture induced by pre-RC assembly. Genes & Development. 2015; 29:212–224. [PubMed: 25593310]
21. Robinson NP, et al. Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus. Cell. 2004; 116:25–38. [PubMed: 14718164]
22. Takahashi T, Ohara E, Nishitani H, Masukata H. Multiple ORC-binding sites are required for efficient MCM loading and origin firing in fission yeast. EMBO J. 2003; 22:964–974. [PubMed: 12574132]
23. Cvetic C, Walter JC. Eukaryotic origins of DNA replication: could you please be more specific? Seminars in Cell & Developmental Biology. 2005; 16:343–353. [PubMed: 15840443]

24. Georgescu R, et al. Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. Proceedings of the National Academy of Sciences. 2017; 114:E697–E706.

25. Bochman ML, Schwacha A. The Mcm2-7 complex has in vitro helicase activity. Molecular Cell. 2008; 31:287–293. [PubMed: 18657510]

26. Shin J-H, Jiang Y, Grabowski B, Hurwitz J, Kelman Z. Substrate Requirements for Duplex DNA Translocation by the Eukaryal and Archaeal Minichromosome Maintenance Helicases. Journal of Biological Chemistry. 2003; 278:49053–49062. [PubMed: 12975364]

27. Dani GM, Zakian VA. Proceedings of the National Academy of Sciences. 1983; 80:3406–3410.

28. Liachko I, Youngblood RA, Keich U, Dunham MJ. High-resolution mapping, characterization, and optimization of autonomously replicating sequences in yeast. Genome Res. 2013; 23:698–704. [PubMed: 23241746]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

One Sentence Summary

The MCM double hexamer, precursor of the replicative DNA helicase, results from a symmetrical loading reaction.

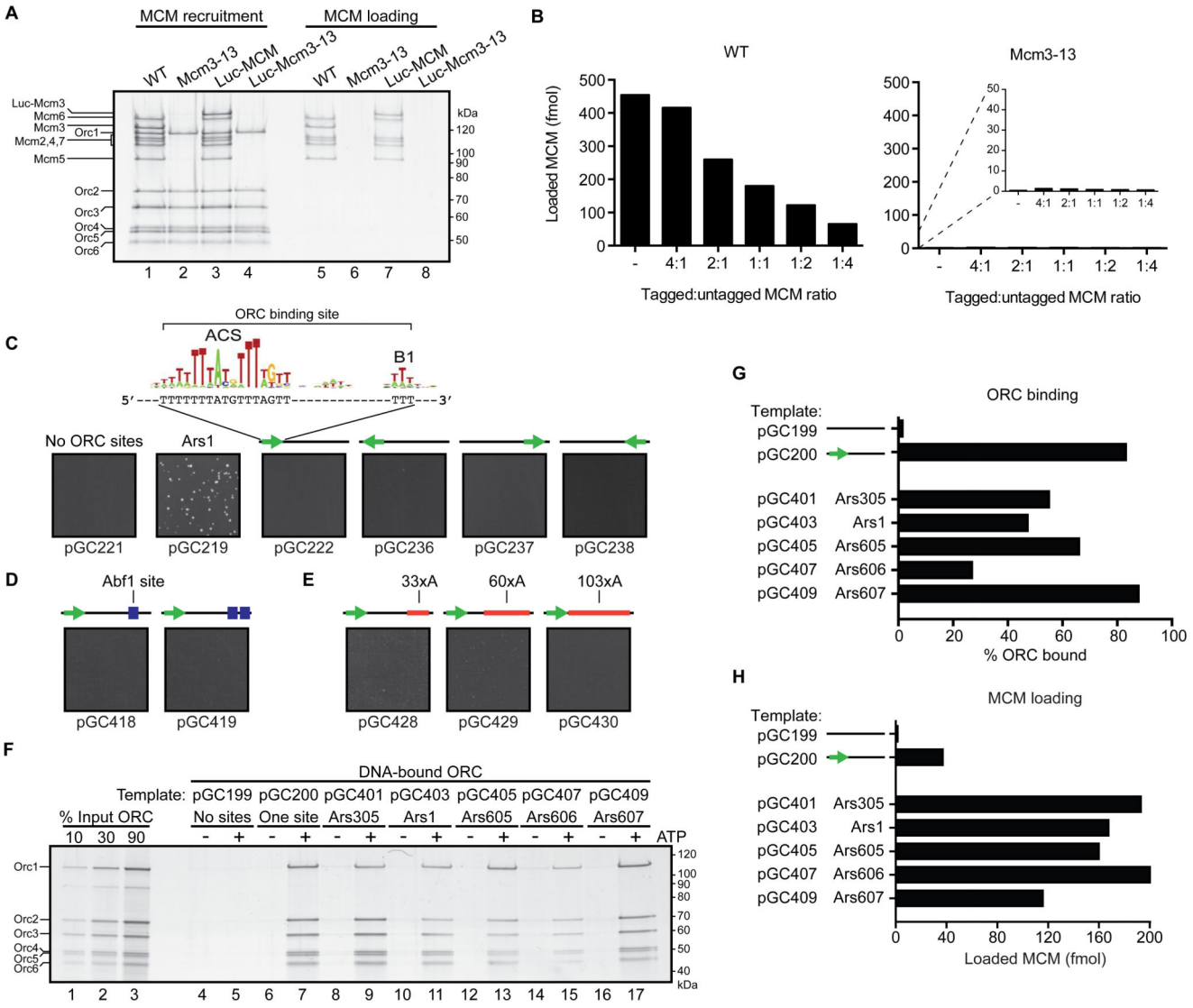

Figure 1. A single ORC binding site is insufficient for optimal MCM loading and origin function

(A) MCM recruitment and loading are unaffected by a luciferase tag and are defective in Mcm3-13. **(B)** Untagged WT MCM does not rescue the loading defect of Mcm3-13. **(C)** A synthetic ORC binding site does not support plasmid replication in vivo, even in the presence of Abf1 binding sites **(D)** or poly(dA) stretches of increasing length **(E)**, assayed by the formation of yeast colonies after transformation. **(F,G)** ORC binding to the synthetic site is ATP dependent and comparable to natural origins. **(H)** The presence of chloride during loading (80 mM KCl) drives sequence specific MCM loading and reveals that natural origins are more efficient than a single synthetic site.

Europe PMC Funders Author Manuscripts

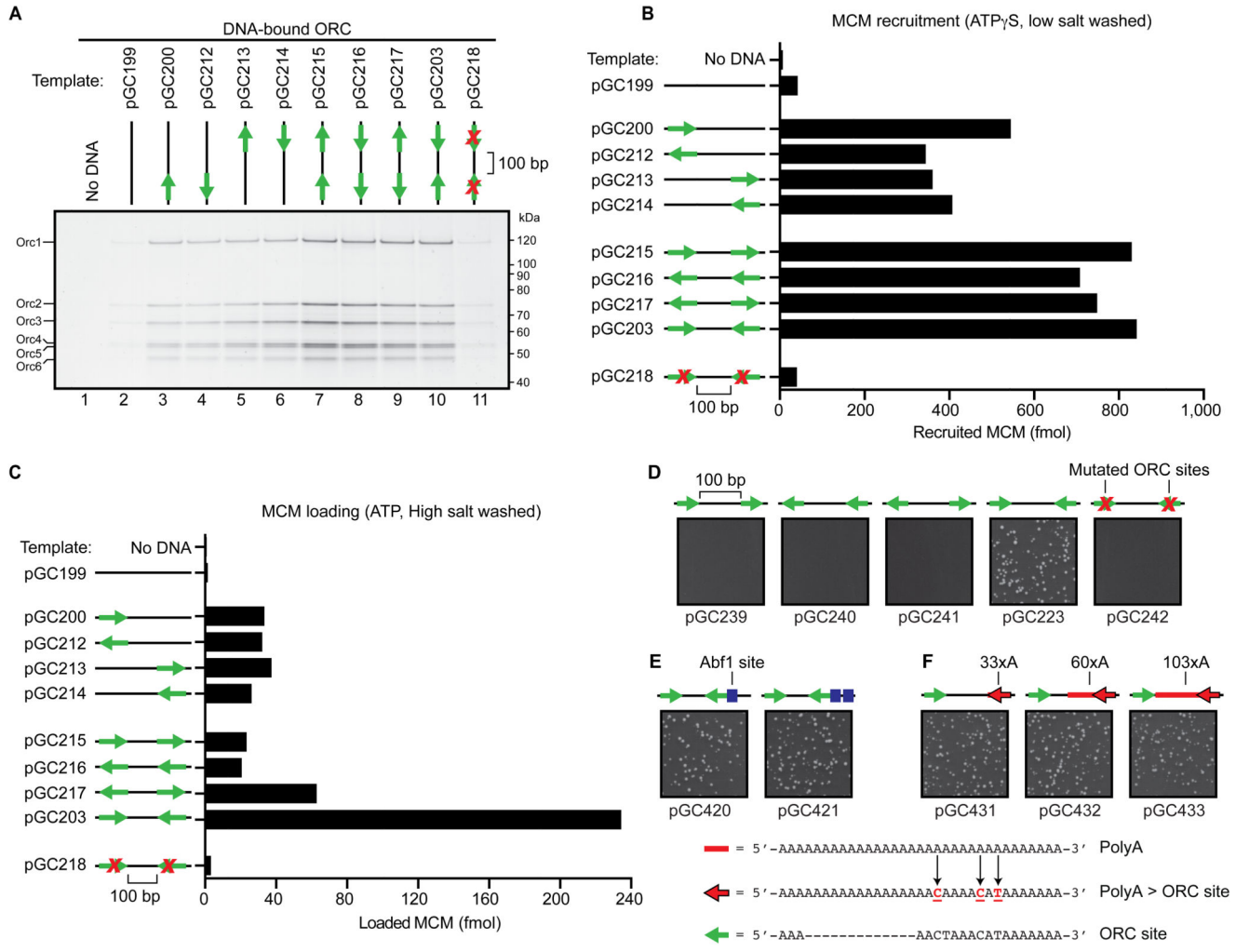

Figure 2. Two ORC sites are necessary and sufficient for maximal MCM loading and origin activity

(A) ORC binding, **(B)** MCM recruitment and **(C)** MCM loading onto synthetic substrates containing one or two ORC binding sites in all possible orientations. in vivo origin activity of synthetic two-site substrates in the absence **(D)** or presence **(E)** of Abf1 binding sites. **(F)** Three base pair changes that create a second ORC site are sufficient to convert the Poly(dA) substrates from Figure 1E into active origins.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

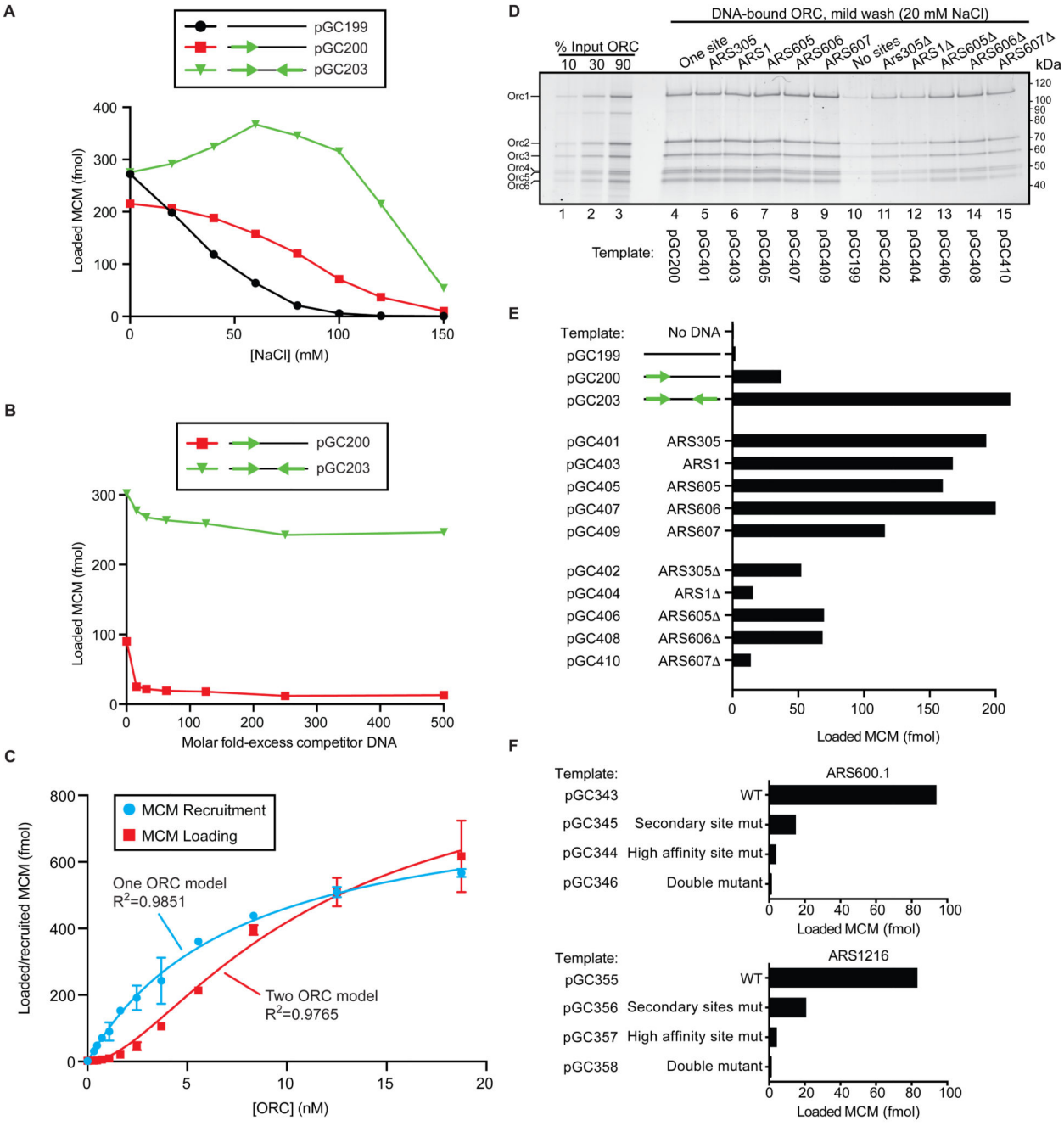

Figure 3. Natural origins employ one high affinity site and additional secondary sites

(A) Salt sensitivity of one-site loading suggests the usage of a second non-specific site. The indicated amount of salt was present during loading. All reactions were subsequently washed with high salt (1M NaCl). **(B)** One-site loading is more sensitive to competitor DNA than two-site loading. **(C)** MCM recruitment and loading as a function of ORC concentration. Data is plotted as mean ±SEM. **(D)** Natural origins harbor secondary ORC binding sites. **(E)** MCM loading with synthetic versus natural origins, as well as deletion

mutants (See Fig. S6). **(F)** Mutations in secondary sites impair MCM loading in ARS600.1 and ARS1216.

Europe PMC Funders Author Manuscripts

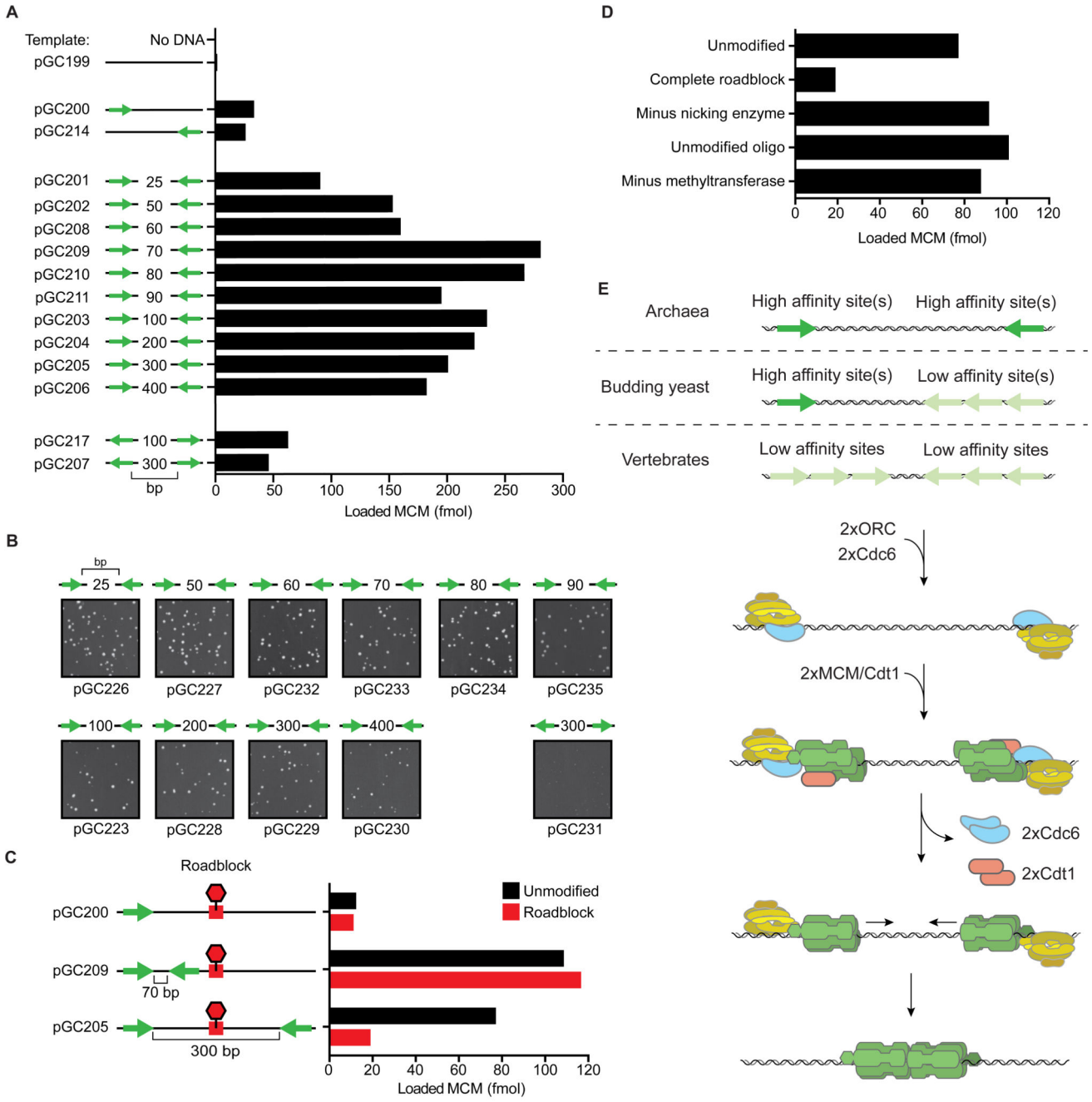

Figure 4. Two-site loading exhibits flexible spacing and is sensitive to an intervening roadblock

(A) Synergistic loading and **(B)** origin activity exhibit flexible inter-site spacing. **(C)** A covalent DNA-protein roadblock between two-sites inhibits synergistic loading. **(D)** Only the complete roadblock reaction leads to a block in loading. **(E)** Proposed model for MCM loading. See text for details.