Specificity and function of Archaeal DNA replication initiator proteins

Samson, Rachel Y.; Xu, Yanqun; Gadelha, Catarina; Stone, Todd A.; Faqiri, Jamal N.; Li, Dongfang; Qin, Nan; Pu, Fei; Liang, Yun Xiang; She, Qunxin; Bell, Stephen D.

Published in:
Cell Reports

DOI:
10.1016/j.celrep.2013.01.002

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):

Download date: 22. feb., 2019
Specificity and Function of Archaeal DNA Replication Initiator Proteins

Rachel Y. Samson,1,2,6 Yanqun Xu,3,4,6 Catarina Gadelha,1,7 Todd A. Stone,2 Jamal N. Faqiri,4 Dongfang Li,5 Nan Qin,5,8 Fei Pu,5 Yun Xiang Liang,3 Qunxin She,3,4,* and Stephen D. Bell1,2,*

1Sir William Dunn School of Pathology, Oxford University, South Parks Road, Oxford OX1 3RE, UK
2Molecular and Cellular Biochemistry Department, Department of Biology, Indiana University, Simon Hall, 212 South Hawthorne Drive, Bloomington, IN 47405, USA
3State Key Laboratory for Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
4Archeal Centre, Department of Biology, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen N, Denmark
5BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen 518083, China
6These authors contributed equally to this work
7Present address: School of Biology, University of Nottingham, Queen’s Medical Centre, Nottingham NG7 2UH, UK
8Present address: State Key Laboratory for Diagnosis and Treatment of Infectious Disease, The First Affiliated Hospital, Zhejiang University, Hangzhou 310003, PRC
*Correspondence: qunxin@bio.ku.dk (Q.S.), stedbell@indiana.edu (S.D.B.)
http://dx.doi.org/10.1016/j.celrep.2013.01.002

SUMMARY

Chromosomes with multiple DNA replication origins are a hallmark of Eukaryotes and some Archaea. All eukaryal nuclear replication origins are defined by the origin recognition complex (ORC) that recruits the repliative helicase MCM(2-7) via Cdc6 and Cdt1. We find that the three origins in the single chromosome of the archaeon *Sulfolobus islandicus* are specified by distinct initiation factors. While two origins are dependent on archaeal homologs of eukaryal Orc1 and Cdc6, the third origin is instead reliant on an archaeal Cdt1 homolog. We exploit the nonessential nature of the *orc1-1* gene to investigate the role of ATP binding and hydrolysis in initiator function in vivo and in vitro. We find that the ATP-bound form of Orc1-1 is proficient for replication and implicates hydrolysis of ATP in downregulation of origin activity. Finally, we reveal that ATP and DNA binding by Orc1-1 remodels the protein’s structure rather than that of the DNA template.

INTRODUCTION

Several archaeal species have multiple origins of replication per chromosome (Lundgren et al., 2004; Norais et al., 2007; Robinson and Bell, 2007; Robinson et al., 2004, 2007). For example, chromosomes of species in the genus *Sulfolobus* are replicated from three origins (Lundgren et al., 2004; Robinson et al., 2004, 2007). All three origins fire once per cell cycle in all cells (Duggin et al., 2008a). Two of the origins, *oriC1* and *oriC2*, lie adjacent to the genes for the candidate initiator proteins *Orc1*-1 and *Orc1*-3, respectively (She et al., 2001). *Orc1*-1 and *Orc1*-3 share 35% amino acid sequence identity and are related to both eukaryotic *Orc1* and Cdc6 (Bell and Dutta, 2002; Robinson and Bell, 2007). There is additionally a third *Orc1/Cdc6* protein in *Sulfolobus*, *Orc1*-2, but it is not encoded adjacent to an origin. The third origin, *oriC3*, lies beside the gene for a crenarchaeal-specific protein, WhiP. WhiP is predicted to have a tandem winged-helix structure reminiscent of bacterial plasmid Rep proteins and has sequence similarity to the eukaryal replication factor Cdt1 (Robinson and Bell, 2007). Biochemical DNA footprinting studies have revealed that complex interactions among the *Sulfolobus solfataricus* *Orc1/Cdc6* proteins, WhiP, and the replication origins are possible in vitro (Robinson and Bell, 2007; Robinson et al., 2004, 2007). More specifically, *Orc1*-1, *Orc1*-2, and WhiP bind to all three origins, and *Orc1*-3 binds to *oriC2* and *oriC3*. However, the precise requirements for origin function either in vivo or in vitro remain unknown. The *Orc1* proteins are AAA+ proteins that undergo a single-turnover hydrolysis of ATP to ADP, but the functional consequences of ATP binding and hydrolysis are not yet known (Singleton et al., 2004). In the current work, we address these fundamental issues. First, we determine the initiator protein requirements for origin function in vivo in the genetically tractable strain *Sulfolobus islandicus* (Sis) (Guo et al., 2011; Zhang et al., 2010). We find that none of the *Orc1/Cdc6* genes are essential for cell viability. Furthermore, we reveal that while *oriC1* requires the *oriC1*-1 gene product for function, *oriC2* is dependent on *Orc1*-3. Finally, we find that *oriC3* functions independently of all three *Orc1/Cdc6* proteins but requires the *Cdt1*-related protein WhiP.

Next, we exploit the nonessential nature of *Orc1*-1 to express mutant forms of the protein and assay their phenotype with respect to *oriC1* firing in vivo. Our data reveal that expression of a Walker B mutant form of the protein that can bind but not hydrolyze ATP leads to an overreplication phenotype, with elevated levels of initiation at *oriC1*. We further observe an enhanced ability of the Walker B mutant protein to recruit MCM to origin DNA in vitro. Finally, we demonstrate that the ATP-bound form of the protein, when bound to DNA, does not...
impinge on the geometry of the DNA but remodels its own conformation.

RESULTS

None of the orci/cdc6 Genes Are Essential for Viability

To address the roles of the individual Orc1/Cdc6 paralogs, we generated mutant S. islandicus cell lines in which the relevant genes were deleted (Figure 1A). The resultant strains were viable and showed growth and flow-cytometry profiles that were essentially indistinguishable from the wild-type (WT) strain (Figures 1B and C). We assayed origin activity using two-dimensional (2D) neutral-neutral agarose gel electrophoresis. As can be seen in Figure 1D, deletion of orci-1 abrogates firing of oriC1 but does not affect initiation at oriC2 or oriC3. Deletion of orci-2 does not affect firing at any of the three origins, but deletion of orci-3 abolishes oriC2 activity.

Next, we constructed double mutants lacking pairs of Orc1/Cdc6 proteins (Figure 2A). All three possible double mutants were generated, but attempts to generate a triple mutant were unsuccessful. The growth rates of all three double mutants were lower than WT or single mutants, and flow cytometry suggested an elevated S phase population of cells in the orci-3 mutant strain (Figure 2B). The growth rates of all three double mutants were generated, but attempts to generate a triple mutant were unsuccessful. The growth rates of all three double mutants were lower than WT or single mutants, and flow cytometry suggested an elevated S phase population of cells in the orci-3 mutant strain (Figure 2B).

Figure 1. Individual Orc1/Cdc6 Genes Are Nonessential

(A) Western analyses of Orc1/Cdc6 levels in wild-type (wt) and strains deleted for orci-1 (Δ1), orci-2 (Δ2), or orci-3 (Δ3). Antiserum against PCNA3 was used as a loading control.

(B) Growth curves of the indicated strains, following growth by absorbance at 600 nm.

(C) Flow-cytometry profiles of WT and mutant strains. Positions of single- and double-chromosome (1C and 2C) content cells are indicated.

(D) 2D gel analysis of origin activity in wild-type (wt) and the indicated mutant strains. The position of the bubble arc, indicative of replication initiation, is indicated by a red arrowhead.

See also Figures S5, S7, and Tables S1, S2, S3, and S4.

The whiP Gene Is Required for oriC3 Function

Interestingly, none of the individual or double Orc1/Cdc6 gene deletions impinged upon firing of oriC3. We therefore speculated that this origin could be functioning independently of Orc1/Cdc6 paralogs. Notably, the Cdt1-related protein WhiP is encoded by the whiP gene. To investigate this further, we generated a strain lacking WhiP by introducing tandem stop codons following the start codon of the open reading frame (ORF). The absence of WhiP in the resultant line was confirmed by western blotting (Figure 3A).

Cell growth and flow cytometry revealed that the mutant cells grew slower than either WT cells or any of the single orci deletion mutants and with a 2-fold increase in sub-1C signal in the flow cytometry (Figures 3B and 3C). Significantly, loss of the WhiP protein correlated with loss of firing at oriC3 but had no effect on initiation at oriC1 and oriC2, as adjudged both by 2D neutral-neutral agarose gel electrophoresis analysis and MFA (Figures 3D and 3E). Thus, while oriC3 is independent of...
individual and pairwise combinations of Orc1/Cdc6 proteins (Figures 1 and 2), WhiP is required for replication initiation at this origin (Figures 3D and 3E).

**Binding of Initiators to Origins In Vivo**

To address origin occupancy by the initiator proteins in vivo, we performed chromatin immunoprecipitation using antisera generated against Orc1-1, Orc1-2, and Orc1-3 and WhiP. In agreement with the lack of requirement for Orc1-2 for firing of any origin, we did not detect binding of Orc1-2 to any of the origins. Also in good agreement with the genetic analyses, we find that oriC1 is bound by Orc1-1, oriC2 shows strongest association with Orc1-3, and while no association of any Orc1/Cdc6 protein is observed at oriC3, the WhiP protein is specifically localized at this origin (Figures 4A–4D).

Archaea are believed to undergo coupled transcription and translation (French et al., 2007). We were, therefore, concerned that the ability to chromatin immunoprecipitate DNA corresponding to the origins located adjacent to the relevant initiator protein gene could reflect crosslinking of nascent poly-

In addition, a broad but shallow peak was observed over the previously described in vitro binding sites (Robinson and Bell, 2007). These lie at the boundaries of the downstream intergenic region, to which initiation at oriC3 has been mapped (Robinson et al., 2007). The whiP ORF contains conserved AT-rich sequence repeats (Figure S1) and footprinting studies reveal that the protein binds to these with nanomolar affinity. We reiterate that the whiP mutation that we employed to inactivate the gene for the genetic studies was by introduction of stop codons at the start of the ORF, rather than by deletion of the ORF (and thus its internal repeats).

In agreement with the 2D gel analyses, we detect Orc1-3 binding at oriC2 but do not observe significant binding of Orc1-1 to this locus. This was surprising to us because previous ChIP analyses and biochemical and structural studies of the closely related organism, Sulfolobus solfataricus (Sso), had demonstrated that both Orc1-1 and Orc1-3 bind to oriC2 in that species (Dueber et al., 2007, 2011; Robinson et al., 2004). However, examination of the sequence of *S. islandicus* oriC2 reveals that one of the characterized mini origin recognition
box (mORB) Orc1-1 binding sites has several base substitutions in conserved positions in the binding site when compared with S. solfataricus oriC2 (Figure S2). Additionally, the second mORB element has substitutions immediately adjacent to the core consensus motif. We therefore tested the S. islandicus Orc1-1 and Orc1-3 proteins’ DNA binding preferences in vitro. Remarkably, even between two closely related species within the Sulfolobus genus, we can observe variation within initiator protein specificity and function.

**Nucleotide Binding and Orc1-1 Function at oriC1 In Vivo**

We wished to exploit the nonessential nature of Orc1-1 to perform structure-function analyses in living cells. We therefore sought to complement the chromosomal Orc1-1 deletion with a plasmid-encoded copy. Initially, we expressed the gene from its own promoter on a low-copy-number plasmid. We could express the Orc1-1 protein from the arabinose promoter in either basal or induced conditions show an aberrant flow-cytometry profile with an elevated population of cells, with between 1C and 2C and greater than 2C genome contents (Figure 5C). Furthermore, fluorescence microscopy of these cells reveals altered cell morphologies, with 14% of cells extruding vesicle-like structures (Figure 5D).

Transcription of the orc1-1 gene in S. acidocaldarius and S. solfataricus is known to be regulated following perturbation of the cell cycle and after a number of cellular stresses (Fröls et al., 2007; Götz et al., 2007; Lundgren and Berman, 2007; Ortmann et al., 2008). We therefore wished to determine the profile of expression of this and the other initiator protein genes during a minimally perturbed cell cycle. We employed our previously described “Baby machine” method for synchronizing cells (Duggin et al., 2008a). For these studies, we utilized S. acidocaldarius cells because they have given the best synchrony in these experiments. We observe a striking cyclic pattern of transcript levels for orc1-1, orc1-3, and whip (Figure 5E), although we have previously demonstrated that Orc1/Cdc6 protein levels remain constant across the cell cycle (Duggin et al., 2008a). Nevertheless, all three transcripts show lowest levels in S phase and peak in or just before G1 phase. Interestingly, orc1-3 mRNA levels peak slightly later than orc1-1, correlating with the broader temporal window of oriC2 firing (Duggin et al., 2008a).

The Orc1/Cdc6 proteins possess a AAA+ domain and previous work has shown that this catalyzes a single turnover ATP hydrolysis event, resulting in stable binding of ADP (Singleton et al., 2004). Indeed, Orc1/Cdc6s from a range of archaeal species, when purified as recombinant proteins from bacteria, are isolated exclusively in the ADP-bound form. By analogy with bacterial DnaA, we assumed that the ATP-bound form of Orc1/Cdc6 proteins is the active form (see also below) (Kaguni, 2011; Mott and Berger, 2007). We therefore speculated that the cell-cycle-dependent wave of orc1-1 transcription
would result in transiently increased levels of Orc1-1-ATP and that this could create a permissive window for replication initiation. Expressing Orc1-1 constitutively during the cell cycle from the arabinose promoter would result in temporally ectopic Orc1-1-ATP and could thus account for the overreplication phenotype observed. To test this, we created mutant derivatives of Orc1-1 with alterations in the Walker A and Walker B motifs. The Walker A K69A is anticipated to be unable to bind ATP and the Walker B E147A mutant should bind but fail to hydrolyze ATP. These predicted properties were confirmed using purified

Figure 4. Origin Binding In Vivo by the Orc1/Cdc6 Proteins and WhiP
(A–D) ChIP analyses to assess the binding of the indicated initiator proteins to the three origin loci in vivo. The percent of input DNA recovered was calculated for each origin locus and a distal control locus (ctrl, Sire2328). Bars represent the SD calculated from the triplicate experiments.
(E–G) ChIP-seq results with Orc1-1, Orc1-3, and WhiP antisera, respectively. Genome coordinates are shown below each panel and the location of each origin indicated in bold. See also Figure S1. The y axis scale is the value obtained when normalizing ChIP read counts to input DNA. See also Figure S2, and Tables S1, S2, S3, and S5.
recombinant proteins (Figure S3). To exclude the possibility of compounding effects due to overexpression, the proteins were then expressed from a plasmid using the native $orc1-1$ promoter in the $Dorc1-1$ strain (Figure 6A). Despite being expressed and binding to $oriC1$ as detected by ChIP, the Walker A mutant failed to support detectable initiation at $oriC1$ (Figures 6A–6C). In contrast, the Walker B mutant shows a bubble arc of greater intensity than seen in WT cells (Figure 6C). Furthermore, this strain shows an aberrant flow cytometry profile, with 22% of the population with greater than 2C genome content (Figure 6D). Additionally, microscopy reveals the presence of large cells with elevated DNA content; 13% of Orc1-1 E147A-expressing cells are greater than 1.5 μm in diameter, compared with less than 1% of WT (Figures 6E and 6F). Interestingly, microscopy of the population expressing the Orc1-1 Walker A mutant also shows aberrant large cells (3.8%); however, in agreement with the flow-cytometry profiles, there is not a discernable population of cells with elevated DNA content (as adjudged by DAPI staining). Taken together, our data therefore indicate that the ATP-bound form of Orc1-1 is proficient for origin firing and, furthermore, implicates the hydrolysis of ATP in regulating origin activity.

**Biochemical Consequences of ATP Binding by Orc1-1**

To investigate the molecular basis of the distinct properties of the proteins, we first performed gel-filtration analyses. These revealed an increase in retention time of the E147A protein compared to WT, suggestive of a conformational alteration (Figure 7A). Next, we performed analytical ultracentrifugation (AUC) with the WT and E147A proteins (Figure 7B; Figure S3). The AUC revealed that the WT protein migrated in two populations, both
corresponding to monomer with distinct conformations. More specifically, 8.3% of the population had a sedimentation coefficient of 2.8 s and frictional ratio of 1.37, revealing a more extended configuration than the majority of the protein that resolved in a peak at 3.7 s with a frictional ratio of 1.07 and thus was more globular. Interestingly, very different behavior was observed with E147A. Although 65% of the E147A protein behaved as globular monomer (frictional ratio of 1.03), there was no peak corresponding to the extended monomer configuration. However, E147A generated a peak with a mass compatible with the formation of a dimer. Furthermore, the sedimentation coefficient and frictional ratio of the monomeric form of the E147A protein were lower than those of the majority WT species, indicating that it adopted a distinct conformation.

Structural studies have revealed that the ADP-bound forms of Orc1/Cdc6 proteins make bidentate interactions with DNA, binding via the C-terminal wH domain and the initiator-specific motif (ISM) in the AAA+ domain. To date, there is no structure available of the ATP-bound form of Orc1/Cdc6 on DNA. However, Wigley and colleagues were successful in denaturing Orc1-2 from *Aeropyrum pernix* (Ape), removing ADP and refolding in the presence of the nonhydrolyzable ATP analog, ADP-NP, prior to crystallization and structure determination (Singleton et al., 2004). The resultant structure revealed a significant alteration in the relative positioning of the wH and AAA+ domain compared to the various conformations adopted by the ADP-bound form (Singleton et al., 2004). Ape Orc1-2 has only 22% sequence identity to *S. islandicus* Orc1-1 and belongs to

Figure 6. Expression of a Walker B Mutant Orc1-1 Causes an Overreplication Phenotype

(A) Western analysis of Orc1-1 levels in WT (lane 1) and Δorc1-1 (lane 2) strains and the latter strain complemented with a plasmid encoding the gene for WT Orc1-1 (lane 3), K69A Walker A mutant (lane 4), and E147A Walker B mutant (lane 5) under the control of the native orc1-1 promoter.

(B) ChIP analysis of the indicated Orc1-1 protein binding to oriC1 and a distal control locus (Sire_2328) in vivo. Error bars are ±SD.

(C) 2D gel analyses of replication initiation at oriC1 in the indicated strains. Red arrowheads indicate structures arising due to replication initiation.

(D) Flow-cytometry analyses of the cell-cycle profile of the indicated strains. Positions of 1C and 2C populations are indicated.

(E) Micrographs of the indicated strains (+wt = Δorc1-1 + pOrc1-1, +K69A = Δorc1-1 + pOrc1-1 [K69A], +E147A = Δorc1-1 + pOrc1-1 [E147A]). Phase contrast, membrane dye FM4-64X (red), and DAPI-stained (blue) images are shown.

(F) Quantitation of aberrant cells (scored as having a diameter greater than 1.5 μm). Note that large cells in the Walker A strain appear to have less DNA than the large cells in the Walker B strain.

See also Figures S3, S4, S7, and Tables S1, S2, and S4.
Figure 7. Biochemical Consequences of the Walker B E147A Mutation

(A) Nucleotide-dependent conformational changes in Orc1-1. Gel filtration analysis of 10 µg of WT Orc1-1-ADP or the E147A-ATP on a Superdex 75 column. The elution positions of size standards are indicated (conalbumin, 75 kDa; ovalbumin, 44 kDa; and ribonuclease A, 13.7 kDa).

(B) Composite three-dimensional contour plot of hydrodynamic parameters obtained from GA-MC50 analysis of WT and Walker B mutant Orc1-1 (E147A). WT analysis reveals the presence of two conformers of the monomer, while the Walker B mutant results indicate the presence of a single monomer conformer in addition to a dimer form (see also Figure S3).

(legend continued on next page)
The structure of the Aeropyrum pernix Orc1-1 ortholog bound to an ORB element from orIC1 is shown. The ISM is shown in red and the WH domain in blue. (Figure generated using Pymol from Protein Data Bank file 2V1U) DNA contact points are correspondingly color-coded. The structure is aligned to the sequence of the S. islandicus oriC1 ORB2 element. The left-hand gel image shows Fe·MPE cleavage patterns in the absence (–) or presence of 1 mM WT or Walker B E147A mutant protein (WB) in the presence of 2 mM ATP (WB). Positions of protection from cleavage are identified at the right of the panel. The right-hand panel shows DNase-footprinting performed in the absence (–) or presence of 1 mM WT or Walker B E147A mutant protein (WB) in the presence of 2 mM ATP (WB).

(d) EMSA analysis of WT or Walker B E147A mutant protein binding to a 231 nt probe containing S. islandicus orIC1. Lanes contain 0, 63, 125, 250, 500, or 1000 nM of the indicated Orc1-1 protein.

(E) Trypsin cleavage products following 22 hr digestion of WT (lanes 2, 4, 6, and 8) and Walker B mutant (lanes 3, 5, 7, and 9) in the absence (lanes 2–5) or presence (lanes 6–9) of ORB2 DNA.

(F) Identity of the terminal trypsin digestion site by determined by liquid-chromatography tandem mass spectrometry.

(G) Western analysis of MCM recruitment to paramagnetic beads containing either control DNA (c) or orIC1-containing DNA (o). Reactions contained DNA at 200 pM and were preincubated with buffer (lanes 3, 4, 9, and 10) or 1 mM WT recombinant protein (lanes 5, 6, 11, and 12) or Walker B mutant E147A Orc1-1 (lanes 7, 8, 13, and 14) in the presence (lanes 9–14) or absence (lanes 3–8) of 2 mM ATP. Following preincubation, extract was added, incubated, and then washed prior to elution of bound proteins by boiling with SDS-PAGE loading buffer. 5 and 0.5% input extract were loaded in lanes 1 and 2. MCM was detected by western blotting. See also Figure S4.

DISCUSSION

Taken together, our data yield insight into the evolution, regulation, and mechanism of replication initiation in the chromosome of Sulfolobus species. We find that the chromosome is a mosaic of a distinct clade of Orc1/Cdc6s; nevertheless, our AUC and gel filtration data suggest that the identity of the nucleotide bound by Orc1-1 also modulates the conformation of this protein in solution.

Superimposition of the ADP-NP-bound Orc1-2 structure onto the Orc1-1·ADP·DNA structure, aligning via the principal DNA-binding WH domain, suggests three possible consequences of ATP binding (Figure S4). First, the ISM could disengage from the DNA, conceivably allowing the AAA+ domain to be accessible for interaction with another protomer or client protein; second, both WH and ISM could remain engaged with DNA with consequent remodeling of the intervening duplex DNA; third, both domains remain engaged with DNA and the protein undergoes an internal conformational alteration without modulating DNA structure. To test for the first two scenarios, we performed methidium propyl EDTA footprinting with the WT and E147A proteins on an isolated DNA binding site, the ORB2 element. This technique allows clear resolution of the DNA sequences contacted by WH and ISM. As can be seen from Figure 7C, no discernable difference is seen between the protection patterns mediated by the two proteins, indicating no significant disengagement from, or remodeling of, the DNA binding site. Similarly, there was no discernable difference in the protection pattern upon DNasel digestion. The same result was observed with ORB1 and ORB3 DNA (data not shown). Furthermore, electrophoretic mobility shift assay (EMSA) analysis with full-length origin DNA fragments (Figure 7D) reveals no significant changes in the affinity for DNA or pattern of shifted complexes generated by either the Orc1-1·ADP or the Orc1-1 E147A·ATP forms of the protein. Additionally, we could not detect melting of origin DNA, by either potassium permanganate or nuclease P1 sensitivity, by either protein in the presence or absence of nucleotide (data not shown). In light of our AUC and gel filtration data revealing altered conformations of the monomeric protein, we therefore speculated that the principal effect of ATP binding was on the conformation of the protein itself rather than on that of the DNA. To investigate this further, we subjected WT and Walker B mutant Orc1-1 to limited proteolysis with trypsin in the presence and absence of saturating levels of ORB2 DNA. Under both conditions, we observe reduced levels of a ~19 kDa species with the Walker B mutant form. In the presence of DNA, the digest patterns of both forms of the protein are altered, with a novel 23 kDa species appearing (Figure 7E). This species was also more abundant with the WT protein. We recovered the gel bands, digested them with chymotrypsin, and subjected the resultant peptides to mass spectrometry (Figure 7F). The upper band yielded two peptides that corresponded to the product of trypsin digestion at R193 and chymotrypsin digestion at Y208 or F185. This is in good agreement with the apparent size of the species of ~23 kDa (trypsin cleavage at R193 would generate a C-terminal fragment of 22.8 kDa and a N-terminal fragment of 22.5 kDa). The ~19 kDa band also yielded a peptide derived from trypsin cleavage at R193 and chymotrypsin at F185. Thus, the ATP-bound form of Orc1-1, both in DNA-bound and unbound form, displays altered sensitivity of R193 to trypsin digestion. Significantly, R193 corresponds to the arginine finger of the protein’s AAA+ site. Thus, in the absence of any observable alterations to the affinity, geometry, or stoichiometry of the protein-DNA interaction between Orc1-1 and DNA, we propose that the differential sensitivity to trypsin reflects a conformational alteration of the ATP-bound form of the protein on DNA that renders it permissive to recruitment of the MCM helicase.

Finally, we established an in vitro assay for MCM recruitment to oriC1. In this assay, we immobilized origin or nonorigin DNA on paramagnetic beads, preincubated with WT (ADP-bound) or E147A Orc1-1 in the presence or absence of supplemental ATP and then added cell extract prepared from logarithmically growing S. islandicus cells. Following incubation at 50°C and washing, beads were recovered and bound proteins eluted by boiling in SDS-PAGE loading buffer prior to western blotting. As can be seen in Figure 7G, neither extract alone nor extract supplemented with WT (ADP-bound) Orc1-1 supports significant MCM recruitment in the absence of ATP. Extract with additional ATP supports recruitment, but interestingly this recruitment is lost by the addition of 1 mM WT Orc1-1·ADP, presumably as a consequence of competition with endogenous Orc1-1 for the 200 pM template. Importantly, however, the E147A mutant supported robust, ATP-stimulated recruitment of MCM in these assays.
of three distinct replicons, each with its own principal initiator protein. In particular, we note that initiation at oriC3 is not impacted upon by individual or combined Orc1/Cdc6 deletions but is abrogated by loss of expression of the Cdt1 homolog WhiP. Intriguingly, parallels can be drawn with metazoa: in *Drosophila*, the origin recognition complex (ORC) is dispensable for endoreduplication, whereas Cdt1 is essential for this process (Park and Asano, 2008). We note that it was not possible for us to generate a triple Orc-1 mutation in which all three paralogs were deleted. Given that Orc1-2 does not bind in vivo to any origin, we speculate that this synthetic lethality may reflect nonorigin roles for the Orc1/Cdc6 proteins. Indeed, our ChIP-seq analyses reveal multiple nonorigin genomic binding sites for the Orc1/Cdc6 proteins. Furthermore, the single mutants show significantly altered transcript profiles compared to WT cells (R.Y.S. and S.D.B., unpublished data). Thus, as with DnaA and eukaryal ORC proteins (Scholefield et al., 2011), it appears that archaeal Orc1/Cdc6 proteins play roles in transcriptional regulation.

We speculate that the mosaic nature of the *Sulfolobus* chromosome has arisen by integration of extrachromosomal elements into an ancestral chromosome that presumably contained a single origin of replication. Given the broad conservation of Orc1-1 orthologs and their DNA binding sites across the archaeal domain of life (Robinson et al., 2004), we suggest that the ancestral chromosome was dependent on Orc1-1 and oriC1. Because WhiP homologs are found in the Desulfurococcales and Sulfohalobiales and Orc1-3 is restricted to the Sulfohalobiales, we propose the sequential acquisition of first WhiP/oriC3 prior to the bifurcation of these two lineage, followed by integration of the Orc1-3/oriC2 replicon in the branch leading to the Sulfohalobiales. Interestingly, the genome replication mode of the bacterium *Vibrio cholera* also appears to have been sculpted by acquisition of extrachromosomal element replication origins. However, in that organism, the genome is split between two chromosomes, each with a single origin: one dependent on a classic bacterial oriC/DnaA system and the second having an oriC with features reminiscent of plasmid origins. Indeed, this partitioning of bacterial genomes into multiple single-origin chromosomes is found in a range of bacterial species (Egan et al., 2005; Egan and Waldor, 2003). In contrast, multiple-origin single-chromosome genomes are found in a number of archaea. Why are archaeal chromosomes permissive in this regard? One explanation may lie in the contrasting mechanisms for replication termination in bacteria versus archaea (Duggin et al., 2008b; Toro and Shapiro, 2010). Whereas termination in bacteria involves polar fork traps, we have recently found that termination in *Sulfolobus* is by random fork collision (Duggin et al., 2011). In support of this, our MFA data on the mutant strains in which only a single origin fires reveal that termination is repositioned to occur 180° across the chromosome from the active origin (Figure 2F). An ectopic origin integrated into one replicochrome of a bacterial chromosome would result in premature arrival and prolonged stalling of a fork at a termination site. This may impact genome stability. In addition, because of the presence of the fork traps, the time taken to replicate a two-origin bacterial chromosome would still be dependent on the length of time it takes the fork arising from oriC to traverse the longest replicochrome. In contrast, the absence of active termination sites in *Sulfolobus* would mean that additional origins would simply result in repositioning of termination events to midway between active origins and a concomitant reduction in the total time spent in S phase, assuming similar timing of firing of the origins. The ability to accommodate and regulate multiple origins will have played a major role in permitting the increase in genome size seen in eukaryotic organisms (Difley, 2011).

Our data also reveal mechanistic insight into the action of the Orc1-1 initiator. As with bacterial DnaA, we find that the ATP-bound form of the initiator is the active form (Kaguni, 2011; Mott and Berger, 2007). Our expression profiling reveals a wave of transcription of the genes for the Orc1-1 and Orc1-3 initiators immediately prior to the onset of replication. We propose that this imposes a temporal window for initiation in which the newly synthesized proteins bind to ATP and associates with origin DNA before hydrolysis of ATP to ADP inactivates the proteins. It may, therefore, be significant that Orc1-1 and Orc1-3 are encoded in the immediate vicinity of their cognate origins. Furthermore, expression of WT Orc1-1 driven by its native promoter from episomes, while resulting in origin occupancy (as adjudged by ChIP), fails to activate the origin, presumably due to hydrolysis of ATP in the cell prior to origin binding. However, expression of the Walker B ATP hydrolysis-defective mutant to the same level results in an overreplication phenotype.

In bacteria, the ATP-bound form of DnaA facilitates the formation of higher-order structures and is able to mediate localized DNA melting at the origin (Bramhill and Komberg, 1988; Duderstadt et al., 2011; Erzberger et al., 2006; Katayama et al., 2010; Ozaki et al., 2008). Although we do observe a subpopulation of the Walker B mutant form of the protein existing as dimer in solution, we find that the ADP- and ATP-bound forms of Orc1-1 show equivalent DNA binding properties. Furthermore, we have been unable to observe any localized sensitivity to nuclease P1 or potassium permanganate, indicative of DNA melting, induced by either form of the protein (data not shown). Perhaps unsurprisingly, in this regard Orc1-1 is behaving in a manner more reminiscent of eukaryal ORC, there is no evidence for eukaryal ORC melting origins; rather, the available data support a model whereby MCM(2-7) is recruited to double-stranded DNA with duplex unwinding occurring at a later stage (Evrin et al., 2009; Heller et al., 2011; Remus et al., 2009). We propose that this is an evolutionarily conserved feature of the Orc1/Cdc6 proteins and that these proteins act principally to define loci for MCM recruitment. In light of our protease sensitivity assays, we suggest that ATP acts to modulate the conformation of Orc1/Cdc6 proteins on the DNA binding site such that the protein is able to effectively recruit MCM. Subsequent hydrolysis of ATP to ADP by Orc1/Cdc6 then renders the protein incapable of recruiting further MCM, thereby contributing to the fidelity of replication control. This simple binary switch model of archaeal Orc1/Cdc6 function is supported by our observation of the overreplication phenotype mediated by the Walker B mutant of Orc1-1. We note that in budding yeast, temporally distinct ATP binding and hydrolysis events by both the ORC complex and Cdc6 are required for iterative MCM(2-7) loading at replication origins (Randell et al., 2006). The increased organizational complexity of the eukaryotic origin definition and MCM recruitment system is likely a consequence of the increased regulatory demands.
imposed by the large number of replication origins found in eukaryotic chromosomes (Diffley, 2011). The simple binary-switch model that we propose for Orc1/Cdc6 function is clearly distinct from the mechanism of action of the bacterial initiator DnaA and provides a paradigm for the function of initiator proteins in archaea and eukaryotes.

**EXPERIMENTAL PROCEDURES**

### Strains, Media, and Growth Conditions

*S. islandicus* strain E2335 (αpyrEF ΔacrS) constructed previously (Deng et al., 2009) was used as the host for genetic manipulations. All archaeal strains used in this work, including the genetic host and its transformants of diverse plasmids as well as the orc1/whip, are summarized in Table S1. These archaeal cell lines were grown in SCVY (0.2% sucrose, 0.2% Casamino acids, 5 ml vitamin solution, and 0.005% yeast extract) or TY5 (0.1% tryptone, 0.05% yeast extract, and 0.2% sucrose) medium at 78°C. If required, uracil was supplemented to 20 μg/ml while 5fluoroorotic acid (5-FOA) was added to 50 μg/ml in SCVY medium. Microscopy was performed as described previously (Samson et al., 2011).

### Protein Purification and Biochemical Analyses

Orc1/Cdc6 and WhiP proteins were expressed in *E. coli* with C-terminal Hexa-His tags and purified as described previously (Robinson and Bell, 2007; Dueber et al., 2007). EMSAs and footprinting were performed as described previously (Dueber et al., 2007; Robinson et al., 2004). The MCM recruitment assay was based on the approach initially established by Seki and Diffley (2000). A total of 10 μg of a 5 kb plasmid (pCR Script containing a 2 kb sequence centered around oriC) or a control locus, Sire, 2328) was linearized by EcoRI and PstI digestion, the EcoRI-generated recessed end was filled in with Klenow fragment (New England Biolabs) and Biotin-14-dATP and dTTP. The DNA was recovered from the reaction mixture using QiaQuick column (Qiagen) and immobilized on Dynabeads M-280 streptavidin beads following the protocols of recombinant proteins (Life Technologies). Binding reactions were assembled in LoBind tubes (Eppendorf) and contained 200 μM DNA in 50 μl of binding buffer (20 mM Tris acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT (pH 7.9) at 25°C) plus 20 ng/μl poly(dIdC). Reactions were heated to 50°C in a water bath and buffer or Orc1-1 protein (1 μM final) were added and supplemented with nucleotide to 2 mM as indicated. Following a 10 min incubation, 20 μl of a 5 mg/ml whole-cell extract of *S. islandicus* was added (see below). Then, 20 μl of water or supplemental ATP to 2mM as indicated was added and the reaction incubated for a further 30 min at 50°C before washing twice with magnet-mediated bead pelleting and resuspension in 100 μl of prewarmed binding buffer, followed by one wash in 100 μl binding buffer containing 500 mM potassium acetate. Following this final wash, beads were resuspended in 1xSDS PAGE loading buffer and subjected to SDS PAGE followed by western blotting. All wash steps employed preheated buffers and were performed in the 50°C water bath. We observed a very high background of recombinant Orc1-1 protein bound to both DNA substrates and even to beads alone (data not shown), but MCM was only efficiently recruited to the origin-containing DNA substrate.

The *S. islandicus* cell extract was prepared by growing cells to midexponential phase (OD 600 nm = 0.4) before harvesting and resuspension in 2x binding buffer (40 mM Tris acetate, 100 mM potassium acetate, 20 mM magnesium acetate, 2 mM DTT (pH 7.9) at 25°C). Cells were lysed using a FrenchPress (Thermo) and the extract was clarified by centrifugation at 23,700 × g at 4°C. Then, 100 μl aliquots were prepared in LoBind tubes and snap frozen on dry ice. Total protein concentration was determined by Bradford assay to be 5 mg/ml. Extracts were stored at −80°C.

### Flow Cytometry

Flow cytometry of *S. islandicus* cells was performed essentially as previously described for *S. acidocaldarius* (Duggin et al., 2008a). Cells were fixed with 70% ethanol and their DNAs were stained with Mithramycin A and analyzed with A-40 Apogee flow cytometer.

### Neutral-Neutral 2D Gel Electrophoresis

*S. islandicus* cells were collected by centrifugation and washed twice with TEN buffer (50 mM Tris-Cl [pH 8], 50 mM EDTA [pH 8], 100 mM NaCl). Cell suspensions were then mixed with low-melting-point agarose and dispensed into Bio-Rad plug molds. Genomic DNA preparation and 2D gel electrophoresis were performed as described previously (Robinson et al., 2004). For further details, please refer to Extended Experimental Procedures.

### Chromatin Immunoprecipitation

ChiP was performed as described by Robinson et al. (2004). For further details, please refer to Extended Experimental Procedures.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.01.002.

**ACKNOWLEDGMENTS**

Work in Q.S.’s laboratory is supported by a special grant from Huazhong Agricultural University and the Danish Independent Research Councils (FTP/11-106683), and work in S.D.B.’s laboratory is funded by the Wellcome Trust (086045/Z/08/Z) and by the College of Arts and Sciences, Indiana University. We thank the staff at the Wellcome Trust Center for Human Genetics for performing the NGS described in Figure 3E.

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

**REFERENCES**


