Europe PMC Funders Group
Author Manuscript
Nature. Author manuscript; available in PMC 2016 December 08.

Published in final edited form as:
Nature. ; 534(7607): 412–416. doi:10.1038/nature17962.

The bacterial DnaA-trio replication origin element specifies ssDNA initiator binding

Tomas T. Richardson, **Omar Harran**, and **Heath Murray**
Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle Upon Tyne, NE2 4AX, UK

Abstract

DNA replication is tightly controlled to ensure accurate inheritance of genetic information. In all organisms initiator proteins possessing AAA+ (ATPases associated with various cellular activities) domains bind replication origins to license new rounds of DNA synthesis1. In bacteria the master initiator protein, DnaA, is highly conserved and has two crucial DNA binding activities2. DnaA monomers recognise the replication origin (*oriC*) by binding double-stranded DNA sequences (DnaA-boxes); subsequently, DnaA filaments assemble and promote duplex unwinding by engaging and stretching a single DNA strand3–5. While the specificity for duplex DnaA-boxes by DnaA has been appreciated for over thirty years, the sequence specificity for single-strand DNA binding remained unknown. Here we identify a new indispensable bacterial replication origin element composed of a repeating 3-mer motif that we term the DnaA-trio. We show that the function of the DnaA-trio is to stabilise DnaA filaments on a single DNA strand, thus providing essential precision to this binding mechanism. Bioinformatic analysis detects DnaA-trios in replication origins throughout the bacterial kingdom, indicating that this element comprises part of the core *oriC* structure. The discovery and characterisation of the novel DnaA-trio extends our fundamental understanding of bacterial DNA replication initiation, and because of the conserved structure of AAA+ initiator proteins these findings raise the possibility of specific recognition motifs within replication origins of higher organisms.

The master bacterial DNA replication initiator, DnaA, is a highly conserved multifunctional protein that utilises distinct domains to achieve its two key DNA binding activities. DnaA recognises double-stranded DNA using a helix-turn-helix motif (domain IV), whereas an ATP-dependent DnaA filament interacts with a single DNA strand using residues within the initiator specific motif (ISM; an α-helical insertion that distinguishes the family of

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding author: Heath Murray, heath.murray@newcastle.ac.uk, Tel: +44 (0) 191 208 3233, Fax: +44 (0) 191 208 7424.

Author Contributions
H.M. and T.T.R. conceived and designed experiments; H.M., T.T.R. and O.H. constructed plasmids and strains; H.M. and O.H. performed growth and marker frequency analysis experiments; H.M. performed microscopy experiments; T.T.R purified proteins, performed the open complex assay, and performed the DnaA filament formation assays; H.M. and T.T.R. interpreted results and wrote the paper.

Author Information
Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.M. (heath.murray@newcastle.ac.uk).

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

replication initiators) of the AAA+ domain (domain III) (Extended Data Fig. 1a-d)4,5. In contrast to DnaA, bacterial replication origins are diverse; they contain variable numbers of DnaA-boxes and seemingly lack a common architecture6,7. Therefore, the sequence information within *oriC* that directs DnaA filament assembly onto a single DNA strand is unknown.

To investigate how DnaA filament formation could be localised to the DNA replication origin of *Bacillus subtilis* we began by characterising site-directed mutants of the DNA unwinding region *in vivo* (Fig. 1a and Extended Data Fig. 1e). In order to enable identification of essential sequences without selecting for suppressor mutations, we generated a strain in which DNA replication could initiate from a plasmid origin (*oriN*) integrated into the chromosome (Fig. 1b and Supplemental Information). Activity of *oriN* requires its cognate initiator protein, RepN; both of these factors act independently of *oriC*/DnaA8. Expression of *repN* was placed under the control of a tightly-regulated inducible promoter, thus permitting both the introduction of mutations into *oriC* and their subsequent analysis following removal of the inducer to shut off *oriN* activity (Fig. 1c and Extended Data Fig. 2).

At the *B. subtilis* replication origin DNA unwinding by DnaA is detected downstream of DnaA-box elements and includes a sequence of 27 continuous A:T base pairs that is thought to facilitate DNA duplex opening (Fig. 1a)9. Surprisingly, we were able to delete the entire AT-rich 27-mer (Δ27) without abolishing origin activity, although the mutant strain did display a slow growth phenotype indicating that the AT-cluster is required for efficient origin function (Fig. 1d). Interestingly, further deletions extending 3 or 6 base pairs (Δ30, Δ33) severely impaired *oriC*-dependent initiation (Fig. 1d) and a deletion series targeting the sequence between the GC-rich and AT-rich clusters confirmed that this region alone was essential for origin function (Fig. 1e). Scrambling this entire region also inhibited cell growth (t1–t6^{Scr}), demonstrating that the specific sequence is required, rather than the spacing between the flanking elements (Fig. 1f). To explore this region in more detail sequences were scrambled three base pairs at a time by exchanging each triplet for its complement. Phenotypic and marker frequency analyses revealed that disruption of sequences closest to the GC-cluster (t1^{Scr} and t2^{Scr}) caused the greatest defect in DNA replication initiation, indicating that the region proximal to the DnaA-boxes is most important for origin activity (Fig. 1f). Although mutagenesis of neither t4 nor t5 alone produced a detectable effect on DNA replication initiation under the conditions tested, they may become important when origin firing is suboptimal as was observed in the AT-cluster deletion mutant (Fig. 1d).

To determine whether this essential DNA sequence between the GC- and AT-clusters has a role in DNA melting *per se*, an open complex formation assay was performed. DnaA was incubated with *oriC* plasmids containing either the wild-type or scrambled sequence (t1-t6^{Scr}), potassium permanganate was added to oxidise distorted bases within the DNA, and base modification was detected by primer extension. Scrambling the sequence inhibited open complex formation, indicating that this region is necessary for DnaA-dependent unwinding (Fig. 1g).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

DnaA monomers are thought to bind DnaA-boxes prior to ATP-dependent filament formation10. Using the strain capable of *oriC*-independent initiation the seven DnaA-box sequences were individually scrambled to abolish DnaA binding11. Culturing these strains in the absence of *oriN* activity revealed that mutation of DnaA-box6 severely inhibited growth and mutation of DnaA-box7 resulted in a significant growth defect, while mutation of the remaining DnaA-boxes had no observable effect (Fig. 2a). Marker frequency analysis confirmed that mutation of DnaA-box7 drastically impaired origin activity, whereas mutation of the remaining DnaA-boxes resulted in only modest decreases in initiation frequency (Fig. 2a). These results indicate that DnaA-boxes proximal to the essential unwinding region are most critical for origin activity.

To directly test whether these DnaA-boxes promote DnaA filament assembly at the essential unwinding region we utilised a previously described DnaA filament formation assay12. Here two cysteine residues are introduced within the AAA+ domain such that the protein remains functional and when the DnaA filament assembles the cysteine residues from interacting protomers come into close proximity. DNA scaffolds were assembled using oligonucleotides and the cysteine-specific crosslinker bis(maleimido)ethane (BMOE; 8 Å spacer arm) was used to capture the oligomeric species formed on each substrate.

Incubation of DnaA with duplex substrates containing DnaA-box6, DnaA-box7 and the GC-rich region produced a dimeric species (Fig 2b-c), whereas incubation of DnaA with a longer duplex substrate containing the unwinding region produced a set of larger oligomeric complexes. We wondered whether the larger species were being formed on the duplex DNA or on a single DNA strand. To test these models scaffolds containing single-stranded (ss) DNA tails were assembled. DnaA filaments readily formed on substrates containing a 5′-tail but were absent when the corresponding 3′-tail was provided (Fig. 2c and Extended Data Fig. 3). Formation of DnaA oligomers on the 5′-tailed substrate was dependent upon both ATP and the ssDNA binding residue Ile190 located within the ISM of the AAA+ domain4, indicating that the assay was capturing DnaA filament formation on ssDNA (Fig. 2c–d, Extended Data Fig. 1d).

Critically, DnaA oligomer formation on the 5′-tailed substrate was specific. DnaA filament assembly was abolished when the DnaA-box sequences within the duplex region were scrambled and it was significantly reduced when the single-stranded region was replaced with its complementary sequence (Fig. 2c). Taken together these results suggest that DnaA filaments are loaded from duplex DnaA-boxes onto ssDNA bearing a 5′-tail. This model is consistent with both biochemical experiments showing that *E. coli* DnaA preferentially interacts with the corresponding single-strand of its DUE and single molecule studies showing that *Aquifex aeolicus* DnaA filaments form with 3′→5′ polarity11,13,14.

DnaA oligomer size was proportional to the length of the 5′-tail up to the formation of a heptamer, after which further DNA extension did not promote longer filaments (Fig. 2e-f). We noted that this limit corresponded to a poly(A) tract in the DNA sequence and wondered whether this sequence inhibited DnaA filament formation. When the poly(A) tract was replaced by sequences from the beginning of the DNA unwinding region, DnaA oligomer length increased beyond a heptamer (Fig. 2e-f). This result suggests that the origin

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

unwinding region is designed to limit DnaA filament formation to a precise position within *oriC*.

To identify a possible single-strand binding motif recognised by DnaA, individual base pairs within the essential unwinding region were inverted and origin activity was analysed *in vivo*. Marker frequency analysis revealed that altering either of two A:T base pairs, which were spaced three nucleotides apart from each other, resulted in the most significant loss of origin activity; in contrast the surrounding mutations had only modest effects (Fig. 3a). Re-examination of the unwinding region shows that A:T base pairs are spaced at three nucleotide intervals throughout this sequence (Fig. 1a). This observation is strikingly congruent with the mechanism proposed for binding of the DnaA filament to ssDNA, where each protomer engages a set of three nucleotides (Fig. 3b)4.

We hypothesised that an array of triplet nucleotide motifs recognised by DnaA are present within the unwinding region and that the motifs proximal to the DnaA-boxes are most important for origin activity. To test this model *in vivo* we created a set of nested deletions that removed either one or three base pairs (Extended Data Fig. 4). All of the single base pair deletions dramatically lowered the replication initiation frequency and several significantly inhibited cell growth (especially the same A:T base pairs noted above), whereas triplet deletions encompassing the single deletions had little or no effect (Fig. 3c-d). These results are consistent with the model that single base pair deletions act both by disrupting a specific 3-mer motif and by shifting the register of downstream 3-mer motifs relative to the DnaA filament start point at the DnaA-boxes.

To test the model that the ssDnaA binding motif is indeed a repeating 3-mer, DnaA filament formation was analysed *in vitro* using tailed substrates which contained either single or triplet base deletions (Fig. 3e). Whereas deletion of one base produced significantly shorter oligomers, deletion of three bases restored formation of full length complexes. Taken together with the *in vivo* deletions, these results indicate that DnaA filaments bind to ssDNA by recognising a specific 3-mer motif found within the unwinding region. We have termed this 3-mer motif the “DnaA-trio”.

To define the precise sequence of the DnaA-trio, DnaA filament formation was observed using a series of DNA scaffolds in which the 5′-tails were extended by increments of one nucleotide. We observed that additional oligomeric species appeared after the following sequences were added: 3′-GAT-5′, 3′-AAT-5′ and 3′-GAA-5′, suggesting that these triplets represent individual DnaA-trio motifs (Fig. 4a).

However, it was surprising that a longer oligomer was not formed following addition of the first 3′-GAT-5′ motif proximal to the GC-cluster, since mutagenesis of this sequence *in vivo* resulted in strong phenotypes (Fig. 3a, 3c-d). In structures of the archaeal initiator Orc1 bound to a replication origin the protein was observed to make two contacts with the DNA, one through its C-terminal DNA binding domain (analogous to DnaA domain IV) and another through its AAA+ motif15,16. We wondered whether DnaA might similarly be capable of contacting both a DnaA-box and the first DnaA-trio, thereby accounting for the absence of a DnaA trimer. Significantly, BMOE crosslinking of cysteines in the AAA+

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

domain would not detect this activity as the assay captures DnaA oligomers formed on either double-stranded or ssDNA12.

In order to test this hypothesis we employed the amine-specific crosslinker bis(sulfosuccinimidyl)suberate (BS^3) which, in contrast to BMOE, only captures DnaA oligomers formed on a single DNA strand (Extended Data Fig. 3 and 5). Crosslinking by BS^3 reveals a DnaA dimer forming in the presence of the first 3′-GAT-5′, indicating that DnaA does recognise this sequence (Extended Data Fig. 5). Taken together with the BMOE crosslinking showing that a DnaA dimer is formed on the double-stranded DNA scaffold containing just DnaA-boxes 6 and 7 and the GC-cluster, the data suggests that the DnaA protein initially bound at DnaA-box7 undergoes a conformational change (detected by BS^3) to engage the first DnaA-trio motif following the GC-cluster. Several lines of evidence support the notion that DnaA adopts distinct conformations when it engages either double-stranded or ssDNA17,18.

To support the assignment of the DnaA-trio we performed a targeted mutagenesis of the proposed sequence. The results indicate that each of the positions (3′-GAT-5′) appears important for DnaA filament formation, specifically the nucleotides at positions 1 and 2, and the deoxyribose group at position 3 (Fig. 4b and Extended Data Fig. 6). Interestingly, in the crystal structure of DnaA bound to a ssDNA substrate the protein makes no base-specific contacts4. These observations suggest either that the sequence of the DnaA-trios is important for an intermediate step in DNA duplex recognition and melting prior to full engagement of the product single-strand, or that the specific base sequence promotes the DNA backbone to adopt a favourable geometry for DnaA binding.

Using this information we first searched for DnaA-trios within other well-characterised origin unwinding elements (Fig. 4c; underlined)19–21. In these cases a set of at least three DnaA-trios could be identified. These DnaA-trios were located proximal to a DnaA-box that shared the same orientation as *B. subtilis* DnaA-box7 and the regions between the DnaA-box and the DnaA-trios were GC-rich. Using these additional criteria we next interrogated predicted bacterial DNA replication origins (DoriC22) for similar patterns. Figure 4c shows that similar elements can be identified within putative *oriC* regions throughout the bacterial kingdom. A sequence logo of the DnaA-trios indicates that the preferred motif is 3′-$^{G}/_{A}$AT-5′ (Fig. 4d), with the central adenine being most highly conserved. We also observed that in the majority of cases a pair of tandem DnaA-boxes preceded the GC-cluster (Extended Data Table 1).

We propose that the DnaA-trio constitutes a new element within bacterial replication origins. Our findings indicate that DnaA-trios play an essential role during DNA replication initiation by providing specificity for DnaA filament formation on a single DNA strand, thereby promoting DNA duplex unwinding. Our analysis also indicates that the arrangement of tandem DnaA-boxes in close proximity to DnaA-trios comprises a widespread strategy used to direct DnaA filament growth onto the unwinding region, with a single DnaA protein likely binding double-strand DNA via domain IV prior to engaging a DnaA-trio via its AAA + motif (Fig. 4e and Extended Data Fig. 5). Together our data is consistent with the two-step DnaA assembly model for DNA melting18.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

We note that the configuration between DnaA-boxes and the DnaA-trios is not strictly required for DnaA to be loaded onto the single DNA strand *in vitro*. Scaffolds containing either a single DnaA-box or containing DnaA-boxes in reverse orientation are competent to promote DnaA filament formation from the duplex DNA onto the 5′-tail, although in the latter situation DnaA filament formation was reduced suggesting that DnaA-box orientation is important (Extended Data Fig. 7). Furthermore, loading does not require the flexibly tethered domains I/II of DnaA, consistent with previous observations suggesting that domains III/IV can adopt multiple conformations (Extended Data Fig. 7)12,17,18. These results suggest that some plasticity can be accommodated between duplex and single-strand DNA binding elements, which is in agreement with recent reanalyses of essential DnaA-boxes in *E. coli* and may also explain the location of atypical origin unwinding sites23–26.

Analysis of replication initiator proteins from both bacteria and archaea shows that the ISM within AAA+ domains is utilised for DNA binding and the recent structure of the *Drosophila* origin recognition complex (ORC) suggests that this is also likely the case for eukaryotes, supporting the model that DNA binding by the ISM is a universal feature of replication initiators4,15,16,27. We find here, for the first time, that the interaction of the *B. subtilis* replication initiator ISM with the origin involves recognition of a specific DNA sequence. We speculate that motifs analogous to the DnaA-trio may be present in replication origins of higher organisms and be recognised by the ISM of ORC proteins. These sites need not be 3-mers, nor would they necessarily share the same spacing observed for the DnaA-trios as they would need to accommodate the arrangement of AAA+ interactions within the respective heterohexameric ORC17,27. The discovery of ISM binding motifs in higher organisms would greatly facilitate origin identification, an elusive problem precluding the understanding of DNA replication control in eukaryotes.

Methods

Media and chemicals

Nutrient agar (NA; Oxoid) was used for routine selection and maintenance of both *B. subtilis* and *E. coli* strains. For experiments in *B. subtilis* cells were grown using Luria-Bertani (LB) medium. Supplements were added as required: chloramphenicol (5 μg/ml), erythromycin (1 μg/ml), kanamycin (5 μg/ml), spectinomycin (50 μg/ml). Unless otherwise stated all chemicals and reagents were obtained from Sigma-Aldrich.

Phenotype analysis of oriC mutants using the inducible oriC-independent strain

Strains were grown for 18-72 hr at 37°C on NA plates either with or without IPTG (1 mM). All experiments were independently performed at least twice and representative data is shown.

Marker frequency analysis

Genomic DNA was harvested from cells during the exponential growth phase and the relative amount of DNA from the replication origin (*ori*) and terminus (*ter*) was determined by qPCR. Strains were grown in LB medium to an A_{600} of 0.3-0.5 whereupon sodium azide (0.5%; Sigma) was added to prevent further metabolism. Chromosomal DNA was isolated

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

using a DNeasy Blood and Tissue Kit (Qiagen). The DNA replication origin (*oriC*) region was amplified using primers 5′-GAATTCCTTCAGGCCATTGA-3′ and 5′-GATTTCTGGCGAATTGGAAG-3′; the region adjacent to *oriN* was amplified using primers 5′-CTTTCTGCCGCAAAGGATTA-3′ and 5′- CCTCTTCATAGCCGTTTTGC-3′; the DNA replication terminus (*ter*) region was amplified using primers 5′-TCCATATCCTCGCTCCTACG-3′ and 5′-ATTCTGCTGATGTGCAATGG-3′. Either Rotor-Gene SYBR Green (Qiagen) or GoTaq (Promega) qPCR mix was used for PCR reactions. qPCR was performed in a Rotor-Gene Q Instrument (Qiagen). By use of crossing points (C_T) and PCR efficiency a relative quantification analysis ($\Delta\Delta C_T$) was performed using Rotor-Gene Software version 2.0.2 (Qiagen) to determine the origin:terminus (*ori:ter*) ratio of each sample. These results were normalised to the *ori:ter* ratio of a DNA sample from *B. subtilis* spores, which only contain one chromosome and thus have an *ori:ter* ratio of 1. Error bars indicate the standard deviation of three technical replicates. All experiments were independently performed at least twice and representative data is shown.

Protein expression

BL21 (DE3)-pLysS cells were transformed with the appropriate expression construct (Supplementary Table 2) and selected on nutrient agar plates containing 100 ng/μl of ampicillin and 34 ng/μl of chloramphenicol. A single transformant colony was used to inoculate an overnight starter culture grown at 37 °C, 180 rpm, in LB medium supplemented with 100 ng/μl of ampicillin and 34 ng/μl of chloramphenicol. The following morning a 1/100 dilution of overnight culture was used to inoculate 1200 ml of LB medium supplemented with 100 ng/μl of ampicillin and grown at 37°C, 180 rpm, to an A_{600} of 0.5. Cells were induced with 1 mM IPTG and cultured for a further 3 hr at 30°C. Cells were pelleted at 3,000 g, 4°C for 10 min before resuspension in 45 ml pf resuspension buffer (25 mM HEPES-KOH (pH7.6); 500 mM potassium glutamate; 10 mM magnesium acetate; 20% sucrose; 30 mM imidazole; 1 x cOmplete EDTA-free protease inhibitor tablet (Roche). The cell suspension was then flash-frozen in liquid nitrogen.

Protein purification

DnaA (WT, WT-CC and I190A-CC) was purified as follows. A frozen 50 ml BL21 cell pellet suspension was thawed on ice with 32 mg of lysozyme and gentle agitation for 1 hr then disrupted by sonication at 20 W for 5 min in 2 s pulses. Cell debris was pelleted by centrifugation at 31,000 xg, 4°C for 45 min and the supernatant further clarified by filtration (0.45 μm). All subsequent steps were performed at 4 °C unless otherwise stated. The clarified lysate was applied at 1 ml/minute to a 1 ml HisTrap HP column (GE), which had previously been equilibrated with Ni binding buffer (25 mM HEPES-KOH [pH7.6]; 250 mM potassium glutamate; 10 mM magnesium acetate; 20% sucrose; 30 mM imidazole). The loaded column was washed with a 10 ml 1 step gradient of 10 % Ni elution buffer (25 mM HEPES-KOH [pH7.6]; 250 mM potassium glutamate; 10 mM magnesium acetate; 20% sucrose; 30 mM imidazole). Specifically-bound proteins were eluted using a 7.5 ml 1-step gradient of 100% Ni elution buffer and the entire fraction collected and diluted into 42.5 ml of Q binding buffer (30 mM Tris-HCl [pH7.6]; 100 mM potassium glutamate; 10 mM magnesium acetate; 1 mM DTT; 20% sucrose). The diluted fraction was then applied at 1 ml/min to a 1 ml HiTrap Q HP column (GE), which had previously been equilibrated with Q

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

binding buffer. The loaded HiTrap Q HP column was washed with 10 ml of Q binding buffer then eluted using a linear 10 ml gradient of 0–100% Q elution buffer (30 mM Tris-HCl [pH7.6]; 1 M potassium glutamate; 10 mM magnesium acetate; 1 mM DTT; 20% sucrose) with 1 ml fractions collected. The peak 3 × 1 ml fractions, based on UV absorbance, were pooled and dialysed into 1 L of FactorXa cleavage buffer (25 mM HEPES-KOH [pH7.6]; 250 mM potassium glutamate; 20% sucrose; 5 mM $CaCl_2$), using 3.5k MWCO SnakeSkin dialysis tubing (Life Technologies) at 4°C overnight. The dialysed protein was diluted to 5 ml total volume in FactorXa cleavage buffer and incubated at 23 °C for 6 hr with 80 μg of FactorXa protease (NEB). The sample was applied at 1 ml/min to a 1 ml HisTrap HP column (GE), which had previously been equilibrated with Factor Xa cleavage buffer. The Factor Xa-cleaved fraction was eluted in 7.5 ml of Ni binding buffer. The eluted fraction was diluted into 42.5 ml of Q binding buffer and purified on a 1 ml HiTrap Q HP column as previously described. Peak fraction(s) were pooled and dialysed into 1 L of final dialysis buffer (40 mM HEPES-KOH [pH7.6]; 250 mM potassium glutamate; 1 mM DTT; 20% sucrose; 20% PEG_{300}), using 3.5k MWCO SnakeSkin dialysis tubing (Life Technologies) at 4°C overnight before aliquoting, flash-freezing in liquid nitrogen and storage at -80 °C. Removal of the N-terminal His-tag, following incubation with FactorXa, was confirmed by anti-pentaHis (Qiagen) Western blotting.

C-terminally His-tagged DnaA (WT-CC and Δ(domainI-II)-CC) purification was performed as for the tag-free variants, except that the protein was dialysed into final dialysis buffer following the first HiTrap Q HP column purification before aliquoting, flash-freezing and storing.

HBsU purification was performed exactly as for DnaA, except that the HiTrap Q HP column was substituted for a 1 ml HiTrap Heparin HP column (GE) and the composition of buffers was modified accordingly. Ni binding buffer (25 mM Tris-HCl [pH8.0]; 400 mM NaCl; 30 mM imidazole). Ni elution buffer (25 mM Tris-HCl [pH8.0]; 400 mM NaCl; 500 mM imidazole). Heparin binding buffer (25 mM Tris-HCl [pH8.0]; 100 mM NaCl; 1 mM EDTA). Heparin elution buffer (25 mM Tris-HCl [pH8.0]; 2 M NaCl; 1 mM EDTA). Factor Xa cleavage buffer (25 mM Tris-HCl [pH8.0]; 100 mM NaCl; 2 mM $CaCl_2$; 20% sucrose). Final dialysis buffer (25 mM Tris-HCl [pH8.0]; 400 mM NaCl; 2 mM $CaCl_2$; 20% sucrose; 20 % PEG_{300}). Peak fractions were determined by SDS-PAGE and Coomassie staining due to the absence of tryptophan, tyrosine and cysteine residues.

Open complex formation assays

$KMnO_4$ footprinting assays were essentially performed as described in 9, except for the following changes. DnaA was not pre-incubated with ATP. The unwinding buffer contained 2 mM ATP, rather than 5 mM, and 500 ng of plasmid pTR541 (wild-type) or pTR542 (t1-$t6^{scr}$) was used per 75 μl-scale reaction. DnaA was added to final concentrations of 0, 100, 250, 500 and 1000 nM. Assembled reactions were incubated at 37 °C for 10 min. $KMnO_4$ treatment was then performed at 37 °C for 10 min. Six μl of β-mercaptoethanol was used to quench reactions; however, EDTA was omitted. $KMnO_4$-treated DNA was immediately purified using a Qiagen PCR clean-up kit, eluting in 20 μl of EB buffer. $KMnO_4$-treated templates were not linearised prior to primer extension. Primer extensions were performed

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

on a 20 μl scale using 0.1 U/μl of Vent exo- DNA polymerase (NEB) in 1x manufacturer's reaction buffer supplemented with 4 mM $MgSO_4$, 200 μM each dNTP, 200 nM Cy5-labelled oligonucleotide (5'-Cy5- AGCTTCAGCAGCATGTAAAAG-3') and 4 μl of PCR-purified template DNA per reaction. Reactions were subjected to thermocycling using a 3Prime thermal cycler (Techne) with 1 min initial denaturation at 98 °C, followed by 35 cycles of (10 s at 98 °C; 30 s at 55 °C; 30 s at 72 °C). Reactions were quenched by addition of an equal volume of stop buffer (95% formamide; 10 mM EDTA; 10 mM NaOH; 0.01% Orange-G) and products subjected to denaturing PAGE (6% acrylamide:bisacrylamide [19:1]; 8 M urea in 1x TBE). Resolved products were visualised using a Typhoon Trio Variable Mode Imager (GE Healthcare). The DnaA-trio marker was generated by primer extension performed under the same conditions as described for $KMnO_4$-treated substrates, but using a PCR product as template generated with a primer corresponding to the end of the first DnaA-trio (5'- TAGGGCCTGTGGATTTGTG -3'). All experiments were independently performed at least twice and representative data is shown.

Filament assembly assays (BMOE)

DNA scaffolds were prepared by mixing each oligonucleotide (50 nM final concentrations) in 10 mM HEPES-KOH (pH7.6), 100 mM NaCl and 1 mM EDTA. Mixed oligonucleotides were heated to 98°C for 5 min in a heat-block and slowly cooled to room-temperature in the heat-block before use. Filament formation was promoted by mixing DnaA-CC proteins (WT, I190A, ΔdomainI-II) (200 nM final concentration) with DNA scaffold (15 nM) on a 20 μl-scale in 30 mM HEPES-KOH (pH7.0), 100 mM potassium glutamate, 100 mM NaCl, 10 mM magnesium acetate, 25% glycerol, 0.01% Tween-20 and 2 mM nucleotide (ADP or ATP). Reactions were incubated at 37°C for 5-12 min before addition of 4 mM BMOE (ThermoFisher Scientific). Reactions were incubated at 37°C for 5-12 min before quenching by addition of 60 mM cysteine. Reactions were incubated once more at 37°C for 10-12 min before fixing in NuPAGE LDS sample buffer (ThermoFisher Scientific) at 98°C for 5 min. Complexes were resolved by running 500 fmol of cross-linked DnaA from each reaction on a NuPAGE Novex 3-8% Tris-Acetate gel (ThermoFisher Scientific) then transferred to Hybond 0.45 μm PVDF membrane (Amersham) in 0.5x NuPAGE Tris-acetate SDS running buffer with 20% MeOH at 35 mA, 4°C overnight using wet transfer apparatus (Biorad). Complexes were visualised by Western blotting using a polyclonal anti-DnaA antibody (Eurogentec). NB/ All filament assembly assays were performed using tag-free proteins with the exception of that shown in Extended Data Figure 9, in which C-terminally His-tagged proteins (ΔdomainI-II-CC and wild-type-CC) were employed. All experiments were independently performed at least twice and representative data is shown.

Filament assembly assays (BS^3)

Filament assembly assays using bis(sulfosuccinimidyl)suberate (BS^3) were performed as described for BMOE, except a tag-free fully wild-type recombinant DnaA protein was employed for Extended Data Figure 5. Tag-free "CC" variants of wild-type and I190A DnaA were used for Extended Data Figure 7a,b. Crosslinking was performed using BS^3 (15 mM final concentration) in place of BMOE and quenching performed by addition of Tris-HCl [pH7.6] (30 mM final concentration). All experiments were independently performed at least twice and representative data is shown.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Microscopy

To visualise GFP-DnaN, starter cultures were grown overnight in defined minimal medium base (Spizizen minimal salts supplemented with Fe-NH_4-citrate (1 μg/ml), $MgSO_4$ (6 mM), $CaCl_2$ (100 μM), $MnSO_4$ (130 μM), $ZnCl_2$ (1 μM), thiamine (2 μM)) supplemented with casein hydrolysate (200 μg/ml) and glycerol (0.5%) with IPTG (1 mM) at 37°C, diluted 1:100 into fresh medium with IPTG (1 mM) and allowed to grow at 37°C for several generations until they reached an A_{600} 0.3. Cells were collected by centrifugation, washed to remove IPTG, and resuspended into fresh medium at an A_{600} 0.1 and allowed to grow until an A_{600} 0.6. Cells were mounted on 1.5% agar pads (0.5X growth media) and a 0.13–0.17 mm glass coverslip (VWR) was placed on top. Microscopy was performed on an inverted epifluorescence microscope (Nikon Ti) fitted with a Plan-Apochromat objective (Nikon DM 100x/1.40 Oil Ph3). Light was transmitted from a 300 Watt xenon arc-lamp through a liquid light guide (Sutter Instruments) and images were collected using a CoolSnap HQ^2 cooled CCD camera (Photometrics). All filters were Modified Magnetron ET Sets from Chroma and details are available upon request. Digital images were acquired and analysed using METAMORPH software (version V.6.2r6). All experiments were independently performed at least twice and representative data is shown.

Strains

Strains are listed in Supplementary Table 1. The genotype of all origin mutants was confirmed by DNA sequencing.

Oligonucleotides

All oligonucleotides were purchased from Eurogentec. Oligonucleotides used for plasmid construction are listed in Extended Data Table 2. Oligonucleotides used to construct DNA scaffolds are listed in Extended Data Table 3.

Plasmids

Plasmids are listed in the Supplementary Table 2 (sequences are available upon request). DH5α (F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$)29 was used for plasmid construction, except where noted. Descriptions, where necessary, are provided below.

pHM327 derivatives were generated by quickchange mutagenesis using the oligonucleotides listed in Extended Data Table 2. Following sequencing to confirm mutated regions, sequences were subcloned using BglII/FspAI.

pHM446 [*bla aprE′ kan lacI* P_{spac}-MCS *′aprE*] is a derivative of pAPNC21330 with a kanamycin resistance cassette replacing the spectinomycin resistance cassette (gift from H. Strahl).

pHM453 [*bla rpnA′ rpmH erm* Δ*incAB* P_{spac}-*dnaA′*] was created in multiple steps. First, pJS1 was generated by ligation with a *Hin*dIII-*Bam*HI PCR product containing 5′ end of *dnaA* and pMUTIN431 cut with *Hin*dIII-*Bam*HI (gift from J. Errington). Second, pHM396 was generated by digestion of pJS1 with *Pvu*II (to remove *lacZ and lacI*) and ligation of the

vector backbone. Finally, pHM453 was generated by ligation of an *Aat*II PCR product containing *rpmH* and the 5′ end of *rpnA* (oHM319 + oHM320 and 168CA genomic DNA as template) with pHM396 cut with *Aat*II.

pHM492 [*bla aprE*′ *kan lacI* P_{spac}-*repN*(*oriN*) ′*aprE*] was generated by ligation of an *Eco*RI-*Xho*I PCR product containing *repN*(*oriN*) (oHM313 + oHM315 and MMB20832 genomic DNA as template) with pHM446 cut with *Eco*RI-*Sal*I.

pHM560 [*bla rpnA*′ *rpmH erm* Δ*incA* P_{dnaA} Δ*incB dnaA*′] was generated by ligation of an *Eco*RV-*Hin*dIII PCR product containing the *dnaA* promoter (oHM510 + oHM511 and 168CA genomic DNA as template) with pHM453 cut with *Eco*RV-*Hin*dIII.

pTR72, pTR73, pTR74, pTR102, pTR168 were generated by quickchange mutagenesis using the oligonucleotides listed in Extended Data Table 2.

pTR208 was generated by two-fragment PCR. oTR384/oTR385 and oTR386/oTR387 were used to amplify products using pTR74 and *B. subtilis* 168CA genomic DNA as templates, respectively. An equal volume of each PCR product was mixed, heated to 98°C and allowed to cool to room temperature before *Dpn*I digestion and transformation.

pTR229 [bla $_{pT7}$(his6-link-Xa-dnaA)] was generated by subcloning a *Hin*dIII-*Xho*I fragment of *dnaA* from pHM239 into the pTR74 backbone.

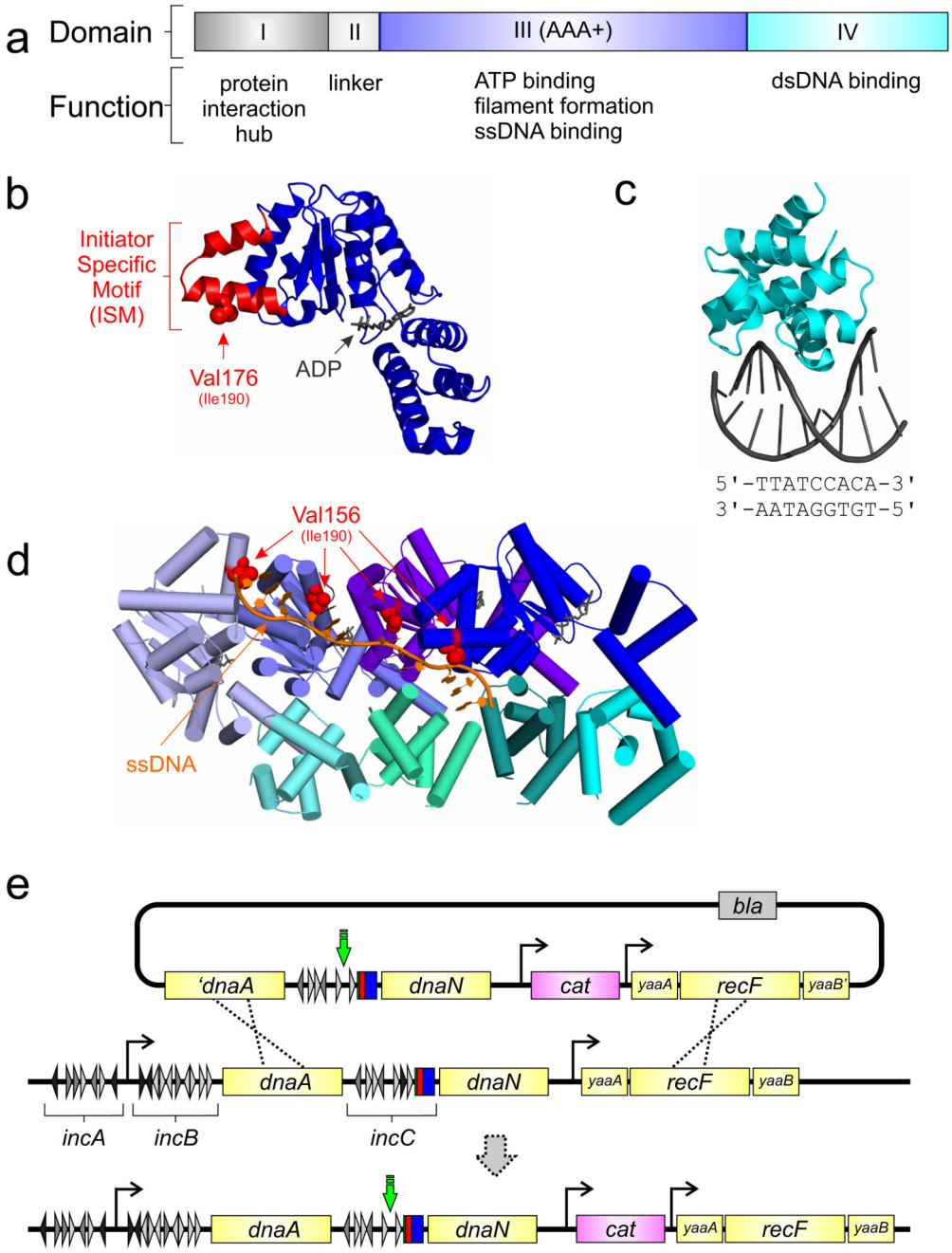
pTR541 and pTR542 were generated by two-fragment PCR. oTR537 and oTR538 were used to amplify the plasmid backbone of pSG1301. oTR535 and oTR536 were used to amplify *incC* with *B. subtilis* 168CA genomic DNA and pTR84 used as the templates for pTR541 and pTR542, respectively. An equal volume of each PCR product was mixed, heated to 98°C and allowed to cool to room temperature before *Dpn*I digestion and transformation into EH3827 (*asnB*32 *relA*1 *spoT*1 *thi*-1 *fuc*-1 *lysA ilv*-192 *zia*::pKN500 Δ*dnaA mad*-1)33. DNA sequencing confirmed the construction of each origin including flanking sequences (>400 bp upstream and downstream).

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

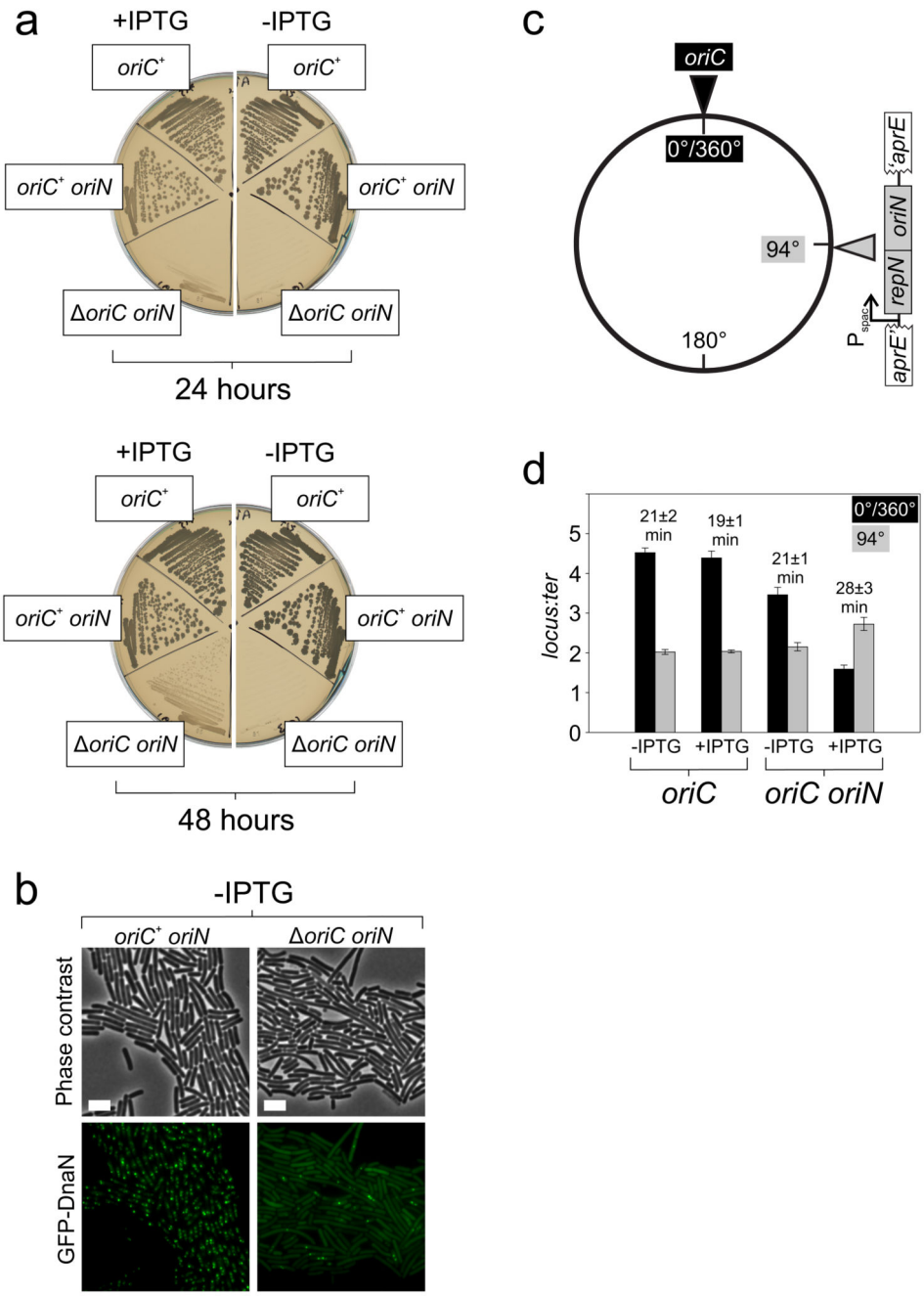
Extended Data



Extended Data Figure 1. Structure of DnaA proteins.

(a) Primary domain structure of DnaA. Key functions are listed below the relevant domain. **(b)** Structure of *Thermatoga maritima* DnaA domain III, highlighting the single-strand binding residue Val176 (Ile190 *B. subtilis*) within the ISM (PDB 2Z4S). **(c)** Structure of *Escherichia coli* DnaA domain IV bound to a DnaA-box (PDB 1J1V). **(d)** Structure of *Aquifex aeolicus* DnaA domain III (blue shades) and domain IV (cyan shades) bound to a

single DNA strand (orange), highlighting the single-strand binding residue Val156 (Ile190 *B. subtilis*) (PDB 3R8F). **(e)** Scheme used to construct mutants within the *B. subtilis* DNA replication origin. The green arrow highlights the location of a DnaA-box mutation.



Extended Data Figure 2. Characterisation of the inducible *repN/oriN* replication initiation system.

Repression of *repN* expression inhibits DNA replication in a Δ*oriC* mutant. A large deletion was introduced into the *B. subtilis* replication origin using a strain harbouring the inducible *oriN/repN* construct. Strain growth was found to be dependent upon addition of the inducer

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

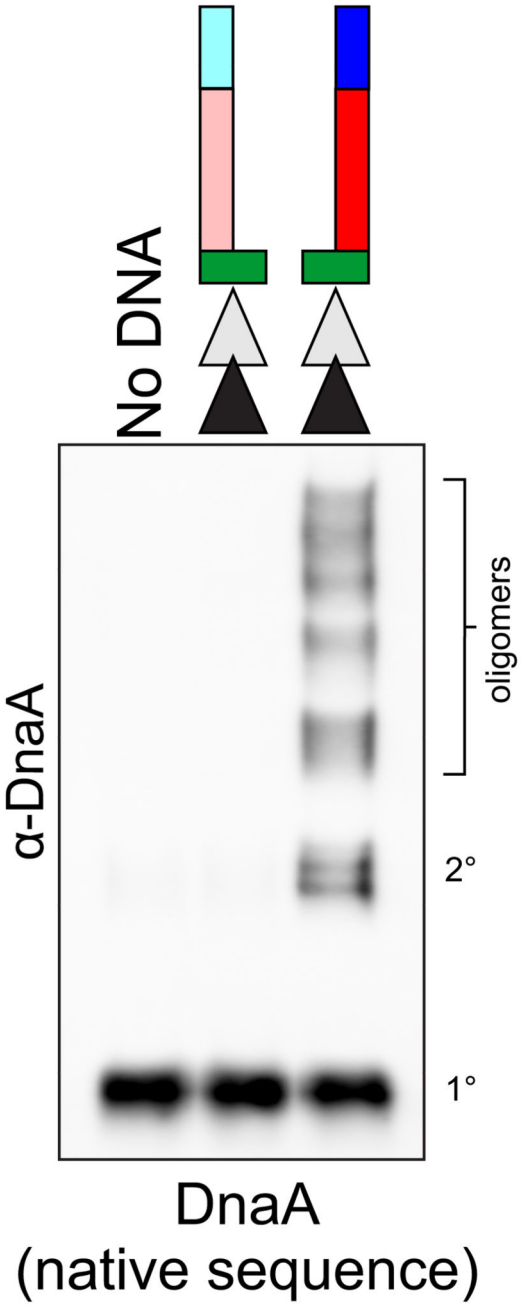
IPTG. **(a)** Strains streaked to resolve single colonies. **(b)** A GFP-DnaN reporter was used to detect DNA replication following removal of IPTG from inducible *oriN/repN* strains. Scale bar = 5 μm. **(c)** Genetic map indicating the location of *oriN* at the *aprE* locus in strain HM1108. **(d)** Analysis of DNA replication initiation at *oriC* and *oriN*. Marker frequency analysis was used to measure the rate of DNA replication initiation in the presence and absence of IPTG (0.1 mM). Genomic DNA was harvested from cells during the exponential growth phase and the relative amount of DNA from either the endogenous replication origin (*ori*C) or the *aprE* locus (*oriN*) compared to the terminus (*ter*) was determined using qPCR (mean and s.d. of 3 technical replicates). Cell doubling times (min) are shown above each data set.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



Extended Data Figure 3. Wild-type DnaA assembles into filaments on 5′-tailed substrates.
DnaA filament formation using amine-specific crosslinking (BS^3) on DNA scaffolds (represented by symbols above each lane). Protein complexes were resolved by SDS-PAGE and DnaA was detected by Western blot analysis.

	Strain	Wild-type sequence indicating deletions	Resulting sequence
Mononucleotide	Wild-type	3′-GATGATAATGAAGAT-5′	3′-GATGATAATGAAGAT-5′
	TR141	3′-GATGATAATGAAGAT-5′	3′-ATGATAATGAAGAT-5′
	TR114	3′-GATGATAATGAAGAT-5′	3′-GTGATAATGAAGAT-5′
	TR143	3′-GATGATAATGAAGAT-5′	3′-GAGATAATGAAGAT-5′
	TR147	3′-GATGATAATGAAGAT-5′	3′-GATATAATGAAGAT-5′
	TR137	3′-GATGATAATGAAGAT-5′	3′-GATGTAATGAAGAT-5′
	TR144	3′-GATGATAATGAAGAT-5′	3′-GATGAAATGAAGAT-5′
	TR145	3′-GATGATAATGAAGAT-5′	3′-GATGATATGAAGAT-5′
		3′-GATGATAATGAAGAT-5′	3′-GATGATATGAAGAT-5′
	TR153	3′-GATGATAATGAAGAT-5′	3′-GATGATAAGAAGAT-5′
Trinucleotide	Wild-type	3′-GATGATAATGAAGAT-5′	3′-GATGATAATGAAGAT-5′
	TR116	3′-GATGATAATGAAGAT-5′	3′-GATAATGAAGAT-5′
		3′-GATGATAATGAAGAT-5′	3′-GATAATGAAGAT-5′
		3′-GATGATAATGAAGAT-5′	3′-GATAATGAAGAT-5′
		3′-GATGATAATGAAGAT-5′	3′-GATAATGAAGAT-5′
	TR139	3′-GATGATAATGAAGAT-5′	3′-GATGATGAAGAT-5′
		3′-GATGATAATGAAGAT-5′	3′-GATGATGAAGAT-5′
		3′-GATGATAATGAAGAT-5′	3′-GATGATGAAGAT-5′

Extended Data Figure 4. DNA sequence of unwinding regions following mononucleotide and trinucleotide deletions.

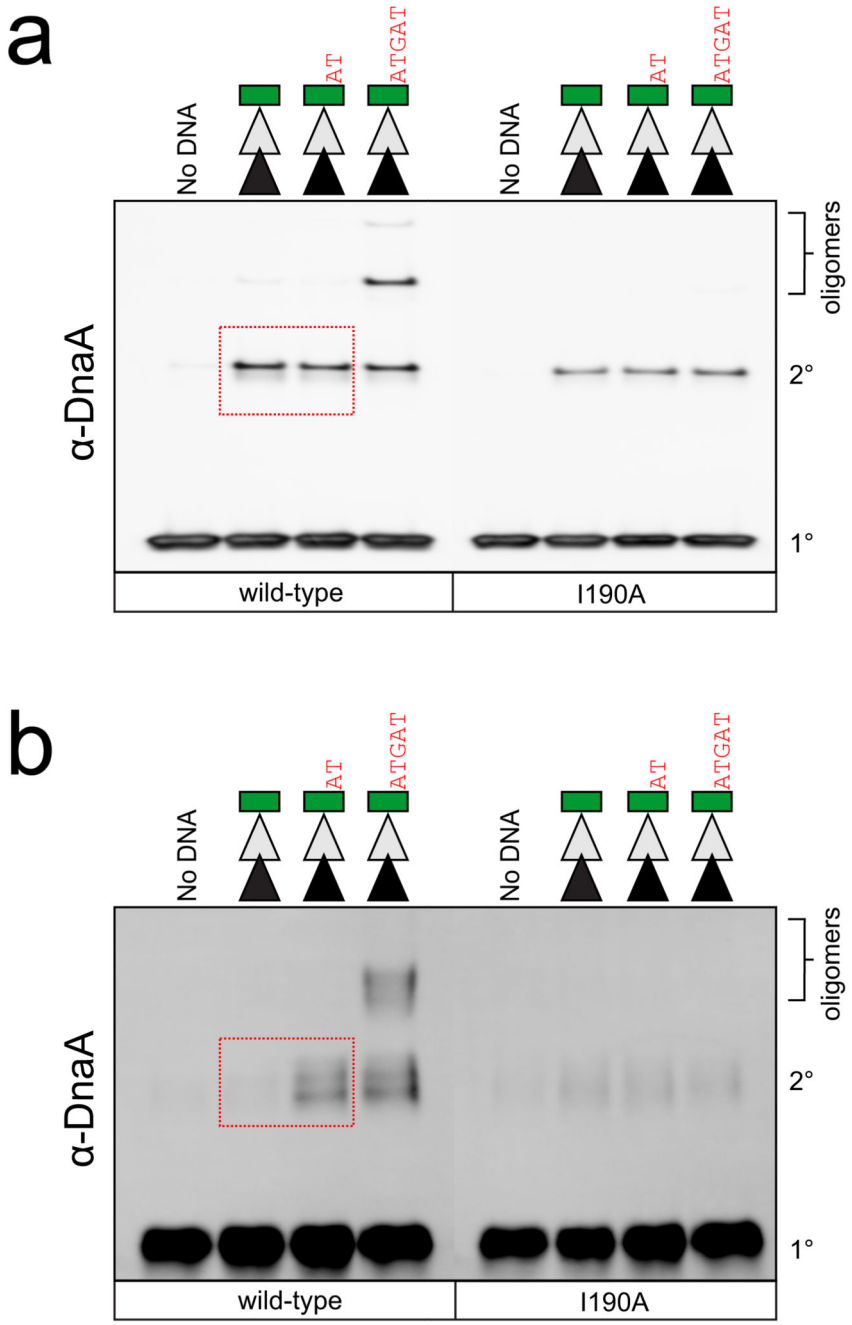
Resulting sequences grouped in boxes are identical for more than one deletion.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

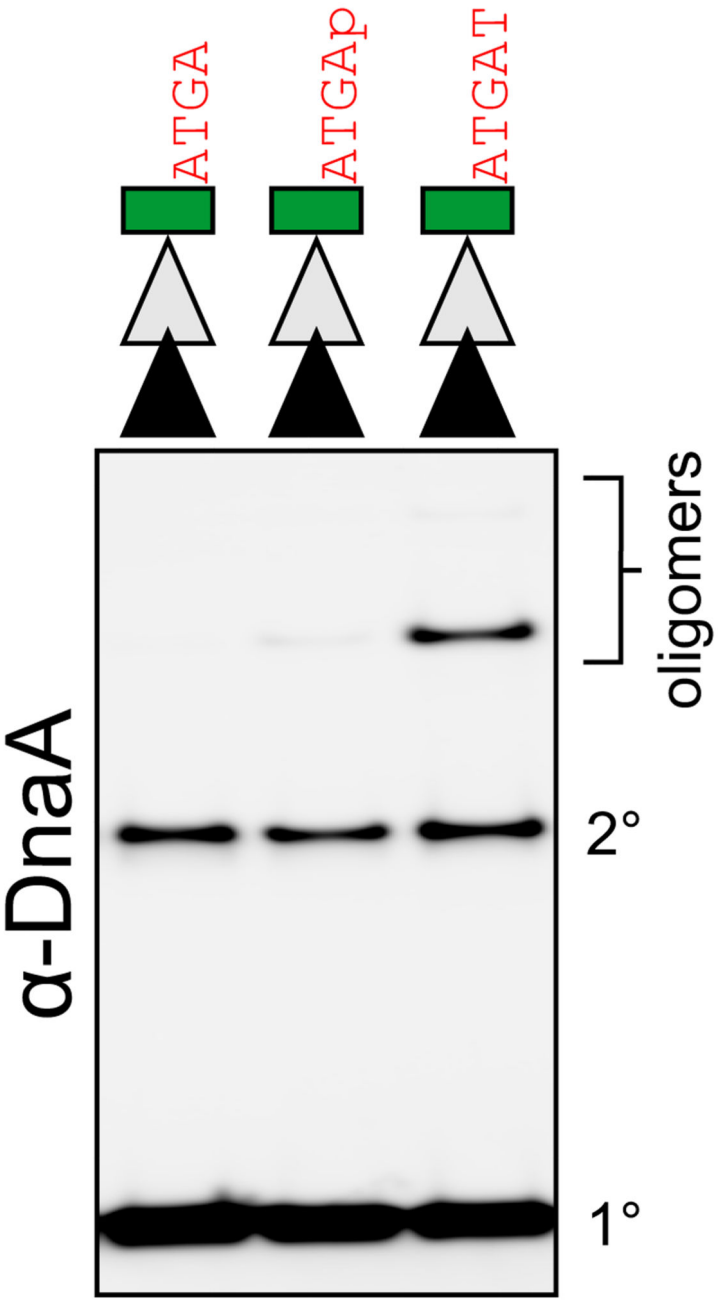


Extended Data Figure 5. Crosslinking with BS^3 captures a distinct DnaA oligomer.

DnaA was incubated with various DNA scaffolds and different crosslinking agents were added to capture distinct DnaA oligomers. **(a)** Crosslinking with BMOE detects DnaA oligomers forming on both duplex and tailed substrates. **(b)** Crosslinking with BS^3 only detects DnaA oligomers forming on tailed substrates, revealing an interaction between DnaA and the first DnaA-trio motif located downstream of the GC-cluster.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

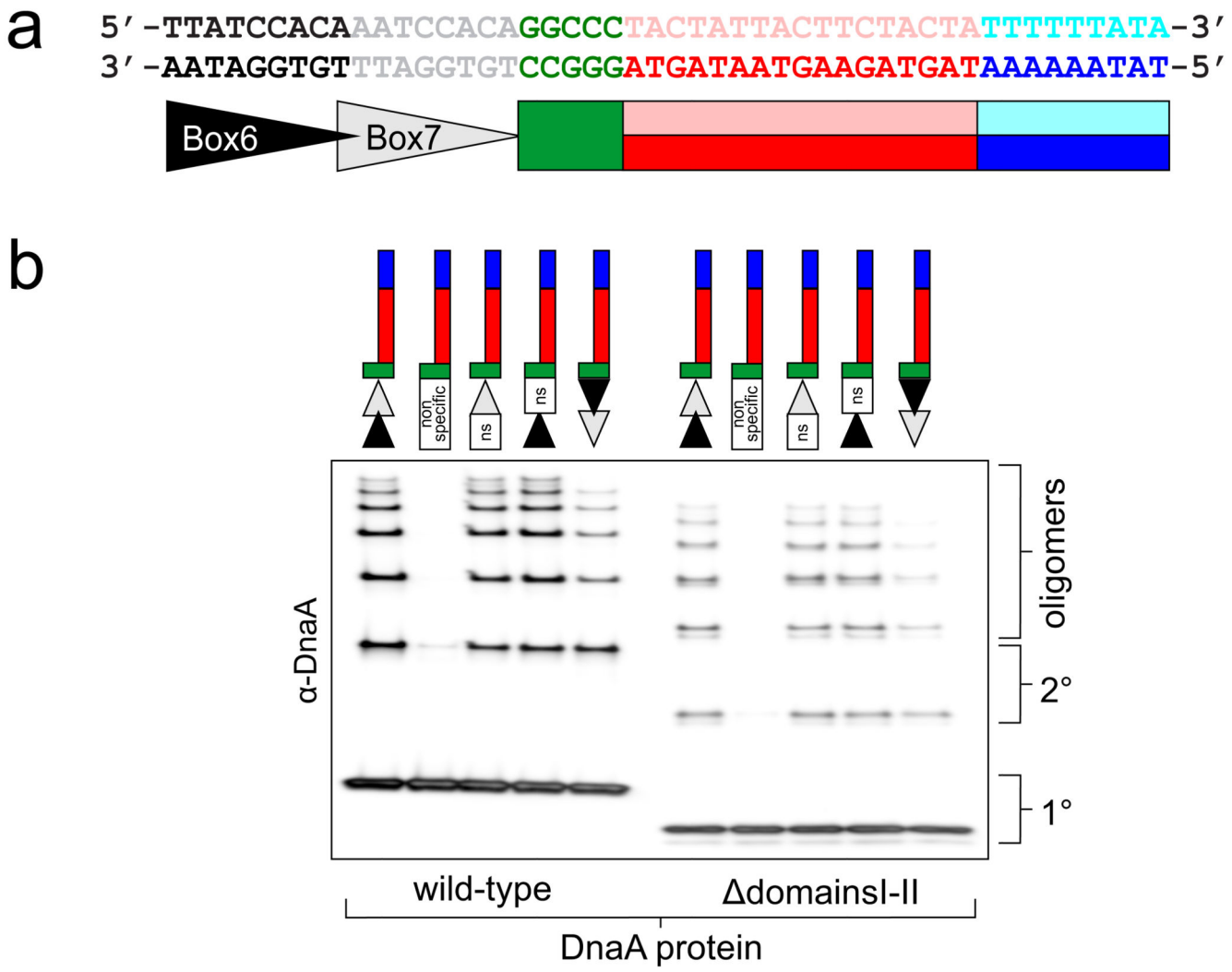


Extended Data Figure 6. The nucleotide at the third position of the DnaA-trio is required to stabilise DnaA.

DNA scaffolds containing the first two nucleotides of a DnaA-trio either with or without a 5'-phosphate are unable to stabilise binding of an additional DnaA protomer, indicating that the nucleotide at the third position is required. Combined with the data shown in Fig. 4b where the position is abasic, the results suggest that the sugar at the third position plays a critical role in DnaA binding.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



Extended Data Figure 7. Relationship between the DnaA-box and the DnaA-trios.

(a) Sequence of the origin region used for constructing DNA scaffolds. Symbols below represent duplex DnaA-boxes (triangles), the GC-rich region (green rectangles), the two strands of the unwinding region (red or pink rectangles), and the AT-rich region (blue rectangle). **(b)** Loading of the DnaA filament onto a single-stranded 5′-tail requires a DnaA-box and DnaA domains III-IV, but the DnaA-box position and orientation are flexible.

Extended Data Table 1

Bacterial replication origin regions in Figure 4c.

Organism	Genome accession #	Reference or Dori C accession # (ORI)	Genome position shown in Figure 4	Tandem DnaA -boxes (spacing)
Aquifex aeolicus VF5	NC_000918	17	166853 to 166897	N
Bacillus subtilis 168	NC_000964	34	1860 to 1902	Y (-1)

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Organism	Genome accession #	Reference or Dori C accession # (ORI)	Genome position shown in Figure 4	Tandem DnaA -boxes (spacing)
Bdellovibrio bacteriovorus HD100	NC_005363	ORI10040030	1569 to 1597	N
Bifidobacterium bifidum PRL2010	NC_014638	ORI94010761	2048 to 2071	Y (0)
Bordetella pertussi s Tohama I	NC_002929	ORI10030012	4084583 to 4084611	Y (+3)
Borrelia afzelii HLJ01	NC_018887	ORI96010684	460118 to 460149	Y? (+1)(+9)
Clostridium botulinum A str. Hall	NC_009698	ORI92010335	1517 to 1552	Y (-1)
Corynebacterium glutamicum ATCC 13032	NC_003450	ORI10010055	1984 to 2010	Y (0)
Enterococcus faecalis V583	NC_004668	ORI10010096	1498 to 1533	Y (0)
Escherichia col i MG1655	NC_000913	35	3925780 to 3925809	N
Helicobacter pylor i 26695	NC_000915	20	1607488 to 1607525	Y (0)
Leuconostoc citreum KM20	NC_010471	ORI92310382	1790 to 1819	Y (0)
Listeria monocytogenes EGD-e	NC_003210	ORI10010047	1773 to 1802	Y (-1)
Oceanobacillu s iheyensi s HTE831	NC_004193	ORI10010074	1746 to 1778	Y (-1)
Staphylococcus aureus NCTC 8325	NC_007795	ORI10010183	2075 to 2104	Y (-1)
Streptococcus pneumoniae R6	NC_003098	ORI10010044	1458 to 1487	Y (0)
Streptomyces coelicolor A3(2)	NC_003888	36	4270070 to 4270096	Y (0)
Synechococcus elongatus PCC 7942	NC_007604	37	2695870 to 2695899	Y (0)
Thermotoga maritim a MSB8	NC_000853	21	157010 to 157040	N
Treponema pallidum Nichols	NC_000919	ORI10010003	1568 to 1588	Y (-1)

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Extended Data Table 2

Oligonucleotides used for plasmid construction.

Product	Template	Primer #1	Sequence (5′ → 3′)	Primer #2	Sequence (5′ → 3′)
pHM453	Genomic DNA	oHM319	AATAATGACGTCGGCAAATTGTTTGAATTTGTC	oHM320	AATAATAGACGTCAGCCCGACACGCAGTTCATC
pHM492	Genomic DNA	oHM313	AATAATGAATTCTTAATTATCTAACCAATTATAAAACGGCAC	oHM315	AATAATCTCGAGCGCTTGGCAGCACCTGAGCAAACC
pHM560	Genomic DNA	oHM510	AATAATGATATCTATAATGGTACCTATATAAGGCCTAGATTGTGACAACCATTG	oHM511	AATATAAAGCTTAGAGGAAAGGTAGGATTAG
pTR54	pHM327	oTR36	CTTCTACCATTATCCGTTAGGAGGATAAAAATGAAATTCAC	oTR37	GGATAATGGTAGAAGTAATAGTAGGGCCTGTGGATTTG
pTR72	pGS43	oTR156	CAGCTTAAATGAGATCCGGCTGCTAACAAAGCCCGAAAGG	oTR157	CGGATCTCATTTAAGCTGTTCTTTAATTTCTTTTACATGC
pTR73	pTR72	oTR158	CATCATCATCATCATCACAGCGAAAATATATTAGACCTGTGGAACCAAGCCCTTGCTCAAATC	oTR159	GCTGTGATGATGATGATGATGCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTC
pTR74	pTR102	oTR189	ATTCAAGGTCGCATGGAAAATATATTAGACCTGTGGAACCAAGCCCTTG	oTR190	TTTCCATGCGACCTTGAATGCCGCTGCTGTGATGATGATGATGATGCATG
pTR83	pHM327	oTR82	CACAGTCTTCCTTGCTGTGGATAGGCTGTGTTTCCTGTCTTTTTC	oTR83	ATCCACAGCAAGGAAGACTGTGTATGACTTCCGAAAAGTTATTC
pTR84	pHM328	oTR66	GCCCATGATAATGAAGATGATATTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR67	AAAATATCATCTTCATTATCATGGGCCTGTGGATTTGTGGATAAGTTG
pTR85	pHM327	oTR118	GAAAGGCAAGGAAGCTTTTCGGAAGTCATACACAGTCTGTC	oTR119	AAAGCTTCCTTGCCTTTCCCCGATTGATCCCCGGTCCTG
pTR86	pHM327	oTR123	GAAGCTTCCTTGCGTCTGTCCACATGTGGATAGGCTGTGTTTCC	oTR124	AGACGCAAGGAAGCTTCCGAAAAGTTATTCACACTTTCCCCGATTG
pTR87	pHM327	oTR125	ACAGCAAGGAAGGCTGTGTTTCCTGTCTTTTTCACAACTTATC	oTR126	CAGCCTTCCTTGCTGTGGACAGACTGTGTATGACTTCCGAAAAG
pTR88	pHM327	oTR86	CCTGTCCTTCCTTGCACTTATCCACAAATCCACAGGCCCTACTATTAC	oTR87	GGATAAGTGCAAGGAAGGACAGGAAACACAGCCTATCCAC
pTR89	pHM327	oTR88	TCACAACCTTCCTTGCAATCCACAGGCCCTACTATTACTTCTAC	oTR89	GTGGATTGCAAGGAAGGTTGTGAAAAAGACAGGAAACACAGCCTATC
pTR90	pHM327	oTR090	CCACATTCCTTGCGGCCCTACTATTACTTCTAC	oTR091	GGGCCGCAAGGAATGTGGATAAGTTGTGAAAAAGACAGGAAAC
pTR102	pTR73	oTR187	CAGCAGCGGCATTCGAAAATATATTAGACCTGTGGAACCAAGCCCTTG	oTR188	TTTCGAATGCCGCTGCTGTGATGATGATGATGATGCATGGTATATC
pTR127	pHM327	oTR219	TTTGTGGATAAGGCCCTACTATTACTTCTAC	oTR220	CCTTATCCACAAATCCACAGTTGTGAAAAAGAC
pTR146	pHM327	oTR247	CAGGCCCATGTATTACTTCTACTATTTTTTATAAATATATATATTAATAC	oTR248	AATACATGGGCCTGTGGATTTGTGGATAAGTTGTG
pTR147	pHM327	oTR249	GCCCTACATATACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR250	AGTATATGTAGGGCCTGTGGATTTGTGGATAAGTTGTG
pTR148	pHM327	oTR251	CCCTACTATATGTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR252	GAACATATAGTAGGGCCTGTGGATTTGTGGATAAGTTGTG
pTR149	pHM327	oTR254	AGTACTTGTAATAGTAGGGCCTGTGGATTTGTGGATAAG	oTR263	CTACTATTACAAGTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAGGAG
pTR150	pHM327	oTR255	TACTTCATGTATTTTTTATAAATATATATATTAATACATTATCCGTTAGGAG	oTR256	AATACATGAAGTAATAGTAGGGCCTGTGGATTTGTG
pTR153	pHM327	oTR261	GATCAATCGGTATCCGTTAGGAGGATAAAAATGAAATTC	oTR262	CTAACGGATACCGATTGATCCCCGGTCCTGCTATTTAAG
pTR168	pTR74	oTR299	GAATTCGCCTGCTCTATCCGAGATAATAAATGC	oTR300	GAGCAGGCGAATTCGTTTGTAAATTTCTCAGAAG
pTR189	pHM327	oTR217	CAACTGTGGATAAAATCCACAGGCCCTACTATTACTTCTAC	oTR218	GATTTTATCCACAGTTGTGAAAAAGACAGGAAACACAG
pTR208	pTR74	oTR384	TCAAGGTCGCATGAACAAAACAGAACTTATCAATG	oTR385	CGGATCTTATTTTCCGGCAACTGCGTCTTTAAGC
pTR208	Genomic DNA	oTR386	TTTTGTTCATGCGACCTTGAATGCCGCTGCTG	oTR387	CGGAAAATAAGATCCGGCTGCTAACAAAGCCCGAAAG
pTR284	pHM327	oTR449	GGCCCAACTATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR450	GTAATAGTTGGGCCTGTGGATTTGTGGATAAG
pTR285	pHM327	oTR451	GGCCCTTCTATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR452	GTAATAGAAGGGCCTGTGGATTTGTGGATAAG
pTR286	pHM327	oTR453	GGCCCTAGTATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATCCG	oTR454	GTAATACTAGGGCCTGTGGATTTGTGGATAAG
pTR287	pHM327	oTR455	CCCTACAATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR456	AGTAATTGTAGGGCCTGTGGATTTGTGG

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Product	Template	Primer #1	Sequence (5′ → 3′)	Primer #2	Sequence (5′ → 3′)
pTR288	pHM327	oTR457	CCCTACTTTTACTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR458	GAAGTAAAAGTAGGGCCTGTGGATTTGTGGATAAG
pTR289	pHM327	oTR459	CCCTACTAATACTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR460	GTATTAGTAGGGCCTGTGGATTTGTGGATAAG
pTR301	pHM327	oTR469	CTTCTACTTTTTATAAATATATATATTAATACATTATCCGTTAGGAGGATAAAAATG	oTR470	TATAAAAAGTAGAAGTAATAGTAGGGCCTGTGG
pTR302	pHM327	oTR471	TTACTTCTTTTTATAAATATATATATTAATACATTATCCGTTAGGAGGATAAAAATG	oTR472	TATAAAAAGAAGTAATAGTAGGGCCTGTGGATTTG
pTR303	pHM327	oTR473	CTATTACTTTTTATAAATATATATATTAATACATTATCCGTTAGGAGGATAAAAATG	oTR474	TATAAAAAGTAATAGTAGGGCCTGTGGATTTG
pTR304	pHM327	oTR475	CAGGCCCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAGGAG	oTR476	AATAGTAGGGCCTGTGGATTTGTGGATAAG
pTR305	pHM327	oTR477	GCCCTACTTTTTATAAATATATATATTAATACATTATCCGTTAGGAGGATAAAAATG	oTR478	TTATAAAAAGTAGGGCCTGTGGATTTGTG
pTR306	pHM327	oTR479	GGCCCTTTTTATAAATATATATATTAATACATTATCCGTTAGGAGGATAAAAATG	oTR480	TATAAAAAGGGCCTGTGGATTTGTGG
pTR346	pHM327	oTR490	GGCCCACTATTACTTCTACTATTTTTTATAAATATATATATTAATAC	oTR491	GAAGTAATAGTGGGCCTGTGGATTTGTGGATAAGTTG
pTR348	pHM327	oTR494	GGCCCTATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR495	AGAAGTAATAGGGCCTGTGGATTTGTGGATAAGTTG
pTR349	pTR54	oTR496	ATTACTTCCATTATCCGTTAGGAGGATAAAAATGAAATTC	oTR497	GATAATGGAAGTAATAGTAGGGCCTGTGGATTTG
pTR350	pTR54	oTR498	ACTATTACCATTATCCGTTAGGAGGATAAAAATG	oTR499	GGATAATGGTAATAGTAGGGCCTGTGGATTTGTG
pTR377	pHM327	oTR539	GGCCCTATATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR540	GTAATATAGGGCCTGTGGATTTGTG
pTR441	pHM327	oTR659	CCCTACATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR660	GTAGAAGTAATGTAGGGCCTGTGGATTTGTG
pTR443	pHM327	oTR663	CCCTACTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR664	GTAGAAGTAGTAGGGCCTGTGGATTTGTG
pTR452	pHM326	oTR698	GGCCGTACTATTACTTCTACTATTTTTTATAAATATATATATTAATAC	oTR699	GTAATAGTACGGCCTGTGGATTTGTGG
pTR453	pHM327	oTR700	AGGCCTACTATTACTTCTACTATTTTTTATAAATATATATATTAATAC	oTR701	GTAATAGTAGGCCTGTGGATTTGTGG
pTR456	pHM327	oTR706	GCCCTCTATTACTTCTACTATTTTTTATAAATATATATATTAATACATTATC	oTR707	AAGTAATAGAGGGCCTGTGGATTTGTG
pTR457	pHM327	oTR708	CCCTACTTTACTTCTACTATTTTTTATAAATATATATATTAATACATTATCC	oTR709	GAAGTAAAGTAGGGCCTGTGGATTTG
pTR458	pHM327	oTR710	CCTACTATACTTCTACTATTTTTTATAAATATATATATTAATACATTATCC	oTR711	TAGAAGTATAGTAGGGCCTGTGGATTTG
pTR475	pHM327	oTR795	CTACTATAACTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR796	TAGAAGTTATAGTAGGGCCTGTGGATTTGTG
pTR476	pHM327	oTR797	TACTATTTCTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR798	GTAGAAGAAATAGTAGGGCCTGTGGATTTGTG
pTR477	pHM327	oTR799	ACTATTAGTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR800	AGTAGAACTAATAGTAGGGCCTGTGGATTTGTG
pTR478	pHM327	oTR801	TACTATTCTTCTACTATTTTTTATAAATATATATATTAATACATTATCCGTTAG	oTR802	GTAGAAGAATAGTAGGGCCTGTGGATTTGTG
pTR541	Genomic DNA	oTR535	ATTACGCCAGCTAGTGCTTTTATTTCTTGCAACCATAATAG	oTR536	ATTAATGCAGTTTTATCCTCCTAACGGATAATG
pTR541	pSG1301	oTR537	AAAGCACTAGCTGGCGTAATAGCGAAGAGG	oTR538	GAGGATAAAACTGCATTAATGAATCGGCCAACG
pTR542	pTR84	oTR535	ATTACGCCAGCTAGTGCTTTTATTTCTTGCAACCATAATAG	oTR536	ATTAATGCAGTTTTATCCTCCTAACGGATAATG
pTR542	pSG1301	oTR537	AAAGCACTAGCTGGCGTAATAGCGAAGAGG	oTR538	GAGGATAAAACTGCATTAATGAATCGGCCAACG

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Extended Data Table 3

Oligonucleotides used to assemble DNA scaffolds.

Oligo 1	Sequence (5′ → 3′)	O ligo 2	Sequence (5′ → 3′)	Figure
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR603	GGGCCTGTGGATTTGTGGATAAGT	2c, 2f, 4a, E5a, E5b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR609	ATATATATTTATAAAAAATAGTAGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	2c, E3, E7b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR619	ATATATATTTATAAAAATATCATCTTCATTATCATGGGCCTGTGGATTTGTGGATAAGT	2c
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR626	AGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR627	TAGGGCCTGTGGATTTGTGGATAAGT	4a, E5a, E5b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR628	GTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR629	AGTAGGGCCTGTGGATTTGTGGATAAGT	4a, E6
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR630	TAGTAGGGCCTGTGGATTTGTGGATAAGT	4a, 4b, E5a, E5b, E6
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR631	ATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR632	AATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR633	TAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a, 2f
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR634	GTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR645	AGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR646	AAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a, 3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR647	GAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR648	AGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a, 3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR649	TAGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	4a
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR754	TATAAAAAATAGTAGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	2f
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR831	TAGTAGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	2f
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR854	TAGTAGTAGTAGTAGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	2f
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR956	AAGAAGTAATATAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR957	AAGAAGTAATGTAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR958	AAGAAGTAAAGTAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR959	AAGAAGTATAGTAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR960	AAGAAGAATAGTAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR961	AAGAAGTAATAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR962	AAGAAGTAGTAGGGCCTGTGGATTTGTGGATAAGT	3d
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR965	pAGTAGGGCCTGTGGATTTGTGGATAAGT	E6
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR970	"AGTAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR971	T^{n}GTAGGGCCTGTGGATTTGTGGATAAGT	4b

Europe PMC Funders Author Manuscripts

Oligo 1	Sequence (5′ → 3′)	O ligo 2	Sequence (5′ → 3′)	Figure
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR972	TA^{n}TAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR995	TAATAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR996	TACTAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR998	TGGTAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR999	TTGTAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR1001	CAGTAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR602	ACTTATCCACAAATCCACAGGCCC	oTR1002	AAGTAGGGCCTGTGGATTTGTGGATAAGT	4b
oTR603	GGGCCTGTGGATTTGTGGATAAGT	oTR608	ACTTATCCACAAATCCACAGGCCCTACTATTACTTCTACTATTTTTTATAAATATATAT	2c, E3
oTR608	ACTTATCCACAAATCCACAGGCCCTACTATTACTTCTACTATTTTTTATAAATATATAT	oTR609	ATATATATTTATAAAAAATAGTAGAAGTAATAGTAGGGCCTGTGGATTTGTGGATAAGT	2c
oTR611	ACCTTCCTTGCTTCCTTGCGGCCC	oTR612	ATATATATTTATAAAAAATAGTAGAAGTAATAGTAGGGCCGCAAGGAAGCAAGGAAGGT	2c, E7b
oTR665	ATATATATTTATAAAAAATAGTAGAAGTAATAGTAGGGCCTGTGGATTGCAAGGAAGGT	oTR667	ACCTTCCTTGCAATCCACAGGCCC	E7b
oTR666	ATATATATTTATAAAAAATAGTAGAAGTAATAGTAGGGCCGCAAGGAATGTGGATAAGT	oTR668	ACTTATCCACATTCCTTGCGGCCC	E7b
oTR946	ACTGTGGATTTGTGGATAAGGCCC	oTR623	ATATATATTTATAAAAAATAGTAGAAGTAATAGTAGGGCCTTATCCACAAATCCACAGT	E7b

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank J. Errington and W. Vollmer for critical review of the manuscript. We thank G. Scholefield for preliminary data, A. Koh for research assistance, and I. Selmes for technical assistance. Research support was provided to H.M. by a Royal Society University Research Fellowship and a Biotechnology and Biological Sciences Research Council (BBSRC) Research Grant (BB/K017527/1), and to O.H. by an Iraqi Ministry of Higher Education and Scientific Research Studentship.

References

1. Duderstadt KE, Berger JM. A structural framework for replication origin opening by AAA+ initiation factors. Curr Opin Struct Biol. 2013; 23:144–153. [PubMed: 23266000]
2. Messer W. The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication. FEMS Microbiol Rev. 2002; 26:355–374. [PubMed: 12413665]
3. Fuller RS, Funnell BE, Kornberg A. The dnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites. Cell. 1984; 38:889–900. [PubMed: 6091903]
4. Duderstadt KE, Chuang K, Berger JM. DNA stretching by bacterial initiators promotes replication origin opening. Nature. 2011; 478:209–213. [PubMed: 21964332]
5. Fujikawa N, et al. Structural basis of replication origin recognition by the DnaA protein. Nucleic Acids Res. 2003; 31:2077–2086. [PubMed: 12682358]
6. Mackiewicz P, Zakrzewska-Czerwinska J, Zawilak A, Dudek MR, Cebrat S. Where does bacterial replication start? Rules for predicting the *oriC* region. Nucleic Acids Res. 2004; 32:3781–3791. [PubMed: 15258248]
7. Wolanski M, Donczew R, Zawilak-Pawlik A, Zakrzewska-Czerwinska J. *oriC*-encoded instructions for the initiation of bacterial chromosome replication. Front Microbiol. 2014; 5:735. [PubMed: 25610430]
8. Hassan AK, et al. Suppression of initiation defects of chromosome replication in *Bacillus subtilis dnaA* and *oriC*-deleted mutants by integration of a plasmid replicon into the chromosomes. J Bacteriol. 1997; 179:2494–2502. [PubMed: 9098044]
9. Krause M, Ruckert B, Lurz R, Messer W. Complexes at the replication origin of *Bacillus subtilis* with homologous and heterologous DnaA protein. J Mol Biol. 1997; 274:365–380. [PubMed: 9405146]
10. Leonard AC, Grimwade JE. Regulation of DnaA Assembly and Activity: Taking Directions from the Genome. Annu Rev Microbiol. 2011; 65:19–35. [PubMed: 21639790]
11. Speck C, Messer W. Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. EMBO J. 2001; 20:1469–1476. [PubMed: 11250912]
12. Scholefield G, Errington J, Murray H. Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. EMBO J. 2012; 31:1542–1555. [PubMed: 22286949]
13. Cheng HM, Groger P, Hartmann A, Schlierf M. Bacterial initiators form dynamic filaments on single-stranded DNA monomer by monomer. Nucleic Acids Res. 2015; 43:396–405. [PubMed: 25477384]
14. Ozaki S, et al. A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. J Biol Chem. 2008; 283:8351–8362. [PubMed: 18216012]
15. Gaudier M, Schuwirth BS, Westcott SL, Wigley DB. Structural basis of DNA replication origin recognition by an ORC protein. Science. 2007; 317:1213–1216. [PubMed: 17761880]
16. Dueber EL, Corn JE, Bell SD, Berger JM. Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. Science. 2007; 317:1210–1213. [PubMed: 17761879]
17. Erzberger JP, Mott ML, Berger JM. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. Nat Struct Mol Biol. 2006; 13:676–683. [PubMed: 16829961]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

18. Duderstadt KE, et al. Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator. J Biol Chem. 2010; 285:28229–28239. [PubMed: 20595381]
19. Krause M, Messer W. DnaA proteins of *Escherichia coli* and *Bacillus subtilis*: coordinate actions with single-stranded DNA-binding protein and interspecies inhibition during open complex formation at the replication origins. Gene. 1999; 228:123–132. [PubMed: 10072765]
20. Donczew R, Weigel C, Lurz R, Zakrzewska-Czerwinska J, Zawilak-Pawlik A. *Helicobacter pylori oriC*--the first bipartite origin of chromosome replication in Gram-negative bacteria. Nucleic Acids Res. 2012; 40:9647–9660. [PubMed: 22904070]
21. Ozaki S, Fujimitsu K, Kurumizaka H, Katayama T. The DnaA homolog of the hyperthermophilic eubacterium *Thermotoga maritima* forms an open complex with a minimal 149-bp origin region in an ATP-dependent manner. Genes Cells. 2006; 11:425–438. [PubMed: 16611245]
22. Gao F, Luo H, Zhang CT. DoriC 5.0: an updated database of *oriC* regions in both bacterial and archaeal genomes. Nucleic Acids Res. 2013; 41:D90–93. [PubMed: 23093601]
23. Kumar S, Farhana A, Hasnain SE. *In vitro* helix opening of *M. tuberculosis oriC* by DnaA occurs at precise location and is inhibited by IciA like protein. PLoS One. 2009; 4:e4139. [PubMed: 19127296]
24. Pei H, et al. Mechanism for the TtDnaA-Tt-*oriC* cooperative interaction at high temperature and duplex opening at an unusual AT-rich region in *Thermoanaerobacter tengcongensis*. Nucleic Acids Res. 2007; 35:3087–3099. [PubMed: 17452366]
25. Kaur G, et al. Building the bacterial orisome: high-affinity DnaA recognition plays a role in setting the conformation of *oriC* DNA. Mol Microbiol. 2014; 91:1148–1163. [PubMed: 24443848]
26. Noguchi Y, Sakiyama Y, Kawakami H, Katayama T. The Arg fingers of key DnaA protomers are oriented inward within the replication origin *oriC* and stimulate DnaA subcomplexes in the initiation complex. J Biol Chem. 2015; 290:20295–20312. [PubMed: 26126826]
27. Bleichert F, Botchan MR, Berger JM. Crystal structure of the eukaryotic origin recognition complex. Nature. 2015; 519:321–326. [PubMed: 25762138]
28. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004; 14:1188–1190. [PubMed: 15173120]
29. Taylor RG, Walker DC, McInnes RR. *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. Nucleic Acids Res. 1993; 21:1677–1678. [PubMed: 8479929]
30. Morimoto T, et al. Six GTP-binding proteins of the Era/Obg family are essential for cell growth in *Bacillus subtilis* . Microbiology. 2002; 148:3539–3552. [PubMed: 12427945]
31. Vagner V, Dervyn E, Ehrlich SD. A vector for systematic gene inactivation in *Bacillus subtilis* . Microbiology. 1998; 144:3097–3104. [PubMed: 9846745]
32. Berkmen MB, Grossman AD. Subcellular positioning of the origin region of the *Bacillus subtilis* chromosome is independent of sequences within *oriC*, the site of replication initiation, and the replication initiator DnaA. Mol Microbiol. 2007; 63:150–165. [PubMed: 17140409]
33. Hansen EB, Yarmolinsky MB. Host participation in plasmid maintenance: dependence upon dnaA of replicons derived from P1 and F. Proc Natl Acad Sci U S A. 1986; 83:4423–4427. [PubMed: 3520571]
34. Moriya S, Atlung T, Hansen FG, Yoshikawa H, Ogasawara N. Cloning of an autonomously replicating sequence (*ars*) from the *Bacillus subtilis* chromosome. Mol Microbiol. 1992; 6:309–315. [PubMed: 1552845]
35. Oka A, Sugimoto K, Takanami M, Hirota Y. Replication origin of the *Escherichia coli* K-12 chromosome: the size and structure of the minimum DNA segment carrying the information for autonomous replication. Mol Gen Genet. 1980; 178:9–20. [PubMed: 6991883]
36. Calcutt MJ, Schmidt FJ. Conserved gene arrangement in the origin region of the *Streptomyces coelicolor* chromosome. J Bacteriol. 1992; 174:3220–3226. [PubMed: 1577691]
37. Watanabe S, et al. Light-dependent and asynchronous replication of cyanobacterial multi-copy chromosomes. Mol Microbiol. 2012; 83:856–865. [PubMed: 22403820]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

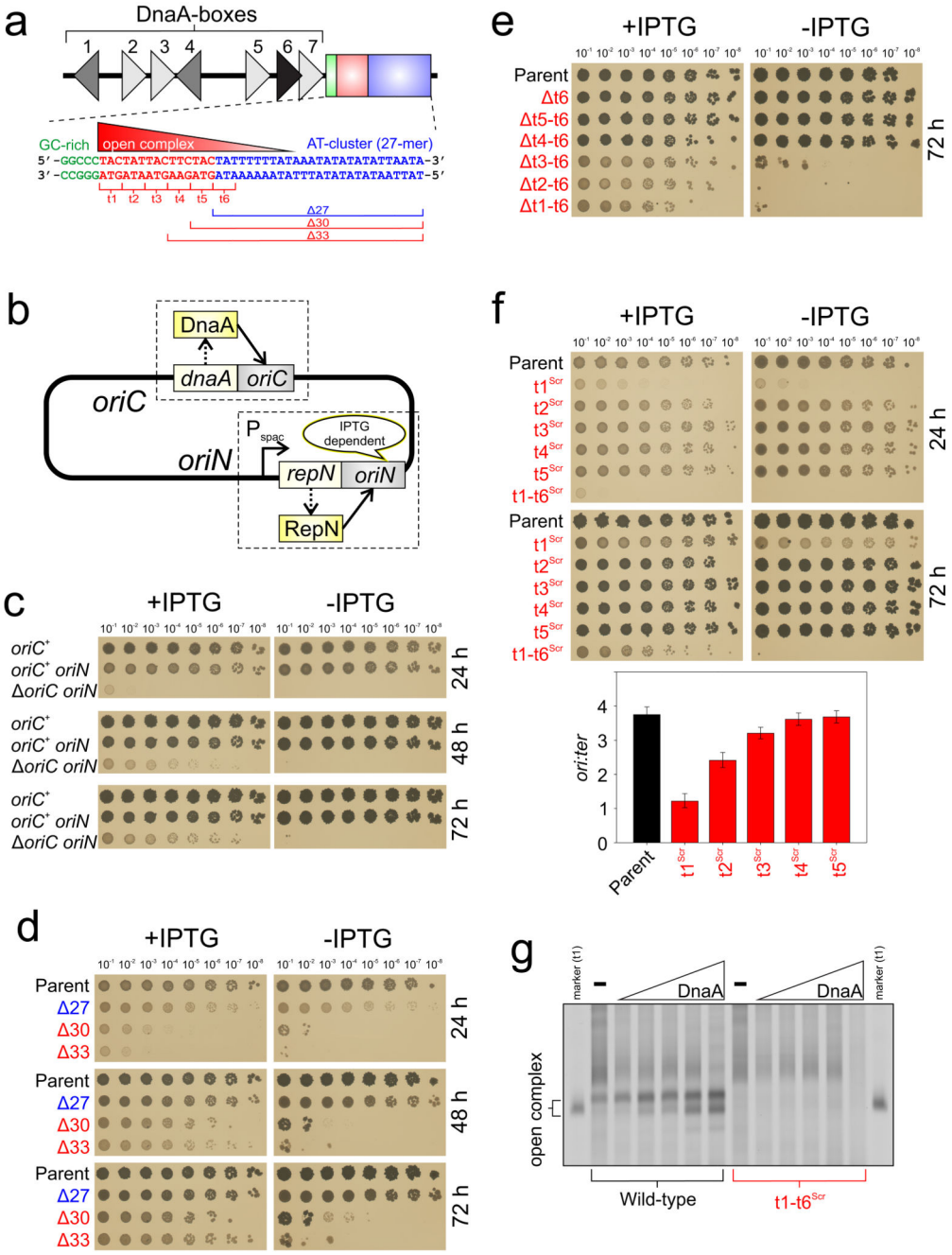


Figure 1. Genetic analysis of the *oriC* DNA unwinding element reveals a critical region required for initiation activity.

(a) *B. subtilis oriC* unwinding region. DnaA-box colouring indicates conservation (consensus 5′-TTATCCACA-3′ in black). **(b)** The *oriC*-independent strain used for constructing replication origin mutations. **(c)** Growth of an *oriC* deletion mutant is dependent upon *oriN* activity. **(d)** Deletions extending beyond the AT-cluster into the initially unwound region inhibit cell growth. **(e)** Sequences between the GC-rich and AT-rich clusters are essential for origin function. **(f)** Sequences proximal to DnaA-boxes are most

important for origin function. Marker frequency analysis was used to measure the rate of DNA replication initiation (mean and s.d. of 3 technical replicates). **(g)** Open complex formation by DnaA requires the native sequence between the GC- and AT-clusters. DNA duplex unwinding was probed by $KMnO_4$ and detected by primer extension.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

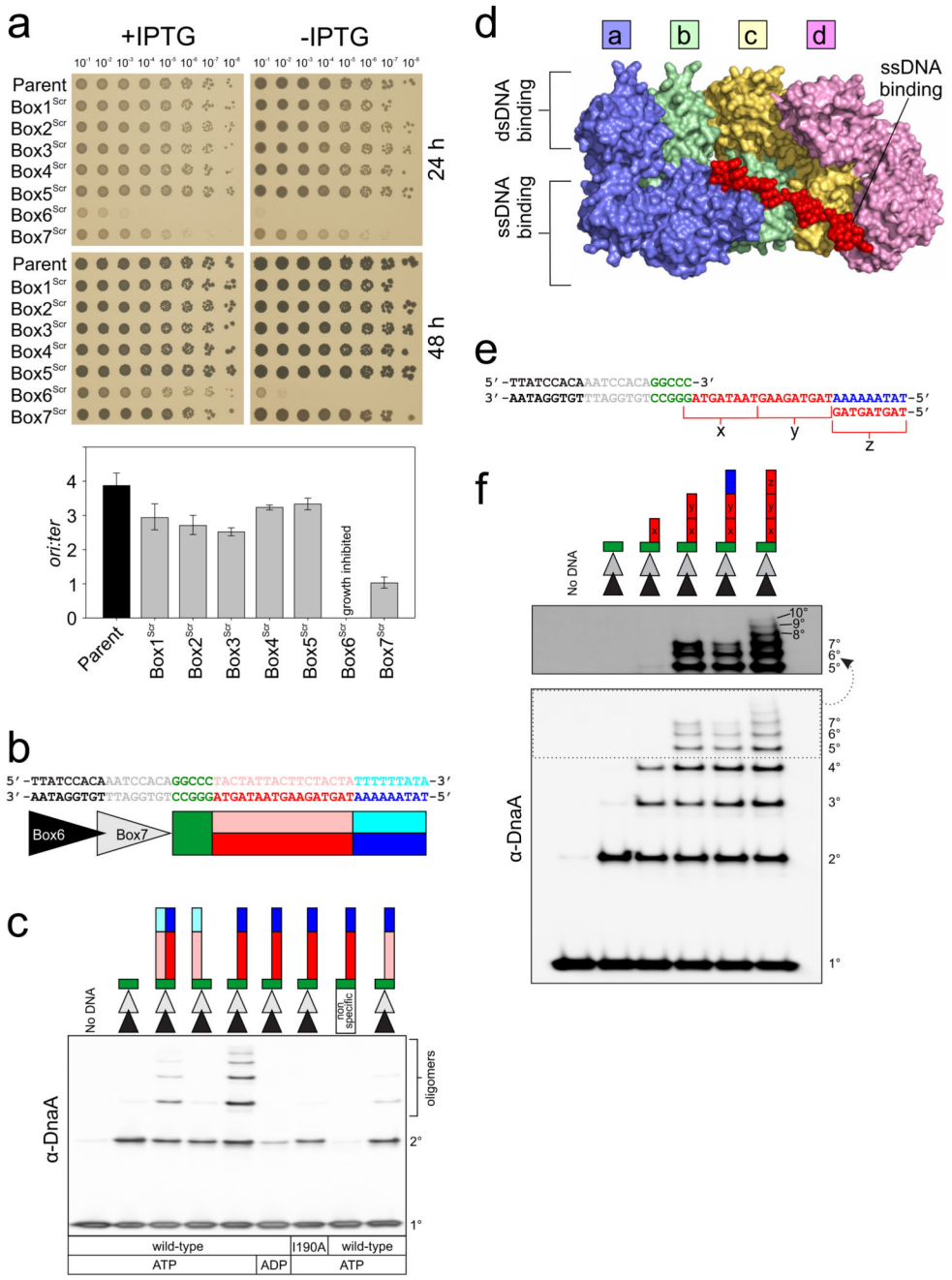


Figure 2. DnaA filaments are loaded from DnaA-boxes onto a specific single-strand sequence within the initially unwound region.

(a) DnaA-boxes proximal to the unwinding region are most important for origin function. Marker frequency analysis was used to measure the rate of DNA replication initiation (mean and s.d. of 3 technical replicates). **(b)** Sequence of the origin region used for constructing DNA scaffolds in (c). **(c)** DnaA filament formation using cysteine-specific crosslinking on DNA scaffolds. DnaA complexes were resolved by SDS-PAGE and detected by Western blot analysis. **(d)** Crystal structure showing ssDNA (dA_{12}) bound to the DnaA filament through

the AAA+ domain (PDB: 3R8F). **(e)** Sequence of the origin region used for constructing DNA scaffolds in (f). **(f)** DnaA filament formation on tailed substrates is arrested by a poly(A) tract. Long oligomers highlighted within the dotted box are shown above with increased contrast.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

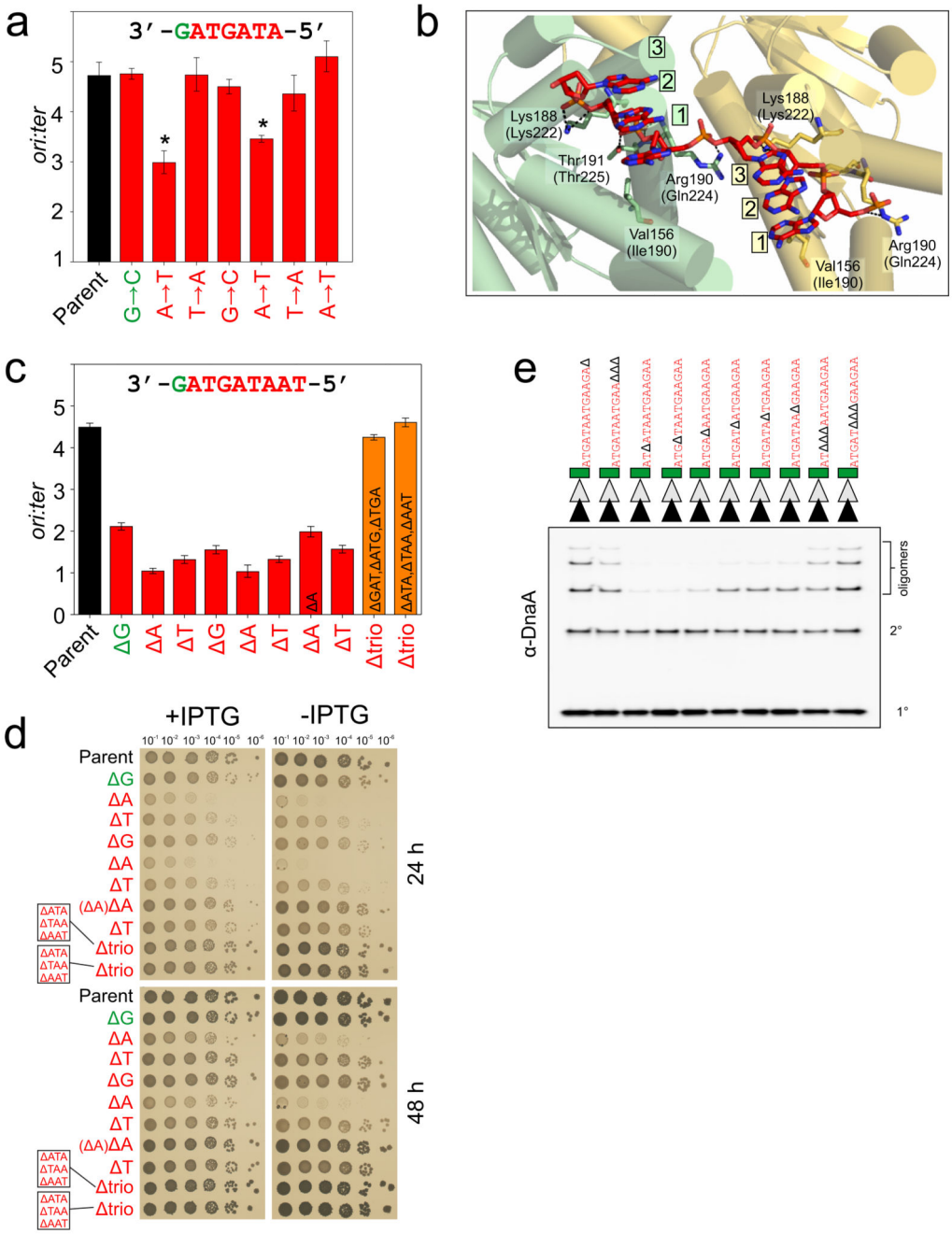


Figure 3. Analysis of the key origin unwinding region provides evidence for functional 3-mer repeats.

(a) Mutagenesis identifies A:T base pairs spaced three nucleotides apart are most critical for origin activity. Marker frequency analysis was used to measure the rate of DNA replication initiation (mean and s.d. of 3 technical replicates). **(b)** Crystal structure showing the interaction of DnaA with sets of three nucleotides (PDB: 3R8F). Residues for *A. aeolicus* indicated above; *B. subtilis* below. **(c)** *In vivo* deletion analysis of the unwinding region. Isogenic deletions indicated in black. Marker frequency analysis was used to measure the

rate of DNA replication initiation (mean and s.d. of 3 technical replicates). **(d)** Growth of mutants used in (**c**). **(e)** *In vitro* deletion analysis of tailed substrates.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

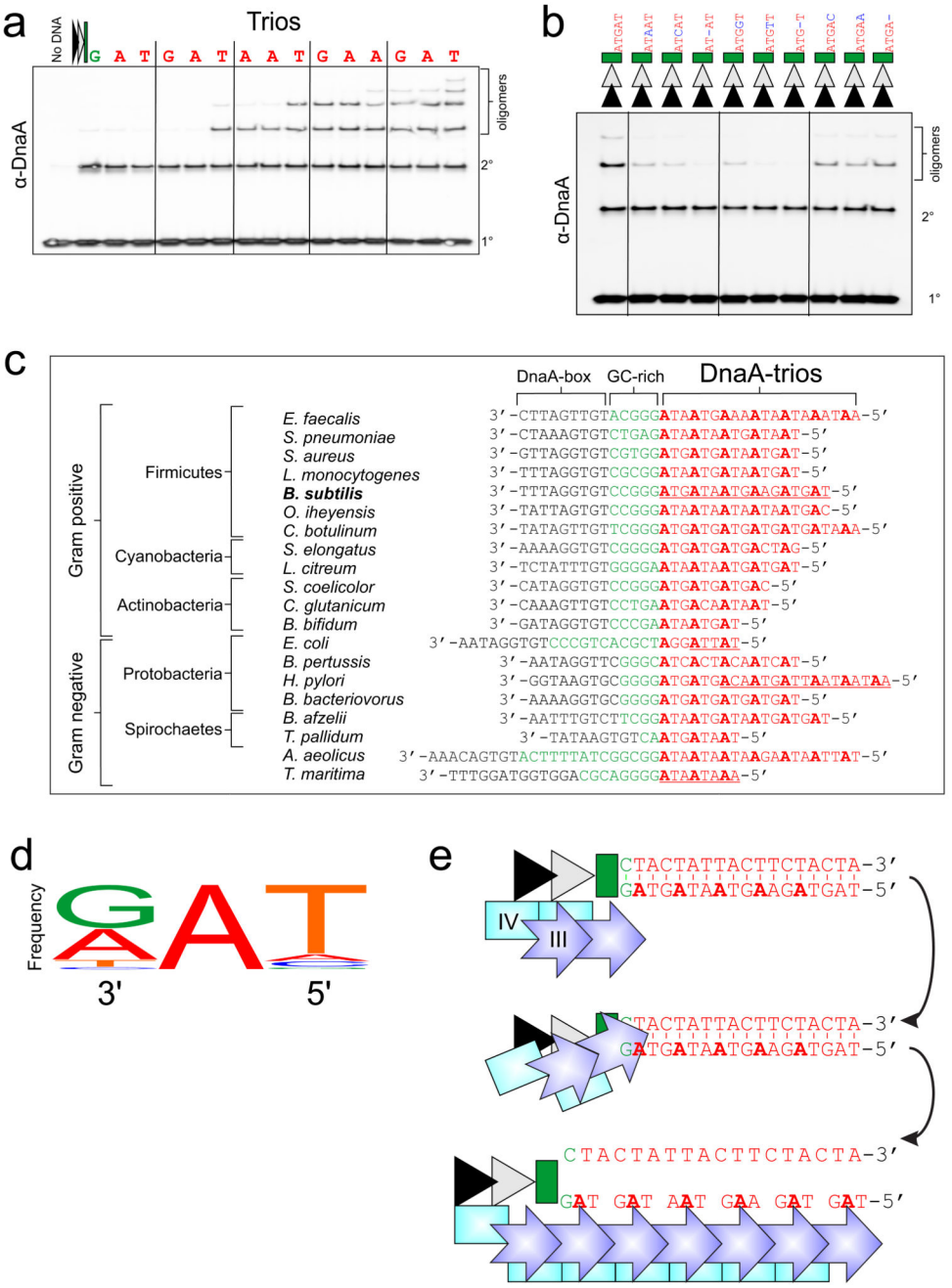


Figure 4. Identification of the DnaA-trio motif.

(a) Varying the length of 5′-tailed substrates identifies the likely DnaA-trio sequence. Lane 2 shows DnaA filament formation on a duplex DNA scaffold (DnaA-box6, DnaA-box7, GC-rich cluster). Letters indicate the nucleotide sequentially added to the 5′-tail. **(b)** Targeted mutagenesis of the proposed DnaA-trio motif. **(c)** Bioinformatic analysis identifies DnaA-trio motifs adjacent to a DnaA-box throughout the bacterial kingdom. Underlined sequences indicate experimentally determined DnaA-dependent unwinding sites. **(d)** DnaA-trios

sequence logo (WebLogo28). **(e)** A schematic of DnaA filament formation from double-stranded DnaA-boxes (triangles) onto a single strand containing the DnaA-trios.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts