

Analyses for mechanisms in formation of the
replication initiation complex and for
functional characteristics of the DNA unwinding
region at the replication origin of
hyperthermophilic bacterium *Thermotoga maritima*

盧, 楚元

<https://hdl.handle.net/2324/7157321>

出版情報 : Kyushu University, 2023, 博士 (創薬科学) , 課程博士
バージョン :
権利関係 :



**Analyses for mechanisms in formation of the replication
initiation complex and for functional characteristics of the
DNA unwinding region at the replication origin of
hyperthermophilic bacterium *Thermotoga maritima***

高度好熱性真正細菌 *Thermotoga maritima* の複製起点に
おける開始複合体の形成メカニズムと開裂部位の特性の解析

Chuyuan Lu

盧 楚元

Department of Molecular Biology

Graduate School of Pharmaceutical Sciences, Kyushu University

九州大学大学院 薬学府 創薬科学専攻 分子生物薬学分野

2023

Table of contents

Abbreviations	3
Abstract	4
Introduction	5
Results	14
1. Identification of a novel motif within <i>tma-oriC</i> required for DUE unwinding	14
2. ATP- <i>tmaDnaA</i> oligomers on DOR bind ssDUE through eTrio	20
3. Individual <i>tmaDnaA</i> boxes within minimal <i>tmaDOR</i> are crucial for unwinding	26
4. Intact DNA helical turn between DUE and DOR is crucial for unwinding	29
5. Involvement of the Arg-finger in open complex formation	32
6. Inverted <i>tmaDnaA</i> boxes permit ATP- <i>tmaDnaA</i> interactions in a head-to-tail manner	38
Discussion	43
Experimental procedures	50
References	59
Publication list	65
Acknowledgement	66

Abbreviations

BSA: bovine serum albumin

AAA+: ATPases associated with diverse cellular activities

DUE: duplex unwinding element

DOR: DnaA oligomerization region

IHF: Integration host factor

IBR: IHF-binding region

ss: single-stranded

ds: double-stranded

EMSA: electrophoretic mobility shift assay

tma: *Thermotoga maritima*

tmaDnaA: *Thermotoga maritima* DnaA homolog

tma-oriC: The origin region from *Thermotoga maritima*

tmaDUE: DUE of *Thermotoga maritima*

tmaDOR: DOR of *Thermotoga maritima*

WT: wild type

Abstract

Initiation of chromosomal replication requires dynamic nucleoprotein complexes. In most eubacteria, the origin *oriC* contains multiple DnaA box sequences to which the ubiquitous DnaA initiators bind. In *Escherichia coli oriC*, DnaA boxes sustain construction of higher-order complexes via DnaA-DnaA interactions, promoting the unwinding of the DNA unwinding element (DUE) within *oriC* and concomitantly binding the single-stranded DUE to install replication machinery. Despite the significant sequence homologies among DnaA proteins, bacterial *oriC* sequences are highly diverse. The present study investigated the design of *oriC* (*tma-oriC*) from *Thermotoga maritima*, an evolutionarily ancient eubacterium. The minimal *tma-oriC* sequence includes a DUE and a flanking region containing five DnaA boxes recognized by the cognate DnaA initiator (*tmaDnaA*). This DUE was comprised of two distinct functional modules, an unwinding module and a *tmaDnaA*-binding module. Three direct repeats of the trinucleotide TAG within DUE were essential for both unwinding and single-stranded DUE binding by *tmaDnaA* complexes constructed on the DnaA boxes. Its surrounding AT-rich sequences stimulated only duplex unwinding. Moreover, head-to-tail oligomers of ATP-bound *tmaDnaA* were constructed within *tma-oriC*, irrespective of the directions of the DnaA boxes. This binding mode was considered to be induced by flexible swiveling of DnaA domains III and IV, which were responsible for DnaA-DnaA interactions and DnaA box binding, respectively. Phasing of specific *tmaDnaA* boxes in *tma-oriC* DNA was also responsible for unwinding. These findings indicate that a single-stranded DUE recruitment mechanism was responsible for unwinding, and would enhance understanding of the fundamental molecular nature of the origin sequences present in evolutionarily divergent bacteria.

Introduction

Unwinding of duplex DNA is fundamental for chromosomal DNA replication in all cellular organisms. The origin of replication *oriC* encodes instructions to form a highly ordered nucleoprotein complex called the initiation complex. In bacteria, this complex is mainly comprised of the ubiquitous family of DnaA initiator proteins (Katayama *et al*, 2017; Grimwade & Leonard, 2021; Wolański *et al*, 2015; Costa *et al*, 2013; Kaguni, 2011; Ozaki, 2019). The minimal *oriC* in *Escherichia coli* consists of an AT-rich DNA unwinding element (DUE) and a flanking DnaA oligomerization region (DOR) containing a cluster of DnaA-binding sequences (DnaA boxes) (**Figure 0-1A**). The ATP-bound DnaA oligomers on the DOR are important for DUE unwinding and loading of DnaB replicative helicase onto the unwound region (Kaguni, 2011; Sakiyama *et al*, 2022; Shimizu *et al*, 2016; Ozaki & Katayama, 2012; Felczak & Kaguni, 2004; Hayashi *et al*, 2020).

Proteins in the DnaA family contain a central domain III with AAA+ (ATPases associated with various cellular activities) motifs (Grimwade & Leonard, 2021; Costa *et al*, 2013; Kaguni, 2011; Ozaki & Katayama, 2009) (**Figure 0-1B**). These motifs play essential roles in ATP/ADP binding, ATP hydrolysis and DnaA-DnaA interactions (Ozaki & Katayama, 2009; Erzberger *et al*, 2006; Felczak & Kaguni, 2004). Similar to other proteins containing AAA+ motifs, head-to-tail oligomerization via this domain underlies the formation of an initiation complex by ATP-DnaA. In *E. coli* DnaA, the AAA+ arginine-finger motif Arg285 predominantly recognizes ATP bound to the adjacent DnaA protomer, promoting co-operative ATP-DnaA binding onto the DOR in a head-to-tail orientation (Kawakami *et al*, 2005; Noguchi *et al*, 2015; Ozaki *et al*, 2012) (**Figure 0-1B**). Moreover, H/B-motifs (hydrophobic Val211 and basic Arg245) in this domain bind single-stranded DUEs (ssDUEs) in a sequence-specific manner (Ozaki & Katayama, 2012; Ozaki *et al*, 2008) (**Figure 0-1B**). The AAA+ domain III of DnaA is linked via a flexible hinge to its C-terminal domain IV, which is responsible for DnaA box-specific DNA binding (Fujikawa, 2003)(**Figure 0-1B**). The N-terminal domain I has multiple sites for interactions with DnaB helicase etc. and for weak domain I-domain I

interactions (Sutton *et al*, 1998; Simmons *et al*, 2003; Abe *et al*, 2007; Hayashi *et al*, 2020; Kawakami *et al*, 2005; Keyamura *et al*, 2009; Felczak *et al*, 2005). Domain II is a flexible linker between domains I and III (Abe *et al*, 2007; Nozaki & Ogawa, 2008) (**Figure 0-1B**).

The arrangement of DnaA boxes within the DOR provides an essential scaffold for formation of the initiation complex (**Figure 0-1A and 0-2**). A canonical DnaA box consists of an asymmetric 9-mer consensus sequence, TTA[T/A]NCACA (Wolański *et al*, 2015; Schaper & Messer, 1995). The DOR in *E. coli* contains no fewer than 12 DnaA boxes as well as a region for specific binding to integration host factor (IHF), a nucleoid-associated protein that introduces a sharp bend in DNA (IBR, IHF-binding region) (**Figure 0-1**) (Katayama *et al*, 2017; Grimwade & Leonard, 2021). Five of these DnaA boxes in the left DOR (R1, R5M, τ 2, and I1-2) share the same orientation, whereas five boxes in the right DOR (R4, C1-3, and I3) share the opposite orientation (Noguchi *et al*, 2015; Shimizu *et al*, 2016; Sakiyama *et al*, 2017; Rozgaja *et al*, 2011; Kawakami *et al*, 2005; McGarry *et al*, 2004). The directional arrangement of the DnaA boxes facilitates head-to-tail oligomerization of ATP-DnaA proteins depending on the ATP-Arg finger interaction. This leads to formation of a pair of pentamers bound to the left and right DORs, with the two facing each other. In the left DOR, the IHF-dependent bending facilitates formation of the DnaA complexes, promoting DUE unwinding activity (Sakiyama *et al*, 2017) (**Figure 0-1B**). DnaA complexes in the right DOR are important for enhancement of the unwound state and efficient DnaB loading (Sakiyama *et al*, 2022; Ozaki & Katayama, 2012).

Experimental determination of the *oriC*s of representative bacterial species has shown that the basic structure of *Vibrio cholerae oriC1*, the origin of chromosome I, is most similar to that of *E. coli oriC*, in that the DUE is flanked by a region containing two DnaA box clusters in the opposite directions (Wolański *et al*, 2015) (**Figure 0-2**). The *oriC* of *Helicobacter pylori* is split into two subsequences, *oriC1* and *oriC2*, by insertion of the *dnaA* gene; the *oriC2* of *H. pylori* may correspond to the region of the *E. coli oriC* spanning DUE to the left DOR, a region containing a single cluster of unidirectional DnaA boxes (Wolański *et al*, 2015) (**Figure 0-2**). However, the *oriC* sequences from

other species vary in the directions of DnaA boxes, despite containing direct repeats of DnaA box sequences (Kumar *et al*, 2009; Pei *et al*, 2007; Jaworski *et al*, 2018). Similar features have also been observed in the bioinformatically predicted *oriC*s of bacterial genomes (Dong *et al*, 2023) (**Figure 0-3**). Thus, the principles underlying the designs of DnaA box arrangements within bacterial DORs that underlie the construction of DnaA oligomers remain unclear.

In *E. coli*, the ATP-DnaA—IHF—Left-DOR complexes are responsible for the unwinding of DUE. The upper strand of the resulting ssDUE is subsequently recruited to the ATP-DnaA pentamers through IHF binding-induced DOR bending, which directs interactions between the H/B-motifs of DnaA and ssDUE, resulting in open complex formation (Sakiyama *et al*, 2022; Ozaki & Katayama, 2012; Ozaki *et al*, 2008; Sakiyama *et al*, 2017) (**Figure 0-1**). This ssDUE recruitment mechanism stabilizes the unwound state of DUE within the open complex, allowing efficient DnaB replicative helicase loading onto the ssDUE region. In-depth analyses have shown that the TTGT/TTATT motifs within DUE bind the DnaA-DOR complex and are crucial for the initiation of replication (Ozaki *et al*, 2008; Hwang & Kornberg, 1992a) (**Figure 0-1A and 0-2**), emphasizing the physiological importance of DnaA-ssDUE interactions.

Functional DnaA-ssDUE interactions have also been implicated in other bacterial species, including *Bacillus subtilis* and *H. pylori* (Richardson *et al*, 2016; Jaworski *et al*, 2021). Both origins display a bilobed structure, with *oriC1* and *oriC2* being separated by insertion of a *dnaA* gene. *B. subtilis oriC2* carries the cognate DOR with seven DnaA boxes, followed by a GC-rich 5-mer and a flanking DUE region that includes 5'-TAG-TAG-AAG-TAA-TAG-TAG-3' sequences (**Figure 0-2**). Of these sequences, the two adenine residues at positions 14 and 17 are crucial for *in vivo* initiation and repeats of the trinucleotide termed DnaA-trios, 5'-TAG-3', 5'-TAA-3', and 5'-AAG-3', are contained (Richardson *et al*, 2016, 2019). Chemical cross-linking has indicated that the *B. subtilis* DnaA-trios on ssDUE promote the oligomerization of cognate ATP-DnaA, depending on the DnaA bound to the duplex DNA DnaA boxes flanking the DUE, thereby supporting DUE unwinding *in vitro* (Richardson *et al*, 2016, 2019). By contrast, in *H. pylori oriC2* carrying the DUE with the cognate DnaA-trios (5'-AAT-AAT-AAT-TAG-TAA-CAG-

TAG-TAG-3'), the cognate DnaA-DOR complexes bind the ssDUE (Jaworski *et al*, 2021) (**Figure 0-2**). Similar mechanisms involving ssDUE interactions of the initiator protein have been observed for the *V. cholerae* chromosome 2 origin (*oriC2*) and its cognate initiator RctB, although RctB is not an AAA+ protein (Chatterjee *et al*, 2020). *oriC2* carries the cognate DUE and the flanking regions with multiple RctB-binding sequences. RctB oligomers within *oriC2* interact with ATCA repeats of ssDUE in a manner stimulated by IHF binding and DNA looping between the two RctB-interacting regions. These mechanisms in *H. pylori* and *V. cholerae* were principally similar to the ssDUE recruitment mechanism in *E. coli* (**Figure 0-1B**).

The ssDUE recruitment mechanism has also been observed at the origin of replication of the bacterium *Thermotoga maritima* (*tma-oriC*). Phylogenetic and biological analyses have placed this organism at a deep branch in the tree of life (Zhaxybayeva *et al*, 2009; Kuwabara & Igarashi, 2020), with its DnaA initiator (*tmaDnaA*) recognizing the noncanonical DnaA box with an asymmetric sequence (*tmaDnaA* box) repeated ten times within *tma-oriC* (**Figure 0-2 and 0-4**). The 149 bp minimal *tma-oriC*, which is responsible for specific unwinding, contains a 24 bp AT-rich *tmaDUE* and a flanking *tmaDOR* consisting of *tmaDnaA* boxes 1–5 (Ozaki *et al*, 2006) (**Figure 1-1**). Although *tmaDnaA* box 2 is oriented in the opposite direction, the *tmaDOR* is associated with the formation of ATP-*tmaDnaA* oligomers responsible for DUE unwinding. Moreover, the upper strand of *tma*-ssDUE binds to ATP-*tmaDnaA* oligomers bound to *tmaDOR*, depending on the *tmaDnaA* residues Val176 and Lys209, which correspond to the H/B-motifs (Ozaki *et al*, 2008). These observations are in good agreement with the ssDUE recruitment mechanism. Notably, *tmaDnaA* boxes 3–5 play a crucial role in formation of ATP-*tmaDnaA* oligomers responsible for ssDUE binding (Ozaki & Katayama, 2012). However, the contribution of the oppositely oriented *tmaDnaA* box 2 to the formation of the initiation complex remains unclear, as do the *tmaDUE* sequences responsible for DUE unwinding and ssDUE binding.

The present study provides evidence showing that the *T. maritima* DUE consists of at least two functional modules, an unwinding module and a DnaA-binding module, and that ATP-*tmaDnaA* head-to-tail oligomers can be constructed on a *tmaDnaA* box cluster

with an inverted box. The *tmaDUE* contained an *E. coli*-type TTATT motif and previously annotated DnaA-trios (5'-AAA-TAA-TA-3') flanking *tmaDnaA* box 1 (Richardson *et al*, 2016) (**Figure 0-2**), with both stimulating unwinding without binding to *tmaDnaA-tmaDOR* complexes. By contrast, binding was strictly dependent on three direct repeats of the trinucleotide TAG located between the two elements. The present study also determined the optimal arrangement of the *tmaDnaA* boxes, including their orientation and spacing. Notably, formation of head-to-tail oligomers of *tmaDnaA* did not always require DnaA box clusters in the same direction. These findings provide insight into the DnaA oligomer-binding mechanisms in *oriCs* of many species that vary in the directions of DnaA boxes.

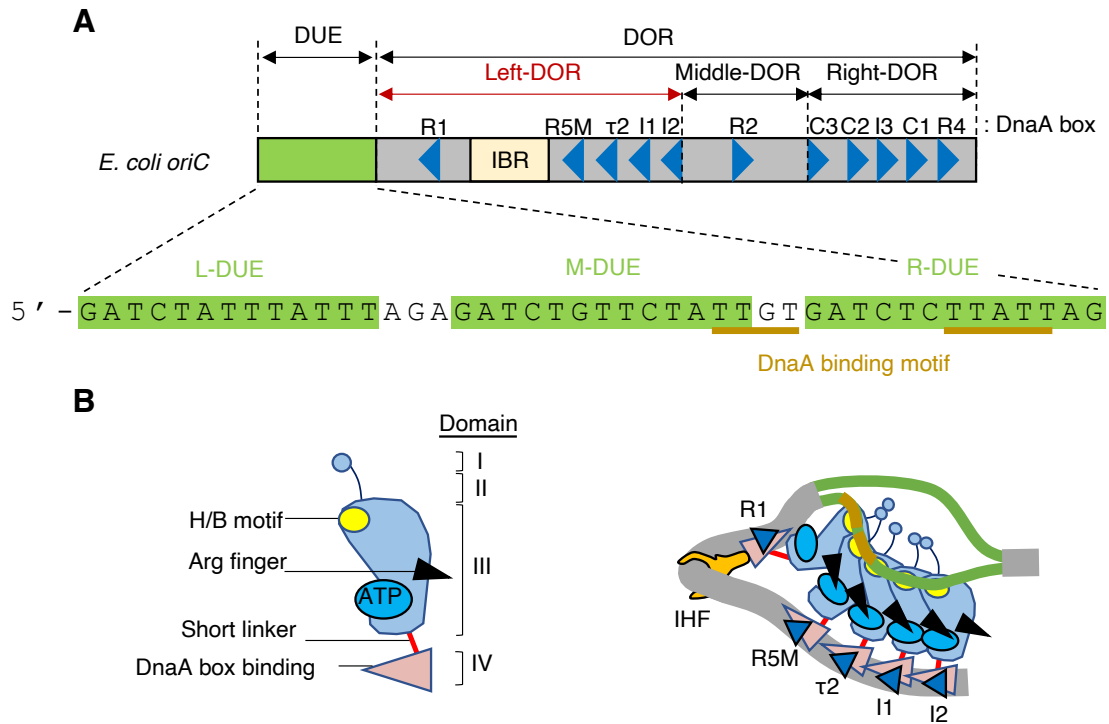


Figure 0-1. The structure of *E. coli oriC*, DnaA and open complex

(A) The structure of *E. coli oriC*. DnaA boxes (R1, R5M, τ 2, I1, I2, C3, C2, I3, C1, and R4) are indicated by arrowheads and the IHF-binding region (IBR) by a rectangle. DnaA box τ 1, which overlaps IBR, is omitted for simplicity. DUE comprises L-, M-, and R-DUE (colored in green). The DnaA-binding motifs, TTGT and TTATT, are highlighted in brown bars. Left DOR spans DnaA boxes R1–I2, middle DOR consists of DnaA box R2, and right DOR spans DnaA boxes C3–R4. (B) A model for open complex formation through the ssDUE recruitment mechanism. Schematic illustration of the domain architecture (I–IV) of a typical protein of the DnaA family. Domains III and IV are connected by a short linker in *E. coli*. The H/B and Arg-finger motifs are also indicated. In *E. coli*, the ATP-DnaA pentamer formed on IHF-bound left DOR unwinds DUE and concomitantly binds the upper strand of the ssDUE in a manner specific to the sequence TT[G/A]T(T).

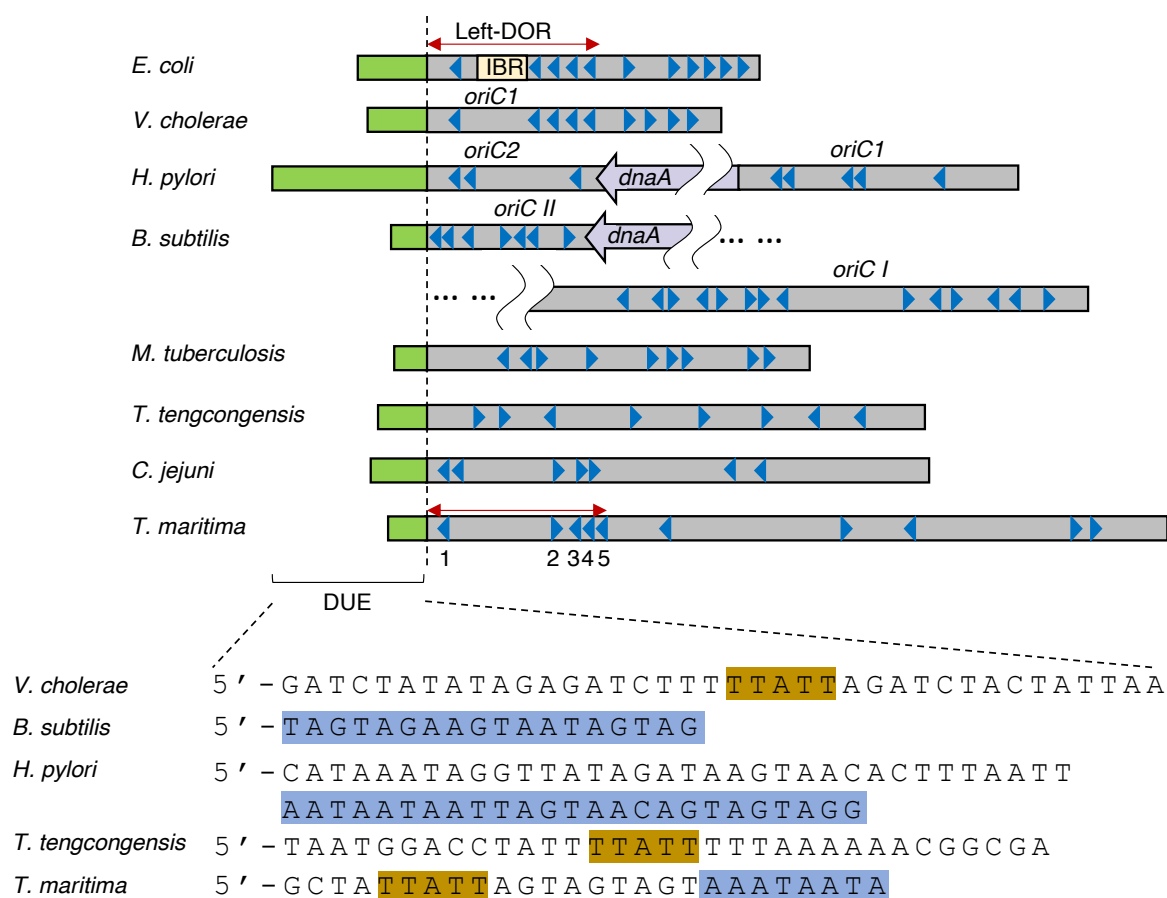


Figure 0-2. Comparison of the overall structures of the replication origins of eubacteria

DORs of the indicated eubacterial organisms are aligned. The bilobed structure of the origins of *Bacillus subtilis* and *Helicobacter pylori* are depicted, with one lobe, *oriC II* for *B. subtilis* and *oriC2* for *H. pylori*, bearing the DUE. The overall structures of the origins of *Mycobacterium tuberculosis* (Kumar *et al*, 2009), *Thermoanaerobacter tengcongensis* (Pei *et al*, 2007) and *Campylobacter jejuni* (Jaworski *et al*, 2018) are also shown. DUE is indicated by green bars, and the sequences of experimentally characterized DUEs from *B. subtilis*, *H. pylori*, *T. tengcongensis*, and *Thermotoga maritima* are shown below, highlighting the TTATT motif in brown and DnaA-trios in blue. DnaA boxes and IBR are shown as in Figure 0-1A. *E. coli* Left-DOR and minimal *tmaDOR* are indicated by red left-right arrows. In *H. pylori* and *B. subtilis*, the origin regions are separated by insertion of the cognate dnaA gene (purple arrows).

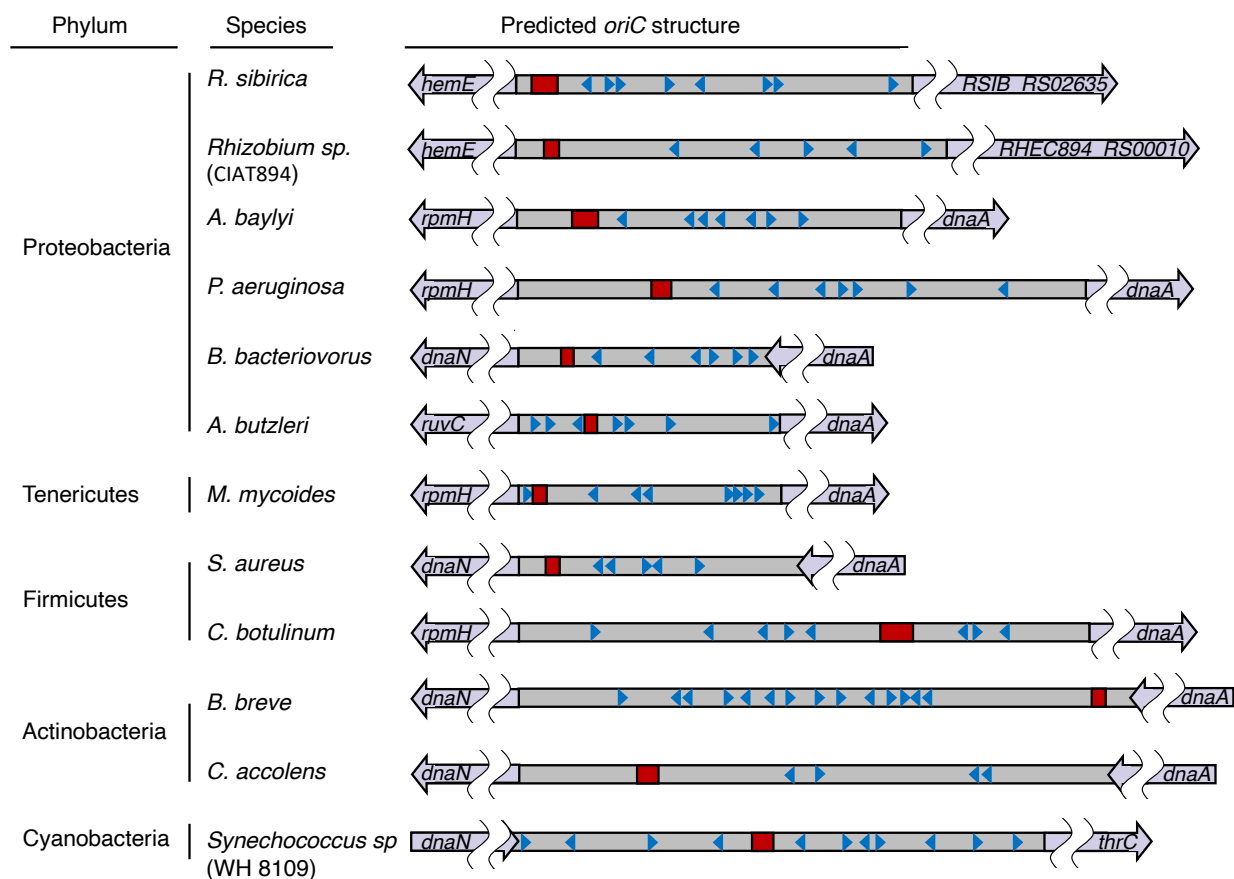


Figure 0-3. Predicted eubacterial *oriC* structures

Comparison of bioinformatically predicted origin sequences from various eubacterial species. Sequence information, including DnaA boxes (triangles), genes (purple arrow box) and AT-rich sequence motifs (red boxes), was obtained from the *DoriC* database (Dong *et al.*, 2023). Arbitrarily selected origins are shown. Abbreviations: *R. sibirica*, *Rickettsia sibirica*; *A. baylyi*, *Acinetobacter baylyi*; *P. aeruginosa*, *Pseudomonas aeruginosa*; *B. bacteriovorus*; *A. butzleri*, *Arcobacter butzleri*; *M. mycoides*, *Mycoplasma mycoides*; *S. aureus*, *Staphylococcus aureus*; *C. botulinum*, *Clostridium botulinum*; *B. breve*, *Bifidobacterium breve*; *C. accolens*, *Corynebacterium accolens*.

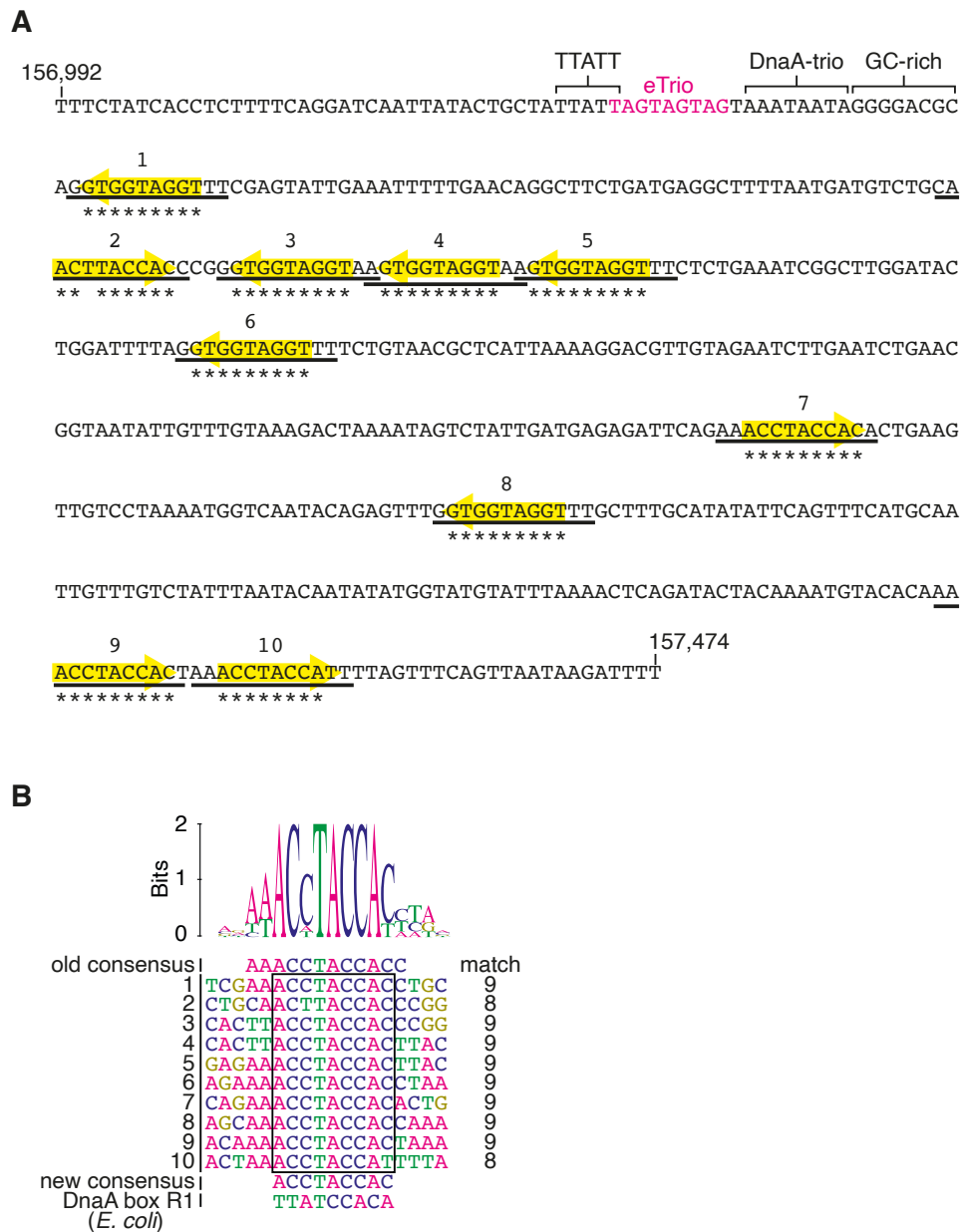


Figure 0-4. The structure of *tma-oriC*

(A) The nucleotide sequence of *tma-oriC* (nt 156,992-157,474 from the *T. maritima* MSB8 genome) is shown, along with previously annotated *tmaDnaA* boxes 1–10 (underlined), TTATT, DnaA-trio, and GC-rich motifs. Newly defined *tmaDnaA* boxes and the motif orientation are indicated by yellow arrows. Sequences identical to the consensus *tmaDnaA* box motif are highlighted by asterisks. (B) Sequence alignment and logo for *tmaDnaA* boxes. For comparison, the *E. coli* DnaA box R1 sequence is also shown.

Results

1. Identification of a novel motif within *tma-oriC* required for DUE unwinding

To gain insights into prototypic structures of the origin, the origin sequences of the deep-branching hyperthermophilic bacterium, *Thermotoga maritima*, were analyzed. The 149 bp minimal *tma-oriC* in the *T. maritima* chromosome consisted of an AT-rich *tmaDUE* and a flanking *tmaDOR* containing the five *tmaDnaA* boxes 1–5 recognized by *tmaDnaA* (**Figure 1-1**). The 12-mer sequence motifs have been named *tmaDnaA* boxes (Ozaki *et al*, 2006; Lopez *et al*, 2000). Based on further sequence analysis indicating that the terminal three bases of each motif are only moderately conserved (**Figure 0-4**), and previous data indicating that all *tmaDnaA* boxes had similar affinity, irrespective of the variations in these three bases (Ozaki *et al*, 2006), the consensus sequence was redefined as the most conserved 9-mer, 5'-ACCTACCAC-3', preserving its asymmetry. Moreover, the right part of *tmaDOR* carrying *tmaDnaA* boxes 3–5 was found to permit formation of distinct ATP-*tmaDnaA* oligomers able to bind ss-*tmaDUE*, most likely through the ssDUE recruitment mechanism (Ozaki & Katayama, 2012). Examination of the *tmaDUE* sequences suggested the importance of two ssDUE-binding motifs: TTATT from *E. coli* and the DnaA-trio from *B. subtilis* (**Figure 1-1**). Although each of these motifs has been implicated in initiation of replication of its respective organism, their functional importance in evolutionarily distant bacteria, such as *T. maritima*, was undetermined.

The DnaA-trios (5'-AAA-TAA-TA-3') are previously annotated within *tmaDUE* at the site proximal to *tmaDOR* (Richardson *et al*, 2016) (**Figure 1-1**). Importance of this sequence for open complex formation was assessed by P1 nuclease assays using *tmaDnaA*, the *E. coli* DNA-binding protein HU, and a 3.1 kb supercoiled plasmid DNA containing the *tma-oriC* (**Figure 1-1**). HU protein is the evolutionarily highly conserved IHF homolog in eubacterial species and can sustain DUE unwinding of *E. coli oriC* instead of IHF which is conserved only in proteobacteria, nitrospirae, and nitrospinae (Swinger & Rice, 2004; Ryan *et al*, 2002; Kamashev *et al*, 2017; Chodavarapu *et al*, 2007; Hwang & Kornberg, 1992b). As we previously demonstrated (Ozaki *et al*, 2006), ATP-bound, but not ADP-bound, *tmaDnaA* promotes open complex formation of *tma-oriC* at

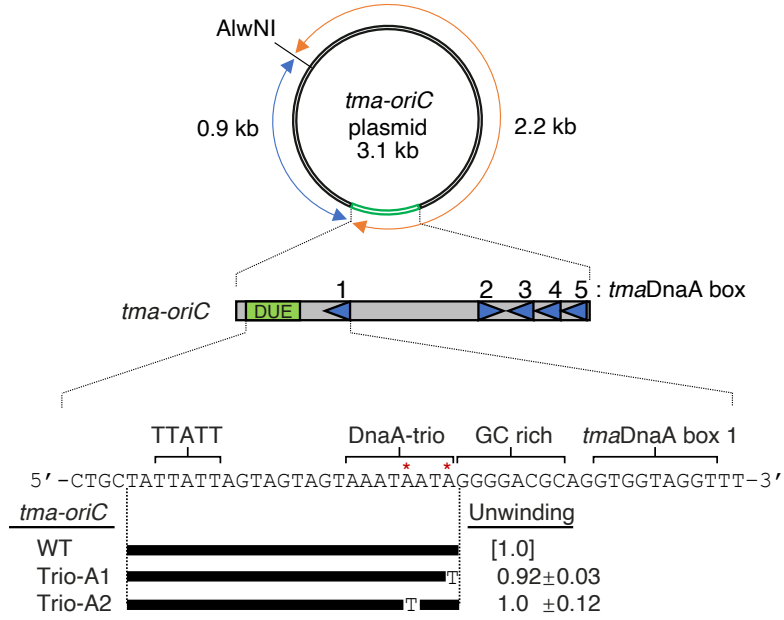
48°C in the presence of *E. coli* HU. In this assay, unwound *tmaDUE* is detected by the endonuclease P1 promoting cleavage of the single-stranded DNA. Further digestion with the restriction enzyme AlwNI yields 2.2 and 0.9 kb DNA fragments (**Figure 1-1**).

First, *tmaDUE* unwinding was assessed using *tma-oriC* bearing mutations in the annotated DnaA-Trios (**Figure 1-1A**). Because a base substitution at either the first or third position from the 3' end of the DnaA-trio sequence motif (5'-TAG-TAG-AAG-TAA-TAG-TA-3', with the corresponding residues underlined), but not at the surrounding positions, was found to lead to severe initiation defects in *B. subtilis* (Jaworski *et al*, 2021), mutant *tma-oriC* plasmids bearing an A-to-T substitution at the corresponding positions (5'-AAA-TAA-TA-3') were analyzed. The unwinding activities of these mutant plasmids were comparable to the activities of WT *tma-oriC* plasmid (**Trio-A1 and Trio-A2 in Figure 1-1BC**). Moreover, moderate unwinding activities remained even when the entire DnaA-trio was scrambled (**Trio in Figure 1-2A**). Taken together, these results suggested that the previously annotated DnaA-trio plays a supportive, but not essential, role in open complex formation.

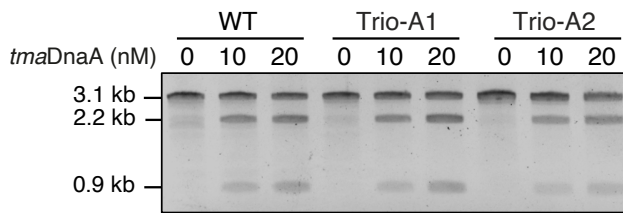
To further analyze the sequences essential for *tmaDUE* unwinding, the TTATT motif was examined similarly (**Figure 1-2A**). Scrambling of this sequence slightly inhibited *tmaDUE* unwinding (**Figure 1-2BC**). Moreover, when the TTATT and DnaA-trio sequences were simultaneously scrambled, the inhibition levels were additive, but slight unwinding activity remained (**U&Trio in Figure 1-2BC**). These observations suggested that both the TTATT motif and the DnaA-trio are required for full *tmaDUE* unwinding activity.

Analyses of the sequences located between the TTATT motif and the DnaA-trio showed three tandem repeats of TAG. Strikingly, scrambling of these sequences completely abolished the *tmaDUE* unwinding activity (**B in Figure 1-2BC**), indicating that these sequences were essential for open complex formation. Based on their homology to DnaA-trios, the three tandem TAG repeats have been named the extended Trio (eTrio).

A



B



C

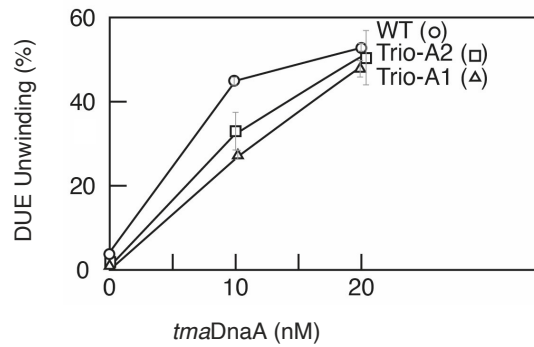


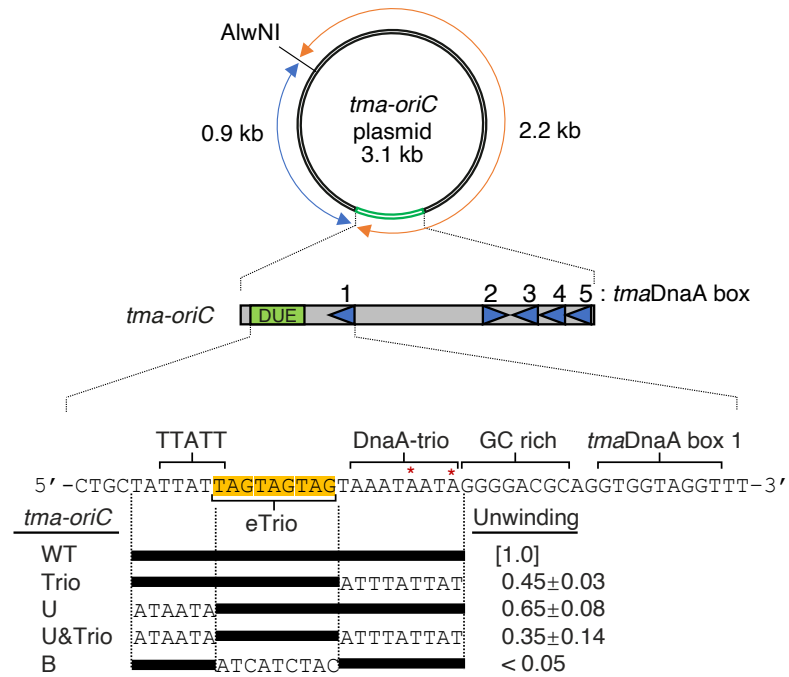
Figure 1-1. DnaA-trio is not essential for open complex formation

(Figure legend is shown in the next page)

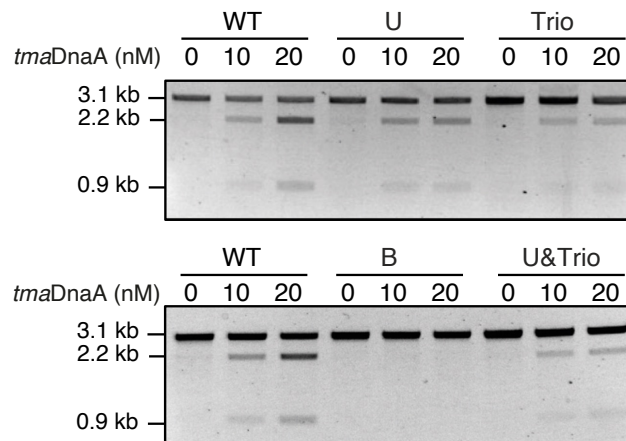
Figure 1-1. DnaA-trio is not essential for open complex formation

(A) The structure of *tma-oriC* plasmids. pOZ14 is a pBluescript derivative with a 149 bp minimal *tma-oriC* DNA. *tma-oriC* is indicated by a grey bar, *tmaDUE* by an green box (DUE), and *tmaDnaA* boxes 1–5 by blue triangles (box 1-5). Sequences of the TTATT, eTrio, DnaA-trio, GC-rich, and *tmaDnaA* box 1 motifs are shown below the *tma-oriC*, as are derivatives of the *tmaDUE* region. Restriction site of AlwNI is indicated. For mutant *tma-oriC* plasmids, intact DUE regions are indicated by bold lines and base substitutions by letters. The percentages of the open complex at 20 nM *tmaDnaA* (unwinding) are shown relative to that of the wild type (WT). The percentages of the open complex at 20 nM *tmaDnaA* (unwinding) are shown relative to that of the wild type (WT). (B and C) Open complex formation. WT and mutant *tma-oriC* plasmids were individually incubated with the indicated concentrations of ATP-*tmaDnaA*, followed by P1 nuclease digestion. After purification, DNA samples were further digested with AlwNI and analyzed by 1% agarose gel electrophoresis and ethidium bromide staining. (B) Gel images. (C) Mean $\pm$ standard deviation percentages of P1 nuclease-digested DNA (n = 2) quantified by FIJI software.

A



B



C

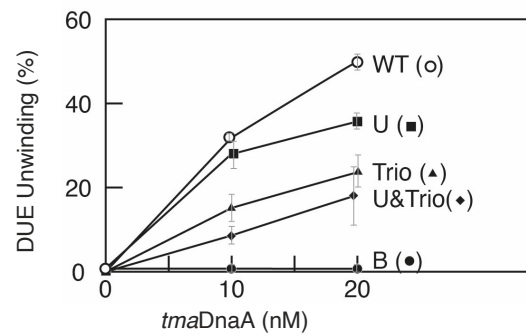


Figure 1-2 The three tandem TAG repeats are essential for DUE unwinding
(Figure legend is shown in the next page)

Figure 1-2 The three tandem TAG repeats are essential for DUE unwinding

(A) The structure of *tma-oriC* plasmids is described in Figure 1-1. The percentages of the open complex at 20 nM *tmaDnaA* (unwinding) of mutant plasmids are shown relative to that of the wild type (WT). (B and C) Open complex formation. WT and mutant *tma-oriC* plasmids were analyzed by P1 assays as described in the legend to Figure 1-1B Gel images. (C) Mean $\pm$ standard deviation percentages of P1 nuclease-digested DNA (n = 2) quantified by FIJI software.

2. ATP-*tmaDnaA* oligomers on DOR bind ssDUE through eTrio

To further dissect the roles of sequence motifs within DUE, interactions between ss-*tmaDUE* and ATP-*tmaDnaA* oligomers constructed on *tmaDOR* were analyzed by EMSA. In these assays, a radio-labeled, 28-mer ss-*tmaDUE* probe was coincubated with ATP-*tmaDnaA* and *tmaDOR*, followed by polyacrylamide gel electrophoresis (**Figure 2-1A**). As we previously reported (Ozaki & Katayama, 2012), ATP-*tmaDnaA* oligomers constructed on *tmaDOR* (ATP-*tmaDnaA-tmaDOR* complexes) specifically interact with the ligand TMA28, an upper strand fragment of ss-*tmaDUE* (**Figure 2-1A-D**). These interactions depend on *tmaDOR* and ATP-bound, but not ADP-bound, *tmaDnaA* (**Figure 2-1E**), as previously reported (Ozaki & Katayama, 2012). DnaA proteins are apt to form irregular aggregates in the absence of DNA binding, remaining in the gel well. Moreover, a right part of *tmaDOR*, including *tmaDnaA* boxes 3–5, is sufficient to construct ATP-*tmaDnaA* complexes capable of interacting with ss-*tmaDUE* (Ozaki & Katayama, 2012).

EMSAs using a set of ss-*tmaDUE* variants with oligo-dC substitutions were performed to determine the minimal sequence required by the upper strand of ss-*tmaDUE* to bind ATP-*tmaDnaA-tmaDOR* complexes. Because eTrio was found essential for DUE unwinding, its ability to bind ATP-*tmaDnaA-tmaDOR* complexes was analyzed, with results showing that the ss-*tmaDUE* variant bearing the eTrio and oligo-dC regions (sB) displayed binding activity comparable to WT TMA28 or variants bearing eTrio and a partial DnaA-trio (sB+6 and sB+3) (**Figure 2-2**). Moreover, all three TAG trinucleotides within eTrio were required for binding as reducing the number of TAG trinucleotides to two (sBΔ3) or one (sBΔ6) completely abolished the binding activity. Similarly, the ss-*tmaDUE* variant bearing only the DnaA-trio and oligo-dC regions (sTrio) was inactive (**Figure 2-2**), indicating that ss-*tmaDUE* binding strongly depends on the three TAG repeats within eTrio. Taken together with the results showing that eTrio is strictly required for DUE unwinding (**Figure 1-2**), these findings indicate that direct interactions between ss-eTrio and ATP-*tmaDnaA-tmaDOR* complexes are important in stabilizing open complexes. Because the DnaA-trio within *tmaDUE* is dispensable for binding to ATP-*tmaDnaA-tmaDOR* complexes but plays only a supportive role in open complex

formation, the DnaA-trio might assist in the process of initial AT-rich DNA-preferential duplex unwinding, reducing the stability of the duplex.

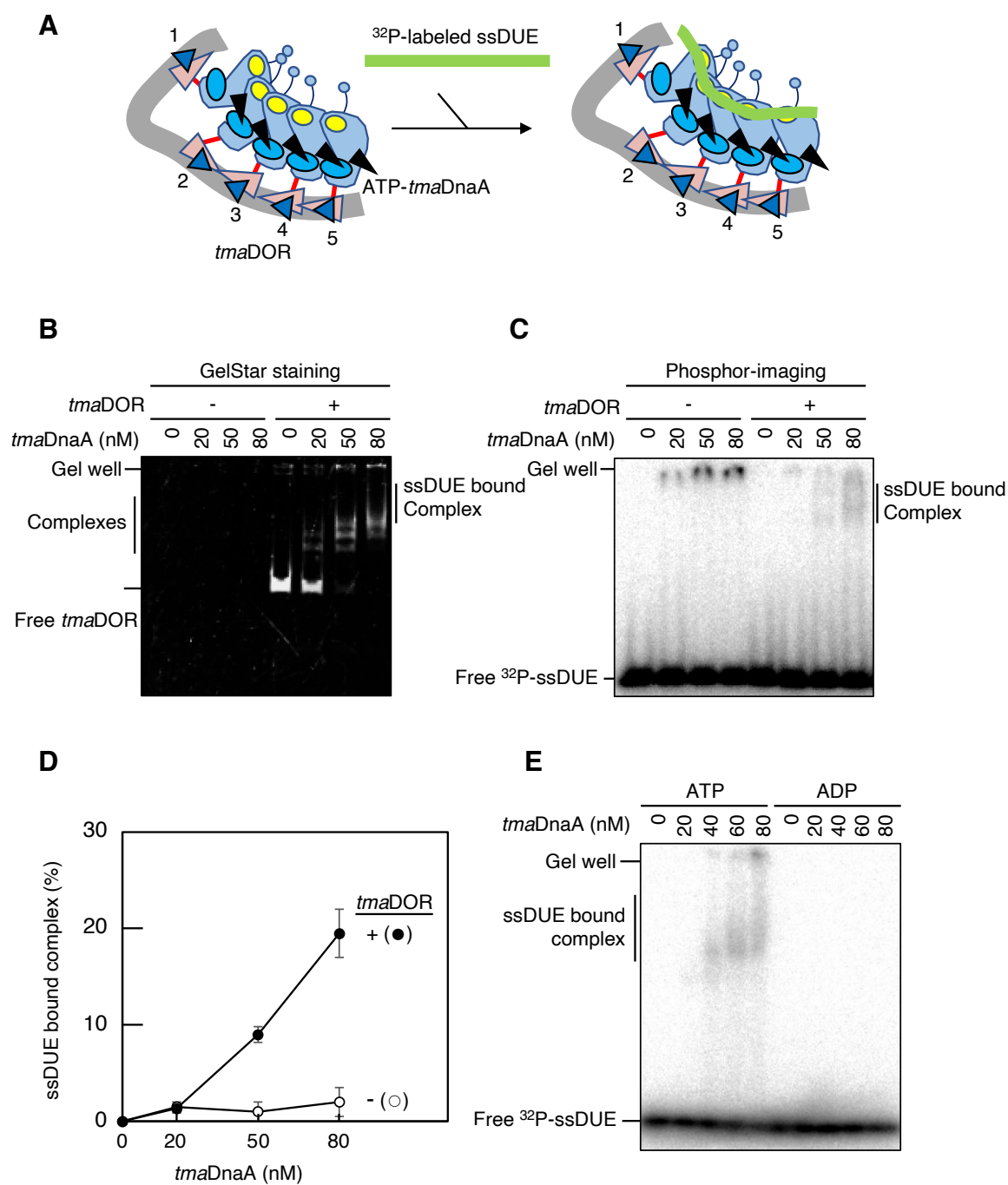


Figure 2-1. ss-*tmaDUE* interactions with ATP-*tmaDnaA*-*tmaDOR* complexes
(Figure legend is shown in the next page)

Figure 2-1. ss-*tma*DUE interactions with ATP-*tma*DnaA-*tma*DOR complexes

(A) Schematic of the ssDUE-binding assay. ATP-*tma*DnaA-*tma*DOR complexes were constructed and incubated with ^{32}P -labeled ssDUE, followed by EMSA. See the other data in this paper for overall structure of ATP-*tma*DnaA-*tma*DOR complexes. (B-E) ATP-*tma*DnaA was incubated with (+) or without (-) *tma*DOR (15 nM) and further incubated with wild-type (WT) ^{32}P -ss-*tma*DUE (1.5 nM), followed by EMSA using a 4% polyacrylamide gel (n=2). (B) Visualization of dsDNA by GelStar staining. (C) Determination of ^{32}P -ss-*tma*DUE by phosphor-imaging. (D) Quantitation of the amounts of ^{32}P -ss-*tma*DUE bound to ATP-*tma*DnaA-*tma*DOR complexes in (C), with the mean $\pm$ standard deviation percentages relative to input *tma*-ssDUE plotted as ssDUE bound complexes. Abbreviations: Free, protein-free DNA; Complexes, ATP-*tma*DnaA-*tma*DOR complexes; Well, gel well; ssDUE bound complex, ATP-*tma*DnaA-*tma*DOR- ^{32}P -ss-*tma*DUE complex. (E) Phosphor-imaging for titration of ATP/ADP-*tma*DnaA in the presence of *tma*DOR and WT ^{32}P -ss-*tma*DUE.

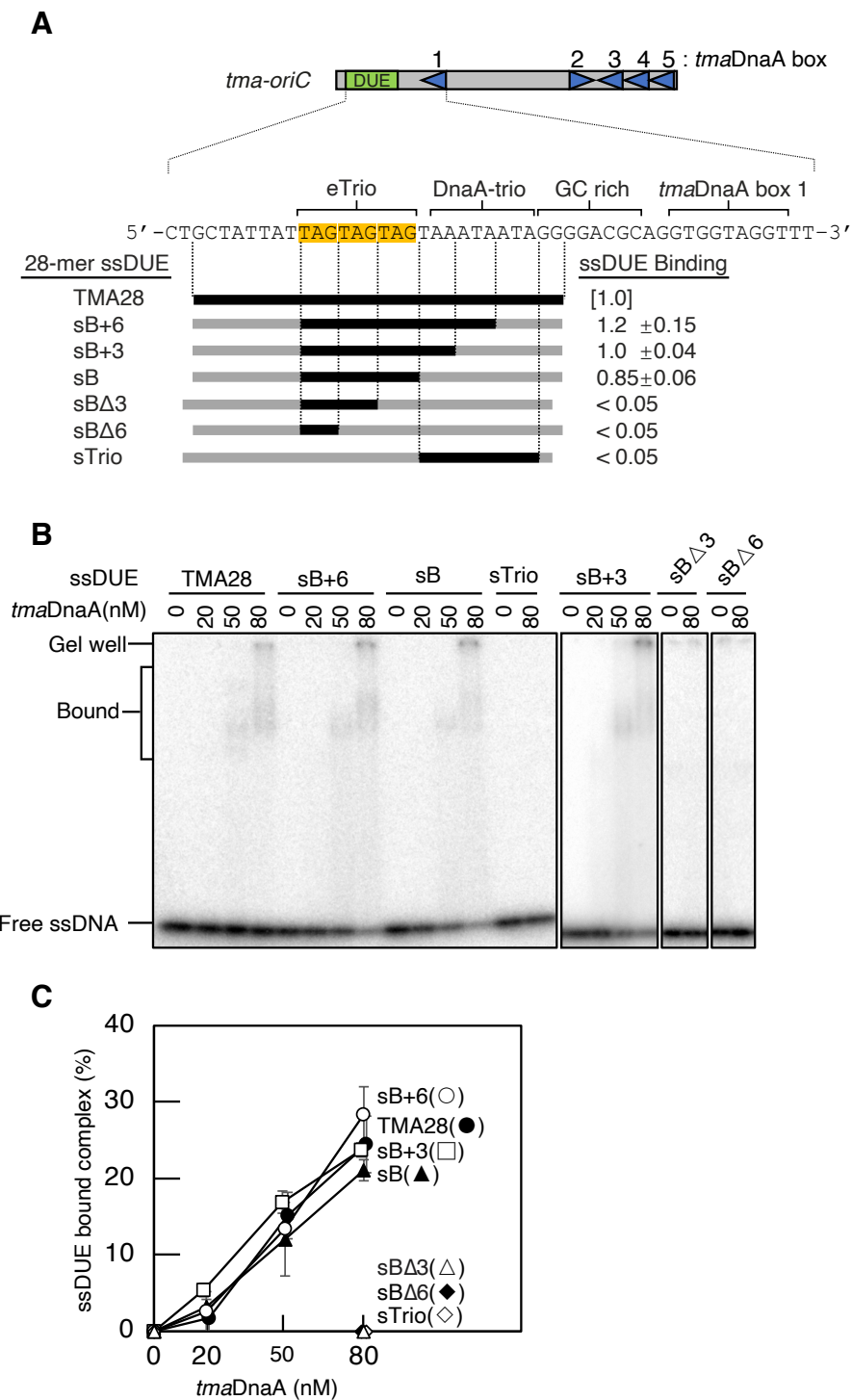


Figure 2-2. The three tandem TAG repeats are essential for ssDUE binding
(Figure legend is shown in the next page)

Figure 2-2. The three tandem TAG repeats are essential for ssDUE binding

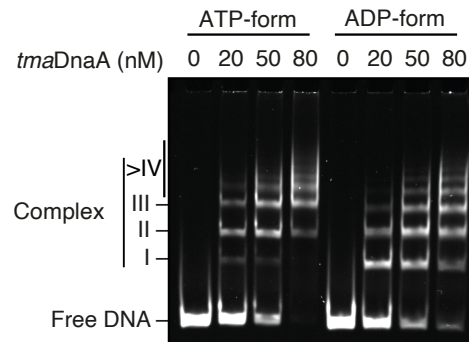
Wild-type ssDUE fragment (TMA28) or its derivatives with oligo-dC-substitution (2.5 nM) were incubated with *tmaDOR* (30 nM) and the indicated amount of ATP-*tmaDnaA*, followed by EMSA. Minimal *tma-oriC* is shown as in Figure 1-1A. The ssDUE sequences used in this assay are illustrated schematically in *panel A*, with intact and oligo-dC-substituted sequences shown in bold and gray lines, respectively. (B) Representative gel images. (C) Quantitation of the results in (B) ($n = 2$) relative to the input ssDUE. Binding activities (ssDUE binding) at 80 nM *tmaDnaA* are shown in *panel B* relative to that of TMA28.

3. Individual *tmaDnaA* boxes within minimal *tmaDOR* are crucial for unwinding

To investigate the mechanisms underlying *tmaDnaA*-complex formation, the roles of individual *tmaDnaA* boxes in DUE unwinding were analyzed. DNase I footprint analyses show binding of ATP-*tmaDnaA* to *tmaDnaA* boxes 1–5 within *tmaDOR* (Ozaki & Katayama, 2012), a finding supported by the EMSA results in the present study (**Figure 3-1**). EMSA showed co-operative binding of ATP-*tmaDnaA* molecules to *tmaDOR*, as well as ADP-*tmaDnaA* binding to *tmaDOR*, the latter resulting from the absence of a competitor DNA and the occurrence of cage effects impeding diffusion, differing from the results of DNase I footprint experiments. In addition, deletion of *tmaDnaA* box 5 from minimal *tma-oriC* is found to reduce the DUE unwinding activity to ~50% of the intact sequence (Ozaki *et al*, 2006). Residual activity is completely abolished by the simultaneous deletion of *tmaDnaA* boxes 4 and 5, suggesting that these *tmaDnaA* boxes are essential for activity (Ozaki *et al*, 2006).

The requirements for individual *tmaDnaA* boxes 1–5 were analyzed by performing P1 nuclease assays using a set of minimal *tma-oriC* plasmid pOZ14 derivatives in which each *tmaDnaA* box sequence was randomized (**Figure 3-2A**). Consistent with previous results, a mutation in *tmaDnaA* box 4 or 5 each inhibited the unwinding activity by ~50% (sub4 and sub5 in **Figure 3-2A-C**). Similarly, randomization of *tmaDnaA* box 2 or 3 inhibited unwinding activity ~50% (sub2 and sub3 in **Figure 3-2A-C**), and randomization of *tmaDnaA* box 1 almost completely abolished unwinding activity (sub1 in **Figure 3-2A-C**). These results are consistent with the hypothesis that each of the five *tmaDnaA* boxes plays a crucial role for full DUE unwinding activity. The stricter requirement of *tmaDnaA* box 1 was also in good agreement with the ssDUE recruitment mechanism, in that ATP-*tmaDnaA* bound to *tmaDnaA* box 1 brings *tmaDUE* close to ATP-*tmaDnaA* complexes on the *tmaDnaA*-complex bound region (**Figure 2-2A and B**).

A



B

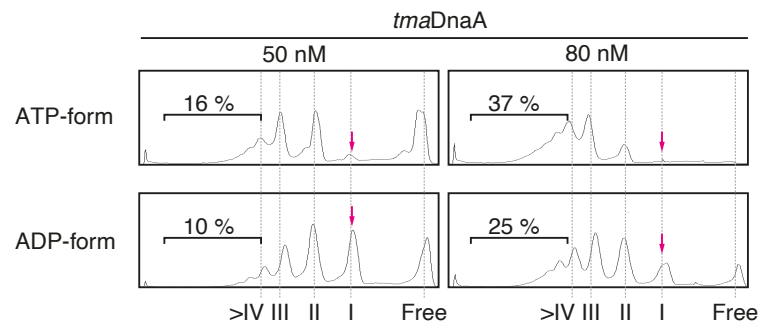


Figure 3-1. ATP-*tmaDnaA* binding of *tmaDOR*

(A) A 203 bp dsDNA (30 nM) containing wild-type *tmaDOR* was incubated with the indicated concentrations of ATP-*tmaDnaA* or ADP-*tmaDnaA*, followed by EMSA. DNA bands were visualized by GelStar staining. (B) Band intensities corresponding to Free DNA and Complexes in (A) at 50 and 80 nM *tmaDnaA* were quantified using FIJI software. The positions of complex I are indicated by red arrowheads. The percentages of intensities corresponding to the complex >IV are shown.

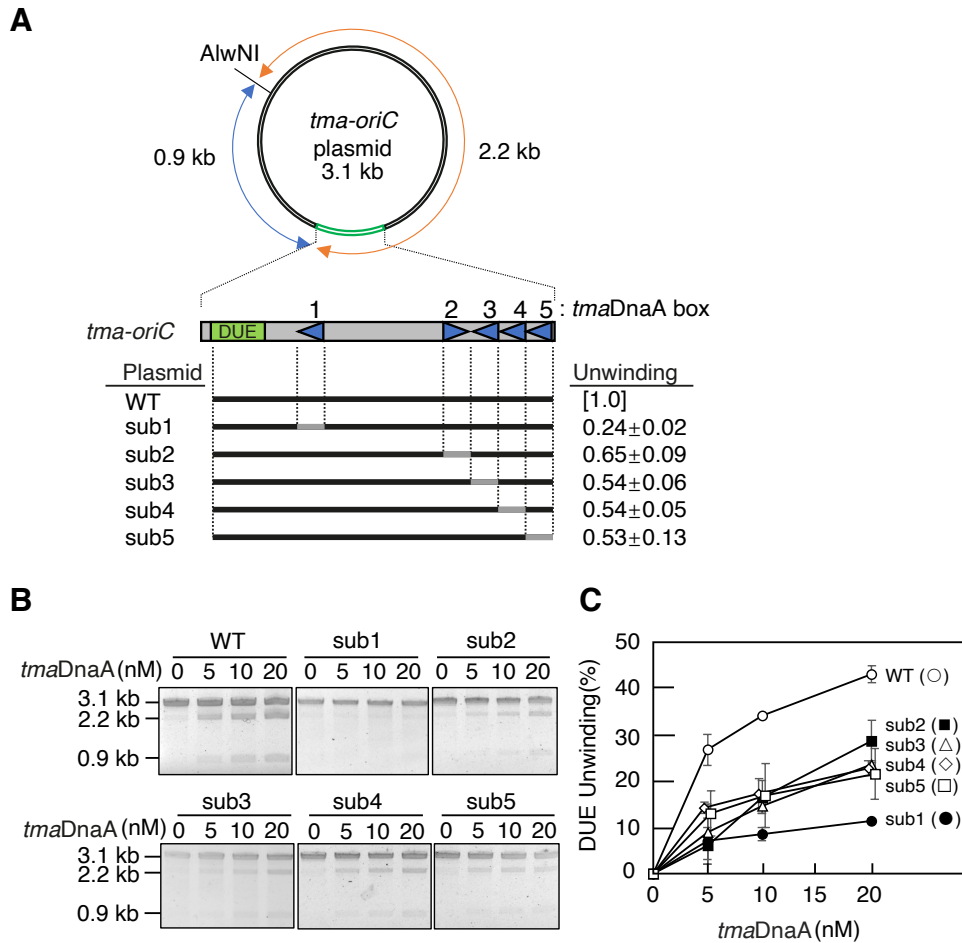


Figure 3-2. Each of the *tmaDnaA* boxes 1–5 within *tmaDOR* is important for DUE unwinding

Open complex formation by wild-type (WT) *tma-oriC* and its mutant derivatives (sub1-5), analyzed by P1 assays as described in the legend to Figure 1-1. (A) Schematic illustration of the *tma-oriC* plasmids. Gray bars indicate base substitutions, in which the 12-mer sequence, including the 9-mer *tmaDnaA* box and its surrounding nucleotides, was replaced by the randomized sequence 5'-CCCAAGCAACAA-3' (Ozaki & Katayama, 2012). The WT sequence is indicated by bold bars. (B) Representative gel images. (C) Quantitative data (n = 2) shown as in Figure 1-1C. The mean ± standard deviation activity relative to that of WT at a *tmaDnaA* concentration of 20 nM are also shown.

4. Intact DNA helical turn between DUE and DOR is crucial for unwinding

To investigate the mechanisms required for open complex formation, importance of DNA helical turn differences in the *tmaDnaA* box 1-box 2 intervening region in *tmaDUE* unwinding was analyzed (**Figure 4A**). Specifically, this study hypothesized that if the ssDUE recruitment mechanism was responsible for open complex formation, then inserting a full (10 bp) turn of a DNA helix would allow this region to retain its phasing, resulting in sustained unwinding activity (**Figure 4B**). Conversely, inserting a half (5 bp) turn of a DNA helix would alter the phasing between the *tmaDUE*-*tmaDnaA* box 1 region and *tmaDnaA* boxes 2–5, altering the interaction modes among the ATP-*tmaDnaA* complexes and impairing interactions between the *tmaDUE* and DnaA complexes.

These possibilities were tested by performing P1 nuclease assay using mutant *tma-oriC* plasmids with either a 5 bp or 10 bp insertion at a position flanking box 1 or box 2. DUE unwinding activity was fully preserved by insertion of a 10 bp fragment but was reduced by insertion of a 5 bp fragment (**Figure 4C and D**). These observations are in good agreement with the ssDUE recruitment mechanism, in that appropriate phasing between the DnaA complexes promotes these interactions via DNA looping, making this conformation crucial for the open complex at *tma-oriC*.

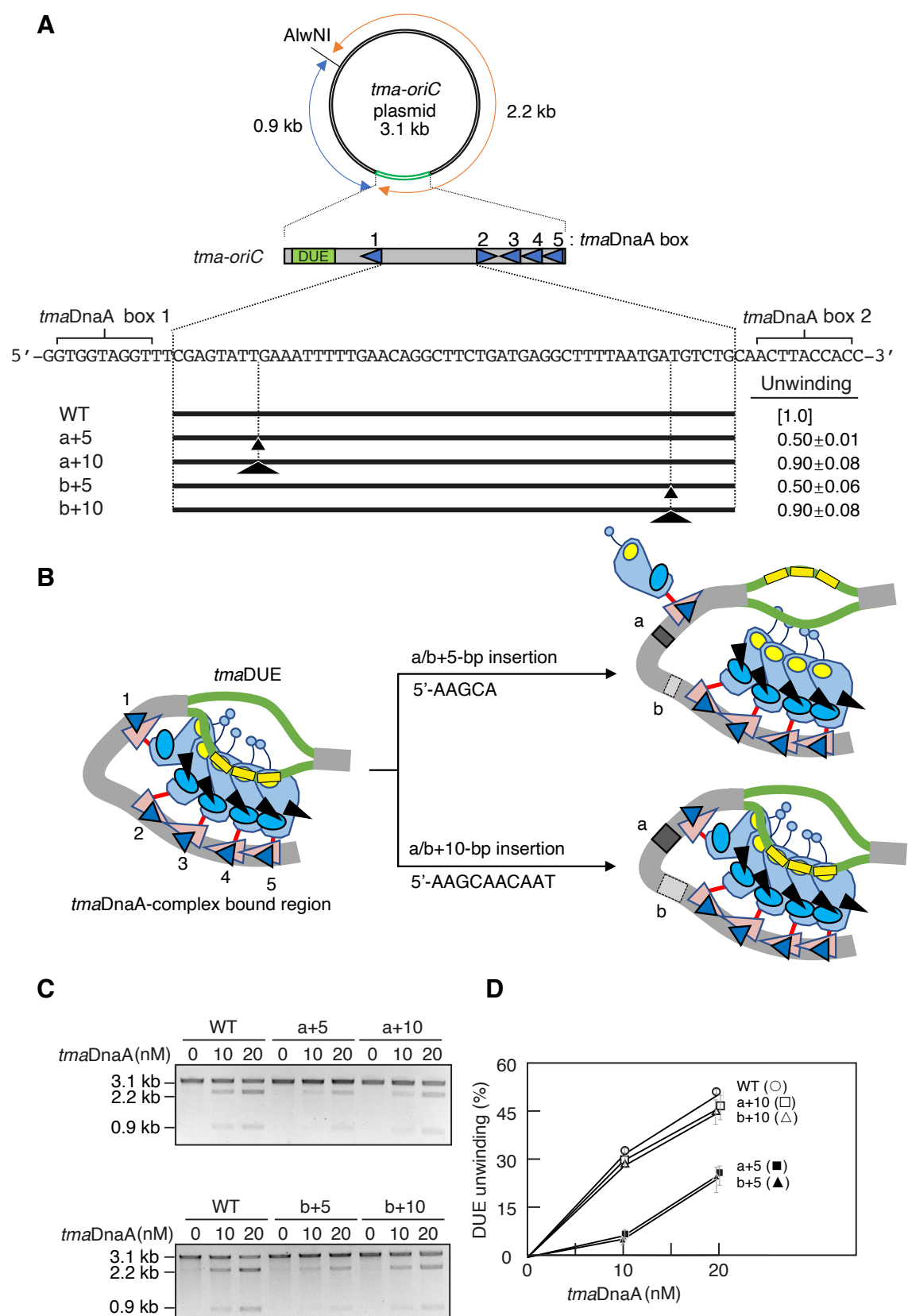


Figure 4. The DNA helical turn between DUE and DOR underlies DUE unwinding
(Figure legend is shown in the next page)

Figure 4. The DNA helical turn between DUE and DOR underlies DUE unwinding

(A) The structure of the *tma-oriC* plasmids. Wild-type (WT) and mutant *tma-oriC* plasmids are shown as described in the legend to Figure 1-1. The positions of base insertions are indicated by small (+5) or large (+10) arrowheads. The amounts of the open complex at 20 nM *tmaDnaA* (unwinding) are shown relative to that of WT.

(B) Schematic presentation of the altered phasing between *tmaDUE* and a *tmaDnaA*-complex-bound region resulting from a 5 or 10 bp insertion within *tmaDOR*. Presumable conformations of the complexes are shown based on the ssDUE recruitment mechanism. Inserted sequences are shown. *tmaDnaA* and *tmaDnaA* boxes are shown similar to that of *E. coli* DnaA in Figure 0-1B. eTrio in *tmaDUE* is indicated by yellow rectangles. (C-D) Open complex formation. WT and mutant *tma-oriC* plasmids were analyzed by P1 nuclease assays, as described in the legend to Figure 1-1. (C) Representative gel images and (D) quantitative data (n = 2) are shown as in Figure 1-1.

5. Involvement of the Arg-finger in open complex formation

It was unclear whether a conventional AAA⁺-family-type head-to-tail interaction was involved in the formation of the open complex at *tma-oriC*. Evaluation of the *E. coli* DOR showed that the direct repeats of DnaA boxes reinforced formation of ATP-DnaA oligomers in a head-to-tail manner via interactions between the DnaA Arg285 Arg-finger and ATP bound to the adjacent protomer (Noguchi *et al*, 2015; Kawakami *et al*, 2005) (**Figure 0-1B**). Although the Arg-finger motif was found to be conserved at position 251 in *tmaDnaA*, the minimal *tma-oriC* includes an inverted *tmaDnaA* box 2. Thus, the mechanism by which the *tmaDnaA* box 2-bound *tmaDnaA* protomer interacts with the flanking *tmaDnaA* protomers within the functional complex was unclear. Because previous DNase I footprint assays showed that ATP-*tmaDnaA*, but not ADP-*tmaDnaA*, binds co-operatively to the region from *tmaDnaA* box 2 to box 5 (Ozaki & Katayama, 2012), then the *tmaDnaA* associated with the inverted *tmaDnaA* box 2 should sustain its ability to interact with the adjacent *tmaDnaA* molecule bound to *tmaDnaA* box 3 through either conventional head-to-tail or nonconventional head-to-head interactions through the AAA⁺ domains.

In *E. coli* DnaA, the C-terminus of domain III (the AAA⁺ domain) is connected to the N-terminus of domain IV (the DnaA box-binding domain) via a short flexible linker consisting of amino acids Leu367-Thr375, with the flexibility of this linker allowing limited swiveling of the two domains, as suggested by a study of the molecular dynamics underlying the construction of DnaA pentamers on left and right DORs (Shimizu *et al*, 2016). AlphaFold2-based structural model consistently suggested that *E. coli* DnaA contains a short flexible linker consisting of only four amino acid residues, Lue373-Ile376 (Jumper *et al*, 2021). By contrast, similar modelling suggested that *tmaDnaA* and other representative DnaA orthologs have corresponding linkers with a longer region consisting of 11–12 amino acid residues, allowing greater swiveling of the domains (**Figure 5-1**). Thus, a *tmaDnaA* protomer bound to the inverted *tmaDnaA* box 2 may use its flexible linker to revolve around its AAA⁺ domain, bringing the Arg-finger close to the ATP at the adjacent protomer bound to *tmaDnaA* box 3 (**Figure 5-2**). The resultant head-to-tail interaction would stabilize the overall structure formed on *tmaDnaA* boxes 2–5.

To confirm this hypothesis, the dependence of the co-operative ATP-*tmaDnaA* interactions within the *tmaDnaA*-complex bound region on the Arg-finger motif was assessed. To investigate the specific role of this motif, a *tmaDnaA* variant containing an Ala residue, rather than an Arg residue, at amino acid 251 was purified (R251A). A filter retention assay demonstrated that the affinity of *tmaDnaA* R251A for ATP or ADP was comparable to that of WT *tmaDnaA* (**Figure 5-3**). Moreover, EMSA using a DNA carrying a single consensus *tmaDnaA* box sequence demonstrated that the *tmaDnaA* box-binding activity of *tmaDnaA* R251A was similar to that of WT *tmaDnaA* (**Figure 5-4A**). Thus, the Arg-finger of *tmaDnaA* is not required for binding to either ATP or a single *tmaDnaA* box. By contrast, a P1 nuclease assay showed that *tmaDnaA* R251A is virtually inactive for unwinding DUE, indicating that the Arg-finger is essential for open complex formation (**Figure 5-4B**). These properties are fully consistent with those of the *E. coli* Arg-finger variant DnaA R285A (Kawakami *et al*, 2005).

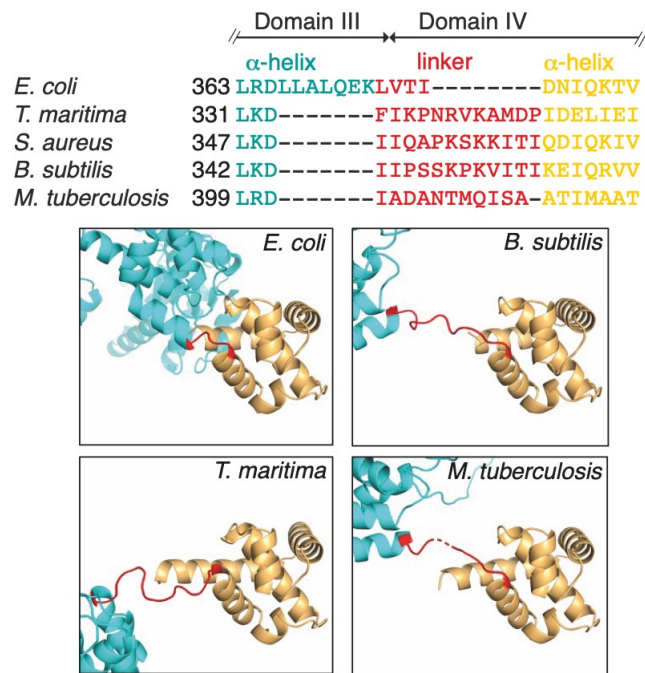


Figure 5-1. Longer linkers between DnaA domains III and IV in DnaA orthologs than in *E. coli* DnaA.

(A) The linker sequences between DnaA domains III and IV in the representative DnaA orthologs are shown. (*upper panel*) Alignment of the amino acid residues (red letters) forming the predicted linkers of *E. coli* DnaA and the indicated DnaA orthologs. (*lower panel*) DnaA structures predicted using AlfaFold2, with the predicted structures spanning the region from the C-terminus of domain III to the N-terminus of domain IV shown for comparison; linker regions are indicated in red.

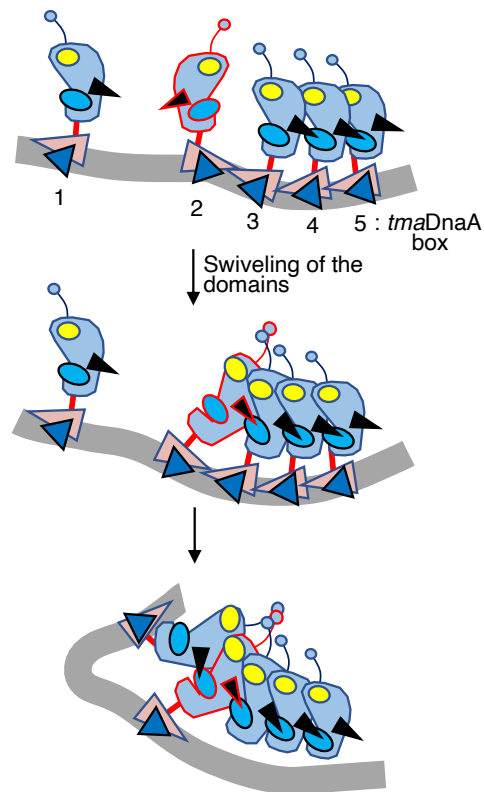


Figure 5-2. Model for the head-to-tail interaction on *tmaDnaA* boxes 2–5.

Model for the Arg-finger-mediated oligomerization of the *tmaDnaA* proteins on *tmaDnaA* boxes 2–5. ATP-*tmaDnaA* binds to each *tmaDnaA* box, with the protomers on *tmaDnaA* boxes 3–5 interacting in a head-to-tail manner through the function of the Arg-finger. Swiveling of the AAA+ domain of *tmaDnaA* on box 2 allows the Arg-finger-mediated interaction of ATP-*tmaDnaA* with the *tmaDnaA* trimer on boxes 2–3, stabilizing the overall structure of the ATP-*tmaDnaA* tetramer constructed on the region encompassing *tmaDnaA* boxes 2–5.

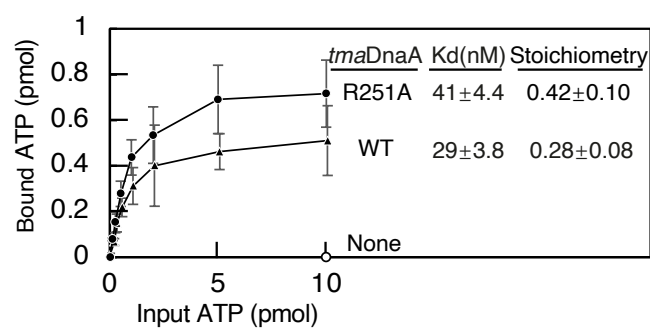
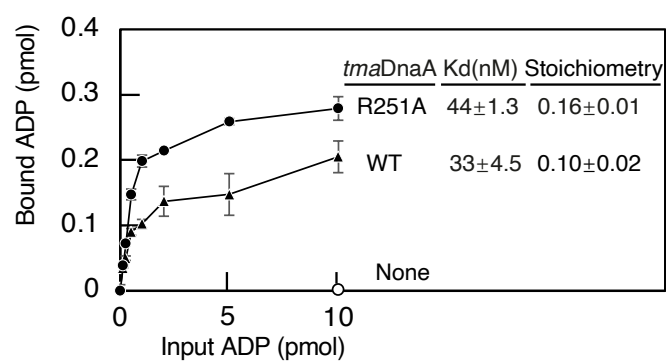
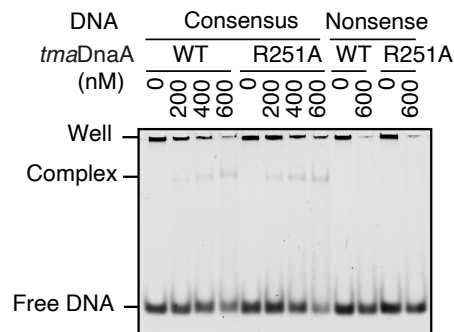
A**B**

Figure 5-3. ATP/ADP-binding activity of WT *tmaDnaA* or *tmaDnaA* R251A

ATP-binding activity. Wild-type (WT) *tmaDnaA* or *tmaDnaA* R251 was mixed with various concentrations of radio-labeled ATP (A) or radio-labeled ADP (B), followed by filter retention assays(n=3). The dissociation constant (K_d) deduced from the Scatcherd plot is shown.

A



B

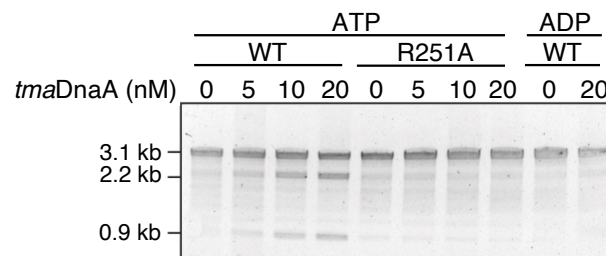


Figure 5-4. The Arg-finger is essential for open complex formation but not for *tmaDnaA* box-binding

(A) DNA-binding activity. ADP-forms of WT *tmaDnaA* or *tmaDnaA* R251A were incubated with 18 bp DNA (300 nM) containing *tmaDnaA* box 1 (5'-AGACCACCTACCACATAA-3'; *tmaDnaA* box is underlined) or a nonsense DNA (5'-AGACCCAAGCAACAATAA-3'), followed by EMSA using an 8% polyacrylamide gel.

(B) Open complex formation. The DUE unwinding activities of ATP-*tmaDnaA*, ADP-*tmaDnaA*, and ATP-*tmaDnaA* R251 proteins were analyzed using pOZ14 DNA, as described in the legend to Figure 1-1.

6. Inverted *tmaDnaA* boxes permit ATP-*tmaDnaA* interactions in a head-to-tail manner

To provide further evidence for this model, EMSA was performed using a DNA fragment with oppositely oriented *tmaDnaA* boxes 2 and 3. Canonical head-to-tail interactions should stabilize the binding of the two *tmaDnaA* molecules on DNA in a manner dependent on both ATP and the Arg-finger. Indeed, when WT ATP-*tmaDnaA* was used, the predominant DNA complex contained two *tmaDnaA* molecules (C2), whereas the complex containing one *tmaDnaA* molecule (C1) was barely detected (**Figure 6-1A and B**). This preference for C2 formation was evident even at limited concentrations of ATP-*tmaDnaA* (50, 100 nM). By contrast, C1 was the major product when WT ADP-*tmaDnaA* or ATP-*tmaDnaA* R251A was used. Thus, both ATP and the Arg-finger are involved in the efficient formation of C2 complexes. The residual amounts of C2 formed by WT ADP-*tmaDnaA* and ATP-*tmaDnaA* R251A suggest that subpopulations of these *tmaDnaA* molecules could interact with each other in a manner dependent on neither ATP nor the Arg-finger: as observed for *E. coli* and *Streptomyces lividans* DnaAs (Sugiyama *et al*, 2019; Majka *et al*, 2001), weak domain I-domain I interaction and head-to-head domain III-domain III interaction might be involved.

The dependence of C2 formation on the spatial arrangement of the two repeated *tmaDnaA* boxes was evaluated by EMSA using a set of direct or inverted repeats of *tmaDnaA* boxes with different interspacing distance. The results suggested that formation of C2 depends on appropriate spacing between two *tmaDnaA* boxes and their orientation (**Figure 6-2A-C**). Specifically, maximum C2 formation by two identically oriented *tmaDnaA* boxes was observed when they were separated by 2 bp, as shown in the arrangements of *tmaDnaA* boxes 3–4 and boxes 4–5. By contrast, maximum C2 formation by two inverted *tmaDnaA* boxes was observed when they were separated by 4 bp interspace, as observed for *tmaDnaA* boxes 2–3. Interestingly, a subpopulation of ATP-*tmaDnaA* showed formation of a higher ordered complex C3 on the inverted *tmaDnaA* repeat with separations of 5 or 6 bp (**Figure 6-2B**). These findings suggest that there may be an as yet uncharacterized mode of inter-ATP-*tmaDnaA* interactions on DNA.

Moreover, the intervening space between two *tmaDnaA* boxes and their orientation are key determinants for efficient formation of these dimeric complexes.

Finally, to confirm the concept that the inverted repeats of *tmaDnaA* boxes allow for head-to-tail ATP-*tmaDnaA* interactions to form an open complex at *tma-oriC*, a P1 nuclease assay was performed using a *tma-oriC* mutant derivative bearing the reversed *tmaDnaA* box 2, aligning *tmaDnaA* boxes 2-4 in the same direction with a 2-bp interspace (**R2 in Figure 6-3C**). As a result, the unwinding activity of this mutant plasmid was comparable to that of the wild-type *tma-oriC* plasmid (**Figure 6-3A-C**). This finding contrasts with the compromised unwinding activity observed in the mutant plasmid with a randomized *tmaDnaA* box 2 (**Figures 3-2 and 6-3**). These observations suggest that *tmaDnaA* box 2 is involved in open complex formation, regardless of its direction, highlighting a tunable nature of DnaA box orientation in the formation of bacterial initiation complexes at the origin.

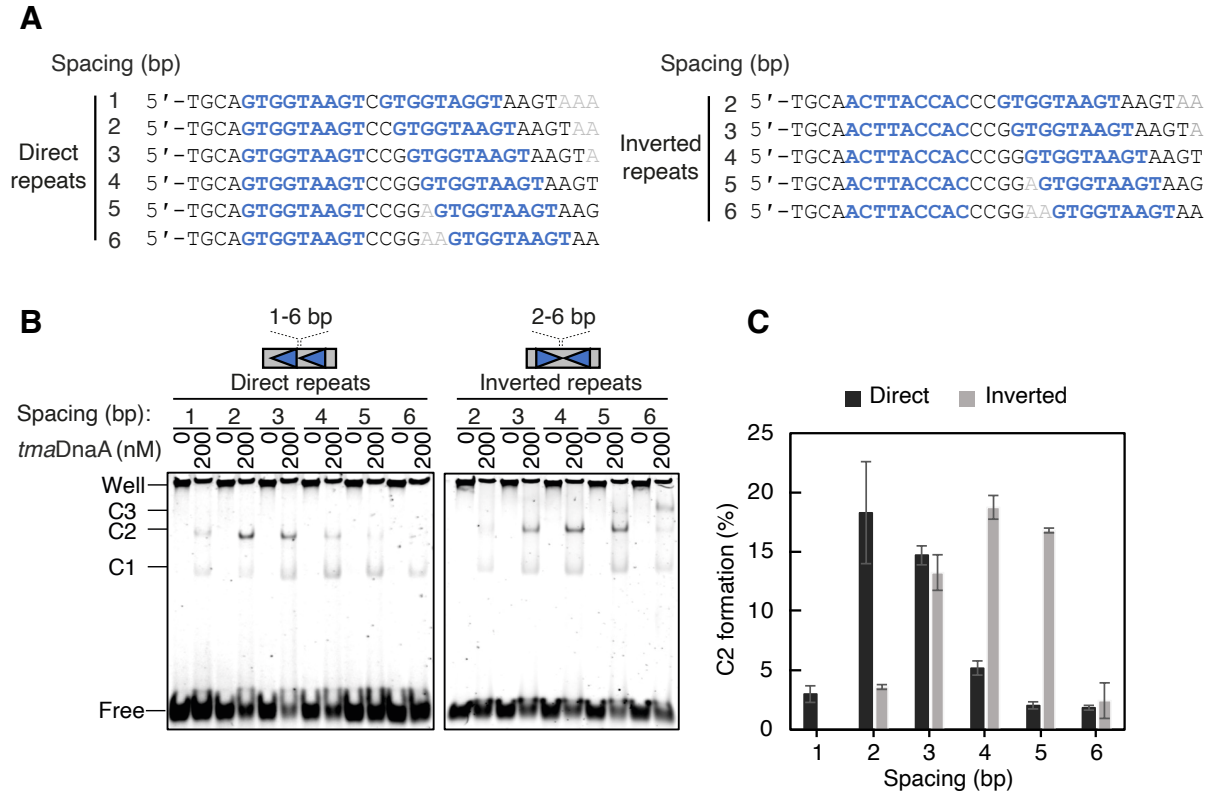


Figure 6-2. Formation of C2 complex was stimulated by appropriate spacing between two *tmaDnaA* boxes and their orientation

(A-C) ATP-*tmaDnaA* (200 nM) was incubated with a 32 bp DNA (100 nM) containing direct or inverted repeats of *tmaDnaA* boxes, with various spaces between boxes. (A) Sequences of the substrate DNA. (B) Visualization of the C2 complexes by 8% polyacrylamide gel electrophoresis and GelStar staining (n=2). (C) Percentages of C2 relative to input DNA.

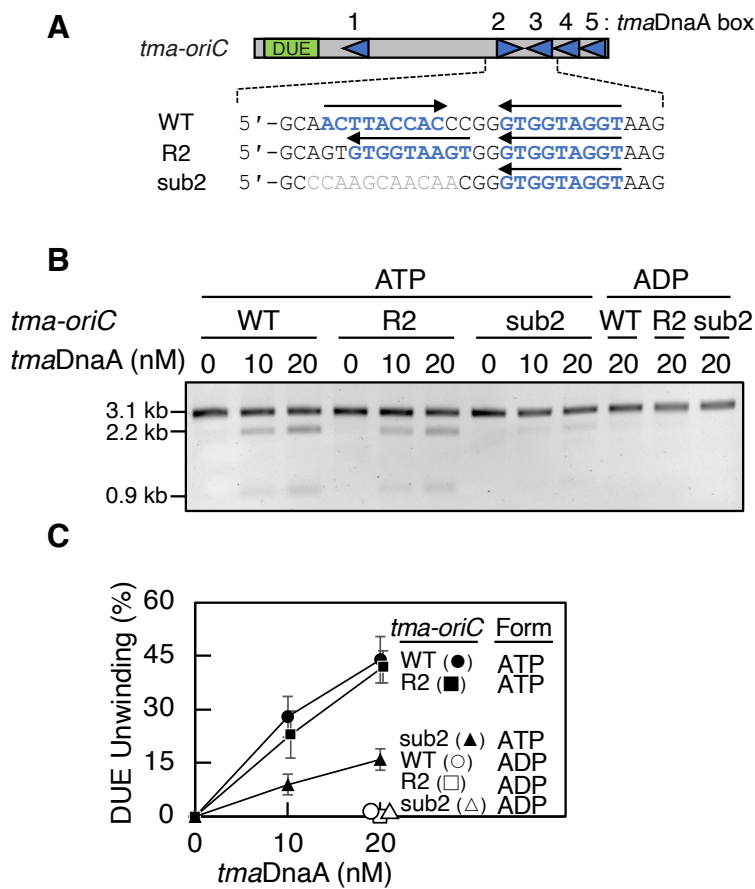


Figure 6-3. The *oriC* mutant variant with reversed *tmaDnaA* box 2 sustained open complex formation.

Open complex formation. The P1 assay was performed as described in the legend to Figure 1-1. The *tma-oriC* mutant derivatives with inverted *tmaDnaA* box 2 (R2) or randomized *tmaDnaA* box 2 (sub2) are illustrated schematically (A). The gel image (B) and mean $\pm$ standard deviation percentages of P1 nuclease-digested DNA ($n = 3$) quantified by FIJI software (C) are shown respectively. Form, ATP- or ADP-forms of *tmaDnaA*.

Discussion

In bacteria, despite the significant sequence homology among the DnaA family of proteins, the origins of bacterial genomes differ substantially in the number of DnaA boxes and their spatial arrangements. Moreover, the unwinding sequences have been insufficiently determined due to the lack of in-depth characterization using *in vitro* reconstituted systems. The present study analyzed *in vitro* reconstituted open complexes of the deep-branching hyperthermophilic bacterium, *T. maritima*, and identified eTrio consisting of three direct repeats of the trinucleotide TAG as a novel functional module within the unwinding sequences (**Figure 7-1**). These *in vitro* reconstituted systems using ATP-*tmaDnaA* and *tma-oriC* provide concrete evidence that single-stranded eTrio interacts directly with ATP-*tmaDnaA* oligomers constructed on *tmaDOR*, thereby facilitating formation of the open complex. Moreover, the full activity of the open complex was dependent on the appropriate phasing between *tmaDUE-tmaDnaA* box 1 and *tmaDnaA* boxes 2–5. These observations strongly suggest that ssDUE recruitment underlies the formation of the tripartite complex consisting of ATP-*tmaDnaA*, *tmaDOR*, and single-stranded eTrio (**Figure 7-2**). Because the right part of *tmaDOR*, spanning *tmaDnaA* boxes 3–5, is responsible for single-stranded DNA binding, and because a single DnaA molecule can be in direct contact with up to three nucleotides (Ozaki & Katayama, 2012; Erzberger *et al*, 2006), each *tmaDnaA* box-bound ATP-*tmaDnaA* protein on the right *tmaDOR* likely senses a single TAG trinucleotide. *tmaDnaA* box 2-bound *tmaDnaA* might enhance this interaction by stimulating conformational changes of the *tma-oriC* complex (also see below).

Scrambling the sequences upstream and downstream of eTrio moderately inhibits duplex DNA unwinding. Although these regions have possible ssDUE-binding motifs, those have activities that are functionally distinct from those of eTrio in the stimulation of DUE unwinding. In papillomavirus, the adenine-thymine tracts flanked by the recognition sites of the initiator protein E2 are crucial for initiation of replication. Although these tracts are not in direct contact with the E2 protein, they display intrinsic DNA bending, thereby facilitating the ability of E2 proteins to recognize their binding

sites (Hizver *et al*, 2001). Similarly, the two AT-rich DNA regions flanking eTrio could contribute to local structural changes by destabilizing the duplex, thereby stimulating efficient unwinding of DUE and subsequent ssDUE binding to *tmaDnaA*.

Taken together, these findings suggest a mechanism for open complex formation in *T. maritima* (**Figure 7-2**). ATP-*tmaDnaA* proteins preferentially form a head-to-tail tetramer on *tmaDnaA* boxes 2–5. The flexible nature of the linker between *tmaDnaA* domains III and IV would allow considerable swiveling of the domains, whereby a *tmaDnaA* protomer bound to the inverted *tmaDnaA* box 2 swivels around its AAA+ domain to bring the Arg-finger close to the ATP on the adjacent protomer bound to *tmaDnaA* box 3. When *tmaDnaA* boxes 1–5 are all occupied by ATP-*tmaDnaA*, the ATP-*tmaDnaA* promoters bound to *tmaDnaA* boxes 1 and 2 would interact transiently as a result of Brownian motion, inducing *tmaDUE* unwinding. The unwound ssDUE would be stabilized through direct interaction of eTrio with the ATP-*tmaDnaA* trimer bound to *tmaDnaA* boxes 3–5. In addition, the interactions between the ATP-*tmaDnaA* protomers bound to *tmaDnaA* boxes 1 and 2 would induce bending of the DNA present in the space between *tmaDnaA* boxes 1 and 2, stimulating ssDUE recruitment and stable unwinding. The site-specific HU binding to this space was recently exemplified using DMS footprint experiments (Yoshida *et al*, 2023). Given HU homologs are ubiquitous in the bacterial kingdom and HU-accessible interspaces between two DnaA boxes are basically present in the predicted bacterial origins (Yoshida *et al*, 2023; Kamashev *et al*, 2017), it is conceivable that the HU-promoted ssDUE recruitment mechanism is prevailing among diverse bacterial species. Besides, the unwound region of *tmaDUE* may further expand over the TTATT and DnaA-trio motifs for stable unwinding and subsequent helicase loading. Expansion of the unwinding regions has been well characterized in open complexes of *E. coli* (Sakiyama *et al*, 2022). This proposed mechanism is also fully consistent with the mechanism underlying ssDUE recruitment.

The observation that single-stranded eTrio is specifically recognized by the DOR-bound *tmaDnaA* complexes also provides evolutionary insight into the functional structures required for eubacterial DUE. The DnaA-trios of *B. subtilis* and *H. pylori* that are crucial for single-stranded DNA binding of their cognate DnaA proteins contain

multiple TAG trinucleotides like eTrio of *tmaDUE*, suggesting that proteins in the DnaA family generally prefer single-stranded TAG repeats (Richardson *et al*, 2016; Jaworski *et al*, 2021). Consistently, the DnaA residues responsible for ssDUE binding are highly conserved in DnaA proteins (Ozaki *et al*, 2008). The molecular mechanisms by which *E. coli* DnaA and *tmaDnaA* proteins recognize different ssDUE motifs remain to be elucidated in future.

Given the overall sequence dissimilarity of the AT-rich regions within the bacterial origins (**Figure 7-3**), it is puzzling how evolutionarily distal bacteria such as *T. maritima*, *B. subtilis*, and *H. pylori* have co-opted TAG for initiation of replication. These bacteria thrive in different environments, leading us to speculate that the selective pressure for the DUE motifs may be determined by the chemical properties of DNA, rather than the environmental conditions. Supporting this hypothesis, the dinucleotide TA has shown flexible base pair morphology, including the ability to roll, tilt, and twist, thereby destabilizing the duplex structure (Young *et al*, 2022; Mack *et al*, 2001; Balaceanu *et al*, 2019). The DnaA proteins may therefore have evolved to bind to the transiently distorted TA, thereby stabilizing the unwound DNA. This is consistent with results showing that the TTATT motif with a central TA dinucleotide, instead of TAG, is present in *E. coli* DUE (Sakiyama *et al*, 2022; Ozaki *et al*, 2008). Intriguingly, TAG repeats have been observed in human telomeric DNA. Flexible TA dinucleotides contribute to pronounced DNA distortions necessary for formation of human telomeric nucleosome core particles (Soman *et al*, 2021).

The finding that inverted repeats of *tmaDnaA* boxes can assist in head-to-tail interactions of ATP-*tmaDnaA* protomers expands the view of tunability in the formation of bacterial initiation complexes at the origin. The predicted origins corresponding to the representative DnaA orthologs contain inverted DnaA boxes (**Figures 0-2 and 0-3**). As in *tmaDnaA*, the central AAA+ domain and the C-terminal DNA-binding domain in most, if not all, of those DnaA orthologs, are structurally predicted to be connected by flexible linkers containing 11–12 amino acid residues, which could engage in head-to-tail DnaA oligomerization on inverted repeats of DnaA boxes to form an initiation complex via the mechanism described in this study. Conversely, the array of direct DnaA box repeats at

the *E. coli* origin represents a head-to-tail oligomerization of ATP-DnaA. Because *E. coli* DnaA carries a relatively short and presumably less flexible linker between the AAA+ and C-terminal DNA-binding domains, this linker likely coevolved with the architecture of the origin to maximize the efficiency of formation of the initiation complex. The *E. coli* chromosome also carries an inverted repeat of the DnaA box motifs within the DnaA-reactivating sequences (*DARS*) 1 and 2, specific chromosomal loci for ADP-DnaA assembly that promote ATP-DnaA production by nucleotide exchange of ADP-DnaA (Katayama *et al*, 2017; Miyoshi *et al*, 2021). This activity is facilitated by a head-to-head interaction of DnaA protomers on the inverted DnaA boxes (Sugiyama *et al*, 2019). Thus, *E. coli* DnaA utilizes the structural constraint of the linker between the AAA+ and C-terminal DNA-binding domains to enhance regulatory systems on DnaA or to expand their repertoire.

In nature, the ssDUE recruitment mechanism has been co-opted, even by non-“DnaA-oriC” replication systems. This is exemplified by a recent study of *V. cholerae* *ori2* where the RctB initiator drives replication initiation (Chatterjee *et al*, 2020). Similarly, the ssDUE recruitment mechanism could underlie initiation of the RK2 plasmid with the TrfA initiator (Wegrzyn *et al*, 2014). In either case, short oligonucleotide repeats are tandemly aligned in their corresponding DUEs, six direct repeats of ATCA in *V. cholerae* *ori2* DUE and four direct repeats of GGTT in RK2 plasmid DUE. These repeated sequences are likely favored by their cognate initiator proteins. Based on the mechanisms of action of different initiator proteins, these origins may have evolved independently from the DnaA-oriC system, thus expanding the mechanisms for ssDUE recruitment.

This study enhances the understanding of the fundamental molecular nature of the origin sequences present in evolutionarily divergent bacteria, including numerous pathogenic bacteria. Consequently, this research has the potential to contribute to the development of novel antibiotics targeting the DNA synthesis of pathogenic bacteria.

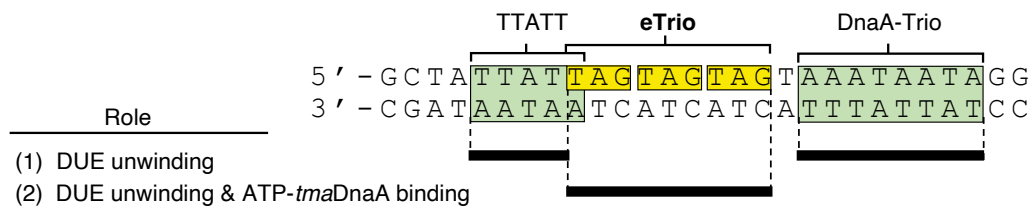


Figure 7-1. The roles of *tmaDUE* sequence motifs

Structure of *tmaDUE*. The three tandem TAG repeats of eTrio are highlighted in yellow, and the TTATT motif and DnaA-trio are highlighted in green. Black bars indicate sites with crucial roles in DUE unwinding and/or ssDUE binding.

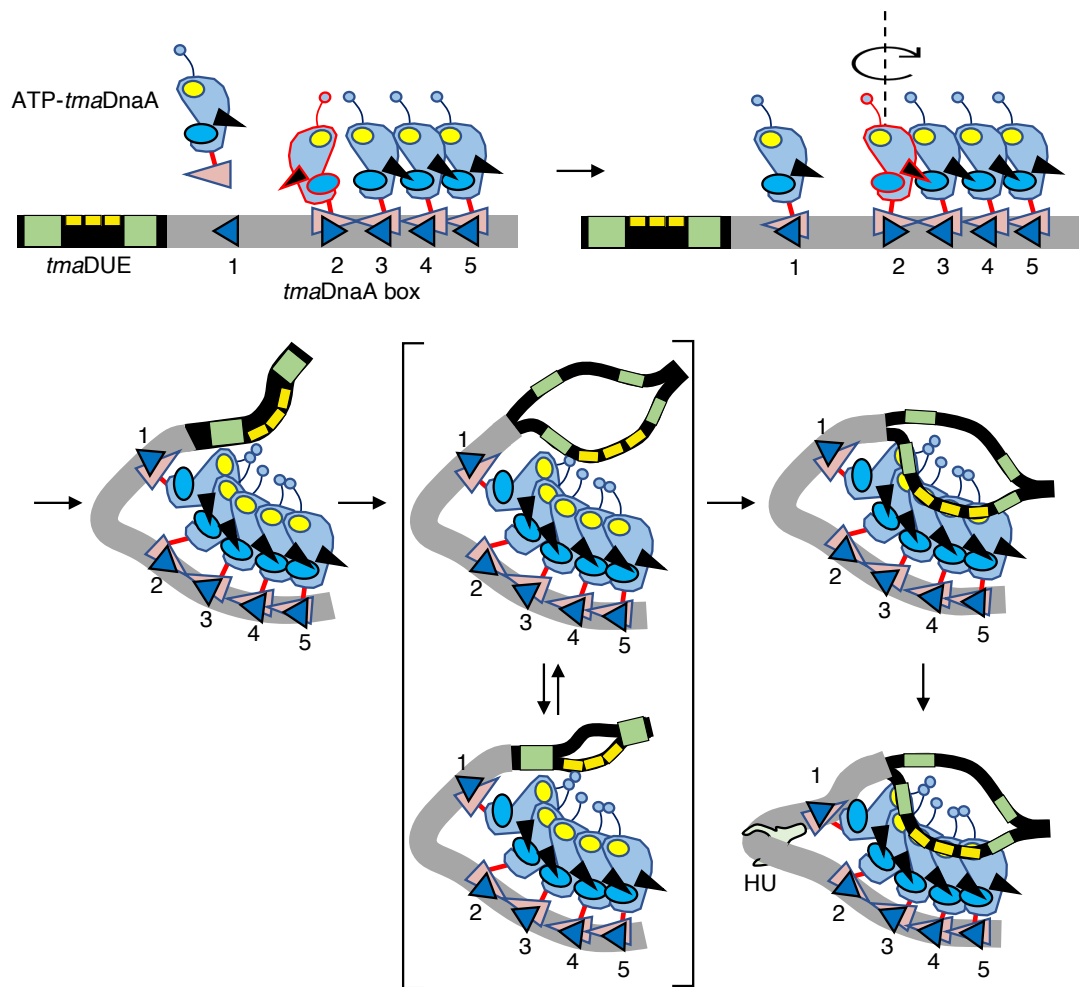


Figure 7-2. Model for ssDUE recruitment mechanism in *T. maritima*

Model for open complex formation. The swiveled ATP-*tmaDnaA* bound to *tmaDnaA* box 2 facilitates formation of ATP-*tmaDnaA* pentamers on *tmaDOR* in a head-to-tail manner, inducing *tmaDUE* unwinding. While TTATT and DnaA-trio engage in efficient DUE unwinding, the single-stranded eTrio directly binds to ATP-*tmaDnaA* trimers bound to *tmaDnaA* boxes 3-5, stabilizing the unwound state. A presumable HU binding to the spacer between *tmaDnaA* boxes 1 and 2 further stabilizes open complex formation.



Figure 7-3. The AT-content profiles of DUEs

The values of AT % for *E. coli*, *B. subtilis*, *H. pylori*, *T. tengcongensis*, and *T. maritima* genomes were plotted in a 10-bp sliding window using the MacVector software. DUE (green bars) and the flanking regions are shown. The x-axis indicates the genomic position of each organism.

Experimental procedures

Strains and proteins

E. coli strain DH5 α was used for cloning. *E. coli* HU, wild-type (WT) *tmaDnaA*, and *tmaDnaA* R251A were purified as described previously (Ozaki *et al*, 2006). Bovine serum albumin (BSA) was purchased from Roche and P1 nuclease from Wako.

Buffers

Buffer E consisted of 60 mM Hepes-KOH (pH 7.6), 130 mM potassium glutamate, 7 mM EDTA, 7.5 mM dithiothreitol, 0.01% Triton X-100, 0.32 mg/mL BSA, 20% glycerol, and 3 μ M ATP. Buffer P consisted of 60 mM Hepes-KOH (pH 7.6), 8 mM magnesium acetate, 0.1 mM zinc acetate, 30% glycerol, 0.32 mg/mL BSA, 100 mM potassium chloride and 5 mM ATP. Buffer G consisted of 20 mM Hepes-KOH (pH 7.6), 1 mM EDTA (pH 8.0), 4 mM dithiothreitol, 5 mM magnesium acetate, 10% (v/v) glycerol, 0.1% Triton X-100, 0.1 mg/mL BSA, and 50 mM ammonium sulfate.

DNA

The plasmids and oligonucleotides used in this study are listed in Tables 1 and 2, respectively.

To construct pOZ14_Trio-A1, pOZ14_Trio-A12, pOZ14_a+5, pOZ14_a+10, pOZ14_b+5, pOZ14_b+10, pOZ14_U, pOZ14_B, pOZ14_Trio, pOZ14_U&Trio, and pOZ14_R2, the inserted DNA fragments were amplified by PCR using pOZ14 and mutagenic primers (929/306 for pOZ14_Trio-A1, 930/306 for pOZ14_Trio-A2, 912/306 for pOZ14_a+5, 913/306 for pOZ14_a+10, 824/823 for pOZ14_b+5, 825/823 for pOZ14_b+10, 799/306 for pOZ14_B, 803/306 for pOZ14_U, 804/306 for pOZ14_Trio, 805/306 for pOZ14_U&Trio, and 1594/823 for pOZ14_R2). The amplified products were digested with BamHI and EcoRI, and ligated into the BamHI and EcoRI sites of pBluescript II (Stratagene).

To construct pTMA R251A, the mutation was introduced into pTHMA-1 using a QuikChange site-directed mutagenesis kit, according to the manufacturer's instructions (Stratagene) (Ozaki *et al*, 2008; Ozaki *et al*, 2012).

P1 nuclease assay

P1 nuclease assays were performed essentially as described Ozaki *et al*, 2008; Ozaki *et al*, 2006). Briefly, ATP-*tmaDnaA* (0–20 nM) and pOZ14 or its mutant derivatives (400 ng; 4 nM) were incubated in buffer P (50 μ L) containing *E. coli* HU (16 ng; 17 nM) for 10 min at 48°C, followed by digestion with 1.5 units P1 nuclease for 200 sec at 48°C. The DNA samples were extracted with phenol/chloroform, precipitated with ethanol, and further digested with the restriction enzyme AlwNI. The products were analyzed using 1% agarose gel electrophoresis and ethidium bromide staining.

P1 assays described in Figure 6-3 were performed using buffer P containing 40 mM ammonium sulfate, instead of 100 mM potassium chloride, which did not affect the specific nuclease activity.

Electrophoretic mobility shift assay

Electrophoretic mobility shift assays (EMSA) were performed essentially as described (Sakiyama *et al*, 2022; Ozaki & Katayama, 2012). Briefly, ATP-*tmaDnaA* (0–80 nM) and a 203 bp DNA fragment containing *tmaDOR* (30 nM) were incubated for 5 min at 48°C, followed by electrophoresis on 4% polyacrylamide gels and visualization of *tmaDOR* using GelStar dye (Lonza). To analyze the formation of ss-*tmaDUE*-ATP-*tmaDnaA*-*tmaDOR* complexes, ATP-*tmaDnaA* (0–80 nM) and the 203 bp DNA fragment containing *tmaDOR* (30 nM) were incubated for 5 min at 48°C; following the addition of ³²P-labeled ss-*tmaDUE* (2.5 nM), the reaction mixtures were incubated for an additional 5 min at 48°C. The mixtures were subjected to 4% polyacrylamide gel electrophoresis, with *tmaDOR* and ³²P-labeled ss-*tmaDUE* visualized using GelStar dye and Typhoon FLA 9500, respectively.

Filter retention assay

Filter retention assays were performed as described previously (40). Briefly, WT *tmaDnaA* or *tmaDnaA* R251A (1.9 pmol) was pre-incubated for 5 min at 38 °C in buffer (25 µL) containing 50 mM HEPES-KOH (pH 7.6), 0.3 mM EDTA, 7 mM dithiothreitol, 20% glycerol, 0.007% Triton X-100, and [α -³²P] ATP or [³H] ADP. Following addition of 5 mM magnesium acetate, the samples were further incubated on ice for 15 min and filtered through nitrocellulose membranes. The retained radioactivity was quantified using a liquid scintillation counter.

Table 1 Plasmids used in this study

Plasmid	Description	Reference
pOZ14	A 3.1 kb pBluescript II derivative bearing a 149 bp minimal <i>tma-oriC</i>	(Ozaki <i>et al</i> , 2006)
pOZ14_Trio-A1	A pOZ14 derivative with modified DUE	This study
pOZ14_Trio-A2	A pOZ14 derivative with modified DUE	This study
pOZ14_U	A pOZ14 derivative with modified DUE	This study
pOZ14_B	A pOZ14 derivative with modified DUE	This study
pOZ14_Trio	A pOZ14 derivative with modified DUE	This study
pOZ14_U&Trio	A pOZ14 derivative with modified DUE	This study
pOZsub1	A pOZ14 derivative in which the sequence of <i>tmaDnaA</i> box 1 is randomized	(Ozaki & Katayama, 2012)
pOZsub2	A pOZ14 derivative in which the sequence of <i>tmaDnaA</i> box 2 is randomized	(Ozaki & Katayama, 2012)
pOZsub3	A pOZ14 derivative in which the sequence of <i>tmaDnaA</i> box 3 is randomized	(Ozaki & Katayama, 2012)
pOZsub4	A pOZ14 derivative in which the sequence of <i>tmaDnaA</i> box 4 is randomized	(Ozaki & Katayama, 2012)
pOZsub5	A pOZ14 derivative in which the sequence of <i>tmaDnaA</i> box 5 is randomized	(Ozaki & Katayama, 2012)
pOZ14_a+5	A pOZ14 derivative with a 5 bp insert between <i>tmaDnaA</i> boxes 1 and 2	This study
pOZ14_a+10	A pOZ14 derivative with a 10 bp insert between <i>tmaDnaA</i> boxes 1 and 2	This study
pOZ14_b+5	A pOZ14 derivative with a 5 bp insert between <i>tmaDnaA</i> boxes 1 and 2	This study
pOZ14_b+10	A pOZ14 derivative with a 10 bp insert between <i>tmaDnaA</i> boxes 1 and 2	This study

pOZ14_R2	A pOZ14 derivative in which the sequence of <i>tmaDnaA</i> box 2 is reversed	This study
pTHMA-1	A plasmid for purification of wild-type <i>tmaDnaA</i>	(Ozaki <i>et al</i> , 2006)
pTMA R251A	A pTHMA-1 derivative carrying the <i>tmaDnaA</i> R251A allele	This study

Table 2 Oligonucleotides used in this study

Oligonucleotide	Sequence (5'-3')	Reference
305_PAT3	GGGGACGCAGGTGGTAGGTTTC	(Ozaki <i>et al</i> , 2006)
306_PBS2	GTAATACGACTCACTATAGGGCGA	(Ozaki <i>et al</i> , 2006)
614_EMSA	CCCCCCCCCTAGCCCCCCCCCCCCCCCC	This study
615_EMSA	CCCCCCCCCTAGTAGCCCCCCCCCCCCCCCC	This study
799_B	AGTGGATCCTGCTATTATATCATCATCTAAATA ATAGGGGACGCAGGTGGTA	This study
803_U	AGTGGATCCTGCATAATATAGTAGTAGTAAATA ATAGGGGACGCAGGTGGTA	This study
804_Trio	AGTGGATCCTGCTATTATTAGTAGTAGATTTAT TATGGGGACGCAGGTGGTAGGTTTCGAG	This study
805_U&Trio	AGTGGATCCTGCATAATATAGTAGTAGATTTAT TATGGGGACGCAGGTGGTAGGTTTCGAG	This study
823_PBS1	AATTAACCCTCACTAAAGGGAAC	(Ozaki <i>et al</i> , 2006)
824_b+5	ATCGAATTCAAACCTACCACTTACCTACCACTT ACCTACCACCCGGGTGGTAAGTTGCAGACATG CTTTCATTAAAAGCCTCATCAGAAGCCT	This study
825_b+10	ATCGAATTCAAACCTACCACTTACCTACCACTT ACCTACCACCCGGGTGGTAAGTTGCAGACAATT GTTGCTTTTCATTAAAAGCCTCATCAGAAGCCT	This study
887_EMSA	CCCCCCCCCTAGTAGTAGCCCCCCCCCCCC	This study
888_EMSA	CCCCCCCCCCCCCCCCCTAAATAATACC	This study
889_EMSA	CCCCCCCCCTAGTAGTAGTAACCCCCCCCC	This study
890_EMSA	CCCCCCCCCTAGTAGTAGTAAATACCCCC	This study

912_a+5	AGTGGATCCTGCTATTATTAGTAGTAGTAAATA ATAGGGGACGCAGGTGGTAGGTTTCGAGTATT AAGCAGAAATTTTTGAACAGGCTTCTGATG	This study
913_a+10	AGTGGATCCTGCTATTATTAGTAGTAGTAAATA ATAGGGGACGCAGGTGGTAGGTTTCGAGTATT AAGCAACAATGAAATTTTTGAACAGGCTTCTGA TG	This study
929_Trio-A1	AGTGGATCCTGCTATTATTAGTAGTAGTAAATT ATAGGGGACGCAGGTGGTAGGTTTC	This study
930_Trio-A2	AGTGGATCCTGCTATTATTAGTAGTAGTAAATA ATTGGGGACGCAGGTGGTAGGTTTCGAG	This study
1031-Consense-1	AGAAAACCTACCACCTAA	(Ozaki <i>et al</i> , 2006)
1032-Consense-2	TTAGGTGGTAGGTTTTCT	(Ozaki <i>et al</i> , 2006)
1033-Nonsense-1	AGACCCAAGCAACAATAA	(Ozaki <i>et al</i> , 2006)
1034-Nonsense-2	TTATTGTTGCTTGGGTCT	(Ozaki <i>et al</i> , 2006)
1035-9mer-1	AGACCACCTACCACATAA	This study
1036-9mer-2	TTATGTGGTAGGTGGTCT	This study
1037-8mer-L-1	AGACCACCTACCAAATAA	This study
1038-8mer-L-2	TTATTTGGTAGGTGGTCT	This study
1039-8mer-R-1	AGACCCCCTACCACATAA	This study
1040-8mer-R-2	TTATGTGGTAGGGGGTCT	This study
1052- DnaABox2+3_1	TGCAACTTACCACCCGGGTGGTGGTAAAGT	This study

1053- DnaABox2+3_2	ACTTTACCACCACCCGGGTGGTAAGTTGCA	This study
1054- DnaABox2+3N_1	TGCAACTTACCACCCGGTGTTGTTGCAAGT	This study
1055- DnaABox2+3N_2	ACTTGCAACAACACCGGGTGGTAAGTTGCA	This study
1056- DnaABox2N+3N _1	TGCACAGGCAACACCGGTGTTGTTGCAAGT	This study
1057- DnaABox2N+3N _2	ACTTGCAACAACACCGGTGTTGCCTGTGCA	This study
1306_DnaABox 2+3(2)_1	TGCAACTTACCACCCGTGGTAGGTAAGTGGAA	This study
1307_DnaABox 2+3(2)_2	TTCCACTTACCTACCACGGGTGGTAAGTTGCA	This study
1308_DnaABox 2+3(3)_1	TGCAACTTACCACCCGGTGGTAGGTAAGTGGAA	This study
1309_DnaABox 2+3(3)_2	TCCACTTACCTACCACCGGGTGGTAAGTTGCA	This study
1310_DnaABox 2+3(4)_1	TGCAACTTACCACCCGGGTGGTAGGTAAGTGG	This study
1311_DnaABox 2+3(4)_2	CCACTTACCTACCACCCGGGTGGTAAGTTGCA	This study
1316_DnaABox 2R+3(1)_1	TGCAGTGGTAAGTCGTGGTAGGTAAGTGGAAA	This study
1317_DnaABox 2R+3(1)_2	TTCCACTTACCTACCACGACTTACCACTGCA	This study

1318_DnaABox 2R+3(2)_1	TGCAGTGGTAAGTCCGTGGTAGGTAAGTGGA	This study
1319_DnaABox 2R+3(2)_2	TTCCACTTACCTACCACGGACTTACCACTGCA	This study
1320_DnaABox 2R+3(3)_1	TGCAGTGGTAAGTCCGGTGGTAGGTAAGTGGA	This study
1321_DnaABox 2R+3(3)_2	TCCACTTACCTACCACCGGACTTACCACTGCA	This study
1322_DnaABox 2R+3(4)_1	TGCAGTGGTAAGTCCGGGTGGTAGGTAAGTGG	This study
1323_DnaABox 2R+3(4)_2	CCACTTACCTACCACCCGGACTTACCACTGCA	This study
1338_2+3(+5)_1	TGCAACTTACCACCCAGGGTGGTAGGTAAGTG	This study
1339_2+3(+5)_2	CACTTACCTACCACCCTGGGTGGTAAGTTGCA	This study
1340_2+3(+6)_1	TGCAACTTACCACCCAAGGGTGGTAGGTAAGT	This study
1341_2+3(+6)_2	ACTTACCTACCACCCTTGGGTGGTAAGTTGCA	This study
1342_2R+3(+5)_ 1	TGCAGTGGTAAGTCCAGGGTGGTAGGTAAGTG	This study
1343_2R+3(+5)_ 2	CACTTACCTACCACCCTGGACTTACCACTGCA	This study
1344_2R+3(+6)_ 1	TGCAGTGGTAAGTCCAAGGGTGGTAGGTAAGT	This study
1345_2R+3(+6)_ 2	ACTTACCTACCACCCTTGGACTTACCACTGCA	This study
1594_2R_R	ATATCGAATTCAAACCTACCACTTACCTACCAC TTACCTACCACCCACTTACCACACTGCAGAC	This study
TMA28	GCTATTATTAGTAGTAGTAAATAATAGG	(Ozaki & Katayama, 2012)

References

- Abe Y, Jo T, Matsuda Y, Matsunaga C, Katayama T & Ueda T (2007) Structure and Function of DnaA N-terminal Domains. *J. Biol. Chem.* 282: 17816–17827
- Balaceanu A, Buitrago D, Walther J, Hospital A, Dans PD & Orozco M (2019) Modulation of the helical properties of DNA: next-to-nearest neighbour effects and beyond. *Nucleic Acids Res.* 47: 4418–4430
- Chatterjee S, Jha JK, Ciaccia P, Venkova T & Chatteraj DK (2020) Interactions of replication initiator RctB with single- and double-stranded DNA in origin opening of *Vibrio cholerae* chromosome 2. *Nucleic Acids Res.* 48: 11016–11029
- Chodavarapu S, Felczak MM, Yaniv JR & Kaguni JM (2007) *Escherichia coli* DnaA interacts with HU in initiation at the *E. coli* replication origin. *Mol. Microbiol.* 67: 781–792
- Costa A, Hood I V. & Berger JM (2013) Mechanisms for Initiating Cellular DNA Replication. *Annu. Rev. Biochem.* 82: 25–54
- Dong M. J., Luo H & Gao F (2023) DoriC 12.0: an updated database of replication origins in both complete and draft prokaryotic genomes. *Nucleic Acids Res.* 51: D117–D120
- Erzberger JP, Mott ML & Berger JM (2006) Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* 13: 676–683
- Felczak MM & Kaguni JM (2004) The Box VII Motif of *Escherichia coli* DnaA Protein Is Required for DnaA Oligomerization at the *E. coli* Replication Origin. *J. Biol. Chem.* 279: 51156–51162
- Felczak MM, Simmons LA & Kaguni JM (2005) An Essential Tryptophan of *Escherichia coli* DnaA Protein Functions in Oligomerization at the *E. coli* Replication Origin. *J. Biol. Chem.* 280: 24627–24633
- Fujikawa N (2003) Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res.* 31: 2077–2086
- Grimwade JE & Leonard AC (2021) Blocking, Bending, and Binding: Regulation of Initiation of Chromosome Replication During the *Escherichia coli* Cell Cycle by

- Transcriptional Modulators That Interact With Origin DNA. *Front. Microbiol.* 12: 732270
- Hayashi C, Miyazaki E, Ozaki S, Abe Y & Katayama T (2020) DnaB helicase is recruited to the replication initiation complex via binding of DnaA domain I to the lateral surface of the DnaB N-terminal domain. *J. Biol. Chem.* 295: 11131–11143
- Hizver J, Rozenberg H, Frolow F, Rabinovich D & Shakked Z (2001) DNA bending by an adenine–thymine tract and its role in gene regulation. *Proc. Natl. Acad. Sci. U.S.A.* 98: 8490–8495
- Hwang DS & Kornberg A (1992a) Opposed actions of regulatory proteins, DnaA and IciA, in opening the replication origin of *Escherichia coli*. *J. Biol. Chem.* 267: 23087–91
- Hwang DS & Kornberg A (1992b) Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. *J. Biol. Chem.* 267: 23083–6
- Jaworski P, Donczew R, Mielke T, Weigel C, Stingl K & Zawilak-Pawlik A (2018) Structure and Function of the *Campylobacter jejuni* Chromosome Replication Origin. *Front. Microbiol.* 9: 1–18
- Jaworski P, Zyla-Uklejewicz D, Nowaczyk-Cieszevska M, Donczew R, Mielke T, Weigel C & Zawilak-Pawlik A (2021) Putative Cooperative ATP–DnaA Binding to Double-Stranded DnaA Box and Single-Stranded DnaA-Trio Motif upon *Helicobacter pylori* Replication Initiation Complex Assembly. *Int. J. Mol. Sci.* 22: 6643
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, *et al* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589
- Kaguni JM (2011) Replication initiation at the *Escherichia coli* chromosomal origin. *Curr. Opin. Chem. Biol.* 15: 606–613
- Kamashev D, Agapova Y, Rastorguev S, Talyzina AA, Boyko KM, Korzhenevskiy DA, Vlaskina A, Vasilov R, Timofeev VI & Rakitina T V. (2017) Comparison of histone-like HU protein DNA-binding properties and HU/IHF protein sequence alignment. *PLoS One* 12: e0188037

- Katayama T, Kasho K & Kawakami H (2017) The DnaA Cycle in *Escherichia coli*: Activation, Function and Inactivation of the Initiator Protein. *Front. Microbiol.* 8: 2496
- Kawakami H, Keyamura K & Katayama T (2005) Formation of an ATP-DnaA-specific Initiation Complex Requires DnaA Arginine 285, a Conserved Motif in the AAA+ Protein Family. *J. Biol. Chem.* 280: 27420–27430
- Keyamura K, Abe Y, Higashi M, Ueda T & Katayama T (2009) DiaA Dynamics Are Coupled with Changes in Initial Origin Complexes Leading to Helicase Loading. *J. Biol. Chem.* 284: 25038–25050
- Kumar S, Farhana A & Hasnain SE (2009) In-Vitro Helix Opening of *M. tuberculosis oriC* by DnaA Occurs at Precise Location and Is Inhibited by IciA Like Protein. *PloS One* 4: e4139
- Kuwabara T & Igarashi K (2020) *Thermotogales* origin scenario of eukaryogenesis. *J. Theor. Biol.* 492: 110192
- Lopez P, Forterre P, le Guyader H & Philippe H (2000) Origin of replication of *Thermotoga maritima*. *Trends Genetics* 16: 59–60
- Mack DR, Chiu TK & Dickerson RE (2001) Intrinsic bending and deformability at the T-A step of CCTTTAAAGG: a comparative analysis of T-A and A-T steps within A-tracts. *J. Biol. Chem.* 312: 1037–49
- Majka J, Zakrzewska-Czerwińska J & Messer W (2001) Sequence Recognition, Cooperative Interaction, and Dimerization of the Initiator Protein DnaA of *Streptomyces*. *J. Biol. Chem.* 276: 6243–6252
- McGarry KC, Ryan VT, Grimwade JE & Leonard AC (2004) Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc. Natl. Acad. Sci. U.S.A.* 101: 2811–2816
- Miyoshi K, Tatsumoto Y, Ozaki S & Katayama T (2021) Negative feedback for *DARS2*–Fis complex by ATP–DnaA supports the cell cycle-coordinated regulation for chromosome replication. *Nucleic Acids Res.* 49: 12820–12835
- Noguchi Y, Sakiyama Y, Kawakami H & Katayama T (2015) 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.* 290: 20295–20312
- Nozaki S & Ogawa T (2008) Determination of the minimum domain II size of *Escherichia coli* DnaA protein essential for cell viability. *Microbiology* 154: 3379–3384
- Ozaki S (2019) Regulation of replication initiation: lessons from *Caulobacter crescentus* *Genes Genet. Syst.* 94: 183–196
- Ozaki S, Fujimitsu K, Kurumizaka H & Katayama T (2006) 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* 11: 425–438
- Ozaki S & Katayama T (2009) DnaA structure, function, and dynamics in the initiation at the chromosomal origin. *Plasmid* 62: 71–82
- Ozaki S & Katayama T (2012) Highly organized DnaA–*oriC* complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.* 40: 1648–1665
- Ozaki S, Kawakami H, Nakamura K, Fujikawa N, Kagawa W, Park S-Y, Yokoyama S, Kurumizaka H & Katayama T (2008) A Common Mechanism for the ATP-DnaA-dependent Formation of Open Complexes at the Replication Origin. *J. Biol. Chem.* 283: 8351–8362
- Ozaki S, Noguchi Y, Hayashi Y, Miyazaki E & Katayama T (2012) Differentiation of the DnaA-*oriC* Subcomplex for DNA Unwinding in a Replication Initiation Complex. *J. Biol. Chem.* 287: 37458–37471
- Pei H, Liu J, Li J, Guo A, Zhou J & Xiang H (2007) 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.* 35: 3087–3099
- Richardson TT, Harran O & Murray H (2016) The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. *Nature* 534: 412–416

- Richardson TT, Stevens D, Pellicciari S, Harran O, Sperlea T & Murray H (2019)
Identification of a basal system for unwinding a bacterial chromosome origin.
EMBO J. 38: e101649
- Rozgaja TA, Grimwade JE, Iqbal M, Czerwonka C, Vora M & Leonard AC (2011) Two
oppositely oriented arrays of low-affinity recognition sites in *oriC* guide
progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol.*
Microbiol. 82: 475–488
- Ryan VT, Grimwade JE, Nievera CJ & Leonard AC (2002) IHF and HU stimulate
assembly of pre-replication complexes at *Escherichia coli oriC* by two different
mechanisms. *Mol. Microbiol.* 46: 113–124
- Sakiyama Y, Kasho K, Noguchi Y, Kawakami H & Katayama T (2017) Regulatory
dynamics in the ternary DnaA complex for initiation of chromosomal replication in
Escherichia coli. *Nucleic Acids Res.* 45: 12354–12373
- Sakiyama Y, Nagata M, Yoshida R, Kasho K, Ozaki S & Katayama T (2022) Concerted
actions of DnaA complexes with DNA-unwinding sequences within and flanking
replication origin *oriC* promote DnaB helicase loading. *J. Biol. Chem.* 298: 102051
- Schaper S & Messer W (1995)
Interaction of the Initiator Protein DnaA of *Escherichia coli* with Its DNA Target.
J. Biol. Chem. 270: 17622–17626
- Shimizu M, Noguchi Y, Sakiyama Y, Kawakami H, Katayama T & Takada S (2016)
Near-atomic structural model for bacterial DNA replication initiation complex and
its functional insights. *Proc. Natl. Acad. Sci. U.S.A.* 113: E8021–E8030
- Simmons LA, Felczak M & Kaguni JM (2003) DnaA Protein of *Escherichia coli*:
oligomerization at the *E. coli* chromosomal origin is required for initiation and
involves specific N-terminal amino acids. *Mol. Microbiol.* 49: 849–858
- Soman A, Liew CW, Teo HL, Berezhnoy N V, Olieric V, Korolev N, Rhodes D &
Nordenskiöld L (2021) The human telomeric nucleosome displays distinct
structural and dynamic properties. *Nucleic Acids Res.* 48: 5383–5396

- Sugiyama R, Kasho K, Miyoshi K, Ozaki S, Kagawa W, Kurumizaka H & Katayama T (2019) A novel mode of DnaA-DnaA interaction promotes ADP dissociation for reactivation of replication initiation activity. *Nucleic Acids Res.* 47: 11209–11224
- Sutton MD, Carr KM, Vicente M & Kaguni JM (1998) *Escherichia coli* DnaA protein. The N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal origin. *J. Biol. Chem.* 273: 34255–62
- Swinger KK & Rice PA (2004) IHF and HU: flexible architects of bent DNA. *Curr. Opin. Struct. Biol.* 14: 28–35
- Wegrzyn K, Fuentes-Perez ME, Bury K, Rajewska M, Moreno-Herrero F & Konieczny I (2014) Sequence-specific interactions of Rep proteins with ssDNA in the AT-rich region of the plasmid replication origin. *Nucleic Acids Res.* 42: 7807–7818
- Wolański M, Donczew R, Zawilak-Pawlik A & Zakrzewska-Czerwińska J (2015) *oriC*-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5: 735
- Yoshida R, Ozaki S, Kawakami H & Katayama T (2023) Single-stranded DNA recruitment mechanism in replication origin unwinding by DnaA initiator protein and HU, an evolutionary ubiquitous nucleoid protein. *Nucleic Acids Res.*
- Young RT, Czapla L, Wefers ZO, Cohen BM & Olson WK (2022) Revisiting DNA Sequence-Dependent Deformability in High-Resolution Structures: Effects of Flanking Base Pairs on Dinucleotide Morphology and Global Chain Configuration. *Life* 12: 759
- Zhaxybayeva O, Swithers KS, Lapierre P, Fournier GP, Bickhart DM, DeBoy RT, Nelson KE, Nesbø CL, Doolittle WF, Gogarten JP, *et al* (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the *Thermotogales*. *Proc. Natl. Acad. Sci. U.S.A.* 106: 5865–5870

Publication list

Lu, C., Yoshida, R., Katayama, T., & Ozaki, S. (2023). *Thermotoga maritima oriC* involves a DNA unwinding element with distinct modules and a DnaA-oligomerizing region with a novel directional binding mode. *J. Biol. Chem.*, 104888. In Press.

Acknowledgement

This study was performed in the laboratory of Molecular Biology under supervision of Assoc. Prof. Shogo Ozaki and Prof. Tsutomu Katayama. I thank Kenya, M., Ryusei, Y. and Yasutaka, W. for their valid suggestions in this study.