Whole-Genome Analysis Reveals That the Nucleoid Protein IHF Predominantly Binds to the Replication Origin oriC Specifically at the Time of Initiation

Kazutoshi Kasho*, Taku Oshima*, Onuma Chumsakul, Kensuke Nakamura, Kazuki Fukamachi and Tsutomu Katayama

1 Department of Molecular Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan, 2 Department of Biotechnology, Toyama Prefectural University, Toyama, Japan, 3 Graduate School of Biological Sciences, Nara Institute of Science and Technology, Nara, Japan, 4 Department of Life Science and Informatics, Maebashi Institute of Technology, Maebashi, Japan

The structure and function of bacterial chromosomes are dynamically regulated by a wide variety of nucleoid-associated proteins (NAPs) and DNA superstructures, such as DNA supercoiling. In Escherichia coli, integration host factor (IHF), a NAP, binds to specific transcription promoters and regulatory DNA elements of DNA replication such as the replication origin oriC: binding to these elements depends on the cell cycle but underlying mechanisms are unknown. In this study, we combined GeF-seq (genomic footprinting with high-throughput sequencing) with synchronization of the E. coli cell cycle to determine the genome-wide, cell cycle-dependent binding of IHF with base-pair resolution. The GeF-seq results in this study were qualified enough to analyze genomic IHF binding sites (e.g., oriC and the transcriptional promoters of ilvG and osmY) except some of the known sites. Unexpectedly, we found that before replication initiation, oriC was a predominant site for stable IHF binding, whereas all other loci exhibited reduced IHF binding. To reveal the specific mechanism of stable oriC–IHF binding, we inserted a truncated oriC sequence in the terC (replication terminus) locus of the genome. Before replication initiation, stable IHF binding was detected even at this additional oriC site, dependent on the specific DnaA-binding sequence DnaA box R1 within the site. DnaA oligomers formed on oriC might protect the oriC–IHF complex from IHF dissociation. After replication initiation, IHF rapidly dissociated from oriC, and IHF binding to other sites was sustained or stimulated. In addition, we identified a novel locus associated with cell cycle-dependent IHF binding. These findings provide mechanistic insight into IHF binding and dissociation in the genome.

Keywords: Escherichia coli, cell cycle, IHF, GeF-seq, oriC
INTRODUCTION

Within bacterial cells, chromosomal DNA forms a dynamic and highly condensed structure called the nucleoid. In *Escherichia coli*, the nucleoid is organized by a wide variety of nucleoid-associated proteins (NAPs), RNA, and DNA supercoiling. Major bacterial NAPs such as integration host factor (IHF), heat unstable (HU), and factor for inversion stimulation (Fis), along with DNA supercoiling, regulate various cellular events such as DNA replication, transcription, recombination, and nucleoid condensation (Dillon and Dorman, 2010; Seah et al., 2014). HU, one of the most abundant NAPs, binds to AT-rich DNA without sequence specificity (Ali Azam et al., 1999; Dillon and Dorman, 2010). IHF, a hetero-dimeric protein that consists of α and β subunits (encoded by *ihfA* and *ihfB*, respectively), binds to DNA in a sequence-specific manner and causes sharp (≥120°) bending (Rice et al., 1996; Aeling et al., 2006). Fis, which is expressed specifically in log-phase cells (Ali Azam et al., 1999; Dillon and Dorman, 2010), also binds DNA in a sequence-specific manner, but bends DNA more moderately (≈60°) (Stella et al., 2010). The chromosomal DNA is not randomly condensed but instead dynamically forms specific structures during DNA replication and segregation (Toro and Shapiro, 2010); however, the regulatory mechanism underlying these structural changes remains unclear.

Initiation of chromosomal DNA replication is rigidly controlled to ensure that it occurs only once during the cell cycle (Katayama et al., 2010, 2017; Skarstad and Katayama, 2013; Riber et al., 2016). The initiator protein DnaA and IHF play crucial roles in initiating replication at the chromosomal origin, oriC (Figures 1A–C). DnaA forms complexes with ATP or ADP, and ATP–DnaA is active in initiation (Shimizu et al., 2016; Sakiyama et al., 2017). oriC contains an AT-rich duplex unwinding element (DUE), a single IHF-binding sequence (IBS), and a DnaA oligomerization region (DOR) that contains two subregions with oppositely oriented clusters of DnaA-binding sites (DnaA boxes; Figure 1B; Ozaki et al., 2012; Noguchi et al., 2015). The high-affinity DnaA boxes R1 and R4 at the outer edges of the DOR are oriented in opposite directions, and cooperative binding of ATP–DnaA molecules to lower affinity sites in the Left and Right DORS results in the formation of two DnaA subcomplexes (Figure 1C). IHF specifically binds to the 13-mer IBS and bends DNA sharply (Shimizu et al., 2016; Sakiyama et al., 2017). At initiation, binding of IHF and ATP–DnaA molecules to oriC induces a conformational change and local unwinding of the oriC DUE (Figure 1C), followed by loading of the DNA replication machinery.

The level of ATP–DnaA in the cell is tightly regulated, peaking at the stage of replication initiation (Kurokawa et al., 1999). During replication, DnaA-bound ATP is hydrolyzed by a complex containing Hda and the DNA-loaded clamp subunit of DNA polymerase III holoenzyme, yielding initiation-inactive ATP–DnaA (Kato and Katayama, 2001; Katayama et al., 2010). This replication-coupled negative feedback system is termed regulatory inactivation of DnaA (RIDA). In addition, the specific chromosomal locus *datA*, which contains four DnaA boxes and an IBS, is required to prevent untimely initiations (Figures 1A,B; Nozaki et al., 2009). Recently, we showed that ATP–DnaA molecules form specific complexes on IHF-bound *datA*, which stimulates hydrolysis of DnaA-bound ATP (Kasho and Katayama, 2013; Kasho et al., 2017). *datA*–IHF binding specifically occurs at the post-initiation stage of the cell cycle. This system for timely inactivation of DnaA, termed DDAH (*data*-dependent DnaA–ATP hydrolysis), plays a supplemental role to RIDA in timely yielding of ATP–DnaA.

In contrast to *datA*, the DnaA-binding chromosomal loci called DARSs (DnaA-reactivating sequences) increase the level of ATP–DnaA by promoting nucleotide exchange of ADP–DnaA (Figures 1A,B; Fujimitsu et al., 2009). The *E. coli* chromosome contains at least two DARSs, *DARS1*, and *DARS2*, which are required for timely initiation of replication during the cell cycle. *DARS1* and *DARS2* share three highly conserved “Core” DnaA boxes that are necessary for nucleotide exchange (Fujimitsu et al., 2009; Sugiyma et al., 2019). In contrast to *DARS1*, *DARS2* requires two activator proteins, IHF and Fis (Kasho et al., 2014). *DARS2*–IHF binding is temporally regulated to occur at the pre-initiation stage of the cell cycle. Thus, IHF bindings to *oriC*, *datA*, and *DARS2* are regulated such that they occur at different times during the cell cycle; however, the regulatory mechanism and the genome-wide dynamics of timely IHF binding during the cell cycle remain to be elucidated.

In this study, we utilized GeF-seq, a unique method for identifying protein-binding sites with base-pair resolution (Chumsakul et al., 2013), to identify genome-wide distribution of IHF-binding sites. Based on results obtained by combining GeF-qPCR, GeF-seq and cell cycle synchronization, we identified novel cell cycle-coordinated IHF dynamics: at the replication initiation stage, IHF specifically binds to *oriC* and dissociates from other genomic IHF-binding loci, whereas after replication initiation, IHF is dissociated from *oriC*, and many IHF molecules stably bind to other binding loci. We analyzed IHF-binding consensus sequences at each cell cycle stage and suggest that IHF can temporarily bind to a secondary IBS on *oriC* at the stage of replication initiation. Further mechanistic analysis of *oriC* revealed that the presence of DnaA box R1, but not the chromosomal location of *oriC*, was required for stable *oriC*–IHF binding at the initiation stage. In addition, we comprehensively analyzed genomic IBS and found novel binding loci in the *ttcA* gene that are likely cell cycle-specific. Based on these findings, we propose a model of the specific mechanism involved in stable IHF binding at *oriC* and hypothesize that the modes of genomic IHF binding drastically change during the cell cycle, potentially having a global effect on the dynamics of nucleoid structures and functions.

MATERIALS AND METHODS

Bacterial Strains and Cultures

For GeF-seq, *E. coli* SH022 (*dnaC2 ihfA-cHis12*) cells were used (Table 1; Kasho et al., 2014; Inoue et al., 2016). To introduce *oriCΔDUE* at the *terC* locus, a pBR322-based plasmid...
FIGURE 1: Role of IHF binding at oriC. (A) The locations of oriC, datA, DARS1/2, and TerC loci are indicated on the chromosome. (B) Schematic representation of the structures of oriC, datA, and DARS1/2 are described. Black arrowheads represent DnaA-binding sites (DnaA boxes) that match the 9-mer consensus sequence completely or with only a single mismatch. Gray arrowheads in oriC indicate low-affinity ATP–DnaA-binding sites R5, I1–3, τ2, and C1–3. For simplicity, a non-essential τ1 site partially overlapping with IHF-interacting region including IHF-binding sequence (IBS; green bars) is omitted (Sakiyama et al., 2017). DnaA boxes 2, 3, 4, and 7 in datA and “Core” DnaA boxes I–III in DARS1/2 are essential for their activities. IBS, Fis-binding sequence (FBS; blue bar), and AT-rich repeats L, M, and R (DUE; purple bars) are also indicated. (C) Role and timing of IHF binding at oriC. IHF binds to a unique IBS and bends DNA during the pre-initiation period. At the initiation stage, ATP–DnaA becomes more abundant and forms specific oligomers on oriC. The left DnaA pentamer with bound IHF unwinds the DUE to initiate replication. After replication initiation, IHF quickly dissociates from oriC. DnaA is shown as pink polygon (domain III) with light yellow ball (domain I) and orange rectangle (domain IV). IHF is also shown as a green ball.

with oriCΔDUE and kan (kanamycin resistant) gene flanking frt sequences (pKX136) was constructed. The DNA fragment bearing the oriCΔDUE-frt-kan with the chromosomal sequence of the terC -proximal locus (intergenic region between pntA and ydgH; Ter-2 locus in Inoue et al., 2016) was amplified using primers TERori1-U and TERori-L (Table 2). Site-directed recombination was performed using SH022 cells, and the frt-kan region was removed using pCP20 to yield KX237 (dnaC2 ihfA-cHis12 TER-oriCΔDUE) and KX238 [dnaC2 ihfA-cHis12 TER-oriCΔ(DUE-R1)] (Table 1; Datsenko and Wanner, 2000).

To introduce oriCΔ(DUE-R1) at the terC locus, oriCΔ(DUE-R1)-frt-kan in pKX136 was amplified using primers TERori2-U and TERori-L, and similarly inserted into the SH022 genome.

Cell cultivation and cell cycle synchronization were performed according to a previously described method with minor modifications (Kasho and Katayama, 2013; Kasho et al., 2014; Inoue et al., 2016). To synchronize the E. coli cell cycle, cells were grown in supplemented M9 medium at 30°C, the permissive temperature for dnaC2, until the A660 of the culture reached 0.03, followed by further incubation at 38°C, the restrictive
temperature, for 90 min. Cells were immediately cooled to 30°C for 90 min. Cells were immediately cooled to 30°C overnight with a gentle rotation. Beads and bound materials were washed seven times with UT buffer, resuspended in 200 µl elution buffer [100 mM Tris–HCl (pH 7.5), 500 mM imidazole, 1% (w/v) SDS, and 10 mM dithiothreitol]. Proteins were degraded by Proteinase K treatment at 42°C for 2 h, followed by further incubation at 65°C for 6 h for de-crosslinking. After removal of proteins by phenol–chloroform–isoamyl alcohol extraction, DNA was recovered by ethanol precipitation in the presence of glycogen and resuspended in 10 µl nuclease-free water.

The DNA library for next-generation sequencing (NGS) was produced using the NEB Next DNA Sample Prep Reagent kit (New England Biolabs) according to the manufacturer’s instructions for “Preparing Samples for Sequencing Genomic DNA” (Illumina). The DNA fragments were then purified using a WIZARD SV Gel and PCR Clean-Up System (Promega), and amplified by 15 cycles of PCR. The sequence of the library was then determined by BioAnalyzer (Agilent Technologies).

Short-read sequencing was performed by BGI (Shenzhen, Guangdong, China) with the paired-end procedure (100 bp × 2) on an Illumina HiSeq 2000 instrument (Illumina, San Diego, California, United States). The fastq files of forward and reverse short reads, were identified precisely using the pmapsr program (see following section).

Determination of Highest IHF-Binding Regions and Determination of IHF-Binding Motifs

The regions protected by IHF in the E. coli genome, which would be sandwiched between the edges of DNase I digestion corresponding to the 5’ and 3’ ends of the forward and reverse short reads, were identified precisely using the pmapsr program (Chumsakul et al., 2013). The DNase I-digested short fragments were estimated to be 70–110 bp long; because we performed 100 bp Illumina sequencing, the reads frequently included the primer sequences added for DNA library construction (see before section), causing severe mismatches for mapping of the reads to the E. coli K-12 MG1655 genome (reference sequence) and decreasing the number of mapped reads. To reduce the number of unmapped reads, we initially mapped reads using the mpmap program permitting 35 bp mismatch in order to be able to map the reads, including the primer sequences (Chumsakul et al., 2013). We then determined the boundaries of the primer sequences and the homologous sequences to the reference genome in the forward and reverse short reads, which
In this study; seq analysis with minor modifications (the original version used regions (see details described below).

We identified the 70–110 bp regions sandwiched between the dnaC2 in exponentially growing cells. In this study, we used dosage bias from the replication origin to the termination site adapted to analyze randomly replicating cells and have a genome

We did not use the options -pbo, -pbt, or -pbs, which have been

We removed probable artifacts of highest IHF-binding regions, i.e., contamination of rRNA operons including rrrC (Figure 3A), which occur frequently in chromatin immunoprecipitation (Waldminghaus and Skarstad, 2010).

To visualize the IHF-binding profiles and highest IHF-binding regions estimated by pmapsr on IGV, which was used for visual inspection of false IHF-binding regions (described above) and to prepare Figures (i.e., Figure 3), we independently mapped the short reads onto the reference genome using Bowtie 2 to prepare sorted BAM files (Langmead and Salzberg, 2012), the format read by IGV. Bowtie 2 was used with the default settings using Illumina short reads with the primer sequence removed by the cutadapt program (Martin, 2011). Although some of highest IHF-binding regions overlapped and were consequently included in one IHF-binding peak, we used all of highest IHF-binding peaks to estimate the IHF-binding consensus sequence in each dataset. The consensus sequences in the highest IBS were estimated using the MEME suite with default settings (Bailey et al., 2009).

Quantitative PCR
Quantitative PCR (qPCR) experiments were performed as previously described (Kasho and Katayama, 2013; Kasho et al., 2014; Inoue et al., 2016). The levels of oriC and ylcC were quantified by real-time qPCR using SYBR Premix Ex Taq II (Perfect Real Time; Takara Bio) and primers ORL_1 and KWoriCRev for oriC; RTYLC-L and RTYLC-R for ylcC in SH022 or KX018 (Figure 2B); ilvG-U and ilvG-L for ilvG, rhb-B and rhb-L for rhb, glnH-U and glnH-L for glnH, osmY-U and osmY-L for osmY in KX018 (Supplementary Figure 7); RTNoriC-U2 and RTNoriC-L2 for native oriC; RTTERoriC-L and RTTERoriC-L for TER-oriC; and RTYLC-L and RTYLC-R for ylcC in KX237 or KX238 (Figures 5B,C).

Chromatin Affinity Precipitation
Chromatin affinity precipitation (ChAP) experiments were performed according to a previously described method (Kasho et al., 2014; Inoue et al., 2016).

3https://software.broadinstitute.org/software/igv/
FIGURE 3 | Cell cycle-specific IHF binding at oriC in GeF-seq. (A) The IHF-binding profile in the region including 3,920–3,970 kb part in the *E. coli* K-12 genome. (Continued)
Previous studies have identified or predicted thousands of IBSs. Before replication initiation, IHF binding is dynamically regulated at two intergenic sites, dnaA and dnaC2, within 5 min in the cells. Previous studies have identified cell-based synchronization, and thus we can synchronize the cell cycle just before replication initiation (0 min). By decreasing the temperature to low temperature (30°C), DnaC is immediately activated to concordantly initiate replication cycle. DnaC is the helicase loader; in the in vitro experiments, the presence of multiple binding regions (components) is possible to be estimated in one large peak. We, therefore, discriminate the “binding peak (peak)” and “binding region (component).” GeF-seq results of the synchronized dnaC2 cells showed a large peak at oriC in the 0 min data set, in which two possible binding regions (components) were estimated by pmapsr (Figure 3A). Those binding regions might indicate the higher-order complex formation by multiple DNA binding proteins and DNA (for instance, DnaA-IHF complexes and higher-order structures made by binding of multiple IHF molecules). In addition, our GeF-seq results included the representative IHF binding sites determined in previous in vitro DNase I footprinting experiments such as those at transcriptional promoters of ilvG.

### RESULTS

#### Specific IHF Binding to the oriC Locus Before Replication Initiation

Previous studies have identified or predicted thousands of IBSs on the *E. coli* chromosome (Grainger et al., 2006; Prieto et al., 2012), and we recently revealed that IHF binding is dynamically regulated at two intergenic sites, *dataA* and *DARS2* (Kasho and Katayama, 2013; Kasho et al., 2014), in addition to *oriC*. These observations suggest that the specific time of IHF binding is crucial for regulation of DNA replication and progression of the cell cycle. In this study, to identify the genome-wide regulation of IHF binding, we applied the GeF-seq method to temperature-sensitive *dnaC2* cells, in which it is possible to synchronize the replication cycle. DnaC is the helicase loader; in the dnaC2 mutant, replication initiation at oriC is specifically inhibited at high temperatures (38–42°C) and thus we can synchronize the cell cycle just before replication initiation (0 min). By decreasing the temperature to low temperature (30°C), DnaC is immediately activated to concordantly initiate replication within 5 min in the cells. Previous studies have identified cell cycle-dependent expression of the genes such as *dnaA*, *mioC*, and *gidA* using dnaC2 cell-based synchronization, and thus this well-established method should be the most suitable for the purpose in this study. By combining DNase I-dependent DNA cleavage, ChAP (a modified ChiP), and NGS, the GeF-seq method identifies protein-bound sites throughout the genome at base-pair resolution (Chumsakul et al., 2013).

Before mapping by GeF-seq, we confirmed by GeF-qPCR that IHF binding was regulated at the *oriC* locus in the dnaC2 mutant, i.e., IHF stably bound to the left part of *oriC* before initiation (0 min; Figure 2) and dissociated after initiation (10 or 20 min; Figure 2), consistent with previous studies (Ryan et al., 2002; Kasho and Katayama, 2013; Kasho et al., 2014).

Next, using the same samples as in Figure 2, we performed genome-wide mapping of IHF binding at specific cell cycle stages using GeF-seq (Figure 3). Unlike standard ChIP-seq, GeF-seq includes the DNase I digestion in the DNA fragmentation process instead of sonication in ChIP-seq. As a result, protein-bound DNA fragments purified in the GeF-seq procedure are shorter than those in standard ChIP-seq and the binding peaks detected by GeF-seq become sharper than those detected by standard ChIP-seq (Chumsakul et al., 2013). In addition, we have the unique program for the GeF-seq analysis, pmapsr which was developed to accurately detect protein-binding regions on genomic DNA using GeF-seq datasets (Supplementary Table 1; Chumsakul et al., 2013). Pmapsr is the program to detect the regions sandwiched by the 5′ and 3′ ends of reads sequenced by illumina sequencing, which are possible to represent the “edge” of the regions protected from the DNase I digestion by protein binding. Previously we succeeded to make the map of the genome wide footprinting of AbrB, the *Bacillus subtilis* global transcriptional regulator, by GeF-seq and the pmapsr analysis with similar resolution of in vitro DNase I footprinting and detected weak binding consensus sequence (Chumsakul et al., 2013). By this method, the presence of multiple binding regions (components) is possible to be estimated in one large peak. We, therefore, discriminate the “binding peak (peak)” and “binding region (component).”

GeF-seq results of the synchronized dnaC2 cells showed a large peak at oriC in the 0 min data set, in which two possible binding regions (components) were estimated by pmapsr (Figure 3A). Those binding regions might indicate the higher-order complex formation by multiple DNA binding proteins and DNA (for instance, DnaA-IHF complexes and higher-order structures made by binding of multiple IHF molecules). In addition, our GeF-seq results included the representative IHF binding sites determined in previous in vitro DNase I footprinting experiments such as those at transcriptional promoters of ilvG.
or glnH (Supplementary Figures 1, 3; Tsui and Freundlich, 1988; Claverie-Martin and Magasanik, 1991), which supports that our GeF-seq experiments were qualified enough to determine the genomic IHF binding regions. As shown in Figure 3, the IHF-binding peak at oriC was prominent relative to all other IHF-binding peaks in 0 min sample (Figure 3A). To determine how specifically IHF binds to oriC before replication initiation, we selected highest IHF-binding peaks and binding regions, and compared the read numbers of IHF binding among those regions. Highest IHF-binding peaks (Peak ID 1-42) and binding regions (Component ID 1-73) were selected by the pmapsr program using high threshold values (see section "Materials and Methods"). Peak ID 32 (Component ID 57, genomic positions 3,925,749 to 3,925,842), highest IHF-binding peak in the 0 min dataset, overlapped with the oriC region (Figures 3A, 4 and Supplementary Table 1), which had the highest read depth in all of highest IHF-binding peaks selected in the 0 min dataset. The average read depth of the oriC locus was 30,601.5, whereas that of Peak ID 1 (Component ID 1, the genomic positions of 16,520 to 16,606), the second highest IHF-binding peak (Figure 4, Supplementary Figure 6, and Supplementary Table 1), was 16,341.3. By contrast, the same oriC locus in the 10 min, 20 min, and Random datasets did not have the highest read depth relative to other IHF-binding regions (Figures 3A, 4 and Supplementary Table 1), consistent with the data shown in Figure 2B and previous studies that IHF binding at oriC locus is increased specifically at pre-initiation period (Kasho and Katayama, 2013; Kasho et al., 2014). This observation clearly indicates that IHF binds to the oriC locus most preferentially before replication initiation; this was a highly distinctive property of IHF binding at the oriC locus, at least among highest binding regions (see next section).
comparable or even larger compared with the IHF binding signals at oriC. In addition, IHF binding at representative non-oriC regions. In previous experiments, IHF binding at those non-oriC loci (Peak ID 34; Component ID 59 and 60) were mostly the same conditions, IHF binding at those non-oriC regions. In addition, IHF binding at representative non-oriC regions. Supplementary Table 1; Colland et al., 2000), which supports Figure 7, in highest IHF binding regions (initiation period (Kasho et al., 2014), oriC). As previously shown (Kasho and Katayama, 2013; Figure 7A), the replication initiation period. FIGURE 5 | Requirement for DnaA box R1 for oriC-IHF binding at the replication initiation period. (A) Deletion derivatives of oriC. Bars indicate the regions of oriCΔDUE and oriCΔ(DUE-R1), which were inserted at the position indicated as TER-oriC in the terC locus (shown in panel (B) of SH022 (dnaC2) strain, resulting in strains KK237 and KK238. Symbols used for the oriC structure are shown as in Figure 1B. (C,D) IHF-ChAP-qPCR with TER-oriC-inserted cells. SH022 (dnaC2), KK237 (dnaC2 TER-oriCΔDUE), and KK238 (dnaC2 TER-oriCΔ(DUE-R1)) cells grown at 30°C (R; Random) were transferred to 38°C and incubated for 90 min (S; Synchronized). The relative levels of native oriC (C) and TER-oriC (D) before and after Ni-affinity purification were determined using real-time qPCR, and yield was calculated (expressed as %). Error bars represent standard deviation from three independent experiments.

In addition, before initiation, the IHF-binding peak was only prominent at the oriC locus, and in contrast, the signals of all other IHF-binding peaks including the rrnC locus, one of the rDNA operons, were relatively modest compared with the IHF binding signals at oriC (Figures 2B, 3A, 0 min sample). In contrast, after initiation, the IHF-binding peak at oriC was relatively modest or discreet, and the peaks at other loci including that at ilvG (Peak ID 33; Component ID 58) or the rhlB loci (Peak ID 34; Component ID 59 and 60) were mostly comparable or even larger compared with the IHF binding signals at oriC (Figure 3A, 10 and 20 min samples; Supplementary Figures 2, 3), consistent with the decreased oriC copies in qPCR analysis (Figure 2B). Notably, the representative IHF binding sites at transcriptional promoter of osmY determined in previous in vitro DNase I footprinting experiments was included in highest IHF binding regions (Supplementary Figure 4 and Supplementary Table 1; Colland et al., 2000), which supports that our methodology was suitable for identifying IHF binding regions. In addition, IHF binding at representative non-oriC loci such as ilvG, rhlB, glhH, and osmY regions was demonstrated in ChIP-qPCR experiments using IHF antibody (Supplementary Figure 7). As previously shown (Kasho and Katayama, 2013; Kasho et al., 2014), oriC-IHF binding was increased only at pre-initiation period (Supplementary Figure 7A). In contrast, under the same conditions, IHF binding at those non-oriC loci was not largely changed or even decreased in the initiation periods (Supplementary Figures 7B–E). These results are basically consistent with the GeF-seq data (see also Discussion).

**Prediction of Specific IHF-Binding Consensus Sequences Before Replication Initiation**

To determine how IHF preferentially binds to oriC before initiation, we investigated the consensus sequences of highest IHF-binding peaks in each dataset using the MEME suite (Bailey et al., 2009). As shown in Figure 3C, the highest consensus sequence in all datasets (0 min: CAnnnnTTT at position 19–27, 10 min: WWTCArnnnTTA at position 13–25, 20 min: WWCArnnnTT at position 5–15, and Random: WAWCAArnnnTT at position 13–24, where S is G or C) are similar and contain essential DNA elements with the known consensus sequence WATCAArnnnTT (W is A or T; n is any nucleotides; R is A or G) (Swinger and Rice, 2004). The 0 min consensus sequence is more enriched in “T/A” at positions 11–16 and 25–29 beside the conserved “CA” at position 19–20 (Figure 3C). In addition, unique “GTTG” and “AAC” elements at positions 1–4 and 31–33, respectively, locate beside these “T/A” elements (Figure 3C). The highest IHF binding region at oriC (Component ID 57) is one of the loci with the best match to the 0 min consensus (Supplementary Figure 5), supporting the idea that these specific DNA elements in the 0 min consensus sequence are relevant to the preferential oriC–IHF binding before initiation.

**Whole Genome Analysis Predicts the Preferential IHF-Binding at oriC Locus Before Replication Initiation**

As shown in Figure 3B, we identified two IHF-protected regions (Component ID 56/57) in the prominent IHF-binding peak at oriC (Peak ID 32; Figure 4, Supplementary Figure 6, and Supplementary Table 1). The right protected region (Component ID 57) at oriC partly overlaps with a known IHF-binding consensus sequence (Figure 3E), which was previously identified as an oriC IBS by in vitro DNase I footprinting analysis (Figure 3D; Sakiyama et al., 2017). This study provides the first direct evidence for in vivo IHF binding to the oriC IBS in E. coli cells. The second protected region (Component ID 56) has weak IHF binding signal. Although no known IBS had been identified in this region, we identified a sequence that matches the new 0 min consensus sequence in Component ID 56 (Figures 3C,E). Alternative possibility is that the DnaA–IHF complex bound to the wide area of the oriC region, and the protected region would be expanded outside oriC. However, the existence of a secondary IBS with unstable IHF binding cannot be ruled out. Taken together, this preferential oriC–IHF binding before initiation requires the additional mechanism to the 0 min consensus sequences; i.e., weak IHF binding to the secondary site and the presence of two IBSs at the oriC locus may support stable oriC–IHF binding at a specific stage of the cell cycle (see below and Discussion).
We did not detect substantial specific signals at the *datA* and *DARS2* loci, although our previous study using ChIP-qPCR indicated that those should appear at 10 and 20 min after initiation. This is probably because of a difference in experimental conditions between ChIP and GeF, e.g., DNase I treatment or NGS sample preparation (see details in Discussion).

**A Role for oriC DnaA Box R1, but Not Chromosomal Position of oriC, in Stable IHF Binding**

To examine the regulatory mechanism by which IHF binding at the oriC locus is stabilized before initiation, we focus on two specific structural features of oriC: (1) overall structure of 1 Mb chromosomal region containing oriC, termed Ori-macromdomain (Niki et al., 2000; Valens et al., 2016), and (2) the local structure of oriC complexed with ATP-DnaA. First, to examine the requirement of the Ori-macromdomain structure, we constructed genome-edited cells carrying an insertion of DUE-deleted oriC sequence at terC-proximal intergenic region between *pntA* and *ydgH* genes (TER-oriCΔDUE; Figures 5A,B), and analyzed IHF-binding patterns before initiation. These cells have intact oriCs at the native position (Native oriC), and as expected, insertion of TER-oriCΔDUE had a minimal effect on stabilization of IHF binding at the native oriC locus before initiation (Figure 5C). Even at the terC locus, the signal of IHF binding was dependent on insertion of TER-oriCΔDUE, and it increased before initiation (Figure 5D). These observations suggest that...
FIGURE 7 | Time-resolved IHF binding at the ttcA-Rac prophage locus. (A) IHF-binding profiles at the ttcA upstream region in the 0 min, 10 min, 20 min, and Random samples. Red arrowhead represents the locus with IHF binding at specific timing (10 min) after replication initiation, which corresponds to Rac prophage attachment site attL. (B) Sequence comparison of Rac prophage attachment sites and IHF-binding consensus sequences in ttcA (top) or ttcC genes (bottom), and characters with yellow highlight is sequences different in ttcA or ttcC genes. The start codons of the ttcA or ttcC genes are shown by red characters. IHF-binding consensus (TAAnnnnTTGATW, where W means T or A) is shown by green characters. Rac prophage attachment sites attL or attR are shown by blue characters. (C) IHF-binding profiles at the ttcC upstream region in the 0 min, 10 min, 20 min, and Random samples. Red arrowhead represents the locus with Rac prophage attachment site attR.
timely stabilization of oriC–IHF binding is independent of Ori-
macromodomain structure. Interestingly, IHF binding signal at
TER-oriCΔDUE might be higher than that at Native oriC,
which could suggest two possibilities that DNA structure at Ter-
macromodomain could be preferable for oriC–IHF binding, or that
oriC DUE could play an inhibitory role for IHF binding by
unidentified mechanism.

Second, to assess the requirement of ATP–DnaA oligomers on
oriC for stabilizing IHF binding before initiation, we further
constructed genome-edited cells carrying an insertion of an oriC
derivative lacking a DUE-R1 region at the TerC locus [TER-
oriCΔ(DUE-R1); Figure 5A]. Deletion of DnaA box R1 should
dramatically change the structure of ATP–DnaA oligomers on
oriC. As with TER-oriCΔDUE, insertion of TER-oriCΔ(DUE-
R1) had little effect on the IHF-binding pattern at the native
oriC locus (Figure 5C); at the terC locus, the signal of IHF
binding was observed in a manner dependent on insertion of
TER-oriCΔ(DUE-R1; Figure 5D). Notably, in contrast to the
TER-oriCΔDUE, the signal of IHF binding at TER-oriCΔ(DUE-
R1) was not increased even before initiation (Figure 5D). These
results indicate that DnaA box R1 is required for stabilization of
IHF binding at oriC at the stage of replication initiation. The
ATP–DnaA molecule bound at DnaA box R1 is predicted to
interact with the one bound at DnaA box R5 via DNA
bending at the primary IBS (IBS1; Figure 6) induced by IHF
binding (Figures 1B,C). The ATP–DnaA oligomer formed on
oriC might stabilize DNA bending by IHF and thus form a rigid
oriC–IHF complex. Another possibility is that the secondary
IBS (IBS2; Figure 6), which partly overlaps with DnaA box R1,
might play a supporting role in stabilizing IHF binding (see
details in Discussion).

An Unique Locus Binds IHF at a Specific
Time
From the genome-wide mapping of IBS by GeF-seq, we identified
a novel cell cycle-dependent IHF binding in the ttcA gene region
(Figure 7). This locus has the unique feature that IHF binds
temporarily 10 min after initiation but dissociates 20 min after
initiation and before initiation, as observed for all other IBSs
except for oriC (Figure 7A). The IHF-binding peak corresponded
to the IHF-binding consensus near the start codon of the
ttcA gene, which encodes a tRNA-thioltransferase (Figure 7B; Liu
et al., 2015). In addition, this locus is adjacent to the Rac prophage
excision site attL (Figure 7B). A region including the other Rac
prophage excision site attR, located in the ttcC gene, has a similar
IHF-binding consensus but does not have IHF-binding peaks (Figures 7B,C). The IHF-binding region in the ttcA gene could
play a major regulatory role in Rac prophage excision; however,
this hypothesis requires further experimental testing.

DISCUSSION

In this study, we sought to characterize the cell cycle-dependent
regulation of the genomic IHF-binding pattern using GeF-
seq at base-pair resolution. The results provided evidence of
unique and specific regulation of strong oriC–IHF binding at
the replication pre-initiation stage. This suggests that at that
stage, IHF preferentially binds to oriC rather than other sites
with affinity for IHF. The mechanisms responsible for this
preference remain unknown (see below). Previous studies have
attempted genome-wide analysis of IHF binding by antibody-
based immunoprecipitation, but no previous study detected
IHF binding at the oriC locus (Grainger et al., 2006; Prieto
et al., 2012). This is consistent with our observations that the
prominent binding peak at oriC was specific for the initiation
stage; in other stages, as well as in random cultures, the IHF-
binding peaks at oriC were smaller than other evident binding
peaks. In addition, formation of bulky oriC-DnaA-IHF superf-
complexes may inhibit the interaction between IHF and the
antibodies used in previous studies. By contrast, our method,
which is based on His-tag affinity purification, improves the yield
of the IHF complex with bulky initiation complexes (Kasho et al.,
2014; Inoue et al., 2016), and successfully detected oriC–IHF
binding in a comprehensive analysis of genomic IHF binding
for the first time.

However, our analyses also have limitations. Some of the
known IHF-binding loci, such as datA and DARS2, were not
substantially detected, which could be explained in three ways.
First, compared with IHF binding to oriC, that to datA and
DARS2 could be more dynamic, with a rapid binding/dissociation
equilibrium to ensure timely interaction during the cell cycle,
making it difficult to detect the binding using the present
methods. Second, in regard to cell sample preparation, our GeF-
seq experiments were performed using cells at early exponential
phase to determine the effect of the cell cycle on actively growing
and replicating cells, whereas previous studies were performed
using cells at mid/late exponential or stationary phases (Grainger
et al., 2006; Prieto et al., 2012). Expression of the IHF protein
is 3–6-fold higher in stationary phase than in exponential phase
(Ali Azam et al., 1999), which might destabilize IHF binding
to datA and DARS2 in early exponential phase. Third, for the
DNA samples used for NGS, protein-bound DNA was treated
with DNase I to precisely determine the protein-binding site;
however, previous studies suggested that DNase I may have
some sequence specificities that would prevent detection of some
chromosomal loci (Herrera and Chaires, 1994; Koohy et al.,
2013), e.g., input read depth at the DARS2 locus was very low
relative to the surrounding regions. In addition, results obtained
by ChIP-qPCR experiments were overall consistent with the GeF-
seq results (Supplementary Figure 7), in certain cases relative
IHF binding levels at 10 or 20 min after replication initiation
were moderately different between GeF-seq and ChIP-qPCR data
(Figures 3A and Supplementary Figures 7B,C). This might be
caused from changes in the local copy number of genomic DNA
during replication. In GeF-seq data, we cannot normalize IHF
signals according to this change. Also, the IHF binding at
gltH locus was lower than the threshold in this experiment, whereas
it was detected by in vitro DNase I footprinting (Claverie-
Martin and Magasanik, 1991). This difference might be because
we set the high threshold value to identify the highest IHF
binding regions which was to enable the identification of the
timepoint-specific IHF binding in GeF-seq results. Therefore,
the relatively lower binding peaks may also represent specific
IHF binding regions. Consistently, in the previous GeF-seq experiment, many of the lower binding peaks were specific and had the consensus sequence of the *B. subtilis* transcriptional regular AbdB (Chumsakul et al., 2013). Another possible reason for the difference in the IHF binding profiles between this and previous experiments is that in experimental conditions.

Our results provide a new model for the specific regulation of local oriC–IHF binding at the replication initiation stage. We successfully reconstituted cell cycle-dependent IHF binding/dissociation by introducing oriC∆DUE at the terC locus and performing oriC mutation analysis. The results suggested that oriC DnaA box R1, but not oriC location, is essential for stable oriC–IHF binding (Figure 5C). In initiation complexes, ATP–DnaA forms specific tight and bulky oligomers, which could stabilize the binding of DNA bound to IHF at the oriC-proximal region (Figure 1C), implying that ATP–DnaA oligomers might prevent IHF from freely dissociating from the oriC region during DNA bending promoted by IHF binding (the process is discussed below). In addition, it should be noted that in the CRISPR–Cas system of *E. coli*, IHF binding and the resultant DNA bending promotes DNA binding by the Cas1–Cas2 integrate complex, and Cas1 directly interacts with IHF (Wright et al., 2017). Similarly, IHF could directly interact with DnaA in the initiation complex. Consistent with this, HU, a structural homolog of IHF, interacts directly with DnaA (Chodavarapu et al., 2008). In addition, the possibility of a DnaA–IHF interaction could explain the appearance of an expanded protection region including DnaA box R1 (Component ID 57 in Figures 3B,E). Another possibility is that specific subcellular localization of IHF by liquid–liquid phase separation (LLPS) could occur before and at the initiation stage, followed by dramatic changes in IHF dynamics after initiation. Given that intrinsically disordered regions (IDRs) of proteins are suggested to stimulate LLPS (Borcherds et al., 2020), this hypothesis is consistent with predictions for IDRs of IHF in the MobiDB database: only α-subunits of proteobacterial IHF (such as *Salmonella typhimurium*, *Myxococcus xanthus*, *Vibrio cholerae*, etc.), but not β-subunits, have IDRs (Potenza et al., 2015). For example, amino acids 49–73 of *E. coli* IHF-α, which forms β-sheets and interacts directly with DNA (Rice et al., 1996), are predicted to be an IDR; however, further analysis is required to prove these possibilities.

Our comprehensive sequence analysis of the IHF-binding consensus suggested that before replication initiation, IHF has a specific consensus sequence that consists of conserved elements “CA” at positions 19–20 and “TT” at positions 26–27, as well as unique elements “GTTG” and “AAC” at positions 1–4 and 31–33, respectively (Figure 3C). Previous studies have tried to address the relationship between the IBS and its binding affinity; however, the requirement of the surrounding AT-rich elements in the IHF-binding consensus remains unclear (Aeling et al., 2006). This study raises the possibility that at a specific cell cycle stage, other NAPs, supercoiling state, or transcriptional profile might change the higher-order genomic structure and strengthen the requirement for surrounding AT-rich elements of the IHF-binding consensus. IHF binding to this new IHF consensus sequence might regulate expression of cell cycle-dependent genes before initiation. DnaA is known to regulate cell cycle-dependent genes such as *nrdAB* and *mioc* (Gon et al., 2006; Hansen et al., 2007). Also, SeqA protein, which specifically recognizes hemi-methylated GATC sequences after initiation, represses transcription of *dnaA* and *gidA* genes after initiation (Bogan and Helmstetter, 1997). Thus in addition to DnaA and SeqA, IHF could be a novel regulator for cell cycle-dependent gene expression.

A previous kinetic study suggested that DNA binding and bending introduced by IHF occur in a stepwise manner, and that IHF–DNA binding is a rapid process, whereas IHF-induced DNA bending is much slower and therefore rate-limiting (Sugimura and Crothers, 2006). This kinetic model and our sequence analysis of IHF-binding consensus at oriC suggested that oriC has a secondary IBS (IBS2) at a region overlapping with R1, which could also explain the mechanism by which oriC–IHF binding is stabilized at the replication initiation stage: first, IHF could bind to either the primary IBS (IBS1) or IBS2, and then if DnaA box R1 is occupied by DnaA, IHF can no longer bind to IBS2; in this case, however, IHF binding to IBS1 would be stabilized by oriC-ATP–DnaA complexes as described (Figure 6). A similar mode of discrimination of NAP–DNA binding was proposed to occur in the regulation of *E. coli* flmA transcription by H-NS and Fur, i.e., H-NS dimers cooperatively bind to the flmA promoter to repress transcription, and Fur expression induces switching from H-NS to Fur to activate flmA transcription (Nandal et al., 2010). Also, as with other DNA binding/bending proteins such as human mitochondrial transcription factor A (TFAM) (Farge et al., 2012), IHF could be mobile and slide along DNA from IBS2 to IBS1. Alternatively, IBS2 could be a reservoir of IHF, i.e., if an IHF molecule bound at the primary site IBS1 is accidentally dissociated, it could re-bind to IBS2, preventing free diffusion and thereby stabilizing overall oriC–IHF binding. These features of IHF-binding dynamics and IBSs in oriC imply that oriC structure is designated in a sophisticated manner to ensure IHF binding for regulation of replication initiation.

IHF dissociates from oriC within 5–10 min after initiation (Figures 2, 3; Kasho and Katayama, 2013; Kasho et al., 2014), although the mechanisms remain unclear. The passage of the replication machinery, which is loaded onto the unwound oriC region, or unknown factors could contribute to IHF dissociation. In addition, oriC contains eleven copies of Dam-dependent methylation sequence GATC and SeqA-dependent DnaA dissociation at post-initiation period can be considered as an analogous mechanism (Nievera et al., 2006; Katayama et al., 2010). Notably, GATC sequence is also present in IHF binding regions identified in this study (Figures 3D,E), suggesting that SeqA sequestrates DnaA and IHF from oriC to inhibit over-initiations. Also, after initiation, ATP–DnaA is converted to ADP–DnaA, causing ADP–DnaA to become predominant in cells (Kurokawa et al., 1999). ADP–DnaA molecules form unstable oligomers on oriC, suggesting that dissociation of DnaA molecules from oriC might impede stable interaction of IHF with oriC in vivo.

In addition, we identified a novel IHF-binding region at the Rac prophage excision site *attL* (Figure 7). In the CRISPR–Cas
system, IHF interacts directly with Cas1 integrase and promotes the interaction of Cas1–Cas2 complexes with DNA, thereby suppressing off-target integration by Cas1–Cas2 (Wright et al., 2017). Also, IHF is a stimulatory factor for both integration and excision of bacteriophage lambda (Segall and Nash, 1996; MacWilliams et al., 1997). In E. coli, Rac prophage plays a role in cellular stress responses and uses a lambda phage-like mechanism for integration and excision by IntR integrase (Liu et al., 2015). Interestingly, the function of Rac prophage is stimulated under specific environmental conditions such as high nutrition or growth in early exponential phase (Liu et al., 2015), suggesting that similar significant role of IHF might be present in Rac prophage excision, and that IHF binding could stimulate the excision at Rac prophage excision site attL at a specific cell cycle stage, which remains to be determined. The significance of these cell cycle-specific interactions remains unknown.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. Genome sequencing data were deposited in the DDBJ Sequence Read Archive (DRA) under BioProject number PRJDB11576 and accession numbers DRA012298 (0 min pulldown), DRA012299 (10 min pulldown), DRA012300 (20 min pulldown), DRA012301 (Random pulldown), DRA012302 (0 min Input), DRA012303 (10 min Input), DRA012304 (20 min Input), and DRA012305 (Random Input).

REFERENCES

Aeling, K. A., Opel, M. L., Steffen, N. R., Treyarchenko-Ladokhina, V., Hatfield, G. W., Lathrop, R. H., et al. (2006). Indirect recognition in sequence-specific DNA binding by Escherichia coli integration host factor: the role of DNA deformation energy. J. Biol. Chem. 281, 39236–39248. doi: 10.1074/jbc.M606363200

Ali Azami, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. J. Bacteriol. 181, 6361–6370. doi: 10.1128/JB.181.20.6361-6370.1999

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. 37, W202–W208. doi: 10.1093/nar/gkp335

Bogan, J. A., and Helmstetter, C. E. (1997). DNA sequestration and transcription in the oriC region of Escherichia coli. Mol. Microbiol. 26, 889–896. doi: 10.1046/j.1365-2958.1997.6221989.x

Borchers, W., Bremer, A., Borgia, M. B., and Mittag, T. (2020). How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation? Curr. Opin. Struct. Biol. 67, 41–50. doi: 10.1016/j.copbio.2020.09.004

Chodavarapu, S., Felczak, M. M., Yaniv, J. R., and Kaguni, J. M. (2008). Escherichia coli Dam2 interacts with HU in initiation at the E. coli replication origin. Mol. Microbiol. 67, 781–792. doi: 10.1111/j.1365-2958.2007.06094.x

Chumsakul, O., Nakamura, K., Kurata, T., Sakamoto, T., Hobman, J. L., Ogasawara, N., et al. (2013). High-resolution mapping of in vivo genomic transcription factor binding sites using in situ DNase I footprinting and ChIP-seq. DNA Res. 20, 325–337. doi: 10.1093/dnares/dst113

Claverie-Martin, F., and Magasanik, B. (1991). Role of integration host factor in the regulation of the araC promoter. Proc. Natl. Acad. Sci. U.S.A. 88, 1631–1635. doi: 10.1073/pnas.88.5.1631

AUTHOR CONTRIBUTIONS

KK and TO performed the experiments. TO, OC, KN, and KF analyzed the IHF binding profile from GeF-seq data. KK, TO, and TK wrote the manuscript. All authors conceived the experiments and analyzed the data.

FUNDING

This study was supported by Grant-in-aid for Scientific Research, JSPS KAKENHI Grant numbers JP21K19233, JP20H03212, JP17H03656, and JP15K18479, and was partly supported by Institute for Fermentation, Osaka (G-2020-2-103 to TO).

ACKNOWLEDGMENTS

We would like to thank Sjoerd Wanrooij for understanding in preparation of this manuscript when KK was a postdoctoral researcher in his laboratory of Umeå University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.697712/full#supplementary-material

Colland, F., Barth, M., Hengge-Aronis, R., and Kolb, A. (2000). Sigma factor selectivity of Escherichia coli RNA polymerase: role for CRP, IHF and Lrp transcription factors. EMBO J. 19, 3028–3037. doi: 10.1093/emboj/19.12.3028

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97, 6640–6645. doi: 10.1073/pnas.120163297

Dillon, S. C., and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat. Rev. Microbiol. 8, 185–195. doi: 10.1038/nrmicro2261

Farge, G., Laurens, N., Broekmans, O. D., Van Den Wildenberg, S. M. J. L., Dekker, L. C. M., Gaspari, M., et al. (2012). Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. Nat. Commun. 3:1013. doi: 10.1038/ncomms2001

Fujimitsu, K., Senruchi, T., and Katayama, T. (2009). Specific genomic sequences of E. coli promote replicational initiation by directly reactivating ADP-DnaA. Genes Dev. 23, 1221–1233. doi: 10.1101/gad.1775809

Gon, S., Camara, J. E., Kluungsoy, H. K., Crooke, E., Skarstad, K., and Beckwith, J. (2006). A novel regulatory mechanism couples deoxynucleotidyl synthesis and DNA replication in Escherichia coli. EMBO J. 25, 1137–1147. doi: 10.1038/sj.emboj.7600990

Grainger, D. C., Hurd, D., Goldberg, M. D., and Busby, S. J. W. (2006). Association of nucleoid proteins with coding and non-coding segments of the E. coli genome. Nucleic Acids Res. 34, 4642–4652. doi: 10.1093/nar/gkl542

Hansen, F. G., Christensen, B. B., and Atlung, T. (2007). Sequence characteristics required for cooperative binding and efficient in vivo titration of the replication initiator protein DnaA in E. coli. J. Mol. Biol. 367, 942–952. doi: 10.1016/j.jmb.2007.01.056

Herrera, J. E., and Chaires, J. B. (1994). Characterization of preferred deoxyribonuclease I cleavage sites. J. Mol. Biol. 236, 405–411. doi: 10.1006/jmbi.1994.1152
Valens, M., Thiel, A., and Boccard, F. (2016). The MaoP/maoS site-specific system organizes the ori region of the *E. coli* chromosome into a macrodomain. *PLoS Genet.* 12:e1006309. doi: 10.1371/journal.pgen.1006309

Waldminghaus, T., and Skarstad, K. (2010). ChIP on chip: surprising results are often artifacts. *BMC Genomics* 11:414. doi: 10.1186/1471-2164-11-414

Wright, A. V., Liu, J.-J., Knott, G. J., Doxzen, K. W., Nogales, E., and Doudna, J. A. (2017). Structures of the CRISPR genome integration complex. *Science* 357, 1113–1118. doi: 10.1126/science.aao0679

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kasho, Oshima, Chumsakul, Nakamura, Fukamachi and Katayama. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.