Iteron Plasmids

IGOR KONIECZNY,1 KATARZYNA BURY,1 ALEKSANDRA WAWRZYCKA,1 and KATARZYNA WEGRZYN1

1Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, Gdansk, Poland

ABSTRACT Iteron-containing plasmids are model systems for studying the metabolism of extrachromosomal genetic elements in bacterial cells. Here we describe the current knowledge and understanding of the structure of iteron-containing replicons, the structure of the iteron plasmid encoded replication initiation proteins, and the molecular mechanisms for iteron plasmid DNA replication initiation. We also discuss the current understanding of control mechanisms affecting the plasmid copy number and how host chaperone proteins and proteases can affect plasmid maintenance in bacterial cells.

INTRODUCTION Iteron plasmids are extrachromosomal genetic elements that can be found in all Gram-negative bacteria. Despite the fact that these plasmids bring antibiotic resistance to host bacterium, they can also bring other features, for example, genes for degradation of specific compounds or toxin production. Iteron plasmids possess characteristic directed repeats located within the origin of replication initiation that are called iterons. These plasmids became model systems for investigation of the molecular mechanisms for DNA replication initiation and for the analysis of mechanisms of control of plasmid copy number in bacterial cells. This research has provided our basic understanding of plasmid biology and the relationship between plasmid DNA and host cells. The control mechanisms utilized by iteron plasmids are based on the nucleoprotein complexes formed by the plasmid-encoded replication initiation protein (Rep). The Rep proteins interact with iterons, which initiates the process of plasmid DNA synthesis, but Rep proteins are also able to form complexes with iterons, which inhibits the replication initiation process. This inhibition is called “handcuffing.” Also, Rep protein can interact with inverted repeated sequences, causing transcriptional auto-repression. Finally, various chaperone protein systems and proteases affect the Rep activity and, therefore, overall plasmid DNA metabolism.

STRUCTURE OF THE ORIGIN OF REPLICATION INITIATION

The origin region is one of the most important sequences within plasmid DNA; it ensures plasmid autonomous replication, independent of replication of the bacterial chromosome. As in other replicons, plasmid origins consist of characteristic motifs recognized by replication initiation proteins. In iteron-containing plasmids (Fig. 1), iterons that are directly repeated sequences play a crucial role during DNA replication initiation and are critical for plasmid copy number control (see also text below). They are quite short sequences, whose lengths vary from 17 bp in plasmid RK2 (1), 19 bp in plasmids F (2) and P1 (3), to 22 bp in R6K (4), pPS10 (5), and plasmids from the IncQ incompatibility group (6). Sometimes, such as in plasmid pXV2 from the IncW incompatibility group, direct repeats within the origin can vary in length. In pXV2 there are two 18-bp and two 19-bp repeats (7). The iteron number and spacing between iterons also can differ among iteron-containing plasmids. From the
The smallest iterons were identified in plasmid pSC101 (8), in which there are three iterons. In plasmids pPS10 (5) and F (2) there are four iterons; in RK2, 5 (1); and up to seven have been identified in R6K (4). In plasmids from the IncQ group there are three or four identical direct repeats, but sometimes the functional origin contains more iterons that are partly deleted, contain point mutations, or are incorrectly spaced (6). Plasmid R478 from the IncHI2 incompatibility group even contains several iterons that differ in length (eight 18-bp and nine 76-bp iterons) and are separated by a sequence of rep genes (9). Iterons are recognized by a plasmid-encoded Rep protein, and they are bound by a Rep monomeric form (10–13) in a cooperative manner (14, 15). Mutations within an iteron sequence can abolish the binding of Rep protein and, in consequence, plasmid replication. This was shown for plasmid R6K, in which changes in a sequence of iterons made impossible the binding of the π protein in vitro and replication of plasmids with mutated origins in vivo (16). Similarly, mutations within an iteron sequence in the origin of plasmid P1 reduced or completely prevented origin activity (17). Negative effects on plasmid replication are also exerted by changes in spacers between iterons. The importance of sequences adjacent to iterons was shown for plasmids P1 (17), RK2 (14), and pSC101 (18). Also, disturbances in the position of iterons in relation to other motifs present in the origin region, especially changes in proper helical phasing, have a negative influence on plasmid replication activity (19).

The binding of the plasmid initiator to double-stranded DNA (dsDNA) containing iterons results in local destabilization of the DNA duplex. Plasmid Rep protein is very often accompanied in its action by host initiator DnaA protein. DnaA protein binds a specific motif called DnaA-box, also localized within the plasmid origin. DnaA-boxes are 9-bp-long sequences with consensus sequences that are varied depending on the host bacteria (20). DnaA-boxes can be localized upstream from iterons (e.g., plasmids RK2 and pPS10), downstream from the region rich in adenine and thymine residues (AT-rich), where local destabilization of the duplex occurs (e.g., plasmids F and pSC101), or in both these positions (e.g., plasmids R6K and P1) (21). In exceptional situations, such as in plasmid pXY2, DnaA-box overlaps with the first iteron (7). In some plasmid origins there is just one DnaA-box (e.g., plasmids pSC101,
pPS10, and pXV2), and in others there are two (e.g., oriC of plasmid R6K and plasmid F), four (e.g., plasmid RK2), or even five (e.g., plasmid P1) such motifs. Their length and sequence usually correspond to the consensus sequence of DnaA-boxes present in the origin of the *Escherichia coli* chromosome (oriC). If there are some deviations from consensus, they usually do not exceed point mutations. Examples include DnaA-boxes from plasmids P1 and RK2, which contain one or two mismatches. The position of DnaA-boxes is as important as their sequence. Insertions of more or less than a helical turn between DnaA-boxes and iterons within the plasmid RK2 origin resulted in inactivation of the origin’s replication activity (19). The binding of DnaA protein to DnaA-boxes in the origin of broad-host-range iteron-containing plasmids can vary in different host bacteria. For instance, in the plasmid RK2 origin, DnaA-boxes 3 and 4 should be present when replication takes place in *E. coli* and *Pseudomonas putida* cells. However, they can be missed during plasmid replication in *Pseudomonas aeruginosa* (22). For *E. coli* chromosome oriC it was shown that beside DnaA-boxes, DnaA protein bound with ATP can interact with ATP-DnaA-boxes localized within AT-rich repeats (23). However, in iteron-containing plasmid origin regions, motifs for ATP-DnaA binding, similar to those observed in oriC, have not been identified to date.

The third motif, in addition to iterons and DnaA-boxes, that can be distinguished within the iteron-containing plasmid origin is the AT-rich region. This is the sequence, usually located near iterons, where local destabilization of the double-stranded helix occurs during the process of replication initiation. This region is therefore considered a DNA unwinding element (DUE) where single-stranded DNA (ssDNA) is created. Although the thermodynamic stability of the AT-rich region can differ in different origins, usually it has much lower free energy (ΔG) than the overall profile of adjacent sequences (21). In the AT-rich region, it is possible to discern short repeated sequences, usually oriented directly. The exception can be the origin of plasmid RK2, where one of the repeated sequences is inverted in relation to the other ones (24). Repeated sequences within the AT-rich region are located tandemly one after the other (e.g., origin of plasmids RK2 [24] and pSC101 [25]), or they are separated with spacers of different length (e.g., 7-, 1-, and 6-bp spacers between AT-rich repeats of plasmid F [25] and 29- and 9-bp spacers in oriC of plasmid R6K [26]). The length of those repeated sequences can preserve 13 nucleotides (13-mers), as are present in the AT-rich region of *E. coli* oriC (e.g., plasmid RK2 [1] and pSC101 [25]). But more often they are shorter, such as in plasmids R6K (10 nucleotides [26, 27]), F (8 nucleotides [25]), and P1 (7 nucleotides [28, 29]). Also, the number of repeats can be different, and there can be two repeats in the AT-rich region of plasmid pSC101, four in plasmids RK2 and F, and up to five in plasmid P1 (21).

Although the consensus sequences for AT-rich repeats in different origins are difficult to identify, the consensus can be established for particular plasmid origins. The presence of all repeats within the AT-rich region, as well as their sequence, is very important for the proper replication activity of the origin. Even point mutations within these sequences can completely abolish plasmid replication (17, 30, 31). Also, substitution of one AT-rich repeat in a plasmid origin into a repeat from a bacterial chromosome origin results in a lack of replication activity *in vitro* and a decrease of activity *in vivo* (30). Although the presence and sequence of AT-rich repeats is critical for plasmid replication, the exact role of these motifs is still ambiguous.

The presence of binding sites for replication initiation proteins, iterons, and DnaA-boxes, as well as the region where duplex opening occurs, is very important for the replication initiation process. However, these motifs are not the only ones that can be distinguished within the origin of iteron-containing plasmids. In some plasmid origins the binding site for integration host factor (IHF) can be identified. Such a situation occurs, for instance, in the plasmid P1 origin, where the IHF binding site is located downstream from the cluster of three DnaA-boxes (32). The binding of IHF protein results in the bending of the DNA molecule; however, not only the bend but also its proper phasing for the downstream DNA is required for the activity of the origin (32). Insertions of less than a helical turn between IHF binding sites and DnaA-boxes in the P1 origin had a negative effect on origin activity. The IHF binding sites were also present in plasmids pSC101 (33) and R6K (34). This motif was identified as well in the plasmid RK2 origin, but the IHF deficiency in *E. coli* seemed not to alter plasmid replication efficiency or plasmid copy number control (35).

Other motifs that can be identified within some plasmid origins but are not directly involved during the replication process are sites, GATC motifs recognized by Dam methylotransferase. They are usually overlapped AT-rich repeated sequences (e.g., in plasmids P1 [36] and pSC101 [31]) or are located adjacent to these repeats (e.g., in plasmid P1 [36]). The methylated GATC sequence becomes hemimethylated during replication and in this form is recognized by the SeqA protein (37,
38), which sequesters newly synthesized DNA (39). SeqA negatively regulates DNA replication by blocking the GATC sites and preventing replication proteins from binding. Apart from the GATC motif, a region rich in guanine and cytidine residues (GC-rich) can be identified in some plasmids’ origins (e.g., plasmids RK2, P1, pPS10, and IncQ). Its exact role is unknown, and in plasmid RK2 it can be deleted without any effects on origin activity (19). In plasmid P1, in which a GT-rich sequence plays the role of a spacer between iterons and AT-rich repeats, the sequence of this region can vary considerably, but its length must be preserved (36).

In a few plasmids identification of motifs other than those described here was reported. For example, in plasmids F and R1 the binding site for the IciA protein was detected (25). The IciA protein, which binds the site located in the AT-rich region of plasmid origins, probably, like in E. coli oriC (40), inhibits the unwinding process at the AT-rich region. In the origin of plasmid R6K, binding sites for other regulatory proteins, Fis (factor for inversion stimulation) were found (41). It was shown that plasmid replication depends on the Fis protein when the gene for the copy-up mutant of the π protein and the penicillin resistance gene were present on plasmid DNA (41).

It could be concluded that for the proper activity of the origin of iteron-containing plasmids, not only the presence and the sequence of essential motifs, such iterons, DnaA-boxes, and AT-rich repeats, is important. The appropriate location of these motifs in relation to each other also has a great impact on replication activity. In particular, changes in proper helical phasing have a negative influence on plasmid replication.

**Rep PROTEIN STRUCTURE**

Although many plasmid Rep proteins have been identified, the crystallographic data are limited to a few replicons. This is due to a high instability of the Rep proteins, so understanding the initiators’ role in the structural context is a challenge. Plasmid replication initiators such as the RepA initiator of plasmid pPS10, RepE of plasmid F, and the π protein of plasmid R6K are best characterized in terms of structure. The RepA initiator of pPS10 was the first Rep protein whose structure was predicted to consist of two winged helix (WH) domains (42). These findings have been confirmed by the crystal structure of the monomer of a homologous RepE initiator of plasmid F, bound to iteron DNA (43). The other crystal structure of a plasmid Rep protein was determined for the monomeric form of the π initiator protein of plasmid R6K as a complex with a single copy of its cognate DNA-binding site (iteron) (44). The crystal structures of both RepE and π proteins are depicted in Fig. 2. Although the crystal structures of RepE and π proteins shed new light on the Rep monomers’ interaction with DNA, the molecular nature of Rep activation remained unknown until the crystal structure of the dimeric N-terminal domain of the plasmid pPS10 initiator (dRepA) was resolved (45). Nonetheless, the crystallographic data obtained for plasmid Rep proteins are limited to the WH domain description. Rep proteins are composed of two WH domains—N-terminal WH1 and C-terminal WH2—that are responsible for interaction with DNA (42) (Fig. 2). The WH2 domain contains a putative helix-turn-helix motif, which is the main determinant of Rep binding to both the iteron sequences and the inverted repeats (partially homologous to the iteron sequence), which was shown for the RepE initiation protein of the mini-F plasmid and RepA of plasmid pPS10 (46, 47). A formation of nucleoprotein complex by Rep protein results in the bending of the DNA molecule. Iteron interaction with the WH1 and WH2 domains of the Rep monomer, or interaction of inverted repeats with both WH2 domains of Rep dimer, induce DNA bending (42, 48). In Rep monomers, the WH2 domain binds to the 3′-half of the iteron, while the WH1 domain changes structure and contacts the 5′-iteron end, through both the phosphodiester backbone and the minor groove (42).

In contrast to initiation proteins of replicons F, R6K, and pPS10, the crystal structures of the TrfA protein of RK2 as well as P1 RepA have not been determined. The structure prediction using fold-recognition homology modeling was carried out in both cases. The N-terminal part of TrfA does not show a unique three-dimensional structure with the absence of stabilizing factors; it seems to be disordered in solution as opposed to the C-terminal part of the protein, which is expected as two copies of WH domains. Helices of both WH structures interact with major grooves of the DNA phosphate backbone (49). A series of mutations located within the WH1WH2 domains have been found to affect the TrfA-DNA interaction (50, 51). The structure predicted for P1 RepA, similar to TrfA, contains WH domains. By means of fold-recognition programs, it was shown that despite the lack of sequence similarity, RepA shares structural homology with plasmid F RepE. The model predicted that RepA binds one half of the binding site through interactions with the N-terminal DNA binding domain (WH1) and the second half through interactions with the C-terminal domain (WH2) (52). Interestingly, the
residues involved in Rep-DNA interactions located outside the WH domains have been determined with the use of RK2 initiator TrfA mutants (50, 51). These results assume the existence of an additional DNA binding motif, apart from WH1WH2 domains.

Like in plasmid-encoded Rep proteins, WH domains responsible for DNA binding were found in Archaea and Eukaryota initiators. However, the AAA+ domain (ATPases associated with various cellular activities) commonly present in Archaea and Eukaryota initiation proteins was not found in plasmid Reps (53) (Fig. 2). Thus, with regard to the DNA binding mechanism, the plasmid Rep proteins are similar to eukaryotic replication initiators. The results of biochemical and spectroscopic experiments revealed functional similarities between pPS10 RepA and archaeal/eukaryal initiators (53). The crystal structure determined for the archael initiator Cdc6 confirmed these findings (54). Interestingly, it was reported that similar to the mammalian proteins PrP and α-synuclein, the WH1 domain of the pPS10 RepA can assemble into amyloid fibers upon binding to DNA in vitro and in E. coli cells (55–58). It opens a direct means to untangle the general pathway(s) for protein amyloidosis in a host with reduced genome and proteome (59).

Plasmid Rep proteins exist in cells mostly as dimers (12, 60). The dissociation of dimers by the action of chaperones or interaction with iteron-containing DNA (see also text below) results in conformational changes in the Rep structure (61). A compact arrangement of the two WH domains, competent for binding to the inversely repeated sequences, becomes a more elongated form, which is suited for iteron binding (42). These conformational changes consist of a significant increase of the overall β-sheet at the expense of the α-helical one (61). The situation is different for the Rep dimers that interact with inversely repeated sequences. Binding of Rep dimers to the inverted repeats does not result in dissociation to monomeric forms or change in the dimers’ conformation (61). Although only the monomeric form of Rep proteins is replication-active, dimers of Rep can bind to an inversely repeated sequence localized close to the promoter region of the rep gene, which results in transcription auto-repression (see text below). This was shown for the RepA initiator of the

### FIGURE 2

Structure of replication initiators. DnaA of *A. aeolicus*, RepE54 from *E. coli* mini-F plasmid, π from R6K, and the C-terminal part of the TrfA protein (190–382 aa) of plasmid RK2 are depicted. Structure of the DnaA, RepE54, and π are derived from crystallographic data (PDB entry 1L8Q, 1REP, and 2NRA, respectively). The TrfA model was developed based on homology modeling. The AAA+ domain is colored in blue, the DNA binding domain (DBD) is shown in red, and Winged-Helix domains (WH1 and WH2) are colored in yellow and green, respectively. References and detailed information for crystallographic data of the DnaA, RepE54, π, and TrfA model are given in the text. doi:10.1128/microbiolspec.PLAS-0026-2014.f2
pSC101 (62) F RepE initiation protein (43) and the π initiator of the plasmid R6K (63). In the dimeric form of Rep, the WH2 domain binds to inverted repeats via the major groove, whereas the WH1 domain acts as the dimerization interface (61). Dimerization of pPS10 RepA is determined by interactions between β-sheets of the monomers that are originated due to a conformational change in the protein that involves a leucine zipper (LZ)-like motif (42). The LZ-like motif, present in several eukaryotic regulatory proteins (64), has also been found in the WH1 domain of RepA of pSC101 (65), RepE of F (43), and π of R6K (66). The dimerization interface is also localized in the WH1 of the model predicted for plasmid RK2 TrfA replication initiation protein (Fig. 2). Similar to the proposal for pPS10 RepA (42), this interface is located on an extended antiparallel β-sheet forming two hairpins (49).

Besides the indirect effect of the LZ motif in Rep protein dimerization, the LZ-like motif was characterized as responsible for Rep interaction with host replication factors. The mutations, described either in pPS10 or in the E. coli chromosome, have revealed evidence of a WH1-mediated interaction between RepA and the chromosomal initiator DnaA (67). Nonetheless, protein-protein interaction of Reps are not restricted to the LZ-like region. The best evidence for this statement is a TrfA initiator of plasmid RK2 existing in two replicationally active forms of different molecular mass. The smaller, 33-kDa protein, TrfA-33, is the result of an independent in-frame translational start in the open reading frame used for the larger, 44-kDa protein, TrfA-44 (68–70). The mutation at the N-terminal end of the trfA gene (resulting in the availability of the TrfA-33 version only) changes the host range of plasmid RK2, but the binding of DNA remains unaffected. These results demonstrate that the N-terminal end of TrfA is involved in interaction with host replication factors (71). With the use of the evolution experiment, IncP1 plasmids were shown to specialize to a novel host due to the single mutations reported at the N-terminal region of replication initiation protein TrfA (72, 73). In P. aeruginosa the TrfA-44 residues between 20 and 30 are responsible for DnaB recruiting (71), and in E. coli TrfA-33 interacts in vitro with DnaB helicase (74). It also acts with the E. coli Hda regulator, which inactivates DnaA and this way prevents overinitiation of RK2 (75). In addition, the specific motif characteristic of proteins interacting with the β clamp of E. coli DNA polymerase III was reported in TrfA and TrfA/RepA orthologues from plasmids related to RK2 and pMLb (76), but the relevance of this interaction needs to be elucidated.

The replication of iteron-containing plasmids requires the plasmid-encoded replication initiator, but the host-encoded initiation protein is also involved. The chromosomal initiator, E. coli DnaA, is composed of four functional domains (77–79). Crystallographic data obtained for the DnaA conserved core domains III/IV of the thermophilic bacterium Aquifex aeolicus revealed that, in contrast to plasmid initiators, this protein is composed of the AAA+ and DBD (DNA binding domain) domains (79) (Fig. 2). These domains are involved in DnaA oligomerization and DNA binding/remodeling functions, which are the critical aspects of origin processing. It is crucial for the interaction with ssDNA DUE at chromosomal replication origins and formation of filament structure (80–82). Since plasmid Rep does not possess an AAA+ domain responsible for nucleotide binding, it could be considered that WH domains, responsible for the binding of iterons within the dsDNA origin, can also bind ssDNA arising after dsDNA melting.

MECHANISM OF ITERON PLASMID DNA REPLICATION INITIATION

Origin Recognition

Models presenting steps of DNA replication initiation of iteron-containing plasmid and bacterial chromosomes are presented in Fig. 3. The first step of replication initiation at the plasmid origin is the formation of an initial complex facilitated by the specific interaction of Rep proteins with iterons. It has been demonstrated that replication initiation of iteron plasmids usually requires cooperative interaction of Rep monomers with iterons. pPS10 RepA as well as RK2 TrfA initiators cooperatively bind iterons at the plasmid replication origins (14, 15). Although the pPS10 RepA dimers and monomers both interact with iterons, only monomers initiate DNA replication. It is noteworthy that the existence of an early transient complex between a dimeric pPS10 RepA and an iteron half has been reported, and based on this, a model for iteron-induced dimeric pPS10 RepA dissociation and conformational activation has been proposed (61). Also, the TrfA protein functionally interacts with plasmid RK2 iterons as a monomer (12). Similar to pPS10 and RK2, the origin of the narrow host range plasmid P1 is recognized by the monomer of the P1 initiation protein RepA (83, 84). The interaction between the Rep protein and iterons has also been shown for RepE of plasmid F (85), RepA of plasmid pSC101 (86), and the π initiation protein of E. coli plasmid R6K (87). The narrow host range plasmid R6K contains
**FIGURE 3** Model of replication initiation: comparison of the processes occurring on the iteron-containing plasmid origin with the replication initiation of bacterial chromosomes. The iteron-containing plasmid origin is recognized by the plasmid-encoded initiator (Rep), which binds cooperatively to the iterons. The interaction of Rep with iterons results in the formation of an open complex and destabilization of the DNA unwinding element (DUE), which creates ssDNA. In RK2, pPS10, F, R6K, P1, and pSC101 the formation of the open complex requires cooperation of the plasmid Rep and host DnaA proteins, while at the chromosomal origin the DnaA protein is sufficient for this process. During the chromosomal origin opening DnaA forms filament on the ssDNA. Helicase delivery and loading requires interaction with the replication initiators; in addition, in *E. coli* the DnaB helicase delivery at the chromosomal oriC, as well as at the plasmid RK2 oriV, requires the DnaC accessory protein. During the RK2 replication initiation in *E. coli* the host-encoded DnaBC helicase complex is delivered to the DnaA-box sequence through interaction with DnaA, and subsequently the plasmid initiator TrfA translocates the helicase to the opened plasmid origin. The interactions between *E. coli* DnaB and the R6K π protein, F RepE, and pSC101 RepA have also been established as essential for helicase complex formation at the plasmids’ origins. The helicase unwinds the DNA double helix, and after a short RNA fragment is synthesized by a primase, a polymerase complex is assembled. Single-stranded DNA binding protein (SSB) is required for replication initiation of both chromosomal and iteron-containing plasmid DNA. The HU/IHF proteins’ contribution in DNA replication initiation was omitted in the scheme. For a detailed description see the text.

three origins of replication, α, β, and γ, but only two elements, the oriR and pir gene product π proteins, are required for a minimal replicon. The binding of seven iterons by the π initiator has been demonstrated as required for proper oriR activity (88–91). The π initiator efficiently binds to oriR iterons but not to the oriA or to the oriβ iterons (92).

**Origin Opening**

It was determined that Rep plasmid interaction with iterons generates a localized strand destabilization of DUE, leading to an open complex formation at the origin of plasmid replication. Although the involvement of the plasmid initiator is essential, the host-encoded DnaA and histone-like proteins are also required for plasmid origin opening. It was demonstrated for pPS10 that mutations within the DnaA-box sequence affect the replication in vivo (5). DnaA is mainly needed for the enhancement or stabilization of the Rep plasmid-induced open complex formation and histone-like protein (HU and/or IHF) interaction with the DNA-enhanced DNA-bending process. It was determined with KMnO4 assay that TrfA interaction with iterons generates a localized strand destabilization, and *E. coli* DnaA protein enhanced the TrfA-induced open complex (24). It was shown that this reaction occurs only in the presence of the *E. coli* HU protein (24). Similar to RK2 initiator TrfA, the binding of the RepE initiator of plasmid F to iterons induces a localized opening in the origin region, with the assistance of HU (93). The addition of DnaA increases the opening of the F plasmid origin (93) and is also required for the pSC101 origin (94) and R6K oriR (88, 89). The open complex formation by pSC101 RepA monomers in cooperation with host DnaA also requires the presence of the IHF protein (33, 95). The open complex at the R6K oriR is formed as a result of cooperative π monomers binding to the iterons and host DnaA interaction with its cognate binding sites (15). KMnO4 footprinting has shown that, in contrast to the RK2 initiator TrfA and F RepE, the PI RepA alone is not sufficient for oriR opening, but in the presence of DnaA, the addition of RepA increased the KMnO4 reactivity of the origin (96). The replication initiation of plasmid RK2 might occur in a DnaA-dependent or DnaA-independent way, depending on the host bacterium. In *E. coli* RK2 efficiently replicates and is maintained in the presence of TrfA and a host DnaA protein, while in *Pseudomonas* the longer form (44 kDa) of the replication initiator is required and DnaA is indispensable (97, 98). In *Caulobacter crescentus* both DnaA-dependent and DnaA-independent models of RK2 plasmid replication initiation are possible (99). Interestingly, the structure of DnaA protein itself might influence the host range of plasmids. Narrow-host-range plasmid pPS10 usually replicates only in the phytopathogen *Pseudomonas savastanoi* cells, due to the ability to bind DnaA-box in the pPS10 origin only by DnaA protein from this bacterium. It has been demonstrated that both the mutation in the LZ motif of pPS10 RepA and mutations in the sequence of *E. coli* DnaA promote the efficient establishment of plasmid pPS10 in the *E. coli* host (67, 100). These results suggest that mutations in plasmid and bacterial initiators that result in expanding the host range of the plasmid probably favor efficient and functional interactions between those proteins. Although the chromosomal initiator, DnaA protein, alone is insufficient for the efficient formation of an open complex at the origin of plasmids F, RK2, pSC101, and R6K (13, 24, 90, 93, 101), it has been shown to be both sufficient and indispensable in opening the AT-rich region at the origin of the bacterial chromosome (see Fig. 3). DnaA interaction with DnaA-box sequences localized within the origin of chromosomal replication (oriC) results in destabilization of the DUE, leading to open complex formation. The histone-like proteins HU and IHF stimulate the assembly of the open complex at oriC (102–104). This nucleoprotein structure formation requires ATP due to *E. coli* DnaA ATP-dependent conformational changes that promote the formation of the DnaA filament on ssDNA of DUE that is essential for the opening of the replication origin (81, 82, 105). The formation of an open complex at the plasmid origin, in contrast to *E. coli* chromosomal replication, is an ATP-independent process (24, 90, 93, 96, 106, 107), but the presence of ATP or its nonhydrolyzable analogue (ATPγS) promotes the extension of the open region (24). It is not known if plasmid Rep proteins can interact with the ssDNA and form filament structures to promote origin opening, like the DnaA replication initiator does.

**Helicase Delivery and Loading**

The origin opening generates ssDNA, which is a key element for replication complex assembly at the replication origin. The first step in the assembly of the replication complex is delivering helicase at the replication origin and loading it on ssDNA. While plasmids belonging to the IncP incompatibility group extensively use the replication proteins from the host cell for their own DNA synthesis, they utilize different host-specific mechanisms for helicase delivery and loading (71, 108, 109). Both in vivo (97, 98) and in vitro (108, 109)
analysis with the use of purified proteins from *E. coli* and *Pseudomonas* sp. revealed different host-dependent requirements for RK2 replication initiation. In *E. coli* the DnaB helicase complex with DnaC is initially recruited by DnaA protein interaction (110). The DnaA bound at DnaA-boxes located at the plasmid origin recruits host helicase (111). Then, as a result of translocation into the AT-rich region of the plasmid origin and interaction with the 33-kDa version of the plasmid replication initiator, the helicase is activated for the unwinding of the plasmid dsDNA template. The mechanism of helicase recruitment and loading during the RK2 plasmid replication in *P. aeruginosa* is DnaA-independent and relies on the 44-kDa TrfA protein, while in *P. putida* cells two variants of TrfA protein can be utilized (108, 109). The helicase complex formation during RK2 replication in *C. crescentus* cells might proceed through two different modes: DnaA-independent employing TrfA-44 and DnaA-dependent relying on the shorter version of the replication initiator (99). *In vitro* activity of *C. crescentus* DnaB helicase on the RK2 DNA template was observed in the presence of TrfA-44, and *C. crescentus* DnaA was not required for this process. *In vivo* the mini-RK2 plasmid encoding only TrfA-33 was as stably maintained as those encoding TrfA-44 or both. In contrast, TrfA-33 in cooperation with *C. crescentus* DnaA in *vitro* was unable to activate *C. crescentus* DnaB. The homologue of the *E. coli* DnaC protein needed for proper helicase loading into the open complex might be required for *C. crescentus* DnaB helicase activation. To date, no data about this kind of protein either in *Pseudomonas* or in *Caulobacter* cells have been reported, and its identification requires further investigation (99).

**Rep–Helicase Interaction**

Similar to the RK2 plasmid initiator TrfA, the interactions between other iteron-containing plasmid Rep proteins and host-encoded helicases have also been reported. *E. coli* DnaB interacts with plasmid replication initiators as was shown for the R6K π protein (112) plasmid F RepE (113) and pSC101 RepA (114). These interactions have been established as essential for helicase complex formation at the mentioned plasmid origins. A DnaB mutant, which does not interact with pSC101 RepA, was unable to activate the replication initiation at the pSC101 origin. Nonetheless, this mutant was able to support *E. coli* chromosomal replication (114). The R6K π protein and pSC101 RepA have also been shown to form complexes with *E. coli* DnaA (90, 101). Similar to R6K and pSC101, the helicase complex formation at the origins of pPS10 and P1 replicons, in addition to the plasmid-encoded initiator, depends on host DnaA protein and requires other host-encoded factors such as DnaC and HU/IFH (67, 115, 116).

The lack of ability for stable complex formation between the plasmid Rep protein and a host helicase might be one of the reasons for plasmid host range restrictions as was shown for *E. coli* plasmid F. The helicase complex at the F origin composed of the replication proteins from the nonnative hosts (*P. aeruginosa* and *P. putida*) might be formed in the presence of F initiator RepE. However, the interactions between RepE and DnaB of *P. aeruginosa* and *P. putida* were unstable, contrary to RepE interaction with *E. coli* DnaB helicase (113).

**Polymerase Complex Assembly**

Synthesis of iteron-containing plasmid DNA depends on the initial activity of a plasmid replication initiator and utilization of host replication machinery. Because plasmids do not encode their own polymerases, the host bacterium polymerase is utilized for the plasmid DNA replication. The mechanism of the events leading to the formation of the polymerase complex at the plasmid origin of replication still needs to be elucidated. Even though the DNA replication of plasmids RK2 (111), R6K (117), and F (118) has been reconstituted in *vitro* with purified proteins, and specific requirements for this reaction have been identified, the molecular mechanism for the assembly of the polymerase complex at plasmid origins is still not known. The *in vitro* analysis showed that in addition to the plasmid Rep protein, the *E. coli* proteins DnaA, HU, DnaB helicase, DnaC, SSB, DnaG primase, DNA gyrase, and Pol III holoenzyme are required for plasmid DNA synthesis. Interestingly, the specific motif (QL[S/D]LF) determining interaction with the β clamp subunit of Pol III has been identified in plasmid Rep proteins (119), though the relevance of the interaction between the β clamp and Rep proteins has not been determined. The loading of the β clamp is a complex reaction involving clamp opening and then positioning around the DNA with the use of the γ-complex (reviewed in reference 120). β clamp interaction with primed DNA is the first of subsequent events leading to polymerase complex assembly at the chromosomal origin of replication (121). Although the direct involvement of a replication initiation protein in the process of polymerase recruitment has not been reported to date, the plasmid Rep protein interaction with specific Pol III holoenzyme subunits might determine the mechanism for an efficient recruitment of host-encoded replication machinery to the plasmid origin.
CONTROL MECHANISMS OF REPLICATION IN ITERON-CONTAINING PLASMIDS

The iteron-containing plasmid replicons have evolved a number of strategies to ensure their hereditary stability and maintenance at the specific copy number. These plasmids occur in a low-copy number per bacterial cell, so their maintenance requires tight regulation of replication. The main elements involved in the regulation of these plasmid replications are iterons.

Control by Handcuffing

“Handcuffing” is a mechanism of replication inhibition observed in iteron-containing plasmids. The handcuff structure formation is based on the ability of the initiator protein to couple two ori regions located on separate plasmid molecules. The ori coupling occurs via binding of the Rep protein to iterons. This pairing of iterons is believed to cause steric hindrance to their function that prevents a new round of replication initiation (Fig. 4) (122) by inhibiting origin melting (123). It is considered that handcuffing is a major mechanism that controls the plasmid copy number.

There are three alternative models of Rep-mediated handcuffing. The first one assumes that the handcuff structures are created by the action of Rep dimers, which can bridge two DNA particles. This model was proposed for the replication protein of plasmid R6K (124, 125). Here, the major role of π dimers in the creation of R6K handcuff complexes was detected by electron microscopy (124) and ligation enhancement assays (66, 126, 127). Both of these techniques enable detection of handcuff structures in reaction to the dimeric form of the π protein. In the ligation assay, the monomeric variant of Rep was less efficient in forming ligated products (125). In contrast, the mutant of the π initiator, which binds iterons exclusively as a dimer (13), handcuffed DNA more efficiently than the wild type of the π protein. To summarize, the π dimers have a greater affinity to participate in handcuff structure creation than π monomers. The indirect evidence supporting this model is the fact of handcuffing being counteracted by molecular chaperones (DnaK-J/GrpE triad), which mediate the dissociation of dimers to monomers (123, 128).

The handcuff structure creation in the second model assumes the participation of Rep monomers in the creation of such structures by direct interactions between two arrays of Rep monomers bound to iterons in two plasmid molecules (56). This model is based on the fact that monomers of Rep initiators have a higher affinity for the iteron repeats than the dimeric forms (42, 124). Moreover, it has been reported for plasmid pPS10 that the dimeric Rep mutant is unable to create handcuff structures (56), and iterons of this plasmid play an active role in displacing the equilibrium between Rep dimers and monomers (61).

The third model of handcuff structure is a combination of the other two models. In this model, two monomers bound to the iterons of two separate plasmid molecules, are bridged by the dimer of the Rep protein. Such a model was proposed for handcuffing of plasmids RK2 (129) and F (123). The evidence for this model was obtained in a purified in vitro replication system (123). The handcuffing was found to be most proficient only when monomeric and dimeric forms of Rep protein were present simultaneously. Models involving participation of Rep protein dimers are also supported by the fact that handcuffing-defective mutants (Rep monomers of RK2 and R6K) were found to have abnormally high copy numbers (130). Therefore, it can be concluded that the handcuffing has a substantial role in iteron-mediated plasmid copy number control.

If the role of the handcuff is to block the origin and inhibit the replication, then there must be a mechanism that acts in an opposite way and “uncuffs” the coupled origin structures, which enables the reinitiation of plasmid replication. However, the mechanism of handcuff reversal is still unclear. There are results suggesting the participation of the chaperones in handcuff structure disruption (128), showing that the efficiency of handcuffing decreases in the presence of chaperones. Those results indicate that an increasing ratio of monomers over dimers is predominantly responsible for handcuffing reversal. It has also been discovered that the efficiency of handcuff structure creation increases with increasing Rep-bound iteron concentration and decreases when the reaction mixture is diluted. However, the dilution did not decrease Rep binding to the iterons (128).

Control by Auto-Repression

A high concentration of Rep protein initiator may result in more frequent, uncontrolled initiation replication events. To prevent this, the control mechanism that limits the amount of Rep initiator in the cell has to exist. Transcriptional auto-repression is a well-known mechanism for maintaining levels of gene product within narrow limits (131, 132). In many plasmid systems (F, R6K, pPS10, and pSC101), auto-repression is mediated by binding of the Rep dimer to inverted repeats located adjacent to the origin region (Fig. 5) (11, 46, 65, 133). A sequence of inverted repeats overlaps with the rep gene promoter. This kind of regulation mechanism inhibits transcription initiation starting from the rep gene promoter, and this effect is promoter-specific (63).
The affinity of the Rep dimers is higher for inverted than direct repeats, so the Reps must have specific, dimeric conformation for binding to these sites (133). Symmetrical motifs in the Rep dimer recognize the symmetry of inverted repeats (134). The mechanism of auto-regulation appears to be one of steric hindrance. When the promoter site is occupied by Rep protein, the RNA polymerase cannot displace it from the binding site. However, it has been shown that the initiator proteins can displace RNA polymerase from the promoter, and the addition of the RNA polymerase before the Rep protein does not prevent binding of Rep protein to its binding site (63, 135, 136). This inhibition of RNA polymerase binding resembles typical repressor-polymerase competition and, in this model, the Rep dimer acts as a repressor. An explanation for this auto-regulation mechanism is a higher affinity of the initiator protein for DNA sequence than that of RNA polymerase for the same sequence (136).

**Activation and Proteolysis of Rep**

As mentioned above, Rep proteins exist in monomer-dimer equilibrium, but only the monomeric form of the proteins can bind specifically to the iterons (12). Saturation of iterons in the replication origin by Rep monomers allows replication initiation. To create such a complex, conformational activation of Rep proteins is required. Dissociation of the Rep dimers into monomers simultaneously changes the conformation of the proteins and makes them competent for the iteron binding. The dissociation may be spontaneous and could occur just by dilution to low/sub-micromolar concentration. This phenomenon has been found for P1 and pSC101 plasmids (62, 83, 137). However, those monomeric forms of Rep proteins require the chaperones for refolding into the active form and for DNA binding (62, 137). The conversion of a dimer to an active monomer can also be mediated by dissociation induced by interaction with iteron-containing DNA. It has been shown that micromolar amounts of DNA, which contain a single iteron, actively induce in vitro the dissociation dimers into both monomers and conformational changes (61, 138).

The monomeric form of Rep may also arise by the action of molecular chaperons, which actively convert regulator of DNA replication by creating "handcuff" structures. Rep proteins couple origins of two separate plasmid particles in a process termed "handcuffing." In the literature suggestions of chaperone proteins' participation in the "uncuffing" process can be found, but the mechanism of the handcuff structures' reversal is still unclear. For details see the text. doi:10.1128/microbiolspec.PLAS-0026-2014.f4
In vitro techniques demonstrated that both the ClpX chaperone (139) and the ClpB/DnaK/DnaJ/GrpE system (140) activate the plasmid RK2 replication initiation protein TrfA by converting inactive dimers to an active monomer form. It has been also shown that DnaK/DnaJ/GrpE heat shock proteins are required for the activation of Rep initiators of F, R6K, and P1 plasmids (10, 60, 83, 138, 142, 143). Monomerization of the P1 plasmid initiator may also occur by the action of the ClpA protein, which alone functions as a molecular chaperone (144, 145).

The proteases are other factors affecting iteron-containing plasmid metabolism. They may influence the replication process by proteolysis of the replication initiator. In E. coli, four cytosolic proteases have been identified to date: ClpXP, ClpAP, ClpYQ, and Lon (146). The proteases limit the half-life of Rep initiator proteins, which is important for replication initiation. It has been shown that initiator proteins of bacteriophages lambda and Mu and of plasmid RK2 are proteolyzed by E. coli ClpXP protease (49, 147, 148) and that ClpAP protease degrades the Rep initiator of plasmid P1 (144). Additionally, it has been described for the TrfA initiator of plasmid RK2 that DNA is a factor that stimulates TrfA proteolysis by ClpAP and Lon proteases (149). Moreover, the Lon protease degrades the TrfA protein only in the nucleoprotein complex, while ClpAP-dependent degradation of TrfA is substantially stimulated in the presence of iteron-containing plasmid DNA (149). This specific stimulation of proteolysis could be important in terms of understanding nucleoprotein complex stability. It may also have an effect on the iteron-containing plasmid copy number, by interaction with the nucleoprotein complex handcuff structure or the other complexes of Rep protein with iteron-containing plasmid DNA.

CONCLUSIONS
All the described mechanisms that affect plasmid metabolism are intended to control the plasmid replication frequency and thereby to control the plasmid copy number. The iteron-containing plasmids, as described above, predominantly use the limitation of Rep protein concentration to control initiation of replication. The limited amount of initiator is achieved by the auto-repression mechanism. This kind of replication regulation was initially proposed to be the sole mechanism of replication control, but subsequent experiments showed the marginal effect of surplus initiator. This proved that this mechanism is insufficient (150–152). In iteron-containing plasmids, origin inactivation by handcuffing is an essential mechanism for effective replication regulation. It assumes that the iteron concentration, rather
than the level of Rep expression, determines the rate of replication. Another critical parameter that influences the replication initiation is the dimer/monomer ratio of the Rep initiator. The efficient control of plasmid replication initiation requires a combination of all the above-mentioned regulatory mechanisms. Furthermore, it has been reported that all these mechanisms need to work in concert and no single mechanism alone is able to regulate plasmid replication effectively (130). Therefore, it seems to be clear why there are multiple modes of control and that all these modes appear to be cooperative rather than mutually exclusive, which explains why they have been conserved.

ACKNOWLEDGMENTS

This work was supported by the Polish National Science Center (grant number 2012/04/A/NZ1/00048) and the Foundation for Polish Science (TEAM/2009-3/5).

Conflict of interest: We disclose no conflicts.

REFERENCES


DNA looping negatively regulates plasmid replication


