Plasmid Rolling-Circle Replication

Ruiz-Masó. J.A.[1], Machón, C[2,3]., Bordanaba-Ruiseco, L., Espinosa, M.[1], Coll, M.[2,3] and del Solar, G.[1,*]

[1] *Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9, 28040 Madrid, Spain.*

[2]*Institute for Research in Biomedicine (IRB-Barcelona), Baldiri Reixac 10-12, 08028 Barcelona, Spain.*

[3]*Institut de Biologia Molecular de Barcelona (CSIC), Baldiri Reixac 10-12, 08028 Barcelona, Spain.*

* Corresponding author: Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9, 28040 Madrid, Spain. E-mail: gdelsolar@cib.csic.es. Tel.:+ 34 91 8373112, Fax: +34 91 5360432

Keywords: rolling-circle replication, leading-strand synthesis, lagging-strand synthesis, plasmid promiscuity, Rep proteins, PcrA helicase,

SUMMARY

Plasmids are DNA entities that undergo controlled replication independent of the chromosomal DNA, a crucial step that guarantees the prevalence of the plasmid in its host. DNA replication has to cope with the incapacity of the DNA polymerases to start de novo DNA synthesis, and different replication mechanisms offer divers solutions to this problem. Rolling-circle replication (RCR) is a mechanism adopted by certain plasmids, among other genetic elements, that represents one of the simplest initiation strategies, that is, the nicking by a replication initiator protein on one parental strand to generate the primer for leading-strand initiation and a single priming site for lagging-strand synthesis. All RCR plasmid genomes consist of a number of basic elements: leading strand initiation and control, lagging strand origin, phenotypic determinants, and mobilization, generally in that order of frequency. RCR has been mainly characterized in Gram-positive bacterial plasmids, although it has also been described in Gram-negative bacterial or archaeal plasmids. Here we aim to provide an overview of the RCR plasmids lifestyle with emphasis on their characteristic traits, promiscuity, stability, utility as vectors, etc. While RCR is one of the better characterized plasmid replication mechanism, there are still many questions left unanswered, which will be pointed out along this review.

GENERAL ASPECTS OF PLASMID ROLLING-CIRCLE REPLICATION

The main features that characterize rolling-circle replication (RCR; see Fig. 1A) derive from its singular initiation mechanism, which relies on the sequence-specific cleavage, at the nick site of the double-strand origin (*dso*), of one of the parental DNA strands by an initiator Rep protein. This cleavage generates a 3'-OH end that allows the host DNA polymerases to initiate the leading strand replication. Therefore, the RCR initiation circumvents the synthesis of a primer RNA that is required in all other modes of replication of circular double-stranded (ds) DNA. Elongation of the leading strand takes place as the parental double helix is unwound by a host DNA helicase and the cleaved non-template strand is covered with the single-stranded DNA binding (SSB) protein. Since the nascent DNA is covalently attached to the parental DNA, termination of a round of leading-strand replication implies a new cleavage event at the reconstituted nick site. This reaction is assumed to be catalyzed by the same Rep molecule that carried out the initiation cleavage and remained bound to the 5' end of the parental strand while traveling along with the replication fork. A trans-esterification then occurs that joins this 5' end to the 3' end generated in the termination cleavage, hence releasing the displaced parental strand as a circular single-stranded (ss) DNA. This replicative intermediate serves as the

template for the synthesis of the lagging strand, which depends solely on host-encoded enzymes and is initiated from a highly-structured region of the ssDNA, termed the single-strand origin (*sso*). Thus, the entire process of asymmetric RCR yields, in two separate steps (this is what asymmetric refers to), two circular dsDNAs containing either the newly-synthesized leading or lagging strand and the complementary parental template strand. The DNA ligase and gyrase of the host cell will convert next the new daughter DNA molecules in supercoiled forms indistinguishable from the rest of the plasmid pool. Generation of the ssDNA replicative intermediates is the hallmark of RCR, and, in fact, detection of intracellular strand-specific plasmid ssDNA provides valuable clues about whether a given plasmid replicates by the rolling-circle mechanism (1, 2).

The basic catalytic mechanism operating in initiation and termination of RCR, i.e. the cleavage and rejoining of ssDNA using an active-site Tyr that forms a transient 5'-phosphotyrosine bond with the cleaved DNA, is involved in a range of processes that take place in mobile genetic elements in all three domains of life. The enzymes that exhibit this catalytic mechanism are mainly included in the widespread HUH endonuclease superfamily and have key roles in replication of plasmids, bacteriophages and plant and animal viruses, in plasmid conjugative transfer and in transposition (3). RCR was discovered in ssDNA coliphage ΦX174 some 45 years ago (4-6). The pioneer characterization of gene A protein made the initiator of ΦX174 RCR the first member of the HUH endonuclease superfamily (7-10).

Plasmid RCR was first evidenced for the *Staphylococcus aureus* plasmid pT181 based on the characterization of the origin-specific nicking-closing activity of the purified pT181-encoded RepC protein (11). Shortly afterwards, several other small plasmids from staphylococci, bacilli, streptococci and streptomyces were also found to replicate by the RCR mechanism (12-14), which led to the assumption that most, if not all, small multicopy plasmids in Gram-positive bacteria used RCR. However, this premise proved inaccurate, as some small plasmids isolated from Gram-positive organisms were later reported to replicate by the theta mode (1). Moreover, although RCR plasmids are particularly abundant in Gram-positive bacteria, they have also been identified in various Gram-negative organisms, in archaea, and in mitochondria of the higher plant *Chenopodium album* (1, 2, 15).

Natural RCR plasmids range in size from as low as the 846 bp of the *Thermotoga* plasmids pRQ7, pMC24 and pRKU1 (16-18) to the almost 30 kb of pCG4 from *Corynebacterium glutamicum* (19). The nearly identical plasmids pRQ7, pMC24 and pRKU1 are the smallest ones found so far and

would consist of only the basic replicon, i.e., the backbone regions involved in replication and copy-number control. The basic replicon of RCR plasmids should include an essential module containing the *dso* and the genes that encode the initiator Rep protein and the replication control element(s), as well as at least one host-recognized *sso*, which, although not strictly essential, provides efficient synthesis of the lagging strand and hence it is present in all natural RCR plasmids (Fig. 2). Homology in the essential module of the basic replicon has actually been the criterion used to classify RCR plasmids into replicon families (see below).

Apart from the basic replicon, some larger RCR plasmids contain additional backbone genes and elements that contribute to their maintenance or help them transfer between host cells (Fig. 2). Of special relevance, because of its frequent presence in RCR plasmids, is the MOB module, which is involved in the conjugative mobilization of the plasmid and consists of the transfer origin (*oriT*) and the *mob* gene(s) encoding the relaxase protein and, in some cases, auxiliary proteins (20). The apparent lack of active partition systems in RCR plasmids is consistent with the medium copy-number (10-30 per chromosome equivalent) that they exhibit in their natural hosts. This feature ensures the stable inheritance of RCR plasmids by only random segregation to the daughter cells, providing that the replication control system efficiently corrects fluctuations of the plasmid copy number in single cells and that the plasmid molecules are maintained as individual copies. In this sense, it is intriguing the presence of homologous to components of toxin-antitoxin (TA) systems in some RCR plasmids (21). Noteworthily, whereas the TA systems were first proposed to play a role in plasmid stability through post-segregational killing of plasmid-free cells, the more recent “competition hypothesis” postulates that acquisition of these modules allows plasmids to exclude competing TA-free plasmids (22-24).

Some RCR plasmids also carry accessory genes encoding functions that can benefit the host cell under special conditions, thus reflecting the adaptation of the bacteria to their environment (Fig. 2). Antibiotic resistance determinants are among the most frequent traits encoded by RCR plasmids isolated from a variety of bacteria (25). Other particular accessory genes have been found to be relatively abundant in RCR plasmids from a given host. This is the case of small heat shock protein (*shsp*) genes carried by *Streptococcus thermophilus* plasmids belonging to the pC194 replicon family (26, 27). The presence of *shsp*-containing plasmids has been reported to increase cell survival at the high temperatures reached during different stages of fermentations in the dairy industry (27). Another striking example is the presence, in some *Bacillus thuringiensis* plasmids, of ORFs encoding collagen-

like proteins that are thought to play a role in aggregation formation or in adherence to other cells or substrates (28).

RCR plasmids are considered to contain promiscuous replicons, as many of them have been shown to replicate in species, genera, or even phyla other than those from which they were isolated (25). The simplicity of the RCR initiation, with only the plasmid-encoded Rep protein participating in recognition of the origin and priming of the leading strand synthesis, may underlie the usual promiscuity of these plasmids. The broadness of the host range of RCR plasmids would depend on the balanced expression of their essential genes involved in initiation and control of replication as well as on the formation of a functional Rep-host helicase complex that can extensively unwind the plasmid DNA in a variety of bacteria (29-31). The broad host range of RCR plasmids is best exemplified by the pMV158-family prototype, which was initially isolated from *Streptococcus agalactiae* and subsequently transferred to a variety of Firmicutes (several *Streptococcus* and *Bacillus* species, *Listeria, S. aureus, Lactococcus lactis, Enterococcus faecalis, Clostridium*), Actinobacteria (*C. glutamicum, Brevibacterium*) and to the γ-proteobacterium *Escherichia coli*. Moreover, the fact that members of each replicon family have been isolated from a variety of bacteria suggests the promiscuity of the ancestors from which these plasmids derive. In turn, plasmid adaptation to a new host can lead to the narrowing of the host range of the adapted plasmid. This seems to be the case of two *Mycoplasma mycoides* plasmids of the pMV158-replicon family, namely pADB201 and pKMK1, whose *rep* genes contain at least one UGA codon, which encodes tryptophan in this bacterium while it is a stop codon in other bacteria, so that the host range of these plasmids is restricted to *Mycoplasma* species (32).

Due to their general smallness, high copy number and promiscuity, RCR plasmids appear to be well suited for the construction of vectors for gene cloning and expression, provided a functional *sso* is present in order to minimize the generation of the recombinogenic ssDNA intermediates, which can lead to structural and segregational plasmid instability (33-37). Nevertheless, it has been reported that cloning of heterologous DNA in RCR plasmid vectors can result in the generation of linear high-molecular-weight (HMW) plasmid multimers in relative amounts that correlate positively with the size of the DNA insert (38, 39). The formation of HMW by RCR plasmids has also been implicated in both structural (40) and segregational (41) instability. The generation of HMW plasmid DNA was at first related to a replication defect, as plasmids lacking *sso* were prone to accumulate HMW DNA (42). Accumulation of HMW plasmid DNA was enhanced in the absence of the ExoV enzyme (RecBCD in Gram-negative or AddAB in Gram-positive bacteria (41). Despite the potential instability problems,

vectors based on RCR plasmids have been developed and successfully used in pneumococci, enterococci, lactococci and corynebacteria (43-45), where genetic and biotechnological tools are scarce and hence welcome. It is worth mentioning that most of the non-integrative plasmid vectors available in *Streptococcus pneumoniae* are based on pMV158 and that inducible expression vector pLS1ROM and recombinant pLS1ROM-GFP (containing the *gfp* gene, encoding the *Aequorea victoria* green fluorescent protein, cloned under control of the maltose-inducible P_M promoter) have proved to be structural and segregationally stable in pneumococcus, even under induction conditions (45). Similarly, most of the autonomously-replicating vectors for the industrial microorganism *C. glutamicum* are based on plasmids pBL1, pCG1 and pGA1 from *C. glutamicum* or on the broad-host-range plasmid pNG2 from *Corynebacterium diphtheriae*, all of them replicating by the rolling circle mode (46). These RCR plasmid vectors were found to be stably maintained in *C. glutamicum* cells grown under nonselective conditions (47).

An aspect of recognized relevance when pursuing the biotechnological use of plasmid vectors is the metabolic cost that carriage of these extrachromosomal elements imposes on the host, since a significant burden can lead to the overgrowth of the culture by plasmid-free cells even though plasmid inheritance is quite stable. Little information is available on the burden caused by RCR plasmids, as this subject has only been analyzed for the pMV158 replicon. Small (4.4 kb), medium-copy-number (~20 copies per chromosome equivalent) pMV158 derivatives that are stably inherited in pneumococcus and harbor an *sso* element efficiently recognized in this host, slightly burden the *S. pneumoniae* cells causing a 7-8% increase in the bacterial doubling time (48). Fitness impairment of pneumococcal cells harboring pMV158-derivatives has not been found, nevertheless, to negatively affect the segregational stability of pLS1ROM and pLS1ROM-GFP (45).

This chapter aims to provide an updated review of the major findings in the study of the RCR plasmids, and to highlight the pending questions and challenges for the detailed understanding of this kind of plasmid replication. Most of these issues have been dealt with in previous reviews on this subject (1, 25, 32, 49, 50).

Apart from the above-referenced asymmetric RCR, which is initiated by the Rep-mediated cleavage of one parental strand, a different, recombination-dependent replication mechanism that also leads to σ-shaped circular intermediates consisting of a circular DNA attached to a growing linear DNA has been reported to play an essential role during the replication cycle of many dsDNA viruses. Single origin-dependent replication of bacterial genomes and of many dsDNA viruses with circular

genomes proceeds by the θ (circle to circle) mechanism. The trade-off between different DNA transactions could lead to the stall or collapse of the replication machinery, so that origin-independent remodeling and assembly of a new replisome at the stalled fork will be required in order to restart the replication process. In dsDNA viruses (e.g. bacteriophage lambda, SPP1, etc), replication restart becomes dependent on recombination proteins with a switch from the origin-mediated θ type to a σ type recombination-dependent replication. The replication shift from θ to σ generates the concatemeric viral DNA substrate needed to produce mature viral particles. This RCR-like σ mode has been reviewed by (51) and will not be addressed here.

THE DOUBLE-STRAND ORIGIN

Replication of the leading strand of RCR plasmids initiates and proceeds in a unidirectional manner from their *dso*, a plasmid DNA region highly specific of its cognate initiator protein that contains the sequences involved in the initiation and termination of the leading strand. The *dso*, along with the *rep* gene and the control elements, is part of an essential module that harbors the functions for plasmid replication. Based on the homologies found in this essential module up to seventeen different RCR plasmid families have been defined. Only three of these plasmid families have been studied in depth, their prototypes being the staphylococcal plasmids pT181/pC221 [(2) and references therein; (52)] and pC194/pUB110 (53), and the streptococcal plasmid pMV158 (1). The following plasmid families have been also studied although less thoroughly: the staphylococcal plasmid pSN2 family (54), the pBL1 and pCG1 plasmid families from *C. glutamicum* [(55) and references therein], the pSTK-1 and pTX14-2 plasmid families from *B. thuringiensis* [(28) and references therein] and the pGRB1 (56) and pGT5 (57) plasmid families from archaea.

The *dso*s of RCR plasmids can be found located upstream of the *rep* gene (pC194, pMV158 and pSN2 families), embedded within the 5' portion (pT181 family) or the 3' portion (pCG1 family) of the sequence coding their respective Rep proteins or even downstream from the *rep* gene stop codon (pTX14-2 family). The *dso* can be physically and functionally divided into two different regions, namely *bind*, which contains the specific binding sequence for the initiator protein, and *nic*, where Rep specifically cleaves the DNA at the nick site. The two loci can be either adjacent to each other (pT181 and pC194 families) or separated by a spacer region of up to 100 bp (pMV158 family) (Fig. 1B). The *dso*s of plasmids of the same family are characterized by a high degree of conservation in the *nic* region and by the presence of a less well-conserved *bind* region. In fact, Rep proteins encoded by

different plasmids of the same family can perform *in vitro* the nicking-closing reaction on the *dso*s of all the plasmids belonging to the same family but there is little or no cross-interaction with the *bind* region, which is indicative of the replicon-specificity of the *bind* locus. Interestingly, the pT181-encoded RepC initiator has been shown to drive *in vitro* replication of plasmid pC221, although this was greatly reduced if a competing pT181-*dso* was present (58). In spite of such an *in vitro* recognition and extensive homologies of the Rep proteins and the *dso*s of pT181 and pC221, there is no cross-reactivity between the Rep proteins and the *dso*s of these plasmids *in vivo*, unless the Rep proteins are overproduced (59).

In the case of the pMV158 family, the DNA sequence of the *bind* locus was reported to consist of two or three direct repeats (DR), whose length ranged from 5 to 21 bp (60), separated from the nick sequence by an intervening sequence of variable length (Fig. 1B). The *dso* of pJB01, a member of the pE194 subfamily, contains as the Rep-binding site three 7-bp non-tandem DR located 77 bp downstream from the nick site (61). Interestingly, the existence of distant DRs has not been elucidated in some plasmids of this subfamily (unpublished observation). The role of the different regions of the pMV158-*dso* in the interaction with the plasmid-encoded RepB initiator protein has been addressed in a systematic study (35, 62-65). RepB binds with high affinity to the *bind* locus, which is constituted by three 11-bp tandem DR located 84 bp downstream from the nick site. These repeats do not constitute an incompatibility determinant toward pMV158 and seem to be essential for plasmid *in vivo* replication but not for *in vitro* relaxation of supercoiled DNA mediated by RepB. A second RepB binding site is located in a region around the nick site, within the *nic* locus. Characterization of the relative affinity of RepB for the *bind* and *nic* loci revealed that the three DR of the *bind* locus constitute the primary binding site, whereas the weaker binding of RepB to the *nic* locus could be involved in recognition of the nick site during initiation of replication (65). In plasmids of the pT181 and pC194 families, the DNA sequences of the *bind* (IRIII) and *nic* (IRII) loci are located in contiguous inverted repeats (Fig. 1B). In pT181, both the spacing and the phasing of IRII to IRIII are crucial for origin functionality (66). In addition, the proximal arm and the central part of the IRIII are important for sequence specific recognition (66). A similar picture is found in plasmids of the pC194 family.

A typical feature of the *nic* regions is the presence of secondary structures as hairpins and cruciform. The Rep nick sequence is generally located on an unpaired region within these hairpins, as exemplified by IRII of pT181 and inverted repeat (IR)-I of pMV158, which accounts for the requirement of plasmid DNA supercoiling in order to render the cleavage sequence a suitable ssDNA substrate for

replication (67-69). The presence of secondary structures is likely to be involved in efficient recruitment and utilization of the initiator protein. Additionally, binding of the initiator protein to the *nic* locus could promote the melting of the substrate nick sequence. This seems to be the case in pMV158, where the extrusion frequency of the cruciform involving IR-I is very low at the growth temperature of the plasmid host (37°C) (70). *In vitro* footprinting experiments performed with supercoiled pMV158 DNA showed that binding of RepB to the *nic* locus promotes the extrusion of the IR-I cruciform, which in turn indicates that initiation of replication would take place only when specific binding of RepB occurs (65). Genetic analysis of the pC194 *dso* pointed to the existence of a hairpin located downstream of the nick site (71) that was shown to be important for replication of the plasmid (72). In contrast, RepU, the initiator protein of pUB110, does not require the presence of hairpins for efficient recognition of the *oriU*. Hairpin II, located downstream from the nick site, seems to be dispensable for initiation of replication of pUB110, although its absence provokes the accumulation of multimers, which is indicative of the involvement of this structure in termination of replication (73).

Out of the three plasmid family prototypes that have been studied in more detail (pT181, pC194 and pMV158), available information regarding the characteristics of the *dso* is limited to a few plasmids of different families. In the *dso* sequences of pJV1, pIJ101 and pSN22, three plasmids belonging to the same subfamily inside the pC194 family, three conserved regions were identified: the conserved region I of about 100 bp, located upstream of the nick sequence, which is essential for replication; the nick sequence in region II; and the region III overlapping with the *rep* start codon. A co-integration experiment between pJV1 and a pIJ101-derived vector allowed the identification of the nick site within the sequence 5'-CTAGGTA-3' of pJV1, located 159 bp upstream of the start codon of the corresponding *rep* gene (74). In the case of pIJ101 and pSN22, the putative nick sequence identified in region II was: 5'-CTTGGGA-3', which is not identical to that of the pC194 group (5'-CTTGATA-3') (75, 76).

Plasmid pGA1 isolated from *C. glutamicum* is the best studied plasmid of the pCG1 family. The location of the *dso* was analyzed using the runoff DNA synthesis assay. The site- and strand-specific breakage of pGA1 dsDNA occurred within the nucleotide sequence 5'-CTGG/AT-3' (where / indicates the nick site) in the distal part of the pGA1 *rep* gene, which is an atypical position among RCR plasmids (77). The group of RCR plasmids isolated from hyperthermophilic archaea is constituted by pGT5, pRT1 and pTN1. Plasmid pGT5 is the first plasmid to be isolated from a hyperthermophilic organism, *Pyrococcus abyssi*, and presents similarities to plasmids from the pC194

family. A sequence of 11 nucleotides identical to that in the *dso* of pC194 and related plasmids is located at the 5' region of the *rep* gene. The presence of pGT5 ssDNA (corresponding to the putative plus strand) in cell extracts of *P. abyssi*, strongly suggests that pGT5 replicates via a RCR mechanism (57). Plasmid pTN1, isolated from *Thermococcus*, shows an identical nick sequence to that of pGT5 (5′-TTATCTTGATA-3′) and is also located in the 5′ region of the *rep74* gene (78). Similarly, the *dso* of pRT1, isolated from *Pyrococcus* sp. strain JT1, exhibited significant identity to the *dso*s of both pGT5 and pC194, further suggesting that the replication mode of this plasmid is via the RCR mechanism, which was confirmed with the detection of ssDNA replication intermediates (79). Position of the *dso* nick site in the plasmid pZMX201, considered as the prototype of a family of plasmids isolated from halophilic archaea, was precisely determined in the sequence 5'-TCTC/GGC-3' (where / denotes the nick site), which is conserved among the members of the family. Although the heptameric sequence is usually located in the stem region of an imperfect hairpin structure that, in turn, could serve as a target for the recognition by the Rep protein, the nick site lies in an unpaired position or near an unpaired nucleotide (56). In addition, the use of a hybrid plasmid system revealed the role of the nucleotides of the conserved nick sequence in the RCR initiation and termination process.

THE REPLICATION INITIATOR REP PROTEINS

As mentioned above, Rep proteins involved in plasmid RCR initiation are mainly included in a vast superfamily of HUH endonucleases (U stands for a bulky hydrophobic residue) that catalyze cleavage and ligation of ssDNA by using particular recognition and reaction mechanisms. Besides the Rep class, which also includes Reps from ssDNA coliphages and animal and plant viruses, proteins involved in conjugative plasmid transfer (Mob class or relaxases) and in DNA transposition (transposases) also belong to this superfamily, and all of them exhibit a familial relationship based on several conserved protein motifs (80). The two most relevant motifs are the metal binding HUH motif, composed of two His residues separated by a bulky hydrophobic residue, and the catalytic motif containing either one or two Tyr residues separated by several amino acids (Fig. 3). Characterization of the biochemical activities of several plasmid RCR initiator proteins has contributed to a better understanding of the molecular events during the initiation and termination of RCR. Initiation of plasmid RCR requires Rep-mediated nicking within the unpaired nick sequence of the *nic* locus in supercoiled DNA. Rep endonucleases exhibit DNA strand-transfer enzymatic activity and catalyze cleavage and re-joining of ssDNA using an active-site Tyr residue to make a transient 5′-

phosphotyrosine bond with the DNA substrate and a free 3′-OH at the cleavage site. Moreover, the resultant 3'-OH not only serves to prime replication but also can act as the nucleophile for strand transfer to resolve the phosphotyrosine intermediate in the termination step of RCR.

The divalent metal ion required for the activity of the HUH enzymes probably coordinates one of the oxygen atoms of the scissile DNA phosphate, polarizing it and facilitating the nucleophilic attack of the hydroxyl group of the catalytic Tyr (81, 82). Curiously, the HUH motif for metal binding is not present in plasmids of the pT181 family or in ssDNA filamentous phages, although in both cases the presence of a divalent metal ion is required for the enzymatic activity. In the case of the pT181 family, it has been shown that the reactive Tyr188 of RepD, the initiator protein of pC221, cleaves the phosphodiester bond 5'-ApT-3' and remains covalently attached to the 5'-P end generated by the cleavage reaction. The importance of the tyrosyl hydroxyl group was confirmed by substitution of Tyr188 by Phe, since this protein variant retains the sequence-specific DNA-binding activities of wild-type RepD but is unable to attach covalently to the replication origin or participate in the nicking-closing reaction *in vitro* (52). Similarly, the initiator protein RepA of pC194 forms a 5'-phosphotyrosyl DNA link at the initiation step and mutations in the catalytic Tyr214 drastically reduced its catalytic activity without affecting RepA binding affinity. RepB of pMV158 does not seem to generate a stable covalent tyrosylphosphodiester bond with its DNA target (83). However, there is experimental evidence showing that RepB, like the filamentous phage gpII, forms a transient covalent complex with the 5'-P end of the cleaved DNA (64, 84). A singular case among Rep proteins of RCR plasmids is represented by Rep75 of pGT5, a protein that exhibits a highly thermophilic nicking-closing activity *in vitro* combined with an unusual site-specific nucleotidyl terminal transferase (NTT) activity which has not been described for proteins of this type (85). Substitution of the catalytic Tyr448 by Phe caused a severe reduction of the nicking-closing activity *in vitro* and prevented the formation of the 5'-phosphotyrosyl DNA link without affecting the dsDNA binding activity of Rep75 (86). A second critical residue for the activity of Rep75 is Arg451, as the protein variant Rep75-Arg451Leu exhibits a reduced closing activity and has completely lost the NTT activity (86).

A modular structure based on the presence of at least two domains involved in origin nicking and specific recognition of the *dso* sequence has been assumed to be a general feature of the initiator proteins of RCR plasmids (87-90). This assumption could be valid for plasmids of the pT181 family as the sequence-specific DNA binding and DNA relaxation activities of RepC, the initiator of pT181, are mutationally separable and lie on distant protein regions (88). Taking advantage of this property, the

identification of the role of individual monomers of pT181-RepC in RCR was addressed by generating heterodimers of the initiator containing a combination of wild type, DNA binding and nicking mutants (91). The results demonstrated that a single monomer of RepC is sufficient for origin-specific binding and nicking. In addition, the monomer involved in sequence-specific binding to the *dso* must also nick the DNA to initiate replication (91). In plasmids of the pMV158 family, a similar assumption was proposed on the basis of the higher degree of amino acid identity found at the N-terminal region than at the C terminus of their Rep proteins, which suggests that their conserved N-terminal moiety would be involved in endonuclease activities whereas the C termini would be involved in specific *dso* recognition (32). However, resolution of the 3D structure of RepB, the initiator protein of pMV158, by X-ray crystallography and by image reconstruction methods showed a different functional domain organization from that initially proposed. RepB is the first and to date the only published example of an atomic structure of a Rep protein from RCR plasmids or bacteriophages (3, 92). Purified full-length native RepB behaves as a hexamer in solution, as observed in analytical ultracentrifugation assays (93), and crystallized in the same oligomeric state, forming a toroidal homohexameric ring (92). Each RepB protomer comprises an N-terminal origin binding domain (OBD), which retains the DNA-binding capabilities as well as the nuclease and strand-transfer activities of RepB, and a C-terminal oligomerization domain (OD) that forms a cylinder with a 6-fold symmetry in the hexamer (Fig. 4A). Separate expression and purification of the RepB OD and OBD domains demonstrated that the enzymatic and dsDNA binding activities and the oligomerization potential can be uncoupled and confirmed the essentiality of OD for hexamerization (D. R. Boer, J. A. Ruiz-Masó, M. Rueda, M. Pethoukov, D. I. Svergun, M. Espinosa, M. Orozco, G. Del Solar and M. Coll, submitted for publication). Resolution of the 3D structure of the catalytic N-terminal domain of RepB confirmed the involvement of the conserved motifs in the enzymatic activity of the protein. Based on the protein sequence alignment of Reps of the pMV158 family, a set of five conserved motifs were designated for this particular family of proteins (32). Motifs I, III and IV would correspond to the conserved motifs 1, 2 and 3 of Ilyina and Koonin (80), whereas motifs II and V were new. A divalent metal ion Mn^{2+}, required for the catalytic activity of RepB, is located at the active site coordinated by four amino acid ligands and a single-solvent molecule. Residues His39 and Asp42, included in motif II and with no function assigned initially, together with the His55 and His57 residues of motif III (HUH), provide the four amino acidic ligands necessary for metal ion coordination. The conserved catalytic Tyr99 (motif IV) and the conserved Tyr115 (motif V), which interacts with the metal ligand Asp42 of motif II, complete the list of

the conserved residues located in the active center. Residues in motif I are outside the active center and seem to play an structural role by forming part of the strand β1.

In spite of the presence of the *bind* locus dsDNA sequence in the co-crystals of pMV158-RepB, it was not possible to extract structural information on the DNA or on the protein elements interacting with it. Nevertheless, the presence of a single and large electropositive region that covers the outer surface and crevice at the N-terminal of the RepB hexamer was consistent with binding to only one *bind* locus-containing DNA molecule (Fig. 4B). In addition, by using EM reconstruction methods, it was possible to observe a density that occludes the N-terminal region crevice in RepB hexameric particles exposed to the *bind* locus (92). These observations point to the location of the DNA-interacting surface in the N-terminal outer region and crevice of the hexamer. In fact, site directed mutagenesis to some positively charged side chains of helix α2 resulted in OBD protein variants with a reduced dsDNA binding affinity but with an intact nicking-closing activity (92).

Two structures of the pMV158-RepB hexamer were obtained from crystals belonging to a trigonal and a tetragonal form, respectively. In both structures, a short hinge region connects the two protein domains. However, the OBDs do not follow the six-fold symmetry of the ODs and the positions of the OBDs with respect to the ODs change significantly when comparing the two crystal forms. In fact, analysis of the two different structures revealed that the N-terminal OBD domains are found in nine distinct orientations relative to the OD. The conformation plasticity of RepB has been explored by combining different techniques like X-ray crystallography, SAXS, sedimentation experiments and molecular simulations (D. R. Boer, J. A. Ruiz-Masó, M. Rueda, M. Pethoukov, D. I. Svergun, M. Espinosa, M. Orozco, G. Del Solar and M. Coll, submitted for publication). These studies revealed that the position and freedom of movement of the OBDs is mainly determined by the OBD-OD hinge region, since contacts between adjacent OBDs do not seem to play an important role in fixing their positions. A consequence of the loosely-coupled domain arrangement observed in the RepB hexamer is the high level of conformational freedom of the OBDs that is probably important for formation of a functional replisome.

The hexameric state of purified RepB is unique among plasmid replication initiators characterized so far. Most of these proteins are purified as monomers, as in the case of the Rep proteins of the pC194 family (53), or dimers, as Reps of the pT181 family (52, 94-96). Interestingly, some of these proteins tend to form multimers upon binding to the *ori* DNA, suggesting that oligomerization could be involved in some of the biochemical activities of the initiators (96, 97).

The existence of tight interactions between the pMV158-RepB all-helical ODs enables the formation of a toroidal ring with near six-fold rotational symmetry and an inner surface narrowing down from a maximum diameter of ~20 Å to a minimum diameter of ~ 13 Å (92). The search for fold similarities in the ODs from different atomic structures, in combination with a comparative analysis of Rep sequences, confirmed the existence of a RepB-like all-helical OD domain responsible for oligomerization in viral Rep proteins and replication initiators from plasmids of the pMV158 family (D. R. Boer, J. A. Ruiz-Masó, M. Rueda, M. Pethoukov, D. I. Svergun, M. Espinosa, M. Orozco, G. del Solar and M. Coll, submitted for publication). The combination of hexameric OD ring and hinge-connected OBDs has been proposed as a general feature of hexameric replication initiators of the pMV158 family, although this configuration can also be found among hexameric initiators of animal and plant viruses (Fig. 3) (D. R. Boer, J. A. Ruiz-Masó, M. Rueda, M. Pethoukov, D. I. Svergun, M. Espinosa, M. Orozco, G. del Solar and M. Coll, submitted for publication). In fact, the domain organization of RepB resembles that of the viral replication initiators, which suggests an evolutionary link between the two protein families. Structural similarities between the α-helical hexamerization domain of RepB and the equivalent domain of the papillomavirus E1 helicase has been reported (92). The crystal structure of E1 comprises the OD and helicase domains and was obtained in complex with a short ssDNA oligonucleotide in the central channel (Fig. 3) (98). Therefore, the resemblance of the E1 and RepB hexameric rings favors a mechanistic model in which the RepB ring might close around a DNA region that could have melted on assembly of the protein and/or cleavage at the nick site, thus encircling one of the plasmid strands within the central channel. Subsequent recruitment of a host helicase (perhaps PcrA) would allow further unwinding of the DNA and the concomitant progression of the hexamer along the plasmid. Enclosure of one plasmid strand may confer high processivity to the RepB/helicase/DNA polymerase replisome complex, thereby allowing replication of pMV158 in a broad range of bacterial hosts.

Ilyina and Koonin (1992) hypothesized that geminiviruses would descend from bacterial replicons according to the limited sequence similarity of the three conserved motifs of the RCR Reps of geminiviruses and certain plasmids of Gram-positive bacteria. However, more recent phylogenetic analyses of various RCR Reps suggest that Rep proteins of geminiviruses share a most recent common ancestor with Reps encoded on plasmids of phytoplasmas (99). Several plasmids from phytoplasma have been sequenced and characterized to some extent. Rep protein of pOYW, a plasmid isolated from onion yellow phytoplasma, is characterized by a chimerical nature containing a

RCR plasmid Rep-like domain in the N-terminal region, which displays four out of the five conserved motifs characteristic of the pMV158-Rep family of proteins, and a virus-like helicase domain in the C-terminal region, which also includes Walker A and B nucleotide-binding motifs and shows great similarity to Reps from eukaryotic small DNA viruses or some RNA viruses (100). Similarly, Rep protein of plasmid pCPa, isolated from *Candidatus* Phytoplasma australiense, also shows four out of the five protein motifs characteristic of the pMV158-Rep family of proteins in its N-terminal region and Walker A and B motifs typically found in geminiviruses in its C-terminal region (Fig. 3) (101).

CONTROL OF ROLLING-CIRCLE REPLICATION

Mechanisms for controlling synthesis of the Rep protein

Control of plasmid RCR is exerted via regulation of the synthesis of the replication initiator. Other general control mechanisms as the origin inactivation by handcuffing, which involves coupling between plasmid molecules through Rep proteins bound to multiple initiator binding sites (iterons), has not been reported in RCR plasmids. Availability of the Rep protein determines the frequency of the leading-strand initiation, which is rate-limiting for the plasmid RCR process, and therefore the *rep* gene encoding the initiator is subjected to strict control. Transcriptional and translational inhibition mechanisms of *rep* expression are not specific of RCR plasmids and have been reviewed in (1, 102, 103). Two different classes of replication control systems have been described in RCR plasmids, those that only use antisense RNAs and those involving an antisense RNA in combination with a transcriptional repressor protein. An example of the first class is found in the replication control of pT181 by an antisense RNA-mediated transcriptional attenuation mechanism (104), which is also expected to control replication in most of the plasmids of the family. In this system, the antisense RNA transcribed from the *cop* region targets the *rep*-mRNA encoding the initiator protein and blocks its expression. The main consequence of the antisense-target interaction is the premature termination of the *rep*-mRNA transcription due to the formation a ρ-independent transcriptional terminator (attenuator) located just 5' to the *repC* start codon. Release of the antisense RNA from the mRNA target permits refolding and, therefore, silencing of the transcriptional terminator (104).

A second class of copy-number control system reported in RCR plasmids involves two trans-acting plasmid elements, namely a transcriptional repressor and an antisense RNA, that are involved in controlling the synthesis of the initiator at the transcriptional and translational level, respectively. This dual mechanism was first discovered in pMV158 and seems to be widespread among plasmids

from Gram-positive bacteria, including the RCR plasmids of the pMV158 family (102, 105). In pMV158, the transcriptional repressor CopG binds to its own promoter and represses the transcription of the *copG-repB* operon. The mechanism of transcriptional repression mediated by CopG has been elucidated by analyzing the interactions between the RNA polymerase (RNAP), CopG and the promoter (106). CopG is able to bind specifically and cooperatively to its operator, which overlaps with the regulated promoter, in such a way that it does not only prevent the binding of the RNAP, but also displaces efficiently the polymerase bound to the promoter. The second control element, the *rnaII* gene encoding RNAII, overlaps the intergenic region of the pMV158 *copG-repB* operon and is transcribed in the opposite direction to it. Therefore, the entire sequence of antisense RNAII is complementary to a *copG-repB* mRNA target region that contains the translation initiation signals for the essential *repB* gene (107). Small antisense RNAII (48-nt long) consists of a single-stranded 5' tail followed by a hairpin and a 3' poly(U) tail, the latter two elements constituting a very efficient intrinsic transcription terminator (108, 109). Additionally, RNAII is able to inhibit plasmid replication *in trans*, it has a short (1-2 min) half-life (Acebo et al., unpublished results), and it determines a strong incompatibility against pMV158, thus matching the features required for an efficient plasmid replication control element (110, 111). In the antisense RNA-mediated systems of RCR plasmid replication control where the entire process has been dissected, the formation of the RNA/RNA duplex seems to be initiated by a "kissing" step that involves reversible base-pairing between complementary hairpin loops (112). However, wild-type antisense RNAII supplied *in trans* retains its inhibitory capacity on derivatives of pMV158 that encode a mutant *copG-repB* mRNA lacking the hairpin complementary to that in RNAII. These findings suggested that formation of a kissing complex was not strictly required for the RNA pairing mechanism of the pMV158 control system (111), and led to the proposal that *copG-repB* mRNA/RNAII binding initiates *via* a loop-linear pairing scheme (113). The existence of an antisense RNA that controls replication of pJB01 (a member of the pMV158 family) has also been proved, and the involvement of different regions of this RNA in translation inhibition of the *rep* gene was studied by mutational analysis. Copy number inspection of the various mutant plasmids suggested that the entire secondary structure of the antisense RNA was important for interaction with the target mRNA (114).

In addition to this dual control of *rep* expression that senses and corrects fluctuations in plasmid copy number, proper availability of Rep also relies on the adequate functionality of the transcription and translation initiation regulatory signals. In pMV158, translation of *repB* was proposed

to be initiated from what was termed an atypical ribosome binding site (ARBS) located in the small intergenic region of the *copG-repB* operon encoding the antisense RNAII. The initiation signals that regulate *repB* translation were identified and characterized in (115). Translation of *repB* relies on its own initiation signals, which rules out a possible mechanism of translational coupling to the upstream *copG* gene. Only changes in the sequences involving the ARBS proximal box, thus named because of its proximity to the *repB* start codon, and the region 3'-adjacent to the ARBS proximal box result in a significant reduction of *in vitro* synthesis of *repB*. The results of this study demonstrate the importance of the region immediately upstream of the *repB* start codon in the efficiency of translation of *repB* and call into question the functionality of the postulated ARBS. Moreover, the conclusions of this study could be applicable to the majority of the plasmids of the pMV158 family due to the high degree of identity found at the *rep* translation initiation regions. The possible involvement of those features of the translation initiation signals of the *rep* genes in an additional mechanism to regulate the level of the Rep protein remains to be determined.

The existence of a singular element that positively influences the plasmid copy number and stability has been reported in plasmid pGA1 from *C. glutamicum*. On one hand, copy number of pGA1 is negatively regulated by an antisense RNA at the translational level (116). Additionally, the IR1 sequence, located in the leader region of the *rep* mRNA, influences negatively the *rep* gene expression. On the other hand, the *per* gene, which encodes a positive effector of replication, affects positively the copy number and segregation stability of pGA1 though it was found to be dispensable for replication (47). In fact, deletion of the *per* gene results in a significant decrease of the pGA1 copy number in *C. glutamicum*. Furthermore, the related plasmid pSR1 encodes a similar gene product that can act *in trans* on pGA1 derivatives (47). Per protein has no effect on the expression from the *rep* promoter and it has been hypothesized that counteracts the inhibitory effect of the antisense RNA on the *rep* gene expression by interacting with it (116). Interestingly, pGA1 still codes for another accessory effector, the small *aes* gene, which was shown to increase the segregational stability of pGA1 derivatives in the presence of *per* (117). Genes *per* and *aes* are located nearby and transcribed convergently. The authors propose that a cooperative regulation of expression of *per* and *aes* genes could be in the basis of the control of the pGA1 plasmid copy number and of its stable maintenance in *C. glutamicum*.

In plasmid pUB110, expression of the replication initiator gene (*repU*) is controlled by two small and unstable antisense RNAs, transcribed from a major incompatibility region, that interfere with

rep-mRNA translation by targeting the *repU* translation initiation signals (118). Interestingly, an additional control of plasmid copy number seems to exist in pUB110. In this plasmid, inactivation of a RepU molecule after a replication event has been proposed to occur by incorporation of a short oligonucleotide into it, in a way analogous to that described for pT181-RepC (see below). The inactive RepU-RepU* hetero-oligomer complex could form a large nucleoprotein structure at the *dso* region that interferes with transcription from the *repU* promoter (97). Therefore, even in the absence of a Cop-like transcriptional repressor, the amount of active RepU protein available for replication initiation could be subjected to a dual control at the translational and transcriptional level.

Mechanisms that restrict the use of Rep molecules to a single replication event

The replication and copy number of RCR plasmids are tightly regulated, and the mechanisms ensuring that the initiator proteins are unable to catalyze multiple rounds of DNA synthesis are critical to achieve such an efficient control. In pT181, the RepC dimer is inactivated after the completion of a round of replication by the attachment of an approximately 10-nt oligonucleotide, representing sequences immediately 3' to the initiation nick site, to the active tyrosine residue of one of its subunits, thus generating a heterodimer termed RepC/C*. Generation of RepC/C* occurs because, once the replication fork reaches the reconstituted origin, the nascent leading strand is extended by ~10 additional nts beyond the Rep nick site. Two site-specific transesterification reactions take place then: the first leads to the release of the circular ssDNA intermediate, and the second yields the RepC/C* heterodimer containing the short 3' leading-strand extension as well as a dsDNA plasmid molecule having the sealed new leading strand paired to the parental template strand (119, 120). Thus, the second transesterification allows completion of the termination process by avoiding recycling, i.e., continuous synthesis of the leading strand driven by the same initiator molecule (121). Analysis of the different interaction pattern of RepC/C and RepC/C* with the pT181-*dso* DNA by *in vitro* footprinting and binding-bending assays revealed that, although RepC/C* retains the ability to bind to DNA, it is unable to promote cruciform extrusion to expose the nick sequence in ssDNA form (122). This may explain why the formation of a RepC/RepC* heterodimer inactivates the protein besides uncoupling termination of leading-strand replication and initiation of a new replicative round.

A different mechanism preventing recycling upon termination of one replication round has been proposed for plasmid pC194. The pC194-encoded RepA protein has two active-site catalytic residues, namely Tyr214 and Glu210, which catalyze the DNA cleavage at the initiation and

termination steps, respectively. In this plasmid, termination reaction catalyzed by the glutamate residue is a hydrolysis that does not generate a covalent complex, thus preventing continuous synthesis of the leading strand (53). By contrast, in gene A protein of phage ΦX174 two catalytic Tyr of a single protomer would perform alternative nicking and nicking-closing reactions following a so called "flip-flop" mechanism that would allow recycling of the initiator protein (123). These two alternative pathways at the termination step of leading-strand synthesis reflect the different life-styles of phages and plasmids, with replication control featuring the latter. Noteworthily, the lack of formation of a Rep-DNA adduct in the cleavage reaction of the termination stage of the pC194 leading-strand replication implies that the yet-unknown mechanism of inactivation of the used RepA molecules would differ from that reported for the Rep proteins of the pT181/pC221-family plasmids. Remarkably, although the catalytic residues Glu and Tyr of RepA are conserved in the RepU initiator of pUB110, a replication-dependent modification of RepU implying the loss of its catalytic activity has been reported, and the modification suggested to consist, as in the case of RepC*, in the covalent attachment of a ss oligonucleotide (97). In pMV158, an inactivation mechanism similar to that described for pT181-RepC does not seem to be applicable, since RepB has not been shown to form a stable covalent complex with the 5'-P end of the cleaved DNA (83). However, generation of the circular ssDNA intermediate during termination of the plasmid leading strand synthesis requires that the 5'-P end to be ligated is covalently attached to RepB (83). Since this is the 5'-P end generated in the initiation cleavage, feasibility of preservation of the RepB-DNA bond through the entire leading-strand replication process under certain circumstances is suggested. Therefore, a mechanism analogous to the "flip-flop" scheme, which involves alternative nicking and nicking-closing of two catalytic residues of a single protomer, is plausible for RepB. As an alternative, the second catalytic residue could be provided by the active site of an adjacent monomer of the RepB hexamer, which would imply that substrates are transferred between OBDs during the termination reactions. In either case, and despite the biochemical and structural data available about RepB, the identification of the second catalytic residue remains undisclosed.

The unusual NTT activity displayed by the pGT5-encoded Rep75 initiator appears to be directly related with the mechanism that prevents plasmid over-replication. The three reactions catalyzed by Rep75, namely nicking, closing and NTT, can be uncoupled *in vitro*, although they share part of their mechanisms. Interestingly, replication initiation activities mediated by Rep75 are inhibited at the concentrations of ATP or dATP that promote the NTT activity of the protein. According to the

proposed model, an adenine residue could be transferred to the 3'-OH generated after specific cleavage by the NTT of another Rep molecule, thus recreating a putative nick site. After cleavage at this site, Rep75 will become attached to an adenine residue through its active Tyr and, therefore, inactivated. Hence, the role of the NTT activity might be to reduce the intracellular level of active Rep molecules. The authors propose that the levels of pGT5 replication could be regulated by an equilibrium between active and inactive Rep proteins, itself determined by the intracellular ATP and dATP pools and the availability of free 3'-OH and host replication proteins (86).

THE SINGLE-STRAND ORIGIN

Synthesis of the lagging strand of RCR plasmids initiates from the so-called single-strand origin (*sso*), a non-coding region containing long and imperfect inverted repeats that form complex secondary structures as the parental leading strand is displaced and becomes a ssDNA replicative intermediate (124, 125). In general, the *sso*s are located shortly upstream of the *dso*s and hence, these elements are among the latest plasmid regions to become ss during replication of the leading strand. This relative position might play a role in preventing run-off synthesis of the lagging strand before closing of the displaced parental strand upon termination of the leading-strand synthesis. The *sso* is recognized by host factors (most usually RNAP) that make a small RNA for priming the lagging-strand synthesis. Functionality of the *sso* is orientation-dependent, which points to a crucial role of unpaired sequences within the secondary structure of these elements.

Five main types of *sso* (*ssoU*, *ssoA*, *ssoT*, *ssoW* and *ssoL*) have been reported that differ from each other in structure, in sequence motifs highly conserved among the members of the same group, and in the host range in which they are functional.

The *ssoA*-type origins consist in a single and long (~150 nts) hairpin structure containing internal and bulge loops, in addition to the terminal loop. These origins display a high degree of sequence heterogeneity, with only two well-conserved regions: the previously termed Recombination Site B (RS_B), which is present in the lower stem of the hairpin of all the *ssoA*s, and a 6-nt consensus sequence (CS-6) that is located in the terminal loop of the hairpin of most but not all of these origins (124-126). The RS_B has been shown to be involved in binding of RNAP to the *ssoA* (127). Mutations in the RS_B abolish almost completely the *ssoA*-dependent synthesis of the lagging strand both *in vitro* and *in vivo*, giving rise to the accumulation of very high amounts of plasmid ssDNA intermediates in the host cell (126, 128). The CS-6-containing terminal loop seems to function as a terminator of the

primer RNA synthesis, and changes in this region lead to a moderate increase in the intracellular amount of ssDNA plasmid forms without impairing *ssoA* binding by RNAP (126-128). Sequences similar to CS-6 have been identified in the terminal loops of hairpins located in DNA regions that have been either proposed or shown to act as *sso*s in plasmids from Actinobacteria (129-133). In addition to the RS_B and CS-6 conserved motifs, sequences resembling the consensus -35 and -10 promoter regions have been identified in various *ssoA*s (134). A given *ssoA* only functions efficiently in its natural host or in a few closely-related species (12, 125, 134).

The *ssoW* was identified in lactococcal plasmid pWVO1 (135, 136). This origin is located in a 250-bp DNA fragment that contains two inverted repeats, IR I and IR II. IR I shows homology to the *ssoA*-type origins since it harbors the CS-6 sequence in the terminal loop as well as a sequence similar to the RS_B at the lower stem. Also, the upper stem of IR I shares remarkable sequence similarity with the ΦX174 minus-strand origin, which is recognized by the primosome for priming synthesis of the complementary strand (137). Full *ssoW* activity requires both IR I and IR II, and conversion of ssDNA into dsDNA from the entire element is only partially inhibited by rifampicin. IR II has not activity on its own, whereas IR I has a partial, RNAP-independent activity for complementary strand synthesis. Thus, priming of the lagging strand from *ssoW* seems to occur through two different pathways: one is catalyzed by RNAP and requires the entire origin, and the other, which was suggested to involve a primosomal complex, only requires IR I. Efficient functionality of the pWVO1-*ssoW* seems to be confined to lactococci (136).

The *ssoT*-type origins have been found in plasmids isolated from *Bacillus* (138-140). In fact, the majority of the RCR plasmids from *B. subtilis* harbor an *ssoT* (141). The minimal *ssoT*, as defined for the *B. subtilis* plasmid pBAA1, spans 120-190 bp encompassing three imperfect palindromes that would give rise to hairpins I, II and III on ssDNA. Results from mutation analysis suggest that the structure of hairpin III and both the structure and sequence of hairpin I are required for full activity of the *ssoT* (140). Comparison of the *ssoT* of pBAA1 with the homologous region of the *B. thuringiensis* pGI2 plasmid showed the existence of three conserved sequence motifs located in the loop of hairpin I (motifs 1 and 2) and in the intervening DNA between hairpins II and III (motif 3). Motif 1 has been shown to play an important role in the activity of the *ssoT*, while the role of motif 3 is controversial and that of motif 2 has not been proved. Initiation of lagging-strand synthesis from the *ssoT* is RNAP dependent, as can be inferred from the intracellular accumulation of ssDNA plasmid forms upon addition of rifampicin (140). It is also worth noting that the pBAA1-*ssoT* functions in both *B. subtilis* and

S. aureus, hence showing a broader host range than the *ssoA*- and *ssoW*-type origins (140). Based on sequence similarity, the *ssoT*s of a number of *B. subtilis* plasmids have been classified into two different groups, *palT1* and *palT2*, each including almost identical *ssoT* origins. The *palT1* group includes the *ssoT*s of plasmids pTA1015, pTA1020, pTA1060, pLS11 and pBAA1, whereas the *palT2* group includes those of pTA1030, pTA1040 and pTA1050 (141). The DNA region involved in initiation of the lagging-strand synthesis in *B. thuringiensis* plasmid pTX14-3 was found to be homologous to the pBAA1 *ssoT* (142). Actually, sequences highly similar to motifs 1, 2 and 3 are present in the pTX14-3 origin. Curiously enough, the activity of the *sso* of pTX14-3 is at least partially resistant to rifampicin, which suggests the existence of a lagging-strand priming mechanism independent of host RNAP (143).

The *ssoU* is the most promiscuous *sso* origin characterized so far, as it seems to be fully functional in many, if not all, Firmicutes (126, 144-147). The *ssoU* origin was first identified in staphylococcal plasmid pUB110 (147), within a ~250 bp DNA fragment with potential to form several hairpin structures containing symmetric and asymmetric internal loops in addition to the terminal one (144). Sequences nearly identical to the pUB110-*ssoU* were subsequently found in the streptococcal pMV158 and *Bacillus* pTB913 plasmids, and the involvement of these elements in lagging-strand synthesis was also proved (148). The high level of sequence identity among the *ssoU*s of plasmids isolated from different bacterial genus is consistent with this kind of *sso* being efficiently recognized in a broad range of hosts. The promiscuous activity of the *ssoU* has been suggested to be accounted for by the proven ability of this origin to bind efficiently to RNAP from different bacteria (149). No sequences with significant similarity to the canonical -35 and -10 promoter regions have been found in the *ssoU*. *In vitro*, RNAP binds to the left one of the two large hairpins of the *ssoU* origin, whereas the transcription initiation of the primer RNA and the transition from RNA to DNA synthesis occur, respectively, at the 3' end of the right large hairpin and at the 3' arm of its stem (149).

In spite of the sequence heterogeneity between the four types of *sso*s described above, they all exhibit an RS_B-like motif partially unpaired in the bottom part of the stem of a hairpin structure. Remarkably, the RS_B-like motif of the *ssoU* origin overlaps the region of the hairpin contacted by RNAP (149).

Apart from the *sso* elements mentioned above, a number of RCR *Lactobacillus* plasmids carry lagging-strand origins showing high levels of similarity among them but not with the *sso*s found in

plasmids from other bacteria genera. This kind of origins (*ssoL*) span ~100 bp, being considerably smaller than those of the other groups (150).

Unlike the *dso*, the *sso* is not essential for plasmid replication provided that an alternative pathway exists for priming lagging-strand synthesis. In general, alternative plasmid-borne signals and/or bacterial mechanisms seem to exist that can partially overcome the lack of the genuine *sso* (12, 126, 144). Especially efficient appears to be the predominant *sso*- and RNAP- independent priming system revealed in *Streptomyces lividans* and proposed to result from the stabilization of RCR plasmids lacking *sso* in this bacterium (151). In a few cases, however, the indispensability of the genuine *sso* origin has been reported, as is the case with the *Nocardia* plasmid pYS1 (132).

Although, as stated above, the *sso*s are usually not strictly required for replication, removal of a DNA region encompassing them not only leads to an increase of the ssDNA replicative intermediates, but frequently results in a reduction of the copy number (measured as dsDNA forms) and in segregational instability of the plasmids (12, 152, 153). The stability function linked to the *sso*s does not seem, however, to rely on the efficient conversion of ssDNA to dsDNA, since stable inheritance is observed for plasmids whose *sso* lacks activity in a given host, in spite of the inefficient conversion of ssDNA replicative intermediates into dsDNA plasmid forms. Consistently, removal of the genuine *sso* region usually leads to unstable plasmid inheritance even in hosts where the lagging-strand origin is not functional at all (153). Be it as it may, and although efficient conversion of ssDNA to dsDNA does not guaranty plasmid stability (154), the presence of an *sso* element appears to contribute to the efficient replication and accurate inheritance of the plasmid molecules and, therefore, to plasmid fitness. Remarkably, all natural RCR plasmids contain at least one *sso* that functions efficiently in their natural host, and plasmids isolated from different bacteria and belonging to the same replicon family (i.e., sharing homology at their *dso*s and *rep* genes) show no conservation at their *sso*s. These observations suggest that acquisition of an active *sso* upon plasmid entrance in a new host may improve not only plasmid fitness itself, but also fitness of the plasmid-containing bacteria, hence enabling them to overgrow cells that contain plasmids without a functional *sso*. In this sense, it is noteworthy that intracellular accumulation of ssDNA, which can arise from the presence of a plasmid lacking a functional *sso*, has been reported to induce bacterial stress responses like the *E. coli* SOS system (155, 156). This could result, in turn, in decreased growth rate of the cells and hence impaired bacterial fitness.

Streptococcal plasmid pMV158 is singular, among many other things that have appeared throughout this chapter, in that it contains two lagging-strand origins, namely streptococcal-specific *ssoA* and promiscuous *ssoU*. It has been shown that whereas the pMV158-*ssoU* participates in the plasmid mobilization between different bacterial species, the pMV158-*ssoA* would be involved mainly in intraspecific transfer (145). The presence of both *sso*s might reflect the evolutionary story and life-style of this mobilizable, highly-promiscuous plasmid.

HOST PROTEINS INVOLVED IN ROLLING-CIRCLE REPLICATION

Participation of DNA and RNA polymerases

In addition to the critical role frequently played by bacterial RNAP in recognition of the *sso* and priming of the lagging-strand synthesis (see above), several experimental evidences point to the direct participation of DNA polymerases (Pol) I and III in plasmid RCR.

Analyses of the involvement of Pol I in the leading- and lagging-strand synthesis during RCR were performed on the pMV158 model system, by employing plasmid derivatives that carried the *ssoA* and lacked the *ssoU*. These studies took advantage of the previous characterization of Pol I of *S. pneumoniae* and the construction of pneumococcal *polA* mutants lacking the polymerase activity of this enzyme (157, 158). Mutant strains depleted in the 5'-to-3' exonuclease function of Pol I could not be obtained since this activity was found to be essential for cell viability of *S. pneumoniae* (157). Therefore, only the involvement of the polymerase function of Pol I in plasmid RCR could be tested.

Participation of the polymerase activity of Pol I in the initiation of the lagging-strand synthesis was inferred from the increased fraction of ssDNA plasmid forms accumulated within the cells of the *polA* mutant strains compared to the wild-type strain (159). The same conclusion was drawn from the analysis of the *in vitro* replication of *ssoA*-containing ssDNA in cell-free extracts prepared from wild-type or *polA*-deficient pneumococcal strains. Wild-type levels of *in vitro* replication of plasmid ssDNA in extracts from the *polA*-depleted strain were obtained only upon complementation with the entire pneumococcal Pol I but not when a protein variant lacking the 5'-to-3' exonuclease activity was added, indicating that both the polymerizing and the exonuclease domains are required for efficient lagging-strand synthesis (127).

A role of Pol I during the termination of the lagging-strand synthesis was suggested from the comparative analysis of single-strand discontinuities detected in the DNA of the *ssoA*-containing pMV158 derivative extracted from the wild-type and *polA* mutant strains of *S. pneumoniae* (159). In

the polymerase-deficient mutants, a discontinuity in the vicinity of the *ssoA* origin was observed that could arise from plasmid molecules in which the conversion of ssDNA to dsDNA was not completed as a consequence of the defective replacement of the primer of the lagging strand.

An additional DNA discontinuity that was only detected in the pMV158 derivative replicating in the pneumococcal *polA* mutants mapped by the nick site of the plasmid *dso*, thus pointing to the participation of the polymerase activity of Pol I in an early step of the synthesis of the leading strand (159).

Finally, a critical role of the host Pol III replicase (the one containing the PolC polymerase) in plasmid RCR was early inferred from the inhibition of the *in vivo* replication of the pT181 DNA in *S. aureus* by hydroxvphenylhvdrazinouracil, an antimicrobial agent that acts specifically on the Gram-positive bacterial PolC polymerase (160).

ROLE OF SF1 HELICASES AND THEIR INTERACTION WITH THE PLASMID-ENCODED REP INITIATOR

As it has been previously described, plasmid RCR starts with the recognition of a *dso* by the replication initiator protein, which binds, cleaves and remains covalently bound to the 5'-end of the nicked DNA strand. After these initial steps, several cellular-encoded proteins are recruited, in order to continue the replication of the plasmid. The versatility of the interaction between these proteins and the replication initiator protein determines whether this plasmid is successfully replicated and maintained in a broad range of hosts or whether is lost. One of the required proteins for the replication of the leading strand is a helicase, capable of unwinding the dsDNA ahead of the replication fork. The helicases responsible for RCR are PcrA for Gram-positive bacteria and UvrD in Gram-negative bacteria, both of which belong to the superfamily group 1 (SF1) (161, 162). These proteins share 42% sequence identity and are structurally similar (163-166), both possessing the seven conserved helicase motifs. In both cases, there is some controversy regarding the active oligomeric state of the protein: some authors have suggested that PcrA and UvrD are monomeric helicases (167, 168), while others have shown they are active only as a dimer (169, 170).

The involvement of PcrA in RCR was first reported by Iordanescu et al. (171), who found that plasmid pT181 was unable to replicate in a strain of *S. aureus* carrying a mutation (*pcrA3*) within the *pcrA* gene. Subsequently, a mutation in the gene encoding the replication initiator protein RepC (Asp57Tyr) was identified, which suppressed the *pcrA3* mutation and restored pT181 replication in the

pcrA3 mutant (172). From these data it was proposed that the *pcrA3* mutation may impede replication by disrupting the PcrA3-RepC wt interaction, thereby preventing separation of the dsDNA, while the interaction is restored by the RepC-Asp57Tyr mutant, hence rescuing plasmid replication. However, it has since been shown by pull-down experiments that the PcrA3 mutant protein was able to interact with both RepC wt and the Asp57Tyr mutant (173). Rather, the authors showed that the *pcrA3* mutation leads to a threonine to isoleucine change in residue 61, which is located in the conserved motif Ia of the SF1 helicases, and results in a weak ATPase activity, which prevents the unwinding of the pT181 DNA in the presence of both RepC wt or Asp57Tyr mutant *in vitro* (173). The authors postulated that the viability of the *pcrA3* mutant and the capability of the double mutant *pcrA3-repC*Asp57Tyr to replicate the pT181 plasmid *in vivo* may be explained by additional cellular factors required for the replication to take place in this strain or by the role of an alternative cellular helicase which may replace PcrA. However, these hypotheses have yet to be confirmed.

PcrA from *S. aureus* is able to hydrolyze ATP, dATP, dGTP, dCTP and TTP and its NTPase activity is increased in the presence of either ssDNA or RepC covalently bound to the *oriC* of pT181 (174). In addition, the unwinding of supercoiled plasmid pT181 is only achieved when RepC is covalently attached to the origin of replication and in the presence of ATP in the reaction (174). PcrA is only capable of unwinding a dsDNA when there is a 3' or a 5' ss-tail exposed. Hence PcrA from *S. aureus* has a bipolar 3'-5' and 5'-3' helicase activity (175, 176).

Two hypotheses have been postulated for the direction of unwinding achieved by PcrA. As mentioned above, PcrA from *S. aureus* may have a dual helicase activity, whereas PcrA from *Bacillus stearothermophilus* has been shown to only unwind dsDNA in a 3'-5' helicase direction (163). This disparity may be explained by several factors, such as, the use of different constructs/protocols for protein purification, the absence *in vitro* of additional cofactors that would favor the unwinding in a specific direction, or the differences in sequence between PcrA from the two bacteria (59% sequence identity), among others.

PcrA from *B. stearothermophilus* is able to bind to a nicked dsDNA containing the *oriD* sequence from pC221 plasmid, then engaging the 3'-OH end and translocating in a 3'-5' direction along the nicked strand, thus unwinding the DNA. However, in the presence of RepD, the helicase is loaded on the opposing strand (the continuous strand), translocating in the same direction (177).The directionality of PcrA translocation was also confirmed to be 3'-5' using atomic force microscopy (AFM) of a linearized plasmid containing the *oriD* sequence at different positions relative to the DNA

ends. The *oriD* was nicked by RepD and the unwinding of the DNA was followed by the appearance of condensed ssDNA (177). Later, the kinetic parameters for PcrA helicase activity were determined in bulk and single molecule experiments, using lineal and supercoiled DNAs containing the *oriD* sequence, in the presence of RepD and saturating concentrations of ATP (178). Under these experimental conditions, the unwinding speed of PcrA was 30 bps^{-1}, whilst the translocation rate on ssDNA was 99 bases s^{-1}. The unwinding rate is dependent on the amount of ATP in the reaction and on the presence of RepD: in the absence of the replication initiator protein the number of unwinding events was reduced by more than ten-fold (167). Moreover, the affinity of PcrA for partial duplex DNA increased by one order of magnitude when RepD was already bound to the DNA, from a K_d of 22 nM to 170 nM (178)

The recruitment of the helicase onto the replication initiation site has also been shown by footprinting experiments using Exonuclease (Exo) III (179). RepD binding to *oriD* creates an area of protection to ExoIII digestion which extends beyond the *oriD* region: ~74-80 bp upstream of ICR I for the continuous strand and ~46-50 bp downstream of ICR III for the nicked strand strand (179), although further resistance points can be found within the *oriD*, after longer digestion times. When PcrA was incorporated into the reaction, the helicase was recruited upstream of ICR I, which served to stabilize the complex. The RepD-PcrA complex covered the region spanning from 80 bp upstream of ICR I on the continuous strand to the limit of ICR III in the nicked strand, and this complex was not displaced by ExoIII. However, a drastic change in the ternary complex was observed when a non-hydrolysable nucleotide (ADPNP) was included in the reaction. Under these circumstances, the resistance to ExoIII digestion was located only on ICR II, indicating that there were important conformational changes at the *oriD* once the helicase had begun to unwind the plasmid. Furthermore, these three stages of protein loading on the DNA have been studied by AFM, where RepD appears as a globular particle which bends the DNA fragment around 90º in 39% of the *oriD* fragments analyzed. However, when both proteins were bound to the DNA, the proportion of bent DNA increased up to 60%, while ADPNP decreased this percentage to 41% (179).

PcrA helicase has also been identified and characterized in other Gram-positive microorganisms such as *S. pneumoniae*, *Bacillus anthracis* or *Bacillus cereus*. In all these strains, the role of PcrA in RCR has been studied using the broad host range pT181 plasmid as a reference (180, 181). A clear interaction of the helicase with the replication initiation protein was successfully observed by pull-down for all three PcrAs. However, they displayed different unwinding activity on the plasmid

DNA in the presence of RepC: while PcrA from *B. anthracis* and *B. cereus* fully unwound the DNA, PcrA from *S. pneumoniae* failed to do so and produced only partial unwinding. This may indicate that the interaction between RepC from *S. aureus* and PcrA from *S. pneumoniae* is not sufficiently stable and, hence, the helicase is unable to continue DNA unwinding. In fact, the authors reported that it was not possible to maintain the pT181 plasmid in the *S. pneumoniae* strain (180).

In Gram-negative bacteria, a role of UvrD in plasmid RCR has been postulated, albeit its role has been much less studied than that of PcrA in Gram-positive bacteria. It has been shown that deletion of *uvrD* in *E. coli* results in the accumulation of nicked pC194 plasmid DNA in the cell and a lack of ssDNA intermediates of plasmid replication (182). Nevertheless, a direct interaction between the replication initiator protein and UvrD or its effect on the helicase activity has yet to be studied.

In summary, the effect of the replication initiator protein on PcrA ATPase/helicase activity has been thoroughly studied at the biochemical and molecular biology level. However, there are still some questions regarding how the replication machinery is loaded onto the replication initiation origin which need answering. For example, it has been reported that PcrA interacts with RNAP through its disordered but highly conserved C-terminal region (183), but it is not known which domains of PcrA and the replication initiator protein are implicated in the interaction between both proteins. In addition, the stability of the helicase-Rep complex is not known, nor whether there are cycles of loading/unloading of the helicase during the entire round of plasmid replication. Furthermore, unlike pT181, plasmids belonging to the pC194 or pMV158 families are able to replicate in *pcrA3* bacterial mutants [(184); Ruiz-Masó et al., unpublished results]. This raises the question of whether the helicase activity of PcrA is responsible for the replication of these plasmids or another helicase can cope with this task. Finally, despite a great deal of work on the helicase activity of UvrD alone (165, 168, 169, 185, 186), as it has been previously indicated, little is known about its role in RCR and on the putative interaction with the replication initiator protein.

Concluding remarks

Here we have tried to compile several aspects of the biology of the small RCR plasmids that constitute, *per se*, an extremely interesting family of replicons. Indeed they link the primitive forms of self-replicating molecules harboring just the information needed for their replication in the host, to more sophisticated beings that are interconnected to the viral world. Although much is known, as we have reflected here, on the replicative and their control mechanisms, still a full horizon expands before

we can truly assert that we do know these molecules. We envisage several approaches to the understanding of the biology of these molecules: i) Mechanistic studies; ii) Solution of 3D structures of Rep proteins; iii) Characterization of regulatory nucleoprotein complexes; iv) Solution of ternary complexes like DNA-Rep-PcrA or DNA-Rep-DNA polymerase; v) Involvement of host-encoded factors, other than the obvious roles of DNA- and RNA- polymerase, DNA helicases and ssDNA binding proteins (with an assumed but not demonstrated participation in plasmid replication); and vi) Comprehension of the mechanisms of adaptation of plasmids to a new host. We could speculate that we have just started uncovering the biology of RCR-plasmids. Their role in the bacterial adaptation to a changing world manipulated by humans, their contribution to the fitness of bacteria to their niches, and to the bio-diversity in the microbial world is totally unknown. Let's do not allow the funding agencies to say that there is no new relevant information in the plasmid world: it is fully untrue.

ACKNOWLEDGEMENTS

We are very grateful to Luis Blanco, Juan Alonso and to the members of the REDEEX consortium for fruitful discussions and help. Funding by the Spanish Ministry of Economy and Competitiveness (grants CSD2008/00013-INTERMODS to M.E.; BFU2011-22588 to M.C.; Ramón and Cajal subprogramme RYC-2011-09071 to C.M.; BFU2010-19597 to G.S. and Complementary Action BFU2008-00179-E-REDEEX to G.S.) are acknowledged.

Conflicts of interest: We disclose no conflicts

REFERENCES

1. **del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, Díaz-Orejas R.** 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. **62:**434-464.
2. **Khan SA.** 2005. Plasmid rolling-circle replication: highlights of two decades of research. Plasmid **53:**126-136.
3. **Chandler M, de la Cruz F, Dyda F, Hickman AB, Moncalian G, Ton-Hoang B.** 2013. Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. Nat Rev Microbiol **11:**525-538.

4. **Dressler D.** 1970. The rolling circle for phiX174 DNA replication. II. Synthesis of single-stranded circles. Proc Natl Acad Sci U S A **67:**1934-1942.

5. **Dressler D, Wolfson J.** 1970. The rolling circle for phiX174 DNA replication. 3. Synthesis of supercoiled duplex rings. Proc Natl Acad Sci U S A **67:**456-463.

6. **Gilbert W, Dressler D.** 1968. DNA replication: the rolling circle model. Cold Spring Harb Symp Quant Biol **33:**473-484.

7. **Eisenberg S, Kornberg A.** 1979. Purification and characterization of phiX174 gene A protein. A multifunctional enzyme of duplex DNA replication. J Biol Chem **254:**5328-5332.

8. **Henry TJ, Knippers R.** 1974. Isolation and function of the gene A initiator of bacteriophage phiX174, a highly specific DNA endonuclease. Proc Natl Acad Sci U S A **71:**1549-1553.

9. **Ikeda JE, Yudelevich A, Hurwitz J.** 1976. Isolation and characterization of the protein coded by *gene A* of bacteriophage phiX174 DNA. Proc Natl Acad Sci U S A **73:**2669-2673.

10. **van Mansfeld AD, van Teeffelen HA, Baas PD, Veeneman GH, van Boom JH, Jansz HS.** 1984. The bond in the bacteriophage phiX174 gene A protein--DNA complex is a tyrosyl-5'-phosphate ester. FEBS Lett **173:**351-356.

11. **Koepsel RR, Murray RW, Rosenblum WD, Khan SA.** 1985. The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities. Proc Natl Acad Sci U S A **82:**6845-6849.

12. **del Solar GH, Puyet A, Espinosa M.** 1987. Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Research **15:**5561-5580.

13. **Deng ZX, Kieser T, Hopwood DA.** 1988. "Strong incompatibility" between derivatives of the Streptomyces multi-copy plasmid pIJ101. Mol Gen Genet **214:**286-294.

14. **te Riele H, Michel B, Ehrlich SD.** 1986. Single-stranded plasmid DNA in *Bacillus subtilis* and *Staphylococcus aureus*. Proc Natl Acad Sci U S A **83:**2541-2545.

15. **Backert S, Meissner K, Borner T.** 1997. Unique features of the mitochondrial rolling circle-plasmid mp1 from the higher plant *Chenopodium album* (L.). Nucleic Acids Res **25:**582-589.

16. **Yu JS, Noll KM.** 1997. Plasmid pRQ7 from the hyperthermophilic bacterium *Thermotoga* species strain RQ7 replicates by the rolling-circle mechanism. J Bacteriol **179:**7161-7164.

17. **Akimkina T, Ivanov P, Kostrov S, Sokolova T, Bonch-Osmolovskaya E, Firman K, Dutta CF, McClellan JA.** 1999. A highly conserved plasmid from the extreme thermophile

Thermotoga maritima MC24 is a member of a family of plasmids distributed worldwide. Plasmid **42:**236-240.

18. **Nesbo CL, Dlutek M, Doolittle WF.** 2006. Recombination in Thermotoga: implications for species concepts and biogeography. Genetics **172:**759-769.

19. **Nesvera J, Hochmannova J, Patek M.** 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. FEMS Microbiol Lett **169:**391-395.

20. **Smith MC, Thomas CD.** 2004. An accessory protein is required for relaxosome formation by small staphylococcal plasmids. J Bacteriol **186:**3363-3373.

21. **Blanco M, Kadlec K, Gutierrez Martin CB, de la Fuente AJ, Schwarz S, Navas J.** 2007. Nucleotide sequence and transfer properties of two novel types of *Actinobacillus pleuropneumoniae* plasmids carrying the tetracycline resistance gene tet(H). J Antimicrob Chemother **60:**864-867.

22. **Cooper TF, Heinemann JA.** 2000. Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. Proc Natl Acad Sci U S A **97:**12643-12648.

23. **Cooper TF, Paixao T, Heinemann JA.** 2010. Within-host competition selects for plasmid-encoded toxin-antitoxin systems. Proc Biol Sci **277:**3149-3155.

24. **Hernández-Arriaga AM, Chan WT, Espinosa M, Díaz-Orejas R.** 2014. Conditional activation of toxin-antitoxin Systems: postsegregational killing and beyond. Microbiology Spectrum **2**.

25. **del Solar G, Fernández-López C, Ruiz-Masó JA, Lorenzo-Díaz F, Espinosa M.** Rolling Circle Replicating Plasmids, Molecular Life Sciences. Springer Science+Business Media New York, 2014.

26. **Petrova P, Miteva V, Ruiz-Masó JA, del Solar G.** 2003. Structural and functional analysis of pt38, a 2.9 kb plasmid of *Streptococcus thermophilus* yogurt strain. Plasmid **50:**176-189.

27. **Petrova PM, Gouliamova DE.** 2006. Rapid screening of plasmid-encoded small hsp-genes in *Streptococcus thermophilus*. Curr Microbiol **53:**422-427.

28. **Andrup L, Jensen GB, Wilcks A, Smidt L, Hoflack L, Mahillon J.** 2003. The patchwork nature of rolling-circle plasmids: comparison of six plasmids from two distinct *Bacillus thuringiensis* serotypes. Plasmid **49:**205-232.

29. **del Solar G, Alonso JC, Espinosa M, Diaz-Orejas R.** 1996. Broad-host-range plasmid replication: an open question. Mol Microbiol **21:**661-666.

30. **Iordanescu S.** 1995. Plasmid pT181 replication is decreased at high levels of RepC per plasmid copy. Mol Microbiol **16:**477-484.

31. **Ruiz-Masó JA, Anand SP, Espinosa M, Khan SA, del Solar G.** 2006. Genetic and biochemical characterization of the *Streptococcus pneumoniae* PcrA helicase and its role in plasmid rolling circle replication. J. Bacteriol. **188:**7416-7425.

32. **del Solar G, Moscoso M, Espinosa M.** 1993. Rolling circle-replicating plasmids from gram-positive and -negative bacteria: a wall falls. Mol. Micobiol. **8:**789-796.

33. **Biswas I, Jha JK, Fromm N.** 2008. Shuttle expression plasmids for genetic studies in Streptococcus mutans. Microbiology **154:**2275-2282.

34. **Bron S, Meijer W, Holsappel S, Haima P.** 1991. Plasmid instability and molecular cloning in *Bacillus subtilis*. Res Microbiol **142:**875-883.

35. **del Solar G, Moscoso M, Espinosa M.** 1993. *In vivo* definition of the functional origin of replication (ori(+)) of the promiscuous plasmid pLS1. Mol. Gen. Genet. **237:**65-72.

36. **O'Sullivan TF, Fitzgerald GF.** 1999. Electrotransformation of industrial strains of *Streptococcus thermophilus*. J Appl Microbiol **86:**275-283.

37. **Turgeon N, Laflamme C, Ho J, Duchaine C.** 2006. Elaboration of an electroporation protocol for *Bacillus cereus* ATCC 14579. J Microbiol Methods **67:**543-548.

38. **Gruss A, Ehrlich SD.** 1988. Insertion of foreign DNA into plasmids from gram-positive bacteria induces formation of high-molecular-weight plasmid multimers. J Bacteriol **170:**1183-1190.

39. **Kiewiet R, Kok J, Seegers JF, Venema G, Bron S.** 1993. The mode of replication is a major factor in segregational plasmid instability in *Lactococcus lactis*. Appl Environ Microbiol **59:**358-364.

40. **Leonhardt H, Alonso JC.** 1991. Parameters affecting plasmid stability in *Bacillus subtilis*. Gene **103:**107-111.

41. **Viret JF, Alonso JC.** 1987. Generation of linear multigenome-length plasmid molecules in *Bacillus subtilis*. Nucleic Acids Res **15:**6349-6367.

42. **Alonso JC, Trautner TA.** 1985. Generation of deletions through a cis-acting mutation in plasmid pC194. Mol Gen Genet **198:**432-436.

43. **Fernández-López C, Bravo A, Ruiz-Cruz S, Solano-Collado V, Garsin DA, Lorenzo-Díaz F, Espinosa M.** 2014. Mobilizable rolling-circle replicating plasmids from gram-positive bacteria: a low-cost conjugative transfer. Microbiology Spectrum **2(5): PLAS-0008-2013**.

44. **Ruiz-Cruz S, Solano-Collado V, Espinosa M, Bravo A.** 2010. Novel plasmid-based genetic tools for the study of promoters and terminators in *Streptococcus pneumoniae* and *Enterococcus faecalis*. J. Microbiol. Methods **83:**156-163.

45. **Ruiz-Masó JA, Lopez-Aguilar C, Nieto C, Sanz M, Buron P, Espinosa M, del Solar G.** 2012. Construction of a plasmid vector based on the pMV158 replicon for cloning and inducible gene expression in *Streptococcus pneumoniae*. Plasmid **67:**53-59.

46. **Patek M, Nesvera J.** 2013. Promoters and Plasmid Vectors of *Corynebacterium glutamicum*. *In* Inui HYaM (ed.), *Corynebacterium glutamicum*. Biology and Biotechnology. Springer Berlin Heidelberg.

47. **Nesvera J, Patek M, Hochmannova J, Abrhamova Z, Becvarova V, Jelinkova M, Vohradsky J.** 1997. Plasmid pGA1 from *Corynebacterium glutamicum* codes for a gene product that positively influences plasmid copy number. J Bacteriol **179:**1525-1532.

48. **Hernandez-Arriaga AM, Espinosa M, del Solar G.** 2012. Fitness of the pMV158 replicon in *Streptococcus pneumoniae*. Plasmid **67:**162-166.

49. **Espinosa M, del Solar G, Rojo F, Alonso JC.** 1995. Plasmid rolling circle replication and its control. FEMS Microbiol Lett **130:**111-120.

50. **Khan SA.** 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. **61:**442-455.

51. **Lo Piano A, Martinez-Jimenez MI, Zecchi L, Ayora S.** 2011. Recombination-dependent concatemeric viral DNA replication. Virus Res **160:**1-14.

52. **Thomas CD, Balson DF, Shaw WV.** 1990. *In vitro* studies of the initiation of staphylococcal plasmid replication. Specificity of RepD for its origin (oriD) and characterization of the Rep-ori tyrosyl ester intermediate. J Biol Chem **265:**5519-5530.

53. **Noirot-Gros MF, Bidnenko V, Ehrlich SD.** 1994. Active site of the replication protein of the rolling circle plasmid pC194. EMBO J **13:**4412-4420.

54. **Novick RP.** 1989. Staphylococcal plasmids and their replication. Annu Rev Microbiol **43:**537-565.

55. **Tauch A, Puhler A, Kalinowski J, Thierbach G.** 2003. Plasmids in *Corynebacterium glutamicum* and their molecular classification by comparative genomics. J Biotechnol **104:**27-40.

56. **Zhou L, Zhou M, Sun C, Han J, Lu Q, Zhou J, Xiang H.** 2008. Precise determination, cross-recognition, and functional analysis of the double-strand origins of the rolling-circle replication plasmids in haloarchaea. J Bacteriol **190:**5710-5719.

57. **Erauso G, Marsin S, Benbouzid-Rollet N, Baucher MF, Barbeyron T, Zivanovic Y, Prieur D, Forterre P.** 1996. Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. J Bacteriol **178:**3232-3237.

58. **Zock JM, Birch P, Khan SA.** 1990. Specificity of RepC protein in plasmid pT181 DNA replication. J Biol Chem **265:**3484-3488.

59. **Iordanescu S.** 1989. Specificity of the interactions between the Rep proteins and the origins of replication of *Staphylococcus aureus* plasmids pT181 and pC221. Mol Gen Genet **217:**481-487.

60. **del Solar G, Moscoso M, Espinosa M.** 1993. *In vivo* definition of the functional origin of replication (*ori*(+)) of the promiscuous plasmid pLS1. Mol Gen Genet **237:**65-72.

61. **Kim SW, Jeong EJ, Kang HS, Tak JI, Bang WY, Heo JB, Jeong JY, Yoon GM, Kang HY, Bahk JD.** 2006. Role of RepB in the replication of plasmid pJB01 isolated from *Enterococcus faecium* JC1. Plasmid **55:**99-113.

62. **de la Campa AG, del Solar GH, Espinosa M.** 1990. Initiation of replication of plasmid pLS1. The initiator protein RepB acts on two distant DNA regions. J Mol Biol **213:**247-262.

63. **Moscoso M, del Solar G, Espinosa M.** 1995. Specific nicking-closing activity of the initiator of replication protein RepB of plasmid pMV158 on supercoiled or single-stranded DNA. J. Biol. Chem. **270:**3772-3779.

64. **Moscoso M, Eritja R, Espinosa M.** 1997. Initiation of replication of plasmid pMV158: mechanisms of DNA strand-transfer reactions mediated by the initiator RepB protein. Journal of Molecular Biology **268:**840.

65. **Ruiz-Maso JA, Lurz R, Espinosa M, del Solar G.** 2007. Interactions between the RepB initiator protein of plasmid pMV158 and two distant DNA regions within the origin of replication. Nucleic Acids Res **35:**1230-1244.

66. **Wang PZ, Projan SJ, Henriquez V, Novick RP.** 1993. Origin recognition specificity in pT181 plasmids is determined by a functionally asymmetric palindromic DNA element. EMBO J **12:**45-52.

67. **del Solar G, Diaz R, Espinosa M.** 1987. Replication of the streptococcal plasmid pMV158 and derivatives in cell-free extracts of *Escherichia coli*. Mol Gen Genet **206:**428-435.

68. **Moscoso M, del Solar G, Espinosa M.** 1995. *In vitro* recognition of the replication origin of pLS1 and of plasmids of the pLS1 family by the RepB initiator protein. J. Bacteriol. **177:**7041-7049.

69. **Noirot P, Bargonetti J, Novick RP.** 1990. Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. Proc Natl Acad Sci U S A **87:**8560-8564.

70. **Puyet A, del Solar G, Espinosa M.** 1988. Identification of the origin and direction of replication of the broad-host-range plasmid pLS1. Nucl. Acids Res. **16:**115-133.

71. **Michel B, Ehrlich SD.** 1986. Illegitimate recombination occurs between the replication origin of the plasmid pC194 and a progressing replication fork. EMBO J **5:**3691-3696.

72. **Gros MF, te Riele H, Ehrlich SD.** 1987. Rolling circle replication of single-stranded DNA plasmid pC194. EMBO J **6:**3863-3869.

73. **Alonso JC, Leonhardt H, Stiege CA.** 1988. Functional analysis of the leading strand replication origin of plasmid pUB110 in Bacillus subtilis. Nucleic Acids Res **16:**9127-9145.

74. **Servin-Gonzalez L.** 1993. Relationship between the replication functions of Streptomyces plasmids pJV1 and pIJ101. Plasmid **30:**131-140.

75. **Kataoka M, Kiyose YM, Michisuji Y, Horiguchi T, Seki T, Yoshida T.** 1994. Complete nucleotide sequence of the Streptomyces nigrifaciens plasmid, pSN22: genetic organization and correlation with genetic properties. Plasmid **32:**55-69.

76. **Suzuki I, Seki T, Yoshida T.** 1997. Nucleotide sequence of a nicking site of the Streptomyces plasmid pSN22 replicating by the rolling circle mechanism. FEMS Microbiol Lett **150:**283-288.

77. **Abrhamova Z, Patek M, Nesvera J.** 2002. Atypical location of double-strand origin of replication (nic site) on the plasmid pGA1 from *Corynebacterium glutamicum*. Folia Microbiol (Praha) **47:**307-310.

78. **Soler N, Justome A, Quevillon-Cheruel S, Lorieux F, Le Cam E, Marguet E, Forterre P.** 2007. The rolling-circle plasmid pTN1 from the hyperthermophilic archaeon *Thermococcus nautilus*. Mol Microbiol **66:**357-370.

79. **Ward DE, Revet IM, Nandakumar R, Tuttle JH, de Vos WM, van der Oost J, DiRuggiero J.** 2002. Characterization of plasmid pRT1 from *Pyrococcus* sp. strain JT1. J Bacteriol **184:**2561-2566.

80. **Ilyina TV, Koonin EV.** 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Res **20:**3279-3285.

81. **Boer R, Russi S, Guasch A, Lucas M, Blanco AG, Pérez-Luque R, Coll M, de la Cruz F.** 2006. Unveiling the molecular mechanism of a conjugative relaxase: the structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. Journal of Molecular Biology **358:**857.

82. **Hickman AB, Ronning DR, Kotin RM, Dyda F.** 2002. Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. Molecular Cell **10:**327.

83. **Moscoso M, del Solar G, Espinosa M.** 1995. Specific nicking-closing activity of the initiator of replication protein RepB of plasmid pMV158 on supercoiled or single-stranded DNA. J Biol Chem **270:**3772-3779.

84. **Asano S, Higashitani A, Horiuchi K.** 1999. Filamentous phage replication initiator protein gpII forms a covalent complex with the 5' end of the nick it introduced. Nucl. Acids Res. **27:**1882-1889.

85. **Marsin S, Forterre P.** 1998. A rolling circle replication initiator protein with a nucleotidyl-transferase activity encoded by the plasmid pGT5 from the hyperthermophilic archaeon *Pyrococcus abyssi*. Molecular Microbiology **27:**1183-1192.

86. **Marsin S, Forterre P.** 1999. The active site of the rolling circle replication protein Rep75 is involved in site-specific nuclease, ligase and nucleotidyl transferase activities. Mol Microbiol **33:**537-545.

87. **Dempsey LA, Birch P, Khan SA.** 1992. Six amino acids determine the sequence-specific DNA binding and replication specificity of the initiator proteins of the pT181 family. J Biol Chem **267:**24538-24543.

88. **Dempsey LA, Birch P, Khan SA.** 1992. Uncoupling of the DNA topoisomerase and replication activities of an initiator protein. Proc Natl Acad Sci U S A **89:**3083-3087.

89. **Thomas CD, Nikiforov TT, Connolly BA, Shaw WV.** 1995. Determination of sequence specificity between a plasmid replication initiator protein and the origin of replication. J. Mol. Biol. **254:**381-391.

90. **Wang PZ, Projan SJ, Henriquez V, Novick RP.** 1992. Specificity of origin recognition by replication initiator protein in plasmids of the pT181 family is determined by a six amino acid residue element. J Mol Biol **223:**145-158.

91. **Chang T-L, Kramer MG, Ansari RA, Khan SA.** 2000. Role of individual monomers of a dimeric initiator protein in the initiation and termination of plasmid rolling circle replication. J. Biol. Chem. **275:**13529-13534.

92. **Boer R, Ruiz-Masó JA, Gomez-Blanco JR, Blanco AG, Vives-Llàcer M, Chacón P, Usón I, Gomis-Ruth FX, Espinosa M, Llorca O, Del Solar G, Coll M.** 2009. Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. EMBO J **28:**1666-1678.

93. **Ruiz-Maso JA, Lopez-Zumel C, Menendez M, Espinosa M, del Solar G.** 2004. Structural features of the initiator of replication protein RepB encoded by the promiscuous plasmid pMV158. Biochim Biophys Acta **1696:**113-119.

94. **Novick RP.** 1989. Staphylococcal plasmids and their replication. Annu. Rev. Microbiol. **43:**537-565.

95. **Rasooly A, Projan SJ, Novick RP.** 1994. Plasmids of the pT181 family show replication-specific initiator protein modification. J Bacteriol **176:**2450-2453.

96. **Zhao AC, Ansari RA, Schmidt MC, Khan SA.** 1998. An oligonucleotide inhibits oligomerization of a rolling circle initiator protein at the pT181 origin of replication. J Biol Chem **273:**16082-16089.

97. **Muller AK, Rojo F, Alonso JC.** 1995. The level of the pUB110 replication initiator protein is autoregulated, which provides an additional control for plasmid copy number. Nucleic Acids Res **23:**1894-1900.

98. **Enemark EJ, Joshua-Tor L.** 2006. Mechanism of DNA translocation in a replicative hexameric helicase. Nature **442:**270.

99. **Krupovic M, Ravantti JJ, Bamford DH.** 2009. Geminiviruses: a tale of a plasmid becoming a virus. BMC Evol Biol **9:**112.

100. **Oshima K, Kakizawa S, Nishigawa H, Kuboyama T, Miyata S, Ugaki M, Namba S.** 2001. A plasmid of phytoplasma encodes a unique replication protein having both plasmid- and virus-like domains: clue to viral ancestry or result of virus/plasmid recombination? Virology **285:**270-277.

101. **Tran-Nguyen LTT, Gibb KS.** 2006. Extrachromosomal DNA isolated from tomato big bud and *Candidatus* Phytoplasma australiense phytoplasma strains. Plasmid **56:**153-166.

102. **Brantl S.** 2014. Plasmid replication control by antisense RNAs. Microbiology Spectrum **2(4):PLAS-0001-2013**.

103. **del Solar G, Espinosa M.** 2000. Plasmid copy number control: an ever-growing story. Mol Microbiol **37:**492-500.

104. **Novick RP, Iordanescu S, Projan SJ, Kornblum J, Edelman I.** 1989. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. Cell **59:**395-404.

105. **del Solar G, Acebo P, Espinosa M.** 1995. Replication control of plasmid pLS1: efficient regulation of plasmid copy number is exerted by the combined action of two plasmid components, CopG and RNA II. Mol Microbiol **18:**913-924.

106. **Hernandez-Arriaga AM, Rubio-Lepe TS, Espinosa M, del Solar G.** 2009. Repressor CopG prevents access of RNA polymerase to promoter and actively dissociates open complexes. Nucleic Acids Res **37:**4799-4811.

107. **del Solar G, Acebo P, Espinosa M.** 1997. Replication control of plasmid pLS1: the antisense RNA II and the compact rnaII region are involved in translational regulation of the initiator RepB synthesis. Mol. Microbiol. **23:**95-108.

108. **del Solar G, Espinosa M.** 2001. *In vitro* analysis of the terminator T(II) of the inhibitor antisense *rna II* gene from plasmid pMV158. Plasmid **45:**75-87.

109. **Lopez-Aguilar C, del Solar G.** 2013. Probing the sequence and structure of in vitro synthesized antisense and target RNAs from the replication control system of plasmid pMV158. Plasmid **70:**94-103.

110. **del Solar G, Acebo P, Espinosa M.** 1995. Replication control of plasmid pLS1: efficient regulation of plasmid copy number is exerted by the combined action of two plasmid components, CopG and RNA II. Mol. Microbiol. **18:**913-924.

111. **del Solar G, Espinosa M.** 1992. The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products: the antisense RNA II and the repressor protein, RepA. Mol. Microbiol. **6:**83-94.

112. **Brantl S, Wagner EG.** 2000. Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181. Mol Microbiol **35:**1469-1482.

113. **Franch T, Petersen M, Wagner EGH, Jacobsen JP, Gerdes K.** 1999. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. J. Mol. Biol. **294:**1115-1125.

114. **Kim SW, Jeong IS, Jeong EJ, Tak JI, Lee JH, Eo SK, Kang HY, Bahk JD.** 2008. The terminal and internal hairpin loops of the ctRNA of plasmid pJB01 play critical roles in regulating copy number. Mol. Cells **26:**26-33.

115. **Lopez-Aguilar C, Ruiz-Maso JA, Rubio-Lepe TS, Sanz M, del Solar G.** 2013. Translation initiation of the replication initiator *repB* gene of promiscuous plasmid pMV158 is led by an extended non-SD sequence. Plasmid **70:**69-77.

116. **Venkova-Canova T, Patek M, Nesvera J.** 2003. Control of *rep* gene expression in plasmid pGA1 from *Corynebacterium glutamicum*. J Bacteriol **185:**2402-2409.

117. **Venkova T, Patek M, Nesvera J.** 2001. Identification of a novel gene involved in stable maintenance of plasmid pGA1 from *Corynebacterium glutamicum*. Plasmid **46:**153-162.

118. **Maciag IE, Viret JF, Alonso JC.** 1988. Replication and incompatibility properties of plasmid pUB110 in *Bacillus subtilis*. Mol Gen Genet **212:**232-240.

119. **Rasooly A, Novick RP.** 1993. Replication-specific inactivation of the pT181 plasmid initiator protein. Science **262:**1048-1050.

120. **Rasooly A, Wang P, Novick R.** 1994. Replication-specific conversion of the *Staphylococcus aureus* pT181 initiator protein from an active homodimer to an inactive heterodimer. EMBO J. **13:**5245-5251.

121. **Novick RP.** 1998. Contrasting lifestyles of rolling-circle phages and plasmids. Trends Biochem. Sci. **23:**434-438.

122. **Jin R, Zhou X, Novick RP.** 1996. The inactive pT181 initiator heterodimer, RepC/C, binds but fails to induce melting of the plasmid replication origin. J Biol Chem **271:**31086-31091.

123. **Noirot-Gros MF, Ehrlich SD.** 1996. Change of a catalytic reaction carried out by a DNA replication protein. Science **274:**777-780.

124. **del Solar G, Puyet A, Espinosa M.** 1987. Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Res. **15:**5561-5580.

125. **Gruss AD, Ross HF, Novick RP.** 1987. Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids. Proc. Natl. Acad. Sci. USA **84:**2165-2169.

126. **Kramer MG, del Solar G, Espinosa M.** 1995. Lagging-strand origins of the promiscuous plasmid pMV158: physical and functional characterization. Microbiology **141 (Pt 3):**655-662.

127. **Kramer MG, Khan SA, Espinosa M.** 1997. Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement for DNA polymerase I for lagging strand synthesis. EMBO J **16:**5784-5795.

128. **Kramer MG, Khan SA, Espinosa M.** 1998. Lagging-strand replication from the *ssoA* origin of plasmid pMV158 in *Streptococcus pneumoniae*: *in vivo* and *in vitro* influences of mutations in two conserved *ssoA* regions. J. Bacteriol. **180:**83-89.

129. **Farrar MD, Howson KM, Emmott JE, Bojar RA, Holland KT.** 2007. Characterisation of cryptic plasmid pPG01 from *Propionibacterium granulosum*, the first plasmid to be isolated from a member of the cutaneous propionibacteria. Plasmid **58:**68-75.

130. **Fernandez-Gonzalez C, Cadenas RF, Noirot-Gros MF, Martin JF, Gil JA.** 1994. Characterization of a region of plasmid pBL1 of *Brevibacterium lactofermentum* involved in replication via the rolling circle model. J Bacteriol **176:**3154-3161.

131. **Nakashima N, Tamura T.** 2004. Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. Appl Environ Microbiol **70:**5557-5568.

132. **Shibayama Y, Dabbs ER, Yazawa K, Mikami Y.** 2011. Functional analysis of a small cryptic plasmid pYS1 from Nocardia. Plasmid **66:**26-37.

133. **Zaman S, Radnedge L, Richards H, Ward JM.** 1993. Analysis of the site for second-strand initiation during replication of the Streptomyces plasmid pIJ101. J Gen Microbiol **139:**669-676.

134. **Kramer MG, Espinosa M, Misra TK, Khan SA.** 1998. Lagging strand replication of rolling-circle plasmids: specific recognition of the *ssoA*-type origins in different gram-positive bacteria. Proc Natl Acad Sci U S A **95:**10505-10510.

135. **Leenhouts KJ, Tolner B, Bron S, Kok J, Venema G, Seegers JFML.** 1991. Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pWVO1. Plasmid **26:**55-66.

136. **Seegers JF, Zhao AC, Meijer WJ, Khan SA, Venema G, Bron S.** 1995. Structural and functional analysis of the single-strand origin of replication from the lactococcal plasmid pWV01. Mol Gen Genet **249:**43-50.

137. **Baas PD, Jansz HS.** 1988. Single-stranded DNA phage origins. Curr Top Microbiol Immunol **136:**31-70.

138. **Chang S, Chang SY, Gray O.** 1987. Structural and genetic analyses of a par locus that regulates plasmid partition in *Bacillus subtilis*. J Bacteriol **169:**3952-3962.

139. **Mahillon J, Seurinck J.** 1988. Complete nucleotide sequence of pGI2, a *Bacillus thuringiensis* plasmid containing Tn4430. Nucleic Acids Res **16:**11827-11828.

140. **Seery L, Devine KM.** 1993. Analysis of features contributing to activity of the single-stranded origin of *Bacillus plasmid* pBAA1. J Bacteriol **175:**1988-1994.

141. **Meijer WJJ, de Boer AJ, van Tongeren S, Venema G, Bron S.** 1995. Characterization of the replication region of the *Bacillus subtilis* plasmid pLS20: a novel type of replicon. Nucleic Acids Res. **23:**3214-3223.

142. **Madsen SM, Andrup L, Boe L.** 1993. Fine mapping and DNA sequence of replication functions of *Bacillus thuringiensis* plasmid pTX14-3. Plasmid **30:**119-130.

143. **Boe L, Nielsen TT, Madsen SM, Andrup L, Bolander G.** 1991. Cloning and characterization of two plasmids from *Bacillus thuringiensis* in *Bacillus subtilis*. Plasmid **25:**190-197.

144. **Boe L, Gros MF, te Riele H, Ehrlich SD, Gruss A.** 1989. Replication origins of single-stranded-DNA plasmid pUB110. J Bacteriol **171:**3366-3372.

145. **Lorenzo-Diaz F, Espinosa M.** 2009. Lagging-strand DNA replication origins are required for conjugal transfer of the promiscuous plasmid pMV158. J Bacteriol **191:**720-727.

146. **Meijer WJ, van der Lelie D, Venema G, Bron S.** 1995. Effects of the generation of single-stranded DNA on the maintenance of plasmid pMV158 and derivatives in *Lactococcus lactis*. Plasmid **33:**91-99.

147. **Viret JF, Alonso JC.** 1988. A DNA sequence outside the pUB110 minimal replicon is required for normal replication in *Bacillus subtilis*. Nucleic Acids Res **16:**4389-4406.

148. **van der Lelie D, Bron S, Venema G, Oskam L.** 1989. Similarity of minus origins of replication and flanking open reading frames of plasmids pUB110, pTB913 and pMV158. Nucleic Acids Res **17:**7283-7294.

149. **Kramer MG, Espinosa M, Misra TK, Khan SA.** 1999. Characterization of a single-strand origin, *ssoU*, required for broad host range replication of rolling-circle plasmids. Mol Microbiol **33:**466-475.

150. **Leer RJ, Luijk N, Posno M, Pouwels PH.** 1992. Structural and functional analysis of two cryptic plasmids from *Lactobacillus pentosus* MD353 and *Lactobacillus plantarum* ATCC 8014. Mol. Gen. Genet. **234:**265-274.

151. **Suzuki I, Kataoka M, Yoshida T, Seki T.** 2004. Lagging strand replication of rolling-circle plasmids in Streptomyces lividans: an RNA polymerase-independent primer synthesis. Arch Microbiol **181:**305-313.

152. **Bron S, Luxen E.** 1985. Segregational instability of pUB110-derived recombinant plasmids in *Bacillus subtilis*. Plasmid **14:**235-244.

153. **del Solar G, Kramer G, Ballester S, Espinosa M.** 1993. Replication of the promiscuous plasmid pLS1: a region encompassing the minus origin of replication is associated with stable plasmid inheritance. Mol Gen Genet **241:**97-105.

154. **Hernandez-Arriaga AM, Espinosa M, del Solar G.** 2000. A functional lagging strand origin does not stabilize plasmid pMV158 inheritance in *Escherichia coli*. Plasmid **43:**49-58.

155. **Gigliani F, Ciotta C, Del Grosso MF, Battaglia PA.** 1993. pR plasmid replication provides evidence that single-stranded DNA induces the SOS system *in vivo*. Mol Gen Genet **238:**333-338.

156. **Higashitani N, Higashitani A, Horiuchi K.** 1995. SOS induction in *Escherichia coli* by single-stranded DNA of mutant filamentous phage: monitoring by cleavage of LexA repressor. J Bacteriol **177:**3610-3612.

157. **Diaz A, Lacks SA, Lopez P.** 1992. The 5' to 3' exonuclease activity of DNA polymerase I is essential for *Streptococcus pneumoniae*. Mol Microbiol **6:**3009-3019.

158. **Lopez P, Martinez S, Diaz A, Espinosa M, Lacks SA.** 1989. Characterization of the *polA* gene of *Streptococcus pneumoniae* and comparison of the DNA polymerase I it encodes to homologous enzymes from *Escherichia coli* and phage T7. J Biol Chem **264:**4255-4263.

159. **Díaz A, Lacks SA, López P.** 1994. Multiple roles for DNA polymerase I in establishment and replication of the promiscuous plasmid pLS1. Mol. Micobiol. **14:**773-783.

160. **Majumder S, Novick RP.** 1988. Intermediates in plasmid pT181 DNA replication. Nucleic Acids Res **16:**2897-2912.

161. **Gorbalenya AE, Koonin EV.** 1993. Helicases: amino acid sequence comparisons and structure-function relationships. Curr Opin Struct Biol **3:**419-429.

162. **Singleton MR, Dillingham MS, Wigley DB.** 2007. Structure and mechanism of helicases and nucleic acid translocases. Annu Rev Biochem **76:**23-50.

163. **Bird LE, Brannigan JA, Subramanya HS, Wigley DB.** 1998. Characterisation of *Bacillus stearothermophilus* PcrA helicase: evidence against an active rolling mechanism. Nucleic Acids Res **26:**2686-2693.

164. **Jia H, Korolev S, Niedziela-Majka A, Maluf NK, Gauss GH, Myong S, Ha T, Waksman G, Lohman TM.** 2011. Rotations of the 2B sub-domain of *E. coli* UvrD helicase/translocase coupled to nucleotide and DNA binding. J Mol Biol **411:**633-648.

165. **Lee JY, Yang W.** 2006. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. Cell **127:**1349-1360.

166. **Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB.** 1999. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. Cell **97:**75-84.

167. **Chisty LT, Toseland CP, Fili N, Mashanov GI, Dillingham MS, Molloy JE, Webb MR.** 2013. Monomeric PcrA helicase processively unwinds plasmid lengths of DNA in the presence of the initiator protein RepD. Nucleic Acids Res **41:**5010-5023.

168. **Fischer CJ, Maluf NK, Lohman TM.** 2004. Mechanism of ATP-dependent translocation of *E.coli* UvrD monomers along single-stranded DNA. J Mol Biol **344:**1287-1309.

169. **Ali JA, Maluf NK, Lohman TM.** 1999. An oligomeric form of *E. coli* UvrD is required for optimal helicase activity. J Mol Biol **293:**815-834.

170. **Yang Y, Dou SX, Ren H, Wang PY, Zhang XD, Qian M, Pan BY, Xi XG.** 2008. Evidence for a functional dimeric form of the PcrA helicase in DNA unwinding. Nucleic Acids Res **36:**1976-1989.

171. **Iordanescu S, Basheer R.** 1991. The *Staphylococcus aureus* mutation *pcrA3* leads to the accumulation of pT181 replication initiation complexes. J Mol Biol **221:**1183-1189.

172. **Iordanescu S.** 1993. Plasmid pT181-linked suppressors of the *Staphylococcus aureus pcrA3* chromosomal mutation. J Bacteriol **175:**3916-3917.

173. **Anand SP, Chattopadhyay A, Khan SA.** 2005. The PcrA3 mutant binds DNA and interacts with the RepC initiator protein of plasmid pT181 but is defective in its DNA helicase and unwinding activities. Plasmid **54:**104-113.

174. **Chang TL, Naqvi A, Anand SP, Kramer MG, Munshi R, Khan SA.** 2002. Biochemical characterization of the *Staphylococcus aureus* PcrA helicase and its role in plasmid rolling circle replication. J Biol Chem **277:**45880-45886.

175. **Anand SP, Khan SA.** 2004. Structure-specific DNA binding and bipolar helicase activities of PcrA. Nucl. Acids Res. **32:**3190-3197.

176. **Chang T-L, Naqvi A, Anand SP, Kramer MG, Munshi R, Khan SA.** 2002. Biochemical characterization of the *Staphylococcus aureus* PcrA helicase and its role in plasmid rolling circle replication. J. Biol. Chem. **277:**45880-45886.

177. **Zhang W, Dillingham MS, Thomas CD, Allen S, Roberts CJ, Soultanas P.** 2007. Directional loading and stimulation of PcrA helicase by the replication initiator protein RepD. J Mol Biol **371:**336-348.

178. **Slatter AF, Thomas CD, Webb MR.** 2009. PcrA helicase tightly couples ATP hydrolysis to unwinding double-stranded DNA, modulated by the initiator protein for plasmid replication, RepD. Biochemistry **48:**6326-6334.

179. **Machón C, Lynch GP, Thomson NH, Scott DJ, Thomas CD, Soultanas P.** 2010. RepD-mediated recruitment of PcrA helicase at the Staphylococcus aureus pC221 plasmid replication origin, oriD. Nucleic Acids Res **38:**1874-1888.

180. **Ruiz-Maso JA, Anand SP, Espinosa M, Khan SA, del Solar G.** 2006. Genetic and biochemical characterization of the Streptococcus pneumoniae PcrA helicase and its role in plasmid rolling circle replication. J Bacteriol **188:**7416-7425.

181. **Anand SP, Mitra P, Naqvi A, Khan SA.** 2004. *Bacillus anthracis* and *Bacillus cereus* PcrA helicases can support DNA unwinding and in vitro rolling-circle replication of plasmid pT181 of *Staphylococcus aureus*. J Bacteriol **186:**2195-2199.

182. **Bruand C, Ehrlich SD.** 2000. UvrD-dependent replication of rolling-circle plasmids in Escherichia coli. Mol Microbiol **35:**204-210.

183. **Gwynn EJ, Smith AJ, Guy CP, Savery NJ, McGlynn P, Dillingham MS.** 2013. The conserved C-terminus of the PcrA/UvrD helicase interacts directly with RNA polymerase. PLoS One **8:**e78141.

184. **Petit MA, Dervyn E, Rose M, Entian KD, McGovern S, Ehrlich SD, Bruand C.** 1998. PcrA is an essential DNA helicase of *Bacillus subtilis* fulfilling functions both in repair and rolling-circle replication. Mol Microbiol **29:**261-273.

185. **Maluf NK, Fischer CJ, Lohman TM.** 2003. A dimer of *Escherichia coli* UvrD is the active form of the helicase *in vitro*. J Mol Biol **325:**913-935.

186. **Tomko EJ, Fischer CJ, Lohman TM.** 2012. Single-stranded DNA translocation of *E. coli* UvrD monomer is tightly coupled to ATP hydrolysis. J Mol Biol **418:**32-46.

Figure Legends.

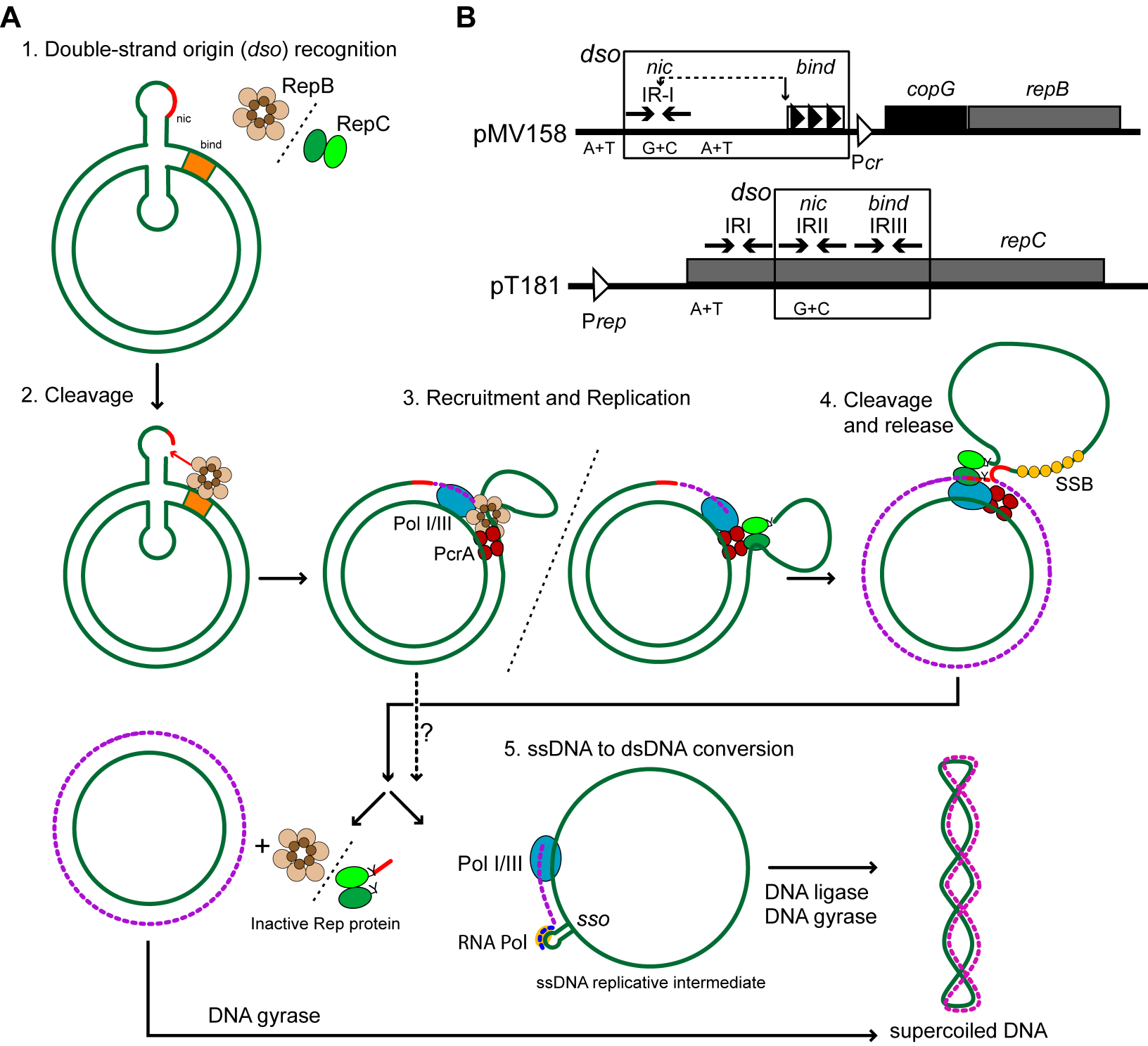
Figure 1. A) A model for plasmid RCR based on pMV158 and pT181 replicons. Detailed information about the RCR process is given in the text. In the pMV158 replication model, a possible mechanism is shown in which, upon assembly and cleavage at the nick site, the hexameric ring of RepB encircles one of the plasmid strands within the central channel. As discussed in the text, the strand enclosure may confer high processivity to the replisome complex. The RepB-mediated mechanism that, at the termination step, yields the dsDNA replication product and the ssDNA intermediate, as well as the mechanism of RepB inactivation, remain undisclosed (dotted arrow plus -?- symbol). **B) Scheme of the *dso*s and of the adjacent regions of pMV158 and pT181 RCR plasmids.** The symbols used are as follows: direct repeats in the replication region are indicated by solid boxed arrows; the inverted arrows represent the two arms of the inverted repeat elements; promoters are indicated as open arrowheads. The AT- and GC-rich sequences (A+T and G+C, respectively) are also indicated. The dotted line above the pMV158 map indicates that the direct repeats of the *bind* locus are separated by 84 bp from the nick site.

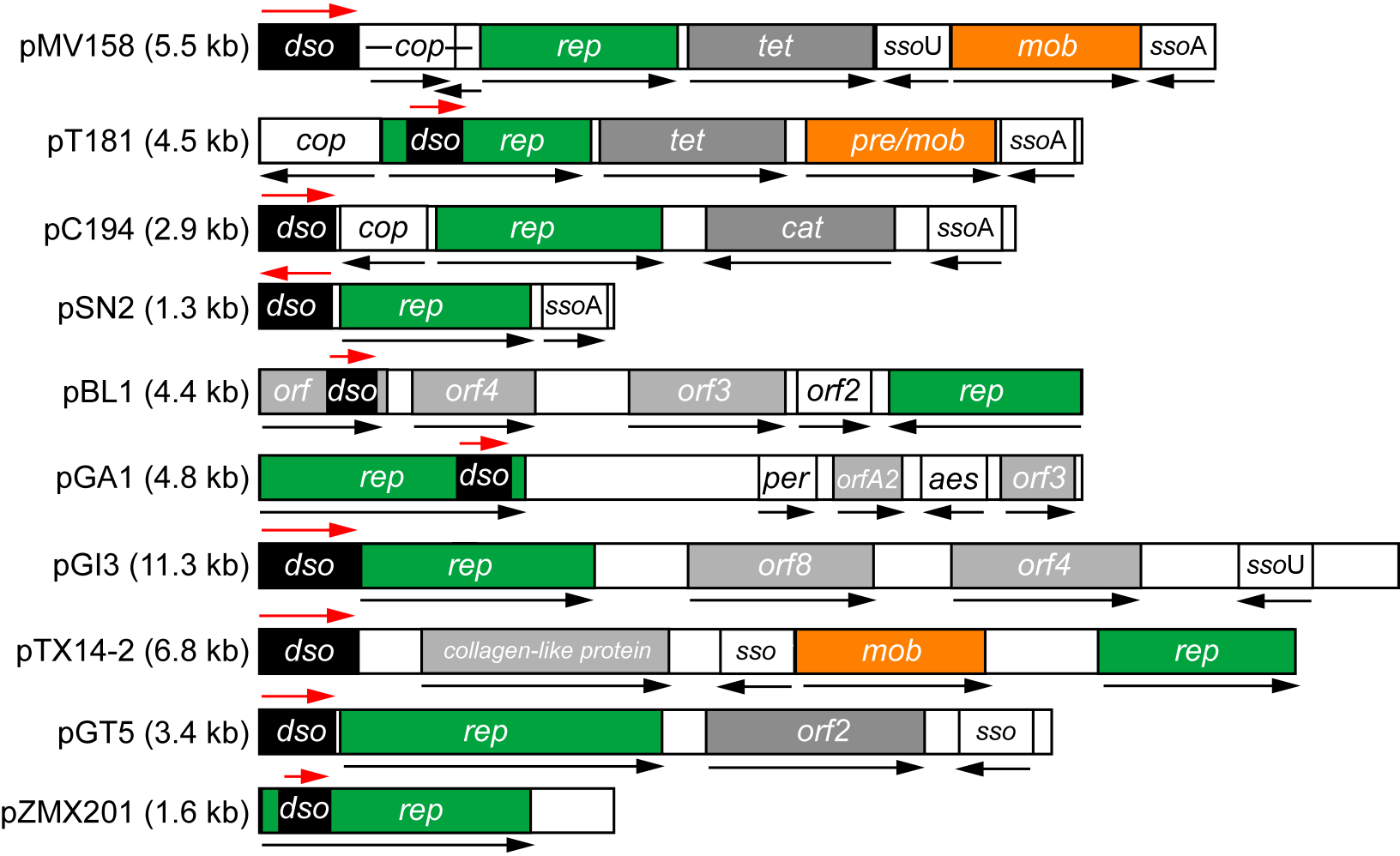
Figure 2. Functional organization of the RCR plasmids. Plasmids representative of the different families are shown. The arrows point to the direction of transcription (black) or the direction of replication (red) from the *dso* (leading strand) and *sso* (lagging strand). Inside the boxes, *rep* is the replication gene; *cop* represents the copy number control gene(s); *dso* is the double-strand origin of replication; *sso* is the single-strand origin of replication; *cat* and *tet* are chloramphenicol and tetracycline resistant genes, respectively; *pre/mob* represents the conjugative mobilization gene; *orf* indicates an open reading frame with unknown homology. The position of the copy number control genes *per* and *aes* of pGA1, and of the *collagen-like protein* gene of pTX14-2 is also indicated.

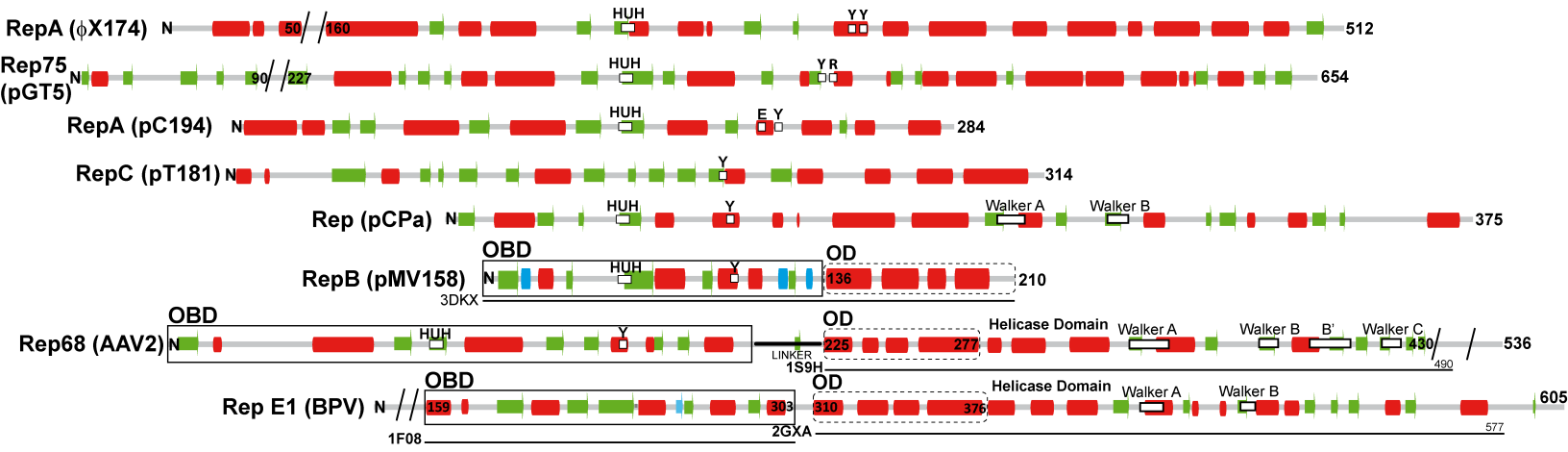
Figure 3. Domain structure of the Rep proteins from RCR plasmids. Predicted and observed secondary structures of the replication proteins of different RCR plasmids and of the Rep proteins from the adeno asociated virus (AAV) and bovine papilomavirus (BPV). The amino-terminal end (N) and the number of amino acids are indicated for each of the proteins analyzed. The predicted or observed α-helices and β-strands are represented as red and green bars, respectively. The 3_{10}-helices are represented as blue bars. Conserved amino acid residues of the active site involved in

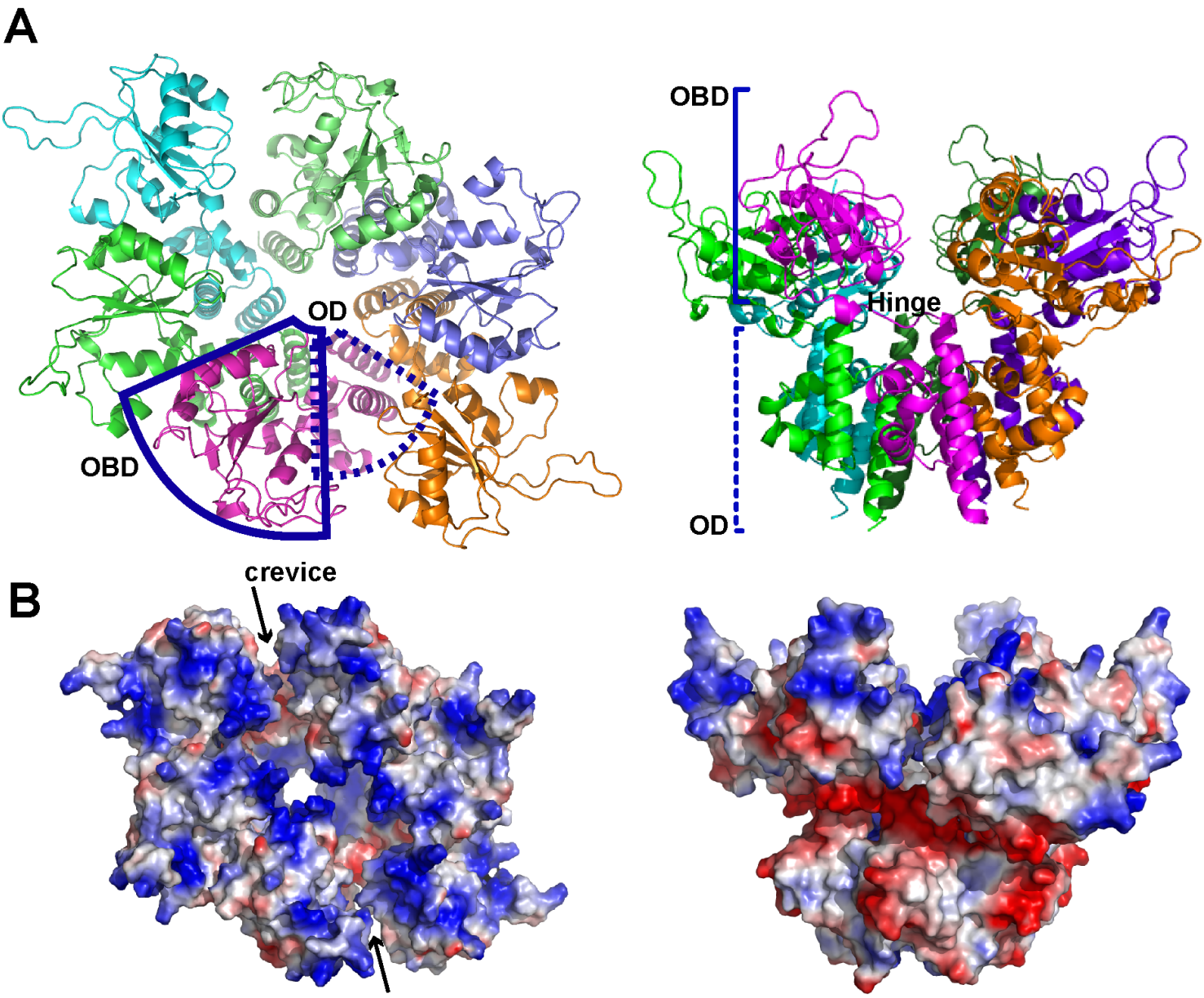
metal binding (HUH) and in the endonucleolytic activity are indicated in the protein maps. The conserved Walker A, B and C motifs are indicated in the proteins with a helicase domain. The limits of the origin binding domain (OBD) and of the oligomerization domain (OD) are indicated in the protein maps of RepB, Rep68 and E1. The additional line below the sequence of RepB, Rep68 and E1 shows the secondary structure present in the crystal structure of the protein (PDB entry code is given in the figure). Plasmidic Rep proteins were aligned by the metal binding HUH motif. However, viral Reps were aligned with RepB by the all-helical OD domain due to the structural similarity found in this region.

Figure 4. Cartoon representation of the structure of RepB obtained by X-ray crystallography. A) Top (left) and side (right) views of the RepB hexamer are shown. Location of the OBD (continuous line) and of the OD (dotted line) is also indicated in the two views. Position of the hinge connecting both domains is indicated in the side view. **B)** Top (left) and side (right) views of the electrostatic potential on the solvent-accessible surface of the RepB hexamer structure. The location of the crevice is indicated.


A
1. Double-strand origin (dso) recognition
RepB
RepC
nic
bind
2. Cleavage
3. Recruitment and Replication
Pol I/III
PcrA
4. Cleavage
and release
SSB
?
5. ssDNA to dsDNA conversion
Pol I/III
RNA Pol
sso
ssDNA replicative intermediate
DNA ligase
DNA gyrase
Inactive Rep protein
DNA gyrase
supercoiled DNA
B
dso
nic
bind
IR-I
copG
repB
pMV158
A+T
G+C
A+T
Pcr
dso
nic
bind
IRI
IRII
IRIII
repC
pT181
Prep
A+T
G+C


pMV158 (5.5 kb)
dso
cop
rep
tet
ssoU
mob
ssoA
pT181 (4.5 kb)
cop
dso
rep
tet
pre/mob
ssoA
pC194 (2.9 kb)
dso
cop
rep
cat
ssoA
pSN2 (1.3 kb)
dso
rep
ssoA
pBL1 (4.4 kb)
orf
dso
orf4
orf3
orf2
rep
pGA1 (4.8 kb)
rep
dso
per
orfA2
aes
orf3
pGI3 (11.3 kb)
dso
rep
orf8
orf4
ssoU
pTX14-2 (6.8 kb)
dso
collagen-like protein
sso
mob
rep
pGT5 (3.4 kb)
dso
rep
orf2
sso
pZMX201 (1.6 kb)
dso
rep


RepA (ϕX174)
N
50
160
HUH
Y Y
512
Rep75
(pGT5)
N
90
227
HUH
Y R
654
RepA (pC194)
N
HUH
E Y
284
RepC (pT181)
N
Y
314
Rep (pCPa)
N
HUH
Y
Walker A
Walker B
375
RepB (pMV158)
OBD
OD
N
HUH
Y
136
210
3DKX
Rep68 (AAV2)
OBD
N
HUH
Y
OD
225
277
Helicase Domain
Walker A
Walker B
B'
Walker C
430
490
536
LINKER
1S9H
Rep E1 (BPV)
N
OBD
159
303
OD
310
376
Helicase Domain
Walker A
Walker B
605
577
1F08
2GXA


A
OD
OBD
OBD
Hinge
OD
B
crevice