Mutual interaction enables the mycobacterial plasmid pAL5000 origin binding protein RepB to recruit RepA, the plasmid replicase, to the origin

Soniya Chatterjee,1 Madhu Manti Patra,1 Sourabh Samaddar,1 Arnab Basu2 and Sujoy K. Das Gupta1,*

Abstract
The Mycobacterium fortuitum plasmid, pAL5000, is the most-studied member of a family of plasmids that are found in Actinobacteria. Its replication is brought about by the combined action of two plasmid-encoded replication proteins, RepA and RepB. RepB has earlier been shown to be a sigma factor homologue that possesses origin-binding activity. The mechanism by which RepA functions, and its relationship with RepB, if any, has not been explored yet. In this study, we show that RepA shares a common catalytic domain, with proteins belonging to the primase-polymerase and DNA polymerase X families. We demonstrate that RepA is functionally a DNA polymerase and that mutations that alter two conserved aspartic acid residues present within the catalytic core lead to inactivation of plasmid replication. Replication of pAL5000 was shown not to depend on the host primase, and thus it is most likely that RepA is responsible for the priming act. We further demonstrate that RepA and RepB function as a pair and that the functional cooperation between the two requires physical contact. The C-terminal domain of RepA, which is structurally a helical bundle, is responsible for unwinding the origin in a site-specific manner and also for the establishment of contacts with RepB. The results presented show that RepB functions by recruiting RepA to the origin in much the same way as sigma factors recruit RNA polymerase core enzyme to promoters.

INTRODUCTION
The evolutionary origin of DNA replication remains an intriguing issue. The lack of conservation in the components of the replication systems belonging to the three domains of life suggests that they may have originated at different times and independently of each other. It has been proposed that DNA replication systems initially evolved in plasmids and phages and were subsequently transferred to the cellular forms of life [1]. Hence, by studying plasmid/phage replication processes it may be possible to obtain an insight into the replication mechanisms that existed in the early stages of evolution of life on Earth. Replication and transcription are fundamental processes in life. There is a basic similarity between the two – both use DNA as a template to synthesize new strands of nucleic acids. Despite the similarity, the proteins involved in executing the processes do not appear to be evolutionarily related [2]. However, in the ancient past the situation could have been different. It is very likely that in the early stages of evolution of life, the processes of replication and transcription were more related to each other than they are at present.

The plasmid pAL5000 belongs to a large family of plasmids that are found in Actinobacteria. The minimal replication region of this plasmid is a 1.8 kb fragment comprising a cis-acting origin sequence and two ORFs, ORF 1 and ORF 2, that code for the replication proteins RepA and RepB, respectively. That RepB is the origin binding protein has been established in several earlier investigations [3, 4]. However, the role of RepA remains enigmatic. The synthesis of RepA and RepB is translationally coupled [5]. Although both proteins are required for plasmid replication, it is unclear as to whether there is any mutualistic relationship between the two [4, 6]. The pAL5000 family is distantly related to the CoE2 plasmids of E. coli [7]. CoE2 family plasmids encode a replicate family protein known as Rep, that shares significant sequence homology with pAL5000 RepA [8, 9]. The major difference, however, lies in the fact that whereas the CoE2 plasmids replicate using a single replication protein – the Rep protein [10], pAL5000 and its

*Correspondence: Sujoy K. Das Gupta, sujoy@jcbose.ac.in
Keywords: Mycobacterium; DNA replication; plasmid; DNA polymerase.
Abbreviations: EMSA, electrophoresis mobility shift assay; HTH, helix-turn-helix; MCS, multi-cloning site; NTA, Ni²⁺-nitrilotriacetic acid; qRT-PCR, quantitative real-time PCR.

Three supplementary figures and two supplementary tables are available with the online Supplementary Material.
The culture media were supplemented with the necessary incubation was done at 37°C (supplemented with 0.25% BSA, 0.2% glycerol and 0.05% plasmid DNA. For overproduction of polyhistidine (His6)-tagged RepA and RepB, BL21 (DE3) was used as the host. In some cases, RepA or its mutated versions, pET28a (Stratagene), with maltose binding protein (MBP) using the vector pMAL-c2X (New England Biolabs). Mycobacterium smegmatis strain mc²155 [13] was used for mycobacterial transformations. The mycobacterial vectors used in this study are pMC2 [14], pMIND [15] and pLAM [16]. For expression of the repA gene in mycobacteria, an integrative version of the acetamide inducible vector pLAM, described below, was used. The mycobacterial vectors confer resistance to kanamycin (KanR) to their hosts except pMIND, which also confers hygromycin (HypR) resistance.

**Bacterial growth conditions**

*M. smegmatis* and *E. coli* cells were cultured in MB7H9 (supplemented with 0.25% BSA, 0.2% glycerol and 0.05% Tween 80) and 2XYT broths, respectively. In both cases, incubation was done at 37°C under shaking conditions and the culture media were supplemented with the necessary antibiotics. Bacterial growth was monitored spectrophotometrically by measuring the optical density of the culture medium at 600 nm (OD₆₀₀). For transformation experiments, cells were grown to mid-log phase (OD₆₀₀ 0.6–0.8). All transformations were done by the method of electroporation as described in earlier studies [14]. Induction of gene expression from recombinant plasmid vectors was achieved by the addition of acetamide (0.2% w/v) in the case of mycobacteria and 1 mM IPTG in the case of *E. coli*.

**Chemicals**

Ni²⁺-nitrilotriacetic acid (NTA) agarose used for the affinity chromatographic preparation of His₆-tagged protein was purchased from Qiagen. Amylose resin used for purifying MBP-tagged proteins was obtained from New England Biolabs. Doxorubicin, an inhibitor of mycobacterial growth, was purchased from Sigma-Aldrich. Restriction enzymes used for cloning were purchased from New England Biolabs. Other chemicals for protein expression, purification and analysis, of the highest grade of purity, were obtained from SRL Laboratories.

**Cloning and expression of repA gene and its mutated versions**

For obtaining RepA in the pure form, a recombinant plasmid (pSOC4) was created by PCR-based cloning of the repA-repB region of pAL5000 in the multi-cloning site (MCS) of pET28a. To create mutated versions of RepA, the aspartate residues located at positions 83 and 85 were changed to alanine, either one at a time (D83A and D85A, respectively) or both simultaneously (D83A+D85A). The resulting constructs were named pSOC4.1, pSOC4.2 and pSOC4.3, in that order. Mutations were introduced into the repA gene using the Quick Change Site-Directed Mutagenesis kit (Stratagene). RepA and its mutated versions that were produced from the pSOC4 series of vectors all carry His₆-tags at their respective N-terminal ends. For the synthesis of the MBP-tagged version of RepA in *E. coli*, the ORF corresponding to RepA was PCR-amplified and cloned into the MCS of pMAL-c2X, resulting in pSOC4.1 and its derivatives (refer to Table S1, available in the online Supplementary Material). The sequences of all the primers used in this study and their applications are mentioned in Table S2.

**Complementation assays**

For complementation experiments, replication-defective versions of pMIND, pMINDrepA1 and pMINDrepA2 were created by introducing mutations into the repA gene of the vector at the codons corresponding to D83 and D85, respectively. The mutated vectors were then introduced into the recombinant strain, *M. smegmatis*::pSOC1-int, that carries an integrated copy of the wild-type repA gene in its genome, the expression of which can be triggered by the addition of acetamide [17]. The strain was constructed by electroporating *M. smegmatis* cells with the integrative vector pSOC1-int. The integrative vector was created by cloning the repA gene of pAL5000 in the MCS of the mycobacterial expression vector pLAM12 to generate pSOC1. The plasmid vector
pSOC1 was then converted into an integrative one by introducing an integration cassette derived from mycobacteriophage L5 into it. The method used to construct the integrative vector pSOC1-int was the same as that used earlier in a different context [18]. The vector pMIND was used in these experiments as it confers hygromycin resistance, a phenotype that could be used to study genetic supplementation. The M. smegmatis :: pSOC1-int host cells were treated with acetamide to induce RepA synthesis prior to making them competent for transformation. Transformation efficiency was expressed as colony-forming units per µg plasmid DNA (c.f.u. µg⁻¹ DNA). The lower limit for detection of transformants was 10 c.f.u. µg⁻¹ DNA.

**Plasmid copy number analysis**

Quantitative real-time PCR (qRT-PCR) was performed to ascertain changes in the copy number of pMC2, a pAL5000-derived vector, in M. smegmatis cells following treatment with doxorubicin (a chromosomal DNA replication inhibitor). Total DNA was isolated from treated and untreated cells by using a genomic DNA isolation kit (Hi-Media), and subjected to qRT-PCR analysis using primers specific for either repA, the plasmid marker, or the 16S rRNA gene, the chromosomal one (refer to Table S2). Equal volumes of the samples (1 µl in each case) were taken and qRT-PCR reactions were set up using Power Syber Green PCR 214 Master Mix (Applied Biosystems) and 2 pmol of each primer. The reactions were carried out in 7500 Fast real-time PCR equipment (Applied Biosystems). The thermal cycling conditions were as follows: 50 °C for 2 min and 95 °C for 8 min, then 45 cycles of denaturation at 94 °C for 10 s, annealing at 59 °C for 15 s and extension at 72 °C for 20 s and the last step consisted of a melting curve analysis. To measure copy number difference between untreated and treated samples, the threshold cycle values (Ct) obtained using plasmid-specific primers were subtracted from the corresponding ones obtained using the chromosome-specific primers (ΔCt). The differences in the ΔCt values (ΔΔCt) between treated and untreated samples were then calculated. The fold increase in copy number is given by the formula 2−ΔΔCt.

**Co-expression of chaperon encoding genes**

RepA was found to be produced as an insoluble protein when it was synthesized in E. coli using pET28a vector. To make it soluble, chaperones GroEL and GroES, which are known to facilitate folding and improve solubility of heterologous proteins in E. coli, were co-produced in the same cells in which RepA was synthesized. E. coli BL21 (DE3) cells were thus first transformed with the recombinant pET28a-based vector expressing repA and then further transformed with pTacSL1 [19]. Selection of doubly transformed cells was done using Kan⁷ (for pET28a) and Spec⁷ (spectinomycin for pTacSL1).

**Purification of pAL5000 replication proteins**

The plasmid constructs made for over-expressing pAL5000-derived replication genes were transformed into E. coli BL21 (DE3), along with the groELS operon expressing plasmid pTacSL1. The transformsants resistant to both kanamycin, 50 µg ml⁻¹ (pET28a derivatives), and spectinomycin, 25 µg ml⁻¹ (for pTacSL1), were cultured in the presence of the two antibiotics to mid-log phase (OD₆₀₀ 0.5). The common inducer IPTG for both plasmids was then added to a final concentration of 1 mM. The bacterial cells were induced for 3 h at 37 °C and then centrifuged at 8000 r.p.m. for 10 min to collect the pellet. The cells, thus harvested, were lysed by sonication. RepA and RepB purification was then performed using Ni-NTA agarose affinity chromatography as described previously [8]. When RepA was produced as an MBP fusion protein, the affinity column chosen was amylose resin. The soluble extracts were passed through a column containing this resin, followed by elution using maltose. The procedure followed to purify MBP-tagged proteins is described in detail in an earlier publication from this laboratory [5]. To confirm that RepA-derived polypeptides were present in the eluted fractions, Western blot analysis was done using an anti-polyhistidine antibody preparation (His-probe Antibody, Santa Cruz Biotechnology) as per the standard protocol. For the purification of RepB, a vector (pTAB3) from which both RepA and RepB are synthesized in a coupled manner was used [8]. The coupled synthesis was necessary as otherwise RepB does not fold efficiently to its native form [5]. The construct pTAB3 is designed so that the RepB, which is produced from it, is His₅-tagged at its C-terminal end. Through such a strategy, native RepB in a highly purified form can be obtained as described previously.

To estimate the concentration of the purified proteins, a semi-quantitative approach was taken in which a 10 µl aliquot of the test sample was loaded on an SDS-polyacrylamide gel along with increasing volumes (1, 2 and 4 µl) of a BSA standard solution. After running the gel, and staining with Coomassie blue, band intensities were determined densitometrically using a Versadoc imaging system (Bio-Rad). An intensity versus protein quantity standard curve was then created for the BSA sample. The amount of protein associated with the band of interest was estimated by interpolating the corresponding intensity value on to the standard curve. The final concentration was expressed in molar terms (µM).

**Antiserum preparations**

For antiserum preparation the replication protein RepA was isolated by performing affinity chromatography under denaturing conditions in the presence of 8 M urea as per the standard protocol [3]. The highly purified protein sample was further gel-purified and injected into rabbits. The use of antiserum against RepB has been documented earlier [3]. The same antiserum was used in this study too. While analysing the specificities of the various antisera, it was found that the sera of pre-immune rabbits showed some background cross-reactivity against RepA, particularly its N-terminal region. Therefore, in experiments involving immune detection of RepA, pre-immune serum was not used as a negative control. Other forms of negative controls were used to ensure specificity of the results obtained. For
probing His$_6$-tagged proteins by Western blotting, anti-His tag antibodies (Santa Cruz Biotechnology) were used as mentioned above.

**DNA polymerase assay**

DNA polymerase activity was measured using 5'-labelled primer-template DNA substrate (primer extension assay). The substrate was prepared by first labelling the 5' end of a 22-mer primer NG13 (5'GTAACCGTGGATCTTCAGAGT3') with [$\gamma$-32P]ATP (BRIT, India) and hybridizing it to a 45-mer oligonucleotide NG15 (5'GTTATGGCATGAAGCCGGC GTAATCAGAGGATCCAGGGAT3'). The annealing reaction was carried out by mixing 100 pmol each of the complementary oligonucleotides in 50 µl annealing buffer [40 mM Tris-HCl (pH 7.5), 80 mM NaCl, 10 mM MgCl$_2$, 1 mM DTT], followed by heating to 95 °C for 2.5 min and gradual cooling to room temperature. The annealed products were subsequently purified by PAGE using the 'crush-and-soak' method as described previously [20]. The gel-purified annealed product thus obtained (primer-template substrate) was directly used for DNA polymerase assay. The assays were done in a reaction mixture (50 µl) containing 1 nM primer-template substrate, 0.2 µM of RepA, 25 mM Tris-HCl (pH 7.5), 1 mM DTT, 10 mM MgCl$_2$ and 0.2 mM dNTPs. After incubation for 30–45 min at 30 °C, the reactions were stopped with 5× stop solution (1.25 % SDS, 75 mM EDTA, 25 % glycerol). The reaction products were purified by using the phenol chloroform extraction method and analysed by performing (10 %) PAGE under denaturing condition (8 M urea). For quantitative analysis, DNA polymerase assays were performed by labelling the extension product with [$\alpha$-32P]dCTP. To measure the quantity of the product formed, the intensity of the autoradiographic band corresponding to the full-length de novo synthesized DNA was determined densitometrically.

**DNA binding assay**

Electrophoretic mobility shift assay (EMSA) was performed using the origin probe as described previously [21, 22]. DNA fragments harbouring the origin sequence were isolated and labelled at their 5' ends using [$\gamma$-32P]ATP. The EMSA reaction mixture (30 µl) contained 3 µl of 10× binding buffer [100 mM Tris (pH8), 600 mM NaCl, 30 mM MgCl$_2$, 1 mM EDTA, 1 mM DTT, 20 % glycerol], 1 µg of salmon sperm DNA, 3 ng of labelled probe (10 000 c.p.m.) and pAL5000-derived replication proteins as specified. The reaction mixtures were incubated in ice for 20 min and the DNA protein complexes formed were resolved by performing 5 % native PAGE at 130 V for 3–4 h at 4 °C after a pre-run for 1 h at 100 V. Electrophoresis was carried out in 0.5× Tris-borate buffer (50 mM Tris-borate, 1 mM EDTA). Following electrophoresis, the gel was dried and bands were visualized by autoradiography. For performing antibody-supershift assays, an optimum amount of anti-RepA antiserum (3 µl) was added while setting up the binding reaction.

DNase I footprinting analysis was performed as described previously [5] with certain modifications. A 200 bp DNA fragment derived from the pAL5000 origin was amplified using primers S1 (5'TCCTGGTTGTACAGGTTTGGG3') and S2 (5'GCTGCTAAATTGCTGGCGG3'). The primer S2 (10 pmol) was labelled with [$\gamma$-32P]ATP using T4 polynucleotide kinase (Thermo Scientific) using the standard protocol and was used directly for PCR. The PCR product was run on a 1.3 % agarose gel and purified using a Qiagen gel extraction kit. This 200 bp pAL5000 origin of replication (nucleotides 4459–4663) encompasses the high- and low-affinity binding sites for RepB [4]. Approximately 0.3 pmol labelled DNA was incubated with desired amounts of pAL5000-derived replication proteins. Binding reactions were carried out under the conditions mentioned above. After the reaction, 50 ng DNase I (Sigma Aldrich) was added and the mixture incubated at 37 °C for 3 min. The digestion was stopped by adding 2 µl of 0.5 mM EDTA. The digested fragments were then purified using phenol-chloroform extraction method and then alcohol-precipitated. Finally, after a 70 % ethanol wash the samples were dried and re-suspended in loading buffer [98 % (v/v) deionized formamide, 10 mM EDTA, 0.025 % (w/v) xylene cyanol and 0.025 % (w/v) bromophenol blue] followed by boiling for 5 min and rapid chilling. The DNase I digested fragments were then separated on an 8 % polyacrylamide (8 M) urea sequencing gel by electrophoresis at 1300 V for 3.5 to 4 h. The gel was dried on Whatman paper and exposed to Kodak BioMax film. An A+G ladder, used as a marker, was prepared by limited exposure of the DNA to formic acid followed by piperidine treatment.

**Protein–protein interaction studies**

The interaction between RepA and RepB was studied by performing ELISA assays. His$_6$-tagged RepB, either wild-type or its mutated version, was immobilized onto the wells of a Ni-NTA-coated microtiter plate (Ni-NTA HisSorb, Qiagen). Following washing and blocking with BSA, MBP-tagged RepA was added. The amount of MBP-RepA that was immobilized in the wells due to specific interaction with RepB was monitored using anti-RepA antiserum. Coulour development was done by adding a secondary antibody tagged with alkaline phosphatase to the wells followed by incubation in the presence of the substrate para-nitrophenylphosphate (pNPP). Readings were taken at 405 nm (OD$_{405}$). All necessary controls were incorporated to eliminate the possibility of non-specific binding of MBP-RepA to the wells. In each ELISA experiment, three technical replicates were always performed and the results expressed as mean OD$_{405}$ ± standard error. The protocol followed is essentially the same as that used in a previous study, where the interaction between mycobacterial DnaK and HspR was investigated [23].

Interaction studies between His$_6$-RepA and RepB were also done by performing pull-down assays using Ni-NTA agarose beads. Extracts of *E. coli* co-producing His$_6$-RepA and RepB were subjected to Ni-NTA agarose affinity chromatography using the same protocol that was used to purify His$_6$-tagged Rep proteins. The polypeptides that were bound to the Ni-NTA agarose beads were eluted and analysed by Western blotting using anti-serum against either RepA or RepB.
Sequence alignment and phylogenetic tree
The alignments were done using CLUSTALW [24] with MEGA 5.0 [25] software. Multiple alignment penalties of 10 and 0.1 were used for gap opening and extension, respectively. The alignments created with MEGA5.0 were subsequently processed using Bioedit Sequence Alignment Editor for convenient schematic representation. Phylogenetic trees were constructed using the neighbour-joining method with the help of MEGA5.

RESULTS
Sequence comparison of RepA with its homologues
The replication region of pAL5000 spans a segment of approximately 1.8 kb length (Fig. 1a), which includes a 500 bp region that functions as the origin (Ori) and two genes, repA and repB, that code for the replication proteins RepA and RepB, respectively. Of the two proteins, the latter is known to bind to the origin, while the function of the former

---

**Fig. 1.** (a) Schematic representation of the replication region (1.8 kb) of pAL5000, showing the relative positions of the origin (Ori) and the replication genes repA and repB. (b) Domain structures of RepA and its homologues: Rep of ColE2, RepB¢ of pRSF1010, ORF904 of pRN1 and rat DNA polymerase β, which belongs to the X family of DNA polymerases (PolX). The N-terminal domains are named as either replicase or primase-polymerase (prim-pol) or trivially as N-terminal, based on the information derived from previous studies. In the case of PolX, the region is referred to as the catalytic domain. In addition, PolX has two other important domains known as finger and thumb. The region that is highly conserved in the N-terminal domains of these proteins is shaded black. The C-terminal regions are named either as primase C-terminal_1 (PriCT_1) or the helical domain. In the case of pRN1 replication protein ORF904, a polypeptide having helicase activity is present at its C-terminal end. The RepB of pAL5000 is shown separately since it is encoded by a separate gene. The HTH motif of ColE2 Rep and the σ3- and σ4-like domains of RepB are indicated. (c and d) Alignment of the replicase and PriCT_1 domains, respectively, of the DNA replication proteins mentioned. The conserved residues are highlighted in black. The two conserved aspartates that are present in the active centers of these polypeptides are indicated by asterisks. (e) Phylogenetic relationship (neighbour-joining) between homologues of RepA. The tree was constructed using an alignment of the N-terminal conserved regions. The scale (horizontal bar below the tree) signifies evolutionary distance in the unit of the number of amino acid substitutions per site.
has not yet been established. To obtain an insight into what could be the function of RepA, we first performed an *in silico* analysis. Conserved-domain analysis revealed that the RepA of pAL5000 possesses two distinct domains: one of which is the N-terminal replicase domain and the other the C-terminal domain which is referred to in the databases as primase C-terminal (PriCT_1). The C-terminal domain is so named as it shares homology with the alpha helical domain found in the C-terminal region of archeo-eukaryotic primases [26]. RepB is produced as a separate entity (Fig. 1b) from an ORF located immediately downstream of the one that codes for RepA (Fig. 1a). In a previous study, it was reported that the N- and C-terminal domains of RepB share sequence homology with regions 3 ($\sigma$3) and 4 ($\sigma$4) of the ECF family sigma factors, respectively [8]. RepA shares homology with the replication protein of CoLE2 class plasmids known as Rep [27]. An important difference between CoLE2 Rep and pAL5000 RepA is that in the case of the former the origin recognizing helix-turn-helix (HTH) domain is an integral part of the protein, whereas, in the latter it is not. Since pAL5000 RepA does not possess a DNA-binding domain, it is unable to bind to the corresponding origin, as noted in an earlier study [4] and here too (refer to Fig. 6b). The N-terminal region of RepA shares homology with not only CoLE2 Rep, but also with replication proteins of other plasmids such as RepB of the broad-host-range plasmid pRSF1010 [28] and ORF904 of the archaeal plasmid pRN1 [29] (Fig. 1b). Also, RepA of pAL5000 appears to be distantly related to the DNA polymerase X family of proteins [30] that are known for their DNA-repair-related activities [1, 31]. Within the N-terminal domains of the proteins mentioned above, there is a region (shaded black in Fig. 1b) which is highly conserved. The most important feature of this region is the presence of three aspartates, two of which (Fig. 1c, indicated by asterisks) are known to be involved in chelating Mg$^{2+}$ ions [32]. These Mg$^{2+}$ ions are required in the two-metal ion mechanism of DNA polymerase for phosphoryl transfer. Unlike the N-terminal region, the C-terminal domain is less conserved, although a low-level sequence homology and conservation of a few amino acid residues was observed (Fig. 1d). Sequences corresponding to the N-terminal conserved regions of RepA and its distant relatives were aligned and the resulting alignment was used to construct a phylogenetic tree (Fig. 1e) The results show that RepA-type plasmid replicases are a part of a larger family of proteins endowed with the ability to replicate DNA.

**Synthesis of recombinant RepA in *E. coli***

To functionally characterize the protein, the operon repA-repB was expressed in *E. coli* using the construct pSOC4. From this construct, both RepA and RepB are produced (Fig. S1) as a result of coupled synthesis (schematically

![Fig. 2](image-url)
illustrated on top of the SDS-PAGE profiles, Fig. 2a–c); however, only RepA (which carries the His_{6} tag) is expected to bind to the Ni^{2+} NTA agarose affinity column. For RepB (Fig. 2f) purification we used a similar approach, except that in this case RepB, and not RepA, was tagged (at the C-terminal end) [5]. RepA is poorly soluble and therefore it tends to be associated with the pellet fraction, and not the soluble supernatant following centrifugation of the cell-lysate. Thus, it was necessary to investigate the conditions under which the protein can be obtained in the soluble form. Initial attempts to purify the protein resulted in low yields. Moreover, in the purified samples we could detect the presence of GroEL. We therefore argued that presence of GroEL may make a difference to the solubility of RepA. To improve the solubility of RepA, GroEL was co-synthesized from a compatible plasmid. The affinity column bound fraction was eluted and then resolved by performing SDS-PAGE. From the SDS-PAGE profile (Fig. 2a), it became apparent that several other proteins co-purify with RepA (the identity of RepA was confirmed by Western blot analysis using anti-His tag antibodies; refer to Fig. 3a, b). Among the additional bands, the major one migrating immediately above RepA was found to be that of GroEL, as determined by performing MALDI-TOF/TOF mass spectrometry. A few other bands corresponding to proteins that co-eluted were found to be present, although their identities have not been established.

Attempts were then made to synthesize the individual domains using the same strategy as used in the case of the full-length protein. When the synthesis and purification of the N-terminal domain alone was attempted, a band corresponding to the desired product was obtained, but in this

![DNA polymerization activity of His_{6}-tagged RepA and its mutated versions](image)

**Fig. 3.** DNA polymerization activity of His_{6}-tagged RepA and its mutated versions, as indicated. (a) SDS-PAGE and (b) Western blot analysis, using His probe antibody, of the indicated affinity purified RepA samples. (c) Primer extension assay using RepA (and its mutated versions). The 5’ labelled annealed primer-template substrate used for the assay is diagrammatically illustrated above. In the lanes marked (−) either protein or dNTP was omitted, as indicated. (d) Determination of the optimal Mg^{2+} ion concentration for the DNA polymerization activity of RepA. The primer-extended product was labelled using [α-^{32}P]dCTP (asterisk in the diagram above). (f) Primer extension assay using RepA, either wild-type or the mutated version D85A, in the presence of suboptimal concentrations of Mg^{2+} ions. (e and g) The intensities of the bands corresponding to the full-length products were estimated and plotted against Mg^{2+} ion concentration (e and g corresponding to d and f, respectively).
case too GroEL was found to co-purify (Fig. 2b). In the case of the C-terminal domain, a few additional minor bands were visible in the high molecular weight region, although none of them appeared to represent GroEL (Fig. 2c). Thus, the association of GroEL with RepA involves the N-terminal domain of RepA and not its C-terminal one.

RepA either fulllength or its N-terminal domain only was also produced in fusion with MBP (Fig. 2d, e). The MBP fusions of RepA were relatively more soluble and co-production of GroEL at the induced level was not necessary to ensure solubility. However, as observed while purifying the His6-tagged version of RepA, in the case of the MBP-tagged version too GroEL was found to co-purify. Thus RepA, particularly its N-terminal domain, appears to have a high affinity for GroEL. Unlike RepA, in the case of RepB, both wild-type and its N-terminal-deleted version (ΔNRepB) were highly soluble and could be purified to homogeneity with relative ease (Fig. 2f).

**DNA replicase activity of RepA**

In the N-terminal domain of RepA, a conserved motif DXD-X_{48-64}-D/E is present, which is also found in the active sites of primase-polymerases and X family DNA polymerases (Fig. 1c). Hence, we speculated that RepA could be functionally a DNA polymerase and proceeded to test this possibility experimentally. For this purpose, we used the affinity-purified RepA samples described above. Prior to using these samples, the authenticity of the His6-tagged RepA band in the SDS-PAGE (Fig. 3a) was confirmed by performing Western blot analysis using anti-His tag antibodies (Fig. 3b). In addition to wild-type RepA protein, we included in the assay RepA(N-terminal) and three mutated versions of the protein, D83A, D85A and (D83A+D85A). A primer extension assay was then performed using a [α-32P]dCTP labelled primer-template substrate (Fig. 3c). As positive control DNA polymerase I Klenow fragment obtained from a commercial source was included. The results of these experiments show that the affinity purified RepA sample was capable of extending the primer, indicating that the enzyme functions as a DNA polymerase (Fig. 3c, lane marked RepA). Interestingly the N-terminal domain of RepA alone (lane marked N-term) failed to give activity. Thus, although the N-terminal catalytic domain is necessary it may not be sufficient – the PriCT_1 domain also has a role to play. The results further indicate that the observed polymerization activity is due to RepA and not due to some other protein that is co-eluted.

The protein variants RepA D83A and 85A continued to show some degree of activity (Fig. 3c; D83A and D85A), which was unexpected. However, the variant in which both the aspartic acid residues were substituted (D83A+D85A) failed to show activity. The role of the D residues is to chelate Mg\(^{2+}\) ions at the active center. We argued that if high Mg\(^{2+}\) ion concentration is used, then the function of the D residues may become mutually redundant. However, at lower Mg\(^{2+}\) ion concentration both the D residues should be essential individually. To test such a hypothesis, we used a quantitative primer extension assay, in which the extended product was internally labelled with [α-32P]dCTP (Fig. 3d). The amount of the product formed was determined by densitometric scanning of the band corresponding to the full-length product that is produced in such assays. When the Mg\(^{2+}\) ion dependence of the primer extension activity of RepA was examined, it was found that the activity increased with Mg\(^{2+}\) ion concentration and reached a saturation value at about 2 mM (Fig. 3d and graphical representation, e). Thus, the 10 mM Mg\(^{2+}\) ion concentration used in Fig. 3(c) was substantially above saturation. To investigate whether the D residues were at all essential, we examined the activity of RepA D85A, using less than optimal Mg\(^{2+}\) ion concentration. Under such conditions, the D85A variant displayed little or no activity (Fig. 3f and graphical representation g). Thus, D85 is essential for the activity of RepA. Similar results were obtained for D83A (data not shown).

Further proof regarding the essentiality of the aspartates (D83 and D85) were obtained from mutational studies in vivo which show that mutations of these residues were indeed deleterious (Table 1, pMINDrepA1 and repA2). Thus, the result shows that the DNA polymerization activity of RepA is essential for plasmid replication. However, the question may arise as to whether the adverse effect of the mutations is due to inactivation of repA or a polar effect on the expression of repB. To resolve this issue, complementation analysis was performed using a recombinant strain of *M. smegmatis*, *M. smegmatis::pSOC1-int* that produces RepA from an integrated copy of the gene. The results (Table 1) show that whereas the mutated vectors

<table>
<thead>
<tr>
<th>Replication genes</th>
<th>Vector</th>
<th>Mutation</th>
<th>Host strain</th>
<th>Transformation efficiency (c.f.u. µg(^{-1}) DNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ori-repA-repB</td>
<td>pMIND</td>
<td>Wild-type</td>
<td><em>M. smegmatis</em></td>
<td>3.5×10^5</td>
</tr>
<tr>
<td>Ori-repA(D83A)-repB</td>
<td>pMIND repA1</td>
<td>repA(D83A)</td>
<td><em>M. smegmatis</em></td>
<td>Nil</td>
</tr>
<tr>
<td>Ori-repA(D85A)-repB</td>
<td>pMIND repA2</td>
<td>repA(D85A)</td>
<td><em>M. smegmatis::pSOC1-int</em></td>
<td>5×10^4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>M. smegmatis</em></td>
<td>4.5×10^4</td>
</tr>
</tbody>
</table>

*Transformation efficiency <10 c.f.u. µg\(^{-1}\) DNA.*
pMINDepA1 and repA2 were unable to multiply in the M. smegmatis strain, they did so efficiently in M. smegmatis :: pSOC1-int. Hence, the adverse effect of the mutations was due to the inactivation of repA and not a polar effect on the expression of repB, the next essential gene in the operon. These results also show that RepA and RepB need each other to bring about plasmid pAL5000 replication.

**RepA as a possible prime**

Since *in silico* analysis predicts that RepA belongs to the PolII group of proteins that are capable of functioning as both DNA polymerases and primases, therefore we addressed the possibility that RepA may also have primase activity in addition to DNA polymerase. To address this issue, we took advantage of the fact that RepA is not related to the host primase either structurally or evolutionarily. It belongs to the archeo-eukaryotic primase family, members of which characteristically possess the ancient RNA recognition motif, instead of the 10prim fold found in the case of eubacterial primases [33]. Therefore, we predicted that inhibitors that inhibit the host primase should not have any effect on the primase activity of RepA. To inactivate the host primase, we used doxorubicin, a drug that has been demonstrated to be capable of inhibiting mycobacterial DNA primase [34]. We argued that if the same primase is involved in the replication of both plasmid and chromosomal DNA, then treatment of plasmid transformed M. smegmatis cells with doxorubicin should lead to the inhibition of replication of both; whereas, if the chromosomal DNA replication is inhibited selectively then the plasmid copy number should increase. To address this issue, qRT-PCR analysis was performed using two markers, one of which (the rRNA gene) was specific to the host chromosome and the other (oriM) to the plasmid. We first examined whether doxorubicin inhibits mycobacterial growth. The results indicate that in the case of untreated culture, OD<sub>600</sub> increased substantially (ΔOD<sub>600</sub>, Fig. 4a) whereas in the doxorubicin-treated cells very little if any increase in OD<sub>600</sub> was observed (Fig. 4a) after 6 h of incubation. Thus, in the presence of doxorubicin, cell growth was inhibited as expected. When the plasmid copy number change following doxorubicin treatment was calculated, a twofold increase was observed (Fig. 4b). The results indicate that the plasmid does not utilize the host primase for its replication. Hence, it is most likely that RepA functions as a primase for the plasmid in vivo.

**Interaction between RepA and RepB**

Since the constructs (pSOC4 series) used to synthesize His<sub>6</sub>-tagged RepA also produce RepB (Fig. S1), therefore, there is a possibility that the latter would co-purify with the former provided that the two proteins interact with each other. We therefore, performed a pull-down assay using the affinity-tagged RepA, either full-length or its truncated versions. The fractions that were eluted from the beads were subjected to SDS-PAGE followed by either Coomassie blue staining (Fig. 5a) or Western blotting to detect the presence of either RepA (using anti-His tag antibodies, Fig. 5b) or RepB, (using anti-RepB antiserum, Fig. 5c). The results indicate that His<sub>6</sub>-tagged RepA, either full-length or the N- and C-terminal halves (indicated as N-term or C-term, respectively), could be detected in the Coomassie blue stained gel (Fig. 5a) and also in the Western blot using His probe antibody (Fig. 5b). However, when probed with anti-RepB antiserum, cross-reactive bands corresponding to RepB were found to be present only in the affinity-purified samples corresponding to RepA and RepA(C-term) but not RepA (N-term) (Fig. 5c). The results indicate that RepB does not co-elute with the N-terminal domain of RepA but does so with the C-terminal one. To further confirm the results, the respective affinity-purified samples were used to perform origin binding assays using [γ-<sup>32</sup>P]labelled probes (EMSA). By using this technique it should be possible to detect trace amounts of RepB if present in the RepA samples. The result (Fig. 5d) indicates that the RepA(full-length) and RepA(C-term) samples gave significant origin binding activity, indicating the presence of RepB in them. However, in the case of the RepA(N-term), no binding was observed, indicating the absence of RepB in the sample. Thus, the results show that RepB-dependent binding can be detected only in affinity-purified RepA and RepA(C-term) samples and not in the RepA(N-term) one. Hence RepB interacts with the C-terminal domain of RepA specifically.

Another question that arises at this stage is whether the binding activity observed in the RepA samples is due to RepB alone or a complex between RepA and RepB. To answer this question, an antibody supershift EMSA experiment was done using an antiserum that specifically recognizes the N-terminal domain of RepA. When this antiserum was used, a positive supershift was observed in the case of RepA (Fig. 5e, lane...
marked RepA). To prove that this super-shift is not a nonspecific one, a parallel experiment was performed using a RepA (C-terminal)–RepB fusion polypeptide, which was also used in footprinting analysis (refer to Fig. 8). This fusion polypeptide was found to bind to the probe, but no super-shift was observed under the same condition (Fig. 5e), which is expected since the antiserum is RepA N-terminal-specific (Fig. S2). Since the negative control did not yield a supershift, therefore we conclude that the shift observed with the RepA full-length protein must be specific in nature. The conclusion from this analysis is that during their synthesis in *E. coli*, the two replication proteins associate forming a complex capable of binding to the origin.

**Origin binding by RepA and RepB complexes formed in vitro**

The observations presented above indicate that RepA interacts with RepB through its C-terminal domain. To further demonstrate the possibility of complex formation between the two, we performed electrophoresis mobility shift assay (EMSA) using combinations of RepB and RepA. In the case of RepA, we preferred to use the MBP-RepA fusion instead of the His₆-tagged protein as the protein sample obtained was relatively more enriched for RepA (refer to Fig. 2d). The result (Fig. 6a) shows that in the presence of RepB, either alone or in combination with MBP-RepA, the desired complex was formed. To investigate whether MBP-RepA becomes a part of the complex when added along with RepB, supershift assay was carried out using anti-RepA antiserum. When this antiserum was added to the binding reaction, where only the RepB antigen was present a relatively minor supershift was observed. This supershift (shift 1) is clearly unrelated to RepA, as it was produced in its absence (Fig. 6a, lane 3). In the case where MBP-RepA was incorporated, supershift 1 could be detected, but in addition a higher-order supershifted band (shift 2) was observed to be present in the autoradiogram (lane 5). When the same experiment was performed with an N-terminal deleted

![Fig. 5. Pull-down assays for RepA–RepB interactions. His₆-tagged RepA or its truncated derivatives comprising either the N- or C-terminal domains, indicated as RepA(N-term) and (C-term), respectively, were synthesized in *E. coli* from pSOC4, pSOC9 and pSOC12 constructs, respectively (refer to supplemental Table S1), and subjected to pull-down assays. (a–c), SDS-PAGE and Western blot analysis of the samples eluted using either His probe antibody (anti-His) or anti-RepB antiserum (anti-RepB). The bands corresponding to either full-length RepA or its truncated versions are indicated in (a) and (b) and that of RepB in (c). Lanes marked M represent Mr marker. (d) EMSA assay of the same samples. (e) Antibody supershift EMSA (using anti-RepA antiserum) performed with the affinity purified RepA sample. A fusion protein RepA(C-terminal)–RepB, indicated as RepA(C)–RepB, that did not cross-react with anti-RepA antiserum (refer to Fig. S2) was used as a negative control. In (d) and (e), the bands corresponding to the free probe, the DNA-protein complexes and the supershift are indicated.
version of RepB, ΔNRepB (amino acid residues 1–32 deleted, Table S1) (refer to Fig. 2f), supershift 2 could not be detected (Fig. 6a, lanes 7 and 9). The results indicate that MBP-RepA becomes a part of the DNA–protein complex when added externally. The results also show that the N-terminal region of RepB has a significant role to play in the establishment of the interaction between RepA and RepB.

To define the region of RepA which interacts with RepB, we used the MBP-RepA and MBP-RepA(N-terminal) polypeptides to perform the supershift assay (Fig. 6b). In this experiment either MBP-RepA or MBP-RepA(N-terminal) was added along with RepB. The result shows that when MBP-RepA was added, the specific supershift (shift 2) induced by anti-RepA antiserum was observed (Fig. 6b lane 6); whereas, when MBP-RepA(N-terminal) was used, no such supershift was detected (lane 10). These results indicate that it must be the C-terminal domain of MBP-RepA that is essential for interaction with RepB. The results obtained from this assay also confirm that RepA does not have origin-binding activity of its own, as is evident from the observation that in the lanes where only MBP-RepA is present, DNA–protein complexes were undetectable (Fig. 6b, lane 3).

**Direct evidence for RepA and RepB interaction in vitro**

The EMSA supershift experiments provide indirect evidence for the interaction between RepA and RepB. To investigate whether such interaction does indeed take place, we performed ELISA-based assays in which either His<sub>6</sub>-tagged RepB or its N-terminal deletion derivative (ΔNRepB) were immobilized on NI-NTA HisSorb plate in triplicate. MBP-RepA was then added to all these wells (Fig. 7a, MBP-RepA). All of these wells were then probed with anti-RepA antiserum. As a loading control, parallel ELISA was carried out with the same amount of RepB, as in the experimental wells, in triplicate. These wells were probed with anti-RepB antiserum (Fig. 7a, RepB and ΔNRepB). The results (Fig. 7a) show that the signal was highest when MBP-RepA was added to RepB coated wells (black bar). In comparison, the signals were significantly lower when only MBP-RepA was present (light shaded bar). This result indicates that MBP-RepA binds to the wells in a RepB-dependent manner, which can be possible only if MBP-RepA interacts with RepB. Interestingly, when MBP-RepA was added to the well in which ΔNRepB was immobilized, little or no binding was observed (checkered bar).

Since there is no significant difference between the amount RepB and ΔNRepB immobilized (refer to Fig. 7a, hatched bar compared to white), therefore the only reason for the lower signals obtained in the case of MBP-RepA added to ΔNRepB must be due to a lack of any significant interaction between the two. The results obtained are consistent with those obtained from antibody-supershift assay, in which it was shown that MBP-RepA does not become a part of the origin complex if ΔNRepB is the mediator instead of RepB.

A similar experiment was performed in which we wanted to investigate whether a truncated RepA, MBP-RepA(N-term), binds RepB or not. The results (Fig. 7b) show that whereas in the case of MBP-RepA (black bar), the optical density (OD<sub>405</sub>) was about three times higher than background (light shaded bar), in the case of MBP-RepA(N-term) the same was almost equal (checkered bar compared to the hatched one). To ensure that the same amount of RepB was used in each case, the RepB antigen alone was immobilized in triplicate and probed with the corresponding antiserum (white bar). The standard error of mean derived was very small, at least not greater than that observed in the experimental set using MBP-RepA and anti-RepA antiserum. Hence, the observed difference in the extent to which MBP-

---

**Fig. 6.** Antibody supershift EMSA to investigate the role of (a) RepB N-terminal and (b) RepA C-terminal, in the interaction between RepA and RepB. Affinity-purified MBP-tagged RepA either full-length or N-terminal domain only (N-term) was incorporated in origin binding assays performed using either RepB or its N-terminal deleted version, ΔNRepB, as indicated. Supershift assay was performed using antiserum against RepA. The bands corresponding to a nonspecific supershift (shift 1) and the RepA specific supershift (shift 2) are indicated by broken and intact arrows, respectively. Bands corresponding to the RepB-bound complex and free probe are shown by dotted and thin arrows, respectively.
RepA and MBP-RepA(N-term) interact with RepB is not due to well-to-well variations in the amount of RepB immobilized. The conclusion from these experiments is that a C-terminal-deficient version of RepA is unable to interact with RepB. Overall the results firmly point to the fact that RepA and RepB interact with each other and that the interaction involves the N-terminal domain of RepB and the C-terminal one of RepA.

Possible consequences of recruitment of RepA by RepB

The results presented above indicate that RepA, specifically its C-terminal region, can associate with RepB. How does this association help in the activation of the pAL5000 origin? In previous crystallographic studies using ColE2 Rep C-terminal domain (PriCT_1-HTH region, refer to Fig. 1b) it was reported that this domain interacts with the origin resulting in it opening up at a specific site [35]. Domain comparison between ColE2 Rep and pAL5000 RepA indicates that whereas in the case of the former, the origin recognizing the HTH domain is contiguous with PriCT_1, in the case of the latter it is not – it is a part of RepB, a separate polypeptide. However, as RepB physically interacts with the C-terminal PriCT_1 domain of RepA, we hypothesized that a RepA(C-terminal)–RepB fusion polypeptide would act in the same way as the C-terminal (PriCT_1-HTH) domain of ColE2 Rep. To address this possibility, we created the chimera RepA(C-term)–RepB (Fig. 8a). The resulting chimeric protein, which had a His$_6$ tag at the N-terminal end, was purified by affinity chromatography and its identity was established by Western blotting (Fig. 8b). Footprinting experiments were then performed using either RepB or the RepA(C-term)–RepB fusion protein. The results show that, as expected, RepB yielded the desired footprint, which was localized over the H site (Fig. 8c, lane marked RepB). The fusion polypeptide RepA(C-term)–RepB, however, gave an extended footprint that covered not just the H site but the adjacent A/T rich sequence as well. Not only that, within the extended region of the footprint DNaseI, hypersensitive bands were observed to be present when RepA(C-term)–RepB was used [Fig. 8c, lane marked RepA(C-term)–RepB]. The hypersensitive sites (indicated by arrows) correspond precisely to a site within the origin (Fig. 8d, corresponding arrows) that was described in an earlier study as highly bendable [8]. The role of the PriCT_1 domains of the two proteins pAL5000 RepA and ColE2 Rep therefore appear to be the same. If recruited to the origin by an origin-binding protein such as RepB, this domain is expected to bring about origin deformation or unwinding.

**DISCUSSION**

The present study reports the molecular characterization of RepA [36], one of the two replication proteins, the other being RepB which the plasmid pAL5000 uses to replicate its...
genome. Although RepB acts as an origin binding protein, RepA has no such ability [4]. The question therefore remains as to what is the role of RepA. Previous attempts to understand the function of RepA failed primarily because the protein is difficult to obtain in the soluble form and also due to its inherent protease sensitivity [4]. Overall, RepA appears to be a difficult protein to work with. However, we succeeded in solubilizing the protein by co-producing GroELS chaperone [19]; although, in the process we ended up in recovering the GroEL-bound form of the protein instead of just the protein itself. It is very likely that GroEL has binding sites on RepA, particularly in its N-terminal region. The purified protein was found to possess the ability to act as a DNA polymerase, as was predicted from the bioinformatic analysis. That the activity observed was due to RepA, and not GroEL or any other co-eluting protein, became evident from the observation that a truncated version of the protein RepA(N-term) failed to show activity. Null activity was also observed when both the conserved aspartates D83 and 85 were substituted with alanine. It appears that under in vitro conditions there is a degree of redundancy in the way the two residues participate in the polymerization activity and this is observed particularly at high Mg\(^{2+}\) ion concentration. If the concentration of the ion is reduced to a suboptimal level, then the observed redundancy disappears. However, in vivo, no such issues were there and substitution of any one of the two D residues abolished plasmid-replication activity. That the loss in activity was specifically due to inactivation of RepA, and not any other reason, was confirmed by performing a complementation experiment using RepA supplied from an integrated copy of the gene in trans. The mutated aspartate residues are conserved in evolution and are present in similar positions in a variety of DNA-synthesizing enzymes. These include DNA polymerase X [30, 37] and the archeo-eukaryotic primases, the latter having the ability to function both as DNA polymerase as well as primase [26, 33]. The fact that RepA is evolutionarily related to the archeo-

![Fig. 8. Origin deformation induced by the RepA C-terminal domain in conjunction with RepB. (a) Schematic representation of the His\(^{6}\)-tagged RepA(C-term)–RepB fusion polypeptide, which was affinity-purified and used for DNA footprinting experiments. (b) SDS-PAGE analysis of the eluted fraction and its corresponding Western blot, using His probe antibody (anti-His). The band corresponding to the fusion polypeptide is indicated by an arrow. The markers used in lane M are the same as mentioned earlier (Fig. 2c). DNaseI footprinting analysis of the lower strand of the origin fragment. The marker lane and the lane without protein are indicated as A+G and (−), respectively. The proteins used were RepB or RepA(C-term)–RepB. The concentration of the protein used (0.5 µM) was the same for both. The high and low affinity binding sites are indicated by black and white bars, respectively. The jagged edge of the white box implies part representation of the L site. The shaded box represents a 15 bp element that is conserved in various pAL5000-related plasmids [38]. (d) DNA sequence of the origin fragment used in the footprinting assays. The hypersensitive sites are indicated by arrows. The H–L sites and the 15 bp conserved region are shaded as in (c).]
eukaryotic primases [26] prompted us to investigate whether RepA can function as a primase too. Attempts to demonstrate primase activity under in vitro conditions using either dNTP or NTP precursors, or both, failed to yield results, and therefore we took an indirect approach to address the issue. We argued that if RepA can act as a primase, then plasmid replication will continue to occur even if the host primase is inhibited. From the literature we found that doxorubicin, a well-known anticancer drug, can inactivate eubacterial DNA primases [34]. In the presence of doxorubicin, therefore, one expects that the plasmid will continue to replicate even if the chromosomal replication is stopped, provided that RepA can substitute for the host primase. The increase in copy number of the E. coli-mycobacteria shuttle vector pMC2 in doxorubicin-treated cells indicates that the plasmid can replicate even if the host enzyme is rendered non-functional. Indirectly, such a result indicates that pAL5000 depends on RepA to prime the synthesis of its replicas.

Another issue that has been addressed is how RepA and RepB cooperate with each other to bring about plasmid replication. We provide several lines of evidence to demonstrate that RepA and RepB physically interact with each other; these include co-purification of RepB with RepA, RepA-specific supershifts in EMSA assays, and positive results obtained using ELISA-based methods. Overall, the results suggest that the two proteins form a complex and that complex formation involves the N-terminal region of RepB and the C-terminal of RepA.

We established in a previous study that RepB is a sigma factor family homologue [8]. Based on the findings presented here, we would like to extend the similarity between RepB and sigma factors, one-step ahead. We propose that the similarity between RepB and sigma factors is not restricted to sequence level, but there is an apparent functional similarity as well. Both function by recruiting their partners, the DNA replicase RepA in the case of RepB and RNA polymerase in the case of sigma factors, to their respective target sites in the DNA. RepB and sigma factors, therefore, appear to be paralogues of each other— they probably evolved from the same ancestor, to accomplish different functions. Considering that plasmids are likely to be extant versions [2] of ancient replicons, it appears that in the evolutionary past sigma factors were involved in not only transcription, but replication as well. Proof of the concept comes from the observation that when the C-terminal domain of RepA is

---

**Fig. 9.** (a) Schematic diagram illustrating how RepB binds to the H site (black box), through its C-terminal HTH domain and recruits RepA(PriCT_1) to the origin by establishing protein–protein contacts (double-edged arrow). The net result being increased DNaseI hypersensitivity at the sites marked by up arrows within the 15 bp conserved region (shaded box). Asterisks indicate the nucleotides that are contacted by RepB (refer to Fig. S3). (b) A model comparing the mechanism of action of RepB with that of sigma factors. In the schema, the C-terminal HTH domain of RepB and the PriCT_1 of RepA are considered to be the functional equivalents of regions 4 and 2 of sigma factors, respectively. Interaction of RepB and sigma factors with their cognate elements in the DNA allows the recruitment of the corresponding nucleotide polymerizing enzymes, RepA (a DNA replicase) and RNA polymerase (RNA pol), respectively, to their sites of action. In both cases, one of the consequences of the recruitment is the opening up of the DNA at specific sites, as shown.
fused to the N-terminus of RepB, the combination develops the ability to distort the DNA, specifically within a 15 bp conserved region found in various plasmids of the pAL5000 family. Our present study, apart from unravelling the recruitment aspect, also gives added insight into the role of the 15 bp conserved region. In previous studies, no footprint was observed over this region, indicating that this segment is always excluded from interaction with the replication proteins. For the first time, in this study, we were able to generate a footprint over this region. Since RepB alone does not protect this region, but the RepA (PriCT)–RepB does, we conclude that the 15 bp conserved segment is the site within the origin where the PriCT_1 domain of RepA interacts with the help of RepB. Not only does PriCT_1 bind this region, it also denatures it at a precise site that coincides with the highly bendable AAATAT sequence located within the 15 bp region. The results obtained thus give added insight into not only how RepA and RepB interact, but the reason behind the conservation of the sequences within the 15 bp region as noted previously [38]. The overall implications of the results obtained in this study may be summarized as shown in the model given (Fig. 9a). RepB binds to the H sites and recruits RepA in such a way that the C-terminal domain of RepA is positioned precisely over the A/T rich origin, so that this region can be melted. We notice that there is significant similarity between the functioning of RepB, on one hand, and sigma (σ) factors, on the other. As shown in the scheme (Fig. 9b), the regions 4 and 2 of sigma factors interact with the −35 and −10 promoter elements, respectively, resulting in the opening up of the DNA at the A/T rich −10 site. A similar phenomenon occurs in case of RepB. It binds to the H site and recruits the PriCT_1 domain of RepA to the A/T rich 15 bp conserved region within the origin of replication (Ori) resulting in it opening up. The results presented in this study further strengthen the proposition made earlier [8] that RepB-like plasmid replication proteins are distantly related to sigma factors.

Funding information
This study was funded by Council of Scientific and Industrial Research (CSIR), India (37/1619/14/EMR-II).

Acknowledgement
We thank Mr Prabir Halder for his technical assistance.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References


Edited by: H. Gramajo and S. V. Gordon

---

**Five reasons to publish your next article with a Microbiology Society journal**

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

**Find out more and submit your article at microbiologyresearch.org.**