

MYMIV replication initiator protein (Rep): Roles at the initiation and elongation steps of MYMIV DNA replication

Dharmendra Kumar Singh, Punjab Singh Malik, Nirupam Roy Choudhury, Sunil Kumar Mukherjee*

Plant Molecular Biology Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), Aruna Asaf Ali Marg, New Delhi–110067, India

ARTICLE INFO

Article history:

Received 23 April 2008

Returned to author for revision 6 June 2008

Accepted 7 July 2008

Available online 15 August 2008

Keywords:

Geminivirus

Mungbean yellow mosaic India virus

Replication initiator protein (Rep)

Common region (CR)

DNase I footprinting

ABSTRACT

In order to explore the mechanism of geminivirus DNA replication, we show that the Replication initiator (Rep) protein encoded by *Mungbean yellow mosaic India virus* (MYMIV), a member of the family *Geminiviridae*, binds specifically to the iterons present in the viral DNA replication origin (CR-A) in a highly ordered manner that might be a prerequisite for the initiation of replication. MYMIV Rep also acts as a helicase during the post-initiation stage and is upregulated in presence of the RPA32 subunit of Replication Protein A. The implication of these findings on the initiation and elongation stages of MYMIV DNA replication has been discussed.

© 2008 Elsevier Inc. All rights reserved.

Introduction

The geminiviruses are a diverse group of plant infecting pathogens that belong to the family *Geminiviridae*. The viruses are characterized by single-stranded, circular DNA genome embedded in the twin icosahedral capsids and are transmitted by an insect vector *Bemisia tabaci* (white fly). The family *Geminiviridae* has been further categorized into four genera, namely *Begomovirus*, *Curtovirus*, *Topocuvirus* and *Mastrevirus*, depending on their genomes, transmissibility and host range (Buck, 1999). *Mungbean yellow mosaic India virus* (MYMIV) belongs to the genus *begomoviridae* and is characterized by bipartite (DNA 'A' and 'B') genomes. This virus causes severe damage predominantly to the leguminous crops such as French bean, Mung bean, *Glycin max*, *Cajans cajan*, etc. in the northern part of the Indian subcontinent (Varma and Malathi, 2003).

The geminiviruses including MYMIV replicate inside the nucleus of the infected host cell, predominantly by rolling-circle replication (RCR) (Gutierrez, 1999). Following the vector mediated entry of the virus inside the plant cell, the first step of MYMIV multiplication is the conversion of the viral ssDNA into a double-stranded (dsDNA) replicative intermediate (RFI form), aided solely by the host-encoded factors. This dsDNA serves as an active template for RCR as well as for the generation of viral transcripts. In the second phase, the viral factors along with the cellular factors synthesize ssDNA using the dsDNA template.

The ~230 bp Common Region (CR), which is common between both the viral DNA components (DNA-A and DNA-B) of most of the *begomoviruses*, harbors the origin for both DNA replication and transcription of the viral genome. The MYMIV-CR-A comprises four highly conserved iteron sequences (CGGTGT), which could predictably be the binding sites of the Replication initiator protein (MYMIV Rep) (Fontes et al., 1992). The presence of bipolar putative Rep-binding sites on both sides of the stem-loop structure of 'CR-A' is a unique feature of the MYMIV origin of DNA replication. It is presumed that one of the earliest events in the initiation stage of RCR is the high order binding of Rep to the iteron sequences, leading to the structural distortions and destabilization of origin. The MYMIV-CR-A also contains a highly conserved secondary structure, which comprises a GC rich stem and AT-rich loop regions. The formation of the secondary structure is supposed to be facilitated by the action of Rep protein. The Rep protein eventually cleaves the viral (+) strand DNA at a conserved TAATATT¹AC sequence between 7th and the 8th position located within the loop (Pant et al., 2001). This nicking makes the 3'-OH end of (+) strand free for extension of the replication fork that assembles at the nicked site.

The extrusion of a cruciform structure in the form of a stem-loop is speculated to be an essential step at the initiation step of RCR and a prerequisite for the subsequent nicking by Rep protein. However, the molecular details of its formation are unknown. In this report we confirm that the binding of MYMIV Rep to its iteron sites is highly ordered and investigate the nature of this binding and consequent structural distortions at the origin using the DNase I and KMnO₄ footprinting assays.

The viral factors alone are not sufficient to carry out the replication of the whole viral genome to its own completion. They require the help

* Corresponding author. Fax: +91 11 26742316.

E-mail address: sunilm@icgeb.res.in (S.K. Mukherjee).

of many host-encoded factors at various stages of viral replication. Rep protein is known to interact with various proteins of the host replication and cell cycle machinery, such as pRBR, PCNA, RFC, RPA32, etc. (Xie et al., 1995, 1996; Bagewadi et al., 2004; Luque et al., 2002; Singh et al., 2007). MYMIV Rep also acts as a helicase during the post-initiation stage of RCR (Choudhury et al., 2006). Recently, we have also shown that Rep protein directly interacts with RPA32, a 32 kDa subunit of Replication protein A (RPA), an eukaryotic, single-stranded DNA binding protein. This interaction modulates the biochemical properties of Rep by downregulating its nicking activity and enhancing its ATPase activity (Singh et al., 2007). Here, we show that MYMIV Rep–RPA32 interaction highly modulated the intrinsic helicase activity of the Rep, too, which might be essential for the Rep protein mediated unwinding of the origin. Such changes might be necessary for the formation as well as the progression of a functional replication fork.

Results

The replication initiator protein (Rep) formed three complexes with the ori containing common region (CR-A) of MYMIV DNA-A

To determine whether MYMIV Rep is able to bind to the MYMIV regulatory region (CR-A), we first carried out filter binding assay with

labeled CR-A DNA and purified His-Rep protein showing that the purified His-Rep bound to the CR-A DNA in a dose dependent manner (Pant et al., 2001; data not shown). In order to understand the detailed nature of this Rep–DNA binding, we carried out electrophoretic mobility shift assays (EMSA) using purified GST-Rep, His-Rep and MBP-Rep fusion proteins, and labeled CR-A. The Rep–DNA complexes were too heavy to enter 5% polyacrylamide gel under the electrophoretic conditions but were resolved on a 1% agarose gel matrix, and were visualized following autoradiography. The results showed that with increasing concentrations (0.25 μ g–1.5 μ g) of differently tagged Rep proteins in the binding reaction the Rep–DNA complexes retarded at three different levels viz., A, B and C (Fig. 1a, lanes 2–7; Fig. 1b, lanes 2–4 and Fig. 1c, lanes 2–3). The formation of higher order complexes (e.g., C complex) required higher dosages of protein. It is obvious from all these figures (Panels a, b and c) that the nature of the mobility shifts of CR–Rep complexes (marked A, B and C) was qualitatively similar with all three versions of recombinant proteins.

The sequence specificity of Rep–CR-A binding was addressed by performing competition assays with unlabelled CR-A (Fig. 1c, lanes 4–5) and with unrelated heterologous sequences (lane 6). The Rep–CR-A complexes were not competed out with DNA lacking CR-A sequences, which suggested the high specificity of the Rep–CR-A interaction. As the recombinant Rep used in EMSA contained MBP as

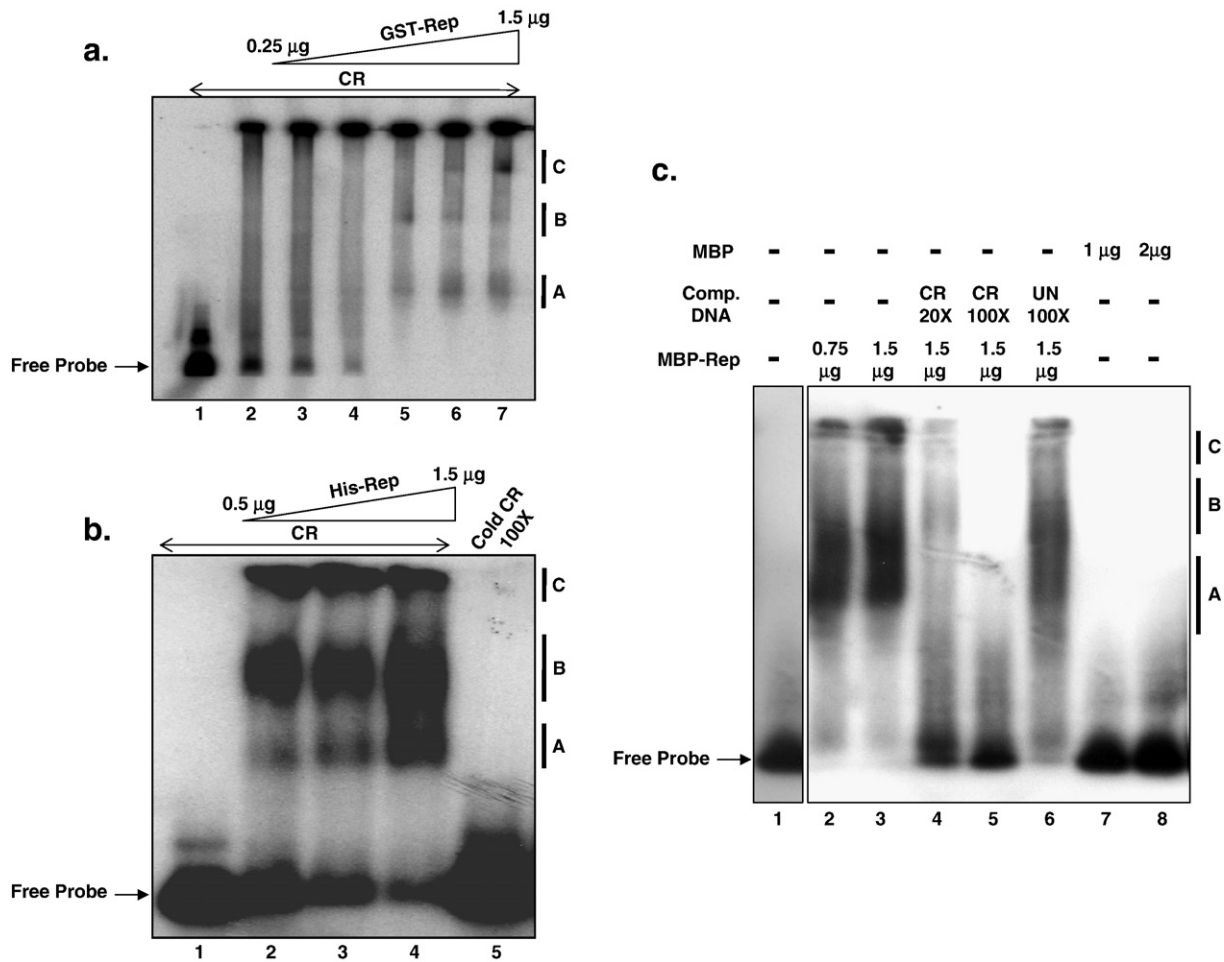


Fig. 1. Electrophoretic mobility shift analysis showing the binding of MYMIV CR-A DNA fragment by various Rep fusion proteins and the effect of competitor DNAs on the binding. (a) $5'$ - 32 P labeled 216 bp CR-A DNA fragment was incubated in the absence (lane 1) or in the presence of increasing amounts (0.25 μ g–1.5 μ g) of GST-Rep protein (lanes 2–7) and resolved on 1% agarose gel followed by autoradiography. (b) The labeled 216 bp CR-A DNA fragment was incubated in the absence (lane 1) or in the presence of increasing amounts (0.5 μ g–1.5 μ g) of His-Rep protein (lanes 2–4) and resolved on 1% agarose gel followed by autoradiography. The result of a competition assay with 100-fold excess of unlabeled CR in presence of 1.5 μ g of His-Rep is shown in lane 5. (c) The labeled CR-A DNA fragment was incubated in the absence (lane 1) or presence of various indicated amounts (lanes 2–6) of MBP-Rep protein. Binding reactions contained 1 μ g of poly (dI:dC) and were carried out without competitor DNA (lanes 2 and 3), with excess of unlabeled homologous DNA (lanes 4 and 5) or with 100-fold excess of unlabeled heterologous (pUC19) DNA (lane 6). MBP protein (1 and 2 μ g) did not show any interaction with CR-A DNA (lanes 7 and 8). Autoradiographs in all the three panels showed three (A, B and C) distinct gel retarded Rep–CR-A complexes with increasing concentrations of Rep protein.

a tag, we checked for CR-A DNA binding ability of MBP. The results clearly demonstrate that MBP does not bind to the CR-A DNA (Fig. 1c, lanes 7 and 8) indicating that the complex formation was dependent on Rep protein alone.

The sequence-specificity and strength of different protein–DNA complexes were further checked by competition with increasing concentrations of unlabelled double-stranded (ds) oligonucleotides. The oligonucleotide 1AB (Fig. 2a), spanning the second iteron of the putative Rep-binding site and positioning distally from the Rep/AL1 initiation codon, competed out the Rep–CR-A interaction at 50× higher concentration (Fig. 2b). A linear increment in the concentration of competing oligonucleotide 1AB in the binding reaction showed that complex A was competed out first, whereas B and C were perhaps higher-affinity complexes (Fig. S1, Supplementary Data). The oligonucleotide 2AB (Fig. 2c) spanning the third iteron also competed with Rep–CR-A binding, albeit to a lesser extent (Fig. 2d). The oligonucleotide 3AB (Fig. 2e), spanning sequences of both the second and third iterons, was most competent for washing out the Rep–CR-A interaction (Fig. 2f). Interestingly, the oligonucleotide 4AB (Fig. 2g), containing the iteron on the right side of the stem-loop region, strongly competed out the Rep–CR-A interaction (Fig. 2h, lanes 4 and 5). These results indicated that MYMIV Rep can form complexes with DNA sequences located both upstream and downstream of the DNA replication initiation site. Moreover, the efficiency of protein binding was also dependent on the sequences

neighboring the iterons, as the formation as well as the competitive dissolution of the complexes was controlled by the neighboring sequences. Significantly, the competitor DNAs shown in Fig. 2c and g competed out the gel shifts shown in Fig. 2b at more than 3000-fold molar excesses of competitor DNA, shown in Fig. 2a. As only a 250-fold molar excess of competitor DNA of Fig. 2a was required to titrate out the shifted bands of Fig. 2b, it would be reasonable to assume that the observed competitions are specific for the DNAs shown in Figs. 2b, d, e and g.

The loss of various types of complexes in the presence of various types of competitor DNA was calculated using ImageQuant software and has been presented in Table S1 in the “Supplementary Data” section. The data clearly reveal that type C complex was most stable and the complexes were titrated out depending on not only the repeat sequences alone, but also on the sequences flanking the repeats.

The repetitive iteron sequence CGGTGT was further dissected to understand the importance of GG dinucleotide and TGT sequences in Rep–CR-A binding. The resulting data showed that both GG dinucleotide and TGT sequences were equally important for Rep binding as mutated oligonucleotides (i.e....CCCTGT... and ... CGGCAC....) did not compete out the Rep–CR-A interaction even with concentrations of 2000× molar excess (data not shown). These results showed that MYMIV Rep protein binds to the origin DNA in a sequence specific and highly ordered manner.

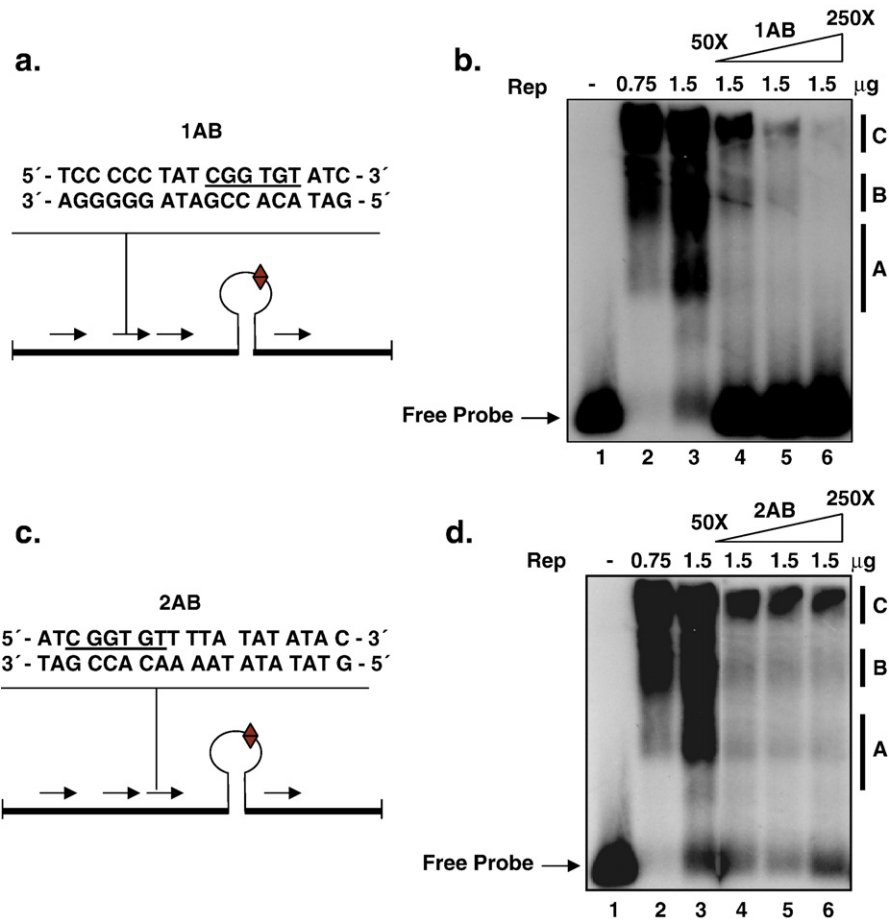


Fig. 2. Competition assays with the iterons 1AB, 2AB, 3AB and 4AB as competitors of high-affinity binding of MYMIV CR-A DNA by MYMIV Rep to form Rep–DNA complexes (A, B and C). The locations and sequences of competitor double-stranded (ds) oligonucleotides 1AB, 2AB, 3AB and 4AB are shown in panels (a), (c), (e) and (g) respectively. Lane 1 in the panels (b), (d), (f) and (h) represents control (no protein), while lanes 2 and 3 represent corresponding results with 0.75 µg and 1.5 µg of MYMIV Rep respectively. The various molar excesses of respective competitor DNAs relative to the labeled probe (CR-A) used are: 50× and 100× (lanes 4 and 5 respectively, in panels b, d, f and h) and 250× (lane 6 in panels b, d and f). The positions of retarded complexes are shown by vertical lines on the right. The binding reactions were performed with indicated amounts of protein in the presence of different concentrations of cold ds oligonucleotides. The reaction products were resolved on 1% agarose gel and subjected to autoradiography.

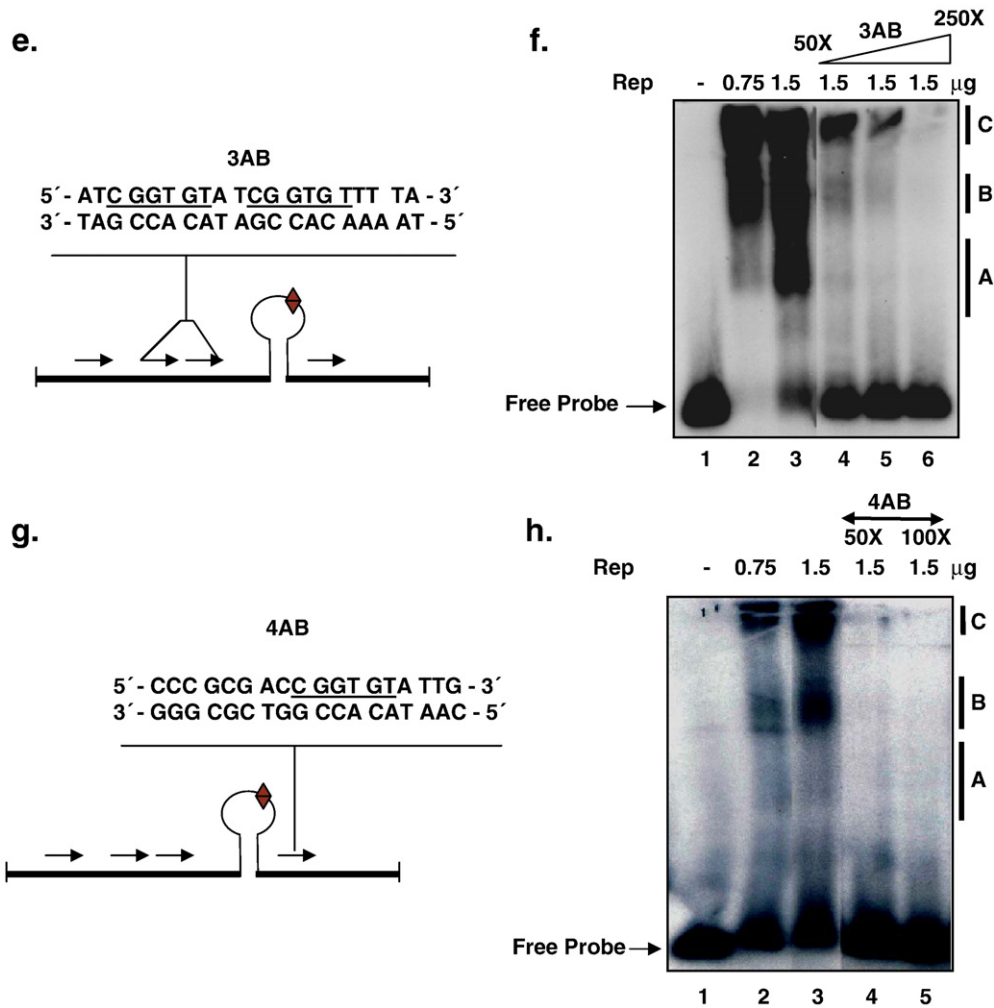


Fig. 2 (continued).

The high-affinity Rep–DNA complexes covered a large DNA region

To determine the Rep-binding site(s) of the CR-A at the nucleotide level, we carried out the DNase I footprinting assays. The 216 bp CR-A DNA fragment containing the *ori* sequences with the labeled sense strand, was used as the substrate for MYMIV Rep binding in the DNase I protection assay. The DNase I concentration was optimized for the reaction and 0.02 U of enzyme with 30 s incubation time resulted in a uniform digestion pattern of naked CR-A DNA (Fig. 3a, lane 4). The naked (for control) as well as Rep protein bound CR-A DNA was subjected to DNase I digestion. The locations of the protected and hypersensitive sites were determined by comparing with the sequencing reaction pattern, shown in parallel (Fig. 3a, left panel). The results revealed that the iterons were protected first at the low amount (0.5 μg) of Rep protein.

However, high amounts (~2.5 μg) of the Rep protein protected a large region of CR-A including the AT-rich loop region (Fig. 3a, lanes 6–9). Such extended zone of DNase I protection probably reflects the cooperative characteristics of Rep binding and the oligomeric nature of MYMIV Rep protein (Fig. 3a, Fig. 4). The middle thymine (T) residue of the CCGTGT showed hypersensitivity to DNase I in all the four iterons present in MYMIV *ori* containing CR-A. Many other prominent hyperactive sites were also observed indicating that Rep binding induced conformational changes at origin sequence. The conformational changes might be important for the downstream activities of Rep, namely, the site-specific nicking at the loop region. A tyrosine residue (Y103) in Rep controls the nicking activity as although the mutant Rep (Y103F) did bind to CR-A and protected it as efficiently as the wild type Rep, it failed to nick the DNA unlike the wt Rep (Raghavan et al., 2004; data not shown). The CR-A was

Fig. 3. Footprinting analysis of Rep-binding sites within the MYMIV-CR-A. (a) 5'-³²P labeled virion sense strand was used as a substrate (lane 1). Effect of different concentrations of DNase I on naked substrate is shown (lanes 2–4). The labeled CR-A DNA was incubated with increasing amounts (0.5 μg–2.5 μg) of His-Rep protein followed by DNase I digestion. Left panel shows the results of sequencing reactions carried out with the same DNA fragment and the primer (CR Sense), separated on a Urea-PAGE gel. The positions of protected regions are indicated by bars. The locations of the major DNase I hypersensitive sites are shown by asterisk. The boundaries of the footprint are uncertain (broken vertical line). (–) sign indicates absence of any DNase I. (b) Analysis of Rep induced structural distortion at MYMIV *ori* of replication using KMnO₄ footprinting. Reaction mixtures containing no Rep protein (lane 1) and 500 ng and 2.0 μg of Rep proteins (lanes 2 and 3 respectively) were incubated with supercoiled pGEMT-CR-A DNA in binding buffer at 37 °C for 30 min followed by an immediate treatment with KMnO₄. After the oxidation reaction was terminated, a primer extension assay with end-labeled primer (CR Antisense) was performed to detect KMnO₄-modified sites in the DNA, and the extended products were resolved on a 6% sequencing gel. Sequencing reactions (lanes G, A, C, and T) with control DNA were electrophoresed on the same gel next to the footprinting assays in order to provide the sequence of the modified regions. The location of the melted loop region is indicated by a bar on the right side. (c) Sequence of the MYMIV *ori* region containing CR-A highlighting the DNase I hypersensitive sites as derived from panels (A) and (B). The symbols ↓ and ↑ denote the DNase I hypersensitive sites of the virion sense and the complementary sense strands respectively, while ★ denotes the KMnO₄ hyperactive sites of the virion sense strand. The iteron sequences have been marked in red.

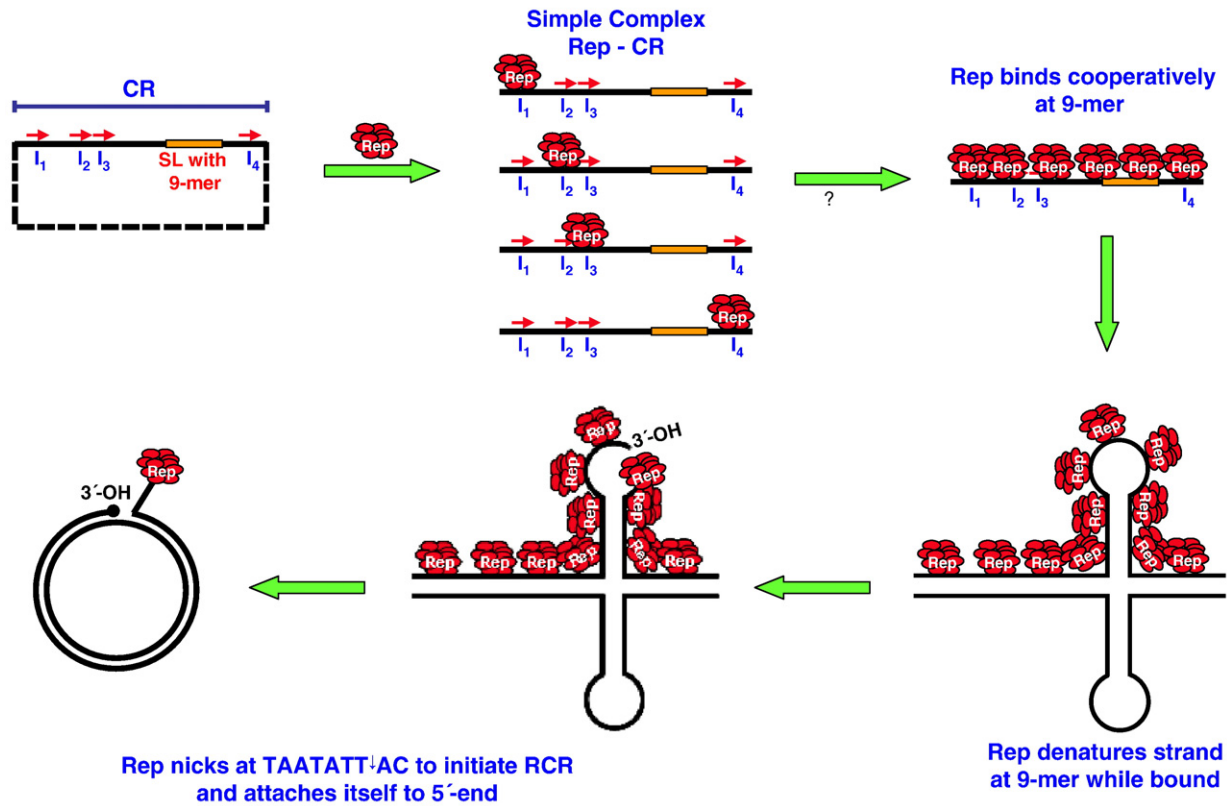


Fig. 4. Model depicting the various steps likely to be involved during the initiation of geminivirus DNA replication.

initiation of RCR. The induction of such structural changes in the CR-A (*ori*) containing plasmid DNA was probed by potassium permanganate (KMnO₄) oxidation assay and the results suggested an altered DNA structure at MYMIV origin following Rep protein binding. Following the Rep-CR-A DNA protein interaction, the CR-A containing plasmid DNA was treated with 5 mM KMnO₄. While KMnO₄ reacts poorly with B form DNA, the DNA that is either melted, apparently bent sharply or significantly untwisted becomes hyper-reactive to KMnO₄ oxidation, predominantly at thymine residues. The neighboring phosphodiester chain becomes labile and snaps upon exposure to alkaline conditions. The affected residues can be precisely located by extending a ³²P-labeled DNA primer across the region of interest with the help of the Klenow fragment of DNA polymerase I. As the polymerase stalled at the affected residue(s), analysis of the extension products by denaturing gel electrophoresis and autoradiography indicated the sites and relative extent of KMnO₄ oxidation. The results revealed that as the amount of Rep was increased up to 1.0 μg, the level of KMnO₄ oxidation also increased at certain regions (Fig. 3b, lanes 1–3). These regions, as identified from the sequencing reaction (Fig. 3b, left panel), corresponded to the loop or the 9-mer region of *ori*. The chemical reactivity of the loop region in CR-A DNA indicated that MYMIV Rep protein had perhaps caused melting of some of the origin sequences including the conserved nonamer sequence. In order to rule out the possibility that a contaminating *E. coli* protein might cause enhanced KMnO₄ sensitivity, we purified His-PCNA protein (Bagewadi et al., 2004) using the same protocol and used this protein for KMnO₄ footprinting assays. With 3 μg of His-PCNA, the footprinting pattern was similar to that shown in lane 1, Fig. 3b. It is noteworthy that the footprinting analyses were carried out with both His-Rep and MBP-Rep; and in both cases the patterns of footprinting were qualitatively identical.

Rep protein binds to the double-stranded DNA origin but can create a nick only in a single-stranded DNA substrate. This requirement of a single-stranded structure at the initiation site can thus be provided by Rep induced denaturation. The KMnO₄ sensitivity revealed that 'T' nucleotides around the initiation site became hyper-reactive when Rep protein was allowed to bind the *ori* sequences prior to the chemical treatment. The results of the two different types of footprinting analyses have been summarized in Fig. 3c, which indicate that the Rep-DNA interaction at the origin goes through a multi-step process. Overall, these results provide new insights into the role of Rep protein in relation to the initiation of MYMIV-viral DNA replication.

RPA32 from pea enhances the intrinsic helicase activity of MYMIV Rep

Following initiation, replication fork formation and its elongation are required for multiplication of viral DNA template. The data from our laboratory suggest that the MYMIV Rep along with some cellular replication factors play the key role in these activities. Recently, we demonstrated that MYMIV Rep has an intrinsic helicase activity, which is indispensable for viral DNA replication, and the cellular factor RPA32 enhances its ATPase activity (Singh et al., 2007; Choudhury et al., 2006). Therefore, we further checked the effect of RPA32 on the intrinsic helicase activity of MYMIV Rep, using the 5'-³²P radiolabeled 23-mer oligonucleotide annealed to M13mp18 ssDNA as substrate. The substrates were incubated either with the Rep protein alone (0.5 μg) or with Rep plus RPA32 (0.5–2.0 μg) proteins at 37 °C for 30 min. Immediately on completion of the reaction, the products were separated on a 15% polyacrylamide gel. The results revealed that RPA32 stimulated the helicase activity of MYMIV Rep by more than 10-fold (Fig. 5). This enhancement of Rep's helicase activity is specific for pea RPA32 as the stimulatory

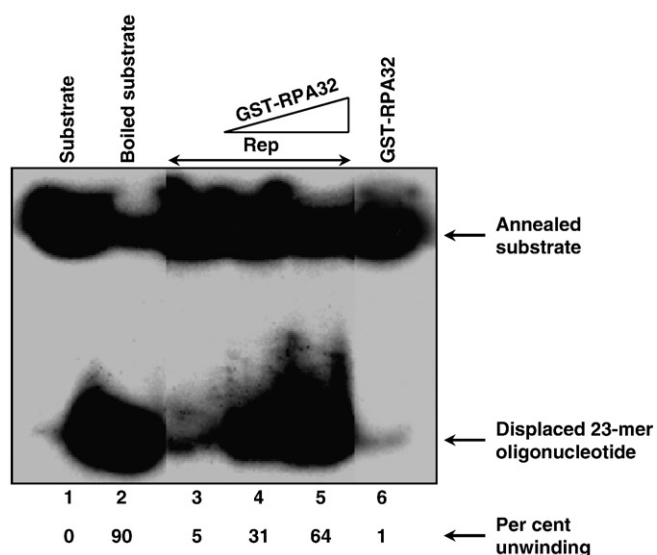


Fig. 5. Intrinsic helicase activity of recombinant MYMIV Rep protein increases in the presence of RPA32. Autoradiogram shows the 5'-³²P labeled 23-mer oligonucleotide annealed to M13mp18 DNA, used as the substrate (lane 1) and boiled substrate (lane 2) as controls. The helicase activities of MYMIV Rep protein (0.25 µg) in the absence (lane 3) and in the presence of increasing amounts (0.5–2.0 µg) of RPA32 (lanes 4–5) are shown. The corresponding helicase activity of RPA32 (2.5 µg) is shown in lane 6. Percent unwinding as quantified using ImageQuant TL software are presented at the bottom of the autoradiogram.

effect was not observed with *E. coli* single-stranded DNA binding protein (SSB) (data not shown).

Discussion

The formation of protein complexes is an important feature of origin recognition and initiation of DNA replication in bacteria, fungi and mammals. In MYMIV, the putative zone of initiation of RCR has been located in a common region (CR) of both the DNA-A and DNA-B components as indicated by *in silico* analyses. During initiation of viral DNA replication, Rep binding to the viral origin (CR) is one of the foremost events and a prerequisite for the subsequent assembly of the replication complex at the viral origin of replication. The repeat elements at *ori* had been previously postulated to be Rep specific-binding sites (Arguello-Astorga et al., 1994b). Fontes et al. have established that the recognition sequence required for Rep binding to the cognate intergenic region of the bipartite *begomoviruses*, viz., *Tomato golden mosaic virus* (TGMV) and *Bean golden mosaic virus* (BGMV), is a 12–13 bp sequence element containing the direct repeat elements GGTAG (for TGMV) and TGGAG (for BGMV) (Fontes et al., 1994a, 1994b). The special feature of MYMIV Rep-binding sites was the presence of the fourth iteron I₄ (CGGTGT) on the right side of the stem-loop structure (Fig. 2 and Fig. 4). This bipolar nature of MYMIV Rep-binding site is unique among *begomoviruses*. Our results with MYMIV further supported the suggestions that the direct repeats functioned as the core elements in Rep-recognition and binding and that the sequence of the repeat element is specific for each geminiviral Rep.

The electrophoretic mobility shift assays suggested that Rep bound to *ori* DNA as a large multimer, rendering Rep–DNA complexes unable to enter the polyacrylamide gel. However, these complexes were suitably resolvable by agarose gel electrophoresis. We have identified at least three distinct DNA–protein complexes of the DNA replication initiator protein (Rep) and the viral regulatory sequences controlling DNA replication and transcription (Fig. 1). The presence of three retarded bands observed in an electrophoretic

mobility shift assay and the extended DNase I protected areas reflected the binding of many Rep molecules to the MYMIV intergenic region. This finding also implies that the DNA-bound Rep–Rep interaction might lead to the formation of a huge multimeric complex consisting of Rep. This speculation corroborates with our earlier report that the recombinant MYMIV Rep forms oligomers even in the native state (Choudhury et al., 2006). Further, it has been shown that the multimerization domain of Rep may be a prerequisite for DNA binding with many *begomoviruses* and this fact might hold true for all geminiviruses (Orozco et al., 1997).

An important parameter for any site-specific DNA interaction is the ability of the enzyme to distinguish the specific from the non-specific DNA. The Rep–CR–A interaction was competed out with cold CR–A sequences whereas an unrelated DNA even at higher concentration could not destabilize these specific protein–DNA complexes, suggesting that the Rep–CR–A binding is highly specific (Fig. 1c, lanes 4–6; Fig. S1, Supplementary Data). However, the data displayed in Figs. 2a–h further suggested that the sequences flanking the repeat regions also played a distinctive role in determining the strength of complex formation with Rep.

The *ori* binding sites of MYMIV Rep have been mapped by combining competitive gel-shift assays with the DNase I footprinting data (Figs. 2a–h and 3a). The competitive EMSAs were carried out to analyze binding of Rep–CR–A DNA using double-stranded oligonucleotides containing either the cognate site or mutated iteron sequences. The ds oligonucleotides spanning iteron sequences of both the sides of stem-loop structure showed the ability to compete out the Rep–CR–A interaction, suggesting their involvement in Rep binding. Of the second and third iterons (CGGTGTATCGGTGT), the third iteron sequence did not compete out the Rep–CR–A interaction as strongly as the second iteron (Figs. 2b and d). This is in contrast with the TGMV Rep-binding site where the 3'-repeat was shown to be essential, the 5'-repeat being the enhancer only. Both the dinucleotide GG and TGT sequences were found out to be equally important for MYMIV Rep binding. The presence of three distinctly retarded protein–DNA complexes in *begomoviruses* is shown for the first time in this study. In other *begomoviruses*, viz., TGMV and *Tomato leaf curl virus* (TLCV), only one type of DNA–protein complex has been identified so far that maps near to the TATA box for complementary sense (c-sense) transcription of viral genes but not around the initiation site located at the stem-loop (Akbar Behjatnia et al., 1998; Christensen and Tattersall, 2002). However, three different Rep protein–DNA complexes have been observed in *Wheat dwarf virus* (WDV), a *Mastrevirus* species. Two of the Rep–DNA complexes (C and the V) are high-affinity complexes, located in the proximity of the two divergent TATA boxes, at 150 and 90 bp, respectively, from the DNA replication initiation site. The third one, the O complex is a low-affinity complex, which can assemble under conditions supporting the DNA cleavage reaction (Castellano et al., 1999). This suggests that formation of some common types of Rep–*ori* complexes might be responsible for the initiation of rolling-circle DNA replication in many geminiviruses. However, there could be dissimilarities in the stoichiometry and other characteristics of the higher order complexes that finally lead to the initiation of site-specific nicking.

The observed protection of sequences bordering the MYMIV repeat motifs in the footprints suggested that Rep binding was not limited to repeat elements alone. The unusually large degree of *ori* DNA protection by MYMIV Rep could be a consequence of cooperative interaction among the bound Rep proteins and the binding of Rep as an oligomer. However, one cannot rule out the possibility that the protection of bordering sequences resulted from exclusion of DNase I due to formation of Rep–DNA complex involving only the repeat elements. Alternative methods of DNA footprinting analysis using smaller chemical cleavage agents would help to resolve this question. Association of a hypersensitive site (middle T in CGGTGT) with all four

iterons is an interesting feature on the virion sense strand that reflects strongly on the alterations of DNA structure following Rep-binding.

The changes in secondary structures following Rep-binding at the MYMIV origin of DNA replication have been analyzed *in vitro* by KMnO₄ footprinting (Figs. 3b and c). As the 9-mer region of CR-A seems to be denatured following Rep-binding, the CR-A region might extrude a cruciform structure keeping the 9-mer in the loop region (Fig. 4). The single-strandedness of the CR-A region is necessary for Rep mediated nicking to provide the initiation site (TAATATT⁺AC) of RCR. Further, the single-stranded region is also necessary for the assembly of the replication complex. Interestingly, the identification of Rep-binding sites on both sides of the stem-loop structure indicates the possibility of Rep mediated DNA loop formation. Since the Rep protein is highly oligomeric in nature, the possibility of Rep–Rep interaction in *cis* might lead to the formation of a DNA loop that might facilitate the cruciform extrusion and commence the RCR initiation.

Based on the above data, a model depicting the various steps in the initiation processes of the replication of geminiviral MYMIV DNA replication can be developed as shown in Fig. 4. The first step in the initiation process would involve highly sequence specific binding of presumably oligomeric MYMIV Rep protein to the iteron(s) present in the CR-A-region forming a Rep–CR-A simple complex. Such binding leads to subsequent cooperative and ordered binding of other Rep proteins to the regions in the vicinity of 9-mer (5'-TAATATTAC-3') sequence, resulting in the formation of a local stem-loop structure. This stem-loop structure forms the substrate for the nicking activity by MYMIV Rep at the conserved 5'-TAATATT⁺AC-3' sequence, marking the initiation of replication event. At the elongation stage, the 3'-OH end is extended by DNA polymerase(s), while Rep protein remains bound to the 5'-end. At this stage, however, a number of other viral and host proteins are likely to be involved, the identities of which have largely remained unexplored.

Recently, a Melting Pot Model has been described to explain the rolling-circle replication initiation of *Porcine circovirus* (PCV1/PCV2) (Cheung, 2004a, 2004b). We wondered if the same model would be applicable for the MYMIV replication initiation. However, at present, there is hardly any option to choose between these two models. There are no mutational data available at the arm(s) of the stem structure of MYMIV replication origin. Moreover, the KMnO₄ sensitivity was more pronounced on one arm of the stem compared to the other arm (Fig. 3c), a fact that cannot be explained by the Melting Pot Model. Thus, a cruciform model might be more relevant for the MYMIV replication initiation.

Upon binding to CR-A, MYMIV Rep causes limited local unwinding of the origin DNA. However, this limited unwinding might not be sufficient for the formation of a proper replication fork and also the ssDNA generated by the unwinding process needs to be stabilized. RPA may act as a stabilizing factor for the nascent single-stranded DNA generated. As reported in parvoviruses (Christensen and Tattersall, 2002), the interaction of RPA with initiator protein NS1 leads to the energy driven extensive unwinding of the nicked origin and suggested the requirement of RPA to the formation of a functional replication fork. A similar scenario could also be envisaged for MYMIV DNA replication, where the Rep–RPA complex would lead to extensive unwinding of the origin, facilitating the formation and progress of the replication fork.

Materials and methods

Isolation and purification of His-Rep, MBP-Rep and GST-Rep proteins

Construction of MYMIV His-Rep1-362 has been discussed elsewhere (Bagewadi et al., 2004). The protein was isolated from overproducing *E. coli* BL21 (DE3) cells and purified following the standard procedures. The cells were induced at 18 °C for 16 h with 0.1 mM IPTG. Induction under such a condition resulted in accumulation of a considerable

amount of soluble Rep proteins. The soluble Rep was chromatographed through Ni-NTA-affinity matrix as per the manufacturer's protocol. The eluted proteins were further purified through Heparin–Sephacrose CL6B and Q–Sephacrose columns (Amersham Biosciences, USA) and finally dialyzed against 50 mM Tris–HCl (pH 8.0), 100 mM NaCl, 1 mM DTT, 0.5 mM each of PMSF and benzamidine, and 40% glycerol. The Rep proteins were also purified with MBP and GST-tags in a similar manner. All biochemical experiments were carried out without removing the tags. These preparations were free of activities of other enzymes, namely DNA polymerase, DNA nuclease, etc. The absence of the contaminating DNA nuclease was confirmed by site-specific nicking (Pant et al., 2001) and helicase assays (Choudhury et al., 2006).

Rep–CR-A DNA binding assay

The reaction mixture containing 1 ng of 5'-³²P labeled MYMIV-A *ori* DNA (216 bp) and the indicated amount of various tagged Rep proteins in 1× binding buffer [50 mM Tris–HCl (pH 8.0), 300 mM NaCl, 5 mM MgCl₂, 2.5 mM DTT, 1 mM EDTA and 3% glycerol] was incubated at 37 °C for 30 min. The DNA–protein complexes were resolved on 1% agarose gel. Electrophoresis was carried out at 50 V in a pre-cooled 1× TBE buffer at 4 °C for 4 h to maintain the integrity of the Rep–CR-A complexes. The gel was dried and exposed to X-ray film. For competition experiments, indicated amounts of the double-stranded (ds) annealed oligonucleotides were added in the binding reaction. About 20–60 ng of unlabeled *ori* DNA and pUC19 (heterologous) DNA was used as cold competitor(s).

DNase I footprinting assay

The MYMIV–CR-A was PCR amplified using MYMIV-A (cloned in pUC19) DNA template and 5'-³²P labeled CR Sense (5'-GGGG-AATCCCTTGCCATATTTGAAGTC-3') and unlabeled CR antisense (5'-ATCGGATCCGATTGAACGACTAAAGATAAG-3') primers to isolate 5'-³²P labeled CR-A. Labeling reaction was performed using T4 polynucleotide kinase (NEB, USA) and [³²P] ATP (6000 Ci/mmol, PerkinElmer Life Sciences, USA). For DNase I protection assay on virion sense strand, the 5'-labeled CR-A served as a substrate for Rep binding. The binding reaction was performed in a total volume of 50 μl containing 5 μl of 10× binding buffer (100 mM Tris–Cl pH 7.5, 3M NaCl, 50 mM MgCl₂, 25 mM DTT, 10 mM EDTA and 30% glycerol), 1 μg poly dI:dC, 2–3 ng of end-labeled CR-A DNA fragment (~15,000 cpm) and His-Rep protein (0.5 μg – 2.5 μg). The reaction was incubated for 30 min at RT. Following incubation, 50 μl of cofactor solution (10 mM MgCl₂, 5 mM CaCl₂) was added to the reaction mix. The DNase I enzyme dilutions were freshly made and added (0.001 to 0.02 U) to each tube. After 30 min of incubation at room temperature the reaction was stopped by the addition of 100 μl of stop solution (1% SDS, 200 mM NaCl, 20 mM EDTA, pH 8.0, 40 μg/ml tRNA) followed by phenol:chloroform extraction and ethanol precipitation. The pellets were resuspended in 4.5 μl of sequencing loading dye. Scintillation counts were checked using Beckman scintillation counter and loaded (~5000 cpm/track) on a standard 6% acrylamide sequencing gel along with standard dideoxy sequencing reaction lanes. The gel was dried and subjected to autoradiography.

The structural distortion assay (KMnO₄ footprinting)

The melting at origin/structural distortion was detected by KMnO₄ oxidation assays as described previously (Sasse-Dwight and Gralla, 1988). The MYMIV origin containing supercoiled plasmid DNA (pGEMT–CR-A) (1–2 μg) was incubated with indicated amounts of Rep proteins in binding buffer for 30 min at 37 °C. KMnO₄ was then added to the reaction mixture to a final concentration of 5 mM to oxidize improperly base-paired thymine residues, followed by an

immediate addition of 2-mercaptoethanol to a final concentration of 1.0 M to terminate the reaction. The modified DNA was purified through a Sephadex G-50 column and using a ^{32}P end-labeled CR antisense primer, we performed the primer extension assay on the purified template. The extended products were subjected to electrophoresis on a standard 6% acrylamide sequencing gel along with sequencing reaction lanes. Gels were dried and autoradiographed. The protected bands in the footprint bands were quantitated using ImageQuant TL software (GE Healthcare).

Helicase assay

The 23-mer oligonucleotide (~3 pmol), having the sequence 5'-CCCAGTCACGACGTTGAAAAACG-3', was 5'-end-labeled using T4 polynucleotide kinase (NEB, USA) and [$\gamma^{32}\text{P}$] ATP (6000 Ci/mmol, PerkinElmer Life Sciences, USA). The radiolabeled oligonucleotide was annealed to M13mp18 ssDNA (5 pmol) by heating at 98 °C and then slowly cooling down the reaction mixture to room temperature. The annealed substrates were separated from free radiolabeled oligonucleotides by purifying through Sepharose CL-6B columns. The substrates were incubated with the desired quantities of proteins in a buffer containing 20 mM Tris-Cl (pH 8.0), 1 mM MgCl_2 , 100 mM KCl, 8 mM DTT, 5 mM ATP and 80 $\mu\text{g/ml}$ BSA in a total reaction volume of 20 μl at 37 °C for 30 min. Immediately after the completion of the reaction, the products were resolved on a 15% polyacrylamide gel. The gel was dried, autoradiographed and the results were analyzed by densitometric scanning using Typhoon 9210 scanner and ImageQuant TL software (Amersham Biosciences, USA).

Acknowledgment

The financial assistance to DKS from CSIR, Govt. of India, New Delhi is duly acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2008.07.010.

References

Akbar Behjatnia, S.A., Dry, I.B., Rezaian, M.A., 1998. Identification of the replication-associated protein binding domain within the intergenic region of tomato leaf curl geminivirus. *Nucleic Acids Res.* 26, 925–931.

- Arguello-Astorga, G.R., Guevara-Gonzalez, R.G., Herrera-Estrella, L.R., Rivera-Bustamante, R.F., 1994b. Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology* 203, 90–100.
- Bagewadi, B., Chen, S., Lal, S.K., Choudhury, N.R., Mukherjee, S.K., 2004. PCNA interacts with Indian mung bean yellow mosaic virus Rep and downregulates Rep activity. *J. Virol.* 78, 11890–11903.
- Buck, K.W., 1999. The *Geminiviridae* family. In: Webster, R.G., Granoff, A. (Eds.), *Encyclopedia of Virology*, second ed. Academic Press, London, pp. 517–524.
- Castellano, M.M., Sanz-Burgos, A.P., Gutierrez, C., 1999. Initiation of DNA replication in an eukaryotic rolling circle replicon: identification of multiple DNA–protein complexes at the geminivirus origin. *J. Mol. Biol.* 290, 639–652.
- Cheung, A.K., 2004a. Detection of template strand switching during initiation and termination of DNA replication of porcine circovirus. *J. Virol.* 78, 4268–4277.
- Cheung, A.K., 2004b. Palindrome regeneration by template strand-switching mechanism at the origin of DNA replication of porcine circovirus via the rolling-circle melting-pot replication model. *J. Virol.* 78, 9016–9029.
- Choudhury, N.R., Malik, P.S., Singh, D.K., Islam, M.N., Kaliappan, K., Mukherjee, S.K., 2006. The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. *Nucleic Acids Res.* 34, 6362–6377.
- Christensen, J., Tattersall, P., 2002. Parvovirus initiator protein NS1 and RPA coordinate replication fork progression in a reconstituted DNA replication system. *J. Virol.* 76, 6518–6531.
- Fontes, E.P.B., Luckow, V.A., Hanley-Bowdoin, L., 1992. A geminiviral replication protein is a sequence-specific DNA binding protein. *Plant Cell* 4, 597–608.
- Fontes, E.P., Eagle, P.A., Sipe, P.S., Luckow, V.A., Hanley-Bowdoin, L., 1994a. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *J. Biol. Chem.* 269, 8459–8465.
- Fontes, E.P., Gladfelter, H.J., Schaffer, R.L., Petty, I.T., Hanley-Bowdoin, L., 1994b. Geminivirus replication origins have a modular organization. *Plant Cell* 6, 405–416.
- Gutierrez, C., 1999. Geminivirus DNA replication. *Cell Mol. Life Sci.* 56, 313–329.
- Luque, A., Sanz-Burgos, A.P., Ramirez-Parra, E., Castellano, M.M., Gutierrez, C., 2002. Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. *Virology* 302, 83–94.
- Orozco, B.M., Miller, A.B., Settlege, S.B., Hanley-Bowdoin, L., 1997. Functional domains of a geminivirus replication protein. *J. Biol. Chem.* 272, 9840–9846.
- Pant, V., Gupta, D., Choudhury, N.R., Malathi, V.G., Varma, A., Mukherjee, S.K., 2001. Molecular characterization of the Rep protein of the blackgram isolate of Indian mung bean yellow mosaic virus. *J. Gen. Virol.* 82, 2559–2567.
- Raghavan, V., Malik, P.S., Roy Choudhury, N., Mukherjee, S.K., 2004. The DNA-A component of a plant geminivirus (Indian mung bean yellow mosaic virus) replicates in budding yeast cells. *J. Virol.* 78, 2405–2413.
- Sasse-Dwight, S., Gralla, J.D., 1988. Probing the *Escherichia coli* glnALG upstream activation mechanism in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 85, 8934–8938.
- Singh, D.K., Islam, M.N., Choudhury, N.R., Karjee, S., Mukherjee, S.K., 2007. The 32 kDa subunit of Replication Protein A (RPA) participates in the DNA replication of *Mungbean yellow mosaic India virus* (MYMIV) by interacting with the viral Rep protein. *Nucleic Acids Res.* 35, 755–770.
- Varma, A., Malathi, V.G., 2003. Emerging geminivirus problems: a serious threat to crop production. *Ann. Appl. Biol.* 142, 145–164.
- Xie, Q., Saurez-Lopez, P., Gutierrez, C., 1995. Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant DNA virus: requirement for efficient viral DNA replication. *EMBO J.* 14, 4073–4082.
- Xie, Q., Sanz-Burgos, A.P., Hannon, G.J., Gutierrez, C., 1996. Plant cell contains a novel member of the retinoblastoma family of growth regulatory proteins. *EMBO J.* 15, 4900–4908.